

UNIVERSIDAD DE LA FRONTERA
FACULTAD DE CIENCIAS AGROPECUARIAS Y FORESTALES



**BIOSÍNTESIS Y DEGRADACIÓN DE CAROTENOIDES EN
BRASSICA SPP.: EVALUACIÓN DE GEN CANDIDATO**

Tesis presentada a la Facultad de Ciencias Agropecuarias y Forestales de la Universidad de La Frontera. Como parte de los requisitos para optar al título de Biotecnólogo.

DANIELA ROXANA QUEZADA MARTÍNEZ
PROFESOR GUÍA: Dra. MARÍA LAURA FEDERICO

TEMUCO – CHILE

2012

“BIOSÍNTESIS Y DEGRADACIÓN DE CAROTENOIDES EN
BRASSICA SPP.: EVALUACIÓN DE GEN CANDIDATO”

PROFESOR GUÍA

: _____

Dra. María Laura Federico

Investigadora del Centro de Genómica
Nutricional Agroacuícola (CGNA)

PROFESOR CONSEJERO

: _____

Dra. Mónica Rubilar Díaz

Profesor Asociado
Universidad de La Frontera.

CALIFICACIÓN PROMEDIO TESIS

: _____

ÍNDICE

1 INTRODUCCIÓN	1
2 REVISIÓN BIBLIOGRÁFICA	3
2.1 Generalidades de las especies cultivables de <i>Brassica</i>	3
2.2 <i>Brassica oleracea</i>	6
2.3 Carotenoides	9
2.4 Familia de las CCDs	13
3 JUSTIFICACIÓN DEL ESTUDIO	16
3.1 Hipótesis	17
3.2 Objetivo general	17
3.3 Objetivos específicos	17
4 METODOLOGÍA	18
4.1 Material vegetal y extracción de ADN	18
4.2 Diseño de oligonucleótidos y clonamiento de gen CCD4	19
4.3 Mapeo genético utilizando polimorfismo de conformación de cadena simple (DNA-SSCP)	21
4.4 Análisis bioinformático de secuencias genómicas CCD4 en <i>Brassica oleracea</i> y predicción de la estructura proteica	22
4.5 Extracción de ARN y síntesis de cDNA	22
4.6 PCR cuantitativa en tiempo real (qRT-PCR)	23
5 RESULTADOS	25
5.1 Identificación del gen candidato causante del fenotipo flor blanca (<i>wf</i>) en <i>B. oleracea</i>	25
5.2 Identificación de las secuencias de <i>BolC.CCD4</i> en el genoma C de <i>Brassica</i>	26
5.3 Análisis de secuencia	31
5.4 Expresión de genes <i>BolC.CCD4</i> en tejidos vegetativos y reproductivos en <i>B. oleracea</i>	37

6 DISCUSIÓN	40
7 CONCLUSIÓN	42
8 REFERENCIAS	43

ÍNDICE DE FIGURAS

- Figura 1.** Relaciones entre las especies cultivadas del género *Brassica*, conocido como el Triángulo de U (U, 1935)
- Figura 2.** Relación evolutiva entre *Arabidopsis thaliana* y *Brassica*
- Figura 3.** Morfotipos de las subespecies de *Brassica oleracea*
- Figura 4.** Partes de la flor de *Brassica oleracea*
- Figura 5.** Grupo de ligamiento C3 para *B. oleracea*
- Figura 6.** Apocarotenoides y las CCDs o NCEDs responsables de su producción (Ohmiya, 2009)
- Figura 7.** Esquema de plegamiento de ACO
- Figura 8.** Método $\Delta\Delta_{Ct}$
- Figura 9.** Mapa comparativo entre *Arabidopsis thaliana* y *Brassica oleracea*
- Figura 10.** Posición relativa de los oligonucleótidos diseñados el mapeo y aislamiento de secuencias CCD4 en *B. oleracea*
- Figura 11.** Patrones de migración de los parentales de la población BolTBDH en gel MDE
- Figura 12.** Patrón de segregación de *BolC.CCD4a* y *BolC.CCD4.b* en una población de mapeo bi-parental desarrollada a partir de la cruce entre una línea flor amarilla (“*Early Big*”) y una línea flor blanca (TO1000DH3)
- Figura 13.** Localización de *BolC.CCD4.a* y *BolC.CCD4.b* en el mapa del genoma C de *Brassica*
- Figura 14.** Alineamiento múltiple de las secuencias aminoacídicas de CCD4 en *A. thaliana* y *B. oleracea*
- Figura 15.** Localización de los oligonucleótidos diseñados para aislar las secuencias codificantes y no codificantes de los genes CCD4 en *B. oleracea*
- Figura 16.** Elementos repetitivos en los genes CCD4
- Figura 17.** Diagrama Ribbon de las enzimas CCD4 de *A. thaliana*, *B. oleracea* TO1000DH3 y “*Early Big*”.

Figura 18. Tejidos analizados mediante qRT-PCR

Figura 19. Genotipos utilizados para el análisis de expresión

Figura 20. Expresión relativa de los genes CCD4 en pétalos y hojas maduras

ÍNDICE DE TABLAS

Tabla 1. Clasificación de los carotenoides

Tabla 2. Dioxigenasas de corte de carotenoides en *Arabidopsis thaliana* (Tan *et al.*, 2003)

Tabla 3. Oligonucleótidos

RESUMEN

En plantas, los carotenoides pueden ser acumulados en flores y frutos, donde sirven como pigmentos y precursores de una serie de compuestos volátiles que atraen polinizadores y aseguran la dispersión de las semillas. El presente estudio fue realizado para investigar la base genética detrás del fenotipo flor blanca (*wf*) presentado por la línea doble haploide de ciclo rápido *Brassica oleracea* L. TO1000DH3. Basados en información sinténica con *Arabidopsis*, se identificó el gen candidato (*CCD4: Carotenoid Cleavage Dioxygenase*) en el cromosoma C3, donde la característica para el color de flor (FC) fue mapeada en estudios previos. Utilizando un set de oligonucleótidos dirigido a regiones conservadas, dos genes *CCD4* (*BolC.CCD4.a* y *BolC.CCD4.b*) fueron mapeados utilizando la población de mapeo BolTBDH. Cabe destacar que el alelo *BolC.CCD4.b* de TO1000DH3 co-segregó con el fenotipo *wf* y mapeó en la localización previamente identificada en el cromosoma C3. Por otra parte, *BolC.CCD4.a* segregó independientemente del fenotipo *wf* y mapeó en el cromosoma C1. Esto fue el primer indicio de que *BolC.CCD4.b* de TO1000DH3 podría estar involucrado en la generación del fenotipo *wf*.

Se realizaron estudios de expresión génica en distintos estadios del desarrollo de pétalos en flores blancas, flores amarillas y hojas maduras. El alelo *BolC.CCD4.b* de TO1000DH3 se expresó diferencialmente en *wf*, teniendo su más alto nivel de expresión cuando los pétalos están completamente abiertos. El gen *BolC.CCD4.a* se expresó preferencialmente en tejidos verdes.

Además, el análisis de la estructura proteica demostró que el alelo *BolC.CCD4.b* de “*Early Big*” codifica para una proteína más corta, cuya estructura proteica es diferente comparada al alelo de TO1000DH3 y a otras enzimas *CCD4*. Este cambio estructural podría estar afectando la funcionalidad de la enzima o cambiando la especificidad al sustrato.

SUMMARY

Plant carotenoids can accumulate in flowers and fruits where they serve as pigments and precursors to a range of scents that attract pollinators and secure seed dispersal. The present study was conducted to investigate the genetic basis underlying the white flower (*wf*) phenotype exhibited by the *Brassica oleracea* L. double haploid (DH) rapid cycling line TO1000DH3. Based on synteny information with *Arabidopsis*, a candidate gene (Carotenoid Cleavage Dioxygenase 4, *CCD4*) was identified in chromosome C3, where the *flower color* (FC) trait had been mapped in previous studies. Using primer pairs targeted to conserved regions, two *CCD4* genes (*BolC.CCD4.a* and *BolC.CCD4.b*) were mapped using the BolTBDH mapping population. Remarkably, the TO1000DH3 *BolC.CCD4.b* allele co-segregated with the *wf* phenotype and mapped to the previously identified in the chromosome C3 location. In contrast, *BolC.CCD4.a* segregated independently of the *wf* phenotype and mapped to chromosome C1. This was the first indication that TO1000DH3 *BolC.CCD4.b* could be involved in the generation of the *wf* phenotype.

Gene expression studies in different stages of petal development in yellow, white flowers and mature leaves were performed. The TO1000DH3 *BolC.CCD4.b* allele is differentially expressed in *wf*, exhibiting the highest level of expression when the petals are fully opened. *BolC.CCD4.a* was preferentially expressed in green tissues.

In addition, a protein structure analysis showed that the “*Early Big*” *BolC.CCD4.b* allele encoded a shortened protein, with a different protein structure compared to the TO1000DH3 allele and other *CCD4* enzymes. This structural change could be impairing the enzyme functionality or changing substrate specificity.

1. INTRODUCCIÓN

Los carotenoides son compuestos isoprenoides sintetizados por todos los organismos fotosintéticos. En flores, los carotenoides son acumulados en los cromoplastos donde sirven como pigmentos y precursores de una serie de aromas producidos para atraer polinizadores y así asegurar la dispersión de las semillas.

Debido a que el color amarillo de los pétalos se debe principalmente a la acumulación de carotenoides en estos tejidos, fue interesante el estudio en *Brassica oleracea* de la existencia de genotipos con fenotipo de flor blanca. Existen dos posibles explicaciones para la aparición de los pétalos color blanco:

- (i) Una deficiencia en la síntesis de carotenoides causada por la mutación de un gen biosintético (por ejemplo: fitoeno sintasa, PSY)
- (ii) Un aumento en la degradación de carotenoides causada principalmente por el cambio expresión de genes que degradan estos compuestos (por ejemplo: CCDs o NCEDs).

Se sabe que el alelo flor blanca (*wf*) es dominante por sobre el alelo silvestre flor amarilla (*wt*) y no se observan fenotipos intermedios (flores cremas) en plantas heterocigotas. Además, durante estadíos tempranos del desarrollo de flores blancas se puede observar que los pétalos son de color amarillo, lo que sugiere que los genes involucrados en la biosíntesis de carotenoides están presentes y funcionales en plantas con fenotipo de flor blanca. Estos resultados preliminares descartan la alternativa (i), ya que deberían observarse fenotipos intermedios en individuos heterocigotos si se tratase de un gen involucrado en la síntesis de carotenoides y a su vez coincide con la alternativa (ii) que señala como posible causante del fenotipo flor blanca a un gen involucrado en la degradación de carotenoides.

Existen reportes donde la expresión diferencial de genes CCD4, pertenecientes a la familia de las *carotenoid cleavage dioxygenases*, causan fenotipos similares (flores, frutos o tubérculos blancos) en diversas especies vegetales como crisantemo, papa y durazno (Ohmiya *et al.*, 2006; Campbell *et al.*, 2010; Brandi *et al.*, 2011). Se ha demostrado que genes CCD4 cortan doble enlaces de diferentes carotenoides produciendo una variedad de apocarotenoides que presentan diversas funciones biológicas (Tan *et al.*, 2003).

2. REVISIÓN BIBLIOGRÁFICA

2.1 Generalidades de las especies cultivables de *Brassica*

El género *Brassica* forma parte de la familia Brassicaceae, la cual consta de aproximadamente 338 géneros y más de 3709 especies (Warwick *et al.*, 2006). Las especies cultivables del género *Brassica*, quienes pertenecen a la misma familia taxonómica de *Arabidopsis thaliana*, son ampliamente utilizadas en la alimentación de varias culturas, en parte, debido a la gran variedad de formas comestibles del género (Paterson *et al.*, 2001). El género *Brassica* incluye especies categorizadas como semillas oleaginosas, vegetales y condimentos, donde *Brassica napus*, *Brassica rapa*, *Brassica juncea* y *Brassica carinata* proveen alrededor del 12% a nivel mundial de los aceites vegetales comestibles (Labana y Gupta, 1993).

Además, las especies de *Brassica* son una valiosa fuente de fibra dietética, vitamina C, y otros factores saludables como compuestos anticancerígenos (Fahey y Talalay, 1995). *Brassica oleracea* y *B. rapa* abarcan muchos de los vegetales que consumimos en nuestra dieta diaria, como la coliflor y el brócoli (Paterson *et al.*, 2001). Finalmente, la semilla de *Brassica nigra* es utilizada como condimento (mostaza).

U (1935) fue el primero en establecer una relación botánica entre las especies de *Brassica* a partir de un estudio taxonómico (Figura 1) donde propuso que las especies con el número de cromosomas más grande *B. juncea* (AABB, $2n = 36$), *B. napus* (AACC, $2n = 38$) y *B. carinata* (BBCC, $2n = 34$), eran anfidiplóides y que éstas derivaban de las especies diploides *B. nigra* (L.) Koch (BB, $2n = 16$), *B. rapa* (sin. *B. campestris*) (AA, $2n = 20$) y *B. oleracea* L. (CC, $2n = 18$).

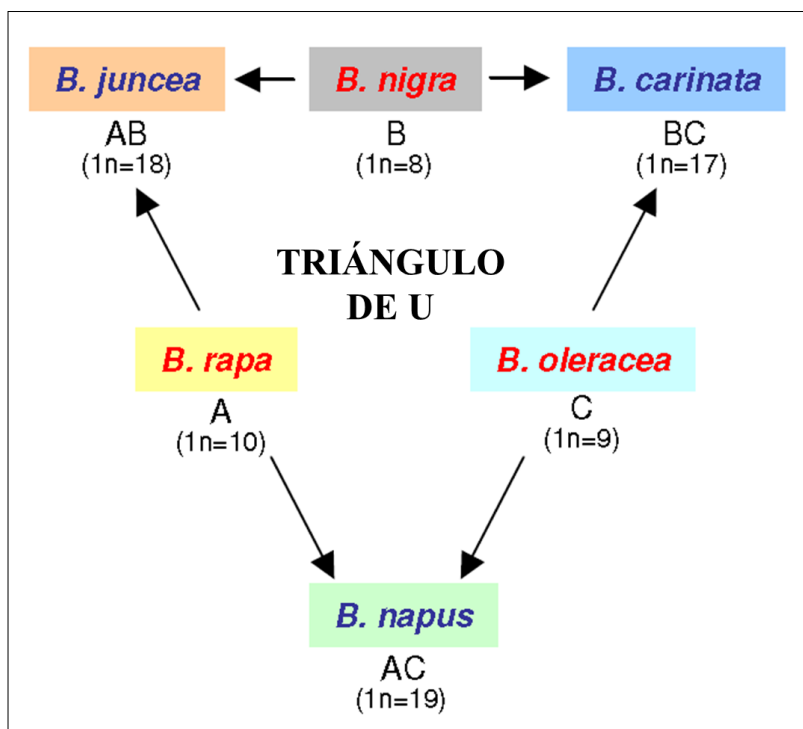


Figura 1. Relaciones entre las especies cultivadas del género *Brassica*, conocido como el Triángulo de U (U, 1935). Los genomas diploides de *B. rapa*, *B. nigra* y *B. oleracea* fueron designados como A, B y C respectivamente.

A lo largo de la evolución de *Brassica*, este género ha sufrido cambios genómicos estructurales que incluyen duplicación cromosómica, rearrreglos y fusiones (Lagercrantz, 1998). El primer evento dentro de la evolución corresponde a la duplicación del genoma ancestral, el cual comparte con *Arabidopsis thaliana*, lo cual sucedió hace aproximadamente 65-90 millones de años (Lynch y Conery, 2000; Simillion *et al.*, 2002; Raes *et al.*, 2003). Posterior a este evento, *Brassica* y *Arabidopsis* divergieron hace aproximadamente 20 millones de años (Koch *et al.*, 2000). Tras esto, *Brassica* sufrió un evento de triplicación del genoma, hace aproximadamente 14-24 millones de años (Parkin *et al.*, 2005). Posteriormente, se produjeron los eventos de especiación (Koch *et al.*, 2001) (Figura 2).

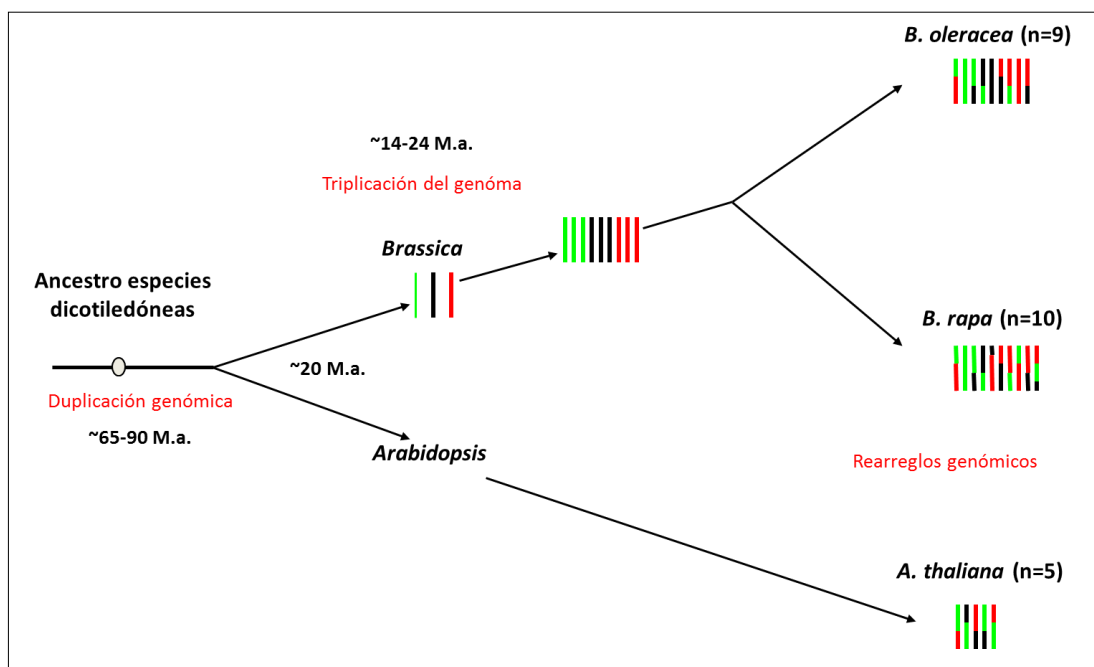


Figura 2. Relación evolutiva entre *Arabidopsis thaliana* y *Brassica*.

Lagercrantz y Lydiat (1996) propusieron que el genoma de *Brassica* se encontraba altamente triplicado, cuya hipótesis ha sido respaldada por estudios posteriores (Lagercrantz, 1998; O'Neill y Bancroft, 2000). En estudios comparativos entre *A. thaliana* y *B. oleracea* se encontraron 15 segmentos dentro del genoma de *A. thaliana* similares pero no necesariamente colineares a más de una región en el genoma de *B. oleracea*. De estas 15 regiones, 5 están presentes en dos copias, 8 están presentes en tres copias y 2 están presentes en más de tres copias (Lukens *et al.*, 2003).

2.2 *Brassica oleracea*

B. oleracea (CC, $2n = 18$) incluye variados morfotipos que han sido seleccionados por meristemas vegetativos alargados en su ápice (repollo, *B. oleracea* subespecie *capitata*) o en hojas axilares (coles de Bruselas, subespecie *gemmifera*), formas con proliferación de meristemas florales (brócoli, subespecie *italica*) o incluso meristemas florales abortados (coliflor, subespecie *botrytis*) y formas con tallos bulbosos prominentes (kohlraabi, subespecie *gongylodes*) o con patrones de hojas rizadas (col rizada, subespecie *acephala*) (Figura 3). Estos genotipos morfológicamente diferentes pueden ser cruzados libremente, lo cual hace a esta especie un excelente modelo para el estudio de la evolución morfológica en plantas (Wang *et al.*, 2011).



Figura 3. Morfotipos de las subespecies de *Brassica oleracea*.

B. oleracea presenta inflorescencia racimosa y en algunos casos con forma de panícula. En el caso de los coliflores, la inflorescencia es abortiva (Pearson, 1933). Las flores de *B. oleracea* L. presentan una coloración amarilla y en algunas ocasiones blancas (Kakizaki, 1930). También existe una subespecie de *B. oleracea* (*botrytis*) que presenta coloración morada, debido a la acumulación de antocianinas (Chiu *et al.*, 2010). La estructura floral consta de 6 estambres, 4 pétalos, 4 sépalos y un ovario (Figura 4).

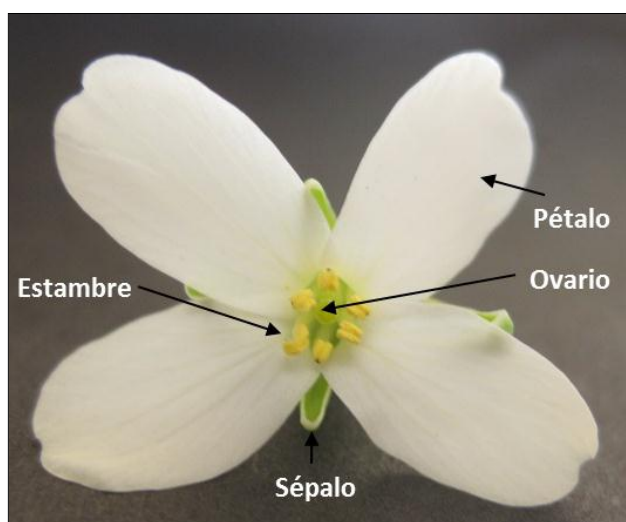


Figura 4. Partes de la flor de *Brassica oleracea*. Pétalo, estambre, ovario y sépalo se encuentran señaladas con flechas.

Estudios iniciales determinaron que el fenotipo flor blanca es dominante sobre el fenotipo flor amarilla en *B. oleracea* L. (Kakizaki, 1930) y que estaría condicionado por un único factor dominante, que dada su poca ocurrencia podría estar ligado a una característica comercial no deseada (Pearson, 1933).

Utilizando una población de mapeo de *B. oleracea* BolTBDH que segrega para el color de flor (FC *flower color*) y que fue originada a partir de una F1 individual proveniente de la cruce entre una línea doble haploide (DH) de brócoli de flor amarilla (“*Early Big*”) y una línea DH de ciclo rápido (TO1000DH3) flor blanca, se pudo establecer que FC mapea en el cromosoma 3 (C3) de *B. oleracea* (Iniguez-Luy *et al.*, 2009) (Figura 5).

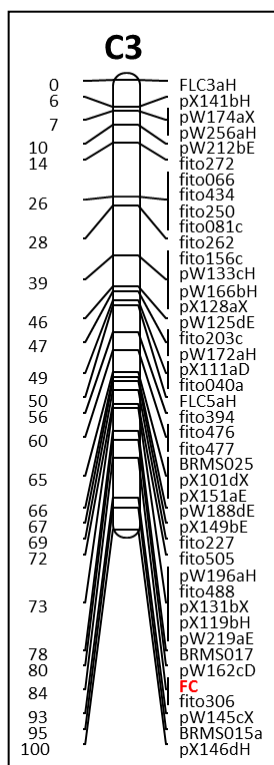


Figura 5. Grupo de ligamiento C3 de *B. oleracea*. La ubicación para el color de flor (FC, *flower color*) se encuentra destacada en rojo.

2.3 Carotenoides

Los carotenoides son compuestos isoprenoides de 40 carbonos que otorgan características de color que van desde el amarillo al rojo y que son sintetizados por todos los organismos fotosintéticamente activos y también por algunos organismos no fotosintéticos como bacterias y hongos (Ahrazem *et al.*, 2010). En la naturaleza, existen cientos de carotenoides, que basados en su estructura, pueden ser clasificados en dos grupos: (i) carotenos (hidrocarburos que pueden ser ciclados en uno o ambos extremos de la molécula) y (ii) xantófilas (derivados oxigenados de carotenos) (Ruíz-Sola y Rodríguez-Concepción, 2012). Existe otro sistema de clasificación basado en la función biológica que cumplen, lo que depende últimamente de la estructura del carotenoide (Namitha y Negi, 2010) (Tabla 1).

Tabla 1. Clasificación de los carotenoides (Delgado-Vargas y Paredes-López, 2003; Schwartz *et al.*, 2008):

Base	Subgrupo	Características
Estructura	Carotenos	Constituidos por carbono e hidrógeno
	Xantófilas	Constituidas por carbono, hidrógeno y oxígeno
Ciclación	Acíclicos	
	a. monocíclicos	Uno de los grupos abiertos y el otro cerrado
	b. bicíclicos	Ambos grupos cerrados
Alteración estructural	Alénico	Dobles enlaces continuos
	Acetilénico	Presencia de un triple enlace
	Apocarotenoide	Menos de 40 átomos de carbono
	Carotenoides mayores	Más de 40 átomos de carbono
Funciones	Primarios	Necesarios para procesos fotosintéticos
	Secundarios	No se relaciona directamente con la sobrevivencia de la planta

Comercialmente, los carotenoides son usados como colorantes para alimentos y suplementos alimenticios en humanos, aditivos en la alimentación para aumentar la pigmentación de peces y huevos, productos farmacéuticos, así como también son utilizados en la industria cosmética y agricultura (Bramley, 2003).

En plantas, las xantófilas, formas oxigenadas de carotenoides, cumplen la función de pigmentos accesorios en la antena recolectora de luz localizada en los cloroplastos, donde tienen la capacidad de transferir energía a la clorofila y también juegan un papel importante en la prevención de daño fotooxidativo (Howitt y Pogson, 2006).

Los carotenoides pueden ser acumulados en todos los tejidos de la planta y estar contenidos en todos los tipos de plástidos, exceptuando los proplastidios (progenitor de los plástidos), los que incluyen cloroplastos (organelos fotosintéticos), cromoplastos (plástidos con color), amiloplastos (plástidos de almacenamiento de almidón), elaioplastos (plástidos de almacenamiento de lípidos), leucoplastos (plástidos sin color) y etioplastos (precursores del crecimiento en oscuridad de los cloroplastos) (Howitt y Pogson, 2006).

En plantas superiores la cantidad y tipo de carotenoides presentes en las membranas fotosintéticas es similar entre las especies, conteniendo principalmente luteína, β -caroteno, violaxantina, neoxantina y zeaxantina (Demming-Adams *et al.*, 1996). Por otra parte, en tejidos no fotosintéticos de la planta (flores, frutos, raíces, tubérculos y semillas, entre otros), el tipo y contenido de carotenoides pueden variar dependiendo de la especie o tejido. Por ejemplo, en pétalos blancos de crisantemo la concentración de carotenoides es baja o nula (Ohmiya *et al.*, 2006), en pétalos amarillos de girasol las concentraciones de carotenoides son abundantes, presentando principalmente ésteres de luteína (Moehs *et al.*, 2001) y en papas dulces el principal carotenoide es β -caroteno (Ben-Amotz y Fishier, 1998).

En flores, los carotenoides son acumulados en cromoplastos, donde los dos grandes mecanismos descritos relacionados con la regulación de la biosíntesis y acumulación de carotenoides en cromoplastos son: (i) un incremento en la biosíntesis de carotenoides, relacionado con el aumento en la abundancia de transcritos de aquellos genes biosintéticos o reguladores de la biosíntesis de carotenoides, y (ii) un incremento en la acumulación de nuevos carotenoides, lo cual ha sido

atribuido a la presencia de estructuras secuestradoras capaces de almacenar carotenoides dentro del plástido (Howitt y Pogson, 2006).

Por otra parte, la regulación del contenido de carotenoides se ha relacionado con cambios en la expresión de genes que codifican para enzimas que catalizan oxidativamente carotenoides a compuestos más pequeños denominados apocarotenoides o norisoprenoides, con características de isoprenoides no-polares que poseen un amplio conjugado de doble enlaces (Kloer y Schulz, 2006), los cuales han sido ampliamente estudiados por sus propiedades antioxidantes y su capacidad de capturar radicales libres (Escribano *et al.*, 1996). La familia enzimática se denomina CCD (por su nombre en inglés, *Carotenoid Cleavage Dioxygenase*). Estas enzimas participan activamente en procesos hormonales de la planta (Booker *et al.*, 2004; Iuchi *et al.*, 2001; Chernys y Zeevaart, 2000; Schwartz *et al.*, 1997), producción de compuestos volátiles (Simkin *et al.*, 2004a; Simkin *et al.*, 2004b) y regulando el color de fenotipos en diversas especies (Ohmiya *et al.*, 2006; Campbell *et al.*, 2010; Brandi *et al.*, 2011) (Figura 6).

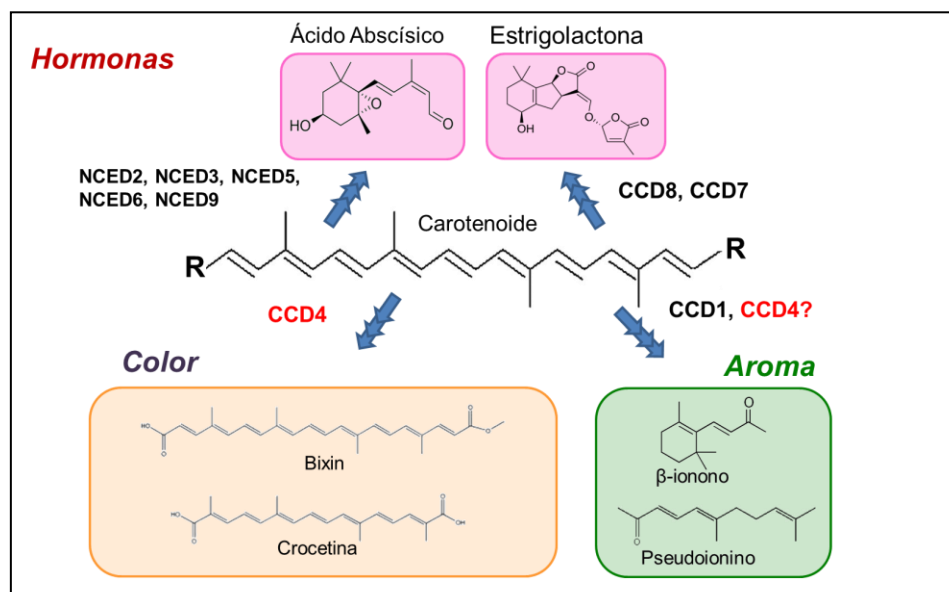


Figura 6. Apocarotenoides y las CCDs o NCEDs responsables de su producción (Ohmiya, 2009). CCD, *carotenoid cleavage dioxygenase*; NCED, *nine cis epoxycarotenoid dioxygenase*.

La única estructura establecida para la familia de las CCDs es la de ACO (por su nombre en inglés, *Apo-carotenoid Cleavage Oxygenase*) de *Synechocystis* PCC 6803 (Kloer *et al.*, 2005). La cadena de plegamiento consiste en siete láminas beta plegadas, cinco de las cuales están constituidas por cuatro láminas beta antiparalelas, mientras que las dos restantes están conformadas por cinco láminas beta antiparalelas (Kloer y Schulz, 2006).

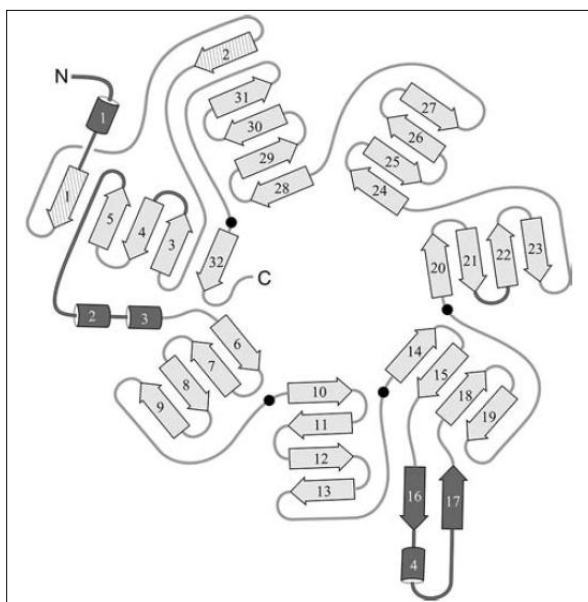


Figura 7. Esquema de plegamiento de ACO. La estructura corresponde a siete láminas beta plegadas como motivo básico. La posición de las cuatro histidinas que coordinan con el hierro al centro de las láminas están marcadas con puntos. Los cuatro bucles y la región N- terminal formando el domo alrededor del centro activo se destacan en gris oscuro, y la extensión N-terminal se encuentra ennegrecida.

2.4 Familia de las CCDs

La familia de las CCDs es antigua, con miembros presentes en bacterias, plantas y animales. En plantas, las enzimas CCDs se encuentran codificadas por familias multigénicas y estudios recientes han comenzado a otorgarle funciones a miembros individuales de cada familia (Auldrige *et al.*, 2006).

La primera enzima CCD de plantas en ser caracterizada fue la *viviparous14* (VP14) aislada desde el maíz mutante *viviparous 14*, el cual presentaba deficiencia en la biosíntesis de ácido abscísico (Schwartz *et al.*, 1997). Análisis enzimáticos de un recombinante VP14 confirmó que esta enzima corta 9-*cis*-epoxicarotenoides para formar C₂₅ apo-aldehídos y xantoxina, un precursor del ácido abscísico en maíz y plantas superiores (Tan *et al.*, 1997; Schwartz *et al.*, 1997).

A partir de análisis basados en homología de secuencia utilizando VP14, se han podido identificar varias enzimas responsables de la síntesis de apocarotenoides de importancia biológica. En *Arabidopsis* se han descrito 9 genes pertenecientes a la familia CCD (Tabla 2), y por lo general, los ortólogos en otras especies son nombrados de acuerdo a su homología con éstos.

Tabla 2. Dioxigenasas de corte de carotenoides en *Arabidopsis thaliana* (Tan *et al.*, 2003).

ND: no determinada.

Nombre	Intrones	Locus	Actividad
AtNCED2	No	At4g18350	11, 12 9- <i>cis</i>
AtNCED3	No	At3g14440	11, 12 9- <i>cis</i>
AtNCED5	No	At1g30100	11, 12 9- <i>cis</i>
AtNCED6	No	At3g24220	11, 12 9- <i>cis</i>
AtNCED9	No	At1g78390	11,12 9- <i>cis</i>
AtCCD1	Si	At3g63520	9,10 (9'-10')
AtCCD4	No	At4g19170	ND
AtCCD7	Si	At2g44990	ND
AtCCD8	Si	At4g32810	ND

Miembros de la familia de las CCDs comparten varias características: (i) requieren Fe^{2+} para su actividad catalítica, (ii) contienen cuatro residuos de histidina altamente conservados, los cuales se cree coordinan el ligando de hierro, (iii) contienen una secuencia peptídica conservada en su extremo carboxil que minoritariamente constituye una secuencia característica de la familia (Auldrige *et al.*, 2006).

Existen un grupo especializado de CCD denominadas 9-*cis*-epoxicarotenoide dioxigenasa (NCED, por su nombre en inglés: *Nine Cis Epoxycarotenoid Dioxygenase*), las cuales se encuentran principalmente implicadas en la biosíntesis de ácido abscísico (Schwartz *et al.*, 1997; Tan *et al.*, 2003; Sun *et al.*, 2012; Chernys y Zeevaart, 2000) como se ilustra en la Figura 6.

Existen fenotipos blancos en flores, frutos y tubérculos que han sido relacionados específicamente con la expresión diferencial de genes CCD4. El primer ejemplo de esto es el gen *CmCCD4a*, el cual se expresa preferencialmente en pétalos blancos de crisantemos, contribuyendo a la formación del color blanco en pétalos mediante el corte enzimático de

carotenoides produciendo compuestos incoloros. Mediante ARN de interferencia (ARNi) del gen *CmCCD4a*, se obtuvo un cambio de color en pétalos desde blanco a amarillo, demostrando que la acumulación de carotenoides en crisantemos solo ocurría en ausencia de la actividad de CCD4 (Ohmiya *et al.*, 2006).

En durazno, se evidenció el efecto del gen CCD4 sobre la regulación del contenido de carotenoides en frutos. Los niveles de expresión del gen CCD4 fueron mayores en el cultivar de pulpa blanca, aumentando los transcritos a medida que el fruto iba madurando (Brandi *et al.*, 2011).

En tubérculos de papa (*Solanum tuberosum*), también se ha investigado el efecto de la actividad del gen CCD4 en el contenido de carotenoides. Se observó una elevada expresión del gen CCD4 en tubérculos maduros de cultivares de pulpa blanca comparados con tubérculos ricos en carotenoides de pulpa amarilla. Junto con el análisis de transcritos, se realizaron líneas ARNi para el gen CCD4 en papas de pulpa blanca, observándose que tras el silenciamiento del gen CCD4, los niveles de carotenoides aumentaron en las flores y tubérculos de la planta (Campbell *et al.*, 2010).

3. JUSTIFICACIÓN DEL ESTUDIO

En plantas superiores, las flores constituyen órganos vitales para la reproducción. En *Brassica*, el fenotipo flor amarilla ha sido seleccionado para el cultivo, dado que su color atrae a los agentes polinizadores facilitando así la polinización de las flores y, por ende, aumenta la obtención de semillas. En *Brassica oleracea*, existe un fenotipo flor blanca que de manera silvestre se encuentra en baja frecuencia y cuya caracterización a nivel molecular no ha sido realizada. Si bien *B. oleracea* es un organismo diploide, a lo largo de su evolución ha sufrido duplicaciones y triplicaciones genómicas que afectan el número de copias de un determinado gen ancestral. El gen analizado en este estudio se presenta en dos copias parálogas dentro del genoma C de *B. oleracea*. El análisis de genes parálogos como CCD4 y los cambios presentes a nivel de secuencia, ayudan a entender el proceso evolutivo en plantas, ya que la existencia de genes duplicados permite que cada parólogo adquiriera diferentes patrones de expresión en distintos tejidos de la planta, muchas veces otorgando plasticidad de respuesta a estímulos y variabilidad fenotípica. Al mismo tiempo, el estudio de fenotipos con eventos de poliploidía que datan desde hace millones de años, son útiles para dilucidar los efectos a largo plazo de la expresión de genes duplicados.

3.1 Hipótesis

El fenotipo flor blanca en *Brassica oleracea* L. TO1000DH3 está determinado por cambios en los patrones de expresión de un gen que cataboliza carotenoides

3.2 Objetivo general

Determinar la base genética detrás del fenotipo flor blanca en la línea doble haploide *Brassica oleracea* TO1000DH3

3.3 Objetivos específicos

1. Mapear genes CCD4 en la población de mapeo de *B. oleracea* BolTBDH utilizando DNA-SSCP.
2. Aislar los genes CCD4 de *B. oleracea* TO1000DH3 (flor blanca) y “*Early Big*” (flor amarilla), secuenciar, analizar y comparar las secuencias nucleotídicas y aminoacídicas correspondientes.
3. Evaluar la expresión génica de *BolC.CCD4.a* y *BolC.CCD4.b* mediante qRT-PCR en 3 estadíos del desarrollo de flores blancas y amarillas de individuos de la población de mapeo BolTBDH y sus respectivos genotipos parentales.

4. METODOLOGÍA

4.1 Material vegetal y extracción de ADN.

Se utilizó la población de mapeo BolTBDH que segrega para el color de flor (FC, *flower color*) la cual fue originada a partir de una F1 individual proveniente de la cruce entre una línea doble haploide (DH) de brócoli de flor amarilla (“*Early Big*”) y una línea DH de ciclo rápido (TO1000DH3) flor blanca (Iniguez-Luy *et al.*, 2009), la cual fue crecida en invernadero con un ciclo día 16/8 horas luz/oscuridad. La extracción de ADN genómico (ADNg) para el mapeo del gen CCD4 fue realizada mediante el método PEX (Jhingan, 1992).

Para el ensayo de expresión se crecieron tres réplicas biológicas de tres individuos dobles haploides (DH) para cada fenotipo (flor blanca y flor amarilla) pertenecientes a la población de mapeo BolTBDH (Iniguez-Luy *et al.*, 2009) junto con las líneas parentales.

Las plantas fueron crecidas en condiciones controladas con un ciclo día 16/8 horas luz/oscuridad, con una temperatura constante de 22 °C.

4.2 Diseño de oligonucleótidos y clonamiento de gen CCD4.

Utilizando la secuencia genómica de *AtCCD4* (At4g19170) se realizó una búsqueda en Blastn, utilizando la base de datos de ESTs (*Expressed sequence tags*) y el organismo *Brassica oleracea* disponibles en NCBI. A partir de esto se identificaron 29 secuencias de ESTs en *B. oleracea* var. *acephala* y *B. oleracea* var. *alboglabra*, provenientes de los siguientes tejidos: hoja, plántula y flor. Estas secuencias fueron utilizadas para diseñar oligonucleótidos para clonar la región codificante del gen (CDS) y para el mapeo de genes CCD4 en *B. oleracea*.

ADN genómico (ADNg) y ADN complementario (cDNA) fueron los templados utilizados para el aislamiento de las regiones codificantes de genes CCD4, utilizándose la enzima KOD Hot Start DNA Polymerase (Novagen, EEUU).

Los oligonucleótidos utilizados en esta tesis fueron diseñados y luego analizados utilizando el server OligoAnalyzer 3.1 (<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>) (Tabla 3).

Tabla 3. Oligonucleótidos. Se muestran las secuencias de todos los oligonucleótidos empleados para mapeo por *DNA-SSCP*, clonamiento de genes *CCD4* y estudios de expresión.

Nombre	Secuencia (5'-3')
Mapeo DNA-SSCP	
Bo.CCD4.1F	CTCCGTATCAACTCCGCCGT
Bo.CCD4.1R	GGAAGGACGTGACGGTGGGT
Clonamiento genes CCD4	
BoCCD4.cdsF	ATGGACTCTCTTTCTTCCTCTTCCT
BoCCD4.cdsR	TTAAAGCTTATTAAGGTCACCTCTCC
Bo.CCD4.a cds1 F	CCCACACTCACACACACTCAT
Bo.CCD4.a cds1 R	CGAACGTCTCTCCGGTTTCG
Bo.CCD4.a cds2 F	GGTTTAGCGAACACGAGCCTC
Bo.CCD4.a cds2 R	GGCAACCTCACTGCGGCTA
Bo.CCD4.a 3prima F	TACGGTCCAGGCTGTTACGGA
Bo.CCD4.a 3prima R	ATCGCCGTCGCCAGAGAAAG
Bo.CCD4.b 5prima F	ACTTACGCAAATATTACATGAGCAA
Bo.CCD4.b 5prima R	ATCGGAGATCGTTCTTCGGA
Bo.CCD4.b cds1 F	CCCACACTGGCACTCATCTTC
Bo.CCD4.b cds1 R	CGAACCGCCCGATCGTTTCA
Bo.CCD4.b cds2 F	GTTTTCCGGATTTAACGGCATGA
Bo.CCD4.b cds2 R	CAGACCGTGATTAAACCTTTGT
Bo.CCD4.b 3prima F	GATGACGGTTACGTGGTCACA
Bo.CCD4.b 3prima R	AAGCTTCTCCCATGCTTCTTTG
Ensayo de expresión	
Bo.CCD4.a F1qpcr	ATTCGACCGGGAAGAAGCGAAA
Bo.CCD4.a R1qpcr	GCATCGAATCCATCCTCATCCGA
Bo.CCD4.a F2qpcr	CGTAGATCCAAAACACGTCCTCTCAGAC
Bo.CCD4.a R2qpcr	GTTACGGATGTAAGCGCCGTTG
Bo.CCD4.b F1qpcr	GGTTTGATTCAACCGGAAAGAAACA
Bo.CCD4.b R1qpcr	CCCTCGAGCACCAGATCCATTATATT
Bo.CCD4.b F2qpcr	CGTAGATCCAAAACACGTTCTCTCAGGT
Bo.CCD4.b R2qpcr	GTTACGGATGTAAGCGCCGTCA
Bol-Ubiquitin F	CGGATCAGCAGAGGTTGATCTT
Bol-Ubiquitin R	CCTGCAGTTGACAGCCCTTGG

4.3 Mapeo genético utilizando polimorfismo de conformación de cadena simple (DNA-SSCP).

Utilizando el set de oligonucleótidos BoCCD41F-1R, se realizó el mapeo de los genes CCD4 con 72 líneas de la población de mapeo de *B. oleracea* BolTBDAH (Iniguez-Luy *et al.*, 2009). Las reacciones de PCR fueron realizadas en un volumen total de 25 μ L, utilizando la enzima ADN polimerasa GoTaq Flexi (Promega, EEUU). La amplificación comenzó con una desnaturalización inicial a 95°C por 5 minutos, seguida de 30 ciclos de amplificación de 1 minuto a 95°C, 30 segundos a 55°C y 1 minuto a 72°C, con una extensión final a 72°C por 7 minutos. El producto de la amplificación a partir de ADNg de los parentales de la población de mapeo fue purificado luego de ser corrido en un gel de agarosa al 1%, para luego ser diluido en 15 μ L. Para los productos purificados, 5 μ L fueron mezclados con 15 μ L de buffer de carga para SSCP (95% formamida, 10 mm NaOH, 0.25% xilencianol, 0,25% azul de bromofenol), mientras que para los productos de la amplificación directa, 3 μ L fueron mezclados con 15 μ L de buffer de carga para SSCP. Posteriormente, las muestras fueron desnaturalizadas a 96 ± 2 °C por 10 minutos e inmediatamente enfriadas en agua/hielo. De éstos, 3 μ L de cada muestra fueron cargados en un gel de aumento de la detección de mutación (*Mutation detection enhancement, MDE*) 0.7 X (Lonza, EEUU). La electroforesis se llevó a cabo durante 16 horas y 35 minutos a 7 watts a temperatura ambiente.

La tinción de las muestras fue realizada inicialmente fijando el gel con ácido acético al 10% por 30 minutos, un lavado posterior con agua destilada por 20 minutos. El gel fue luego sumergido en una solución de nitrato de plata (1g/L) y formaldehído (0,06%), posteriormente lavado con agua destilada y revelado finalmente en una solución fría de carbonato de sodio (30g/L), formaldehído (0,06%) y tiosulfato de sodio (1 mg/L). El revelado del gel fue detenido con una solución de ácido acético 10%.

El análisis de ligamiento y la construcción del mapa fue realizado utilizando JoinMap® v. 4.0 (Van Ooijen y Voorrips, 2001). Los loci ligados fueron agrupados usando un LOD de 4.0. Las distancias en centiMorgans (cM) fueron calculadas usando la función de mapeo Kosambi.

4.4 Análisis bioinformático de secuencias genómicas CCD4 en *Brassica oleracea* y predicción de la estructura proteica.

Los alineamientos de nucleótidos y aminoácidos fueron realizados usando ClustalW2 (Chenna *et al.*, 2003). La predicción de la estructura de los genes fue realizada en FGENESH (<http://linux1.softberry.com/berry.phtml?topic=fgenes&group=programs&subgroup=gfind>) utilizando “Plantas Dicotiledoneas (*Arabidopsis*)” como organismo de referencia. La presencia de péptido señal transitorio fue predicha utilizando el servidor ChloroP 1.1 (Emanuelsson *et al.*, 1999). La búsqueda de secuencias repetitivas se realizó utilizando el programa RepeatMasker Web Server (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>) y Plant Repeat Databases (<http://plantrepeats.plantbiology.msu.edu/search.html>). Los modelos proteicos fueron realizados utilizando el software SWISS-MODEL (Arnold *et al.*, 2006) en modo automático. El cálculo del peso molecular de cada proteína se realizó en el software Compute pI/MW Tool (http://web.expasy.org/compute_pi/).

4.5 Extracción de ARN y síntesis de cDNA

La extracción de ARN total se realizó a los siguientes tejidos vegetativos y reproductivos: hoja madura #11 y 3 estadíos de desarrollo en pétalos (P1: brote floral verde, P2: brote floral amarillo, P3: pétalos expandidos). Los 3 estadíos de pétalos fueron recolectados el mismo día en paralelo. Todas las muestras fueron recolectadas entre las 10 AM y 12 PM del día. El ARN total fue extraído utilizando Absolutely RNA RT-PCR Miniprep Kit (Agilent Technologies, EEUU) de acuerdo a las indicaciones del fabricante. La cuantificación de los ácidos nucleicos se realizó utilizando el espectrofotómetro ScanDrop nano-volumen (Analytik Jena, Alemania). A partir de 1 µg de ARN total se sintetizó el cDNA utilizando SuperScript® VILO cDNA Synthesis Kit (Invitrogen, EEUU) de acuerdo a las indicaciones del fabricante (volumen final de reacción 20 µL). A partir de los 20 µL de cDNA sintetizados, se realizó una dilución 1:10, la cual fue utilizada como templado para el ensayo de expresión.

4.6 PCR cuantitativa en tiempo real (qRT-PCR).

La eficiencia de los oligonucleótidos utilizados en el experimento fue determinada mediante la amplificación de diluciones seriadas (10^{-9} - 10^{-1}) de plásmidos que contenían las secuencias a analizar. La determinación de la pendiente de la recta y la eficiencia de los oligonucleótidos se realizó utilizando el software MxPRO-Mx3000p versión 4.10 (Stratagene, EEUU).

Un microlitro de cDNA diluido (1:10) fue agregado bajo las siguientes condiciones de reacción: 10 μ L de Brilliant II SYBR® Green QPCR Master Mix (Agilent Technologies, EEUU), 30 nM de el colorante de referencia (ROX), 200nM de cada oligonucleótido y H₂O grado PCR hasta un volumen final de 20 μ L. Estas condiciones de reacción fueron utilizadas para los tres genes evaluados en el experimento (*BolC.CCD4.a*, *BolC.CCD4.b* y *ubiquitina*). Cada templado fue amplificado con dos réplicas técnicas.

El qRT-PCR fue realizado en el equipo Mx3000p (Stratagene, EEUU), donde la amplificación comenzó con una desnaturalización inicial de 10 minutos a 95°C, seguido de 40 ciclos de amplificación a 95°C por 20 segundos, 60°C (*BolC.CCD4.a* y *BolC.CCD4.b*) ó 56°C (*ubiquitina*) por 20 segundos y 72°C por 20 segundos, tomando la medición de fluorescencia al final de cada ciclo. Finalizados los 40 ciclos, se realizó la curva de disociación bajo las siguientes condiciones 10 segundos a 95°C, 1 segundo a 55°C y 1 segundo a 95°C, las mediciones de la fluorescencia fueron tomadas en cada punto desde que la temperatura estaba en 55°C, hasta que alcanzó los 95°C.

El análisis de la expresión relativa de los genes *BolC.CCD4* fue llevado a cabo utilizando el software qbasePLUS versión 2.4 (Helleman *et al.*, 2007) mediante el método $\Delta\Delta C_t$ (Figura 8). Todas las muestras fueron normalizadas utilizando el gen *ubiquitina*.

$$\Delta\Delta C_t = \frac{\text{Eficiencia}_{\text{gen de interés}}^{\Delta C_t_{\text{gen de interés (control-muestra)}}}}{\text{Eficiencia}_{\text{gen de referencia}}^{\Delta C_t_{\text{gen de referencia (control-muestra)}}}}$$

Figura 8. Método $\Delta\Delta C_t$. Fórmula del método utilizado para el cálculo de las expresiones relativas de los genes CCD4 utilizando el gen *ubiquitina* como referencia.

5. RESULTADOS

5.1 Identificación del gen candidato causante del fenotipo flor blanca (*wf*) en *B. oleracea*.

Mediante genómica comparativa con *Arabidopsis thaliana*, se determinó que el gen candidato CCD4 podría encontrarse en regiones genómicas colineares al intervalo donde mapea el fenotipo flor blanca (*wf*) en el cromosoma C3 de *B. oleracea* (Figura 9).

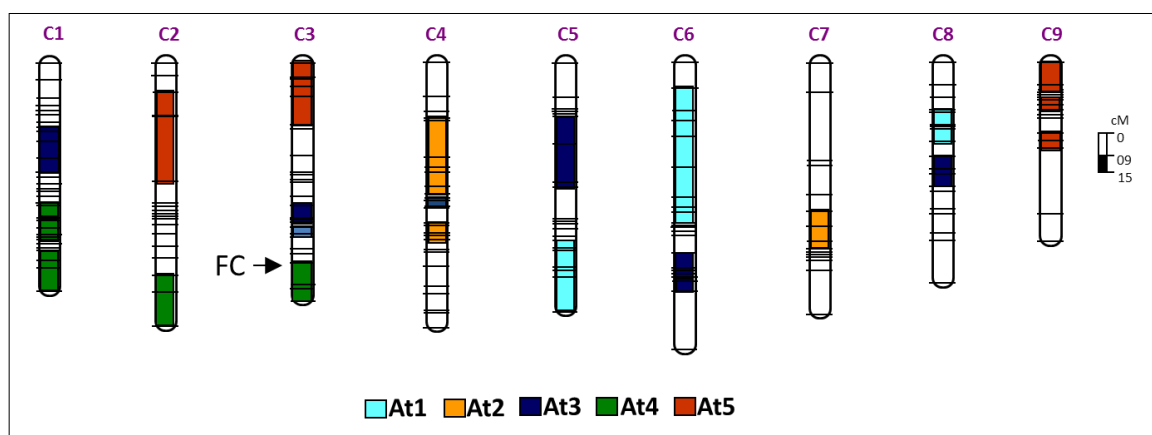


Figura 9. Mapa comparativo entre *Arabidopsis thaliana* y *Brassica oleracea*. Cada color representa a un cromosoma de *A. thaliana* localizado en el genoma de *B. oleracea*. La posición donde mapea el marcador fenotípico para el color de flor (FC, *flower color*) en el cromosoma C3 de *B. oleracea* se encuentra indicado con una flecha.

5.2 Identificación de las secuencias *BolC.CCD4* en el genoma *C* de *Brassica*.

El gen CCD4 se encuentra como copia única en el genoma de *Arabidopsis thaliana*, cuya secuencia genómica no posee intrones, codificando para una proteína de 595 aminoácidos. Utilizando la secuencia genómica de *AtCCD4* (Locus: At4g19170), se realizó un Blastn, en la base de datos EST (*Expressed Sequence Tags*), seleccionando *Brassica oleracea* como organismo. A partir de esto, se identificaron 29 secuencias ESTs con una alta identidad de secuencia (>80%) con el gen *AtCCD4*. Una vez identificadas las secuencias, se realizó un alineamiento, identificando regiones génicas conservadas para el diseño de oligonucleótidos (Tabla 3) tanto para capturar la secuencia codificante (CDS) como para mapear genes CCD4 en la población BolTBDH de *B. oleracea* (Figura 10).

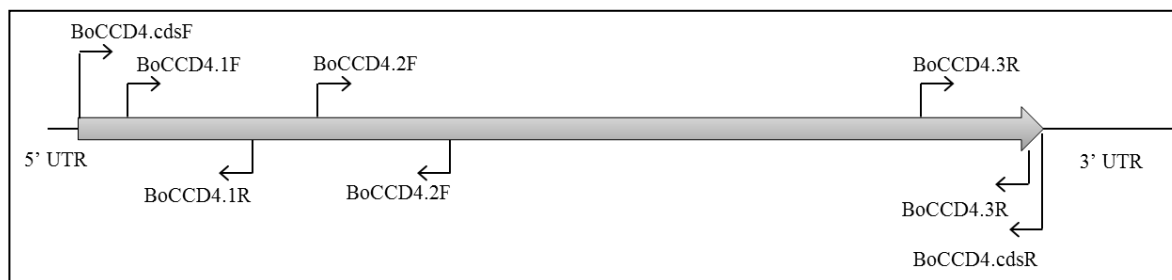


Figura 10. Posición relativa de los oligonucleótidos diseñados para el mapeo y aislamiento de secuencias CCD4 en *B. oleracea*. Los oligonucleótidos indicados son los que fueron utilizados para el mapeo de los genes CCD4 en la población BolTBDH (BoCCD4.1F-1R, BoCCD4.2F-2R, BoCCD4.3F-3R) y los utilizados para el aislamiento de la secuencia codificante (BoCCD4.cdsF-R)

El set de oligonucleótidos realizado para el mapeo de genes CCD4 fue utilizado inicialmente en los parentales de la población de mapeo BolTBDH (TO1000DH3 y “*Early Big*”) para analizar si existían diferencias entre los patrones de migración en geles MDE. Los productos purificados de ambas amplificaciones fueron corridos en un gel MDE 0.7 X (Figura 11). Se observó que los productos amplificados utilizando el set de oligonucleótidos BoCCD4.1F-1R presentó una diferencia en el patrón de migración, lo cual permitiría posteriormente el mapeo de los genes CCD4.

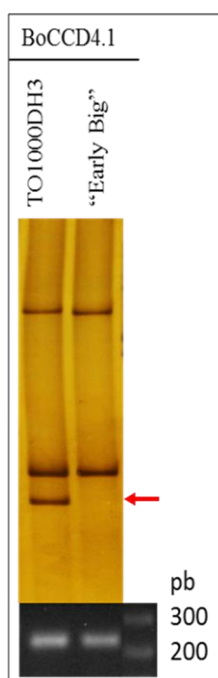


Figura 11. Patrones de migración de los parentales de la población BolTBDH en gel MDE. El set de oligonucleótidos BoCCD4.1F-1R presenta un patrón de migración diferencial entre los parentales de la población, presentándose una banda extra en el amplicón proveniente de ADNg de TO1000DH3 (flecha).

A partir de estos resultados, se procedió a la amplificación mediante PCR del ADNg de 72 individuos de la población BolTBDH utilizando el set de oligonucleótidos BoCCD4.1F-1R.

Una vez obtenidos los patrones de migración de los individuos en el gel MDE, se observó claramente que existían dos patrones independientes de migración, uno de los cuales co-segregó con el fenotipo flor blanca/amarilla y uno que segregó independientemente del fenotipo que presentaba la flor (Figura 12). Esto demostró la existencia de dos genes CCD4 presentes en el genoma C de *Brassica*. El nombre asignado a cada gen se realizó según la nomenclatura establecida para nombrar los genes en *Brassica* spp. (Ostergaard y King, 2008), la cual sugiere la siguiente sintaxis (< > indica las categorías): <GÉNERO 1letra> [<especie 2 letras>] <GENOMA 1 LETRA>|<X>.<NOMBRE 3-6 LETRAS>.<locus 1 letra>.

El gen que segregó independiente del fenotipo de la flor se denominó *BolC.CCD4.a* y mapeó en el cromosoma C1 de *B. oleracea*. El gen que co-segregó con el fenotipo flor blanca/amarilla fue denominado *BolC.CCD4.b* y mapeó en el cromosoma C3 de *B. oleracea* junto con el marcador fenotípico previamente identificado para el color de flor FC (*flower color*) (Figura 13).

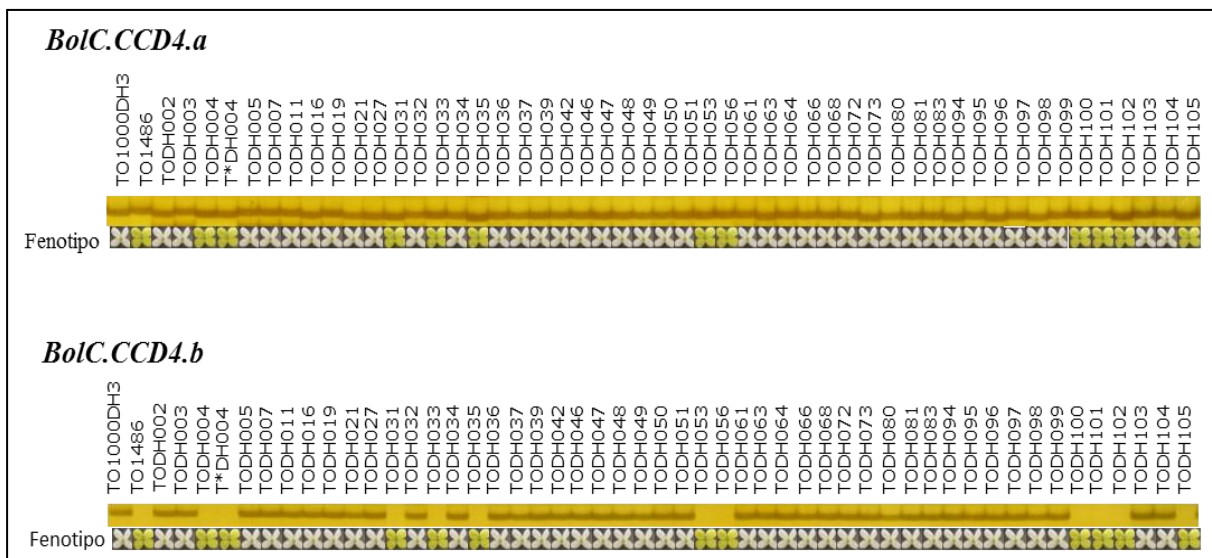


Figura 12. Patrón de segregación de *BolC.CCD4a* y *BolC.CCD4.b* en una población de mapeo bi-parental desarrollada a partir de la cruce entre una línea flor amarilla (“*Early Big*”) y una línea flor blanca (TO1000DH3). Los patrones de amplificación corresponden a una muestra de 50 de los 72 individuos de la población evaluados, utilizando los oligonucleótidos BoCCD4.1F-1R.

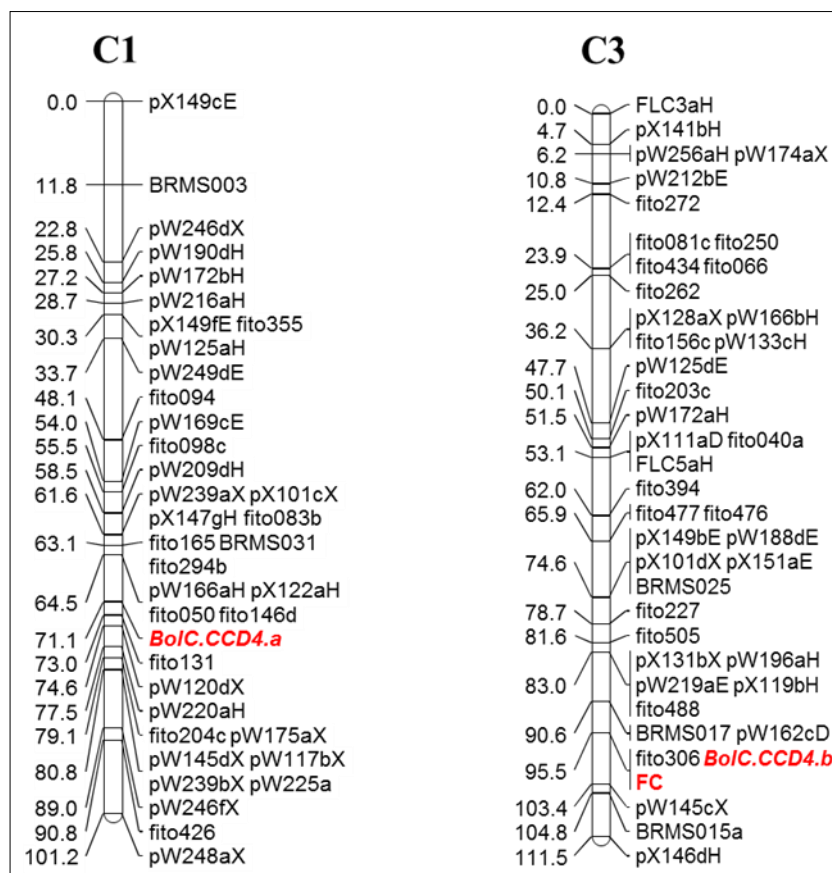


Figura 13. Localización de *BolC.CCD4.a* y *BolC.CCD4.b* en el mapa del genoma C de *Brassica*. El nombre de cada locus y las distancias (cM) aparecen ubicados a la izquierda y derecha de la barra vertical, respectivamente. Los dos marcadores moleculares basados en secuencias de los genes CCD4 y el marcador fenotípico para el color de flor (FC) se encuentran destacados en rojo.

5.3 Análisis de secuencia

A partir de ADNg y cDNA (pétalo estadio 3) de *B. oleracea* TO1000DH3 y ADNg de *B. oleracea* “*Early Big*” utilizando los oligonucleótidos diseñados para capturar la secuencia codificante (CDS) (BoCCD4cds1F-1R), se procedió a aislar y enviar a secuenciar 20 clones. En *B. oleracea* TO1000DH3 se obtuvo el CDS para ambos genes CCD4. En cambio, para *B. oleracea* “*Early Big*” sólo se obtuvo la secuencia correspondiente al CDS de *BolC.CCD4.a* utilizando este primer set de oligonucleótidos.

A partir de la combinación de oligonucleótidos Bo.CCD4.b F2qpcr/Bo.CCD4.b R1qpcr se obtuvo la secuencia parcial del gen *BolC.CCD4.b* de “*Early Big*”. Esta secuencia fue posteriormente completada gracias a la colaboración existente con el Agriculture and Agri-Food Canada-Saskatoon Research Centre (AAFC-SRC).

El primer análisis realizado fue comparar las secuencias obtenidas de ADNg y cDNA de *B. oleracea* TO1000DH3, donde se obtuvo 100% de identidad entre las secuencias, lo que nos indicó la ausencia de intrones en las secuencias de los genes *BolC.CCD4.a* y *b*.

El análisis de secuencias codificantes (CDS) reveló que *BolC.CCD4.a* y *BolC.CCD4.b* en TO1000DH3 codifican proteínas de 593 y 596 aminoácidos respectivamente, exhibiendo un 87% de identidad entre ambos parálogos y que, al igual que el gen *AtCCD4*, no poseen intrones. Los dominios de proteínas presentes indican que estos genes pertenecen a la familia de los β -caroteno dioxigenasas y a la subfamilia de 9-*cis*-epoxidioxigenasas. Mediante el análisis de los aminoácidos presentes en la proteína, se estableció que ambas proteínas CCD4 en TO1000DH3 presentan cuatro residuos conservados de histidina (H) necesarios para coordinar la unión a hierro (cofactor requerido para la actividad catalítica de la proteína CCD4). Además, ambas proteínas contienen un péptido señal en su región N-terminal, lo que predice su localización subcelular en los plástidos, al igual que para *AtCCD4* (Figura 14).

Por otra parte, el análisis de las secuencias obtenidas para “*Early Big*” correspondientes al gen *BolC.CCD4.a* que codifica para una proteína de 593 aminoácidos, sin intrones, con dominios

conservados correspondientes a la familia de los β -caroteno dioxigenasas y subfamilia 9-*cis*-epoxicarotenoides. En esta proteína también están presentes los cuatro residuos de histidina, anteriormente mencionados (Figura 14).

La secuencia codificante predicha para el gen *BolC.CCD4.b* de “*Early Big*” codificó para una proteína más corta, la cual consta de 455 residuos aminoácidos, sin la presencia de intrones. En esta proteína sólo se observaron tres de los cuatro residuos de histidina, lo cual presumiblemente podría afectar su actividad catalítica. Al igual que su parálogo, contiene una secuencia peptídica transitoria en su región N-terminal que la dirigiría hacia los cloroplastos (Figura 14).

AtCCD4	MDSVSSSSFLSSTFSLHHSLLRRSSSPTLLRINS	AVVEERSPITNPSDNNDRR-NKPKTLHNRTNHTLVSSPKLRPEM	79
BolC.CCD4.a_Early.Big	MDSLSSSSFISTFSPASSLLLPSPSSRRLLRINS	AVVEERSPITNPSQNNDLPPSKPKKLYTRTNHTAVSSPAKPRQOT	80
BolC.CCD4.a_To1000DH3	MDSLSSSSFISTFSPASSLLLPSPSSRRLLRINS	AVVEERSPITNPSQNNDLPPSKPKKLYTRTNHTAVSSPAKPRQOT	80
BolC.CCD4.b_Early.Big	-----MSVRIPLLLSCSS---	VPSPKHKKLYSRTNRNSVASPAKLRPET	40
BolC.CCD4.b_To1000DH3	MYSVSSSSSFIISTFSPKPSLHLHSSSSRLLPRI	NSTVSEERSPISNPSENNVPFSPKLYKLYTRTNRNSVASPAKLRPET	80
AtCCD4	TLATALFTTVEDVINTFIDPPSRPSVDPKHVLS	DNFAPVLDLPPTDCEIIHGTLPLSLNGAYIRNGPNPQFLPRGPYHL	159
BolC.CCD4.a_Early.Big	SLATALFTTMEVINTFIDPPSRPSVDPKHVLS	DNFAPVLDLPPTDCEIIHGSLPPSLNGAYIRNGPNPQFLPRGPYHL	160
BolC.CCD4.a_To1000DH3	SLATALFTTMEVINTFIDPPSRPSVDPKHVLS	DNFAPVLDLPPTDCEIIHGSLPPSLNGAYIRNGPNPQFLPRGPYHL	160
BolC.CCD4.b_Early.Big	TLVTALFTTVEDVINTFIDPPSRPSVDPKHVLS	SGNFAPVLDLPPTDCEIIHGSLPPFLDGAYIRNGPNPQFLPRGPYHL	120
BolC.CCD4.b_To1000DH3	TLVTALFTTVEDVINTFIDPPSRPSVDPKHVLS	SGNFAPVLDLPPTDCEIIHGSLPPSLDGAYIRNGPNPQFLPRGPYHL	160
AtCCD4	FDGDGMLHAIRIRNGKATLC	SRVYKTYKYNVEKQTGPVMPNVFSGFNGVTASVARGALTAARVLTGQYNPNVNGIGLANT	239
BolC.CCD4.a_Early.Big	FDGDGMLHAIRIRNGKATLC	SRVYKTYKYNVEKQTGPVMPNVFSGFNGVTASVARGALTAARVLTGQYNPNVNGIGLANT	240
BolC.CCD4.a_To1000DH3	FDGDGMLHAIRIRNGKATLC	SRVYKTYKYNVEKQTGPVMPNVFSGFNGVTASVARGALTAARVLTGQYNPNVNGIGLANT	240
BolC.CCD4.b_Early.Big	FDGDGMLHAIRIRNGKATLC	SRVYKTYKYNVEKQAGAVIPNVFSGFNGVTASVARGALTAARVLTGQYNPNVNGIGLANT	200
BolC.CCD4.b_To1000DH3	FDGDGMLHAIRIRNGKATLC	SRVYKTYKYNVEKQAGAVIPNVFSGFNGVTASVARGALTAARVLTGQYNPNVNGIGLANT	240
AtCCD4	SLAFFSNRLFALGESDLPYAVRLTESGDIETIG	RYDFDGLKAMSMTAHPKTDPIETGETFAFRYGPVPPFLTFFRFDSAGK	319
BolC.CCD4.a_Early.Big	SLAFFCNRLFALGESDLPYAVRLTESGDIETIG	RHDFDGLKAMSMTAHPKTDPIETGETFAFRYGPVPPFLTFFRFDSAGK	320
BolC.CCD4.a_To1000DH3	SLAFFCNRLFALGESDLPYAVRLTESGDIETIG	RHDFDGLKAMSMTAHPKTDPIETGETFAFRYGPVPPFLTFFRFDSAGK	320
BolC.CCD4.b_Early.Big	SLAFFCNRLFALGESDLPYAVRLTESGDIETIG	RHDFDGLKAMSMTAHPKTDPIETGETFAFRYGPVPPFLTFFRFDSAGK	280
BolC.CCD4.b_To1000DH3	SLAFFCNRLFALGESDLPYAVRLTESGDIETIG	RHDFDGLKAMSMTAHPKTDPIETGETFAFRYGPVPPFLTFFRFDSAGK	320
AtCCD4	KQRDVPIFSMTSPSFIHDFAITKRHAIFAEIQ	LMRMNIMDLVLEGGSPVGTDNGKTPRLGVIPIRYAGDESEMKEWFEVPG	399
BolC.CCD4.a_Early.Big	KRKDVPIFSMTSPSFIHDFAITKRHAIFAEIQ	LMRMNIMDLVLEGGSPVGADNGKTPRLGVIPIRYAGDESEMKEWFEVPG	397
BolC.CCD4.a_To1000DH3	KRKDVPIFSMTSPSFIHDFAITKRHAIFAEIQ	LMRMNIMDLVLEGGSPVGADNGKTPRLGVIPIRYAGDESEMKEWFEVPG	397
BolC.CCD4.b_Early.Big	KQRDVPIFSMTSPSFIHDFAITKRHAIFAEIQ	LMRMNIMDLVLEGGSPVGADNRKTPRLGVIPIRYAGDESEMKEWFEVPG	360
BolC.CCD4.b_To1000DH3	KQRDVPIFSMTSPSFIHDFAITKRHAIFAEIQ	LMRMNIMDLVLEGGSPVGADNRKTPRLGVIPIRYAGDESEMKEWFEVPG	400
AtCCD4	FNIIHAINAWDEDDGNSVLIAPNIMSIEHTLER	MDLVHALVEKVKIDLVGTGIVTRHPISARNLDFAVINPAFLGRCSRY	479
BolC.CCD4.a_Early.Big	FNIIHAINAWDEDDGNTIVLIAPNIMSIEHTLER	MDLVHALVEKVKIDLVGTGIVTRHPISARNLDFAVINPAFLGRCSRY	477
BolC.CCD4.a_To1000DH3	FNIIHAINAWDEDDGNTIVLIAPNIMSIEHTLER	MDLVHALVEKVKIDLVGTGIVTRHPISARNLDFAVINPAFLGRCSRY	477
BolC.CCD4.b_Early.Big	FNIIHAINAWDDR-----FDRSERY		380
BolC.CCD4.b_To1000DH3	FNIIHAINAWDEDDGNTIVLIAPNIMSIEHTLER	MDLVHSLVEKVKIDLVGTGIVTRHPISARNLDFAVINPAFLGRCSRY	480
AtCCD4	VYAAIGDPMPIKISGVVKLDVSKGDRDDCTV	ARRMYGSGCYGGEFFVARDPGNPEAEEDDGYYVTVVHDEVAGESKFLVM	559
BolC.CCD4.a_Early.Big	VYAAIGDPMPIKISGVVKLDVTRGDRDDCTV	ARRMYGSGCYGGEFFVARDPGNPEAEEDDGYYVTVVHDEVAGESKFLVM	557
BolC.CCD4.a_To1000DH3	VYAAIGDPMPIKISGVVKLDVTRGDRDDCTV	ARRMYGSGCYGGEFFVARDPGNPEAEEDDGYYVTVVHDEVAGESKFLVM	557
BolC.CCD4.b_Early.Big	VHRRAY-----VREDGSRSR-----LGGE-----	SDRSRHRDYRDTSSDLREPRFRGD	423
BolC.CCD4.b_To1000DH3	VYAAIGDPMPIKISGVVKLDVSKGDRDDCTV	ARRMYGSGCYGGEFFVARDPQDQAEEDDGYYVTVVHDEVAGESKFLVM	560
AtCCD4	DAKSPELEIVAAVRLPRRPYGFHGLFVKESD	LNKL	595
BolC.CCD4.a_Early.Big	DAKSPELEVAAVRLPRRPYGFHGLFVKESD	LNKL	593
BolC.CCD4.a_To1000DH3	DAKSPELEVAAVRLPRRPYGFHGLFVKESD	LNKL	593
BolC.CCD4.b_Early.Big	QSGVSREREPLRVGRNRS--DAEDLRSGESR	--	455
BolC.CCD4.b_To1000DH3	DAKSPELEIAAAVRLPRRPYGFHGLFVKESD	LNKL	596

Figura 14. Alineamiento múltiple de las secuencias aminoácidas de CCD4 en *A. thaliana* y *B. oleracea*. Los parálogos de CCD4 en *B. oleracea* TO1000DH3 tienen un 87% de identidad a nivel de secuencia proteica. Los cuatro residuos de histidina (H) se encuentran encerrados en rectángulos, destacado con una estrella se encuentra la H faltante en *BolC.CCD4.b*. La secuencia correspondiente al péptido señal se encuentra subrayada en rojo.

Las secuencias río arriba y río abajo de los genes *BolC.CCD4* de TO1000DH3 y “*Early Big*” fueron obtenidas gracias a la colaboración existente con el Agriculture and Agri-Food Canada-Saskatoon Research Centre (AAFC-SRC), quienes forman parte del consorcio internacional que se encuentra secuenciando el genoma de *B. oleracea* (genoma C) utilizando TO1000DH3 como strain de referencia. Estas secuencias fueron confirmadas por PCR, aislación y secuenciación utilizando cuatro sets de oligonucleótidos por gen, diseñados a fin de abarcar tanto regiones codificantes como no codificantes de estos (Figura 15).

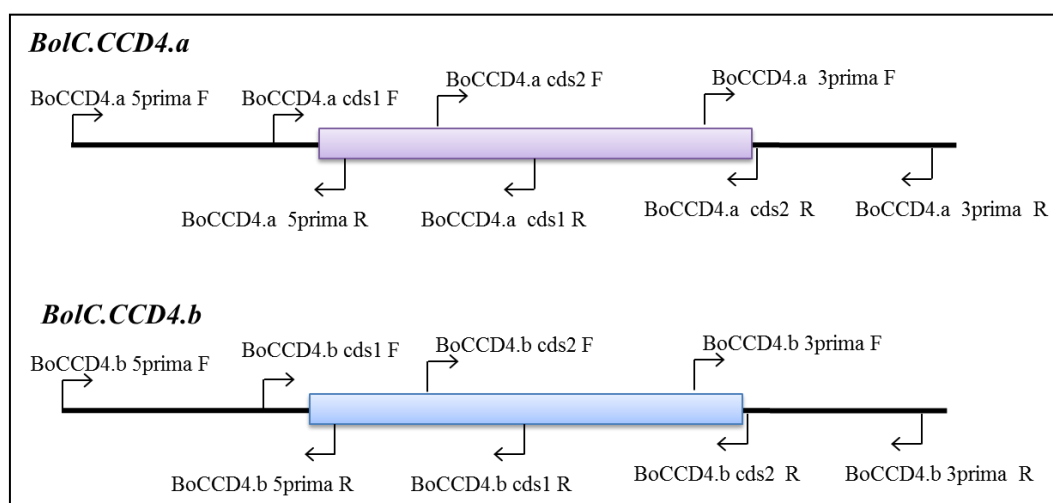


Figura 15. Localización de los oligonucleótidos diseñados para aislar las secuencias codificantes y no codificantes de los genes CCD4 en *B. oleracea*.

El análisis de secuencias no codificantes para los genes CCD4 en *B. oleracea* reveló elementos repetitivos tanto en regiones 5' como 3' de los genes (Figura 16). El elemento repetitivo SINE (por su nombre en inglés *Short Interspersed Nuclear Elements*) se encuentra presente en el gen *BolC.CCD4.b* de TO1000DH3 en regiones no codificantes colindantes al promotor del gen. Si bien este tipo de elemento se encuentran altamente metilado (Kochanek *et al.*, 1993) por lo que se relacionan con la no actividad transcripcional, estos elementos puede ser activados en un tejido específico y/o de manera específica en el desarrollo (Hellmann-Blumberg *et al.*, 1993, Kochanek *et al.*, 1993, Deragon *et al.*, 1996).

Otro elemento encontrado en los genes *BolC.CCD4* corresponde a los retrotransposones, (*Ty1-copia*) los cuales constituyen la mayor división de los elementos móviles eucarióticos que se transponen a través de un transcrito de ARN que mediante la transcriptasa reversa codificada por el retrotransposon, es transcrito en reversa a ADN (Kumar y Bennetzen, 1999). Estas diferencias podrían estar asociadas a cambios en los patrones de expresión de estos genes.

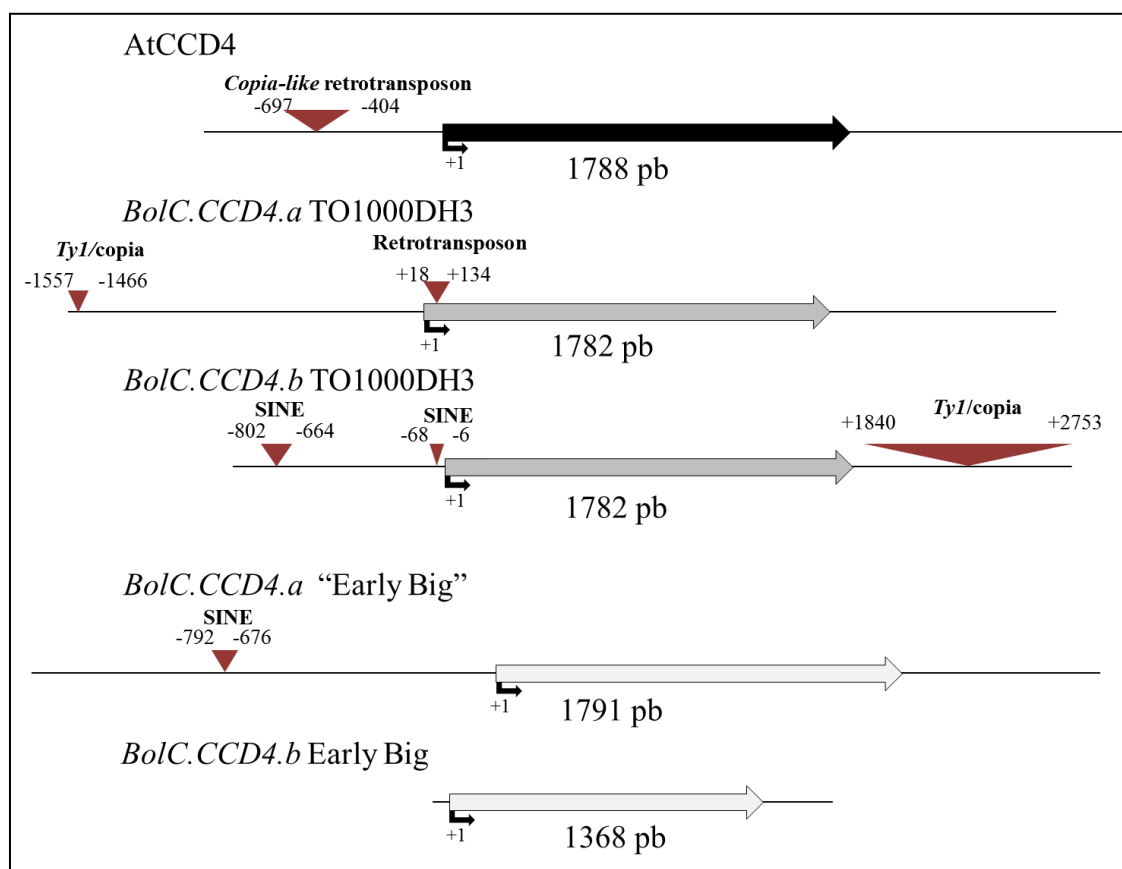


Figura 16. Elementos repetitivos en los genes CCD4. Los elementos repetitivos de cada secuencia nucleotídica se encuentran marcados con triángulos rojos, indicando su posición y tipo de elemento. Los tamaños (pb) correspondientes a las regiones codificantes se encuentran bajo cada gen.

A partir de la secuencia aminoacídica de los genes CCD4 en *B. oleracea* TO1000DH3 y “*Early Big*” se realizó una modelación de la estructura proteica utilizando el software SWISS-MODEL (Arnold *et al.*, 2006). Donde se obtuvo una estructura proteica similar a la descrita para la enzima ACO de *Synechocystis* PCC 6803 (Kloer *et al.*, 2005) para *BolC.CCD4.a* de ambos genotipos y *BolC.CCD4.b* de TO100DH3. Por otra parte, el modelamiento estructural para la proteína *BolC.CCD4.b* de “*Early Big*” reveló cambios evidentes a nivel conformacional (Figura 17).

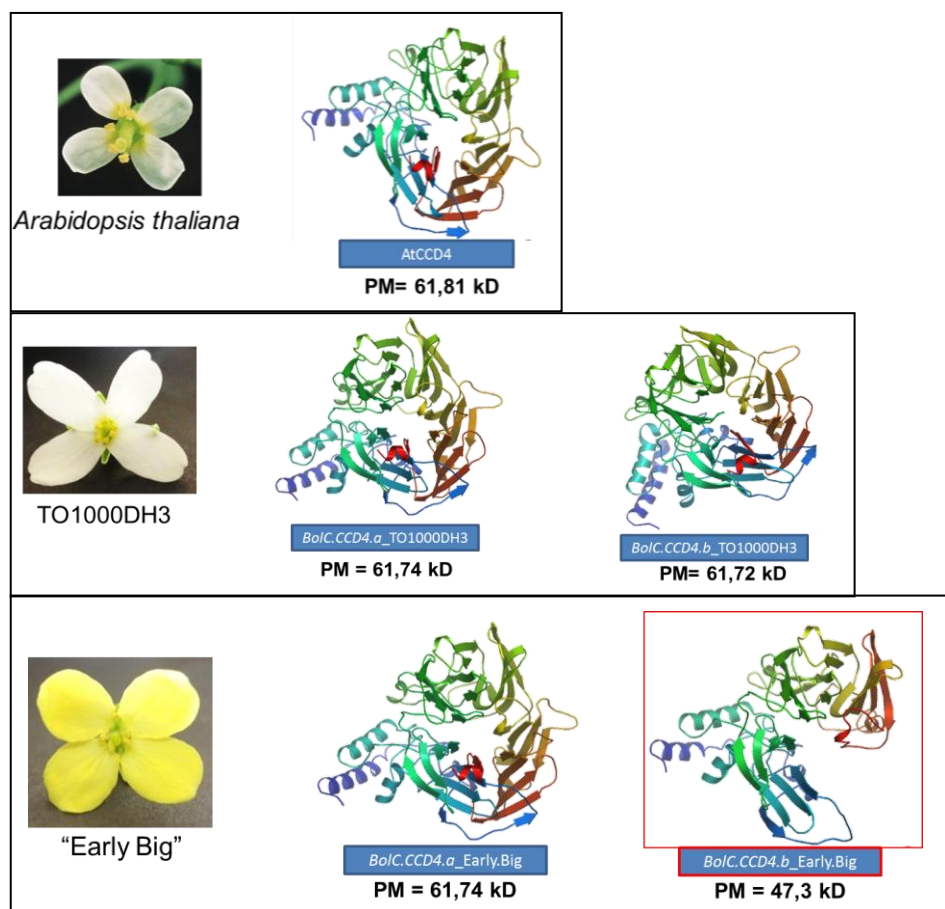


Figura 17. Diagrama Ribbon de las enzimas CCD4 de *A. thaliana*, *B. oleracea* TO1000DH3 y “*Early Big*”. Los pesos moleculares (PM) de cada enzima se encuentran indicados bajo cada figura. En un cuadrado rojo se encuentra destacada la estructura proteica de *BolC.CCD4.b* de “*Early Big*” que presenta cambios conformacionales.

5.4 Expresión de genes *BolC.CCD4* en tejidos vegetativos y reproductivos en *B. oleracea*

Inicialmente se realizó la evaluación de la eficiencia de amplificación de los set de oligonucleótidos específicos diseñados para cada parólogo de CCD4 y *ubiquitina* (gen normalizador). Utilizando una dilución seriada de plásmidos (10^{-9} – 10^{-1}) clonados con la secuencia codificante para los genes *BolC.CCD4.a*, *BolC.CCD4.b*, y una secuencia parcial del gen *ubiquitina* se procedió a la amplificación de cada dilución. Las eficiencias obtenidas para cada set de oligonucleótidos fueron: (i) Bo.CCD4.a F1qpcr/ Bo.CCD4.a R1qpcr: 97%, (ii) Bo.CCD4.a F2qpcr/ Bo.CCD4.a R2qpcr: 112,4 %, (iii) Bo.CCD4.b F1qpcr/ Bo.CCD4.b R1qpcr: 98%, (iv) Bo.CCD4.b F2qpcr/ Bo.CCD4.b R2qpcr: 97%, (v) Bol-UbiquitinF/Bol-UbiquitinR: 95%.

A partir de estos resultados, se decidió emplear los oligonucleótidos Bo.CCD4.a F1qpcr/ Bo.CCD4.a R1qpcr y Bo.CCD4.b F1qpcr/ Bo.CCD4.b R1qpcr para la amplificación específica de los transcritos de los parálogos de CCD4 y, el set Bol-UbiquitinF/Bol-UbiquitinR para el gen normalizar *ubiquitina*.

Para determinar si el fenotipo flor blanca podría estar determinado por cambios en los patrones de expresión del gen *BolC.CCD4.b* se realizó mediante qRT-PCR la cuantificación relativa de los transcritos en tres estadios de desarrollo de pétalos (Figura 18A) y en hoja madura número 11 (Figura 18B), utilizando ocho genotipos (Figura 19).

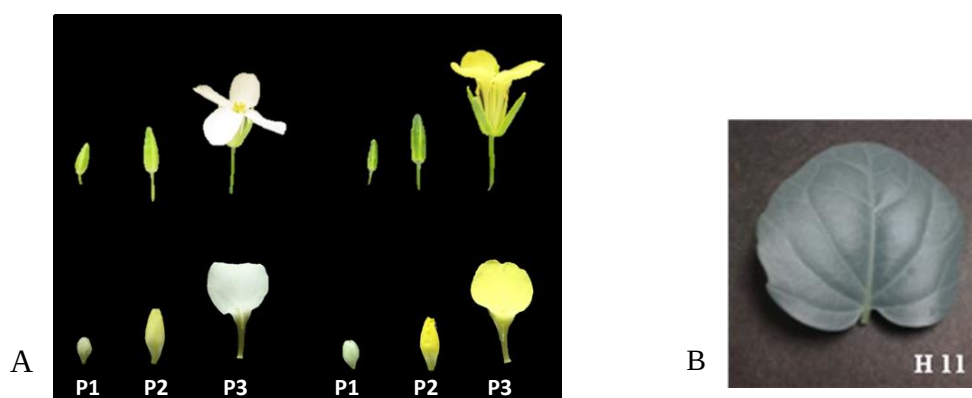


Figura 18. Tejidos analizados mediante qRT-PCR. A, estadios de desarrollo en flores blancas y amarillas, P1: brote floral verde, P2: brote floral amarillo, P3: pétalos expandidos. B, hoja madura.

Parentales	Individuos DH población BolTBDH		
 TO1000DH3	 TODH334	 TODH342	 TODH271
 Early Big	 TODH276	 TODH345	 TODH329

Figura 19. Genotipos utilizados para el análisis de expresión. Los parentales de la población BolTBDH *B. oleracea* TO1000DH3 y “Early Big” junto con tres individuos doble haploides (DH) para cada fenotipo (flor blanca o flor amarilla).

El análisis de la expresión relativa se llevó a cabo utilizando tres réplicas biológicas por genotipo analizado. A partir de los resultados obtenidos mediante qRT-PCR, se observó que el gen *BolC.CCD4.a* se expresó preferencialmente en tejido verde (hoja madura) de todos los genotipos analizados. Por otra parte, el gen *BolC.CCD4.b* se expresó preferencialmente en pétalos blancos, cuyo nivel de transcrito fue aumentando a medida que el pétalo fue perdiendo coloración, alcanzando el máximo del nivel de expresión en el estadio tres de pétalo (P3), cuando estos se encuentran completamente expandidos (Figura 20).

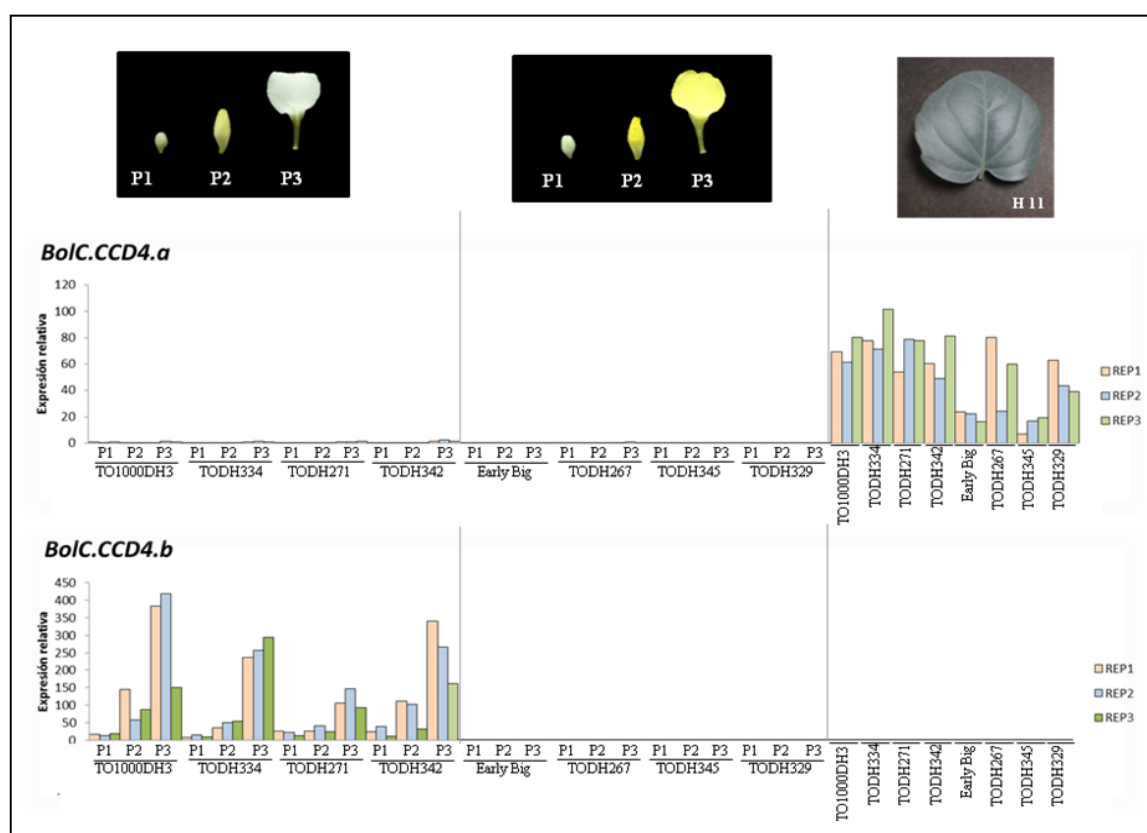


Figura 20. Expresión relativa de los genes CCD4 en pétalos y hojas maduras. El genotipo correspondiente a cada muestra se encuentra en la leyenda del eje X. El gráfico se encuentra dividido en tres partes, donde el primer bloque corresponde a las muestras obtenidas de fenotipo flor blanca, el segundo comprende a los fenotípos flor amarilla y el tercero a las muestras recolectadas de hojas maduras. Cada réplica biológica se encuentra graficada en forma independiente (REP1, REP2, REP3).

6. DISCUSIÓN

El objetivo primordial de este estudio fue determinar la base genética detrás del fenotipo flor blanca en la línea doble haploide *Brassica oleracea* TO1000DH3. Utilizando la genómica comparativa, identificamos un gen candidato con homología a CCD4 de *A. thaliana* en el intervalo del cromosoma C3 de *B. oleracea*, donde previamente se había mapeado la característica color de flor (FC, *flower color*) (Figura 9). Mediante mapeo genético, utilizando ADN-SSCP, logramos identificar dos genes parálogos en el genoma C de *B. oleracea* (Figura 12), los cuales fueron denominados *BolC.CCD4.a* y *BolC.CCD4.b*. Cabe destacar que el gen *BolC.CCD4.b* co-segregó con el fenotipo flor blanca/amarilla y mapeó en la misma posición que el loci previamente identificado para FC (Iniguez-Luy *et al.*, 2009) (Figura 13). Este fue el primer indicio que el gen *BolC.CCD4.b* se encontraría relacionado con la generación del fenotipo flor blanca en *B. oleracea*.

Aunque este resultado no es concluyente, condujo a la formulación de la principal hipótesis. A fin de testearla, se realizó un análisis de expresión para evaluar los genes parálogos en genotipos portadores de flor blanca y flor amarilla. En este análisis de expresión se observó que *BolC.CCD4.b* se encuentra preferencialmente expresado en los pétalos blancos, y que la expresión relativa de este gen va aumentando a medida que el pétalo va perdiendo coloración durante el desarrollo, teniendo su más alta expresión cuando los pétalos se encuentran completamente abiertos (Figura 20). Este efecto no sólo fue observado en el genotipo TO1000DH3, sino que también en individuos de la población doble haploide BolTBDH analizados, que presentaban el mismo fenotipo (Figura 20). Este efecto de los genes CCD4 sobre la acumulación de carotenoides también ha sido observado en flores (Ohmiya *et al.*, 2006), tubérculos (Campbell *et al.*, 2010) y frutos (Brandi *et al.*, 2011) de otras especies vegetales.

Además de los cambios de expresión observados en *BolC.CCD4.b* de TO1000DH3, el análisis estructural de las proteínas CCD4 de *B. oleracea* reveló que el gen *BolC.CCD4.b* de “*Early Big*” codifica para una proteína más corta, con la ausencia de uno de los cuatro residuos de histidina

necesarios para la actividad catalítica de esta familia enzimática (Figura 17). Si bien esta proteína contiene una secuencia peptídica en su región N-terminal (Figura 14) la cual la dirigiría hacia los cloroplastos donde se encontraría su sustrato, la modificación estructural observada podría estar afectando la función de la proteína, haciéndola ya sea una proteína no funcional o modificando la especificidad del sustrato. Esta variación conformacional de la proteína observada en el alelo *BolC.CCD4.b* de “*Early Big*” pudiera estar influenciada por la presión evolutiva, ya que los pétalos amarillos, dado su atractivo color, atraen mucho más a los agentes polinizadores. Estudios futuros serán realizados a fin de evaluar la funcionalidad enzimática de esta proteína y la presión de selección a la cual estuvo sujeta esta familia de genes en *Brassica*.

7. CONCLUSIONES

- El gen *BolC.CCD4.b* mapeó con el marcador fenotípico previamente descrito para el color de flor (FC), lo cual sugiere su participación en el fenotipo flor blanca/amarilla, convirtiéndose en gen candidato.
- El análisis de secuencias nucleotídicas de genes CCD4 de dos genotipos de *B. oleracea*, TO1000DH3 y “*Early Big*”, detectó la presencia de elementos repetitivos en regiones no codificantes tanto 5’ como 3’, lo cual podría estar relacionado con distintos patrones de expresión observados entre parálogos (*BolC.CCD4a* y *BolC.CCD4b*) y entre alelos de un mismo gen.
- La presencia de color amarillo en estadios tempranos de desarrollo en pétalos blancos sugiere que la ruta de biosíntesis de carotenoides está activa en este tejido, y que posteriormente los carotenoides son consumidos para la producción de apocarotenoides incoloros mediante la actividad del gen TO1000DH3 *BolC.CCD4.b* que se encuentra altamente expresado en este tejido, observándose una expresión diferencial en pétalos blancos comparados con pétalos amarillos.
- El análisis de secuencias aminoacídicas reveló que los genes CCD4 de *B. oleracea* presentan un péptido señal en su extremo N-terminal, el cual dirigiría las proteínas a los cromoplastos de los pétalos donde se encontraría su sustrato, catabolizando carotenoides a compuestos incoloros pudiendo generar así el fenotipo flor blanca.
- A su vez, la comparación de secuencias nucleotídicas entre los alelos presentes en TO1000DH3 y “*Early Big*” para ambos genes CCD4 en *B. oleracea* reveló que “*Early Big*” *BolC.CCD4.b* codifica para una proteína más corta comparada con el alelo presente en TO1000DH3 y ambos parálogos *BolC.CCD4.a*. El análisis estructural de la proteína *BolC.CCD4.b* de “*Early Big*”, sugiere que este gen podría ser no funcional o poseer distinta especificidad al sustrato.

8. REFERENCIAS

- AHRAZEM, O., TRAPERO, A., GOMEZ, M. D., RUBIO-MORAGA, A. y GOMEZ-GOMEZ, L.** 2010. Genomic analysis and gene structure of the plant carotenoid dioxygenase 4 family: a deeper study in *Crocus sativus* and its allies. *Genomics*, 96, 239-50.
- ARNOLD, K., BORDOLI, L., KOPP, J. y SCHWEDE, T.** 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22, 195-201.
- AULDRIDGE, M. E., MCCARTY, D. R. y KLEE, H. J.** 2006. Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Curr Opin Plant Biol*, 9, 315-21.
- BEN-AMOTZ, A. y FISHIER, R.** 1998. Analysis of carotenoids with emphasis on 9-cis β -carotene in vegetables and fruits commonly consumed in Israel. *Food Chemistry*, 62, 515-520.
- BOOKER, J., AULDRIDGE, M., WILLS, S., MCCARTY, D., KLEE, H. y LEYSER, O.** 2004. MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. *Curr Biol*, 14, 1232-8.
- BRAMLEY, P. M.** 2003. The genetic enhancement of phytochemicals: the case of carotenoids. In: JOHNSON, I. y WILLIAMSON, G. (eds.) *Phytochemical Functional Foods*. . Cambridge: Woodhead Publishing Ltd.
- BRANDI, F., BAR, E., MOURGUES, F., HORVATH, G., TURCSI, E., GIULIANO, G., LIVERANI, A., TARTARINI, S., LEWINSOHN, E. y ROSATI, C.** 2011. Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biol*, 11, 24.
- CAMPBELL, R., DUCREUX, L. J., MORRIS, W. L., MORRIS, J. A., SUTTLE, J. C., RAMSAY, G., BRYAN, G. J., HEDLEY, P. E. y TAYLOR, M. A.** 2010. The metabolic and developmental roles of carotenoid cleavage dioxygenase4 from potato. *Plant Physiol*, 154, 656-64.
- CHENNA, R., SUGAWARA, H., KOIKE, T., LOPEZ, R., GIBSON, T. J., HIGGINS, D. G. y THOMPSON, J. D.** 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*, 31, 3497-500.
- CHERNYS, J. T. y ZEEVAART, J. A.** 2000. Characterization of the 9-cis-epoxycarotenoid dioxygenase gene family and the regulation of abscisic acid biosynthesis in avocado. *Plant Physiol*, 124, 343-53.
- CHIU, L. W., ZHOU, X., BURKE, S., WU, X., PRIOR, R. L. y LI, L.** 2010. The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol*, 154, 1470-80.
- DELGADO-VARGAS, F. y PAREDES-LÓPEZ, O.** 2003. Natural Colorants for Food and Nutraceutical Uses, CRC PressINC.

- DEMMING-ADAMS, B., GILMORE, A. M. y ADAMS, W. W.** 1996. Carotenoids 3: In vivo functions of carotenoids in higher plants *The Federation of American Societies for Experimental Biology* 10, 403-412.
- DERAGON, J. M., GILBERT, N., ROUQUET, L., LENOIR, A., ARNAUD, P. y PICARD, G.** 1996. A transcriptional analysis of the S1Bn (*Brassica napus*) family of SINE retrotransposons. *Plant Mol Biol*, 32, 869-78.
- EMANUELSSON, O., NIELSEN, H. y VON HEIJNE, G.** 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci*, 8, 978-84.
- ESCRIBANO, J., ALONSO, G. L., COCA-PRADOS, M. y FERNANDEZ, J. A.** 1996. Crocin, safranal and picrocrocin from saffron (*Crocus sativus* L.) inhibit the growth of human cancer cells in vitro. *Cancer Lett*, 100, 23-30.
- FAHEY, J. W. y TALALAY, P.** 1995. The role of crucifers in cancer chemoprotection. In: GUSTINE, D. L. y FLORES, H. E. (eds.) *Phytochemicals and Health*. Rockville: American Society of Plant Physiologists.
- HELLEMANS, J., MORTIER, G., DE PAEPE, A., SPELEMAN, F. y VANDESOMPELE, J.** 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol*, 8, R19.
- HELLMANN-BLUMBERG, U., HINTZ, M. F., GATEWOOD, J. M. y SCHMID, C. W.** 1993. Developmental differences in methylation of human Alu repeats. *Mol Cell Biol*, 13, 4523-30.
- HOWITT, C. A. y POGSON, B. J.** 2006. Carotenoid accumulation and function in seeds and non-green tissues. *Plant Cell Environ*, 29, 435-45.
- INIGUEZ-LUY, F. L., LUKENS, L., FARNHAM, M. W., AMASINO, R. M. y OSBORN, T. C.** 2009. Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa* and *B. oleracea*. *Theor Appl Genet*, 120, 31-43.
- IUCHI, S., KOBAYASHI, M., TAJI, T., NARAMOTO, M., SEKI, M., KATO, T., TABATA, S., KAKUBARI, Y., YAMAGUCHI-SHINOZAKI, K. y SHINOZAKI, K.** 2001. Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. *Plant J*, 27, 325-33.
- JHINGAN, J. L.** 1992. A Novel Technology For DNA Isolation. *Methods in Molecular and Cellular Biology*, 3, 15-22.
- KAKIZAKI, Y.** 1930. A dominant white-flowered mutant of *Brassica oleracea* L. . *Japanese Journal of Genetics*, 6, 55-60.
- KLOER, D. P., RUCH, S., AL-BABILI, S., BEYER, P. y SCHULZ, G. E.** 2005. The structure of a retinal-forming carotenoid oxygenase. *Science*, 308, 267-9.
- KLOER, D. P. y SCHULZ, G. E.** 2006. Structural and biological aspects of carotenoid cleavage. *Cell Mol Life Sci*, 63, 2291-303.
- KOCH, M., HAUBOLD, B. y MITCHELL-OLDS, T.** 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *Am J Bot*, 88, 534-44.

- KOCH, M. A., HAUBOLD, B. y MITCHELL-OLDS, T.** 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Mol Biol Evol*, 17, 1483-98.
- KOCHANKE, S., RENZ, D. y DOERFLER, W.** 1993. DNA methylation in the Alu sequences of diploid and haploid primary human cells. *EMBO J*, 12, 1141-51.
- KUMAR, A. y BENNETZEN, J. L.** 1999. Plant retrotransposons. *Annu Rev Genet*, 33, 479-532.
- LABANA, K. S. y GUPTA, M. L.** 1993. Importance and Origin. In: LABANA, K. S. y BANGA, S. S. (eds.) *Breeding Oilseed Brassicas*. Berlin: Springer-Verlag.
- LAGERCRANTZ, U.** 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics*, 150, 1217-28.
- LAGERCRANTZ, U. y LYDIATE, D. J.** 1996. Comparative genome mapping in *Brassica*. *Genetics*, 144, 1903-10.
- LUKENS, L., ZOU, F., LYDIATE, D., PARKIN, I. y OSBORN, T.** 2003. Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. *Genetics*, 164, 359-72.
- LYNCH, M. y CONERY, J. S.** 2000. The evolutionary fate and consequences of duplicate genes. *Science*, 290, 1151-5.
- MOEHS, C. P., TIAN, L., OSTERYOUNG, K. W. y DELLAPENNA, D.** 2001. Analysis of carotenoid biosynthetic gene expression during marigold petal development. *Plant molecular biology*, 45, 281-293.
- NAMITHA, K. K. y NEGI, P. S.** 2010. Chemistry and biotechnology of carotenoids. *Crit Rev Food Sci Nutr*, 50, 728-60.
- O'NEILL, C. M. y BANCROFT, I.** 2000. Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J*, 23, 233-43.
- OHMIYA, A.** 2009. Carotenoid cleavage dioxygenases and their apocarotenoid products in plants. *Plant Biotechnology*, 26, 351-358.
- OHMIYA, A., KISHIMOTO, S., AIDA, R., YOSHIOKA, S. y SUMITOMO, K.** 2006. Carotenoid cleavage dioxygenase (CmCCD4a) contributes to white color formation in chrysanthemum petals. *Plant Physiol*, 142, 1193-201.
- OSTERGAARD, L. y KING, G. J.** 2008. Standardized gene nomenclature for the *Brassica* genus. *Plant Methods*, 4, 10.
- PARKIN, I. A., GULDEN, S. M., SHARPE, A. G., LUKENS, L., TRICK, M., OSBORN, T. C. y LYDIATE, D. J.** 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics*, 171, 765-81.
- PATERSON, A. H., LAN, T. H., AMASINO, R., OSBORN, T. C. y QUIROS, C.** 2001. *Brassica* genomics: a complement to, and early beneficiary of, the *Arabidopsis* sequence. *Genome Biol*, 2, REVIEWS1011.
- PEARSON, O. H.** 1933. Study of the Life History of *Brassica oleracea*. *Botanical Gazette*, 94, 534-550.

- RAES, J., VANDEPOELE, K., SIMILLION, C., SAEYS, Y. y VAN DE PEER, Y.** 2003. Investigating ancient duplication events in the Arabidopsis genome. *J Struct Funct Genomics*, 3, 117-29.
- RUÍZ-SOLA, M. A. y RODRÍGUEZ-CONCEPCIÓN, M.** 2012. Carotenoid biosynthesis in Arabidopsis: a colorful pathway. *Arabidopsis Book*, 10, e0158.
- SCHWARTZ, S. H., TAN, B. C., GAGE, D. A., ZEEVAART, J. A. y MCCARTY, D. R.** 1997. Specific oxidative cleavage of carotenoids by VP14 of maize. *Science*, 276, 1872-4.
- SCHWARTZ, S. J., JOACHIM, H., ELBE, V. y GIUSTI, M. M.** 2008. Colorants. In: DAMODARAN, S., PARKIN, K. L. y FENNEMA, O. R. (eds.) *Food Chemistry*. London: CRC Press.
- SIMILLION, C., VANDEPOELE, K., VAN MONTAGU, M. C., ZABEAU, M. y VAN DE PEER, Y.** 2002. The hidden duplication past of Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, 99, 13627-32.
- SIMKIN, A. J., SCHWARTZ, S. H., AULDRIDGE, M., TAYLOR, M. G. y KLEE, H. J.** 2004a. The tomato carotenoid cleavage dioxygenase 1 genes contribute to the formation of the flavor volatiles beta-ionone, pseudoionone, and geranylacetone. *Plant J*, 40, 882-92.
- SIMKIN, A. J., UNDERWOOD, B. A., AULDRIDGE, M., LOUCAS, H. M., SHIBUYA, K., SCHMELZ, E., CLARK, D. G. y KLEE, H. J.** 2004b. Circadian regulation of the PhCCD1 carotenoid cleavage dioxygenase controls emission of beta-ionone, a fragrance volatile of petunia flowers. *Plant Physiol*, 136, 3504-14.
- SUN, L., SUN, Y., ZHANG, M., WANG, L., REN, J., CUI, M., WANG, Y., JI, K., LI, P., LI, Q., CHEN, P., DAI, S., DUAN, C., WU, Y. y LENG, P.** 2012. Suppression of 9-cis-epoxycarotenoid dioxygenase, which encodes a key enzyme in abscisic acid biosynthesis, alters fruit texture in transgenic tomato. *Plant Physiol*, 158, 283-98.
- TAN, B. C., JOSEPH, L. M., DENG, W. T., LIU, L., LI, Q. B., CLINE, K. y MCCARTY, D. R.** 2003. Molecular characterization of the Arabidopsis 9-cis epoxycarotenoid dioxygenase gene family. *Plant J*, 35, 44-56.
- TAN, B. C., SCHWARTZ, S. H., ZEEVAART, J. A. y MCCARTY, D. R.** 1997. Genetic control of abscisic acid biosynthesis in maize. *Proc Natl Acad Sci U S A*, 94, 12235-40.
- U, N.** 1935. Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Japan J Bot*, 7, 389-452.
- VAN OOIJEN, J. y VOORRIPS, R.** 2001. JoinMap ® Version 3.0, software for the calculation of genetic linkage maps. Wageningen: Plant Research International.
- WANG, X., TORRES, M. J., PIERCE, G., LEMKE, C., NELSON, L. K., YUKSEL, B., BOWERS, J. E., MARLER, B., XIAO, Y., LIN, L., EPPS, E., SARAZEN, H., ROGERS, C., KARUNAKARAN, S., INGLES, J., GIATTINA, E., MUN, J. H., SEOL, Y. J., PARK, B. S., AMASINO, R. M., QUIROS, C. F., OSBORN, T. C., PIRES, J. C., TOWN, C. y PATERSON, A. H.** 2011. A physical map of Brassica oleracea shows complexity of chromosomal changes following recursive paleopolyploidizations. *BMC Genomics*, 12, 470.
- WARWICK, S. I., FRACIS, A. y AL-SHEHBAZ, I. A.** 2006. Brassicaceae: Species checklist and database on CD-Rom. *Plant Systematics and Evolution*, 259, 249-258.