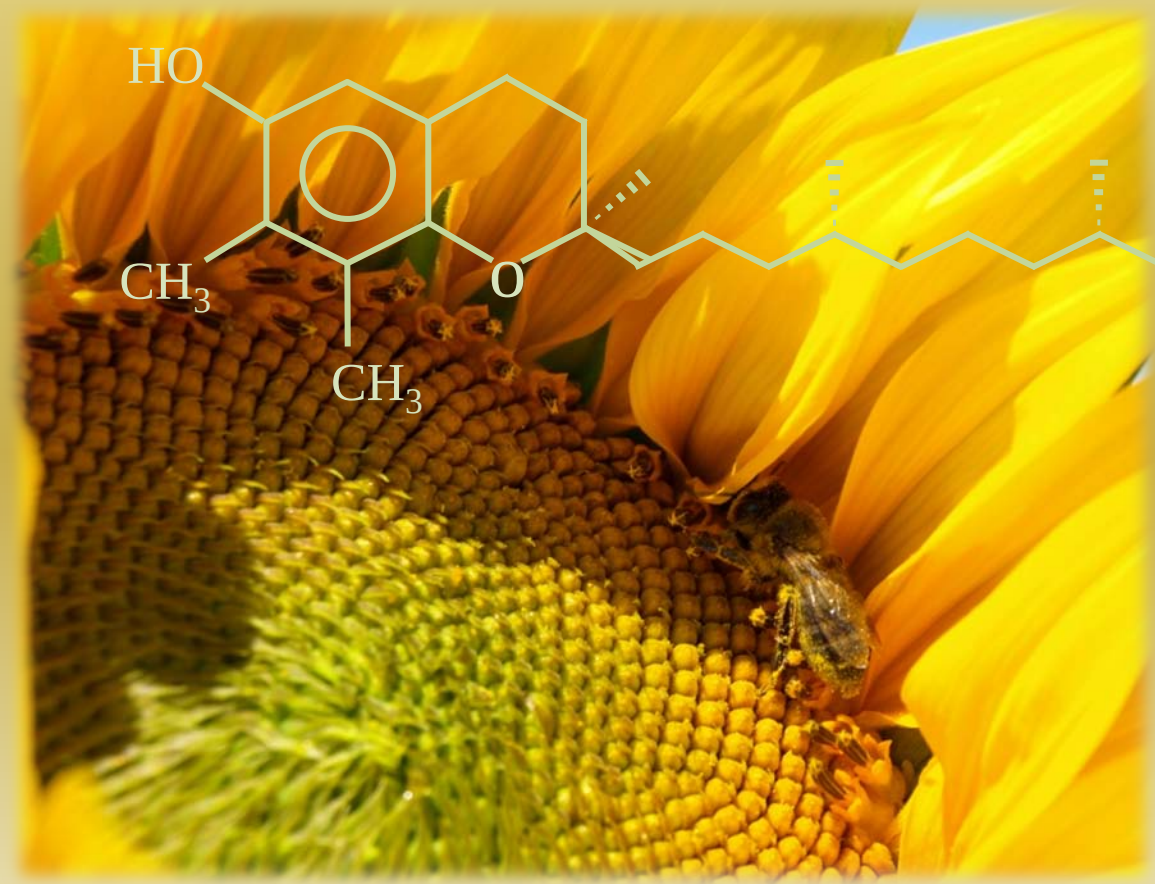
TESIS DOCTORAL

CARACTERIZACIÓN MOLECULAR DE LA ACUMULACIÓN DE GAMMA-TOCOFEROL EN SEMILLAS DE GIRASOL Y CÁRTAMO

MARÍA JOSÉ GARCÍA-M. PÉREZ



GENÉTICA

TÍTULO: *Caracterización molecular de la acumulación de gamma-tocoferol en semillas de girasol y cártamo*

AUTOR: *María José García-Moreno Pérez*

© Edita: Servicio de Publicaciones de la Universidad de Córdoba. 2012
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

UNIVERSIDAD DE CÓRDOBA

DEPARTAMENTO DE GENÉTICA

TESIS DOCTORAL

Caracterización molecular de la acumulación de gamma-tocoferol en semillas de girasol y cártamo

Autor

María José García-Moreno Pérez

Dirigida por

Dra. Begoña Pérez Vich

Dr. Leonardo Velasco Varo

Córdoba, Febrero 2012

UNIVERSIDAD D CÓRDOBA

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
IAS
Instituto de Agricultura Sostenible

Dra. Begoña Pérez Vich, Científico Titular.

INFORMA

Que el trabajo titulado “Caracterización molecular de la acumulación de gamma-tocoferol en semillas de girasol y cártamo” realizado por Dña. María José García-Moreno Pérez, bajo su dirección, se considera finalizado y puede ser presentado para su exposición y defensa como Tesis Doctoral del Departamento de Genética de la Universidad de Córdoba.

Considerando que se encuentra concluida, doy el VºBº para su presentación y lectura.

Fdo: Begoña Pérez Vich

Córdoba, Febrero 2012

UNIVERSIDAD D CÓRDOBA

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
IAS
Instituto de Agricultura Sostenible

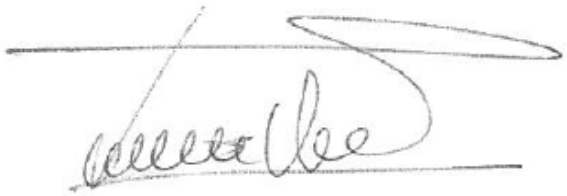
Dr. Leonardo Velasco Varo, Investigador Científico.

INFORMA

Que el trabajo titulado "Caracterización molecular de la acumulación de gamma-tocoferol en semillas de girasol y cártamo" realizado por Dña. María José García-Moreno Pérez, bajo su dirección, se considera finalizado y puede ser presentado para su exposición y defensa como Tesis Doctoral del Departamento de Genética de la Universidad de Córdoba.

Considerando que se encuentra concluida, doy el VºBº para su presentación y lectura.

Fdo: Leonardo Velasco Varo



Córdoba, Febrero 2012

TÍTULO DE LA TESIS: Caracterización molecular de la acumulación de gamma-tocoferol en semillas de girasol y cártamo

DOCTORANDO/A: María José García Moreno Pérez

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

La Tesis Doctoral que se presenta se enmarca en un programa de caracterización genética y molecular de germoplasma de girasol y cártamo con modificaciones en los perfiles de tocoferoles en semillas. En concreto, la tesis se centra en varios aspectos relacionados con el carácter alto gamma-tocoferol en ambas especies. La Tesis incluye cinco capítulos que se centran en i) estudio genético comparativo de diversas fuentes de alto gamma-tocoferol en girasol y mapeo del gen responsable del carácter; ii) estudio fenotípico de segregaciones transgresivas observadas en cruzamientos con una de las líneas alto gamma-tocoferol; iii) caracterización a nivel molecular de los genes modificadores identificados en el capítulo anterior; iv) estudio de la transferencia a cártamo de marcadores moleculares de girasol; y v) identificación y mapeo del gen responsable de alto contenido en alto gamma-tocoferol en cártamo. El desarrollo experimental de la Tesis Doctoral ha requerido el trabajo en campos muy diversos como la realización de cruzamientos, análisis del perfil de tocoferoles en semillas mediante HPLC, puesta a punto y aplicación de protocolos de marcadores moleculares, realización de mapas de ligamiento y análisis QTL, clonación y secuenciación de genes, etc. Este trabajo experimental se ha desarrollado de forma ejemplar por parte de la Doctoranda, lo que ha conducido a unos resultados a nuestro juicio excepcionales que han permitido avanzar de forma muy significativa en el conocimiento sobre la genética de la acumulación de gamma-tocoferol en semillas de girasol y cártamo. Tres capítulos han dado lugar a artículos ya publicados en tres prestigiosas revistas internacionales en el ámbito de la mejora genética vegetal (Crop Science, Euphytica, Molecular Breeding), otro capítulo ha dado lugar a un artículo que está actualmente en evaluación en otra revista de alto impacto como es BMC Plant Biology, mientras que otro capítulo corresponde a un trabajo presentado en la 17th International Sunflower Conference, celebrada en Córdoba en 2008, y publicado en las actas de la misma. Los Directores consideramos que se han cumplido con creces los objetivos propuestos y que el desarrollo de Tesis ha sido plenamente satisfactorio.
Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 9 de Febrero de 2012

Firma del/de los director/es

Fdo.: Begoña Pérez Vich Fdo.: Leonardo Velasco Varo

A mis padres y hermanos

A Carmen

No es valiente el que no tiene miedo,
sino el que sabe conquistarlo.

Nelson Mandela

AGRADECIMIENTOS

Quiero agradecer en primer lugar a mis directores de tesis la Dra. Begoña Pérez Vich y el Dr. Leonardo Velasco Varo, la confianza depositada en mí para llevar a cabo este trabajo, así como por su orientación, apoyo y empatía en ciertos momentos en los que las fuerzas ya flaqueaban. Agradecerles también el estímulo y motivación constante que han supuesto para mí desde que empecé mi formación con ellos en la Mejora Genética Vegetal. También quiero mostrar un gran agradecimiento al Profesor de Investigación Dr. José María Fernández Martínez por su apoyo y ánimos en todos estos años.

Un especial agradecimiento al Dr. Fernando Pistón y la Dra. Mª José Giménez por su amistad y por compartir conmigo muchos de sus conocimientos que han sido de gran ayuda. También agradecer su apoyo y ayuda a mis compañeros de despacho el Dr. Javier Gil Humanes, la Dra. María José Cobos y la Dra. Sara Fondevilla.

Muchas gracias a todo el equipo que ha vivido conmigo muchos ratos de laboratorio haciendo posible el desarrollo de esta amistad, entre ellos compañeros que ya no se encuentran entre nosotros como: Dr. Yamen Hamdan, Dr. Ahmed Yassein, Dra. Paula Caetano, Dr. Behailu Guta, y Paqui Rodríguez; y las que aún continúan: Lidia del Moral, Elena García y Rocío Pineda. Así como a los compañeros y amigos del laboratorio de arriba: Angustias Jiménez, Plácida Nieto, Álvaro Fernández y Antonia Escobar. Todos habéis ido dejando una huella imborrable en mí.

Una mención especial a Lucía Medina que compartió trabajo y casa conmigo durante varios años y de lo que ha quedado una gran amistad.

Gracias también a los chicos del campo, Alberto Merino, Benito Gallardo y Manolo por su inestimable dedicación en los aspectos menos reconocidos de este trabajo de investigación.

Un cariñoso agradecimiento a mis amigos del IAS: Dra. María Ángeles Castillejo, Dr. Mahmoud Eldeeb, Javier Sánchez, Estefanía Carrillo, Carmen Ozuna, Racha Auni y recientemente a la Dra. Gema Guzmán por sus ánimos y por tantos buenos momentos compartidos. No quiero olvidar a Rafa, el vigilante de seguridad de la mañana, sin tus detalles florales, las largas tardes de despacho se habrían hecho mucho menos amenas.

Y como olvidar a todas aquellas personas que ajenas al mundo científico me han mostrado su cariño y estima. Quiero destacar a Teresa Rudilla, mi “mamy” postiza y a toda su familia, por ofrecerme su casa y su encantadora personalidad en aquellos primeros años de tesis. También a José Melero Verdejo, mi cuasi-psicólogo, por ayudarnos a mi familia y a mí y sobre todo por nuestra fuerte amistad. Y a la Dra. Corona Alonso, del departamento de Inmunología, a la que aún, después de tanto tiempo, me une una incondicional amistad. Y no… no me olvido de mis sufridores particulares derivados de esta tesis: Juan Carlos (primo), Inés, Loli, Fermín, Carmen,

Catherine y las chicas brasileñas. Gracias por haber sabido comprender y entender mis ausencias de la vida social.

En el plano más afectivo me gustaría agradecer el apoyo de toda mi familia. Gracias a mis padres (Amalia y Fernando), hermanos, abuelos, tíos… por creer siempre en mi y por poner todo vuestro empeño para que yo consiguiera hacer realidad mi sueño. Espero que estéis tan orgullosos de mí como yo lo estoy de vosotros.

Agradecer al Consejo Superior de Investigaciones Científicas (CSIC) y a la empresa Dow Agrosciences por proporcionarnos la financiación necesaria para culminar este y otros proyectos de investigación.

Por último, pero en el lugar más importante, muchísimas gracias a Carmen (*Little safflower* L.), y a mis aliados Dana, Chica y Moisés. Ha sido un largo camino, lleno de altibajos, de cambios bruscos y siempre habéis estado ahí al pie del cañón dándome la alegría, el cariño y la motivación que necesitaba para seguir adelante.

ÍNDICE GENERAL

INTRODUCCIÓN 1

1. El girasol 1

1.1. Origen e historia 1

1.2. Taxonomía 1

1.3. Descripción botánica 4

2.-El cártamo 5

2.1. Origen e historia 5

2.2. Taxonomía 7

2.3. Descripción botánica 8

3. Importancia económica y distribución geográfica del girasol y del cártamo 9

4. Los tocoferoles en la calidad del aceite de las semillas oleaginosas 11

4.1. Estructura química de los tocoferoles 12

4.2. Biosíntesis de tocoferoles 14

5. Funciones de los tocoferoles 15

5.1. Actividad antioxidante de los tocoferoles 15

5.2. Funciones de los tocoferoles *in vivo* 17

5.2.1. La vitamina E en el organismo 17

5.2.2. Funciones de los tocoferoles en plantas 19

5.3. Funciones de los tocoferoles *in vitro* 21

6. Mejora genética del contenido total y perfil de tocoferoles 22

6.1. Objetivos 22

6.2. Variabilidad para el contenido total y perfil de tocoferoles en girasol 23

6.3. Variabilidad para el contenido total y perfil de tocoferoles en girasol 23

6.4. Estudios genéticos del perfil de tocoferoles en girasol y cártamo 24

6.5. Genes modificadores 25

7. Estudios moleculares en girasol y cártamo 26

7.1. Marcadores moleculares y mapas genéticos en girasol 26

7.2. Marcadores moleculares y mapas genéticos en cártamo 29

7.3. Mejora molecular de caracteres de calidad del aceite de girasol y cártamo 30

7.3.1. Ácidos grasos 30

7.3.2. Tocoferoles 31

REFERENCIAS 33

OBJETIVOS DE LA TESIS 51

RESUMEN GLOBAL DE LOS RESULTADOS Y DISCUSIÓN 52

REFERENCIAS 64

CONCLUSIONES FINALES 70

CAPÍTULO I:

Genetic and molecular analysis of high gamma-tocopherol content in sunflower 71

Abstract 71

Introduction 71

Materials and Methods 72

Plant Material 72

Genetic Study 72

Bulked Segregant Analysis 72

F_2 SSR Genotyping, Map Construction, and Tph2 Mapping 72

Analysis of Tocopherols by High-Performance Liquid Chromatography 73

Results and Discussion 73

Comparative Genetic Study of High Gamma-Tocopherol Lines 73

Molecular Mapping of the Tph2 Gene 73

Acknowledgments 76

References 76

CAPÍTULO II:

A modifying gene affecting gamma-tocopherol content in sunflower 78

Abstract 78

Introduction 78

Materials and Methods 79

Results y Discussion 79

Acknowledgments 81

References 81

CAPÍTULO III:

Genetic basis of unstable expression of high gamma-tocopherol content in sunflower seeds 82

Abstract 83

Background 83

Results 83

Conclusions 84

Background 84
Results 86
Phenotypic segregations 86
Map construction and candidate gene mapping 87
QTL analyses 89
Sequence analysis of gamma-TMT loci 91
Full-length gamma-TMT genomic DNA sequences 92
Discussion 93
Conclusions 97
Methods 98
Plant material, phenotypic analyses and DNA extraction 98
Map construction and molecular analysis 100
Sequence analysis gamma-TMT loci 103
Full-length sequence analysis of the IAST-1 and nmsT2100 gamma-TMT gene 104
Autors´ contributions 105
Acknowledgments and funding 105
References 107
Figure legends 113
Tables 115
Additional files 122

CAPÍTULO IV:

Transferability of non-genic microsatellite and gene-based sunflower markers to safflower 157
Abstract 157
Introduction 157
Materials and Methods 158
Results 159
Discussion 160
Acknowledgments 161
References 161
Supplementary material 163

CAPÍTULO V:

Molecular tagging and candidate gene analysis of the high gamma-tocopherol trait in safflower (Carthamus tinctorius L.) 166
Abstract 166

Introduction 166
Materials and methods 167
Plant material and phenotypic analyses 167
DNA extraction and bulked segregant analysis 167
Cloning and sequencing of RAPD fragments and SCAR development 169
Candidate gene analysis 169
Results 171
Bulked segregant analysis 171
Candidate gene analyses 172
Partial sequencing of the safflower gamma-TMT gene and development of gene-specific markers 174
Discussion 174
Acknowledgments 177
References 177
Supplementary material 179
INFORME DEL FACTOR DE IMPACTO 181
ACEPTACIÓN DE LOS COAUTORES DE LOS TRABAJOS 185

INTRODUCCIÓN GENERAL

1. El girasol

1.1. Origen e historia

El origen geográfico del girasol (*Helianthus annuus* L.) se sitúa en el oeste de los territorios actuales de los Estados Unidos y norte de México (Heiser, 1954). Las tribus indígenas americanas cultivaban el girasol con fines alimenticios (fabricación de pan y dulces), medicinales (efecto diurético, expectorante y propiedades suavizantes e hidratantes) y rituales (de las flores liguladas se extraía un tinte de color amarillo para la decoración del cuerpo en ceremonias religiosas o con fines guerreros). La evidencia más temprana del girasol domesticado data del año 4625 antes de J.C. (Crites, 1993). Los colonizadores españoles introdujeron la semilla de girasol en España desde donde pasó al resto de Europa cultivándose con fines ornamentales hasta el siglo XVIII (Fick, 1989). A finales del siglo XIX, el girasol se empezó a cultivar en Rusia donde se llevó a cabo una fuerte selección para incrementar su contenido en aceite y llegó a alcanzar una superficie de más de 150.000 hectáreas cultivadas (Alba y Llanos, 1989). Estas variedades oleaginosas fueron introducidas en Norteamérica a finales del siglo XIX (Putt, 1997). Una vez finalizada la Primera Guerra Mundial, el girasol empieza a ocupar un lugar destacado a nivel mundial entre los cultivos oleaginosos. En España, a partir del año 1964 se produjo un incremento espectacular de su cultivo y en el año 1984 se superó el millón de hectáreas cultivadas (Alba y Llanos, 1989).

El origen del girasol cultivado ha sido estudiado mediante el empleo de técnicas moleculares. Burke et al. (2002) realizaron un análisis QTL (Quantitative Trait Loci) para estudiar caracteres potencialmente asociados con la domesticación del cultivo cuya conclusión fue que la fuerte selección direccional sobre el aumento del tamaño del aquenio fue decisiva en la domesticación del girasol. Harter et al. (2004) estudiaron la variación genética existente en girasol silvestre y cultivado concluyendo que los girasoles cultivados surgieron de poblaciones silvestres en la parte central de Estados Unidos.

1.2. Taxonomía

El girasol cultivado (*Helianthus annuus* L.) pertenece a la familia *Asteracea*e (Compuestas), tribu *Heliantheae* y subtribu *Helianthineae* (Panero y Funk, 2002). Su número básico de cromosomas es n = 17. El género *Helianthus* incluye especies diploides, tetraploides y hexaploides (Fernández-Martínez et al., 2009).

Los primeros investigadores que clasificaron el género *Helianthus* fueron Heiser et al. (1969). Posteriormente dicha clasificación fue reformulada por Schilling y Heiser (1981) y recientemente ha sido modificada por Jan y Seiler (2007) (Tabla 1 y Tabla 2).

Tabla 1. Clasificación infragenérica de las especies anuales de *Helianthus* (n =17).

Sección[a]	Especie
Helianthus	*H. annuus* L.
	H. anomalus S.F. Blake
	H. argophyllus Torr. y A. Gray
	H. bolanderi A. Gray
	H. debilis Nutt.
	subsp. *debilis*
	subsp. *cucumerifolius* (Torr. y A. Gray) Heiser
	subsp. *silvestris* Heiser
	subsp. *tardiflorus* Heiser
	subsp. *vestitus* (E. Watson) Heiser
	H. deserticola Heiser
	H. exilis A. Gray
	H. neglectus Heiser
	H. niveus (Benth.) Brandegee
	subsp. *canescens* (A. Gray) Heiser
	subsp. *niveus*
	subsp. *tephrodes* (A. Gray) Heiser
	H. paradoxus Heiser
	H. petiolaris Nutt.
	subsp. *fallax* Heiser
	subsp. *petiolaris*
	H. praecox Engelm. y A. Gray
	subsp. *hirtus* (Heiser) Heiser
	subsp. *praecox*
	subsp. *runyonii* (Heiser) Heiser
Agrestes	*H. agrestis* Pollard
Porteri	*H. porteri* (A. Gray) Pruski

[a] Schilling y Heiser (1981); Jan y Seiler (2007)

Tabla 2. Clasificación infragenérica de las especies perennes de *Helianthus.*

Sección[a]	Serie	Especie	Nº de cromosomas (n)
Ciliares	*Ciliares*	*H. arizonensis* R.C. Jacks	17
		H. ciliaris DC.	34, 51
		H. laciniatus A. Gray	17
Ciliares	*Pumili*	*H. cusickii* A. Gray	17
		H. gracilentus A. Gray	17
		H. pumilus Nutt.	17
Atrorubens	*Coronasolis*	*H. californicus* DC.	51
		H. decapetalus L.	17, 34
		H. divaricatus L.	17
		H. eggertii Small	51
		H. giganteus L.	17
		H. grosseserratus M. Martens	17
		H. hirsutus Raf.	34
		H. maximiliani Schrad.	17
		H. mollis Lam.	17
		H. nuttallii Torr. y A. Gray	
		subsp. *nuttallii*	17
		subsp. *parishii* (A. Gray) Heiser	17
		subsp. *rydbergii* (Britton) R. Long	17
		H. resinosus Small	51
		H. salicifolius A. Dietr.	17
		H. schweinitzii Torr. y A. Gray	51
		H. strumosus L.	34, 51
		H. tuberosus L.	51
Atrorubens	*Microcephali*	*H. glaucophyllus* D.M. Sm.	17
		H. laevigatus Torr. y A. Gray	34
		H. microcephalus Torr. y A. Gray	17
		H. smithii Heiser	17, 34
Atrorubens	*Atrorubentes*	*H. atrorubens* L.	17
		H. occidentalis Riddell	
		subsp. *occidentalis*	17
		subsp. *plantagineus* (Torr. y A. Gray) Heiser	17
		H. pauciflorus Nutt.	
		subsp. *pauciflorus*	51
		subsp. *subrhomboideus* (Rydb.) O. Spring y E.E. Schill.	51
		H. silphioides Nutt.	17
Atrorubens	*Angustifolii*	*H. angustifolius* L.	17
		H. carnosus Small	17
		H. floridanus A. Gray ex Chapm.	17
		H. heterophyllus Nutt.	17
		H. longifolius Pursh	17
		H. radula (Pursh) Torr. y A. Gray	17
		H. simulons E. Watson	17
		H. verticillatus Small	17

[a] Schilling y Heiser (1981); Jan y Seiler (2007)

1.3. Descripción botánica

El girasol es una planta anual, de gran desarrollo en todos sus órganos. Su potente sistema radicular presenta una gran capacidad de adaptación a los recursos hídricos de los distintos niveles del suelo. El tallo es erecto y robusto. Las hojas son alternas, grandes, trinervadas, largamente pecioladas, acuminadas, dentadas, de áspera vellosidad en ambas caras y de forma variable dependiendo de su posición en el tallo. La inflorescencia, denominada capítulo, está formada por numerosas flores situadas en un receptáculo discoidal. Al contrario que las especies silvestres, el girasol cultivado presenta un único capítulo rodeado por unas brácteas involucrales. Los capítulos jóvenes presentan heliotropismo, que cesa en el momento en el que se desarrollan las flores, orientándose únicamente hacia la dirección de salida del sol. La forma del receptáculo puede ser aplanado, cóncavo o convexo, y está formado por dos tipos de flores: liguladas y tubulosas. Las flores liguladas oscilan en un número de entre 30 y 70, tienen forma lanceolada, son asexuadas y están dispuestas radialmente. El color de estas flores varía de amarillo a anaranjado. Las flores tubulosas, situadas en el centro del capítulo, son hermafroditas y están dispuestas en arcos espirales que parten desde el exterior. La polinización en su mayor parte es entomófila. Las abejas son atraídas por el néctar segregado por los nectarios situados en la base de las flores (Vrânceanu, 1977) (Figura 1).

Figura 1: Inflorescencia de girasol visitada por abejas.

El fruto es de tipo aquenio y se caracteriza por presentar un pericarpio duro y fibroso dentro del cual está la semilla. El contenido medio en aceite en los aquenios de las variedades oleaginosas comerciales de girasol es de alrededor del 50%, con un

contenido en cáscara que oscila del 20-25% (Fernández-Martínez et al., 2009) (Figura 2).

Figura 2: Aquenios de girasol (izquierda) y esquema de un fruto de girasol (derecha).

2. El cártamo

2.1. Origen e historia

Vavilov (1951) planteó tres posibles centros de origen del cártamo (*Carthamus tinctorius* L). Un primer centro localizado en la India (centro II), basado en la variabilidad observada y en el estudio de la producción de cártamo en la cultura antigua, un segundo centro limitado a Afganistán (centro III), establecido en base a la diversidad del cártamo encontrada y a la proximidad a especies silvestres, y por último, un tercer centro de origen situado en Etiopía (centro VI), en base a la presencia de especies silvestres de cártamo. Posteriormente Ashri y Knowles (1960) y Hanelt (1961), partiendo de la hipótesis de la existencia de una estrecha similitud entre el cártamo cultivado y las especies silvestres *C. flavescens* (= *C. oxyacanthus* Bieb.), distribuido en Turquía, Siria y Líbano y *C. palaestinus* Eig, localizado en el desierto de Irak y en Palestina, concluyeron que el centro de origen del cártamo debió ser Oriente Medio. Finalmente, Knowles (1969) propuso el término centros de similaridad en base a la observación de semejanzas entre los tipos de planta cultivados, estableciendo siete áreas como origen del cártamo cultivado: el Lejano Oriente, India-Paquistán, Oriente Medio, Egipto, Sudán, Etiopía y Europa.

Se han llevado a cabo estudios moleculares para precisar el origen de este cultivo. Chapman y Burke (2007), mediante análisis filogenéticos, señalaron a *C.*

palaestinus como único progenitor silvestre del cártamo cultivado mientras que *C. oxyacanthus* y *C. persicus* estaban evolutivamente más distanciados. Sehgal et al. (2008), analizando el ADN cloroplástico concluyeron la contribución tanto de *C. flavescens* (= *C. oxyacanthus*) como de *C. palaestinus* en la evolución del cártamo cultivado. Recientemente se ha llevado a cabo otro estudio en el que se emplearon marcadores RAPD (Random Amplified Polymorphic DNA), SSR (Simple Sequence Repeats) y AFLP (Amplified Fragment Length Polymorphism) para estudiar los centros de similaridad de cártamo, encontrando que la diversidad genética de *C. tinctorius* detectada en Afganistán es muy alta si se la compara con la del Lejano Oriente (Sehgal et al., 2009). Bowles et al. (2010) realizaron un análisis filogenético en el género *Carthamus* y revelaron, al igual que Chapman y Burke (2007), que *C. palaestinus* es la especie más cercana al cártamo cultivado. Chapman et al. (2010) realizaron un estudio de variabilidad genética en esta especie revelando un origen en oriente próximo y cinco centros de diversidad localizados en: 1) Europa; 2) Turquía, Irán, Irak y Afganistán; 3) Palestina, Jordania y Siria; 4) Egipto y Etiopía y 5) India y Paquistán.

El cártamo ha sido cultivado desde la antigüedad principalmente por el tinte rojo-anaranjado (cartamina) extraído de sus flores y por la calidad de su aceite, rico en ácidos grasos polinsaturados (Singh y Nimbkar, 2007). En Egipto fue empleado tanto para colorear tejidos (algodón y seda) como en ceremonias religiosas (embalsamamiento y momificación). Se han encontrado semillas y guirnaldas con flores de cártamo en momias de más de 4000 años de antigüedad. Además, su aceite fue muy apreciado como ungüento y como combustible de lámparas. En Oriente Medio, India y África, el cártamo fue muy solicitado con fines medicinales por su efecto purgativo, antipirético y como antídoto para ciertos venenos (Weiss, 1971). En el siglo XVIII el cártamo se extendió a Italia, Francia y Gran Bretaña, donde se empleó el tinte extraído de sus flores como colorante alimentario. Además, fue extensivamente utilizado como colorante en las fábricas de alfombras de Europa del Este, Oriente Medio y el subcontinente Indio hasta el siglo XIX. La disponibilidad de nuevos tintes de anilina sintética, más económicos, hizo que el cártamo pasara a cultivarse principalmente por el aceite comestible extraído de sus semillas, con altos niveles de ácido oleico o linoleico según los cultivares (Dajue y Mündel, 1996).

2.2. Taxonomía

El género *Carthamus* pertenece a la familia *Asteraceae* (Compuestas), subfamilia *Tubulifloreae* y tribu *Cynareae* (Sehgal y Raina, 2011). La especie cultivada es *C. tinctorius* L. (Ashri y Knowles, 1960). Tanto *C. tinctorius* como la mayoría de las especies silvestres son diploides (Sehgal y Raina, 2011). Los estudios de clasificación realizados sobre el género *Carthamus* presentan la dificultad de distinguir taxonómicamente dicho género del género *Carduncellus* (Hanelt, 1963). López-González (1990) clasificó las especies de estos dos géneros en cuatro géneros en base a estudios morfológicos, anatómicos, corológicos y biosistemáticos. Los grupos se muestran a continuación:

I. *Carthamus* L. (lectotipo *Carthamus tinctorius* L.).
II. *Phonus* J. Hill (lectotipo *Phonus arborescens* (L.) G. López = *Carthamus arborescens* L.).
III. *Lamottea* Pomel (lectotipo *Lamottea caerulea* (L.) Pomel = *Carthamus caeruleus* L.).
IV. *Carduncellus* Adanson (lectotipo *Carduncellus monspelliensium* All.).

Asimismo López-González (1990) dividió el género *Carthamus* en tres secciones dependiendo del número de cromosomas:

I. Sección *Carthamus* (n=12). Especies que comprende: *C. curdicus* Hanelt, *C. gypsicola* Ilj., *C. oxyacanthus* Bieb., *C. palaestinus* Eig, *C. persicus* Willd. y *C. tinctorius* L.

II. Sección *Odonthagnathius* (DC.) Hanelt (n=10, 11). Especies que comprende: *C. boissieri* Halácsy, *C. dentatus* Vahl, *C. divaricatus* Beguinot y Vacc. (n=11), *C. glaucus* Bieb., *C. leucocaulos* Sm. y *C. tenuis* (Boiss. y Bl.) Bornm.

III. Sección *Atractylis* Reichenb. (n=11). Especies que comprende: *C. lanatus* L., *C. creticus* L. [= *C. baeticus* (Boiss. y Reuter) Nyman] y *C. turkestanicus* M. Popov.

Además, López-González (1990) indicó que la especie *C. nitidus* Boiss. no es posible encajarla en ninguna de estas tres secciones por lo que sugirió que podría constituir una sección independiente.

2.3. Descripción botánica

El cártamo es una planta anual. Posee un sistema de raíces pivotantes bastante profundas que le permite extraer el agua y los nutrientes de las capas más profundas del suelo, por lo que es adecuado como cultivo de secano. La planta se caracteriza por ser muy ramificada y cada rama puede alojar de uno a cinco capítulos rodeados por unas brácteas espinosas (Singh y Nimbkar, 2007). El capítulo es de forma globular y puede contener de 20 a 250 flores sobre un receptáculo plano o convexo. Las hojas son generalmente muy espinosas (Figura 3) y su tamaño varía mucho entre variedades y en la misma planta. Las brácteas son cortas y rígidas (Dajue y Mündel, 1996).

Figura 3: Hojas espinosas e inflorescencia de cártamo.

El color de las flores del capítulo puede ser amarillo (Figura 3), anaranjado o blanco. La floración dura alrededor de un mes y se inicia en forma de espiral en las flores más exteriores (centrípetamente). Las flores son tubulares, hermafroditas y su polinización en su mayor parte ocurre por autofecundación, pero también puede darse polinización cruzada, llevada a cabo principalmente por abejas. El fruto es de tipo aquenio y puede presentar vilano (Figura 4). El contenido medio en aceite en los aquenios de las variedades oleaginosas comerciales de cártamo se encuentra entre 35-39%, con un contenido en cáscara en torno al 40% (Fernández-Martínez, 1997). Cada capítulo produce normalmente entre 15 y 30 semillas, a veces incluso más, que maduran de 4 a 8 semanas después de la floración (Dajue y Mündel, 1996).

Figura 4: Fruto (aquenio) y detalle de un vilano de cártamo.

3. Importancia económica y distribución geográfica del girasol y del cártamo

Actualmente, el girasol ocupa la cuarta posición en producción mundial de aceite, mientras que el cártamo es un cultivo oleaginoso cuya importancia en la producción y comercio mundial de aceites vegetales es limitada (FAOSTAT, 2011). La producción media anual de los aceites obtenidos de semillas de girasol y cártamo en el periodo comprendido entre el año 2004 y 2009 fue de 11.256.000 y 144.000 toneladas, respectivamente. Dichas cifras quedan muy lejos de los 38 millones de toneladas de aceite de palma producidos de media anualmente, siendo éste el principal aceite producido en el mundo (Tabla 3).

Tabla 3. Producción mundial de aceite ($x10^3$ toneladas) en las principales especies de oleaginosas desde 2004 a 2009 (FAOSTAT, 2011).

Producto	2004	2005	2006	2007	2008	2009	Promedio
Aceite de palma	31525	34406	37286	38847	43428	45084	38429
Aceite de soja	30472	33979	34757	37279	36633	35873	34832
Aceite de colza	15013	16759	17934	17914	19164	21175	17993
Aceite de girasol	9807	10648	11600	11193	11060	13228	11256
Aceite de cacahuete	5278	5589	4883	5710	5523	5561	5424
Aceite de nuez de palma	3950	4424	4824	4977	5502	5734	4902
Aceite de coco	3581	3413	3449	3691	3395	3502	3505
Aceite de oliva	2914	2552	2911	2762	2620	2911	2778
Aceite de sésamo	920	921	971	966	961	1006	958
Aceite de lino	610	619	710	691	610	580	637
Aceite de cártamo	144	132	130	150	158	150	144

La Tabla 4 muestra la superficie cultivada a nivel mundial de las principales oleaginosas en el periodo 2004-2009. La superficie media anual cultivada de girasol y de cártamo en este periodo fue de 23.066.000 y 787.000 hectáreas, respectivamente, mientras que la soja alcanzó un promedio de 94 millones de hectáreas.

Entre los principales países donde se cultiva el girasol destaca Rusia, que en el periodo 2004-2009 representó un porcentaje de 24% en promedio de producción mundial, lo que lo coloca como el primer país productor. Le siguen Ucrania, que ha contribuido con un 20%, Argentina, con un 14%, China, con un 7% y Francia, con un 6% (Tabla 5).

Tabla 4. Superficie cultivada a nivel mundial ($x10^3$ ha) de las principales especies de oleaginosas desde 2004 a 2009 (FAOSTAT, 2011).

Especie	2004	2005	2006	2007	2008	2009	Promedio
Soja	91593	92524	95308	90156	96481	99501	94261
Colza	25317	27694	27441	29888	30660	31121	28687
Cacahuete	23740	24048	21615	22465	23958	23951	23296
Girasol	21407	22985	23975	21281	25031	23717	23066
Palma	12270	12893	13275	13889	14702	14921	13659
Coco	11097	11191	11196	11529	11546	11864	11404
Oliva	9156	9654	9995	9559	10045	9207	9603
Sésamo	7454	7530	7380	7053	7407	7700	7421
Lino	2438	2785	2784	2050	2197	2112	2394
Ricino	1431	1586	1267	1491	1554	1481	1468
Cártamo	966	837	705	755	710	752	787
Mostaza	972	721	606	604	690	861	742

Tabla 5. Promedio de producción de semillas ($x10^3$ toneladas), superficie cultivada ($x10^3$ ha) y rendimiento de girasol (kg/ha), desde 2004 a 2009 en los principales países productores (FAOSTAT, 2011).

País	Producción	Superficie	Rendimiento
Rusia	6243	5441	1143
Ucrania	5024	3819	1299
Argentina	3536	2111	1671
China	1703	931	1820
Francia	1504	630	2392
EE.UU.	1326	840	1561
India	1229	1977	625
Hungría	1210	520	2323
Rumania	1207	864	1376
Bulgaria	1066	664	1592
Turquía	983	569	1724
España	724	679	1056

La India es el principal país productor de semillas de cártamo. En el periodo 2004-2009 la producción de este país representó un porcentaje del 33% de la producción mundial. Le siguen México, que ha contribuido con un 19%, EE.UU. con un 17%, Kazajstán con un 9% y Argentina con un 7% (Tabla 6).

Tabla 6. Promedio de producción de semillas (toneladas), superficie cultivada (ha) y rendimiento de cártamo (kg/ha), desde 2004 a 2009 en los principales países productores (FAOSTAT, 2011).

País	Producción	Superficie	Rendimiento
India	198517	348740	573
México	117262	97281	1267
EE.UU.	103598	69708	1480
Kazajstán	56698	101107	570
Argentina	44172	53136	799
China	31750	21833	1455
Kirguistán	15050	17631	850
Australia	13471	22178	527
Tanzania	8868	22951	386
Etiopía	7293	9020	836
Turquía	5031	4894	1110
Uzbekistán	4810	10403	512

El promedio del rendimiento mundial de girasol para el periodo 2004-2009 fue de 1548 kg/ha. Francia (2392 kg/ha) registró el mayor rendimiento para este período, seguido por Hungría (2323 kg/ha) y China (1820 kg/ha). España registró un rendimiento de 1056 kg/ha en este periodo (Tabla 5).

El promedio del rendimiento mundial de cártamo para el periodo 2004-2009 fue de 864 kg/ha. EE.UU. (1480 kg/ha) registró el mayor rendimiento para este período, seguido por China (1455 kg/ha) y México (1267 kg/ha) (Tabla 6).

4. Los tocoferoles en la calidad del aceite de las semillas oleaginosas

Los aceites vegetales están constituidos en su mayor parte por triacilgliceroles (92-98%), formados por un esqueleto de glicerol esterificado con tres moléculas de ácidos grasos, y en menor cantidad por monoacilgliceroles y diacilgliceroles, lípidos polares, ácidos grasos libres y por una serie de compuestos liposolubles entre los que se encuentran los tocoferoles (Åppelqvist, 1989). Debido a su composición química, el aceite extraído de las semillas oleaginosas es un componente vital de la dieta humana ya

que constituye una de las fuentes de energía más importantes, actúa como transportador de ciertas vitaminas y proporciona al organismo ácidos grasos esenciales (Vles y Gottenbos, 1989). A partir del aceite se elaboran numerosos productos como aderezos de ensaladas, margarinas, etc., pero también se destina a fines no alimenticios, como la fabricación de lubricantes, biodiesel, detergentes, jabones, surfactantes, emulsionantes, cosméticos, etc. (Velasco et al., 2004a).

La calidad del aceite está determinada principalmente por su composición en ácidos grasos, el patrón de distribución de estos ácidos grasos en las moléculas de triacilgliceroles y por la presencia de compuestos menores entre los que se encuentran los tocoferoles. Los tocoferoles son unas sustancias con alta capacidad antioxidante tanto *in vivo* como *in vitro*. Estas moléculas contribuyen a las propiedades físicas, químicas y nutricionales del aceite (Somerville, 1991; Padley et al., 1994; Velasco et al., 2004a).

4.1. Estructura química de los tocoferoles

Los tocoferoles son unos compuestos liposolubles que junto con los tocotrienoles y el plastocromanol-8 forman parte de un grupo más amplio de moléculas químicas denominadas genéricamente tococromanoles (Kamal-Eldin y Åppelqvist, 1996). Los tococromanoles son compuestos viscosos a temperatura ambiente, de color amarillento, insolubles en agua pero solubles en disolventes no polares como el hexano. Bioquímicamente, son moléculas anfipáticas cuya estructura consta de dos partes: un anillo cromano polar formado por dos anillos, uno fenólico y otro heterocíclico, y una cadena lateral de naturaleza isoprenoide (Caretto et al., 2010). Los tocoferoles y los tocotrienoles se diferencian en la saturación de esta cadena lateral. Los tocoferoles poseen una cadena saturada mientras que los tocotrienoles presentan una cadena insaturada con 3 dobles enlaces (Brigelius-Flohé et al., 2002) (Figura 5).

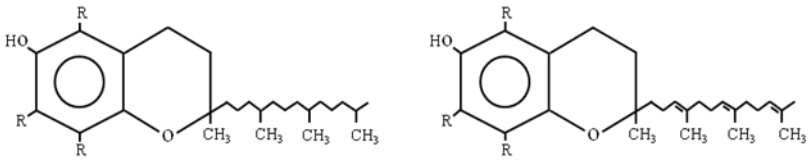


Figura 5: Fórmula química de los tocoferoles y los tocotrienoles.

Tanto los tocoferoles como los tocotrienoles pueden presentar cuatro formas diferentes denominadas alfa- (α-), beta- (β-), gamma- (γ) y delta-tocoferol (δ-tocoferol), que difieren en el número de grupos metilo y en su posición en el anillo fenólico (Kamal-Eldin y Åppelqvist, 1996; Brigelius-Flohé et al., 2002) (Figura 6).

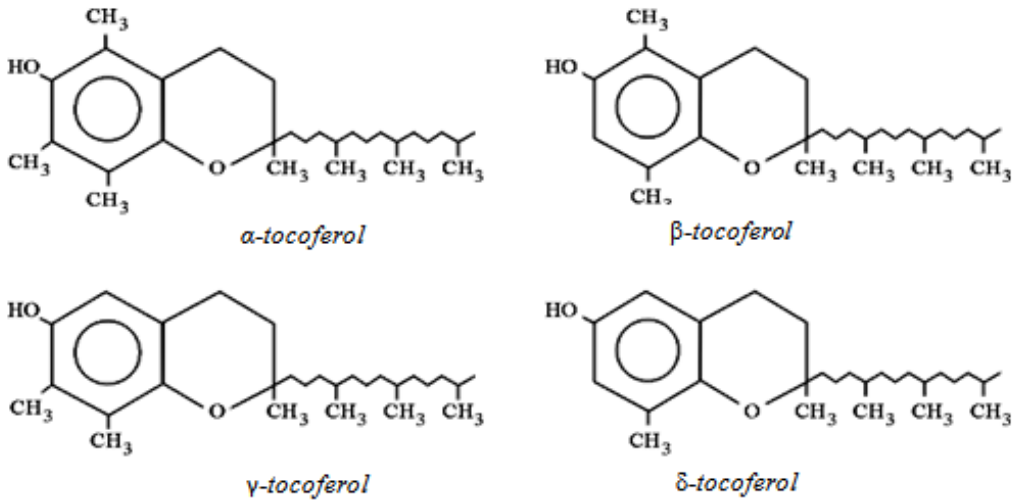


Figura 6: Estructura química de los distintos tipos de tocoferoles.

Los tocoferoles se sintetizan exclusivamente en los plastidios de los organismos fotosintéticos y en algunas cianobacterias (Schultz, 1990). Se han detectado tocoferoles en hojas, semillas, raíces, tubérculos, frutos, tallos, hipocótilos y cotiledones de las plantas superiores, si bien su contenido y composición es muy heterogénea. La semilla es el tejido que presenta mayor contenido en tocoferoles (Mène-Saffrané y DellaPenna, 2010). Los tejidos verdes fotosintéticos generalmente contienen un alto porcentaje deα -tocoferol, mientras que las semillas contienen un menor porcentaje de α-tocoferol, siendo γ-tocoferol y δ-tocoferol los más abundantes (Grusak y DellaPenna, 1999). En la Tabla 7 se muestra el contenido en tocoferoles de diferentes aceites y grasas.

Tabla 7. Contenido en tocoferoles (mg/kg) de aceites y grasas (Padley et al., 1994)

	Tocoferoles			
Grasas y aceites	α	β	γ	δ
Colza	202	65	490	9
Ricino	28	29	111	310
Coco	-	-	-	4
Algodón	338	17	429	3
Lino[a,b]	4	-	407	-
Maíz	282	54	1034	54
Oliva	93	-	7	-
Palma[c]	89	-	18	-
Nuez de palma	62	-	-	-
Cacahuete	178	9	213	8
Cártamo	477	-	44	10
Soja	100	8	1021	421
Girasol	670	27	11	-

[a] Velasco y Goffman, 2000.

[b] Contiene también plastocromanol-8.

[c] Contiene también tocotrienoles.

4.2. Biosíntesis de tocoferoles

Las enzimas que participan en la biosíntesis de tocoferoles en plantas se encuentran asociadas a la envoltura plastidial (Schultz, 1990). La cadena lateral isoprenoide deriva de fitil difosfato. Éste a su vez se puede sintetizar por dos vías, a partir del geranilgeranil difosfato (GGDP) procedente de la ruta plastidial de biosíntesis de compuestos isoprenoides (vía 1-deoxi-D-xilulosa-5-fosfato) (DellaPenna y Last, 2006), o bien a partir de fitol libre procedente de la degradación de la clorofila (Valentin y Qungang, 2005). El anillo fenólico procede del ácido homogentísico (HGA) que es sintetizado a partir del 4-hidroxifenilpiruvato, producto del catabolismo de la tirosina, por medio de una enzima citosólica denominada 4-hidroxifenilpiruvato dioxigenasa (Grusak y DellaPenna, 1999; Ajjawi y Shintani, 2004). El HGA se condensa después con fitil difosfato para producir 2-metil-6-fitil-1,4-benzoquinona (MPBQ) en una reacción catalizada por la enzima homogentisato fitiltransferasa. El sustrato MPBQ es metilado por la enzima 2-metil-6-fitil-1,4-benzoquinona/2-metil-6-solanil-1,4-benzoquinona metiltransferasa (MPBQ/MSBQ-MT) para producir 2,3-dimetil-6-fitil-1,4-benzoquinona (DMPBQ). La enzima que convierte estas quinonas en tocoferoles es la tocoferol ciclasa. Así, MPBQ es convertido en δ -tocoferol y DMPBQ da lugar a γ-

tocoferol, que posteriormente son metilados a β-tocoferol y α-tocoferol respectivamente, por medio de la enzima γ-tocoferol metiltransferasa (γ-TMT) (DellaPenna y Pogson, 2006) (Figura 7).

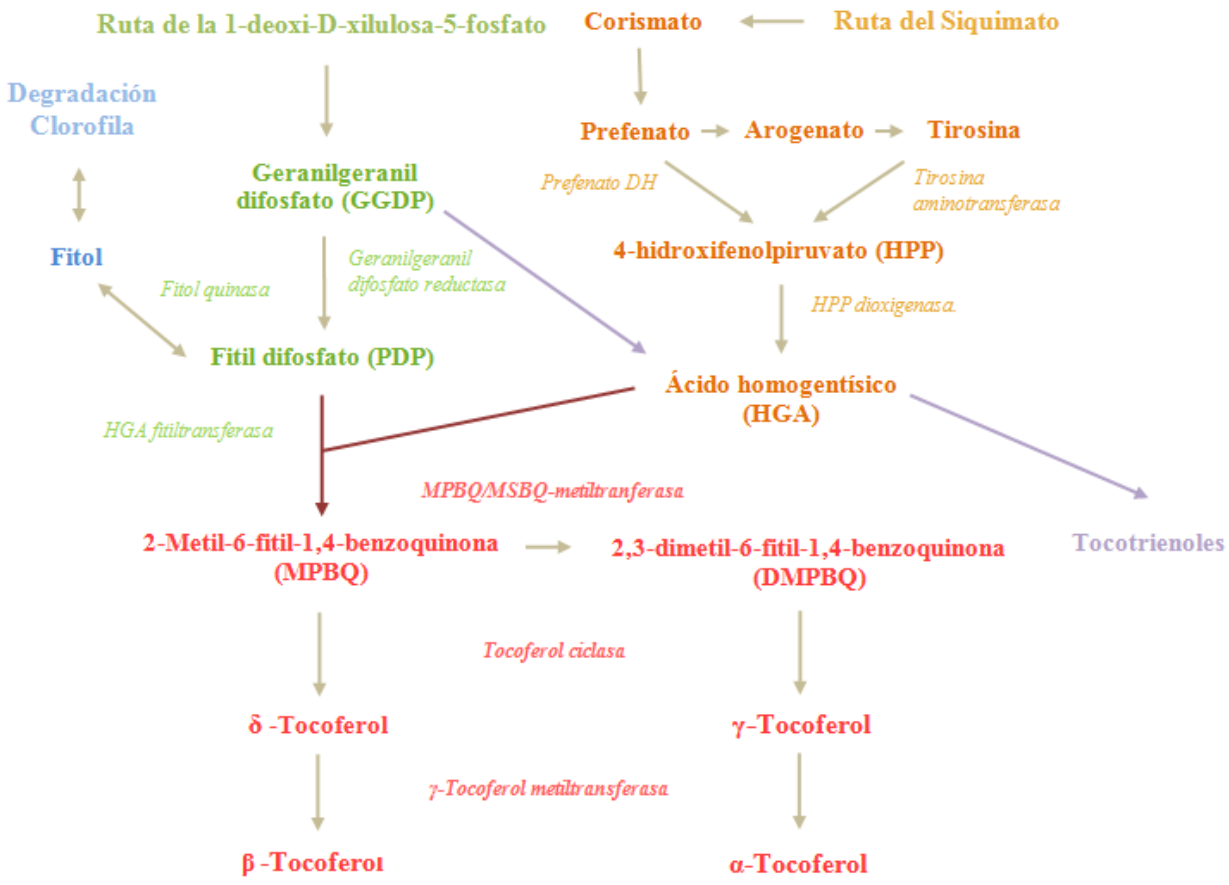


Figura 7: Ruta biosintética de los tocoferoles en el plastidio.

5. Funciones de los tocoferoles

5.1. Actividad antioxidante de los tocoferoles

Se denomina peroxidación lipídica a un conjunto de reacciones de oxidación que afectan a los dobles enlaces de los ácidos grasos presentes tanto en los sistemas biológicos como en los aceites y grasas y en los alimentos que los contienen. *In vivo*, la oxidación de los ácidos grasos insaturados de las membranas provoca daños celulares y cambios patológicos (Halliwell y Gutteridge, 1989; Muggli, 1994). *In vitro*, se

desarrollan olores y sabores desagradables y se generan compuestos nocivos para la salud (Schultz et al., 1962; Aruoma, 1991; Kamal-Eldin y Åppelqvist, 1996). Los compuestos con actividad antioxidante como los tocoferoles inhiben eficazmente estas reacciones (Hess, 1993; Kamal-Eldin y Åppelqvist, 1996). Esta actividad antioxidante de los tocoferoles radica principalmente en la habilidad para donar sus hidrógenos fenólicos a radicales libres lipídicos (Pokorny, 1987; Kamal-Eldin y Åppelqvist, 1996). α-tocoferol presenta el mayor potencial antioxidante *in vivo* (actividad como vitamina E), seguido por β-tocoferol, γ-tocoferol y δ-tocoferol, mientras que *in vitro* γ-tocoferol muestra mejor actividad antioxidante, seguido por δ-tocoferol, β-tocoferol y α-tocoferol (Pongracz et al., 1995). Chow (2000) estimó una actividad biológica relativa de 100% para α-tocoferol, 15 a 27% para β-tocoferol, 3 a 20% para γ-tocoferol y 0.3 a 2% para δ-tocoferol.

El proceso de oxidación lipídica se divide en tres etapas (Kamal-Eldin y Åppelqvist, 1996):

a) Iniciación: se generan radicales alquilo (L•) de un ácido graso insaturado (LH). Son reacciones muy lentas y dependen del tipo de iniciador (*I*). Estas reacciones pueden ser catalizadas por calor, luz, trazas de metales y/o ciertas enzimas.

$$I + LH \rightarrow L\bullet + IH$$

b) Propagación: Los radicales alquilo (L•) formados durante la etapa anterior son muy reactivos y se combinan fácilmente con el oxígeno, originando radicales peróxido (LOO•). Estos radicales peróxido desaparecen mediante reacciones más lentas reaccionando con nuevos LH dando lugar a hidroperóxidos lipídicos (LOOH) y a nuevos radicales alquilo que propagan la reacción en cadena.

$$L\bullet + O_2 \rightarrow LOO\bullet$$

$$LOO\bullet + LH \rightarrow LOOH + L\bullet$$

c) Terminación: Se produce cuando todos los lípidos son consumidos. Los radicales tienden a dimerizarse y finalizará la reacción en cadena.

$$LOO\bullet \rightarrow \text{Productos no radicales}$$
$$LO\bullet \rightarrow \text{Productos no radicales}$$

El mecanismo de acción de los tocoferoles fue propuesto por Nagaoka et al. (1992). La molécula de tocoferol (TOH) cede un átomo de hidrógeno al radical peróxido lipídico formando radicales tocoferoxil (TO•) que se caracterizan por ser muy estables debido a la deslocalización de su electrón libre.

$$LOO\bullet + TOH \rightarrow LOOH + TO\bullet$$

5.2. Funciones de los tocoferoles "in vivo"

5.2.1. La vitamina E en el organismo.

Evans y Bishop (1922) fueron los primeros investigadores que introdujeron el término vitamina E para describir un importante factor en la dieta necesario en la reproducción animal. Posteriormente, esta vitamina se relacionó con ciertas propiedades antioxidantes (Epstein et al., 1966). Desde entonces numerosas investigaciones se han enfocado en el estudio del papel fisiológico de la vitamina E en humanos y animales.

Como se ha mencionado anteriormente, los distintos tocoferoles difieren en su actividad biológica o actividad de vitamina E. α-tocoferol es la forma más común en sangre y tejidos humanos y la que posee más alta actividad biológica. La vitamina E tiene que ser ingerida unida a lípidos en la dieta debido a que el organismo no puede sintetizarla (Brigelius-Flohé et al., 2002). Las principales fuentes de vitamina E son los aceites vegetales y otros productos derivados de las plantas tales como nueces y semillas (Sheppard et al., 1993; Dial y Eitenmiller, 1995). La cantidad diaria recomendada (CDR) de vitamina E en adultos es de 15 mg/día, incluidos en los alimentos, alimentos enriquecidos y suplementos (Food and Nutrition Board, 2000). La detección de bajos niveles de esta vitamina en plasma está asociada con un incremento del riesgo de enfermedades cardiovasculares, algunos cánceres y disminución de la respuesta inmune (Knekt et al., 1994; Kushi et al., 1996; Wright et al., 2006; DellaPenna y Mène-Saffrané, 2011). La dieta en gran parte de Europa es rica en α-tocoferol debido al consumo extendido de aceite de girasol mientras que en EE.UU. se consume principalmente aceite de soja rico en γ-tocoferol (Hass et al., 2006).

La vitamina E ingerida en la dieta junto con lípidos será posteriormente absorbida en el tracto digestivo, donde formará parte de los quilomicrones (lipoproteínas) que serán los encargados de transportar dichos lípidos desde el intestino

hasta el hígado. En el hígado, la distribución de la vitamina E está regulada por la proteína de transferencia deα -tocoferol (α-TTP), que presenta una baja afinidad por el resto de tocoferoles. La mayoría de la vitamina E está presente en el plasma formando parte de las lipoproteínas de baja densidad (LDL) y de alta densidad (HDL). La vitamina E tiende a acumularse en tejidos adiposos (Packer y Obermüler-Jevic, 2002).

La vitamina E actúa como antioxidante lipofílico en lipoproteínas y membranas celulares. El estrés oxidativo provoca cambios en el estado redox normal de los tejidos y produce un aumento de la cantidad de radicales libres (Síes, 1997). Estos radicales libres son continuamente generados durante el metabolismo, pero también se originan como consecuencia de la exposición a estrés medioambiental (luz ultravioleta, contaminación y sustancias químicas, etc.), infecciones por microorganismos, virus, parásitos y durante el envejecimiento (Packer y Obermüler-Jevic, 2002). Los tocoferoles forman parte de las defensas celulares frente al estrés oxidativo protegiendo las lipoproteínas y las membranas celulares frente a la peroxidación lipídica (Burlón y Ingold, 1983). Las funciones antioxidantes de la vitamina E son: 1) destruir los radicales peróxido protegiendo a los ácidos grasos polinsaturados, 2) neutralizar una gran variedad de radicales libres derivados del oxígeno, y otras especies reactivas de oxígeno (ROS) y 3) reaccionar con especies de nitrógeno (Wang y Quinn, 1999). Además, la vitamina E participa en la denominada red antioxidante. Brevemente, el tocoferol al ser oxidado por el radical libre, se transforma en radical tocoferoxil que es muy estable. Este radical puede ser posteriormente regenerado a vitamina E mediante la aceptación de un electrón proveniente de agentes reductores como la vitamina C (DellaPenna y Mène-Saffrané, 2011). Asimismo, la vitamina E actúa como agente estabilizador de la fluidez la membrana celular (Bender, 1992; Lemaire-Ewing et al., 2010).

Además de estas funciones, se han encontrado otras funciones biológicas de la vitamina E que incluyen la modulación de la señalización celular, expresión génica y proliferación celular. Varios estudios indican que la vitamina E actúa como una molécula señal inhibiendo la actividad de la enzima proteína quinasa C en muchos tipos de células (Freedman et al., 1996; Studer et al., 1997). Esta quinasa interviene en procesos de señalización celular, modulando la expresión de genes durante el crecimiento, proliferación y diferenciación celular (Nishizuka, 2001).

Aunque α-tocoferol es la principal forma de vitamina E debido a su fuerte potencial antioxidante *in vivo*, existen estudios que sugieren que γ-tocoferol podría ser igualmente importante en cuanto a salud y protección frente a enfermedades se refiere. En los últimos años se han llevado a cabo investigaciones en animales que asocian γ-tocoferol con ciertas propiedades antioxidantes y anti-inflamatorias (Christen et al., 1997; Jiang et al., 2000, 2002, 2003). γ-tocoferol se encuentra en nueces, semillas y aceites vegetales, y es especialmente abundante en la dieta de Amurica del Norte donde se estima que el 70% de la vitamina E ingerida es en forma de γ-tocoferol, debido al alto consumo de soja y otros aceites vegetales como la colza (McLaughlin et al., 1979; Lehmann et al., 1986). Se ha encontrado que γ-tocoferol neutraliza radicales libres, principalmente óxidos de nitrógeno (Cooney et al., 1993; Christen et al., 1997), reduce la inflamación (Wagner et al., 2004), protege frente a ciertos tipos de cáncer (Stone et al., 2004; Jiang et al., 2004a, 2004b; Weinstein et al., 2005), y activa la expresión de genes involucrados en la protección frente a la enfermedad de Alzheimer (Morris et al., 2005; Rota et al., 2005; Usoro y Mousa, 2010; Mangialasche et al., 2010).

5.2.2. Funciones de los tocoferoles en las plantas

En las plantas, α-tocoferol está localizado en la envoltura del cloroplasto y en las membranas tilacoidales (Munné-Bosch, 2007). Las principales funciones de los tocoferoles, principalmente de α-tocoferol, son:

a) *Regulación del balance de las especies reactivas de oxígeno (ROS) y los antioxidantes.* En plantas, las ROS son continuamente producidas como consecuencia del metabolismo aeróbico, estrés biótico (patógenos) y abiótico (luz, sequía, salinidad y temperaturas extremas). Su función es reducir los niveles de estas especies, principalmente el oxígeno singlete y los radicales hidroxilo, y limitar el grado de peroxidación lipídica en las membranas fotosintéticas (Apel y Hirt, 2004). Además, interviene en el ciclo ascorbato-glutatión interaccionando con otros antioxidantes, por ejemplo la vitamina C, asociados a los fotosistemas para proteger la fotosíntesis (Li et al., 2008).

El efecto antioxidante de los tocoferoles frente a la peroxidación lipídica es especialmente importante durante la germinación, en el establecimiento de plántulas y en la longevidad de la semilla (Sattler et al., 2004, 2006; DellaPenna y Pogson, 2006; Mène-Saffrané y DellaPenna, 2010; DellaPenna y Mène-Saffrané, 2011). Por el

contrario, no se ha publicado ningún estudio que indique que los tocoferoles son esenciales en plantas adultas, debido probablemente a que los tejidos fotosintéticos maduros poseen otras defensas frente a ROS (Porfirova et al., 2002; Maeda et al., 2006; Mène-Saffrané y DellaPenna, 2010; DellaPenna y Mène-Saffrané, 2011).

b) *Transducción de señales celulares:* Se han publicado estudios que indican que los tocoferoles podrían actuar como moléculas señal en las plantas (Munné-Bosch, 2007; Li et al., 2008; DellaPenna y Mène-Saffrané, 2011). Este hecho ocurre a varios niveles: 1) la biosíntesis de tocoferoles podría estar regulada mediante señales ROS y fitohormonas, y viceversa (Norris et al., 1995; Falk et al., 2002; Sandorf y Holländer-Czytko, 2002; Welsch et al., 2003; Havaux et al., 2003), 2) regulación del metabolismo de carbohidratos (DellaPenna y Mène-Saffrané, 2011) y 3) inducción de los mecanismos de señalización celular provocando la muerte celular programada (Evans et al., 2005; Li et al., 2008).

c) *Estabilidad de la fluidez de la membrana celular.* Los tocoferoles pueden estabilizar la estructura de la membrana interaccionando con los ácidos grasos polinsaturados presentes en ella (Wang y Quinn, 1999).

El único papel atribuido específicamente a γ-tocoferol en plantas se ha asociado con sus propiedades antioxidantes. El óxido nítrico (NO) es un gas inestable que tiene varias funciones en animales y plantas (Mène-Saffrané y DellaPennna, 2010). Estudios en animales han demostrado que γ-tocoferol pero no α-tocoferol podrían reaccionar químicamente con NO formando 5-nitroso-γ-tocoferol (Cooney et al., 1993). Basados en estas observaciones, se concluyó que γ-tocoferol podría estar involucrado en la regulación de los niveles de NO en células animales y más recientemente este concepto ha sido extendido a plantas. Un primer estudio analizando diferentes variedades de cebada con diferentes niveles de γ-tocoferol en semillas mostró una correlación inversa entre los niveles de γ-tocoferol y el crecimiento temprano de la raíz (Desel y Krupinska, 2005). Basado en el hecho de que NO estimula la germinación y la elongación de la raíz, se ha especulado que γ-tocoferol podría neutralizar NO en semillas y por tanto retrasar la emergencia de la raíz. Desel y Krupinska (2005) indicaron que las variedades de cebada con valores más altos de NO mostraron bajas cantidades de γ-tocoferol en semillas.

5.3. Funciones de los tocoferoles "in vitro"

El enranciamiento oxidativo de los aceites consiste en la oxidación irreversible de los dobles enlaces presentes en los ácidos grasos insaturados por acción del oxígeno del aire dando lugar a compuestos no radicales, aldehídos y cetonas de bajo peso molecular, responsables del olor a rancio (Mataix y Ochoa, 2002). Los tocoferoles reducen los efectos de la oxidación y contribuyen a la estabilidad del aceite (Seppanen et al., 2010).

En relación con el efecto de los distintos tocoferoles sobre la estabilidad oxidativa del aceite de girasol, Demurin et al. (1996) demostraron que una sustitución parcial de α-tocoferol por otros tocoferoles como β-tocoferol, γ-tocoferol o δ-tocoferol está relacionada con una mejora significativa de la estabilidad oxidativa del aceite. Concretamente, aceites con niveles elevados de γ-tocoferol y ácido oleico mostraron mayor estabilidad oxidativa, comparada con la del aceite de girasol común que está caracterizado por poseer altos niveles de α-tocoferol y ácido linoleico. Existen estudios que indican que la presencia de γ-tocoferol durante los procesos de fritura (Warner et al. 2003) y cuando se emplea en combinación δon -tocoferol (Warner, 2005) mejora significativamente la estabilidad oxidativa del aceite de girasol. Warner et al. (2008) encontraron que los niveles elevados de γ-tocoferol y δ-tocoferol en el aceite "NuSun® medio oleico" provocaron un aumento en la estabilidad oxidativa del mismo comparado con el aceite "NuSun® medio oleico" con bajos niveles de γ-tocoferol. Marmesat et al. (2008) investigaron cómo afecta la temperatura durante los procesos de fritura a dos nuevos tipos de aceites de girasol caracterizados por poseer alto contenido en ácido oleico y ácido palmítico y por contener -tocoferol o γ-tocoferol como antioxidantes naturales más abundantes. Los resultados obtenidos a 120ºC mostraron que el aceite que contenía γ -tocoferol mostró mayor estabilidad oxidativa. Experimentos a alta temperatura (180ºC) simulando los procesos de fritura demostraron que la degradación del aceite y la pérdida de los tocoferoles naturales fueron significativamente menores en el aceite que contenía γ-tocoferol.

Seppanen et al. (2010), sintetizando todos los estudios publicados sobre las funciones antioxidantes relativas de los tocoferoles en aceites, grasas y ciertos alimentos, concluyeron que cuando α-tocoferol actúa a bajas concentraciones en la mayoría de los aceites y grasas analizados mostró mejor actividad antioxidante que -tocoferol, pero a altas concentraciones γ-tocoferol resultó ser un antioxidante más activo

que α-tocoferol. También indicaron que las concentraciones óptimas de los tocoferoles dependen del tipo de grasas y aceites usadas en el experimento en cuestión, concluyendo que los estudios realizados sobre la actividad antioxidante de los tocoferoles en los alimentos son muy variados y no pueden ser uniformemente evaluados.

6. Mejora genética del contenido total y del perfil de tocoferoles en girasol y cártamo

6.1. Objetivos

Tanto los aceites vegetales dirigidos a la alimentación humana (aderezo de ensaladas, frituras, margarinas, etc.) como los empleados con fines no alimentarios (biocombustibles, biolubricantes, fabricación de plásticos, industria cosmética, farmacéutica, etc.) requieren de una elevada estabilidad oxidativa para evitar la pérdida de su calidad. Tradicionalmente se había considerado que la estabilidad oxidativa del aceite estaba determinada casi exclusivamente por su perfil de ácidos grasos. Actualmente se está dando importancia a la presencia en el aceite vegetal de compuestos con actividad antioxidante como los tocoferoles (Fernández-Martínez et al., 2006). Los efectos beneficiosos que provocan los tocoferoles en sistemas biológicos y sobre la estabilidad oxidativa del aceite *in vitro* han hecho que tanto el aumento del contenido total de tocoferoles como la modificación del perfil de tocoferoles mediante sustitución parcial de α -tocoferol por otros tocoferoles sean importantes objetivos para desarrollar aceites con mejores propiedades nutricionales y/o alta estabilidad oxidativa (Fernández-Martínez et al., 2004; Velasco et al., 2004a).

6.2. Variabilidad para el contenido total y perfil de tocoferoles en girasol

Las semillas de girasol contienen principalmente α-tocoferol, que representa más del 90% del total de tocoferoles. β-tocoferol y γ-tocoferol pueden estar presentes pero en cantidades muy bajas (<2%), mientras que δ-tocoferol está ausente (Demurin, 1993; Dolde et al., 1999). Demurin (1993) aisló mediante evaluación de germoplasma de girasol dos líneas con perfiles modificados de tocoferoles: la línea LG-15, con alta concentración de β -tocoferol (50%) y la línea LG-17 con elevada concentración de γ -

tocoferol (95%). Cruzamientos entre estas líneas dieron lugar a la línea LG-24, con 84% de γ-tocoferol y 8% de δ-tocoferol (Demurin, 1993). Velasco et al. (2004b) aislaron las líneas T589 y T2100 con alto contenido en β-tocoferol (>30%) y γ-tocoferol (>85%), respectivamente, a partir de una evaluación de germoplasma. Mediante mutagénesis química e hibridación se han conseguido nuevas líneas con variaciones en el perfil de tocoferoles. Así, Velasco et al. (2004c) mediante un tratamiento con metil sulfonato de etilo (EMS) desarrollaron las líneas IAST-1 e IAST-540 con alto contenido en γ-tocoferol en las semillas (>85%). Cruzamientos entre las líneas IAST-1 y T589 permitieron el desarrollo de las líneas IAST-5 con 30% de α-tocoferol y 69% de β-tocoferol, e IAST-4 con 4% α-tocoferol, 3% β-tocoferol, 34% de γ-tocoferol y 58% de δ-tocoferol.

Otro importante objetivo de mejora es el aumento del contenido total de tocoferoles en semillas de girasol. Velasco et al. (2010) identificaron las líneas IAST-413 e IAST-306 en una colección de germoplasma, caracterizadas por poseer un alto contenido total de tocoferoles en semillas de 467 mg/kg y 463 mg/kg, respectivamente, frente a 251 mg/kg presente en la línea control HA89. Asimismo, en este mismo estudio se desarrolló la línea IAST-522, con contenido reducido en tocoferoles (73 mg/kg).

6.3. Variabilidad para el contenido total y perfil de tocoferoles en cártamo

Al igual que ocurre en girasol, las semillas de cártamo contienen predominantemente α-tocoferol, acumulándose más del 90% del contenido total en tocoferoles (Johnson et al., 1999). En consecuencia, los aspectos más importantes dentro de la mejora genética de tocoferoles en cártamo son el incremento del contenido y la modificación del perfil de tocoferoles (Velasco et al., 2005b). Se ha encontrado muy poca variabilidad en el perfil de tocoferoles en germoplasma de *C. tinctorius*. Johnson et al. (1999) no encontró prácticamente ninguna variabilidad en el perfil de tocoferoles, y Velasco y Fernández-Martínez (2001) sólo identificaron una pequeña variación en el contenido de γ-tocoferol (<10%). En cambio, si se ha encontrado variabilidad para el perfil de tocoferoles en germoplasma de especies silvestres de cártamo (*Carthamus* spp). Velasco et al. (2005a) identificaron entradas de germoplasma de *C. lanatus* subsp. *turkestanicus* y *C. oxyacanthus* con niveles elevados de γ-tocoferol, 9.8% y 36% del total de tocoferoles, respectivamente. La entrada de *C. oxyacanthus* segregó ampliamente tanto para contenido en γ-tocoferol como para caracteres morfológicos, presentando un alto grado

de introgresión de caracteres de *C. tinctorius*. Mediante selección simultánea para alto contenido en γ-tocoferol y caracteres morfológicos se desarrolló la línea IASC-1, con alto contenido en γ-tocoferol (>85%), y tipo de planta próximo a *C. tinctorius*. Varios estudios han identificado variabilidad para el contenido total de tocoferoles en cártamo. Velasco y Fernández-Martínez (2004d) desarrollaron las líneas CR-34 y CR-81 de cártamo con niveles elevados de tocoferoles en semilla, principalmente α -tocoferol. El contenido total de tocoferoles promedio de diferentes ambientes evaluados fue de 679 mg/kg en la línea CR-34, y 718 mg/kg en la línea CR-81, respectivamente, comparado con 488 mg/kg de la línea control Rancho.

6.4. Estudios genéticos del perfil de tocoferoles en girasol y cártamo

En girasol, los estudios de caracterización genética realizados sobre las líneas LG-15 y LG-17 concluyeron que los niveles elevados de β-tocoferol y γ-tocoferol son el resultado de alelos recesivos en los loci *Tph1* y *Tph2*, respectivamente (Demurin et al., 1996). Por otro lado, Velasco y Fernández-Martínez (2003) concluyeron que el alto contenido en γ-tocoferol y β-tocoferol en las líneas T589 y T2100 estaba determinado por alelos recesivos en único gen. Demurin et al. (2004) y Vera-Ruiz et al. (2005) realizaron estudios genéticos comparativos concluyendo que los alelos *tph1* estaban presentes tanto en LG-15 como T589. Demurin et al. (2004) realizaron estudios genéticos comparativos sobre las líneas LG17 y T2100 y concluyeron que el alto contenido en γ-tocoferol en ambas líneas se encuentra controlado por los alelos *tph2*, antes mencionados.

En cártamo, Velasco et al. (2005b) realizaron el estudio genético de la línea mutante natural IASC-1 con alto contenido en γ-tocoferol. Para ello, se realizó el cruzamiento de dicha línea con plantas estériles de la línea CL-1 con contenido estándar en tocoferoles. Las semillas F_1 mostraron un contenido en γ-tocoferol que oscilaba entre 0 y 7.8% indicando que se trata de un carácter parcialmente recesivo. El análisis del contenido en γ-tocoferol en seis poblaciones de semillas F_2 reveló una distribución bimodal para dicho carácter, con tres cuartas partes de las semillas mostrando un contenido de γ-tocoferol que osciló entre 0 y 8.5%, y una cuarta parte de las semillas con un alto contenido de γ-tocoferol oscilando entre 75.7 y 97.0% indicando segregación de un gen parcialmente recesivo que se denominó *Tph2*, al igual que en

girasol. El examen de las familias $F_{2:3}$ confirmó el control monogénico del carácter alto contenido en γ-tocoferol.

6.5. Genes modificadores

Tanto en girasol como en cártamo se han publicado estudios fenotípicos que indican que determinados caracteres pueden verse afectados por la presencia de uno o varios genes modificadores. Dichos genes modificadores se definen como genes que no tienen efecto conocido, salvo intensificar o disminuir la expresión de un gen con efecto mayor (Briggs y Knowles, 1967).

En girasol, se ha sugerido la presencia de genes modificadores que afectan al alto contenido en ácido oleico originando una fuerte distorsión en los patrones de segregación (Urie, 1985; Fernández-Martínez et al., 1989; Velasco et al., 2000) o incluso provocando una reversión completa del carácter alto oleico (Lacombe et al., 2001). También se ha publicado la existencia de genes modificadores que afectan a genes mayores de resistencia a jopo (*Orobanche cumana* Wallr.) (Velasco et al., 2007).

En cártamo, Knowles (1972) postuló la posible existencia de genes modificadores que afectan al alto contenido en ácido oleico sugiriendo que los valores no esperados de alto contenido en ácido oleico de algunos cruces estudiados podrían explicarse por la presencia de uno o más genes modificadores. Hamdan et al. (2009) demostraron la presencia de genes modificadores responsables de efectos positivos y negativos sobre el alto contenido en ácido oleico.

7. Estudios moleculares en girasol y cártamo

Los programas de mejora de los cultivos oleaginosos más importantes se apoyan en el desarrollo y uso de marcadores moleculares. En el caso del girasol, estas herramientas moleculares comenzaron a desarrollarse tempranamente por lo que en la actualidad existe un gran número de estudios y de marcadores moleculares disponibles en programas de selección de este cultivo. Por el contrario, las investigaciones moleculares en cártamo, especie filogenéticamente cercana a girasol, son más recientes por lo que actualmente los recursos son limitados.

7.1. Marcadores moleculares y mapas genéticos en girasol

En girasol se han desarrollado y empleado diferentes tipos de marcadores de ADN.

1) Marcadores de ADN anónimos o neutros. Derivan de regiones polimórficas sin función conocida, amplificadas a lo largo de un genoma. Dentro de ellos se han desarrollado y empleado cuatro tipos de marcadores:

a) Marcadores RAPD (Random Amplified Polymorphic DNA). Este tipo de marcadores se usó en los primeros estudios moleculares de girasol. Los principales inconvenientes son su naturaleza dominante y baja reproducibilidad. Se emplearon en análisis de grupos segregantes uniformes (BSA o Bulked Segregant Analysis) para la identificación de marcadores asociados con la resistencia a roya (Lawson et al., 1996) y a jopo (Lu et al., 2000). La principal ventaja de este tipo de marcadores es su posible conversión en marcadores SCAR (Sequence Characterized Amplified Region), altamente específicos (Lawson et al., 1998; Lu et al., 2000). Asimismo, se han empleado en la construcción de mapas genéticos de girasol en etapas iniciales, principalmente de las especies silvestres *H. anomalus* (Rieseberg et al., 1993) y *H. annuus* y *H. petiolaris* (Rieseberg et al., 1995).

b) Marcadores AFLP (Amplified Fragment Length Polymorphism). Aunque se trata de marcadores de tipo dominante, su alta reproducibilidad hace que sean muy adecuados para el mapeo genético (Peerbolte y Peleman, 1996), estudios de variabilidad y estudio de la huella de ADN (fingerprint) (Hongtrakul et al., 1997). Se han desarrollado diferentes mapas genéticos basados en combinaciones de marcadores AFLP con otros tipos de marcadores empleando poblaciones F_2 (Peerbolte y Peleman, 1996; Gedil et al., 1999) o poblaciones de líneas puras recombinantes (Recombinant inbred line, RIL) (Flores-Berrios et al., 2000; Rachid Al-Chaarani et al., 2002; Langar et al., 2003). Estos mapas presentan un alto grado de saturación y cobertura del genoma.

c) Marcadores RFLP (Restriction Fragment Length Polymorphism). Estos marcadores se desarrollaron a partir de clones anónimos de cDNA de girasol, con un bajo número de copias y con presencia de sitios de restricción polimórficos (Berry et al., 1994; Gentzbittel et al., 1994; Jan et al., 1993). Se han

construido mapas de ligamiento empleando estos marcadores en poblaciones F_2 (Berry et al., 1995, 1996; Jan et al., 1998; Gentzbittel et al., 1999) o en una combinación de poblaciones F_2 y poblaciones BC_1F_1 (Gentzbittel et al., 1995).

d) Marcadores SSR (Simple Sequence Repeats). También denominados microsatélites, consisten en repeticiones de motivos de nucleótidos simples, muy abundantes en genomas eucarióticos (Tóth et al., 2000). Las principales ventajas de estos marcadores son su abundancia, altos niveles de polimorfismo, naturaleza multialélica, herencia codominante, amplia dispersión en genomas, facilidad de ensayo empleando PCR y fácil difusión entre laboratorios (Powell et al., 1996). En girasol se dispone de un elevado número de marcadores SSR (Gedil, 1999; Paniego et al., 2002; Yu et al., 2002; Tang et al., 2002). Los primeros mapas de SSR de girasol fueron construidos por Tang et al. (2002) usando RIL, y por Yu et al. (2003) empleando poblaciones F_2. Tang et al. (2003), basándose en estos dos mapas, desarrollaron un mapa de ligamiento compuesto, que permitió la selección de un esqueleto de 95 marcadores SSR, distribuidos uniformemente a lo largo de los 17 grupos de ligamiento de girasol y que en la actualidad se emplea como una primera aproximación a la exploración del genoma de girasol**.** Además, los marcadores SSR han permitido la saturación de mapas de ligamiento construidos con otros tipos de marcadores (Lai et al., 2005; Rachid Al-Chaarini et al., 2004; Paniego et al., 2007).

2) Marcadores basados en genes. También llamados marcadores de genes candidatos, se caracterizan por derivar de sitios polimórficos en genes. En girasol se han desarrollado los siguientes:

a) Marcadores basados en sondas de RFLP-cDNA secuenciadas. A partir 81 marcadores RFLP de girasol se desarrollaron marcadores INDEL (Insertion/Delection) mediante la secuenciación de clones de cDNA y posterior alineamiento con secuencias de ADN genómicas de *Arabidopsis*. Sobre estas secuencias se diseñaron parejas de cebadores que flanquean intrones, para amplificar estas regiones polimórficas (Yu et al., 2003). Estos marcadores se integraron en el mapa de SSR desarrollado por Tang et al. (2002) (Yu et al., 2003).

2) Marcadores basados en EST (Expressed Sequence Tags). Las EST son secuencias derivadas de cDNA sin editar y procesadas automáticamente, asociadas con regiones codificantes del genoma. Existen diferentes programas de secuenciación de EST de girasol (Fernández et al., 2003; Tamborindeguy et al., 2004; Ben et al., 2005) entre los que destaca el Compositae Genome Project (http://compgenomics.ucdavis.edu/). En girasol, esta base de datos de EST se ha empleado principalmente para desarrollar marcadores SNP (Single Nucleotide Polymorphism), SSR (Pashley et al., 2006), y SNP/INDEL (Lai et al., 2005). En girasol, las EST han permitido el desarrollo de marcadores moleculares útiles en mapeo comparativo y análisis filogenéticos en la familia *Asteraceae* (Chapman et al., 2007).

3) Marcadores funcionales. Derivan de sitios polimórficos de genes conocidos que se encuentran asociados a variaciones de un determinado carácter fenotípico. En girasol se han desarrollado marcadores funcionales para caracteres que determinan la calidad del aceite (Tang et al., 2006; Hass et al., 2006) o resistencia a herbicidas (Kolkman et al., 2004). En el apartado "Mejora molecular de caracteres de calidad del aceite de girasol y cártamo" de la presente tesis se describe con más detalle este tipo de marcadores y su funcionalidad.

Recientemente, Kane et al. (2011), utilizando técnicas de secuenciación masiva, han conseguido importantes progresos en la secuenciación del genoma en girasol. Además, este estudio ha permitido el desarrollo de un mapa genético y un mapa físico con una alta densidad de marcadores. El mapa físico y el genético representan aproximadamente el 85% y el 80% del genoma de girasol, respectivamente. Los análisis preliminares indican que el 78% del genoma de girasol está constituido por secuencias repetitivas. La disponibilidad de la secuencia del genoma de girasol permitirá realizar estudios de resecuenciación y potenciará la utilización de EST de girasol, revelando localizaciones de genes, el tamaño de familias génicas, secuencias de promotores e intrones, relaciones entre genes parálogos y splicing alternativo (Kane et al., 2011).

7.2. Marcadores moleculares y mapas genéticos en cártamo

Los marcadores moleculares empleados inicialmente en este cultivo se usaron para realizar estudios sobre diversidad genética y fueron RAPD, ISSR (Inter-Simple Sequence Repeat) y AFLP (Patzak, 2001). En la actualidad se siguen dos estrategias

para potenciar el desarrollo de marcadores disponibles en cártamo. La primera estrategia consiste en la transferencia de marcadores moleculares a cártamo desde especies relacionadas, como el girasol (Heesacker et al., 2008; Chapman et al., 2007). Heesacker et al. (2008) estudiaron la transferencia a cártamo de marcadores SSR e INDEL desarrollados a partir de EST de girasol encontrando que de los 466 marcadores SSR-EST o INDEL-EST probados, 67 (14.4%) amplificaron alelos en cártamo. Por otro lado, Chapman et al. (2007) desarrollaron una serie de marcadores universales basados en secuencias conservadas de girasol, lechuga y *Arabidopsis* para estudios filogenéticos y para el desarrollo de mapas genéticos en la familia *Asteraceae*. Algunos de estos marcadores amplificaron con éxito en cártamo. La segunda estrategia tiene como objetivo el desarrollo de nuevos marcadores SSR más específicos de cártamo. Así, Chapman et al. (2009) desarrollaron 104 marcadores SSR basados en EST de cártamo. Naresh et al. (2009) desarrollaron otros cinco marcadores SSR basados en EST que resultaron ser muy útiles en la identificación de híbridos de cártamo. Finalmente, Mayerhofer et al. (2008, 2010) desarrollaron 1412 SSR y 75 RFLP a partir de librerías genómicas y de cDNA y llevaron a cabo el primer análisis de ligamiento en el género *Carthamus*. Estos investigadores desarrollaron los dos primeros mapas genéticos realizados en esta especie a partir de una población F_2 intraespecífica de *C. tinctorius* y de otra población F_2 interespecífica obtenida del cruce entre *C. tinctorius* y *C. oxyacanthus*. Para ello caracterizaron y mapearon 153 marcadores SSR basados en EST y 32 marcadores SSR polimórficos generados a partir de librerías genómicas y de cDNA. Los mapas genéticos de ambas poblaciones están formados por 13 grupos de ligamiento con una longitud total de 580 y 954 cM, respectivamente.

7.3. Mejora molecular de caracteres de calidad del aceite en girasol y cártamo.

El uso de herramientas moleculares en la mejora genética de la calidad del aceite de girasol y cártamo se ha centrado en caracteres asociados a modificaciones en la composición en ácidos grasos y tocoferoles. Las técnicas empleadas han sido el análisis BSA, QTL y aproximaciones de genes candidatos (Pérez-Vich y Berry, 2010).

7.3.1. Ácidos grasos

Se ha descrito el desarrollo de marcadores moleculares asociados a genes mayores que determinan un alto contenido en ácido esteárico en el aceite y se ha procedido a su mapeo en el mapa genético de girasol. Pérez-Vich et al. (2002, 2006) identificaron marcadores moleculares RFLP y SSR asociados con el carácter alto contenido en ácido esteárico y describieron que los genes *Es1* y *Es3* implicados en este carácter mapean en los grupos de ligamiento 1 y 8, respectivamente. Asimismo estos autores identificaron mediante una aproximación de genes candidatos que el gen *Es1* corresponde a un locus estearato desaturasa que codifica la enzima responsable de la desaturación de ácido esteárico a ácido oleico (Pérez-Vich et al., 2002). Por otro lado, el análisis QTL ha permitido la caracterización de genes menores que provocan un incremento moderado en el contenido en ácido esteárico. De esta manera, se han identificado los QTL *st2.1*, *st2.2* y *st2.3* con un efecto menor sobre el contenido en este ácido graso en los grupos de ligamiento 3, 7 y 13, respectivamente, del mapa genético de girasol (Pérez-Vich et al., 2004).

El carácter alto oleico también ha sido objeto de estudio en girasol. Las primeras investigaciones publicadas permitieron la identificación de marcadores RAPD asociados con el gen *Ol1* responsable en gran medida de este carácter (Dehmer y Friedt, 1998). Posteriormente, se demostró que dicho gen *Ol1* cosegrega con un locus oleato desaturasa (*FAD2-1*), que codifica la enzima que determina la desaturación de ácido oleico a ácido linoleico (Hongtrakul et al., 1998; Lacombe et al., 2001; Martínez-Rivas et al., 2001; Pérez-Vich et al., 2002), y se caracterizaron a nivel molecular genes modificadores que alteraban la expresión del gen *Ol1* (Lacombe et al., 2001, 2002; Pérez-Vich et al., 2002). Pérez-Vich et al. (2002) describieron la existencia de un QTL menor en el grupo de ligamiento 8 que mostró un efecto epistático con el QTL mayor para el ácido oleico en el locus *FAD2-1* del grupo de ligamiento 14. Lacombe et al. (2001, 2002) identificaron un locus que suprimió el efecto del locus *FAD2-1*.

En cártamo, se ha realizado un estudio molecular empleando un análisis BSA que ha permitido la identificación de marcadores RAPD asociados a los genes *Li* y *Ms*, responsables del contenido elevado en ácido linoleico y de la androesterilidad nuclear, respectivamente. Estos marcadores RAPD se convirtieron en marcadores SCAR y se construyó un grupo de ligamiento que incluyó cinco marcadores SCAR y los genes *Li* y

Ms. La distancia entre ambos genes es de 11.8 cM y se encuentran flanqueados por los marcadores SCAR IASCA37 e IASCA39, a distancias mínimas de 15.7 cM de *Li* y 3.7 cM de *Ms* (Hamdan et al., 2008).

7.3.2. Tocoferoles

Se han identificado tres loci (*m=Tph1*, *g=Tph2* y *d),* implicados en la síntesis de α-tocoferol, que producen nuevos perfiles de tocoferoles en semillas de girasol (Demurin, 1993; Demurin et al., 1996; Hass et al., 2006). Los loci *m* y *d* se relacionan con la enzima 2-metil-6-fitil-1,4-benzoquinona/2-metil-6-solanil-1,4-benzoquinona metiltransferasa (*MPBQ/MSBQ-MT*) (Tang et al., 2006). El locus *g* se relaciona con la enzima γ-tocoferol metiltransferasa (*γ-TMT*) (Hass et al., 2006).

Tang et al. (2006) identificaron dos genes parálogos *MPBQ/MSBQ-MT* (*MT-1* y *MT-2*) relacionados con las mutaciones *m* y *d*, respectivamente. La mutación *m* incrementó el porcentaje de β-tocoferol en semillas de girasol. *MT-1* cosegregó con la mutación *m* (Tang et al., 2006; Vera-Ruiz et al., 2006) mientras que *MT-2* cosegregó con la mutación *d* (Tang et al., 2006). Ambos loci fueron mapeados en los grupos de ligamiento 1 y 4, respectivamente (Hass et al., 2006; Tang et al., 2006). *MT-2* mostró epistasia con el locus *MT-1*. La mutación *d* incrementó significativamente los porcentajes de β-tocoferol en homocigotos mutantes *mm*. El análisis molecular reveló que la mutación *m* era causada por la inserción de un retrotransposón de tipo *Ty3/gypsy* (5175 bp) en el exón 1 del gen *MT-1*.

Hass et al. (2006) mapearon el gen *Tph2* que determina alto contenido en γ-tocoferol en semillas de girasol en el grupo de ligamiento 8 del mapa genético de esta especie. Estos investigadores aislaron y caracterizaron dos genes parálogos que codifican la enzima γ-tocoferol metiltransferasa (*γ-TMT-1* y *γ-TMT-2*). Ambos genes fueron mapeados en el grupo de ligamiento 8 y cosegregaron con el locus *Tph2*. El análisis de las secuencias de ADN genómico en cinco líneas de girasol reveló la existencia de cinco haplotipos distintos. La mutación *g* disminuyó la transcripción del gen parálogo *γ-TMT-1* e interrumpió la transcripción del gen parálogo *γ-TMT-2*, originó la pérdida de un sitio de inicio de transcripción y provocó splicing alternativo. Estos efectos son característicos de mutaciones provocadas por la inserción de transposones en intrones y secuencias reguladoras. En este sentido, estos investigadores encontraron

varios fragmentos homólogos a transposones de tipo I y II en los intrones de *γ-TMT*, pero ninguno fue asociado con el fenotipo mutante.

Hasta la fecha, no se han publicado estudios moleculares sobre la composición de tocoferoles en cártamo.

Referencias

Ajjawi I y Shintani D (2004) Engineered plants with elevated vitamin E: a nutraceutical success story. Trends Biotechnol. 22:104-107.

Alba OA y Llanos M (1989) Origen, expansión e importancia actual del girasol en el mundo. p. 13-19. *En*: El cultivo del girasol. Mundi-Prensa. Madrid.

Apel K y Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. 55:373-99.

Åppelqvist LA (1989) The chemical nature of vegetables oils. p. 22-37. *En*: RK Downey et al. (eds.). Oil crops of the world. McGraw-Hill. EE.UU.

Aruoma OI (1991) Prooxidant properties: an important consideration for food additives and/or nutrient components. p. 173-194. *En:* OI Aruoma y B Halliwell (eds.). Free radicals and food additives. Taylor y Francis. London.

Ashri A y Knowles PF (1960) Cytogenetics of safflower (*Carthamus* L.) species and their hybrids. Agron. J. 52:11-17.

Ben C, Hewezi T, Jardinaud M, Bena F, Ladouce N, Moretti S, Tamborindeguy C, Liboz T, Petitprez M y Gentzbittel L (2005) Comparative analysis of early embryonic sunflower cDNA libraries. Plant Mol. Biol. 57:255-270.

Bender DA (1992) Vitamin E: tocopherols and tocotrienols. p. 87-105. *En*: Nutritional biochemistry of the vitamins. Cambridge University Press.

Berry ST, Leon AJ, Challis P, Livini C, Jones R, Hanfrey CC, Griffiths S y Roberts A (1996) Construction of a high density, composite RFLP linkage map for cultivated sunflower (*Helianthus annuus* L.). p. 1155-1160. *En*: Proc. 14th Int. Sunflower Conf., Pekín, China.

Berry ST, Leon AJ, Hanfrey CC, Challis P, Burkholz A, Barnes SR, Rufener GK, Lee M y Caligari PDS (1995) Molecular-marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 91:195-199.

Berry ST, Allen RJ, Barnes SR y Caligari PDS (1994) Molecular-marker analysis of *Helianthus annuus* L. 1. Restriction fragment polymorphism between inbred lines of cultivated sunflower. Theor. Appl. Genet. 89:435-441.

Bowles VG, Mayerhofer R, Davis C, Good AG y Hall JC (2010) A phylogenetic investigation of *Carthamus* combining sequence and microsatellite data. Plant Syst. Evol. 287:85-97.

Brigelius-Flohé R, Kluth D, Landes N, Pfluger P y Birringer M (2002) Mechanisms of vitamin E metabolism. p. 171-179. *En*: L Packer et al. (eds.). The antioxidant vitamins C and E. AOCS Press, California.

Briggs FD y Knowles PF (1967) Introduction to plant breeding. Reinhold Publishing Corporation, Nueva York.

Burke JM, Tang S, Knapp S y Rieseberg LH (2002) Genetic analysis of sunflower domestication. Genetics 162:1257-1267.

Burlón G e Ingold K (1983) Is vitamin E the only lipid-soluble, chain-breaking antioxidant in human blood plasma and erythrocyte membranes? Arch. Biochem. Biophys. 221:281-290.

Caretto S, Nisi1 R, Paradiso A y De Gara L (2010) Tocopherol production in plant cell cultures. Mol. Nutr. Food Res. 54:726-730.

Chapman MA, Hvala J, Strever J y Burke JM (2010) Population genetic analysis of safflower (*Carthamus tinctorius*; *Asteraceae*) reveals a near eastern origin and five centers of diversity. Am. J. Bot. 97:831-840.

Chapman MA, Hvala J, Strever J, Matvienko M, Kozik A, Michelmore RW, Tang S, Knapp SJ y Burke JM (2009) Development, polymorphism, and cross-taxon utility of EST-SSR markers from safflower (*Carthamus tinctorius* L.). Theor. Appl. Genet. 120, 85-91.

Chapman MA, Chang JC, Weisman D, Kesseli RV y Burke JM (2007). Universal markers for comparative mapping and phylogenetic analysis in the *Asteraceae* (Compositae). Theor. Appl. Genet. 115:747-755.

Chapman MA y Burke JM (2007) DNA sequence diversity and the origin of cultivated safflower (*Carthamus tinctorius* L.; *Asteraceae*). BMC Plant Biol. 7:60.

Chow CK (2000) Vitamin E. p. 584-598. *En*: MH Stipanuk (ed.). Biochemical and physiological aspects of human nutrition. Saunders, PA, USA.

Christen S, Woodall AA, Shigenaga MK, Southwell-Keely PT, Duncan MW y Ames BN (1997) γ-Tocopherol traps mutagenic electrophiles such as NOx and complements α-tocopherol: physiological implications. Proc. Natl. Acad. Sci. U. S. A. 94:3217-3222.

Cooney RV, Franke AA, Harwood PJ, Hatch-Pigott V, Custer LJ y Mordan LJ (1993) Gamma-tocopherol detoxification of nitrogen dioxide: superiority to alpha-tocopherol. Proc. Natl. Acad. Sci. U.S.A. 90:1771-1775.

Crites GD (1993) Domesticated sunflower in fifth millennium b.p. temporal context: new evidence from middle Tennessee. Am. Antiq. 58:146-148.

Dajue L y Mündel HH (1996) Safflower, *Carthamus tinctorius* L.: Promoting the conservation and use of underutilized and neglected crops. IPK, Gatersleben, Alemania y IPGRI, Roma, Italia.

Dehmer KJ y Friedt W (1998) Development of molecular markers for high oleic acid content in sunflower. Ind. Crop. Prod. 7:311-315.

DellaPenna D y Mène-Saffrané L (2011) Vitamin E. p. 180-216. *En*: JC Kader et al. (eds.). Advances in botanical research, Vol. 59, Elsevier Ltd.

DellaPenna D y Pogson BJ (2006) Vitamin synthesis in plants: tocopherols and carotenoids. Annu. Rev. Plant Biol. 57:711-738.

DellaPenna D y Last RL (2006) Progress in the dissection and manipulation of plant vitamin E biosynthesis. Physiol. Plant 126:356-368.

Demurin Y, Efimenko SG y Peretyagina TM (2004) Genetic identification of tocopherol mutations in sunflower. Helia 27:113-116.

Demurin Y, Škorić D y Karlovic D (1996) Genetic variability of tocopherol composition in sunflower seeds as a basis of breeding for improved oil quality. Plant Breed. 115:33-36.

Demurin Y (1993) Genetic variability of tocopherol composition in sunflower seeds. Helia 16:59-62.

Desel C y Krupinska K (2005) The impact of tocochromanols on early seedling development and NO release. J. Plant Physiol. 162:771-776.

Dial S y Eitenmiller R (1995) Tocopherols and tocotrienols in key foods in the U.S. diet. p. 327-342. *En*: A Ong et al. (eds.).Nutrition, lipids, health, and disease. AOCS Press. Champaign, IL.

Dolde D, Vlahakis C y Hazebroek J (1999) Tocopherols in breeding lines and effects of planting location, fatty acid composition, and temperature during development. J. Am. Oil Chem. Soc. 76:349-355.

Epstein SS, Forsyth J, Saporoschetz IB y Mantel N (1966) An exploratory investigation on the inhibition of selected photosensitizers by agents of varying antioxidant activity. Radiat. Res. 28:322-335.

Evans HM y Bishop KS (1922) On the existence of a hitherto unrecognized dietary factor essential for reproduction. Science 56:650.

Evans NH, McAinsh MR, Hetherington AM y Knight MR (2005) ROS perception in *Arabidopsis thaliana*: The ozone-induced calcium response. Plant J. 41:615-626.

Falk J, Krauß N, Dähnhardt D y Krupinska K (2002) The senescence associated gene of barley encoding 4-hydroxyphenylpyruvate dioxygenase is expressed during oxidative stress. J. Plant Physiol. 159:1245-1253.

FAOSTAT (2011) Database of the food and agriculture organization of the United Nations (FAO). http://www. fao.org.

Fernández-Martínez JM, Pérez-Vich B y Velasco L (2009) Sunflower. p. 155-232. *En*: J Vollmann e I Rajcan (eds.). Oil crops. Handbook of plant breeding. Springer.

Fernández-Martínez JM, Pérez-Vich B y Velasco L (2006) Mejora de la calidad del girasol. p. 449-471. *En*: G Llácer et al. (eds.). Mejora Genética de la Calidad en Plantas. Universidad Politécnica de Valencia.

Fernández-Martínez JM, Velasco L y Pérez-Vich B (2004) Progress in the genetic modification of sunflower oil quality. p. 1-14. *En:* Proc. 16th Int. Sunflower Conf., Fargo, ND, USA.

Fernández-Martínez J (1997). Update on safflower genetic improvement and germplasm resources. p. 187-195. *En:* A Corleto y HH Mundel (eds.). Proc. 4th Int. Safflower Conf., Bari, Italia.

Fernández-Martínez JM, Jiménez A, Domínguez J, García JM, Garcés R y Mancha M (1989) Genetic analysis of the high oleic content in cultivated sunflower (*Helianthus annuus* L.). Euphytica 41:39-51.

Fernández P, Paniego N, Lew S, Hopp H y Heinz RA (2003) Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project. BMC Genomics 4:40.

Fick GN (1989) Sunflower. p. 22-37. *En*: RK Downey et al. (eds.). Oil crops of the world. McGraw-Hill, EE.UU.

Flores-Berrios E, Gentzbittel L, Kayyal H, Alibert G y Sarrafi A (2000) AFLP mapping of QTLs for in vitro organogenesis traits using recombinant inbred lines in sunflower (*Helianthus annuus* L.). Theor. Appl. Genet. 101:1299-1306.

Food and Nutrition Board (2000) Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids. p. 186-283. National Academy Press, Washington, DC.

Freedman JE, Farhat JH, Loscalzo J y Keaney JF Jr (1996) α-Tocopherol inhibits aggregation of human platelets by a protein kinase C-dependent mechanism. Circulation 94:2434-2440.

Gedil MA (1999) Marker development, genome mapping, and cloning of candidate disease resistance genes in sunflower, *Helianthus annuus* L. Ph.D. Thesis, Oregon State Univ., Corvallis, OR, USA.

Gentzbittel L, Mestries E, Mouzeyar S, Mazeyrat F, Badaoul S, Vear F, Tourvieille de Labrouhe D y Nicolas P (1999) A composite map of expressed sequences and phenotypic traits of the sunflower (*Helianthus annuus* L.) genome. Theor. Appl. Genet. 99:218-234.

Gentzbittel L, Vear F, Zhang YX y Bervillé A (1995) Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annuus* L.). Theor. Appl. Genet. 90:1079-1086.

Gentzbittel L, Zhang YX, Vear F, Griveau B y Nicolas P (1994) RFLP studies of genetic relationships among inbred lines of the cultivated sunflower, *Helianthus annuus* L.: evidence for distinct restorer and maintainer germplasm pools. Theor. Appl. Genet. 89:419-425.

Grusak MA y DellaPenna D (1999) Improving the nutrient composition of plants to enhance human nutrition and health. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50:133-161.

Halliwell B y Gutteridge JMC (1989) Free radicals in biology and medicine. Oxford University Press, Oxford.

Hamdan YAS, Pérez-Vich B, Velasco L y Fernández-Martínez JM (2009) Inheritance of high oleic acid content in safflower. Euphytica 168:61-69.

Hamdan YAS, Velasco L y Pérez-Vich B (2008) Development of SCAR markers linked to male sterility and very high linoleic acid content in safflower. Mol. Breed. 22:385-393.

Hanelt P (1963) Monographische Übersicht der gattung *Carthamus* L. (*Compositae*). Feddes Repert. 67:41-180.

Hanelt P (1961) Information on *Carthamus tinctorius* L. Die Kulturpflanze. 9:114-145.

Harter AV, Gardner KA, Falush D, Lentz DL, Bye RA y Rieseberg LH (2004) Origin of extant domesticated sunflowers in eastern North America. Nature 430:201-205.

Hass CG, Tang S, Leonard S, Traber M, Miller J y Knapp S (2006) Three non-allelic epistatically interacting methyltransferase mutations produce novel tocopherol (vitamin E) profiles in sunflower. Theor. Appl. Genet. 113:767-782.

Havaux M, Lütz C y Grimm B (2003) Chloroplast membrane stability in *chlP* transgenic tobacco plants deficient in tocopherols. Plant Physiol. 132:300-310.

Heesacker A, Kishore VK, Gao W, Tang S, Kolkman JM, Gingle A, Matvienko M, Kozik A, Michelmore RM, Lai Z, Rieseberg LH y Knapp SJ (2008) SSRs and INDELs mined from the sunflower EST database: abundance, polymorphisms, and crosstaxa utility. Theor. Appl. Genet. 117:1021-1029.

Heiser CB, Smith DM, Clevenger SB y Martin WC (1969) The North American sunflowers (*Helianthus*). Mem. Torrey Bot. Club 22:1-218.

Heiser CB (1954) Variation and subspeciation in the common sunflower, *Helianthus annuus*. Am. Midl. Nat. 51:287-305.

Hess JL (1993) Vitamin E: α-Tocopherol, in antioxidants in higher plants. p. 111-134 *En*: RG Alscher y JL Hess (eds.). CRC Press, Boca Ratón.

Hongtrakul V, Slabaugh MB y Knapp SJ (1998) A seed specific Δ-12 oleate desaturase gene is duplicated, rearranged and weakly expressed in high oleic acid sunflower lines. Crop Sci. 38:1245-1249.

Hongtrakul V, Huestis GM y Knapp SJ (1997) Amplified fragment length polymorphism as a tool for DNA fingerprinting sunflower germplasm: genetic diversity among oilseed inbred lines. Theor. Appl. Genet. 95:400-407.

Jan CC y Seiler G (2007) Sunflower. p. 103-165. *En*: RJ Singh (ed.). Oilseed crops, genetic resources, chromosome engineering, and crop improvement. Vol. 4. CRC Press, Boca Raton, FL, USA.

Jan CC, Vick BA, Miller JF, Kahler AL y Butler ET (1998) Construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 96:15-22.

Jan CC, Vick BA, Miller JF, Kahler AL y Butler ET (1993) Progress in the development of a genomic RFLP map of cultivated sunflower (*Helianthus annuus*). p. 125-128. *En*: Proc. 15th Sunflower Res. Workshop. Natl. Sunflower Assoc. Bismarck, ND, USA.

Jiang Q, Wong J, Fyrst H, Saba JD y Ames BN (2004a) γ-Tocopherol, or combinations of vitamin E forms, induce cell death in human prostate cancer cells by interrupting sphingolipid synthesis. Proc. Natl. Acad. Sci. U.S.A. 101:17825-17830.

Jiang Q, Wong J y Ames BN (2004b) Gamma-tocopherol induces apoptosis in androgen-responsive LNCaP prostate cancer cells via caspase-dependent and independent mechanisms. Ann. NY Acad. Sci. 1031:399-400.

Jiang Q y Ames BN (2003) Gamma-tocopherol, but not alpha-tocopherol, decreases proinflammatory eicosanoids and inflammation damage in rats. Faseb J. 17:816-822.

Jiang Q, Lykkesfeldt J, Shigenaga MK, Shigeno ET, Christen S y Ames BN (2002) Gamma-tocopherol supplementation inhibits protein nitration and ascorbate oxidation in rats with inflammation. Free Radic. Biol. Med. 33:1534-1542.

Jiang Q, Elson-Schwab I, Courtemanche C y Ames BN (2000) Gamma-tocopherol and its major metabolite, in contrast to alpha-tocopherol, inhibit cyclooxygenase activity in macrophages and epithelial cells. Proc. Natl. Acad. Sci. U. S. A. 97:11494-11499.

Johnson RC, Bergman JW y Flynn CR (1999) Oil and meal characteristics of core and non-core safflower accessions from the USDA collection. Genet. Resour. Crop Evol. 46:611-618.

Kamal-Eldin A y Åppelqvist LÅ (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31:671-701.

Kane NC, Gill N, King MG, Bowers JE, Berges H, Gouzy J, Bachlava E, Langlade NB, Lai Z, Stewart M, Burke JM, Vincourt P, Knapp SJ y Rieseberg LH (2011) Progress towards a reference genome for sunflower. Botany 89:429-437.

Knekt P, Reunanen A, Jarvinen R, Seppanen R, Heliovaara M y Aromaa A (1994) Antioxidant vitamin intake and coronary mortality in a longitudinal population study. Am. J. Epidemiol. 139:1180-1189.

Knowles PF (1972) The plant geneticist's contribution towards changing lipid and amino acid composition of safflower. J. Am. Oil Chem. Soc. 49:27-29.

Knowles PF (1969) Centers of plant diversity and conservation of crop germplasm: Safflower. Econ. Bot. 23:34-329.

Kolkman JM, Slabaugh MB, Bruniard JM, Berry S, Bushman BS, Olungu C, Maes N, Abratti G, Zambelli A, Miller JF, Leon A y Knapp SJ (2004) Acetohydroxyacid synthase mutations conferring resistance to imidazolinone or sulfonylurea herbicides in sunflower. Theor. Appl. Genet. 109:1147-1159.

Kushi LH, Folsom AR, Prineas RJ, Mink PJ, Wu Y y Bostick RM (1996) Dietary antioxidant vitamins and death from coronary heart disease in postmenopausal women. N. Engl. J. Med. 334:1156-1162.

Lacombe S, Léger S, Kaan F y Bervillé A (2002) Genetic, molecular and expression features of the Pervenets mutant leading to high oleic acid content of seed oil in sunflower. OCL 9:17-23.

Lacombe S, Kaan F, Léger L y Bervillé A (2001) An oleate desaturase and a suppressor loci direct high oleic acid content of sunflower (*Helianthus annuus* L.) oil in the Pervenets mutant. C.R. Acad. Sci. 324:1-7.

Lai Z, Livingstone K, Zou Y, Church SA, Knapp SJ, Andrews J y Rieseberg LH (2005) Identification and mapping of SNPs from ESTs in sunflower. Theor. Appl. Genet. 111:1532-1544.

Langar K, Lorieux M, Desmarais E, Griveau Y, Gentzbittel L y Bervillé A (2003) Combined mapping of DALP and AFLP markers in cultivated sunflower using F9 recombinant inbred lines. Theor. Appl. Genet. 106:1068-1074.

Lawson WR, Goulter KC, Henry RJ, Kong GA y Kochman JK (1998) Marker assisted selection for two rust resistance genes in sunflower. Mol. Breed. 4:227-234.

Lawson WR, Goulter KC, Henry RJ, Kong GA y Kochman JK (1996) RAPD markers for a sunflower rust resistance gene. Aust. J. Agric. Res. 47:395-401.

Lehmann J, Martin HL, Lashley EL, Marshall MW y Judd JT (1986) Vitamin E in foods from high and low linoleic acid diets. J. Am. Diet. Assoc. 86:1208-1216.

Lemaire-Ewing S, Desrumaux C, Néel D y Lagrost L (2010) Vitamin E transport, membrane incorporation and cell metabolism: Is α-tocopherol in lipid rafts an oar in the lifeboat? Mol. Nutr. Food Res. 54:631-640.

Li Y, Wang Z, Sun X y Tang K (2008) Current opinions on the functions of tocopherol based on the genetic manipulation of tocopherol biosynthesis in plants. J. Integr. Plant Biol. 50:1057-1069.

Lopez-Gonzalez G (1990) Acerca de la clasificación natural del genero *Carthamus* L., s.l. Anales Jardín Bot. Madrid 47:11-34.

Lu YH, Melero-Vara JM, García-Tejada JA y Blanchard P (2000) Development of SCAR markers linked to the gene Or5 conferring resistance to broomrape (*Orobanche cumana* Wallr.) in sunflower. Theor. Appl. Genet. 100:625-632

Maeda H, Song W, Sage TL y DellaPenna D (2006) Tocopherols play a crucial role in low-temperature adaptation and phloem loading in *Arabidopsis*. Plant Cell 18:2710-2732.

Mangialasche F, Kivipelto M, Mecocci P, Rizzuto D, Palmer K, Winblad B y Fratiglioni L (2010) High plasma levels of vitamin E forms and reduced Alzheimer's Disease risk in advanced age. J. Alzheimers Dis. 20:1029-1037.

Marmesat S, Velasco L, Ruiz-Méndez MV, Fernández-Martínez JM y Dobarganes C (2008) Thermostability of genetically modified sunflower oils differing in fatty acid and tocopherol compositions. Eur. J. Lipid Sci. Technol. 110:776-782.

Martínez-Rivas JM, Sperling P, Lühs W y Heinz E (2001) Spatial and temporal regulation of three oleate desaturase (*FAD2*) genes from normal-type and high-oleic varieties of sunflower (*Helianthus annuus* L.). Mol. Breed. 8:159-168.

Mataix J y Ochoa J (2002) Vitaminas. p. 137-203. *En*: J Mataix Verdú (ed.). Nutrición y alimentación humana. Ergón, Majadahonda (Madrid).

Mayerhofer R, Archibald C, Bowles V y Good AG (2010) Development of molecular markers and linkage maps for the *Carthamus* species *C. tinctorius* and *C. oxyacanthus*. Genome 53:266-276.

Mayerhofer R, Bowles V, Mayerhofer M y Good AG (2008) Genetic linkage maps of *Carthamus* species based on SSR and RFLP markers. *En*: SE Knights y TD Potter (eds.). Proc. 7th Int. Safflower Conf., Wagga Wagga, Australia.

McLaughlin PJ y Weihrauch JL (1979) Vitamin E content of foods. J. Am. Diet. Assoc. 75:647-665.

Mène-Saffrané L y DellaPenna D (2010) Biosynthesis, regulation and functions of tocochromanols in plants. Plant Physiol. Biochem. 48:301-309.

Morris MC, Evans DA, Tangney CC, Bienias JL, Wilson RS, Aggarwal NT y Scherr PA (2005) Relation of the tocopherol forms to incident Alzheimer Disease and to cognitive change. Am. J. Clin. Nutr. 81:508-514.

Muggli R (1994) Vitamin E-Bedarf bei Zufuhr von Polyenfettsäuren. Eur. J. Lipid Sci. Technol. 96:17-19.

Munné-Bosch S (2007) Alpha-tocopherol: a multifaceted molecule in plants. Vitam. Horm. 76:375-92.

Nagaoka S, Kuranaka A, Tsuboi H, Nagashima U y Mukai K (1992) Mechanism of antioxidant reaction of vitamin E. Charge transfer and tunneling effect in proton-transfer reaction. J. Phys. Chem. 96:2754-2761.

Naresh V, Yamini KN, Rajendrakumar P y Kumar VD (2009) EST-SSR marker-based assay for the genetic purity assessment of safflower hybrids. Euphytica 170:347-353.

Nishizuka Y (2001) The protein kinase C family and lipid mediators for transmembrane signaling and cell regulation. Alcoholism (NY) 25:3S-7S.

Norris SR, Barette TR y DellaPenna D (1995) Genetic dissection of carotenoid synthesis in *Arabidopsis* defines plastoquinone as an essential component of phytone desaturation. Plant Cell 7:2139-2149.

Packer L y Obermüler-Jevic UC (2002) Vitamin E: an introduction. p. 133-151. *En:* L Packer et al. (eds.). The antioxidant vitamins C and E. AOCS Press, Champaign, Illinois.

Padley FB, Gunstone FD y Harwood JL (1994) Occurrence and characteristics of oils and fats. p. 47-223. *En*: FD Gunstone et al. (eds.). The lipid Handbook. London, Chapman and Hall.

Panero JL y Funk VA (2002) Toward a phylogenetic subfamilial classification for the *Compositae* (*Asteraceae*). Proc. Biol. Soc. Wash. 115:909-922.

Paniego N, Heinz R, Fernández P, Talia P, Nishinakamasu V y Hopp HE (2007) Sunflower. p. 153-177. *En*: C. Kole (ed.). Genome mapping and molecular breeding in plants, Vol. 2, Oilseeds, Springer-Verlag, Berlin-Heidelberg, Germany.

Paniego N, Echaide M, Muñoz M, Fernández L, Torales S, Faccio P, Fuxan I, Carrera M, Zandomeni R, Suárez EY y Hopp HE (2002) Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.). Genome 45:34-43.

Pashley CH, Ellis JR, McCauley DE y Burke JM (2006) EST databases as a source for molecular markers: lessons from *Helianthus*. J. Hered. 97:381-388.

Patzak J (2001) Comparison of RAPD, STS, ISSR and AFLP molecular methods used for assessment of genetic diversity in hop (*Humulus lupulus* L.). Euphytica 121:9-18.

Peerbolte RP y Peleman J (1996) The CARTISOL sunflower RFLP map (146 loci) extended with 291 AFLP markers. p. 174-178. *En*: Proc. 18th Sunflower Res. Forum, Fargo, ND, USA, Natl. Sunflower Assoc, Bismarck, ND, USA.

Pérez-Vich B y Berry ST (2010) Molecular Breeding. p. 221-252. *En*: J. Hu et al. (eds.). Genetics, Genomics and Breeding of Sunflower. Science Publishers.

Pérez-Vich B, Velasco L, Grondona M, Leon AJ y Fernández-Martínez JM (2006) Molecular analysis of the high stearic acid content in sunflower mutant CAS-14. Theor. Appl. Genet. 112:867-875.

Pérez-Vich B, Knapp SJ, Leon AJ, Fernández-Martínez JM y Berry ST (2004) Mapping minor QTL for increased stearic acid content in sunflower seed oil. Mol. Breed. 13:313-322.

Pérez-Vich B, Fernández-Martínez JM, Grondona M, Knapp SJ y Berry ST (2002) Stearoyl-ACP and oleoyl-PC desaturase genes cosegregate with quantitative trait loci underlying stearic and oleic acid mutant phenotypes in sunflower. Theor. Appl. Genet. 104:338-349.

Pokorny J (1987) Major factors affecting the autoxidation of lipids. p. 141-206. *En*: HWS Chan (ed.).Autoxidation of unsaturaled lipids. Academic Press, London.

Pongracz G, Weiser H y Matzinger D (1995) Tocopherole. Antioxidanten der Natur. Eur. J. Lipid Sci. Technol. 97:90-104.

Porfirova S, Bergmüller E, Tropf S, Lemke R y Dormann P (2002) Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. Proc. Natl. Acad. Sci. U. S. A. 99:12495-12500.

Powell W, Machray GC y Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 7:215-222.

Putt ED (1997) Sunflower early history. p. 1-21. *En*: AA Schneiter (ed.). Sunflower production and technology. Agronomy Monograph 35. ASA-CSSA-SSSA, Madison, WI, USA.

Rachid Al-Chaarani GR, Gentzbittel L, Huang X y Sarrafi A (2004) Genotypic variation and identification of QTLs for agronomic traits using AFLP and SSR in recombinant inbred lines of sunflower (*Helianthus annuus* L.). Theor. Appl. Genet. 109:1353-1360.

Rachid Al-Chaarani GR, Roustaee A, Gentzbittel L, Mokrani L, Barrault G, Dechamp-Guillaume G y Sarrafi A (2002) A QTL analysis of sunflower partial resistance to downy mildew (*Plasmopara halstedii*) and black stem (*Phoma macdonaldii*) by the use of recombinant inbred lines (RILs). Theor. Appl. Genet. 104:490-496.

Rieseberg LH, Van Fossen C y Desrochers A (1995) Hybrid speciation accompanied by genomic reorganization in wild sunflowers. Nature 375:313-316.

Rieseberg LH, Choi H, Chan R y Spore C (1993) Genomic map of a diploid hybrid species. Heredity 70:285-293.

Rota C, Rimbach G, Minihane AM, Stoecklin E y Barella L (2005) Dietary vitamin E modulates differential gene expression in the rat hippocampus: potential implications for its neuroprotective properties. Nutr. Neurosci. 8:21-29.

Sandorf I y Holländer-Czytko H (2002) Jasmonate is involved in the induction of tyrosine aminotransferase and tocopherol biosynthesis in *Arabidopsis thaliana*. Planta 216:173-149.

Sattler SE, Mène-Saffrané L, Farmer EE, Krischke M, Mueller MJ y DellaPenna D. (2006) Nonenzymatic lipid peroxidation reprograms gene expression and activates defense markers in *Arabidopsis* tocopherol-deficient mutants. Plant Cell 18:3706-3720.

Sattler SE, Gilliland LU, Magallanes-Lundback M, Pollard M y DellaPenna D (2004) Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. Plant Cell 16:1419-1432.

Schilling EE y Heiser CB (1981) Infrageneric classification of *Helianthus* (*Compositae*). Taxon 30:393-403.

Schultz G (1990) Biosynthesis of α-tocopherol in chloroplasts of higher plants. Eur. J. Lipid Sci. Technol. 92:86-91.

Schultz HW, Day EA y Sinnhuber RD (1962) Symposium on foods: lipids and their oxidation. AVI Publishing Co., Westport.

Sehgal D y Raina SN (2011) *Carthamus*. p. 63-95. *En*: C Kole (ed.). Wild crop relatives: genomic and breeding resources. Springer Berlin Heidelberg.

Sehgal D, Rajpal V, Raina SN, Sasanuma T y Sasakuma T (2009) Assaying polymorphism at DNA level for genetic diversity diagnostics of the safflower (*Carthamus tinctorius* L.) world germplasm resources. Genetica 135:457-470.

Sehgal D, Rajpal VR y Raina SN (2008) Chloroplast DNA diversity reveals the contribution of two wild species to the origin and evolution of cultivated safflower. Genome 51:638-643.

Seppanen CM, Song Q y Csallany AS (2010) The antioxidant functions of tocopherol and tocotrienol homologues in oils, fats, and food systems. J. Am. Oil Chem. Soc. 87:469-481.

Sheppard A, Pennington J y Weihrauch J (1993) Analysis and distribution of vitamin E in vegetable oils and foods. p. 9-31. *En:* L Packer y J Fuchs (eds.). Vitamin E in health and disease. Marcel Dekker, New York.

Síes H (1997) Oxidative stress: oxidants and antioxidants. Exp. Physiol. 82:291-295.

Singh V y Nimbkar N (2007) Safflower (*Carthamus tinctorius* L.). p. 167-194. *En*: R Singh (ed.). Genetic resources, chromosome engineering, and crop improvement. Oilseed crops. CRC Press, Boca Raton, FL, EE.UU.

Somerville C (1991) Prospects for genetic modification of the composition of edible oils from higher plants. Food Biotechnol. 5:217-228.

Stone WL, Krishnan K, Campbell SE, Qui M, Whaley SG y Yang H (2004) Tocopherols and the treatment of colon cancer. Ann. NY Acad. Sci. 1031:223-233.

Studer RK, Graven PA y Derubertis FR (1997) Antioxidant inhibition of protein kinase C-signaled increases in transforming growth factor-Beta in mesangial cells. Metabolism 46:918-925.

Tamborindeguy C, Ben C, Liboz T y Gentzbittel L (2004) Sequence evaluation of four specific cDNA libraries for developmental genomics of sunflower. Mol. Genet. Genomics 271:367-375.

Tang S, Hass CG y Knapp SJ (2006) *Ty3/gypsy*-like retrotransposon knockout of a 2-methyl-6-phytyl-1,4-benzoquinone methyltransferase is non-lethal, uncovers a cryptic paralogous mutation, and produces novel tocopherol (vitamin E) profiles in sunflower. Theor. Appl. Genet. 113:783-799.

Tang S, Kishore VK y Knapp SJ (2003) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor. Appl. Genet. 107:6-19.

Tang S, Yu JK, Slabaugh MB, Shintani DK y Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. 105:1124-1136.

Tóth G, Gáspari Z y Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res. 10:967-981.

Urie AL (1985) Inheritance of high oleic acid in sunflower. Crop Sci. 25:986-989.

Usoro OB y Mousa SA (2010) Vitamin E forms in Alzheimer's Disease: a review of controversial and clinical experiences. Crit. Rev. Food Sci. Nutr. 50:414-419.

Valentin HE y Qungang Q (2005) Biotechnological production and application of vitamin E: current state and prospects. Appl. Microbiol. Biotechnol. 68:436-444.

Vavilov NI (1951) The origin, variation, immunity and breeding of cultivated plants. Ronald Press Company, Nueva York.

Velasco L, Del Moral L, Pérez-Vich B y Fernández-Martínez JM (2010) Selection for contrasting seed tocopherol content in sunflower seeds. J. Agric. Sci. 148:393-400.

Velasco L, Pérez-Vich B, Jan CC y Fernández-Martínez JM (2007) Inheritance of resistance to broomrape (*Orobanche cumana* Wallr.) race F in a sunflower line derived from wild sunflower species. Plant Breed. 126:67-71.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2005a) Identification and genetic characterization of a safflower mutant with a modified tocopherol profile. Plant Breed. 124:459-463.

Velasco L, Pérez-Vich B, Hamdan YAS y Fernández-Martínez JM (2005b) Genetic study of several seed oil quality traits in safflower. p. 74-79. *En*: E Essendal (ed.). Proc. 6th Int. Safflower Conf., Estambul, Turquía.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2004a) Grain quality in oil crops. p. 389-427. *En*: R Benech-Arnold y RA Sánchez (eds.). Handbook of seed physiology: applications to agriculture. The Haworth Press, Nueva York.

Velasco L, Domínguez J y Fernández-Martínez JM (2004b) Registration of T589 and T2100 sunflower germplasms with modified tocopherol profiles. Crop Sci. 44:361-362.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2004c) Novel variation for tocopherol profile in a sunflower created by mutagenesis and recombination. Plant Breed. 123:490-492.

Velasco L y Fernández-Martínez JM (2004d) Registration of CR-34 and CR-81 safflower germplasms with increased tocopherol. Crop Sci. 44:2278.

Velasco L y Fernández-Martínez JM (2003) Identification and genetic characterization of new sources of beta- and gamma-tocopherol in sunflower germplasm. Helia 26:17-24.

Velasco L y Fernández-Martínez JM (2001) Breeding for oil quality in safflower. p. 133-137. *En*: JW Bergman y HH Mündel (eds.). Proc. 5th Int. Safflower Conf. Williston, ND y Sidney, MO, USA.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2000) Inheritance of oleic acid content under controlled environment. *En:* Proc. 15th Int. Sunflower Conf., Toulouse, Francia.

Velasco L y Goffman F (2000) Tocopherol, plastochromanol and fatty acid patterns in the genus *Linum*. Plant Syst. Evol. 221:1-2.

Vera-Ruiz EM, Velasco L, Leon AJ, Fernández-Martínez JM y Pérez-Vich B (2006) Genetic mapping of the *Tph1* gene controlling beta-tocopherol accumulation in sunflower seeds. Mol. Breed. 17:291-296.

Vera-Ruiz EM, Pérez-Vich B y Fernández-Martínez JM, Velasco L (2005) Comparative genetic study of two sources of beta-tocopherol in sunflower. Helia 28:1-8.

Vles RO y Gottenbos JJ (1989) Nutritional characteristics and food uses of vegetable oils. p. 63-86. *En*: RK Downey et al. (eds.). Oil crops of the world. New York, McGraw-Hill.

Vrânceanu AV (1977) Morfología. p. 61-72. El girasol. Mundi-Prensa, Madrid.

Wagner KH, Kamal-Eldin A y Elmadfa I (2004) Gamma-tocopherol - an underestimated vitamin? Annu. Nutr. Metab. 48:169-88.

Wang X y Quinn P (1999) Vitamin E and its function in membranes. Prog. Lipid Res. 38:309-336.

Warner K, Miller JF y Demurin Y (2008) Oxidative stability of crude mid-oleic sunflower oils from seeds with high γ- and δ- tocopherols levels. J. Am. Oil Chem. Soc. 85:529-533.

Warner K (2005) Effects of the flavor and oxidative stability of stripped soybean and sunflower oils with added pure tocopherols. J. Agric. Food Chem. 53:9906-9910.

Warner K, Neff WE y Eller FJ (2003) Enhancing quality and oxidative stability of aged fried food with gamma-tocopherol. J. Agric. Food Chem. 51:623-627.

Weinstein SJ, Wright ME, Pietinen P, King I, Tan C, Taylor PR, Virtamo J y Albanes D (2005) Serum alpha-tocopherol and gamma-tocopherol in relation to prostate cancer risk in a prospective study. J. Natl. Cancer Inst. 97:396-399.

Weiss EA (1971) Castor, sesame and safflower. p. 529-744. Barnes y Noble. Inc., New York.

Welsch R, Medina J, Giuliano G, Beyer P y von Litig J (2003) Structural and functional characterization of the phytoene synthase promoter from *Arabidopsis thaliana*. Planta 216:523-534.

Wright ME, Lawson KA, Weinstein SJ, Pietinen P, Taylor PR, Virtamo J y Albanes D (2006) Higher baseline serum concentrations of vitamin E are associated with lower total and cause-specific mortality in the alpha-tocopherol, beta-carotene cancer prevention study. Am. J. Clin. Nutr. 84:1200-1207.

Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu WC, Webb DM, Thompson L, Edwards KJ, Berry S, Leon A, Olungu C, Maes N y Knapp SJ (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci. 43:367-387.

Yu JK, Mangor J, Thompson L, Edwards KJ, Slabaugh MB y Knapp SJ (2002) Allelic diversity of simple sequence repeat markers among elite inbred lines in cultivated sunflower. Genome 45:652-660.

OBJETIVOS

OBJETIVOS DE LA TESIS

1. Realización de un análisis genético comparativo de cuatro líneas de girasol con alto contenido en γ-tocoferol. Identificación de marcadores moleculares ligados al gen *Tph2* que controla la acumulación de γ-tocoferol en semillas de girasol y mapeo de dicho gen en el mapa genético de referencia de girasol.

2. Caracterización fenotípica y molecular de genes modificadores del alto contenido en γ-tocoferol en girasol, responsables de la aparición de niveles intermedios de γ-tocoferol en cruzamientos donde está involucrada la línea IAST-1 caracterizada por poseer alto contenido en γ-tocoferol.

3. Transferencia de marcadores moleculares específicos de regiones génicas y no génicas de girasol a cártamo.

4. Desarrollo de marcadores moleculares asociados al gen que determina alto contenido en γ-tocoferol en cártamo e identificación y caracterización de dicho gen mediante una estrategia de genes candidatos.

RESULTADOS GLOBALES Y DISCUSIÓN

RESULTADOS GLOBALES Y DISCUSIÓN

Los tocoferoles son los principales compuestos antioxidantes presentes en semillas oleaginosas. Su actividad antioxidante tiene lugar tanto en sistemas biológicos (actividad *in vivo* o vitamina E) como en aceites y productos derivados (*in vitro*) (Kamal-Eldin y Appelqvist, 1996). De los cuatro tocoferoles existentes, α-tocoferol ejerce la mayor actividad antioxidante *in vivo*, mientras que el resto de los tocoferoles (β-, γ-, y δ-tocoferol) poseen menor actividad como vitamina E pero mayor actividad antioxidante *in vitro* (Warner y Moser, 2009). α-tocoferol es la forma más abundante en semillas de girasol (>90% del total de tocoferoles) (Demurin, 1993) y de cártamo (>95%) (Johnson et al., 1999). En ambos cultivos se han desarrollado líneas con perfiles de tocoferoles modificados. En concreto, se ha identificado germoplasma de girasol con alto contenido en β-tocoferol (Demurin, 1993; Velasco et al., 2004a, 2004b), γ-tocoferol (Demurin, 1993; Velasco et al., 2004a, 2004b), y δ-tocoferol (Velasco et al., 2004b; Hass et al., 2006), mientras que en cártamo se ha desarrollado una línea con alto contenido en γ-tocoferol (>85%) (Velasco et al., 2005). Varios estudios indican que la sustitución deα -tocoferol por γ-tocoferol mejora considerablemente la estabilidad oxidativa del aceite de girasol (Demurin, 1996; Fuster et al., 1998; Yanishlieva et al., 2002; Warner et al. 2003, 2008; Warner, 2005; Marmesat et al., 2008). Por otro lado, existen investigaciones médicas que sugieren que γ-tocoferol podría tener un efecto positivo en la salud y proteger frente a enfermedades. Concretamente, varios estudios indican que es eficaz neutralizando óxidos de nitrógeno (Cooney et al., 1993; Christen et al., 1997), reduciendo la inflamación (Wagner et al., 2004), protegiendo frente a varios tipos de cáncer (Stone et al., 2004; Jiang et al., 2004a, 2004b; Weinstein et al., 2005) y en prevención de la enfermedad de Alzheimer (Morris et al., 2005; Rota et al., 2005; Usoro y Mousa, 2010; Mangialasche et al., 2010). En plantas, no existen muchas investigaciones donde se haya estudiado el efecto biológico de γ-tocoferol. Desel y Krupinska (2005) basándose en las observaciones obtenidas previamente en animales (Cooney et al., 1993; Christen et al., 1997) indicaron que γ-tocoferol podría controlar los niveles de óxido nítrico en semillas, retrasando la emergencia de la raíz.

En este contexto se plantea la presente tesis, cuyo objetivo global es la profundización en las bases genéticas del alto contenido en γ-tocoferol en semillas de girasol y cártamo. Para ello se plantearon los siguientes objetivos específicos:

1. Realización de un análisis genético comparativo de cuatro líneas de girasol con alto contenido en γ-tocoferol. Identificación de marcadores moleculares ligados al gen *Tph2* que controla la acumulación de γ-tocoferol en semillas de girasol y mapeo de dicho gen en el mapa genético de referencia de girasol.

2. Caracterización fenotípica y molecular de genes modificadores del alto contenido en γ-tocoferol en girasol, responsables de la aparición de niveles intermedios de γ-tocoferol en cruzamientos donde está involucrada la línea IAST-1 caracterizada por poseer alto contenido en γ-tocoferol.

3. Transferencia de marcadores moleculares específicos de regiones génicas y no génicas de girasol a cártamo.

4. Desarrollo de marcadores moleculares asociados al gen que determina alto contenido en γ-tocoferol en cártamo e identificación y caracterización de dicho gen mediante una estrategia de genes candidatos.

Estudios genéticos sobre el carácter alto γ-tocoferol en girasol

En girasol se han desarrollado cuatro líneas con alto contenido en γ-tocoferol (>85%) en semillas. Dichas líneas se identificaron a partir de una evaluación de germoplasma: LG-17 (Demurin et al., 1993) y T2100 (Velasco et al., 2004a), y mediante mutagénesis: IAST-1 e IAST-540 (Velasco et al., 2004b). Los primeros estudios genéticos acerca de este carácter se llevaron a cabo en la línea de girasol LG-17. Se concluyó que el alto contenido en γ-tocoferol en esta línea era el resultado de alelos recesivos en un único gen que se denominó *Tph2* (Demurin et al., 1993, 1996). Velasco y Fernández-Martínez (2003) corroboraron la presencia de alelos recesivos en un único locus para el carácter alto γ-tocoferol en la línea T2100. Posteriormente, Demurin et al. (2004) realizaron estudios genéticos comparativos entre las líneas LG-17 y T2100 confirmando que el carácter alto contenido en γ-tocoferol se encuentra controlado en ambas líneas por los alelos *tph2*.

En la presente tesis, se ha llevado a cabo el estudio genético comparativo de las líneas LG-17, T2100, IAST-1 e IAST-540. Este análisis confirmó la existencia de alelos recesivos en estas líneas en el locus *Tph2* responsable de este carácter de acuerdo a lo

descrito anteriormente por Demurin et al. (2004) para las líneas LG-17 y T2100. La línea IAST-1 se seleccionó de una familia que mostró una amplia segregación para el contenido en γ-tocoferol (Velasco et al., 2004b), pero en el presente estudio no se encontraron niveles intermedios de γ-tocoferol en esta línea ni en las generaciones segregantes derivadas de sus cruzamientos con otras líneas con alto contenido en γ-tocoferol. Asimismo se ha mapeado el gen *Tph2* en el grupo de ligamiento 8 del mapa genético de girasol a partir de la población F_2 obtenida del cruzamiento entre la línea CAS-12, con bajo contenido en γ-tocoferol (<5%), y la línea IAST-540. Para ello se realizó un análisis de grupos segregantes uniformes (Bulked Segregant Analysis, BSA) empleando los datos fenotípicos tanto de las semillas F_2 como de sus correspondientes familias F_3, a partir de los cuales se determinaron los alelos del gen *Tph2* presentes en cada individuo de la población. Se construyó un "bulk bajo γ-tocoferol" mezclando ADN de 12 individuos homocigotos dominantes para el gen *Tph2* (*Tph2Tph2*) y otro "bulk alto γ-tocoferol" de igual forma, con la salvedad de que en este caso se mezcló ADN de individuos homocigotos recesivos para el gen *Tph2* (*tph2tph2*). Los bulks junto con muestras de las líneas parentales se analizaron con un esqueleto de 95 marcadores SSR (Simple Sequence Repeats) distribuidos uniformemente a lo largo de los 17 grupos de ligamiento de girasol (Tang et al., 2003). Se identificaron dos marcadores SSR del grupo de ligamiento 8 polimórficos entre los bulks y las líneas parentales. El cribado de todos los marcadores SSR (prefijo ORS y CRT) (Tang et al., 2002; 2003), e INDEL (Insertion/Delection) (prefijo ZVG) (Yu et al., 2003) del grupo de ligamiento 8 del mapa de ligamiento de girasol permitió la construcción de un grupo de ligamiento en el que se integró el gen *Tph2* y 8 loci de marcadores (ORS830, ZVG35, ORS312, ORS599, ORS70, ORS243, CRT35 y ORS456) distribuidos en una distancia genética de 44.6 cM. El análisis de mapeo por intervalos identificó un único QTL (Quantitative Trait Loci) para el contenido en γ-tocoferol, entre los loci ORS312 y ORS599, que explicó el 90.2% de la variación de este carácter. Aunque no se encontró ningún marcador que cosegregara con *Tph2*, estos dos marcadores (ORS312 y ORS599) flanquearon dicho gen a unas distancias genéticas de 3.6 y 1.9 cM, respectivamente. La proximidad de estos marcadores moleculares al gen *Tph2* permite que puedan ser herramientas útiles en programas de selección asistida por marcadores (Marker Assisted Selection, MAS) para la selección de alelos *tph2* en la introgresión del carácter recesivo alto contenido en γ-tocoferol en líneas de girasol. El gen *Tph2* se mapeó asimismo en un

estudio paralelo realizado por Hass et al. (2006), siendo su posición en el grupo de ligamiento 8 similar a la identificada en nuestro estudio.

A continuación se abordó el estudio genético y molecular de genes modificadores que afectan la expresión del alto contenido en γ-tocoferol en la línea IAST-1, derivada de un programa de mutagénesis e hibridación donde se identificó una familia que exhibió una gran variación para este carácter, desde 0 a 84.5% (Velasco et al., 2004b). En una primera etapa se realizaron cruzamientos entre la línea HA89, con bajo contenido en γ-tocoferol con las líneas con alto contenido en γ-tocoferol T2100 e IAST-1, así como el cruzamiento entre T2100 e IAST-1. Al analizar las semillas F_2 derivadas del cruzamiento entre HA89 y T2100 se observó la segregación esperada para un gen recesivo (proporción 3:1; <5% γ-tocoferol:>90% γ-tocoferol). Sin embargo en el cruzamiento entre HA89 e IAST-1 se observaron dos patrones de segregación diferentes. Uno de ellos similar al observado para el cruzamiento HA89 x T2100, y un segundo patrón donde se identificaron niveles intermedios de γ-tocoferol (5 a 80%) (proporción 13:3; <80% γ-tocoferol:>90% γ-tocoferol). Por último, al evaluar las semillas F_2 derivadas del cruzamiento entre T2100 e IAST-1 se observó que algunas familias $F_{1:2}$ mostraron, como era esperable, alto contenido en γ-tocoferol, mientras que en otras familias se observaron niveles intermedios transgresivos para este carácter. Estos resultados sugerían la presencia de genes modificadores que produjeron una reducción en el alto contenido en γ-tocoferol hasta niveles intermedios en semillas con la configuración alélica esperada *tph2tph2*.

En una segunda etapa se llevó a cabo el estudio molecular de los genes responsables de la aparición de estos niveles transgresivos. Para ello se analizó una población procedente del cruzamiento entre nmsT2100 e IAST-1 que mostró una segregación continua para el contenido en γ-tocoferol (1.0 a 99.7%), sin que se pudieran separar clases fenotípicas discretas. A partir de esta población se desarrolló un mapa genético completo. Se cribaron 437 marcadores SSR (Tang et al., 2002, 2003) e INDEL (Yu et al., 2003) en las líneas parentales y cuatro individuos F_2 seleccionados al azar. El mapa de ligamiento se construyó con 129 loci (SSR e INDEL) distribuidos en 17 grupos de ligamiento (792.4 cM). Siguiendo una estrategia de genes candidatos se emplearon marcadores INDEL basados en el gen que codifica la enzima gamma-tocoferol metiltransferasa (γ-TMT) responsable del carácter alto contenido en γ-tocoferol en girasol (Hass et al., 2006). Se mapearon distintos loci *γ-TMT*. El locus principal mapeó

en el grupo de ligamiento 8 (*γ-TMTa-Tph2*) en la misma posición descrita previamente para el gen *Tph2* (Hass et al., 2006). Otros loci secundarios mapearon en los grupos de ligamiento 16 (*γ-TMTb*) y *1* (*γ-TMTc*). Asimismo, otros loci secundarios no polimórficos en esta población se mapearon en otras poblaciones del programa de mejora de tocoferoles (IAST-413 x HA89 y CAS-12 x IAST-540). De esta forma se mapeó el locus *γ-TMTd* en el grupo de ligamiento 14.

Las bandas correspondientes a los loci *γ-TMTa-Tph2, γ-TMTb, γ-TMTc* y *γ-TMTd* se aislaron, clonaron y secuenciaron en las respectivas líneas parentales para determinar su naturaleza. El locus *γ-TMTa-Tph2* de nmsT2100 e IAST-1 contenía alelos que mostraron ser idénticos a ambos genes parálogos *γ-TMT-1* (haplotipo 4) y *γ-TMT-2* (haplotipo 5) descritos por Hass et al. (2006). Los loci *γ-TMTb, γ-TMTc* y *γ-TMTd* mostraron una similitud significativa con los haplotipos 2 y 3 ($3e^{-58}$), haplotipos 4 y 5 ($4e^{-14}$) y haplotipos 1, 3, 4 y 5 ($8e^{-18}$), respectivamente, del gen *γ-TMT* . El alineamiento múltiple de las secuencias de nucleótidos de estos loci reveló una gran semejanza entre los loci *γ-TMTa-Tph2* y *γ-TMTb*, y los loci *γ-TMTc* y *γ-TMTd*. Estos resultados indicaron la posible existencia de loci *γ-TMT* duplicados. La duplicación de loci en girasol no se trata de un fenómeno raro. Barker et al. (2008) estudiaron la evidencia de paleopoliploidía en la familia *Asteracea*e empleando EST y revelaron la existencia de duplicación génica como mecanismo de evolución y diversificación en la tribu *Heliantheae*, a la que pertenece el girasol.

El análisis de los genes modificadores se realizó en varias etapas. En primer lugar, se llevó a cabo un análisis de varianza de un factor que reveló un fuerte efecto sobre los niveles de γ-tocoferol del locus *γ-TMTa-Tph2* mapeado en el grupo de ligamiento 8 tanto en F_2 como en F_3. El análisis de los niveles de γ-tocoferol en las distintas clases genotípicas del locus *γ-TMTa-Tph2* del grupo de ligamiento 8 reveló un contenido intermedio de γ-tocoferol (52.3±29.6%) en la clase homocigota para el alelo procedente de la línea IAST-1, mientras que la clase homocigota para el alelo de T2100 presentó niveles elevados (98.9±0.8%). Estos resultados sugieren la existencia de dos alelos diferentes en el locus *γ-TMTa-Tph2* en las líneas nmsT2100 e IAST-1. El alelo inestable de IAST-1 se designó $tph2^a$, y en homocigosis produjo una variación para el contenido en γ-tocoferol que osciló de 0.97 a 89.2% en la generación de semillas F_2, mientras que el alelo *tph2* de nmsT2100 fue muy estable (>97%).

En segundo lugar, se realizó un análisis de mapeo por intervalos compuesto que permitió la identificación de un QTL principal, denominado *Tph2-γ-TMT*, centrado en el locus *γ-TMTa-Tph2* en el grupo de ligamiento 8 que explicó el 41.2% y el 44.4% de la varianza fenotípica deγ -tocoferol en F_2 y F_3, respectivamente. Además, se detectaron otros cuatro QTL con efectos moderados sobre el contenido en γ-tocoferol, localizados en los grupos de ligamiento 1 (*GamT1.1*), 9 (*GamT9.1*), 14 (*GamT14.1*) y 16 (*GamT16.1*), que explicaron individualmente del 2.3 al 4.1% de la variación del carácter. El efecto conjunto de todos los QTL explicó el 44.4% y 61.2% de la variación fenotípica para el contenido en γ-tocoferol de las generaciones F_2 y F_3, respectivamente. Los picos de los QTL *GamT1.1* y *GamT16.1* se centraron en los loci *γ-TMTc* y *γ-TMTb* en los grupos de ligamiento 1 y 16, respectivamente. Asimismo se determinó que el QTL *GamT14.1* del grupo de ligamiento 14 podría coincidir con el locus *γ-TMTd* mapeado en otras poblaciones, por comparación de su posición con dicho locus.

Finalmente, mediante un análisis de varianza de dos factores del contenido emγ -tocoferol en la generación F_2, se estimaron las interacciones entre el locus *γ-TMTa-Tph2* del grupo de ligamiento 8 y todos los marcadores genotipados en este estudio revelando interacciones muy significativas para el contenido en γ-tocoferol entre dicho locus *γ-TMTa-Tph2* y los loci *γ-TMT* de los grupos de ligamiento 1, 14 y 16. Los análisis de epistasia indicaron que los QTL de los grupos de ligamiento 14 y 16 sólo tuvieron un efecto significativo cuando el alelo de IAST-1 estaba presente en el locus *γ-TMTa-Tph2* del grupo de ligamiento 8, pero no se detectó ninguna interacción cuando el alelo de nmsT2100 estaba presente en este mismo locus. Sin embargo, cuando el locus *γ-TMTa-Tph2* fue homocigoto para el alelo de IAST-1, la presencia de alelos de nmsT2100 en los loci *γ-TMT* de los grupos de ligamiento 14 y 16, provocó una reducción drástica en el contenido en γ-tocoferol.

Asimismo, se llevó a cabo el aislamiento y la secuenciación completa de dos genes parálogos *γ-TMT-1* y *γ-TMT-2* identificados en las líneas IAST-1 y T2100. El análisis de las secuencias no ha revelado ninguna mutación que pudiera explicar por qué los alelos $tph2^a$ presentes en la línea IAST-1 son inestables y se ven afectados por genes modificadores. La respuesta podría estar relacionada con una posible mutación en las secuencias reguladoras. Hasta el momento no se ha identificado la mutación asociada al carácter alto γ-tocoferol en girasol (Hass et al., 2006).

Los caracteres heredados de forma mendeliana pueden presentar fenotipos que difieren en mayor o menor medida de los fenotipos de los parentales. Entre las causas probables de la aparición de estos fenotipos variables están la presencia de alelos alternativos, factores medioambientales y genes modificadores (Nadeau, 2001). Briggs y Knowles (1967) definieron el concepto de genes modificadores como genes que no tienen efecto conocido, salvo intensificar o disminuir la expresión de un gen con efecto mayor. Varios estudios han descrito la aparición de patrones de segregación diferentes debido a la presencia de genes modificadores en girasol, relacionados con el alto contenido en ácido oleico (Urie, 1985; Fernández-Martínez et al., 1989; Velasco et al., 2000) y con resistencia a jopo (Velasco et al., 2007). En el caso concreto del carácter alto contenido en ácido oleico, se ha observado que estos genes provocan una fuerte distorsión en los patrones de segregación (Urie, 1985; Fernández-Martínez et al., 1989; Velasco et al., 2000) o incluso pueden llegar a revertir completamente el carácter (Lacombe et al., 2001). En la presente tesis se ha descrito la presencia de loci *γ-TMT* duplicados cuya interacción podría revertir el fenotipo alto γ-tocoferol a niveles intermedios y bajos, tal como se ha descrito anteriormente respecto a genes modificadores que afectan al alto contenido en ácido oleico en girasol (Fernández-Martínez et al., 1989; Lacombe, 2001).

La presencia de genes modificadores que afectan al carácter alto contenido enγ -tocoferol representa una limitación importante en los programas de mejora dirigidos a la introgresión de este carácter en líneas élite de girasol, debido a que reducen la probabilidad de recuperar el fenotipo alto γ-tocoferol. No obstante su presencia podría tener un efecto beneficioso para el desarrollo de perfiles alternativos de tocoferoles como por ejemplo niveles intermedios de α y γ-tocoferol. En este sentido, los genes modificadores que limitan la expresión del alto contenido en ácido oleico (>80%) en girasol han permitido el desarrollo de híbridos con contenido medio en ácido oleico (55-75%) (Miller y Vick, 2002).

Transferencia de marcadores moleculares de girasol a cártamo

Cuando abordamos el estudio del carácter alto contenido enγ -tocoferol en cártamo, los recursos moleculares disponibles para este cultivo eran limitados. Por este motivo, se inició un estudio acerca de la posibilidad de transferir marcadores moleculares a cártamo desde una especie relacionada filogenéticamente, el girasol. Peakall et al.

(1998) y Rossetto (2001) describieron que el éxito de la amplificación de marcadores moleculares entre distintas especies de plantas aumenta a media que las distancias filogenéticas disminuyen. El girasol pertenece a la misma familia (*Asteraceae*) que el cártamo por lo que se trata de una fuente de herramientas moleculares muy valiosa para su empleo en cártamo. Se evaluaron 119 marcadores SSR de girasol basados en regiones no génicas (Tang et al., 2002) así como 48 marcadores IFLP (Intron Fragment Length Polymorphism) (Yu et al., 2003) y 19 marcadores RGC (Resistance Gene Candidate) (Radwan et al. 2008), ambos basados en regiones génicas. Adicionalmente se estudió la calidad y polimorfismo de 69 marcadores SSR derivados de secuencias EST (Expressed Sequence Tags) de girasol ensayados previamente en cártamo por Heesacker et al. (2008). Los marcadores se clasificaron según la intensidad de la amplificación y la facilidad de evaluación en señal fuerte y evaluación fácil (+++), señal moderada pero permite su evaluación (++), señal débil y evaluación dificultosa (+) y ninguna señal (-). Los marcadores que amplificaron como +++ o ++ se consideraron transferibles. El valor de transferencia obtenido en el caso de los marcadores SSR basados en regiones no génicas fue de 17.6 %. Este resultado se ajusta a lo descrito anteriormente por Whitton et al. (1997) que indicaron que las regiones que flanquean las repeticiones en el ADN de esta familia de plantas no están muy conservadas. Este valor es muy similar al encontrado en las tasas de amplificación entre géneros diferentes en otras familias de plantas (Peakall et al., 1998; Kuleung et al., 2004). Por otro lado, se encontró una alta tasa de transferencia para los marcadores basados en regiones génicas IFLP (56.2%) y RGC (73.7%). En el caso de los marcadores EST-SSR, se obtuvo buena calidad de amplificación en el 49.3% de ellos. El porcentaje de marcadores transferidos que mostraron polimorfismo en cártamo fue de 66.6% para SSR de regiones no génicas, 55.5% para IFLP, 71.4% para marcadores basados en RGC, y 70.6% para marcadores EST-SSR. Estos resultados corroboraron los mayores niveles de transferencia encontrados en marcadores basados en regiones génicas respecto a los basados en regiones no génicas en diferentes estudios debido a la conservación de las regiones codificantes (Varshney et al., 2005; Gutierrez et al., 2005; Pashley et al., 2006). No existen muchos trabajos acerca del análisis de la transferencia de marcadores moleculares de girasol a cártamo. Heesacker et al. (2008) no encontraron diferencias en la tasa de amplificación entre marcadores moleculares génicos y no génicos cuando se testaron en especies distintas del género *Helianthus*, e indicaron que sólo el 14.8% de los marcadores probados amplificaron en una entrada de cártamo. Este porcentaje es

similar al obtenido en la presente investigación en los marcadores basados en regiones no génicas (17.6%). Chapman et al. (2007) indicaron que los marcadores basados en secuencias EST de girasol, lechuga y *Arabidopsis* mostraron altos niveles de transferencia y polimorfismo en cártamo. En cuanto al estudio del grado de polimorfismo de los marcadores transferidos basados en regiones génicas y no génicas, nuestros resultados mostraron un menor grado que el encontrado anteriormente en cártamo empleando marcadores RAPD, AFLP (Amplified Fragment Length Polymorphism), ISSR (Inter-Simple Sequence Repeat) (Johnson et al., 2007; Yang et al., 2007; Amini et al., 2008) o SSR basados en EST (Chapman et al., 2009).

Estudios moleculares sobre el carácter alto γ-tocoferol en cártamo

Hasta la fecha sólo se ha encontrado variabilidad para el perfil de tocoferoles en germoplasma de especies silvestres de cártamo. Velasco et al. (2005) identificaron entradas de *C. lanatus* subsp. *turkestanicus* y *C. oxyacanthus* con niveles elevados de γ-tocoferol, 9.8% y 36% del total de tocoferoles, respectivamente. La entrada de *C. oxyacanthus* segregó ampliamente tanto para contenido en γ-tocoferol como para caracteres morfológicos, presentando un alto grado de introgresión de caracteres de *C. tinctorius*. Mediante selección simultánea para alto contenido en γ-tocoferol y caracteres morfológicos se desarrolló la línea IASC-1, con alto contenido en γ-tocoferol (>85%), y tipo de planta próximo a *C. tinctorius* (Velasco et al., 2005).

La caracterización del gen que determina el alto contenido en γ-tocoferol en cártamo se ha llevado a cabo combinando un análisis BSA y el desarrollo de marcadores SCAR con una estrategia de genes candidatos. Se realizó el cruzamiento de la línea CL-1 con bajo contenido en γ-tocoferol (<1%) y la línea IASC-1 con alto contenido en γ-tocoferol (>85%) (Velasco et al., 2005). El análisis de las semillas F_2 reveló una proporción 3:1 (<3.66% γ-tocoferol:>87.65% γ-tocoferol) indicando segregación de un único gen parcialmente recesivo, denominado *Tph2*, que determina alto contenido en γ-tocoferol en cártamo. Para el análisis BSA se construyó un "bulk bajo γ-tocoferol" mezclando ADN de 10 individuos F_2 seleccionados por ser homocigotos dominantes para el gen *Tph2* (*Tph2Tph2*) en base a los análisis de semillas F_2 y F_3 y otro "bulk alto γ-tocoferol" de igual forma, con la diferencia que en este caso se mezcló ADN de individuos F_2 homocigotos recesivos para el gen *Tph2* (*tph2tph2*). Se cribaron 122

marcadores RAPD, 88 marcadores SSR de cártamo, y 20 marcadores SCAR desarrollados anteriormente por Hamdan et al. (2008) en los parentales y los bulks. Entre ellos, 8 marcadores RAPD (OPAA11, OPAA12, OPG10, OPJ9, OPB8, OPM10, OPL20, OPH12) y un marcador SSR (CAT49) fueron polimórficos en los bulks y mostraron una fuerte asociación con el contenido en γ-tocoferol. El análisis de ligamiento agrupó estos marcadores en el mismo grupo de ligamiento que *Tph2*, en una distancia genética de 40.3 cM. Los loci que flanquearon el gen *Tph2* se clonaron y secuenciaron para el desarrollo de marcadores SCAR. De los diferentes marcadores SCAR desarrollados, tres de ellos se seleccionaron por amplificar bandas polimórficas del tamaño esperado en los parentales y los bulks. El análisis de ligamiento agrupó estos tres loci SCAR y el locus *Tph2* en el mismo grupo de ligamiento. Por otro lado, se testaron los marcadores SCAR ligados a los genes *Li* y *Ms*, responsables del muy alto contenido en ácido linoleico y la androesterilidad nuclear de cártamo, previamente mapeados por Hamdam et al. (2008), para determinar la posible asociación entre estos genes y el gen *Tph2*. No se encontró ningún ligamiento entre estos genes.

En este trabajo se han identificado por primera vez marcadores moleculares ligados a genes que controlan niveles modificados de tocoferoles en cártamo. Los escasos estudios moleculares existentes en este cultivo se han centrado en caracteres hereditarios simples asociados con el perfil de ácidos grasos, androesterilidad génica nuclear y compuestos flavonoides. Así, Hamdan et al. (2008) desarrollaron marcadores SCAR ligados a los genes *Li* y *Ms* responsables de muy alto contenido en ácido linoleico y androesterilidad nuclear, respectivamente. Asimismo, Zhang et al. (2009) desarrollaron marcadores SCAR ligados al gen *HSya* que controla la acumulación de un compuesto flavonoide de cártamo empleado en la medicina tradicional china.

La identificación del gen responsable del carácter altoγ -tocoferol en semillas de cártamo se realizó mediante una estrategia de genes candidatos a través de la transferencia de marcadores basados en el gen *γ-TMT* de girasol (Hass et al., 2006) a cártamo, puesto que habíamos demostrado previamente una elevada tasa de transferencia de marcadores basados en regiones génicas entre estas dos especies (Capítulo 4). Uno de los marcadores probados amplificó dos fragmentos. Uno de ellos, de mayor peso molecular (1000 pb), no fue polimórfico en los parentales y en los bulks, mientras que otro fragmento de menor peso molecular (650 pb) presentó un polimorfismo dominante entre los parentales y los bulks. Este locus cosegregó con

Tph2. Las bandas amplificadas en los parentales se clonaron y secuenciaron. Las secuencias consenso de los fragmentos de mayor tamaño en ambos parentales fueron idénticas. El análisis BLAST reveló que ambos loci mostraron similitud al gen *γ-TMT* de girasol, siendo más significativa en el caso del locus de menor peso molecular asociado a *Tph2* ($3e^{-18}$) indicando que el gen responsable de este carácter está asociado al gen *γ-TMT*, como se ha descrito anteriormente en girasol (Hass et al., 2006) y en *Arabidopsis* (Bergmüller et al., 2003). El análisis de alineamiento clasificó estas secuencias consenso en dos grupos distintos entre sí, por lo que podría tratarse de dos copias diferentes del gen *γ-TMT* de cártamo. Con el objeto de obtener una secuencia de mayor longitud del gen *γ-TMT* de cártamo, se realizó el alineamiento de las secuencias consenso obtenidas con la secuencia genómica del gen *γ-TMT* de girasol. La conservación de la secuencia en las regiones codificantes entre cártamo y girasol nos llevó a desarrollar nuevas parejas de cebadores basados en estas regiones que fueron testados en los parentales y los bulks. Dos parejas de cebadores produjeron fragmentos polimórficos que se mapearon, aislaron, clonaron y secuenciaron. Las secuencias genómicas consenso parciales derivadas de los parentales CL-1 e IASC-1 se enviaron al GenBank, cuyos números de entrada son HM028671 y HM028672, respectivamente. Finalmente, el alineamiento de estas dos secuencias y la secuencia del gen *γ-TMT* de girasol mostró una alta similitud revelando que la secuencia HM028671 abarca desde el exón 1 al exón 4 de los 6 exones presentes en el gen *γ-TMT* de girasol. La comparación de las secuencias parciales de CL-1 e IASC-1 reveló presencia de polimorfismos pero ninguno de ellos se asoció con la mutación responsable de la aparición de niveles elevados de γ-tocoferol en semillas de cártamo. Sin embargo, no se puede excluir que la(s) mutacion(es) puedan encontrarse en el resto del gen no secuenciado o en las secuencias reguladoras. En este sentido, Hass et al. (2006) describieron la ausencia de mutaciones en la secuencia codificante del gen *γ-TMT* de girasol y sugirieron que la mutación podría estar implicada en la interrupción de secuencias reguladoras.

Basándonos en los polimorfismos encontrados en la secuencia parcial del gen *γ-TMT* de cártamo, se diseñaron marcadores INDEL e IFLP específicos que se cribaron en la F_2 y cosegregaron con *Tph2*. Asimismo se testaron en 22 entradas de germoplasma, cultivares y líneas de mejora de cártamo. Dos de los marcadores INDEL y un marcador IFLP produjeron productos de amplificación polimórficos. Los marcadores INDEL amplificaron dos alelos diferentes, siendo uno de ellos exclusivo de la línea IASC-1. El

marcador IFLP también amplificó dos alelos, y mostró un valor de heterocigosidad de 0.43.

Bergman y Flynn (2001) indicaron que la producción de aceites de cártamo de gran calidad podría tener un importante impacto en el cultivo de esta especie. Varios autores han indicado que aceites con alto contenido en γ-tocoferol poseen mayor resistencia a la oxidación durante los procesos de fritura o el almacenamiento (Marmesat et al., 2008; Warner et al., 2008). En este sentido, el desarrollo de variedades de cártamo con alto contenido en γ-tocoferol abre un campo de nuevas aplicaciones para el aceite de cártamo (Hamdan et al., 2009). Asimismo, las robustas herramientas moleculares desarrolladas en el presente trabajo de investigación facilitarán la introgresión de los alelos *tph2* en líneas de cártamo con diferentes perfiles de ácidos grasos en programas de selección asistida por marcadores con el objetivo de desarrollar variedades de este cultivo con nuevas calidades de aceite.

REFERENCIAS

Amini F, Saeidi G y Arzani A (2008) Study of genetic diversity in safflower genotypes using agro-morphological traits and RAPD markers. Euphytica 163:21-30.

Barker MS, Kane NC, Kozik A, Michelmore RW, Matvienko M, Knapp SJ y Rieseberg L (2008). Multiple paleopolyploidizations during the evolution of the *Asteraceae* reveal parallel patterns of duplicate gene retention after millions of years. Mol. Biol. Evol. 25:2445-2455.

Bergman JW y Flynn CR (2001) High oleic safflower as a diesel fuel extender. p 289-293. A potential new market for Montana safflower. En: JW Bergman y HH Mündel (eds.). Proceedings of 5th Int. Safflower Conf., Williston y Sidney.

Bergmüller E, Porfirova S y Dörmann P (2003) Characterization of an *Arabidopsis* mutant deficient in γ-tocopherol methyltransferase. Plant Mol. Biol. 52:1181-1190.

Briggs FD y Knowles PF (1967) Introduction to plant breeding. Reinhold Publishing Corporation, Nueva York.

Chapman MA, Hvala J, Strever J, Matvienko M, Kozik A, Michelmore RW, Tang S, Knapp SJ y Burke JM (2009) Development, polymorphism, and cross-taxon utility of EST-SSR markers from safflower (*Carthamus tinctorius* L.). Theor. Appl. Genet. 120:85-91.

Chapman MA, Chang JC, Weisman D, Kesseli RV y Burke JM (2007). Universal markers for comparative mapping and phylogenetic analysis in the *Asteraceae* (Compositae). Theor. Appl. Genet. 115:747-755.

Christen S, Woodall AA, Shigenaga MK, Southwell-Keely PT, Duncan MW y Ames BN (1997) γ-Tocopherol traps mutagenic electrophiles such as NOx and complements α-tocopherol: physiological implications. Proc. Natl. Acad. Sci. U. S. A. 94:3217-3222.

Cooney RV, Franke AA, Harwood PJ, Hatch-Pigott V, Custer LJ y Mordan LJ (1993) Gamma-tocopherol detoxification of nitrogen dioxide: superiority to alpha-tocopherol. Proc. Natl. Acad. Sci. U.S.A. 90:1771-1775.

Demurin Y, Efimenko SG y Peretyagina TM (2004) Genetic identification of tocopherol mutations in sunflower. Helia 27:113-116.

Demurin Y, Škorić D y Karlovic D (1996) Genetic variability of tocopherol composition in sunflower seeds as a basis of breeding for improved oil quality. Plant Breed. 115:33-36.

Demurin Y (1993) Genetic variability of tocopherol composition in sunflower seeds. Helia 16:59-62.

Desel C y Krupinska K (2005) The impact of tocochromanols on early seedling development and NO release. J. Plant Physiol. 162:771-776.

Fernández-Martínez JM, Jiménez A, Domínguez J, García JM, Garcés R y Mancha M (1989) Genetic analysis of the high oleic content in cultivated sunflower (*Helianthus annuus* L.). Euphytica 41:39-51.

Fuster MD, Lampi AM, Hopia A y Kamal-Eldin A (1998) Effects of alpha- and gamma-tocopherols on the autoxidation of purified sunflower triacylglycerols. Lipids 33:715-722.

Gutierrez MV, Vaz Patto MC, Huguet T, Cubero JI, Moreno MT y Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theor. Appl. Genet. 110:1210-1217.

Hamdan YAS, Pérez-Vich B, Fernádez-Martínez JM y Velasco L (2009) Novel safflower germplasm with increased saturated fatty acid content. Crop Sci. 49:127-132.

Hamdan YAS, Velasco L y Pérez-Vich B (2008) Development of SCAR markers linked to male sterility and very high linoleic acid content in safflower. Mol. Breed. 22:385-393.

Hass CG, Tang S, Leonard S, Traber M, Miller J y Knapp SJ (2006) Three non-allelic epistatically interacting methyltransferase mutations produce novel tocopherol (vitamin E) profiles in sunflower. Theor. Appl. Genet.113:767-782.

Heesacker A, Kishore VK, Gao W, Tang S, Kolkman JM, Gingle A, Matvienko M, Kozik A, Michelmore RM, Lai Z, Rieseberg LH y Knapp SJ (2008) SSRs and

INDELs mined from the sunflower EST database: abundance, polymorphisms, and crosstaxa utility. Theor. Appl. Genet. 117:1021-1029.

Jiang Q, Wong J, Fyrst H, Saba JD y Ames BN (2004a) γ-Tocopherol, or combinations of vitamin E forms, induce cell death in human prostate cancer cells by interrupting sphingolipid synthesis. Proc. Natl. Acad. Sci. U.S.A. 101:17825-17830.

Jiang Q, Wong J y Ames BN (2004b) Gamma-tocopherol induces apoptosis in androgen-responsive LNCaP prostate cancer cells via caspase-dependent and independent mechanisms. Ann. NY Acad. Sci. 1031:399-400.

Johnson RC, Kisha TJ y Evans MA (2007) Characterizing safflower germplasm with AFLP molecular markers. Crop Sci. 47:1728-1736.

Johnson RC, Bergman JW y Flynn CR (1999) Oil and meal characteristics of core and non-core safflower accessions from the USDA collection. Genet. Resour. Crop Evol. 46:611-618.

Kamal-Eldin A y Åppelqvist LÅ (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31:671-701.

Kuleung C, Baenziger PS y Dweikat I (2004) Transferability of SSR markers among wheat, rye, and triticale. Theor. Appl. Genet. 108:1147-1150.

Lacombe S, Kaan F, Léger L y Bervillé A (2001) An oleate desaturase and a suppressor loci direct high oleic acid content of sunflower (*Helianthus annuus* L.) oil in the Pervenets mutant. C.R. Acad. Sci. 324:1-7.

Mangialasche F, Kivipelto M, Mecocci P, Rizzuto D, Palmer K, Winblad B y Fratiglioni L (2010) High plasma levels of vitamin E forms and reduced Alzheimer's Disease risk in advanced age. J. Alzheimers Dis. 20:1029-1037.

Marmesat S, Velasco L, Ruiz-Méndez MV, Fernández-Martínez JM y Dobarganes C (2008) Thermostability of genetically modified sunflower oils differing in fatty acid and tocopherol compositions. Eur. J. Lipid Sci. Technol. 110:776-782.

Miller JF y Vick BA (2002) Registration of four mid-range oleic acid sunflower genetic stocks. Crop Sci. 42:994.

Morris MC, Evans DA, Tangney CC, Bienias JL, Wilson RS, Aggarwal NT y Scherr PA (2005) Relation of the tocopherol forms to incident Alzheimer Disease and to cognitive change. Am. J. Clin. Nutr. 81:508-514.

Nadeau JH (2001) Modifier genes in mice and humans. Nat. Rev. Genet. 2:165-174.

Pashley CH, Ellis JR, McCauley DE y Burke JM (2006) EST databases as a source for molecular markers: Lessons from *Helianthus*. J. Hered. 97:381-388.

Peakall R, Gilmore S, Keys W, Morgante M y Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Mol. Biol. Evol. 15:1275-1287.

Radwan O, Gandhi S, Heesacker A, Whitaker B, Taylor C, Plocik A, Kesseli R, Kozik A, Michelmore RW y Knapp SJ (2008) Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. Mol. Genet. Genomics 280:111-125.

Rossetto M (2001) Sourcing of SSR markers from related plant species. p. 211-224. *En*: RJ Henry (ed.). Plant genotyping: The DNA fingerprinting of plants. CABI, Wallingford.

Rota C, Rimbach G, Minihane AM, Stoecklin E y Barella L (2005) Dietary vitamin E modulates differential gene expression in the rat hippocampus: potential implications for its neuroprotective properties. Nutr. Neurosci. 8:21-29.

Stone WL, Krishnan K, Campbell SE, Qui M, Whaley SG y Yang H (2004) Tocopherols and the treatment of colon cancer. Ann. NY Acad. Sci. 1031:223-233.

Tang S, Kishore VK y Knapp SJ (2003) PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor. Appl. Genet. 107:6-19.

Tang S, Yu JK, Slabaugh MB, Shintani DK y Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. 105:1124-1136.

Urie AL (1985) Inheritance of high oleic acid in sunflower. Crop Sci. 25:986-989.

Usoro OB y Mousa SA (2010) Vitamin E forms in Alzheimer's Disease: a review of controversial and clinical experiences. Crit. Rev. Food Sci. Nutr. 50:414-419.

Varshney RK, Graner A y Sorrells M (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23:48-55.

Velasco L, Pérez-Vich B, Jan CC y Fernández-Martínez JM (2007) Inheritance of resistance to broomrape (*Orobanche cumana* Wallr.) race F in a sunflower line derived from wild sunflower species. Plant Breed. 126:67-71.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2005) Identification and genetic characterization of a safflower mutant with a modified tocopherol profile. Plant Breed. 124:459-463.

Velasco L, Domínguez J y Fernández-Martínez JM (2004a) Registration of T589 and T2100 sunflower germplasms with modified tocopherol profiles. Crop Sci. 44:361-362.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2004b) Novel variation for tocopherol profile in a sunflower created by mutagenesis and recombination. Plant Breed. 123:490-492.

Velasco L y Fernández-Martínez JM (2003) Identification and genetic characterization of new sources of beta- and gamma-tocopherol in sunflower germplasm. Helia 26:17-24.

Velasco L, Pérez-Vich B y Fernández-Martínez JM (2000) Inheritance of oleic acid content under controlled environment. *En:* Proc. 15th Int. Sunflower Conf., Toulouse, Francia.

Wagner KH, Kamal-Eldin A y Elmadfa I (2004) Gamma-tocopherol-an underestimated vitamin? Annu. Nutr. Metab. 48:169-88.

Warner K y Moser J (2009) Frying stability of purified mid-oleic sunflower oil triacylglycerols with added pure tocopherols and tocopherol mixtures. J. Am. Oil Chem. Soc. 86:1199-1207.

Warner K, Miller JF y Demurin Y (2008) Oxidative stability of crude mid-oleic sunflower oils from seeds with high γ- and δ- tocopherols levels. J. Am. Oil Chem. Soc. 85:529-533.

Warner K (2005) Effects of the flavor and oxidative stability of stripped soybean and sunflower oils with added pure tocopherols. J. Agric. Food Chem. 53:9906-9910.

Warner K, Neff WE y Eller FJ (2003) Enhancing quality and oxidative stability of aged fried food with gamma-tocopherol. J. Agric. Food Chem. 51:623-627.

Weinstein SJ, Wright ME, Pietinen P, King I, Tan C, Taylor PR, Virtamo J y Albanes D (2005) Serum alpha-tocopherol and gamma-tocopherol in relation to prostate cancer risk in a prospective study. J. Natl. Cancer Inst. 97:396-399.

Whitton J, Rieseberg LH y Ungerer MC (1997) Microsatellite loci are not conserved across the *Asteraceae*. Mol. Biol. Evol. 14:204-209.

Yang YX, Wu W, Zheng YL, Chen L, Liu RJ y Huang CY (2007) Genetic diversity and relationships among safflower (*Carthamus tinctorius* L.) analyzed by inter-simple sequence repeats (ISSRs). Genet. Resour. Crop Evol. 54:1043-1051.

Yanishlieva NV, Kamal-Eldin A, Marinova EM y Toneva AG (2002) Kinetics of antioxidant action of alpha- and gamma-tocopherols in sunflower and soybean triacylglycerols. Eur. J. Lipid Sci. Technol. 104:262-270.

Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu WC, Webb DM, Thompson L, Edwards KJ, Berry S, Leon A, Olungu C, Maes N y Knapp SJ (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci. 43:367-387.

Zhang Z, Guo M y Zhang J (2009) Identification of AFLP fragments linked to hydroxysafflor yellow A in *Flos Carthami* and conversion to a SCAR marker for rapid selection. Mol. Breed. 23:229-237.

CONCLUSIONES FINALES

CONCLUSIONES FINALES

1) Las líneas alto γ-tocoferol de girasol LG-17, T2100, IAST-1 e IAST-540 presentan alelos recesivos en el locus *Tph2* responsable del carácter alto contenido en γ-tocoferol.

2) Se ha mapeado el gen *Tph2* en el grupo de ligamiento 8 del mapa genético de girasol, y se han identificado marcadores SSR e INDEL asociados a este gen que ha permitido la construcción de un grupo de ligamiento que incluye 8 loci de marcadores y el gen *Tph2*.

3) Los niveles intermedios de γ-tocoferol observados en cruzamientos donde interviene la línea IAST-1 de girasol se explican por la presencia de genes modificadores que alteran la expresión del gen *Tph2*.

4) Existen dos alelos recesivos diferentes en el locus *Tph2* que determina alto contenido en γ-tocoferol en girasol. El alelo *tph2*^{a} presente en la línea IAST-1 es sensible a la acción de genes modificadores, mientras que el alelo *tph2* presente en la línea T2100 no se ve afectado por genes modificadores.

5) Los genes modificadores son loci duplicados del gen *γ-tocoferol metiltransferasa*, localizados en los grupos de ligamiento 1, 14 y 16. Estos loci interaccionan con los alelos *tph2*^{a} produciendo una reducción de los niveles esperados de γ-tocoferol.

6) Los marcadores IFLP y RGC de girasol, basados en regiones génicas, presentaron una alta tasa de transferencia a cártamo de 56.2% y 73.7%, respectivamente, mientras que la tasa mostrada por los marcadores SSR derivados de regiones no génicas fue únicamente de 17.6%. El porcentaje de marcadores transferidos que mostraron polimorfismo en cártamo fue de 66.6% para SSR de regiones no génicas, 55.5% para IFLP, y 71.4% para marcadores basados en RGC.

7) Se han identificado 8 marcadores RAPD y un marcador SSR asociados al gen *Tph2*, responsable del carácter alto contenido en γ-tocoferol de cártamo. Tres de los marcadores RAPD se transformaron en marcadores SCAR y se construyó un grupo de ligamiento que incluyó los marcadores SCAR y el gen *Tph2*. Asimismo, se demostró la asociación del carácter alto contenido en γ-tocoferol con un gen *γ-tocoferol metiltransferasa*.

CAPÍTULO 1

Análisis genético y molecular del alto contenido en gamma-tocoferol en girasol

María J. García-Moreno, Elsa M. Vera-Ruiz, José M. Fernández-Martínez, Leonardo Velasco, y Begoña Pérez-Vich

Genetic and Molecular Analysis of High Gamma-Tocopherol Content in Sunflower

Crop Science 46: 2015-2021 (2006)

Published online September 8, 2006

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

Genetic and Molecular Analysis of High Gamma-Tocopherol Content in Sunflower

María J. García-Moreno, Elsa M. Vera-Ruiz, José M. Fernández-Martínez, Leonardo Velasco, and Begoña Pérez-Vich*

ABSTRACT

Sunflower (*Helianthus annuus* L.) seeds contain alpha-tocopherol as the major tocopherol derivative, which accounts for more than 900 g kg^{-1} total tocopherols. However, four sources of high gamma-tocopherol content (>850 g kg^{-1}) have been developed. First studies on the lines LG-17 and T2100 concluded that the trait in both lines was determined by recessive alleles at the Tph2 locus. The objectives of the present research were (i) to conduct an allelic study on the other two lines, IAST-1 and IAST-540, (ii) to identify markers linked to the Tph2 gene, and (iii) to map this gene. Plants of T2100 were crossed with plants of the other three lines, which resulted in F_1 and F_2 populations with uniformly high gamma-tocopherol content in the seeds, indicating the presence of *tph2* alleles in the four lines. Genetic mapping of the Tph2 gene was conducted with an F_2 population from the cross between CAS-12, with standard tocopherol profile, and IAST-540. F_2 seeds segregated following a 3 low to 1 high gamma-tocopherol ratio. Bulked segregant analysis identified two simple sequence repeats (SSR) markers on linkage group (LG) 8 linked to Tph2. A large linkage group was constructed by genotyping additional markers. Tph2 mapped between markers ORS312 (3.6 cM proximal) and ORS599 (1.9 cM distal). The availability of closely linked PCR-based markers and the location of the Tph2 gene on the sunflower genetic map will be useful for marker-assisted selection and further characterization of tocopherol biosynthesis in sunflower seeds.

Tocopherols are the most important compounds having antioxidant activity in sunflower seeds. In vivo, they exert vitamin E activity, protecting cellular membrane lipids against oxidative damage (Muggli, 1994). In vitro, they inhibit lipid oxidation in oils and fats, as well as in foods and feeds containing them (Kamal-Eldin and Appelqvist, 1996). Alpha-tocopherol exerts the most active biological activity (Traber and Sies, 1996), but it shows the weakest antioxidant potency in vitro. Conversely, beta-, gamma-, and delta-tocopherol possess a lower vitamin E value, but they exert a considerably greater in vitro antioxidant protection than alpha-tocopherol (Pongracz et al., 1995).

Conventional sunflower seeds mainly contain alpha-tocopherol, which accounts for more than 900 g kg^{-1} total tocopherols. Beta- and gamma-tocopherol can be present in sunflower seeds, usually in amounts below 20 g kg^{-1} of the total tocopherols (Demurin, 1993; Dolde et al., 1999). Sunflower germplasm with modified tocopherol profile has been developed. Demurin (1993) reported the line LG-15, with increased concentration of beta-tocopherol (500 g kg^{-1} tocopherols), and the line LG-17, with increased concentration of gamma-tocopherol (950 g kg^{-1} tocopherols), both of them developed from segregating accessions identified in the evaluation of a germplasm collection. Genetic characterization of both lines concluded that the increased levels of beta-tocopherol were produced by recessive alleles at the Tph1 locus, whereas the increased levels of gamma-tocopherol were the result of recessive alleles at the Tph2 locus (Demurin et al., 1996). Also through the evaluation of the natural variability existing in germplasm collections, Velasco et al. (2004a) developed the line T589, with a beta-tocopherol content above 300 g kg^{-1} total tocopherols, and the line T2100, with a gamma-tocopherol content above 850 g kg^{-1}. Velasco and Fernández-Martínez (2003) reported the presence of recessive alleles at a single locus underlying each of the modified tocopherol profiles, i.e., the increased beta-tocopherol concentration in seeds of T589 and the high gamma-tocopherol content in seeds of T2100. Comparative genetic studies concluded that *tph1* alleles were present in both LG-15 and T589 lines (Demurin et al., 2004; Vera-Ruiz et al., 2005), and *tph2* alleles were present in both LG-17 and T2100 lines (Demurin et al., 2004).

Additional variation for gamma-tocopherol content was created in sunflower by using chemical mutagenesis (Velasco et al., 2004b). The authors isolated the lines IAST-1 and IAST-540, with gamma-tocopherol content above 850 g kg^{-1} total tocopherols. No comparative genetic studies have been conducted to determine whether the high gamma-tocopherol lines developed by mutagenesis are allelic to those developed through germplasm evaluation.

Recent advances in molecular marker technologies in sunflower, especially the development of public SSRs (microsatellites) (Tang et al., 2002), SNPs (single nucleotide polymorphisms) (Lai et al., 2005), and integrated genetic linkage maps (Gedil et al., 2001; Yu et al., 2003; Lai et al., 2005) have made possible the genetic mapping and dissection of quantitative and qualitative traits in this crop and the application of this technology to sunflower breeding. Genetic mapping of tocopherol biosynthesis genes and identification of molecular markers linked to them would provide important tools for increased selection efficiency and for investigating the function and organization of these genes. Currently, only the Tph1 gene, conferring increased beta-tocopherol content to sunflower seeds, has been mapped in the sunflower genetic map (Vera-Ruiz et al., 2006). This gene mapped to the upper end of linkage group 1 and cosegregated with the SSR markers ORS1093, ORS222, and ORS598.

The objectives of the present research were (i) to conduct a comparative genetic analysis of the four sources

Instituto de Agricultura Sostenible (CSIC). Apartado 4084. E-14080 Córdoba, Spain. Received 25 Oct. 2005. *Corresponding author (bperez@cica.es).

Published in Crop Sci. 46:2015–2021 (2006).
Crop Breeding & Genetics
doi:10.2135/cropsci2005.10.0388
© Crop Science Society of America
677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: HPLC, high-performance liquid chromatography; INDEL, insertion-deletion polymorphisms; SSR, simple sequence repeats.

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

of high gamma-tocopherol developed so far in sunflower, (ii) to identify PCR-based molecular markers linked to the Tph2 gene controlling gamma-tocopherol accumulation in sunflower seeds, and (iii) to map Tph2 in the sunflower genetic map.

MATERIALS AND METHODS

Plant Material

The genetic study included four sunflower lines with high gamma-tocopherol content and a standard line used as check. The line LG-17 was developed from a germplasm entry of the VIR world germplasm collection (Demurin, 1993). The line T2100 was developed from an accession of the open-pollinated cultivar Peredovik (Velasco et al., 2004a). The lines IAST-1 and IAST-540 were isolated after chemical mutagenesis on seeds of several Peredovik accessions (Velasco et al., 2004b). Seeds of the four lines are characterized by a high gamma-tocopherol content above 850 g kg^{-1} total tocopherols, the rest being mainly alpha-tocopherol. HA89 is an oilseed maintainer line with standard tocopherol profile released by the Texas Agricultural Experiment Station and the USDA-ARS in 1971. For the molecular study, plants of the line IAST-540 were crossed with plants of the line CAS-12, with modified fatty acid profile but standard tocopherol profile (Fernández-Martínez et al., 1997).

Genetic Study

Twenty-four half seeds of HA89, CAS-12, LG-17, T2100, IAST-1, and IAST-540 were nondestructively analyzed for tocopherol profile as described below, germinated and planted in pots in a field screenhouse in spring 2003. Plants of T2100, used in all cases as females, were crossed with plants of LG-17, IAST-1, and IAST-540. Plants of CAS-12, used as females, were crossed with plants of IAST-540. Crossing was done by emasculating florets of the female parent followed by pollination of their stigmas with pollen from the male parent. Half seeds of the parents as well as F_1 half-seeds were analyzed for tocopherol profile. F_1 and parent half seeds were sown in September 2003 and the corresponding plants were grown in the greenhouse. F_1 plants were self-pollinated to obtain the F_2 generation.

F_2 half seeds from one (T2100 × LG17; CAS-12 × IAST-540) to four (rest of the crosses) F_1 plants were analyzed for tocopherol profile. Forty-eight F_2 half seeds per F_1 plant were analyzed in crosses involving high gamma-tocopherol parents, whereas 294 F_2 half seeds from a single F_1 plant were analyzed in the cross CAS-12 × IAST-540. F_2 half-seeds from the latter cross were germinated and the corresponding plants were transplanted to the field in spring 2004. Germination was low in this population, which resulted in a population of 145 F_2 plants. F_2 plants were selfed and ninety of them each produced more than 12 F_3 seeds, which was the minimum number of seeds used for genotypic classification of the F_2 individuals. Twelve to twenty-four F_3 seeds from each of the 90 F_2 plants were analyzed for tocopherol profile. F_2 individuals were classified as *Tph2Tph2* if their F_3 seeds had a uniform low gamma-tocopherol content (<20 g kg^{-1} total tocopherols), *Tph2tph2* if their F_3 seeds segregated for low and high (>850 g kg^{-1} total tocopherols) gamma-tocopherol content, and *tph2tph2* if their F_3 seeds showed a uniform high gamma-tocopherol content.

Bulked Segregant Analysis

Two fully expanded leaves were cut from each of the 145 F_2 plants from the mapping population CAS-12 × IAST-540 and frozen at −80°C. The leaf tissue was lyophilized and ground to a fine powder in a laboratory mill. DNA was isolated from ground leaf tissue from each F_2 plant as described in Berry et al. (1995). DNA was also isolated from three plants of the CAS-12 and IAST-540 parents. For bulked segregant analysis (Michelmore et al., 1991), two bulks were constructed by pooling aliquots (20 μL) of DNA from two sets of individuals with contrasting genotypes. The low gamma-tocopherol bulk was made up from 12 F_2 individuals classified as *Tph2Tph2*, and the high gamma-tocopherol bulk was constructed from 12 individuals classified as *tph2tph2*. Homozygosity of F_2 individuals included in the bulks was verified through the analysis of their respective F_3 seeds. Two replicate samples of each bulk and the parental lines were screened with a genome-wide framework of 95 sunflower SSRs (Tang et al., 2003). For SSRs analyses, PCRs were performed as described by Pérez-Vich et al. (2004), and the amplification products were resolved on 3% (w/v) Metaphor (BMA, Rockland, ME) agarose gels in 1× TBE buffer with ethidium bromide incorporated in the gel.

Linkage between Tph2 and the SSR markers polymorphic between the low gamma-tocopherol and the high gamma-tocopherol bulks was verified by genotyping these SSR markers on the 145 F_2 individuals from CAS-12 × IAST-540. The significance of each marker's association with the gamma-tocopherol content was determined by one-way analysis of variance (ANOVA) using the statistical package SPSS v 12.0 (SPSS for Windows; SPSS Inc., Chicago, IL), with marker genotypes being classes. Additionally, linkage of these markers and Tph2 was also verified by running a preliminary linkage analysis with MAPMAKER/EXP v 3.0b (Whitehead Institute, Cambridge, MA; Lander et al., 1987) using segregation data from the markers and Tph2. The genotypes for the Tph2 gene were inferred from gamma-tocopherol phenotypes in F_2 and F_3 seeds. On the basis of the F_3 analyses, F_2 plants were classified as *Tph2Tph2*, *Tph2tph2*, or *tph2tph2* as described above. F_2 individuals not producing the minimum number of seeds for F_3 analyses (55 of a total of 145), were classified as *Tph2__* if they had a low F_2 gamma-tocopherol content (<20 g kg^{-1}) and *tph2tph2* if they had a high F_2 gamma-tocopherol content (>850 g kg^{-1}). Linkage was considered significant if the LOD score was >8.0. For consideration of the positions of the SSR marker loci relative to the target locus Tph2, linkage distances were calculated as two-point data.

F_2 SSR Genotyping, Map Construction, and Tph2 Mapping

Once the Tph2 linkage group location was identified, all SSR markers known to map to the same linkage group (Tang et al., 2002, 2003; identified by prefixes ORS and CRT), excluding those already used for BSA, were screened for polymorphisms between the parental lines CAS-12 and IAST-540 to construct a complete genetic map including the Tph2 gene. Additionally, INDEL (insertion-deletion polymorphisms) markers mapping to the same linkage group were also used, and they are identified by ZVG prefixes (Yu et al., 2003). Primer sequences from nonpublished markers were kindly provided by Dr. S.J. Knapp (Center for Applied Genetic Technologies, University of Georgia, Athens, Georgia) and Dr. A.J. Leon (Advanta Seeds, Buenos Aires, Argentina). SSR marker analyses were performed as described above. INDEL analyses were performed following Yu et al. (2003).

The SSR and INDEL polymorphic markers were then genotyped in the 145 F_2 individuals from CAS-12 × IAST-540, and a linkage map including Tph2 was constructed with MAPMAKER. The genotypes for the Tph2 gene were deduced as described above, and mapped accordingly. Two-point analysis was used to group all SSR marker loci and Tph2 at a LOD

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

score of 4 and a maximum recombination frequency of 0.35. Three-point and multi-point analyses were used to determine the order and interval distances between the markers. The map distances, expressed in centimorgans (cM), were calculated by means of the Kosambi mapping function. Linkage group maps were drawn by the MapChart software (Voorrips, 2002). Chi-square analyses were performed on each locus to detect deviations from the expected Mendelian ratios for co-dominant (1:2:1) or dominant (3:1) markers.

A regression interval mapping analysis by the PLABQTL 1.1 software (Utz and Melchinger, 1996) was performed using the genetic map constructed to assess the effect of Tph2 on the tocopherol content. For this analysis, a new genetic map was constructed removing Tph2 segregation data and calculating map distances using the Haldane mapping function. The phenotypic data consisted of trait values (gamma-tocopherol content) for each F_2 half-seed. The genetic map was scanned for the presence of QTLs at a LOD threshold of 3.0 at every 2.0-cM interval. Gene action was tested by fitting QTLs to dominant and additive genetic models.

Analysis of Tocopherols by High-Performance Liquid Chromatography

The analysis of tocopherol profile followed the method of Goffman et al. (1999). Half seeds were placed into 10-mL tubes with 2 mL iso-octane. The half seeds were then crushed with a stainless steel rod as finely as possible. The samples were stirred and extracted overnight at room temperature in darkness (extraction time about 16 h). After extraction, the samples were stirred again, centrifuged, and filtered. Five microliters of the extract were analyzed by HPLC using a fluorescence detector at 295-nm excitation and 330-nm emission and iso-octane/tert-butylmethylether (94:6) as eluent at an isocratic flow rate of 0.8 mL min^{-1}. Chromatographic separation of the tocopherols was performed on a LiChrospher 100 diol column (250- × 2-mm I.D.) with 5-μm spherical particles, connected to a silica guard column (LiChrospher Si 60, 5- × 4-mm I.D.). The peak areas of the individual tocopherols were corrected according to their previously calculated response factors, which follow: alpha-tocopherol = 1.0; beta-tocopherol = 1.80; gamma-tocopherol = 1.85; delta-tocopherol = 2.30. The relative content of individual tocopherols in a single half seed was calculated based on corrected peak areas and expressed as g kg^{-1} of the total tocopherols present in the half seed.

RESULTS AND DISCUSSION

Comparative Genetic Study of High Gamma-Tocopherol Lines

Seeds of the high gamma-tocopherol lines LG-17, T2100, IAST-1, and IAST-540 had a uniformly high gamma-tocopherol content of 942 ± 40 (mean ± SD), 952 ± 24, 959 ± 32, and 941 ± 30 g kg^{-1}, respectively, and a low alpha-tocopherol content of 55 ± 40, 40 ± 20, 40 ± 32, and 58 ± 30 g kg^{-1}, respectively. In contrast, seeds of the standard line HA89 used as check had a uniformly high alpha-tocopherol content of 994 ± 6 g kg^{-1} and a low gamma-tocopherol content of 2 ± 4 g kg^{-1}.

F_1 seeds from crosses of LG-17, IAST-1, and IAST-540 with T2100 had a uniformly high gamma-tocopherol content of 947 ± 32 g kg^{-1} (T2100 × LG-17), 944 ± 23 g kg^{-1} (T2100 × IAST-1), and 941 ± 21 g kg^{-1} (T2100 × IAST-540). These results were confirmed in the analysis of F_2 seeds, which also showed high gamma-tocopherol contents of 960 ± 25 g kg^{-1} (T2100 × LG-17), 970 ± 34 g kg^{-1} (T2100 × IAST-1), and 980 ± 16 g kg^{-1} (T2100 × IAST-540). Demurin et al. (2004) evaluated the F_1 from a cross between the line VK 175, a line with *tph2tph2* genotype derived from LG-17, and T2100, concluding that T2100 was allelic to *tph2*. The results of the present research suggested that *tph2* alleles are present in the four lines.

The line IAST-1 was selected from a population that exhibited a wide segregation for gamma-tocopherol content, from zero to 845 g kg^{-1} total tocopherols, including intermediate levels (Velasco et al., 2004b). This had not been observed in the other three high gamma-tocopherol lines, which were selected from populations that only showed segregation for low and high gamma-tocopherol content but not for intermediate values (Demurin, 1993; Velasco et al., 2004a, 2004b). The different segregation pattern in the population from which IAST-1 was selected initially suggested that the line might carry an allele different to *tph2* (Velasco et al., 2004b). This view has been discarded in the present research. However, the occurrence of intermediate levels of gamma-tocopherol (between 50 and 850 g kg^{-1}) in the original population from which IAST-1 was isolated cannot be completely explained taking into account only the presence of *tph2* alleles, and its characterization will require additional specific research.

Molecular Mapping of the Tph2 Gene

Seeds of the line CAS-12 had a low gamma-tocopherol content of 5 ± 2 g kg^{-1}. Seeds of the line IAST-540 had a high gamma-tocopherol content of 946 ± 32 g kg^{-1}. F_1 seeds from the CAS-12 × IAST-540 cross exhibited a low gamma-tocopherol phenotype of 12 ± 3 g kg^{-1}, which is in agreement with previous reports on the recessive character of the trait (Demurin et al., 1996; Velasco and Fernández-Martínez, 2003). F_2 seeds segregated following a bimodal distribution that was not significantly different ($\chi^2 = 2.40$, $p = 0.12$) from a 3:1 (low: high gamma-tocopherol content) ratio (Fig. 1), which indicated segregation of a single, recessive gene. The analysis of 90 $F_{2:3}$ families allowed the classification of F_2 genotypes into three classes, characterized by uniformly low gamma-tocopherol content ($n = 23$; genotype *Tph2Tph2*), segregating for gamma-tocopherol content ($n = 53$; *Tph2tph2*), and uniformly high gamma-tocopherol content ($n = 14$; *tph2tph2*). Such a distribution fit the expected 1:2:1 segregation ratio ($\chi^2 = 4.60$, P = 0.10) that confirms monogenic inheritance of gamma-tocopherol content in IAST-540. According to the allelic study described above, the altered locus in IAST-540 is the Tph2 locus reported by Demurin et al. (1996).

Twenty-eight out of 91 SSR markers that produced amplification products were polymorphic between the parental lines CAS-12 and IAST-540. Two markers from linkage group (LG) 8 (ORS70 and ORS456) were also polymorphic between the low gamma-tocopherol and the high gamma-tocopherol bulks (Fig. 2). The CAS-12 allele only amplified in the low gamma-tocopherol bulk, and the IAST-540 allele only amplified in the high gamma-tocopherol bulk. These results indicated that

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

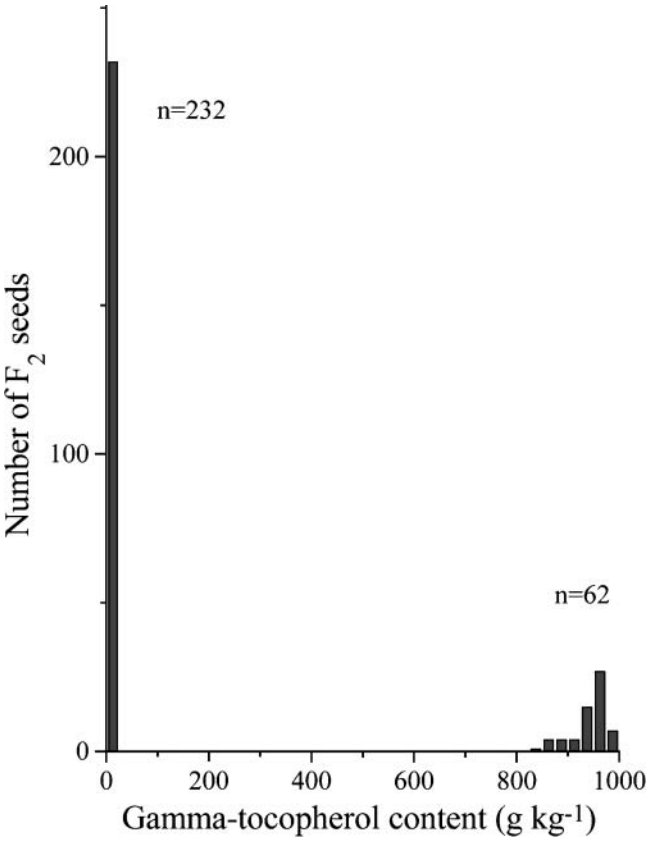


Fig. 1. Histogram of gamma-tocopherol content content (g kg^{-1} total tocopherols) in an F_2 population from the cross between the sunflower lines CAS-12, with standard low gamma-tocopherol content, and IAST-540, with high gamma-tocopherol content.

Tph2 might reside on LG 8. The other three SSR marker loci on LG 8 from the genome-wide framework (ORS166, ORS1161, and ORS894) were monomorphic between CAS-12 and IAST-540.

Linkage of ORS70 and ORS456 with Tph2 was verified by genotyping these SSR markers on 145 F_2 individuals from the mapping population. ANOVA analyses revealed clear significant differences between the marker class means for gamma-tocopherol content ($p < 0.001$). Additionally, a preliminary linkage analysis was run using segregation data from Tph2, ORS70, and ORS456. Two-point analysis showed ORS70 and ORS456 to be 5.7 and 17.4 cM, respectively, from Tph2. These data confirmed linkage of ORS70 and ORS456 with Tph2.

All ORS-SSR, CRT-SSR, and ZVG-INDEL markers known to map to LG 8 (Tang et al., 2002, 2003; Yu et al., 2003), excluding those already used for BSA, were screened for polymorphisms between CAS-12 and IAST-540 to construct a complete linkage map of LG 8. Two codominant (ORS243, and CRT35) and four dominant (ORS830, ORS312, ORS599, and ZVG35) marker loci were then genotyped on the 145 F_2 individuals from the CAS-12 × IAST-540 population. Linkage analysis was performed, including segregation data from Tph2. All markers were grouped together. LG 8 comprised 9 marker loci, including the Tph2 gene, and was 44.6 cM long (Fig. 3). The locus order for the SSR markers and the reference linkage maps (Tang et al., 2002, 2003) was identical, except for the ORS456 locus. The Tph2 gene mapped 26.9 cM downstream from the upper end of LG 8, between markers ORS312 and ORS599. The ORS312 and the ORS599 markers were 3.6 cM proximal and 1.9 cM distal, respectively, of the Tph2 locus (Fig. 3).

Since no Tph2 cosegregating markers were found, a regression interval mapping analysis was performed using the marker data and the F_2 gamma-tocopherol pheno-

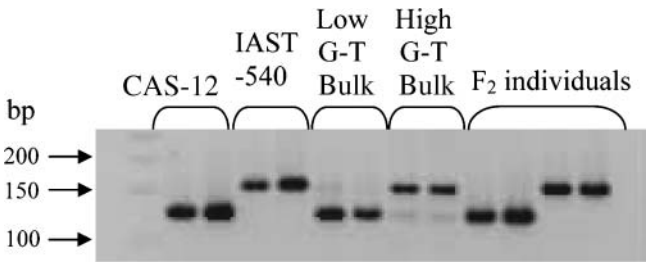


Fig. 2. Amplification products of the SSR marker ORS70. Replicate samples of the low gamma-tocopherol parental line CAS-12, the high gamma-tocopherol parental line IAST-540, the low gamma-tocopherol (G-T) bulk, the high gamma-tocopherol bulk, and four F_2 individuals from CAS-12 × IAST-540 are shown.

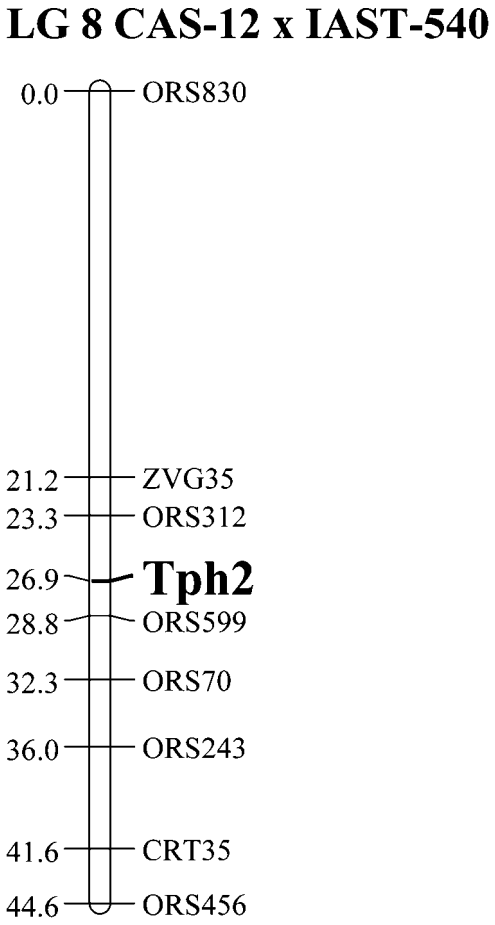


Fig. 3. Molecular map of sunflower linkage group (LG) 8 containing the Tph2 gene determining high gamma-tocopherol content. The ORS and CRT prefixes denote SSR marker loci, and the ZVG prefix denotes INDEL marker loci. The cumulative distances in centimorgans are shown at the left of the map.

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

Table 1. QTL affecting F_2 gamma-tocopherol content in the CAS-12 × IAST-540 population.

QTL	Position‡	Support interval§	Marker interval	LOD	R^2 (%)	Effect†	
						a	*d*
Tph2.8	29	28–36	ORS312-ORS599	73.12	90.2	49.3**	−51.4**

* Significant at the 0.05 probability level.
** Significant at the 0.01.
† *a* = Additive effect. A positive sign means an increase of the mean value of the trait because of IAST-540 alleles; *d* = Dominant effect.
‡ Absolute position from the top of the LG in centimorgans (cM).
§ Support interval in centimorgans; Refers to the region flanking the QTL peak in which LOD scores decline by one.

typic data to assess Tph2 effect on gamma-tocopherol content more accurately (Table 1 and Fig. 4). The marker locus ORS456 was removed from the analysis because its map position was not coincident with that already reported (Tang et al., 2003). Interval mapping analysis identified a single QTL for gamma-tocopherol content on LG 8, between the marker loci ORS312 and ORS599 (Table 1). The QTL had a very large effect, explaining a 90.2% of the phenotypic variation of this trait (Table 1). The LOD peak was directly centered on the Tph2 locus (Fig. 4). Hass et al. (2003) mapped the Tph2 gene in a population derived from the LG-24 line, which was developed from crosses between LG-15 and LG-17 (Demurin, 1993). This gene was mapped on LG 8 of the sunflower genetic map. Despite these authors did not report its position on LG 8, results from the comparative genetic study indicate that the Tph2 gene mapped by Hass et al. (2003) is the same gene reported here, confirming results from our mapping approach.

Since alleles determining high gamma-tocopherol content at the Tph2 locus are recessive, the heterozygote is indistinguishable from the wild-type homozygote. This fact will complicate the selection of appropriate plants carrying the *tph2* allele in backcross programs focused on introgressing the high gamma-tocopherol trait into elite lines. The use of marker-assisted selection can contribute overcoming such a limitation. In the present research, the markers ORS312 and ORS599 flanking the Tph2 gene at a distance of 3.6 cM and 1.9 cM, respectively, have been identified. In the genetic maps reported by Tang et al. (2002, 2003) and Yu et al. (2003), no additional markers mapped in the ORS312-ORS599 marker interval. However, new marker sets are being developed from sunflower expressed sequence tags (ESTs) (Lai et al., 2005), increasing the chance of finding polymorphic markers that map closer to the Tph2 gene, both in the present population or in populations developed from different combinations of standard and high gamma-tocopherol parents.

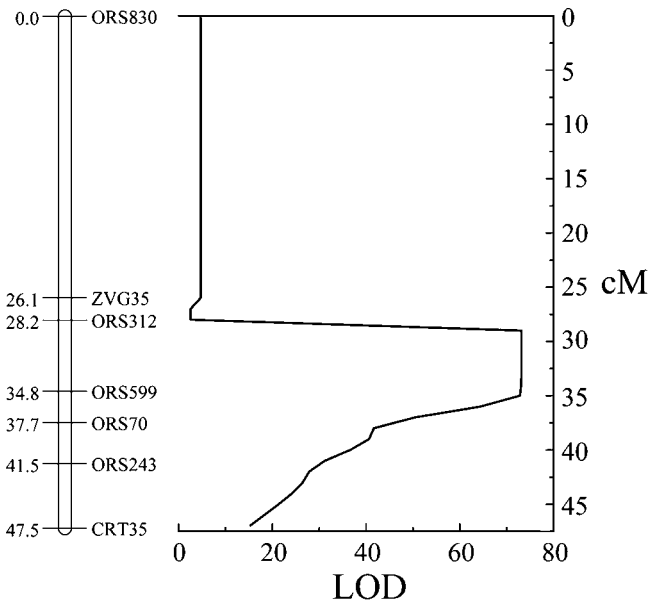


Fig. 4. Likelihood odds (LODs) for F_2 gamma-tocopherol QTL on linkage group (LG) 8 in CAS-12 × IAST-540.

Other genes controlling tocopherol biosynthesis have been mapped in sunflower. The Tph1 gene determining increased beta-tocopherol content in sunflower lines LG-15 and T589 (Demurin et al., 2004) was mapped to LG 1 of the sunflower genetic map (Vera-Ruiz et al., 2006). The results of the present research demonstrated that both the Tph1 and the Tph2 genes will segregate independently, since they are located in different linkage groups, and the identification of molecular markers linked to them provides an efficient system to select the *tph1-tph1tph2tph2* genotype. Recombination of *tph*1 and *tph*2 alleles produces novel tocopherol profiles of great potential value for sunflower oil quality. Thus, Demurin et al. (1996) reported the occurrence of 220 g kg^{-1} delta-tocopherol in segregants from the cross between the line LG-15 (*tph1tph1*) and the high gamma-tocopherol line LG-17 (*tph2tph2*), whereas Velasco et al. (2004b) reported levels of 700 g kg^{-1} beta-tocopherol and 580 g kg^{-1} delta-tocopherol, respectively, in two lines developed from the cross between the line T589 (*tph1tph1*) and the high gamma-tocopherol line IAST-1 (*tph2tph2*).

The development of different combinations of fatty acid and tocopherol profiles for specific end uses of the oil is now an interesting possibility in sunflower, since a wide range of fatty acid and tocopherol profiles are available (Fernández-Martínez et al., 2004). For that purpose, absence of linkage between genes controlling fatty acid and tocopherol profiles is desired. So far, only the Es3 gene associated with increased stearic acid levels in the CAS-14 mutant has been located on LG 8 of the sunflower genetic map (Pérez-Vich et al., 2006). This gene mapped between the ORS243 and the ORS1161 markers, and genetic distance between Es3 and Tph2 was estimated to be 11.5 cM. Even though this distance is large enough to obtain recombinants *es3es3tph2tph2* expressing both an increased stearic acid and gamma-tocopherol content, breeding for this phenotype would be easier with other sources of high stearic acid content determined by genes located at different linkage groups, for exam-

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

ple, CAS-3 with a major gene located at LG 1 (Pérez-Vich et al., 2002).

Alpha-tocopherol is synthesized from gamma-tocopherol by a methylation reaction mediated by the enzyme gamma-tocopherol methyl transferase, which also mediates the conversion of delta- to beta-tocopherol (DellaPenna, 2005). The gene encoding the enzyme has been cloned in *Arabidopsis* through a genomics-based approach and overexpressed in the *Arabidopsis* genome, which led to a drastic alteration in tocopherol profile, from 50 g kg^{-1} alpha-tocopherol and 950 g kg^{-1} gamma-tocopherol to 950 g kg^{-1} alpha-tocopherol and 50 g kg^{-1} gamma-tocopherol (Shintani and DellaPenna, 1998). Such an alteration of tocopherol profile is of similar magnitude but opposite direction to that occurring in the sunflower germplasm with high gamma-tocopherol content carrying *tph2* alleles, which suggests that gamma-tocopherol methyl transferase activity might be altered in this germplasm. The location of the Tph2 gene in the sunflower genetic map and the identification of molecular markers associated with it opens up the possibility of testing this hypothesis through map-based cloning or candidate gene strategies.

In summary, the present research concluded that the four sources of high gamma-tocopherol content identified so far in sunflower are allelic to each other, with recessive alleles at the Tph2 locus determining the trait. The gene has been mapped to LG 8 of the sunflower genetic map and molecular markers flanking the gene have been identified, which will facilitate marker-assisted selection in breeding programs focused on introgressing the trait into elite germplasm. Additionally, the results of the present research provide a basis for determining the function of the Tph2 gene in the tocopherol biosynthesis pathway.

ACKNOWLEDGMENTS

The authors thank Dr. Yakov Demurin (All-Russia Research Institute of Oil Crops, Krasnodar, Russia) for kindly providing seeds of LG-17, Dr. Steven J. Knapp (Center for Applied Genetic Technologies, Athens, GA), and Dr. Alberto J. Leon (Advanta Seeds, Buenos Aires, Argentina) for kindly providing non-published primer sequences of molecular markers, as well as Angustias Jiménez and Cristóbal Prieto for technical assistance. The research was supported by research project AGL2004-01765 and by a post-doctoral contract to Begoña Pérez-Vich from the Spanish Ramón y Cajal program (MEC-FEDER).

REFERENCES

Berry, S.T., A.J. Leon, C.C. Hanfrey, P. Challis, A. Burkholz, S.R. Barnes, G.K. Rufener, M. Lee, and P.D.S. Caligari. 1995. Molecular marker analysis of *Helianthus annuus* L.2. Construction of an RFLP linkage map for cultivated sunflower. Theor. Appl. Genet. 91: 195–199.

DellaPenna, D. 2005. A decade of progress in understanding vitamin E synthesis in plants. J. Plant Physiol. 162:729–737.

Demurin, Y. 1993. Genetic variability of tocopherol composition in sunflower seeds. Helia 16:59–62.

Demurin, Y., D. Škorić, and D. Karlovic. 1996. Genetic variability of tocopherol composition in sunflower seeds as a basis of breeding for improved oil quality. Plant Breed. 115:33–36.

Demurin, Y., S.G. Efimenko, and T.M. Peretyagina. 2004. Genetic identification of tocopherol mutations in sunflower. Helia 27:113–116.

Dolde, D., C. Vlahakis, and J. Hazebroek. 1999. Tocopherols in breeding lines and effects of planting location, fatty acid composition, and temperature during development. J. Am. Oil Chem. Soc. 76:349–355.

Fernández-Martínez, J.M., M. Mancha, J. Osorio, and R. Garcés. 1997. Sunflower mutant containing high levels of palmitic acid in high oleic background. Euphytica 97:113–116.

Fernández-Martínez, J.M., L. Velasco, and B. Pérez-Vich. 2004. Progress in the genetic modification of sunflower oil quality. p. 1–14. *In* Proc. 16th Int. Sunflower Conf., Fargo, ND, USA, 29 August–2 September 2004. International Sunflower Association, Paris.

Gedil, M.A., C. Wye, S.T. Berry, B. Seger, J. Peleman, R. Jones, A. Leon, M.B. Slabaugh, S.J. Knapp. 2001. An integrated RFLP-AFLP linkage map for cultivated sunflower. Genome 44:213–221.

Goffman, F.D., L. Velasco, and W. Thies. 1999. Quantitative determination of tocopherols in single seeds of rapeseed (*Brassica napus* L.). Fett/Lipid 101:142–145.

Hass, C.G., S.W. Leonard, J.F. Miller, M.B. Slabaugh, M.G. Traber, and S.J. Knapp. 2003. Genetics of tocopherol (Vitamin E) composition of mutants in sunflower. *In* Abstr. XI Plant Animal Genome Conf. San Diego, CA, USA. 11–15 Jan. 2003 http://www.intl-pag.org/11/abstracts/P7b_P821_XI.html; verified 6 July 2006).

Kamal-Eldin, A., and L.Å. Appelqvist. 1996. The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31:671–701.

Lander, E., P. Green, J. Abrahanson, A. Barlow, M. Daley, S. Lincoln, and L. Newburg. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

Lai, Z., K. Livingstone, Y. Zou, S.A. Church, S.J. Knapp, J. Andrews, and L.H. Rieseberg. 2005. Identification and mapping of SNPs from ESTs in sunflower. Theor. Appl. Genet. 111:1532–1544.

Michelmore, R.W., I. Paran, and V. Kesseli. 1991. Identification of markers linked to disease resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA 88:9828–9832.

Muggli, R. 1994. Vitamin E-Bedarf bei Zufuhr von Polyenfettsäuren. Fat Sci. Technol. 96:17–19.

Pérez-Vich, B., J.M. Fernández-Martínez, M. Grondona, S.J. Knapp, and S.T. Berry. 2002. Stearoyl-ACP and oleoyl-PC desaturase genes cosegregate with quantitative trait loci underlying high stearic and high oleic acid mutant phenotypes in sunflower. Theor. Appl. Genet. 104:338–349.

Pérez-Vich, B., B. Akhtouch, S.J. Knapp, A.J. Leon, L. Velasco, J.M. Fernández-Martínez, and S.T. Berry. 2004. Quantitative trait loci for broomrape (*Orobanche cumana* Wallr.) resistance in sunflower. Theor. Appl. Genet. 109:92–102.

Pérez-Vich, B., A. J. Leon, M. Grondona, L. Velasco, and J. M. Fernández-Martínez. 2006. Molecular analysis of the high stearic acid content in sunflower mutant CAS-14. Theor. Appl. Genet. 112: 867–875.

Pongracz, G., H. Weiser, and D. Matzinger. 1995. Tocopherole. Antioxidanten der Natur. Fat Sci. Technol. 97:90–104.

Shintani, D., and D. DellaPenna. 1998. Elevating the Vitamin E content of plants through metabolic engineering. Science 282:2098–2100.

Tang, S., J.K. Yu, M. Slabaugh, D.K. Shintani, and S.J. Knapp. 2002. Simple sequence repeat map of the sunflower genome. Theor. Appl. Genet. 105:1124–1136.

Tang, S., V.K. Kishore, and S.J. Knapp. 2003. PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower. Theor. Appl. Genet. 107:6–19.

Traber, M.G., and H. Sies. 1996. Vitamin E in humans: Demand and delivery. Annu. Rev. Nutr. 16:321–347.

Utz, H.F., and A.E. Melchinger. 1996. PLABQTL: A program for composite interval mapping of QTL. J. Quant. Trait Loci 2 (1) [Online]. Available at: www.uni-hohenheim.de/~ipspwww/soft.html; verified 12 July 2006.

Velasco, L., and J.M. Fernández-Martínez. 2003. Identification and genetic characterization of new sources of beta- and gamma-tocopherol in sunflower germplasm. Helia 26:17–24.

Velasco, L., J. Domínguez, and J.M. Fernández-Martínez. 2004a. Registration of T589 and T2100 sunflower germplasms with modified tocopherol profiles. Crop Sci. 44:361–362.

Velasco, L., B. Pérez-Vich, and J.M. Fernández-Martínez. 2004b. Novel variation for tocopherol profile in a sunflower created by mutagenesis and recombination. Plant Breed. 123:490–492.

Vera-Ruiz, E.M., B. Pérez-Vich, J.M. Fernández-Martínez, and L. Velasco. 2005. Comparative genetic study of two sources of beta-tocopherol in sunflower. Helia 28:1–8.

Vera-Ruiz, E.M., L. Velasco, A.J. Leon, J.M. Fernández-Martínez, and B. Pérez-Vich. 2006. Genetic mapping of the Tph1 gene controlling beta-tocopherol accumulation in sunflower seeds. Mol. Breed. 17:291–296.

Voorrips, R.E. 2002. MapChart: Software for the graphical presentation of linkage maps and QTL. J. Hered. 93:77–78.

Yu, J.K., S. Tang, M.B. Slabaugh, A. Heesacker, G. Cole, M. Herring, J. Soper, F. Han, W.C. Chu, D.M. Webb, L. Thompson, K.J. Edwards, S. Berry, A.J. Leon, C. Olungu, N. Maes, and S.J. Knapp. 2003. Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci. 43:367–387.

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

CAPÍTULO 2

Un gen modificador afecta al alto contenido en gamma-tocoferol en girasol

María J. García-Moreno, José M. Fernández-Martínez, Begoña Pérez-Vich y Leonardo Velasco

A modifying gene affecting gamma-tocopherol content in Sunflower

17th Int. Sunflower Conf., Córdoba, España, 8-12 Junio 2008

A modifying gene affecting gamma-tocopherol content in sunflower

María J. García-Moreno, José M. Fernández-Martínez, Begoña Pérez-Vich, Leonardo Velasco
[1]Instituto de Agricultura Sostenible (CSIC), Alameda del Obispo s/n, 14004 Córdoba, Spain.
E-mail: ia2veval@uco.es

ABSTRACT

High levels of gamma-tocopherol confer greater oxidative stability to sunflower oil. Several lines with high gamma-tocopherol content have been developed and in all of them the trait has been found to be controlled by recessive alleles at the *Tph2* locus, underlying a gamma-tocopherol methyltransferase. Genetic studies involving crosses between standard lines and lines with high gamma-tocopherol content reported bimodal segregation patterns with clear-cut classes including low gamma-tocopherol (<10%) and high gamma-tocopherol (>85%) levels, respectively but not intermediate levels. Intermediate gamma-tocopherol content ranging from 10 to 85% has only been reported in the isolation process of the line IAST-1, derived from a mutagenic treatment. The objective of this research was to investigate the occurrence of intermediate gamma-tocopherol content in crosses involving the high gamma-tocopherol line IAST-1. Plants of the high gamma-tocopherol lines T2100 and IAST-1, and the standard line HA89 were crossed and the F_2 seed generation was analysed for seed tocopherol profile. F_2 seeds from all the $F_{1:2}$ families from the crosses between HA89 and T2100 followed bimodal distributions with clear-cut classes fitting a 3:1 (<5%:>90%) ratio, corresponding to the expected segregation of the recessive alleles *tph2*. In addition to the 3:1 ratio, a 13:3 (<80%:>90%) ratio was identified in $F_{1:2}$ families from the crosses between HA89 and IAST-1, which included F_2 seeds with intermediate levels of gamma-tocopherol (5 to 80%). Intermediate levels of gamma-tocopherol were also observed in some $F_{1:2}$ families derived from the crosses between T2100 and IAST-1. The results suggested the presence of a modifying gene that produced intermediate gamma-tocopherol levels in combination with the *tph2* alleles.

Key words: gamma-tocopherol – modifying gene – oil quality – tocopherols

INTRODUCTION

Conventional sunflower seeds mainly contain alpha-tocopherol, which accounts for more than 90% of the total tocopherols. Several lines with modified tocopherol profiles have been developed. Demurin (1993) reported the lines LG-15 and LG-17, with increased levels of beta-tocopherol (50%) and gamma-tocopherol (95%), respectively. Both lines were developed from segregating accessions identified in a germplasm collection. Also in the course of germplasm evaluation, Velasco et al. (2004a) identified variations for beta- and gamma-tocopherol content, which allowed the development of the lines T589 and T2100, with increased levels of beta-tocopherol (>30%) and gamma-tocopherol (>85%), respectively. Additional variation for gamma-tocopherol content was created in sunflower by using chemical mutagenesis (Velasco et al., 2004b). The authors identified two M_2 seeds, derived from different M_1 plants, with increased gamma-tocopherol contents of 19.2% and 96.7%, respectively. M_3 progenies from the M_2 seed with 96.7% bred true for high gamma-tocopherol content, containing more than 90% gamma-tocopherol, which led to the development of the line IAST-540. M_3 progenies from the M_2 seed with 19.2% gamma-tocopherol segregated from zero to 84.6% gamma-tocopherol. Selection for high gamma-tocopherol content produced the line IAST-1, with stable high gamma-tocopherol content.

Genetic studies conducted by Demurin et al. (1996) concluded that the increased levels of beta-tocopherol were produced by recessive alleles at the *Tph1* locus, whereas the increased levels of gamma-tocopherol were the result of recessive alleles at the *Tph2* locus (Demurin et al., 1996). Similarly, Velasco and Fernández-Martínez (2003) reported the presence of recessive alleles at a single locus underlying the increased beta-tocopherol content in T589 and the high gamma-tocopherol content in T2100 seeds. Comparative genetic studies concluded that *tph1* alleles were present in both LG-15 and T589 lines (Demurin et al., 2004; Vera-Ruiz et al., 2005), and *tph2* alleles were present in the high gamma-tocopherol lines LG-17, T2100, IAST-540, and IAST-1 (Demurin et al., 2004; García-Moreno et al., 2006). The *Tph2* gene underlies a gamma-tocopherol methyltransferase (Hass et al., 2006). Genetic studies involving crosses between lines with high gamma-tocopherol content and lines with wild-type high alpha-tocopherol content have reported bimodal segregation patterns with clear-cut classes including low gamma-tocopherol (<10%) and high gamma-tocopherol (>85%) levels, respectively (Demurin et al., 1996; Velasco and Fernández-Martínez, 2003). Intermediate gamma-tocopherol content ranging from 10

to 85% in germplasm segregating for *tph2* alleles has only been reported so far in the isolation process of IAST-1. The objective of this research was to investigate the occurrence of intermediate gamma-tocopherol content in crosses involving the high gamma-tocopherol line IAST-1.

MATERIALS AND METHODS

The study included the sunflower lines T2100 and IAST-1, with high gamma-tocopherol content (>85%), and the standard line HA89, with high alpha-tocopherol content (>95%). T2100 was developed from an accession of the open pollinated cultivar 'Peredovik' (Velasco et al., 2004a). IAST-1 was isolated after chemical mutagenesis on seeds of several 'Peredovik' accessions (Velasco et al., 2004b). HA89 is an oilseed maintainer line released by the Texas Agricultural Experiment Station and the USDA-ARS in 1971.

Twenty-four half seeds of HA89, T2100, and IAST-1 were nondestructively analyzed for tocopherol profile, germinated and planted in pots under open air conditions in spring 2005. Plants of the three lines were crossed following an incomplete diallel design. Half seeds of the parents as well as F_1 half seeds were analysed for tocopherol profile. F_1 and parent half seeds were sown in March 2006 and the corresponding plants were grown in pots under open air conditions. F_1 plants were self-pollinated to obtain the F_2 generation.

Twenty-four to 96 F_2 half seeds from 12 to 24 F_1 plants from each cross were analysed for tocopherol profile following the procedure reported by Velasco et al. (2004b).

RESULTS AND DISCUSSION

Seeds of the high gamma-tocopherol lines T2100 and IAST-1 showed uniformly high gamma-tocopherol content (>95% of the total tocopherols). Seeds of the standard line HA89 showed uniformly high alpha-tocopherol content (>95%). F_2 seeds from all the $F_{1:2}$ families from the crosses between HA89 and T2100 followed bimodal distributions with clear-cut classes characterized by low (<5%) and high (>90%) gamma-tocopherol content that fitted a 3:1 (low:high) segregation ratio (Fig. 1), corresponding to the expected segregation of the recessive alleles *tph2* (Demurin et al., 1996; Velasco and Fernández-Martínez, 2003).

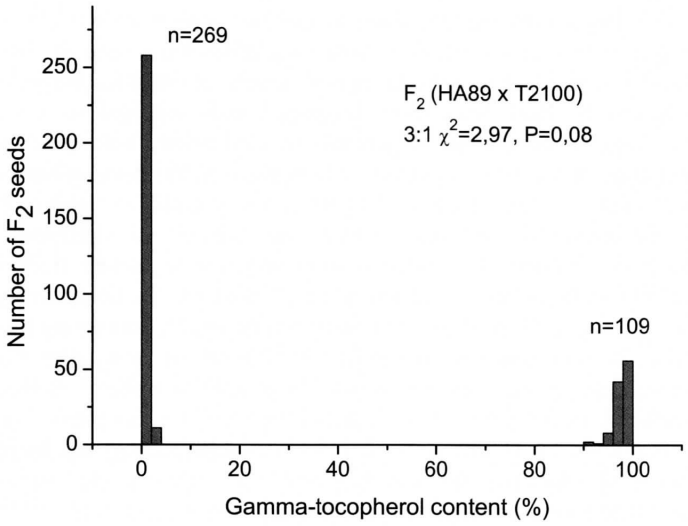


Fig. 1. Gamma-tocopherol content (% of the total tocopherols) in F_2 seeds from the cross between the high gamma-tocopherol line T2100 and the standard line HA89.

Two different segregation patterns were identified in $F_{1:2}$ families from the crosses between the standard line HA89 and the high gamma-tocopherol line IAST-1. The first segregation pattern was similar to that observed for the cross between HA89 and T2100, with F_2 seeds distributed into low and high gamma-tocopherol classes that fitted a 3:1 segregation ratio (Fig. 2A). The second segregation pattern showed the particularity of the presence of F_2 seeds with intermediate levels of gamma-tocopherol content (5 to 80%). The high gamma-tocopherol (>90%) class included 3 out of every 16 F_2 seeds (Fig.

2B), suggesting the presence of a second recessive gene that produced intermediate gamma-tocopherol levels in combination with the *tph2* alleles in a homozygous condition.

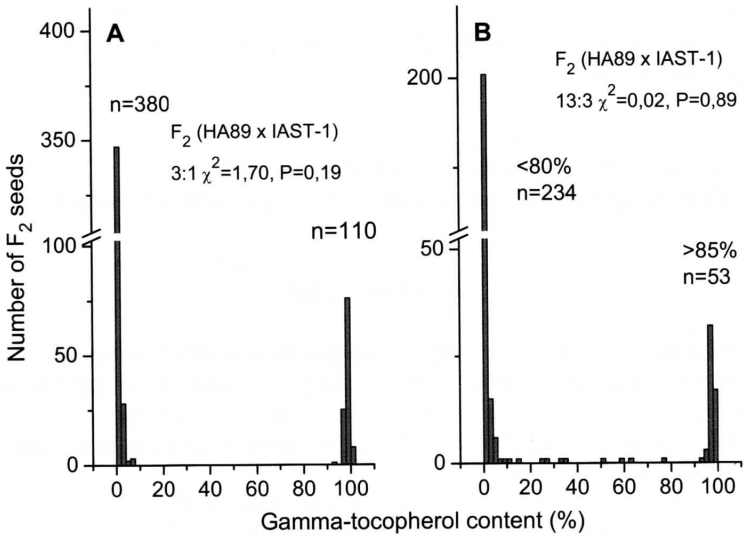


Fig. 2. Gamma-tocopherol content (% of the total tocopherols) in F_2 seeds from the cross between the high gamma-tocopherol line IAST-1 and the standard line HA89.

Two different patterns of gamma-tocopherol distribution were also identified in F_2 seeds from $F_{1:2}$ families derived from the crosses between the high gamma-tocopherol lines T2100 and IAST-1. F_2 seeds had uniformly high gamma-tocopherol content in some $F_{1:2}$ families (Fig. 3A), whereas other families showed segregation for a wide range of intermediate gamma-tocopherol levels (Fig. 3B).

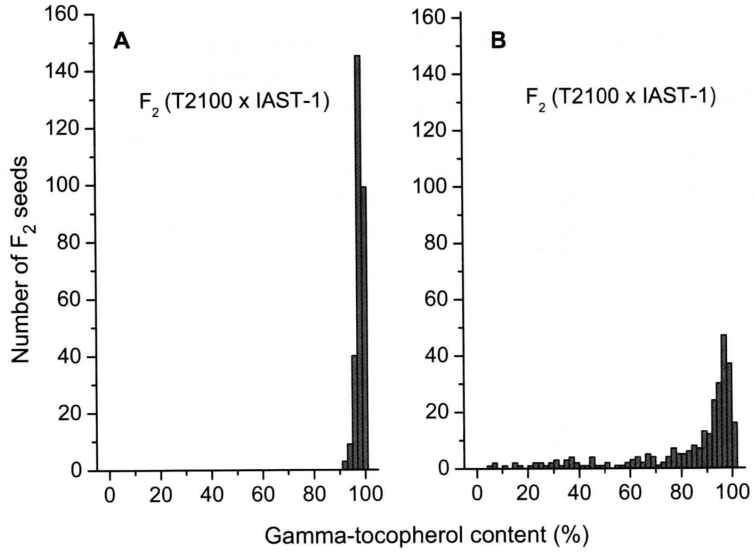


Fig. 3. Gamma-tocopherol content (% of the total tocopherols) in F_2 seeds from the cross between the high gamma-tocopherol lines T2100 and IAST-1.

A previous genetic study concluded that the high gamma-tocopherol lines T2100 and IAST-1 shared the alleles *tph2*, as both the F_1 and F_2 seed generations from crosses between them showed uniformly a high gamma-tocopherol content (García-Moreno et al., 2006). The present research work suggested the presence of a modifying gene affecting gamma-tocopherol content in IAST-1. The modifying gene produced a reduction in gamma-tocopherol content from high (>90%) to intermediate (5 to 80%) levels in

seeds with expected allelic configuration *tph2tph2*. This effect was observed in some F_1 plants from the crosses of IAST-1 with HA89 and T2100, but not in others. Additionally, the genetic effect of the modifying gene was not expressed in seeds of the IAST-1 parent grown in the same environment. Modifying genes affecting high oleic acid content have been reported in sunflower, leading to supression of the trait (Lacombe et al., 2001) or a strong distortion of segregation patterns (Fernández-Martínez et al., 1989). Further characterization of the modifying gene affecting high gamma-tocopherol content in sunflower is currently under way.

ACKNOWLEDGEMENTS

The research was supported by Ministerio of Educación y Ciencia, research project AGL2007-62834.

REFERENCES

Demurin, Y. 1993. Genetic variability of tocopherol composition in sunflower seeds. Helia 16:59-62.

Demurin, Y., D. Škorić, and D. Karlovic. 1996. Genetic variability of tocopherol composition in sunflower seeds as a basis of breeding for improved oil quality. Plant Breed. 115:33-36.

Demurin, Y., S.G. Efimenko, and T.M. Peretyagina. 2004. Genetic identification of tocopherol mutations in sunflower. Helia 27(40):113-116.

Fernández-Martínez, J.M., A. Jiménez, J. Domínguez, J.M. García, R. Garcés, and M. Mancha, 1989. Genetic analysis of the high oleic content in cultivated sunflower (*Helianthus annuus* L.). Euphytica 41:39-51.

García-Moreno, M., E.M. Vera-Ruíz, J.M. Fernández-Martínez, L. Velasco, and B. Pérez-Vich. 2006. Genetic and molecular analysis of high gamma-tocopherol content in sunflower. Crop Sci. 46:2015-2021.

Hass, C.G., S. Tang, S. Leonard, M.G. Traber, J.F. Miller, and S.J. Knapp. 2006. Three non-allelic epistatically interacting methyltransferase mutations produce novel tocopherol (vitamin E) profiles in sunflower. Theor. Appl. Genet. 113:767-782.

Lacombe, S., F. Kaan, L. Léger, and A. Bervillé, 2001. An oleate desaturase and a suppressor loci direct high oleic acid content of sunflower (*Helianthus annuus* L.) oil in the Pervenets mutant. C.R. Acad. Sci. Paris, Life Sci. 324, 839-845.

Velasco, L., and J.M. Fernández-Martínez. 2003. Identification and genetic characterization of new sources of beta- and gamma-tocopherol in sunflower germplasm. Helia 26(38):17-24.

Velasco, L., J. Domínguez, and J.M. Fernández-Martínez. 2004a. Registration of T589 and T2100 sunflower germplasms with modified tocopherol profiles. Crop Sci. 44:361-362.

Velasco, L., B. Pérez-Vich, and J.M. Fernández-Martínez. 2004b. Novel variation for tocopherol profile in a sunflower created by mutagenesis and recombination. Plant Breed. 123:490-492.

Vera-Ruiz, E.M., B. Pérez-Vich, J.M. Fernández-Martínez, and L. Velasco. 2005. Comparative genetic study of two sources of beta-tocopherol in sunflower. Helia 28(42):1-8.

CAPÍTULO 3

Bases genéticas de la expresión inestable del alto contenido en gamma-tocoferol en semillas de girasol

María J. García-Moreno, José M. Fernández-Martínez, Leonardo Velasco y Begoña Pérez-Vich

Genetic basis of unstable expression of high gamma-tocopherol content in sunflower seeds

Enviado a BCM Plant Biology

Genetic basis of unstable expression of high gamma-tocopherol content in sunflower seeds

María J García-Moreno, José M Fernández-Martínez, Leonardo Velasco, Begoña Pérez-Vich[§]

Instituto de Agricultura Sostenible (IAS-CSIC), Alameda del Obispo s/n, 14004 Córdoba, Spain

[§]Corresponding author

E-mail addresses:

MJG-M: mjgarcia@ias.csic.es

JMF-M: jfernandezm@ias.csic.es

LV: lvelasco@ias.csic.es

BP-V: bperez@ias.csic.es

Abstract

Background

Tocopherols are natural antioxidants with both *in vivo* (vitamin E) and *in vitro* activity. Sunflower seeds contain predominantly alpha-tocopherol (>90% of total tocopherols), with maximum vitamin E effect but lower *in vitro* antioxidant action than other tocopherol forms such as gamma-tocopherol. Sunflower germplasm with stable high levels of gamma-tocopherol (>85%) has been developed. The trait is controlled by recessive alleles at a single locus *Tph2* underlying a gamma-tocopherol methyltransferase (gamma-TMT). Additionally, unstable expression of increased gamma-tocopherol content in the range from 5 to 85% has been reported. The objective of this research was to determine the genetic basis of unstable expression of high gamma-tocopherol content in sunflower seeds.

Results

Male sterile plants of nuclear male sterile line nmsT2100, with stable high gamma-tocopherol content, were crossed with plants of line IAST-1, with stable high gamma-tocopherol content but derived from a population that exhibited unstable expression of the trait. F_2 seeds showed continuous segregation for gamma-tocopherol content from 1.0 to 99.7%. Gamma-tocopherol content in F_2 plants (average of 24 individual F_3 seeds) segregated from 59.4 to 99.4%. A genetic linkage map comprising 17 linkage groups (LGs) was constructed from this population using 109 SSR and 20 INDEL marker loci, including INDEL markers for tocopherol biosynthesis genes. QTL analysis revealed a major QTL on LG 8 that corresponded to the gamma-TMT *Tph2* locus, which suggested that high gamma-tocopherol lines nmsT2100 and IAST-1 possess different alleles at this locus. Modifying genes were identified at LGs 1, 9, 14 and 16, corresponding in most cases with gamma-TMT duplicated loci.

Conclusions

Unstable expression of high gamma-tocopherol content is produced by the effect of modifying genes on *tph2*^{a} allele at the gamma-TMT *Tph2* gene. This allele is present in line IAST-1 and is different to allele *tph2* present in line nmsT2100, which is not affected by modifying genes. No sequence differences at the gamma-TMT gene were found associated to allelic unstability. Our results suggested that modifying genes are mostly epistatically interacting gamma-TMT duplicated loci.

Background

Tocopherols are the main antioxidants present in seed oils. They form a family of four fat-soluble compounds with vitamin E activity named alpha-, beta-, gamma-, and delta tocopherol. Tocopherols contain a 6-chromanol ring structure methylated to varying degrees at positions 5, 7, and 8, and an isoprenoid-derived C16 saturated side chain at position 2. The four tocopherols differ by the number and positions of the methyl groups on the 6-chromanol ring. Alpha-tocopherol is trimethylated, beta- and gamma-tocopherol are dimethylated, and delta-tocopherol is monomethylated [1].

Tocopherols are important antioxidants operating both *in vivo*, protecting cells from oxidative damage, as well as *in vitro*, protecting oils and oil-based products from oxidation [2]. The relative *in vivo* and *in vitro* antioxidant properties of the specific tocopherols is determined by their chemical structure. The relative biological activity of the tocopherols is estimated as 100% for alpha-tocopherol, 15 to 27% for beta-tocopherol, 3 to 20% for gamma-tocopherol, and 0.3 to 2% for delta-tocopherol [3]. However, there is a widespread confusion concerning their relative potency *in vitro* [4], though it is generally accepted that alpha-tocopherol shows better antioxidant activity in fats and oils at lower concentrations, but at

higher concentrations gamma-tocopherol is a more active antioxidant [5]. Studies conducted in sunflower seed oil, in which alpha-tocopherol accounts for more than >90% of the total tocopherols, concluded that substitution of alpha-tocopherol by gamma-tocopherol has a positive impact on the stability of the oil [6-9].

Four sunflower germplasms named LG-17, T2100, IAST-1, and IAST-540 in which alpha-tocopherol in the seeds was almost completely replaced by gamma-tocopherol have been developed [6, 10-11]. Gamma-tocopherol in the novel germplasm accounts for more than 85% of the total seed tocopherols, compared with more than 90% alpha-tocopherol in conventional sunflower seeds. The increased gamma-tocopherol levels are produced by recessive alleles at the *Tph2* locus [6, 12-13], which encodes a gamma-tocopherol methyltransferase (gamma-TMT) enzyme [14]. This enzyme catalyzes the methylation of delta- and gamma-tocopherol to yield beta- and alpha-tocopherol, respectively [15]. Gamma-TMT mutation in sunflower disrupts the synthesis of alpha-tocopherol and causes the accumulation of gamma-tocopherol [14]. In a detailed sequence analysis of the gamma-TMT gene in sunflower, Hass et al. [14] identified two gamma-TMT paralogs (*gamma-TMT-1* and *gamma-TMT-2*) and five different haplotypes (haplotypes 4 and 5 corresponding to paralogs 1 and 2, respectively). Both gamma-TMT paralogs 1 and 2 cosegregated with *Tph2* and were mapped to linkage group (LG) 8 of the sunflower linkage map. Even though the *Tph2* mutation reduced or disrupted the expression of the two paralogs in developing sunflower seeds, none of the DNA polymorphisms found within the gamma-TMT *Tph2* gene were associated with the high gamma-tocopherol phenotype [14]. The authors suggested that the *Tph2* mutation must be very tightly linked to the gamma-TMT locus on LG 8 and may disrupt regulatory sequences.

Phenotypic studies of Demurin et al. [16] and García-Moreno et al. [13] concluded that the four high gamma-tocopherol lines LG-17, T2100, IAST-1, and IAST-540 possess the

same allele at *Tph2*, as no transgressive segregations were observed in crosses involving the four lines. Lines LG-17, T2100 and IAST-540 were isolated from germplasm or M_1 mutant plants that segregated for high alpha-tocopherol (>90%) and high gamma-tocopherol (>85%), with no intermediate levels of both tocopherol forms being observed [6, 10-12]. Conversely, IAST-1 was isolated from an M_2 plant that exhibited large variation for alpha- and gamma-tocopherol levels, with gamma-tocopherol content in M_3 seeds showing a continuous variation from zero to 84.6% and M_4 seeds from selected M_3 plants showing a variation from 60.4 to 97.4% gamma-tocopherol, which was uniformly high (>90%) in M_5 seeds from selected M_4 plants [11]. García-Moreno et al [17] suggested that the intermediate gamma-tocopherol levels observed during the isolation of IAST-1 might be produced by the presence of modifying genes that determined unstable expression of mutated alleles at *Tph2* locus. Modifying genes have been found to influence important traits in sunflower such as high oleic acid content [18] or broomrape resistance [19].

In this study, the genetic analysis of a population that showed segregation from low to high gamma-tocopherol values obtained from the cross between the two high gamma-tocopherol lines IAST-1 and nmsT2100 has been carried out. In the course of this analysis, we identified two different alleles at the gamma-TMT *Tph2* locus at LG 8. The allele *tph2* was present in line nmsT2100, whereas the allele $tph2^a$ was identified in line IAST-1. Additionally, we found four modifying genes at LGs 1, 9, 14 and 16 that in most cases corresponded to duplicated gamma-TMT loci. Modifying genes influenced the expression of $tph2^a$ alleles, but did not affect *tph2* alleles.

Results

Phenotypic segregations

Seeds of sunflower lines nmsT2100 (fertile plants) and IAST-1 showed uniformly high gamma-tocopherol content, from 91.2 to 99.8% in nmsT2100 and from 92.3 to 99.4% in IAST-1. F_1 seeds from the cross between nmsT2100 and IAST-1 had also high gamma-tocopherol content, from 92.6 to 97.2%. However, large segregation for gamma-tocopherol content was observed in F_2 seeds from some F_1 plants. Particularly, the analysis of 192 F_2 seeds from an F_1 plant derived from an F_1 seed with 95.5% gamma-tocopherol revealed large segregation for the trait, from 1.0 to 99.7%, with no discrete phenotypic classes being observed (Figure 1). Variation for gamma-tocopherol content was smaller at the F_2 plant generation (average of 24 individual F_3 seeds), which segregated from 59.4 to 99.4% (Figure 2). However, examination of variation of individual F_3 seeds within each F_2 plant showed that minimum gamma-tocopherol content in individual F_3 seeds from the different F_3 families ranged from 0.0 to 98.4%, whereas maximum gamma-tocopherol content was in all cases above 94%. No discrete classes could be distinguished for minimum gamma-tocopherol content in the $F_{2:3}$ population (data not shown).

Map construction and candidate gene mapping

One hundred and sixty three out of 437 (37.3%) simple sequence repeat (SSR) and insertion-deletion (INDEL) markers were polymorphic in the screening of the high gamma-tocopherol lines nmsT2100 and IAST-1 and four randomly selected F_2 individuals from their cross. A set of 128 high quality, evenly spaced, and preferably co-dominant markers were genotyped in the F_2 population. Final linkage map for QTL analyses comprised 129 marker loci (109 SSR and 20 INDEL) grouped on 17 linkage groups (Table 1). The total genetic distance covered by these marker loci was 792.4 cM, with an average marker interval of 7.0 cM (Table 1). Linkage groups contained between 3 and 16 marker loci, with 99.6% of the mapped genome

being within 20 cM to the nearest marker. No marker loci had significantly distorted segregation ratios ($P<0.001$).

INDEL markers for the gamma-TMT gene described by Hass et al. [14] were not polymorphic between nmsT2100 and IAST-1 or showed polymorphism in secundary loci. Accordingly, new primer combinations based on the full-lenght sequence of the gamma-TMT gene in nmsT2100 and IAST-1 (see below) were designed and tested. A primer combination (gamma-TMT-F1/F2/R24) based on the forward primers F1 and F2 placed upstream and dowstream, respectively, of the T39 transcription initiation site [14], and a reverse primer R24 complementary to the DNA sequence in exon 2 amplified a primary gamma-TMT locus that was polymorphic between IAST-1 and nmsT2100 (Figure 3). The polymorphic band was about 1460 bp in IAST-1 and 1420 bp in nmsT2100 (Figure 3). When screened against the individuals from the F_2 population, the gamma-TMT-F1/F2/R24 locus was co-dominantly mapped to LG 8 at the previously described position of the gamma-TMT *Tph2* gene [14]. Other INDEL markers for the gamma-TMT gene were also tested. The F9/R24 gamma-TMT INDEL marker showed three loci in the nmsT2100 and IAST-1 parental lines (Figure 4). A primary locus (gamma-TMT-F9/R24a) corresponding to a band of about 1200 bp was not polymorphic with this primer combination (Figure 4). This locus, however, was demostrated to co-segregate with *Tph2* and was co-dominantly mapped to LG 8 in the population CAS-12 x IAST-540, also segregating for gamma-tocopherol content [13]. A second locus (gamma-TMT-F9/R24b) that showed a dominant polymorphism, with a band of about 750 bp present in nmsT2100 and absent in IAST-1, was mapped to LG 16 (Figure 4). The locus was also mapped to LG 16 in the CAS-12 x IAST-540 population. Finally, a third locus (gamma-TMT-F9/R24c) that also showed a dominant polymorphism, with a band of 375 bp absent in nmsT2100 and present in IAST-1, was mapped to LG 1 (Figure 4). In addition to the gamma-TMT-F9/R24a, b and c loci, a fourth gamma-TMT locus was identified in populations other

than nmsT2100 x IAST-1. This locus was named gamma-TMT-F9/R24d and was mapped to LG 14 in populations CAS-12 x IAST-540 [13] and IAST-413 x HA-89 (Del Moral L, unpublished data) (Figure 4). Finally, the locus *MT-2* of the 2-methyl-6-phytyl-1,4-benzoquinone/2-methyl-6-solanyl-1,4-benzoquinone methyltransferase (MPBQ/MSBQ-MT) gene was mapped to LG 4 using the F24/R25 INDEL marker [20]. None of the other markers for tocopherol biosynthesis genes tested were polymorphic between nmsT2100 and IAST-1.

QTL analyses

One-factor analysis of variance revealed that the gamma-TMT-F1/F2/R24 locus on LG 8 was underlying a major QTL affecting gamma-tocopherol content. This locus had a significant effect on gamma-tocopherol content of the F_2 and the F_3 generations (Table 2) and corresponded to the gamma-TMT *Tph2* gene also mapped to the same position on LG 8 by Hass et al. [14]. The mean and the standard deviation for gamma-tocopherol content in the F_2 seed generation of plants homozygous for the IAST-1 allele were 52.3±29.6%, whereas those for plants homozygous for the nmsT2100 allele were 98.9±0.8% (Table 2). These results suggested the existence to two different alleles at the gamma-TMT *Tph2* locus in lines nmsT2100 and IAST-1. The allele at IAST-1 was unstable, producing in homozygous condition a broad distribution ranging from 0.97 to 89.02% in the F_2 seed generation, whereas the allele at nmsT2100 was highly stable and resulted in gamma-tocopherol values above 97%. Similar results were observed in the F_3 seed generation (Table 2). It is interesting to note that only three F_2 plants homozygous for the nmsT2100 allele produced sufficient number of F_3 seeds for tocopherol analyses, compared to 27 F_2 plants homozygous for the IAST-1 allele (Table 2). This was probably caused by a close linkage between *Tph2* and *Ms11* loci at LG 8, the latter responsible for male sterility in the nuclear male sterile (NMS) line P21 [21] from which *Ms11* was introgressed to nmsT2100.

Composite interval mapping analyses confirmed the existence of a QTL with a main effect centered on the gamma-TMT-F1/F2/R24 locus on LG 8 (Table 3 and Figure 5). This QTL was named *Tph2-gamma-TMT* and explained 41.2% and 44.4% of the F_2 and F_3 gamma-tocopherol phenotypic variance, respectively (Table 3). A second QTL peak on LG 8 16 cM apart from *Tph2-Gamma-TMT* was identified only in the QTL analysis of the F_3 data (Figure 5). This adjacent peak might be a ghost QTL resulting from the distorted F_3 data in this region and was not taken into consideration to protect against type I errors in declaring QTLs.

Besides *Tph2-Gamma-TMT*, four other QTL with a moderate effect on gamma-tocopherol content were observed at the F_2 level. These QTL were located at LG 1 (*GamT1.1*), 9 (*GamT9.1)*, 14 (*GamT14.1*), and 16 (*GamT16.1*). They individually accounted for 2.3-4.1% of trait variation (Table 3 and Figure 5). The QTL on LG 9, 14 and 16 were also detected in the F_3 analyses at practically identical positions (Table 3 and Figure 5). The model with all the QTL explained 44.4% and 61.2% of the phenotypic variance for gamma-tocopherol content in the F_2 and the F_3 generations, respectively. The QTL peaks of *GamT1.1* and *GamT16.1* were centered on the gamma-TMT loci mapped at these linkage groups (gamma-TMT-F9/R24c in LG1 and gamma-TMT-F9/R24b in LG 16) (Figure 5). Additionally, the QTL peak of *GamT14.1* was likely to correspond to the gamma-TMT-F9/R24d locus mapped to LG 14 in populations CAS-12 x IAST-540 [13] and IAST-413 x HA-89 (Del Moral L, unpublished data), located 1.5 cM from the ZVG64 marker locus at this LG.

Analysis for epistasis in the F_2 generation through two-way ANOVA revealed significant interactions for gamma-tocopherol content involving the *Tph2-gamma-TMT* marker locus on LG 8 and markers on LG 1, LG 14 and LG 16. The most significant interactions involved the ORS185 marker locus on LG 14 (F=10.5, P<0.0001), the gamma-

TMT_F9R24b locus on LG 16 (*F*=4.22, *P*=0.017), and the gamma-TMT_F9R24c locus on LG 1 (*F*=3.12, *P*=0.048), which were associated to the *GamT14.1*, *GamT16.1*, and *GamT1.1* QTL, respectively. Significant interactions involving the *Tph2-Gamma-TMT* QTL on LG 8, the *gamT14.1* on LG 14, and the *gamT16.1* QTL on LG 16 were also detected in the composite interval mapping analyses, with the percentage of explained phenotypic variance increasing from 44.4% to 52.4% in the F_2 and from 61.2% to 71.3% in the F_3 when epistatic interactions were included in the multiple-locus model (Table 3). Genotypic means for F_2 gamma-tocopherol content in allelic combinations of the epistatically interacting marker loci were calculated to detail these interactions. Since ORS185 and gamma-TMT_F9R24b marker loci were dominant, genotypic means were computed using close co-dominant markers, ORS578 on LG 14 and ORS126 on LG 16 (Table 4). The results showed that QTL regions on LG 14 and LG 16 only had a significant phenotypic effect when the IAST-1 allele was present at the *Tph2-gamma-TMT* locus on LG 8 (gamma-TMT-F1/F2/R24). Thus, no significant differences were observed between different genotypes at ORS578 on LG 14 or ORS126 on LG 16 in presence of the nmsT2100 allele at *Tph2*. However, when the *Tph2* locus was homozygous for the IAST-1 allele, the presence of nmsT2100 alleles at ORS578 or ORS126 resulted in a drastic reduction of gamma-tocopherol content as compared to the presence of IAST-1 alleles (Table 4).

Sequence analysis of gamma-TMT loci

Several loci amplified by the gamma-TMT F9/R24 INDEL marker in IAST-1 and nmsT2100 (Figure 4) were cloned and sequenced. A band of about 1200 bp from the non-polymorphic locus *a* and a band of about 375 bp from the dominant locus *c* that mapped to LG 1 were sequenced from IAST-1. A band of about 1200 bp from the non-polymorphic locus *a* and a band of about 750 bp from the dominant locus *b* that mapped to LG 16 were sequenced from

nmsT2100. The locus *d* that mapped to LG 14 was sequenced from lines IAST-540 and HA-89.

The locus *a* fragment isolated from nmsT2100 harboured alleles from both gamma-TMT paralogs, which showed showed 100% sequence identity to gamma-TMT haplotype 4 (paralog 1) and 5 (paralog 2) from Hass et al. [14]. For IAST-1, the locus *a* sequence showed 100% sequence identity to gamma-TMT haplotype 4. The locus *b* consensus sequence was 785 bp long and showed a significant homology to gamma-TMT haplotypes 2 and 3 (GenBank accessions nos. DQ229829 and DQ229830, $3e^{-58}$), with 94% maximum sequence identity spanning 61% of the locus *b* fragment coverage. The consensus sequence for locus *c* (388 bp long) showed a significant homology to gamma-TMT haplotypes 4 and 5 (GenBank accessions nos. DQ229831 to DQ229834, $4e^{-14}$), with 86% maximum sequence identity spanning 19% of the locus *c* fragment coverage. The sequences of locus *d* from IAST-540 (405 bp) and HA-89 (406 bp) were very similar, with a 96% of sequence identity between them. The consensus sequence obtained from this alignment showed significant homology to gamma-TMT haplotypes 1, 3, 4 and 5 (GenBank accession nos. DQ229828, DQ229830, and DQ229831 and DQ229834; $8e^{-18}$), with 89% maximum sequence identity spanning 18% of the fragment coverage.

Nucleotide sequences from IAST-1 locus *a*, nmsT2100 locus *a*, nmsT2100 locus *b*, IAST-1 locus *c*, HA-89 locus *d*, and IAST-540 locus *d* were aligned. The sequence alignment tree (Figure 6) revealed great similarity between loci *a* and *b* (85 to 87% of sequence identity) and between loci *c* and *d* (92-93% of sequence identity). Both groups were more distantly related, with 56 to 58% of sequence identity between locus *a* and the cluster of loci *c* and *d*.

Full-length gamma-TMT genomic DNA sequences

Agarose gels from PCR products obtained with the F1 and R92 primers flanking the gamma-TMT gene mapped to LG 8 [14] revealed two bands in nmsT2100 and IAST-1 lines that were identified as the two gamma-TMT paralogs 1 and 2 described by Hass et al. [14]. Consensus nucleotide sequences for IAST-1 gamma-TMT paralog 1 (4126 bp) and paralog 2 (4280 bp) were identical to gamma-TMT haplotype 4 (GenBank accessions nos. DQ229831 and DQ229832) and 5 (GenBank accessions nos. DQ229833 to DQ229834), respectively, from Hass et al. [14] (Figure S1). Consensus nucleotide sequence for nmsT2100 gamma-TMT paralog 1 (4126 bp) was identical to gamma-TMT haplotype 4 [14] (Figure S1). Consensus nucleotide sequence for nmsT2100 gamma-TMT paralog 2 (4281 bp) was almost identical to gamma-TMT haplotype 5 [14], with the exception of five SNPs, one in the 5´UTR (G to A), one in intron 1 (C to T), and three in intron 4 (A to T) (Figure S1).

Discussion

The results of this research suggest that high gamma-tocopherol lines IAST-1 and nmsT2100 possess different mutated alleles at the gamma-TMT *Tph2* locus on LG 8. The gamma-TMT catalyses the methylation step from gamma-tocopherol to alpha-tocopherol in sunflower seeds [14]. Mutated *tph2* alleles disrupt the activity of gamma-TMT, which results in accumulation of gamma-tocopherol [14]. The allele at nmsT2100 (*tph2*) is stable and not affected by modifying genes, whereas the allele at IAST-1 ($tph2^{a}$) is unstable and affected by modifying genes. Putative modifying genes located at LG 1, 9, 14 and 16 were identified in the present research, being the effect of those on LG 1, 14 and 16 highly significant in the epistatic interaction with the *Tph2* locus on LG 8. Studies on expression of *Tph2* mutations in plant tissues other than seeds also pointed to differences between T2100 and IAST-1 lines, as the latter showed lower gamma-tocopherol content in leaves, roots, and pollen [22].

High gamma-tocopherol lines T2100 and IAST-1 were isolated following different strategies. T2100 derived from an open-pollinated cultivar that showed variation for high gamma-tocopherol content (>85%) at the single-seed level [9]. Genetic study of the trait in T2100 seeds indicated that the trait was controlled by recessive alleles at a single locus *Tph2* in such a way that *Tph2Tph2* and *Tph2tph2* genotypes produced low levels of gamma-tocopherol (<5%) and *tph2tph2* genotypes showed high levels of gamma-tocopherol (>85%), with no intermediate levels being observed [12]. On the contrary, IAST-1 derived from a mutagenesis program in which an M_2 seed with intermediate gamma-tocopherol content (19.2%) was identified. The trait showed continuous variation for gamma-tocopherol levels (zero to 84.5%) in the M_3 generation, which allowed selection of genotypes with stable high gamma-tocopherol content in the M_4 generation [11]. A comparative genetic study between both lines concluded that they shared the same allele at *Tph2*, as no transgressive segregation was observed in the F_1 and F_2 generations from crosses between them [13]. However, the evaluation of several F_2 populations from crosses of IAST-1 with T2100 and the conventional sunflower line HA89 showed that intermediate gamma-tocopherol values (5 to 85%) occurred in some F_2 populations, whereas other F_2 populations showed uniformly high gamma-tocopherol content (IAST-1 x T2100) or segregated into well-defined low and high gamma-tocopherol classes (IAST-1 x HA89) [17]. The results of this research suggest that intermediate gamma-tocopherol levels are produced by the effect of modifying genes on $tph2^a$ alleles from IAST-1. We hypothesize that modifying genes were present in the mutagenized population from which IAST-1 was developed, where selection against negative alleles at modifying genes led to the isolation of IAST-1, and also in other lines such as HA89 and T2100. Modifying genes have no phenotypic effect in absence of alteration in the major gene [23]. In this research, it was also found that modifying genes have no phenotypic effect on the mutated allele at T2100 (*tph2*). Accordingly, they are expected to be segregating in HA89 and

T2100 lines, which produced different segregation patterns in crosses with IAST-1 [17]. The occurrence of different segregation patterns depending on allelic configuration of modifying genes has been previously reported in sunflower for high oleic acid content [24-25] and broomrape resistance [19].

Full length genomic gamma-TMT sequences corresponding to the *Tph2* locus on LG 8 have been obtained in this study for both IAST-1 and nmsT2100. Both lines have two gamma-TMT paralogs. Paralog 1 is identical between the two lines and also to the gamma-TMT haplotype 4 from Hass et al. [14], which included sequences from both high and low gamma-tocopherol lines. Paralog 2 from IAST-1 is identical to gamma-TMT haplotype 5 from Hass et al. [14], which also included sequences from both high and low gamma-tocopherol lines, while paralog 2 from nmsT2100 carries slight SNP differences. Consequently, we have not found any sequence difference in the gamma-TMT gene from IAST-1 that would explain why the $tph2^a$ allele present in this line is unstable and affected by modifying genes. Previous research did not identify sequence differences in the gamma-TMT gene related to the high gamma-tocopherol phenotype [14]. The authors found identical gamma-TMT paralog 1 and 2 sequences in both high and low gamma-tocopherol lines and reported that mutation leading to the high gamma-tocopherol trait in the sunflower material derived from the high gamma-tocopherol line LG-17 [6] may disrupt regulatory sequences of the gamma-TMT gene. Relating the nmsT2100 line, the slight sequence differences in gamma-TMT paralog 2 were changes to those nucleotides also present in paralog 1 sequences from IAST-1 and nmsT2100 (2 SNPs in the 5´UTR region and in intron 1) or found in a repetitive sequence within an intron (3 SNPs in intron 4) and were not likely to be involved in determining a more stable gamma-TMT allele.

Modifying genes affecting important traits for sunflower breeding such as high oleic acid content [24-25] and broomrape resistance [19] have been reported. The discovery of

modifying genes affecting high gamma-tocopherol content confirms that the occurrence of modifying genes is not a rare phenomenon in sunflower genetics. Thus far there was no indication on the nature and mode of action of modifying genes. The results of the present research suggest that *Tph2* modifying genes are in most cases additional gamma-TMT loci duplicated in the sunflower genome. This was confirmed for modifying genes at LG 1 and 16, whereas there is also indication that the modifying gene at LG 14 might be an additional copy of a gamma-TMT. Duplicated gamma-TMT loci have also been found in safflower [26] and rapeseed [27]. In fact, the two gamma-TMT loci from safflower (one of them co-segregating with the safflower *Tph2* gene determining high gamma-tocopherol content in this crop) were identified by using the F9/R24 primer combination from the sunflower gamma-TMT [26], the same one used in this study to map different gamma-TMT loci. These results suggested that the F9 and R24 primer sequences are within a highly conserved region of the gamma-TMT gene. In sunflower, Hass et al. [14] mapped, in addition to *gamma-TMT-Tph2* on LG 8, another gamma-TMT locus on LG 16, although the authors considered this locus a randomly amplified polymorphic DNA (RAPD). In this study, the locus on LG 16 (gamma-TMT locus *b*) showed significant homology to the sunflower gamma-TMT gene and underlay a gamma-tocopherol QTL. Our results indicated that interaction between duplicated gamma-TMT loci revert the high gamma-tocopherol phenotype to intermediate-low gamma-tocopherol values. This effect has also been reported for modifying genes affecting the high oleic acid phenotype in sunflower, where modifying genes suppressed the effect of the *Ol-FAD2-1* allele that is essential for accumulating high oleic acid content [18, 28]. Also, epistatically interacting duplicated MPBQ/MSBQ-MT genes that modify beta-tocopherol levels have been reported in sunflower [20]. It is well known that duplicate gene pairs can form negative epistasis due to their overlapping functions [29] and that suppression of a mutant phenotype can be altered by

gene dosage [30]. However, additional biochemical, functional, and sequence analyses are required to determine the mode of action of duplicated gamma-TMT loci in sunflower.

From a breeding perspective, the existence of several modifying genes influencing high gamma-tocopherol content should not be a constraint for the development of cultivars with seeds rich in gamma-tocopherol, as this research revealed that modifying genes do not alter the phenotypic expression of the mutated allele *tph2* present at nmsT2100. On the other hand, the influence of modifying genes on expressivity of mutated $tph2^a$ alleles opens up the possibility of selection for balanced levels of alpha- and gamma-tocopherol content, provided that stable combinations of modifying genes can be obtained. The development of mid oleic acid hybrids of sunflower was based on genetic stocks carrying modifying genes that limit the expression of high oleic acid content produced by mutated alleles at *Ol* locus (>80%) to the mid-range level (55-75%) [31]. In this sense, determination of allelic diversity at gamma-TMT loci underlying gamma-tocopherol modifying genes present in hybrid parental lines might be a useful tool to predict the presence of adequate allelic combinations giving rise to the desired levels of gamma-tocopherol content. Further studies should elucidate whether the other two gamma-tocopherol sources developed thus far, IAST-540 and LG-17, possess stable or unstable alleles at *Tph2*. Demurin et al. [6] reported differences in the expressivity of the recessive homozygotes of the *Tph2* gene from LG-17 in different genetic backgrounds, which might be indicative of allelic instability.

Conclusions

The results of this research suggest that high gamma-tocopherol in sunflower lines IAST-1 and nmsT2100 is determined by different mutated alleles at the gamma-TMT *Tph2* locus on LG 8. The allele at nmsT2100 is not affected by modifying genes, whereas the allele at IAST-1 is unstable and affected by modifying genes. Putative modifying genes have been located at

LG 1, 9, 14 and 16, being the effect of those on LG 1, 14 and 16 highly significant in their epistatic interaction with the *Tph2* locus on LG 8. Finally, our results suggest that modifying genes correspond to gamma-TMT loci duplicated in the sunflower genome. Phenotypic effects of modifying genes altering the expression of important genes in crop plants are known [23], but in most cases the genetic basis for modification remains unclear. In this study, we shed light into the mode of action and nature of modifier genes in sunflower, and suggest the relevance of duplicated loci affecting the expression of seed quality specific mutations.

Methods

Plant material, phenotypic analyses and DNA extraction

The study included the sunflower lines nmsT2100 and IAST-1, both with high gamma-tocopherol content (>85%). nmsT2100 is a NMS line developed by introgressing monogenic recessive NMS from line P21 [32] into high gamma-tocopherol line T2100 [10]. IAST-1 was isolated in the course of a chemical mutagenesis program on seeds of an accession of 'Peredovik' [11]. Twenty-four half seeds of nmsT2100 and IAST-1 were nondestructively analyzed for tocopherol profile as described below, germinated and planted in pots under open air conditions in the spring of 2005. NMS plants of nmsT2100 were pollinated with pollen of IAST-1 plants. Half seeds of the parents as well as F_1 half seeds were analysed for tocopherol profile. F_1 and parent half seeds were sown in March 2006 and the corresponding plants were grown in pots under open air conditions. F_1 plants were bagged before flowering to obtain the F_2 generation. F_2 seeds from several F_1 plants were analysed for seed tocopherol profile. In most cases, F_2 seeds had uniformly high gamma-tocopherol content (>90%). However, some F_1 plants showed segregation for gamma-tocopherol content at the F_2 seed level, which indicated the expression of modifier genes. One population of 192 F_2 seeds from

a single F_1 plant that showed large segregation for gamma-tocopherol content was selected for the molecular analyses. F_2 half seeds were germinated and the corresponding plants were grown in pots under open air conditions in 2007. F_2 plants were bagged before flowering to produce the F_3 seed. Five fully expanded leaves from each F_2 plant (135 F_2 plants) were cut, frozen at -80ºC, lyophilised and ground to a fine powder in a laboratory mill. DNA was isolated from ground leaf tissue from each F_2 plant and from five plants of nmsT2100 and IAST-1 parental lines as described in Berry et al. [33]. Twenty four F_3 seeds from each F_2 plant that produced sufficient amount of seeds were analysed for tocopherol profile.

Two additional F_2 populations were also used in this study with the objective of mapping tocopherol biosynthesis loci. One population, described in García-Moreno et al. [13], derived from a cross between CAS-12, with wild-type tocopherol profile mainly made up of alpha-tocopherol, and IAST-540, with high gamma-tocopherol content. The other population derived from a cross between lines IAST-413 and HA-89, both with wild-type tocopherol profile mainly made up of alpha-tocopherol, though IAST-413 is characterized by increased total tocopherol content [34].

The analysis of tocopherol profile was made for all analyzed generations and populations in half seeds. The half-seed technique is a common technique used in sunflower breeding consisting in cutting a small seed piece from the seed part distal to the embryo, which is used for nondestructive analysis of seed quality traits, as the remaining seed containing the embryo can be germinated after the corresponding analysis [35]. Individual half seeds were analysed for tocopherol profile following the method of Goffman et al. [36]. Half seeds were placed into 10-ml tubes with 2 ml iso-octane. The half seeds were then crushed with a stainless steel rod as fine as possible. The samples were stirred and extracted overnight at room temperature in darkness (extraction time about 16 h). After extraction, the samples were stirred again, centrifuged, and filtered. Twenty-five µl of the extract were

analysed by HPLC using a fluorescence detector at 295 nm excitation and 330 nm emission and iso-octane/tert-butylmethylether (94:6) as eluent at an isocratic flow rate of 1 ml min^{-1}. Chromatographic separation of the tocopherols was performed on a LiChrospher 100 diol column (250 mm x 2 mm I.D.) with 5-µm spherical particles, connected to a silica guard column (LiChrospher Si 60, 5mm x 4 mm I.D.). The peak areas of the individual tocopherols were corrected according to their previously calculated response factors: alpha-tocopherol=1.0; beta-tocopherol=1.80; gamma-tocopherol=1.85; delta-tocopherol=2.30.

Map construction and molecular analysis

A complete linkage map for the nmsT2100 x IAST-1 population was constructed to scan the genome for modifier genes affecting the expression of the *Tph2* gene. For this, the parental lines IAST-1 and nmsT2100 were initially screened for polymorphisms in two replicate samples together with four F_2 individuals, using a genome-wide framework of 95 sunflower SSRs [37]. A preliminary genetic linkage map from this population was constructed. A set of INDEL markers [38], identified by ZVG prefixes, and an additional set of SSR markers mapped by Tang et al. [39] and Yu et al. [38], identified by ORS and CRT prefixes, were additionally screened for polymorphisms between these parental lines to complete the linkage map. INDEL markers for the tocopherol biosynthesis genes gamma-TMT, MPBQ/MSBQ-MT and tocopherol cyclase described by Hass et al. [14] and Tang et al. [20] were also screened for polymorphisms between nmsT2100 and IAST-1. PCRs for SSRs analyses were performed as described by Pérez-Vich et al. [40]. INDEL analyses were carried out following Yu et al. [38] and Hass et al. [14]. SSR and INDEL amplification products were separated on 3% (w/v) Metaphor® (BMA, Rockland, ME, USA) and 1.5% agarose gels, respectively, in 1x TBE buffer with ethidium bromide incorporated in the gel. SSR and INDEL markers revealing

polymorphisms were then genotyped in the nmsT2100 x IAST-1 F_2 population, following the protocols mentioned above.

Chi-square statistics were computed on each genotyped locus to detect deviations from the expected Mendelian ratios for codominant (1:2:1) or dominant (3:1) markers. The nmsT2100 x IAST-1 linkage map was constructed using the software MAPMAKER/EXP version 3.0b (Whitehead Institute, Cambridge, MA, USA) [41]. Two-point analysis was used to identify linkage groups (LGs) at a LOD score of 3 and a maximum recombination frequency of 0.40. Three-point and multi-point analyses were used to determine the order and interval distances between the markers at each LG. The Haldane mapping function was used to compute the map distances in centiMorgans (cM) from the recombination fractions. Multiple loci detected by a single maker were coded with the marker name plus the suffix a, b, c, or d to indicate each duplicate locus. Linkage group maps were drawn using the MapChart software [42].

Genetic analysis of modifier genes was performed in several stages. In the first stage, the significance of each marker's association with the phenotypic trait [gamma-tocopherol content at the F_2 seed and F_2 plant (average value of 24 F_3 seeds per F_2 plant) generations] was determined by one-way analysis of variance (ANOVA) using the statistical package SPSS Statistics v. 19, with marker genotypes being classes. In this analysis, we identified an unexpected macromutation on LG 8 at the *Tph2* locus. The effects of the macromutations, if ignored, could dramatically reduce the power for identifying other genes or QTL affecting the studied trait.

In a second stage, composite interval mapping (CIM) [43-44] was used to scan the genome for QTL affecting gamma-tocopherol content, in order to strengthen and corroborate the results of the analyses of variance, evaluate the existence of additional QTL, and estimate the interaction and global effect of all the detected QTL. Computations were carried out using

the software PLABQTL Version 1.1 [45]. The phenotypic data consisted on gamma-tocopherol content in the F_2 seed and F_2 plant generations. Additional analyses were carried out by using other parameters calculated from the F_3 seed data such as the minimum, maximum and the standard deviation of gamma-tocopherol content in F_3 seeds per F_2 plant, and the number of F_3 seeds within each F_2 plant with less than 90% of gamma-tocopherol content. Since these analyses gave similar results to those obtained with the mean gamma-tocopherol value per F_3 family, only results based on the F_3 mean value are shown. Analyses were made initially with the "first" statement to check the database for errors and outliers. Next, simple interval mapping (SIM) was carried out for an initial scan and detection of QTL with main effects. Finally, CIM was performed with markers closest to the main QTL as co-factors. Genome-wide threshold values (α=0.05) for declaring the presence of QTL were estimated from 1000 permutations of each phenotypic trait [46]. The thresholds of the LOD score (and their 0.95 confidence intervals) were 2.57 (2.48 - 2.70) and 2.65 (2.58 - 2.76) for gamma-tocopherol content in the F_2 and the F_3 generations, respectively. Estimates of QTL positions were obtained at the point where the LOD score reaches its maximum in the region under consideration. One-LOD support limits for the position of each QTL were also calculated [47]. The proportion of phenotypic variance explained by each individual QTL was calculated as the square of the partial correlation coefficient (R^2). Estimates of the additive (a_i) and dominance (d_i) effects, as defined by Falconer [48], for the ith putative QTL, the total LOD score, as well as the total proportion of the phenotypic variance explained by all QTL, were obtained by fitting a multiple regression model including all putative QTL for the respective trait simultaneously [47]. The occurrence of QTL x QTL interactions was tested by adding digenic epistatic effects to the model.

QTL software such as PLABQTL estimate epistatic interactions among previously identified QTL. Since modifying genes are defined as genes having no known effect except to

intensify or diminish the expression of a major gene [23], their effect as individual loci and subsequently their interaction with major loci may be undetectable with this type of analyses. In consequence, two-way interactions between the *Tph2* major locus and all the marker loci genotyped in this study were also tested. Two-way interactions were analyzed at a significance threshold of $P \leq 0.05$ by analysis of variance using the general linear model (GLM) of SPSS Statistics v. 19. Statistical significance of differences for gamma-tocopherol content in different genotypes combining two marker loci were also computed using Duncan's multiple range test. The significant epistatic interaction terms were combined with those of the previously identified QTL in multiple locus models using the "seq" statement of PLABQTL.

Sequence analysis gamma-TMT loci

In the course of the genetic analyses of the nmsT2100 x IAST-1 population, we identified different gamma-TMT loci amplified with INDEL marker gamma-TMT-F9/R24 associated to gamma-tocopherol QTL. In order to confirm their nature, these loci were sequenced as follows. F9/R24 INDEL fragments amplified from the IAST-1 and nmsT2100 parental lines were separated on a 1.5% agarose gel, excised and purified by means of the QIAquick gel extraction kit (Qiagen GmbH, Hilden, Germany). The purified fragments were ligated into the T/A vector (pCR2.1) and the recombinants were transformed to TOP10 Chemically Competent *E. coli* using the TOPO-TA cloning kit (Invitrogen, San Diego, CA, USA) as described by the manufacturer. Five recombinant bacterial colonies (white) per isolated band were picked from the plate containing ampicillin and X-gal as selective media and cultured overnight at 37°C. Plasmids were extracted and purified using QIAprep Spin Miniprep Kit (Qiagen GmbH, Hilden, Germany). PCR with M13 forward and reverse vector primers and F9 and R24 primers, and restriction enzyme digestion was performed to confirm the presence

and size of the insert. Sequencing in both forward and reverse orientations of the cloned fragments (two clones per locus) was performed at GATC Biotechnology (Konstanz, Germany) using the M13 forward and reverse sequencing primers. Sequence analysis was conducted with the aid of the software Vector NTI Advance 10.3.0 (Invitrogen, San Diego, CA, USA).

Full-length sequence analysis of the IAST-1 and nmsT2100 gamma-TMT gene

Full-length gamma-TMT genomic DNA sequences were isolated from the high gamma-tocopherol lines nmsT2100 and IAST-1 by long distance PCR using primers developed at the 5´end (forward primer gamma-TMT-F1 from Hass et al. [14]) and the 3´end (reverse primer R92: TAATTCCTTGGGATGCCATT) of the sunflower gamma-TMT gene (GenBank accessions nos. DQ229828 to DQ229834). AccuPrime High Fidelity Taq DNA Polymerase (Invitrogen Life Technologies, Carlsbad, CA, USA) was used for PCR amplification in three individuals of nmsT2100 and five of IAST-1 as described by the manufacturer. The amplified products from each individual were separated on 1.5% agarose gels, showing in both lines two bands of a size higher than 4 kb that corresponded to the two gamma-TMT paralogs described by Hass et al. [14]. The upper (paralog 2) and the lower (paralog 1) bands were independently purified in each individual and cloned using the TOPO-TA cloning kit (Invitrogen Life Technologies, Carlsbad, CA, USA) as described above, with the exception that twenty recombinant bacterial colonies (white) per isolated band were picked from the plate. Restriction enzyme digestion was performed to confirm the presence of the insert and the restriction patterns characteristic for each gamma-TMT paralog. Sequencing in both forward and reverse orientations of the cloned fragments (a total of 10 clones for IAST-1 paralog 1, 13 clones for IAST-1 paralog 2, 5 clones for nmsT2100 paralog 1, and 6 clones for nmsT2100 paralog 2) was performed at GATC Biotechnology (Konstanz, Germany) using the

universal M13 forward and reverse primers and internal primers (Gamma-TMT-F9, F27, F67, R10, R24, R35 and R78 from Hass et al. [14], and primers from Table 5) designed at 500 to 1000 bp intervals from the sunflower gamma-TMT gene. Sequence analysis was conducted using Vector NTI Advance 10.3.0 (Invitrogen, San Diego, CA, USA). A consensus sequence for IAST-1 gamma-TMT paralog 1, IAST-1 gamma-TMT paralog 2, nmsT2100 gamma-TMT paralog 1, and nmsT2100 gamma-TMT paralog 2 was made from the analysis of 10, 13, 5, and 6 sequenced clones, respectively. Changes in the nucleotide sequence were only included in the consensus sequence when they were conserved among the different clones and individuals sequenced from each line.

List of abbreviations

Gamma-TMT – gamma-tocopherol methyltransferase; INDEL – insertion-deletion; LG – linkage group; MPBQ/MSBQ-MT – 2-methyl-6-phytyl-1,4-benzoquinone/2-methyl-6-solanyl-1,4-benzoquinone methyltransferase; NMS – nuclear male sterile; SSR – simple sequence repeat

Authors' contributions

MJG-M carried out genetic map construction and all other molecular analyses. JMF-M and LV selected the sunflower lines, crossed them, produced and studied the phenotypic data, and participated in the design of the study. BP-V conceived and designed the study and supervised molecular analyses and interpretation of results. All authors contributed to the manuscript preparation, and read and approved the final manuscript.

Acknowledgements and funding

The research was funded by the Spanish Ministry of Science and Innovation and the European Union FEDER (research project AGL2007-62834) and Dow Agrosciences LLC.

References

1. Eitenmiller R, Lee J: *Vitamin E. Food Chemistry, Composition, and Analysis*. New York: Marcel Dekker, New York; 2004.
2. Hunter SC, Cahoon EB: **Enhancing vitamin E in oilseeds: Unravelling tocopherol and tocotrienol biosynthesis**. *Lipids* 2007, **42:**97-108.
3. Chow CK: **Vitamin E**. In *Biochemical and Physiological Aspects of Human Nutrition*. Edited by Stipanuk MH. Philadelphia: Saunders; 2000:584-598.
4. Kamal-Eldin A, Appelqvist L-Å: **The chemistry and antioxidant properties of tocopherols and tocotrienols**. *Lipids* 1996, **31:**671-701.
5. Seppanen CM, Song Q, Csallany AS: **The antioxidant functions of tocopherol and tocotrienol homologues in oils, fats, and food systems**. *Journal of the American Oil Chemists' Society* 2010, **87:**469-481.
6. Demurin Y, Skoric D, Karlovic D: **Genetic variability of tocopherol composition in sunflower seeds as a basis of breeding for improved oil quality**. *Plant Breeding* 1996, **115:**33-36.
7. Fuster MD, Lampi AM, Hopia A, Kamal-Eldin A: **Effects of alpha- and gamma-tocopherols on the autoxidation of purified sunflower triacylglycerols**. *Lipids* 1998, **33:**715-722.
8. Yanishlieva NV, Kamal-Eldin A, Marinova EM, Toneva AG: **Kinetics of antioxidant action of alpha- and gamma-tocopherols in sunflower and soybean triacylglycerols**. *European Journal of Lipid Science and Technology* 2002, **104:**262-270.
9. Marmesat S, Velasco L, Ruiz-Méndez MV, Fernández-Martínez JM, Dobarganes C: **Thermostability of genetically modified sunflower oils differing in fatty acid and**

tocopherol compositions. *European Journal of Lipid Science and Technology* 2008, **110:**776-782.

10. Velasco L, Dominguez J, Fernández-Martínez JM: **Registration of T589 and T2100 sunflower germplasms with modified tocopherols**. *Crop Science* 2004, **44:**362-363.
11. Velasco L, Pérez-Vich B, Fernández-Martínez JM: **Novel variation for tocopherol profile in a sunflower created by mutagenesis and recombination**. *Plant Breeding* 2004, **123:**490-492.
12. Velasco L, Fernández-Martínez JM: **Identification and genetic characterization of new sources of beta- and gamma-tocopherol in sunflower germplasm**. *Helia* 2003, **26:**17-24.
13. García-Moreno MJ, Vera-Ruiz EM, Fernández-Martínez JM, Velasco L, Pérez-Vich B: **Genetic and molecular analysis of high gamma-tocopherol content in sunflower**. *Crop Science* 2006, **46:**2015-2021.
14. Hass CG, Tang S, Leonard S, Traber M, Miller JF, Knapp SJ: **Three non-allelic epistatically interacting methyltransferase mutations produce novel tocopherol (vitamin E) profiles in sunflower**. *Theoretical and Applied Genetics* 2006, **113:**767-782.
15. Bergmüller E, Porfirova S, Dörmann P: **Characterization of an *Arabidopsis* mutant deficient in γ-tocopherol methyltransferase**. *Plant Molecular Biology* 2003, **52:**1181-1190.
16. Demurin Y, Efimenko SG, Peretyagina TM: **Genetic identification of tocopherol mutations in sunflower**. *Helia* 2004, **27:**113-116.
17. García-Moreno MJ, Fernández-Martínez JM, Pérez-Vich B, Velasco L: **A modifying gene affecting gamma-tocopherol content in sunflower**. In *Proceedings of the*

Seventeenth International Sunflower Conference: 8-12 June 2008; Córdoba, Spain. Edited by Velasco L. Paris: International Sunflower Association; 2008:601-604.

18. Lacombe S, Kaan F, Léger L, Bervillé A: **An oleate desaturase and a suppressor loci direct high oleic acid content of sunflower (*Helianthus annuus* L.) oil in the Pervenets mutant**. *Comptes Rendus de l'Academie des Sciences, Series III, Sciences de la Vie* 2001, **324:**839-845.
19. Velasco L, Pérez-Vich B, Jan CC, Fernández-Martínez JM: **Inheritance of resistance to broomrape (*Orobanche cumana* Wallr.) race F in a sunflower line derived from wild sunflower species**. *Plant Breeding* 2007, **126:**67-71.
20. Tang S, Hass C, Knapp S: **Ty3/gypsy-like retrotransposon knockout of a 2-methyl-6-phytyl-1,4-benzoquinone methyltransferase is non-lethal, uncovers a cryptic paralogous mutation, and produces novel tocopherol (vitamin E) profiles in sunflower**. *Theoretical and Applied Genetics* 2006, **113:**783-799.
21. Pérez-Vich B, Berry ST, Velasco L, Fernández-Martínez JM, Gandhi S, Freeman C, Heesacker A, Knapp SJ, Leon AJ: **Molecular mapping of nuclear male sterility genes in sunflower**. *Crop Science* 2005, **54:**1851–1857.
22. Del Moral L, Fernández-Martínez JM, Pérez-Vich B, Velasco L: **Expression of modified tocopherol content and profile in sunflower tissues**. *Journal of the Science of Food and Agriculture*, in press (DOI: 10.1002/jsfa.4585).
23. Briggs FD, Knowles PF: *Introduction to Plant Breeding*. New York: Reinhold Publishing Corporation; 1967.
24. Urie AL: **Inheritance of high oleic acid in sunflower**. *Crop Science* 1985, **25:**986-989.
25. Velasco L, Pérez-Vich B, Fernández-Martínez JM: **Inheritance of oleic acid content under controlled environment**. In *Proceedings of the Fifteenth International*

Sunflower Conference: 12-15 June 2000; Toulouse, France. Paris: International Sunflower Association; 2000:A31-A36.

26. García-Moreno MJ, Fernández-Martínez JM, Velasco L, Pérez-Vich B: **Molecular tagging and candidate gene analysis of the high gamma-tocopherol trait in safflower (*Carthamus tinctorius* L.)**. *Molecular Breeding* 2011, **28**:367-379.
27. Endrigkeit J, Wang X, Cai D, Zhang C, Long Y, Meng J, Jung C: **Genetic mapping, cloning, and functional characterization of the *BnaX.VTE4* gene encoding a gamma-tocopherol methyltransferase from oilseed rape**. *Theoretical and Applied Genetics* 2009, **119:**567-575.
28. Fernández-Martínez JM, Jiménez A, Domínguez J, García JM, Garcés R, Mancha M: **Genetic analysis of the high oleic content in cultivated sunflower (*Helianthus annuus* L.)**. *Euphytica* 1989, **41:**39-51.
29. Jiang H, Gu Z: **Growth of novel epistatic interactions by gene duplication**. *Genome Biology and Evolution* 2011, **3:**295-301.
30. Prelich G: **Suppresion mechanisms**. *Trends in Genetics* 1999, **15:**261-266.
31. Miller JF, Vick BA: **Registration of four mid-range oleic acid sunflower genetic stocks**. *Crop Science* 2002, **42:**994.
32. Jan CC: **Inheritance and allelism of mitomycin C- and streptomycin-induced recessive genes for male sterility in cultivated sunflower**. *Crop Science* 1992, **32:**317-320.
33. Berry ST, Leon AJ, Hanfrey CC, Challis P, Burkholz A, Barnes SR, Rufener GK, Lee M, Caligari PDS: **Molecular marker analysis of *Helianthus annuus* L. 2. Construction of an RFLP linkage map for cultivated sunflower**. *Theoretical and Applied Genetics* 1995, **91:**195-199.

34. Del Moral L, Pérez-Vich B, Fernández-Martínez JM, Velasco L: **Inheritance of increased seed tocopherol content in sunflower line IAST-413**. *Plant Breeding*, in press (DOI: 10.1111/j.1439-0523.2011.01865.x).

35. Fernández-Martínez JM, Perez-Vich B, Velasco L: **Sunflower**. In *Oil Crops*. Edited by Vollmann J, Rajcan I. New York: Springer; 2009: 155-232.

36. Goffman FD, Velasco L, Thies W: **Quantitative determination of tocopherols in single seeds of rapeseed (*Brassica napus* L.)**. *Fett/Lipid* 1999, **101:**142-145.

37. Tang S, Kishore VK, Knapp SJ: **PCR-multiplexes for a genome-wide framework of simple sequence repeat marker loci in cultivated sunflower**. *Theoretical and Applied Genetics* 2003, **107:**6-19.

38. Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu W-C, Webb DM, Thompson L, Edwards KJ, Berry S, Leon AJ, Grondona M, Olungu C, Maes N, Knapp SJ: **Towards a saturated molecular genetic linkage map for cultivated sunflower**. *Crop Science* 2003, **43:**367-387.

39. Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ: **Simple sequence repeat map of the sunflower genome**. *Theoretical and Applied Genetics* 2002, **105:**1124-1136.

40. Pérez-Vich B, Akhtouch B, Knapp SJ, Leon AJ, Velasco L, Fernández-Martínez JM, Berry ST **Quantitative trait loci for broomrape (*Orobanche cumana* Wallr.) resistance in sunflower**. *Theoretical and Applied Genetics* 2004, **109:**92-102.

41. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L: **MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations**. *Genomics* 1987, **1:**174-181.

42. Voorrips RE: **MapChart: Software for the graphical presentation of linkage maps and QTL**. *Journal of Heredity* 2002, **93**:77-78.

43. Jansen RC, Stam P: **High resolution of quantitative traits into multiple loci via interval mapping**. *Genetics* 1994, **136:**1447-1455.

44. Zeng ZB: **Precision mapping of quantitative trait loci**. *Genetics* 1994, **136:**1457-1468.

45. Utz HF, Melchinger AE: **PLABQTL: A program for composite interval mapping of QTL**. *Journal of Quantitative Trait Loci* 1996, **2:**1-5.

46. Doerge RW, Churchill GA: **Permutation tests for multiple loci affecting a quantitative character**. *Genetics* 1996, **142:**285-294.

47. Bohn M, Khairallah MM, González-de-León D, Hoisington DA, Utz HF, Deutsch JA, Jewell DC, Mihm JA, Melchinger AE: **QTL mapping in tropical maize: I. Genomic regions affecting leaf feeding resistance to sugarcane borer and other traits**. *Crop Science* 1996, **36:**1352-1361.

48. Falconer DS: *Introduction to quantitative genetics, 3rd edn*. London: Longman; 1989.

Figure Legends

Figure 1 Phenotypic segregation of the F_2 generation

Histograms of gamma-tocopherol content (% of total tocopherols) in sunflower lines nmsT2100, IAST-1, and the F_2 generation from their cross.

Figure 2 Phenotypic segregation of the F_2 vs F_3 generations

Gamma-tocopherol content (% of total tocopherols) in F_2 seeds from the cross nmsT2100 x IAST-1 and their corresponding F_2 plants (24 F_3 seeds averaged per F_2 plant).

Figure 3 Amplification profile of gamma-TMT_F1/F2/R24 INDEL marker

Amplification profile of gamma-TMT_F1/F2/R24 INDEL marker in replicate samples of the high gamma-tocopherol parent line IAST-1, the high gamma-tocopherol parent line nmsT2100 and four F_2 individuals. Lane 1-2, replicate samples of IAST-1; lanes 3-4, replicate samples of nmsT2100 and lanes 5-8, F_2 individuals.

Figure 4 Amplification profile of gamma-TMT_F9/R24 INDEL marker

Amplification profile of gamma-TMT_F9/R24 INDEL marker in single samples of the high gamma-tocopherol IAST-1, nmsT2100, IAST-540 lines and the standard gamma-tocopherol IAST-413, HA-89 and CAS-12 lines. Lane 1, IAST-413; lane 2, HA-89; lane 3, 50 bp DNA ladder; lane 4, IAST-1; lane 5, nmsT2100; lane 6, IAST-540 and lane 7, CAS-12. Polymorphic loci are highlighted by boxes, and the linkage group and the population in which they have been mapped are shown at the right of the figure.

Figure 5 LODs for gamma-tocopherol QTL

Likelihood odds (LODs) for F_2 (black line) and F_3 (grey line) gamma-tocopherol QTL on linkage groups (LGs) 1, 8, 9, 14 and 16 in nmsT2100 x IAST-1. Gamma-tocopherol-TMT (γ-TMT) loci are highlighted in bold.

Figure 6 Sequence alignment tree for gamma-TMT loci

Sequence alignment tree obtained by aligning DNA genomic sequences from the different loci amplified with primer combination gamma-TMT F9/R24 and constructed using the AlignX program in the Vector NTI software suite. Calculated distance based on distances between all pairs of sequence values are shown in parenthesis following the molecule name. The locus *a* fragment isolated from nmsT2100 and IAST-1, the locus *b* isolated from nmsT2100, the locus *c* isolated from IAST-1, and the locus *d* isolated from HA-89 and IAST-540, together with the F9 to R24 region from gamma-TMT haplotypes 1 (DQ DQ229828), 2 (DQ DQ229829), 3 (DQ DQ229830), 4 (DQ DQ229831 and DQ DQ229832), and 5 (DQ DQ229833 and DQ DQ229834) are included.

Table 1 nmsT2100 x IAST-1 linkage map coverage

Genome coverage offered by the marker set used for QTL analysis in the nmsT2100 x IAST-1 population. Distances are expressed in Haldane cM

	LG coverage[2]				
			Coverage (cM)		
LG[1]	SSR	INDEL	Mean	Largest interval	Total
1	13	3	4.3	19.3	64.0
2	6	0	4.4	8.2	22.1
3	6	1	16.6	27.8	99.7
4	5	1	4.2	8.1	20.9
5	5	1	5.5	15.5	27.6
6	5	2	3.6	13.4	21.5
7	7	1	10.7	43.4	74.9
8	8	3	4.5	14.9	44.6
9	6	0	6.6	22.8	32.9
10	7	4	10.1	29.8	100.9
11	3	0	6.6	8.9	13.2
12	4	2	4.3	7.5	21.3
13	7	0	8.3	31.2	50.0
14	7	1	4.6	10.2	32.4
15	6	0	10.1	36.5	50.5
16	8	1	9.6	32.1	76.8
17	6		7.8	17.1	39.1
Total	109	20			792.4

[1]LG= Linkage group

[2]SSR= Number of SSR loci; INDEL= Number of INDEL loci

Table 2 Effect of gamma-TMT-F1/F2/R24 on seed gamma-tocopherol content

Association between the gamma-TMT-F1/F2/R24 locus on LG 8 and gamma-tocopherol content (% of the total tocopherols) determined by variance analysis in the nmsT2100 x IAST-1 population. Mean gamma-tocopherol ± standard deviation (SD) are presented in different genotypic classes: A=homozygous with respect to the allele derived from IAST-1, B=homozygous with respect to the allele derived from nmsT2100, H=heterozygous.

Seed generation	No. individuals within each marker class			Mean ± SD for gamma-tocopherol content (% total tocopherols) within each marker class			ANOVA analysis	
	A	H	B	A (IAST-1)	H	B (nmsT2100)	*F*	*P*
F_2	32	73	27	52.3a±29.6	72.2b±19.1	98.9c±0.8	38.6	<0.001
F_3	26	58	3	86.8a±11.8	96.9b±2.6	98.6b±0.9	20.4	<0.001

Mean values within rows followed by the same letter are not significantly different at 0.05 level of probability (Duncan´s multiple range test).

Table 3 QTL affecting gamma-tocopherol content in the nmsT2100 x IAST-1 population

Generation	QTL	LG[1]	Pos.[2]	Supp int.[3]	Marker interval[4]	LOD	R^2 (%)	Significant gene effects[5]	
								a	*d*
F_2	*GamT1.1*	1	44	29-63	ORS552 to Gamma_TMT_F9/R24c	2.4	7.8		-2.96*
	Tph2_Gamma-TMT	8	12	7-15	ZVG34 to Gamma_TMT_F1/F2/R24	15.0	41.2	20.9**	
	GamT9.1	9	32	19-32	ORS887 to ORS176	2.3[6]	7.5	-7.7**	
	GamT14.1	14	16	6-17	ORS185 to ORS307	4.1	13.1	-11.5**	
	GamT16.1	16	13	11-21	Gamma_TMT_F9/R24b to ORS700	3.4	10.9		
					Total	17.2	44.4		
					add*Tph2_Gamma-TMT**add*GamT14.1*			9.9*	
					add*Tph2_Gamma-TMT**add*GamT16.1*			14.0**	
					Total epistasis	21.7	52.4		

Table 3 (continued)

Generation	QTL	LG[1]	Pos.[2]	Supp int.[3]	Marker interval[4]	LOD	R^2 (%)	Significant gene effects[5]	
								a	*d*
F_3	*Tph2_Gamma-TMT*	8	20	16-23	Gamma_TMT_F1/F2/R24 to ORS70	11.1	44.0	10.9**	
	GamT9.1	9	31	13-32	ORS887 to ORS176	2.04[6]	10.1	-2.6**	
	GamT14.1	14	16	6-27	ORS185 to ORS307	2.7	13.0	-3.5**	2.5*
	GamT16.1	16	14	13-16	ORS700 to ORS757	3.3	15.9	10.6**	-22.9**
					Total	18.1	61.2		
					add*Tph2_Gamma-TMT**add*GamT9.1*			6.8**	
					add*Tph2_Gamma-TMT**add*GamT14.1*			6.7**	
					add*Tph2_Gamma-TMT**add*GamT16.1*			8.5**	
					Total epistasis	23.8	71.3		

**=significant at the 0.01 and *=significant at the 0.05 probability level.

[1]LG=Linkage group

[2]Absolute position from the top of the LG in centiMorgans (cM)

118

Table 3 (continued)

[3]One-LOD support interval in centiMorgans: Refers to the region flanking each QTL peak in which LOD scores decline by one

[4]Markers flanking the likelihood peak for a putative QTL

[5]*a*=additive effect. A positive sign means an increase of the trait value due to nmsT2100 alleles. *d*=dominant effect. *a* and *d* estimates, as well as total R^2 and LOD score values were obtained from a simultaneous fit of all putative QTL using multiple regression

[6]QTL detected below the LOD threshold

Table 4 Effect of epistatically interacting marker loci on seed gamma-tocopherol content

Genotypic means for F_2 gamma-tocopherol (gamma-T) content (% of the total tocopherols) in allelic combinations at epistatically interacting marker loci. A=homozygous with respect to the allele derived from IAST-1, B=homozygous with respect to the allele derived from nmsT2100, H=heterozygous.

Genotype of γ-TMT-F1/F2/R24 (LG 8)	Genotype of ORS578 (LG 14)	F_2 gamma-T[1] Mean	SD	n	Genotype of ORS126 (LG 16)	F_2 gamma-T[1] Mean	SD	n
A (IAST-1)	B (nmsT2100)	23.48a	24.10	9	B (nmsT2100)	51.68ab	31.89	14
	H	64.60b	20.27	12	H	44.55a	27.08	14
	A (IAST-1)	67.48b	22.69	10	A (IAST-1)	81.26cd	6.96	4
H	B (nmsT2100)	71.61b	18.93	18	B (nmsT2100)	77.84c	16.68	22
	H	71.00b	19.19	42	H	69.17bc	18.71	36
	A (IAST-1)	76.91b	19.60	13	A (IAST-1)	71.23c	22.38	15
B (nmsT2100)	B (nmsT2100)	98.92c	0.53	9	B (nmsT2100)	98.59d	0.97	7
	H	98.55c	0.87	9	H	98.96d	0.71	16
	A (IAST-1)	99.15c	0.86	9	A (IAST-1)	99.03d	0.82	4

Values within columns followed by the same letter are not significantly different at 0.05 level of probability (Duncan´s multiple range test).

[1]Mean and standard deviation (SD) for F_2 gamma-T content and number of F_2 individuals per genotypic class (n).

Table 5 Sequencing primers designed each 500-1000 bp into the gamma-TMT gene.

Primer	Sequence (5´-3´)
F90	GGATGAATCGTTTGTTATTG
F91	GTCAATGGAGAGTGGAGAGC
F92	AGGAAGAAAAAATCTTGAATAA
F93	ATCGCTTCATCATCATCATA
F94	CACTAAATTTGACATCCACAAC
F95	GCCACTAATGATTGAAGGATT
R94	ACCACAACGTAAAAATGTTT
R95	CCACTACGTAGCAATGAAGT
R96	CCTTTAGTTTGCCAATTCAC
R97	CCGAGTCAACTCACTAACAA
R98	TCATTCACAAACTGCAGTAG

Additional files

File name: FigS1

File format: .pdf

Title: Gamma-TMT sequence alignment

Description: Sunflower gamma-tocopherol methyltransferase genomic DNA sequence alignment for the high gamma-tocopherol lines IAST-1 and nmsT2100 as well as genbank sequences from gamma-TMT haplotypes 1 (DQ DQ229828), 2 (DQ DQ229829), 3 (DQ DQ229830), 4 (DQ DQ229831 and DQ DQ229832), and 5 (DQ DQ229833 and DQ DQ229834) and the cDNA gamma-TMT EF495161 sequence.

Additional files provided with this submission:

Additional file 1: Figure S1.pdf, 554K
http://www.biomedcentral.com/imedia/3086055785941371/supp1.pdf

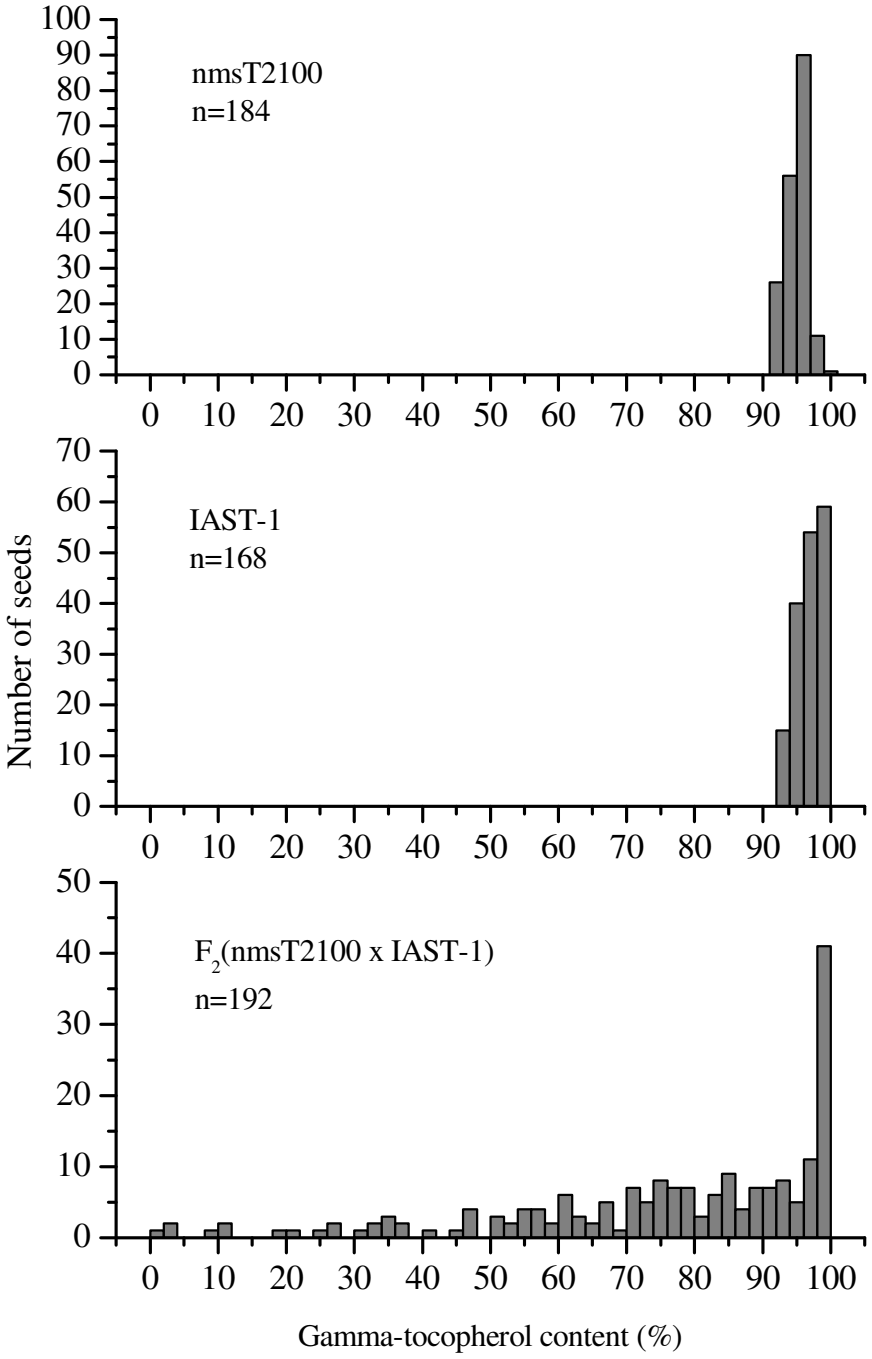

nmsT2100
n=184
IAST-1
n=168
F2(nmsT2100 x IAST-1)
n=192
Number of seeds
Gamma-tocopherol content (%)

Figure 1

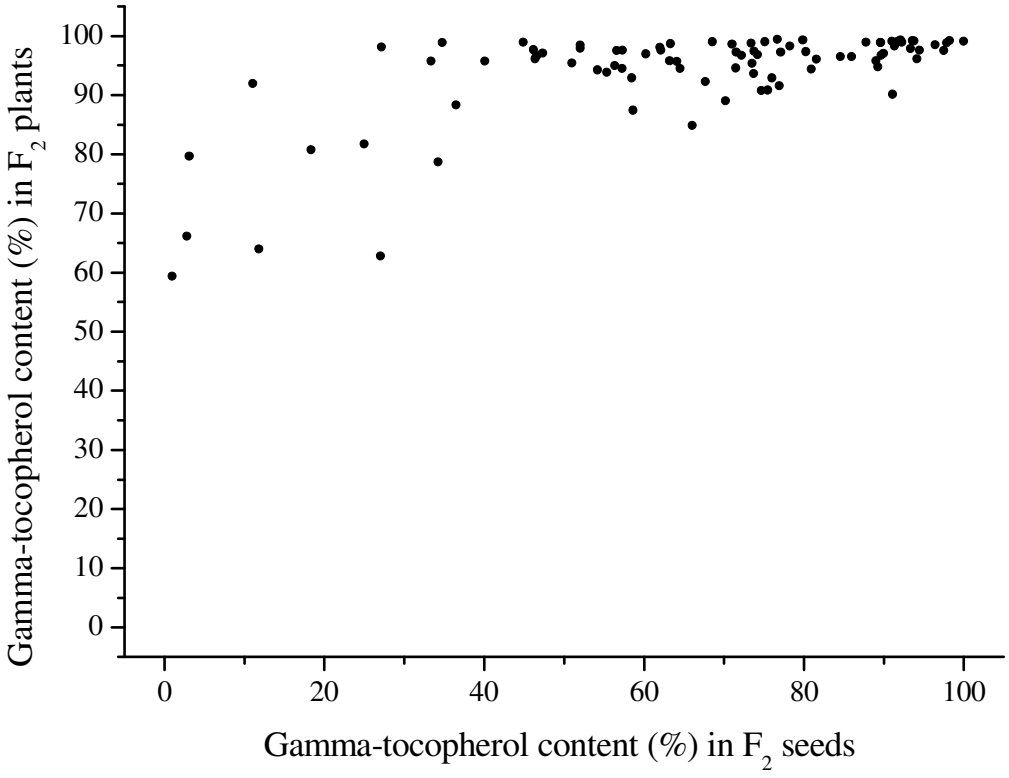

Gamma-tocopherol content (%) in F2 plants
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100
Gamma-tocopherol content (%) in F2 seeds

Figure 2

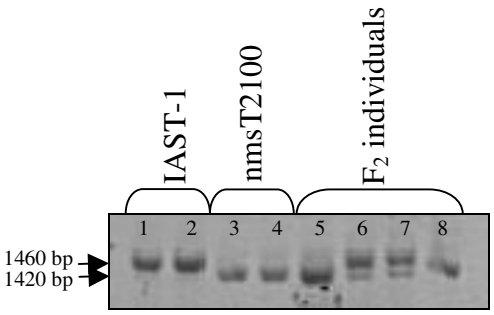


Figure 3

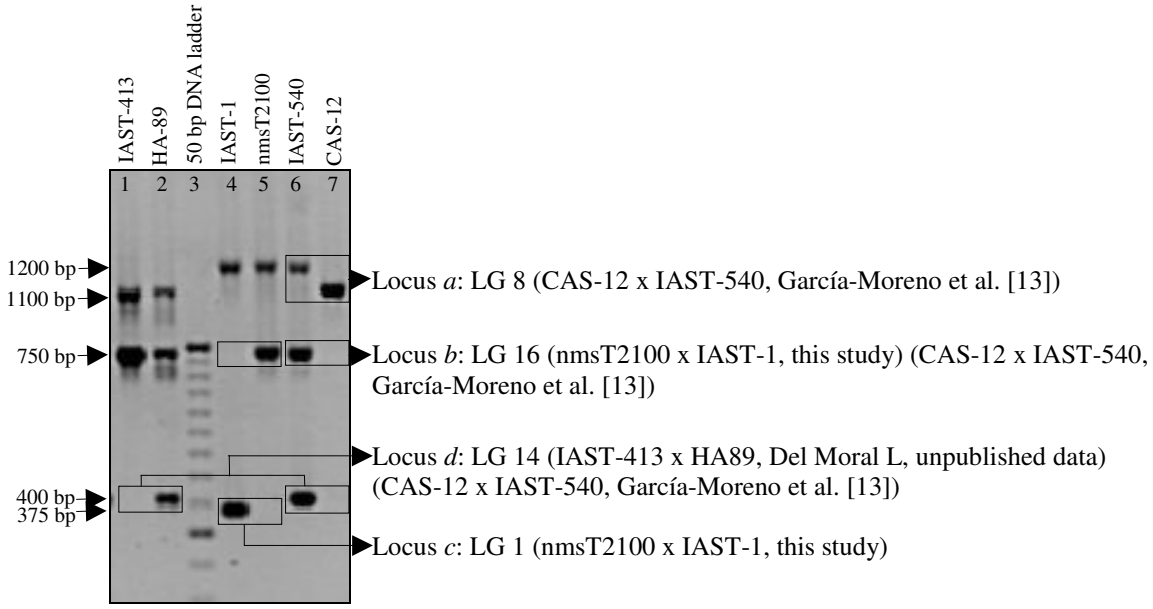

IAST-413
HA-89
50 bp DNA ladder
IAST-1
nmsT2100
IAST-540
CAS-12
1 2 3 4 5 6 7
1200 bp
1100 bp
750 bp
400 bp
375 bp
Locus a: LG 8 (CAS-12 x IAST-540, García-Moreno et al. [13])
Locus b: LG 16 (nmsT2100 x IAST-1, this study) (CAS-12 x IAST-540, García-Moreno et al. [13])
Locus d: LG 14 (IAST-413 x HA89, Del Moral L, unpublished data) (CAS-12 x IAST-540, García-Moreno et al. [13])
Locus c: LG 1 (nmsT2100 x IAST-1, this study)

Figure 4

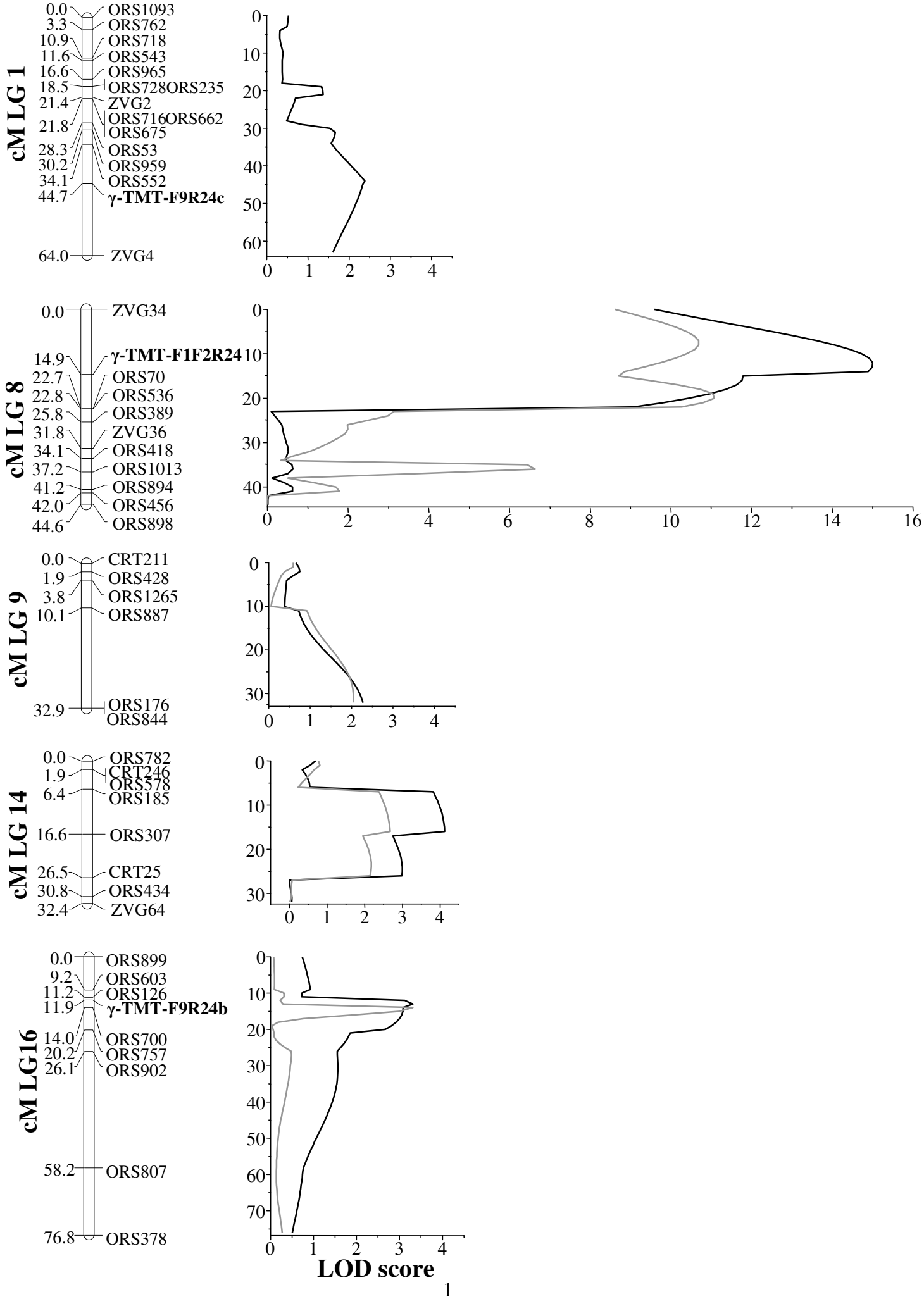

cM LG 1
0.0 ORS1093
3.3 ORS762
10.9 ORS718
11.6 ORS543
16.6 ORS965
18.5 ORS728ORS235
21.4 ZVG2
ORS716ORS662
21.8 ORS675
28.3 ORS53
30.2 ORS959
34.1 ORS552
44.7 γ-TMT-F9R24c
64.0 ZVG4
cM LG 8
0.0 ZVG34
14.9 γ-TMT-F1F2R24
22.7 ORS70
22.8 ORS536
25.8 ORS389
31.8 ZVG36
34.1 ORS418
37.2 ORS1013
41.2 ORS894
42.0 ORS456
44.6 ORS898
cM LG 9
0.0 CRT211
1.9 ORS428
3.8 ORS1265
10.1 ORS887
32.9 ORS176
ORS844
cM LG 14
0.0 ORS782
1.9 CRT246
ORS578
6.4 ORS185
16.6 ORS307
26.5 CRT25
30.8 ORS434
32.4 ZVG64
cM LG16
0.0 ORS899
9.2 ORS603
11.2 ORS126
11.9 γ-TMT-F9R24b
14.0 ORS700
20.2 ORS757
26.1 ORS902
58.2 ORS807
76.8 ORS378
LOD score

Figure 5

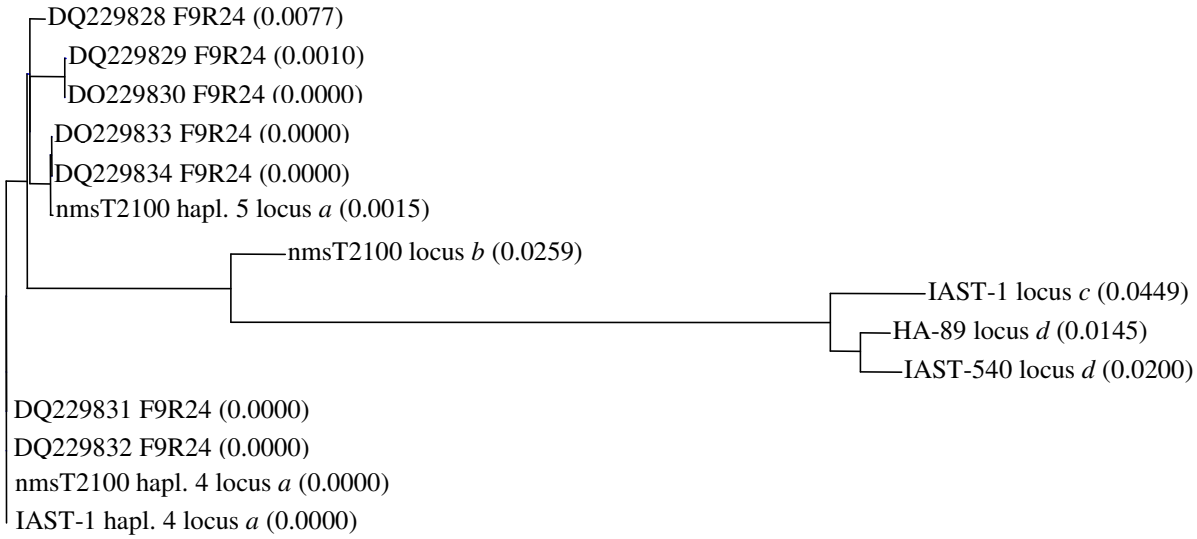

DQ229828 F9R24 (0.0077)
DQ229829 F9R24 (0.0010)
DO229830 F9R24 (0.0000)
DO229833 F9R24 (0.0000)
DQ229834 F9R24 (0.0000)
nmsT2100 hapl. 5 locus a (0.0015)
nmsT2100 locus b (0.0259)
IAST-1 locus c (0.0449)
HA-89 locus d (0.0145)
IAST-540 locus d (0.0200)
DQ229831 F9R24 (0.0000)
DQ229832 F9R24 (0.0000)
nmsT2100 hapl. 4 locus a (0.0000)
IAST-1 hapl. 4 locus a (0.0000)

Figure 6

Section 1
(1) 1 10 20 30 40 50 60 79
EF495161 HELIANT-g-TMT_ mRNA (1) ----------ACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
DQ229828_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
DQ229829_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACC------GCCAAA
DQ229830_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACGTCACCACCACCACCGCCAAA
DQ229831_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
DQ229832_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
IAST-1_haplotype_4 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
nmsT2100_haplotype_4 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
DQ229833_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACGTCACCACCACCACCGCCAAA
DQ229834_F1R92 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACGTCACCACCACCACCGCCAAA
IAST-1_haplotype_5 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACGTCACCACCACCACCGCCAAA
nmsT2100_haplotype_5 (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
Consensus (1) ATGTCTATTTGAACACCACGTAAACTCTGTGAAGCTTTACGTGCCATTGTTGACACACATCACCACCACCACCGCCAAA
Section 2
(80) 80 90 100 110 120 130 140 158
EF495161 HELIANT-g-TMT_ mRNA (42) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229828_F1R92 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229829_F1R92 (74) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229830_F1R92 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229831_F1R92 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229832_F1R92 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
IAST-1_haplotype_4 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
nmsT2100_haplotype_4 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229833_F1R92 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
DQ229834_F1R92 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
IAST-1_haplotype_5 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
nmsT2100_haplotype_5 (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC
Consensus (80) TTCACCACTCACTCACACAACCTGCTATGGCTACGACGGCAGTTGGCGTATCGGCGACGCCGATGACGGAGAAGCTGAC

Section 3

		159	170	180	190	200	210	220	237
EF495161 HELIANT-g-TMT_ mRNA	(121)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229828_F1R92	(159)	GGCGGCAGATGATGACCAGCAACAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229829_F1R92	(153)	GGCGGCAGATGATGACCAGCAACAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229830_F1R92	(159)	GGCGGCAGATGATGACCAGCAACAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229831_F1R92	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229832_F1R92	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
IAST-1_haplotype_4	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
nmsT2100_haplotype_4	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229833_F1R92	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
DQ229834_F1R92	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
IAST-1_haplotype_5	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
nmsT2100_haplotype_5	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG
Consensus	(159)	GGCGGCAGATGATGACCAGCAGCAGCAGAAGCTCAAAAAAGGAATCGCAGAGTTCTACGACGAATCCTCAGGTATGTGG

Section 4

		238	250	260	270	280	290	300	316
EF495161 HELIANT-g-TMT_ mRNA	(200)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229828_F1R92	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229829_F1R92	(232)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229830_F1R92	(238)	GAGAACATTTGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229831_F1R92	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229832_F1R92	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
IAST-1_haplotype_4	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
nmsT2100_haplotype_4	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229833_F1R92	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
DQ229834_F1R92	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
IAST-1_haplotype_5	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
nmsT2100_haplotype_5	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG
Consensus	(238)	GAGAACATATGGGGAGAACACATGCATCACGGATATTATAACTCCGACGACGTCGTTGAACTCTCCGATCACCGTTCTG

Section 5

```
                                  (317) 317       330       340       350       360       370       380       395
EF495161 HELIANT-g-TMT_ mRNA      (279) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAG-------------------------------
DQ229828_F1R92                    (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
DQ229829_F1R92                    (311) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTATGTTTCAGGTAGTTATCAGTTGATTTATTCTGTTATTGT
DQ229830_F1R92                    (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTTGATTTATTCTGTTATTGT
DQ229831_F1R92                    (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
DQ229832_F1R92                    (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
IAST-1_haplotype_4                (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
nmsT2100_haplotype_4              (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
DQ229833_F1R92                    (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
DQ229834_F1R92                    (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
IAST-1_haplotype_5                (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
nmsT2100_haplotype_5              (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
Consensus                         (317) CTCAGATCCGTATGATTGAACAAGCCCTAACGTTCGCCTCTGTTTCAGGTAGTTATCAGTGGATTTATTCTGTTATTGT
```

Section 6

```
                                  (396) 396       410       420       430       440       450       460       474
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------------
DQ229828_F1R92                    (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAATTATGAAAA-GCGCTTTCTGAACGGAGTTTTC
DQ229829_F1R92                    (390) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAGTTATGAAAAAGCGGTTTCTGAACGGAGTTTTC
DQ229830_F1R92                    (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAGTTATGAAAAAGCGGTTTCTGAACGGAGTTTTC
DQ229831_F1R92                    (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAATTATGAAAA-GCGCTTTCTGAACGGAGTTTTC
DQ229832_F1R92                    (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAATTATGAAAA-GCGCTTTCTGAACGGAGTTTTC
IAST-1_haplotype_4                (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAATTATGAAAA-GCGCTTTCTGAACGGAGTTTTC
nmsT2100_haplotype_4              (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAAATTATGAAAA-GCGCTTTCTGAACGGAGTTTTC
DQ229833_F1R92                    (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTAT---------TATGAAAA-GCGCTTTCTGAACGGAGTTTTC
DQ229834_F1R92                    (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTAT---------TATGAAAA-GCGCTTTCTGAACGGAGTTTTC
IAST-1_haplotype_5                (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTAT---------TATGAAAA-GCGGTGTCTGAACGGAGTTTTC
nmsT2100_haplotype_5              (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTAT---------TATGAAAA-GCGGTGTCTGAACGGAGTTTTC
Consensus                         (396) TTAATTGAATCTGACGGTTTAGAACTTTTAGATTTGTATAGATGAA TTATGAAAA GCGCTTTCTGAACGGAGTTTTC
```

Section 7

```
                                  (475) 475     480         490         500         510         520         530         540         553
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------------------------
DQ229828_F1R92                    (474) GATGTTTCGATGCGGTTTTTGAGTTTTATAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
DQ229829_F1R92                    (469) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
DQ229830_F1R92                    (475) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
DQ229831_F1R92                    (474) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTCTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
DQ229832_F1R92                    (474) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTCTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
IAST-1_haplotype_4                (474) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTCTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
nmsT2100_haplotype_4              (474) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTCTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
DQ229833_F1R92                    (465) GA--------TGCGGTTTTTGAGTTTTATAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
DQ229834_F1R92                    (465) GA--------TGCGGTTTTTGAGTTTTATAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
IAST-1_haplotype_5                (465) GA--------TGCGGTTTTTGAGTTTTATAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
nmsT2100_haplotype_5              (465) GA--------TGCGGTTTTTGAGTTTTATAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
Consensus                         (475) GATGTTTCGATGCGGTTTTTGAGTTTTAAAATTTGTACTCACTTTTTCTGAATGATTTTATTTGCTATTAGTTTCGATG
```

Section 8

```
                                  (554) 554     560         570         580         590         600         610         620         632
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------------------------
DQ229828_F1R92                    (553) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
DQ229829_F1R92                    (548) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
DQ229830_F1R92                    (554) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
DQ229831_F1R92                    (553) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
DQ229832_F1R92                    (553) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
IAST-1_haplotype_4                (553) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
nmsT2100_haplotype_4              (553) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
DQ229833_F1R92                    (536) GTGCTATAGTTGATTGTGAATGGTATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
DQ229834_F1R92                    (536) GTGCTATAGTTGATTGTGAATGGTATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
IAST-1_haplotype_5                (536) GTGCTATAGTTGATTGTGAATGGTATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
nmsT2100_haplotype_5              (536) GTGCTATAGTTGATTGTGAATGGTATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
Consensus                         (554) GTGCTATAGTTGATTGTGAATGATATCCTTACTGATTATGTGTTTGTTATAAGATTTGTAATTGTTAATTAGGTTTTTG
```

Section 9

```
                                  (633) 633       640       650       660       670       680       690       700       711
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------------------
DQ229828_F1R92                    (632) AGAACTAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGATTTCGAAG
DQ229829_F1R92                    (627) AGAACTAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGGTTTCGAAG
DQ229830_F1R92                    (633) AGAACTAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGGTTTCGAAG
DQ229831_F1R92                    (632) AGAACTATGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAATGGAGATTTCGAAG
DQ229832_F1R92                    (632) AGAACTATGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAATGGAGATTTCGAAG
IAST-1_haplotype_4                (632) AGAACTATGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAATGGAGATTTCGAAG
nmsT2100_haplotype_4              (632) AGAACTATGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAATGGAGATTTCGAAG
DQ229833_F1R92                    (615) AGAACCAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGATTTCGAAG
DQ229834_F1R92                    (615) AGAACCAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGATTTCGAAG
IAST-1_haplotype_5                (615) AGAACCAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGATTTCGAAG
nmsT2100_haplotype_5              (615) AGAACTAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGATTTCGAAG
Consensus                         (633) AGAACTAAGTTTGAGTTATAGCTGAATAATCTATGCCGATTTTGAATTAGACGTGTGTTTTCGAACGGAGATTTCGAAG
```

Section 10

```
                                  (712) 712       720       730       740       750       760       770       780       790
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------------------
DQ229828_F1R92                    (711) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAGTGC
DQ229829_F1R92                    (706) TTTGATGCGGTTTCTGATTTACAAATACTTTAGGCTCGCTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
DQ229830_F1R92                    (712) TTTGATGCGGTTTCTGATTTACAAATACTTTAGGCTCGCTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
DQ229831_F1R92                    (711) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
DQ229832_F1R92                    (711) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
IAST-1_haplotype_4                (711) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
nmsT2100_haplotype_4              (711) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
DQ229833_F1R92                    (694) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCCTATGC
DQ229834_F1R92                    (694) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCCTATGC
IAST-1_haplotype_5                (694) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCCTATGC
nmsT2100_haplotype_5              (694) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCCTATGC
Consensus                         (712) TTTGATGCGGTTTCTGATTTACAAATATTTTATACTCACTTTTGGTGGATGAATCGTTTGTTATTGAATTTGCTAATGC
```

Section 11

```
                                  (791) 791       800       810       820       830       840       850       869
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------
DQ229828_F1R92                    (790) TACAATAGTTTGGGAATGATTGATTACAGGTTATGTGTTT-------GTTGTT---AGATTCATGACTGTCAATTAGGT
DQ229829_F1R92                    (785) TACAATAGTTTGGGAATGATTGGTTACAGGTTATGTGTTTTTTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
DQ229830_F1R92                    (791) TACAATAGTTTGGGAATGATTGGTTACAGGTTATGTGTTTTTTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
DQ229831_F1R92                    (790) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTT-------GTTGTTGTTAGATTCATGACTGTCAATTAGGT
DQ229832_F1R92                    (790) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTT-------GTTGTTGTTAGATTCATGACTGTCAATTAGGT
IAST-1_haplotype_4                (790) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTT-------GTTGTTGTTAGATTCATGACTGTCAATTAGGT
nmsT2100_haplotype_4              (790) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTT-------GTTGTTGTTAGATTCATGACTGTCAATTAGGT
DQ229833_F1R92                    (773) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTTGTTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
DQ229834_F1R92                    (773) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTTGTTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
IAST-1_haplotype_5                (773) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTTGTTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
nmsT2100_haplotype_5              (773) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTTGTTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
Consensus                         (791) TACAATAGTTTGTGAATGATTGGTTACAGGTTATGTGTTT TTGTTAGTTGTTGTTAGATTCATGACTGTCAATTAGGT
```

Section 12

```
                                  (870) 870       880       890       900       910       920       930       948
EF495161 HELIANT-g-TMT_ mRNA      (327) ------------------------------------------------------------------------
DQ229828_F1R92                    (859) TTTTGAAAATTGAAGTTGAATTCATGGCTGAAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAGTG
DQ229829_F1R92                    (864) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAATG
DQ229830_F1R92                    (870) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAATG
DQ229831_F1R92                    (862) TTTTGATAATTGAAGTTGAATTCATGGCTGGAATCTCACAAGGATTGAAATAAAC-GAAAATTGTACTTTTACTGAGTG
DQ229832_F1R92                    (862) TTTTGATAATTGAAGTTGAATTCATGGCTGGAATCTCACAAGGATTGAAATAAAC-GAAAATTGTACTTTTACTGAGTG
IAST-1_haplotype_4                (862) TTTTGATAATTGAAGTTGAATTCATGGCTGGAATCTCACAAGGATTGAAATAAAC-GAAAATTGTACTTTTACTGAGTG
nmsT2100_haplotype_4              (862) TTTTGATAATTGAAGTTGAATTCATGGCTGGAATCTCACAAGGATTGAAATAAAC-GAAAATTGTACTTTTACTGAGTG
DQ229833_F1R92                    (852) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAGTG
DQ229834_F1R92                    (852) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAGTG
IAST-1_haplotype_5                (852) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAGTG
nmsT2100_haplotype_5              (852) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAGTG
Consensus                         (870) TTTTGAAAATTGAAGTTGATTTCATGGCTGTAATCTCACAAGGATTGAAATTAAATGAAAATTGTACTTTTACTGAGTG
```

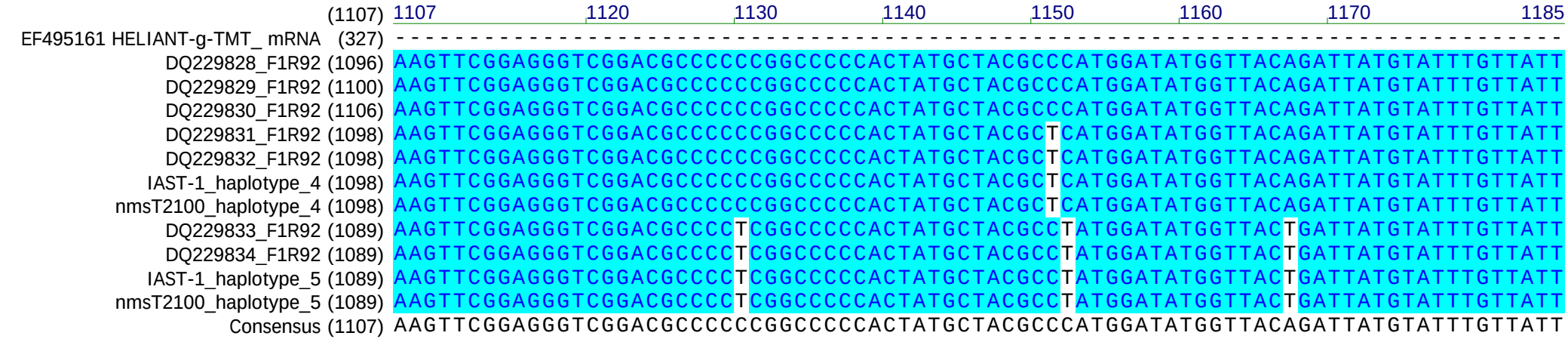
Section 13

```
                                (949) 949       960       970       980       990       1000      1010      1027
EF495161 HELIANT-g-TMT_ mRNA    (327) ------------------------------------------------------------------------
DQ229828_F1R92                  (938) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGCTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAT
DQ229829_F1R92                  (943) AATTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
DQ229830_F1R92                  (949) AATTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
DQ229831_F1R92                  (940) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAATGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
DQ229832_F1R92                  (940) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAATGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
IAST-1_haplotype_4              (940) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAATGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
nmsT2100_haplotype_4            (940) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAATGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
DQ229833_F1R92                  (931) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
DQ229834_F1R92                  (931) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
IAST-1_haplotype_5              (931) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
nmsT2100_haplotype_5            (931) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
Consensus                       (949) AACTCATTTGCTAATGCTACTGCAGTTTGTGAATGATATGGTCAGTGGCGAAGCTTGACCTAAAATTATACCTAAAAAA
```

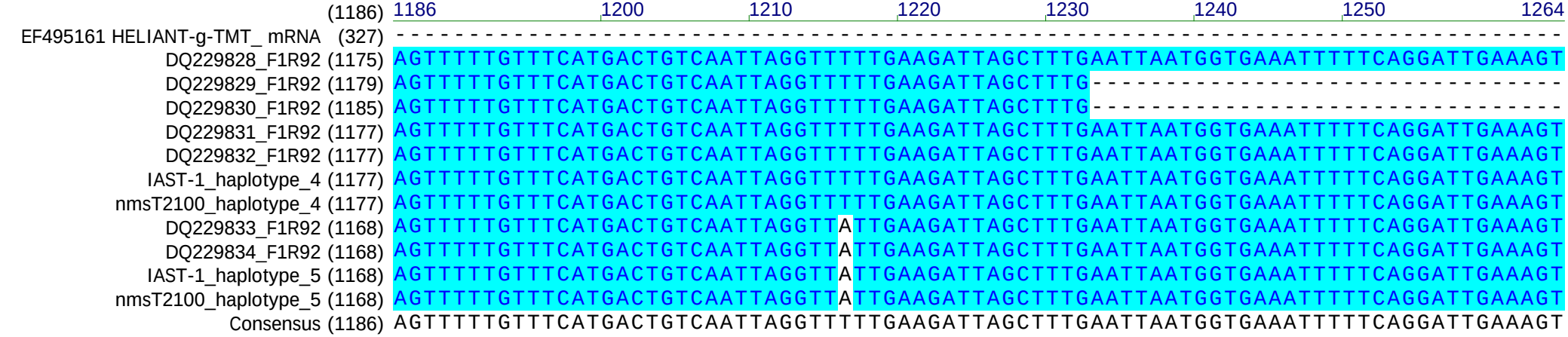
Section 14

```
                                (1028) 1028      1040      1050      1060      1070      1080      1090      1106
EF495161 HELIANT-g-TMT_ mRNA    (327) ------------------------------------------------------------------------
DQ229828_F1R92                  (1017) TTATAAAACCGGGGGGACGAAAACGTATATACCTAAAAAATTATATACGAAAACTACATACCGGACACTACTGAGCGAA
DQ229829_F1R92                  (1022) TTATAAAACC-GGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
DQ229830_F1R92                  (1028) TTATAAAACC-GGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
DQ229831_F1R92                  (1019) TTATAAAACTGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
DQ229832_F1R92                  (1019) TTATAAAACTGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
IAST-1_haplotype_4              (1019) TTATAAAACTGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
nmsT2100_haplotype_4            (1019) TTATAAAACTGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
DQ229833_F1R92                  (1010) TTATAAAACCGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACTGGACACTACTGAGCGAA
DQ229834_F1R92                  (1010) TTATAAAACCGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACTGGACACTACTGAGCGAA
IAST-1_haplotype_5              (1010) TTATAAAACCGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACTGGACACTACTGAGCGAA
nmsT2100_haplotype_5            (1010) TTATAAAACCGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACTGGACACTACTGAGCGAA
Consensus                       (1028) TTATAAAACCGGGGGGACGAAAACGTATATACCTAAAAAATTCTATACGAAAACTACATACCGGACACTACTGAGCGAA
```

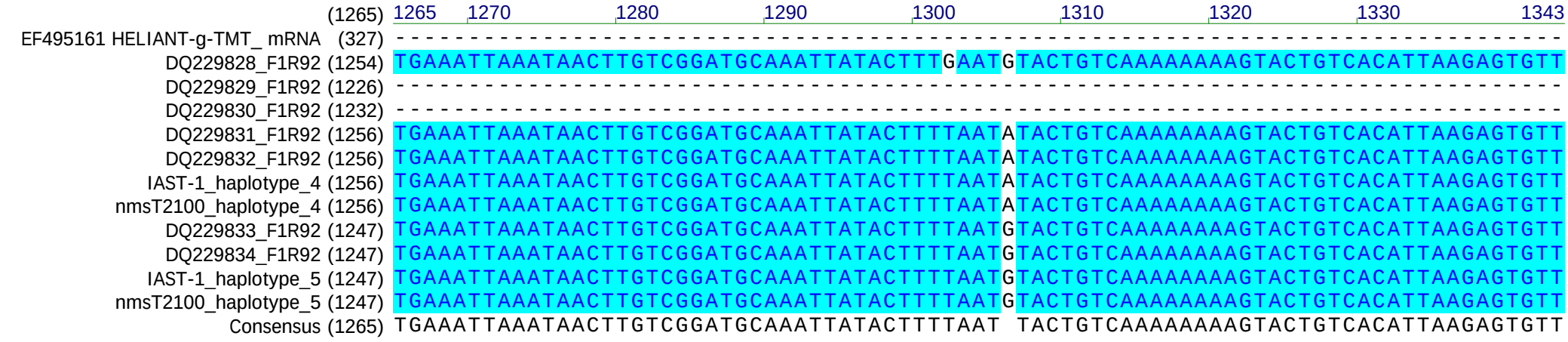
Section 15



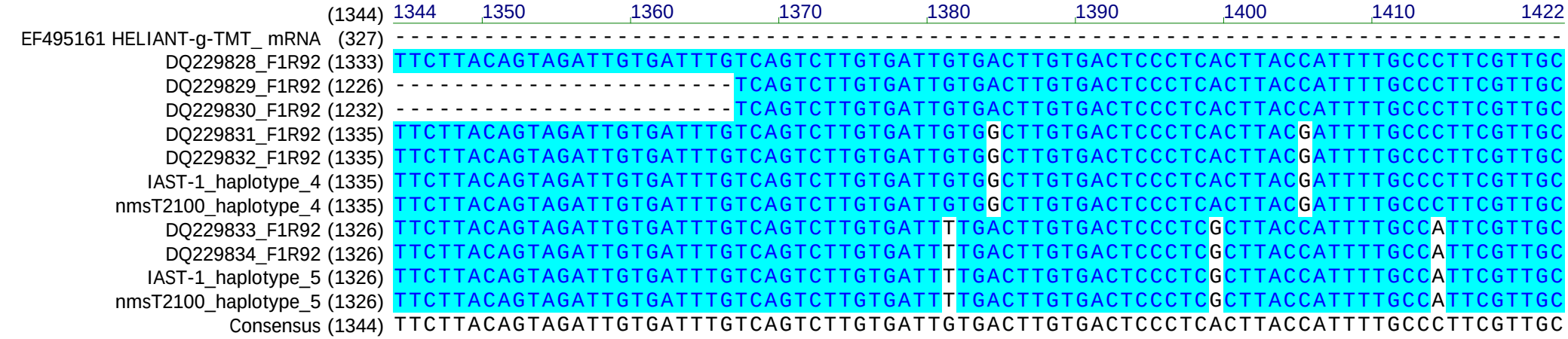
Section 16

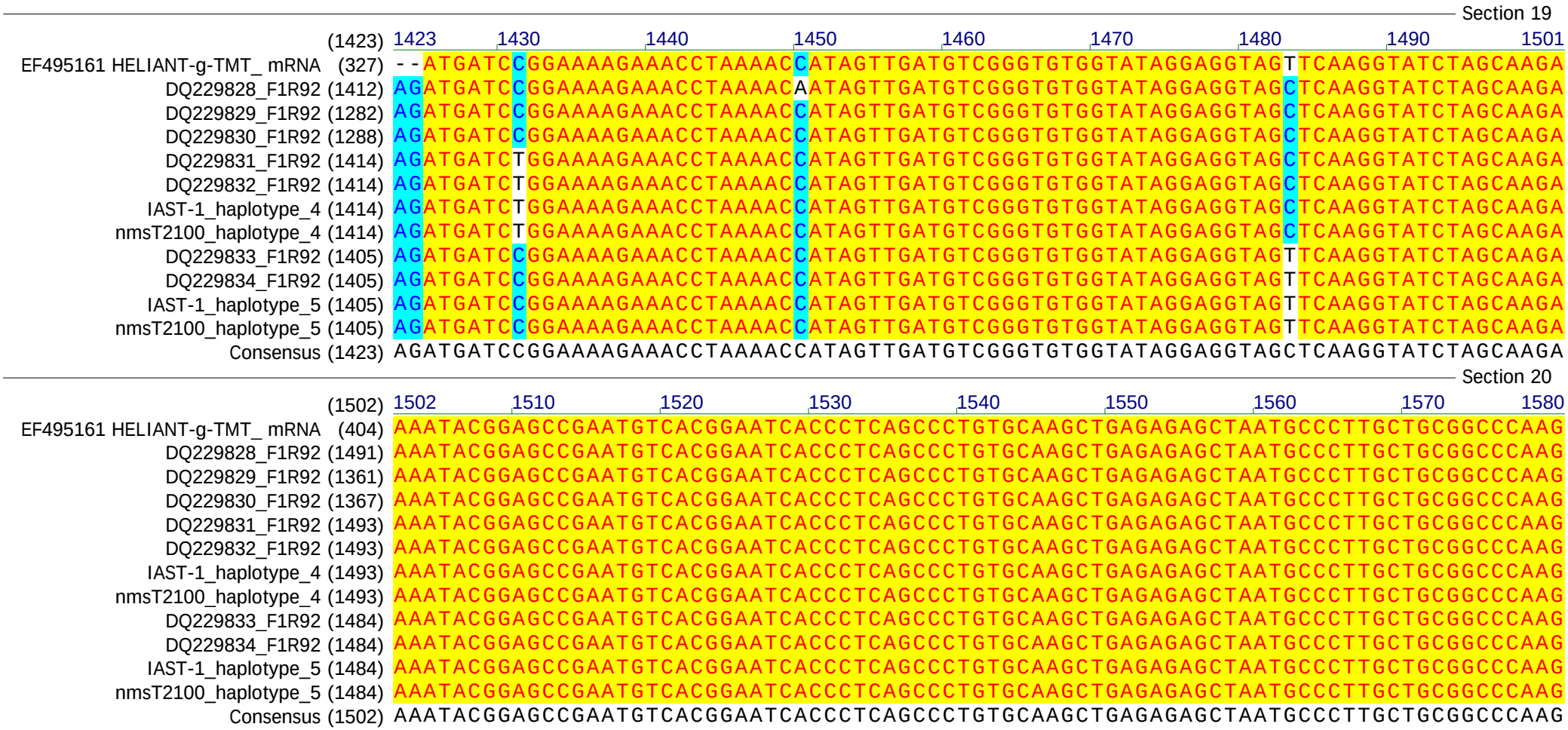


Section 17

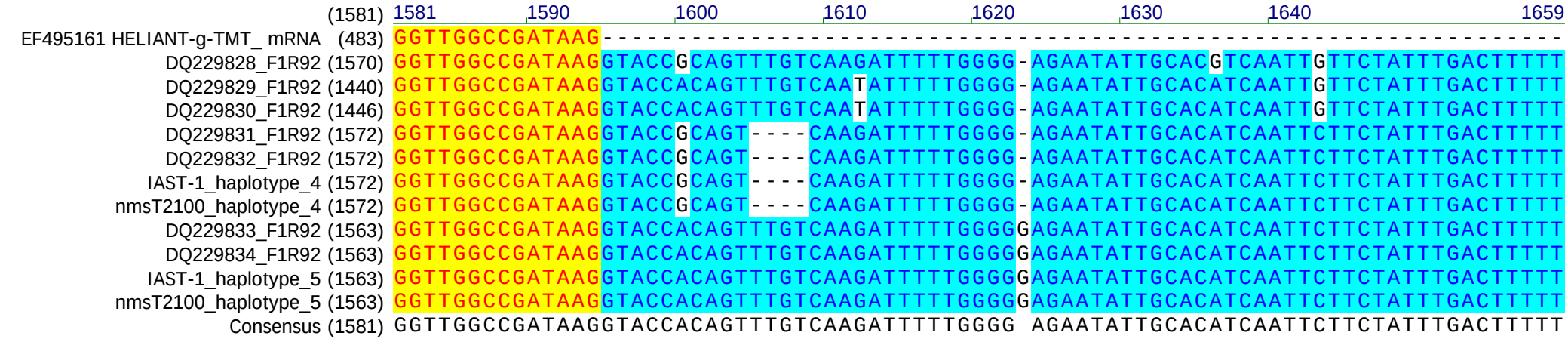


Section 18

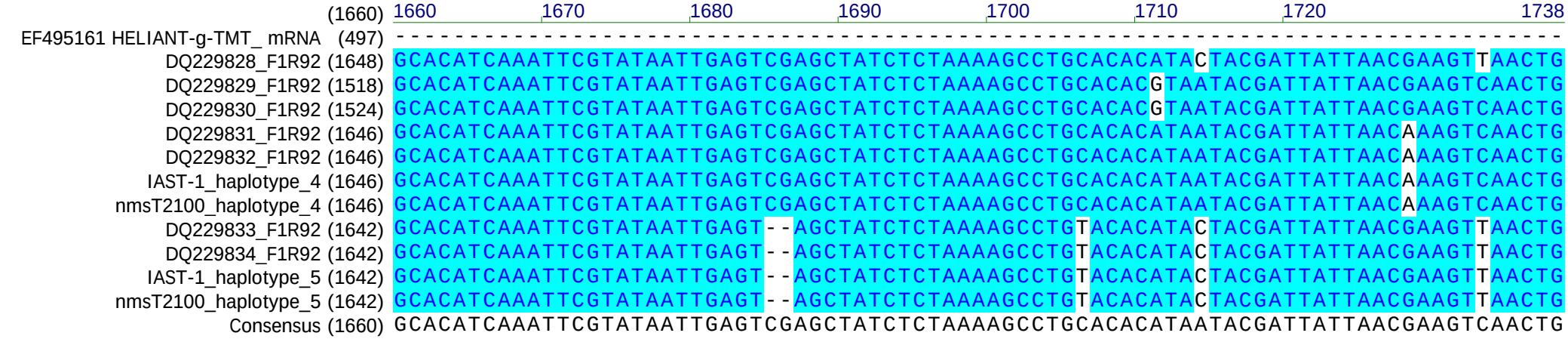



Section 19
(1423) 1423 1430 1440 1450 1460 1470 1480 1490 1501
EF495161 HELIANT-g-TMT_ mRNA (327) --ATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGTTCAAGGTATCTAGCAAGA
DQ229828_F1R92 (1412) AGATGATCCGGAAAAGAAACCTAAAACAATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
DQ229829_F1R92 (1282) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
DQ229830_F1R92 (1288) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
DQ229831_F1R92 (1414) AGATGATCTGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
DQ229832_F1R92 (1414) AGATGATCTGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
IAST-1_haplotype_4 (1414) AGATGATCTGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
nmsT2100_haplotype_4 (1414) AGATGATCTGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
DQ229833_F1R92 (1405) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGTTCAAGGTATCTAGCAAGA
DQ229834_F1R92 (1405) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGTTCAAGGTATCTAGCAAGA
IAST-1_haplotype_5 (1405) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGTTCAAGGTATCTAGCAAGA
nmsT2100_haplotype_5 (1405) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGTTCAAGGTATCTAGCAAGA
Consensus (1423) AGATGATCCGGAAAAGAAACCTAAAACCATAGTTGATGTCGGGTGTGGTATAGGAGGTAGCTCAAGGTATCTAGCAAGA
Section 20
(1502) 1502 1510 1520 1530 1540 1550 1560 1570 1580
EF495161 HELIANT-g-TMT_ mRNA (404) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229828_F1R92 (1491) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229829_F1R92 (1361) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229830_F1R92 (1367) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229831_F1R92 (1493) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229832_F1R92 (1493) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
IAST-1_haplotype_4 (1493) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
nmsT2100_haplotype_4 (1493) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229833_F1R92 (1484) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
DQ229834_F1R92 (1484) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
IAST-1_haplotype_5 (1484) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
nmsT2100_haplotype_5 (1484) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG
Consensus (1502) AAATACGGAGCCGAATGTCACGGAATCACCCTCAGCCCTGTGCAAGCTGAGAGAGCTAATGCCCTTGCTGCGGCCCAAG

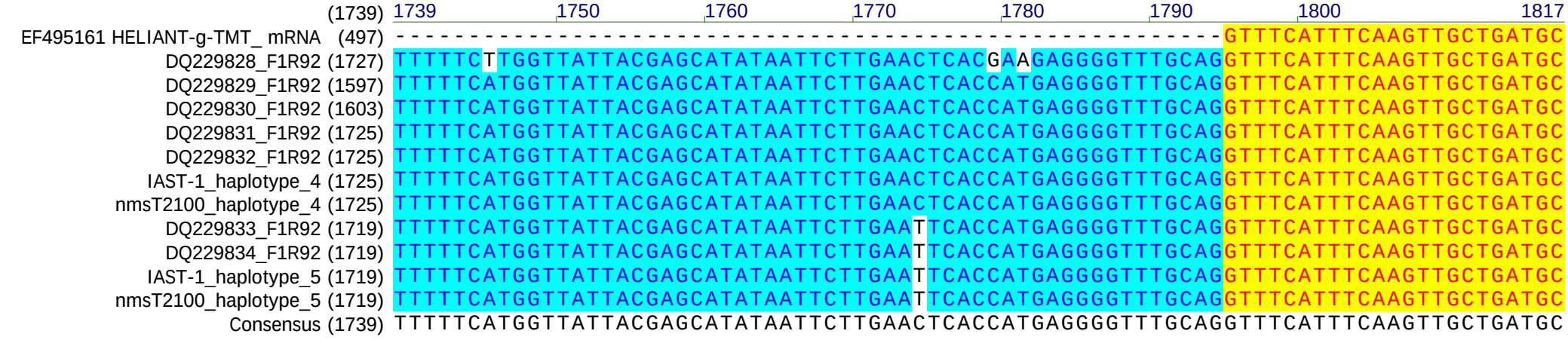
Section 21



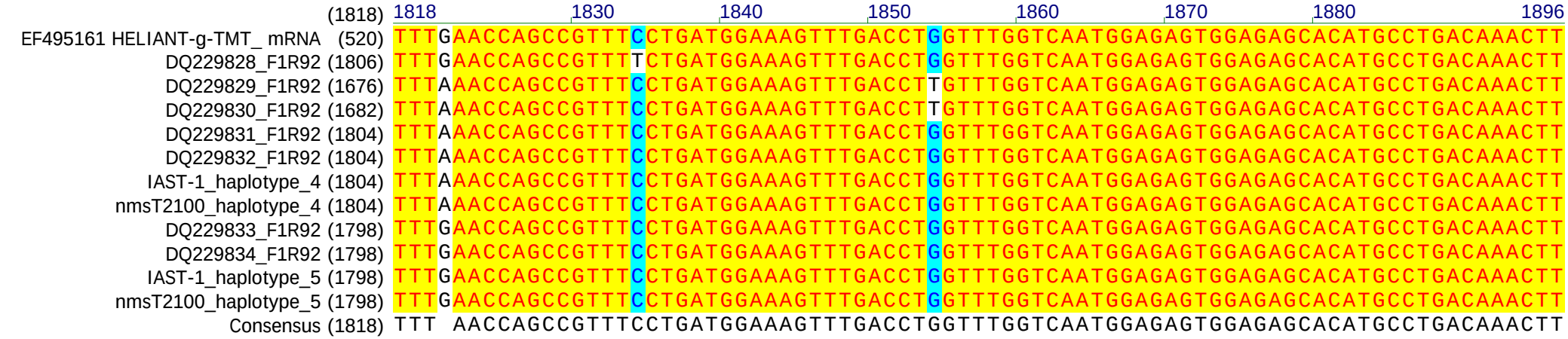
Section 22



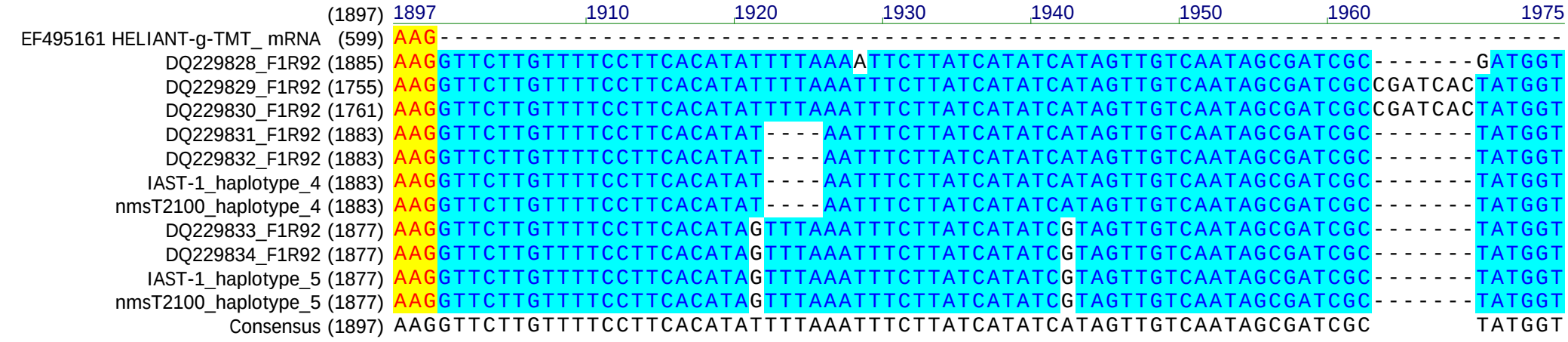
Section 23



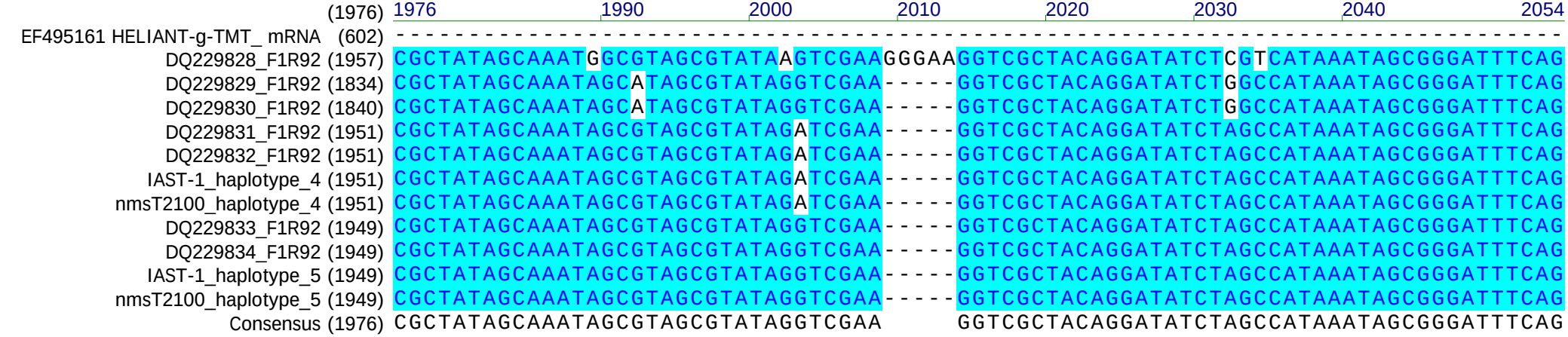
Section 24



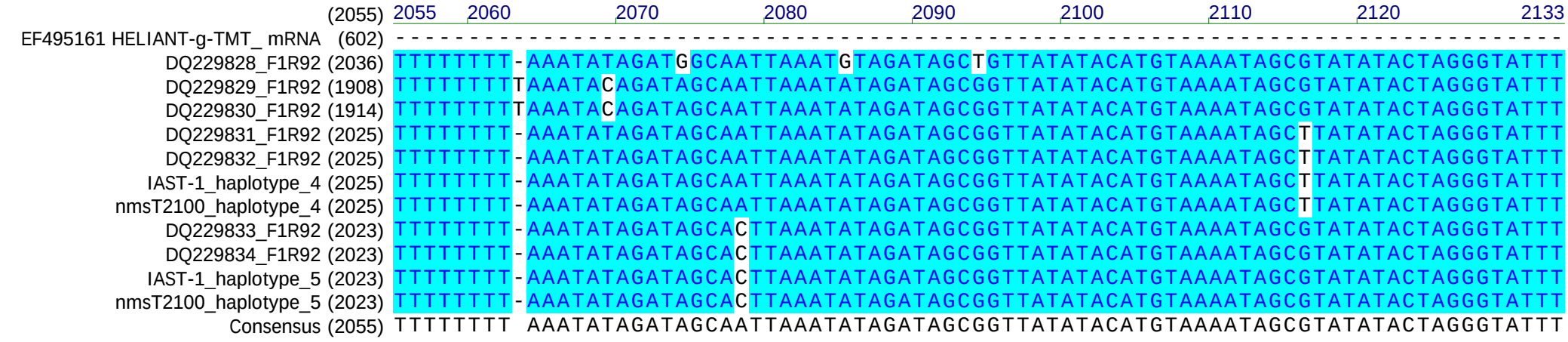
Section 25



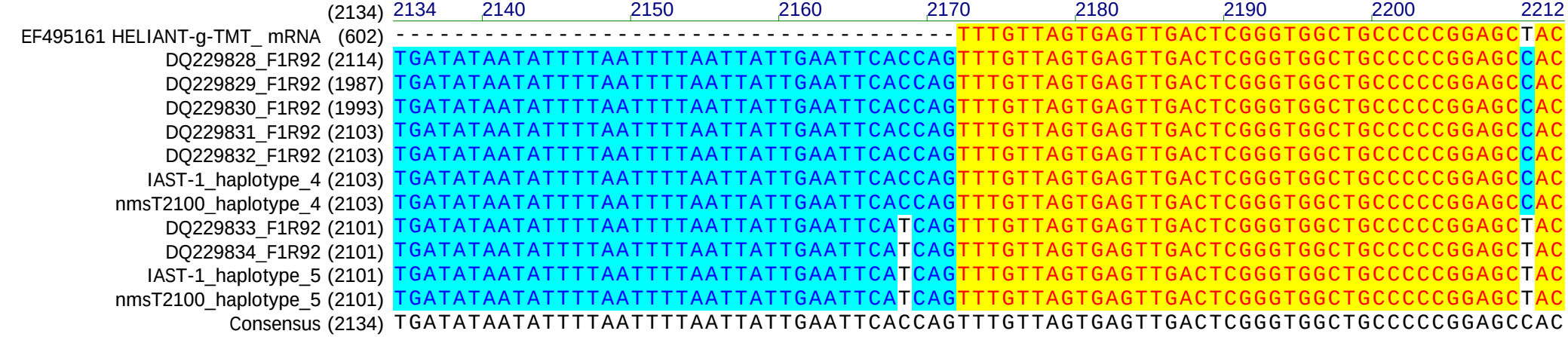
Section 26



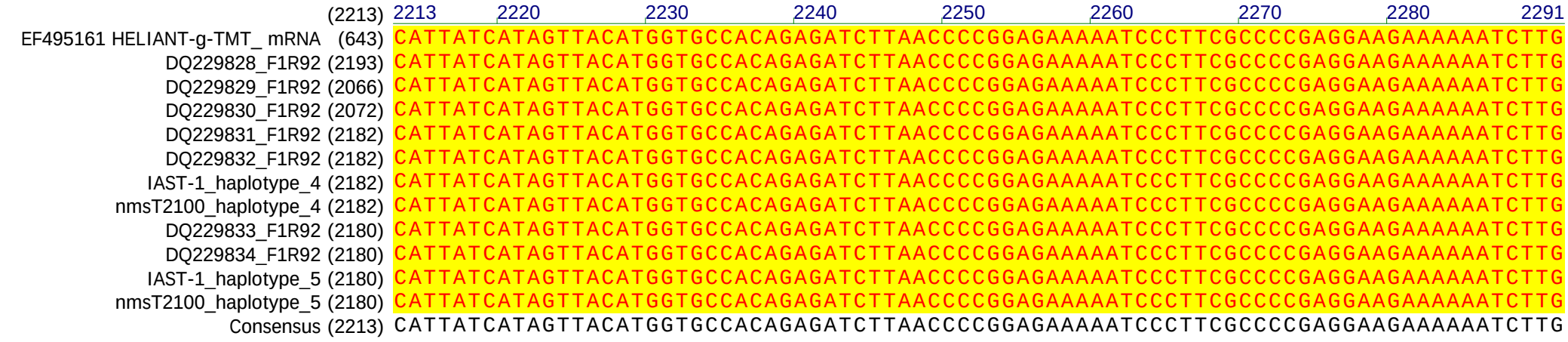
Section 27



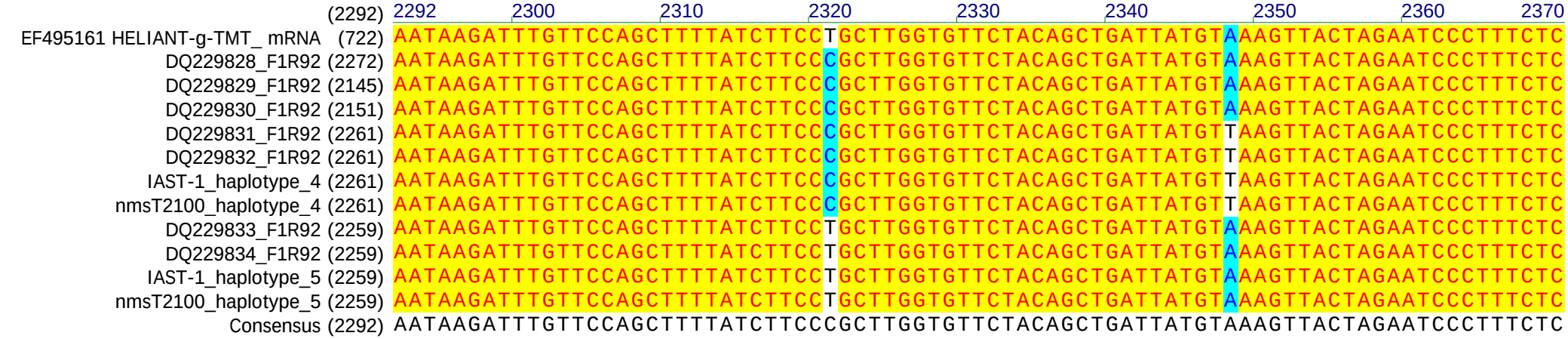
Section 28



Section 29



Section 30



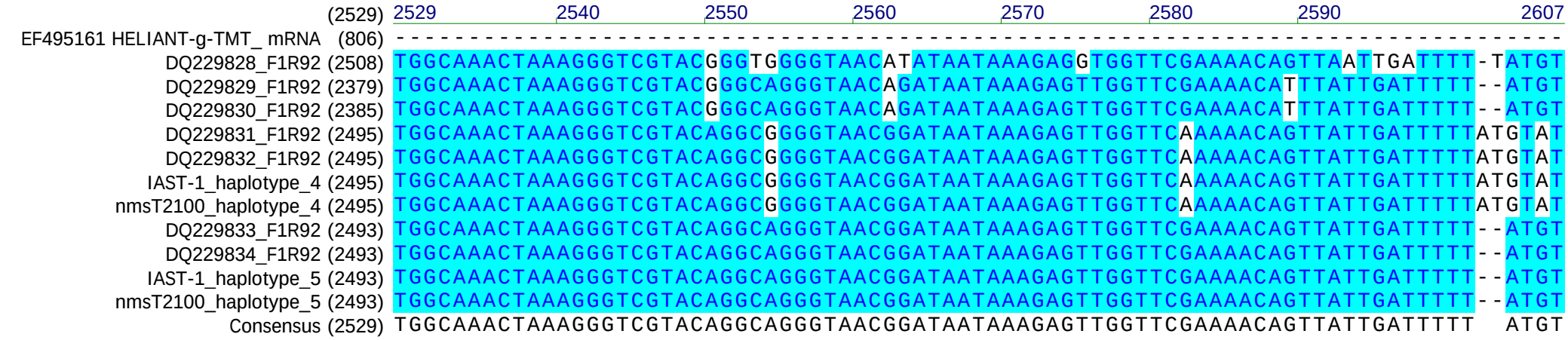
Section 31

```
                                (2371) 2371      2380      2390      2400      2410      2420      2430                2449
EF495161 HELIANT-g-TMT_ mRNA    (801) TTCAG-------------------------------------------------------------------------
DQ229828_F1R92                  (2351) TTCAGGTAAACTTCATTATTAATCGCCCAAAACCTAGTTTAATTTCGGAGTTTATATTAATACTATTTGCAAACATTTA
DQ229829_F1R92                  (2224) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTGATA---TTTGCAAACATTTA
DQ229830_F1R92                  (2230) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTGATA---TTTGCAAACATTTA
DQ229831_F1R92                  (2340) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTAATA---TTTGCAAACATTTA
DQ229832_F1R92                  (2340) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTAATA---TTTGCAAACATTTA
IAST-1_haplotype_4              (2340) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTAATA---TTTGCAAACATTTA
nmsT2100_haplotype_4            (2340) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTAATA---TTTGCAAACATTTA
DQ229833_F1R92                  (2338) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCAGAGTTTATATTAATA---TTTGCAAACATTTA
DQ229834_F1R92                  (2338) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCAGAGTTTATATTAATA---TTTGCAAACATTTA
IAST-1_haplotype_5              (2338) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCAGAGTTTATATTAATA---TTTGCAAACATTTA
nmsT2100_haplotype_5            (2338) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCAGAGTTTATATTAATA---TTTGCAAACATTTA
Consensus                       (2371) TTCAGGTAAACTTCATTATTAATCACCCAAAACCTAGTTTAATTTCGGAGTTTATATTAATA    TTTGCAAACATTTA
```

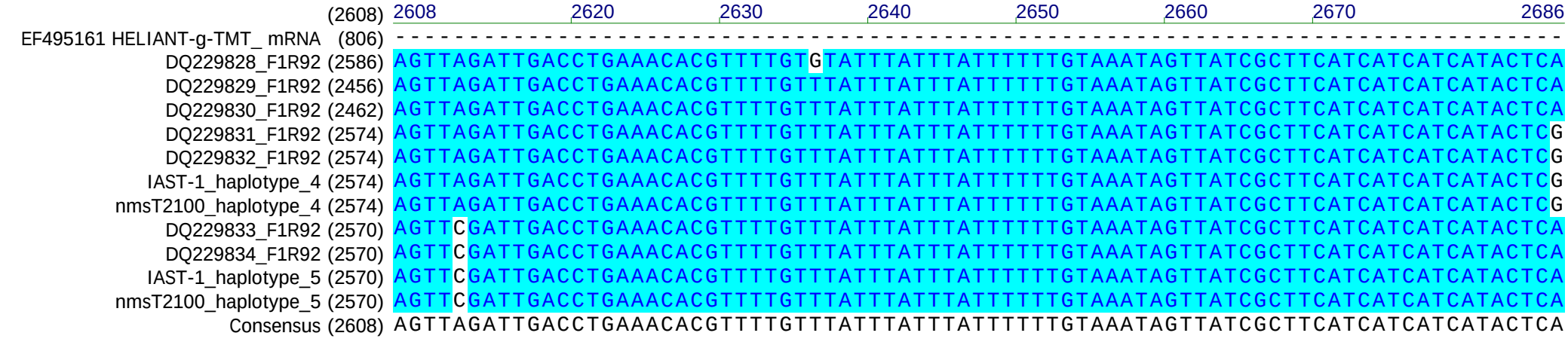
Section 32

```
                                (2450) 2450      2460      2470      2480      2490      2500      2510                2528
EF495161 HELIANT-g-TMT_ mRNA    (806) -------------------------------------------------------------------------------
DQ229828_F1R92                  (2430) GTTACCTTTAGGAATCTTGAGGGGG-CTACCAACTATCATACACTCCAACCACCTCTCTATCTTTGTCGGCCTGTGAAT
DQ229829_F1R92                  (2300) GTTACCTTTAGGAATCTTGAGGGGGGCTACCAACTATCATACACTCCAACCACCTCTCTATCTTTGTCGGTCTGTGAAT
DQ229830_F1R92                  (2306) GTTACCTTTAGGAATCTTGAGGGGGGCTACCAACTATCATACACTCCAACCACCTCTCTATCTTTGTCGGTCTGTGAAT
DQ229831_F1R92                  (2416) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
DQ229832_F1R92                  (2416) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
IAST-1_haplotype_4              (2416) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
nmsT2100_haplotype_4            (2416) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
DQ229833_F1R92                  (2414) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
DQ229834_F1R92                  (2414) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
IAST-1_haplotype_5              (2414) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
nmsT2100_haplotype_5            (2414) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
Consensus                       (2450) GTTAACTTTAAGAATCTTGAGGGGGGCTACCAACTACCGTACACTCCAACCACCTCTCTATGTCTGTCGGTCTGTGAAT
```

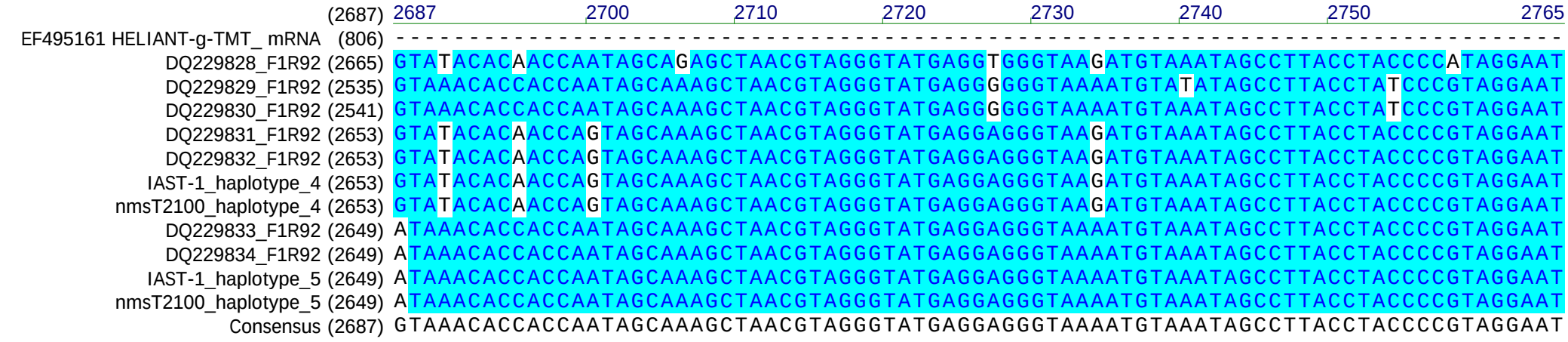
Section 33



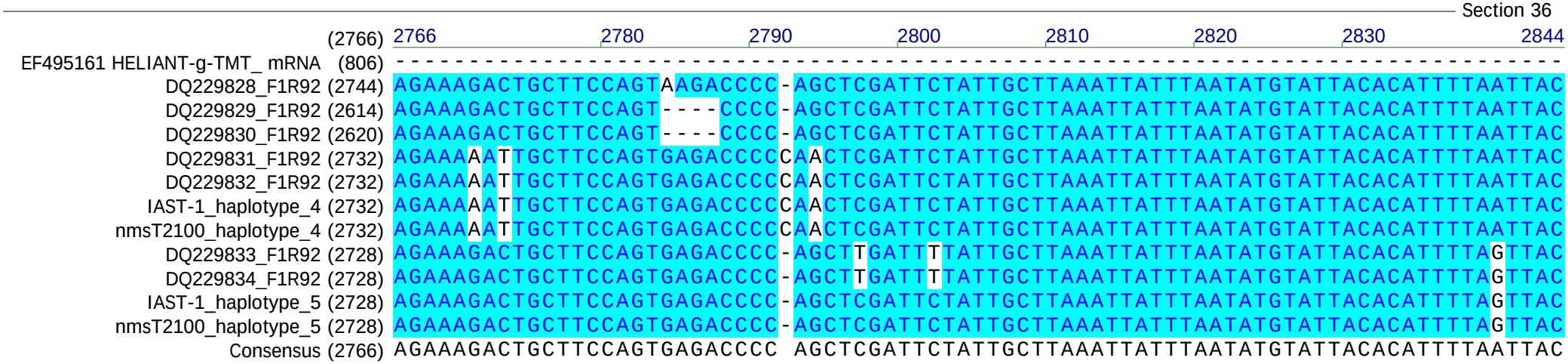
Section 34



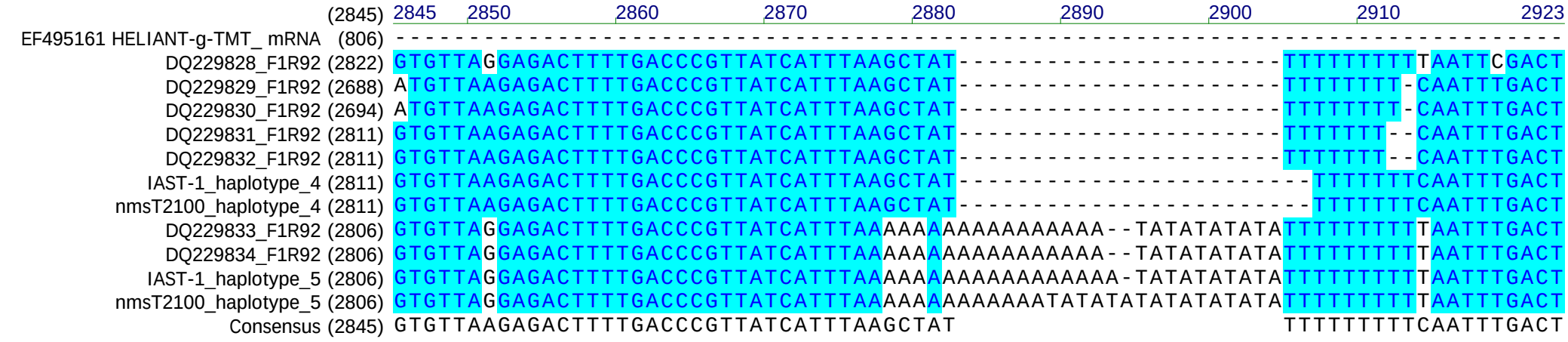
Section 35



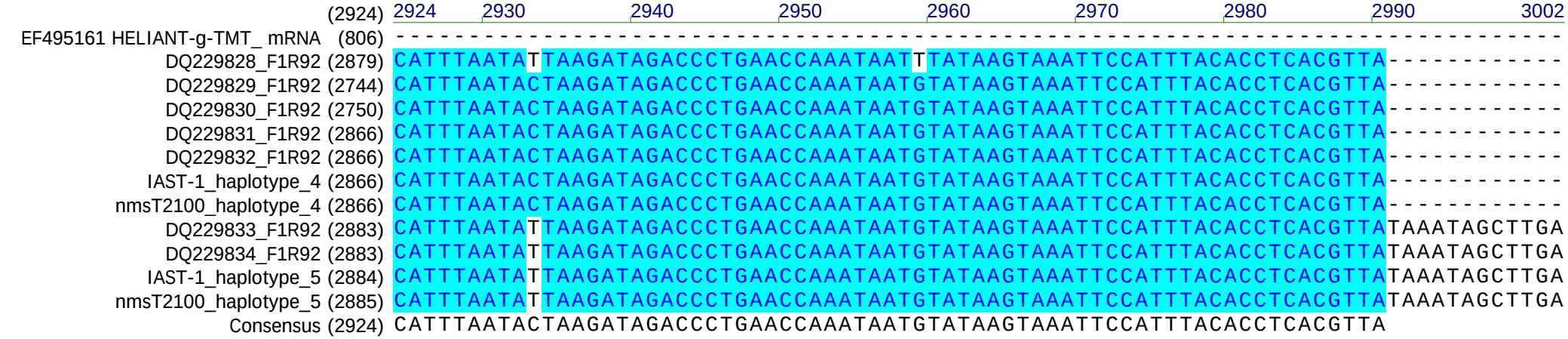
Section 36



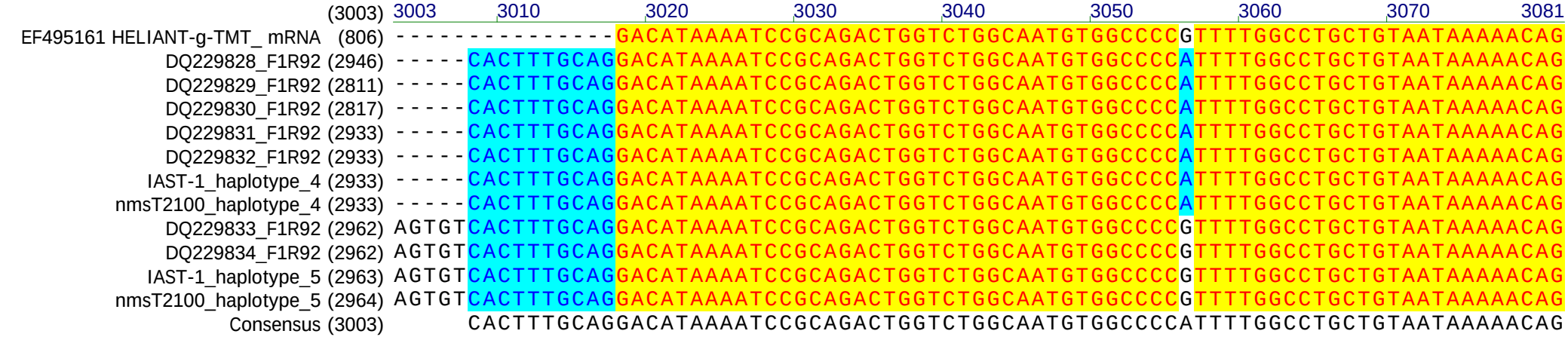
Section 37



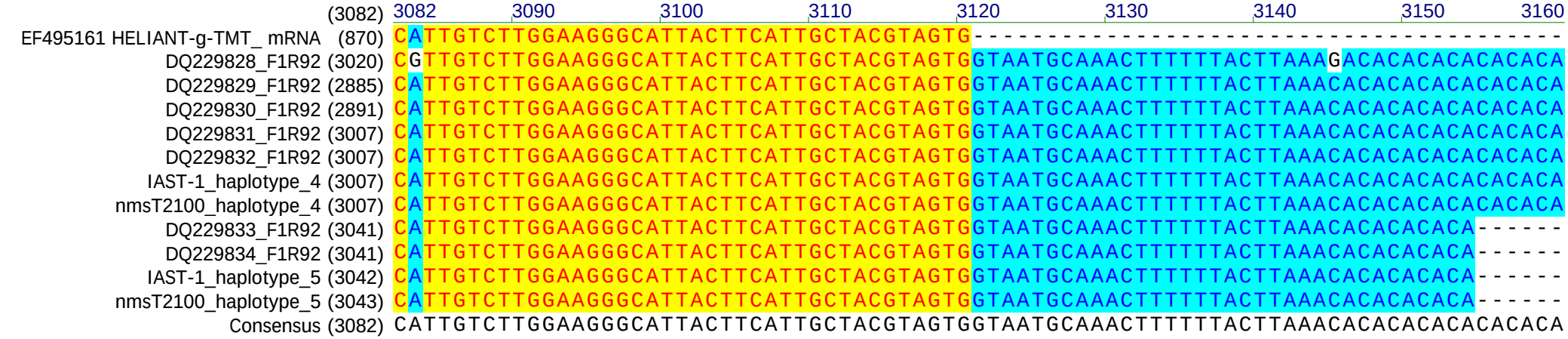
Section 38



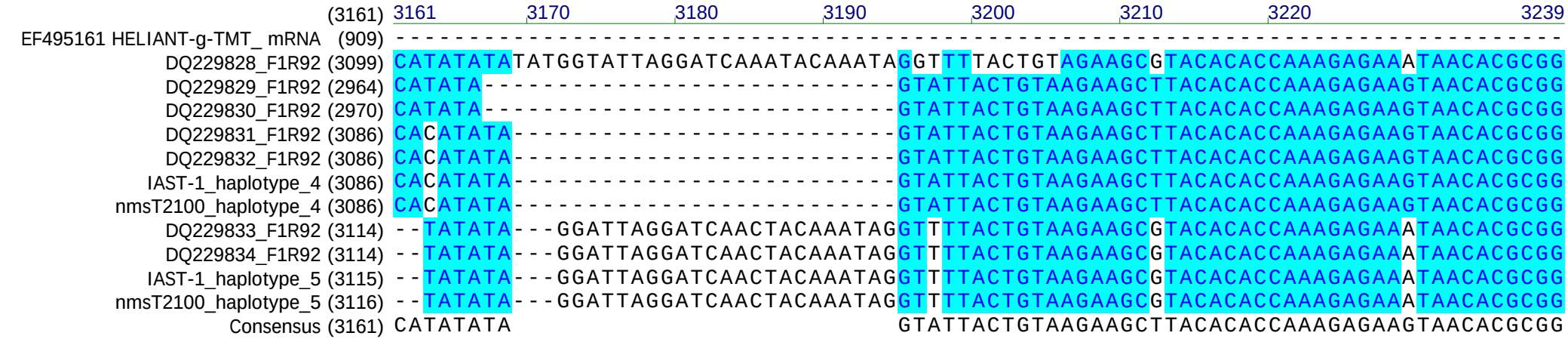
Section 39



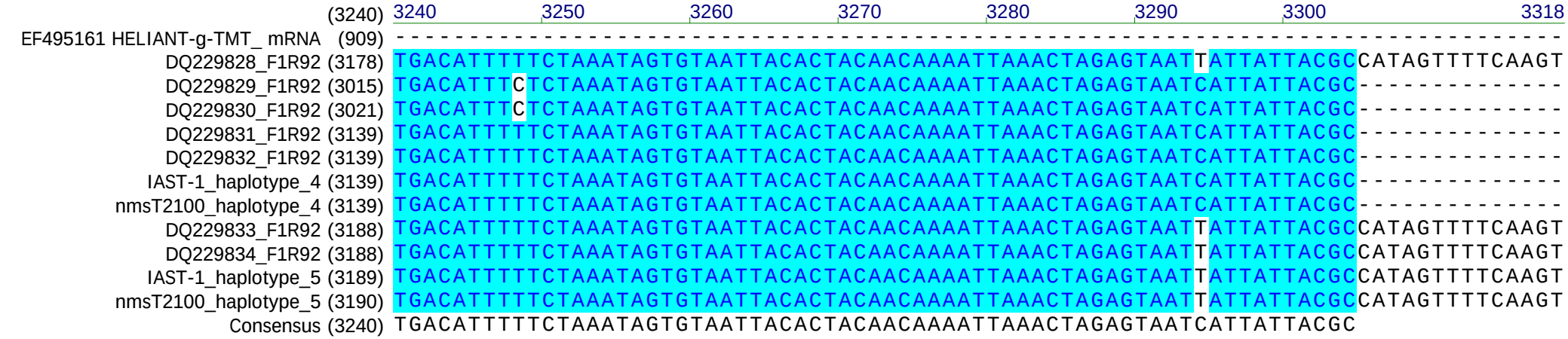
Section 40



Section 41



Section 42



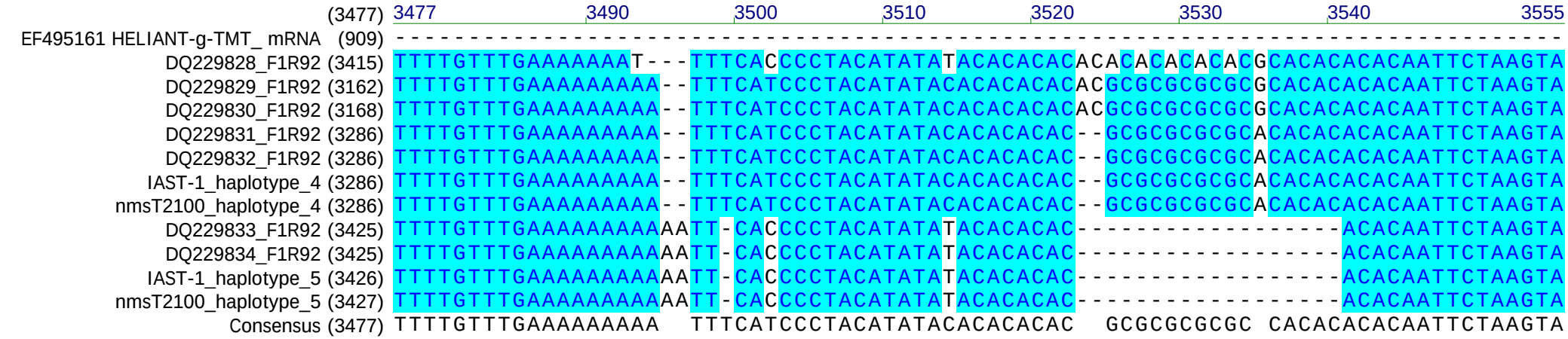
Section 43

```
                                (3319) 3319        3330        3340        3350        3360        3370        3380        3397
EF495161 HELIANT-g-TMT_ mRNA   (909) ----------------------------------------------------------------------------------------
DQ229828_F1R92                (3257) AGCGTAACGATAGATATGCTACTTCATTTGTTGTATTGTATCACACGCTACAAGAAAATTGAACTAGAGTAATTATTAT
DQ229829_F1R92                (3080) ----------------------------------------------------TACTTGAAAACTATGACGCAATAATAATTAC
DQ229830_F1R92                (3086) ----------------------------------------------------TACTTGAAAACTATGACGCAATAATAATTAC
DQ229831_F1R92                (3204) ----------------------------------------------------TACTTGAAAACTATGACGCAATAATAATTAC
DQ229832_F1R92                (3204) ----------------------------------------------------TACTTGAAAACTATGACGCAATAATAATTAC
IAST-1_haplotype_4            (3204) ----------------------------------------------------TACTTGAAAACTATGACGCAATAATAATTAC
nmsT2100_haplotype_4          (3204) ----------------------------------------------------TACTTGAAAACTATGACGCAATAATAATTAC
DQ229833_F1R92                (3267) AGCGTAATGATAGATATGCTACTTCATTTGTTGTAGTGTATCATACGCTACAACAAAATTGAACTAGAGTAATTATTAT
DQ229834_F1R92                (3267) AGCGTAATGATAGATATGCTACTTCATTTGTTGTAGTGTATCATACGCTACAACAAAATTGAACTAGAGTAATTATTAT
IAST-1_haplotype_5            (3268) AGCGTAATGATAGATATGCTACTTCATTTGTTGTAGTGTATCATACGCTACAACAAAATTGAACTAGAGTAATTATTAT
nmsT2100_haplotype_5          (3269) AGCGTAATGATAGATATGCTACTTCATTTGTTGTAGTGTATCATACGCTACAACAAAATTGAACTAGAGTAATTATTAT
Consensus                     (3319)                                                  TACTTGAAAACTATGACGCAATAATAATTAC
```

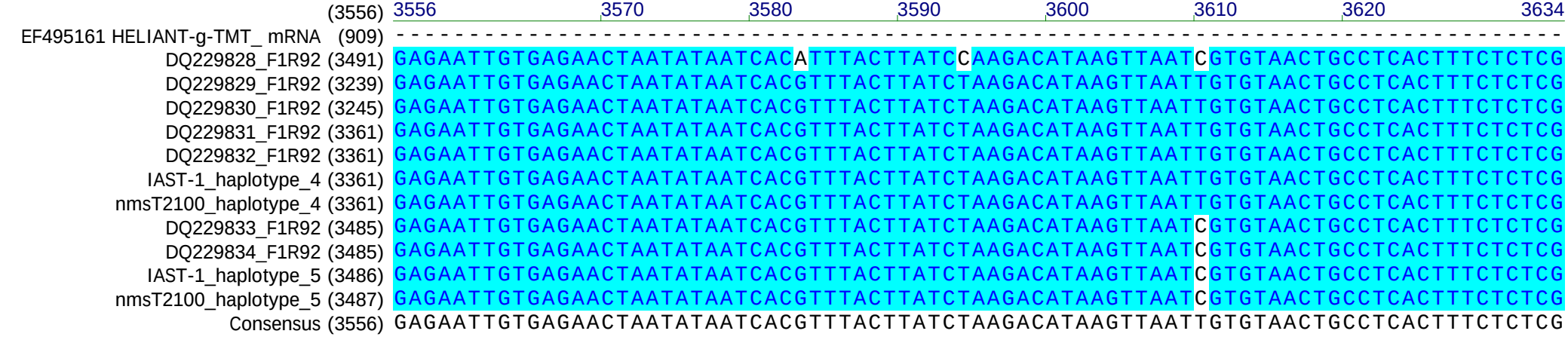
Section 44

```
                                (3398) 3398        3410        3420        3430        3440        3450        3460        3476
EF495161 HELIANT-g-TMT_ mRNA   (909) ----------------------------------------------------------------------------------------
DQ229828_F1R92                (3336) TACGCTACTTGAAAACTATGACGCAATAATAATTACTCTAGTTCAATTTTTAGCAGCGTAATTTCTTAGCACGCTAAAG
DQ229829_F1R92                (3111) T-----------------------------CTAGTAGTCTAGTTCAATTTTTAGCAGCGTGATTTCTCAGCACGCTAAAG
DQ229830_F1R92                (3117) T-----------------------------CTAGTAGTCTAGTTCAATTTTTAGCAGCGTGATTTCTCAGCACGCTAAAG
DQ229831_F1R92                (3235) T-----------------------------CTAGTAGTCTAGTTCAATTTTTAGCAGCGTAATTTCTCAGCACGCTAAAG
DQ229832_F1R92                (3235) T-----------------------------CTAGTAGTCTAGTTCAATTTTTAGCAGCGTAATTTCTCAGCACGCTAAAG
IAST-1_haplotype_4            (3235) T-----------------------------CTAGTAGTCTAGTTCAATTTTTAGCAGCGTAATTTCTCAGCACGCTAAAG
nmsT2100_haplotype_4          (3235) T-----------------------------CTAGTAGTCTAGTTCAATTTTTAGCAGCGTAATTTCTCAGCACGCTAAAG
DQ229833_F1R92                (3346) TACGCTACTTGAAAACTATGACGCAATAATAATTACTCTAGCTCAATTTTTAGCAGCGTAATTTCTTAGCACGCTAAAG
DQ229834_F1R92                (3346) TACGCTACTTGAAAACTATGACGCAATAATAATTACTCTAGCTCAATTTTTAGCAGCGTAATTTCTTAGCACGCTAAAG
IAST-1_haplotype_5            (3347) TACGCTACTTGAAAACTATGACGCAATAATAATTACTCTAGCTCAATTTTTAGCAGCGTAATTTCTTAGCACGCTAAAG
nmsT2100_haplotype_5          (3348) TACGCTACTTGAAAACTATGACGCAATAATAATTACTCTAGCTCAATTTTTAGCAGCGTAATTTCTTAGCACGCTAAAG
Consensus                     (3398) T                             CTAGTAGTCTAGTTCAATTTTTAGCAGCGTAATTTCTCAGCACGCTAAAG
```

Section 45



Section 46



Section 47

```
                                (3635) 3635  3640      3650      3660      3670      3680      3690      3700      3713
EF495161 HELIANT-g-TMT_ mRNA   (909) ------------------------------------------------------------------------------
DQ229828_F1R92 (3570) TCCACTAAATTTGACATCCACAACATACAATTATTTAAGTATAAATTGAGGTTCCCGACTTAATAATTGTTATGTATTA
DQ229829_F1R92 (3318) TCCACTAAATTTGACATCCACAACATACAATTGTTT----------TAAG--------------------TATATTTTG
DQ229830_F1R92 (3324) TCCACTAAATTTGACATCCACAACATACAATTGTTT----------TAAG--------------------TATATTTTG
DQ229831_F1R92 (3440) TCCACTAAATTTGACATCCACAACATACAATTGTTT----------TAAG--------------------TATATTTTG
DQ229832_F1R92 (3440) TCCACTAAATTTGACATCCACAACATACAATTGTTT----------TAAG--------------------TATATTTTG
IAST-1_haplotype_4 (3440) TCCACTAAATTTGACATCCACAACATACAATTGTTT----------TAAG--------------------TATATTTTG
nmsT2100_haplotype_4 (3440) TCCACTAAATTTGACATCCACAACATACAATTGTTT----------TAAG--------------------TATATTTTG
DQ229833_F1R92 (3564) TCCACTAAATTTGACATCCACAACATACAATTATTTAAGTATAAATTGAGGTTCCCGACTTAATAATTGTTATGTATTA
DQ229834_F1R92 (3564) TCCACTAAATTTGACATCCACAACATACAATTATTTAAGTATAAATTGAGGTTCCCGACTTAATAATTGTTATGTATTA
IAST-1_haplotype_5 (3565) TCCACTAAATTTGACATCCACAACATACAATTATTTAAGTATAAATTGAGGTTCCCGACTTAATAATTGTTATGTATTA
nmsT2100_haplotype_5 (3566) TCCACTAAATTTGACATCCACAACATACAATTATTTAAGTATAAATTGAGGTTCCCGACTTAATAATTGTTATGTATTA
Consensus (3635) TCCACTAAATTTGACATCCACAACATACAATTGTTT           TAAG                      TATATTTTG
```

Section 48

```
                                (3714) 3714  3720      3730      3740      3750      3760      3770      3780      3792
EF495161 HELIANT-g-TMT_ mRNA   (909) ------------------------------------------------------------------------------
DQ229828_F1R92 (3649) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAACAAT
DQ229829_F1R92 (3367) AACGTGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
DQ229830_F1R92 (3373) AACGTGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
DQ229831_F1R92 (3489) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTGAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
DQ229832_F1R92 (3489) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTGAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
IAST-1_haplotype_4 (3489) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTGAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
nmsT2100_haplotype_4 (3489) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTGAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
DQ229833_F1R92 (3643) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCACGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
DQ229834_F1R92 (3643) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCACGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
IAST-1_haplotype_5 (3644) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCACGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
nmsT2100_haplotype_5 (3645) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCACGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
Consensus (3714) AACGCGTCCATATACATGTAAATAGAGAGATAGGGCGCGCTCAGTATACTATCAATTTTAAGGAGAAAACTACAAGAAT
```

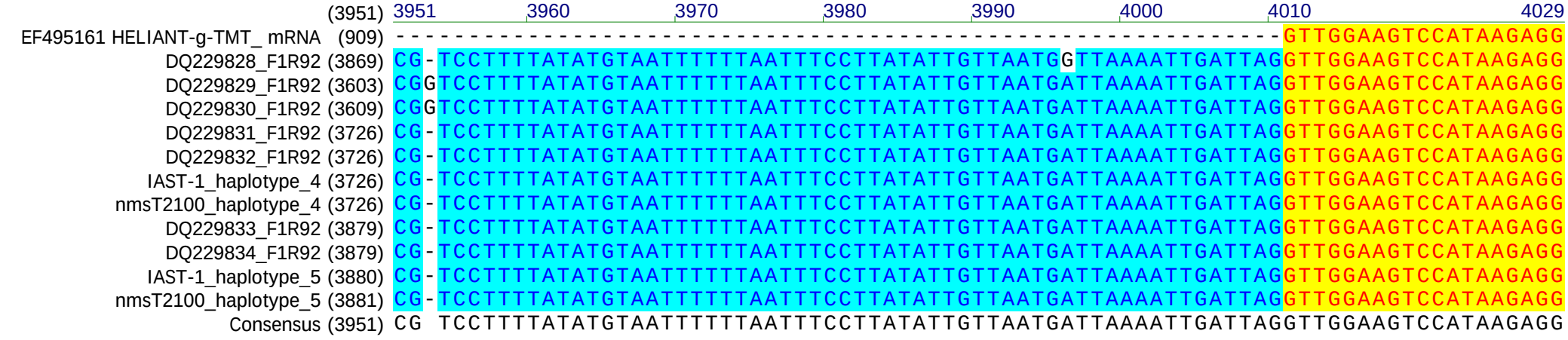
Section 49

```
                                   (3793) 3793      3800      3810      3820      3830      3840      3850      3860      3871
EF495161 HELIANT-g-TMT_ mRNA       (909) ------------------------------------------------------------------------------------
DQ229828_F1R92                     (3728) CATATAAACATTTTTACGTTGTGGTTATGTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCAGTTAGATAGC
DQ229829_F1R92                     (3446) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCGGTTAGATAGC
DQ229830_F1R92                     (3452) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCGGTTAGATAGC
DQ229831_F1R92                     (3568) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTTTATGCGGTTAGATAGC
DQ229832_F1R92                     (3568) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTTTATGCGGTTAGATAGC
IAST-1_haplotype_4                 (3568) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTTTATGCGGTTAGATAGC
nmsT2100_haplotype_4               (3568) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTTTATGCGGTTAGATAGC
DQ229833_F1R92                     (3722) CATATAAACATTTTTACGTTGTGGTTATGTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCGGTTAGATAGC
DQ229834_F1R92                     (3722) CATATAAACATTTTTACGTTGTGGTTATGTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCGGTTAGATAGC
IAST-1_haplotype_5                 (3723) CATATAAACATTTTTACGTTGTGGTTATGTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCGGTTAGATAGC
nmsT2100_haplotype_5               (3724) CATATAAACATTTTTACGTTGTGGTTATGTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT-ATGCGGTTAGATAGC
Consensus                          (3793) CATATAAACATTTTTACGTTGTGGTTATCTTATAAAGTGGAAGAGAGAGATAGTGGAGCGGTT ATGCGGTTAGATAGC
```

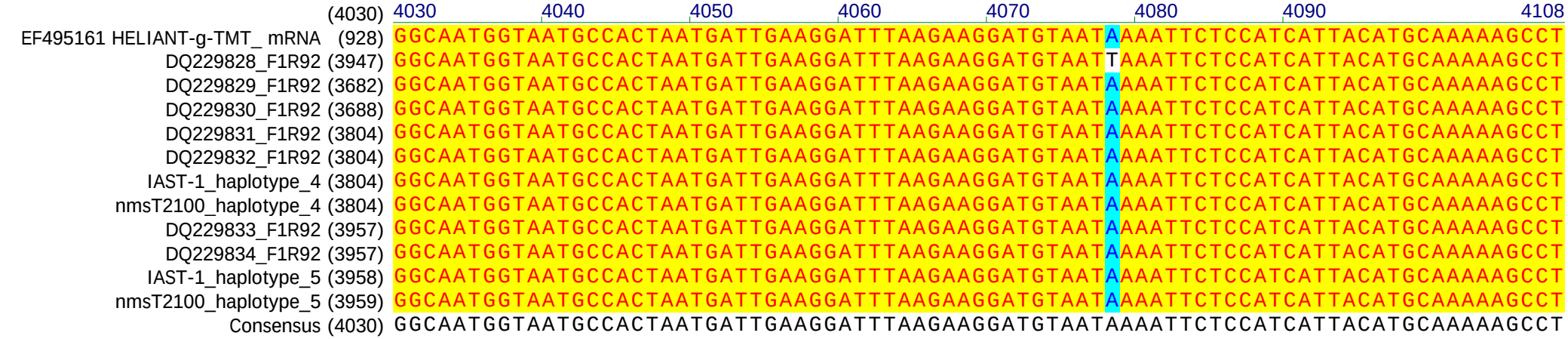
Section 50

```
                                   (3872) 3872      3880      3890      3900      3910      3920      3930      3940      3950
EF495161 HELIANT-g-TMT_ mRNA       (909) ------------------------------------------------------------------------------------
DQ229828_F1R92                     (3806) CGTC----------------AAGCAAAATGTACCTGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
DQ229829_F1R92                     (3524) CGTCTTTTGCTTTGAGTTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
DQ229830_F1R92                     (3530) CGTCTTTTGCTTTGAGTTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
DQ229831_F1R92                     (3647) CGTCTTTTGCATTGAATTTCAGGCAAAATGTACATGTGTTTACGCGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
DQ229832_F1R92                     (3647) CGTCTTTTGCATTGAATTTCAGGCAAAATGTACATGTGTTTACGCGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
IAST-1_haplotype_4                 (3647) CGTCTTTTGCATTGAATTTCAGGCAAAATGTACATGTGTTTACGCGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
nmsT2100_haplotype_4               (3647) CGTCTTTTGCATTGAATTTCAGGCAAAATGTACATGTGTTTACGCGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
DQ229833_F1R92                     (3800) CGTCTTTTGCATTGAATTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGGTTTGAGCA
DQ229834_F1R92                     (3800) CGTCTTTTGCATTGAATTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGGTTTGAGCA
IAST-1_haplotype_5                 (3801) CGTCTTTTGCATTGAATTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGGTTTGAGCA
nmsT2100_haplotype_5               (3802) CGTCTTTTGCATTGAATTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGGTTTGAGCA
Consensus                          (3872) CGTCTTTTGCATTGAATTTCAAGCAAAATGTACATGTGTTTACGTGGTTCTCACAGTTTTCTAGATTCTGTTTTGAGCA
```

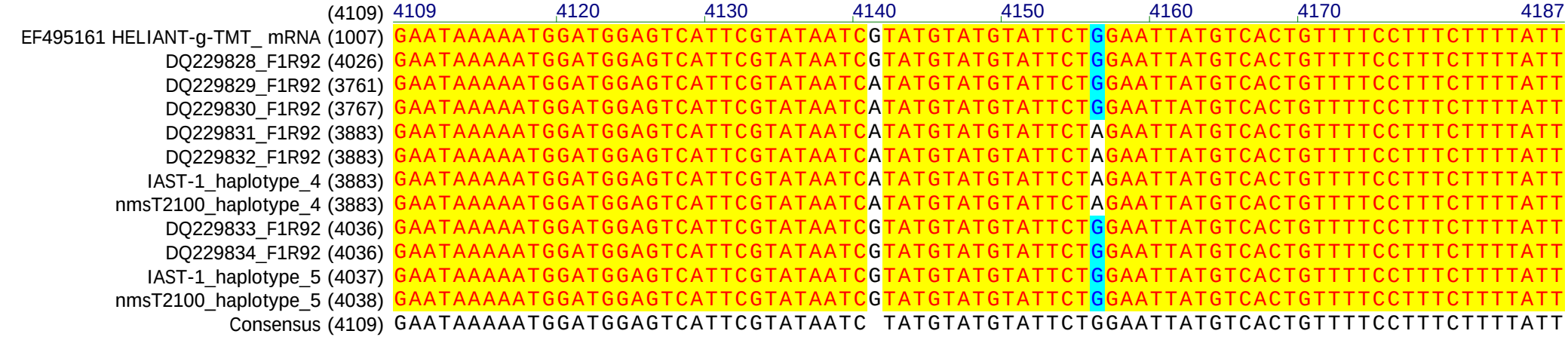
Section 51



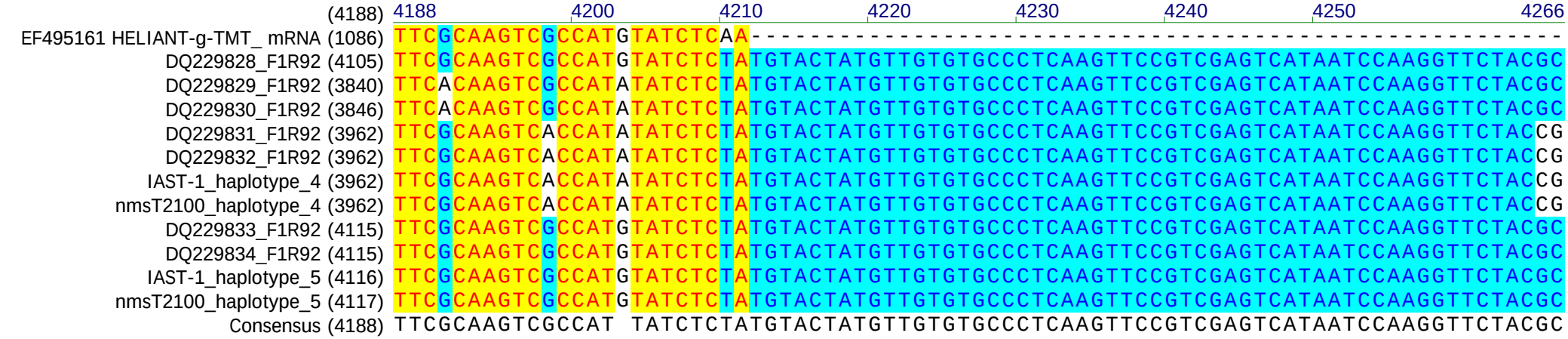
Section 52



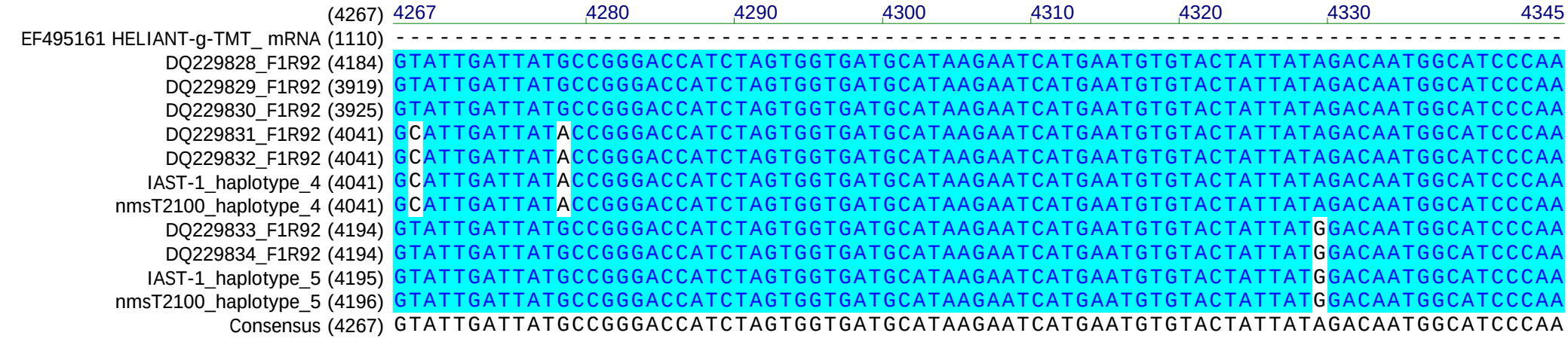
Section 53



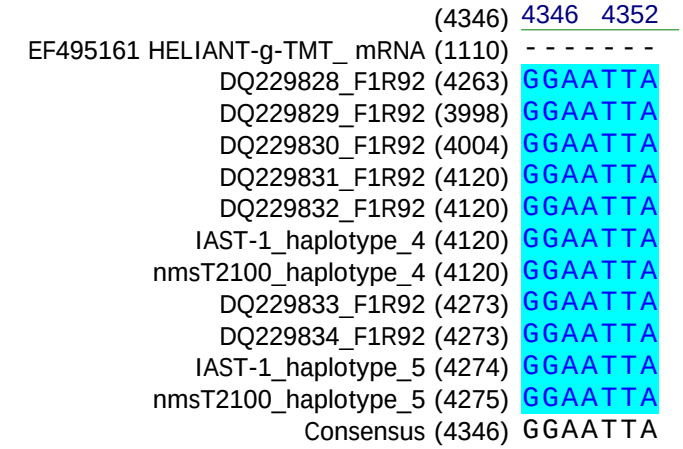
Section 54



Section 55



Section 56



CAPÍTULO 4

Transferencia de marcadores microsatélites basados en regiones no génicas y génicas de girasol a cártamo

María J. García-Moreno, Leonardo Velasco y Begoña Pérez-Vich

Transferability of non-genic microsatellite and gene-based sunflower markers to safflower

Euphytica 175:145-150 (2010)

Euphytica (2010) 175:145–150
DOI 10.1007/s10681-010-0139-6

Transferability of non-genic microsatellite and gene-based sunflower markers to safflower

M. J. García-Moreno · L. Velasco · B. Pérez-Vich

Received: 26 July 2009 / Accepted: 22 January 2010 / Published online: 6 February 2010
© Springer Science+Business Media B.V. 2010

Abstract Safflower (*Carthamus tinctorius* L.) DNA marker resources are currently very limited. The objective of this study was to determine the feasibility of transferring non-genic microsatellite (SSR) markers and gene-based markers, including intron fragment length polymorphism (IFLP) and resistance gene candidates (RGC)-based markers from sunflower (*Helianthus annuus* L.) to safflower, both species belonging to the Asteraceae family. Cross-species amplification of 119 non-genic SSRs, 48 IFLPs, and 19 RGC-based sunflower markers in 22 lines and germplasm accessions of safflower was evaluated. Additionally, 69 EST-SSR markers previously reported to amplify in safflower were tested. The results showed that 17.6% of the non-genic SSR, 56.2% of the IFLP, and 73.7% of the RGC-based markers were transferable to safflower. The percentage of transferable markers showing polymorphic loci was 66.6% for non-genic SSR, 70.6% for EST-SSR, 55.5% for IFLP, and 71.4% for RGC-based markers. The highest polymorphism levels were found for non-genic SSR. The average number of alleles per polymorphic locus and mean heterozygosity values were 2.9 and 0.46, respectively, for non-genic SSR, 2.2 and 0.35 for EST-SSR, 2.1 and 0.24 for IFLP, and 2.0 and 0.34 for RGC-based markers. The results of this study revealed a low rate of transferability for non-genic SSR sunflower markers and a better rate of transferability for IFLP and RGC-based markers. Transferable genic and non-genic sunflower markers can have utility for trait and comparative mapping studies in safflower.

Keywords Cross-species marker transferability · *Helianthus annuus* L. · *Carthamus tinctorius* L. · Microsatellites · Gene-based markers · Molecular markers

Introduction

Safflower is a minor crop currently regarded as a promising alternative for oilseed production in many areas of the world. Its adaptation to a wide range of environments, the development of hybrids with larger seed and oil yields, and the production of oil types highly demanded by industry are important factors that are contributing to this promising position (Mündel and Bergman 2008).

Nowadays, breeding programs on the major oilseed crops are based to a large extent on the development and use of molecular tools. However, DNA marker resources available to safflower breeders are very limited. The main molecular markers currently being used in safflower are RAPDs (random amplified

Electronic supplementary material The online version of this article (doi:10.1007/s10681-010-0139-6) contains supplementary material, which is available to authorized users.

M. J. García-Moreno · L. Velasco · B. Pérez-Vich (✉)
Institute for Sustainable Agriculture (CSIC), Alameda del Obispo s/n, 14004 Cordoba, Spain
e-mail: bperez@ias.csic.es

polymorphic DNA), ISSRs (intersimple sequence repeats), and AFLPs (amplified fragment length polymorphism). These are the markers of choice for crops with inadequate genomic resources, since they do not require previous sequence information, and they have been used mainly for assessing genetic diversity in this species (Sujatha 2008). The development of more robust markers, such as microsatellites and gene-based markers has not been faced at a large scale in safflower, mainly due to the absence of large public or private research programmes on this crop. Alternatively, several studies have shown the feasibility of transferring these types of molecular markers between phylogenetically related plant species, for example in cereals (Kuleung et al. 2004), legumes (Peakall et al. 1998), and Cruciferae (Plieske and Struss 2001).

Vast genomic resources are currently available for several Asteraceae species, including sunflower (The Compositae Genome Project, http://compgenomics.ucdavis.edu/), the closest oilseed crop to safflower. Sunflower genomic resources might be of great value for safflower breeding if they prove to be transferable. Within this general goal, the objective of the present study was to determine the feasibility of transferring non-genic microsatellite (SSR) markers and gene-based markers, including intron fragment length polymorphism (IFLP) markers and resistance gene candidates (RGC)-based markers, from sunflower to safflower. Marker informativeness for the transferable non-genic SSR, IFLP, and RCG-based markers and an additional set of genic SSR sunflower markers developed from expressed sequence tags (ESTs) was also evaluated.

Materials and methods

To evaluate the transferability of sunflower SSR and gene-based markers, a total of 22 safflower lines and germplasm accessions were used. These included the cultivars and breeding lines CL1, IASC-1, CR-6, CR-9, CR-142, CR-50, CR-58, CR-69, CR-34, CR-81, and Rancho (Velasco and Fernández-Martínez 2000, 2004; Velasco et al. 2005), and the germplasm accessions PI-259994, PI-401584, PI-537598, PI-537607, PI-537637, PI-537643, PI-534657, PI-537695, PI-537707, PI-560166, and PI-572471. Germplasm accessions were provided by the Western Regional Plant Introduction Station of the US Department of Agriculture. The sunflower lines P-21, P-96, and R-96 (Jan 1992; Fernández-Martínez et al. 2004) were used as a positive control. For DNA extraction, two fully expanded leaves were cut from five plants of each line and frozen at −80°C. The leaf tissue was lyophilized and ground to a fine powder in a laboratory mill. DNA was isolated from ground leaf tissue from two individual plants per line using a modified version of the protocol described by Rogers and Bendich (1985).

The following sets of molecular markers developed in sunflower were used to evaluate marker transferability to safflower: (a) A set of 119 non-genic SSR sunflower markers selected from its public map (Tang et al. 2002; ORS markers), (b) a set of 48 IFLP sunflower markers based on previously developed cDNA-RFLP probes (Yu et al. 2003; ZVG markers), and (c) a set of 19 RGC markers, based on recognition-dependent disease resistance genes encoding nucleotide binding site (NBS) leucine-rich repeat (LRR) proteins (Radwan et al. 2008; RGC markers). Additionally, a set of 69 sunflower EST-SSR (genic SSR) markers previously reported to amplify alleles in one safflower accession (Heesacker et al. 2008) was tested. Details of the SSR, IFLP, and RGC marker sets used are provided in supplementary Table 1.

All markers were initially screened in a smaller set of samples (six safflower DNA samples from three lines, and two sunflower DNA samples from two lines) using PCR conditions described for the source species (sunflower). PCR reactions were performed in a final volume of 30 µl containing 1× PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Invitrogen, San Diego, CA, USA), 0.3 µM of primers, 0.7 U of *Taq* DNA polymerase (BioTaqTM DNA Polymerase, Bioline, London, UK), and 50 ng of template DNA. Amplifications were run on a GeneAmp PCR System 9700 (Applied Bosystems, Foster City, CA, USA). For SSR markers, amplification conditions were: initial denaturation at 94°C for 2 min, followed by 1 cycle of 94°C for 30 s, annealing temperature (T_a) recommended for the source-species SSR amplification +10°C for 30 s, and 72°C for 30 s; during each of the nine following cycles the annealing temperature was decreased 1°C per cycle, at which time the products were amplified for 32 cycles at 94°C for 30 s, T_a for 30 s, and 72°C for 30 s with a final extension of 20 min at 72°C. For IFLP and RGC markers the amplification conditions were: initial denaturation at 94°C for 3 min, followed by 34 cycles of 94°C for 30 s, T_a recommended for the source-species IFLP and

RGC amplification for 45 s, and 72°C for 1 min; with a final extension of 7 min at 72°C. The amplification products were resolved on 3% Metaphor (BMA, Rockland, ME) and/or on 1.5% agarose gels in 1× TBE buffer with ethidium bromide incorporated in the gel. For those markers that showed very weak amplification products, PCR conditions were optimized by adjusting reaction mixtures to varying concentrations of $MgCl_2$ (1.5, 2, 2.5, and 3 mM), primers (0.4, 0.8, and 1.2 μM), *Taq* DNA polymerase (1 unit), and DNA (100 ng), and by lowering the annealing temperature between 3 and 5°C and/or using a non-touchdown PCR program. Once PCR mixture and amplification conditions were optimized, those sunflower markers showing amplification in safflower were tested in a larger set of samples, which included the above-mentioned 22 safflower and two sunflower lines.

To evaluate the transferability of the SSR, IFLP and RGC-based markers, the amplification products observed in safflower were classified into four classes based on the band intensity and ease of scoring: strong signal and easy score (+++); moderate signal but able to score (++); weak signal and difficult to score (+); and no signal (−) (Table 1). Markers classified as +++ and ++ were considered as transferable.

For each transferable marker, the specificity of the amplification was determined by comparing the band size of the safflower products to that of the sunflower products. The amplified fragments were classified as specific when they produced intense amplification products with a similar size (within 100 bp) to that of sunflower. Additionally, the total number of loci, the number of polymorphic loci, and the number of specific polymorphic loci were recorded for each transferable marker. Two measures of marker informative values were calculated: (i) observed number of alleles per locus, and (ii) heterozygosity. Heterozygosity (H) at each microsatellite locus was estimated according to the formula $H = 1 - \Sigma p_i^2$, where p_i is the frequency of the ith allele (Nei 1978).

Results

Twenty-one (17.6%) of the 119 non-genic sunflower SSR markers produced consistent cross-amplification in safflower (classes +++ and ++; Table 1) and accordingly they were considered as transferable. Higher percentages of marker transferability were recorded for IFLP and RGC-based markers. For IFLP markers, 27 out of 48 markers (56.2%) were transferable, whereas 14 out of 19 RGC-based markers (73.7%) were transferable. Within transferable markers, 19% of the non-genic SSR, 63% of the IFLP, and 50% of the RCG-based markers displayed strong amplification products (class +++, Table 1). Transferability details for every marker are indicated in supplementary Table 1. From the 69 genic SSR sunflower markers reported to amplify alleles in one safflower germplasm accession (Heesacker et al. 2008), only 34 produced consistent cross-amplifications showing an acceptable quality (classes +++ and ++) in the set of 22 safflower lines and germplasm accessions used in this study (Supplementary Table 1).

In most cases, transferable markers showed specific loci, i.e., their amplification products in safflower were of similar size (within ±100 bp) to those in sunflower. The percentage of transferable markers showing specific loci was 100% for non-genic SSR, 79.4% for genic SSR, 85.2% for IFLP, and 85.7% for RGC-based markers (Table 2). Multiple products were observed in the four marker types, mainly in IFLP

Table 1 Quality and number of markers transferable from sunflower to safflower

Type of sunflower marker	No. of markers within safflower amplification quality classes[a]				No. and percentage (%) of markers transferable from sunflower to safflower[b]
	+++	++	+	−	
Non-genic SSR	4	17	14	84	21 (17.6%)
IFLP	17	10	8	13	27 (56.2%)
RGC	7	7	2	3	14 (73.7%)

[a] (+++) Strong signal and easy score; (++) moderate signal but able to score; (+) weak signal and difficult to score; and (−) no signal

[b] Markers classified as +++ and ++ based on the quality of the amplification were considered as transferable

markers (3.5 loci per transferable marker) but also in RGC (3.1 loci per marker) and SSR (2.1 loci per non-genic SSR marker and 2.4 per genic SSR marker) (Table 2).

More than half of the transferable markers showed at least one polymorphic locus (Table 2). However, the number of alleles per polymorphic locus was in general low in the four marker types, averaging 2.92 for non-genic SSR, 2.25 for genic SSR, 2.06 for IFLP, and 2.00 for RGC markers (Table 2). Heterozygosities for the non-genic SSR, genic SSR, IFLP, and RGC markers averaged 0.46, 0.35, 0.26, and 0.34, respectively (Table 2). Polymorphism details for every marker are indicated in supplementary Table 1.

Discussion

Cross-amplification of molecular markers across plant species typically increase as phylogenetic distances decrease (Peakall et al. 1998; Rossetto 2001). For example, the range of cross-species amplification of soybean (*Glycine max*) SSRs among genera within the Fabaceae family has been reported to be between 3 and 13%, compared to 65% within species of the *Glycine* genus (Peakall et al. 1998). In this study, we obtained a transferability value for non-genic SSR markers of 17.2%, which is similar to those previously reported for cross-species amplification among different genera in other plant families (Peakall et al. 1998; Kuleung et al. 2004). The low transferability value suggests that the regions flanking the repeats were not highly conserved. An overall lack of conservation of SSR primer sequences in the Asteraceae family has been reported by Whitton et al. (1997) in the analysis of a set of SSR loci in 26 species representing eight divergent tribes of this family.

Transferability of sunflower markers to safflower was considerably improved by using markers based on genes (IFLP and RGC-based markers). Higher levels of transferability of genic markers as compared to non-genic markers reflect the conserved nature of coding sequences compared to non-coding genomic DNA (Varshney et al. 2005). Similar results have been obtained in other studies in different plant species in which transferability of non-genic SSR and EST-derived SSR has been evaluated. Thus, Gutierrez et al. (2005) found that transferability of EST-derived SSR was twofold higher than that of non-genic SSR in different genus within the Fabaceae family. Pashley et al. (2006) reported that EST-derived SSR were three times more transferable across species of the genus *Helianthus* than anonymous SSR. Conversely, Heesacker et al. (2008) found no differences between genic and non-genic sunflower SSR markers when amplified across species of the genus *Helianthus*.

Studies on the transferability of molecular markers to safflower have been scarce. Heesacker et al. (2008) evaluated the transferability of sunflower genic markers (EST-derived SSR and EST-INDEL markers). They found that only 14.8% of the markers amplified alleles in one safflower accession. This percentage is similar to that obtained in the present research for non-genic SSR (17.2%) and much lower

Table 2 Characteristics of markers transferable from sunflower to safflower (classified in classes +++ and ++ based on their quality) tested in a set of 22 safflower entries

Type of sunflower marker	Total number of transferable markers	Total number of loci per transferable marker		Number of transferable markers showing specific loci[a]	Number of transferable markers showing polymorphisms		Number of alleles per polymorphic locus		Heterozygosity
		Mean ± SD	Range		Total	In specific loci	Mean ± SD	Range	Mean ± SD
Non-genic SSR	21	2.1 ± 1.2	1–5	21 (100%)	14 (66.7%)	13 (61.9%)	2.92 ± 0.9	2–4	0.46 ± 0.23
Genic SSR	34	2.4 ± 1.6	1–7	27 (79.4%)	24 (70.6%)	17 (50.0%)	2.25 ± 0.7	2–5	0.35 ± 0.16
IFLP	27	3.5 ± 2.3	1–10	23 (85.2%)	18 (66.7%)	13 (48.1%)	2.06 ± 0.2	2–3	0.26 ± 0.13
RGC	14	3.1 ± 1.5	1–6	12 (85.7%)	10 (71.4%)	4 (40.0%)	2.00 ± 0.0	2	0.34 ± 0.11

[a] Specific loci: those within ±100 bp of the sunflower product

to the transferability of genic markers (56.2% for IFLP and 73.7% for RGC-based markers). Chapman et al. (2007) developed universal markers based on EST sequences conserved across sunflower, lettuce and Arabidopsis, which showed a high degree of successful amplification and polymorphism in safflower. These markers have been successfully used for phylogenetic studies in *Carthamus* (Chapman and Burke 2007).

Although there are no previous reports about the degree of polymorphism in safflower of non-genic SSR markers or IFLP and RGC-based markers, other marker systems such as RAPDs, AFLPs or ISSRs have shown higher levels of polymorphism that those obtained in the present study for non-genic SSR, IFLP and RGC-based transferable markers (Johnson et al. 2007; Yang et al. 2007; Amini et al. 2008). For the genic-SSR marker system, Chapman et al. (2009) reported also higher levels of polymorphism than those found in the present study. Several studies have shown that the transference of markers from one species to another produces a reduction of polymorphism and allelic diversity, which is more pronounced when the transference occurs between phylogenetically distant species (Peakall et al. 1998; Decroocq et al. 2003).

The results of the present research revealed a limited utility of transferring non-genic SSR markers from sunflower to safflower, which agrees with the results of Heesacker et al. (2008) for EST-SSR and EST-INDEL markers. Conversely, other sunflower genic markers such as IFLP and RGC-based markers demonstrated to be more transferable and they constitute a valuable source for increasing marker resources available for safflower molecular studies.

Acknowledgments The authors would like to thank Dr. Alberto J. Leon (Advanta Semillas, Balcarce, Buenos Aires, Argentina) and Dr. Steven J. Knapp (The University of Georgia, Athens, USA) for kindly providing IFLP and RGC markers, respectively. The research was funded by the Spanish Ministry of Science and Innovation, research project AGL-2007-62834.

References

Amini F, Saeidi G, Arzani A (2008) Study of genetic diversity in safflower genotypes using agro-morphological traits and RAPD markers. Euphytica 163:21–30

Chapman MA, Burke JM (2007) DNA sequence diversity and the origin of cultivated safflower (*Carthamus tinctorius* L.; Asteraceae). BMC Plant Biol 7:60. doi:10.1186/1471-2229-7-60

Chapman MA, Chang JC, Weisman D, Kesseli RV, Burke JM (2007) Universal markers for comparative mapping and phylogenetic analysis in the Asteraceae (Compositae). Theor Appl Genet 115:747–755

Chapman MA, Hvala J, Strever J, Matvienko M, Kozik A, Michelmore RW, Tang S, Knapp SJ, Burke JM (2009) Development, polymorphism, and cross-taxon utility of EST–SSR markers from safflower (*Carthamus tinctorius* L.). Theor Appl Genet 120:85–91

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912–922

Fernández-Martínez JM, Pérez-Vich B, Akhtouch B, Velasco L, Muñoz-Ruz J, Melero-Vara JM, Domínguez J (2004) Registration of four sunflower germplasm lines resistant to race F of broomrape. Crop Sci 44:1033–1034

Gutierrez MV, Vaz Patto MC, Huguet T, Cubero JI, Moreno MT, Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theor Appl Genet 110:1210–1217

Heesacker A, Kishore VK, Gao W, Tang S, Kolkman JM, Gingle A, Matvienko M, Kozik A, Michelmore RM, Lai Z, Rieseberg LH, Knapp SJ (2008) SSRs and INDELs mined from the sunflower EST database: Abundance, polymorphisms, and cross-taxa utility. Theor Appl Genet 117:1021–1029

Jan CC (1992) Inheritance and allelism of mitomycin C- and streptomycin-induced recessive genes for male sterility in cultivated sunflower. Crop Sci 32:317–320

Johnson RC, Kisha TJ, Evans MA (2007) Characterizing safflower germplasm with AFLP molecular markers. Crop Sci 47:1728–1736

Kuleung C, Baenziger PS, Dweikat I (2004) Transferability of SSR markers among wheat, rye, and triticale. Theor Appl Genet 108:1147–1150

Mündel HH, Bergman JW (2008) Safflower breeding. In: Vollmann J, Rajcan I (eds) Oil crop breeding. Springer Series Handbook of Plant Breeding, Springer, New York/Berlin. Reprinted for the 7th international safflower conference with permission from the Special Licensing Department, Springer, Dordrecht, The Netherlands

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Pashley CH, Ellis JR, McCauley DE, Burke JM (2006) EST databases as a source for molecular markers: lessons from *Helianthus*. J Hered 97:381–388

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Mol Biol Evol 15:1275–1287

Plieske J, Struss D (2001) Microsatellite markers for genome analysis in *Brassica*. I. Development in *Brassica napus* and abundance in Brassicaceae species. Theor Appl Genet 102:689–694

Radwan O, Gandhi S, Heesacker A, Whitaker B, Taylor C, Plocik A, Kesseli R, Kozik A, Michelmore RW, Knapp SJ (2008) Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. Mol Genet Genomics 280:111–125

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76

Rossetto M (2001) Sourcing of SSR markers from related plant species. In: Henry RJ (ed) Plant genotyping: the DNA fingerprinting of plants. CABI, Wallingford, pp 211–224

Sujatha M (2008) Biotechnological interventions for genetic improvement of safflower. In: Knights SE, Potter TD (eds) Proceedings of the 7th international safflower conference, 3–6 November 2008, Wagga Wagga, New South Wales, Australia

Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor Appl Genet 105:1124–1136

Varshney RK, Graner A, Sorrells M (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48–55

Velasco L, Fernández-Martínez JM (2000) Isolation of lines with contrasting seed oil fatty acid profiles from safflower germplasm. Sesame Safflower Newsl 15:104–108

Velasco L, Fernández-Martínez JM (2004) Registration of CR-34 and CR-81 safflower germplasms with increased tocopherol. Crop Sci 44:2278

Velasco L, Pérez-Vich B, Hamdan Y, Fernández-Martínez JM (2005) Genetic study of several seed oil quality traits in safflower. In: Proceedings of the 6th international safflower conference, Istanbul, Turkey, 6–10 June 2005, pp 74–79

Whitton J, Rieseberg LH, Ungerer MC (1997) Microsatellite loci are not conserved across the Asteraceae. Mol Biol Evol 14:204–209

Yang YX, Wu W, Zheng YL, Chen L, Liu RJ, Huang CY (2007) Genetic diversity and relationships among safflower (*Carthamus tinctorius* L.) analyzed by inter-simple sequence repeats (ISSRs). Genet Resour Crop Evol 54: 1043–1051

Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring M, Soper J, Han F, Chu WC, Webb DM, Thompson L, Edwards KJ, Berry ST, Leon A, Olungu C, Maes N, Knapp SJ (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci 43:367–387

Characteristics of markers transferred from sunflower to safflower

Sunflower marker	Quality of the amplificación in safflower (+++, ++, +, -)	No. of loci	No. of specific loci (within ± 100 bp of the sunflower product)	No. of polymorphic loci	Polymorphism in specific loci	No. of alleles per locus
Non-genic SSR						
ORS 7	++	4	1	2	Yes	2 alleles/locus
ORS 16	++	2	1	1	Yes	2 alleles/locus
ORS 70	-					
ORS 90	++	4	1	1	No	2 alleles/locus
ORS 154	-					
ORS 166	-					
ORS 185	+++	5	1	2	Yes	2 alleles/locus
ORS 202	+++	2	1	0		
ORS 229	+					
ORS 230	-					
ORS 243	+					
ORS 297	-					
ORS 299	-					
ORS 307	-					
ORS 309	-					
ORS 311	++	2	1	0		
ORS 312	-					
ORS 316	-					
ORS 317	-					
ORS 328	+					
ORS 329	++	1	1	0		
ORS 331	-					
ORS 342	-					
ORS 366	-					
ORS 371	-					
ORS 381	+					
ORS 407	-					
ORS 420	++	2	1	0		
ORS 423	-					
ORS 428	-					
ORS 437	-					
ORS 442	-					
ORS 453	+					
ORS 456	-					
ORS 457	+++	4 (2 strong)	1 (weak)	0		
ORS 483	-					
ORS 502	-					
ORS 505	-					
ORS 533	-					
ORS 534	-					
ORS 536	+++	2	1	0		
ORS 543	-					
ORS 561	-					
ORS 565	-					
ORS 595	+					
ORS 599	+					
ORS 610	++	1	1	0		
ORS 613	++	1	1	0		
ORS 621	-					
ORS 630	-					
ORS 650	-					
ORS 656	-					
ORS 665	-					
ORS 666	-					
ORS 668	-					
ORS 674	++	2	1	0		
ORS 687	-					
ORS 691	-					
ORS 694	++	2	2	1	Yes	2 alleles/locus
ORS 695	-					
ORS 716	-					
ORS 733	-					
ORS 735	-					

Sunflower marker	Quality of the amplificación in safflower (+++, ++, +, -)	No. of loci	No. of specific loci (within ± 100 bp of the sunflower product)	No. of polymorphic loci	Polymorphism in specific loci	No. of alleles per locus
ORS 750	++	2	2	0		
ORS 761	-					
ORS 762	+					
ORS 774	-					
ORS 778	-					
ORS 780	+					
ORS 785	-					
ORS 810	-					
ORS 826	++	1	1	0		
ORS 830	+					
ORS 832	-					
ORS 837	-					
ORS 844	-					
ORS 852	-					
ORS 857	-					
ORS 878	-					
ORS 885	-					
ORS 887	-					
ORS 894	-					
ORS 898	-					
ORS 899	-					
ORS 925	-					
ORS 938	-					
ORS 949	-					
ORS 963	-					
ORS 966	-					
ORS 993	-					
ORS 1013	++	1	1	0		
ORS 1024	-					
ORS 1030	-					
ORS 1036	-					
ORS 1041	-					
ORS 1043	++	1	1	0		
ORS 1065	-					
ORS 1079	-					
ORS 1085	-					
ORS 1108	+					
ORS 1114	+					
ORS 1120	-					
ORS 1141	-					
ORS 1143	-					
ORS 1146	-					
ORS 1152	++	2	2	0		
ORS 1161	++	1	1	0		
ORS 1178	-					
ORS 1179	-					
ORS 1222	-					
ORS 1227	-					
ORS 1231	+					
ORS 1245	-					
ORS 1248	-					
ORS 1256	-					
ORS 1260	-					
ORS 1265	-					
CRT-35	++	1	1	0		
CRT-650	+					
IFLPs						
ZVG1	-					
ZVG2	+++	1	1	0		
ZVG3	-					
ZVG4	+++	3	3	1	Yes	2 alleles/locus
ZVG5	+++	3	2	2	Yes	2 alleles/locus
ZVG11	-					
ZVG14	-					
ZVG15	+++	3	2	2	Yes	2 alleles/locus

Sunflower marker	Quality of the amplificación in safflower (+++, ++, +, -)	No. of loci	No. of specific loci (within ± 100 bp of the sunflower product)	No. of polymorphic loci	Polymorphism in specific loci	No. of alleles per locus
ZVG23	-					
ZVG26	++	3	1	1	Yes	2 alleles/locus
ZVG28	-					
ZVG30	+++	3	2	2	No	2 alleles/locus
ZVG34	+					
ZVG35-N1	+					
ZVG35-N2	++	3	2	0		
ZVG36	+					
ZVG37	+					
ZVG40-I1	+					
ZVG40-I2	+++	1	1	1	Yes	2 alleles/locus
ZVG43	+++	3	3	0		
ZVG44	+++	1	0	0		
ZVG45	++	1		0		
ZVG48	+					
ZVG52	-					
ZVG55	++	4	3	0		
ZVG59	++	6	5	0		
ZVG62	-					
ZVG64	+++	1	0	0		
ZVG64	+++	4	1	3	Yes	2 alleles/locus
ZVG69	+					
ZVG70	++	10	1	1	No	2 alleles/locus
ZVG71	-					
ZVG73	++	3	2	1	Yes	2 alleles/locus
ZVG75	++	1	1	1	No	2 alleles/locus
ZVG80	+++	2	1	1	No	2 alleles/locus
ZVG83	+++	2	2 (secundary loci)	1	Yes	2 alleles/locus
ZVG127	-					
ZVG135	+++	6	4	4	Yes	2 alleles/locus and 3 alleles/locus
ZVG302	++	2	2	0		
ZVG433	+	1	1			
K0131	+++	9	Not clear	2	Yes	2 alleles/locus
ZVG375-I1	+++	6	3 within 100 bp	2	Yes	2 alleles/locus
ZVG375-I2	-					
IN030	++	6	2	1	Yes	2 alleles/locus
K0177	-					
ZVG541	-					
ZVG650	+++	1	1	1	Yes	2 alleles/locus
ZVG819	+++	2	2	1	No	2 alleles/locus
RGCs						
RGC1	++	2	1	0		
RGC2	+					
RGC3	-					
RGC4	-					
RGC6	++	4	4	1	No	2 alleles/locus
RGC7	+++	2	2	1	No	2 alleles/locus
RGC8	++	2	1	0		
RGC9	+					
RGC10	+++	3	1	2	No	2 alleles/locus
RGC11	+++	3	1	2	Yes	2 alleles/locus
RGC14	++	6	2	1	No	3 alleles/locus
RGC17	+++	4	2	2	No	2 alleles/locus
RGC18	+++	3	2	2	Yes	2 alleles/locus
RGC19	++	3	0	1	Yes	2 alleles/locus
RGC21	+++	1	1	1	Yes	2 alleles/locus
RGC23	-					
RGC24	++	1	1	0		
RGC30	+++	1	0	0		
RGC35	++	2	0	1	No	2 alleles/locus

CAPÍTULO 5

Etiquetado molecular y análisis de genes candidatos del carácter alto contenido en gamma-tocoferol en cártamo (*Carthamus tinctorius* L.)

María J. García-Moreno, José M. Fernández-Martínez, Leonardo Velasco y Begoña Pérez-Vich

Molecular tagging and candidate gene analysis of the high gamma-tocopherol trait in safflower (*Carthamus tinctorius* L.)

Molecular Breeding 28:367–379 (2011)

Mol Breeding (2011) 28:367–379
DOI 10.1007/s11032-010-9489-y

Molecular tagging and candidate gene analysis of the high gamma-tocopherol trait in safflower (*Carthamus tinctorius* L.)

M. J. García-Moreno · J. M. Fernández-Martínez · L. Velasco · B. Pérez-Vich

Received: 15 April 2010 / Accepted: 23 July 2010 / Published online: 7 August 2010
© Springer Science+Business Media B.V. 2010

Abstract Genetic control of the synthesis of high gamma-tocopherol (gamma-T) content in the seed oil of safflower (*Carthamus tinctorius* L.) and development of highly reliable molecular markers for this trait were determined through molecular tagging and candidate gene approaches. An F2 population was developed by crossing the high gamma-T natural mutant IASC-1 with the CL-1 line (standard, high alpha-T profile). This population segregated for the partially recessive gene *Tph2*. Bulked segregant analysis with random amplified polymorphic DNA (RAPD) and microsatellite (SSR) markers revealed linkage of eight RAPD and one SSR marker loci to the *Tph2* gene and allowed the construction of a *Tph2* linkage map. RAPD fragments closest to the *Tph2* gene were transformed into sequence-characterized amplified region markers. A gamma-T methyltransferase (gamma-TMT) locus was found to co-segregate with *Tph2*. The locus/band was isolated, cloned and sequenced and it was confirmed as a gamma-TMT gene. A longer partial genomic DNA sequence from this gene was obtained. IASC-1 and CL-1 sequence alignment showed one non-synonymous and two synonymous nucleotide mutations. Intron fragment length polymorphism and insertion-deletion markers based on the gamma-TMT sequence diagnostic for the *Tph2* mutation were developed and tested across 22 safflower accessions, cultivars, and breeding lines. The results from this study provide strong support for the role of the gamma-TMT gene in determining high gamma-T content in safflower and will assist introgression of *thp2* alleles into elite safflower lines to develop varieties with improved tocopherol composition for specific market niches.

Keywords Candidate gene · *Carthamus tinctorius* · Gamma-tocopherol · Gamma-tocopherol methyltransferase · Safflower · SCAR

M. J. García-Moreno · J. M. Fernández-Martínez · L. Velasco · B. Pérez-Vich (✉)
Instituto de Agricultura Sostenible (CSIC), Apartado 4084, 14080 Córdoba, Spain
e-mail: bperez@ias.csic.es

Introduction

Tocopherols are the main naturally occurring substances with antioxidant activity in oil seeds and derived seed oils. They exist as four forms named alpha-, beta-, gamma-, and delta-tocopherol (alpha-, beta-, gamma-, and delta-T). Their antioxidant protective action is exhibited both in biological systems (in vivo or vitamin E activity) as well as in oils and oil-based products (in vitro antioxidant activity) (Kamal-Eldin and Appelqvist 1996). Alpha-T exerts the maximum in vivo antioxidant activity (Trabber and Sies 1996), but it shows the weakest antioxidant efficiency in vitro. Conversely, the other tocopherol

forms possess lower vitamin E activity but they are more powerful in vitro antioxidants (Warner and Moser 2009).

Alpha-T is the predominant tocopherol in safflower seeds, accounting for more than 95% of total tocopherols (Johnson et al. 1999). This makes safflower oil an excellent source of dietary vitamin E, but a poor oil for high temperature applications, such as frying or lubrication, where maximum resistance to oxidation is required (Schneider 2006; Marmesat et al. 2008). Breeding for modified tocopherol profile in safflower led to the development of the natural mutant line IASC-1, in which alpha-T was partially replaced by gamma-T (85% of total tocopherols) (Velasco et al. 2005a). The high gamma-T trait in this line was controlled by partially recessive alleles at a single locus (Velasco et al. 2005a). Other high gamma-T natural mutants have been identified in sunflower (Demurin 1993; Velasco et al. 2004), in which the high gamma-T trait was also found to be controlled by a partially recessive gene named *Tph2* (Demurin et al. 1996; García-Moreno et al. 2006). Even though the locus controlling gamma-T accumulation in safflower was initially named *Tph1* (Velasco et al. 2005a), we will follow sunflower nomenclature to avoid confusion and will refer to it as *Tph2*.

Safflower molecular tools for trait mapping are very scarce. To date, the only reported studies for mapping simply inherited traits are based on bulked segregant analyses using anonymous RAPD (random amplified polymorphic DNA) and AFLP (amplified fragment length polymorphism) markers to be converted into more robust SCAR (sequence-characterized amplified region) markers (Hamdan et al. 2008; Zhang et al. 2009). When assumptions can be made regarding the biological function of the gene(s) of interest, a candidate gene approach is an alternative strategy for developing more efficient tools for molecular breeding and for characterizing and cloning those factors underlying Mendelian loci (Pflieger et al. 2001). This is the case for the high gamma-T trait in safflower. Gamma-T methyltransferase (gamma-TMT) has been shown to be the candidate gene for high gamma-T mutant phenotypes in other plants, for example in sunflower seeds (Hass et al. 2006) and *Arabidopsis* leaves (Bergmüller et al. 2003). The product from this gene catalyzes the conversion from delta- to beta- and from gamma- to alpha-T (Bergmüller et al. 2003). In oilseed crops with naturally occurring high alpha-T content in the seeds, such as sunflower, gamma-TMT mutations disrupt the synthesis of alpha-T and cause a build-up of gamma-T (Hass et al. 2006).

The objectives of the present research were: (1) to tag the safflower *Tph2* gene determining high gamma-T content with molecular markers, including RAPD markers to be converted into SCARs, and microsatellite (SSR) markers, and (2) to carry out a candidate gene approach for the characterization of the *Tph2* gene and to develop highly reliable markers based on the gene sequence.

Materials and methods

Plant material and phenotypic analyses

IASC-1 is a safflower line with high gamma-T content in the seeds (>85% of total tocopherols) selected from *Carthamus oxyacanthus* M. Bieb. germplasm accession PI 426472 that showed segregation for the trait and strong introgression of cultivated safflower traits (Velasco et al. 2005a). High gamma-T content in this line is controlled by partially recessive alleles at the *Tph2* gene (Velasco et al. 2005a). CL-1 is a nuclear male sterile (NMS) line characterized by a standard seed tocopherol profile (>95% alpha-T, <1% gamma-T, the rest being beta-T) isolated from the USDA-ARS germplasm accession PI560161, which in turn is derived from the germplasm line UC-148 (Heaton and Knowles 1980). Male sterility in this line is controlled by recessive alleles at the *Ms* locus (Heaton and Knowles 1982). Plants of both lines were crossed in 2003 and produced F1 seeds with a gamma-T content ranging from 0 to 7.8% (Velasco et al. 2005a). F1 plants were grown and self-pollinated in 2004. F2 populations from the cross CL-1 × IASC-1 and phenotypic segregation for gamma-T content have been previously described in Velasco et al. (2005a). A population of 105 F2 seeds from the cross CL-1 × IASC-1 was selected for the molecular analyses and analysed for tocopherol composition by the half-seed technique as described below. The F2 half-seeds were germinated and the corresponding F2 plants were grown in the field and self-pollinated to obtain the F3 seeds. Since about one-quarter of the F2 plants showed a NMS phenotype due to the

segregation of the *Ms* gene, their genotypic configuration for gamma-T content was determined by testcrossing with plants of IASC-1. A set of 12–48 individual F3 or testcross seeds per F2 plant was analysed for tocopherol profile to determine presence or absence of segregation for gamma-T content. F2 plants were classified as *Tph2Tph2* if they had a low gamma-T content similar to CL-1 in their respective F2 half-seed and showed uniformly low gamma-T levels in their F3 or testcross seeds (<5%), *Tph2tph2* if their F3 or testcross seeds segregated for gamma-T content, and *tph2tph2* if their F2 half-seed had a high gamma-T content similar to IASC-1 and their F3 or testcross seeds showed uniformly high gamma-T content (>85%). F2 plants not producing a minimum number of F3 or testcross seeds for tocopherol analyses were classified as *Tph2_* if gamma-T content in the F2 half-seed was <5% and *tph2tph2* if gamma-T in the F2 half-seed was >85%.

Tocopherol analyses were conducted using the procedure reported by Goffman et al. (1999) with slight modifications. Half-seeds (small seed pieces excised from the seed part distal to the embryo) were placed into a 10 ml tube. After addition of 2 ml of iso-octane, the half-seeds were crushed as finely as possible with a stainless-steel rod. The samples were stirred and extracted overnight at room temperature in darkness (extraction time about 16 h). After extraction, the samples were stirred again, centrifuged and filtered, and 25 µl of the extract were analysed by HPLC using a fluorescence detector (Waters 474, Waters Corporation, Milford, MA, USA) at 295 nm excitation and 330 nm emission and iso-octane/tert-butylmethylether (94:6) as eluent at an isocratic flow rate of 1 ml/min. Chromatographic separation of the tocopherols was performed on a LiChrospher 100 diol column (250 × 3 mm I.D.) with 5 µm spherical particles, connected to a silica guard column (LiChrospher Si 60, 5 × 4 mm I.D.). The peak areas of the individual tocopherols were corrected according to their previously calculated response factors: alpha-T = 1.0; beta-T = 1.80; gamma = 1.85; delta-T = 2.30.

DNA extraction and bulked segregant analysis

One hundred and five F2 plants were used for molecular analysis. Ten fully expanded leaves were cut from each plant and frozen at −80°C. The leaf tissue was lyophilized and ground to a fine powder in a laboratory mill. DNA was isolated from ground leaf tissue from each F2 plant using a modified version of the protocol described by Rogers and Bendich (1985). DNA was also isolated from three individuals of CL-1 and IASC-1 parental lines. For bulked segregant analysis (Michelmore et al. 1991), bulks were constructed by pooling aliquots (30 µl) of DNA from two sets of individuals with contrasting genotypes for the *Tph2* locus based on tocopherol analyses of both F2 and F3 or testcross seed generations. The low gamma-T bulk was made up from 10 F2 individuals classified as *Tph2Tph2*, and the high gamma-T bulk was made up from 10 individuals classified as *tph2tph2*.

Two replicate samples of each bulk and the parental lines were screened with a set of 122 RAPD primers (Operon Technologies, Alameda, CA, USA; complete kits G, H, L, M, S, and two to six primers from kits A, B, F, J, and AA), a set of 88 safflower simple sequence repeat (SSR) markers, and a set of 20 SCAR markers previously developed by Hamdan et al. (2008). The PCR reaction mixture (30 µl) contained 1× PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP (Invitrogen, San Diego, CA, USA), 0.6 µM of RAPD primer or 0.3 µM of each SSR primer, 0.7 U of *Taq* DNA polymerase (Biotaq™ DNA Polymerase, Bioline, London, UK), and 50 ng of template DNA. DNA amplification was performed in a GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA). The RAPD markers were amplified using an initial denaturation step at 91°C for 5 min, followed by 40 cycles of 1 min at 91°C, 1 min at 36°C, and 2 min at 72°C, and ending with an extension period of 5 min at 72°C. The SSR markers were amplified using a touchdown PCR with an initial denaturation at 94°C for 2 min, followed by 1 cycle of 94°C for 30 s, final annealing temperature (T_A) + 10°C for 30 s, and 72°C for 45 s. The annealing temperature was decreased by 1°C per cycle during each of the nine following cycles, at which time the products were amplified for 32 cycles at 94°C for 30 s, T_A for 30 s, and 72°C for 45 s with a final extension of 20 min at 72°C. Final annealing temperatures varied from 46 to 62°C. Amplification products were resolved by electrophoresis on 1.5% (RAPD markers) or 3% Metaphor® (BMA, Rockland, ME, USA) (SSR markers) agarose gels in 1× TBE buffer with ethidium bromide incorporated in the gels

and visualized under UV light. A 50-bp DNA ladder (Invitrogen, San Diego, CA, USA) was used as a standard molecular weight marker to get an approximate size of DNA fragments. SCAR marker analyses were conducted as described in Hamdan et al. (2008).

Linkage between the *Tph2* locus and the RAPD and SSR markers polymorphic between the bulks was verified by genotyping these markers on the 105 F2 individuals. Chi-square tests were performed to examine the goodness of fit between the expected Mendelian ratio and the segregation data for RAPD and SSR marker loci. The significance of each marker's association with the phenotypic trait was determined by one-way analysis of variance (ANOVA) using the statistical package SPSS Statistics v.17.0 (SPSS for Windows; SPSS Inc., Chicago, IL, USA), with marker genotypes being classes. In addition, linkage analysis was run with MAPMAKER v.3.0 (Whitehead Institute, Cambridge, MA; Lander et al. 1987) using segregation data for RAPD and SSR marker loci and for the *Tph2* locus. The genotypes for the *Tph2* gene were inferred from gamma-T phenotypes, as described above. Two-point analysis was used to group the RAPD and SSR marker loci and *Tph2*. A LOD threshold of 11 and a maximum recombination fraction of 0.3 were used as linkage criteria. Three-point and multi-point analyses were used to determine the order and interval distances between the markers. Recombination fractions were converted to centiMorgans (cM) using the Kosambi mapping function. Linkage group maps were drawn using MapChart software (Voorrips 2002).

Cloning and sequencing of RAPD fragments and SCAR development

The RAPD fragments closest to the *Tph2* locus were transformed into more consistent SCAR markers (Paran and Michelmore 1993). *Tph2* locus-linked RAPD fragments amplified from the low or high gamma-T parental lines were separated on a 1.5% agarose gel before being excised and purified by means of the QIAquick gel extraction kit (Qiagen GmbH, Hilden, Germany). The purified fragments were re-amplified with the respective primers under the same PCR conditions, with the only change being a longer final extension step of 20 min. The re-amplified product was verified for fragment size and gel purified again as described above. The purified products were ligated in a T/A vector (pCR2.1) and cloned with the TOPO-TA cloning kit (Invitrogen, San Diego, CA, USA) as described by the manufacturer. Transformed *E. coli* were plated on selective media containing ampicillin and X-gal. Ten white colonies per marker were picked from the plate and cultured overnight. The QIAprep Spin Miniprep Kit (Qiagen GmbH, Hilden, Germany) was used for plasmid DNA extraction. Restriction enzyme digestion was performed to confirm the presence of the insert. Sequencing in both forward and reverse orientations of the cloned fragments (two clones per marker) was performed at GATC Biotechnology (Konstanz, Germany) using the M13 forward and reverse sequencing primers. Sequence analysis was conducted with the aid of the software Vector NTI Advance 10.3.0 (Invitrogen, San Diego, CA, USA). The consensus sequence for the two clones sequenced for each RAPD fragment was used to design the SCAR markers. A first set of primer pairs (F1 and R1) was designed containing the original 10 bases of the RAPD primer plus the next 8–11 internal bases (Table 1). A second set of primer pairs (F2 and R2) located more internally was designed with the software Vector NTI Advance 10.3.0. SCAR markers were genotyped in the 105 F2 individuals. SCAR-PCR amplifications were done as reported in Hamdan et al. (2008). Chi-square, ANOVA, and linkage analyses for the SCAR and *Tph2* loci were performed as described above.

Candidate gene analysis

For candidate gene analysis, BLAST searches of the Compositae Genome Program Database (CGPdb: http://cgpdb.ucdavis.edu/) and National Center for Biotechnology Information (NCBI) GenBank databases (http://www.ncbi.nlm.nih.gov/) were performed using cDNA for sunflower gamma-TMT (Hass et al. 2006) and *Arabidopsis* gamma-TMT (Shintani and DellaPenna 1998; Bergmüller et al. 2003) as query templates to search for safflower homologs. Since no putative homologs were identified, INDEL (insertion–deletion polymorphisms) markers based on the gamma-TMT gene from sunflower (Hass et al. 2006) were used. Both safflower and sunflower belong to the Asteraceae family, and a good

Table 1 SCAR markers linked to the *Tph2* gene, their original RAPD fragments (primer name_fragment size), primer sequences and their melting temperature (T_m)

SCAR marker	RAPD fragment	Primer	Primer sequence (5′–3′)	T_m
IASTO-257	$OPL20_{1050}$	F1	TGGTGGACCATACACGAC	55.9
		R1	TGGTGGACCAGTGGGTGG	64.1
		F2	GGAACTTGTGCCTTGAATGA	58.7
		R2	CGACGGTGGATATGTTGAAGT	59.9
IASTO-259	$OPAA12_{475}$	F1	GGACCTCTTGTAGTTATCTTT	49.5
		R1	GGACCTCTTGCCAAGAAG	55.8
		F2	GACCTCTTGTAGTTATCTTTTGT	51.5
		R2	ACTCGTCAATTAGATAGATCC	50.2
IASTO-302	$OPH12_{1100}$	F1	ACGCGCATGTAAAACATTT	56.8
		R1	ACGCGCATGTTTAGATGGA	60.2
		F2	GGAGACCCAAAACCATCG	58.9
		R2	TGCAACGGCGGTGATGGT	67.9

The underlined nucleotides are derived from RAPD primers

transferability rate of markers based on genes has been reported between the species (García-Moreno et al. 2010). The γ-TMT_F9 and γ-TMT_R24 primers complementary to conserved DNA sequences in the first and second exons, respectively, of the sunflower gamma-TMT gene (Hass et al. 2006) were tested in two replicate samples of each gamma-T bulk and the parental lines CL-1 and IASC-1. PCR reaction mixture (30 µl) contained 1× PCR buffer, 2 mM $MgCl_2$, 0.2 mM dNTP (Invitrogen, San Diego, CA, USA), 0.8 µM of primer, 1 U of *Taq* DNA polymerase (Biotaq™ DNA Polymerase, Bioline, London, UK), and 50 ng of template DNA. γ-TMT_F9/γ-TMT_R24 INDEL-PCR amplification was carried out using an initial denaturation step at 94°C for 3 min, followed by 34 cycles of 30 s at 94°C, 45 s at 58°C, and 1 min at 72°C, and ending with an extension period of 7 min at 72°C. PCR products were electrophoresed on a 1.5% agarose gel in 1× TBE buffer with ethidium bromide incorporated in the gels and visualized under UV. Since the γ-TMT_F9/γ-TMT_R24 INDEL marker yielded a fragment that was polymorphic between the gamma-T bulks (i.e., it was present in CL-1 and the low gamma-T bulk, and absent in IASC-1 and the high gamma-T bulk; see Fig. 3), it was genotyped in the 105 F2 individuals. Chi-square, ANOVA, and linkage analyses for the γ-TMT_F9/γ-TMT_R24 INDEL and *Tph2* loci were performed as described above.

Partial sequencing of the safflower gamma-TMT gene and development of gene-specific markers

The γ-TMT_F9/γ-TMT_R24 fragment present in the low gamma-T line CL-1, absent in the IASC-1 and polymorphic between the gamma-T bulks (Fig. 3) was isolated and cloned as described above. DNA sequencing of two clones was performed in both directions at GATC Biotechnology (Konstanz, Germany) using the M13 forward and reverse sequencing primers. The γ-TMT_F9/γ-TMT_R24 DNA sequence was used as query template in BLAST searches of the NCBI GenBank nucleotide databases, which resulted in homology to the gamma-TMT gene, the closest sequence being that of the sunflower gamma-TMT gene (GenBank accessions nos. DQ229828–DQ229834). The sequence conservation in the coding regions between safflower and sunflower prompted us to design new primers based on conserved coding regions of the sunflower gamma-TMT gene outside (upstream and downstream) the safflower fragment from γ-TMT_F9 to γ-TMT_R24 (external primers, a total of 19). These primers together with internal specific primers (a total of six) based on the safflower γ-TMT_F9/γ-TMT_R24 gamma-TMT sequence were used to amplify DNA from the parental lines CL-1, IASC-1 and the gamma-T bulks in order to obtain a longer genomic fragment from the safflower gamma-TMT gene. Amplification reactions were done as described above for the INDEL marker

γ-TMT_F9/γ-TMT_R24, excepting that annealing temperatures were adjusted for each primer pair. A longer fragment obtained with the safflower internal primer Cart_F6 (GATTCAGATGCTTTTCGTAATTAT) and the external downstream primer Heli_TMT_R54 (CGGGAAGATAAAAGCTGG), based on the sunflower gamma-TMT sequence, that was found to co-segregate with *Tph2*, was isolated from the IASC-1 and CL-1 lines and cloned and sequenced, as described above. The γ-TMT_F9/Heli_TMT_R54 nucleotide sequence from CL-1 was also obtained by isolating, cloning and sequencing the 1,800-bp γ-TMT_F9/Heli_TMT_R54 PCR fragment from CL-1. Sequence analyses and alignments and design of primers were performed using the software Vector NTI Advance 10.3.0. Intron splicing site sequences were predicted through sequence alignments with gamma-TMT cDNAs and verified with SplicePort Athaliana (http://spliceport.cs.umd.edu/Athaliana/Athaliana.html; Dogan et al. 2007).

Allele-specific markers for the safflower gamma-TMT gene were developed from sequence polymorphisms (INDEL and intron fragment length polymorphisms [IFLP]) identified between IASC-1 and CL-1 lines. The different INDEL and IFLP markers were tested in six F2 individuals and those showing unclear amplification patterns were discarded. The PCR reaction mixture and conditions for INDEL and IFLP markers were as described above, with the exception that annealing temperatures were adjusted for each primer pair. Linkage between the *Tph2* gene and polymorphic INDEL and IFLP markers was verified by genotyping the markers on the 105 F2 individuals. Chi-square, ANOVA and linkage analysis for the INDEL markers and the *Tph2* locus were performed as described above.

Safflower gamma-TMT INDEL and IFLP markers were also tested against a panel of 22 safflower cultivars, breeding lines and germplasm accessions. The cultivars and breeding lines were CR-6, CR-9, CR-142, CR-50, CR-58, CR-69, CR-34, CR-81, and Rancho (Velasco and Fernández-Martínez 2000, 2004; Velasco et al. 2005b), and the germplasm accessions were PI-259994, PI-401584, PI-537598, PI-537607, PI-537637, PI-537643, PI-534657, PI-537695, PI-537707, PI-560166, and PI-572471, provided by the Western Regional Plant Introduction Station of the US Department of Agriculture. DNA extraction and PCR analyses were conducted as described above. Heterozygosity (H) at each locus was estimated according to the formula $H = 1 - \Sigma p_i^2$, where p_i is the frequency of the ith allele (Nei 1978).

Results

Bulked segregant analysis

The 105 individuals of the half-seed F2 mapping population from the cross CL-1 × IASC-1 showed a bimodal distribution, with around three-quarters of the seeds having gamma-T content between 0 and 3.66% ($n = 80$) and around one-quarter of them having gamma-T content between 87.65 and 95.48% ($n = 25$), fitting the expected 3:1 ratio ($\chi^2 = 0.08$, $P = 0.78$). Fifteen out of 122 RAPD markers that produced clear and scorable amplification products showed one polymorphic fragment between the two parents CL-1 and IASC-1 and seemed to be polymorphic between the contrasting DNA bulks. Additionally, one out of the 88 SSRs tested was also polymorphic between the DNA bulks (co-dominant polymorphism). The RAPD markers and the SSR marker were screened against the 105 individuals in the F2 population. According to ANOVA analyses, eight of the polymorphic fragments revealed by RAPD primers OPAA11, OPAA12, OPG10, OPJ9, OPB8, OPM10, OPL20, and OPH12, and the SSR marker CAT49, showed a strong association with gamma-T content. These eight RAPD polymorphic fragments were approximately 1,650, 475, 350, 1,500, 750, 1,650, 1,050, and 1,100 bp long, and they were named $OPAA11_{1650}$, $OPAA12_{475}$, $OPG10_{350}$, $OPJ9_{1500}$, $OPB8_{750}$, $OPM10_{1650}$, $OPL20_{1050}$, and $OPH12_{1100}$, respectively. Linkage analyses grouped the eight RAPD markers and the SSR marker at the same linkage group as the *Tph2* locus (Fig. 1a).

The $OPL20_{1050}$ RAPD band from the low gamma-T line CL-1, and the $OPH12_{1100}$ and $OPAA12_{475}$ RAPD bands from the high gamma-T line IASC-1, were cloned and sequenced. The size of the consensus sequence for the two clones analysed for each of the RAPD fragments $OPL20_{1050}$, $OPH12_{1100}$, and $OPAA12_{475}$, was 926, 1,010, and 446 bp, respectively. BLAST search of these sequences did not reveal a strong significant homology to any known sequences in the databases. Two pairs of SCAR primers were developed from the

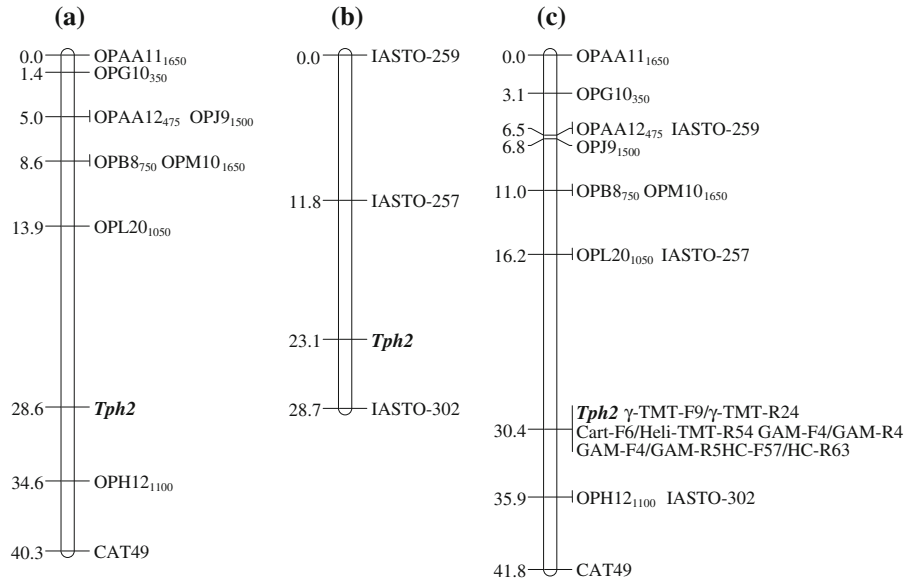


Fig. 1 Linkage maps containing the safflower *Tph2* gene for high gamma-tocopherol content. **a** Linkage map for *Tph2* and RAPD and SSR marker loci; **b** linkage map for *Tph2* and RAPD converted to SCAR, **c** complete linkage map for *Tph2* and RAPD, SCAR, SSR, candidate gene and gamma-TMT-based INDEL and IFLP marker loci. The cumulative distances in centiMorgans (Kosambi) are shown at the left of the map

consensus sequence for each cloned RAPD fragment. A first pair (F1 and R1) was designed including the ten bases of the initial RAPD primer sequence, and a second pair (F2 and R2) was located more internally. The sequence of the SCAR primers, their names, and the RAPD fragment from which they were developed are presented in Table 1. Each primer pair was used for PCR amplification of DNA from the parental lines CL-1 and IASC-1 and the gamma-T bulks. For SCAR markers IASTO-257 and IASTO-259, both F1/R1 and F2/R2 primer pairs gave a similar and clear amplification pattern. For IAST0-302, the primer combination F2/R2 failed to amplify. The three SCAR markers IASTO-259_F1/R1, IASTO-257_F1/R1, and IASTO-302_F1/R1, developed from OPAA12$_{475}$, OPL20$_{1050}$, and OPH12$_{1100}$, respectively, gave robust and polymorphic amplification patterns, both in the parental lines as well as in the *Tph2* bulks, and amplified specific fragments of the expected size (Fig. 2). These three polymorphic SCAR markers were genotyped in the whole F2 mapping population. Segregation data for the three SCAR loci gave a good fit to the expected 3:1 ratio (IASTO-259_F1/R1, $\chi^2 = 0.326$, $P = 0.567$; IASTO-257_F1/R1, $\chi^2 = 0.013$, $P = 0.908$; and IASTO-302_F1/R1, $\chi^2 = 0.003$, $P = 0.955$). Linkage analysis was performed with the three SCAR marker loci and the *Tph2* locus, with all the loci being grouped together (Fig. 1b).

Previously developed SCAR markers linked to the *Li* and *Ms* genes determining very high linoleic acid content and nuclear male sterility (Hamdan et al. 2008), respectively, were also tested between the two parents CL-1 and IASC-1 and the contrasting DNA bulks. The IASCA39 and the IASCA73 markers showed polymorphism between the parents, but not between the contrasting bulks. In order to confirm their lack of association with *Tph2*, they were genotyped in the CL-1 × IASC-1 mapping population. Linkage analysis did not integrate the IASCA39 or the IASCA73 marker loci within the *Tph2* linkage map.

Candidate gene analyses

Since no putative homologs for the safflower gamma-TMT were identified, INDEL markers based on the gamma-TMT gene from sunflower (Hass et al. 2006) were used. The γ-TMT_F9 and γ-TMT_R24 primers complementary to conserved DNA sequences in the first and second exons, respectively, of the sunflower gamma-TMT gene (Hass et al. 2006) amplified two DNA fragments in safflower: (1) a 1,000-bp band that was not polymorphic between the parental lines CL-1 and IASC-1, and (2) a 650-bp fragment that showed a dominant polymorphism between these two parents and the low and high gamma-T bulks (Fig. 3). The γ-TMT_F9/γ-TMT_R24 INDEL marker was

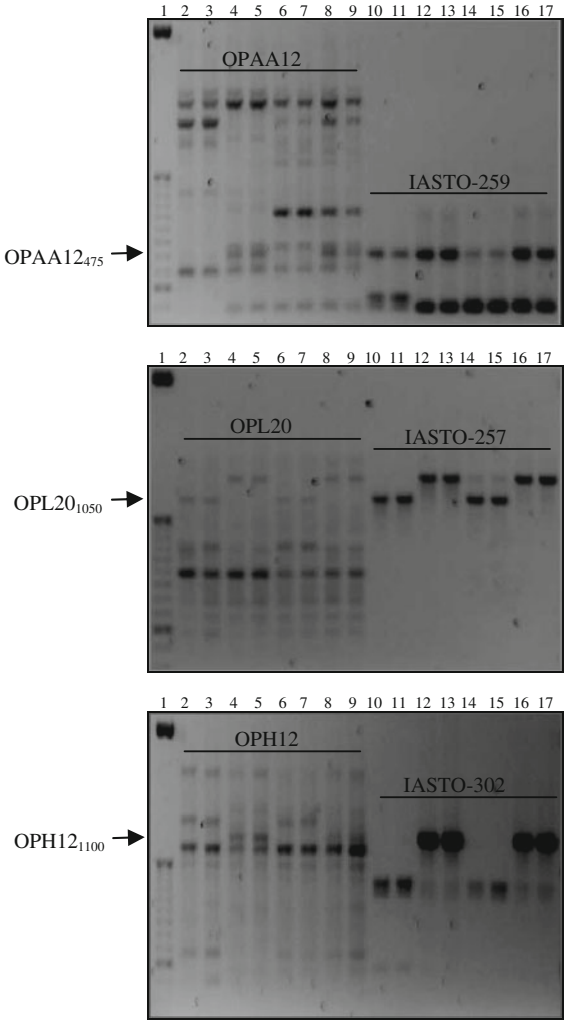


Fig. 2 Amplification profiles of RAPD markers OPAA12, OPL20, and OPH12 and their corresponding SCAR markers IASTO-259, IASTO-257, and IASTO-302, respectively, in replicate samples of the low gamma-tocopherol content parental line CL-1, the high gamma-tocopherol content parental line IASC-1, the low gamma-tocopherol bulk, and the high gamma-tocopherol bulk. RAPD polymorphic fragments from which SCAR markers were developed are indicated by arrows. *Lane 1* 50-bp DNA ladder; *lanes 2–3* and *10–11* replicate samples of CL-1; *lanes 4–5* and *12–13* replicate samples of IASC-1; *lanes 6–7* and *14–15* replicate samples of low gamma-tocopherol bulk; *lanes 8–9* and *16–17* replicate samples of the high gamma-tocopherol bulk

genotyped in the F2 CL-1 × IASC-1 population and the 650-bp marker locus co-segregated with *Tph2* (Fig. 1c). ANOVA analyses with genotypes at the 650-bp marker locus used as classes indicated clear significant differences ($P < 0.001$) between the

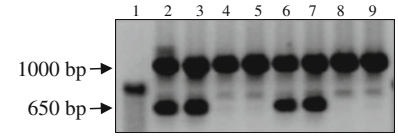


Fig. 3 Amplification profile of γ-TMT_F9/γ-TMT_R24 INDEL marker in replicate samples of the low gamma-tocopherol parental line CL-1, the high gamma-tocopherol parental line IASC-1, the low gamma-tocopherol bulk, and the high gamma-tocopherol bulk. *Lane 1* 50-bp DNA ladder; *lanes 2–3* replicate samples of CL-1; *lanes 4–5* replicate samples of IASC-1; *lanes 6–7* replicate samples of low gamma-tocopherol bulk; *lanes 8–9* replicate samples of high gamma-tocopherol bulk

marker class means for gamma-T content at the F2 and the F3 generations (Table 2).

Both the 1,000-bp and the *Tph2*-associated 650-bp fragments were cloned and sequenced, the first one in both the CL-1 and the IASC-1 parental lines, and the second one only in CL-1. The 1,000-bp fragment nucleotide sequence was identical between the CL-1 and the IASC-1 lines (two clones were sequenced per line). A 1,006-bp consensus sequence was used as query template for BLAST searches. The six most significant best hits corresponded to the sunflower gamma-TMT gene (DQ229828 and DQ229830 to DQ229834; $4e^{-13}$). For the *Tph2*-associated 650-bp fragment, a 663-bp consensus sequence was obtained for two clones from CL-1. BLAST search of this sequence also revealed a strong significant homology to the gamma-TMT gene of sunflower (DQ229828 and DQ229830 to DQ229834; $3e^{-18}$). The 1,006 and the 663-bp nucleotide fragments did not assembled in a contig. Sequence alignment indicated classification of the sequences into two groups, which were distinct from each other at several positions and could represent two different gene copies of the safflower gamma-TMT (data not shown).

Partial sequencing of the safflower gamma-TMT gene and development of gene-specific markers

The Cart-F6/Heli_TMT_R54 primer combination produced a unique and clear amplification product showing co-dominant polymorphism between the two parents CL-1 and IASC-1 and the gamma-T bulks. The fragments from CL-1 and IASC-1 were approximately 1,550 and 1,600 bp long, respectively. The Cart-F6/Heli_TMT_R54 primer pair was screened against the 105 individuals in the F2 population and co-segregated with *Tph2* (Table 2; Fig. 2c). The Cart-F6/Heli_

Table 2 Association between marker loci co-segregating with the *Tph2* gene and gamma-tocopherol content (% of total tocopherols) determined by variance analysis in the CL-1 × IASC-1 population

Marker	Seed generation	No. of individuals within each marker class				Mean ± SD for gamma-tocopherol content within each marker class[a]				ANOVA analyses	
		A	H	B	D	A	H	B	D	*F*	*P*
γ-TMT_F9/γ-TMT_R24	F2			25	80			93.2 ± 1.8	0.8 ± 0.9	116,184.2	<0.001
	F3			12	41			95.5 ± 2.0	17.3 ± 16.0	282.5	<0.001
Cart-F6/Heli_TMT_R54	F2	24	56	25		0.03 ± 0.1	1.13 ± 0.9	93.2 ± 1.8		67,002.1	<0.001
	F3	13	28	12		0.14 ± 0.3	25.3 ± 13.0	95.5 ± 2.0		337.9	<0.001

[a] Mean gamma-tocopherol content ± standard deviation (SD) is presented in different genotypic classes: A, homozygous with respect to the allele derived from CL-1; B, homozygous with respect to the allele derived from IASC-1; D, H or A

TMT_R54 band from CL-1 and that from IASC-1 were cloned and sequenced. The size of the consensus sequence for the two clones analysed per line was 1,558 bp for CL-1 and 1,593 bp for IASC-1. BLAST search of CL-1 and IASC-1 sequences revealed a strong significant homology to the sunflower gamma-TMT gene ($2e^{-55}$ and $7e^{-56}$, respectively). The Cart-F6/Heli_TMT_R54 and γ-TMT_F9/γ-TMT_R24 DNA sequences from CL-1 were aligned, and showed an identical sequence in the common fragment from Cart-F6 to γ-TMT_R24.

The γ-TMT_F9/Heli_TMT_R54 nucleotide sequence from CL-1 (1,875 bp, GenBank accession number HM028671), the Cart-F6/Heli_TMT_R54 from IASC-1 (1,593 bp, GenBank accession number HM028672), and the sequence fragment from sunflower gamma-TMT gene (DQ229830.1) were aligned. Sequence conservation between safflower and sunflower was high and concentrated on coding regions, with 91% maximum sequence identity spanning 30% of the CL-1 fragment coverage. The 1,875-bp partial sequence from CL-1 spanned from exon 1 to exon 4 of the six exons present in the sunflower gamma-TMT gene, and was predicted to contain about one-half of the gene. Nucleotide sequence alignment between the 1,875-bp fragment from CL-1 and the 1,593-bp fragment from IASC-1 revealed several differences within intron sequences, including nucleotide mutations, insertions, and deletions (Fig. 4). In the coding regions, one non-synonymous nucleotide mutation in exon 3 and two synonymous nucleotide mutations in exon 4 were identified in IASC-1 compared to CL-1 (Fig. 4). The non-synonymous nucleotide mutation (C to T) implied a change of alanine in CL-1 to valine in IASC-1 (Figs. 4, 5). In that position, valine was also found in the sunflower gamma-TMT amino acid sequences, both in standard and in seed tocopherol modified lines (Fig. 5) (Hass et al. 2006).

INDEL and IFLP allele-specific markers were developed based on gamma-TMT sequence polymorphisms identified in CL-1 and IASC-1. Two INDEL markers (GAM_F4/GAM_R4 and GAM_F4/GAM_R5, Table 3) and one IFLP marker (HC_F57/HC_R63, Table 3) produced clear and co-dominant amplification products that co-segregated with *Tph2* when genotyped in the mapping population (Fig. 2c). These markers were screened against a panel of 22 safflower cultivars, breeding lines, and germplasm accessions. Two different alleles were observed for the INDEL markers, with that present in the high gamma-T line IASC-1 being unique among the 22 lines tested (Fig. 6 for GAM_F4/GAM_R5). The IFLP marker also revealed two alleles, and showed a heterozygosity value of 0.43.

Discussion

Safflower has not been the subject of trait mapping studies until recently. Previous studies involving simply inherited traits have been carried out using anonymous markers, mainly bulked segregant analysis with RAPD or AFLP markers to be converted into more robust SCAR markers. In this sense, Hamdan et al. (2008) reported the development of SCAR markers linked to the *Li* and *Ms* genes determining very high linoleic acid content and nuclear male sterility, respectively, and Zhang et al. (2009) described the identification of a SCAR marker

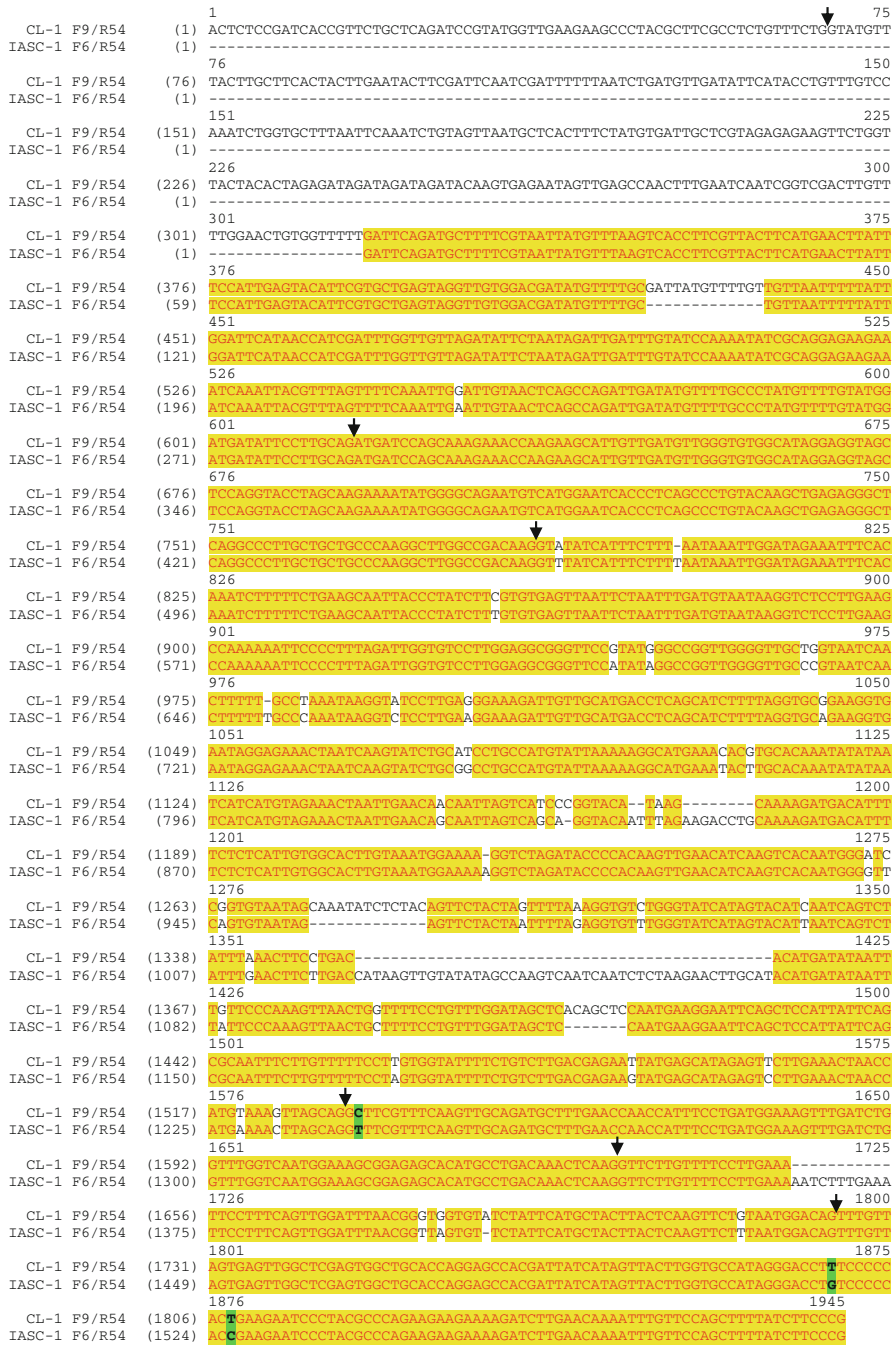


Fig. 4 Alignment of safflower gamma-tocopherol methyltransferase (*gamma-TMT*) partial genomic DNA sequence in CL-1 and IASC-1. Predicted intron splicing sites are indicated by *arrows*. Single nucleotide polymorphisms (*SNPs*) in coding regions are highlighted

associated with the *HSya* gene controlling the accumulation of hydroxysafflor yellow A, a flavonoid component of a crude drug from traditional Chinese medicine. Our study combined bulked segregant analysis and the development of SCAR markers with a candidate gene approach, and allowed the development of robust and efficient molecular markers for the gene determining high gamma-T content in safflower seeds and the integration of this gene into a SCAR-SSR-RAPD linkage map. These results represent a valuable advance towards the implementation of molecular breeding programs in safflower.

Candidate gene analysis revealed polymorphisms located in the safflower gamma-TMT gene that

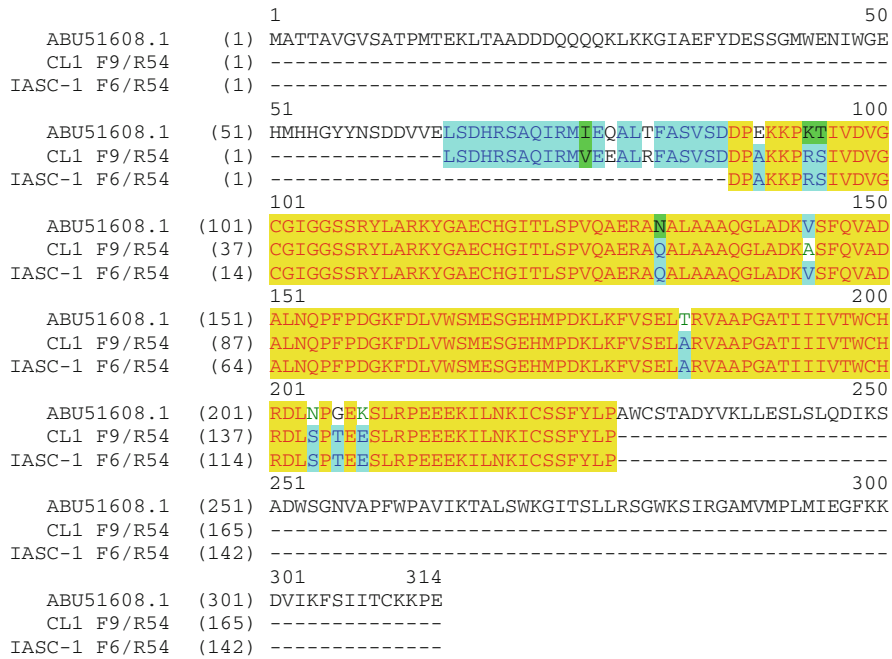


Fig. 5 Amino acid sequence alignment of safflower gamma-tocopherol methyltransferase (*gamma-TMT*) from CL-1 (CL1 F9/R54) and IASC-1 (IASC-1 F6/R54) and sunflower gamma-tocopherol methyltransferase (ABU51608.1)

Table 3 Sequences, locations, orientations, and melting temperature (T_m) of INDEL and IFLP primers based on the safflower gamma-TMT sequence

Primer name	Type of marker	Location nt[a]	Orientation	Sequence (5′–3′)	T_m
GAM_F4	INDEL	1210	F	GTGGCACTTGTAAATGGAA	54.5
GAM_R4	INDEL	1,414	R	TGGGAACAAATTATATCATG	51.0
GAM_R5	INDEL	1,440	R	CTATCCAAACAGGAAAAGCAGT	57.6
HC_F57	IFLP	47	F	CTACGCTTCGCCTCTGTTT	58.3
HC_R63	IFLP	606	R	TGGATCATCTGCAAGGAAT	56.5

[a] The nucleotide (nt) locations of forward (F) and reverse (R) primers are identified as shown in reference genomic DNA sequence alignments for gamma-TMT (Fig. 4)

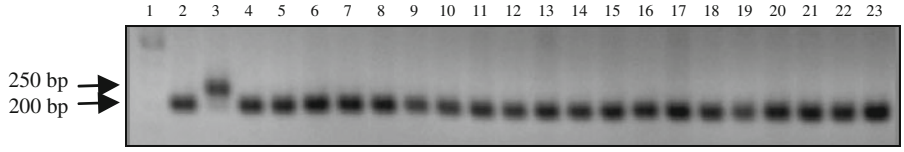


Fig. 6 Amplification profile of INDEL marker GAM_F4/GAM_R5 among several safflower cultivars, breeding lines, and germplasm accessions. *Lane 1* 50-bp DNA ladder; *lane 2* CL-1; *lane 3* IASC-1; *lane 4* Rancho; *lane 5* CR-9; *lane 6* CR-142; *lane 7* CR-50; *lane 8* CR-58; *lane 9* CR-69; *lane 10* CR-6; *lane 11* CR-34; *lane 12* CR-81; *lane 13* PI-259994; *lane 14* PI-401584; *lane 15* PI-537598; *lane 16* PI-537607; *lane 17* PI-537637; *lane 18* PI-537643; *lane 19* PI-534657; *lane 20* PI-537695; *lane 21* PI-537707; *lane 22* PI-560166; *lane 23* PI-572471

co-segregated with the gamma-T content. These results strongly indicate that gamma-TMT underlies the high gamma-T mutation in safflower, and they are in agreement with previous studies on sunflower and *Arabidopsis* mutants with high gamma-T content in the seeds (Hass et al. 2006) or the leaves (Bergmüller et al. 2003), respectively, in which high gamma-T levels were also associated with the gene encoding

the gamma-TMT. However, despite the fact that safflower gamma-TMT DNA polymorphisms were diagnostic for the *Tph2* mutation, none of those found in the coding regions were identified as responsible for altered gamma-TMT function leading to the phenotypic change from low to high seed gamma-T content. Two of the mutations detected were synonymous changes, and the non-synonymous change produced an amino acid in the high gamma-T safflower line IASC-1 identical to that found in standard-low gamma-T lines in sunflower (Hass et al. 2006). These authors also reported the lack of mutations in the coding sequence of the gamma-TMT gene associated with the mutant phenotype in sunflower, and speculated that they could disrupt regulatory sequences. Nevertheless, because only one-half of the safflower gamma-TMT gene was sequenced, the presence of mutations elsewhere in the IASC-1 gamma-TMT gene cannot be excluded.

The key components required for an efficient system for molecular breeding are the identification and characterization of suitable genetic markers and the development of a reference genetic map using both an internationally accepted linkage group nomenclature system and publicly available markers, for cross-referencing maps and mapped gene locations. Although safflower lags behind other oilseed crops in terms of development of these necessary molecular breeding tools, there are ongoing projects with this objective (Mayerhofer et al. 2008; Ravikumar et al. 2008; Chapman et al. 2009. Integration of *Tph2* in an SCAR-SSR-RAPD linkage map would accelerate locating this gene in a complete safflower linkage map and cross-referencing with other trait mapping studies. In fact, this has already been possible for the *Li* and *Ms* genes, which are likely unlinked to *Tph2*, since none of the polymorphic SCAR markers previously reported to be linked to these genes (Hamdan et al. 2008) were integrated within the *Tph2* linkage map.

The production of high-value oils for specific market niches has been proposed for the enhancement of safflower cultivation (Bergman and Flynn 2001). Oils with high gamma-T content exhibit resistance to oxidation processes that occur during long-term storage or high temperature operations of vegetable oils (Marmesat et al. 2008; Warner et al. 2008). In this way, the incorporation of the novel, non-genetically-engineered high gamma-T trait into the currently existing safflower types with contrasting fatty acid profiles (Knowles 1989; Hamdan et al. 2009) opens up new potential applications for safflower oil. Molecular markers developed in this research will greatly support introgression of *thp2* alleles into elite lines with different fatty acid profiles through marker-assisted breeding programs. Gamma-TMT-based markers are highly predictive for the phenotype and optimal for selection, since recombination between the gene associated with the modified phenotype and the marker has been found to be null. Additionally, the INDEL and IFLP markers developed are co-dominant, and are therefore the most useful for marker-assisted backcrossing because selection among backcross progeny involves identification of heterozygous individuals. It is noteworthy that the gamma-TMT-based INDEL allele from the high gamma-T line IASC-1 was unique among a diverse array of cultivars, breeding lines, and safflower accessions tested, identifying this line unambiguously, and therefore it will have a wide application in marker-assisted breeding. Additionally, the linkage map in which the *Tph2* gene has been integrated, with robust SCAR and SSR markers flanking the gene, is an efficient tool for selecting against donor parent alleles around the *Tph2* gene.

In conclusion, the results from this study have shed light on the genetic mechanism underlying the high gamma-T mutation in safflower, and allowed the development of efficient tools for the establishment of molecular breeding programs in this crop aimed at the development of high gamma-T cultivars.

Acknowledgments The research was funded by the Spanish Ministry of Science and Innovation and the European Union FEDER (research project AGL2007-62834).

References

Bergman JW, Flynn CR (2001) High oleic safflower as a diesel fuel extender. A potential new market for Montana safflower. In: Bergman JW, Mündel HH (eds) Proceedings of 5th international safflower conference, Williston and Sidney, pp 289–293

Bergmüller E, Porfirova S, Dörmann P (2003) Characterization of an *Arabidopsis* mutant deficient in gamma-tocopherol methyltransferase. Plant Mol Biol 52:1181–1190

Chapman M, Hvala J, Strever J, Matvienko M, Kozik A, Michelmore R, Tang S, Knapp S, Burke J (2009) Development, polymorphism, and cross-taxon utility of

EST–SSR markers from safflower (*Carthamus tinctorius* L.). Theor Appl Genet 120:85–91
Demurin Y (1993) Genetic variability of tocopherol composition in sunflower seeds. Helia 16:59–62
Demurin Y, Škoric D, Karlovic D (1996) Genetic variability of tocopherol composition in sunflower seeds as a basis of breeding for improved oil quality. Plant Breed 115:33–36
Dogan RI, Getoor L, Wilbur WJ, Mount SM (2007) SplicePort—an interactive splice-site analysis tool. Nucleic Acids Res (Web Server issue) 35:W285–W291
García-Moreno MJ, Vera-Ruiz EM, Fernández-Martínez JM, Velasco L, Pérez-Vich B (2006) Genetic and molecular analysis of high gamma-tocopherol content in sunflower. Crop Sci 46:2015–2021
García-Moreno MJ, Velasco L, Pérez-Vich B (2010). Transferability of non-genic microsatellite and gene-based sunflower markers to safflower. Euphytica (in press, doi: 10.1007/s10681-010-0139-6)
Goffman FD, Velasco L, Thies W (1999) Quantitative determination of tocopherols in single seeds of rapeseed (*Brassica napus* L.). Lipid Fett 101:142–145
Hamdan YAS, Velasco L, Pérez-Vich B (2008) Development of SCAR markers linked to male sterility and very high linoleic acid content in safflower. Mol Breed 22:385–393
Hamdan YAS, Pérez-Vich B, Fernández-Martínez JM, Velasco L (2009) Novel safflower germplasm with increased saturated fatty acid content. Crop Sci 49:127–132
Hass C, Tang S, Leonard S, Traber M, Miller J, Knapp S (2006) Three non-allelic epistatically interacting methyltransferase mutations produce novel tocopherol (vitamin E) profiles in sunflower. Theor Appl Genet 113:767–782
Heaton TC, Knowles PF (1980) Registration of UC-148 and UC-149 male-sterile safflower germplasm. Crop Sci 20:554
Heaton TC, Knowles PF (1982) Inheritance of male sterility in safflower. Crop Sci 22:520–522
Johnson RC, Bergman JW, Flynn CR (1999) Oil and meal characteristics of core and non-core safflower accessions from the USDA collection. Genet Resour Crop Evol 46:611–618
Kamal-Eldin A, Appelqvist LÅ (1996) The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31:671–701
Knowles PF (1989) Safflower. In: Downey RK, Röbbelen G, Ashri A (eds) Oil crops of the world. McGraw-Hill, New York, pp 363–374
Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181
Marmesat S, Velasco L, Ruiz-Méndez MV, Fernández-Martínez JM, Dobarganes C (2008) Thermostability of genetically modified sunflower oils differing in fatty acid and tocopherol compositions. Eur J Lipid Sci Technol 110:776–782
Mayerhofer R, Bowles V, Mayerhofer M, Good AG (2008) Genetic linkage maps of *Carthamus* species based on SSR and RFLP markers. In: Proceedings of 7th international safflower conference, Wagga Wagga, 3–6 November 2008
Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832
Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590
Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993
Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Mol Breed 7: 275–291
Ravikumar RL, Roopa VK, Soregaon CD, Satish D (2008) Molecular diversity in *Carthamus* species and development of inter-specific mapping population toward development of the first molecular map in safflower. In: Proceedings of 7th international safflower conference, Wagga Wagga, 3–6 November 2008
Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76
Schneider MP (2006) Plant-oil-based lubricants and hydraulic fluids. J Sci Food Agric 86:1769–1780
Shintani D, DellaPenna D (1998) Elevating the vitamin E content of plants through metabolic engineering. Science 282:2098–2100
Trabber MG, Sies H (1996) Vitamin E in humans: demand and delivery. Annu Rev Nutr 16:321–347
Velasco L, Fernández-Martínez JM (2000) Isolation of lines with contrasting seed oil fatty acid profiles from safflower germplasm. Sesame Safflower Newsl 15:104–108
Velasco L, Fernández-Martínez JM (2004) Registration of CR-34 and CR-81 safflower germplasms with increased tocopherol. Crop Sci 44:2278
Velasco L, Domínguez J, Fernández-Martínez JM (2004) Registration of T589 and T2100 sunflower germplasms with modified tocopherol profiles. Crop Sci 44:361–362
Velasco L, Pérez-Vich B, Fernández-Martínez JM (2005a) Identification and genetic characterization of a safflower mutant with a modified tocopherol profile. Plant Breed 124:459–463
Velasco L, Pérez-Vich B, Hamdan Y, Fernández-Martínez JM (2005b) Genetic study of several seed oil quality traits in safflower. In: Proceedings of 6th international safflower conference Istanbul, 6-–10 June 2005, pp 74–79
Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78
Warner K, Moser J (2009) Frying stability of purified mid-oleic sunflower oil triacylglycerols with added pure tocopherols and tocopherol mixtures. J Am Oil Chem Soc 86:1199–1207
Warner K, Miller J, Demurin Y (2008) Oxidative stability of crude mid-oleic sunflower oils from seeds with high gamma and delta tocopherol levels. J Am Oil Chem Soc 85:529–533
Zhang Z, Guo M, Zhang J (2009) Identification of AFLP fragments linked to hydroxysafflor yellow A in *Flos Carthami* and conversion to a SCAR marker for rapid selection. Mol Breed 23:229–237

Carthamus tinctorius cultivar CL-1 gamma-tocopherol methyltransferase gene, partial sequence

```
LOCUS       HM028671               1875 bp    DNA     linear   PLN 06-OCT-2011
DEFINITION  Carthamus tinctorius cultivar CL-1 gamma-tocopherol
            methyltransferase gene, partial sequence.
ACCESSION   HM028671
VERSION     HM028671.1  GI:301072475
KEYWORDS    .
SOURCE      Carthamus tinctorius (safflower)
  ORGANISM  Carthamus tinctorius
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; campanulids; Asterales; Asteraceae; Carduoideae;
            Cardueae; Centaureinae; Carthamus.
REFERENCE   1  (bases 1 to 1875)
  AUTHORS   Garcia-Moreno,M.J., Fernandez-Martinez,J.M., Velasco,L. and
            Perez-Vich,B.
  TITLE     Molecular tagging and candidate gene analysis of the high
            gamma-tocopherol trait in safflower (Carthamus tinctorius L.)
  JOURNAL   Mol. Breed. 28 (3), 367-379 (2011)
REFERENCE   2  (bases 1 to 1875)
  AUTHORS   Garcia-Moreno,M.J., Fernandez-Martinez,J.M., Velasco,L. and
            Perez-Vich,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-2010) Plant Breeding, Instituto de Agricultura
            Sostenible (CSIC), Alameda del Obispo s/n, Cordoba 14004, Spain
FEATURES             Location/Qualifiers
     source          1..1875
                     /organism="Carthamus tinctorius"
                     /mol_type="genomic DNA"
                     /cultivar="CL-1"
                     /db_xref="taxon:4222"
     gene            <1..>1875
                     /gene="gamma-tocopherol methyltransferase"
                     /note="coding region not determined"
ORIGIN
        1 actctccgat caccgttctg ctcagatccg tatggttgaa gaagccctac gcttcgcctc
       61 tgtttctggt atgtttactt gcttcactac ttgaatactt cgattcaatc gatttttaa
      121 tctgatgttg atattcatac ctgtttgtcc aaatctggtg ctttaattca aatctgtagt
      181 taatgctcac tttctatgtg attgctcgta gagagaagtt ctggttacta cactagagat
      241 agatagatag atacaagtga gaatagttga gccaactttg aatcaatcgg tcgacttgtt
      301 ttggaactgt ggtttttgat tcagatgctt ttcgtaatta tgtttaagtc accttcgtta
      361 cttcatgaac ttatttccat tgagtacatt cgtgctgagt aggttgtgga cgatatgttt
      421 tgcgattatg ttttgttgtt aatttttatt ggattcataa ccatcgattt ggttgttaga
      481 tattctaata gattgatttg tatccaaaat atcgcaggag aagaaatcaa attacgttta
      541 gttttcaaat tggattgtaa ctcagccaga ttgatatgtt ttgccctatg ttttgtatgg
      601 atgatattcc ttgcagatga tccagcaaag aaaccaagaa gcattgttga tgttgggtgt
      661 ggcataggag gtagctccag gtacctagca agaaaatatg gggcagaatg tcatggaatc
      721 accctcagcc ctgtacaagc tgagagggct caggcccttg ctgctgccca aggcttggcc
      781 gacaaggtat atcatttctt taataaattg gatagaaatt tcacaaatct ttttctgaag
      841 caattaccct atcttcgtgt gagttaattc taatttgatg taataaggtc tccttgaagc
      901 caaaaaattc ccctttagat tggtgtcctt ggaggcgggt tccgtatggg ccggttgggg
      961 ttgctggtaa tcaacttttt gcctaaataa ggtatccttg agggaaagat tgttgcatga
     1021 cctcagcatc ttttaggtgc ggaaggtgaa taggagaaac taatcaagta tctgcatcct
     1081 gccatgtatt aaaaaggcat gaaacacgtg cacaaatata taatcatcat gtagaaacta
     1141 attgaacaac aattagtcat cccggtacat aagcaaaaga tgacattttc tctcattgtg
     1201 gcacttgtaa atggaaaagg tctagatacc ccacaagttg aacatcaagt cacaatggga
     1261 tccggtgtaa tagcaaatat ctctacagtt ctactagttt taaaggtgtc tgggtatcat
     1321 agtacatcaa tcagtctatt taaacttcct gacacatgat ataatttgtt cccaaagtta
     1381 actggttttc ctgtttggat agctcacagc tccaatgaag gaattcagct ccattattca
     1441 gcgcaatttc ttgtttttcc ttgtggtatt ttctgtcttg acgagaatta tgagcataga
     1501 gttcttgaaa ctaaccatgt aaagttagca ggcttcgttt caagttgcag atgctttgaa
     1561 ccaaccattt cctgatggaa agtttgatct ggtttggtca atggaaagcg gagagcacat
     1621 gcctgacaaa ctcaaggttc ttgttttcct tgaaattcct ttcagttgga tttaacgggt
     1681 ggtgtatcta ttcatgctac ttactcaagt tctgtaatgg acagtttgtt agtgagttgg
     1741 ctcgagtggc tgcaccagga gccacgatta tcatagttac ttggtgccat agggaccttt
     1801 cccccactga agaatcccta cgcccagaag aagaaaagat cttgaacaaa atttgttcca
     1861 gcttttatct tcccg
//
```

Carthamus tinctorius cultivar IASC-1 gamma-tocopherol methyltransferase gene, partial sequence

```
LOCUS       HM028672               1593 bp    DNA     linear   PLN 06-OCT-2011
DEFINITION  Carthamus tinctorius cultivar IASC-1 gamma-tocopherol
            methyltransferase gene, partial sequence.
ACCESSION   HM028672
VERSION     HM028672.1  GI:301072476
KEYWORDS    .
SOURCE      Carthamus tinctorius (safflower)
  ORGANISM  Carthamus tinctorius
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; campanulids; Asterales; Asteraceae; Carduoideae;
            Cardueae; Centaureinae; Carthamus.
REFERENCE   1  (bases 1 to 1593)
  AUTHORS   Garcia-Moreno,M.J., Fernandez-Martinez,J.M., Velasco,L. and
            Perez-Vich,B.
  TITLE     Molecular tagging and candidate gene analysis of the high
            gamma-tocopherol trait in safflower (Carthamus tinctorius L.)
  JOURNAL   Mol. Breed. 28 (3), 367-379 (2011)
REFERENCE   2  (bases 1 to 1593)
  AUTHORS   Garcia-Moreno,M.J., Fernandez-Martinez,J.M., Velasco,L. and
            Perez-Vich,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAR-2010) Plant Breeding, Instituto de Agricultura
            Sostenible (CSIC), Alameda del Obispo s/n, Cordoba 14004, Spain
FEATURES             Location/Qualifiers
     source          1..1593
                     /organism="Carthamus tinctorius"
                     /mol_type="genomic DNA"
                     /cultivar="IASC-1"
                     /db_xref="taxon:4222"
     gene            <1..>1593
                     /gene="gamma-tocopherol methyltransferase"
                     /note="coding region not determined"
ORIGIN
        1 gattcagatg cttttcgtaa ttatgtttaa gtcaccttcg ttacttcatg aacttatttc
       61 cattgagtac attcgtgctg agtaggttgt ggacgatatg ttttgctgtt aatttttatt
      121 ggattcataa ccatcgattt ggttgttaga tattctaata gattgatttg tatccaaaat
      181 atcgcaggag aagaaatcaa attacgttta gttttcaaat tgaattgtaa ctcagccaga
      241 ttgatatgtt ttgccctatg ttttgtatgg atgatattcc ttgcagatga tccagcaaag
      301 aaaccaagaa gcattgttga tgttgggtgt ggcataggag gtagctccag gtacctagca
      361 agaaaatatg gggcagaatg tcatggaatc accctcagcc ctgtacaagc tgagagggct
      421 caggcccttg ctgctgccca aggcttggcc gacaaggttt atcatttctt ttaataaatt
      481 ggatagaaat ttcacaaatc tttttctgaa gcaattaccc tatctttgtg tgagttaatt
      541 ctaatttgat gtaataaggt ctccttgaag ccaaaaaatt ccccttaga ttggtgtcct
      601 tggaggcggg ttccatatag gccggttggg gttgcccgta atcaactttt ttgcccaaat
      661 aaggtctcct tgaaggaaag attgttgcat gacctcagca tcttttaggt gcagaaggtg
      721 aataggagaa actaatcaag tatctgcggc ctgccatgta ttaaaaaggc atgaaatact
      781 tgcacaaata tataatcatc atgtagaaac taattgaaca gcaattagtc agcaggtaca
      841 atttagaaga cctgcaaaag atgacatttt ctctcattgt ggcacttgta aatggaaaaa
      901 ggtctagata ccccacaagt tgaacatcaa gtcacaatgg ggttcagtgt aatagagttc
      961 tactaatttt agaggtgttt gggtatcata gtacattaat cagtctattt gaacttcttg
     1021 accataagtt gtatatagcc aagtcaatca atctctaaga acttgcatac atgatataat
     1081 ttattcccaa agttaactgc ttttcctgtt tggatagctc caatgaagga attcagctcc
     1141 attattcagc gcaatttctt gtttttccta gtggtatttt ctgtcttgac gagaagtatg
     1201 agcatagagt ccttgaaact aaccatgaaa acttagcagg tttcgtttca agttgcagat
     1261 gctttgaacc aaccatttcc tgatggaaag tttgatctgg tttggtcaat ggaaagcgga
     1321 gagcacatgc ctgacaaact caaggttctt gttttccttg aaaaatcttt gaaattcctt
     1381 tcagttggat ttaacggtta gtgttctatt catgctactt actcaagttc tttaatggac
     1441 agtttgttag tgagttggct cgagtggctg caccaggagc cacgattatc atagttactt
     1501 ggtgccatag ggacctgtcc cccaccgaag aatccctacg cccagaagaa gaaaagatct
     1561 tgaacaaaat ttgttccagc ttttatcttc ccg
//
```

INFORME DEL FACTOR DE IMPACTO

María J. García-Moreno, Elsa M. Vera-Ruiz, José M. Fernández-Martínez, Leonardo Velasco, and Begoña Pérez-Vich (2006) **Genetic and Molecular Analysis of High Gamma-Tocopherol Content in Sunflower.** Crop Science, 46:2015-2021.

Factor de impacto en 2006

Cites in 2006 to items published in:	2005 =	437	Number of items published in:	2005 =	467
	2004 =	545		2004 =	385
	Sum:	982		Sum:	852
Calculation:	Cites to recent items	982	= **1.153**		
	Number of recent items	852			

Journal Ranking en 2006

For **2006**, the journal **CROP SCIENCE** has an Impact Factor of **1.153**.

This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
AGRONOMY	49	17	Q2

Reproduced from Crop Science. Published by Crop Science Society of America. All copyrights reserved.

Published online September 8, 2006

Genetic and Molecular Analysis of High Gamma-Tocopherol Content in Sunflower

María J. García-Moreno, Elsa M. Vera-Ruiz, José M. Fernández-Martínez, Leonardo Velasco, and Begoña Pérez-Vich*

ABSTRACT

Sunflower (*Helianthus annuus* L.) seeds contain alpha-tocopherol as the major tocopherol derivative, which accounts for more than 900 g kg^{-1} total tocopherols. However, four sources of high gamma-tocopherol content (>850 g kg^{-1}) have been developed. First studies on the lines LG-17 and T2100 concluded that the trait in both lines was determined by recessive alleles at the Tph2 locus. The objectives of the present research were (i) to conduct an allelic study on the other two lines, IAST-1 and IAST-540, (ii) to identify markers linked to the Tph2 gene, and (iii) to map this gene. Plants of T2100 were crossed with plants of the other three lines, which resulted in F_1 and F_2 populations with uniformly high gamma-tocopherol content in the seeds, indicating the presence of *tph2* alleles in the four lines. Genetic mapping of the Tph2 gene was conducted with an F_2 population from the cross between CAS-12 with standard tocopherol profile, and IAST-540. F_2 seeds segregated following a 3 low to 1 high gamma-tocopherol ratio. Bulked segregant analysis identified two simple sequence repeats (SSR) markers on linkage group (LG) 8 linked to Tph2. A large linkage group was constructed by genotyping additional markers. Tph2 mapped between markers ORS312 (3.6 cM proximal) and ORS599 (1.9 cM distal). The availability of closely linked PCR-based markers and the location of the Tph2 gene on the sunflower genetic map will be useful for marker-assisted selection and further characterization of tocopherol biosynthesis in sunflower seeds.

Tocopherols are the most important compounds having antioxidant activity in sunflower seeds. In vivo, they exert vitamin E activity, protecting cellular membrane lipids against oxidative damage (Muggli, 1994). In vitro, they inhibit lipid oxidation in oils and fats, as well as in foods and feeds containing them (Kamal-Eldin and Appelqvist, 1996). Alpha-tocopherol exerts the most active biological activity (Traber and Sies, 1996), but it shows the weakest antioxidant potency in vitro. Conversely, beta-, gamma-, and delta-tocopherol possess a lower vitamin E value, but they exert a considerably greater in vitro antioxidant protection than alpha-tocopherol (Pongracz et al., 1995).

Conventional sunflower seeds mainly contain alpha-tocopherol, which accounts for more than 900 g kg^{-1} total tocopherols. Beta- and gamma-tocopherol can be present in sunflower seeds, usually in amounts below 20 g kg^{-1} of the total tocopherols (Demurin, 1993; Dolde et al., 1999). Sunflower germplasm with modified tocopherol profile has been developed. Demurin (1993) reported the line LG-15, with increased concentration of beta-tocopherol (500 g kg^{-1} tocopherols), and the line LG-17, with increased concentration of gamma-tocopherol (950 g kg^{-1} tocopherols), both of them developed from segregating accessions identified in the evaluation of a germplasm collection. Genetic characterization of both lines concluded that the increased levels of beta-tocopherol were produced by recessive alleles at the Tph1 locus, whereas the increased levels of gamma-tocopherol were the result of recessive alleles at the Tph2 locus (Demurin et al., 1996). Also through the evaluation of the natural variability existing in germplasm collections, Velasco et al. (2004a) developed the line T589, with a beta-tocopherol content above 300 g kg^{-1} total tocopherols, and the line T2100, with a gamma-tocopherol content above 850 g kg^{-1}. Velasco and Fernández-Martínez (2003) reported the presence of recessive alleles at a single locus underlying each of the modified tocopherol profiles, i.e., the increased beta-tocopherol concentration in seeds of T589 and the high gamma-tocopherol content in seeds of T2100. Comparative genetic studies concluded that *tph1* alleles were present in both LG-15 and T589 lines (Demurin et al., 2004; Vera-Ruiz et al., 2005), and *tph2* alleles were present in both LG-17 and T2100 lines (Demurin et al., 2004).

Additional variation for gamma-tocopherol content was created in sunflower by using chemical mutagenesis (Velasco et al., 2004b). The authors isolated the lines IAST-1 and IAST-540, with gamma-tocopherol content above 850 g kg^{-1} total tocopherols. No comparative genetic studies have been conducted to determine whether the high gamma-tocopherol lines developed by mutagenesis are allelic to those developed through germplasm evaluation.

Recent advances in molecular marker technologies in sunflower, especially the development of public SSRs (microsatellites) (Tang et al., 2002), SNPs (single nucleotide polymorphisms) (Lai et al., 2005), and integrated genetic linkage maps (Gedil et al., 2001; Yu et al., 2003; Lai et al., 2005) have made possible the genetic mapping and dissection of quantitative and qualitative traits in this crop and the application of this technology to sunflower breeding. Genetic mapping of tocopherol biosynthesis genes and identification of molecular markers linked to them would provide important tools for increased selection efficiency and for investigating the function and organization of these genes. Currently, only the Tph1 gene, conferring increased beta-tocopherol content to sunflower seeds, has been mapped in the sunflower genetic map (Vera-Ruiz et al., 2006). This gene mapped to the upper end of linkage group 1 and cosegregated with the SSR markers ORS1093, ORS222, and ORS598.

The objectives of the present research were (i) to conduct a comparative genetic analysis of the four sources

Instituto de Agricultura Sostenible (CSIC). Apartado 4084. E-14080 Córdoba, Spain. Received 25 Oct. 2005. *Corresponding author (bperez@ias.csic.es).

Published in Crop Sci. 46:2015–2021 (2006).
Crop Breeding & Genetics
doi: 10.2135/cropsci2005.10.0368
© Crop Science Society of America
677 S. Segoe Rd., Madison, WI 53711 USA.

Abbreviations: HPLC, high-performance liquid chromatography; INDEL, insertion-deletion polymorphism; SSR, simple sequence repeats.

M. J. García-Moreno, L. Velasco and B. Pérez-Vich (2010) **Transferability of non-genic microsatellite and gene-based sunflower markers to safflower.** Euphytica 175:145–150.

Factor de impacto en 2010

Cites in 2010 to items published in:	2009 =	348	Number of items published in:	2009 =	240
	2008 =	511		2008 =	298
	Sum:	859		Sum:	538
Calculation:	Cites to recent items	859	= **1.597**		
	Number of recent items	538			

Journal Ranking en 2010

For **2010**, the journal **EUPHYTICA** has an Impact Factor of **1.597**.

This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
AGRONOMY	75	22	Q2
HORTICULTURE	30	7	Q1
PLANT SCIENCES	188	74	Q2

Euphytica (2010) 175:145–150
DOI 10.1007/s10681-010-0139-6

Transferability of non-genic microsatellite and gene-based sunflower markers to safflower

M. J. García-Moreno · L. Velasco · B. Pérez-Vich

Received: 26 July 2009 / Accepted: 22 January 2010 / Published online: 6 February 2010
© Springer Science+Business Media B.V. 2010

Abstract Safflower (*Carthamus tinctorius* L.) DNA marker resources are currently very limited. The objective of this study was to determine the feasibility of transferring non-genic microsatellite (SSR) markers and gene-based markers, including intron fragment length polymorphism (IFLP) and resistance gene candidates (RGC)-based markers from sunflower (*Helianthus annuus* L.) to safflower, both species belonging to the Asteraceae family. Cross-species amplification of 119 non-genic SSRs, 48 IFLPs, and 19 RGC-based sunflower markers in 22 lines and germplasm accessions of safflower was evaluated. Additionally, 69 EST-SSR markers previously reported to amplify in safflower were tested. The results showed that 17.6% of the non-genic SSR, 56.2% of the IFLP, and 73.7% of the RGC-based markers were transferable to safflower. The percentage of transferable markers showing polymorphic loci was 66.6% for non-genic SSR, 70.6% for EST-SSR, 55.5% for IFLP, and 71.4% for RGC-based markers. The highest polymorphism levels were found for non-genic SSR. The average number of alleles per polymorphic locus and mean heterozygosity values were 2.9 and 0.46, respectively, for non-genic SSR, 2.2 and 0.35 for EST-SSR, 2.1 and 0.24 for IFLP, and 2.0 and 0.34 for RGC-based markers. The results of this study revealed a low rate of transferability for non-genic SSR sunflower markers and a better rate of transferability for IFLP and RGC-based markers. Transferable genic and non-genic sunflower markers can have utility for trait and comparative mapping studies in safflower.

Keywords Cross-species marker transferability · *Helianthus annuus* L. · *Carthamus tinctorius* L. · Microsatellites · Gene-based markers · Molecular markers

Introduction

Safflower is a minor crop currently regarded as a promising alternative for oilseed production in many areas of the world. Its adaptation to a wide range of environments, the development of hybrids with larger seed and oil yields, and the production of oil types highly demanded by industry are important factors that are contributing to this promising position (Mündel and Bergman 2008).

Nowadays, breeding programs on the major oilseed crops are based to a large extent on the development and use of molecular tools. However, DNA marker resources available to safflower breeders are very limited. The main molecular markers currently being used in safflower are RAPDs (random amplified

Electronic supplementary material The online version of this article (doi:10.1007/s10681-010-0139-6) contains supplementary material, which is available to authorized users.

M. J. García-Moreno · L. Velasco · B. Pérez-Vich (✉)
Institute for Sustainable Agriculture (CSIC), Alameda del Obispo s/n, 14004 Cordoba, Spain
e-mail: bperez@ias.csic.es

M. J. García-Moreno, J. M. Fernández-Martínez, L. Velasco and B. Pérez-Vich (2011) **Molecular tagging and candidate gene analysis of the high gamma-tocopherol trait in safflower (Carthamus tinctorius L.).** Mol Breeding 28:367–379.

Factor de impacto en 2010

Cites in 2010 to items published in:	2009 =	187	Number of items published in:	2009 =	95
	2008 =	245		2008 =	102
	Sum:	432		Sum:	197
Calculation:	Cites to recent items	432	= **2.193**		
	Number of recent items	197			

Journal Ranking en 2010

For **2010**, the journal **MOLECULAR BREEDING** has an Impact Factor of **2.193**.

This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
AGRONOMY	75	14	Q1
GENETICS & HEREDITY	156	96	Q3
HORTICULTURE	30	5	Q1
PLANT SCIENCES	188	52	Q2

Mol Breeding (2011) 28:367–379
DOI 10.1007/s11032-010-9489-y

Molecular tagging and candidate gene analysis of the high gamma-tocopherol trait in safflower (*Carthamus tinctorius* L.)

M. J. García-Moreno · J. M. Fernández-Martínez · L. Velasco · B. Pérez-Vich

Received: 15 April 2010 / Accepted: 23 July 2010 / Published online: 7 August 2010
© Springer Science+Business Media B.V. 2010

Abstract Genetic control of the synthesis of high gamma-tocopherol (gamma-T) content in the seed oil of safflower (*Carthamus tinctorius* L.) and development of highly reliable molecular markers for this trait were determined through molecular tagging and candidate gene approaches. An F2 population was developed by crossing the high gamma-T natural mutant IASC-1 with the CL-1 line (standard, high alpha-T profile). This population segregated for the partially recessive gene *Tph2*. Bulked segregant analysis with random amplified polymorphic DNA (RAPD) and microsatellite (SSR) markers revealed linkage of eight RAPD and one SSR marker loci to the *Tph2* gene and allowed the construction of a *Tph2* linkage map. RAPD fragments closest to the *Tph2* gene were transformed into sequence-characterized amplified region markers. A gamma-T methyltransferase (gamma-TMT) locus was found to co-segregate with *Tph2*. The locus/band was isolated, cloned and sequenced and it was confirmed as a gamma-TMT gene. A longer partial genomic DNA sequence from this gene was obtained. IASC-1 and CL-1 sequence alignment showed one non-synonymous and two synonymous nucleotide mutations. Intron fragment length polymorphism and insertion-deletion markers based on the gamma-TMT sequence diagnostic for the *Tph2* mutation were developed and tested across 22 safflower accessions, cultivars, and breeding lines. The results from this study provide strong support for the role of the gamma-TMT gene in determining high gamma-T content in safflower and will assist introgression of *tph2* alleles into elite safflower lines to develop varieties with improved tocopherol composition for specific market niches.

Keywords Candidate gene · *Carthamus tinctorius* · Gamma-tocopherol · Gamma-tocopherol methyltransferase · Safflower · SCAR

Introduction

Tocopherols are the main naturally occurring substances with antioxidant activity in oil seeds and derived seed oils. They exist as four forms named alpha-, beta-, gamma-, and delta-tocopherol (alpha-, beta-, gamma-, and delta-T). Their antioxidant protective action is exhibited both in biological systems (in vivo or vitamin E activity) as well as in oils and oil-based products (in vitro antioxidant activity) (Kamal-Eldin and Appelqvist 1996). Alpha-T exerts the maximum in vivo antioxidant activity (Trabber and Sies 1996), but it shows the weakest antioxidant efficiency in vitro. Conversely, the other tocopherol

M. J. García-Moreno · J. M. Fernández-Martínez · L. Velasco · B. Pérez-Vich (✉)
Instituto de Agricultura Sostenible (CSIC),
Apartado 4084, 14080 Córdoba, Spain
e-mail: bperez@ias.csic.es

María J García-Moreno, José M Fernández-Martínez, Leonardo Velasco, Begoña Pérez-Vich (2011) **Genetic basis of unstable expression of high gammatocopherol content in sunflower seeds.** BCM PLANT BIOLOGY, *In review*.

Factor de impacto en 2010

Cites in 2010 to items published in:	2009 = 448		Number of items published in:	2009 = 152	
	2008 = 712			2008 = 132	
	Sum: 1160			Sum: 284	
Calculation:	Cites to recent items	1160 = **4.085**			
	Number of recent items	284			

Journal Ranking en 2010

For **2010**, the journal **BMC PLANT BIOLOGY** has an Impact Factor of **4.085**.

This table shows the ranking of this journal in its subject categories based on Impact Factor.

Category Name	Total Journals in Category	Journal Rank in Category	Quartile in Category
PLANT SCIENCES	188	17	Q1

Genetic basis of unstable expression of high gamma-tocopherol content in sunflower seeds

María J García-Moreno, José M Fernández-Martínez, Leonardo Velasco, Begoña Pérez-Vich[§]

Instituto de Agricultura Sostenible (IAS-CSIC), Alameda del Obispo s/n, 14004 Córdoba, Spain

[§]Corresponding author

E-mail addresses:

MJG-M: mjgarcia@ias.csic.es

JMF-M: jfernandezm@ias.csic.es

LV: lvelasco@ias.csic.es

BP-V: bperez@ias.csic.es

ACEPTACIÓN POR ESCRITO DE LOS COAUTORES DE LOS TRABAJOS

MINISTERIO
DE CIENCIA
E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
IAS
INSTITUTO DE AGRICULTURA SOSTENIBLE

Dña. Elsa María Vera Ruiz, coautora del artículo titulado "Genetic and Molecular Analysis of High Gamma-Tocopherol Content in Sunflower" (remitido a "Crop Science"), doy mi total consentimiento a Dña. María José García-Moreno Pérez para que sea incluido como parte de su tesis doctoral.

Y para que conste, firmo la presente en La Laguna a 07 de noviembre de dos mil once.

Elsa María Vera Ruiz

ALAMEDA DEL OBISPO, S/N
APDO. 4084 14080 CORDOBA
TEL.: 957 499200
FAX.: 957 499252

MINISTERIO
DE CIENCIA
E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
IAS.
INSTITUTO DE AGRICULTURA SOSTENIBLE

Dr. José María Fernández Martínez, Profesor de Investigación del Instituto de Agricultura Sostenible (IAS-CSIC), coautor del artículo titulado "Genetic and molecular analysis of high gamma-tocopherol content in sunflower" (remitido a "Crop Science"), doy mi total consentimiento a Dña. María José García-Moreno Pérez para que sea incluido como parte de su tesis doctoral.

Y para que conste, firmo la presente en Córdoba a 6 de Febrero de dos mil doce.

José María Fernández Martínez

ALAMEDA DEL OBISPO, S/N
APDO. 4084 14080 CORDOBA
TEL.: 957 499200
FAX.: 957 499252

MINISTERIO
DE CIENCIA
E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
IAS.
INSTITUTO DE AGRICULTURA SOSTENIBLE

Dr. José María Fernández Martínez, Profesor de Investigación del Instituto de Agricultura Sostenible (IAS-CSIC), coautor del artículo titulado "Molecular tagging and candidate gene analysis of the high gamma-tocopherol in safflower (*Carthamus tinctorius*)" (remitido a "Molecular Breeding"), doy mi total consentimiento a Dña. María José García-Moreno Pérez para que sea incluido como parte de su tesis doctoral.

Y para que conste, firmo la presente en Córdoba a seis de Febrero de dos mil doce.

José María Fernández Martínez

ALAMEDA DEL OBISPO, S/N
APDO. 4084 14080 CORDOBA
TEL.: 957 499200
FAX.: 957 499252

MINISTERIO
DE CIENCIA
E INNOVACIÓN

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
IAS
INSTITUTO DE AGRICULTURA SOSTENIBLE

Dr. José María Fernández Martínez, Profesor de Investigación del Instituto de Agricultura Sostenible (IAS-CSIC), coautor del artículo titulado "Genetic basis of unstable expresión of high gamma-tocopherol content in sunflower seeds" (remitido a "BCM Plant Biology"), doy mi total consentimiento a Dña. María José García-Moreno Pérez para que sea incluido como parte de su tesis doctoral.

Y para que conste, firmo la presente en Córdoba a seis de Febrero de dos mil doce.

José María Fernández Martínez

ALAMEDA DEL OBISPO, S/N
APDO. 4084 14080 CORDOBA
TEL.: 957 499200
FAX.: 957 499252