



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Detección de una translocación cromosómica en pepino dulce (*Solanum muricatum*) mediante el empleo de marcadores moleculares

TRABAJO FINAL DE GRADO

GRADO EN BIOTECNOLOGIA

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Autor: Miguel López Aparicio
Tutor académico: Santiago Vilanova Navarro
Director experimental: Francisco Javier Herraiz García
Valencia, noviembre de 2020

RESUMEN

Título: Detección de una translocación cromosómica en pepino dulce (*Solanum muricatum*) mediante el empleo de marcadores moleculares.

El pepino dulce (*Solanum muricatum*) es un cultivo de procedencia andina cultivado por sus frutos, los cuales son apreciados por su sabor dulce, su aroma, así como por sus propiedades bioactivas (Herraiz et al., 2016b). Aunque filogenéticamente es una especie muy cercana tanto al tomate como a la patata, no ha sido tan estudiada a nivel genómico debido a su menor importancia económica. En este sentido, el grupo de mejora genética de solanáceas del COMAV, entre otros trabajos, llevó a cabo la secuenciación del transcriptoma de pepino dulce (Herraiz et al., 2016a) y un mapa genético de alta resolución mediante 1625 marcadores SNP obtenidos por GBS (Enshire et al., 2011). Este mapa genético determinó la presencia de 12 grupos de ligamiento, y su comparación con genomas de otras especies próximas sacó a la luz varias anomalías cromosómicas. Una de estas anomalías es una translocación no presente en ninguna otra solanácea, la cual requería de su confirmación. Este es el objetivo del presente trabajo final de grado. Para ello se sembraron semillas obtenidas por autofecundación de una variedad de pepino dulce con el fin de obtener una población de plantas de las que se extrajo ADN. A partir de secuencias del transcriptoma, se seleccionaron varios marcadores SNPs situados en los fragmentos implicados en la translocación. Estos marcadores fueron seleccionados con la condición de ser heterocigotos en pepino dulce. El estudio de la segregación de dichos marcadores y su frecuencia de recombinación o transmisión independiente nos permitirá confirmar o descartar la presencia de dicha translocación en el genoma de esta especie. Para ello se realizó un ensayo HRM (Li et al., 2010) cuyos resultados fueron analizados en un estudio de ligamiento genético.

Palabras clave: Mapas genéticos, pepino dulce, *Solanum muricatum*, HRM, SNPs, translocación.

Alumno: Miguel López Aparicio

Valencia, noviembre 2020

Prof. D. Santiago Vilanova Navarro

D. Francisco Javier Herraiz García

ABSTRACT

Title: Detection of a chromosomal translocation in sweet cucumber (*Solanum muricatum*) using molecular markers.

The sweet cucumber (*Solanum muricatum*) is a crop of Andean origin grown for its fruits, which are appreciated for their sweet taste, aroma, as well as for their bioactive properties (Herraiz et al., 2016b). Although phylogenetically it is a species very close to both tomato and potato, it has not been as well studied at the genomic level due to its lesser economic importance. In this sense, the Solanaceae genetic improvement group of COMAV, among other works, carried out the sequencing of the sweet cucumber transcriptome (Herraiz et al., 2016a) and a high-resolution genetic map using 1625 SNP markers obtained by GBS. (Enshire et al., 2011). This genetic map determined the presence of 12 linkage groups, and its comparison with genomes of other nearby species brought to light several chromosomal abnormalities. One of these anomalies is a translocation not present in any other nightshade, which required its confirmation. This is the objective of this final degree project. For this, seeds obtained by self-fertilization of a variety of sweet cucumber were sown in order to obtain a population of plants from which DNA was extracted. From transcriptome sequences, several SNP markers located in the fragments involved in the translocation were selected. These markers were selected with the condition of being heterozygous in sweet cucumber. The study of the segregation of these markers and their frequency of recombination or independent transmission will allow us to confirm or rule out the presence of these translocation in the genome of this species. For this, an HRM trial was carried out (Li et al., 2010) whose results were analyzed in a genetic linkage study.

Keywords: Genetic maps, sweet cucumber, *Solanum muricatum*, HRM, SNPs, translocation.

Student: Miguel López Aparicio

Valencia, November 2020

Prof. Mr. Santiago Vilanova Navarro

Mr. Francisco Javier Herraiz García

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1. El pepino dulce.....	1
1.2. Taxonomía.....	2
1.3. Valor nutricional y nutracéutico.....	2
1.4. La biotecnología y la mejora genética de plantas.....	3
1.5. Introducción a los estudios de ligamiento y a los mapas genéticos.....	4
1.6. Introducción al genotipado por secuenciación.....	5
1.7. Introducción al análisis de alta resolución de fusión.....	6
1.8. Marco general del proyecto.....	7
2. OBJETIVOS.....	8
3. MATERIALES Y MÉTODOS.....	8
3.1. Diseño del experimento.....	8
3.2. Material vegetal.....	9
3.3. Toma de muestras y extracción de ADN.....	10
3.4. Electroforesis en gel de agarosa.....	11
3.5. Análisis de alta resolución de fusión.....	11
3.6. Posicionamiento de los marcadores.....	12
3.7. Estudio de ligamiento y mapa genético.....	12
4. RESULTADOS.....	12
4.1. Extracción de ADN.....	12
4.2. Posicionamiento de los marcadores.....	14
4.3. Estudio de ligamiento y mapa genético.....	18
5. CONCLUSIONES.....	20
6. BIBLIOGRAFÍA.....	21
7. ANEXO.....	25

ÍNDICE DE FIGURAS

Figura 1. Frutos de una variedad de pepino dulce.....	1
Figura 2. Pasos en la construcción de la biblioteca GBS.....	6
Figura 3. Genotipado de plantas F2 de arroz mediante HRM con marcadores SNP.....	7
Figura 4. Esquema de los cruzamientos realizados para la realización del mapa genético de pepino dulce.....	7
Figura 5. Esquema de los cruzamientos realizados para la obtención de las plantas con las que se realizó el trabajo.....	9
Figura 6. Fotografía del proceso de germinación de las semillas.....	9
Figura 7. Fotografías del proceso de crecimiento de las plantas.....	10
Figura 8. Gel de agarosa al 0,8% para la evaluación visual de la integridad del ADN extraído.....	14
Figura 9. Alineamiento de la secuencia que contiene al marcador 20530 contra la base de datos de cromosomas de tomate (versión SL4.0).....	14
Figura 10. Alineamiento de la secuencia que contiene al marcador 28042 contra la base de datos de cromosomas de tomate (versión SL4.0).....	15
Figura 11. Alineamiento de la secuencia que contiene al marcador 30276 contra la base de datos de cromosomas de tomate (versión SL4.0).....	15
Figura 12. Alineamiento de la secuencia que contiene al marcador 26784 contra la base de datos de cromosomas de tomate (versión SL4.0).....	16
Figura 13. Alineamiento de la secuencia que contiene al marcador 26025 contra la base de datos de cromosomas de tomate (versión SL4.0).....	16
Figura 14. Alineamiento de la secuencia que contiene al marcador 28075 contra la base de datos de cromosomas de tomate (versión SL4.0).....	17
Figura 15. Alineamiento de la secuencia que contiene al marcador 20530 contra la base de datos de cromosomas de tomate (versión SL4.0).....	17
Figura 16. Genotipado de plantas de pepino dulce mediante HRM con marcadores SNP.....	19
Figura 17. Mapa genético de pepino dulce.....	19

ÍNDICE DE TABLAS

Tabla 1. Clasificación taxonómica de pepino dulce.....	2
Tabla 2. Valor nutricional de tres cultivos de pepino dulce.....	3
Tabla 3. Composición del tampón de carga para electroforesis en gel de agarosa.....	11
Tabla 4. Composición de carga de la placa de HRM por cada muestra.....	12
Tabla 5. Valores obtenidos con el equipo NanoDrop™ ND-1000 (NanoDrop™ Technologies, Wilmington, Delaware, EEUU).....	13
Tabla 6. Posicionamiento físico de los marcadores estudiados respecto al genoma de tomate (versión SL4.0).....	18
Tabla 7. Secuencias, tamaños y temperaturas de fusión de los cebadores diseñados mediante el software Primer3 (v. 0.4.0) (Rozen, 2007).....	18
Tabla 8. Se muestran los marcadores estudiados (filas) agrupados según los bloques de ligamiento obtenidos, la posición de los diferentes marcadores en el genoma de tomate (versión SL4.0) y los datos de genotipado de las muestras (columnas).....	19

ABREVIATURAS

ADN	Ácido desoxirribonucleico.
AFLP	<i>Amplified fragment length polymorphism.</i>
BLAST	<i>Basic local alignment search tool.</i>
CAPS	<i>Cleaved amplified polymorphic sequences.</i>
COMAV	Instituto de Conservación y Mejora de la Agrodiversidad Valenciana.
CTAB	<i>Cetyl Trimethyl Ammonium Bromide</i>
DALP	<i>Direct amplification of length polymorphisms.</i>
GBS	<i>Genotyping by sequencing.</i>
HRM	<i>High resolution melting.</i>
LOD	<i>Logarithm of the odds.</i>
MAS	<i>Marker assisted selection.</i>
NGS	<i>Next generation sequencing.</i>
PCR	<i>Polymerase chain reaction.</i>
RADP	<i>Randomly amplified polymorphic DNA.</i>
RFLP	<i>Restriction fragment length polymorphism.</i>
SCAR	<i>Sequence characterized amplified region.</i>
SNPs	<i>Single nucleotide polymorphisms.</i>
SSR	<i>Simple sequence repeat.</i>
TAE	Tampon tris, acetato y EDTA.
TE	Tampon tris y EDTA.
Tm	Temperatura de <i>melting</i> o temperatura de fusión

1. INTRODUCCIÓN

1.1. EL PEPINO DULCE

El pepino dulce (*Solanum muricatum* Aiton) también conocido como pera melón es una especie vegetal fanerógama perteneciente a la familia de las solanáceas.

Originaria del continente sudamericano, más concretamente de la región andina, se cultiva en países, como Colombia, Bolivia, Perú y Ecuador (Nuez y Ruiz, 1996). El interés del cultivo reside en el fruto, una baya refrescante de sabor agradable, que presenta una gran variación tanto en tamaño como en forma y color (Herráiz, 2015) (Figura 1).



Figura 1. Frutos de una variedad de pepino dulce.

Si se recoge maduro puede usarse como postre o para la elaboración de compotas o mermeladas, mientras que si se recoge verde puede consumirse en ensalada (Nuez y Ruiz, 1996).

Por poseer multitud de características atractivas el pepino dulce se ofrece como posible nuevo cultivo prometedor en Europa. Ejemplos de ello son, su sabor o calidad organoléptica (Nuez y Ruiz, 1996) y su valor nutraceútico (Hsu et al., 2011; Shathish & Guruvayoorappan, 2014).

1.2. TAXONOMÍA

Como ya se ha dicho, el pepino dulce pertenece a la familia de las solanáceas, y dentro de ella al género *Solanum*. A este mismo género también pertenecen cultivos de gran importancia económica como la patata, la berenjena y el tomate. A continuación se recoge la clasificación taxonómica de *S. muricatum* (Nuez y Ruiz, 1996) (Tabla 1).

Tabla 1. Clasificación taxonómica de pepino dulce.

Reino	<i>Plantae</i>
División	<i>Magnoliophyta</i>
Clase	<i>Magnoliopsida</i>
Orden	<i>Solanales</i>
Familia	<i>Solanaceae</i>
Subfamilia	<i>Solanoideae</i>
Tribu	<i>Solaneae</i>
Género	<i>Solanum</i>
Subgenero	<i>Potatoe</i>
Sección	<i>Basarthrum</i>
Serie	<i>Muricata</i>

Hablando en términos evolutivos el pepino dulce está estrechamente emparentado con el tomate y la patata, habiéndose sugerido incluso que forman un grupo monofilético que podría establecer un género aparte (Lester, 1991).

1.3. VALOR NUTRICIONAL Y NUTRACÉUTICO

El fruto del pepino dulce tiene alto contenido en agua (>92%) y bajo contenido calórico (250 kcal/kg) (Rodríguez-Burruezo et al., 2011). Se trata de un alimento de alto valor nutracéutico ya que posee alto contenido en potasio y vitamina C (Hsu et al., 2011). Además existen varios estudios acerca de las propiedades beneficiosas que este fruto tiene sobre la salud como por ejemplo su efecto antitumoral (Shathish & Guruvayoorappan, 2014), antioxidante (Sudha et al., 2011) y antidiabético (Hsu et al., 2011).

A continuación se recoge la información nutricional de tres cultivos de pepino dulce en diferentes países: Nueva Zelanda (Redgwell & Turner, 1986), Perú (Delgado de la Flor, 1988) y Chile (Bravo y Arias, 1983) (Tabla 2).

Tabla 2. Valor nutricional de tres cultivos de pepino dulce.

Componente (Valores por 100 g)	Nueva Zelanda	Perú	Chile
Calorías (kcal)	-	26,0	25,0
Agua (g)	91,8-93,2	92,3	92,4
Fibra (g)	-	0,5	0,5
Proteínas (g)	0,10-0,13	0,30	0,40
Lípidos y pigmentos (mg)	25-44	-	100
Azúcares solubles (g)	4,9-6,4	7,0	6,3
Almidón (mg)	20,0-90,0	-	-
Vitamina A (mg)	-	317	-
Vitamina C (mg)	46,0-68,8	29,7	26,0
Ácidos orgánicos no volátiles (mg)	119-153	-	-

El aroma del pepino dulce se debe a una compleja mezcla de compuestos volátiles, los diferentes grupos varietales pueden clasificarse en función de su aroma predominante en tres grupos (Rodríguez-Burruezo et al., 2004a):

- Aroma a fruta madura: constituido por acetatos y prenol.
- Aroma a vegetal verde: constituido por aldehídos C6 y C9.
- Aroma a fruta exótica: constituido por lactonas, mesifurano y β -damascenona.

1.4. LA BIOTECNOLOGÍA Y LA MEJORA GENÉTICA DE PLANTAS

Tradicionalmente, la mejora genética de plantas ha sido llevada a cabo realizando cruzamientos entre individuos seleccionados para producir nuevos cultivares con rasgos deseables, esto se conoce como mejora genética de plantas (mejora vegetal) tradicional o mejora mediante métodos convencionales. Largos períodos de tiempo y varias generaciones son necesarios para seleccionar y evaluar genotipos útiles mediante los métodos tradicionales. Gracias a la aplicación de las herramientas proporcionadas por la biotecnología y la biología molecular, en las últimas décadas, han sido desarrolladas multitud de técnicas novedosas para obtener nuevos cultivares y mejorar los ya existentes (He et al., 2014).

La selección asistida por marcadores o MAS (por sus siglas en inglés "*marker-assisted selection*") es un proceso mediante el cual los marcadores moleculares se utilizan para la selección indirecta de rasgos de interés en los cultivos. Este método ha sido ampliamente aplicado de manera efectiva para mejorar la calidad, el rendimiento y la tolerancia al estrés biótico o abiótico de diversos cultivos. Otra de las técnicas empleadas frecuentemente para la mejora de plantas es el genotipado por secuenciación o GBS (por sus siglas en inglés "*genotyping-by-sequencing*"). Esta técnica ofrece una herramienta definitiva a este método de selección para acelerar el proceso de mejora de plantas y cultivos (He et al., 2014). Múltiples tipos de marcadores moleculares basados en PCR han sido utilizados para la mejora genética de especies, ejemplos de ello son, RFLP (Botstein et al., 1980), RAPD (Williams et al., 1990), SCAR (Paran & Michelmore, 1993), CAPS (Konieczny Y Ausubel, 1993), SSR (Salimath et al., 1995), AFLP (Vos et al., 1995) y DALP (Desmarais et al., 1998).

Entre la amplia gama de marcadores moleculares existentes, los SNPs han demostrado ser los más abundantes y los más adecuados para el análisis de genomas de diversos tamaños (Rafalski, 2002). Multitud de esfuerzos se han centrado en el desarrollo de técnicas y plataformas que permitan identificar un gran número de sitios polimórficos en los genomas en tiempos relativamente cortos. En los últimos 15 años, las técnicas de marcadores basadas en SNPs han sentado las bases del genotipado de alto rendimiento, reduciendo considerablemente los costes en comparación con los enfoques de genotipado anteriores (He et al., 2014).

La creciente demanda de plataformas que permitan la secuenciación de ADN a un bajo coste ha impulsado el desarrollo de tecnologías de alto rendimiento también conocidas como tecnologías de secuenciación masiva (NGS, por sus siglas en inglés “*next-generation sequencing*”). Han sido muchas las plataformas de secuenciación desarrolladas a lo largo de estos últimos años. Algunas de ellas son, Roche 454 FLX™ (Rothberg et al., 2008), Illumina MiSeq® o Ion Torrent™ Personal Genome Machine™ (PGM) (Quail et al., 2012). Un punto en común en la mayoría de estas tecnologías NGS es la división de forma aleatoria el ADN genómico en tamaños más pequeños que son inmovilizados en una superficie sólida. Millones de moléculas de ADN amplificadas en paralelo son inmovilizadas en una superficie sintética para generar hasta varios miles de millones de secuencias. La secuenciación se realiza de manera iterativa, donde la incorporación de los nucleótidos es monitorizada por la emisión de una señal (puede ser de diferente naturaleza) y su detección por el secuenciador (Metzker, 2010).

1.5 INTRODUCCIÓN A LOS ESTUDIOS DE LIGAMIENTO Y MAPAS GENÉTICOS

Por definición dos loci están ligados cuando existe entre ellos una asociación física, es decir, estos loci se encuentran cercanos en una misma hebra de ADN y por lo tanto la frecuencia de recombinación entre los mismos es baja dándose una mayor probabilidad de herencia conjunta (William et al., 2006).

Es sabido que durante la formación de gametos en el proceso de división celular que conocemos como meiosis los cromosomas homólogos se emparejan e intercambian material genético, este proceso es denominado recombinación genética. El resultado del mismo es un entrecruzamiento o “barajado” de los alelos. Por lo tanto, podemos afirmar que el proceso de recombinación es una importante característica para la generación de variabilidad en la reproducción sexual. Mediante cruces entre individuos genéticamente definidos, determinando las frecuencias con las que la descendencia porta rasgos diferentes a los parentales, es posible medir la frecuencia a la que se produce la recombinación entre dos loci del mismo cromosoma. A mayor distancia genética entre los loci, mayor frecuencia de recombinación. El proceso de estimación de dicha distancia genética es conocido como mapeo (William et al., 2006).

Múltiples avances han sido introducidos en el proceso de mapeo genético desde que Morgan y Bridges publicaron el primer mapa de ligamiento del cromosoma X de *Drosophila melanogaster* allá por 1916 (Bridges & Morgan, 2016). Uno de ellos es la utilización de funciones de mapeo (ajustes matemáticos en la distancia genética calculada en base a las frecuencias de recombinación). Estos ajustes son necesarios

debido a diferentes fenómenos que provocan que la relación entre las frecuencias de recombinación y la distancia genética se desvíe de la linealidad, como por ejemplo eventos de recombinación múltiples o el fenómeno conocido como interferencia de entrecruzamientos (la ocurrencia de un evento de recombinación hace menos probable que ocurra otro cerca en el mismo cromosoma). Algunas de estas funciones de mapeo son la función de mapeo Haldane (Haldane, 1919) y la función de mapeo de Kosambi (Kosambi, 1944).

Entre las numerosas aplicaciones de los mapas genéticos cabe mencionar las siguientes, identificación de genotipos específicos, estudios de evolución de genomas y establecimiento de homologías cromosómicas entre especies (estudios de conservación de la sintenia) (Benavides y Guénet, 2003).

En la actualidad existen multitud de programas informáticos desarrollados para procesar los datos obtenidos de cruces experimentales y el posterior análisis de los marcadores en la realización de un estudio de ligamiento. Ejemplos de ello son JoinMap® (Stam, 1993) y el paquete de R ASmap (Taylor & Butler, 2017), programas que fueron utilizados en la elaboración del mapa de ligamiento de pepino dulce que dio lugar a la realización de este trabajo.

1.6. INTRODUCCIÓN AL GENOTIPADO POR SECUENCIACIÓN

El genotipado por secuenciación o GBS (por sus siglas en inglés “genotyping-by-sequencing”) hace posible la secuenciación del genoma de especies con genomas grandes y de alta variabilidad genética de una forma relativamente barata. El procedimiento se basa en la reducción de la complejidad del genoma mediante el uso de enzimas de restricción. Tras someterse al proceso de digestión por los enzimas, estos fragmentos son etiquetados y secuenciados mediante tecnologías de nueva generación. Este método rápido, simple, altamente específico y reproducible, es adecuado para estudios de poblaciones, caracterización de germoplasma, mapeo de rasgos y mejora genética de especies (Elshire et al., 2011).

Utilizando enzimas de restricción sensibles a la metilación pueden evitarse las regiones repetitivas de los genomas y dirigir el análisis a las regiones de copia más baja con una mayor eficiencia, simplificando en gran medida los problemas de alineación en especies con gran diversidad genética. Es crítica la selección de un enzima que no tenga alta frecuencia de corte en la fracción repetitiva del genoma. Una de las ventajas de este enfoque es que solo es necesario desarrollar un mapa de referencia alrededor de los sitios de restricción, realizándose el mismo durante el proceso de genotipado de la muestra. Por lo tanto no es necesario disponer de la secuencia completa del genoma de la especie previamente. En estos casos, el consenso de los clusters de lecturas en los sitios etiquetados de la secuencia se convierte en la referencia (Elshire et al., 2011) (Figura 2).

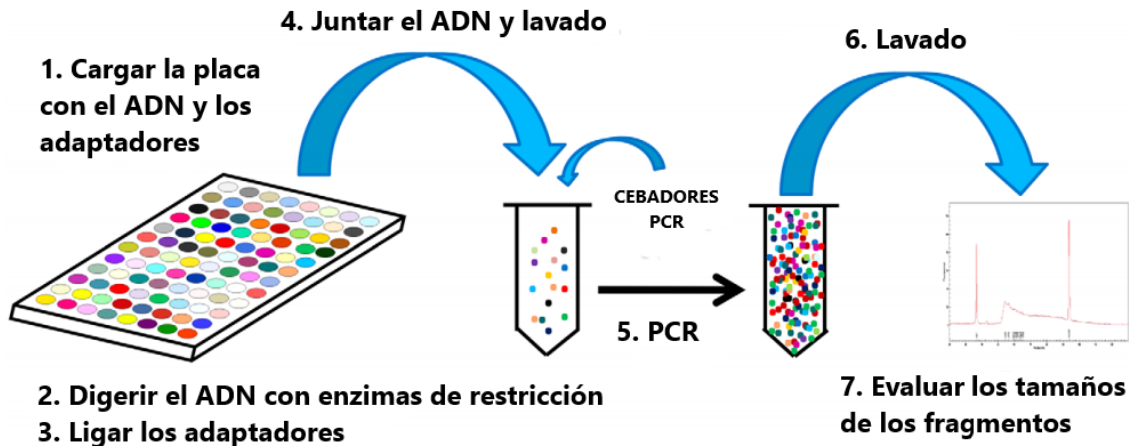


Figura 2. Pasos en la construcción de la biblioteca GBS. Hasta 96 muestras de ADN pueden ser procesadas simultáneamente. (1) las muestras de ADN, el código de barras y los pares de adaptadores se cargan en la placa y se secan; (2–3) las muestras se digieren con enzimas de restricción y los adaptadores se ligan a los extremos de los fragmentos de ADN genómico; (4) la ligasa T4 se inactiva por calentamiento, se agrupa una parte alícuota de cada muestra y se aplica a una columna de exclusión por tamaño para eliminar los adaptadores sin reaccionar; (5) se añaden cebadores apropiados con sitios de unión en los adaptadores ligados y se realiza PCR para aumentar el conjunto de fragmentos; (6–7) Los productos de PCR se lavan y los tamaños de fragmentos de la biblioteca resultante se verifican en un analizador de ADN (BioRad Experion® o un instrumento similar). Las bibliotecas sin dímeros adaptadores se retienen para la secuenciación del ADN (Elshire et al., 2011).

1.7. INTRODUCCIÓN AL ANÁLISIS DE ALTA RESOLUCIÓN DE FUSIÓN

El análisis de alta resolución de fusión o HRM (por sus siglas en inglés “*high-resolution melting*”) es una técnica de biología molecular utilizada para la detección de polimorfismos, mutaciones y diferencias epigenéticas en muestras de ADN de doble cadena (Migheli et al., 2013; Wittwer et al., 2003).

Las ventajas del HRM incluyen entre otras, mayor simplicidad, alta eficiencia y ahorro sustancial en costes y tiempo de trabajo. Desde que esta tecnología fue desarrollada por la compañía Idaho Technology Inc. y la Universidad de Utah en 2003 ha sido progresivamente aplicada en diferentes campos como la patología humana (Dagar et al., 2009) o la biología vegetal (Sun et al., 2010).

Esta técnica se basa en el análisis y comparación de curvas de fusión para discernir pequeñas diferencias en la secuencia de ADN de amplicones obtenidos mediante PCR. La temperatura de fusión del amplicón (temperatura a la cual las dos hebras de ADN se separan) depende de su secuencia.

La realización del ensayo es sencilla, la muestra se somete a un calentamiento mientras se monitoriza la temperatura y fluorescencia de la misma. Para ello es imprescindible el uso de marcadores fluorescentes que se unan al ADN de doble cadena haciendo posible su detección. De este modo estamos determinando el grado de hibridación pues está directamente relacionado con la fluorescencia, a mayor fluorescencia mayor hibridación. Esta técnica permite detectar diferencias de secuencia de una sola base y por ello se emplea para genotipado de SNPs (Figura 3).

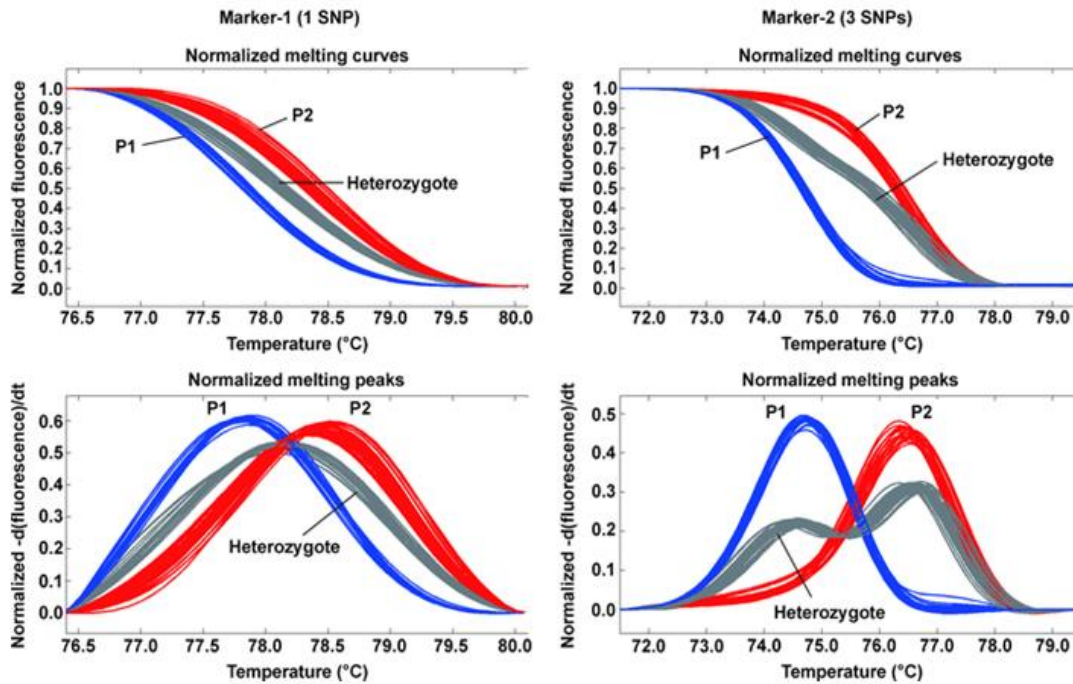


Figura 3. Genotipado de plantas F2 de arroz mediante HRM con marcadores SNP (variaciones G/C>T/A). En la parte de arriba se muestran las curvas de fusión normalizadas de los amplicones de 96 plantas F2, en las que se identifican los genotipos P1, P2, y heterocigoto. En la parte de abajo se muestran los picos de fusión, incremento negativo de la fluorescencia respecto a la temperatura (Li et al., 2010).

1.8. MARCO GENERAL DEL PROYECTO

Este trabajo se enmarca dentro del proyecto de secuenciación del transcriptoma y elaboración de un mapa genético de alta resolución de *S. muricatum* llevado a cabo por el grupo de mejora genética de solanáceas del COMAV.

Para obtener dicho mapa genético se realizó un cruce entre una variedad de *S. muricatum* (Sweet Long) y una especie silvestre *S. caripense* (accesión EC-40.9). Esta especie se ha propuesto como uno de los posibles ancestros del cual (o los cuales) se domesticó el pepino dulce (Blanca et al. 2007). A partir de ese cruzamiento fue obtenida una descendencia (F1) que resultó no ser homogénea ya que las variedades empleadas en el cruzamiento son muy heterocigotas. Una de las plantas de esta descendencia fue seleccionada y autofecundada para obtener la siguiente generación (F2). Se realizó una extracción de

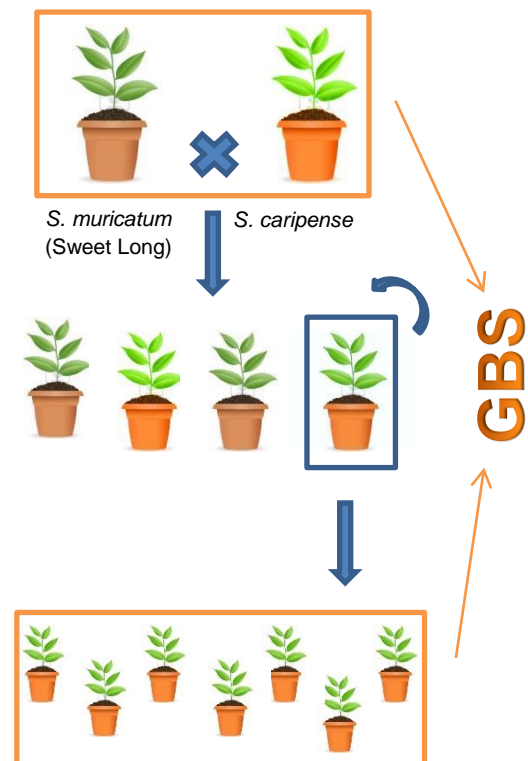


Figura 4. Esquema de los cruzamientos realizados para la realización del mapa genético de pepino dulce.

ADN de las plantas de esta F2 y las de sus parentales (*S. muricatum* y *S. caripense*). Este ADN se envió a una empresa de secuenciación (Macrogen, Inc.) donde fue analizado mediante un ensayo GBS que proporcionó un total de 1625 marcadores SNP (Figura 4). Seguidamente los datos fueron procesados mediante programas de mapeo (JoinMap® (Stam, 1993) y ASmap (Taylor & Butler, 2017)).

Este mapa genético, determinó la presencia de doce grupos de ligamiento y su comparación con genomas de otras especies próximas, sacó a la luz varias anomalías cromosómicas. Una de estas anomalías es una translocación no presente en ninguna otra solanácea, la cual requería de su confirmación, siendo este el objeto del presente trabajo final de grado. Para ello se dispuso de una población segregante de pepino dulce y se evaluó la segregación de marcadores localizados en la translocación y en regiones adyacentes.

2. OBJETIVOS

El objetivo de este trabajo final de grado es la confirmación de una translocación cromosómica en pepino dulce mediante el empleo de marcadores moleculares SNPs. Detectada gracias al mapa genético de esta especie, realizado por el grupo de mejora genética de solanáceas del COMAV, esta translocación consiste en una aparente fusión cromosómica. Respecto al genoma de tomate, los cromosomas 1 y 2 parecen estar fusionados en el genoma de pepino dulce. Dicha translocación no ha sido encontrada en ninguna otra especie de la misma familia y por lo tanto ha de ser comprobada.

3. MATERIALES Y MÉTODOS

3.1. DISEÑO DEL EXPERIMENTO

Diecisiete SNPs habían sido previamente seleccionados (consultar anexo) para la realización de este estudio de ligamiento. En principio se abordó el proyecto utilizando marcadores moleculares basados en PCR y digestión por enzimas de restricción (CAPS (Konieczny & Ausubel, 1993)). Para ello fue proporcionada una lista de cebadores para la amplificación de regiones que incluían dichos SNPs y de enzimas de restricción para el análisis de los mismos. Una vez fueron probados los cebadores se descartó esta técnica para realizar el análisis ya que los enzimas necesarios no estaban disponibles en el laboratorio en ese momento. Seguidamente se planteó el análisis HRM como técnica más adecuada para la realización del trabajo. Para ello se diseñaron nuevos cebadores y a continuación se procedió a la realización del ensayo.

3.2. MATERIAL VEGETAL

Como material de partida aproximadamente 90 semillas de *S. muricatum* fueron proporcionadas por el banco de germoplasma del Instituto de Conservación y Mejora de la Agrodiversidad Valenciana (COMAV) (Valencia, España).

Estas semillas se obtuvieron de la autofecundación de la variedad de pepino dulce Valencia que a su vez procede del cruzamiento de la variedad Sweet Long (Sm-29) con otra entrada del banco conocida como Sm-26 (Rodríguez-Burruezo et al., 2004b) (Figura 5).

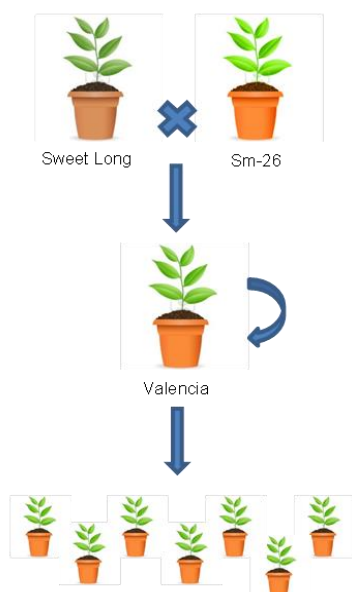


Figura 5. Esquema de los cruzamientos realizados para la obtención de las plantas con las que se realizó el trabajo.

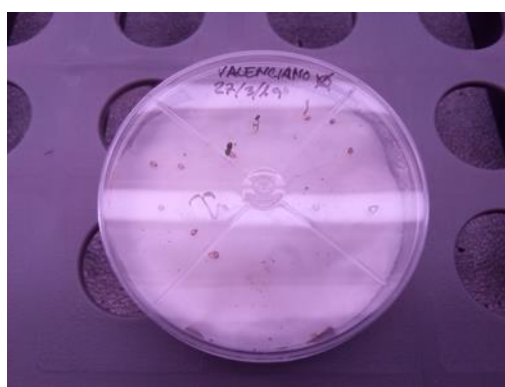


Figura 6. Fotografía del proceso de germinación de las semillas.

Para el germinado, las semillas fueron desinfectadas mediante un lavado de 20 minutos con una solución de hipoclorito de sodio al 30% (preparada previamente). Después se realizaron otros tres lavados con agua, manteniendo las semillas durante 10 minutos en el último lavado. A continuación las semillas fueron distribuidas en tres placas de Petri fabricadas en poliestireno de 90x14 mm en un lecho de algodón humedecido sobre el que se colocó un papel de filtro para evitar que las raíces quedasen enredadas en el algodón y prevenir así posibles daños al retirar las plántulas de la placa. Las placas se mantuvieron en la cámara de cultivo durante 13 días (Figura 6).

Una vez germinadas, las plántulas fueron trasplantadas a macetas de 0,23 litros de capacidad rellenas de sustrato que fueron mantenidas en la cámara de cultivo durante 52 días siendo regadas con agua cada 7 días aproximadamente. Finalmente un total de 43 plantas consiguieron superar el proceso (Figura 7).



Figura 7. Fotografías del proceso de crecimiento de las plantas.

3.3. TOMA DE MUESTRAS Y EXTRACCIÓN DE ADN

Se realizó una extracción de ADN genómico de las hojas de cada una de las 43 plantas y de sus 2 parentales de acuerdo con una modificación (revisada el 20/03/2019 por D. Alonso) del método CTAB (Vilanova et al., 2020; Doyle & Doyle, 1987) tal y como se recoge a continuación:

1. Preparar un tubo Eppendorf de 2 mL con una bola de cristal de 5mm de diámetro en su interior por cada una de nuestras muestras.
2. Con la tapa del Eppendorf obtener dos discos de material vegetal. Congelar en nitrógeno.
3. Triturar las muestras durante 30 segundos.
4. Devolver las muestras al nitrógeno líquido.
5. En la campana de gases añadir a cada muestra 1000 μ L de buffer de extracción y 14 μ L de β -mercaptoetanol. Vortex durante 30 s o hasta completar la homogenización.
6. Añadir 2 μ L de RNasa (1 mg/ml) a cada muestra. Incubar en el termobloque 30 min a 65 $^{\circ}$ C.
7. Mantener en hielo 5 min.
8. Añadir 700 μ L de cloroformo-alcohol isoamílico (24:1).
9. Mezclar bien con el vortex hasta que las dos fases desaparezcan y una especie de nube blanca sea visible (entre 30 s y 1 min).
10. Centrifugar: 11.000 rpm, 5 min.
11. Cuidadosamente recuperar 800 μ L de la fase acuosa sin tocar la interfase y transferirlos a un nuevo tubo Eppendorf de 2 mL.
12. Añadir 1200 μ L de solución de unión. Es crítico seguir el orden correcto. Primero añadir 480 μ L de NaCl 2,5M + 20% PEG y agitar suavemente. Después añadir 720 μ L de etanol absoluto y agitar de nuevo.
13. Añadir 20 μ L de matriz de sílice.
14. Remover en agitador durante 5 min.
15. Centrifugar durante 5 o 6 segundos.
16. Descartar el sobrenadante por decantación y añadir 700 μ L de EtOH 70%.
17. Resuspender completamente el pellet de sílice agitando a mano (no utilizar vortex).
18. Centrifugar durante 5 o 6 segundos.

19. Descartar el sobrenadante por decantación y dejar secar 5 min en el termobloque o hasta que todo el EtOH se haya evaporado.
20. Añadir 50 µL de TE pH 8, mezclar suavemente a mano aproximadamente 30 s e incubar 10 min a 65 °C en el termobloque.
21. Centrifugar: 14.000 rpm, 9 min.
22. Recuperar el sobrenadante y transferirlo a tubos Eppendorf de 1,5 mL nuevos.

Una vez realizada la extracción el DNA fue visualizado para valorar su integridad en un gel de agarosa al 0,8% y su concentración estimada utilizando un espectrofotómetro NanoDrop™ ND-1000 (NanoDrop™ Technologies, Wilmington, Delaware, EEUU) (Consultar anexo). El DNA extraído fue diluido a una concentración de trabajo de 50 ng/µL.

3.4. ELECTROFORESIS EN GEL DE AGAROSA

Las electroforesis se realizaron utilizando gel (26 pocillos) de agarosa al 0,8% en tampón TAE 1X fijando la fuente de corriente a 100 voltios durante 30 minutos. En cada ensayo se utilizó un marcador de pesos moleculares adecuado para 100 pb (FastGene® DNA ladders, NIPPON Genetics®). El tampón de carga por cada muestra se componía de (Tabla 3):

Tabla 3. Composición del tampón de carga para electroforesis en gel de agarosa.

H ₂ O Milli-Q®	6 µL
GelRed®	2 µL
Producto de PCR	2 µL

3.5. ANÁLISIS DE ALTA RESOLUCIÓN DE FUSIÓN

Un equipo LightCycler® 480 (Roche Molecular Systems Inc., Pleasanton, California, EEUU) y el software proporcionado por el fabricante (LightCycler® 480 Software, Version 1.5) fueron utilizados para la realización del ensayo HRM. Los análisis “*Tm Calling*” y “*Gene Scanning*” fueron ejecutados para cada marcador. Las condiciones del programa seleccionado se recogen a continuación:

- Preincubación: Un ciclo, 95 °C 10 min.
- Amplificación: 55 ciclos, 95 °C 10 s, 55 °C 15 s, 72 °C 15 s.
- Fusión: Un ciclo, 95 °C 1 min, 40 °C 1 min, 60 °C 1 s, 95 °C.
- Enfriamiento: Un ciclo, 40 °C 10 min.

El diseño de los cebadores a utilizar en el ensayo se realizó con la aplicación web Primer3 (v. 0.4.0) (Rozen, 2007) fijando el tamaño del producto en un rango de 90 a 120 pares de bases y seleccionando el SNP correspondiente con un corchete, el resto de las opciones se dejaron por defecto.

Se utilizó una placa de 96 pocillos para cada uno de los SNPs a analizar. Optando por una preparación lista para usar proveída por una casa comercial, en este caso MasterMix qPCR No-ROx PyroTaq EvaGreen® 5x (Cultek®, Dutscher group, Brumath, Francia), que contiene todo lo necesario para realizar el ensayo; dNTPs ultrapuros, MgCl₂ y tinción EvaGreen® (Biotium Inc., Fremont, EEUU).

Cada pocillo de la placa se cargó con (Tabla 4):

Tabla 4. Composición de carga de la placa de HRM por cada muestra.

PyroTaq5x	2 µL
Cebador directo [100µM]	0,3 µL
Cebador reverso [100µM]	0,3 µL
DNA [≈50ng/µL]	2 µL
H ₂ O ultrapura	5,4 µL

3.6 POSICIONAMIENTO DE LOS MARCADORES

Se realizó un alineamiento de las secuencias (previamente obtenidas del transcriptoma de pepino dulce, recogidas en el anexo) que incluían a los marcadores estudiados contra la base de datos de cromosomas de tomate (versión SL4.0). Este alineamiento fue llevado a cabo utilizando la herramienta BLAST (Altschul et al., 1990), seleccionando el algoritmo “blastn” con un tamaño de palabra de 11 y un valor de corte de 1e-10.

3.7. ESTUDIO DE LIGAMIENTO Y MAPA GENÉTICO

Los datos de genotipado obtenidos mediante el ensayo HRM fueron analizados con el software JoinMap® (Stam, 1993). Se realizó un estudio de ligamiento y un mapa genético utilizando el algoritmo de alta verosimilitud considerando ligados los marcadores para los que se obtuvo un “LOD score” mayor o igual que tres.

4. RESULTADOS

4.1. EXTRACCIÓN DE ADN

En la siguiente tabla (Tabla 5) se reflejan los valores de concentración y pureza obtenidos con espectrofotómetro NanoDrop™ ND-1000 (NanoDrop™ Technologies, Wilmington, Delaware, EEUU). En algunas muestras se observa un ratio 260/280 ligeramente mayor de 2,1 lo que podría indicar una posible contaminación por ARN. Por otra parte, podemos apreciar algunos valores del ratio 260/230 menores de 1,8 lo que nos indica posible contaminación por sales caotrópicas, fenoles o carbohidratos (Williams et al., 1990).

Tabla 5. Valores obtenidos con el equipo NanoDrop™ ND-1000 (NanoDrop™ Technologies, Wilmington, Delaware, EEUU). Nº de muestra. ng/ul: Valor de concentración de la muestra en ng/ul. A260: Absorbancia de la muestra a 260 nm. A280: Absorbancia de la muestra a 280 nm. A230: Absorbancia de la muestra a 230 nm. 260/280: Ratio de absorbancia 260/280 ($\geq 1,8$ -2,1 se considera ADN de pureza optima). 230/280: Ratio de absorbancia 230/280 ($> 1,8$ se considera ADN de pureza aceptable) (Williams et al., 1990).

Nº Muestra	ng/ul	A260	A280	A230	260/280	260/230
1	53,25	1,065	0,508	0,879	2,1	1,21
2	193,72	3,874	1,816	2,22	2,13	1,75
3	145,74	2,915	1,412	1,681	2,06	1,73
4	221,81	4,436	2,13	2,644	2,08	1,68
5	109,01	2,18	1,022	1,347	2,13	1,62
6	336,53	6,731	3,135	3,556	2,15	1,89
7	330,83	6,617	3,187	3,963	2,08	1,67
8	169,55	3,391	1,585	1,999	2,14	1,7
9	305,23	6,105	2,903	3,299	2,1	1,85
10	133,03	2,661	1,31	1,763	2,03	1,51
11	382,36	7,647	3,636	4,051	2,1	1,89
12	173,59	3,472	1,633	2,145	2,13	1,62
13	103,82	2,076	1,029	1,411	2,02	1,47
14	515,55	10,311	4,847	4,899	2,13	2,1
15	499,11	9,982	4,705	4,816	2,12	2,07
16	63,49	1,27	0,657	0,897	1,93	1,42
17	375,25	7,505	3,495	3,946	2,15	1,9
18	175,09	3,502	1,63	1,854	2,15	1,89
19	487,61	9,752	4,504	4,735	2,17	2,06
20	311,2	6,224	2,973	3,071	2,09	2,03
21	298,3	5,966	2,892	3,143	2,06	1,9
22	170,59	3,412	1,594	2,004	2,14	1,7
23	252,32	5,046	2,446	2,866	2,06	1,76
24	79,39	1,588	0,795	0,951	2	1,67
25	769,11	15,382	7,279	7,328	2,11	2,1
26	451,66	9,033	4,199	4,643	2,15	1,95
27	235,39	4,708	2,225	2,425	2,12	1,94
28	205,94	4,119	1,988	2,387	2,07	1,73
29	172,02	3,44	1,56	1,75	2,21	1,97
30	308,48	6,17	2,884	3,261	2,14	1,89
31	182,37	3,647	1,768	2,059	2,06	1,77
32	223,42	4,468	2,108	2,512	2,12	1,78
33	236,7	4,734	2,239	2,579	2,11	1,84
34	100,98	2,02	0,977	1,307	2,07	1,55
35	302,38	6,048	2,977	3,244	2,03	1,86
36	293,49	5,87	2,807	3,12	2,09	1,88
37	113,77	2,275	1,095	1,323	2,08	1,72
38	238,57	4,771	2,266	2,491	2,11	1,92
39	349,49	6,99	3,448	3,6	2,03	1,94
40	329,64	6,593	3,082	3,543	2,14	1,86
41	122,94	2,459	1,197	1,415	2,05	1,74
42	66,16	1,323	0,686	1,13	1,93	1,17
43	273,97	5,479	2,57	3,029	2,13	1,81
SL	72,89	1,458	0,842	1,183	1,73	1,23
VAL	80,6	1,612	0,843	1,275	1,91	1,26


```

>SL4.0ch01
Length=90863682

Score = 486 bits (263), Expect = 6e-136
Identities = 295/311 (95%), Gaps = 0/311 (0%)
Strand=Plus/Plus

Query 1          TCATCAATGAATGCCTGGAGATTCTGCTGGGAGGAGCCACCATCAGCCACCGCATCCTCC 60
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 87375289    TCATCGATGAATGCTTGGAGATTCTGCTGGGATGAGCCACCATCAGCCACCGCATCCTCC 87375348

Query 61         GCCACTTTCTTCCACTTCAACGCGTTTTCTTTCATCTCGGCGGCTTTGCTCCGGTGGTT 120
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 87375349    GCCACTTTCTTCCACTTCAACGCGTTTTCTTTCATCTCGGCGCTTTGCTCCACTCGTC 87375408

Query 121        GCCTCCCTAACGCACTTCTCCACTTCGTCTCGTGGGAATGATTCTATTTTCATCCTCACCT 180
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 87375409    GCCTCCCTCACGCACTTCTCCACTTCTCTCGTGGGAATGATTCTATTTTCATCCTCACCT 87375468

Query 181        CTACACAATTGAACCCCACTTTAAACACGTCCTAAGTACTTAGCATCAAGAACTTGA 240
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 87375469    CTACATAATTGAATCCCACTTTAAACACATCCACTAAGTACTTAGCGTCAAGAACTTGA 87375528

Query 241        TCACCCATTGTGGAACGCAATAATTGGCGTACCAATGGATATTGCTTCCATAGTTGAA 300
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 87375529    TCACCCATTGTGGAATGCAATAATTGGCGTACCAATAGATATTGCTTCCATAGTTGAA 87375588

Query 301        TTCCATCCACA 311
                ||||| |||||
Sbjct 87375589    TTCCATCCACA 87375599

```

Figura 10. Alineamiento de la secuencia que contiene al marcador 28042 contra la base de datos de cromosomas de tomate (versión SL4.0) .

```

>SL4.0ch02
Length=53473368

Score = 449 bits (243), Expect = 8e-125
Identities = 287/309 (93%), Gaps = 0/309 (0%)
Strand=Plus/Plus

Query 1          TTCCCTCAAATTTTCTCTTCACTTCCACAACCTTCCAAGAAAACCCACATCATACT 60
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 41121549    TTCCCTCAAATTTTCTCTCTCACTTCCACAACCTTCCAAGAAAACCCACATAAACA 41121608

Query 61         CATCCATTTCCAAGAAAATTTCTATCTACTGCCAAGACTCTAACTTTATGCTCTGTTACA 120
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 41121609    CATCCATTTCCAAGAAAATTTCTATTACTGCCAAGACTCTAACTTTATGCTCTGTTGCA 41121668

Query 121        GCTTATGAGTCATTATCTTGTGGTGGCTGGGATGATCCAAGTTTCATTGATGATCCAATT 180
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 41121669    GCTTCTGAATCATTGTCTTGTGGTGGCTGGGATGACCCAAGTTTCATTGGTATCCAATT 41121728

Query 181        AATCCTGGTGAGTCCAACAGCTTCACAATTTTCTTAATAAACTAGGAATTAATGATAAA 240
                | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 41121729    AGCCCTGGTGAGTCCAACAAGCTTCAATAATTTCTTAATAAACTAGGAATTAATGATAAA 41121788

Query 241        AAGTATATCTTTGTGTACCTATTGGGGTTTGTGTTGTGCTTTGGCTATTCAAGGGTGAAA 300
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 41121789    AAGTATGTCTTTGTTGCCCTTGGGGTTTGTGTTGTGCTTTGGCTATTCAAGGGTGAAA 41121848

Query 301        GTTTCCTCA 309
                ||||| |||||
Sbjct 41121849    GTTTCCTCA 41121857

```

Figura 11. Alineamiento de la secuencia que contiene al marcador 30276 contra la base de datos de cromosomas de tomate (versión SL4.0) .


```

>SL4.0ch02
Length=53473368

Score = 289 bits (156), Expect = 2e-76
Identities = 170/177 (96%), Gaps = 0/177 (0%)
Strand=Plus/Plus

Query 138      CAAGAACTTGACAGAACTGGACATTCAGGAGAATGGCTTGGATGATATTTCTGGTTGTTG 197
                || ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Sbjct 41787388 CAGGAACTTGACAGAGCTGGACATTCAGGAGAATGGCATGGATGATATTTCTGGTAGTTG 41787447

Query 198      GCTGAGTTGTTTTCCGGATGACTTTACATCACTGGAGGTGCTTAACTTTGCCAGTATGAA 257
                ||| ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Sbjct 41787448 GCTAAGTTGTTTTCCGGATGACTTTACATCACTGGAGGTGCTTAACTTTGCCAGTATGAA 41787507

Query 258      CACTGAGATCAGTAAAGATGCTTTAGAGAGACTTGTGAGGTGCAAATCACTGAG 314
                ||||||||||| ||||||||||| ||| |||||||||||
Sbjct 41787508 CACTGAGATCAGTAAAGATGCTTTAGAGAGACTTGTGAGGTGCAAATCACTGAG 41787564

```

Figura 12. Alineamiento de la secuencia que contiene al marcador 26784 contra la base de datos de cromosomas de tomate (versión SL4.0) .

```

>SL4.0ch02
Length=53473368

Score = 555 bits (300), Expect = 2e-156
Identities = 324/336 (96%), Gaps = 0/336 (0%)
Strand=Plus/Minus

Query 1        TATGGGGCAATGGTTGACTTAATTGGTCGAGCTGGTCGGTTGAGTGAGGCTTGGAACTTC 60
                ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||
Sbjct 41949876 TATGGGGCAATGGTTGACCTAATTGGTCGAGCTGGTCGGCTGAGTGAGGCTTGGAACTTC 41949817

Query 61       ATTGATAATATGCCTATCAGACCAGGACTCAATGTGTATGGAGCAATGTTAGGAGCTTGC 120
                ||||||||||| ||||||||||| ||| ||||||| |||||||
Sbjct 41949816 ATTGATAATATGCCTACCAGACCAGGACTCAATGTCTATGGTGCAATGTTGGGAGCTTGC 41949757

Query 121      AAGATTCACAAAAATGTTGATCTGGGGGAGAAGGCAGCAGACAAGCTATTTGAGCTAGAT 180
                ||||||||||| ||| ||||||||||| |||||||||||
Sbjct 41949756 AAGATTCACAAAAATGTAGATTTGGGGGAGAAGGCAGCAGACAAGCTATTTGAGCTAGAT 41949697

Query 181      CCCGATGATGGTGGATATCATGTGCTTCTTGCTAATATGTACGCCAGGGCTTCAATATGG 240
                ||||||||||| ||||||||||| ||||||||||| |||||||||||
Sbjct 41949696 CCCGATGATGGTGGATATCATGTGCTTCTTGCTAATATGTACGCCAGGGCTTCAATATGG 41949637

Query 241      CATAAAGTTGCTAAAGTCAGAACTATGATGGAGAGAAAAGGTATTCAAAGACACCTGGA 300
                ||||||||||| ||||||||||| ||||||||||| |||||||||||
Sbjct 41949636 CATAAAGTTGCTAATGTAGAAGTATGATGGAGAGAAAAGGTATTCAAAGACACCTGGA 41949577

Query 301      TGGAGTCTAGTTGATTTGAGAAACGAGGTCCACACA 336
                ||||||||||| |||||||||||
Sbjct 41949576 TGGAGTCTAGTTGATTTAAGAAACGAGGTCCACACA 41949541

```

Figura 13. Alineamiento de la secuencia que contiene al marcador 26025 contra la base de datos de cromosomas de tomate (versión SL4.0) .

```

>SL4.0ch02
Length=53473368

Score = 444 bits (240), Expect = 4e-123
Identities = 254/261 (97%), Gaps = 0/261 (0%)
Strand=Plus/Minus

Query 4      TTGATCTTTCTTGTTCATTTGCAATTCTGCGGCTCCAACCATTCAATGACAGAGTTTAT 63
            |||
Sbjct 43081345 TTGATCTTTCTTGTTCATTTGCAATTTTGC GGCTCCAACCATTCAATGACAGAGTTTAT 43081286

Query 64     TTCCCCAAATGGTATCTTAAGGGGTTGAGACACAGTCCCACCCGCTCTGGGGCATTGTC 123
            |||
Sbjct 43081285 TTCCCCAAATGGTATCTTAAGGGGTTGAGACACAGTCCCACCCACTCTGGGGCATTGTC 43081226

Query 124    GCGAAATTTGTTAATGTTGATTGGAGGGCATATATAAGGTTTCTGAACTGGATACCAGAT 183
            ||
Sbjct 43081225 GCGAAATTTGTTAATGTTGATTGGAGAGCATATATAAGGTTTCTGAACTGGATACCAGAT 43081166

Query 184    GCACTGAAAATGCCTGAACCTGAGCTTATTGACCATGCAGGGTTGGACTCTGCTGTTTAT 243
            ||
Sbjct 43081165 GCTCTGAAAATGCCTGAACCTGAGCTTATTGACCATGCAGGGTTGGACTCTGCTGTTTAT 43081106

Query 244    TTGCGCATATACTTGCTAGGG 264
            |||
Sbjct 43081105 TTGCGCATATACTTGCTAGGG 43081085

```

Figura 14. Alineamiento de la secuencia que contiene al marcador 28075 contra la base de datos de cromosomas de tomate (versión SL4.0) .

```

>SL4.0ch02
Length=53473368

Score = 490 bits (265), Expect = 5e-137
Identities = 297/313 (95%), Gaps = 0/313 (0%)
Strand=Plus/Plus

Query 1      CCTGAGTTTCGTTTCAGTGCTTATCCAAACATTACATCCATCCGACCACATACTCTCAAG 60
            |||
Sbjct 43249751 CCTGAGCTTCCTTCAGTGCTTATCCGAAACATTACGTCATCTGACCACATACTGTCAAG 43249810

Query 61     AGTGTACTACTCCAAACAATTCCTCATATAAATCTGTGTTACAACACTATATCAGAAACCA 120
            |||
Sbjct 43249811 AGTGTACTACTCCAAACAATTCCTCATATAAAGTCTGTGCTACAACACTATATCAGAAACCA 43249870

Query 121    GAGGTTCAATAGATCAACCACTCCAAAACCAGACATTATTTTCACTCCGTTACAAGAATC 180
            |||
Sbjct 43249871 GAGGTTCAATAGATCAACCACTCCAAAACCAGACATTATTTTCACTCCGTTACAAGAATC 43249930

Query 181    GCATGTCCAGGTTGCTGTCATTTGCTGCAGGAAAACGACAAGTACCTAAAGATCCGAAG 240
            |||
Sbjct 43249931 CCATGTTCAAGTTGCTGTCATTTGCTGCAGGAAAACGACAAGTACCTAAAGATCCGAAG 43249990

Query 241    TGGAGGGCATGACTATGAAGGCATTTCTTATGTCTCCAACATCAGTTTTGTTCTTCTTGA 300
            |||
Sbjct 43249991 TGGAGGCCATGACTACGAAGGCATTTCTTATGTCTCCAACATCAGTTTTGTTCTTCTTGA 43250050

Query 301    CATGTCTAACCTG 313
            |||
Sbjct 43250051 CATGTCTAACCTG 43250063

```

Figura 15. Alineamiento de la secuencia que contiene al marcador 18376 contra la base de datos de cromosomas de tomate (versión SL4.0) .

Tabla 6. Posicionamiento físico de los marcadores estudiados respecto al genoma de tomate (versión SL4.0).

SNP ID	Cromosoma	Posición
20530	1	81506359
28042	1	87375289
30276	2	41121549
26784	2	41787388
26025	2	41949876
28075	2	43081345
18376	2	43249751
26365	-	-

4.2. ESTUDIO DE LIGAMIENTO Y MAPA GENÉTICO

A partir de secuencias del transcriptoma previamente obtenidas (anexo), se seleccionaron diecisiete marcadores SNPs situados en los fragmentos implicados en la translocación cromosómica objeto de estudio de este trabajo. Esta translocación consiste en una aparente fusión cromosómica, respecto al genoma de tomate, los cromosomas uno y dos parecen estar fusionados en el genoma de pepino dulce. Fueron diseñados cebadores para amplificar zonas del genoma que incluyesen a estos marcadores SNPs obteniendo amplicones de entre 90 y 120 pb. A continuación se recogen las secuencias, tamaños y temperaturas de fusión de los cebadores diseñados mediante el software Primer3 (v. 0.4.0) (Rozen, 2007) (Tabla 7).

Tabla 7. Secuencias, tamaños y temperaturas de fusión de los cebadores diseñados mediante el software Primer3 (v. 0.4.0) (Rozen, 2007).

SNP ID	Cebador directo			Cebador reverso		
	Secuencia (5'→3')	Tamaño (pb)	Tm (°C)	Secuencia (5'→3')	Tamaño (pb)	Tm (°C)
26450	ACGGAAGTGGTGGTTGAAGT	20	59	CAGCTGCCAGTTCCTTTTCT	20	60
28099	CCCATTTCAAAATAGCCAGAG	21	59	GGAGAGGGGACAATCAAAGA	20	59
13325	AATGATGTTTTGGCCAACCT	20	59	TTCTTCCAAAATCATTCTTCACAA	24	60
06146	GGCCTGATGAAGAGAGAGTCC	21	57	TGATACAGAAGAAACGCAGGTG	22	45
13632	TGCATCAGCATCACTGTTTG	20	59	GCATTGATATTGGTGTGGTTG	21	59
28677	TTTTGCCCAACTTACAAGC	20	60	GGTTTCAGTCATTTTCAGAGGA	21	59
26365	TGCACTTCTGAGGCATTTTG	20	60	AGCACCAATACCAACTGCAA	20	59
30276	CCAAGAAAACCCACATCAT	20	60	CCCAGCCACCACAAGATAAT	20	60
26784	TGGCAAAATCATTCTTGGA	20	59	AGTTCTTGCAAGTGCAGCA	20	60
18376	TCAACCACTCCAAAACAGCA	20	59	CCCTCCACTTCGGATCTTTA	20	59
26025	GGAGCTTGCAAGATTCACAA	20	59	CCGTGGCGTACATATTAGCA	20	60
28075	CGTCGCGAAATTTGTTAATG	20	59	ACCCTGCATGGTCAATAAGC	20	60
20530	TGAATTATGTAATTTTCAGAAAAGTTGG	27	59	CGCATGAAAGCATTTCGTTTT	20	61
28042	CACTTCGTCTCGTGAATGA	20	60	GCCAATTATTGCGTTTCCAC	20	60

5. CONCLUSIONES

Según los datos obtenidos en el estudio de ligamiento realizado, existen indicios de una posible translocación cromosómica en pepino dulce no encontrada hasta la fecha en ninguna otra especie de la familia *Solanaceae*. El ligamiento de marcadores posicionados en los cromosomas uno y dos en el genoma de tomate (versión SL4.0) podría indicar una posible fusión cromosómica en pepino dulce respecto a otras especies de la misma familia. Sin embargo, no es posible confirmar de forma definitiva esta translocación, para ello sería necesario realizar un nuevo estudio utilizando un mayor número de marcadores que dotasen al mismo de una mayor robustez.

Cabe destacar que el orden de los marcadores y las distancias genómicas no son coherentes respecto al genoma de tomate (versión SL4.0) lo que podría indicar una posible reordenación en esa región cromosómica.

6. BIBLIOGRAFIA

Aiton, W. T. (1789). *Hortus Kewensis; Or a Catalogue of the Plants Cultivated. in the Royal Botanic Garden at Kew.(with Figures)*. George Nicol.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410. Visto el 20 de noviembre de 2020.
<https://solgenomics.net/tools/blast/>

Benavides, F., & Guénet, J. L. (2003). Manual de genética de roedores de laboratorio. *Cantoblanco: Universidad de Alacala de Henares*.

Bravo, A. y Arias, E. (1983). Cultivo del pepino dulce: antecedentes agronómicos y económicos. *El campesino*, 114(3). 15-34.

Bridges, C. B., & Morgan, T. H. (2016). *Sex-linked inheritance in Drosophila*. Good Press.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*, 32(3), 314.

Dagar, V., Chow, C. W., Ashley, D. M., & Algar, E. M. (2009). Rapid detection of SMARCB1 sequence variation using high resolution melting. *BMC cancer*, 9(1), 437.

Delgado de la Flor, B. (1988). Pepino, en: *Cultivos hortícolas. Datos básicos*. Universidad Agraria La Molina, Lima, Perú. 73-75.

Desmarais, E., Lanneluc, I., & Lagnel, J. (1998). Direct amplification of length polymorphisms (DALP), or how to get and characterize new genetic markers in many species. *Nucleic acids research*, 26(6), 1458-1465.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19(1). 11-15.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, 6(5).

Haldane, J. B. S. (1919). The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet*, 8(29), 299-309.

He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H., & Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in plant science*, 5, 484.

Herraiz, F. J., Blanca, J., Ziarolo, P., Gramazio, P., Plazas, M., Anderson, G. J., ... & Vilanova, S. (2016a). The first de novo transcriptome of pepino (*Solanum muricatum*):

assembly, comprehensive analysis and comparison with the closely related species *S. caripense*, potato and tomato. *BMC genomics*, 17(1), 321.

Herraiz, F. J., Villaño, D., Plazas, M., Vilanova, S., Ferreres, F., Prohens, J., & Moreno, D. A. (2016b). Phenolic profile and biological activities of the pepino (*Solanum muricatum*) fruit and its wild relative *S. caripense*. *International journal of molecular sciences*, 17(3), 394.

Herráiz, F. J. (2015). Desarrollo de herramientas morfológicas y genómicas para el estudio del pepino dulce (*Solanum muricatum*) y especies relacionadas. Caracterización de su valor nutracéutico. Tesis doctoral. *Universidad Politécnica de Valencia, Valencia, España*.

Hsu, C. C., Guo, Y. R., Wang, Z. H., & Yin, M. C. (2011). Protective effects of an aqueous extract from pepino (*Solanum muricatum* Ait.) in diabetic mice. *Journal of the Science of Food and Agriculture*, 91(8), 1517-1522.

Konieczny, A., & Ausubel, F. M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *The plant journal*, 4(2), 403-410.

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Annals of Eugenics*, Vol.12, 172–175.

Lester, R. N. (1991). Evolutionary relationships of tomato, potato, pepino and wild species of *Lycopersicon* and *Solanum*. *Solanaceae III: taxonomy, chemistry, evolution*, 283-301.

Li, Y. D., Chu, Z. Z., Liu, X. G., Jing, H. C., Liu, Y. G., & Hao, D. Y. (2010). A cost-effective high-resolution melting approach using the EvaGreen dye for DNA polymorphism detection and genotyping in plants. *Journal of integrative plant biology*, 52(12), 1036-1042.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature reviews genetics*, 11(1), 31-46.

Migheli, F., Stoccoro, A., Coppede, F., Omar, W. A. W., Failli, A., Consolini, R., ... & Migliore, L. (2013). Comparison study of MS-HRM and pyrosequencing techniques for quantification of APC and CDKN2A gene methylation. *PloS one*, 8(1).

Nuez, F., y Ruiz Martínez, J. J. (1996). *El pepino dulce y su cultivo*. FAO, Roma (Italia).

Paran, I., & Michelmore, R. W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. *Theoretical and applied genetics*, 85(8), 985-993.

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., ... & Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics*, 13(1), 341.

- Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current opinion in plant biology*, 5(2), 94-100.
- Redgwell, R. J., & Turner, N. A. (1986). Pepino (*Solanum muricatum*): Chemical composition of ripe fruit. *Journal of the Science of Food and Agriculture*, 37(12), 1217-1222.
- Rodríguez-Burruezo, A., Kollmannsberger, H., Prohens, J., Nitz, S., & Nuez, F. (2004a). Analysis of the volatile aroma constituents of parental and hybrid clones of pepino (*Solanum muricatum*). *Journal of agricultural and food chemistry*, 52(18), 5663-5669.
- Rodríguez-Burruezo, A., Prohens, J., & Fita, A. M. (2011). Breeding strategies for improving the performance and fruit quality of the pepino (*Solanum muricatum*): A model for the enhancement of underutilized exotic fruits. *Food Research International*, 44(7), 1927-1935.
- Rodríguez-Burruezo, A., Prohens, J., & Nuez, F. (2004b). Valencia': A New Pepino (*Solanum muricatum*) Cultivar with Improved Fruit Quality. *HortScience*, 39(6), 1500-1502.
- Rothberg, J. M., & Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nature biotechnology*, 26(10), 1117-1124.
- Rozen S., Remm M., Koressaar T. and Skaletsky H. Primer3 (v. 0.4.0) (2007). Whitehead Institute for Biomedical Research. Cambridge, Massachusetts, EEUU. Visto el 29 de marzo de 2020. <https://bioinfo.ut.ee/primer3-0.4.0/>
- Salimath, S. S., Oliveira, A. C. D., Bennetzen, J. L., & Godwin, I. D. (1995). Assessment of genome origins and genetic diversity in the genus *Eleusine* with DNA markers. *Genome*, 38(4), 757-763.
- Shathish, K., & Guruvayoorappan, C. (2014). *Solanum muricatum* Ait. inhibits inflammation and cancer by modulating the immune system. *Journal of cancer research and therapeutics*, 10(3), 623.
- Stam, P. (1993). Construction of integrated genetic linkage maps by means of a new computer package: Join Map. *The plant journal*, 3(5), 739-744.
- Sudha, G., Sangeetha, P. M., Indhu, S. R., & Vadivukkarasi, S. (2011). Antioxidant activity of ripe pepino fruit (*Solanum muricatum* Aiton). *Int j pharm pharm sci*, 3(3), 257-61.
- Sun, H., Wang, H. T., Kwon, W. S., In, J. G., Lee, B. S., & Yang, D. C. (2010). Development of Molecular Markers for the Determination of the New Cultivar 'Chunpoong' in *Panax ginseng* C.A. MEYER Associated with a Major Latex-Like Protein Gene. *Biological and Pharmaceutical Bulletin*, 33(2), 183-187.
- Taylor, J., & Butler, D. (2017). R package ASMap: efficient genetic linkage map construction and diagnosis. *arXiv preprint arXiv:1705.06916*.

Vilanova, S., Alonso, D., Gramazio, P., Plazas, M., Ferrante, P., Schmidt, M., Diez-Niclós, M. J., ... & Prohens, J. (2020). SILEX: A fast and inexpensive high-quality DNA extraction method suitable for multiple sequencing platforms and recalcitrant plant species. *Plant Methods* 16, 110.

7. ANEXO

En este anexo se recogen las secuencias obtenidas del transcriptoma de pepino dulce que incluyen los marcadores SNPs (señalados en rojo) analizados en este estudio de ligamiento genético.

20530 (308pb)

```
CGGCAGCACACATAATCAATGCTATTTATGGAGAAGATTTAGTGTCTAGTGGTTGG
TTTTTCAGACATGTCAAAGATCCAAGTATATCAGCCAAGAAGAGATGTATAACA
CTGTAATAATAGCTGTTGAATTATGTAATTTTCAGAAAAGTTGGATATAACCCTGTA
AATAGCTGCAGAATCAAGTAATTTCAATC[G]ATTAACGAATGCTTTCATGCGCATA
TGTTCTTTGTTCTATGAAAAACGAAAACTTGACAAAATGCTTTCATGTGCATATGT
TTTTGTTCCAGAAAAATGAAATGA
```

24694 (309pb)

```
TCCTTCTTCATCACTGAATTGAAAAAGATGGGTTTATCATTCCGGAAGCTTTTCA
GTCGGCTATTTGCCAAGAAGGAGATGCGTATTCTGATGGTTGGTCTTGATGCAGC
TGGTAAACCACCATATTGTACAAGCTCAAGCTGGGAGAGATTGTTACCACTATTC
CTACCATTGGATTCAATGTGGAGACTGTTGA[G]TACAAGAACATAAGCTTCACTGT
CTGGGATGTCGGGGGTCAGGACAAGATCAGACCATTGTGGAGGCATTACTTCCAA
AACACACAAGGACTTATCTTTGTGGTTGACAG
```

29632 (303pb)

```
TCTTCCCTCACATGCAGCTATTGTCTGTGGCTTAGTAAAGAATCAGGTATATTCTG
ATCTGGAGGTAGCTTTTGTATATCAGAAGACGAACTATCTATAAAGCACTATCC
AAAACCATGGAATTATTACAAGGAAATACTGCTTTATAAGCCAATATTATTATGGC
TGTCTTTTCCATGAGCTGCTA[T]CTTTTCAGCTTGCTGATCGACGTCCACCTGCTG
AGAGAGAAAGTGAGAAGTTGAAAGCACCCAAGATGATTGGCTTTCCTAGCTCTGT
CAACTCAGTACTCAATCATGCAGA
```

28677 (316pb)

```
TTCTCCAGAATTCCTACTGACCAGATAATTGGTTTATATTTCTTCCCTCACTTGATGC
CCCCTTATATGAGCCTGAATTTTAAACAATTTTTTGTGCAATTAAGGAAGACTTTTT
CTCTTATTCCAGCCACGGAACTTTTTCTGGATTTGAATGGCAGCCGCATGAGCTAC
ACCATTGTTTTGCCCAACTTACAAGC[T]CTAGAAGCTACAATAGAAAGAGCATTTC
TCATCAGATGATAACTCATTGTCACAATGCTCAATAATTTGCTTCCCTCTGAAATGAC
TGAACCCTGAAAATTTGATGTATCCGAGCAGCTG
```


28042 (311pb)

TCATCAATGAATGCCTGGAGATTCTGCTGGGAGGAGCCACCATCAGCCACCGCAT
CCTCCGCCACTTTCTTCCACTTCAACGCGTTTTCTTTCATCTCGGCGGCCTTTGCT
CCGGTGGTTGCCTCCCTAACGCACTTCTCCACTTCGTCTCGTGGAATGATTCTATT
TTCATCCTCACCTCTACACAAT[T]GAACCCCAACTTTAAACACGTCCACTAAGTACT
TAGCATCAAGAACTTGATCACCCATTGTGGAAACGCAATAATTGGCGTACCAATG
GATATTGCTTCCATAGTTGAATTCCATCCACA

26450 (305pb)

GCAGCAGTAAGCGAAAATCCAGGTGCAGAAGATACGGAAGTGGTGGTTGAAGTA
CTTAACAATGAAAATGCCGGGGCAGAAGAAGCAGAAGCTTGCGGGTGAAGCACCA
GGGAAAGAAAATACAG[A]TGAAGAAGAAAAGGAACTGGCAGCTGCCGGGCTGGC
AAACGAAAGTGCAGGTGAAGAAGAAGAAGCTGAATTCGAAGCCGACTGGAACGA
AAGTGAGGCCGCGGATGACGGAAAACCAAATGAAGGAGCTGTTGAAGATGATCC
AAAAGGGGAAGATGACGCTGAAGAAGACGACCCAAA

28099 (333pb)

TGTCAACCCTCAATAACAGGTGGTAGCTCTGAGCTCTTACACATTACATTAAGAA
TACAGTGACATGACCAACTGAGGACATCATCTTAATACCCATTTCAAATAGCCAG
AGTTTACACATCATTCAAGGATATAACC[G]GTGTTTTTTCAACTGTATTATGATCA
AAAGGACGATATGAACCTTTGAATCTTTGATTGTCCCCTCTCCTTCTCTTCTGGTC
CTTCCAGATCCCCTATGTTGTCATGATACTCCTGGAGTATATCTTTTACATCTTGCT
TAGTCAGCTTTTGTGCTTTAGCAAAAGGTCAATCCTTGCTCTGATTCCGGGG

18376 (314pb)

CCTGAGTTTCGTTCAAGTGCCTTATCCAAACATTCACATCCATCCGACCACATACTCT
CAAGAGTGTAACCTCAAACAATTCCTCATATAAATCTGTGTTACAACACTATATCA
GAAACCAGAGGTTCAATAGATCAACCACTCCAAAACCAGACATTATTTTCACTCCG
TTACAAGAATCGCATGTCCAGGTTGC[T]GTCATTTGCTGCAGGAAAACCTGACAAGT
ACCTAAAGATCCGAAGTGGAGGGCATGACTATGAAGGCATTTCTTATGTCTCCAA
CATCAGTTTTGTTCTTCTTGACATGTCTAACCTG

28075 (301pb)

GCGTTGATCTTTCTTGTGCAATTTGCAATTCTGCGGCTCCAACCATTCAATGACAG
AGTTTATTTCCCAAATGGTATCTTAAGGGGTTGAGACACAGTCCCACCCGCTCTG
GGGCATTCGTCGCGAAATTTGTTAATGTTGATTGGAGGGCATATATAAGGTTTCTG
AACTGGATAACAGATGCA[C]TGAAAATGCCTGAACCTGAGCTTATTGACCATGCAG
GGTTGGACTCTGCTGTTTATTTGCGCATATACTTGCTAGGGCTGAAAATATTTGTT
CCCATAACGTTGCTTGCATGGG

26784 (317pb)

TTGAGATTGAAAAGAATGGCTGTTAGTGATGAGAGTTTGGAGTTTTTGGCAAATC
ATTTCTTGGATTTAAGGCTCTCTCATTGTTGAGTTGTGATGGTTTTAGCACTGATG
GGAT[A]AGTAGCATTGCTGCTCACTGCAAGAACTTGACAGAACTGGACATTCAGG
AGAATGGCTTGGATGATATTTCTGGTTGTTGGCTGAGTTGTTTTCCGGATGACTTT
ACATCACTGGAGGTGCTTAACTTTGCCAGTATGAACACTGAGATCAGTAAAGATGC
TTTAGAGAGACTTGTCTAGGTGCAAATCACTGAGACA

06146 (327pb)

GGGAGTGCATAATGTGATGGCTTCTTGGCTAATCTTTAGTTGAATTGGTCTGTTAA
GTTGTATTTGCTAAAAAGATTTCTATCACGGGATGAAGTGACAATTTAGTAGGGCC
TGATGAAGAGAGAGTCCAG[G]AACAATCTTCAATTTTGGAGCTAATTATCCACAA
TTTCAATAGTATTATATTACACCTGCGTTTCTTCTGTATCAAGATGAAGCCTAGAGC
AAAATGGCTAACACAAAATTGAAATATGATACACACACAATTTCTATACCCAAAGGA
AAGTAAAACCAAGGATGATGGAATCAGCAAATCACTTTTGCCGAG

30276 (309pb)

TTCCCTCAAATTTTCTCTTTCACTTTCCACAACCTTCCCAAGAAAACCCACATCAT
ACTCATCCATTTCCAAGAAAATTTCTATCTACTGCCAAGACTCTAACTTTATGCTCT
GTTACAGCTT[A]TGAGTCATTATCTTGTGGTGGCTGGGATGATCCAAGTTTCATTG
ATGATCCAATTAATCCTGGTGAGTCCAACCAGCTTCACAATTTTCTTAATAAACTAG
GAATTAATGATAAAAAGTATATCTTTGTGTACCTATTGGGGTTTGTGTTGTGCTTTGG
CTATTTCAAGGGTGAAAGTTTCTCA

13632 (306pb)

CACAACCTAACCTTGATCCAAAATTTACGTAGCTCTAAACGACGTTGAAGCTTTA
CTACTGCTTGTATAAACTGCATCAGCATCACTGTTTGACCAACTTTTAAAATCACAA
CTCAAGT[C]ATCACTTTTCAAGATGGACTTGGAGTTCAACCACACCAATATCAATGCA
TTAATTGATCAATGGTTCTTCGTATGCATGCAAGGAACTTTCTCAAGCTGTATTTG
GCTATCTTCCCCAGTATCCTAGGATTCAACTCCATTGCAGAGCGAAAATCCAGCA
TTGAAACAGAAAAGAGATCACTCGG

26365 (311pb)

TCGCTGATGTACTTGCAGATGTTACAATTTTCACTGCTACTGGTCAGTCCGTCAAG
TTTGAGGACCTTTGGGATCAAATGAGGGAGTAGCTGTTGTTGCACTTCTGAGGC
ATTTTGGCTGTCTTGGCTGCTGGGAACTTGCTCAAGCACTCAAAGAGACAAAGGA
GAGATTTGATTCTGC[T]GGTGTAAGCTAATTGCAGTTGGTATTGGTGCTCCTAAC

AAAGCACGTATTCTTGCTGAAAGGTTACCATTTCATTGGATTGCTTGTATGCTGA
TCCGGACCGCAGGGCATAACGATGTTCTAGGCCTG

27482 (303pb)

GAGCAAACACCATGGCAATATTCCTTGCATTCATCCTTGTTGTAACGCTCATTATGC
ACAACGTCTGCCATTAATTGATTGCCAGTCAAGCAGTGCAGCTTCAGTTGGAG
GAAGTAAGTTCACTAGTTGGGTGCACTCTTCTTCAGTGTGCAATTCATTACCTGC
TCTGGGGTTAGAGAATCTAG[T]ACACCTGTAGGCAGCTCCCTAAACCATGCCTTAA
TCAATCCTGCCAAACAATGAACGTCAATTCCATATGGAACAACGCCTCTACTATTG
AGCTGGTTCCTCACAATTTCTCCTT

13325 (304pb)

ATCTGGACCATGCAGCAAAGACCAAGAACAGTGCTGAAGCACAGAAGTACTATGC
TGAAACTGTATCTACCTTAAATGATGTTTTGGCCAACCTTGGCTAGAAAGCTTTTCT
GTA[C]TATGTTACTTCTGTAAGTTAACGAATACTTCTTGAATCAATTGTGAAG
AATGATTTTGGAGAATGTTTAAATCTGAAACAATAAGCGCTGCATCTGAGAGC
ATCAGCTTGAGTTTTGCATGTGATTTTATTGGGAATTTCACTTGATTGAGGCCTG
GAAACAATAAACATGGTTGCACCA

26025 (336pb),

TATGGGGCAATGGTTGACTTAATTGGTTCGAGCTGGTTCGGTTGAGTGAGGCTTGA
ACTTCATTGATAATATGCCTATCAGACCAGGACTCAATGTGTATGGAGCAATGTTA
GGAGCTTGCAAGATTCACAAAAATGT[T]GATCTGGGGGAGAAGGCAGCAGACAAG
CTATTTGAGCTAGATCCCGATGATGGTGGATATCATGTGCTTCTTGCTAATATGTA
CGCCACGGCTTCAATATGGCATAAAGTTGCTAAAGTCAGAACTATGATGGAGAGA
AAAGGTATTCAAAGACACCTGGATGGAGTCTAGTTGATTTGAGAAACGAGGTCC
ACACA