

Análisis Genético de Caracteres de Interés agronómico en pepino



Autor: Víctor Villalobos Ramírez

Tutor: Juan Capel Salinas

Departamento de Biología y Geología

Empresa colaboradora: Rijk Zwaan Ibérica S.A.

Cursos Académico 2020/2021: Máster en Biotecnología Industrial y Agroalimentaria por la Universidad de Almería (UAL)

Análisis Genético de Caracteres de Interés agronómico en pepino

Resumen

La gran extensión del cultivo de pepino (*Cucumis sativus* L.) hacen de esta especie una de las hortícolas de mayor importancia económica a nivel mundial. Las estrategias aplicadas para lograr un aumento de la producción y de la calidad de *C. sativus* son muy variadas, pero las mejoras obtenidas nutricionalmente y por el manejo de los cultivos sólo explotan el potencial genético disponible en los individuos. Por este motivo, los programas de mejora persiguen la selección de aquellas poblaciones genéticamente más ventajosas o las líneas que porten mecanismos de resistencia frente a agentes de estrés (abiótico o biótico), causantes tanto de un descenso considerable de la producción como de la calidad. Estudiar las relaciones genéticas existentes para la expresión de estos caracteres en la descendencia puede facilitar el proceso de mejora. Una de las estrategias de mayor eficacia consiste en la identificación de elementos fenotípicos (análisis de correlación) o genotípicos (análisis de QTL por marcadores moleculares) que puedan tener una relación directa con los caracteres de interés. Para el presente proyecto se ha trabajado con 2 poblaciones BC1 de la variedad “Dutch” para el análisis de 14 caracteres agronómicos (8 en planta y 6 en fruto) y estudiar su posible ligamiento a un conjunto de marcadores moleculares tipo SNP facilitados por la empresa Rijk Zwaan S.L. En conjunto se detectaron 17 asociaciones entre marcadores y QTLs para algunos de los caracteres de estudio (precocidad, número y longitud de entrenudos, número de frutos por nudo, longitud, diámetro, cavidad seminal y forma del cuello). De entre estos, 12 QTLs se han detectado en la población A y 5 QTLs en la población B, siendo los caracteres “número de frutos por nudo”, “longitud” y “forma de cuello” únicos para la población A, mientras que para el carácter “diámetro” sólo se encontró asociación en la población B. Afinar la localización de estos QTL con respecto del genoma de pepino y poder identificar o estimar los genes implicados con precisión, son los objetivos de los posteriores ensayos de investigación que deberán llevarse a cabo. Además, el análisis de correlación pone de manifiesto asociaciones de covarianza entre algunos de los caracteres estudiados para ambas poblaciones, lo que contribuye como conocimiento base para futuros programas de mejora.

Palabras clave: correlación, marcador molecular, mejora genética, pepino, QTL, MAS, SNP.

Genetic Analysis of Agronomic Interest Characters in Cucumber

Abstract

The cucumber cultivation (*Cucumis sativus L.*) represents one of the most economically important horticultural species worldwide. The strategies to achieve an increase in production and the quality of its cultivation are very varied. This is because the improvements obtained through nutrition and by crop management can only use the genetic potential available in individuals. For this reason, breeding programs seek to select genetically most advantageous populations or lines that carry resistance mechanisms against stress agents (abiotic or biotic), causing both a considerable decrease in production and quality. Studying the existing genetic relationships for the expression of these traits in the offspring can facilitate the improvement process. One of the most efficient strategies is to identify phenotypic (correlation analysis) or genotypic (QTL analysis by molecular markers) elements that may have a direct relationship with the traits of interest. For this project, we have worked with 2 BC1 populations of the "Dutch" variety for the analysis of 14 agronomic characters (8 in the plant and 6 in the fruit) and study their possible linkage to a set of SNP-type molecular markers provided by the company Rijk Zwaan S.L. Overall 17 associations between markers and QTLs were detected for some of the study characters (precocity, number and size of internodes, number of fruits per joint, length, diameter, seminal cavity and neck). Of these, 12 have been detected in population A and 5 in population B, being the characters of number of fruits by joint, length and neck unique for population A, while the diameter was associated with population B. The correlation analysis reveals covariance associations between some of the characters studied for both populations, which contributes as a base knowledge for future improvement programs based on MAS. The objectives of subsequent research trials should refine the location of these QTL with respect to the cucumber genome and be able to identify or estimate the genes involved with precision.

Key words: cucumber, correlation, molecular marker, genetic breeding, QTL, MAS, SNP.

Indice

1. Introducción	6
1.1) <i>Cucumis sativus L.</i>	6
1.1.1) Grupos varietales y tipos de pepino cultivados en España.....	8
1.1.2) Importancia socio-económica del pepino	9
1.2) Mejora genética.....	10
1.2.1) Importancia de la mejora genética	11
1.2.2) Marcadores genéticos.....	11
1.2.3) Aplicaciones de los marcadores moleculares	12
1.2.4) Mapas genéticos y QTLs	13
1.2.5) Selección asistida por marcadores moleculares (MAS).....	13
2) Objetivos	14
3) Material y métodos	15
3.1) Material vegetal empleado	15
3.2) Condiciones de cultivo.....	16
3.4) Evaluación y caracterización fenotípica	17
3.5) Análisis de datos (MapQTL6)	19
3.6) Análisis de regresión y correlación.....	21
4) Resultados	22
4.1) Evaluación fenotípica.....	22
4.2) Mapa de Ligamiento	24
4.3) Análisis de QTLs	25
4.4) Análisis de correlación.....	33
5) Discusión.....	35
5.1) Evaluación fenotípica.....	35
5.2) Mapa genético	37
5.3) Análisis de QTLs	38

5.4) Análisis de correlación.....	41
6) Conclusiones	44
7) Bibliografía	45

1. Introducción

Actualmente la humanidad se encuentra en un periodo de expansión demográfica que demanda una creciente cantidad de alimento para poder sustentarla. Además, los efectos evidentes del cambio climático son cada vez más radicales y agresivos, lo que afecta a los cultivos directamente a causa del estrés abiótico en las áreas de producción e indirectamente a través de los agentes patógenos que se desarrollan favorablemente en estas condiciones. Frente a este panorama se hace necesario desarrollar nuevas variedades agrícolas que permitan incrementar tanto su producción como la calidad de sus frutos.

Entre las especies hortícolas de mayor interés económico destaca el pepino (*Cucumis sativus L.*), siendo usado fundamentalmente como alimento primario, aunque también se emplea como materia prima para la elaboración de extractos y procesados. En la actual sociedad preocupada por la salud y bienestar, este alimento está siendo cada vez más demandado por sus propiedades beneficiosas (diurético, depurativo, demulcente, nutritivo, emoliente, antipruginoso, fuente de vitaminas y minerales esenciales) (Sakata et al., 2010).

1.1) *Cucumis sativus L.*

El pepino es una especie angiosperma dicotiledónea encuadrada taxonómicamente dentro del orden *Cucurbitales*, familia *Cucurbitaceae*, género *Cucumis*, especie *Cucumis sativus* (Tabla 1) (Sebastian et al., 2010). Esta familia comprende 90 géneros y más de 750 especies, dentro de los cuales el género *Cucumis* constituye uno de los más importantes a nivel económico al abarcar también otros cultivos como el de melón. Se compone de más de 30 especies originarias de África y Asia (Tatlioglu, 1993).

Tabla 1. Clasificación taxonómica del pepino (*Cucumis sativus L.*)

Clasificación taxonómica	
Orden	<i>Cucurbitales</i>
Familia	<i>Cucurbitaceae</i>
Clase	<i>Dicotiledóneas</i>
Tipo	<i>Fanerógamas</i>
Subtipo	<i>Angiospermas</i>
Género	<i>Cucumis</i>
Especie	<i>Cucumis sativus</i>

El origen de *C. sativus* se encuentra en las regiones tropicales del sur de Asia, considerándose la India el epicentro de su origen y centro primario de su diversificación, respaldado por la amplia diversidad de variedades locales y silvestres en las regiones del Himalaya. La ruta de diversificación pasa por las regiones de Asia Central y Occidental como centro secundario, a partir del cual paso a Europa y posteriormente a los continentes africano y americano. A partir de la especie *C. sativus var. Hardwickii* se llevó a cabo un proceso de domesticación que ha culminado en las distintas variedades que conocemos hoy día. (Tatlioglu, 1993; Sebastian et al., 2010; Qi et al., 2013)

Se trata de una planta herbácea anual considerada como rastrera o trepadora con un crecimiento de tipo indeterminado. Las plantas de pepino presentan un crecimiento en nudos de los cuales parten una hoja y un zarcillo (estructura que favorece la fijación de la planta a distintas superficies aéreas). En la región axilar se observa el desarrollo de un meristemo lateral y de un meristemo floral, el cual puede dar lugar a una o varias flores (masculinas o femeninas). La mayoría de las variedades de pepino suelen ser monoicas (flores masculinas y femeninas en el mismo pie de planta), aunque también pueden encontrarse variedades con plantas andromonoicas (flores hermafroditas y masculinas), e incluso dioicas (flores masculinas y femeninas en distinto pie de planta). Presentan una reproducción de tipo alogámica, lo que favorece la mezcla genética y la heterocigosidad de la especie (Tatlioglu, 1993; Valcárcel, 2018).

Actualmente, todas las variedades con interés comercial son de tipo ginoicas, es decir, las plantas solo desarrollan flores femeninas. Además, estas son partenocárpicas, desarrollándose el ovario sin necesidad de que haya polinización (Sun et al., 2006). Este carácter está determinado por el gen "P" con herencia por dominancia incompleta y cuya presencia está ligada a un aumento de la producción y calidad del fruto. (Pike & Peterson., 1969)

Las condiciones climáticas adecuadas para su crecimiento coinciden con las del clima templado. En términos generales crecen de forma óptima en el rango entre 20-30°C, observándose desequilibrios perjudiciales fuera de este intervalo (López-Elías et al., 2011). En cuanto a la humedad relativa, el óptimo oscila entre 60-70% siendo tanto el déficit como el exceso perjudiciales para el desarrollo de la planta (López-Elías et al., 2011). Gracias a los invernaderos se ha logrado su cultivo durante todo el año simulando dichas condiciones.

El fruto o pepónida se conoce comúnmente como pepino y está compuesto principalmente por agua ($\pm 97\%$), siendo el 3% restante una composición equilibrada de carbohidratos, grasas,

proteínas, vitaminas y minerales (Ca, Fe, Mg y Mn, principalmente). En cuanto al contenido vitamínico destaca el contenido en vitaminas K y C, así como el ácido pantoténico, que en conjunto confieren algunas de las propiedades beneficiosas de interés de la especie (Tatlioglu, 1993; Feng et al., 2020).

Fenotípicamente los frutos pueden presentar un aspecto superficial liso o rugoso de tamaño variable dependiendo de la variedad. El color es verde y su tonalidad varía en función de la variedad y de su etapa de desarrollo. El fruto comercial del pepino se recolecta en estado inmaduro, ya que el fruto maduro presenta características organolépticas que no son atractivas para los consumidores (exocarpo duro, sabor amargo, semillas muy duras y calibre de fruto inapropiado) (Valcárcel, 2018).

La calidad externa del fruto viene reflejada por un conjunto de caracteres como son el color, la forma, el tamaño y la presencia o ausencia de cuello, entre otros. Estos indicadores de calidad varían en función del mercado donde nos encontremos, pudiendo encontrar demandas por frutos con color oscuro, sin cuello y de tamaño corto en una región, mientras que en otra buscan tonos claros con cuello pronunciado y de gran envergadura (Imam et al., 1977; Smith et al., 1978).

1.1.1) Grupos varietales y tipos de pepino cultivados en España

Según el lugar de origen se puede reconocer una diversidad de subespecies y variedades (comerciales y silvestres sin interés comercial). Las diversas variedades comerciales se establecen en base a características fenotípicas del fruto como son el tamaño (longitud-diámetro) color, grosor de la piel y presencia de protuberancias o verrugas con espinas. En 1998, Shetty y Wehner establecieron una clasificación de los grupos varietales que son cultivados y que ha sido revisada posteriormente por diversos autores, concretando los siguientes grupos: “Gherkin”, “Slicer”, “Dutch”, “Beit Alpha” o “Persian”, “Oriental”, “Blanco”, “Moteado” y “Manzano” o “limón”. En España los morfotipos más cultivados son los de tipo Dutch o Holandés, Slicer, Gherkin y tipo Persian (Shetty y Wehner., 1998; Sakata et al., 2010; Valcárcel, 2018).

1.1.2) Importancia socio-económica del pepino

El cultivo de pepino representa un cultivo de gran importancia económica, teniendo un elevado índice de consumo tanto fresco como a través de procesados. Se encuentra entre las especies hortícolas con mayor relevancia a nivel global consolidándose con un gran avance y desarrollo en las últimas décadas. Mundialmente se ha alcanzado una superficie destinada a su cultivo de 2.231.402 ha con una producción de 87.805.086 toneladas en 2019, contrastando con las 1.055.605 ha de 1980, las cuales producían 12.758.034 de toneladas (FAOSTAT, 2021) .

Asia constituye el mayor productor a nivel mundial representando el 89% de la producción. Por su parte Europa produjo el 6,84%, destinado la mayor parte al comercio intracomunitario. Dentro de Europa España se posiciona como el tercer mayor productor abasteciendo el 12,29% del mercado Europeo, lo que refleja el papel tan relevante de este producto para el comercio y la economía de nuestro país (Figura 1) (FAOSTAT, 2021).

El desarrollo y avance científico-tecnológico en el área agroalimentaria ha posibilitado la mejora de las variedades hortícolas en términos de productividad, adaptación y resistencia a las condiciones ambientales y las enfermedades incidentes. De forma conjunta estos avances han posibilitado un incremento de la capacidad productiva con un requerimiento menor del área de cultivo. Como evidencia de ello podemos observar el incremento del rendimiento a nivel mundial, pasando de 120.860 hg/ha en 1980 a 393.497 hg/ha en 2019 (Figura 2) (FAOSTAT, 2021).

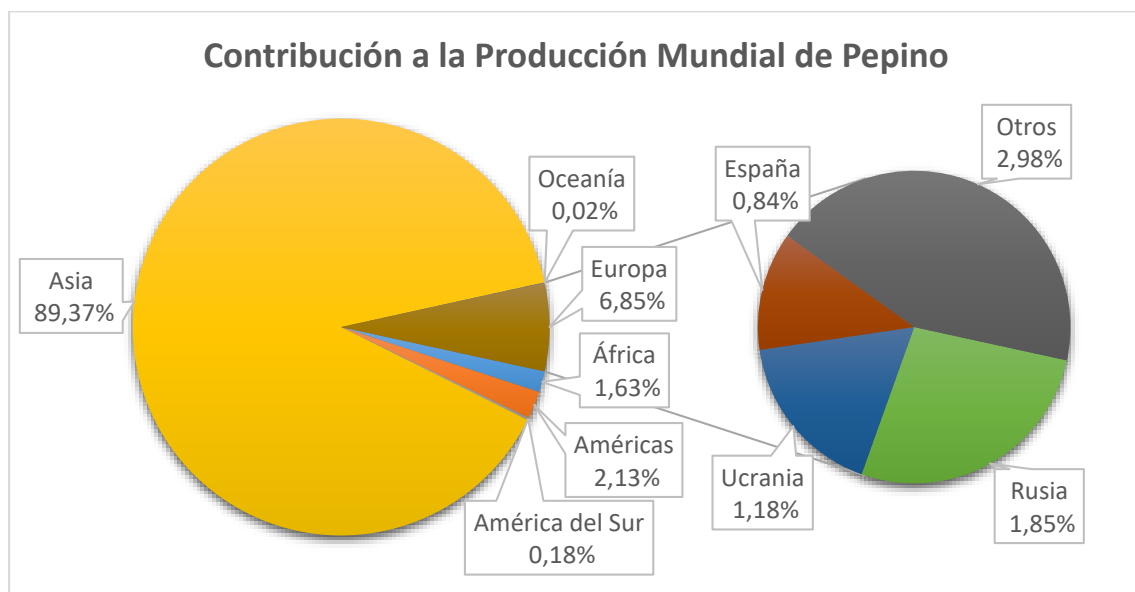


Figura 1. Porcentaje de contribución de cada continente a la producción mundial de pepino y de forma desglosada los principales productores a nivel europeo y su contribución a nivel mundial. Datos extraídos de FAOSTAT., 2021.

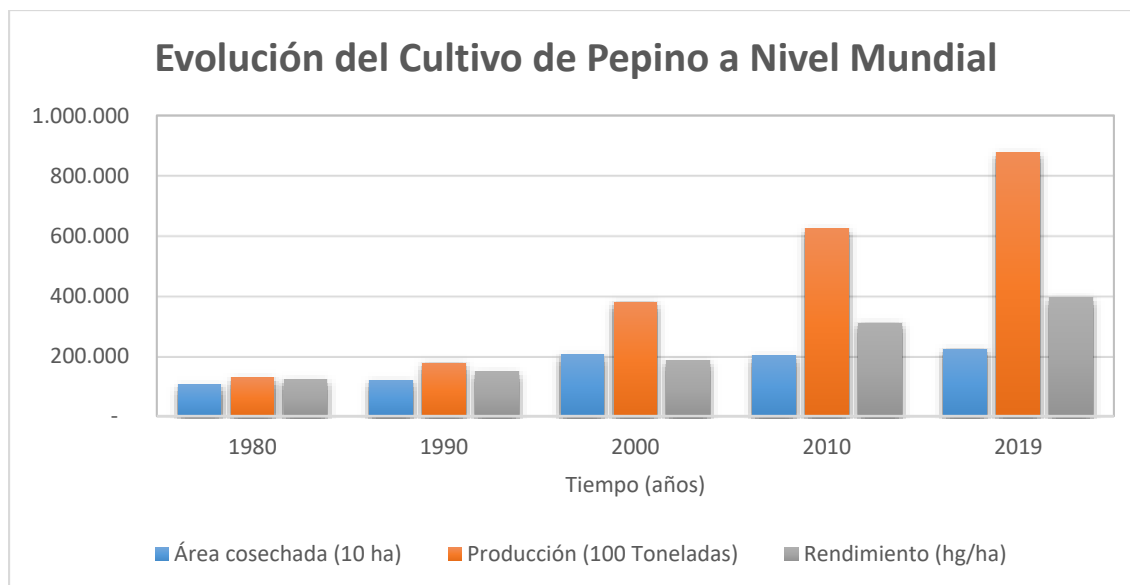


Figura 2. Evolución del cultivo de pepino a nivel mundial, atendiendo a valores de área de cultivo (10 ha), producción (100 toneladas) y rendimiento (hg/ha), desde 1980 hasta los datos más recientes 2019. Datos extraídos de FAOSTAT, 2021.

1.2) Mejora genética

La mejora genética es una disciplina que emplea la elección de los individuos reproductores para intentar obtener una descendencia que muestre características deseadas o favorables, modificando durante el proceso la estructura genética de las poblaciones. Se entiende por características favorables aquellas que logran un incremento en la productividad, la calidad o la adaptación al medio físico y económico en que se cultivan dichas plantas (Cubero, 2013). Los estudios genómicos han posibilitado la reforma de la mejora genética tradicional, haciendo posible la aplicación de técnicas de mayor precisión, lo que ha conseguido un aumento considerable de los resultados (De Giovanni et al., 2004; Doebley y col., 2006).

Los logros a pequeña escala de la mejora genética de las poblaciones se suelen traducir en grandes ganancias cuando se aplican a las poblaciones a gran escala. Esto se debe a que las mejoras obtenidas a través del componente nutricional y el manejo de los cultivos únicamente puede explotar el potencial genético disponible en los individuos, el cual solo puede incrementarse en los programas de mejora (Valcárcel, 2018).

Dentro de los diversos estudios y análisis genéticos que pueden ser realizados en cualquier especie, destaca la obtención de mapas genéticos por el valor base que aportan. Resulta de especial relevancia disponer de un mapa genético bien desarrollado y preciso para poder evaluar y determinar de forma selectiva las estrategias de los programas de mejora genética (De Giovanni et al., 2004).

1.2.1) Importancia de la mejora genética

El aumento de la producción de los cultivos y de la calidad de los mismos son objetivos de investigación del sector hortícola. Las estrategias para lograr estos fines son muy variadas, ya que pueden implicar la selección de aquellas poblaciones genéticamente más productivas o enfocarse en las líneas que porten mecanismos de resistencia frente a agentes patogénicos o estreses abióticos (temperatura, salinidad, déficit hídrico,...). Los estreses abióticos y bióticos son causantes tanto de un descenso considerable de la producción como de la merma de las características organolépticas (Llácer, 2007; Espejo, 2006).

En la mejora genética de especies vegetales se trabaja con el objetivo de obtener variedades que reúnan todos estos caracteres de interés. El lograr aunar todos ellos es una labor de suma complejidad, debido a que suelen estar controlados por varios genes e influenciados por las condiciones ambientales. Estudiar las relaciones genéticas existentes para la expresión de estos caracteres en la descendencia puede facilitar el proceso.

Una de las estrategias de mayor eficacia consiste en la identificación de elementos fenotípicos o genotípicos que puedan estar relacionados de forma directa con la expresión del carácter de estudio. En el caso de los marcadores fenotípicos se trabaja con análisis de correlación, permitiendo determinar una posible asociación entre dos parámetros o caracteres. Por su parte los genotípicos están ligados a elementos o señales moleculares ligadas al carácter de interés.

1.2.2) Marcadores genéticos

Los marcadores genéticos son señales reconocibles de forma fenotípica o molecularmente (proteínas o DNA) dentro de un conjunto de genotipos polimórficos (O'Brien, 1990). Son susceptibles de ser utilizados para la selección genotípica, ya sea de forma individual o combinada. Se puede distinguir entre marcadores morfológicos y moleculares, siendo los primeros los más empleados en la mejora genética clásica y los segundos en la mejora genética moderna.

Si nos centramos en los marcadores moleculares, estos pueden ser cualquier secuencia de DNA (marcadores de DNA) o de proteínas (marcadores bioquímicos) que pueda detectarse por técnicas moleculares y que ponga de manifiesto polimorfismos entre individuos. En la mejora genética vegetal principalmente se emplean los marcadores de DNA (Lefebvre y Chèvre, 1995). Los marcadores de DNA se definen como cambios en la secuencia de DNA que no llevan asociados cambios fenotípicos, pudiendo presentar varios alelos y regirse por las mismas leyes que regulan la segregación y heredabilidad en los genes (Lefebvre y Chèvre, 1995).

Desde su aparición se han desarrollado una amplia variedad de tipos, entre los cuales, podemos destacar los más utilizados en la elaboración de mapas genéticos y en la mejora vegetal: (Tanksley, 1993; Lefebvre y Chèvre, 1995)

- Marcadores RFLP (Restriction Fragment Length Polymorphism)
- Marcadores CAPS (Cleaved Amplified Polymorphic Sequences)
- Marcadores RAPD (Random Amplified Polymorphic DNA)
- Marcadores SCAR (Sequence Characterized Amplified Regions)
- Marcadores AFLP (Amplified Fragment Length Polymorphism)
- Marcadores VNTRs (Variable Number of Tandem Repeats)
- Marcadores SNP (Single-Nucleotide Polymorphisms)

Los marcadores moleculares que se han empleado en este estudio están basados en polimorfismos de un solo nucleótido en la secuencia de DNA (SNP). En los SNP el polimorfismo se debe a cambios en un solo nucleótido de la secuencia genotípica, siendo cada cambio un *locus* con 4 alelos (A, T, G, C). La distribución de estos *loci* en el genoma de cualquier especie es aleatoria, pudiendo estar o no en un gen (región codificante o reguladora). En términos generales estas diferencias no llevan asociado cambios en el fenotipo y sólo se consideran marcadores aquellos que están presentes en más del 1% de una población (Martín, 2002).

1.2.3) Aplicaciones de los marcadores moleculares

La aparición de los marcadores moleculares ha constituido una revolución en el campo de la genética, permitiendo la obtención de las conocidas como “huellas genéticas”, las cuales detectan la variabilidad existente tanto entre especies, poblaciones, variedades e incluso individuos.

Entre los usos más destacados cabe mencionar la realización de análisis de asociación de genoma completo (GWAS), elaboración de mapas genéticos de elevada precisión, la localización de genes ligados a caracteres de herencia cuantitativa (QTL, Quantitative Trait *Loci*) y la selección asistida por marcadores (MAS, Marker Assisted Selection).

El empleo de los marcadores moleculares como herramienta en los programas de mejora de variedades agrícolas, como el pepino, es de gran interés, debido en parte a la gran influencia del ambiente en el desarrollo fenotípico de las plantas (Yuan et al., 2008).

1.2.4) Mapas genéticos y QTLs

Los mapas genéticos determinan la posición relativa de los distintos genes o marcadores moleculares de referencia dentro de un grupo de ligamiento determinado, así como la distancia existente entre ellos. Representan la ordenación lineal de un conjunto de marcadores asociados a los distintos grupos de ligamiento y constituyendo en su conjunto el genoma de una especie.

Los mapas genéticos se elaboran en base a las frecuencias de recombinación estimadas y existentes entre los distintos componentes de estudio, ya sean genes o marcadores, reflejando la distancia genética que los separa (medida en centimorgan, cM) (Lee, 1995).

Los mapas genéticos son una herramienta de gran importancia en múltiples áreas de conocimiento base y tecnológico. En lo que respecta a la mejora genética ha logrado facilitar y acelerar la identificación y posterior incorporación de genes en variedades nuevas. Además, la información que proporcionan permite conocer más en profundidad la biología de los caracteres complejos, facilitando el proceso de clonación posicional de genes (Jones et al., 1997).

La gran mayoría de los caracteres con potencial interés para la mejora genética son complejos (varían de forma continua), estando determinados por multitud de genes segregantes entre la población (poligenes), con efecto cuantitativo e influenciados por el componente ambiental (González, 2000). Normalmente se desconoce los genes o regiones reguladoras implicadas, pero sí podemos acotar su posición en el genoma basándonos en el ligamiento con marcadores moleculares. Por este motivo se buscan marcadores con una baja frecuencia de recombinación con el QTL, de forma que exista una correlación significativa entre ambos (ligamiento) (Tanksley, 1993; Kearsey y Farquhar, 1998). De forma general se emplea la puntuación LOD (LOD-scores) para la cuantificación del ligamiento entre un QTL y un marcador. La puntuación LOD compara la probabilidad de que dos *loci* se encuentren ligados, con la probabilidad de que no lo estén y que por tanto la frecuencia de recombinación sea debida al azar.

1.2.5) Selección asistida por marcadores moleculares (MAS).

El empleo de estas herramientas permite la introgresión génica en variedades de interés comercial o de estudio, aumentando la velocidad y eficacia del proceso al reducir el número de cruzamientos necesarios (Liu et al., 2000). Este proceso consiste en la introducción de nuevos alelos o genes en una población mediante retrocruzamientos repetidos del híbrido con el parental de interés. Su utilización es especialmente relevante cuando se trabaja con caracteres difíciles de observar, que presentan baja heredabilidad o expresión tardía en el ciclo de cultivo (Mohan et al., 1997).

Por otra parte, los marcadores moleculares pueden identificar los genotipos de interés en estadios de desarrollo muy tempranos. Este fenómeno permite analizar las poblaciones segregantes de gran tamaño reduciendo gastos de tiempo, cultivo y personal.

La Empresa Rijk Zwaan S.L. dispone de un banco de germoplasma compuesto por multitud de líneas autóctonas de distintas partes del mundo, así como líneas consanguíneas recombinantes (RIL), líneas casi isogénicas (NIL) y líneas de introgresión (IL). Esto junto con el desarrollo de estudios de caracterización de nuevos marcadores moleculares permite a los mejoradores o Breeders seleccionar las plantas parentales para la realización de programas de cruzamiento que puedan culminar en la obtención de nuevos híbridos comerciales.

Para el presente trabajo se ha trabajado con 2 poblaciones BC1 de la variedad “Dutch” resultantes del cruce entre un individuo F1 y uno de sus parentales (retrocruce). Estas poblaciones proceden de un cruce inicial entre un parental de interés distinto y un parental común o compartido (PC). Los equipos de pre-breeding de la empresa Rijk Zwaan determinaron 67 marcadores de tipo SNP como posibles marcadores moleculares divergentes (polimórficos) de interés entre las poblaciones F1 de ambos cruces. Una vez realizado el retrocruzamiento el número de marcadores se redujo a 42 para la población A y 37 para la población B, siendo 16 de ellos compartidos en el estudio de ambas poblaciones.

2) Objetivos

El objetivo principal de este trabajo ha sido el análisis del control genético de 14 caracteres de interés agronómico (8 caracteres relacionados con el desarrollo vegetativo y 6 relacionados con la calidad del fruto) en 2 poblaciones segregantes de pepino. Con tal fin se han generado dos mapas genéticos con los datos del genotipado de un conjunto de marcadores moleculares tipo SNP realizados por la empresa Rijk Zwaan Ibérica, para la identificación de QTL para algunos de dichos caracteres. Adicionalmente se ha analizado la existencia de correlaciones, y si son debidas a ligamiento genético o causas fisiológicas, entre los caracteres analizados.

3) Material y métodos

Esquema temporal del procedimiento experimental:

- Preparación de las poblaciones de estudio (5 semanas): Siembra y trasplante.
- Evaluación fenotípica para los 14 caracteres de estudio (2 semanas).
- Cálculo de mapa genético y análisis de QTLs empleando marcadores moleculares tipo SNP (1 semana).
- Análisis y discusión de los resultados obtenidos (4 semanas)

3.1) Material vegetal empleado

A partir de 2 poblaciones F1 obtenidas del cruce entre un parental común (PC) y dos parentales desconocidos (PD1 y PD2, información confidencial de la empresa) se ha realizado un retrocruce hacia el parental común portador de unas cualidades fenotípicas de interés agronómico a fin de obtener líneas consanguíneas recombinantes (RIL) (Figura 3).

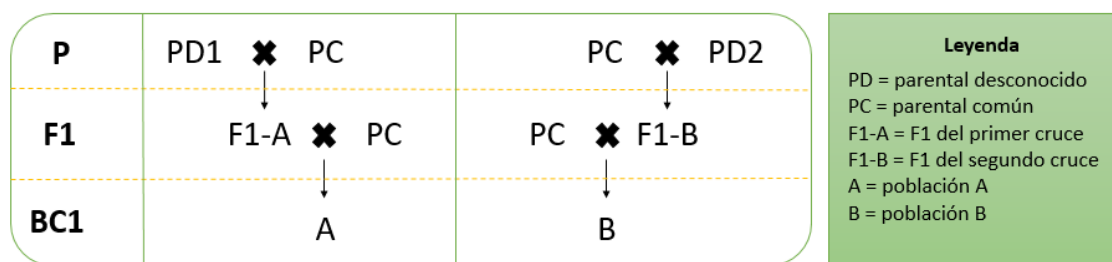


Figura 3. Esquema de los cruzamientos realizados para la obtención de las poblaciones de estudio dentro del programa de mejora de la empresa Rijk Zwaan Iberica.

Fenotípicamente, la F1-A es una planta multipistilar, con un fruto corto de color pálido y cuello no muy pronunciado. Por su parte la F1-B se trata de una planta monopistilar con un fruto largo de color relativamente claro, caracterizándose por un marcado “cuello de botella” (región anterior del fruto estrecha) y por ser de piel lisa. El objetivo del retrocruzamiento se centra en los caracteres mencionados, siendo el PC una planta dipistilar de tamaño intermedio de fruto, presentando un tono más oscuro y un cuello redondeado.

La siembra del ensayo se realizó el 30 de Mayo del presente año (2021) y consistió en 64 plantas de cada una de las poblaciones BC1 (A y B), 12 plantas para cada una de las F1 del retrocruce (8 de estudio y 4 como plantas polinizadoras del programa) y un número similar para los 4 controles del ensayo (líneas estables genéticamente y confidenciales de la empresa). Tras 19 días estas plantas se trasplantaron a un invernadero de tipo multitúnel con regulación térmica (15-35°C) y fotoperiodo natural de aproximadamente 14:10 horas (luz solar: oscuridad) con una incidencia lumínica media de 349,19 W/m² s. La densidad del invernadero de ensayo era

muy elevada (2,2 plantas/m²) lo que permitió mantener una elevada humedad ambiental (50-90%), la cual es imprescindible para el cultivo de pepino como se ha comentado anteriormente. Los controles se dispusieron de forma aleatoria para identificar un comportamiento ambiental homogéneo en toda el área de invernadero, mientras que las F1 macho o polinizadoras se ubicaron en el extremo del invernadero y separadas físicamente del ensayo por plantas control. El número de 64 plantas por población se estimó como número suficiente para la observación de la segregación genética para los caracteres de interés para la empresa, es decir, tamaño mínimo de la familia para conseguir al menos un individuo con un genotipo determinado.

$$n = \frac{\log(1 - S)}{\log(1 - p)}$$

Siendo “*n*” el número de individuos, “*S*” el factor de riesgo, nivel de seguridad o probabilidad con la que queremos que aparezca al menos un individuo de un determinado genotipo que tiene “*p*” probabilidad de aparecer. A modo de ejemplo, puede decirse que siendo *n* = 64, la probabilidad (*S*) de que se observe al menos un individuo que muestre un fenotipo debido a dos genes recesivos (*p* = 1/4 x 1/4 = 1/16) es igual al 98,4 %.

3.2) Condiciones de cultivo

El invernadero presenta unos sistemas automatizados de ventilación, calefacción-refrigeración, iluminación artificial y sistemas de riego por goteo. Además, dispone de sensores que recopilan la información relativa a la temperatura, humedad y radiación del ensayo en todo momento (software “Mithra clima”). Esta monitorización constante permitió mantener unas condiciones de cultivo óptimas durante la realización del presente trabajo.

Durante la etapa inicial del cultivo tras el trasplante, los individuos se cubrieron con el sistema de protección denominado “manta térmica”, el cual evita la pérdida excesiva de calor, favorece el crecimiento y otorga una protección frente a estreses bióticos (plagas e insectos vectores). Posteriormente se llevó un manejo de poda continuada de los meristemas secundarios a fin de estimular el desarrollo primario de la planta

Como sustrato de cultivo se empleó una mezcla estéril, entre la que destaca la fibra de coco. Este elemento es esencial gracias a sus cualidades que permiten una rápida rehidratación y retención de agua. Junto con los riegos se le suministro una solución nutritiva (Nutrialbor amarilla-naranja) cuya composición nutricional consistía en 1,6-3,3 nitrógeno en su forma nítrica (NO₃), 0-4% fósforo soluble (P₂O₅), 3,5-5% potasio soluble (K₂O), 0-0,5% óxido de magnesio (MgO), 0-5% óxido de calcio (CaO) y 0-1,1% de sulfatos (S).

3.4) Evaluación y caracterización fenotípica

La evaluación fenotípica se realizó 33 días después del trasplante (DDT) coincidiendo con el estado fenológico apropiado para la evaluación de los caracteres. La evaluación se llevó a cabo en dos periodos: 1º) Se analizaron los 8 caracteres vegetativos de la planta (21-27/06/2021) y 2º) estudio de los caracteres asociados al fruto (28/06-04/07/2021). Esta subdivisión permitió obtener un número representativo de frutos para su evaluación, ofreciendo unos datos más precisos.

Los caracteres analizados en el ensayo, así como los criterios seguidos para su evaluación se detallan a continuación. Estos criterios se basan en los trabajos previos de los investigadores de la empresa Rijk Zwaan Iberica y apoyados por estudios previos (Dijkhuizen & Staub, 2002; Yuan et al., 2008a)

Caracteres vegetativos

- 1) **Precocidad.** Tiempo requerido por la planta para desarrollar el primer fruto (comercial). Se estima en días después del trasplante (DDT).
- 2) **Número de estructuras nodales** desarrolladas hasta alcanzar una longitud determinada, en este caso la altura útil del invernadero (2,25m).
- 3) **Longitud media de los entrenudos desarrollados.**
- 4) **Cabezas quemadas.** Susceptibilidad del ápice apical de la planta a una exposición intensa a la luz y/o altas temperaturas. También puede estar relacionado con la humedad relativa presente en el invernadero.
- 5) **Desarrollo de flor macho.** Capacidad de aparición de flores macho en los distintos genotipos expuestos a las mismas condiciones ambientales.
- 6) **Número de frutos en un nudo.** Este carácter se calculó como la media de frutos en cada nudo observada hasta la altura útil del invernadero (2,25m).
- 7) **Vigor de planta.** Crecimiento vegetativo primario de la planta en altura. Se trata de una valoración 1-5 tomando como referencia una altura determinada (2,25m) balizada por un marcador (alambre o rafia). El criterio de evaluación puede observarse en la Figura 4.
- 8) **Longitud del pedicelo.** Medida media de la longitud del pedicelo (cm) que conecta el fruto con el nudo y responsable del suministro floemático al fruto.

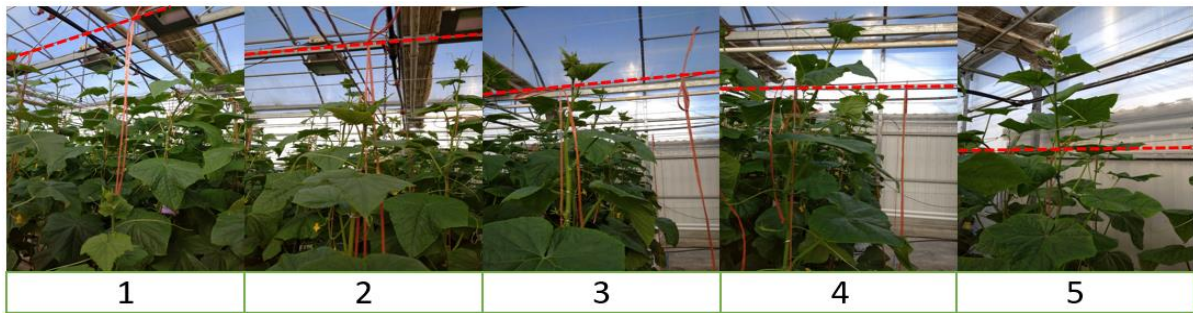


Figura 4. Niveles fenotípicos de vigor en planta de pepino medidos visualmente a 35 días tras el trasplante tomando como marca de referencia una altura de 2.25m balizado con una alambre (destacado en línea roja discontinua). Nivel 1, escaso crecimiento; nivel 2, crecimiento retrasado; nivel 3, crecimiento medio; nivel 4, crecimiento avanzado; nivel 5 crecimiento muy avanzado. En los niveles 1 y 2 el desarrollo vegetativo es más lento, mientras que en los niveles 4 y 5 este es muy acelerado.

Caracteres de fruto

9) **Peso.** Peso medio del fruto (g).

10) **Longitud.** Longitud media del fruto (cm).

11) **Diámetro.** Diámetro medio del fruto (cm).

12) **Cavidad seminal.** Diámetro medio de la cavidad seminal del fruto (cm).

13) **Color.** La escala de colores se ha tomado en base a una herramienta suministrada por la empresa hospedadora. Los valores de la escala van de 1-5, siendo de más claro a oscuro, y pueden apreciarse en la Figura 5.

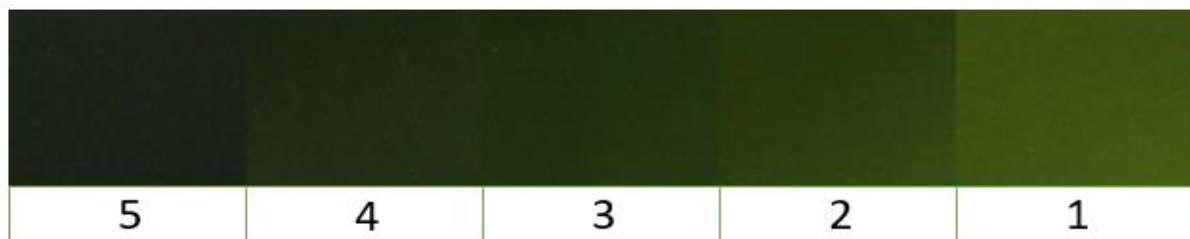


Figura 5. Representación gráfica de la escala colorimétrica facilitada por la empresa Rijk Zwaan Iberica para la evaluación fenotípica del color en fruto de pepino variedad Dutch.

14) **Cuello.** Se refiere a una cualidad anatómica caracterizada por el desarrollo del extremo anterior del fruto con un fenotipo alargado (estrecho en diámetro) o más redondeado (con mayor grosor). La valoración se tomó siguiendo una escala usada por la empresa y que puede visualizarse en la Figura 6.

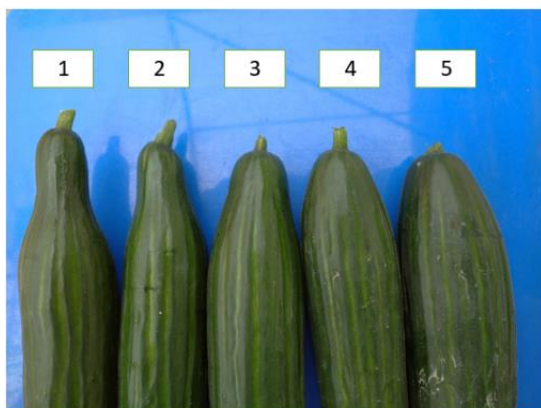


Figura 6. Representación gráfica del criterio de evaluación del carácter fenotípico cuello de fruto. El nivel 1 se corresponde con el fenotipo extremo conocido como "cuello de botella", mientras que el nivel 5 se corresponde con el fenotipo redondeado.

3.5) Análisis de datos (MapQTL6)

El mapeo de caracteres cuantitativos o QTL se emplea para el estudio de la herencia de caracteres complejos gobernados por varios genes en los que cada gen puede contribuir en distinta proporción al fenotipo final. En primer lugar es necesario disponer de un análisis fenotípico detallado de una población segregante, descendiente del cruzamiento entre dos líneas parentales con suficiente variación genética entre ellas o al menos para los caracteres que se quieren mapear. El tamaño poblacional debe ser lo mayor posible para incrementar la eficacia y la capacidad de detección de QTL.

Además del análisis fenotípico, para el mapeo de los QTL se hace necesario disponer de un mapa genético en el que se represente de forma ordenada y lineal un grupo de genes y/o marcadores genéticos dispuestos a lo largo del genoma y siendo la distancia entre los *loci* adyacentes proporcional a la frecuencia de recombinación entre los mismos. Para la identificación de QTLs es recomendable el uso de marcadores codominantes, es decir, aquellos capaces de discernir entre un genotipo homocigótico dominante, recesivo o una heterocigoto. Los marcadores moleculares utilizados para este estudio fueron marcadores del tipo SNP seleccionados por los mejoradores de la empresa Rijk Zwaan Iberica, los cuales cumplen estos requisitos. El genotipado se realizó en las instalaciones centrales y el genotipo de un número significativo de marcadores fue proporcionado para la realización de este TFM. Con los datos proporcionados se realizaron los análisis correspondientes usando el programa JoinMap y se pudieron generar mapas de cada población.

Los pasos llevados a cabo para poder realizar el análisis de QTLs fueron:

1. Se genotipo a los individuos para el marcador y en función del genotipo darles un valor genético respecto al carácter de interés.

2. Elaboración del mapa genético mediante el software “JoinMap5” identificando los distintos grupos de ligamiento donde se ubican los marcadores empleados y su posicionamiento.
3. Identificación del ligamiento entre marcadores y *loci* mediante el cálculo de las frecuencias de recombinación y su significancia estadística (puntuación LOD). Para ello se empleó el software “MapQTL6” mediante un análisis de permutación (1000 test de permutación para un nivel de confianza del 95%) que permite estimar la puntuación LOD umbral a partir de la cual se considera significativa la presencia de ligamiento y un análisis de mapeo de marcadores por intervalos (*Interval Mapping*).
4. Estimación de la proporción de la variabilidad fenotípica que es determinada por el QTL. Para un mismo carácter es frecuente que existan varios QTL y por ello habría que escoger el de mayor peso para que la MAS sea efectiva.
5. A fin de precisar los resultados obtenidos se procedió a emplear los potenciales QTL como cofactores (*Automatic Cofactor Selection*) y realizar un análisis por mapeo de intervalos compuestos (*MQM Mapping*).

En el mapeo por intervalos o “*Interval Mapping*” se calcula el conocido como perfil de verosimilitud QTL. Para cada posición del genoma (cada centiMorgan) se determina la probabilidad de la presencia de un QTL segregante respecto de la probabilidad de la ausencia de segregación de QTL en dicha posición (Hipótesis Nula). Esta razón de verosimilitud (LOD) es el logaritmo en base 10 del cociente de las dos probabilidades respectivas. Cuando este valor excede el umbral de significancia, se concluye una segregación de QTL, correspondiéndose la posición con el LOD más grande en el grupo de vinculación con la posición estimada para el QTL en el mapa. El programa usa el llamado “parámetro de tamaño de paso de mapeo” para ir de una posición a la siguiente entre *loci*, establecida en 1 cM (Van Ooijen, 2009; Van Ooijen, 2018).

El método *MQM Mapping* se basa en múltiples modelos, aunque para el análisis actual en MapQTL se limita su uso al de marcadores como cofactores de un modelo de aproximación de múltiples QTL con acciones genéticas aditivas (y dominantes opcionalmente). La aplicación completa de este modelo no es factible computacionalmente debido a que trataría de realizar una búsqueda multidimensional sobre los grupos de enlace. Este modelo se basa en buscar primero el posible QTL y posteriormente seleccionar los marcadores cercanos como cofactores, de forma que asuman el rol de QTL para el análisis posterior. Se lleva a cabo una búsqueda unidimensional del genoma sobre un QTL segregante como en el mapeo de intervalos, pero de

forma simultánea se van ajustando los cofactores. De esta manera, los cofactores reducen la varianza residual y aumenta subsecuentemente el poder de búsqueda de otros QTL segregantes (Van Ooijen, 2009; Van Ooijen, 2018).

3.6) Análisis de regresión y correlación

El análisis de regresión y correlación tiene como objetivo la construcción de modelos matemáticos capaces de describir la relación existente entre las variables de estudio (Ahmad et al., 2016).

El coeficiente de correlación (r) es una medida empleada para determinar las relaciones entre ciertos caracteres de interés (Ahmad et al., 2016). Resulta de especial interés cuando se trata de caracteres poligénicos que pudiesen encontrarse ligados entre sí (Augustina et al., 2013). Los principales componentes de la correlación son:

- La fuerza o grado que explica los resultados observados. Se mide como Nula ($r = 0$), débil ($0 < r < 0,3$), moderada ($0,3 < r < 0,6$) y fuerte ($0,6 < r \leq 1$).
- El sentido mide la variación que presentan los resultados respecto a los de las otras variables. Esta relación puede ser directa (positivo) o inversa (negativo).
- La forma de la línea de mejor ajuste define el trazado que mejor se ajusta a la nube de puntos generada al analizar distintas variables. Puede ser de tipo línea recta, curva monotónica o no monotónica.

Existen diversos coeficientes, los cuales se ajustan a la naturaleza de los datos. No obstante, el más conocido y empleado es el de Pearson ($\rho_{x,y}$). Este coeficiente es una medida de la dependencia lineal entre 2 variables aleatorias cuantitativas y siendo independiente de la escala de medida de las variables. Puede ser empleado siempre que las variables sean de tipo cuantitativo y continuo (Benesty et al., 2009), por lo que en este trabajo se han excluido los caracteres no continuos “cabezas quemadas” y “flor macho”. Se considera una correlación estadísticamente significativa con un nivel de confianza del 95% ($P\text{-valor} < 0,05$) y se ha empleado el software “Statgraphics” para su estimación.

4) Resultados

Con el fin de determinar el modo de herencia de caracteres de interés para los programas de mejora genética de pepino de Rijk Zwaan Iberica, los mejoradores han desarrollado poblaciones segregantes que han sido caracterizadas fenotípica y molecularmente tal y como se describe en los siguientes apartados.

4.1) Evaluación fenotípica

La evaluación de los caracteres vegetativos se realizó una vez que las plantas habían alcanzado el estado óptimo de desarrollo (33-37 días tras trasplante, DDT), mientras que los caracteres de relacionados con el desarrollo del fruto fueron evaluados entre los 40-45 DDT. Los valores medios para los distintos caracteres analizados en ambas poblaciones y sus respectivos controles son mostrados en la Tabla 2.

En las poblaciones analizadas se cumple que el tiempo medio para el desarrollo del primer fruto (comercial), se produce entre los 30-35 días después del trasplante. Entre los individuos de las poblaciones estudiadas se observaron diferencias, en su mayoría entre 1-2 días de más o de menos, pero destaca la población B con una desviación de casi 4 días y medio, reflejando una gran variabilidad genética para este carácter en dicha población.

El siguiente carácter con una mayor divergencia en las poblaciones es el número de entrenudos y consecuentemente la longitud de los entrenudos, pues parece existir una relación inversamente proporcional entre ambos. Esta deducción se alcanza de la observación que las plantas más compactas (mayor número de entrenudos) son las mismas que presentan una menor longitud media de los entrenudos. Las poblaciones F1-B y C1 son las que presentaron los mayores valores de divergencia para este carácter.

El vigor de planta, entendido como la capacidad de crecimiento vegetativa de la planta para un mismo periodo de tiempo permitió apreciar que las poblaciones A y F1-A mostraban valores ligeramente inferiores a la media general. Para el caso de la población F1-B su susceptibilidad al daño en el ápice apical afecto en gran medida a la estimación de este carácter, lo que casi hace a esta población poco representativa para el presente estudio.

Los datos correspondientes al número de frutos por estructura nodal se comportaron según lo esperado tanto para las poblaciones de estudio como los controles. Puede apreciarse que los valores superiores se corresponden a poblaciones genéticamente multipistilares (A, F1-A y C1).

Tabla 2. Evaluación fenotípica de los caracteres de estudio en las poblaciones BCI (A y B), en las F1 (F1-A y F1-B) y en los controles (C1, C2, C3 y C4). Se ha estimado el valor medio para los distintos caracteres, así como la desviación estándar.

Carácter	A	B	F1-A	F1-B	C1	C2	C3	C4
<i>Precocidad</i>	31,94 ± 1,99	34,50 ± 4,36	33 ± 0	34 ± 1,15	34 ± 2	35 ± 0	35,50 ± 1,77	35 ± 1,07
<i>Cabeza quemada (0-1)</i>	0,02 ± 0,13	0,11 ± 0,31	-	0,75 ± 0,50	-	-	-	-
<i>Nº de entrenudos</i>	24 ± 1,55	22,86 ± 1,32	23,25 ± 0,71	22,50 ± 2,12	23 ± 3,37	22,50 ± 0,58	22,75 ± 0,71	21,50 ± 0,53
<i>Longitud de los entrenudos (cm)</i>	9,41 ± 0,63	9,87 ± 0,57	9,68 ± 0,30	10,04 ± 0,95	9,96 ± 1,70	10 ± 0,26	9,89 ± 0,31	10,47 ± 0,26
<i>Vigor (1-5)</i>	3,47 ± 1,15	4,37 ± 0,96	3,88 ± 0,35	2,25 ± 1,26	4 ± 0	4,38 ± 0,48	4,38 ± 0,52	4,50 ± 0,53
<i>Flores macho</i>	0,23 ± 0,5	0,22 ± 0,42	0,13 ± 0,35	-	0,25 ± 0,5	0,25 ± 0,50	-	-
<i>Frutos en un nudo</i>	2,23 ± 0,68	1,05 ± 0,15	2,56 ± 0,42	1,00 ± 0	1,63 ± 0,25	1,13 ± 0,25	1 ± 0	1 ± 0
<i>Longitud de pedicelo (cm)</i>	7,66 ± 0,71	7,37 ± 0,56	8,19 ± 0,59	6,50 ± 0,41	7,38 ± 0,25	7,75 ± 0,65	7,50 ± 0,71	9,25 ± 0,53
<i>Peso (gr)</i>	377,48 ± 40,39	408,94 ± 42,46	349,46 ± 24,95	356,04 ± 29,19	411,98 ± 50,86	384,33 ± 47,09	383,95 ± 21,21	365,95 ± 22,51
<i>Longitud (cm)</i>	29,44 ± 1,55	30,74 ± 1,73	27,88 ± 0,80	29,67 ± 1,65	30,38 ± 1,38	30,75 ± 2,02	30,81 ± 0,84	31 ± 1,25
<i>Diámetro (cm)</i>	4,56 ± 0,20	4,55 ± 0,22	4,54 ± 0,15	4,25 ± 0,19	4,58 ± 0,30	4,48 ± 0,22	4,39 ± 0,17	4,38 ± 0,10
<i>Cavidad seminal (1-5)</i>	2,06 ± 0,17	1,98 ± 0,22	2,13 ± 0,13	1,65 ± 0,24	2 ± 0,16	1,85 ± 0,17	1,78 ± 0,20	1,90 ± 0,18
<i>Color (1-5)</i>	3,57 ± 0,78	2,83 ± 0,68	2,38 ± 0,52	2,50 ± 0,58	4,75 ± 0,50	4,50 ± 0,58	3,63 ± 0,52	4 ± 0
<i>Cuello (1-5)</i>	2,94 ± 0,81	1,94 ± 0,71	3 ± 0,80	1,38 ± 0,25	2,50 ± 0,41	1,25 ± 0,50	1,81 ± 0,26	2,19 ± 0,26

Por lo que respecta al carácter de longitud del pedicelo debemos mencionar la diferencia de tamaño entre las F1 parentales (siendo A superior y B inferior a la media general), y la escasa diferencia entre las poblaciones BC1. Esto puede sugerir que la heterocigosis resultante para ciertas regiones genómicas en el retrocruce tiende a estabilizar dicho carácter.

Tanto para el peso como para la longitud de los frutos la media poblacional se encuentra dentro del calibrado fijado por el mercado Europeo para la comercialización de esta variedad de pepino (Dutch) y por consiguiente los valores se consideran representativos para el estudio de la herencia del carácter. Podemos apreciar diferencias entre los frutos de las poblaciones que llegan hasta los 20-40 g y casi los 2 cm. Por su parte el diámetro de los frutos y el tamaño de la cavidad seminal apenas muestran divergencia entre las poblaciones, siendo por tanto caracteres muy estables dentro del fondo genético de esta variedad (Dutch). Sin embargo, se aprecia que las poblaciones A, F1-A y C1 desarrollan una mayor cavidad seminal coincidiendo esta diferencia con el menor grosor (diámetro) observado en las poblaciones F1-B, C2, C3 y C4.

En lo que respecta al color y a la presencia de cuello en los frutos, los controles han respondido de la forma esperada, lo que ha facilitado el fenotipado de las plantas F1 y BC1. Puede apreciarse una ganancia relativa de coloración con respecto a las F1, mientras que el cuello parece mejorar únicamente para la población B y manteniéndose estable para A.

Para este ensayo se analizaron dos caracteres cualitativos, correspondientes a la sintomatología de “cabeza quemada” y a la capacidad de desarrollo de flores macho. Para el primer carácter sólo se apreció incidencia en tres poblaciones (A, B y F1-B), mientras que el segundo se manifestó en 5 de ellas (A, B, F1-A, C1 y C2). Su expresión se dio en menos del 25 % de los individuos, salvo la sintomatología de cabeza quemada para la población F1-B, para la cual el 75% de sus individuos se vieron afectados.

4.2) Mapa de Ligamiento

El mapa de ligamiento para cada población fue elaborado con los marcadores polimórficos con potencial asociación a los caracteres estudiados en las líneas de pepino de la variedad Dutch. Ambos mapas son mostrados en la Figura 7. Se establecen 7 grupos de ligamiento, lo que se corresponde con la dotación cromosómica para la especie ($2n = 14$) (Fazio et al., 2003; Staub et al., 2007; Yuan et al., 2008a).

El mapa genético de la población A cubre 420,3 cM del genoma de pepino, mientras que B abarca una distancia de 411,5 cM distribuidos en 7 y 6 grupos de ligamiento, respectivamente, y con una distancia media entre marcadores de 10 cM para A y de 11,1 cM para B. Podemos

observar diferencias en la composición de los grupos de ligamiento para ambas poblaciones en especial si prestamos atención a los 16 marcadores comunes. Así se observa que los marcadores 12, 13, 14 y 15 forman parte de un mismo grupo de ligamiento en la población A, mientras que componen dos grupos de ligamiento en la población B, hecho que probablemente es debido a que la distancia genética entre los marcadores 14 y 15 es muy grande.

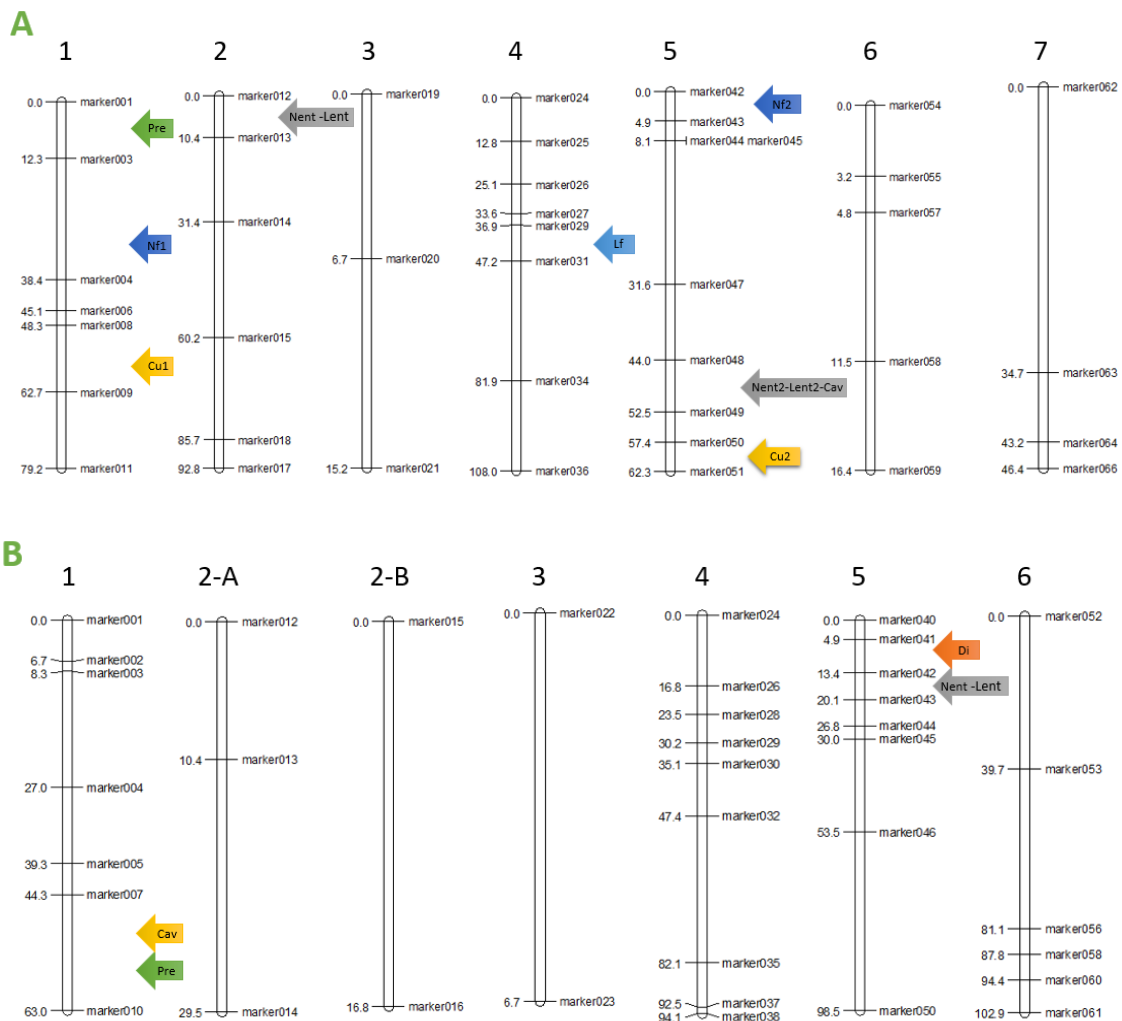


Figura 7. Mapa genético basado en las poblaciones BCI A (arriba) y B (abajo) del cruce entre la F1-A y F1-B con el parental común (PC). El nombre de los marcadores y su posición son mostrados a la derecha e izquierda del grupo de ligamiento, respectivamente. Las flechas situadas a la derecha de los grupos de ligamiento representan las posiciones de los QTL para los caracteres analizados. El mapa genético ha sido calculado usando el programa JoinMap5.

4.3) Análisis de QTLs

Una vez obtenido los dos mapas genéticos, se realizó una búsqueda de QTLs en dichos mapas usando los datos fenotípicos y genotípicos de cada una de las plantas que componían las poblaciones. Con tal fin se realizó un análisis de búsqueda de QTLs usando intervalos simples (*Interval Mapping*). Los resultados obtenidos permitieron la detección de 17 asociaciones entre marcadores y QTLs para algunos de los caracteres de estudio (precocidad, número y longitud

de entrenudos, número de frutos en un nudo, longitud, diámetro, cavidad seminal y cuello). De entre estos, 12 de los QTLs se han detectado en la población A y sólo 5 en la población B, siendo los caracteres de número de frutos en un nudo, longitud y cuello únicos para la población A, mientras que para el diámetro sólo se encontró asociación en la población B. Para la población B estos se concentran en los grupos de ligamiento 1 y 5, mientras que para la población A aparecen implicados otros grupos (1, 2, 4 y 5), pero siendo en el grupo 1 y 5 donde se aprecia una mayor agrupación (Tabla 2 y 3).

Los porcentajes de varianza fenotípica explicados por los QTLs identificados se encuentran entre 15-30%, salvo el carácter Precocidad (Pre) en la población B y el carácter Número de frutos en un nudo en la población A (Nf1), los cuales logran justificar un 60,7% y un 45,7%, respectivamente. Para los rasgos estudiados se identificaron valores aditivos variables (-1,66 – 6,77), lo que indica que los alelos de cada parental aportan de forma positiva o negativa al efecto agronómico determinado por el QTL. Así, los valores positivos indican que la presencia de esos alelos incrementa los valores para el carácter de estudio, indicando los negativos el efecto contrario.

Debe destacarse la existencia de asociaciones con una puntuación LOD muy superior al umbral establecida. Esto pone de manifiesto una gran probabilidad de ligamiento entre los marcadores y el QTL identificado. Los resultados más significativos se han obtenido para el carácter precocidad en la población B (Pre) y el número de frutos en un nudo en la población A (Nf1). Otros resultados de valor LOD muy elevado se han obtenido para el número y longitud de los entrenudos en la población A (Nent1, Nent2, Lent1 y Lent2).

Tabla 3. Análisis de locus simples ligados a QTL para los caracteres de estudio en la población A resultantes del cruce entre F1-A con el parental común (PC). Estimación calculada mediante el software MapQTL6.

Población A						
QTL	Intervalo marcador	Grupo de ligamiento (posición)	LOD umbral	LOD	Efecto aditivo	Variación fenotípica (%)
Precocidad (Pre)	Marker001-003	1 (0,0 – 12,3)	2,3	3,51	1,89	22,6
	Marker009-011	1 (62,6 – 79,2)	2,3	2,49	1,44	12,9
Número de entrenudos (Nent)	Marker048-049	5 (43,9 – 52,4)	2,2	4,72	-1,66	29,2
	Marker012-013	2 (0,0 – 10,3)	2,2	3,94	1,54	25
	Marker048-049	5 (43,9 – 52,4)	2,3	4,49	0,66	28

Longitud de entrenudos (Lent)	Marker012-013	2 (0,0 – 10,3)	2,3	3,71	-0,61	23,8
Número de frutos (Nf)	Marker003-004	1 (12,3 – 38,4)	2,2	8,49	0,91	45,7
	Marker042-043	5 (0,0 – 4,9)	2,2	3,59	0,65	22,8
Longitud de fruto (LF)	Marker029-031	4 (36,8 – 47,2)	2,3	4	1,56	25,4
Cavidad (Cav)	Marker048-049	5 (43,9 – 52,4)	2,4	2,5	-0,13	16,7
Cuello (Cu)	Marker008-009	1 (48,3 – 62,6)	2,3	2,69	-0,69	18,1
	Marker050-051	5 (57,4-62,3)	2,3	2,45	-0,66	16,6

Tabla 4. Análisis de locus simples ligados a QTL para los caracteres de estudio en la población B resultantes del cruce entre F1-A con el parental común (PC). Estimación calculada mediante el software MapQTL6.

Población B						
QTL	Intervalo marcador	Grupo de ligamiento (posición)	LOD umbral	LOD	Efecto aditivo	Variación fenotípica (%)
Precocidad (Pre)	Marker007-010	1 (44,2 – 62,9)	2,3	12,99	6,77	60,7
Número de entrenudos (Nent)	Marker042-043	6 (13,4 – 20,1)	2,4	2,89	-1,14	18,8
Longitud de entrenudos (Lent)	Marker042-043	6 (13,4 – 20,1)	2,3	2,99	0,49	19,3
Diámetro (Di)	Marker041-042	6 (4,9 – 13,4)	2,2	2,38	-0,17	15,7
Cavidad (Cav)	Marker007-010	1 (44,2 – 62,9)	2,3	2,38	-0,17	15,7

Con el objetivo de precisar aún más los resultados obtenidos se procedió a utilizar los marcadores identificados como QTL como cofactores (después de realizar un análisis del tipo *Automatic Cofactor Selection* fueron esos mismos marcadores los seleccionados) y realizar un análisis por mapeo de intervalos compuestos (*MQM Mapping*). En las Figuras (8-18) se representa gráficamente la puntuación LOD obtenida para los marcadores con respecto al límite

establecido por la LOD umbral. Observándose los reajustes de cálculo de forma sencilla en la comparación de métodos con y sin cofactor. De forma general se validan los resultados obtenidos en el análisis de intervalos sencillos, pero en el caso del carácter “Precocidad” para la población A se detecta la presencia de otro QTL en el mismo grupo de ligamiento.

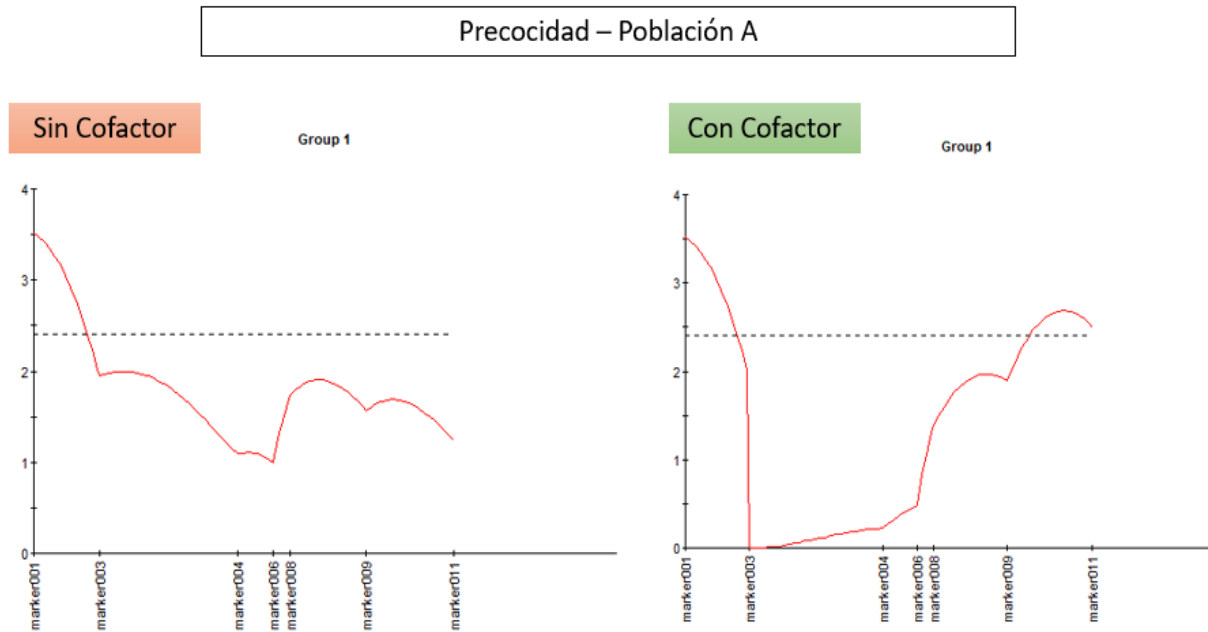


Figura 8. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter precocidad en la población A. Gráfica obtenida mediante el software MapQTL6.

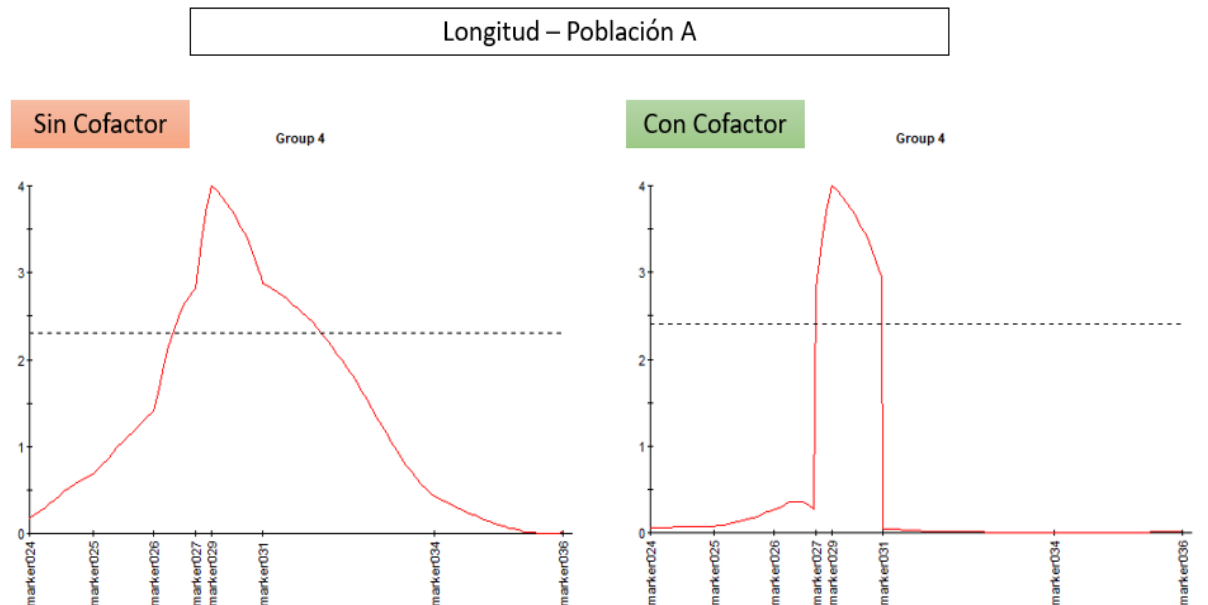


Figura 9. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter longitud en la población A. Gráfica obtenida mediante el software MapQTL6.

Número de entrenudos – Población A

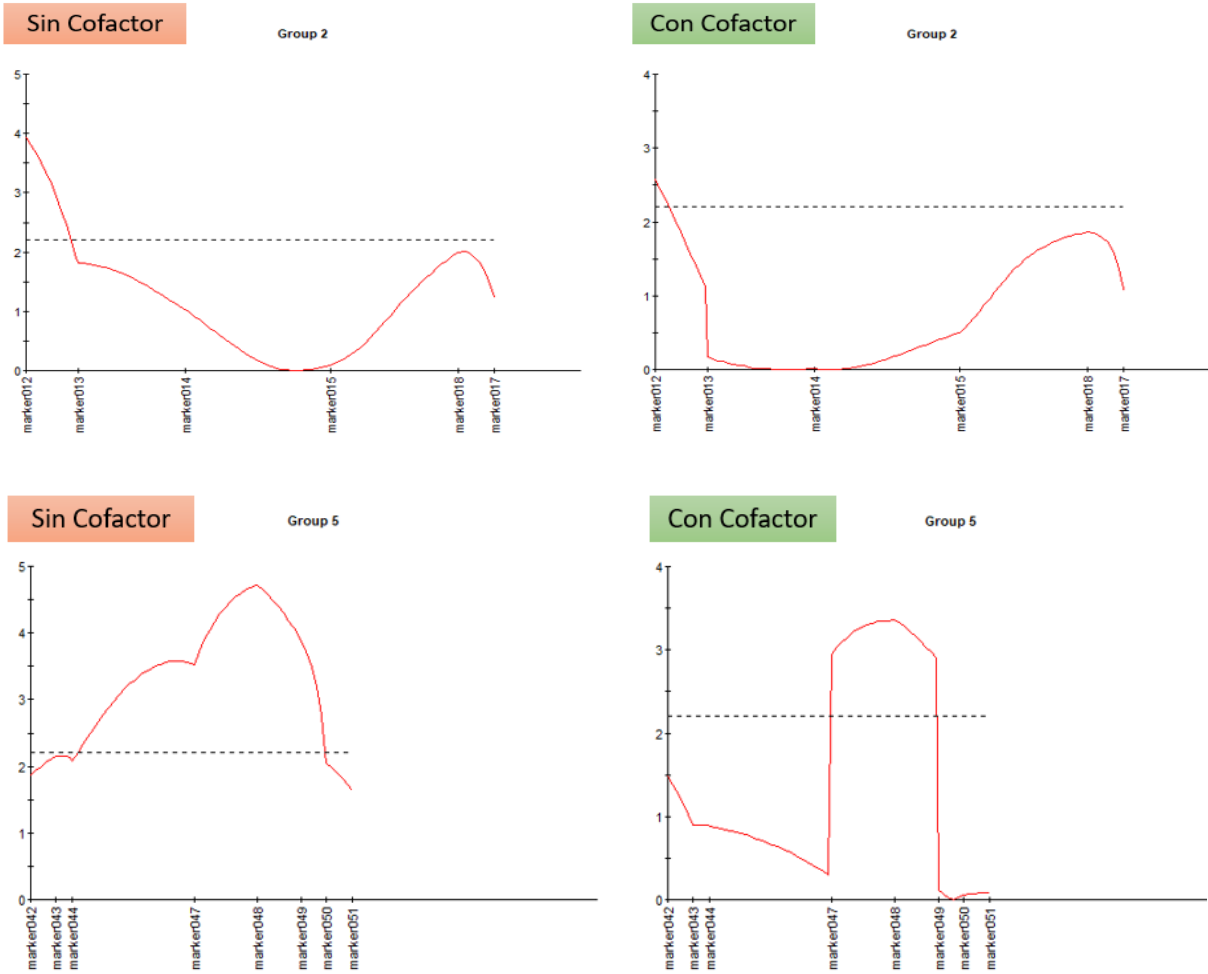
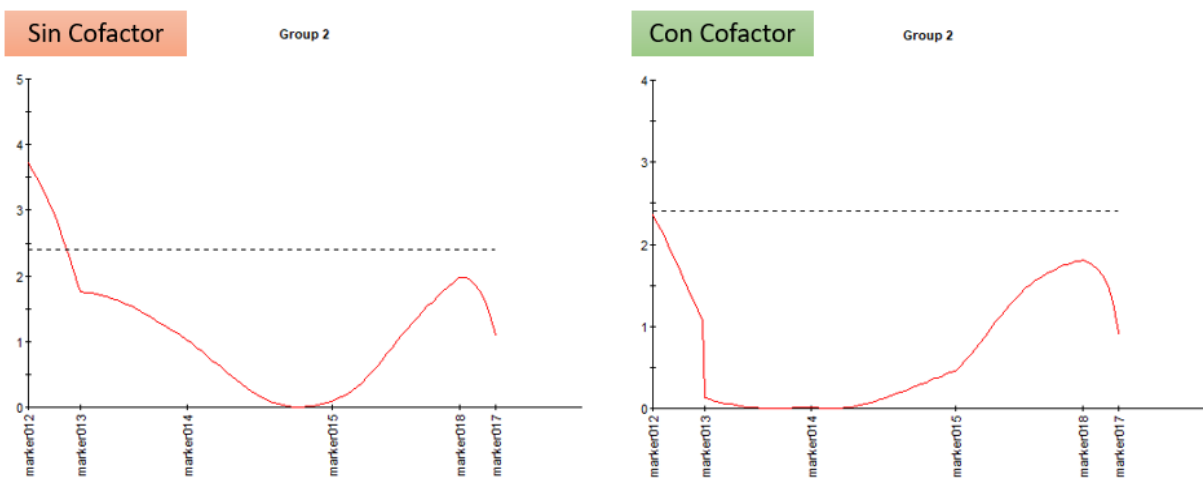


Figura 10. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter número de entrenudos en la población A. Gráfica obtenida mediante el software MapQTL6.

Longitud de entrenudos – Población A



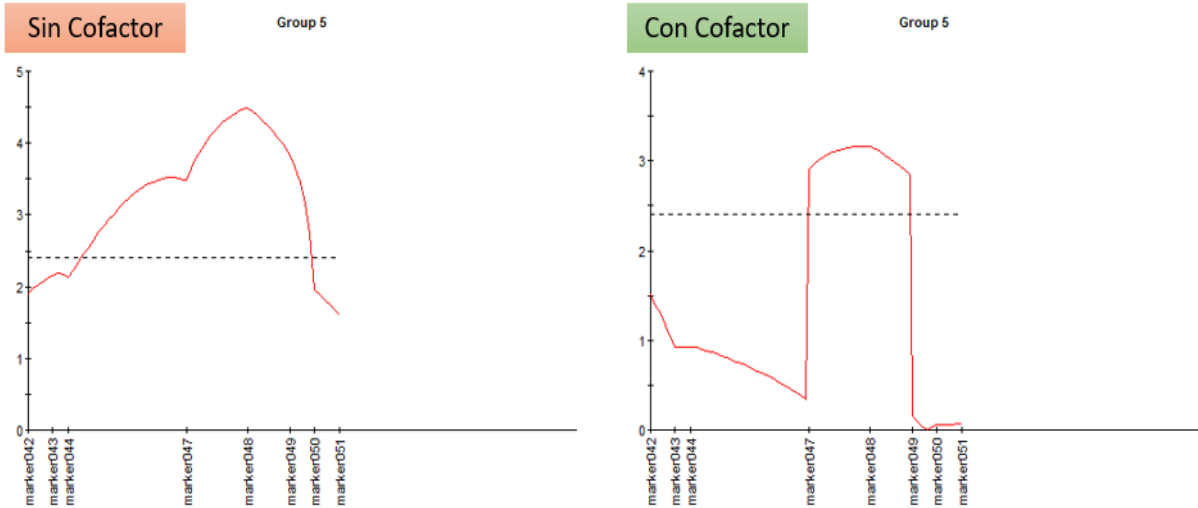


Figura 11. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter longitud de entrenudos en la población A. Gráfica obtenida mediante el software MapQTL6.

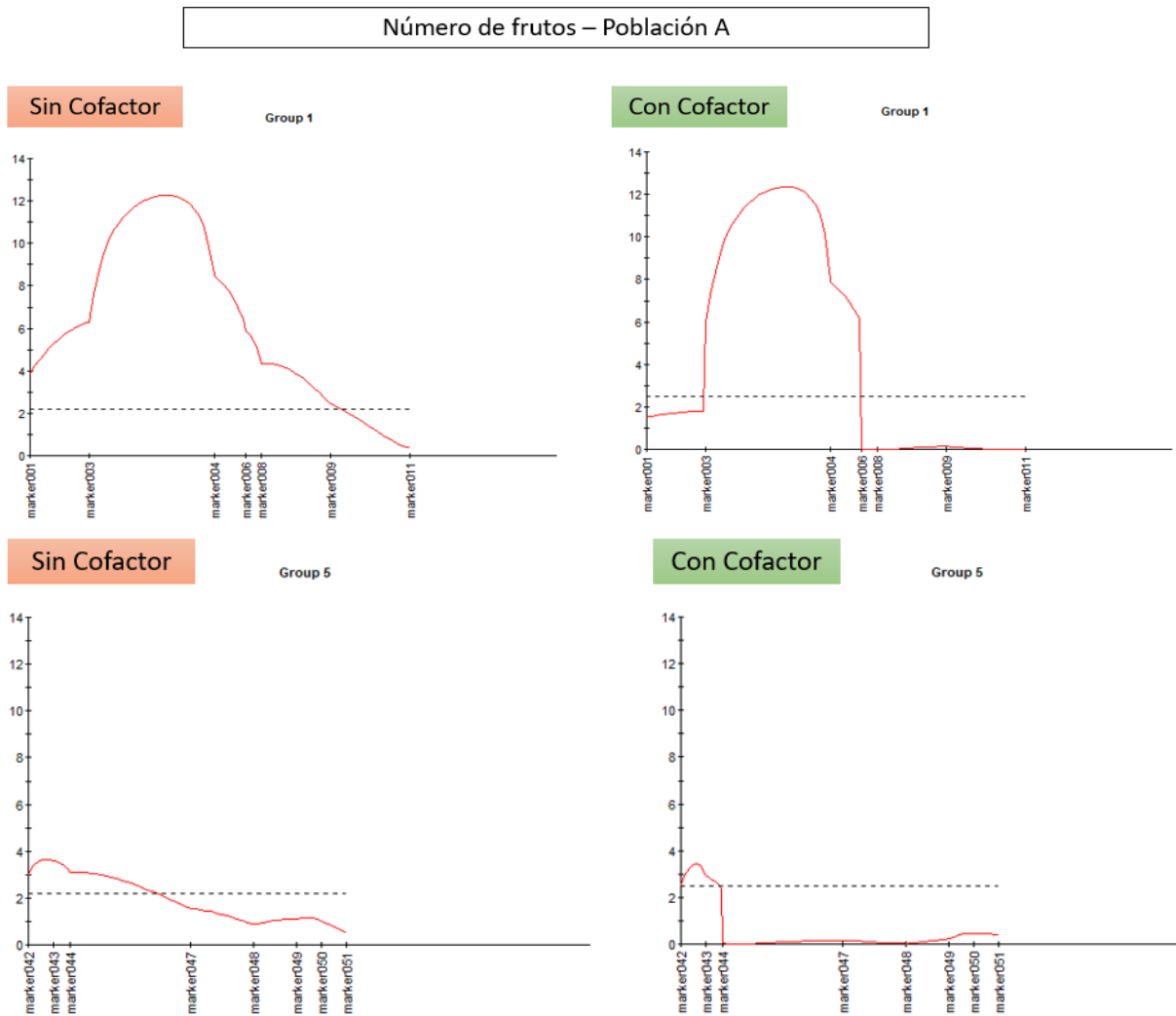


Figura 12. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter número de frutos en la población A. Gráfica obtenida mediante el software MapQTL6.

Cavidad seminal – Población A

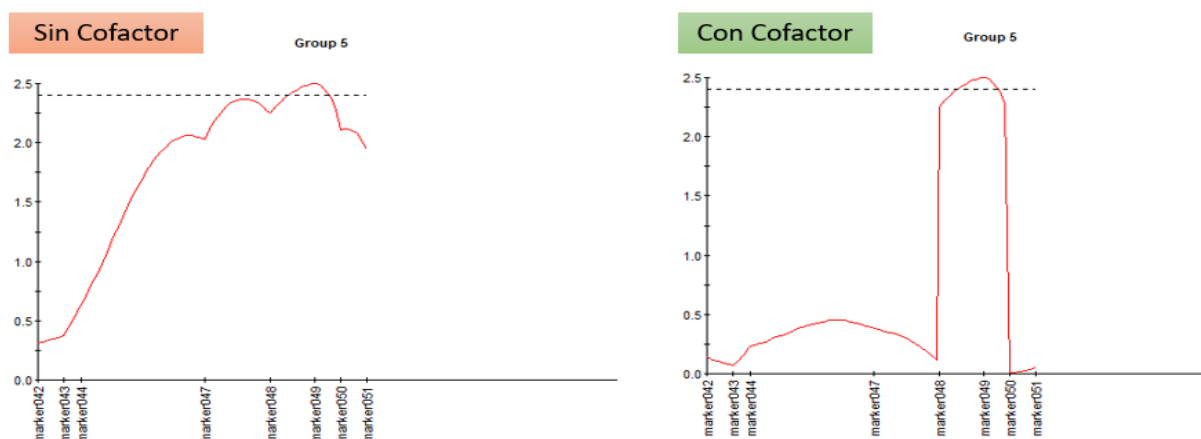


Figura 13. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter cavidad seminal en la población A. Gráfica obtenida mediante el software MapQTL6.

Precocidad – Población B

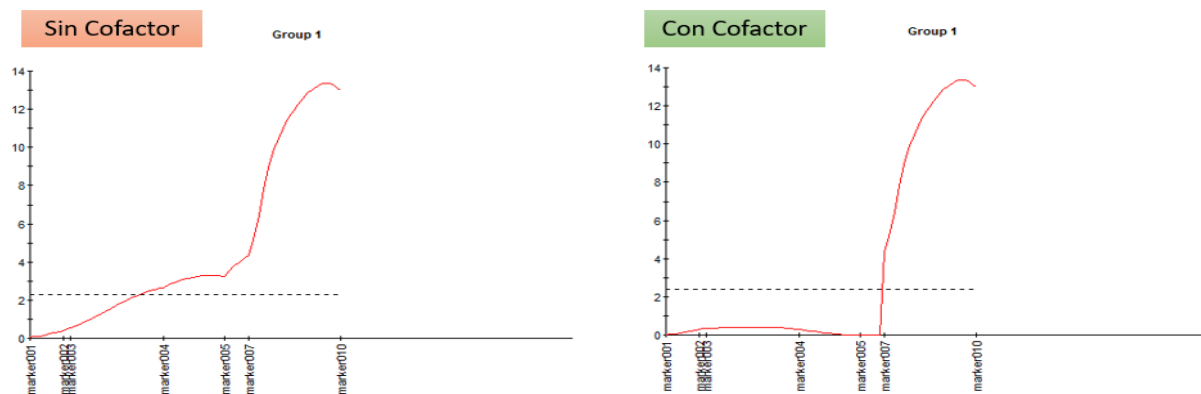


Figura 14. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter precocidad en la población B. Gráfica obtenida mediante el software MapQTL6.

Número de entrenudos – Población B

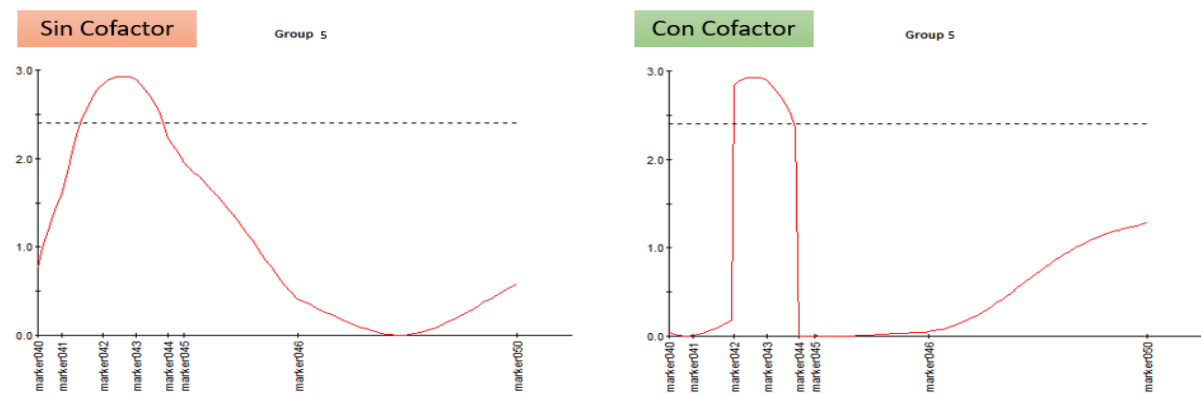


Figura 15. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter número de entrenudos en la población B. Gráfica obtenida mediante el software MapQTL6.

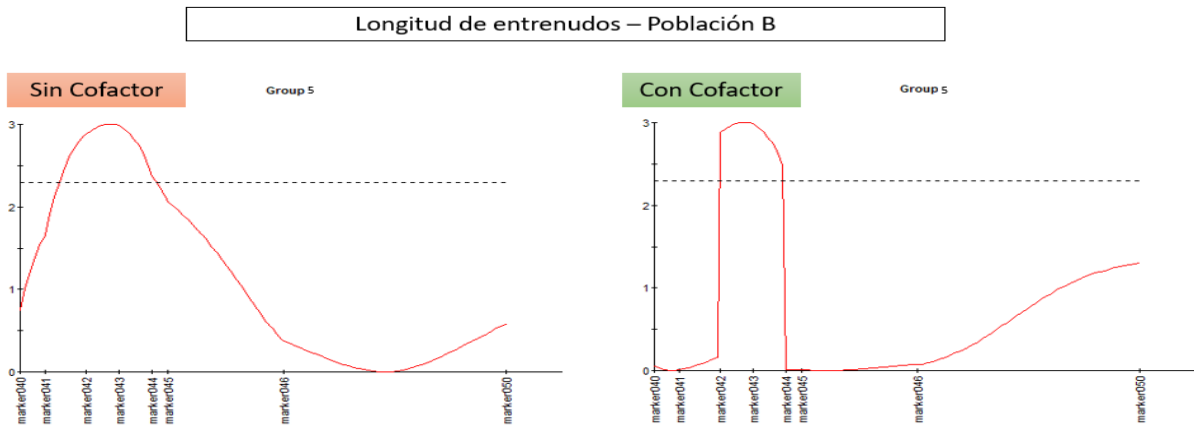


Figura 14. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter longitud de entrenudos en la población B. Gráfica obtenida mediante el software MapQTL6.

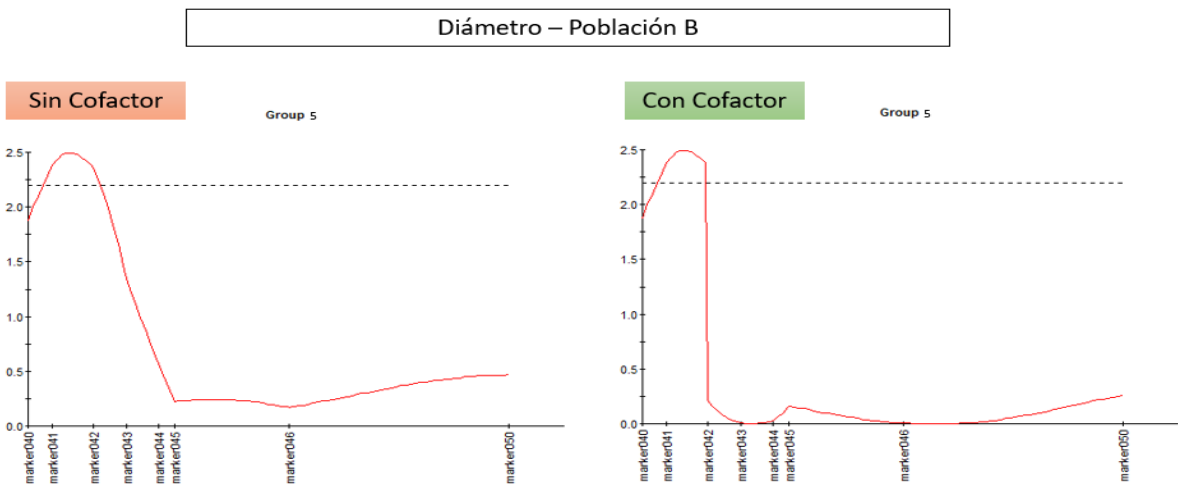


Figura 15. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter diámetro en la población B. Gráfica obtenida mediante el software MapQTL6.

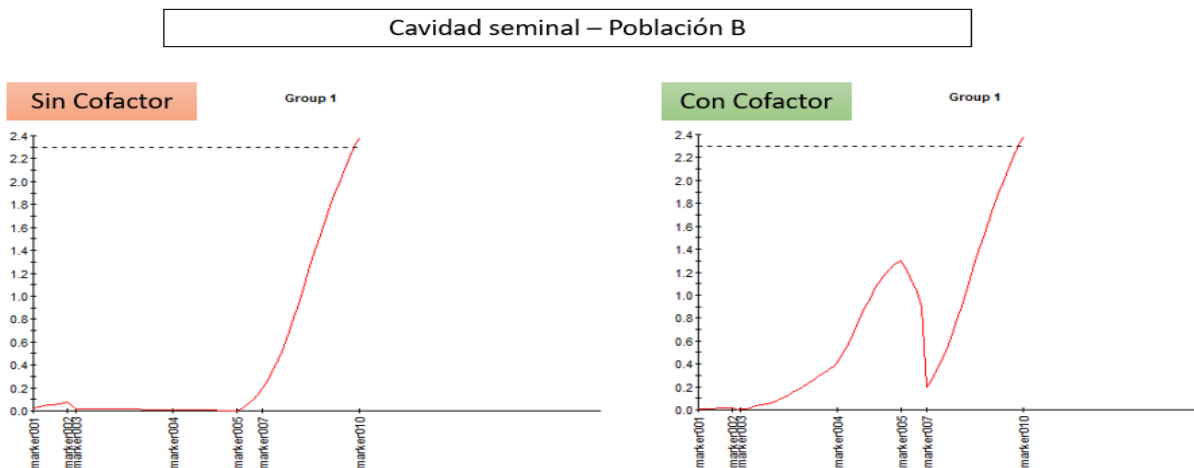


Figura 16. Representación gráfica del análisis de QTL mediante el método de mapeo por intervalos o “Interval Mapping” (izquierda) y el método MQM Mapping (derecha) para el carácter cavidad seminal en la población B. Gráfica obtenida mediante el software MapQTL6.

4.4) Análisis de correlación

Las Tablas 5 y 6 muestran los resultados obtenidos del estudio de las correlaciones existentes entre los distintos caracteres analizados en las dos poblaciones caracterizadas. Las Tablas muestran los coeficientes de correlación obtenidos ($\rho_{x,y}$) entre todas las combinaciones posibles resultantes del análisis de los distintos fenotipos analizados como variables cuantitativas. Para aquellos análisis cuyo P-valor fuese inferior al nivel de significación establecido (α) se considera un resultado estadísticamente significativo para la existencia de correlación. Esto representa que el resultado obtenido es más inusual que el esperado según nuestra hipótesis nula (no hay correlación) y se denota con ”*”.

Tabla 5. Análisis de correlación entre los diferentes caracteres de estudio para la población A. Muestra el coeficiente de correlación de Pearson entre dos caracteres obtenido mediante el software “Statgraphics Centurion XVI.I.

Población A	Precocidad	Nº entrenudos	Long. entrenudos	Vigor	Frutos/nudo	Long. pedicelo	Peso	Longitud	Diámetro	Cavidad seminal	Color	Cuello
Precocidad												
Nº entrenudos	* -0,256 *											
Long. entrenudos	0,235	* -0,996 *										
Vigor	0,226	* -0,431 *	* 0,409 *									
Frutos nudo	* 0,306 *	* -0,547 *	* 0,537 *	0,190								
Long. pedicelo	0,173	* -0,461 *	* 0,469 *	* 0,447 *	0,148							
Peso	* -0,306 *	0,160	-0,140	-0,144	-0,162	0,025						
Longitud	-0,209	-0,025	0,041	-0,153	-0,126	0,077	* 0,650 *					
Diámetro	* -0,315 *	0,252	-0,243	-0,165	-0,068	-0,038	* 0,749 *	0,154				
Cavidad seminal	-0,200	0,293 *	-0,299 *	0,016	-0,215	-0,114	0,200	-0,205	* 0,481 *			
Color	* -0,298 *	* 0,446 *	* -0,439 *	* -0,417 *	-0,217	* -0,356 *	0,188	0,187	0,193	* 0,286 *		

Cuello	-0,104	0,106	-0,094	-0,180	* -0,315 *	-0,002	0,036	-0,171	0,189	0,221	-0,003	
--------	--------	-------	--------	--------	------------------	--------	-------	--------	-------	-------	--------	--

Tabla 6. Análisis de correlación entre los diferentes caracteres de estudio para la población A. Muestra el coeficiente de correlación de Pearson entre dos caracteres obtenido mediante el software "Statgraphics Centurion XVI.I.

Población B	Precocidad	Nº entrenudos	Long. entrenudos	Vigor	Frutos nudo	Long. pedicelo	Peso	Longitud	Diámetro	Cavidad seminal	Color	Cuello
Precocidad												
Nº entrenudos	-0,118											
Long. entrenudos	0,122	* -0,997 *										
Vigor	-0,073	* -0,283 *	* 0,275 *									
Frutos nudo	-0,054	-0,104	0,096	0,219								
Long. pedicelo	0,146	-0,119	0,120	* 0,250 *	* 0,367 *							
Peso	-0,171	0,084	-0,084	0,104	-0,117	-0,054						
Longitud	-0,061	-0,073	0,078	0,144	-0,059	0,105	* 0,706 *					
Diámetro	* -0,336 *	0,104	-0,107	0,122	0,012	-0,022	* 0,742 *	0,224				
Cavidad seminal	* -0,446 *	0,056	-0,060	0,185	0,115	-0,067	* 0,407 *	0,172	* 0,623 *			
Color	-0,047	0,243	* -0,249 *	-0,187	-0,004	0,229	0,118	0,086	0,167	0,172		
Cuello	-0,173	0,064	-0,059	0,044	-0,009	* -0,257 *	0,007	-0,211	0,196	0,127	-0,240	

Como cabría esperar, no existen correlaciones entre la mayoría de caracteres analizados en las dos poblaciones caracterizadas. Sin embargo, se han detectado 24 correlaciones entre los caracteres de la población A (5 débiles, 16 moderados y 3 fuertes) y 13 para la B (5 débiles, 3 moderados y 5 fuertes), siendo estos resultados mayores de los esperados. Como ya se mencionó previamente, la mayor correlación detectada ha sido entre los caracteres "longitud de entrenudos" y "número de entrenudos" que muestran correlaciones negativas pero superiores a

0,99 en ambas poblaciones. También se ha detectado una correlación positiva y superior a 0,65 en ambas poblaciones entre los caracteres “longitud”, “diámetro” y “peso” de los frutos.

5) Discusión

5.1) Evaluación fenotípica

Si atendemos a la variabilidad a nivel de familias segregantes, todos los descendientes de un cruzamiento son el resultado de la combinación de la información genética que portan sus progenitores. Por ello, como consecuencia de los procesos de segregación independiente de los cromosomas y de la recombinación genética, cada individuo resulta en una entidad genética única. La variabilidad a nivel de individuos representa que todos los miembros de una especie difieren entre sí. Estas características se refieren tanto a características morfológicas como a aspectos fisiológicos. El origen de esta variabilidad puede ser muy variado, siendo en última instancia la combinación de los siguientes factores: genética, epigenética e interacciones de estos factores con el ambiente.

La diversidad genética de una población puede ser calculada como la variedad de alelos y genotipos presentes en una población y que, frecuentemente, se refleja en las diferencias morfológicas, fisiológicas y de comportamiento entre los individuos. Su importancia reside en ser la base de la adaptabilidad de la especie, es decir, determina la capacidad de responder ante los cambios ambientales (potencial evolutivo y eficacia biológica). Diversos estudios han determinado que el grado de diversidad genética en *C. sativus* es relativamente bajo (3-8%) en comparación con las otras especies del género *Cucumis* (10-25%). Este hecho es debido a la intensa selección artificial que se vienen realizando sobre un reducido material genético durante la mejora de esta especie (Cavagnaro et al., 2010; Behera et al., 2011).

No puede dejar de considerarse que las poblaciones evaluadas en este TFM forman parte de programas de mejora genética de una empresa. Por ello, la evaluación de los caracteres estudiados son de interés para el programa de mejora del que forman parte las líneas y que está focalizado para un mercado concreto, en este caso el Europeo, al que se destinan los pepinos tipo Dutch. Dicho mercado valora de forma positiva un fruto de color oscuro, de cuello redondeado, con un peso comprendido entre 350-450 g y una longitud entre 25-30 cm. Con respecto a los caracteres vegetativos, las plantas precoces, compactas y monopistilares son de especial interés para los productores por su mayor capacidad productiva relativa y el abaratamiento de las labores de recolección con respecto a las multipistilares (Yuan et al., 2008b).

Un análisis de los resultados del fenotipado de la Tabla 2 muestra que la población A es más precoz, compacta y menos multipistilar que su parental (F1-A). Además, genera frutos de mayor longitud y peso, los cuales son de una coloración más oscura. En su conjunto esta población ha respondido según lo esperado tras el cruzamiento con el PC, el cual presentaba cualidades más deseables para los caracteres mencionados. Por su parte, la población B presenta unos valores ligeramente más favorables a los de su parental (F1-B), salvo para el vigor, traduciéndose en una mayor velocidad de desarrollo. Hay que resaltar la varianza observada para el carácter precocidad de la población B, que debe estar ocasionado por la existencia de una gran divergencia entre los individuos de la población, lo que a su vez podría ser debido a la segregación de numerosos genes que controlan este carácter en esta población.

En ambas poblaciones, y para todos los caracteres analizados, es apreciable la existencia de un rango de desviación estándar, lo cual quiere decirnos que para los 64 individuos de las poblaciones la intensidad de la expresión de los caracteres analizados no ha sido la misma y han dado lugar a combinaciones únicas. Este fenómeno de segregación es la base para la identificación de los potenciales QTL que rigen estos caracteres cuando se analiza la correlación existente entre el fenotipo de un individuo, el genotipo de ese individuo para cada marcador analizado y las relaciones de ligamiento existentes entre los marcadores analizados y visibles en el mapa genético de cada población.

Por otro lado, la sensibilidad de los meristemas apicales a las condiciones ambientales extremas y la influencia de estas mismas a la hora de estimular el desarrollo de flores macho, son dos características que deben eliminarse del fondo genético ya que conllevan una reducción de producción. Estas cualidades parecen ser aportadas en parte por el PC, ya que la población A manifiesta el carácter “cabeza quemada” y la población B el carácter “flores macho”, estando estas ausentes para sus respectivos parentales F1. Sin embargo estas deducciones deberán ser contrastadas en futuros ensayos, tal y como se ha demostrado previamente para esos caracteres (Mibus & Tatlioglu., 2004).

Finalmente, al estudiar los controles y las poblaciones F1 podemos determinar que pese a existir influencia ambiental en la determinación de estos caracteres, esta no parece haberse ejercido con un peso significativo en el presente trabajo. Esta conclusión ha sido posible por la inclusión en el trabajo de poblaciones estables genéticamente como controles. En dichas poblaciones, la variación observada debe estar ocasionada únicamente por las condiciones ambientales. Sin embargo, tanto la media poblacional como la varianza observada en esas poblaciones es

equiparable con las poblaciones segregantes de estudio. Estos datos nos permiten estimar la heredabilidad de los caracteres estudiados, proporcionando una idea del grado de control genético para la expresión de un rasgo particular y la confiabilidad del fenotipo para predecir su valor genético. Estos resultados podrán aportar información base para los futuros programas de mejora, relacionándose altos valores de heredabilidad (h^2) con una baja influencia ambiental lo que facilita la labor de los mejoradores (Dijkhuizen & Staub, 2002; Ullah et al., 2012).

En ambas poblaciones ha sido posible obtener individuos con características fenotípicas más convenientes o deseables que los parentales, lo que se conoce como segregación transgresiva (Cubero, 2013). Cabe mencionar que las plantas alógamas al reproducirse mediante fecundación cruzada son fuertemente heterocigóticas (elevada frecuencia media de heterocigotos por *locus*). La autofecundación artificial suele conllevar una pérdida de vigor por degeneración por consanguinidad al reducir la variabilidad genética de la población, llegando en casos extremos a generar plantas estériles. Estas poblaciones hay que mejorarlas tratando de modificar su composición genética a favor de los caracteres considerados como favorables y manteniéndolas en polinización cruzada. Estos cruces se traducen en híbridos de interés que muestran lo que se conoce como vigor híbrido o heterosis. La búsqueda de marcadores que permitan predecir la aparición del vigor híbrido o heterosis es del mayor interés básico y aplicado (Fan et al., 2006).

5.2) Mapa genético

Con los resultados disponibles ha sido posible la generación de un mapa genético de cada una de las poblaciones analizadas. Ambos mapas presentan diferencias significativas en la distribución de los marcadores en los distintos grupos de ligamiento. Esto puede generar una cobertura no uniforme del genoma del pepino para la búsqueda de potenciales QTL. Sin embargo, la selección de los marcadores SNPs que se han podido incluir en el mapa había llevada a cabo basándose en resultados previos de la empresa, que indicaban posibles relaciones con los caracteres estudiados.

Cuanto mayor sea el número de marcadores empleados en el mapa genético, mayor será la resolución del mismo y la probabilidad de lograr identificar marcadores ligados a los caracteres de interés (Yuan et al., 2008a). Si consideramos como más preciso o veraz el mapa de la población A por contener el mayor número de marcadores, y ser el que mayor distancia genética abarca, podemos observar que el mapa genético de la población B contiene únicamente 6 grupos de ligamiento (1, 2-A, 2-B, 3, 4, 5 y 6), encontrándose el grupo de ligamiento 2 dividido en dos

grupos. Para este grupo de ligamiento encontramos 4 marcadores comunes (markers012, 013, 014, 015), lo cual nos puede llevar a pensar que el polimorfismo de estos *loci* no es lo suficientemente significativo como para determinar ligamiento para esta población. Por su parte la población A presenta un grupo de ligamiento único, el 7, conformado por 4 marcadores (markers062, 063, 064, 066), del que no se encuentra equivalente en la población B. Este hecho puede ser debido a la relativa escasez de marcadores analizados o a la falta de marcadores de referencia que facilitasen el establecimiento consensuado de los grupos de ligamiento (Yuan et al., 2008a). Sin embargo, el número de marcadores moleculares que la empresa ha puesto a disposición de este trabajo ha sido limitado. En cualquier caso, los resultados obtenidos aunque positivos deberán ser contrastados en estudios posteriores para, con un mayor número de marcadores e individuos, establecer de forma más precisa mapas genéticos de ambas poblaciones. No obstante, los mapas obtenidos en este TFM podrían ser utilizados como primera versión de los futuros mapas genéticos definitivos.

5.3) Análisis de QTLs

La variación observada para algunos de estos caracteres depende de factores ambientales, tales como las técnicas agrícolas, pero también en parte por los factores genéticos aun no definidos. La dificultad del proceso reside principalmente en que la variabilidad fenotípica que observamos se debe a diferencias genéticas entre los individuos de la población y a las diferencias por efecto del ambiente en el que se encuentran. Por este motivo, la evaluación fenotípica del carácter se considera tan importante y debe realizarse en condiciones ambientales controladas para reducir su impacto sobre el fenotipo o en diferentes condiciones ambientales para estimar la proporción de variabilidad fenotípica debida al ambiente. Mediante estas estrategias se logra identificar con precisión los factores genéticos que regulan el carácter de interés.

En este estudio, la evaluación de 2 poblaciones BC1 en condiciones de cultivo controladas bajo invernadero permitieron la detección de 17 asociaciones entre marcadores y QTLs para algunos de los caracteres de estudio (precocidad, número y longitud de entrenudos, número de frutos en un nudo, longitud, diámetro, cavidad seminal y cuello). De entre estos, 12 QTLs se han detectado en la población A y 5 en la población B, siendo los caracteres de número de frutos en un nudo, longitud y cuello únicos para la población A, mientras que para el diámetro sólo se encontró asociación en la población B.

La identificación de QTL ligados a marcadores compartidos en distintas poblaciones sugiere que los QTLs analizados son más confiables (Fan et al., 2006). La asociación con uno de estos marcadores sin repetirse en la otra población nos puede llevar a pensar que el fondo genético entre estas poblaciones presenta diferencias para las regiones analizadas. Es decir, la evolución seguida por las poblaciones puede haber generado eventos de inserción o delección de regiones genómicas o mutaciones puntuales. Otra posibilidad sería la falta de datos suficientes como para lograr encontrar dicha asociación, como por ejemplo que el número de individuos analizados no es lo suficientemente extenso como para determinar con exactitud ese resultado (Mansur et al., 1996)

La precocidad es un carácter muy estudiado por los mejoradores puesto que, cuando una planta empieza a ser productiva antes, suele conllevar una mayor capacidad productiva con respecto a otras con similar dinámica de desarrollo de frutos. Para ambas poblaciones se han detectado QTL en el primer grupo de ligamiento (Marker001, 010 y 011), alcanzando un valor LOD seis veces superior al umbral para la población B. Esto confirma la existencia de un QTL implicado en el desarrollo de los frutos en el grupo de ligamiento 1. Esto se apoya en los resultados obtenidos para la población A, la cual también encuentra asociaciones para ese grupo, siendo una de ellas próxima a la posición génica detectada en B. Sin embargo, podemos observar una diferencia de variación fenotípica explicable, esto podría deberse a que para la población A este carácter ha pasado a estar regulado por más regiones (aumentando su complejidad y reduciendo el peso genético). Por otra parte, el estudio del Marker010 en la población B puede ayudar a comprender mejor la regulación genética para dicho carácter, gracias a que por sí mismo es capaz de explicar el 60% de la variación fenotípica observada (Dijkhuizen & Staub, 2002; Yuan et al., 2008b).

Para el número de entrenudos se valora positivamente un mayor número con respecto a una longitud determinada (planta compacta). En general, a mayor número de estructuras nodales mayor es la producción, ya que por cada nudo la planta genera una o varias flores. Este carácter está ligado tanto a la capacidad productiva de la planta como al correcto desarrollo de los frutos y la calidad de los mismos. Los resultados indican la presencia de 2 asociaciones para la población A (Marker012 y 048) y una para la población B (Marker043). Estos resultados son respaldados por el análisis de la longitud de entrenudos, debido a la relación existente entre ambos caracteres. Para ambas poblaciones aparece un QTL en el grupo 5, lo que puede sugerir una región génica involucrada en el desarrollo vegetativo en dicho grupo conservada para ambas poblaciones.

Otro carácter para el que se ha encontrado asociación en ambas poblaciones es la cavidad seminal de los frutos. Este carácter determina la cantidad de endo/mesocarpo del fruto y está ligado a características organolépticas. En este caso se han localizado en grupos de ligamiento distinto 1 (Marker010) y 5 (Marker049). Esto puede sugerir que el desarrollo del fruto es un proceso complejo y en el que participan distintos genes, lo que contribuye a su regulación. Podemos ver que estos marcadores alcanzan a explicar un 15-18% de la varianza, por lo que existe un 70% de la varianza del carácter que desconocemos. Cabe la posibilidad de que los QTLs identificados en este trabajo se correspondan con alguno de los 8 identificados por el grupo de Yuan (2008a, 2008b), más concretamente con 2 de ellos que presentan un porcentaje de variación (r^2) aproximado, aunque la falta de marcadores moleculares comunes nos impide acotar más su identificación.

El número de frutos por nudo parece estar regulado principalmente por dos QTL que explicarían en su conjunto el 68% de la varianza fenotípica observada en la población A: Marker004 (1) y Marker043 (5). Estos resultados pueden encontrar concordancia con los obtenidos por los investigadores Dijkhuizen & Staub (2002). En la población B, la falta de variabilidad significativa para este carácter dificulta enormemente su análisis, ya que como se ha comentado se requiere de variación genética para los caracteres de estudio en la población segregante. La identificación de esta asociación resulta de utilidad, ya que generalmente se prefiere plantas monopistilares debido al ahorro que supone en mano de obra agrícola y el beneficio de redistribución de recursos para la planta. Por tanto, es más que probable que la empresa use esta información en sus futuros programas de mejora genética de pepino, para evitar en las nuevas variedades los alelos del QTL que controlan la formación de este carácter.

El resto de caracteres para los que se han identificado QTLs están implicados en el desarrollo y morfología del fruto (longitud, diámetro y cuello). Así, se han encontrado QTLs en 3 grupos de ligamiento distintos: 1 (Marker009), 4 (Marker029) y 5 (Marker041 y 051). Estos marcadores explican entre el 15-25% de la varianza fenotípica observada, lo que nos ayuda a comprender la cantidad de genes y las complejas rutas de regulación involucradas para el desarrollo y maduración de un fruto. Algunos autores han observado una menor consistencia en la determinación de QTL asociado a la forma del fruto al realizar repeticiones de sus ensayos. Estos caracteres están muy condicionados a la etapa fisiológica de desarrollo en el que se encuentren los frutos, presentando una dificultad añadida para su identificación (Dijkhuizen & Staub, 2002). Los QTLs identificados en este TFM pueden corresponderse con los identificados por Serquen y colaboradores (1997) o los del equipo de Yuan (2008a, 2008b). Sin embargo, de

nuevo nos encontramos con la limitación impuesta por los marcadores analizados y que no se han incluido marcadores moleculares de los mapas genéticos previamente publicados (Fazio et al., 2003; Staub et al., 2007; Yuan et al., 2008a).

Puede concluirse por tanto que para el presente estudio la falta de información limita afinar la localización de los QTL localizados con respecto del genoma de pepino y poder identificar o estimar los genes implicados con precisión. El número de asociaciones (carácter-marcador) detectables y significativas fue limitado a causa del tamaño poblacional y probablemente puede generar una subestimación del número total de *loci* implicados en la expresión de los caracteres de estudio, tal y como se ha demostrado previamente (Staub et al., 1996). No obstante, algunas de las líneas pueden ser consideradas como genotipos potenciales para futuros programas de mejora genética profundizando en la obtención y fijación de los caracteres buscados.

5.4) Análisis de correlación

La naturaleza codominante de los marcadores identificados en este estudio pone de manifiesto su potencial como herramienta molecular para los futuros programas de mejora en los que se usen los caracteres fenotípicos asociados a los fondos genéticos analizados. La selección en base a un solo QTL constituye un riesgo a la hora de producir un efecto “arrastre” de los genes próximos. Esto provocaría homocigosis también para esos *loci*, los cuales pueden ser desfavorables o incluso deletéreos. Por este motivo se hace necesario combinar la selección utilizando información molecular y también fenotípica para otros caracteres de interés.

Si observamos los resultados obtenidos en el análisis de regresión y correlación, podemos observar interacciones muy interesantes y con gran valor para su consideración en los programas de mejora. El 72,97% de estas interacciones son de tipo moderado o superior, lo que da una idea de su peso como estimadores fenotípicos. Para lograr la mejora de algún carácter debe estudiarse sus posibles relaciones con otros, en términos de fuerza y sentido, ya que puede conducir tanto a fijar varios caracteres de interés al mismo tiempo como a la fijación de uno en detrimento de otros. Esto deberá de sopesarse en función de los intereses particulares del programa de mejora.

El hecho de disponer de 2 poblaciones de estudio ha facilitado la identificación de correlaciones comunes en ambas poblaciones y que por consiguiente tienen potencial para su estudio de extrapolación a otras líneas poblacionales de pepino Dutch. Estas interacciones son las siguientes y su análisis nos lleva a distintas conjeturas:

- Precocidad – Diámetro (-)

- Número entrenudos – Longitud entrenudos (-)
- Número entrenudos – Vigor (-)
- Longitud entrenudos – Vigor (+)
- Longitud entrenudos – color (-)
- Vigor – Longitud pedicelo (+)
- Peso – Longitud fruto (+)
- Peso – Diámetro (+)
- Diámetro – Cavidad seminal (+)

Para ir en orden, la precocidad se ha visto relacionada de forma negativa (inversamente proporcional) al diámetro para ambas poblaciones y con otros caracteres de fruto (cavidad seminal, peso y color) pero en una sola de las poblaciones analizadas. De forma general podemos estimar que las plantas que presentan una mayor velocidad para generar frutos en etapas tempranas, presentan carencias en el desarrollo de los mismos. Esto puede deberse a la distribución de recursos nutricionales que tiene lugar en la planta, teniendo que abastecer tanto al desarrollo primario de la planta como a los frutos (sumidero de fotosintatos, vitaminas y minerales esenciales) al mismo tiempo. (Tatlioglu, 1993; Dijkhuizen & Staub, 2002)

Con respecto a la interacción entre el número de entrenudos y la longitud de los mismos podemos ver como se verifica la teoría comentada anteriormente. Se trata de una interacción fuerte de sentido negativo. Podemos ver que las distintas interacciones con los otros caracteres tienen su réplica pero de sentido opuesto entre ambas. El vigor es una medida de la velocidad de desarrollo vegetativo de la planta y se aprecia que está muy relacionada con ambos caracteres, de forma que las plantas más vigorosas tendrán menor número de entrenudos y de mayor tamaño. Esto puede ser de gran utilidad debido a que el tamaño de los primeros entrenudos permitiría una selección rápida en primera instancia del resto de caracteres correlacionados. También se ha visto que estos caracteres están relacionados con el color del fruto para ambas poblaciones y, adicionalmente, con la cavidad seminal y la longitud del pedicelo en la población A. Las características del fruto parecen incrementarse junto con el número de los entrenudos, mientras que la longitud del pedicelo lo hace con la longitud de los entrenudos. Las plantas más compactas crecen de forma más pausada (menos vigorosas) y en consecuencia tienen una mayor capacidad para la gestión de los recursos disponibles. Por su parte, las plantas más vigorosas tienen a favorecer el crecimiento vegetativo, estando la estructura del pedicelo incluida. Esto último se refuerza por la correlación moderada y positiva existente entre vigor y longitud de pedicelo en ambas poblaciones. (Ullah et al., 2012)

Continuado con la longitud del pedicelo, podemos observar distintas correlaciones en las poblaciones: con el color en A (moderada y negativa) y con la forma del cuello en B (leve y negativa). Una mayor longitud del pedicelo parece conllevar aspectos morfológicos no deseables en los frutos de pepino.

Por último, si prestamos atención a los caracteres del fruto podemos percatarnos de correlaciones fuertes y positivas entre el peso de los frutos con su longitud y diámetro. Estos resultados son totalmente esperados debido a que se tratan de caracteres involucrados en el proceso de desarrollo y maduración de los frutos, de forma que su evolución es proporcional con el tiempo hasta su estado óptimo de recolección (fruto inmaduro) o para la obtención de semillas (fruto maduro). Además, hemos observado correlaciones de este carácter con la cavidad seminal (población B) o de forma indirecta a través de la correlación fuerte y positiva entre el diámetro y la cavidad seminal, en ambas poblaciones (Dijkhuizen & Staub, 2002; Ullah et al., 2012).

Debido al carácter multipistilar segregante en la población A, se encontraron correlaciones únicas con respecto al número de frutos. Este carácter se correlaciona de forma moderada y positiva con la precocidad y la longitud de los entrenudos, mientras que negativamente con el número de entrenudos y el cuello. La dinámica de comportamiento de las plantas precoces parece estar implicada en la capacidad para desarrollar un mayor número de frutos. Sin embargo, este carácter sería más conveniente que lo presentasen plantas con menor número de entrenudos, ya que en caso contrario sería difícil para la planta la producción de múltiples frutos en un gran número de coyunturas nodales. Finalmente, se observa que el morfotipo no deseable de forma del cuello del fruto está ligado a plantas multipistilares, lo que puede encontrar explicación en el desarrollo de distintos frutos al mismo tiempo para la misma coyuntura, como han descrito previamente otros autores (Ullah et al., 2012).

6) Conclusiones

En este Trabajo de Fin de Master se ha llevado a cabo la caracterización fenotípica y molecular de dos poblaciones segregantes de pepino para caracteres de interés agronómico. A partir del conjunto de resultados descritos se ha llegado a las siguientes conclusiones.

La caracterización fenotípica y molecular realizada ha hecho posible la identificación de 17 QTLs para caracteres de interés agronómico. Estos resultados han de ser considerados como preliminares hasta ser corroborados en ensayos posteriores con un mayor número de individuos por población, así como con mapas genéticos con mayor densidad de marcadores. Sin embargo, los QTLs identificados pueden ser ya usados en los futuros programas de mejora genética de pepino para la selección de plantas con los caracteres buscados en genotipos de referencia.

Adicionalmente, los datos obtenidos han de ser considerados como la base de proyectos de localización e identificación de los genes implicados en los QTL de interés agronómico en pepino y el polimorfismo (alelos) responsables de los fenotipos diferenciales entre las líneas parentales. Estos proyectos se están llevando a cabo por el grupo de investigación donde se ha realizado este TFM.

7) Bibliografía

Ahmad, S., Fiaz, S., Riaz, A., Bashir, I., & Zeb, A. (2016). Correlation analysis of morphological and fiber quality traits in upland Cotton (*Gossypium hirsutum* L.). *International Journal of Biosciences*, 9(4), 200-208.

Augustina, U. A., Iwunor, O. P., & Ijeoma, O. R. (2013). Heritability and character correlation among some rice genotypes for yield and yield components. *Journal of plant breeding and Genetics*, 1(2), 73-84.

Behera, T. K., Staub, J. E., Behera, S., Delannay, I. Y., & Chen, J. F. (2011). Marker-assisted backcross selection in an interspecific *Cucumis* population broadens the genetic base of cucumber (*Cucumis sativus* L.). *Euphytica*, 178(2), 261-272.

Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing* (pp. 1-4). Springer, Berlin, Heidelberg.

Cavagnaro, P. F., Senalik, D. A., Yang, L., Simon, P. W., Harkins, T. T., Kodira, C. D., ... & Weng, Y. (2010). Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC genomics*, 11(1), 1-18.

Cubero Salmerón, J. I. (2013). *Introducción a la mejora genética vegetal*. Mundi-Prensa Libros.

De Giovanni, C., Dell'Orco, P., Bruno, A., Ciccarese, F., Lotti, C., & Ricciardi, L. (2004). Identification of PCR-based markers (RAPD, AFLP) linked to a novel powdery mildew resistance gene (ol-2) in tomato. *Plant Science*, 166(1), 41-48.

Dijkhuizen, A., & Staub, J. E. (2002). QTL conditioning yield and fruit quality traits in cucumber (*Cucumis sativus* L.) effects of environment and genetic background. *Journal of New Seeds*, 4(4), 1-30.

Doebley, J. F., Gaut, B. S., & Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell*, 127(7), 1309-1321.

Espejo, J.R., y Lozano, R.R. Cartografía de los genes implicados en el determinismo del sexo en pepino (*Cucumis sativus* L.). [Departamento de Biología y Geología]: Universidad de Almería; 2006.

Fan, Z., Robbins, M. D., & Staub, J. E. (2006). Population development by phenotypic selection with subsequent marker-assisted selection for line extraction in cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics*, 112(5), 843-855.

FAOSTAT, Fao.org.(2021). Food and Agriculture Organization of the United Nations, FAOSTAT.[online] Available at: <http://www.fao.org/faostat/es/#data/QC> [Accessed 29 Jul. 2021].

Fazio, G., Staub, J. E., & Stevens, M. R. (2003). Genetic mapping and QTL analysis of horticultural traits in cucumber (*Cucumis sativus* L.) using recombinant inbred lines. *Theoretical and Applied Genetics*, 107(5), 864-874.

Feng, S., Zhang, J., Mu, Z., Wang, Y., Wen, C., Wu, T., ... & Wang, H. (2020). Recent progress on the molecular breeding of *Cucumis sativus* L. in China. *Theoretical and Applied Genetics*, 133(5), 1777-1790.

González, J. M. (2000). Marcadores moleculares para la selección de caracteres cuantitativos. In *Los marcadores genéticos en la mejora vegetal* (pp. 383-408). Servicio de Publicaciones.

Imam, M. K., Abobakr, M. A., & Yacoub, H. M. (1977). Inheritance of some characters in cucumber, *Cucumis sativus* L. II. Some quantitative characters. *Libyan journal of agriculture*.

Jeffrey, C. (1980). A review of the *Cucurbitaceae*. *Botanical Journal of the Linnean society*, 81(3), 233-247.

Jones, C. J., Edwards, K. J., Castaglione, S., Winfield, M. O., Sala, F., Van de Wiel, C.,... & Brettschneider, R. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular breeding*, 3(5), 381-390.

Kearsey, M. J., & Farquhar, A. G. L. (1998). QTL analysis in plants; where are we now?. *Heredity*, 80(2), 137.

Kirkbride, J. H. (1993). Biosystematic monograph of the genus *Cucumis* (*Cucurbitaceae*): botanical identification of cucumbers and melons. Parkway Publishers, Inc..

Lee, M. (1995). DNA markers and plant breeding programs. In *Advances in agronomy* (Vol. 55, pp. 265-344). Academic Press.

Lefebvre, V., & Chèvre, A. M. (1995). Tools for marking plant disease and pest resistance genes: a review

Liu, J., Liu, D., Tao, W., Li, W., Wang, S., Chen, P., ... & Gao, D. (2000). Molecular marker-facilitated pyramiding of different genes for powdery mildew resistance in wheat. *Plant Breeding*, 119(1), 21-24.

- Llácer, G. (2007). Fruit breeding in Spain. In XII EUCARPIA Symposium on Fruit Breeding and Genetics 814 (pp. 43-56).
- López-Elías, J., Rodríguez, J. C., Huez, M. A., Garza, S., Jiménez, J., & Leyva, E. I. (2011). Producción y calidad de pepino (*Cucumis sativus* L.) bajo condiciones de invernadero usando dos sistemas de poda. *Idesia (Arica)*, 29(2), 21-27.
- Mansur, L. M., Orf, J. H., Chase, K., Jarvik, T., Cregan, P. B., & Lark, K. G. (1996). Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Science*, 36(5), 1327-1336.
- Mibus, H., & Tatlioglu, T. (2004). Molecular characterization and isolation of the F/f gene for femaleness in cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics*, 109(8), 1669-1676.
- Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R., & Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular breeding*, 3(2), 87-103.
- O'Brien, S. J. (1990). Genetic maps: *locus* maps of complex genomes (pp. 4-92). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press
- Pike, L. M., & Peterson, C. E. (1969). Inheritance of parthenocarpy in the cucumber (*Cucumis sativus* L.). *Euphytica*, 18(1), 101-105.
- Qi, J., Liu, X., Shen, D. I., Miao, H., Xie, B., Li, X., ... & Huang, S. (2013). A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nature genetics*, 45(12), 1510-1515.
- Sakata, Y., Sugiyama, M., Yoshioka, Y., & Ohara, T. (2010). Morphological characteristic and yields of five major cucumber types under cultivation in Japan. *Bulletin of the National Institute of Vegetable and Tea Science*, 9, 113-123.
- Sebastian, P., Schaefer, H., Telford, I. R., & Renner, S. S. (2010). Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. *Proceedings of the National Academy of Sciences*, 107(32), 14269-14273.

- Serquen, F. C., Bacher, J., & Staub, J. E. (1997). Mapping and QTL analysis of horticultural traits in a narrow cross in cucumber (*Cucumis sativus* L.) using random-amplified polymorphic DNA markers. *Molecular Breeding*, 3(4), 257-268.
- Shetty, N. V., & Wehner, T. C. (1998). Evaluation of oriental trellis cucumbers for production in North Carolina. *HortScience*, 33(5), 891-896.
- Smith, O. S., OS, S., RL, L., & RH, M. (1978). Estimates of Heritabilities and variance components pickling cucumber. *Journal American Society for Horticultural Science*, 103, 222-225.
- Staub, J. E., Serquen, F. C., & Gupta, M. (1996). Genetic markers, map construction, and their application in plant breeding. *HortScience*, 31(5), 729-741.
- Staub, J. E., Sun, Z., Chung, S. M., & Lower, R. L. (2007). Evidence for colinearity among genetic linkage maps in cucumber. *HortScience*, 42(1), 20-27.
- Sun, Z. H. A. A. N. G. Y. O. U. N. G., Lower, R. L., & Staub, J. E. (2006). Analysis of generation means and components of variance for parthenocarpy in cucumber (*Cucumis sativus* L.). *Plant breeding*, 125(3), 277-280.
- Tanksley, S. D. (1993). Mapping polygenes. *Annual review of genetics*, 27(1), 205-233.
- Tatlioglu, T. (1993). Cucumber: *Cucumis sativus* L. In *Genetic improvement of vegetable crops* (pp. 197-234). Pergamon.
- Ullah, M. Z., Hasan, M. J., Chowdhury, A. Z. M. K. A., Saki, A. I., & Rahman, A. H. M. A. (2012). Genetic variability and correlation in exotic cucumber (*Cucumis sativus* L.) varieties. *Bangladesh Journal of Plant Breeding and Genetics*, 25(1), 17-23.
- Valcárcel Germes, J. V. (2018). Racionalización de la colección de pepino (*Cucumis sativus* L.) del banco de germoplasma del COMAV (Doctoral dissertation, Universitat Politècnica de València).
- Van Ooijen, J.W, 2018. JoinMap ® 5, Software for the calculation of genetic linkage maps in experimental populations of diploid species. Kyazma B.V., Wageningen, Netherlands.
- Van Ooijen, J.W., 2009. MmapQTL ® 6, software for the mapping of quantitative trait *loci* in experimental populations of diploid species. Kyazma B.V., Wageningen, Netherlands.

Yuan, X. J., Li, X. Z., Pan, J. S., Wang, G., Jiang, S., Li, X. H., ... & Cai, R. (2008a). Genetic linkage map construction and location of QTLs for fruit-related traits in cucumber. *Plant Breeding*, 127(2), 180-188.

Yuan, X. J., Pan, J. S., Cai, R., Guan, Y., Liu, L. Z., Zhang, W. W., ... & Zhu, L. H. (2008b). Genetic mapping and QTL analysis of fruit and flower related traits in cucumber (*Cucumis sativus L.*) using recombinant inbred lines. *Euphytica*, 164(2), 473-491.

