



Universitat de Lleida



Estación Experimental de Aula Dei

La colección nuclear española de cebada: diversidad genética y potencial agronómico

Memoria que para optar al grado de Doctor Ingeniero Agrónomo

Presenta:

Samia Yahiaoui

Tutor: Dr. Ignacio Romagosa Clariana

**Directores: Dra. Ana Maria Casas Cendoya
Dr. Ernesto Igartua Arregui**

Zaragoza, Abril 2006

D. ERNESTO IGARTUA ARREGUI, Científico Titular del Consejo Superior de Investigaciones Científicas y Dña ANA MARIA CASAS CENDOYA, Científico Titular del mismo organismo

CERTIFICAN

Que la Tesis Doctoral titulada “**La colección nuclear española de cebada: diversidad genética y potencial agronómico**” ha sido realizada por la Ingeniero Agrónomo Dña SAMIA YAHIAOUI, en el Departamento de Genética y Producción Vegetal, de la Estación Experimental de Aula Dei del Consejo Superior de Investigaciones Científicas bajo su dirección y reúne, a su juicio, las condiciones requeridas para optar al Grado de Doctor Ingeniero Agrónomo

Zaragoza, Abril de 2006

Fdo: Ernesto Igartua Arregui

Fdo: Ana María Casas Cendoya

D. IGNACIO ROMAGOSA CLARIANA, Profesor de la Universidad de Lleida

CERTIFICA

Que la Tesis Doctoral titulada “**La colección nuclear española de cebada: diversidad genética y potencial agronómico**” ha sido realizada por la Ingeniero Agrónomo Dña SAMIA YAHIAOUI, en el Departamento de Genética y Producción Vegetal de la Estación Experimental de Aula Dei del Consejo Superior de Investigaciones Científicas bajo su tutela y reúne, a su juicio, las condiciones requeridas para optar al Grado de Doctor Ingeniero Agrónomo

Lleida, Abril de 2006

Fdo: Ignacio Romagosa Clariana

Agradecimientos

Este trabajo ha sido el fruto de la colaboración de todo un equipo. Por eso quisiera agradecer a todas aquellas personas que lo han hecho posible, aunque será difícil nombrarlas a todas.

Mis agradecimientos van en primer lugar a mis directores de tesis

Dra. Ana Maria Casas Cendoya, por la dedicación que ha tenido a este trabajo, sus orientaciones y ayuda han sido inestimables.

Dr. Ernesto Igartua Arregui, a pesar de sus responsabilidades como Director de la Estación Experimental de Aula Dei, ha estado siempre disponible cuando fue necesario. El reconocimiento que le tengo por su ayuda, y sus consejos nunca será suficiente.

Al Dr. Ignacio Romagosa Clariana, le agradezco de haber aceptado la tutoría de esta Tesis.

A la Estación Experimental de Aula Dei del Consejo Superior de Investigación Científicas por haberme dado la oportunidad de realizar esta Tesis.

Agradezco también la colaboración de todo el equipo de “mejora de cebada” del Departamento de Genética y Producción Vegetal. Les agradezco su acogida y simpatía que hicieron mi estancia en Aula Dei más agradable.

A Pili Asín, le agradezco su amabilidad y su ayuda siempre que se lo he pedido.

Al Dr. J.L Molina-Cano del IRTA-UdL (Lleida), Dra. M. Moralejo UdL (Lleida) y Dr. F. Ciudad y P. Codesal del ITA-JCyL (Valladolid) por suministrar parte de los datos utilizados en este trabajo.

Al Dr. Antonio Martinez-Cob, le agradezco su ayuda con los datos climáticos.

A todos mis amigos y amigas que me apoyaron en los momentos difíciles. Les recordaré siempre con mucho cariño.

A mi familia, que me apoyaron siempre para seguir adelante.

A la administración de l'Institut National de la Recherche Agronomique d'Algérie, por darme la oportunidad de terminar mis estudios de post grado.

Esta tesis se realizó gracias a una beca de la Agencia Española de Cooperación Internacional, y a la financiación de los proyectos del Ministerio de Educación y Ciencia:- Programa Nacional de Recursos y Tecnología Agroalimentaria (INIA) RTA01-088-C3-3 y RF02-016-C2-1;- Plan Nacional, AGL2004-05311.

A mes Parents

Resum

La col·lecció nuclear d'ordis espanyols és un recurs fitogenètic posat a punt gràcies a l'esforç de la col·laboració entre l'IRTA de Catalunya, l'ITA de Castella-Lleó i la EEAD-CSIC de Saragossa. És una representació sistematizada de la diversitat genètica existent en la col·lecció de referència guardada en el Centre de Recursos Fitogenéticos de l'INIA, que conserva més de 2.000 entrades de *Hordeum*, la majoria d'ells ordis autòctons. Aquest treball presenta l'avaluació de la diversitat genètica i agronòmica d'aquesta col·lecció.

La diversitat genètica, d'aquesta col·lecció va ser evaluada mitjançant marcadors moleculars, concretament seixanta-quatre microsatèlites, distribuïts en el set cromosomes. En total, van ser examinats 225 genotips d'ordi. D'ells, 159 línies pures procedents de varietats locals i 16 varietats antigues de llarga tradició a Espanya, conformen la col·lecció nuclear. A més a més, van ser avaluades altres 50 entrades testimonis d'altres orígens, majoritàriament europees. El conjunt de microsatèlites empleat va detectar un alt grau de polimorfisme general. Així mateix, es van identificar nombrosos alels específics de grups de germoplasma, especialment en les entrades espanyoles de 6 carreres.

Els anàlisis multivariants (cluster i coordenades principals) duts a terme sobre una matriu de distàncies genètiques, així com l'anàlisi de l'estructura de població, realitzat mitjançant el programa STRUCTURE, van separar clarament els genotips de 6 carreres espanyols dels europeus. Les entrades de 2 carreres espanyoles es van separar menys de les varietats de 2 carreres de primavera del conjunt de referència. Tots aquestes anàlisis van apuntar a l'existència de dos grups principals d'entrades espanyoles. Aquests dos grups van mostrar diferències en distribució geogràfica i semblen tenir orígens diferents.

L'associació entre els alels de 73 marcadors analitzats i les paràmetres eco-geogràfics del lloc de recol·lecció de les entrades de la col·lecció nuclear va mostrar que la distribució geogràfica d'alguns alels està influenciada per aquests paràmetres, el que suggereix que els patrons de diversitat observats poden haver estat causats en part per forces adaptatives.

Les entrades de la col·lecció nuclear, juntament amb 26 varietats comercials, van ser agronòmicament avaluades en deu assajos al llarg de tres anys i en un total de cinc localitats. Aquesta avaluació va posar de manifest la major variabilitat existent en les entrades autòctones per a molts caràcters, d'acord amb la notable diversitat detectada a nivell molecular. A més a més, va permetre ressaltar, en general, la superioritat de les varietats comercials enfront de les entrades de la col·lecció nuclear, però va mostrar també l'estabilitat d'aquestes últimes en condicions desfavorables.

Un anàlisi de “mapeo” per associació a escala del genoma va ser dut a terme en la totalitat de les entrades de la col·lecció nuclear i també en un dels grups genètics. Aquest anàlisi va detectar un gran nombre d'associacions, que va disminuir dràsticament al tenir en compte l'estructura de poblacions en l'anàlisi. La majoria de les associacions més probables revelades van coincidir amb els QTLs detectats en poblacions biparentales, o amb posicions de gens coneguts.

Resumen

La colección nuclear de cebadas españolas es un recurso fitogenético puesto a punto gracias a un esfuerzo de colaboración entre el IRTA de Cataluña, el ITA de Castilla-León y la EEAD-CSIC de Zaragoza. Es una representación sistematizada de la diversidad genética existente en la colección de referencia guardada en el Centro de Recursos Fitogenéticos del INIA, que conserva más de 2.000 entradas de *Hordeum*, la mayoría de ellas cebadas autóctonas. Este trabajo presenta la evaluación de la diversidad genética y agronómica de esta colección.

La diversidad genética de esta colección fue evaluada mediante marcadores moleculares, concretamente sesenta y cuatro microsatélites, distribuidos por los siete cromosomas. En total, fueron examinados 225 genotipos de cebada. De ellos, 159 líneas puras (148 de seis carreras y 11 de dos carreras) procedentes de variedades locales y 16 variedades antiguas de larga tradición en España, conforman la colección nuclear. Además fueron evaluadas otras 50 entradas testigos de otros orígenes, mayoritariamente europeos. El conjunto de microsatélites empleado detectó un alto grado de polimorfismo general. Asimismo, se identificaron numerosos alelos específicos de grupos de germoplasma, especialmente en las entradas españolas de 6 carreras.

Los análisis multivariantes (cluster y coordenadas principales) llevados a cabo sobre una matriz de distancias genéticas, así como el análisis de la estructura de población, realizado mediante el programa STRUCTURE, separaron claramente los genotipos de 6 carreras españoles de los europeos. Las entradas de 2 carreras españolas se separaron menos de las variedades de 2 carreras de primavera del conjunto de referencia. Todos estos análisis apuntaron la existencia de dos grupos principales dentro de las entradas españolas de seis carreras. Estos dos grupos mostraron diferencias en distribución geográfica y parecen tener orígenes distintos.

La asociación entre los alelos de 73 marcadores analizados y los parámetros eco-geográficos del lugar de recolección de las entradas de la colección nuclear, mostró que la distribución geográfica de algunos alelos está influenciada por dichos parámetros, lo que sugiere que los patrones de diversidad observados pueden haber sido causados en parte por fuerzas adaptativas.

La evaluación agronómica de las entradas autóctonas de la colección nuclear, junto con 26 variedades comerciales fue llevada a cabo en diez ensayos, a lo largo de tres años, y en un total de cinco localidades. Esta evaluación puso de manifiesto la mayor variabilidad existente en las entradas autóctonas para muchos caracteres, en consonancia con la notable diversidad detectada a nivel molecular. Esta evaluación permitió también resaltar en general la superioridad de las variedades comerciales frente a las entradas de la colección nuclear, pero mostró también la estabilidad de estas últimas en condiciones desfavorables.

Un análisis de mapeo por asociación a escala del genoma fue llevado a cabo en la totalidad de las entradas de la colección nuclear y también en uno de los grupos genéticos. Este análisis detectó un gran número de asociaciones, que disminuyó drásticamente al tener en cuenta la estructura de poblaciones en el análisis. La mayoría de las asociaciones más probables reveladas coincidieron con QTLs detectados en poblaciones biparentales, o con posiciones de genes conocidos.

Summary

The Spanish barley core collection is a phylogenetic resource developed in collaboration by the IRTA of Cataluña, the ITA of Castilla-León, and the EEAD-CSIC of Zaragoza. It represents, in a systematic manner, the genetic diversity of the collection of over 2,000 accessions of *Hordeum* kept at the Center of Phylogenetic Resources of the INIA, most of them native barleys. This work presents the evaluation of the genetic and agronomical diversity of this collection.

The genetic diversity of this collection was evaluated by means of 64 microsatellites, distributed along the seven chromosomes. Altogether, 225 genotypes of barley were examined. Out of them, 159 inbred lines coming from local landraces (148 6-row and 11 2-row), and 16 old varieties of long tradition in Spain, constitute the nuclear collection. In addition, another 50 accessions of other origins, mainly European cultivars, were evaluated. The set of microsatellites detected a high degree of overall polymorphism. Also, numerous private alleles of each germplasm group were identified, specially in the group of 6 row Spanish accessions.

Multivariate analyses (cluster and principal coordinate analysis) carried out on a matrix of genetic distances, as well as an analysis of population structure (by means of the software package STRUCTURE), clearly separated the Spanish 6-row genotypes from the European ones. On the other hand, the Spanish set of 2-row accessions was classified relatively closer to the 2-row spring European cultivars. These analyses indicate the existence of two main groups among the Spanish 6-row accessions. These two groups showed differences in geographic distribution and seem to have different origins.

The association between the alleles of 73 markers, and the eco-geographic data of the collection sites of the Spanish accessions, showed an influence of these parameters on the geographic distribution of some alleles, suggesting that the patterns of distribution of diversity observed could be partly caused by adaptive forces.

The landrace accessions of the core collection, along with 26 commercial varieties, were evaluated for agronomic traits in ten field trials, carried out over three years and five locations. This evaluation showed a great variability for most traits among the Spanish landraces, in keeping with the diversity detected at the molecular level. This evaluation revealed an overall superiority of the commercial varieties over the landraces of the core collection, but the latter showed greater stability under unfavourable conditions.

A genome-wide association mapping was carried out at two different levels: one for the entire set of landraces of the core collection, and another one for the landraces of the genetic group which showed a lesser amount of linkage disequilibrium. These analyses detected many apparently significant associations, that were drastically diminished when the structure of populations was considered in the analysis. Many of the most likely associations detected agreed with locations of QTLs found in biparental populations studies, or with positions of well-known genes.

Résumé

La collection nucléaire espagnole est une ressource phytogénétique mise au point grâce à l'effort de collaboration entre l'IRTA de Catalogne, l'ITA de Castille-Leon, et la EEAD-CSIC de Saragosse. C'est une représentation systématisée de la diversité génétique existante au niveau de la collection de référence maintenu dans le Centre de Ressources Phytogénétiques de L'INIA, que conserve plus de 2000 accessions de *Hordeum*, la majorité d'entre-elles autochtones. Cette étude se base sur l'évaluation de la diversité génétique y agronomique de cette collection.

La diversité génétique de cette collection a été évaluée par marqueurs moléculaires, concrètement soixante quatre microsattellites, distribué sur les sept chromosomes. Au total, 225 génotypes d'orge ont été examinés. Entre eux, 159 lignées pures provenant de variétés locales y 16 variétés anciennes de longue tradition en Espagne. En outre, 50 autres accessions témoins de différente origine, en majorité européenne ont été évaluées. L'ensemble des microsattellites utilisés ont détecté un degré de polymorphisme général élevé. Également, beaucoup d'allèles spécifiques des groupes de germplasm ont été identifiés, spécialement au niveau des accessions espagnoles à six rangs.

Les analyses multivariantes (cluster y coordonnées principales) réalisée a partir d'une matrice de distances génétiques, ainsi que l'analyse de la structure de population, effectuée par le programme STRUCTURE, ont clairement séparé les génotypes a 6 rangs espagnols des européens. Les accessions de 2 rangs espagnoles ont été moins séparées des variétés printanières de 2 rangs du lot de référence. Toutes ces analyses ont signalé l'existence de deux groupes principaux dans les génotypes à six rangs espagnoles. Ces deux groupes ont montré des différences en distribution géographique et paraissent avoir des origines différentes.

L'association entre les allèles de 73 marqueurs analysés et les paramètres éco-géographiques du lieu de collecte des accessions de la collection nucléaire, a montré que que la distribution géographique de quelques allèles est influencée par ces paramètres, ce qui suggère que les patrons de diversité observés puissent avoir été causés en partie par des forces adaptatives.

L'évaluation agronomique des accessions autochtones la collection nucléaire, et 26 variétés commerciales a été menée dans dix essais, au long de trois années, dans un total de cinq localités. Cette évaluation a mis en évidence la grande variabilité existante dans les génotypes autochtones pour beaucoup de caractères, en accord avec la diversité remarquable détectée au niveau moléculaire. Cette évaluation a également permis de souligner en général la supériorité des variétés commerciales face aux accessions de la collection nucléaire, mais a montré aussi la stabilité de ces dernières dans des conditions défavorables.

Une analyse de cartographie génétique par association, à l'échelle du génome a été effectuée en considérant la totalité des accessions de la collection nucléaire, y aussi un des groupes génétiques. Cette analyse a détecté un grand nombre d'associations, qui a diminué fortement en tenant compte la structure de populations dans l'analyse. La majorité des associations probables révélées ont coïncidé avec des QTLs détectés dans des populations biparentales, ou avec les positions de gènes connus.

Índice de Figuras

Figura 1.1	Distribución por Comunidades Autónomas de la superficie dedicada al cultivo de la cebada en España (Fuente: Anuario de Estadística Agroalimentaria M.A.P.A., 2004.)	6
Figura 3.1	Dendrograma elaborado según el procedimiento UPGMA de los 225 genotipos analizados (españoles y referencia), empleando una matriz de distancias calculada según el método de Manhattan (Cain y Harrison, 1958), con 64 microsatélites. Los grupos (ver texto) se diferencian por colores distintos.	49
Figura 3.2	Análisis de coordenadas principales realizado a partir de la matriz de distancias de Manhattan (Cain y Harrison, 1958), de los 225 genotipos analizados con 64 marcadores. Los genotipos están representados según los grupos establecidos a priori.	51
Figura 3.3	Evolución del valor del ' <i>log-likelihood</i> ' (LnP(D)) en función del número K de poblaciones (promedio de 25 repeticiones del análisis para cada valor de K).	52
Figura 3.4	Clasificación de los 225 genotipos en diferentes poblaciones según su estructura genética a partir de la información procurada por 64 microsatélites, analizados con el programa STRUCTURE.	53
Figura 3.5A	Evaluación del desequilibrio de ligamiento entre pares de loci, en los 225 genotipos analizados: 1-Representación de la variación de r^2 (cuadrado del coeficiente de correlación entre alelos) en función de las distancias genéticas entre pares de loci; 2- Evolución de la probabilidad de significación (-log p) de los valores de r^2 con respecto a las distancias genéticas entre loci.	58
Figura 3.5B	Evaluación del desequilibrio de ligamiento entre pares de loci, en las 4 poblaciones obtenidas por STRUCTURE: 1-Representación de la variación de r^2 (cuadrado del coeficiente de correlación entre alelos) en función de las distancias genéticas entre pares de loci; 2- Evolución de la probabilidad de significación (-log p) de los valores de r^2 con respecto a las distancias genéticas entre loci.	59
Figura 3.6	Promedios de los valores de r^2 , según las 3 clases de distancias definidas (0-10 cM, > 10 cM, no ligados), en los 225 genotipos analizados y las diferentes poblaciones.	60
Figura 4.1	Análisis en componentes principales de las variables climáticas y geográficas descritas en el texto, correspondientes a las zonas de recolección de las 159 <i>landraces</i> de la colección nuclear.	83

Figura 4.2	Representación de la distribución de los alelos: A) MWG699-A (negro), MWG699-D (azul); B) Bmac316-6 (azul), Bmac316-9 (negro) según las variaciones de altitud y la representación de la distribución de los alelos: C) Bmac399-16 (negro), D) Bmac134-5 (negro) según las variaciones respectivamente de la pluviometría de invierno (enero) y de primavera (abril). Los alelos de color gris en las Figuras A, B, C y D, son los que no presentaron asociación con la variable considerada o son alelos raros que no entraron en el análisis	93
Figura 4.3	Representación de la distribución de los alelos: A) HvM062-3 (negro), HvM062-4 (azul), B) HvBM5A-2 (azul) HvBM5A-4 (negro) según las variaciones de temperaturas de invierno (enero) y la representación de la distribución de los alelos: C) scssr07759-4 (negro), scssr07759-5 (azul), D) HvLTPPB-3 (negro) y HvLTPPB-2 (azul) según las variaciones de temperaturas de primavera (abril). Los alelos de color gris en las figuras A, C y D, son los que no presentaron asociación con la variable considerada o son alelos raros que no entraron en el análisis	95
Figura 4.4	A) Representación de la distribución de los alelos 2 (color azul) y 4 (color negro) del locus HvBM5A en la Península Ibérica, que presentaron asociación con la clasificación climática de Papadakis. Los puntos de color gris representan los alelos no asociados. B) Diferentes tipos de clima que se encuentran en la Península Ibérica según la clasificación climática de Papadakis.	99
Figura 5.1	Esquema de los ensayos de campo, mostrando la disposición de los sub-ensayos y las repeticiones.	118
Figura 5.2	Distribución de los valores (en porcentajes) de los caracteres: espigado, altura, encamado, rendimiento, rendimiento en ambiente de alta productividad, rendimiento en ambiente de baja productividad y peso hectolítrico, para las variedades comerciales y <i>landraces</i> . Cada clase de los histogramas es de una amplitud de 0,5 desviaciones típicas a partir de la media para las variables cuantitativas y de 1 unidad para la variable encamado	136
Figura 5.3	Distribución de los valores (en porcentajes) de los caracteres: peso de mil granos, espigas m ⁻² , granos m ⁻² , granos por espiga, susceptibilidad a oídio, moteado de las hojas y roya para las variedades comerciales y <i>landraces</i> . Cada clase de los histogramas es de una amplitud de 0,5 desviaciones típicas a partir de la media para las variables cuantitativas y de 1 unidad para las variables reacción a oídio, moteado de las hojas y roya.	139

Figura 5.4	Distribución de los valores (en porcentajes) de los caracteres calibre 2,2 mm, calibre superior a 2,5 mm, materia seca en extracto y proteína total en la harina de malta, para las variedades comerciales y <i>landraces</i> . Cada clase de los histogramas es de una amplitud de 0,5 desviaciones típicas a partir de la media.	143
Figura 5.5	Representación del PCA1 del AMMI versus rendimiento medio ($q\ ha^{-1}$) de 182 genotipos divididos en 7 grupos, y evaluados en 9 ensayos.	157
Figura 5.6	Representación de los dos componentes principales del modelo AMMI2, para 182 genotipos divididos en 7 grupos evaluados en 9 ensayos. Están incluidas las variables genotípicas y climáticas explicativas del GxE. Los vectores indican en que dirección se da el mayor cambio de las variables genotípicas (líneas enteras) y climáticas (líneas discontinuas). El tamaño de los vectores es proporcional al R^2 de cada variable explicado por los dos ejes principales.	159
Figura 6.1	Promedios de los coeficientes de determinación ajustados resultantes de los análisis de regresión entre los 73 marcadores y las variables agronómicas, de tolerancia a enfermedades y calidad maltera según los 3 modelos SG, CG y G4. Solo se representaron los promedios de las variables que mostraron alguna asociación significativa en los modelos CG y G4.	193
Figura 6.2	Promedios de los coeficientes de determinación ajustados resultantes de los análisis de regresión entre los 73 marcadores y las variables morfológicas según los 3 modelos SG, CG y G4. Solo se representaron los promedios de las variables que mostraron alguna asociación significativa en los modelos CG y G4.	194
Figuras 6.3a - 6.9a	Representación de un mapa consenso de <i>bins</i> de los cromosomas 1H-7H. Los marcadores en negrita corresponden a los analizados en este estudio. Los loci subrayados marcan el inicio del <i>bin</i> . Las tablas corresponden a la separación de medias de los alelos de loci que mostraron asociaciones positivas con los caracteres analizados en el modelo multilocus de las 156 <i>landraces</i> con el modelo con grupos (CG).	217
Figuras 6.3b - 6.9b	Representación de un mapa consenso de <i>bins</i> de los cromosomas 1H-7H. Los marcadores en negrita corresponden a los analizados en este estudio. Los loci subrayados marcan el inicio del <i>bin</i> . Las tablas corresponden a la separación de medias de los alelos de loci que mostraron asociaciones positivas con los caracteres analizados en el modelo multilocus de las entradas del grupo 4 (G4).	218

**Figuras
6.3c- 6.9c**

Mapa consenso de *bins* de los cromosomas 1H-7H, Los marcadores en negrita corresponden a los analizados en este estudio. Los loci subrayados marcan el inicio del *bin*. Las tablas representan las asignaciones de posiciones de QTLs encontrados en otras poblaciones, solamente para los caracteres con asociaciones significativas en este estudio. Los números corresponden a las poblaciones donde se localizaron los QTLs, los números marcados en negrita representan el pico del QTL en dicha población, las poblaciones están descritas en la Tabla 6.7. Las letras corresponden a la referencia de descripción de los genes de los caracteres correspondientes, que se presentan en la Tabla 6.8

Índice de Tablas

Tabla 3.1	Distribución de los 225 genotipos en 4 grupos de germoplasma.	24
Tabla 3.2	Distribución geográfica, por región agro-ecológica y por provincia de las entradas de la colección nuclear.	25
Tabla 3.3	Pedigrí y origen de las variedades analizadas.	26
Tabla 3.4	Descripción de los 64 microsatélites analizados, incluyendo localización cromosómica, motivo y tamaño en pb y fuente de referencia.	30
Tabla 3.5	Origen de los microsatélites.	32
Tabla 3.6	Esquema del análisis molecular de varianza.	34
Tabla 3.7	Número total de alelos detectados, rango de tamaño de los alelos de cada locus; número de alelos comunes y alelos específicos de los microsatélites, en cada uno de los 4 grupos de germoplasma establecidos a priori.	39
Tabla 3.8	Índices de diversidad: total (H_T), de los genotipos de 2 carreras (H_{T2C}) y 6 carreras (H_{T6C}) y de los 4 grupos de germoplasma establecidos a priori (2c Españolas, 2c Referencia, 6c Españolas, y 6c Referencia).	41
Tabla 3.9	Resultados del análisis molecular de varianza (AMOVA), comparando los grupos de genotipos establecidos a priori, basado sobre los modelos de evolución de los microsatélites IAM y SMM.	44
Tabla 3.10.	AMOVA (análisis molecular de varianza) locus por locus, en orden cromosómico (1H a 7H). Índices F_{ST} y R_{ST} que comparan la divergencia genética entre los grupos de genotipos españoles y de referencia.	45
Tabla 3.11	Índices de Fijación F_{ST} y R_{ST} de la comparación entre poblaciones y sub-poblaciones derivados del análisis con el programa STRUCTURE.	56
Tabla 3.12	Distribución de los haplotipos del STS MWG699 dentro de las poblaciones y sub-poblaciones	57
Tabla 3.13	Comparación mediante análisis de varianza y contrastes entre poblaciones y clases de distancias y interacción entre las dos variables.	61
Tabla 4.1	Variables geográficas y climáticas analizadas.	81

Tabla 4.2	Número de alelos detectados en las 159 <i>landraces</i> de la colección nuclear, y resultado global de los análisis de regresión de los alelos frecuentes frente a las 13 variables climáticas y geográficas analizadas. El valor del índice F_{ST} entre los dos grupos principales de entradas se ofrece a efectos comparativos.	86
Tabla 4.3	Coefficientes de determinación de los análisis de regresión entre alelos (variables dependientes) y variables geográficas y climáticas (variables independientes).	89
Tabla 5.1	Coordenadas geográficas y características climáticas de las localidades experimentales, y fechas de siembra de los diferentes ensayos.	116
Tabla 5.2	Caracteres agronómicos registrados en cada ensayo.	120
Tabla 5.3	Descripción de los 10 ensayos empleados en este estudio: promedios y desviaciones estándar de las variables agronómicas (rendimiento, espigado, altura, encamado, peso hectolítrico y peso de mil granos) de las variedades testigo por ensayo y sub-ensayo. C: coeficiente de corrección de los sub-ensayos.	128
Tabla 5.4	Promedio y desviación estándar de las variables de componentes de rendimiento (espigas m^{-2} , número de granos m^{-2} , granos por espiga y peso de mil granos), de las variedades testigo, por sub-ensayo; C: coeficiente de corrección de los sub-ensayos.	130
Tabla 5.5	Promedios y desviación estándar de las enfermedades (moteado de las hojas y oídio) de los testigos observados en los sub-ensayos de los ensayos de Artesa2002, Bell-lloc 2002, Sádaba 2002 y Foradada 2003).	130
Tabla 5.6	Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables agronómicas (rendimiento, espigado, altura, encamado, peso hectolítrico y peso de mil granos) observados en los dos grupos estudiados: <i>landraces</i> y variedades comerciales, en los diferentes ensayos.	135
Tabla 5.7	Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables de componentes de rendimiento (espigas m^{-2} , granos m^{-2} , granos por espiga y peso de mil granos) observados en los dos grupos estudiados: <i>landraces</i> y variedades comerciales, en el ensayo VED03.	138
Tabla 5.8	Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables de enfermedades (oídio, moteado de las hojas y roya) observados en los dos grupos estudiados: <i>landraces</i> y variedades comerciales, en los diferentes ensayos.	140

Tabla 5.9	Parámetros estadísticos (Media, Ssd: desviación estándar, Min: mínimo, Max: máximo) de las variables de calidad maltera (calibre <2,2 mm, calibre >2,5mm, materia seca en el extracto de malta y proteína total en la malta) observados en los dos grupos estudiados: <i>landraces</i> y variedades comerciales, en el ensayo de Foradada 2003.	142
Tabla 5.10a	Cuadrados medios de los análisis de varianza para las variables espigado (días desde el 1 de enero), altura de planta (cm) y encamado (valoración visual) evaluados en los diferentes ensayos de campo.	144
Tabla 5.10b	Separación de medias de los grupos analizados, para las variables espigado (días desde 1 de enero), altura de planta (cm) y encamado.	145
Tabla 5.11a	Cuadrados medios de los análisis de varianza para las variables rendimiento ($q\ ha^{-1}$), peso hectolítrico ($kg\ hl^{-1}$) y peso de mil granos (g), evaluados en los diferentes ensayos de campo.	145
Tabla 5.11b.	Separación de medias de los grupos analizados, para las variables rendimiento ($q\ ha^{-1}$), peso hectolítrico ($kg\ hl^{-1}$) y peso de mil granos (g).	146
Tabla 5.12a	Cuadrados medios de los análisis de varianza para la variable rendimiento ($q\ ha^{-1}$), evaluada en los diferentes ensayos de los 2 ambientes de alta y baja productividad.	147
Tabla 5.12b	Separación de medias de los grupos analizados, para la variable rendimiento ($q\ ha^{-1}$), en ambientes de alta y baja productividad.	147
Tabla 5.13a	Cuadrado medios de los análisis de varianza para las variables espigas m^{-2} , granos m^{-2} , y granos por espiga y peso de mil granos evaluados en el ensayo del Vedado 2003.	148
Tabla 5.13b	Separación de medias de los grupos analizados, para las variables espigas m^{-2} , granos m^{-2} , y granos por espiga evaluados en el ensayo del Vedado 2003.	148
Tabla 5.14a	Cuadrado medios de los análisis de varianza para las variables oídio, moteado de las hojas y roya evaluadas en los diferentes ensayos.	149
Tabla 5.14b	Separación de medias de los grupos analizados, para las variables oídio, moteado de las hojas y roya.	149
Tabla 5.15a	Cuadrado medios de los análisis de varianza para las variables calibre <2,2 mm, calibre > 2,5 mm, materia seca en el extracto de malta, proteína total en harina de malta evaluados en el ensayo de FOR03.	150
Tabla 5.15b	Separación de medias de los grupos analizados, para las variables calibre <2,2 mm, calibre >2,5 mm, materia seca en el extracto de malta, proteína total en harina de malta evaluados en el ensayo de Foradada 2003.	150

Tabla 5.16.	Parámetros estadísticos (Media, Sd: desviación estándar) de las variables rendimiento ($q\ ha^{-1}$), peso hectolítrico ($kg\ hl^{-1}$), peso de mil granos (g), espigas m^{-2} , granos m^{-2} , granos por espiga, calibre $<2,2\ mm$ y calibre $>2,5\ mm$), en los grupos <i>landraces</i> y variedades comerciales, según el tipo de espiga 2 ó 6 carreras, considerando el número de ensayos en los que fue tomado cada carácter.	152
Tabla 5.17	Análisis de varianza del AMMI para rendimiento	154
Tabla 5.18	Regresión factorial para el factor principal “Genotipo”, incluyendo covariables genotípicas.	161
Tabla 5.19	Repartición de la interacción GxE, mediante análisis de regresión factorial con las covariables genotípicas.	162
Tabla 5.20	Repartición de la interacción GxE, mediante análisis de regresión factorial con las covariables ambientales.	163
Tabla 5.21	Repartición de la interacción GxE, mediante análisis de regresión factorial con las covariables genotípicas y ambientales.	163
Tabla 5.22	Clasificación de los genotipos según el promedio de rendimiento en los 9 ensayos (sin ART02), promedio de rendimiento en ambiente de alta productividad, promedio de rendimiento en ambiente de baja productividad e índice agronómico. Los números corresponden a las entradas de la colección nuclear, entre paréntesis se indica a que grupo genético pertenecen. Las variedades comerciales aparecen en negro, y los testigos además están escritos en cursiva.	165
Tabla 6.1	Caracteres analizados	187
Tabla 6.2.	Número de loci que presentaron asociaciones positivas con las variables estudiadas, según el modelo de análisis empleado (sin grupos, con grupos o grupo 4); número de loci comunes entre los diferentes modelos.	192
Tabla 6.3	Asociaciones positivas de los loci con las variables estudiadas, según el modelo de análisis empleado (A: sin estructura de población o SG; B: considerando la estructura de población o CG; C: grupo 4 ó G4); * y X corresponden respectivamente a un nivel de significación α de 0,05 ó de 0,1, empleando la corrección de Bonferroni para 73 tests independientes.	199
Tabla 6.4	Distribución fenotípica de las variables morfológicas en las 159 <i>landraces</i> , de acuerdo con la escala aplicada en la descripción de cada carácter.	201

Tabla 6.5	Coeficientes de determinación mostrados por los diferentes loci que revelaron asociaciones positivas con las variables analizadas en el modelo con grupos, en los análisis locus por locus y en el modelo multilocus.	212
Tabla 6.6	Coeficientes de determinación mostrados por los diferentes loci que revelaron asociaciones positivas con las variables analizadas en el modelo G4, en los análisis locus por locus y en el modelo multilocus.	214
Tabla 6.7	Descripción de las poblaciones numeradas en las figuras 6.3c-6.9c y sus correspondientes referencias.	244
Tabla 6.8	Descripción de los genes citados en las figuras 6.3c - 6.9c, con los códigos asignados a las referencias	245

ÍNDICE

Resum.....	vii
Resumen.....	ix
Summary.....	xi
Résumé.....	xiii
Índice Figuras.....	xv
Índice Tablas.....	xix
Capítulo 1. Introducción general.....	1
1. Introducción.....	3
1-1. La especie <i>Hordeum vulgare</i> . L.	3
1-1-1. Origen y domesticación	3
1-1-2. Taxonomía y citología	4
1-1-3. Tipos de cebada:	5
1-2. Importancia económica y distribución geográfica	5
1-3. Mejora de la cebada.....	7
1-4. La importancia de la conservación de los recursos fitogenéticos	8
1-5. Referencias bibliográficas.....	11
Capítulo 2. Objetivos.....	15
2. Objetivos.....	17
Capítulo 3. Diversidad genética	19
3-1. Introducción	21
3-2. Material y Métodos	23
3-2-1. Evaluación de la diversidad por marcadores moleculares	23
3-2-1-1. Material vegetal	23
3-2-1-2. Microsatélites analizados.....	29
3-2-1-3. Condiciones de amplificaciones y electroforesis.....	32
3-2-1-4. Análisis estadísticos.....	33
a. Diversidad genética.....	33
b. Matriz de distancias.....	33
c. Análisis molecular de varianza	34
d. Análisis de la estructura genética	36
e. Evaluación del desequilibrio de ligamiento.....	36
3-3. Resultados	38
3-3-1. Polimorfismo	38
3-3-2. Comparación de los grupos establecidos a priori	38
3-3-2-1. Polimorfismo e índice de diversidad	38
3-3-2-2. Análisis molecular de varianza.....	43
3-3-2-3. Análisis multivariantes	47
a. Análisis de cluster.....	47
b. Análisis de coordenadas principales.....	48
3-3-3. Estructura genética.....	52
3-3-4. Análisis molecular de varianza	56

3-3-5. Distribución del STS MWG699	57
3-3-6. Desequilibrio de ligamiento.....	58
3-4. Discusión.....	61
3-4-1. Polimorfismo e índices de diversidad.....	61
3-4-2. Análisis molecular de varianza	63
3-4-3. Análisis multivariantes	64
3-4-4. Estructura genética.....	65
3-4-5. Desequilibrio de ligamiento.....	67
3-5. Referencias bibliográficas.....	69

Capítulo 4. Distribución de la diversidad en relación con los datos eco-geográficos..... 75

4-1. Introducción	77
4-2. Material y Métodos	78
4-2-1. Material vegetal	78
4-2-2. Marcadores moleculares	78
4-2-3. Variables geográficas y climáticas.....	79
4-2-4. Análisis estadísticos.....	81
4-3. Resultados	82
4-3-1. Análisis en componentes principales de las variables geográficas y climáticas	82
4-3-2. Regresiones entre alelos y variables geográficas y climáticas.....	84
4-3-3. Relación entre número de asociaciones detectadas por loci y los F_{ST} locus por locus de comparación entre los grupos 3 y 4 determinados por STRUCTURE.....	98
4-4. Discusión.....	101
4-5. Referencias bibliográficas.....	105

Capítulo 5. Caracterización agronómica..... 109

5-1. Introducción	111
5-2. Material y Métodos	112
5-2-1. Material vegetal	112
5-2-2. Diseño experimental y localidades	113
5-2-3. Caracteres estudiados.....	118
5-2-3-1. Espigado (ESP).....	118
5-2-3-2. Altura (ALT).....	118
5-2-3-3. Encamado (ENC).....	119
5-2-3-4. Rendimiento (RDT).....	119
5-2-3-5. Peso hectolítrico (PHL)	119
5-2-3-6. Peso de 1000 granos (PMG).....	119
5-2-3-7. Componentes de rendimiento	120
5-2-3-8. Enfermedades	120
5-2-3-9. Calidad maltera.....	121
5-2-4. Análisis estadísticos.....	122
5-2-4-1. Análisis de cada ensayo.....	122
5-2-4-2. Análisis combinado de los ensayos	123
a. Estudio descriptivo de las variables.....	123
b. Análisis de varianza (interacción genotipo x ambiente).....	123
c. AMMI (<i>Additive Main effects and Multiplicative Interaction</i>)	124

d. Regresión factorial.....	125
5-3. Resultados	126
5-3-1. Análisis de cada ensayo	126
5-3-2. Comparación entre las <i>landraces</i> y las variedades comerciales	131
5-3-2-1. Espigado	131
5-3-2-2. Altura de planta	131
5-3-2-3. Encamado	132
5-3-2-4. Rendimiento	132
5-3-2-5. Peso hectolítrico	133
5-3-2-6. Peso de mil granos.....	134
5-3-2-7. Componentes de rendimiento	137
5-3-2-8. Enfermedades	140
5-3-2-9. Calidad maltera.....	141
5-3-3. Análisis combinado de los ensayos	143
5-3-3-1. Caracteres agronómicos.....	143
5-3-3-2. Enfermedades	148
5-3-3-3. Calidad maltera.....	149
5-3-4. Comparación entre genotipos de 2 y 6 carreras.....	150
5-3-5. Análisis de los factores G y GxE del rendimiento.....	153
5-3-5-1. AMMI.....	153
5-3-5-2. Análisis de regresión factorial	161
5-3-6. Clasificación de los genotipos	163
5-4. Discusión.....	166
5.5 Referencias bibliográficas	175

Capítulo 6. Mapeo por asociación.....179

6-1. Introducción	181
6-2. Material y Métodos	183
6-2-1. Material vegetal	183
6-2-2. Marcadores moleculares	184
6-2-3. Caracteres estudiados.....	184
6-2-4. Análisis de asociación.....	188
6-2-4-1. Regresión marcador a marcador	188
6-2-4-2. Regresión multilocus	190
6-3. Resultados	191
6-3-1. Análisis de regresión locus por locus	191
6-3-1-1. Asociaciones con caracteres agronómicos	195
6-3-1-2. Asociaciones con tolerancia a enfermedades	198
6-3-1-3. Asociaciones con calidad maltera.....	200
6-3-1-4. Asociaciones con variables morfológicas.....	201
6-3-2. Análisis de regresión multilocus.....	205
6-3-2-1. Análisis multilocus con caracteres agronómicos.....	205
6-3-2-2. Análisis multilocus de tolerancia a enfermedades.....	207
6-3-2-3. Análisis multilocus con las variables de calidad maltera	208
6-3-2-4. Análisis multilocus con las variables morfológicas	208
6-4. Discusión.....	246
6-5. Referencias bibliográficas	264

Capítulo 7. Discusión General.....275

7-1. Discusión general	277
7-1-1. Diversidad y adaptación.....	277
7-1-2. Relación entre la diversidad genética y comportamiento agronómico	278
7-1-3. Utilidad de los recursos fitogenéticos de la colección nuclear española.	280
7-1-4. Desequilibrio de ligamiento y mapeo por asociación.....	281
7-1-5. Referencias bibliográficas.....	283
Capítulo 8. Conclusiones.....	285
8. Conclusiones.....	287
Anexos.....	289

Capítulo 1

Introducción general

1. Introducción

1-1. La especie *Hordeum vulgare*. L.

1-1-1. Origen y domesticación

La cebada (*H. vulgare*. L.) fue una de las primeras especies domesticadas. Es aparentemente originaria del suroeste de Asia. Restos arqueológicos de granos de cebada encontrados en el Creciente Fértil indican que el cultivo fue domesticado hace por lo menos 10.000 años (Zohary y Hopf, 1993 en Yadun *et al.*, 2000). Los primeros tipos que aparecieron fueron los de 2 carreras. Algo más tarde aparecieron los tipos de 6 carreras, aproximadamente hace 9.500 años (Bothmer *et al.*, 2003).

Existen diferentes teorías sobre la domesticación de la cebada a partir de un pariente silvestre. De Candolle (1959, en Bothmer *et al.*, 2003) consideró que la cebada cultivada de 2 carreras procede directamente de la especie silvestre *H. vulgare* ssp. *spontaneum*, debido a la fuerte semejanza que existe entre ellas. Sin embargo no explicó la aparición de las cebadas de 6 carreras. Debido a las grandes diferencias morfológicas que existen entre las cebadas de 2 y 6 carreras, se creyó firmemente que los 2 tipos representaban eventos de domesticación filogenéticamente separados. El descubrimiento de formas de 6 carreras con un raquis frágil en el Oeste de China al principio de la década de los años 1930 parecía resolver el problema (Bothmer *et al.*, 2003). Åberg (1938) describió este material como *H. agriocrithon*, y fue asumido como el ancestro de las cebadas de 6 carreras (Åberg 1940, en Bothmer *et al.*, 2003). Sin embargo, las opiniones más recientes vuelven a defender la presencia de una sola línea evolutiva que conduce a la cebada de 2 carreras, y teniendo como único progenitor a *H. vulgare* ssp. *spontaneum* (Bothmer *et al.*, 2003).

El ancestro silvestre *H. vulgare* ssp. *spontaneum* C. Koch de la cebada cultivada sobrevive en los hábitats primarios del Creciente Fértil, desde Israel y Jordania, al sur de Turquía, Kurdistan iraquí y suroeste de Irán (Harlan, 1968). De hecho, ocupa todavía amplios espacios en dicha región (Zohary, 1964; Harlan y Zohary, 1966).

Aunque el Creciente Fértil sea frecuentemente citado en la bibliografía como centro de origen de la cebada, diferentes estudios han propuesto la posible existencia de otros centros de origen o de diversificación tales como el Tibet, Etiopía y el oeste de la cuenca Mediterránea.

Etiopía es uno de los centros de origen propuestos por Vavilov (1926, en Molinacano *et al.*, 2002), en parte por la amplia diversidad de los tipos morfológicos de la

cebada existentes en la región. Más tarde se aceptará que es poco probable que la cebada fuera domesticada en una zona donde no se ha encontrado ningún ancestro no cultivado (Vavilov, 1940 en Molina-Cano *et al.*, 2002). Sin embargo, aún en fechas relativamente recientes se reclamaba todavía el estatus de centro de origen para Etiopía (Bekele, 1983). A propósito del Tibet, Takahashi (1955) demostró que las cebadas espontáneas de esta región poseen características morfológicas y bioquímicas distintas de las del Oriente Próximo.

En cuanto al oeste del Mediterráneo, Molina-Cano *et al.* (1982) detectaron la presencia de *H. spontaneum* C. Koch en el sur de Marruecos. Este hallazgo provocó un debate científico sobre la posibilidad de que Marruecos pudiera ser otro centro independiente de domesticación o un centro de diversificación (Molina-Cano *et al.*, 1999; Zohary, 1999).

1-1-2. Taxonomía y citología

El genero *Hordeum* pertenece a la tribu de las *Triticeae*, familia *Poaceae* (*Gramineae*), e incluye 32 especies conocidas, ampliamente distribuidas en todo el mundo.

En la taxonomía moderna *H. vulgare* L., *H. spontaneum* C. Koch y *H. agriocrithon* Åberg están considerados como sub-especies de *H. vulgare* L. (Bothmer y Jacobsen, 1985 en Badr *et al.*, 2000).

El genero *Hordeum* contiene especies diploides ($2n=2x=14$), tetraploides ($2n=4x=28$), y hexaploides ($2n=6x=42$), aunque la mayoría sean de la primera categoría (Briggs, 1978). La especie *H. vulgare* L. es diploide.

El genoma de *H. vulgare* L. contiene alrededor de 5.5 pg de DNA por núcleo haploide, lo que equivale aproximadamente a 5×10^9 pb (Arumuganathan y Earle, 1991). El genoma de la cebada ha sido muy bien caracterizado por métodos de genética clásica y citogenética, que han proporcionado un catálogo de más de 1000 genes conocidos (von Wettstein-Knowles, 1992).

La cartografía comparativa entre el genoma de la cebada y otras especies de la familia de *Triticeae* muestra que los cromosomas 1 a 7 de la cebada son homeólogos de los cromosomas 7H, 2H, 3H, 4H, 1H, 6H, 5H, respectivamente, de las otras especies de esa familia (Linde-Laursen *et al.*, 1997).

1-1-3. Tipos de cebada:

Existen diferentes tipos de cebada, según se consideren criterios morfológicos, agronómicos, genéticos o de uso final. Los criterios de clasificación más relevantes para el presente trabajo son:

La morfología. En las cebadas de dos carreras, en la tríada de espiguillas presentes en cada entrenudo del raquis, sólo se desarrolla la espiguilla central, las laterales abortan. Sin embargo en las cebadas de seis carreras se desarrollan las tres espiguillas (Briggs, 1978).

La adaptación agronómica. Existen cebadas de invierno o ciclo largo y cebadas de primavera o ciclo corto. Las cebadas de invierno requieren vernalización para florecer, por eso se siembran en otoño. Las cebadas de primavera no necesitan un periodo de vernalización.

En latitudes altas y regiones montañosas, la cebada se suele sembrar en primavera, para escapar a los daños causados por el frío del invierno. En estas regiones, se requieren tipos de primavera para crecer y florecer normalmente. Los tipos de primavera prevalecen también en altitudes bajas, donde las temperaturas son demasiado altas para inducir vernalización en los tipos de invierno. En las regiones de latitudes medias, incluyendo el norte de África, el sur de Europa, Nepal, China y Japón, los dos tipos de cebadas (primavera e invierno) se siembran generalmente en otoño (Bothmer *et al.*, 2003).

1-2. Importancia económica y distribución geográfica

La cebada es una de las especies cultivadas más antiguas del mundo. Hoy en día es ampliamente cultivada y crece en condiciones ambientales muy diversas, incluyendo latitudes y altitudes extremas (Ullrich, 2002). Puede ser cultivada desde el Ecuador hasta Escandinavia, y desde las regiones más húmedas de Europa hasta los oasis del Sahara y Asia (Briggs, 1978).

A nivel mundial, la cebada ocupa el cuarto lugar en importancia entre los cereales, después del trigo, el arroz y el maíz. La superficie dedicada a su cultivo es de 57 millones de hectáreas, o sea, el 8,4% de la superficie total cultivada con cereales (FAO, 2004).

En España, la cebada se clasifica en el primer lugar entre los cereales desde el punto de vista de producción total, ocupando el 48% de la superficie total dedicada a los

mismos. A nivel mundial, se considera como el sexto país productor de cebada después de Rusia, Canadá, Alemania, Ucrania y Francia (FAO, 2004).

La mayoría de la producción de la cebada en España se localiza en las regiones de Castilla y León, Castilla-La Mancha y Aragón. En estas 3 comunidades se encuentran el 80% de las tierras dedicadas al cultivo de la cebada (Figura 1.1)

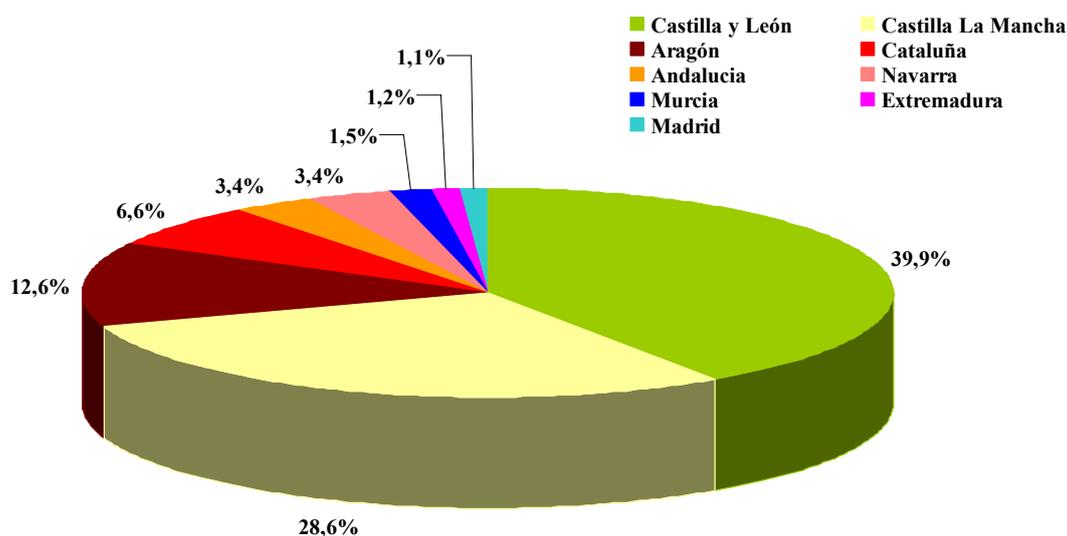


Figura 1.1 Distribución por Comunidades Autónomas de la superficie dedicada al cultivo de la cebada en España (Fuente: Anuario de Estadística Agroalimentaria M.A.P.A., 2004.)

La superficie dedicada al cultivo de cebada de 2 carreras tiende a aumentar cada vez más, en comparación con la dedicada a las cebadas de 6 carreras. Esto se debe probablemente a la introducción de un mayor número de variedades de cebada de 2 carreras obtenidas en Europa, más productivas y de mejor calidad.

El cultivo de cebada está destinado a diferentes usos. Su utilización en la alimentación humana persiste en las regiones montañosas del Asia Central y del Suroeste, así como en los países del norte de África y Etiopía (Bothmer *et al.*, 2003). Su uso en la alimentación animal crece cada vez más. Hoy en día, se estima que el 85% de la producción mundial, está destinada a este último propósito (Fischbeck, 2002). La producción de malta es el segundo uso más importante de la cebada. Alrededor de 18 millones de toneladas de cebada se destinan cada año a la fabricación de cerveza (Fischbeck, 2002). En España, la cebada se cultiva principalmente para la producción de pienso.

1-3. Mejora de la cebada

El acervo genético de este cultivo ha sufrido importantes cambios durante los procesos de domesticación y de mejora. Durante las últimas décadas, la mejora de la cebada en Europa se ha centrado sobre diferentes aspectos como rendimiento, calidad del grano y adaptación a estreses bióticos y abióticos. Según Fischbeck (1992), en Europa Central y del Noroeste, la mejora del último siglo se centró fundamentalmente sobre la cebada de dos carreras de primavera. La mejora moderna comenzó con una primera etapa de selección en variedades tradicionales. Posteriormente, a partir de la 3ª década del siglo XX, se hicieron cruzamientos entre variedades de alto rendimiento, que procedían de zonas geográficas distintas. En una fase posterior, se llevaron a cabo cruzamientos entre variedades que salieron del primer ciclo de cruzamientos. En los últimos años de desarrollo de cultivares de cebada de primavera en Europa, ha habido introgresión de germoplasma exótico para introducir genes de resistencia a enfermedades, como el oídio y la roya. Para no afectar a la calidad maltera de las variedades, este proceso de introgresión de genes, que ha empleado como donantes variedades locales exóticas, se hizo con la precaución de realizar cruzamientos intermedios con un material mejorado. Esta estrategia de introgresión controlada ha dado buenos resultados.

Las cebadas de invierno ocuparon poca superficie comparándolas con las de primavera en Europa Central y del Oeste hasta la generalización de la mecanización, lo que permitió la preparación de grandes superficies para siembras precoces (Fischbeck, 1992). Eso explica el reducido número de variedades de invierno autóctonas disponibles para la mejora. Las cebadas de invierno de hoy en día, provienen de dos fuentes de variedades tradicionales, unas originarias de los Países Bajos y otras de Canadá, en particular la variedad *Mammuth* (Aufhammer, 1982). La escasez de fuentes de germoplasma europeo de cebada de invierno llevó a los mejoradores a hacer cruzamientos con las cebadas de dos carreras de primavera. El empleo de material exótico para la introducción de genes de resistencia ha sido también importante en la mejora de la cebada de invierno de seis carreras durante los últimos ciclos de mejora. La variedad *Dura*, que proviene de la recombinación entre *Ragusa* (línea tradicional de Croacia resistente al oídio) y *Friedrichswerther Berg*, dominó la producción de cebada de invierno durante dos décadas en Europa Central. Del mismo modo, tuvo gran difusión la variedad *Vogelsanger Gold*, que deriva de un cruzamiento entre una línea de

cebada de primavera resistente al oidio, retrocruzada con diferentes cebadas de invierno de seis carreras, entre ellas *Mammuth* y *Hauter*. Las variedades *Vogelsanger Gold* y *Dura* se encuentran en el pedigrí de la mayoría de los cultivares recientes de cebada de invierno de seis carreras. Por otra parte, el cultivar *Hauter*, cultivado durante un tiempo en Francia, contribuyó al desarrollo de variedades de invierno de seis carreras en Francia. Los cultivares *Ager* y *Astrix*, que provienen de un doble cruzamiento entre *Hauter* y *Bordia*, así como *Kenia* y *Hatif de Grignon*, ocupan un lugar importante en el desarrollo de cultivares de cebada de invierno en la región Sur de Europa. Estas dos variedades aparecen en el pedigrí de casi todos los cultivares desarrollados en esta zona. En general, se puede concluir que unas pocas variedades han dominado la mejora de la cebada de invierno de seis carreras en Europa.

La mejora de cebadas de invierno de dos carreras se desarrolló especialmente después de la Segunda Guerra Mundial, con el fin de desarrollar variedades de invierno destinadas a la producción de malta. El cultivar *Malta* fue un parental importante en los cruzamientos para la selección de cultivares de cebadas de dos carreras de invierno. *Sonja* e *Igri*, derivados ambos del cultivar *Malta*, ocuparon durante mucho tiempo grandes superficies en Europa Central y del Oeste, sin llegar a un nivel de producción de calidad maltera aceptable.

En resumen, en las últimas décadas de la mejora de cebada en Europa, la base genética empleada para la obtención de nuevas variedades ha sido cada vez más estrecha. Hay unos pocos cultivares que se encuentran en la mayoría de los pedigríes de cultivares de cebadas modernas. Graner *et al.* (1994) confirmaron este hecho, al analizar 48 cultivares europeos con RFLPs, demostrando que estos cultivares incluyen en su pedigrí, por lo menos, uno de los 4 progenitores principales empleados en la mejora.

1-4. La importancia de la conservación de los recursos fitogenéticos

El desarrollo de la agricultura mecanizada es una de las causas de la erosión genética. Con la aparición de las variedades modernas muy productivas, uniformes y adaptadas a las técnicas agrícolas modernas, los agricultores se deshicieron de las variedades tradicionales. La consecuencia fue la desaparición de un patrimonio genético importante en muchos países europeos. La FAO estima en un 75% la pérdida de la diversidad genética de las plantas cultivadas desde el principio del siglo XX. Esto nos hace depender cada vez más de un número reducido de variedades cultivadas (FAO, 1993). Esta uniformidad genética, favorecida por la industrialización de la agricultura,

es una invitación al desastre, por el aumento de la vulnerabilidad de los cultivos frente a las enfermedades y plagas.

En el caso de la cebada, el reemplazo de las variedades originales empezó gradualmente en el Reino Unido durante la primera mitad del siglo XIX, donde un stock de semillas “mejoradas” fue seleccionado por los agricultores. Cuando su rendimiento era más alto o de mejor calidad, atraían el interés general y se distribuían ampliamente. Durante la segunda mitad del siglo XIX, estas semillas alcanzaron el resto del continente (Fischbeck, 2003). De este modo, las variedades tradicionales fueron rápidamente sustituidas en casi toda Europa Central y del Noroeste. En las primeras décadas del siglo XX, las variedades tradicionales desaparecieron, por el hecho de que en esta época no se tomaron iniciativas para la conservación de semillas de estas variedades (Schachl, 1975). Solamente las variedades tradicionales cultivadas en las regiones montañosas, donde la agricultura intensiva fue menos aplicada, sufrieron menor erosión. Aunque probablemente, estas variedades locales que restan no representan la totalidad de la diversidad genética de las poblaciones originales (Fischbeck, 2003). En España e Italia, la sustitución de las variedades tradicionales se retrasó hasta la segunda mitad del siglo XX, lo que produjo una ventaja adicional, pues en esta época ya existía la preocupación de prevenir la extinción de estas variedades, por lo que muchas de ellas fueron recolectadas con objeto de guardarlas en bancos de germoplasma (Fischbeck, 2003). En el Banco Nacional de Germoplasma español se encuentran 2.155 entradas de cebada, de las cuales el 91% son variedades tradicionales españolas. Muchas de ellas fueron recogidas antes de 1942, o sea, antes de la introducción masiva de variedades foráneas. La recolección de estas variedades cubrió la totalidad de las zonas geográficas del cultivo de la cebada en el país (Igartua *et al.*, 1998).

Las regiones del WANA (*West Asia, North Africa*) y Etiopía fueron menos afectadas por la agricultura intensiva. En esta parte del mundo, las variedades tradicionales de cebada aún ocupan el 85% de la superficie cultivada (Fischbeck, 2003). Las condiciones difíciles a las que están sometidas estas regiones, son la principal razón para la subsistencia de estas variedades. En efecto, los cultivares modernos seleccionados para condiciones óptimas no se adaptan a la aridez del clima de estas regiones, y presentan unos rendimientos peores que las variedades tradicionales locales.

Siria es uno de los países que sufre el problema de adaptación de las variedades modernas. Para solucionar el problema, el ICARDA elaboró un programa de mejora

basado en la selección de nuevas variedades a partir del material local adaptado. Ceccarelli y Grando (1989) demostraron que la selección por el rendimiento en condiciones de estrés hídrico no es eficiente para el germoplasma élite convencional. Las líneas seleccionadas a partir de variedades tradicionales locales mostraron una mayor estabilidad de producción y un rendimiento superior a los de las variedades modernas, en condiciones desfavorables (Ceccarelli *et al.*, 1991). Este enfoque ha sido aplicado con éxito, como lo demuestra la aceptación de las variedades procedentes del programa de mejora de ICARDA en la región WANA.

En España, no es extraño que los agricultores de las zonas semi-áridas prefieran sembrar cultivares antiguos o incluso variedades tradicionales, en lugar de cultivares modernos, por problemas de adaptación de estos últimos (Lasa *et al.*, 2001). Aunque las primeras no tienen una buena producción, son más estables y aseguran al agricultor un rendimiento mínimo, en caso de estrés hídrico severo, lo que ocurre frecuentemente en estas zonas. En Aragón, el 69,4% de la superficie dedicada al cultivo de la cebada de 6 carreras, está ocupada por el cultivar Albacete (DGA, 2002), una variedad española derivada de una población local, en la Estación Experimental de Aula Dei. Siguiendo la experiencia del ICARDA, un mayor empleo del germoplasma autóctono podría dar un nuevo impulso a la selección de variedades adaptadas a las condiciones mediterráneas. En este sentido, Tolbert *et al.* (1979) pusieron de manifiesto la singularidad genética de las cebadas del Oeste del mediterráneo, incluyendo las españolas. Lo que sugiere la posible existencia de caracteres de adaptación en el material autóctono que no fue empleado en su totalidad en la mejora de la especie. Igartua *et al.* (1998) sugirieron que las variedades tradicionales españolas, que evolucionaron independientemente durante un periodo que pudo ser de hasta miles de años bajo las condiciones mediterráneas, serían probablemente una fuente de alelos para adaptación a estas condiciones. Comparando las variedades españolas antiguas de 2 y 6 carreras con otro germoplasma europeo, mediante análisis de RFLPs, Casas *et al.* (1998) confirmaron la singularidad de algunas entradas españolas. Sin embargo, este germoplasma local fue durante mucho tiempo infrautilizado en los programas de mejora, aunque demostró su valor en la obtención de nuevas variedades tales como Candela (Estación Experimental de Aula Dei del Consejo Superior de Investigaciones Científicas, EEAD-CSIC) y Kika (Instituto Técnico Agronómico Provincial de Albacete, ITAP). La primera se obtuvo a partir del cruzamiento del cultivar tradicional español Pané con la variedad francesa Barbarrosa,

mientras que la segunda se obtuvo del cruzamiento de la variedad española Albacete con la variedad inglesa Kym.

Una de las líneas de trabajo que se venía ejecutando en el programa de mejora de cebada de la EEAD-CSIC era la evaluación agronómica de las entradas del Banco Nacional de Germoplasma. Sin embargo la utilización práctica de un número tan alto de entradas es bastante complejo en un programa de mejora. Como solución a este problema, se creó una colección nuclear representativa de la diversidad genética existente en la totalidad de las entradas del Banco de Germoplasma (Igartua *et al.*, 1998) con la colaboración del IRTA de Catalunya y el ITA de Castilla y León. La evaluación de esta colección nuclear permitiría su uso más eficaz en programas de mejora.

1-5. Referencias bibliográficas.

- Åberg E. 1938. *Hordeum agriocrithon* nova sp., a wild six rowed barley. Annals Agricultural College of Sweden, 6: 159-216
- Anuario de Estadística Agroalimentaria. 2004. Ministerio de Agricultura, Pesca y Alimentación.
- Arumuganathan K, Earle ED. 1991. Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter, 9: 208-218.
- Aufhammer G. 1982. Abstammung der Wintergersten-Sorten – Ahnentafel. Bayer. Landw. Jb. 59: 536-544.
- Badr A, Müller K, Schäfer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F. 2000. On the origin and domestication history of barley (*Hordeum vulgare*). Molecular Biology and Evolution, 17: 499-510.
- Bekele E. 1983. A differential rate of regional distribution of barley flavonoid patterns in Ethiopia, and a view on the center of origin of barley. Hereditas, 98: 269-280.
- Bothmer R von, Sato K, Komatsuda T, Yasuda S, Fischbeck G. 2003. The domestication of cultivated barley. In Diversity in Barley. Ed. Bothmer R von. Hintum T van, Knüpffer H, Sato K. Elsevier Science B.V., Amsterdam, The Netherlands, 9-27.
- Briggs DE. 1978. Barley, Chapman Hall, London, 612 p.
- Casas AM, Igartua E, Vallés MP, Molina-Cano JL. 1998. Genetic diversity of barley cultivars grown in Spain, estimated by RFLP, similarity and coancestry coefficients. Plant Breeding, 117: 429-435.
- Ceccarelli S, Grandó S. 1989. Efficiency of empirical selection under stress conditions in barley. Journal Genetics & Breeding, 43: 25-31.

- Ceccarelli S, Acevedo E, Grandó S. 1991.** Breeding for yield stability in unpredictable environments: single traits, interaction between traits, and architecture of genotypes. *Euphytica*, 56: 169-185.
- DGA. 2002.** Encuestas productivas de cultivos herbáceos de invierno. Campaña 1999-2000. Departamento de Agricultura del Gobierno de Aragón. <http://portal.aragob.es>
- FAO. 1993.** Harvesting nature's diversity. FAO October, 1993.
- FAO. 2004.** FAOSTAT Database. <http://www.fao.org/>.
- Fischbeck G. 1992.** Barley cultivar development in Europe-success in the past and possible changes in the future. *In* Barley Genetics VI. Ed. L. Munck, Munksgaard International Publishers, Copenhagen, Vol. II: 885-901.
- Fischbeck G. 2002.** Contribution of barley to agriculture. A brief overview. *In* Barley Science. Recent Advances from Molecular Biology to Agronomy of Yield and Quality. Ed. GA Slafer, JL Molina-Cano, R Savin, JL Araus, I Romagosa, Haworth Press, New York, 1-10.
- Fischbeck. G. 2003.** Diversification through breeding. *In* Diversity in barley (*Hordeum vulgare*). Ed. Bothmer R von. Hintum T van, Knüpffer H, Sato K. Elsevier Science B.V., Amsterdam, The Netherlands, 29-52.
- Graner A, Ludwig WF, Melchinger AE. 1994.** Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. *Crop Science*, 34: 1199-1205.
- Harlan JR. 1968.** On the origin of barley. *USDA Agriculture Handbook* 338: 9-31.
- Harlan JR, Zohary D. 1966.** Distribution of wild wheats and barley. *Science*, 153: 1074-1080.
- ICARDA. 1997.** ICARDA: Saving the genes of our future food. February 1997. www.icarda.org.
- Igartua E, Gracia MP, Lasa JM, Medina B, Molina-Cano JL, Montoya JL, Romagosa I. 1998.** The Spanish barley core collection. *Genetic Resources and Crop Evolution*, 45: 475-481.
- Lasa JM, Igartua E, Ciudad FJ, Codesal P, García EV, Gracia MP, Medina B, Romagosa I, Molina-Cano JL, Montoya JL. 2001.** Morphological and agronomical diversity patterns in the Spanish barley core collection. *Hereditas*, 135: 217-225.
- Linde-Laursen I, Heslop-Harrison JS, Shepherd KW, Taketa S. 1997.** The barley genome and its relationship with the wheat genomes. A survey with an internationally agreed recommendation for barley chromosome nomenclature. *Hereditas*, 126: 1-16.
- Molina-Cano JL, Gómez-Campo C, Conde J. 1982.** *Hordeum spontaneum* C. Koch as a weed of barley fields in Morocco. *Zeitschrift für Pflanzenzüchtung*, 88: 161-167.

- Molina-Cano JL, Moralejo M, Igartua E, Romagosa I. 1999.** Further evidence supporting Morocco as centre of origin of barley. *Theoretical & Applied Genetics*, 98: 913-918.
- Molina-Cano JL, Igartua E, Casas AM, Moralejo M. 2002.** New views on the origin of cultivated barley. *In Barley Science. Recent Advances from Molecular Biology to Agronomy of Yield and Quality*. Ed. GA Slafer, JL Molina-Cano, R Savin, JL Araus, I Romagosa, Haworth Press, New York, 15-29.
- Schachl R. 1975.** Refugia of Austrian land varieties in the subalpine region. *In Barley Genetics III. Proc. 3rd Int. Barley Genetics Symposium*, Thiemig, München, 182-190.
- Takahashi R. 1955.** The origin of cultivated barley. *Advances in Genetics*, 7: 227-266.
- Tolbert DM, Qualset CO, Jain SK, Craddock JC. 1979.** A diversity analysis of a world collection of barley. *Crop Science*, 19: 789-794.
- Ullrich SE. 2002.** Genetics and breeding of barley feed quality attributes. *In Barley Science. Recent Advances from Molecular Biology to Agronomy of Yield and Quality*. Ed. GA Slafer, JL Molina-Cano, R Savin, JL Araus, I Romagosa, Haworth Press, New York, 115-142.
- von Wettstein-Knowles P. 1992.** Cloned and Mapped Genes: Current Status. *In Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology*. Ed. PR Shewry. CAB International, Wallingford, UK, 73-98.
- Yadun SL, Gopher A, Abbo S. 2000.** The cradle of agriculture. *Science*, 288: 1602-1603.
- Zohary D. 1964.** Spontaneous brittle six-row barleys, their nature and origin. *In Barley Genetics I*. Ed. S Broekhuizen, G Dantuma, H Lamberts, W Lange. Pudoc, Wageningen, 27-31.
- Zohary D. 1999.** Monophylitic vs polyphylitic origin of the crops on which agriculture was founded in the Near East. *Genetic Resources and Crop Evolution*, 46: 133-142.

Capítulo 2

Objetivos

2. Objetivos

El objetivo general de este trabajo se centra en la evaluación de la colección nuclear de cebadas españolas. Los objetivos específicos son:

- Caracterización de la diversidad genética de la colección nuclear española mediante marcadores moleculares, y su emplazamiento en el panorama de la diversidad genética europea (**Capítulo 3**).
- Identificación de relaciones entre patrones de variabilidad genética y variabilidad geográfica y climática (**Capítulo 4**).
- Evaluación del comportamiento de las entradas de la colección nuclear de cebadas españolas en zonas representativas del cultivo de cebada en España (**Capítulo 5**).
- Analizar los factores genotípicos y ambientales que determinan el rendimiento en condiciones españolas (**Capítulo 5**).
- Análisis de asociación o *linkage disequilibrium mapping* en las entradas de la colección nuclear de cebadas españolas. Se analizará la asociación de marcadores moleculares con treinta caracteres agronómicos y morfológicos (**Capítulo 6**).

Capítulo 3

Diversidad genética

3-1. Introducción

En muchos cultivos, la introducción de la agricultura moderna provocó el reemplazo de las variedades tradicionales por cultivares mejorados de alto rendimiento. Este fenómeno, a su vez, produjo una reducción de la base genética de los cultivares, especialmente en países desarrollados (Briggs, 1978). No está claro si éste ha sido el caso de la cebada en Europa. Algunos autores (por ejemplo, Koebner *et al.*, 2003) no encontraron evidencia de erosión genética como consecuencia de la mejora moderna de la cebada en Europa. La mayoría de estudios en este sentido, sin embargo, sí ofrecen datos que sustentan esa hipótesis (Graner, *et al.*, 1994; Melchinger *et al.*, 1994; Ellis *et al.*, 1997; Russell *et al.*, 2000). Una de las causas que han llevado a esta erosión genética pudo ser el limitado número de variedades locales que se emplearon en las primeras etapas de la mejora moderna de la cebada. Por tanto, es posible que la diversidad genética presente en las variedades locales europeas originales no haya sido explotada en su totalidad.

En España, el Centro Nacional de Recursos Fitogenéticos mantiene una colección de más de 2000 entradas de cebada cultivada. La mayoría son variedades locales recolectadas en la primera mitad del siglo XX (Lasa *et al.*, 2001). Algunos de estos materiales han sido empleados con éxito en programas de mejora, pero en menor medida que las variedades de otros países europeos no mediterráneos. La cebada ha sido cultivada en España desde tiempos prehistóricos. Dada su historia de adaptación y selección bajo condiciones mediterráneas, los materiales autóctonos españoles pueden albergar genes y alelos de adaptación que probablemente no han entrado aún en programas de mejora (Igartua *et al.*, 1998).

La evaluación de la utilidad de estos recursos genéticos para la obtención de variedades constituye un objetivo de investigación en sí mismo. Con este objeto, se sintetizó una colección nuclear de cebadas españolas, representativa de la variabilidad guardada en el banco nacional (Igartua *et al.*, 1998).

A lo largo del tiempo, los métodos de detección y de análisis de la diversidad genética han evolucionado gradualmente, desde los análisis basados sobre caracteres morfológicos, a las evaluaciones electroforéticas de variantes bioquímicas y, más recientemente, al análisis molecular de variantes de las secuencias del ADN.

En la cebada, los marcadores morfológicos han sido utilizados por Cross (1994) para la identificación de una colección de cebadas de primavera, y por Bahrman *et al.*

(1999), para el análisis de la diversidad de una colección de variedades antiguas francesas. Los marcadores morfológicos tienen sus límites en los estudios de diversidad, fundamentalmente por el número de caracteres disponible. Por eso, la mayoría de los estudios recientes utilizan marcadores moleculares. Estos marcadores permiten una caracterización de la diversidad más potente y precisa. Los RFLPs son los primeros marcadores moleculares que se emplearon para este fin. Diferentes autores utilizaron estos marcadores para estudios de diversidad en cebada con éxito, como Saghai-Marooif *et al.* (1994a) en la evaluación de cultivares de cebada de calidad maltera, o Casas *et al.* (1998), quienes emplearon los RFLPs para estudiar la diversidad genética de un conjunto de variedades de cebada cultivadas en España.

El descubrimiento de las técnicas basadas sobre la PCR, hizo que los investigadores se dirigieran más a este tipo de marcadores, de uso más fácil. Los RAPDs y los AFLPs, son los primeros métodos basados en esta técnica. Los dos se manejaron mucho en estudios de diversidad en cebada. Manninen y Nissilä (1997) usaron RAPDs para la evaluación de la diversidad de la cebada de seis carreras de primavera. Baum *et al.* (1998) aplicaron los RAPDs en la identificación de las cebadas canadienses. Ellis *et al.* (1997) utilizaron AFLPs para evaluar la diversidad genética de las cebadas europeas de primavera, y estimar la relación de parentesco entre ellas. También Hayes *et al.* (1997) usaron estos marcadores para la caracterización y explotación de la diversidad genética de las cebadas americanas y el estudio de los caracteres cuantitativos presentes en ellas.

Recientemente, los investigadores recurren cada vez más a los microsatélites, por su facilidad de uso, el alto grado de polimorfismo que presentan y su fiabilidad. Morgante y Olivieri (1993) los describieron como la herramienta más eficaz para la detección de la variación genética inter e intra-específica.

La eficacia de los microsatélites se ha demostrado en diferentes estudios de diversidad genética, Struss y Plieske (1998) aplicaron estos marcadores para detectar la diversidad existente en cebadas espontáneas y cultivadas. Russell *et al.* (1997) usaron microsatélites para evaluar la diversidad genética existente en algunas variedades europeas de cebada. Pillen *et al.* (2000) y Maestri *et al.* (2002) estudiaron respectivamente la variación genética en cebadas cultivadas alemanas e italianas. Matus y Hayes (2002) estudiaron la diversidad en tres grupos de germoplasma, incluyendo variedades y descendencias de cruzamientos entre líneas élite de germoplasma americano. De forma similar, Karakousis *et al.* (2003) identificaron variedades del

germoplasma australiano, y Russell *et al.* (2003) variedades locales de Siria y Jordania, mediante microsatélites. Otros autores como Dávila *et al.* (1999), Ivandic *et al.* (2002) usaron estos marcadores, para estudiar la diversidad en *Hordeum vulgare ssp. spontaneum*.L.

El objetivo de este capítulo es la caracterización de la diversidad genética de la colección nuclear de cebadas españolas mediante marcadores moleculares, y su emplazamiento en el panorama de la diversidad genética europea. Por las ventajas que muestran los microsatélites, particularmente su alta capacidad de detección de polimorfismos, fueron elegidos para llevar a cabo este análisis.

3-2. Material y Métodos

3-2-1. Evaluación de la diversidad por marcadores moleculares

3-2-1-1. Material vegetal

Se analizaron un total de 225 genotipos de cebadas de 2 y 6 carreras, de los cuales 159 son líneas puras de la colección nuclear española de cebada (Igartua *et al.*, 1998). Estas líneas puras provienen de poblaciones locales, probablemente heterogéneas. Para aislar líneas puras, se eligió una espiga de cada población, a partir de la cual se hicieron autofecundaciones durante cuatro generaciones en parcelas totalmente aisladas (Lasa *et al.*, 2001).

Los 225 genotipos están divididos en grupos de germoplasma, según la información disponible *a priori*, resultando cuatro grupos según los genotipos sean de 2 ó 6 carreras, españoles o de otras procedencias (Tabla 3.1).

Las líneas puras de la colección nuclear proceden de diferentes provincias del territorio español, agrupadas en regiones agro-ecológicas, definidas por Igartua *et al.* (1998), según las zonas de cultivo de la cebada y los rendimientos históricos registrados en las diferentes provincias del territorio español durante 47 años (1948-1994). El número de entradas por provincias está representado en la Tabla 3.2.

Tabla 3.1. Distribución de los 225 genotipos en 4 grupos de germoplasma.

Grupo	Descripción	Nº de entradas
1	Seis carreras, origen español	152
	- Líneas puras de la colección nuclear	148
	- Cultivares procedentes de poblaciones españolas	4
2	Seis carreras de otras procedencias (variedades de referencia)	33
3	Dos carreras, origen español (líneas puras de la colección nuclear)	11
4	Dos carreras de otras procedencias (variedades de referencia)	29
Total		225

El nombre, origen y pedigrí de las variedades de referencia, así como el de los 4 cultivares españoles de 6 carreras están representados en la Tabla 3.3.

El conjunto de variedades de referencia se introdujo en el estudio para servir de base de comparación de la diversidad genética entre el germoplasma español y foráneo, principalmente europeo.

Estos 225 genotipos se sembraron en invernadero. Se cosecharon hojas de las plántulas 14 días después de la siembra, para la extracción del DNA. Esta extracción se hizo según el protocolo descrito por Saghai-Marroof *et al.* (1984).

Tabla 3.2. Distribución geográfica, por región agro-ecológica y por provincia de las entradas de la colección nuclear.

Grupos	Número de entradas	
	6 carreras	2 carreras
1- Zona de alto rendimiento		
Álava	2	
Barcelona	3	2
Cuenca	7	
Gerona	1	
Lérida	5	1
Navarra	2	
Rioja	3	
Tarragona	2	
2- Zona del Centro y valle del Ebro		
Albacete	8	
Castellón	1	
Guadalajara	3	1
Huesca	2	
Madrid	2	
Soria	3	1
Teruel	4	2
Toledo	8	
Zaragoza	6	
3- Zona de Castilla y León		
Ávila	3	
Burgos	5	
León	1	1
Palencia	3	
Salamanca	3	
Segovia	5	
Valladolid	4	1
Zamora	3	1
4- Zona Sur		
Alicante	4	
Almería	3	
Badajoz	7	1
Baleares	1	
Cáceres	3	
Cádiz	3	
Ciudad Real	7	
Córdoba	3	
Granada	4	
Jaén	4	
Huelva	3	
Málaga	4	
Murcia	4	
Sevilla	5	
Valencia	1	
5- Tenerife	3	
Total	148	11

Tabla 3.3 Pedigrí y origen de las variedades analizadas.

Nombres	Origen	Número de carreras	Pedigrí
<u>Variedades españolas</u>			
Albacete*	España	6	Selección a partir de una variedad local española
Almunia*	España	6	Selección a partir de una variedad local española
Pané*	España	6	Selección a partir de una variedad local española
Candela	España	6	Barbarrosa x Pané
<u>Variedades de referencia de 6 carreras</u>			
Ager*	Francia	6	Bordia x Kenia x LBP 259/711
Barbarrosa*	Francia	6	(Hauter x (Hatif de Grignon x Ares)) x Ager
Dobla*	Francia	6	Union x Nympe
Hatif de Grignon*	Francia	6	Variedad tradicional francesa
Monlón*	Francia	6	Breusted Schladener x Hatif de Grignon
Asplund	Suecia	6	Selección en una población indefinida
Athenais	Grecia	6	Variedad tradicional griega
Athene	Reino unido- Alemania	6	Herfordia x H204 (<i>Hordeum sp nigrum</i>)x Mädru x Weissenhaus-Stamm
Banteng	Holanda	6	((Platen 23/49 x Vinesco)x Dea)) x Jumbo
Bordia	Bélgica	6	(Lignée 66x Lignée 125)
Dea	Alemania	6	((Ragusa x Peragis 12) x (Heils Franken x Friedrichswerther Berg)) x (Ragusa x Mahnd. Viktoria) x (Bolivia x Ragusa))
Dura	Alemania	6	(Friedrichswerther Berg x Ragusa) x Doria
Frisia	Alemania	6	Granat x Pyrthyjärven x Eckendorfer (WiG) x Kalkreuther
Gerbel	Francia	6	(Ager x Jumbo) x FDE 244/95 (Ager x Asterix x Mana)
Hauter	Alemania-Francia	6	((Ragusa x Peragis 12) x (Heils Franken x Friedrichswerther. Bergx Tschermaks 2z)x (Ragusa x Mahndofer Viktoria x (Bolivia x Ragusa)
Herfordia	Alemania	6	Peragis Stamm x Schladener I
Juli	Dinamarca	6	Selección en una variedad primitiva danesa
Maskin	Noruega	6	Selección en una variedad local noruega
Mirco	Italia	6	Selección en una variedad local italiana

Tabla 3.3. (Continuación)

Nombres	Origen	Número de carreras	Pedigrí
<u>Variedades de referencia de 6 carreras</u>			
Olli	Finlandia	6	Selección en una variedad local finlandesa
Plaisant	Francia	6	Ager x Nymphe
Ragusa B	Yugoslavia	6	Selección en una variedad local de Dalmacia
Senta	Alemania	6	(Firlbecks 4Zlg x Ungarische) x Dea
Tapir	Holanda	6	DSGW167 x Pella
Vega Svalöf	Suecia	6	Selección en una variedad local sueca
Vindicat	Holanda	6	Selección en una variedad local holandesa
Vogelsanger Gold	Alemania	6	(Isaria x H204 (<i>Hordeum. sp nigrum</i>) x WG5
S-36	Siria	6	Selección ICARDA
S-45	Siria	6	Selección ICARDA
Morex	EEUU	6	Cree x Bonanza
Steptoe	EEUU	6	Washington sel. 3564 x Unitan
Mammuth	Canadá	6	Variedad local canadiense
Orria	España	6	Selección a partir de material CIMMYT
<u>Variedades de referencia de 2 carreras</u>			
Alexis	Alemania	2	Br.1622 x Triumph
Angora	Alemania	2	Br.301a x LBP 5907 (Carsten2z x Herfodia x LBP 409/250 (SG) x Isaria
Camelot	Alemania	2	Aramir x (Magnif.105 X Universe)
Cameo	Reino-Unido	2	RPB 713-77 x Kym
Clarine	Francia	2	Igri x Mogador
Gaelic	Francia	2	FDO 8109.136 x Marinka
Graphic	Reino-Unido	2	Casino x Dandy
Alpha*	Francia	2	Ager x (Ager x Ceres)
Beka*	Francia	2	Bethge XIII x Kneifel
Hassan*	Holanda	2	Delta x (Agio x Kenia 3 x var. Arabische)
Kym*	Alemania	2	Georgie x Hanna

Tabla 3.3. (Continuación)

Nombres	Origen	Número de carreras	Pedigrí
<u>Variedades de referencia de 2carreras</u>			
Pallas*	Suecia	2	Mutante de Bonus
Union*	Alemania	2	Weihenstephaner Melhtauresistente II x Firlbeck 621
Wisa*	Alemania	2	(Weih. Melh. I x Breun IN 2511) x Isaria
Zaida*	España	2	Adora x Union
Hispanic	Francia	2	Mosar x (Flika x Lada)
Labea	Alemania	2	Br.269cc x LBP 5907
Nevada	Reino-Unido	2	R 26/73 E x Igri
Tipper	Reino-Unido	2	Alpha x (Dayton x Jotun)
Tremois	Francia	2	(Dram x Aramir) x Berar
Triumph	Alemania	2	Tschemnakuy Hana x Kargyn
Volga	Francia	2	(Tosca x Kiebitz x Marion x 3873D) x Athos x Aurea x 4923D x 7375D)
Seira	Suecia	2	Selección a partir de una población compuesta sueca
Igri	Alemania	2	Malta x LBP 1427 x Ingrid
Mogador	Francia	2	Alpha x Sonja
Albaicin	España	2	Triumph x Hassan
Logan	EEUU	2	(ND7085 x Bowman sib) x ND7556
PC-4	España	2	Desconocido
S-7	Siria	2	Selección ICARDA

Fuente: El pedigrí de la mayoría de estas variedades ha sido extraído de Baumer y Cais (2000).

*Variedades introducidas en la colección nuclear española

3-2-1-2. Microsatélites analizados

La caracterización molecular de los 225 genotipos se hizo con microsatélites. Se eligieron 64 SSRs distribuidos a lo largo del genoma (Tabla 3.4).

Los microsatélites marcados con un asterisco en la Tabla 3.4, pertenecen al conjunto de los 48 descritos por Macaulay *et al.* (2001) como “*genotyping set*” para la cebada.

Los cromosomas de cebada deben ser designados de acuerdo con la homeología con las otras especies Triticeae (Linde-Laursen *et al.*, 1997). Esta recomendación fue seguida en este estudio.

Las distancias en cM presentadas en la Tabla 3.4 pertenecen al mapa genético de la población derivada del cruzamiento del cultivar Lina x *H. spontaneum* Canada Park, presentado por Ramsay *et al.* (2000).

Los microsatélites empleados en este estudio se seleccionaron a partir de secuencias públicas. Del total, 19 están ligados a genes. En la Tabla 3.5, se describe el origen de estos microsatélites.

Tabla 3.4. Descripción de los 64 microsatélites analizados, incluyendo localización cromosómica, motivo y tamaño en pb y fuente de referencia.

Microsatélite	Cr	cM	Repetición	Tamaño	Fuente
Bmac399*	5 (1H)	25	(ac)21	145	Ramsay <i>et al.</i> , 2000
HvGLUEND	5 (1H)		(t)11	262	Pillen <i>et al.</i> , 2000
Bmac032*	5 (1H)	55	(ac)7t(ca)15(at)9	215	Russell <i>et al.</i> , 1997
HvALAAT	5 (1H)	57	(c)12(tc)9	183	Pillen <i>et al.</i> , 2000
HvM020	5 (1H)	58	(ga)19	151	Liu <i>et al.</i> , 1996
Bmag211*	5 (1H)	62	(ct)16	174	Ramsay <i>et al.</i> , 2000
Bmag382*	5 (1H)	97	(ag)7aa(ag)7	109	Ramsay <i>et al.</i> , 2000
HvHVA1*	5 (1H)	112	(acc)5	151	Ramsay <i>et al.</i> , 2000
WMC1E8*	5 (1H)	164	(ac)24	197	Ramsay <i>et al.</i> , 2000
Bmac134*	2 (2H)	5	(ac)28	148	Ramsay <i>et al.</i> , 2000
HvM036*	2 (2H)	17	(ga)13	114	Liu <i>et al.</i> , 1996
Bmag378*	2 (2H)	44	(ag)14	147	Ramsay <i>et al.</i> , 2000
Bmac132	2 (2H)	50	(ac)15	187	Ramsay <i>et al.</i> , 2000
HvBKASI	2 (2H)	50	(c)10(a)11	197	Becker y Heun, 1995
Bmac093*	2 (2H)	50	(ac)24	151	Ramsay <i>et al.</i> , 2000
Bmag125*	2 (2H)	63	(ag)19	134	Ramsay <i>et al.</i> , 2000
HvM054*	2 (2H)	103	(ga)14	159	Liu <i>et al.</i> , 1996
HvLTPPB*	3 (3H)	25	(ac)10(at)5	221	Pillen <i>et al.</i> , 2000
Bmag136*	3 (3H)	50	(ag)6-(ag)10-(ag)6 (ag)6	197	Ramsay <i>et al.</i> , 2000
Bmac067*	3 (3H)	54	(ac)18	171	Ramsay <i>et al.</i> , 2000
Bmag006	3 (3H)	54	(ag)17	174	Ramsay <i>et al.</i> , 2000
Bmac209*	3 (3H)	55	(ac)13	176	Ramsay <i>et al.</i> , 2000
Bmag225*	3 (3H)	74	(ag)26	162	Ramsay <i>et al.</i> , 2000
Hv13GEIII	3 (3H)		(at)5	175	Pillen <i>et al.</i> , 2000
Bmag013*	3 (3H)	141	(ct)21	155	Ramsay <i>et al.</i> , 2000
HvM062*	3 (3H)	154	(ga)11	251	Liu <i>et al.</i> , 1996
Bmac029	3 (3H)		(ac)20	169	L. Ramsay (com. personal)
HvM040*	4 (4H)	14	(ga)6(gt)4(ga)7	160	Liu <i>et al.</i> , 1996
HvOLE	4 (4H)	21	(gcct)4	198	Ramsay <i>et al.</i> , 2000
Bmag353*	4 (4H)	45	(ag)21	119	Ramsay <i>et al.</i> , 2000
HvM003*	4 (4H)	45	(at)29	188	Saghai-Marroof <i>et al.</i> , 1994b
Bmag384*	4 (4H)	49	(ag)18	116	Ramsay <i>et al.</i> , 2000
HvBTAI3	4 (4H)		(gcc)4	234	Pillen <i>et al.</i> , 2000
EBmac701*	4 (4H)	76	(ac)23	149	Ramsay <i>et al.</i> , 2000
HvMLOH1A	4 (4H)	88	(ga)6	175	Ramsay <i>et al.</i> , 2000
HvM067*	4 (4H)	118	(ga)11	116	Liu <i>et al.</i> , 1996
HvBAMY	4 (4H)	123	(a)17	132	Pillen <i>et al.</i> , 2000
HdAMYB	4 (4H)		(tg)10(g)16	219	Pillen <i>et al.</i> , 2000

Tabla 3.4. (Continuación).

Microsatélite	Cr	cM	Repetición	Tamaño	Fuente
Bmag337	7 (5H)	35	(ag)22	145	Ramsay <i>et al.</i> , 2000
Bmac113*	7 (5H)	41	(at)7(ac)18	187	Ramsay <i>et al.</i> , 2000
Bmac096	7 (5H)	41	(at)6(ac)16	173	Ramsay <i>et al.</i> , 2000
EBmac970*	7 (5H)	54	(ac)8	112	Ramsay <i>et al.</i> , 2000
Bmag223*	7 (5H)	69	(ag)16	197	Ramsay <i>et al.</i> , 2000
HvHEMH1	7 (5H)		(t)8	194	Pillen <i>et al.</i> , 2000
HvLEU*	7 (5H)	150	(att)4	166	Becker y Heun, 1995
HvDHN7	7 (5H)		(ttg)5	177	Becker y Heun, 1995
Bmag222*	7 (5H)	162	(ac)9(ag)17	179	Ramsay <i>et al.</i> , 2000
Bmac316*	6 (6H)	6	(ac)19	135	Ramsay <i>et al.</i> , 2000
Bmag500	6 (6H)	38	(ag)6cg(ag)29 (agagg)3(ag)6	150	Ramsay <i>et al.</i> , 2000
Bmag173*	6 (6H)	79	(ct)29	150	Ramsay <i>et al.</i> , 2000
Bmac018*	6 (6H)	103	(ac)11	138	Russell <i>et al.</i> , 1997
Bmag009*	6 (6H)	103	(ag)13	172	Ramsay <i>et al.</i> , 2000
EBmac806*	6 (6H)	119	(ca)4ga(ca)8,(ca)5	168	Ramsay <i>et al.</i> , 2000
Bmac040*	6 (6H)	151	(ac)20	236	Russell <i>et al.</i> , 1997
Bmag206*	1 (7H)	19	(gt)5(ag)14	239	Ramsay <i>et al.</i> , 2000
HvM004	1 (7H)	27	(at)9	198	Saghai-Marooof <i>et al.</i> , 1994b
HvSS1	1 (7H)		(t)20	240	Pillen <i>et al.</i> , 2000
HvCMA*	1 (7H)	85	(at)9	141	Becker y Heun, 1995
Bmac064	1 (7H)	98	(ac)21	155	Russell <i>et al.</i> , 1997
Bmag369	1 (7H)	100	(ct)16	191	Ramsay <i>et al.</i> , 2000
Bmag120*	1 (7H)	118	(ag)15	230	Ramsay <i>et al.</i> , 2000
Bmac156*	1 (7H)	156	(ac)22(at)5	139	Ramsay <i>et al.</i> , 2000
Bmag135*	1 (7H)	165	(ag)10gg(ag)12	161	Ramsay <i>et al.</i> , 2000
HvGLB2	1 (7H)		(gt)5(g)12	210	Pillen <i>et al.</i> , 2000

Además de estos 64 microsatélites, se empleó el STS MWG699. Este marcador ligado al locus *vrs1*, que es el gen de determinación de tipo de espiga 2 ó 6 carreras, fue propuesto por Tanno *et al.* (1999, 2002) como un marcador de la domesticación de la cebada cultivada. Estos autores propusieron que las cebadas de 6 carreras se subdividen en dos grupos, según el haplotipo presentado en ese marcador. Según esos autores, el primer haplotipo (A) está ampliamente distribuido, mientras que el segundo (D) se encuentra exclusivamente en la región Mediterránea.

Tabla 3.5. Origen de los microsatélites.

Microsatélites	Origen	Descripción “EMBL gene”
HvGLUEND	Genbank	Intrón; ‘1,3; 1,4 β-glucanase’
HvALAAT	Genbank	Extremo 3’ no codificante; ‘Alanine aminotransferase’
HvHVA1	Genbank	Exón; ‘Gene for abscisic acid’ (estrés hídrico)
HvBKASI	Genbank	Intrón; ‘β-ketoacyl-hacer carrier protein synthasel
HvLTPPB	Genbank	Extremo 3’ no codificante; ‘Lipid transfer protein
Hv13GEIII	Genbank	‘1,3-Glucan endohydrolase III’
HvOLE	Genbank	Extremo 3’ no codificante; ‘Olesin’
HvM003	Genbank	Intrón; ‘Rubisco activase’
HvBTAI3	Genbank	‘α-amylase inhibitor’
HvMLOH1A	Genbank	Extremo 5’ no codificante; ‘Mlo-hi gene’
HvBAMY	Genbank	Intrón; ‘β-amylase’
HdAMYB	Genbank	Intrón; ‘β-amylase’
HvHEMH1	Genbank	Extremo 3’ no codificante; ‘Ferrochelatase’
HvLEU	Genbank	Intrón; ‘Thiol protease aleurin’
HvDHN7	Genbank	Extremo 3’ no codificante; ‘gene for dehydrin-7’
HvM004	Genbank	‘Starch synthase (waxy)’
HvSS1	Genbank	Intrón; ‘gene for sucrose synthase’
HvCMA	Genbank	Extremo 3’ no codificante; ‘α-amylase inhibitor’
HvGLB2	Genbank	1,3-1,4- ‘β-D glucan 4-glucanohydrolase’

3-2-1-3. Condiciones de amplificaciones y electroforesis

Las secuencias de los cebadores y las condiciones de las amplificaciones están extraídas de Pillen *et al.* (2000), Ramsay *et al.* (2000) y Macaulay *et al.* (2001).

Las amplificaciones de las PCR se llevaron a cabo en un volumen final de 15 µl: 1x PCR Buffer [75mM Tris-HCl (pH 9.0), 2 mM MgCl₂, 50 mM KCl, 20mM (NH₄)₂SO₄] (Biotools), 0.15 µM de cada cebador, 0.2 mM de dNTPs, 0.4 U de Tth DNA Polymerase (Biotools) y 50 ng de DNA genómico.

Las reacciones de PCR se realizaron con dos termocicladores: PCR Express (Hybaid) y GeneAmp 2700 (Applied Biosystems).

Antes de la electroforesis, se procedió a la desnaturalización de las muestras, después de añadirles un tampón de carga que contiene 95% formamida.

La detección del polimorfismo se realizó por electroforesis en geles de acrilamida desnaturalizantes al 5 % (35 x 43 cm), para permitir la separación de fragmentos que se diferencien de 1 a 2 pares de bases.

Las electroforesis se corrieron en una cubeta Modelo S 2001 (Life Technologies), a 65 W, en presencia de un tampón de electroforesis TBE 1X, durante un tiempo de 2 a

3 horas, según el tamaño del microsatélite analizado. El revelado de los geles se hizo por tinción en una solución de nitrato de plata (Bassam *et al.*, 1991).

La lectura de los geles obtenidos se efectuó por densitometría mediante el aparato Molecular Imager FX (BIO-RAD) y el programa *Diversity Database*. El tamaño de las bandas en pb de cada microsatélite se determinó haciendo referencia a la escala de peso molecular (*30-330 pb AFLP Ladder, Invitrogen*). El número de repeticiones del motivo correspondiente a cada banda fue deducido a partir de su tamaño.

Doce microsatélites (Bmag136, Bmag013, Bmag384, Bmag353, EBmac701, EBmac970, Bmac113, Bmag223, EBmac806, Bmag206, Bmag120 y Bmag135) fueron analizados con el secuenciador automático ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Para estos microsatélites, las amplificaciones se llevaron a cabo en un volumen final de 10 μ l: 1x PCR Buffer [10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl₂, 50 mM KCl), 0.1 μ M de cada cebador, 0.2 mM de dNTPs, 0.4 U de AmpliTaq Gold (Applied Biosystems) y 50 ng de DNA genómico.

Estos productos de PCR se corrieron junto con un marcador de peso molecular GeneScan 500 [TAMRA] siguiendo las instrucciones del manual del ABI 310 (Applied Biosystems). Los resultados fueron analizados con el programa GeneScan.

Para el STS MWG699 la amplificación por PCR y la restricción con el enzima *TaqI* fueron realizados según el procedimiento descrito por Komatsuda *et al.* (1998). Las condiciones de electroforesis de este marcador están descritas en Casas *et al.* (2005).

3-2-1-4. Análisis estadísticos

a. Diversidad genética

El nivel de polimorfismo de cada locus fue calculado mediante el índice de diversidad génica de Nei (1973):

$$H_T = 1 - \sum P_i^2$$

donde P_i es la frecuencia del alelo i en el locus considerado.

b. Matriz de distancias

Las distancias genéticas entre individuos fueron calculadas según el método *Manhattan* o *City block* (Cain y Harrison, 1958), con el programa NTYSYS pc v. 2.1. (Rohlf, 2000).

$$M_{ij} = \frac{1}{n} \sum_k |x_{ki} - x_{kj}|$$

donde x_{ki} y x_{kj} son el número de repeticiones de los alelos i y j del locus k , y n es el número de loci analizados.

A partir de la matriz de distancias calculada, se construyó un dendrograma (método UPGMA) con el programa MEGA version 3.0 (Kumar *et al.*, 2004). La misma matriz fue empleada para la representación de los 225 genotipos en coordenadas principales. Este último análisis fue realizado mediante el programa NTSYS pc v. 2.1. (Rohlf, 2000).

c. Análisis molecular de varianza

Se realizó un análisis molecular de varianza para comparar los grupos de germoplasma establecidos *a priori* (Tabla 3.1), o sea, grupos de genotipos españoles y de referencia, respectivamente, para 2 y 6 carreras. Este análisis se efectuó mediante el programa AMOVA (Arlequín software, Shneider *et al.*, 2000).

El análisis se realizó basándose sobre dos modelos teóricos que explican la evolución del polimorfismo de los microsatélites.

El modelo llamado *infinite alleles model* (IAM) considera que el cambio en el tamaño de los alelos puede ocurrir a saltos grandes o pequeños, con la misma probabilidad. El índice de fijación estadístico F_{ST} (Slatkin, 1995) calculado de acuerdo con este modelo, refleja diferencias que existen entre clases de alelos sin tener en cuenta la amplitud del tamaño real de la diferencia. El índice F_{ST} se define como:

$$F_{ST} = \frac{\sigma^2_a}{\sigma^2_T};$$

donde σ^2_a y σ^2_T se calculan a partir de la tabla del AMOVA (Tabla 3.6)

Tabla 3.6. Esquema del análisis molecular de varianza.

Fuente de variación	Grados de libertad*	Suma de cuadrados	Cuadrado medio esperado**
Entre grupos	P-1	Entre grupos	$n\sigma^2_a + \sigma^2_b$
Dentro de grupos	N-P	Dentro de grupos	σ^2_b
Total	N-1	Total	σ^2_T

* P: número total de grupos analizados

* N: número total de individuos

$$** n = \frac{N - \sum_P \frac{N_P^2}{N}}{P - 1}$$

El modelo conocido como *stepwise mutation model* o SMM (Ohta y Kimura, 1973), propone que los cambios en el tamaño de los microsatélites ocurren por ganancias o pérdidas de un número pequeño de repeticiones del nucleótido (1 a 2 unidades). El índice de fijación estadístico R_{ST} (Slatkin, 1995), calculado según los postulados de este modelo, refleja las diferencias genéticas que existen entre poblaciones de un modo proporcional a la diferencia en el número de repeticiones de sus alelos. Este parámetro se define como:

$$R_{ST} = \frac{\sigma^2 a}{\sigma^2 T} = \frac{\bar{S} - S_W}{\bar{S}}$$

donde

$$\bar{S} = \frac{2n-1}{2nd_s-1} S_W + \frac{2n(d_s-1)}{2nd_s-1} S_B$$

siendo n el número de individuos de cada d_s poblaciones.

S_W es la media de la suma de cuadrados de diferencias en el tamaño de los alelos en cada grupo, se calcula según la formula siguiente:

$$S_W = \frac{1}{d_s} \sum_{j=1}^{d_s} \frac{2}{2n(2n-1)} \sum_{i < i'} (a_{ij} - a_{i'j})$$

siendo a_{ij} el tamaño del alelo i ($i = 1, \dots, 2n$) en el grupo j ($j = 1, \dots, d_s$).

S_B simboliza la diferencia cuadrada media entre todos los pares de alelos. Se calcula según la formula siguiente:

$$S_B = \frac{2}{(2n)^2 d_s (d_s - 1)} \sum_{j < j'} \sum_{i < i'} (a_{ij} - a_{i'j'})$$

Empleando estos dos métodos, se calcularon índices de fijación estadísticos (F_{ST} y R_{ST}) locus por locus, para los grupos de germoplasma considerados (Tabla 3.1). El nivel de significación elegido fue de 0.0008, lo que corresponde a un nivel de 0.05 para 64 marcadores independientes, aplicando una corrección de Bonferroni.

d. Análisis de la estructura genética

La posible estructura en poblaciones de la muestra de genotipos se investigó mediante un análisis bayesiano de Monte Carlo con cadenas Markov (MCMC), implementado en el programa STRUCTURE (Falush *et al.*, 2003), empleando la opción de “modelo de ligamiento”. Este método permite clasificar los genotipos en poblaciones, mediante el procedimiento de minimizar el desequilibrio de ligamiento dentro de las mismas.

Se hicieron 25 ejecuciones del programa, proporcionando cada vez un número distinto de poblaciones (K), de 1 a 10. Para cada ejecución del programa tanto el número de repeticiones de entrenamiento como el de repeticiones MCMC fue de 10.000. El número de poblaciones más probable se decide según el valor del ‘*log-likelihood*’. Falush *et al.* (2003), explican que el número óptimo de poblaciones que componen una muestra corresponde al valor del ‘*log-likelihood*’ más alto, que se distingue como un punto de inflexión de la curva producida por los distintos valores del ‘*log-likelihood*’ en función de los valores de K.

Para verificar que el número final de poblaciones elegidas fue el más adecuado, se hizo un análisis molecular de varianza (AMOVA), para cada valor de K, comparando las poblaciones 2 a 2, para cuantificar la contribución a la varianza genética entre las poblaciones y dentro de las mismas.

e. Evaluación del desequilibrio de ligamiento

La evaluación del desequilibrio de ligamiento entre pares de marcadores se estimó en la totalidad del material examinado y en las poblaciones derivadas del análisis de STRUCTURE. Este análisis se llevó a cabo mediante el programa TASSEL (Zhang *et al.*, 2006). Los alelos raros presentes en frecuencias inferiores al 5%, en cada grupo analizado, fueron considerados como datos perdidos, para reducir el efecto desproporcionado de los alelos raros sobre el desequilibrio de ligamiento.

Este programa evalúa el desequilibrio de ligamiento entre pares de loci empleando los dos estadísticos más comunes: r^2 y D' . Considerando un par de loci con los alelos A y a en un locus, y B y b en el segundo locus, con las frecuencias π_A , π_a , π_B y π_b respectivamente. Las frecuencias de los haplotipos resultantes son π_{AB} , π_{Ab} , π_{aB} y π_{ab} . El componente básico de todos los estadísticos para medir desequilibrio de ligamiento es la diferencia entre las frecuencias haplotípicas observadas y esperadas.

$$D_{ab} = \pi_{AB} - \pi_A \pi_B$$

y

$$r^2 = \frac{(D_{ab})^2}{\pi_A \pi_a \pi_B \pi_b}$$

Alternativamente, D' se calcula como sigue:

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A \pi_b, \pi_a \pi_B)} \text{ por } D_{ab} < 0;$$

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A \pi_B, \pi_a \pi_b)} \text{ por } D_{ab} > 0.$$

En este estudio se eligió el estadístico r^2 (cuadrado del coeficiente de correlación) como medida del desequilibrio de ligamiento, por ser el más correcto. En este sentido Flint-Garcia *et al.* (2003) comentan que la D' considera solamente la historia de recombinación mientras que la r^2 tiene en cuenta al mismo tiempo la historia de mutación. Además, aunque ni la r^2 como la D' son potentes con muestras pequeñas o con frecuencias alélicas bajas, esta última se ve relativamente más afectada por el tamaño de muestra (muestras demasiado pequeñas), resultando en un comportamiento altamente errático cuando la comparación ocurre entre loci que tienen alelos con frecuencias bajas, como es nuestro caso.

El nivel de significación del desequilibrio de ligamiento entre loci se examina con el test exacto de Fisher, para loci bialélicos. Para pares de marcadores que tienen más de dos alelos en uno o ambos loci, la significación del desequilibrio del ligamiento se probó usando 100.000 permutaciones. Se eligió este número a la vista de la estabilidad de los resultados después de ejecutar el programa con 1000, 10.000 50.000 y 100.000 permutaciones.

En cada uno de los grupos genotípicos analizados, se evaluó además el nivel de desequilibrio de ligamiento entre loci clasificados según su distancia de recombinación. Así, se calculó el desequilibrio de ligamiento entre loci posiblemente ligados (se estimó como tales los loci a menos de 10 cM según la información de mapas disponible); loci que están en el mismo cromosoma pero a una distancia mayor que 10 cM y, por último, entre loci presentes en diferentes cromosomas.

3-3. Resultados

3-3-1. Polimorfismo

Los microsatélites empleados en este estudio detectaron una única banda por genotipo, con unas pocas excepciones. De los genotipos autóctonos de la colección nuclear, se encontraron bandas dobles en seis SSRs, en un sólo genotipo en cada caso. Esto confirma la homocigosidad de las entradas y la eficacia del procedimiento de autofecundación empleado para su obtención. Para las variedades comerciales, se registró la presencia de doble banda en 13 casos en total, para 6 genotipos. En todos estos casos, se repitió la PCR para verificar los resultados, y se anotó tan sólo la banda más intensa en la base de datos. Por otra parte se detectaron 3 duplicados en la colección nuclear que fueron reemplazados por otras 3 entradas del banco de germoplasma.

Se detectaron un total de 669 alelos en los 225 genotipos analizados, con un tamaño variable de 71 a 287 pb. El número de alelos por locus varió de 2 (HvHVA1, WMC1E8, HvBTAI3, HvLEU y HvCMA) a 38 (Bmac399), con una media de 10,45 alelos por locus (Tabla 3.7). Se registró también la presencia de alelos nulos en cinco microsatélites (HvLTPPB, HvDHN7, EBmac806, Bmag135, y HvGLB2).

En general, los SSRs derivados de genes fueron menos polimórficos que los provenientes de librerías genómicas (6,53 y 12,09 alelos por locus de promedio, respectivamente), con la excepción entre los primeros de HvM003 y HvM004, que detectaron 25 y 23 alelos respectivamente.

3-3-2. Comparación de los grupos establecidos a priori

3-3-2-1. Polimorfismo e índice de diversidad

De los 699 alelos detectados, 228 fueron detectados únicamente en uno sólo de los grupos de germoplasma establecidos a priori (seis carreras españoles, seis carreras de referencia, dos carreras españoles y dos carreras de referencia). El 80,7 % de estos alelos específicos aparecieron en el grupo español de 6 carreras. Muchos de ellos son raros, presentes en frecuencias muy bajas, pero 9 de ellos se encontraron en más del 10% de los individuos de dicho grupo. En el grupo español de 2 carreras, se encontraron sólo 11 alelos específicos, mientras que en los grupos de referencia de 2 y 6 carreras, se localizaron respectivamente 12 y 21 alelos específicos, todos ellos raros (Tabla 3.7).

Tabla 3.7. Número total de alelos detectados, rango de tamaño de los alelos de cada locus; número de alelos comunes y alelos específicos de los microsatélites, en cada uno de los 4 grupos de germoplasma establecidos *a priori*. Los SSR están listados en orden dentro de cada cromosoma, del 1H al 7H.

Locus	Número de alelos					Tamaño (pb)	Número de alelos específicos			
	Total	6 C		2 C			6 C Esp	6 C Ref	2 C Esp	2 C Ref
		Esp	Ref	Esp	Ref					
Bmac399	38	36	12	9	6	71-189	21	0	1	0
HvGLUEND	5	5	4	2	2	259-264	1	0	0	0
Bmac032	28	25	16	3	2	209-255	12	2	0	0
HvM020	6	6	4	3	3	133-167	2	0	0	0
HvALAAT	7	7	6	2	2	197-205	1	0	0	0
Bmag211	10	8	6	3	4	166-192	4	2	0	0
Bmag382	5	3	4	2	3	93-111	0	1	0	0
HvHVA1	2	1	2	2	2	133-136	0	0	0	0
WMC1E8	2	2	2	2	2	191-238	0	0	0	0
Bmac134	13	12	10	3	4	128-172	2	1	0	0
HvM036	9	8	7	3	4	108-140	2	1	0	0
Bmag378	5	3	2	1	4	133-149	0	0	0	2
Bmac132	5	5	5	2	3	183-193	0	0	0	0
HvBKASI	4	3	3	3	3	181-200	0	0	0	0
Bmac093	5	5	4	1	4	153-161	0	0	0	0
Bmag125	11	10	7	3	4	118-146	3	0	0	1
HvM054	7	6	3	2	2	145-163	3	1	0	0
HvLTPPB	4	3	3	1	3	217-221	0	0	0	0
Bmag136	5	5	4	2	3	195-203	0	0	0	0
Bmac067	14	13	10	5	6	149-223	4	0	0	0
Bmag006	19	18	9	4	3	162-206	9	1	0	0
Bmac209	12	9	8	7	3	178-202	1	1	1	0
Bmag225	15	14	9	6	7	132-172	3	0	0	0
Hv13GEIII	3	3	3	3	2	169-175	0	0	0	0
Bmag013	20	18	7	7	6	129-185	7	0	1	1
HvM062	5	5	4	3	3	231-261	1	0	0	0
Bmac029	8	8	2	4	2	149-181	3	0	0	0
HvM040	7	4	4	5	4	142-162	0	1	1	0
HvOLE	4	4	4	2	2	190-210	0	0	0	0
Bmag353	18	17	9	8	4	90-138	6	1	0	0
HvM003	25	21	11	8	6	160-218	10	1	1	1
Bmag384	7	6	4	4	3	106-118	1	0	1	0
HvBTAI3	2	2	1	1	2	228-234	0	0	0	0
EBmac701	13	12	6	5	6	115-159	4	0	0	1
HvMLOH1A	3	2	2	2	2	175-183	0	0	1	0
HvM067	5	4	4	3	3	106-118	0	0	0	1
HvBAMY	6	4	5	4	4	127-134	0	0	1	0
HdAMYB	13	12	8	7	7	186-221	2	0	0	0

Tabla 3.7. (Continuación.).

Locus	Número de alelos					Tamaño (pb)	Número de alelos específicos			
	Total	6 C		2 C			6 C		2 C	
		Esp	Ref	Esp	Ref		Esp	Ref	Esp	Ref
Bmag337	12	11	6	4	5	125-153	4	0	1	0
Bmac113	15	13	10	7	8	167-207	4	1	0	0
Bmac096	11	9	9	4	7	161-215	2	0	0	0
EBmac970	4	2	3	3	3	188-194	0	1	0	0
HvHEMH1	3	2	3	2	3	194-200	0	0	0	0
Bmag223	22	21	14	6	11	121-179	5	0	0	1
HvLEU	2	2	2	2	2	167-171	0	0	0	0
HvDHN7	3	2	3	3	3	177-191	0	0	0	0
Bmag222	6	5	5	2	3	149-181	1	1	0	0
Bmac316	10	8	6	4	3	131-191	4	1	0	0
Bmag500	23	19	8	8	7	84-190	8	1	0	1
Bmag173	15	12	7	6	8	122-158	1	0	1	1
Bmac018	5	5	3	3	2	138-146	2	0	0	0
Bmag009	5	5	4	3	4	172-180	1	0	0	0
EBmac806	7	6	7	3	6	160-174	0	0	0	0
Bmac040	24	21	9	7	10	174-252	5	1	0	0
Bmag206	22	21	9	5	7	233-287	9	0	0	1
HvM004	23	21	10	7	5	194-280	12	0	0	0
HvSS1	5	5	4	2	3	238-250	1	0	0	0
HvCMA	2	1	2	2	2	133-141	0	0	0	0
Bmac064	7	6	5	4	4	143-159	0	1	0	0
Bmag369	7	6	4	4	4	175-191	2	0	0	0
Bmag120	18	17	10	4	4	230-270	6	1	0	0
Bmac156	30	26	14	8	9	107-205	14	0	1	0
Bmag135	10	8	5	4	8	138-162	1	0	0	1
HvGLB2	8	8	7	5	6	210-220	0	0	0	0
Total	669	591	383	249	272		184	21	11	12

Aunque no se puede comparar el polimorfismo alélico entre los grupos, debido al desequilibrio en el tamaño de muestra que existe entre ellos, se percibe una cierta tasa de especificidad superior en las cebadas de la colección nuclear, principalmente en el grupo de 6 carreras.

El índice de diversidad total (H_T) calculado para los 225 genotipos mostró valores que variaron entre 0,08 (HvHVA1) y 0,93 (HvM003), con una media de 0,66 (Tabla 3.8).

Tabla 3.8. Índices de diversidad: total (H_T), de los genotipos de 2 carreras (H_{T2c}) y 6 carreras (H_{T6c}) y de los 4 grupos de germoplasma establecidos *a priori* (2c Españolas, 2c Referencia, 6c Españolas, y 6c Referencia).

Cr	Locus	H_T	H_{T2c}	H_{T2c}		H_{T6c}	H_{T6c}	
				Españolas	Referencia		Españolas	Referencia
1H	Bmac399	0,77	0,73	0,64	0,73	0,77	0,77	0,71
1H	HvGLUEND	0,65	0,29	0,30	0,29	0,66	0,64	0,66
1H	Bmac032	0,91	0,41	0,43	0,40	0,92	0,90	0,88
1H	HvM020	0,68	0,51	0,43	0,35	0,63	0,57	0,69
1H	HvALAAT	0,71	0,18	0,17	0,19	0,77	0,77	0,72
1H	Bmag211	0,73	0,68	0,56	0,69	0,71	0,72	0,65
1H	Bmag382	0,46	0,16	0,17	0,15	0,32	0,30	0,40
1H	HvHVA1	0,08	0,35	0,30	0,37	0,01	0,00	0,06
1H	WMC1E8	0,45	0,35	0,30	0,37	0,37	0,31	0,50
2H	Bmac134	0,83	0,45	0,31	0,48	0,80	0,75	0,81
2H	HvM036	0,76	0,70	0,56	0,65	0,72	0,65	0,69
2H	Bmag378	0,42	0,38	0,00	0,49	0,42	0,46	0,06
2H	Bmac132	0,49	0,43	0,17	0,49	0,37	0,39	0,27
2H	HvBKASI	0,48	0,48	0,53	0,45	0,47	0,50	0,28
2H	Bmac093	0,46	0,50	0,00	0,62	0,45	0,47	0,32
2H	Bmag125	0,80	0,69	0,66	0,58	0,81	0,78	0,77
2H	HvM054	0,61	0,45	0,30	0,45	0,54	0,57	0,32
3H	HvLTPPB	0,43	0,14	0,00	0,19	0,47	0,50	0,22
3H	Bmag136	0,68	0,47	0,46	0,47	0,68	0,62	0,37
3H	Bmac067	0,87	0,74	0,74	0,65	0,87	0,86	0,73
3H	Bmag006	0,90	0,65	0,61	0,51	0,90	0,89	0,72
3H	Bmac209	0,81	0,70	0,84	0,55	0,79	0,77	0,82
3H	Bmag225	0,82	0,78	0,66	0,69	0,77	0,74	0,83
3H	Hv13GEIII	0,44	0,37	0,60	0,07	0,44	0,46	0,35
3H	Bmag013	0,86	0,80	0,83	0,74	0,86	0,88	0,63
3H	HvM062	0,60	0,50	0,56	0,42	0,62	0,60	0,59
3H	Bmac029	0,47	0,27	0,55	0,13	0,51	0,55	0,22
4H	HvM040	0,58	0,67	0,71	0,56	0,50	0,41	0,53
4H	HvOLE	0,35	0,20	0,17	0,20	0,37	0,40	0,18
4H	Bmag353	0,91	0,77	0,84	0,71	0,90	0,89	0,86
4H	HvM003	0,93	0,82	0,86	0,70	0,93	0,92	0,87
4H	Bmag384	0,75	0,42	0,64	0,19	0,76	0,76	0,65
4H	HvBTAI3	0,17	0,18	0,00	0,24	0,17	0,20	0,00
4H	EBmac701	0,82	0,76	0,78	0,72	0,82	0,80	0,62
4H	HvMLOH1A	0,19	0,49	0,30	0,49	0,10	0,03	0,37
4H	HvM067	0,61	0,50	0,51	0,49	0,54	0,53	0,48
4H	HvBAMY	0,72	0,71	0,61	0,68	0,67	0,59	0,52
4H	HdAMYB	0,88	0,86	0,84	0,81	0,85	0,81	0,74

Tabla 3.8. (Continuación.).

Cr	Locus	H_T	$H_{T\ 2c}$	$H_{T\ 2c}$	$H_{T\ 2c}$	$H_{T\ 6c}$	$H_{T\ 6c}$	$H_{T\ 6c}$
				Españolas	Referencia		Españolas	Referencia
5H	Bmag337	0,77	0,73	0,64	0,73	0,77	0,77	0,71
5H	Bmac113	0,76	0,81	0,79	0,72	0,74	0,74	0,71
5H	Bmac096	0,83	0,80	0,71	0,76	0,82	0,78	0,81
5H	EBmac970	0,52	0,55	0,34	0,51	0,51	0,49	0,54
5H	HvHEMH1	0,42	0,51	0,17	0,53	0,39	0,31	0,60
5H	Bmag223	0,92	0,87	0,78	0,86	0,91	0,90	0,89
5H	HvLEU	0,47	0,35	0,30	0,37	0,49	0,49	0,48
5H	HvDHN7	0,46	0,57	0,64	0,49	0,40	0,43	0,12
5H	Bmag222	0,60	0,52	0,46	0,44	0,58	0,58	0,59
6H	Bmac316	0,72	0,65	0,68	0,61	0,71	0,65	0,62
6H	Bmag500	0,89	0,85	0,84	0,81	0,88	0,88	0,82
6H	Bmag173	0,85	0,83	0,78	0,78	0,81	0,79	0,66
6H	Bmac018	0,68	0,53	0,64	0,40	0,69	0,70	0,62
6H	Bmag009	0,57	0,67	0,63	0,60	0,47	0,35	0,66
6H	EBmac806	0,66	0,68	0,58	0,70	0,65	0,58	0,77
6H	Bmac040	0,92	0,90	0,83	0,88	0,91	0,91	0,79
7H	Bmag206	0,86	0,80	0,78	0,68	0,87	0,88	0,76
7H	HvM004	0,90	0,80	0,83	0,73	0,90	0,90	0,82
7H	HvSS1	0,63	0,27	0,17	0,30	0,61	0,52	0,64
7H	HvCMA	0,29	0,29	0,46	0,19	0,06	0,00	0,30
7H	Bmac064	0,59	0,59	0,58	0,40	0,54	0,49	0,57
7H	Bmag369	0,68	0,44	0,61	0,35	0,68	0,68	0,57
7H	Bmag120	0,87	0,79	0,78	0,71	0,85	0,84	0,84
7H	Bmac156	0,92	0,83	0,86	0,76	0,91	0,90	0,82
7H	Bmag135	0,62	0,81	0,73	0,83	0,54	0,56	0,45
7H	HvGLB2	0,84	0,82	0,71	0,81	0,82	0,84	0,70
Promedio		0,66	0,57	0,54	0,53	0,64	0,62	0,58

Comparando los grupos españoles con los de referencia, se observa que existen más diferencias dentro del conjunto de cebadas de 6 carreras que del de 2 carreras (Tabla 3.8). De hecho, para las cebadas de 6 carreras, los índices de diversidad calculados para cada loci son en general mayores para las cebadas españolas (0,62) que para las de referencia (0,58). En el conjunto de cebadas de 2 carreras, no se notó esta diferencia entre las cebadas españolas y las de referencia (promedios de 0,54 y 0,53, respectivamente). Por otra parte se observó mayor diversidad en general en las cebadas de 6 carreras (promedio general 0,64) que en las de 2 carreras (promedio 0,57).

3-3-2-2. Análisis molecular de varianza

Se realizó un análisis molecular de varianza para comparar los grupos establecidos a priori, para las cebadas de 2 y 6 carreras. Este análisis fue efectuado según los dos modelos teóricos del proceso de evolución del polimorfismo de los microsatélites 'IAM' y 'SMM' definidos en el apartado 3-2-1-4 (c). Los resultados de este análisis revelaron diferencias altamente significativas entre los grupos comparados (2 carreras españolas vs 2 carreras de referencia, 6 carreras españolas vs 6 carreras de referencia). La divergencia genética, expresada por los índices de fijación estadísticos F_{ST} (IAM) y R_{ST} (SMM), entre los grupos españoles y grupos de referencia fueron respectivamente de 10,4% y 21,4% para las cebadas de 6 carreras, mientras que se registraron los valores de 11,9% y 7,5% en la cebadas de 2 carreras (Tabla 3.9). Se observa que los índices de fijación estadísticos F_{ST} son parecidos para las cebadas de 2 y 6 carreras mientras que el R_{ST} fue bastante mayor para las cebadas de 6 carreras.

El análisis molecular de varianza locus por locus, realizado con los dos métodos descritos, permitió identificar los microsatélites que discriminan mejor el germoplasma español del material de referencia. Los loci que presentan valores F_{ST} y R_{ST} altamente significativos a la vez, para las cebadas de 6 carreras fueron: Bmac032, WMC1E8, Bmac134, Bmag125, Bmag136, HvM040, Bmag353, EBmac701, HvBAMY, HdAMYB, Bmag223, Bmac316, Bmag173, Bmag009, HvSS1 y Bmac156. Dentro del grupo de cebadas de 2 carreras, los microsatélites que mejor separan las entradas españolas de las variedades de referencia fueron Bmag225, Hv13GEIII y Bmac064 (Tabla 3.10).

Tabla 3.9. Resultados del análisis molecular de varianza (AMOVA), comparando los grupos de genotipos establecidos *a priori*, basado sobre los modelos de evolución de los microsatélites IAM y SMM.

Modelo	Grupos de germoplasma	Fuente de variación	Grados de libertad	Suma de cuadrados	Componentes de varianza		Porcentaje de variación
Infinite alleles model (IAM)	6 carreras españolas <i>versus</i> 6 carreras referencia	entre grupos	1	137,4	2,19	(σ^2_a)	10,43
		dentro de grupos	183	3.438,3	18,79	(σ^2_b)	89,57
		Total	184	3.575,7	20,98	(σ^2_T)	
		Índice de fijación (%)	F_{ST}	10,43	Probabilidad		0,0000
Infinite alleles model (IAM)	2 carreras españolas <i>versus</i> 2 carreras referencia	entre grupos	1	53,3	2,28	(σ^2_a)	11,93
		dentro de grupos	38	641,1	16,87	(σ^2_b)	88,07
		Total	39	694,4	19,16	(σ^2_T)	
		Índice de fijación (%)	F_{ST}	11,93	Probabilidad		0,0000
Stepwise mutation model (SMM)	6 carreras españolas <i>versus</i> 6 carreras referencia	entre grupos	1	18.886,7	326,23	(σ^2_a)	21,43
		dentro de grupos	183	218.903,5	1196,19	(σ^2_b)	78,57
		Total	184	237.790,2	1522,42	(σ^2_T)	
		Índice de fijación (%)	R_{ST}	21,43	Probabilidad		0,0000
Stepwise mutation model (SMM)	2 carreras españolas <i>versus</i> 2 carreras referencia	entre grupos	1	2.065,7	72,92	(σ^2_a)	7,47
		dentro de grupos	38	34.301,8	902,68	(σ^2_b)	92,53
		Total	39	36.367,5	975,59	(σ^2_T)	
		Índice de fijación (%)	R_{ST}	7,47	Probabilidad		0,0088

Tabla 3.10. AMOVA (análisis molecular de varianza) locus por locus, en orden cromosómico (1H a 7H). Índices F_{ST} y R_{ST} que comparan la divergencia genética entre los grupos de genotipos españoles y de referencia.

Locus	6c españolas vs 6c referencia		2c españolas vs 2c referencia	
	F_{ST}	R_{ST}	F_{ST}	R_{ST}
Bmac399	4,4	-1,6	4,1	-6,0
HvGLUEND	8,2	13,6	-6,7	-6,7
Bmac032	7,0 *	30,3 *	-4,5	-2,1
HvM020	18,3 *	4,7	46,3 *	55,0
HvALAAT	-0,4	-0,4	-1,2	-5,4
Bmag211	1,8	3,7	1,9	-6,6
Bmag382	-0,9	2,1	-0,9	-6,9
HvHVA1	6,4	6,4	-5,6	-5,6
WMC1E8	22,2 *	22,2 *	-5,6	-5,6
Bmac134	14,7 *	32,9 *	1,8	-4,2
HvM036	21,9 *	6,2	16,5	-6,5
Bmag378	20,7 *	10,4	10,1	9,7
Bmac132	1,0	-1,8	10,7	-5,8
HvBKASI	5,5	11,8	-2,0	-1,3
Bmac093	2,6	12,8	18,1	-0,1
Bmag125	10,1 *	38,1 *	21,6	30,0
HvM054	6,9	1,6	12,6	10,2
HvLTPPB	10,9	16,5	-0,5	-6,1
Bmag136	36,9 *	36,1 *	-6,0	-5,1
Bmac067	9,0	3,1	13,6	8,6
Bmag006	10,7 *	15,2	28,4	6,9
Bmac209	1,2	1,6	14,5	-5,0
Bmag225	2,6	9,0	22,3 *	59,0 *
Hv13GEIII	1,7	2,4	62,6 *	63,3 *
Bmag013	7,8 *	1,7	2,9	20,0
HvM062	6,9	4,2	14,7	17,3
Bmac029	7,8	-1,0	17,6	2,4
HvM040	32,6 *	36,9 *	15,1	18,7
HvOLE	3,8	0,6	-1,0	-2,5
Bmag353	6,9 *	65,1 *	-0,2	16,8
HvM003	3,0 *	0,6	13,4	26,4
Bmag384	8,5	18,5	40,9 *	-2,3
HvBTAI3	5,4	5,4	4,5	4,5
EBmac701	16,3 *	31,5 *	2,2	-0,1
HvMLOH1A	36,7 *	36,7	18,0	23,0
HvM067	11,7	-0,1	-5,1	-6,7
HvBAMY	34,9 *	22,2 *	9,2	14,1
HdAMYB	16,6 *	63,7 *	5,1	1,6

Tabla 3.10. (Continuación).

Locus	6c españolas vs 6c referencia		2c españolas vs 2c referencia	
	F _{ST}	R _{ST}	F _{ST}	R _{ST}
Bmag337	2,2	16,6	0,0	6,7
Bmac113	-0,4	-1,4	13,1	24,4
Bmac096	11,6 *	0,6	8,7	3,1
EBmac970	2,7	8,0	26,3	30,4
Bmag223	3,9 *	26,0 *	3,4	9,2
HvHEMH1	23,1 *	8,4	27,2	22,7
HvLEU	-1,7	-1,7	-5,6	-5,6
HvDHN7	15,0	15,6	9,6	28,8
Bmag222	0,6	13,1	26,6	-6,5
Bmac316	24,8 *	24,8 *	-0,2	9,8
Bmag500	1,5	0,0	4,4	-5,8
Bmag173	17,0 *	30,5 *	6,9	7,3
Bmac018	1,9	2,3	20,2	-5,3
Bmag009	31,3 *	55,6 *	16,4	35,7
EBmac806	14,2	20,7 *	-2,6	-6,4
Bmac040	5,3 *	-0,3	4,6	-1,6
Bmag206	2,7	2,0	18,9 *	21,0
HvM004	4,5 *	0,7	5,3	37,9 *
HvSS1	27,2 *	32,6 *	-4,4	-4,6
HvCMA	37,4	37,4 *	15,5	15,5
Bmac064	16,5 *	2,7	42,4 *	39,9 *
Bmag369	7,0	6,8	3,9	-6,6
Bmag120	3,8	-1,0	12,1	28,5
Bmac156	7,4 *	21,6 *	7,1	-6,6
Bmag135	0,9	-1,8	-2,1	3,7
HvGLB2	2,4	2,9	3,9	30,0

* F_{ST} y R_{ST} significativos al nivel de significación de 0.0008.

3-3-2-3. Análisis multivariantes

a. Análisis de cluster

A partir de una matriz de distancias calculada con el método de Manhattan para el conjunto de los 225 genotipos analizados, se realizó un análisis de cluster según el procedimiento UPGMA, que se representó con un dendrograma (Figura 3.1). Eligiendo subjetivamente el punto de corte a la distancia de 1,6, se pueden distinguir 8 grupos:

- Grupo I: se encuentran 71 genotipos españoles, entre ellos las variedades Albacete, Pané y Candela. La mayoría de las entradas de este grupo son del Sur de España y de la región del Valle del Ebro. Se observan también algunos genotipos de la Meseta así como las 3 entradas de las Islas Canarias.
- Grupo II: se agrupan 32 variedades locales de la colección nuclear, de los cuales el 84% provienen de la Meseta. El resto de genotipos son del valle del Ebro, entre ellos la variedad Almunia.
- Grupo III: agrupa solamente 4 entradas procedentes del Sur de España, que parecen distantes del resto de los genotipos de esta región.
- Grupo IV: en este grupo se observan 21 genotipos originarios de distintas regiones del país. La variedad francesa Ager se agrupa con estas entradas.
- Grupo V: reúne 10 entradas de la colección nuclear y la variedad griega Athenais.
- Grupo VI: se encuentran 3 variedades de 6 carreras, dos son materiales del ICARDA y la otra es la variedad española Orria, seleccionada a partir de material CIMMYT. Se agrupa con estas entradas la variedad francesa de 2 carreras Hispanic.
- Grupo VII: en este grupo se juntan las variedades de 2 carreras de primavera con 9 entradas de 2 carreras de la colección nuclear.
- Grupo VIII: reúne 52 genotipos, que se subdividen en dos sub-grupos. Por una parte, se encuentran las variedades de 6 carreras de primavera junto con algunas entradas de 6 carreras de la colección nuclear, y por otra parte se agrupan las variedades de invierno de 2 y 6 carreras, con algunas otras entradas de 6 carreras de la colección nuclear.

Se observa un agrupamiento de entradas del mismo origen geográfico (Figura 3.1). Esto se aprecia particularmente en el grupo II, que reúne muchas entradas de Castilla-La Mancha (CLM) y Castilla y León (CLE). También se percibe este fenómeno en el grupo I, donde se observa un agrupamiento de muchas entradas de Andalucía, como las 7 entradas presentes al final del grupo I (AND034, AND128, AND089, AND092, AND018, AND075 y AND017). También se observan agrupamientos de entradas de orígenes geográficos distintos, como es el caso de algunas entradas de los grupos I, IV y V (Figura 3.1).

b. Análisis de coordenadas principales

La matriz de distancias genéticas, calculada anteriormente, se examinó también mediante un análisis de coordenadas principales. Este análisis permite ilustrar mejor los ejes principales de variabilidad entre los genotipos, reflejando su ordenación en el espacio (Figura 3.2). Las dos primeras coordenadas resultantes de éste análisis explicaron un porcentaje pequeño de la totalidad de la diversidad genética (18,9%). El eje 1 explicó el 12,7% de la variabilidad y separa principalmente el grupo de cebadas de 6 carreras españolas de los otros grupos establecidos *a priori*. El eje 2, que explicó el 5,6% de la variabilidad, separa la mayoría de las cebadas de 2 carreras de las de 6 carreras de referencia, y divide las españolas de 6 carreras en dos grandes grupos. Las entradas de 2 carreras españolas están más próximas a las variedades de 2 carreras de primavera del conjunto de referencia, mientras que algunas variedades de 2 carreras de invierno, tales como Alpha, Mogador y Tipper (las dos situadas en torno a Alpha en la Figura 3.2), se encuentran más cercanas a las variedades de 6 carreras de invierno (como Barbarrosa o Plaisant). Esto no es nada sorprendente por el hecho que estas variedades tienen cultivares comunes en sus pedigrís.

Por otra parte, se observa que la variedad griega Athenais se junta con las españolas, manifestando así su relación relativamente próxima con el material mediterráneo, en comparación con el de otros países europeos.

Veintinueve de las variedades analizadas en este estudio pertenecen a la colección nuclear europea de cebada (<http://www.barley.ipk-gatersleben.de/ebdb.php3>), incluyendo las españolas Albacete y Candela. Estas variedades están indicadas por su nombre en la Figura 3.2, para servir de puntos de referencia para la interpretación de los resultados.

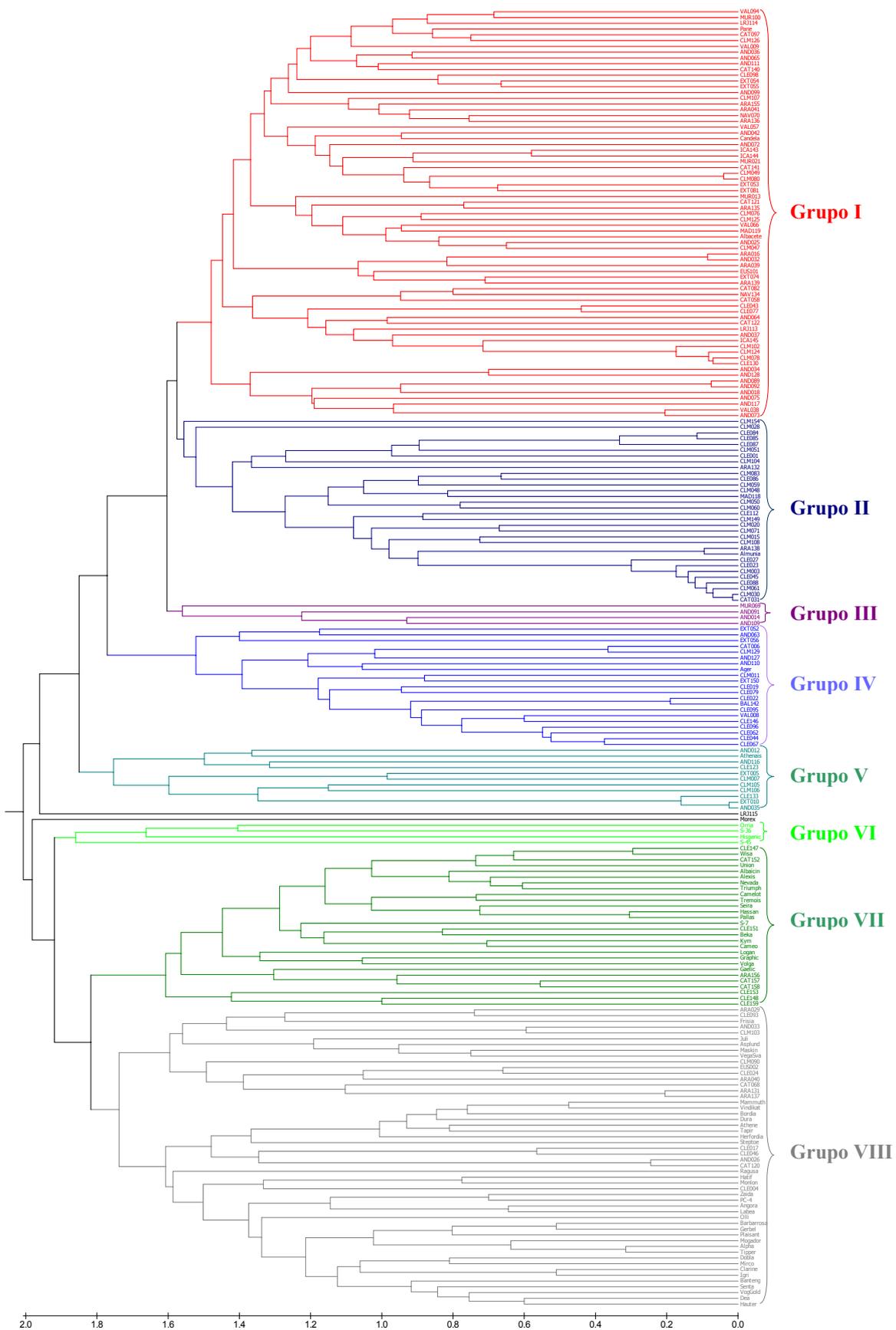


Figura 3.1. Dendrograma elaborado según el procedimiento UPGMA de los 225 genotipos analizados (españoles y referencia), empleando una matriz de distancias calculada según el método de Manhattan (Cain y Harrison, 1958), con 64 microsatélites. Los grupos (ver texto) se diferencian por colores distintos.

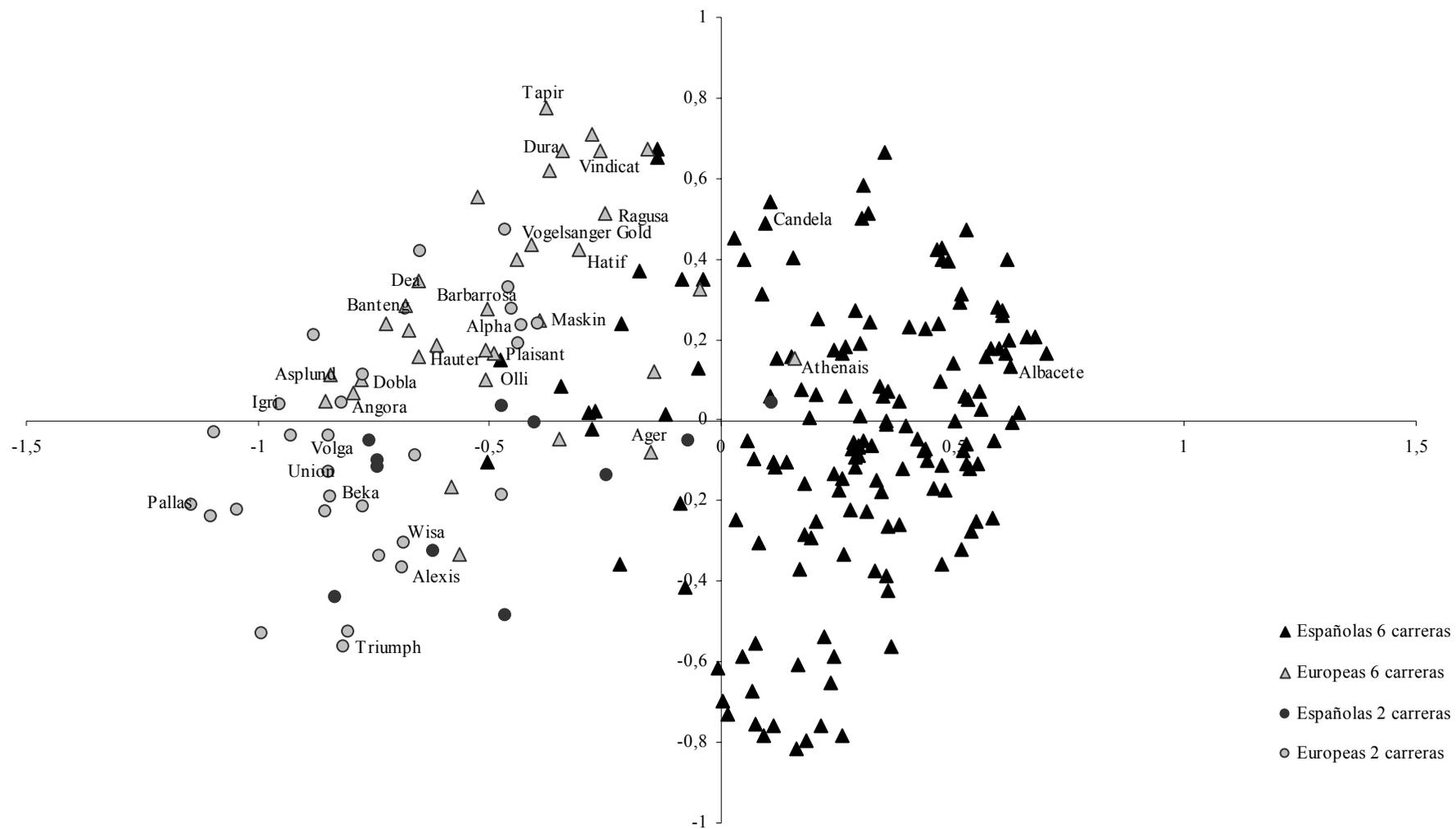


Figura 3.2. Análisis de coordenadas principales realizado a partir de la matriz de distancias de Manhattan (Cain y Harrison, 1958), de los 225 genotipos analizados con 64 marcadores. Los genotipos están representados según los grupos establecidos *a priori*.

Tanto el dendrograma como el análisis de coordenadas principales permitieron ver que los genotipos no se agrupan según la distribución establecida *a priori*. Se nota que las entradas de 6 carreras españolas no constituyen un único grupo homogéneo. También parece que los genotipos españoles de 2 carreras no son tan distantes genéticamente de las variedades de referencia. Por tanto, parece apropiado estudiar la estructura de poblaciones que existe dentro del germoplasma analizado.

3-3-3. Estructura genética

La estructura genética del germoplasma analizado fue estudiada con el programa STRUCTURE (Falush *et al.*, 2003). Este análisis clasificó los 225 genotipos en diferentes poblaciones (o clusters) de individuos próximos genéticamente, minimizando el desequilibrio de ligamiento dentro de los clusters. Ensayando un número K de poblaciones de 1 a 10, el programa sugirió dividir los 225 genotipos en 4 poblaciones. Este número se eligió según el valor del ‘*log-likelihood*’. En nuestro caso, el punto de inflexión descrito por Falush *et al.* (2003), producido por la curva de los distintos valores del ‘*log-likelihood*’ en función de los valores de K, no se manifestó claramente. Sin embargo se observó que las poblaciones obtenidas cada vez que se ejecutaba el programa eran estables hasta K=4. A partir de este valor, los clusters variaban de una ejecución a otra. Por otra parte, el valor del ‘*log-likelihood*’ aumenta considerablemente de K=1 a K=4, estabilizándose posteriormente (Figura 3.3). Basándose sobre estos dos criterios, se optó por un número de poblaciones de K=4.

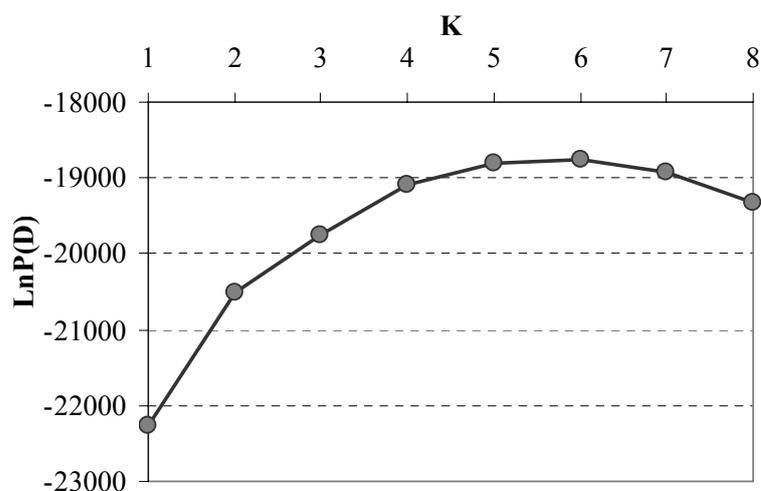


Figura 3.3. Evolución del valor del ‘*log-likelihood*’ (LnP(D)) en función del número K de poblaciones (promedio de 25 repeticiones del análisis para cada valor de K).

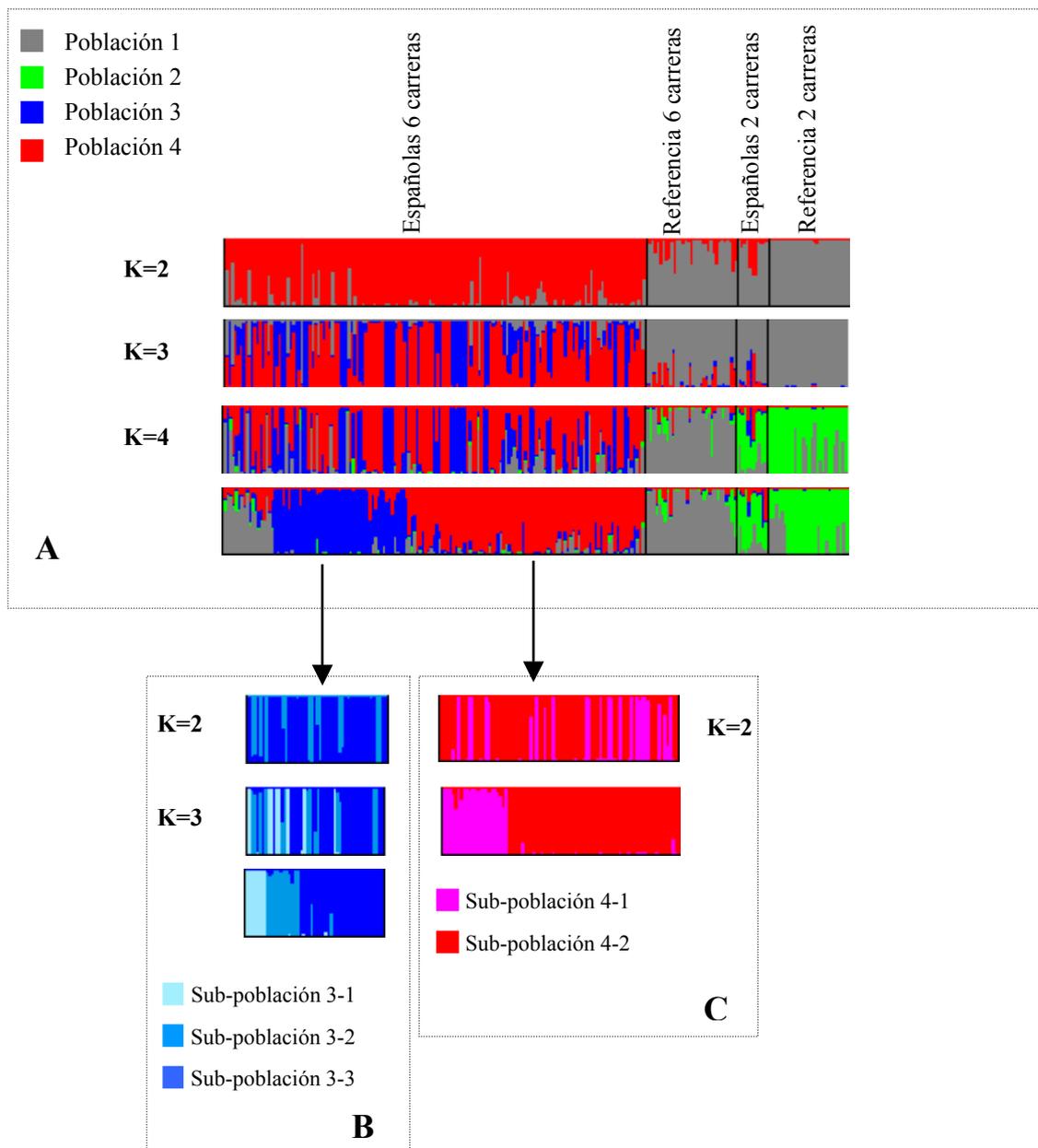


Figura 3.4. Clasificación de los 225 genotipos en diferentes poblaciones según su estructura genética a partir de la información procurada por 64 microsatélites, analizados con el programa STRUCTURE. Los genotipos están representados en columnas. K es el número de poblaciones (A) o sub-poblaciones (B y C). En los 3 primeros gráficos de la Figura A, los genotipos están representados por orden ascendente según los grupos de germoplasma al que pertenecen. En el gráfico 4 Figura A), los genotipos están clasificados según la probabilidad Q de pertenencia a una población. En la Figura B y C están representados, respectivamente, las sub-poblaciones de las poblaciones 3 y 4. En los 2 primeros gráficos de la Figura B, así como el primer gráfico de la Figura C, los genotipos están representados por orden de número de entrada. En el tercer gráfico de la Figura B así como el segundo de la Figura C, los genotipos están clasificados según la probabilidad Q de pertenencia a una sub-población. Las poblaciones y sub-poblaciones están simbolizadas por diferentes colores.

La primera separación que ocurrió entre los 225 genotipos separó las entradas españolas de 6 carreras del resto del material. (Figura 3.4A). La siguiente división ($K=3$) sucedió dentro de las entradas españolas de 6 carreras (colores rojo y azul, Figura 3.4A). En la tercera etapa ($K=4$), se separaron las variedades de 6 carreras de referencia, junto con algunas entradas de 6 carreras de la colección nuclear española, de las entradas de 2 carreras (colores gris y verde en la Figura 3.4A). Se observa la agrupación de la mayoría de las entradas de 2 carreras de invierno con las variedades de referencia de 6 carreras. Al final, se obtuvieron 4 poblaciones, con el 68% de los genotipos asignados a uno de ellas con una probabilidad de pertenencia superior a 0,75. Los 225 genotipos se repartieron en las 4 poblaciones de la forma siguiente:

- Población 1: contiene 55 genotipos de los que 31 son variedades de referencia de 6 carreras. Se agrupan con ellas 18 *landraces* de 6 carreras de la colección nuclear y 6 variedades de invierno de 2 carreras.
- Población 2: reúne 34 entradas, entre ellas se encuentran todas las variedades de 2 carreras de primavera y 9 *landraces* de 2 carreras de la colección nuclear. Se encuentran también en esta población, 4 variedades de 2 carreras de invierno y 2 variedades de 6 carreras (Olli y Dobla).
- Población 3: agrupa 50 *landraces*, la mayoría provenientes de la Meseta central de España.
- Población 4: contiene 86 entradas, originarias fundamentalmente del Sur del país, la costa mediterránea y el Valle del Ebro.

La posible existencia de sub-poblaciones en las poblaciones 3 y 4 se estudió mediante el análisis individual de cada una de ellas con el programa STRUCTURE, como recomiendan Pritchard y Wen (2002). La población 3 se subdividió en 3 sub-poblaciones (Figura 3.4B). La sub-población 1 reúne 8 entradas (4 de Castilla y León, 3 de Castilla-La Mancha y 1 del Valle del Ebro); en la sub-población 2 se encuentran 12 entradas, la mayoría de Castilla y León. En la sub-población 3 se agrupan 30 genotipos, siendo el 60% de los genotipos originarios de Castilla-La Mancha.

La población 4 se subdividió en 2 sub-poblaciones de 24 y 62 entradas respectivamente (Figura 3.4C). La sub-población 4-1 se compone mayoritariamente de entradas procedentes del Valle del Ebro, mientras que la sub-población 4-2 agrupa entradas originarias del Sur de España.

3-3-4. Análisis molecular de varianza

La divergencia genética entre las poblaciones encontradas en el apartado 3-3-3 fue estimada con un análisis molecular de varianza (AMOVA), realizado según los dos modelos del proceso de evolución de los microsatélites *IAM* y *SMM*. Los índices de fijación F_{ST} y R_{ST} resultado de este análisis son altamente significativos. Fueron particularmente altos cuando la comparación implicaba la población 2 (genotipos de 2 carreras de referencia y españoles) con las poblaciones 3 y 4, que contienen las entradas de 6 carreras españolas (Tabla 3.11).

Tabla 3.11. Índices de Fijación F_{ST} y R_{ST} de la comparación entre poblaciones y sub-poblaciones derivados del análisis con el programa STRUCTURE.

F_{ST}					R_{ST}				
Poblaciones	1	2	3	4	Poblaciones	1	2	3	4
1	-				1	-			
2	13,0	-			2	12,6	-		
3	14,0	23,3	-		3	20,9	32,0	-	
4	11,9	23,0	11,8	-	4	19,9	32,8	10,5	-

Sub-poblaciones				Sub-poblaciones			
	3-1	3-2	3-3		3-1	3-2	3-3
3-1	-			3-1	-		
3-2	51,3	-		3-2	70,0	-	
3-3	25,4	21,5	-	3-3	25,7	29,8	-

Sub-poblaciones			Sub-poblaciones		
	4-1	4-2		4-1	4-2
4-1	-		4-1	-	
4-2	8,3	-	4-2	7,3	-

La diferencia entre las poblaciones 1 y 2 no fue tan acentuada. Por otra parte, los genotipos de la población 1 parecen relativamente más divergentes de los genotipos de la población 3 (Meseta) que de los de la población 4 (Sur y Valle del Ebro).

Por otro lado, cuando se comparan poblaciones que contienen genotipos de referencia entre si (población 1 y población 2) o poblaciones españolas entre si (poblaciones 3 y 4), se observa que los valores de F_{ST} son relativamente más altos que los de R_{ST} . Sin embargo, cuando la comparación implica las poblaciones 1 y 2 con las poblaciones 3 y 4, los valores del índice R_{ST} son más altos que los del F_{ST} .

La estimación de la divergencia genética entre las sub-poblaciones de las poblaciones 3 y 4, respectivamente manifestó valores F_{ST} y R_{ST} significativos. Estos

valores fueron más altos entre las 3 sub-poblaciones de la población 3 en comparación con los valores registrados entre las sub-poblaciones 4-1 y 4-2.

3-3-5. Distribución del STS MWG699

El análisis de los 225 genotipos con el STS MWG699, propuesto por Tanno *et al.* (1999, 2002) como marcador de domesticación de la cebada, presenta tres haplotipos A, D y K. Este último solo se encuentra en genotipos de 2 carreras. Los otros dos haplotipos están ampliamente distribuidos en la colección nuclear española (Casas *et al.*, 2005). La distribución de los haplotipos de este marcador en las poblaciones y sub-poblaciones de genotipos definidos en este estudio se presentan en la Tabla 3.12.

Tabla 3.12. Distribución de los haplotipos del STS MWG699 dentro de las poblaciones y sub-poblaciones

Poblaciones y sub-poblaciones	Nº de genotipos	Haplotipos		
		A	D	K
1	55	23	26	6
2	34	4	2	28
3	50	9	40	1
3-1	8	8	0	0
3-2	12	0	12	0
3-3	30	1	28	1
4	86	51	35	0
4-1	24	4	20	0
4-2	62	47	15	0
Total	225	87	103	35

Los haplotipos A y D se distribuyen de forma relativamente similar en la población 1, constituido en su mayoría por las variedades de referencia de 6 carreras. Los 6 genotipos del haplotipo K presentes en esta población corresponden a las entradas de 2 carreras. El haplotipo K parece ser el más frecuente, en general, en las entradas de 2 carreras. De hecho en la población 2, el 82% de los genotipos son de este tipo. En las poblaciones 3 y 4, se observa una distribución interesante de los haplotipos A y D. La frecuencia del haplotipo A está reducida en la población 3 (9 entradas), mientras que el haplotipo D es muy frecuente. Curiosamente en la sub-división de la población 3, las entradas con el haplotipo A se separan del resto de los genotipos para constituir la sub-población 3-1, con la excepción de una *landrace* que se agrupa con las entradas de la sub-población 3-3. En la población 4, el haplotipo A es relativamente más frecuente que

el D (59% y 41% respectivamente). Estos haplotipos se distribuyen de forma distinta en las dos sub-poblaciones de la población 4. El haplotipo D caracteriza casi todas las entradas de la sub-población 4-1 (genotipos del Valle del Ebro) mientras que las entradas de la sub-población 4-2 (Sur) son principalmente de tipo A (Tabla 3.12)

3-3-6. Desequilibrio de ligamiento

El desequilibrio de ligamiento entre pares de loci fue evaluado en la totalidad del material analizado y en las 4 poblaciones derivadas de STRUCTURE, mediante el programa TASSEL (Zhang *et al.*, 2006)

Este análisis reveló la existencia de desequilibrio de ligamiento en la totalidad de los 225 genotipos considerados, así como en las distintas poblaciones analizadas, pero con unas tasas variables. En la totalidad de los genotipos se detectaron unos valores de r^2 que varían en su mayoría (98,5% de los casos) entre 0 y 0,1. Aunque los valores de r^2 detectados son pequeños, su probabilidad de significación en la mayoría de los casos (66%) es altamente significativa (Figura 3.5A), indicando la presencia de un desequilibrio de ligamiento importante entre los pares de loci

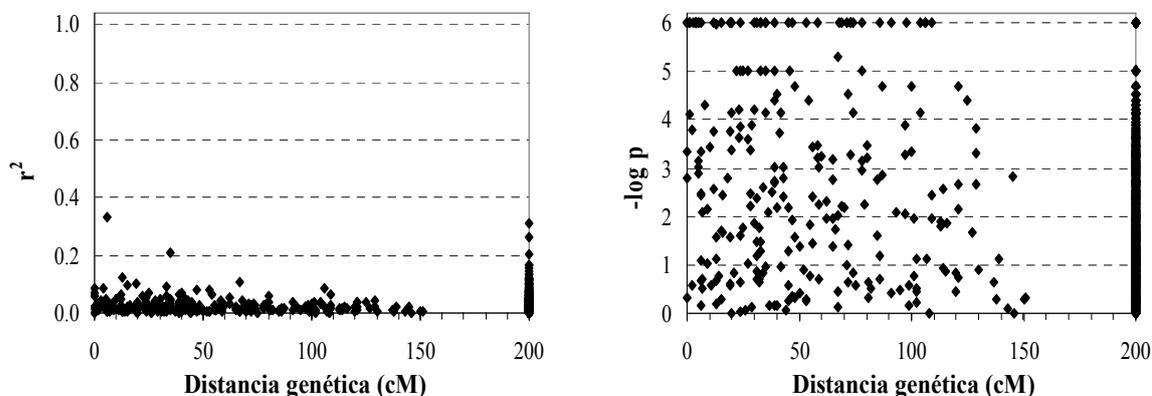


Figura 3.5A. Evaluación del desequilibrio de ligamiento entre pares de loci, en los 225 genotipos analizados: 1-Representación de la variación de r^2 (cuadrado del coeficiente de correlación entre alelos) en función de las distancias genéticas entre pares de loci (izquierda); 2- Evolución de la probabilidad de significación ($-\log p$) de los valores de r^2 con respecto a las distancias genéticas entre loci (derecha).

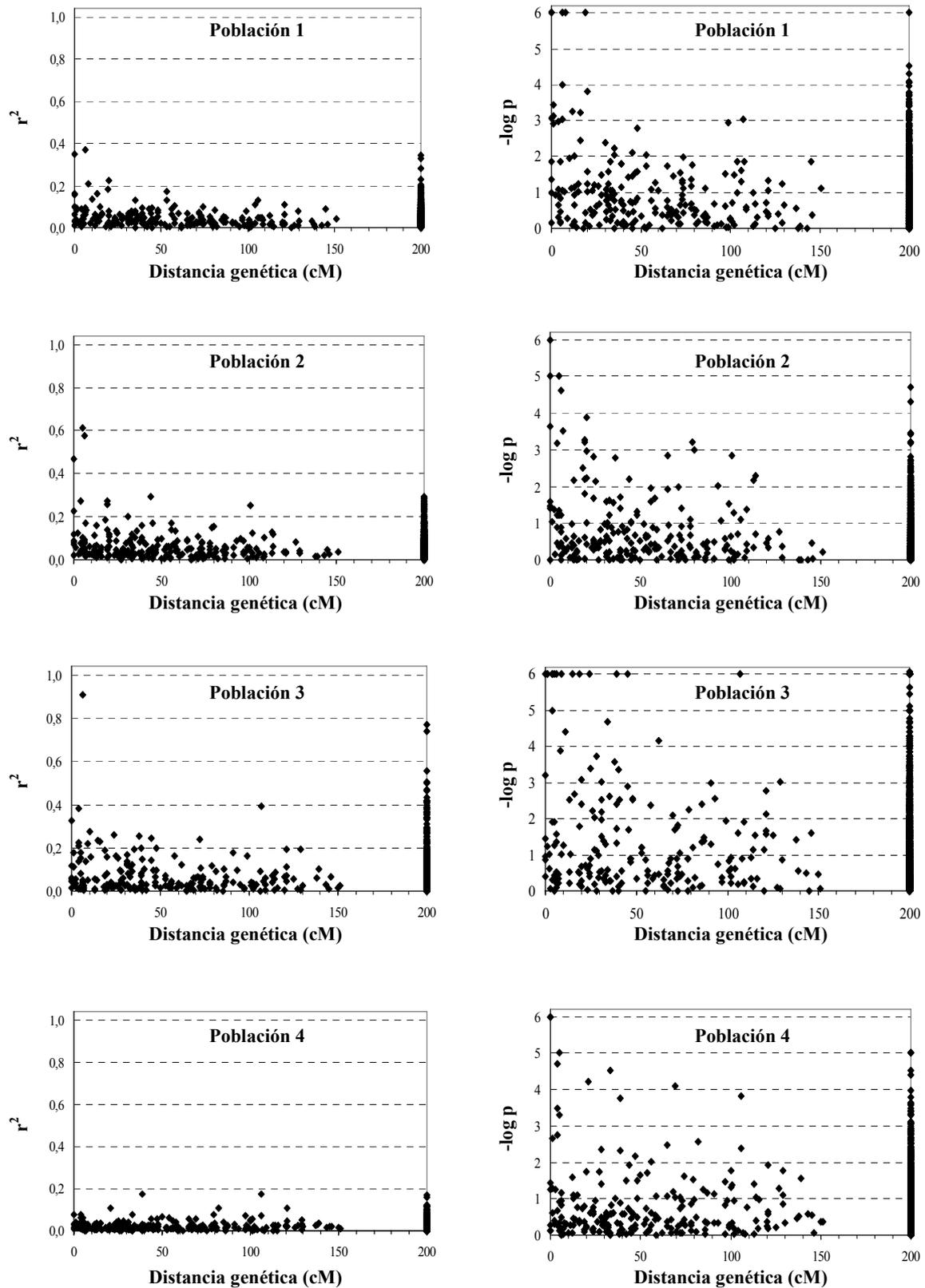


Figura 3.5B. Evaluación del desequilibrio de ligamiento entre pares de loci, en las 4 poblaciones obtenidas por STRUCTURE: 1-Representación de la variación de r^2 (cuadrado del coeficiente de correlación entre alelos) en función de las distancias genéticas entre pares de loci (izquierda); 2-Evolución de la probabilidad de significación ($-\log p$) de los valores de r^2 con respecto a las distancias genéticas entre loci (derecha)

En las poblaciones 1, 2 y 3, los valores de r^2 son relativamente más altos que los observados en la totalidad de los genotipos. Estos valores tienden a disminuir al aumentar la distancia genética entre los pares de marcadores aunque se observaron también valores de r^2 altos entre pares de loci no ligados. Este fenómeno se aprecia particularmente en la población 3, que es la que mayor desequilibrio de ligamiento presentó. Por el contrario, la población 4 muestra valores de r^2 muy bajos, que no superan el valor 0,2, tanto entre loci ligados como no ligados. Esta población es la que menos desequilibrio de ligamiento exhibió. En las Figuras 3.5A y 3.5B se presentan las variaciones de la r^2 en función de la distancia genética entre los pares de marcadores considerados, así como las probabilidades de significación (-log p) de estos valores.

Se calcularon los promedios de los valores de r^2 entre pares de loci que se encuentran a diferentes distancias genéticas: inferior a 10 cM, superior a 10 cM dentro del mismo cromosoma y entre loci no ligados (localizados en distintas cromosomas).

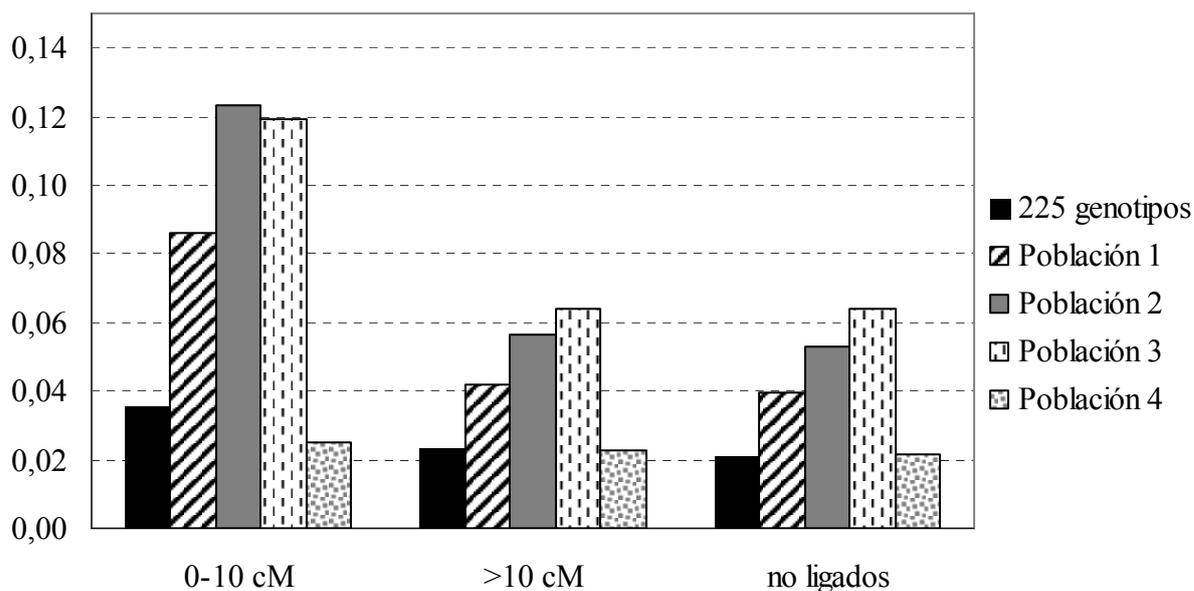


Figura 3.6. Promedios de los valores de r^2 , según las 3 clases de distancias definidas (0-10 cM, >10 cM, no ligados), en los 225 genotipos analizados y las diferentes poblaciones.

Este análisis permitió ver claramente que el desequilibrio de ligamiento se concentra, en general, entre los pares de loci que están a una distancia inferior o igual a 10 cM (Figura 3.6). Más allá de esta distancia se observa una disminución considerable

de los valores de r^2 . Por otra parte, se nota también que los valores de r^2 entre pares de loci del mismo cromosoma, presentes a una distancia superior a 10 cM, son parecidos a los valores observados entre pares de loci de distintos cromosomas.

El análisis de varianza realizado para comparar por un lado las poblaciones entre sí, y por otro las distintas clases de distancias (0-10 cM, >10 cM y loci no ligados), permitió ver que las diferencias son altamente significativas tanto entre poblaciones como entre clases (Tabla 3.13). Las diferencias entre poblaciones se deben esencialmente al contraste de la población 4 frente a las demás, lo que explica el 80% de la varianza. Las diferencias observadas entre clases corresponden fundamentalmente al contraste de la clase 1 (distancias ≤ 10 cM) frente a las otras dos. Esto representa el 93% de la varianza observada entre las clases.

Tabla 3.13. Comparación mediante análisis de varianza y contrastes entre poblaciones y clases de distancias y interacción entre las dos variables.

Fuente de variación	GI	Suma de cuadrados	F	Significación
Clase	2	0,243	48,07	<0,0001
Población	3	0,553	72,95	<0,0001
Clase*Población	6	0,081	5,33	<0,0001
Contrastes				
Población 1 vs poblaciones 2-3-4	1	0,005	1,91	0,1670
Población 2 vs poblaciones 1-3-4	1	0,100	39,51	<0,0001
Población 3 vs poblaciones 1-2-4	1	0,176	69,74	<0,0001
Población 4 vs poblaciones 1-2-3	1	0,446	176,28	<0,0001
Clase 1 vs clases 2-3	1	0,227	89,76	<0,0001
Clase 2 vs clase 3	1	0,002	0,79	0,3738
(Clase 1 vs clases 2-3) * (Población 4 vs poblaciones 1-2-3)	1	0,068	26,96	<0,0001
Error	6595	16,673		

Los pares de marcadores en desequilibrio de ligamiento, en los 225 genotipos y en cada población, están representados en el anexo 1 (Figuras 1, 2, 3, 4 y 5).

3-4. Discusión

3-4-1. Polimorfismo e índices de diversidad

Con los 64 microsatélites empleados se pudieron identificar los 225 genotipos estudiados. Fueron detectados un total de 669 alelos, con una media de 10,45 alelos por locus. Empleando marcadores comunes, otros autores revelaron un número de alelos por locus más bajos en muestras pequeñas (20-40 genotipos) de *Hordeum vulgare* (Russell *et al.*, 1997; Macaulay *et al.*, 2001; Karakousis *et al.*, 2003; Sjakste *et al.*, 2003). La

media encontrada en este trabajo es similar a la observada en poblaciones de *H. spontaneum* (Ivandic *et al.*, 2002; Baek *et al.*, 2003) y en estudios con mayor tamaño de muestra (100-150 individuos) de *H. vulgare* (Russell *et al.*, 2000; Matus and Hayes, 2002).

Los microsatélites derivados de genes resultaron ser menos polimórficos que los que provienen de librerías genómicas, con la excepción de HvM003 y HvM004, que detectaron cada uno más de 20 alelos. Esto concuerda con los resultados de Matus y Hayes (2002), quienes encontraron que los SSRs derivados de genes son menos polimórficos que los derivados de librerías genómicas. Sin embargo no son menos informativos en la descripción de la diversidad en este estudio.

Los índices de diversidad calculados para cada microsatélite, mostraron valores altos, parecidos a los encontrados por Struss y Plieske (1998), en una colección mundial de cebadas. La alta diversidad observada en el germoplasma estudiado, se concentra especialmente en las cebadas de 6 carreras españolas, donde se encuentra el mayor número de alelos detectados (591 del total de 669). Se observó además que el índice de diversidad aumenta al aumentar el número de alelos detectados por locus. Se encontró un coeficiente de correlación de 0,78 entre estos dos parámetros. También, Huang *et al.* (2002) encontraron un coeficiente de correlación significativo entre valores de diversidad y número de alelos detectados por diferentes loci.

El examen de la distribución de los alelos en los 4 grupos considerados puso de manifiesto la presencia de alelos específicos en cada grupo, principalmente en las cebadas de 6 carreras españolas (184 alelos). Este alto número de alelos específicos, en el germoplasma español de 6 carreras, sugiere una posible pérdida de alelos durante la creación de variedades élite, como consecuencia de una fuerte presión de selección, o que estos alelos nunca han estado presentes en los genotipos progenitores empleados para la creación de dichas variedades. En este contexto, Russell *et al.* (2000) interpretaron la depresión de la variabilidad genética en los recientes cultivares como resultante de una fuerte selección. Allard (1996) apoyaba esta idea, pero indicando también que la reducción de la variabilidad alélica sería debida sobre todo a una selección purificante por la eliminación de alelos indeseables, más que a una erosión de la variabilidad genética útil.

Por otra parte, la presencia de alelos específicos tanto en las variedades de referencia que no se encuentran en la colección nuclear como en las españolas, apunta la

posibilidad de que los dos grupos de germoplasma tengan orígenes distintos, o se hayan separado en una época relativamente lejana.

3-4-2. Análisis molecular de varianza

El análisis molecular de varianza permitió ver que existe una divergencia genética entre el germoplasma español y el de referencia, tanto para las cebadas de 6 carreras como para las de 2 carreras. Considerando el modelo 'infinite allele model', la diferencia entre los grupos de referencia y los grupos españoles es similar para las cebadas de 2 y 6 carreras. Los índices estadísticos de fijación (F_{ST}) que resultaron de este análisis son parecidos a los encontrados por Maestri *et al.* (2002) y Koebner *et al.* (2003) en la comparación entre cebadas de invierno y primavera. Sin embargo, el índice estadístico de fijación R_{ST} , derivado del modelo 'stepwise mutation model', que es el modelo de evolución más apropiado para los microsatélites, muestra que la divergencia que existe entre el germoplasma español y el de referencia es mucho más importante entre los grupos de cebadas de 6 carreras que entre los grupos de 2 carreras. Esto se debe al hecho de que en las cebadas de 6 carreras, además de las diferencias que existen en las frecuencias de los alelos en los dos grupos (español y referencia), existe una diferencia de tamaño importante entre los alelos más frecuentes en los dos grupos. Este fenómeno no se aprecia tanto entre los grupos de 2 carreras. Esto sugiere que el germoplasma español de 6 carreras es más distante genéticamente del germoplasma de referencia, en comparación con el de dos carreras.

El análisis molecular de varianza locus por locus permitió destacar los microsatélites que diferencian los grupos españoles de los de referencia. Los loci que diferencian entre los grupos de 2 carreras son distintos de los que separan los grupos de 6 carreras. Se observó también en algunos loci la presencia de valores altamente significativos de F_{ST} , pero con un R_{ST} no significativo (ej. HvM036 o Bmag378, en los grupos de 6 carreras). Esto se debe al motivo explicado anteriormente, que la diferencia de tamaño entre los alelos frecuentes en un grupo o en el otro es pequeña. Ivandic *et al.* (2002) también encontraron resultados distintos para los índices F_{ST} y R_{ST} al analizar microsatélites que diferenciaban grupos de genotipos de *H. spontaneum* del Creciente Fértil. Estos autores recomendaron emplear el método SMM para los microsatélites que presentan un número de alelos limitado mientras que el método IAM sería más adecuado para marcadores muy polimórficos. Sin embargo, sugirieron que el uso de los dos métodos era la forma más acertada de buscar SSRs informativos.

3-4-3. Análisis multivariantes

El dendrograma elaborado a partir de la matriz de distancias permitió ver que los genotipos estudiados no se agrupan según la división establecida a priori (entradas españolas y entradas de referencia, de otros orígenes). Las entradas españolas de 6 carreras de la colección nuclear se repartieron aparentemente en 5 grupos, mientras que las de 2 carreras se juntaron con las variedades de dos carreras de referencia.

Dentro de los grupos de 6 carreras españoles, se perciben agrupamientos de individuos que provienen de la misma zona geográfica. Esto no es sorprendente, porque dos individuos que provienen de la misma región, o de zonas próximas tienen más probabilidades de ser próximos genéticamente. También se observaron agrupaciones entre entradas de zonas geográficas alejadas. Dentro de este grupo de germoplasma, la división más importante ocurre entre las entradas del grupo I (predominantemente del Sur de España) y las entradas del grupo II (predominantemente de la zona central, de las Mesetas). Los genotipos del Valle del Ebro parecen intermedios entre los dos grupos, dividiéndose entre los grupos I y II. Por otra parte, en el grupo VIII, se observó la agrupación de algunas entradas de 6 carreras españolas con las variedades de 6 carreras de referencia. Éstas se subdividen en dos sub-grupos, separando las variedades de primavera de las de invierno. Junto a estas últimas se encuentran también las variedades de dos carreras de invierno.

Curiosamente, las entradas españolas de 2 y algunas de las de 6 carreras, que se agrupan respectivamente con variedades de 2 y de 6 carreras, parecen estar cercanas al material de primavera de referencia.

El análisis en coordenadas principales confirmó los resultados obtenidos con el dendrograma. Se observó una separación evidente de los genotipos de 6 carreras españoles de los otros grupos establecidos *a priori*, con la excepción de algunos genotipos que se juntaron con las variedades de 6 carreras. Los genotipos de este primer grupo formaron una nube de puntos muy dispersos, lo que revela la gran diversidad que existe en este grupo. Las entradas de 2 carreras españolas no parecen distintas de las variedades de referencia, tal como se pensaba. Se nota también que las variedades de 2 carreras de invierno están más próximas genéticamente a las variedades de 6 carreras; de hecho en el dendrograma se apartaban de las variedades de primavera. Esto no es sorprendente, por el hecho que muchas de estas variedades tienen cultivares comunes en su pedigrí.

3-4-4. Estructura genética

El análisis de los 225 genotipos con el programa STRUCTURE permitió ver la posible estructura de poblaciones que existe en el germoplasma estudiado. Este programa se propone subdividir la muestra en poblaciones en las que se alcance el equilibrio de Hardy-Weinberg y se minimice el desequilibrio de ligamiento. El primer criterio no tiene aplicación en una muestra de líneas puras de una especie autógama, pero el segundo es plenamente aplicable. Se trata de un nuevo algoritmo de análisis de cluster, basado en criterios de genética de poblaciones, y con la ventaja de que ofrece criterios con un cierto grado de objetividad, por lo que la agrupación final no depende sólo del criterio del investigador.

Este análisis confirmó en gran manera los resultados obtenidos por el dendrograma y por el análisis de coordenadas principales, en lo referente a la separación de las entradas de 6 carreras españolas de las variedades de referencia, y al agrupamiento de los genotipos de 2 carreras españolas con las variedades de 2 carreras de referencia. Es interesante destacar que, después de la separación de las entradas españolas de 6 carreras del resto de los genotipos analizados (al nivel $K=2$), la segunda partición sucede dentro de este mismo grupo ($K=3$), desgajando la mayoría de los genotipos de la zona central (Mesetas) de los del resto de España (principalmente Sur, costa mediterránea y Valle del Ebro). Curiosamente, ésta división ocurre antes de que ocurra ninguna separación en el otro gran grupo, que reúne a casi todas las variedades comerciales, de dos y seis carreras, de primavera y de invierno. La última división considerada ($K=4$) establece la división entre variedades de dos carreras de primavera y variedades de seis carreras y de invierno de dos carreras. Por otro lado, se observó la posible existencia de sub-poblaciones dentro de las dos principales poblaciones españolas de 6 carreras. La subdivisión observada en la población 3 (fundamentalmente, de las Mesetas) parecía seguir un gradiente geográfico. Así, se observó la predominancia de genotipos originarios de Castilla y León en una sub-población y genotipos de Castilla-La Mancha en otra sub-población. Por otra parte, en la población 4, se obtuvieron dos sub-poblaciones, separando los genotipos del Valle del Ebro de los del Sur.

La divergencia genética entre las poblaciones creadas por STRUCTURE fue confirmada por el análisis molecular de varianza. Esta diferencia resultó mayor comparando las poblaciones de las entradas españolas (3 y 4) con las poblaciones de las

variedades de referencia (1 y 2). La divergencia genética entre la población 2 (genotipos de 2 carreras) y las poblaciones 3 y 4 (genotipos de 6 carreras españoles) está basada tanto sobre diferencias en clases de alelos que existen en los dos grupos de germoplasma (2 y 6 carreras) como en diferencias de tamaño grandes entre estas clases de alelos. Sin embargo la divergencia entre las poblaciones 1 (variedades de referencia de 6 carreras) y 2 (genotipos de 2 carreras) se debe exclusivamente a diferencias entre clases de alelos.

La comparación de los grupos españoles entre sí reveló que la divergencia genética que existe entre estos grupos es cierta. Los valores de F_{ST} y R_{ST} que resultaron de este análisis son similares a los observados por Russell *et al.* (2003) en la comparación de *landraces* de Jordania y Siria.

Aunque la población 3 parece más homogénea en cuanto a su constitución geográfica, la divergencia genética entre sus sub-poblaciones es mayor que la que existe entre los dos sub-poblaciones 4-1 (Valle del Ebro) y 4-2 (Sur).

Los agrupamientos de genotipos que provienen de regiones geográficas similares en una misma población y/o sub-población pueden estar relacionados con caracteres de adaptación.

La separación del material español en dos poblaciones mayoritarias, se presenta antes que la de los grandes grupos de germoplasma de cebada (2/6 carreras, invierno/primavera), lo que hace pensar que esa divergencia es más remota en el tiempo. De hecho la distribución de los haplotipos A y D del STS MWG699 en los grupos de entradas de 6 carreras españolas no es aleatoria. La dominancia del tipo A en el grupo del Sur y del tipo D en los grupos de la Meseta y del Valle del Ebro es remarcable. En su estudio, Tanno *et al.* (2002) mostraron que el STS MWG699 divide los genotipos de 6 carreras en dos grupos distintos que corresponden a los haplotipos A y D. El haplotipo A está ampliamente distribuido mientras que el D se encuentra solamente en la región Mediterránea. Asimismo indicaron que el haplotipo A se encuentra en genotipos de *H. spontaneum* originarios de Turkmenistán mientras que el haplotipo D se localiza también en algunas cebadas cultivadas de 2 carreras del norte de África y entradas de *H. spontaneum* de Marruecos. Casas *et al.* (2005) comentan la amplia distribución del haplotipo D en la colección nuclear española, y su presencia en genotipos del Sur y Centro de Europa, sin embargo no se observa en genotipos escandinavos. Todo esto sugiere la posible existencia de orígenes distintos en las cebadas españolas y que luego su distribución en la península ibérica se hizo según su

adaptación a las condiciones agro-ecológicas. En este sentido, Moralejo *et al.* (1994), también sugirieron la posible existencia de dos orígenes distintos en cebadas españolas de 2 carreras.

3-4-5. Desequilibrio de ligamiento

En la estimación del desequilibrio de ligamiento en los 225 genotipos, se detectaron valores de r^2 más bajos que en tres de las poblaciones, pero con significación más alta (Figuras 3.5 y 3.6). Esta no es una situación anómala, pues también ha sido descrita en estudios de líneas puras de *Arabidopsis* (otra planta autógama) por Nordborg *et al.* (2002) y por Schmid *et al.* (2006), y en trigo duro por Maccaferri *et al.* (2005). Los parámetros de medida de desequilibrio de ligamiento están influenciados por el tamaño de muestra y se ven especialmente afectados por tamaños de muestra pequeños (Flint-García *et al.*, 2003), así que no es posible comparar directamente los valores de estos parámetros para muestras de tamaños distintos. En cualquier caso, el programa STRUCTURE trabaja buscando grupos de genotipos dentro de los cuales se minimice el desequilibrio de ligamiento, por lo que las poblaciones resultantes de este análisis deberían mostrar menor desequilibrio de ligamiento que la muestra total. Aunque la significación de los parámetros de desequilibrio de ligamiento es superior para la muestra total que para las cuatro poblaciones, la comparación directa es impracticable, pues la significación de r^2 (y de cualquier otra medida de desequilibrio de ligamiento) depende del tamaño de la muestra (Ardlie *et al.*, 2002).

Los niveles de desequilibrio de ligamiento observados en las diferentes poblaciones no fueron semejantes. Las poblaciones 1 y 2 mostraron tasas de desequilibrio de ligamiento aparentemente mayores que las de la población 4 y algo menores que las de la población 3. Las poblaciones 1 y 2 están constituidas en su mayoría por cultivares, mientras que las poblaciones 3 y 4 están compuestas muy mayoritariamente por *landraces*. Caldwell *et al.* (2006) compararon medidas de desequilibrio de ligamiento en tres grupos de germoplasma de la cebada: cultivares, *landraces*, y cebada silvestre. Encontraron que el desequilibrio de ligamiento (r^2) era consistentemente superior en los cultivares frente a los otros dos grupos. Esta observación concuerda con lo observado para la comparación de las poblaciones 1 y 2 frente a la 4. La estructura que persiste en la población 3 (ver los elevados valores de los índices de fijación F_{ST} y R_{ST} entre las sub-poblaciones 3-1, 3-2 y 3-3 en la Tabla 3.11) puede ser una de las razones del alto desequilibrio de ligamiento observado en ella. De

hecho, en la población 4 también se observó una estructura subyacente, aunque la divergencia entre las sub-poblaciones 4-1 y 4-2 no fue tan acentuada como las que se observaron entre las sub-poblaciones 3-1, 3-2 y 3-3. Aparte de la estructura de población, el alto desequilibrio de ligamiento observado en la población 3 podría ser el resultado de un efecto fundador o de “cuello de botella”, que hubiera ocurrido probablemente en los ancestros de esta población. Por el contrario, en la población 4 pareció no haber restricciones a la panmixia (dentro de lo limitado de los cruzamientos de una especie fundamentalmente autógena), lo que permitió romper el desequilibrio de ligamiento entre loci ligeramente o no ligados.

Caldwell *et al.* (2006) afirman que los distintos niveles de desequilibrio de ligamiento que encontraron para los diferentes grupos de germoplasma de cebada fueron probablemente debidos a las diferentes historias de las poblaciones, asociadas con la ocurrencia de cuellos de botella y selección en el germoplasma domesticado. En nuestro caso, los distintos niveles de desequilibrio de ligamiento pueden deberse a fenómenos de mezcla incompleta de poblaciones de distintos orígenes, o de efecto fundador en la población 3, y a distintas historias filogenéticas en general entre las poblaciones. Así, uno de los factores que más afectan al desequilibrio de ligamiento es la recombinación. Cabe pensar que haya habido más meiosis efectivas, y por tanto mayor recombinación, en la larga historia de selección y adaptación de las *landraces* que en los programas de mejora que han llevado a la formación de los cultivares modernos. Aunque es imposible afirmarlo con seguridad.

En las poblaciones se observa una cierta disminución de r^2 al aumentar la distancia entre loci comparados (Figura 3.5B). Este fenómeno no fue observado en la muestra total (Figura 3.5A). Maccaferri *et al.* (2005) describieron una situación similar en un estudio de una colección de trigo duro.

El nivel del desequilibrio de ligamiento en las poblaciones, parece ser más alto entre pares de loci que se encuentran a una distancia menor que 10 cM. A partir de esta distancia, el valor de r^2 tendía a disminuir, de modo que los valores para marcadores del mismo cromosoma a más de 10 cM fueron similares a los observados entre loci en distintos cromosomas. De todos modos, el número de valores considerados para loci cercanos (a menos de 10 cM) es relativamente escaso por lo que no se puede establecer con certeza la relación entre el desequilibrio de ligamiento y la distancia. Kraakman *et al.* (2004), con un número mayor de marcadores analizados, sí que encontraron un

umbral de aproximadamente 10 cM a partir del cual disminuye el desequilibrio de ligamiento en variedades de cebada.

En el presente estudio, se observaron también tanto casos de desequilibrio de ligamiento significativos entre pares de loci no ligados, como casos de equilibrio entre pares de loci estrechamente ligados. Estos resultados parecen coincidir con los obtenidos por Nordborg *et al.* (2002), en *Arabidopsis*, y Kraakman *et al.* (2004), en un estudio de una colección de 146 variedades de cebada de primavera europeas de 2 carreras con 236 AFLPs.

La cebada cultivada parece tener un desequilibrio de ligamiento más extendido que otras especies. Kraft *et al.* (2000) indicaron que el desequilibrio de ligamiento, en líneas de remolacha, se reduce a una distancia inferior a 3 cM. En maíz, el desequilibrio de ligamiento disminuye a partir de 1500 pb (Remington *et al.*, 2001). Garris *et al.* (2003) mostraron que el desequilibrio de ligamiento en arroz disminuye a partir de 100 kb. En *Arabidopsis*, el análisis del desequilibrio de ligamiento en la región del gen de floración *FRIGIDA*, mostró que éste disminuye a partir de 250 kb, equivalente a 1 cM (Nordborg *et al.*, 2002). Sin embargo, la evaluación del desequilibrio de ligamiento alrededor del locus *RPM1* (resistencia a una enfermedad), en poblaciones fundadoras recolectadas recientemente, reveló que éste se extiende a unas distancias de 50-100 cM. Este desequilibrio de ligamiento elevado fue explicado por los autores como consecuencia del número limitado de recombinaciones que ocurrió entre estas poblaciones durante los últimos 200 años.

La comparación de la extensión del desequilibrio de ligamiento entre diferentes especies parece indicar que en las plantas alógamas, el desequilibrio de ligamiento decae más rápidamente que en las autógamas. Allard (1988), describió que en poblaciones autógamas se observan asociaciones multilocus estadísticamente significativas entre muchos loci polimórficos, mientras que estas asociaciones se limitan a loci estrechamente ligados en poblaciones alógamas. La explicación que daba este autor es que las asociaciones desarrolladas entre loci poco o no ligados se rompieron en cada generación, como consecuencia de la recombinación en plantas alógamas.

3-5. Referencias bibliográficas

Allard RW. 1988. Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. *Journal of Heredity*, 79: 225-238.

- Allard RW. 1996.** Genetic basis of the evolution of adaptedness in plants. *Euphytica*, 92: 1-11.
- Ardlie KG, Kruglyak L, Seielstad M. 2002.** Patterns of linkage disequilibrium in the human genome. *Nature Reviews Genetics*, 3: 299-309.
- Briggs DE. 1978.** *Barley*, Chapman Hall, London, 612 p.
- Bahrman N, Le Gouis J, Hariri D, Guilbaud L, Jestin L. 1999.** Genetic diversity of old French six-rowed winter barley varieties assessed with molecular, biochemical and morphological markers and its relation to BaMMV resistance. *Heredity*, 83: 568-574.
- Baek HJ, Beharav A, Nevo E. 2003.** Ecological-genomic diversity of microsatellites in wild barley, *Hordeum spontaneum*, populations in Jordan. *Theoretical & Applied Genetics*, 106: 397-410
- Bassam BJ, Caetano-Anolles G, Gresshoff PM. 1991.** Fast and sensitive silver staining of DNA in polyacrilamide gels. *Analytical Biochemistry*, 196: 80-83.
- Baum BR, Mechanda S, Penner GA, Ednie AB. 1998.** Establishment of a scheme for the identification of Canadian barley (*Hordeum vulgare* L.) six row cultivars using RAPD diagnostic bands. *Seed Science & Technology*, 26: 449-462.
- Baumer M, Cais R. 2000.** Abstammung der Gerstensorten. Ed Bayerische Landesanstalt für BodenKultur und Pflanzenbau. 109 p.
- Becker J, Heun M. 1995.** Barley microsatellites: allele variation and mapping. *Plant Molecular Biology*, 27: 835-845.
- Cain AJ, Harrison GA. 1958.** An analysis of the taxonomists' judgement of affinity. *Proceedings of the Zoological Society of London*, 131: 85-98.
- Caldwell K.S, Russell J, Langridge P, Powell W. 2006.** Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. *Genetics* 172: 557-567.
- Casas AM, Igartua E, Vallés MP, Molina-Cano JL. 1998.** Genetic diversity of barley cultivars grown in Spain, estimated by RFLP, similarity and coancestry coefficients. *Plant Breeding*, 117: 429-435.
- Casas AM, Yahiaoui S, Ciudad F, Igartua E. 2005.** Distribution of MWG699 polymorphism in Spanish European barleys. *Genome*, 48: 41-45.
- Cross RJ. 1994.** Geographical trends within a diverse spring barley collection as identified by agro/morphological and electrophoretic data. *Theoretical & Applied Genetics*, 88: 597-603.
- Dávila JA, Loarce Y, Ramsay L, Waugh R, Ferrer E. 1999.** Comparison of RAMP and SSR markers for study of wild barley genetic diversity. *Hereditas*, 131: 5-13.
- Ellis RP, McNicol JW, Baird E, Booth A, Lawrence P, Powell W. 1997.** The use of AFLPs to examine genetic relatedness in barley. *Molecular Breeding*, 3: 359-369.

- Falush D, Stephens M, Pritchard JK. 2003.** Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, 164: 1567–1587
- Flint-Garcia SA, Thornsberry JM, Buckler IV ES. 2003.** Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, 54: 357-374.
- Garris A, McCouch S, Kresovich S. 2003.** Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (*Oryza sativa* L.). *Genetics*, 165: 759–769.
- Graner A, Ludwig WF, Melchinger AE. 1994.** Relationships among European barley germplasm: II. Comparison of RFLPs and pedigree data. *Crop Science*, 34: 1199-1205.
- Hayes PM, Cereno J, Witsenboer H, Kuiper M, Zabeau M, Sato K, Kleinhofs AK, Kudrna D, Kilian A, Saghai Maroof M, Hoffman D and the NABGMP. 1997.** Characterizing and exploiting genetic diversity and quantitative traits in barley (*Hordeum vulgare*) using AFLP markers. *Journal of Agricultural Genomics*, Vol. 3, art. 2. En: <http://www.ncgr.org/jag/papers97/paper297/indexp297.html>
- Huang XQ, Börner A, Röder MS, Ganai MW. 2002.** Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. *Theoretical & Applied Genetics*, 105: 531-536.
- Igartua E, Gracia MP, Lasa JM, Medina B, Molina-Cano JL, Montoya JL, Romagosa I. 1998.** The Spanish barley core collection. *Genetic Resources and Crop Evolution*, 45: 475-481.
- Ivandic V, Hackett CA, Nevo E, Keith R, Thomas WTB, Forster BP. 2002.** Analysis of simple sequence repeats (SSRs) in wild barley from the Fertile Crescent: Associations with ecology, geography and flowering time. *Plant Molecular Biology*, 48: 511-527.
- Karakousis A, Barr AR, Chalmers KJ, Ablett GA, Holton TA, Henry RJ, Lim P, Langridge P. 2003.** Potential of SSR markers for plant breeding and variety identification in Australian barley germplasm. *Australian Journal of Agricultural Research*, 54:1197-1210.
- Koebner RMD, Donini P, Reeves JC, Cooke RJ, Law JR. 2003.** Temporal flux in the morphological and molecular diversity of UK barley. *Theoretical & Applied Genetics*, 106: 550–558.
- Komatsuda T, Nakamura I, Takaiwa F, Oka S. 1998.** Development of STS markers closely linked to the vrs1 locus in barley, *Hordeum vulgare*. *Genome*, 41: 680-685.
- Kraakman ATW, Niks RE, van der Berg P, Stam P, van Eeuwijk FA. 2004.** Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, 168: 435-446.
- Kraft T, Hansen M, Nilson N-O. 2000.** Linkage disequilibrium and fingerprinting in sugar beet. *Theoretical & Applied Genetics*, 101: 323–326.

- Kumar S, Tamura K, Nei M. 2004.** MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics*, 5: 150-163.
- Lasa JM, Igartua E, Ciudad FJ, Codesal P, García EV, Gracia MP, Medina B, Romagosa I, Molina-Cano JL, Montoya JL. 2001.** Morphological and agronomical diversity patterns in the Spanish barley core collection. *Hereditas*, 135: 217-225.
- Linde-Laursen I, Heslop-Harrison JS, Shepherd KW, Taketa S. 1997.** The barley genome and its relationship with the wheat genomes. A survey with an internationally agreed recommendation for barley chromosome nomenclature. *Hereditas*, 126: 1-16.
- Liu ZW, Biyashev RM, Saghai-Marroof MA. 1996.** Development of simple sequence repeat DNA markers and their integration into a barley linkage map. *Theoretical & Applied Genetics*, 93: 869-876.
- Macaulay M, Ramsay L, Powell W, Waugh R. 2001.** A representative, highly informative ‘genotyping set’ of barley SSRs. *Theoretical & Applied Genetics*, 102: 801-809.
- Maccaferri M, Sanguineti MC, Noli E, Tuberosa R. 2005.** Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Molecular Breeding*, 15: 271-289.
- Maestri E, Malcevski A, Massari A, Marmioli N. 2002.** Genomic analysis of cultivated barley (*Hordeum vulgare*) using sequence-tagged molecular markers. Estimates of divergence based on RFLP and PCR markers derived from stress-responsive genes, and simple sequence repeats (SSRs). *Molecular Genetics and Genomics*, 267: 186-201.
- Manninen O, Nissilä E. 1997.** Genetic diversity among Finnish six rowed barley cultivars based on pedigree information and DNA markers. *Hereditas*, 126: 87-93.
- Matus IA, Hayes PM. 2002.** Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. *Genome*, 45: 1095-1106.
- Melchinger AE, Graner A, Singh M, Messmer MM. 1994.** Relationships among European barley germplasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs. *Crop Science*, 34: 1191-1199.
- Moralejo M, Romagosa I, Salcedo G, Sánchez-Monge R, Molina-Cano JL. 1994.** On the origin of Spanish two-rowed barleys. *Theoretical & Applied Genetics*, 87: 829-836.
- Morgante M, Olivieri AM. 1993.** PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal*, 3: 175-182.
- Nei M. 1973.** Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA*, 70: 3321-3323.
- Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D. 2002.** The

extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, 30: 190-193.

Ohta T, Kimura K. 1973. The model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genetic Research*, 22: 201-204.

Pritchard JK, Wen W. 2002. Documentation for *structure* software: version 2. <http://pritch.bsd.uchicago.edu>.

Pillen K, Binder A, Kreuzkam B, Ramsay L, Waugh R, Förster J, León J. 2000. Mapping new EMBL-derived barley microsatellites and their use in differentiating German barley cultivars. *Theoretical & Applied Genetics*, 101: 652-660.

Ramsay L, Macaulay M, degli Ivanissevich S, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuvešson S, Morgante M, Massari A, Maestri E, Marmioli N, Sjakste T, Ganai M, Powell W, Waugh R. 2000. A simple sequence repeat based linkage map of barley. *Genetics*, 156: 1997-2005.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler IV ES. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences, USA*, 98: 11479-11484.

Rohlf FJ. 2000. NTSYSpc ver. 2.1. Numerical taxonomy and multivariate analysis system, Exeter Software, New York.

Russell JR, Fuller J, Young G, Thomas WTB, Taramino G, Macaulay M, Waugh R, Powell W. 1997. Discriminating between barley genotypes using microsatellite markers. *Genome*, 40: 442-450.

Russell JR, Ellis RP, Thomas WTB, Waugh R, Provan J, Booth A, Fuller J, Lawrence P, Young G, Powell W. 2000. A retrospective analysis of spring barley germplasm development from ‘foundation genotypes’ to currently successful cultivars. *Molecular Breeding*, 6: 553-568.

Russell JR, Booth A, Fuller JD, Baum M, Ceccarelli S, Grando S, Powell W. 2003. Patterns of polymorphism detected in the chloroplast and nuclear genomes of barley landraces sampled from Syria and Jordan. *Theoretical & Applied Genetics*, 107: 413-421.

Saghai-Maroo MA, Soliman KM, Jorgensen RA, Allard RW. 1984. Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proceedings of the National Academy of Sciences, USA*, 81: 8014-8018.

Saghai-Maroo MA, Zhang Q, Chojecki J. 1994a. RFLPs in cultivated barley and their application in the evaluation of malting quality cultivars. *Hereditas*, 121: 21-29.

Saghai-Maroo MA, Biyashev RM, Yang GP, Zhang Q, Allard RW. 1994b. Extraordinarily polymorphic microsatellite DNA in barley: species diversity,

chromosomal locations and populations dynamics. Proceedings of the National Academy of Sciences, USA, 91: 5466-5470.

SAS Institute. 1988. SAS/STAT User's Guide, Release 6.03. Edition. Cary, NC: SAS Institute Inc., 1028p

Schneider S, Roessli D, Excoffier L. 2000. Arlequin ver2.000. A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva. Switzerland.

Schmid KJ, Törjèk O, Meyer R, Schmutz H, Hoffman MH, Altmann T. 2006. Evidence for a large-scale population structure of *Arabidopsis thaliana* from genome-wide single nucleotide polymorphism markers. Theoretical & Applied Genetics, DOI 10.1007/s00122-006-0212-7

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139: 457-462.

Sjakste T, Rashal I, Röder M. 2003. Inheritance of microsatellite alleles in pedigrees of Latvian varieties and related European ancestors. Theoretical & Applied Genetics, 106: 539-549.

Struss D, Plieske J. 1998. The use of microsatellite markers for detection of genetic diversity in barley populations. Theoretical & Applied Genetics, 97:308-315.

Tanno K, Takaiwa F, Oka S, Komatsuda T. 1999. A nucleotide sequence linked to the *vrs1* locus for studies of differentiation in cultivated barley (*Hordeum vulgare L.*). Hereditas, 130: 77-82.

Tanno K, Taketa S, Takeda K, Komatsuda T. 2002. A DNA marker closely linked to the *vrs 1* locus (row type gene) indicates multiple origins of six rowed cultivated barley (*Hordeum vulgare L.*). Theoretical & Applied Genetics, 104: 54-60.

Zhang Z, Bradbury PJ, Kroon DE, Casstevens TM, Buckler ES. 2006. A software package for association and diversity analyses in plants and animals. PAGXIV, San Diego. <http://www.maizegenetics.net/bioinformatics/tassel/index.html>.

Capítulo 4

*Distribución de la diversidad
en relación con los datos
eco-geográficos*

4-1. Introducción

La cebada, *Hordeum vulgare* L., es uno de los principales cereales desde el punto de vista de producción mundial, y es cultivada en todas las áreas templadas del mundo (Bothmer *et al.*, 1995). Muestra una gran capacidad de adaptación a un rango amplio de condiciones climáticas, por lo que probablemente cuenta con una dotación genética flexible y diversa para caracteres de adaptación.

La cebada silvestre, *Hordeum vulgare* ssp. *spontaneum* C. Koch, es el ancestro de la cebada cultivada *H. vulgare* ssp. *vulgare*. En el Creciente Fértil, se encuentran poblaciones de esta especie en hábitats muy diversos, con distintas situaciones de disponibilidad de agua (de pantanos a desiertos) y temperatura (desde cerca de 0° hasta 30°C). Varios estudios de diversidad genética llevados a cabo con esta especie (Baek *et al.*, 2003; Baum *et al.*, 1997; Chalmers *et al.*, 1992; Huang *et al.*, 2002; Ivandic *et al.*, 2002, 2003; Kalendar *et al.*, 2000; Liviero *et al.*, 2002; Morrell *et al.*, 2005; Nevo 1992; Nevo *et al.*, 1979, 1986, 1997, 1998, 2005; Owuor *et al.*, 1997, 2003; Saghai Maroof *et al.*, 1990; Supronova *et al.*, 2004; Turpeinen *et al.*, 2001), han identificado líneas de divergencia en la diversidad genética (de isoenzimas y marcadores de ADN) a escalas micro y macroadaptativa.

A lo largo del tiempo, el proceso de domesticación de la cebada a partir de los centros de origen llevó a la aparición de variedades locales, también adaptadas a un rango amplio de ambientes. Estas variedades locales ocupan cinco millones de hectáreas de zonas marginales con alta incidencia de estrés hídrico (Ceccarelli, 1994). Russell *et al.* (2003) estudiaron la diversidad existente en una colección de variedades locales de cebada provenientes de cinco regiones ecogeográficas diferentes de Siria y Jordania con microsatélites nucleares y cloroplásticos. Estos autores encontraron que las cinco regiones difieren considerablemente entre sí en composición y grado de variabilidad genética. Zhang *et al.* (1993) propusieron la hipótesis de que este grado de diferenciación entre regiones geográficas y taxones (*H. vulgare* y *H. vulgare* ssp. *spontaneum*) indicaría que la domesticación y la adaptación subsiguiente han ocurrido no como un proceso de selección de alelos en loci individuales, sino que esos procesos habrían reorganizado el genoma en haplotipos multilocus adaptados a los distintos ambientes. En este mismo sentido, Allard (1999) indicó que es esperable que la autogamia sea especialmente eficaz en la reorganización de alelos favorables en

haplotipos que permitan un comportamiento favorable en respuesta al ambiente, aunque los alelos se encuentren en cromosomas distintos.

La cebada ha sido cultivada en España desde tiempos prehistóricos. Dada su historia de adaptación y selección bajo condiciones climáticas locales, es lógico pensar que pueda poseer genes de adaptación. La caracterización molecular de la colección nuclear española ha mostrado la existencia de un polimorfismo importante y la presencia de muchos alelos específicos no encontrados en los genotipos europeos. La búsqueda de una posible estructura de población en estas entradas permitió ver que el polimorfismo detectado no está distribuido de forma aleatoria. La separación bastante clara de los genotipos de la Meseta de los del Sur, Levante y Valle del Ebro sugiere una adaptación a las condiciones climáticas.

El objetivo de este capítulo es averiguar si existe una relación entre el polimorfismo observado en las entradas de la colección nuclear y las condiciones eco-geográficas del sitio de origen, y ver si estas condiciones influyeron en la separación de estas entradas en los grupos mayoritarios.

4-2. Material y Métodos

4-2-1. Material vegetal

Tan sólo las 159 líneas puras procedentes de *landraces* de la colección nuclear de cebadas españolas, definidas en el capítulo anterior, fueron analizadas en este estudio. Este conjunto está constituido por 11 entradas de 2 carreras y 148 entradas de 6 carreras. La procedencia de estos genotipos esta definida en la Tabla 3.2 del capítulo 3.

4-2-2. Marcadores moleculares

Un conjunto de setenta y tres marcadores moleculares fueron empleados en este análisis: 68 microsatélites, 3 STS y 2 genes candidatos.

De los 68 microsatélites, 64 están descritos en el apartado 3-2-1-2 del capítulo 3. Los 4 marcadores nuevos introducidos en este capítulo son scsr10477 (antes denominado Ctig5280, cromosoma 1H) y scsr07759 (antes denominado Ctig7981, cromosoma 2H), incluidos en el trabajo de Moralejo *et al.* (2004), EBmac640, descrito por Ramsay *et al.* (2000), en el cromosoma 2H y Hv635P2.4 descrito por Francia *et al.* (2004) en el cromosoma 5H. Los tres STSs utilizados fueron MWG699 (Casas *et al.*, 2005) y los marcadores MWG518 y MWG858. Estos últimos se obtuvieron a partir de

las sondas de RFLPs, amplificando y secuenciando los fragmentos de los genotipos Beka y Mogador, parentales de una población de mapeo. En ambos casos, se identificó un polimorfismo de tipo SNP dentro de la diana de un enzima de restricción, por lo que se pudieron convertir en marcadores de tipo CAPS: MWG518/NheI y MWG858/MspI (datos no publicados). Se detectaron 2 alelos distintos para cada marcador. Se han incluido también los marcadores HvZCCT y HvBM5A, candidatos a genes de vernalización. La secuencia de los cebadores utilizados y el procedimiento seguido para amplificar están descritos por von Zitzewitz *et al.* (2005). El cluster de genes HvZCCT se amplificó con los cebadores HvZCCT.01F y HvZCCT.02R, dando lugar a dos alelos, uno nulo, por ausencia de producto y otro de 1400 pb. El intrón 1 del gen HvBM5A se amplificó con los cebadores HvBM5.55F y HvBM5.56R produciendo 4 alelos distintos: 4850 pb, 1200 pb, 150 pb y otro nulo. El alelo nulo, como consecuencia de una delección en el intrón, se amplificó con los cebadores HvBM5.42F y HvBM5.43R, generando un producto de 1900 pb. Para simplificar, en el texto se utilizará la clasificación en cuatro alelos obtenida con los cebadores 55F/56R.

La introducción de estos 8 nuevos marcadores en este capítulo se debe a su localización en regiones genómicas de gran importancia en el desarrollo del ciclo vegetativo de la cebada. En el cromosoma 1H, scssr10477 y MWG518 se localizan en la zona del gen de respuesta a fotoperiodo corto (*Ppd-H2*). MWG858 y scssr07759 flanquean la zona del gen de respuesta a fotoperiodo largo (*Ppd-H1*) del cromosoma 2H. En este mismo cromosoma, el microsatélite EBmac640 se sitúa en la zona del gen de precocidad *Eam6* (*early maturity*), de importancia capital en la determinación de la fecha de floración en condiciones mediterráneas (Boyd *et al.*, 2003). Por otra parte, HvZCCT y HvBM5A, son marcadores candidatos de los genes de respuesta a vernalización *Vrn-H2* (4H) y *Vrn-H1* (5H), respectivamente. El microsatélite Hv635P2.4, es también un marcador flanqueante al gen de respuesta a vernalización del cromosoma 5H.

Las amplificaciones y la electroforesis se realizaron de acuerdo al protocolo descrito en cada caso.

4-2-3. Variables geográficas y climáticas

Las entradas de la colección nuclear disponen de datos de pasaporte. Estos incluyen principalmente las coordenadas geográficas del sitio de recolección, latitud, longitud y altitud. Estas entradas fueron clasificadas en grupos de tipos potencialmente

distintos según una clasificación agro-ecológica de zonas de cultivo de la cebada. Para ello, se hizo un análisis de cluster de los valores de rendimiento históricos durante 47 años (1948-1994) de las provincias españolas (Igartua *et al.*, 1998).

Los datos climáticos correspondientes a los sitios de recolección de cada entrada fueron extraídos a partir de la base de datos SIGA (Sistema de Información Geográfica de Agricultura), del Ministerio de Agricultura, Pesca y Alimentación. (<http://www.mapa.es/es/sig/pags/siga/intro.htm>). Los datos climáticos de la base de datos fueron elaborados entre 1960 y 1996, sobre 4186 localidades del país, y suministrados por el Instituto Nacional de Meteorología.

Para 77 entradas se encontraron estaciones climáticas que se ajustan a las coordenadas geográficas de la localidad de recolección. Para el resto de las entradas, se eligió la estación climática más próxima con características geográficas similares. Para 34 entradas, se tomaron datos climáticos de estaciones que se encuentran a una distancia menor de 10 km, para otras 35 entradas, de estaciones que se encuentran dentro de un radio de 20 km del sitio de recolección y para otras 13, de estaciones que se localizan a más de 20 km de la zona de la recolección.

Los datos climáticos recogidos de este modo fueron:

- Pluviometría mensual y anual
- Evapotranspiración mensual y anual
- Temperaturas mensuales y promedio de temperatura anual
- Índice de Turc
- Clasificación de Papadakis

El índice de Turc se define como la productividad agronómica de una región, basada en correlaciones entre factores climáticos y la producción sobre un periodo de tiempo largo. Se expresa en toneladas de materia seca por hectárea de una planta adaptada bajo condiciones de cultivo estándar (Ruimy *et al.*, 1996).

La clasificación de Papadakis es una clasificación climática internacional que clasifica los climas según los factores que afectan a la producción de un cultivo, principalmente la temperatura y la humedad (Papadakis, 1975).

A partir de los datos mensuales de pluviometría, temperatura y evapotranspiración, se calcularon las pluviometrías, temperaturas y evapotranspiración de otoño (octubre, noviembre y diciembre), invierno (enero, febrero y marzo) y primavera (abril, mayo y junio), respectivamente. Del mismo modo se calculó una nueva variable: la relación entre pluviometría y evapotranspiración, a la que

denominaremos *ratio*. El *ratio* es el cociente entre la precipitación registrada en una zona considerada en un periodo dado y la evapotranspiración registrada en el mismo periodo.

En la Tabla 4.1 se enumeran las variables geográficas y climáticas analizadas en este estudio.

Tabla 4.1 Variables geográficas y climáticas analizadas.

Variables	Código
Región agro-ecológica	Ragroec
Latitud	LAT
Altitud	ALT
Pluviometría de otoño	PLVoto
Pluviometría de invierno	PLVinv
Pluviometría de primavera	PLVpri
Pluviometría total	PLVtot
Evapotranspiración de otoño	ETPoto
Evapotranspiración de invierno	ETPinv
Evapotranspiración de primavera	ETPpri
Evapotranspiración total	ETPtot
<i>Ratio</i> de otoño	RAToto
<i>Ratio</i> de invierno	RATinv
<i>Ratio</i> de primavera	RATpri
Temperatura de otoño	TMPoto
Temperatura de invierno	TMPinv
Temperatura de primavera	TMPpri
Índice de Turc	Turc
Clasificación de Papadakis	Ppkis

4-2-4. Análisis estadísticos

Las variables climáticas calculadas anteriormente, junto con las variables de pluviometría total, temperatura anual, evapotranspiración anual, altitud y latitud fueron analizadas mediante componentes principales, para reducir el número de variables a introducir en los análisis de regresiones con los marcadores, y evitar la redundancia de información. Este análisis se llevó a cabo mediante el procedimiento *proc princomp* (SAS 1988).

La asociación entre los marcadores moleculares y las variables geográficas y climáticas fue evaluada mediante un análisis de regresión para las variables cuantitativas y análisis de varianza para las variables cualitativas (región agro-ecológica y clasificación de Papadakis). Solamente los alelos presentes en una frecuencia igual o

superior a 5% en las entradas, fueron considerados en el análisis. La información de los marcadores se codificó creando una variable por cada alelo, expresando mediante un “1” la presencia de un alelo en un individuo, y con un “0” su ausencia en el resto de variables. Por ejemplo, un marcador con dos alelos quedaba plenamente descrito con dos variables.

Estos análisis se llevaron a cabo mediante el procedimiento *proc glm* (SAS 1988). Los alelos fueron introducidos en el modelo como variables dependientes mientras que las variables climáticas y geográficas se consideraron como variables independientes. El nivel de significación elegido para el modelo fue de 0,01.

La distribución de alelos de algunos loci que mostraron coeficientes de determinación relativamente elevados en los análisis de regresión con algunas de las variables geográficas o climáticas analizadas, fue representada gráficamente sobre mapas climáticos mediante el programa DIVA-GIS (Hijmans *et al.*, 2005). Estas representaciones gráficas fueron realizadas empleando los datos climáticos y geográficos de la base de datos de dicho programa.

4-3. Resultados

En este capítulo se analiza la relación de la distribución alélica de una serie de marcadores moleculares con la climatología y otros factores geográficos. El análisis de diversidad de 64 marcadores está presentado en la Tabla 3.7 del capítulo 3. Como se indica en el apartado de Material y Métodos, se han incluido ocho marcadores nuevos en los análisis de asociación (capítulos 4 y 6). Como referencia, y para facilitar la interpretación de los análisis de asociación, se incluyen en el anexo 2 los resultados del análisis de diversidad de estos nuevos marcadores.

4-3-1. Análisis en componentes principales de las variables geográficas y climáticas

Las relaciones entre las variables climáticas y geográficas continuas fueron analizadas mediante un análisis de componentes principales, pues se sospechaba de la existencia de multicolinealidad entre estas variables. Los dos primeros ejes resultantes de éste análisis explicaron un 80,5% de la variación. Tan sólo el primer eje del análisis ya explicó un 53% de la variación. Este eje muestra una agrupación de las variables de temperatura y evapotranspiración a un lado, opuestas a las variables de pluviometría de primavera y *ratio*, y a las variables geográficas latitud y altitud (Figura 4.1). El eje 2,

explicó un 27,5% de la variación, y está dominado por el contraste entre las variables de pluviometría, índice de productividad de Turc y *ratio* frente a la altitud y latitud.

Este análisis refleja las variaciones geográficas y climáticas existentes en la Península Ibérica. La agrupación de la altitud y latitud en un mismo cuadrante denota la existencia de una correlación positiva entre estas variables. Al estar excluida del análisis toda la zona costera del Cantábrico, las zonas de mayor altitud corresponden en general a las zonas de mayor latitud (zona de la Meseta Central), y éstas se caracterizan por temperaturas más bajas y una evapotranspiración más moderada, contrariamente a las zonas de altitud baja (en este estudio, fundamentalmente el sur de España y la costa mediterránea). Estas últimas se caracterizan por pluviometrías más elevadas que las zonas de altitudes altas. Las *ratios* son también más elevadas en las zonas de baja altitud, debido a las temperaturas relativamente más altas de estas zonas, que provocan mayor evapotranspiración.

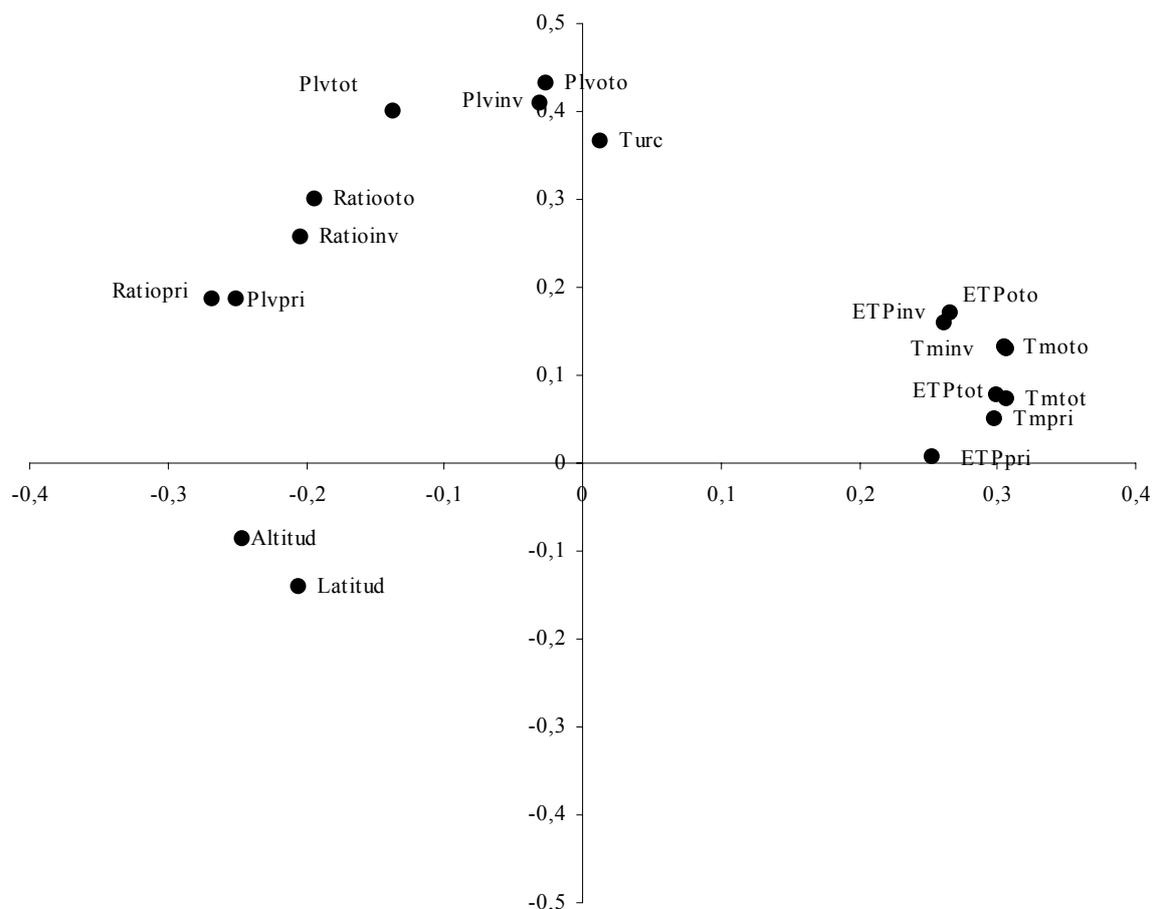


Figura 4.1. Análisis en componentes principales de las variables climáticas y geográficas descritas en el texto, correspondientes a las zonas de recolección de las 159 *landraces* de la colección nuclear.

Este análisis permitió ver que las diferentes variables de pluviometría, temperatura, evapotranspiración y *ratio* correspondientes a las 3 estaciones (otoño, invierno y primavera), presentan respectivamente una correlación positiva entre sí. Esta correlación es más elevada entre las variables correspondientes a los datos de otoño e invierno (Figura 4.1). Los datos de primavera son los más distintos de las tres estaciones para todas las variables, especialmente para la pluviometría. Los datos climáticos que corresponden a los totales anuales se sitúan siempre entre las variables climáticas de invierno y primavera. Por todo lo cual, se eligieron los valores de invierno y primavera como los más representativos para llevar a cabo los análisis de relación de la distribución de los marcadores con las variables climáticas. La inclusión de las variables correspondientes al otoño y los totales anuales en los análisis de regresión con los marcadores solo provocaría una redundancia de información.

En consecuencia, de todas las variables geográficas y climáticas disponibles se consideraron 13 en los análisis de regresión con los marcadores moleculares: latitud, altitud, región agro-ecológica, pluviometría, evapotranspiración, temperatura y *ratio* de invierno y de primavera, índice de Turc y la clasificación de Papadakis.

4-3-2. Regresiones entre alelos y variables geográficas y climáticas

Los 73 marcadores analizados detectaron un total de 651 alelos en las 159 entradas. Solamente 304 estuvieron presentes en una frecuencia igual o superior al 5%, el límite que se fijó para que fueran considerados en los análisis de regresión. Ciento ochenta y seis de estos alelos (61,2%) no mostraron asociación con ninguna de las 13 variables geográficas y climáticas, mientras que ciento dieciocho de estos alelos (38,8%) mostraron por los menos una asociación con alguna de las variables analizadas (Tabla 4.2). En total, fueron detectadas doscientas setenta y nueve asociaciones. El número de asociaciones detectadas por locus varió entre 0 y 12 (Tabla 4.2).

Los loci HvGLUEND, Bmac134, scssr07759, Bmag378, MWG699, HvLTPPB, Bmag136, Bmag006, HvM062, HvM003, HvBTAI3, HvM067, Bmac113, HvBM5A, y Bmag369 son los que más asociaciones presentaron con las variables analizadas. Trece loci no mostraron ninguna asociación y quince loci mostraron asociación con una única variable.

A continuación se presentan los resultados detallados de las asociaciones con cada una de las variables.

Región agro-ecológica

Cincuenta y siete alelos, que se distribuyen entre 37 loci, mostraron asociación con la variable región agro-ecológica. Estos revelaron unos coeficientes de determinación que variaron entre 0,08 y 0,28. Los alelos de los loci HvALAAT, scssr07759, MWG699, HvBM5A y Bmag136 son los que presentaron una asociación relativamente más intensa con esta variable. Los valores R^2 más altos, 0,27 y 0,28, correspondieron a los alelos 4 de scssr07759 y de HvBM5A, respectivamente.

Latitud

Los análisis de regresión detectaron la asociación de 34 alelos, que pertenecen a 26 loci, con esta variable. Los coeficientes de determinación que presentaron estos alelos variaron entre 0,04 y 0,19. Los coeficientes más altos correspondieron al alelo 2 de HvGLUEND (0,13), el alelo 3 de HvLTPPB (0,19) y HvM062 (0,16) y el alelo 4 de HvBM5A (0,13).

Altitud

Veintinueve alelos mostraron asociación con la altitud. Estos se distribuyeron entre 22 loci. La mayoría de estos alelos presentaron coeficientes de determinación que oscilaron entre 0,05 y 0,08. Los alelos que presentaron coeficientes relativamente más altos, entre 0,10 y 0,15, corresponden a los loci MWG699, HvM062, HvBTAI3, Bmac113, HvBM5A y Bmac316 (Tabla 4.3). La representación de la distribución de los alelos A y D de MWG699 con el programa DIVA-GIS (Hijmans *et al.*, 2005) muestra claramente la presencia del alelo A en zonas costeras (alelo de color negro, Figura 4.2A) mientras que el alelo D abunda en las regiones de mayor altitud (alelo de color azul, Figura 4.2A) Lo mismo se observa con los alelos 6 y 9 del microsatélite Bmac316. El primero se encuentra en mayor frecuencia en las zonas de altitudes elevadas (alelo de color azul, Figura 4.2B) mientras que el segundo se distribuye más en regiones de altitudes bajas (alelo de color negro, Figura 4.2B)

Tabla 4.2. Número de alelos detectados en las 159 *landraces* de la colección nuclear, y resultado global de los análisis de regresión de los alelos frecuentes frente a las 13 variables climáticas y geográficas analizadas. El valor del índice F_{ST} entre los dos grupos principales de entradas se ofrece a efectos comparativos.

Locus	N.º alelos	Alelos frecuentes (>5%)	N.º de alelos no significativos	N.º de alelos significativos	N.º asociaciones	F_{ST} Grupos 3 y 4
Bmac399	38	6	3	3	3	0,04
HvGLUEND	5	3	0	3	10	0,45
Bmac032	26	8	7	1	5	0,00
HvM020	6	3	1	2	6	0,23
HvALAAAT	7	6	3	3	7	0,13
Bmag211	8	4	2	2	2	0,03
scsr10477	3	3	1	2	4	0,01
MWG518	2	2	1	1	1	0,08
Bmag382	3	2	2	0	0	-0,02
HvHVA1	2	1	0	1	1	-
WMC1E8	2	2	2	0	0	0,04
Bmac134	12	9	4	5	12	0,27
HvM036	8	4	3	1	1	0,09
MWG858	2	2	0	2	4	0,00
scsr07759	5	3	1	2	10	0,12
Bmag378	3	2	0	2	9	0,20
EBmac640	5	2	1	1	1	-0,01
Bmac132	5	3	2	1	1	-0,01
HvBKASI	3	3	1	2	3	-0,01
Bmac093	5	3	2	1	1	0,03
MWG699	3	3	0	3	11	0,24
Bmag125	10	4	2	2	7	0,15
HvM054	6	4	4	0	0	0,03
HvLTPPB	3	3	1	2	8	0,11
Bmag136	5	4	1	3	9	0,48
Bmac067	14	6	4	2	2	0,13
Bmag006	17	7	2	5	8	0,16
Bmac209	11	6	5	1	1	0,02
Bmag225	15	5	3	2	7	0,04
Hv13GEIII	3	3	0	3	2	0,19
Bmag013	19	7	5	2	4	0,09
HvM062	5	3	1	2	11	0,35
Bmac029	8	3	3	0	0	0,03

Tabla 4.2. (Continuación)

Locus	N.º alelos	Alelos frecuentes (>5%)	N.º de alelos no significativos	N.º de alelos significativos	N.º asociaciones	F _{ST} Grupos 3 y 4
HvM040	6	2	0	2	6	0,28
HvOLE	4	3	0	3	4	0,08
Bmag353	17	7	4	3	3	0,09
HvM003	22	11	9	2	11	0,09
Bmag384	7	5	4	1	2	0,09
HvBTAI3	2	2	0	2	10	0,16
EBmac701	12	5	4	1	4	0,00
HvMLOH1A	3	1	1	0	0	0,01
HvM067	4	3	0	3	8	0,32
HvZCCT	2	2	0	2	5	0,01
HvBAMY	6	4	2	2	1	0,19
HdAMYB	12	5	5	0	0	0,05
Bmag337	12	6	4	2	2	0,01
Bmac113	14	5	3	2	8	0,29
Bmac096	9	5	5	0	0	0,03
EBmac970	3	2	1	1	1	0,07
Bmag223	20	7	6	1	1	0,03
HvHEMH1	2	2	0	2	1	0,21
HvLEU	2	2	2	0	0	0,00
Hv635P24	8	6	4	2	4	-0,01
HvBM5A	4	4	2	2	9	0,44
HvDHN7	3	2	2	0	0	0,00
Bmag222	5	3	3	0	0	0,03
Bmac316	8	5	2	3	9	0,22
Bmag500	20	5	4	1	1	0,00
Bmag173	13	4	4	0	0	0,04
Bmac018	5	4	1	3	3	0,21
Bmag009	5	4	2	2	2	0,05
EBmac806	6	2	0	2	3	0,19
Bmac040	22	8	7	1	1	0,01
Bmag206	21	8	7	1	2	0,03
HvM004	23	7	7	0	0	0,05
HvSS1	5	4	3	1	6	0,00
HvCMA	2	1	0	1	3	-
Bmac064	6	2	2	0	0	0,17
Bmag369	6	3	1	2	9	0,30
Bmag120	17	9	7	2	2	0,04
Bmac156	28	5	3	2	5	0,02
Bmag135	8	4	3	1	1	0,22
HvGLB2	8	6	5	1	1	0,01
Total	651	304	186	118	279	

Pluviometría de invierno

Se detectaron asociaciones entre la pluviometría de invierno y 23 alelos, que se distribuyeron entre 18 loci. Los coeficientes de determinación de estos alelos variaron entre 0,04 y 0,11 (Tabla 4.3). El alelo 16 de Bmac399 y el alelo 6 de Bmag369 son los que mostraron una asociación relativamente más fuerte con esta variable, con unos coeficientes de determinación de 0,10 y 0,11, respectivamente. La Figura 4.2C muestra claramente que las *landraces* que tienen el alelo 16 de Bmac399 se localizan preferentemente en las zonas de mayor pluviometría invernal.

Pluviometría de primavera

Dieciocho alelos, que pertenecen a 15 loci mostraron asociación con la pluviometría de primavera. Los valores de los coeficientes de determinación de estos alelos variaron entre 0,04 y 0,12 (Tabla 4.3). El alelo 5 del locus Bmac134 y el alelo 3 de HvM062 son los que mostraron una asociación relativamente más fuerte con esta variable, con unos valores R^2 de 0,12 y 0,10, respectivamente. La Figura 4.2D muestra la distribución de las entradas que tienen el alelo 5 de Bmac134, que aparece con mayor frecuencia en zonas de mayor pluviometría primaveral.

Evapotranspiración de invierno

Los análisis de regresión detectaron la asociación de 38 alelos, que pertenecen a 25 loci, con esta variable. Los coeficientes de determinación de estos alelos oscilaron entre 0,04 y 0,28 (Tabla 4.3). Alelos de los loci HvGLUEND, HvALAAT, HvBTAI3 y HvBM5A son los que mostraron una asociación relativamente más fuerte con esta variable, con unos valores R^2 entre 0,18 y 0,28.

Tabla 4.3. Coeficientes de determinación de los análisis de regresión entre alelos (variables dependientes) y variables geograficas y climáticas (variables independientes).

Cr	Alelos	Tamaño (pb)	N. de entradas	Ragroec	LAT	ALT	PLVinv	PLVpri	ETPinv	ETPpri	RATinv	RATpri	TMPinv	TMPpri	TURC	Ppkis
1H	Bmac399-8	131	9													0,26***
	Bmac399-16	141	23				0,10***								0,08***	
	Bmac399-25	151	8			0,05**										
	HvGLUEND-2	260	68	0,18***	0,13***	0,06**		0,06**	0,18***	0,06**		0,06**	0,20***	0,13***		0,23***
	HvGLUEND-4	262	61						0,07***				0,06**			
	HvGLUEND-5	264	24	0,09**					0,04**				0,06**	0,04**		
	Bmac032-5	217	13		0,04**	0,05**			0,04**				0,06**	0,06**		
	HvM20-3	151	53	0,12***	0,06**		0,05**						0,05**	0,05**	0,06**	
	HvM20-4	153	86				0,04***								0,10***	
	HvALAAT-2	198	14	0,23***	0,16***	0,05**			0,18***				0,12***	0,04**		0,23***
	HvALAAT-3	199	27	0,09**												
	HvALAAT-5	202	66	0,17***		0,05**			0,07**				0,07***			0,15**
	Bmag211-7	186	42	0,10**												
	Bmag211-8	188	10													
	scssr10477-1	126	131											0,05**	0,04**	
	scssr10477-5	nulo	18		0,05**				0,05**					0,07***	0,07**	
MWG518-1		36		0,09**												
HvHVA1-1	151	157				0,05**				0,14***						
2H	Bmac134-2	130	70												0,06**	
	Bmac134-5	148	11		0,05**			0,12***	0,04**		0,07**	0,08***	0,06**			
	Bmac134-9	158	11			0,05**			0,05**	0,05**			0,08***	0,07**		0,21***
	Bmac134-10	162	8													0,14**
	Bmac134-13	172	10	0,15***						0,08***				0,05**		
	HVM36-7	136	16	0,11***												
	MWG858-1		51		0,08***				0,05**				0,08***	0,06**		
	MWG858-2		108		0,08***				0,05**				0,08***	0,06**		
	scssr07759-4	213	39	0,27***	0,09***	0,08***	0,07***		0,10***	0,08***			0,15***	0,12***	0,05**	0,22**
	scssr07759-5	221	108	0,16***			0,07**		0,06**				0,08***	0,06**	0,05**	0,13**
	Bmag378-2	137	51	0,14***	0,06**	0,05**				0,08***				0,10***		0,19***
	Bmag378-3	139	107	0,15***	0,06**	0,06***		0,05**	0,05**	0,08***			0,09***	0,11***		0,20***
	EBmac640-1	174	125								0,05**					
	Bmac132-3	189	31										0,10***			
	HvBKASI-3	199	105				0,05**									
	HvBKASI-4	200	32	0,09**	0,06**											
	Bmac093-1	153	12	0,13***												
	MWG699-A		63	0,22***	0,09***	0,10***	0,05**			0,11***				0,12***	0,05**	0,07***
MWG699-D		87	0,14***		0,06**	0,05**			0,06**				0,06**		0,10***	0,16***
MWG699-K		9						0,07***				0,05**	0,05**			
Bmag125-5	134	39			0,06**	0,04**			0,06**				0,06**	0,05**	0,07***	
Bmag125-7	138	43		0,06**		0,06**										

** P<0,01; *** P<0,001.

Evapotranspiración de primavera

Dieciséis alelos que se distribuyen entre 11 loci mostraron asociación con esta variable, con unos coeficientes de determinación que variaron entre 0,04 y 0,11 (Tabla 4.3). Los alelos del locus HvBTAI3 son los que presentaron una asociación relativamente más intensa con esta variable.

Ratio de invierno

Doce alelos de 11 loci mostraron asociación con esta variable. Los coeficientes de determinación fueron bajos, variando en su mayoría entre 0,04 y 0,09 (Tabla 4.3). El alelo 1 del locus HvHVA1 es el único que mostró un R^2 relativamente más alto, de 0,14.

Ratio de Primavera

Quince alelos, de 12 loci mostraron asociación con esta variable. Los coeficientes de determinación que mostraron estos alelos variaron entre 0,04 y 0,08 (Tabla 4.3). Este último valor corresponde al alelo 5 del locus Bmac134 y al alelo 7 del locus Bmag225.

Los alelos 5 de Bmac134, 11 de HvM003 y 1y 2 de HvZCCT son los únicos que mostraron a la vez asociación con la *ratio* de invierno y primavera.

Temperatura de invierno

Los análisis de regresión detectaron la asociación de 51 alelos con la temperatura de invierno. Estos alelos se distribuyen entre 32 loci. Los coeficientes de determinación que mostraron estos alelos variaron entre 0,04 y 0,23 (Tabla 4.3). El alelo 2 de HvGLUEND, los 2 alelos de HvBTAI3 y el alelo 4 de HvBM5A fueron los que presentaron una asociación más fuerte con esta variable, con unos valores R^2 de 0,20, 0,21 y 0,23 respectivamente. La representación grafica de la distribución de los alelos del locus HvBM5A, en la Península Ibérica muestra claramente la abundancia del alelo 4 (alelo de color negro, Figura 4.3B) en las zonas de temperaturas relativamente altas mientras que el alelo 2 (alelo de color azul, Figura 4.3B) tiende a extenderse en las regiones de temperaturas más bajas. También se representaron la distribución de los alelos 3 y 4 del locus HvM062, que presentaron unos valores R^2 de 0,14 y 0,15 respectivamente. El alelo 4 (alelo de color azul, Figura 4.3A) es más frecuente en las zonas de bajas temperaturas mientras que el alelo 3 (alelo de color negro, Figura 4.3A) se encuentra más en las regiones de temperaturas relativamente más altas.

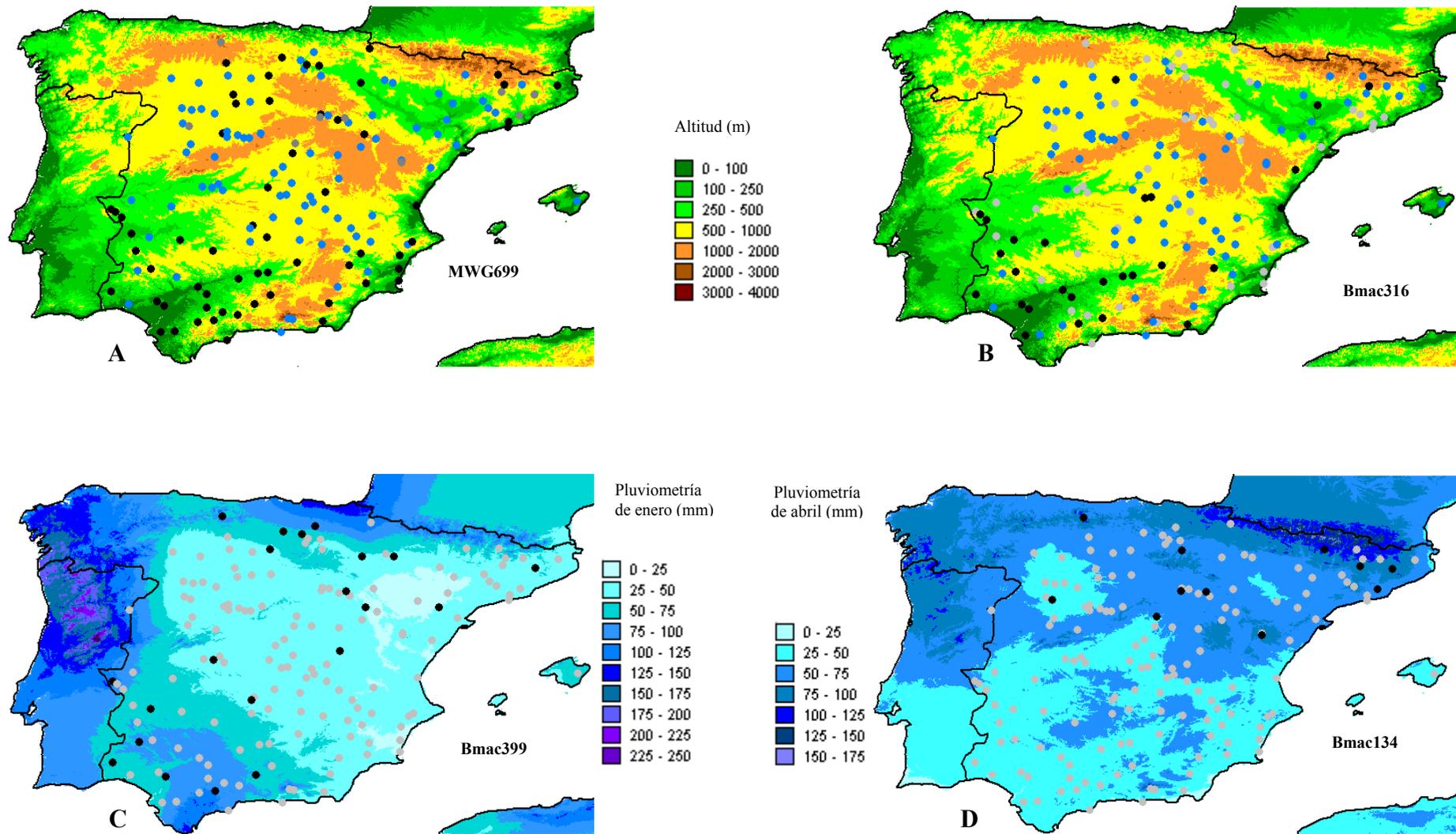


Figura 4.2. Representación de la distribución de los alelos: A) MWG699-A (negro), MWG699-D (azul); B) Bmac316-6 (azul), Bmac316-9 (negro) según las variaciones de altitud y la representación de la distribución de los alelos: C) Bmac399-16 (negro), D) Bmac134-5 (negro) según las variaciones respectivamente de la pluviometría de invierno (enero) y de primavera (abril). Los alelos de color gris en las Figuras A, B, C y D, son los que no presentaron asociación con la variable considerada o son alelos raros que no entraron en el análisis

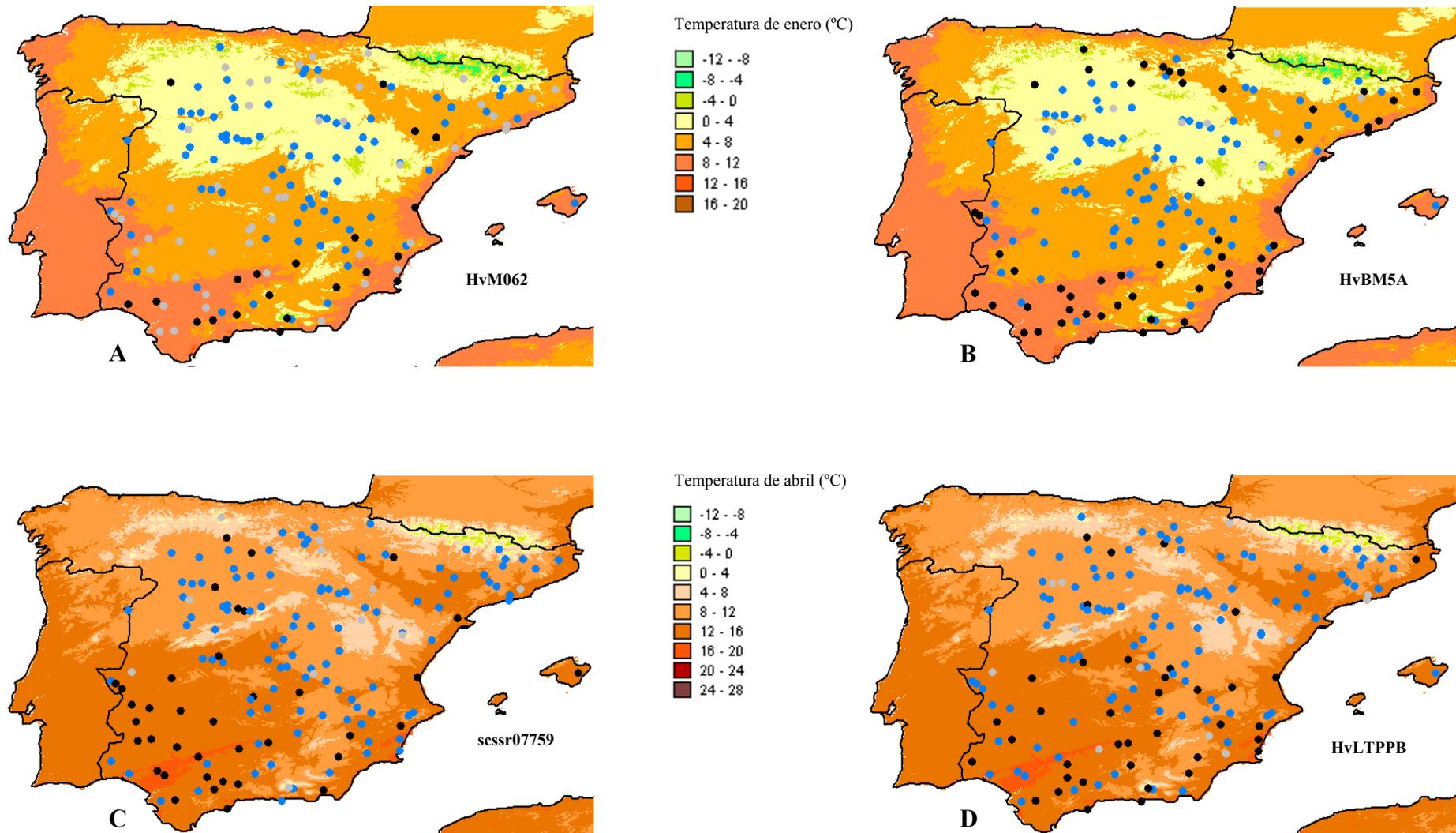


Figura 4.3. Representación de la distribución de los alelos: A) HvM062-3 (negro), HvM062-4 (azul), B) HvBM5A-2 (azul) HvBM5A-4 (negro) según las variaciones de temperaturas de invierno (enero) y la representación de la distribución de los alelos: C) scssr07759-4 (negro), scssr07759-5 (azul), D) HvLTPPB-3 (negro) y HvLTPPB-2 (azul) según las variaciones de temperaturas de primavera (abril). Los alelos de color gris en las figuras A, C y D, son los que no presentaron asociación con la variable considerada o son alelos raros que no entraron en el análisis

Temperatura de primavera

Treinta y ocho alelos, que se distribuyen entre 26 loci, mostraron asociación con la temperatura de primavera. Estos mostraron coeficientes de determinación que oscilaron entre 0,04 y 0,15 (Tabla 4.2). Este último valor corresponde a los alelos 1 y 2 del locus HvBTAI3. El alelo 2 de HvGLUEND y el alelo 4 de scssr07759 también mostraron coeficientes de determinación relativamente altos, de 0,13 y 0,12 respectivamente. La distribución de los alelos 4 y 5 del locus scssr07759 fue representada con respecto a las variaciones de la temperatura de primavera en la Península Ibérica. La Figura 4.3C muestra la abundancia del alelo 4 (color negro) en las zonas de altas temperaturas, mientras que el alelo 5 (color azul) es más frecuente en zonas de temperaturas relativamente más bajas. También se representaron las distribuciones de los alelos 2 y 3 del locus HvLTPPB, que mostraron un R^2 de 0,08. La Figura 4.3D muestra claramente la abundancia del alelo 3 (color negro) en las zonas de altas temperaturas, contrariamente al alelo 2 (color azul) que tiende a estar más presente en regiones de temperaturas relativamente más bajas.

De los alelos que mostraron asociación con temperatura de invierno y primavera, 34 fueron comunes entre las dos variables.

Índice de Turc

Los análisis de regresión mostraron la asociación de 30 alelos con el índice de Turc, distribuidos en 20 loci. Los valores de los coeficientes de determinación variaron entre 0,04 y 0,17 (Tabla 4.2). Los alelos 2 y 4 de HvBM5A y el alelo 6 de Bmag369 fueron los que presentaron valores R^2 relativamente más altos, de 0,12, 0,14 y 0,17, respectivamente. El alelo 2 de HvBM5A se distribuye en zonas donde el índice de Turc es más bajo contrariamente el alelo 4. También el alelo 6 de Bmag369 está más presente en regiones de más alta productividad.

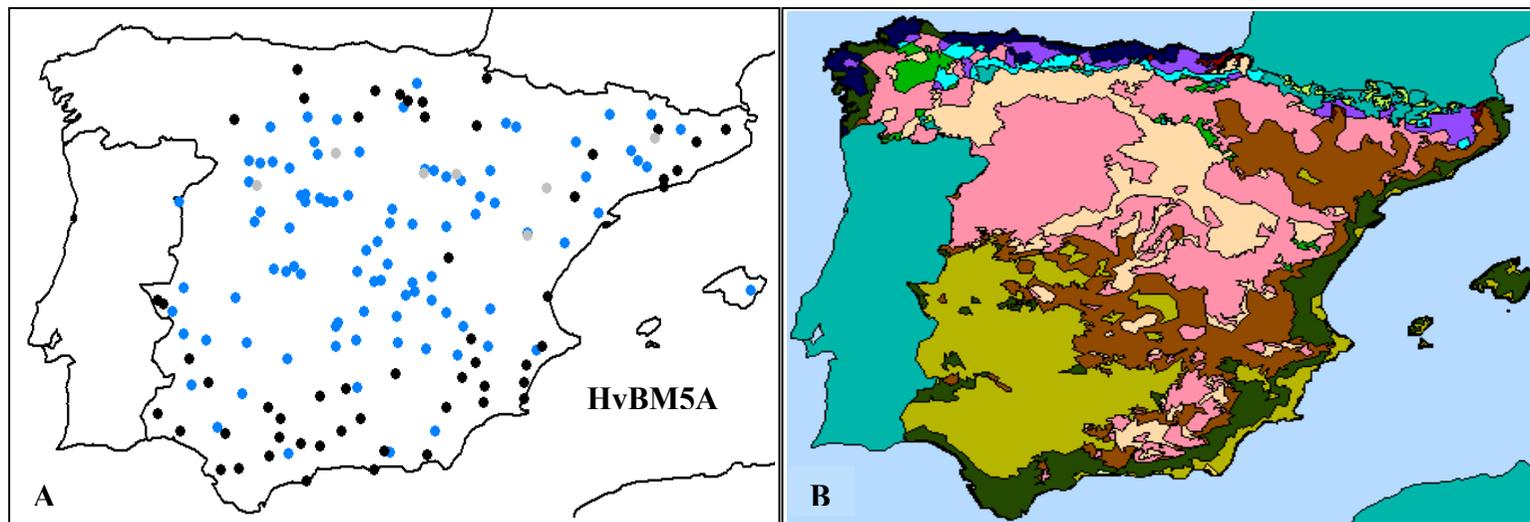
Clasificación de Papadakis

Treinta y ocho alelos distribuidos entre 26 loci presentaron asociación con la clasificación de Papadakis. Estos mostraron unos coeficientes de determinación que variaron de 0,12 a 0,29 (Tabla 4.3). Los alelos 7 de Bmag225, 8 de Bmac399 y 4 de HvBM5A son los que mostraron valores R^2 más altos, de 0,25, 0,26 y 0,29, respectivamente. En la Figura 4.4A, se representan la distribución de los alelos 2 y 4 del

locus HvBM5A; la superposición de la Figura 4.4A sobre la Figura 4.4B, mostraría que el alelo 4 (color negro) está más distribuido en las zonas de clima tipo mediterráneo marítimo a mediterráneo subtropical mientras que el alelo 2 (color azul) se extiende más en las regiones caracterizadas por un clima mediterráneo templado, mediterráneo templado fresco a mediterráneo continental.

4-3-3. Relación entre número de asociaciones detectadas por loci y los F_{ST} locus por locus de comparación entre los grupos 3 y 4 determinados por STRUCTURE

Las distribuciones de los alelos según las variaciones climáticas o geográficas parecen seguir la distribución de los grupos o poblaciones 3 y 4, definidos por similitud genética en el apartado 3-3-3 del capítulo 3. Para comprobarlo, se compararon los resultados de la regresión entre marcadores moleculares y caracteres climáticos y geográficos, con los de un análisis molecular de varianza locus por locus para los grupos 3 y 4. El análisis molecular de varianza (AMOVA, Arlequín software, Shneider *et al.*, 2000) produce un índice, F_{ST} , que indica la divergencia entre poblaciones para cada locus. La comparación entre los índices estadísticos F_{ST} resultantes de este análisis y el número de asociaciones detectados por locus con variables climáticas y geográficas permitió ver que, en la mayoría de los casos, los loci que presentaron un F_{ST} mayor son los que detectan también un número mayor de asociaciones con las variables climáticas y geográficas (Tabla 4.2). De hecho el coeficiente de correlación entre los F_{ST} y el número de asociaciones es de 0,64.



- Mediterráneo templado
- Mediterráneo templado fresco
- Mediterráneo subtropical
- Mediterráneo marítimo
- Mediterráneo continental

Figura 4.4. A) Representación de la distribución de los alelos 2 (color azul) y 4 (color negro) del locus HvBM5A en la Península Ibérica, que presentaron asociación con la clasificación climática de Papadakis. Los puntos de color gris representan los alelos no asociados. B) Diferentes tipos de clima que se encuentran en la Península Ibérica según la clasificación climática de Papadakis.

4-4. Discusión

Diferentes estudios llevados a cabo en *H. spontaneum* han mostrado que la diversidad encontrada en esta especie no está distribuida de forma aleatoria, sino que se encuentra influida por parámetros eco-geográficos (Baek *et al.*, 2003, Chalmers *et al.*, 1992, Huang *et al.*, 2002, Ivandic *et al.*, 2002, Kalendar *et al.*, 2000, Liviero *et al.*, 2002, Morrell *et al.*, 2005, Nevo *et al.*, 2005, Supronova *et al.*, 2004 y Turpeinen *et al.*, 2001). En el presente estudio, los análisis de regresión y de varianza realizados entre los marcadores moleculares y las variables climáticas y geográficas sugieren también que la diversidad genética encontrada en la colección nuclear española tampoco está distribuida de forma aleatoria en la Península Ibérica.

El número más alto de asociaciones entre marcadores y caracteres geográficos y climáticos ocurrió con la variable región agro-ecológica. Esto se explicaría por el hecho de que esta variable integra la información de diferentes factores ambientales (climáticos, edáficos y geográficos). Aunque los coeficientes de determinación que mostraron los análisis con esta variable (y con la clasificación climática de Papadakis), fueron altos en comparación con los demás factores analizados, estos no son directamente comparables, por haberse realizado con dos procedimientos distintos (regresión para las variables continuas y análisis de varianza para las discretas). La clasificación climática de Papadakis integra también diferentes parámetros climáticos, lo que explicaría la similitud de los coeficientes de determinación observados entre esta variable y la variable región agro-ecológica, aunque el número de asociaciones detectadas entre los alelos y la variable clasificación climática de Papadakis es menor que el observado con la variable región agro-ecológica.

En cuanto a las demás variables climáticas y geográficas sencillas, se observaron muchas más asociaciones con las temperaturas que con las pluviometrías, las evapotranspiraciones o las *ratios* descritas. También se observaron (excepto para las *ratios*) un mayor número de asociaciones significativas para las variables de invierno que para las de primavera. De modo general, las variables con más asociaciones fueron también las que presentaron valores de R^2 más elevados para alelos individuales.

En general, estos resultados parecen indicar que la distribución de la diversidad genética es el resultado de procesos de adaptación a una variedad de factores ambientales, que aparecen mejor descritos en clasificaciones integradoras como la de regiones agro-ecológicas o climas. El factor individual que más parece haber influido en

la distribución de la diversidad genética es la temperatura, especialmente la invernal. Las asociaciones con altitud y latitud pueden ser en realidad reflejo de las asociaciones con temperatura (de hecho, el 81 % de las asociaciones de marcadores con latitud y altitud y con la temperatura de invierno son comunes). Esto se debería a la correlación negativa existente entre temperaturas y altitud y latitud en estos datos.

El alelo 4 (1200 pb) del locus HvBM5A fue el que mostró el coeficiente de determinación más alto con la variable temperatura de invierno. Este alelo es más frecuente en el sur de la Península Ibérica, y se ha encontrado también en las variedades españolas de invierno Albacete y Pané (von Zitzewitz *et al.*, 2005). Aunque estas variedades están clasificadas agrónomicamente como de invierno, sus requerimientos de vernalización son menores que los genotipos europeos típicos de invierno, como Igri. Esto explicaría la abundancia de este alelo en los genotipos de la colección nuclear provenientes de regiones donde las temperaturas de invierno no son muy bajas (Figura 4.3B). El alelo 2 (4850 pb) del locus HvBM5A también reveló una asociación con la temperatura de invierno. Este alelo es más frecuente en genotipos de zonas caracterizadas por temperaturas invernales bajas. Fu *et al.* (2005) y von Zitzewitz *et al.* (2005) describieron que los genotipos de primavera presentaban grandes deleciones en el primer intrón del locus HvBM5A. Los genotipos de invierno se caracterizaban por la presencia de un fragmento de 5,2 kb, mientras que en los genotipos de primavera bien se amplificaba un producto de 150 pb (Harrington o Morex), o bien se observaba un genotipo nulo, que podía ser la consecuencia de una deleción mayor. Al analizar la colección nuclear no se ha encontrado el alelo de 5,2 kb, sino que aparece un alelo ligeramente menor, de 4850 pb. Por el tamaño que tiene este alelo, se podría especular que se trata de genotipos de invierno, por lo cual se explicaría su abundancia en las entradas provenientes de regiones de bajas temperaturas invernales de la Península Ibérica. Este alelo mostró asociación con 8 de las variables analizadas y se caracterizó por valores R^2 relativamente altos. El locus HvBTAI3, que es el gen inhibidor de α -amilasa, también mostró muchas asociaciones con las variables geográficas y climáticas analizadas, y presentó valores R^2 relativamente altos. Esto parece indicar que el ambiente influye en la actividad de este gen. Algo similar fue observado por Chalmers *et al.* (1992) con la enzima β -amilasa en poblaciones de *H. spontaneum* de Israel. Estos autores encontraron 7 isoformas de esta enzima, una de ellas restringida a localidades desérticas. El 78% de la variación de la frecuencia de esta isoforma fue explicada por la pluviometría y temperatura del sitio de origen.

Aparte de los loci HvBM5A y HvBTAI3, los alelos de los loci HvGLUEND, Bmac134, scssr07759, Bmag378, MWG699, HvLTPPB, Bmag136, Bmag006, HvM062, HvM003, HvM067, Bmac113, Bmac316 y Bmag369 son los que mostraron un número importante de asociaciones con las variables analizadas (8 a 12). También se observaron asociaciones para alelos de otros loci, pero con menos variables. Algunas de las asociaciones encontradas en este estudio coinciden con las observadas en otros trabajos. Turpeinen *et al.* (2001) mostraron que la diversidad encontrada por los microsatélites en *H. spontaneum* de Israel está asociada a factores ecológicos. Estos autores establecieron correlaciones entre la temperatura y el locus HvM036; la evaporación y los loci HvBKASI, HvM062, HvM067, BMS90 y HvM014, o la humedad y los loci HvBKASI, HvM062, HvM067 y HvM014. También encontraron correlaciones entre la latitud y los microsatélites HvM036 y HvM014. Ivandic *et al.* (2002), encontraron asociaciones entre la temperatura y los loci Bmag211, WMC1E8, EBmac970, EBmac806, HvM036, Bmag378, HvLTPPB, Bmac029, HvLDEX y Bmac273; la disponibilidad de agua y los loci Bmag211, Bmag378, Bmac029, Bmac273, Bmag120, HvM040 y HvM003 y por último entre localidades geográficas y los loci Bmac399, HvM054, Bmag009 y Bmac040. Baek *et al.* (2003) encontraron correlaciones entre los loci BMS64, HvBKASI, HvCSG, HvM062, HvM067 y HvM068 y la altitud, que a su vez está correlacionada con la temperatura, en poblaciones de *H. spontaneum* de Jordania. También, encontraron asociaciones entre los loci HvM036 y HvM006 y las variables temperatura, disponibilidad de agua y latitud. En el presente estudio, los loci HvM062 y HvM067 mostraron asociación con la evapotranspiración, temperatura y altitud, lo que coincide con las correlaciones observadas por Turpeinen *et al.* (2001) y Baek *et al.* (2003). También se observó en este estudio la asociación del locus HvM003 (intrón, ‘Rubisco activase’) con la pluviometría, lo cual concuerda con los resultados de Forster *et al.* (1997) e Ivandic *et al.* (2002). Otros resultados encontrados en el presente estudio, análogos a los encontrados por Ivandic *et al.* (2002), son las asociaciones del locus Bmag378 con la pluviometría y temperatura y la del locus HvLTPPB con esta última. Por otra parte, la asociación observada entre el locus Bmag369 y la altitud concuerda con la correlación encontrada por Baek *et al.* (2003) entre el locus BMS64 (otra denominación de Bmac064), marcador contiguo a Bmag369, y dicha variable.

La mayoría de los loci que mostraron muchas asociaciones con las variables climáticas y geográficas coinciden con loci que presentaron valores de F_{ST} altos, al

comparar los grupos genéticos 3 y 4 deducidos con el programa STRUCTURE (Tabla 4.2). Esto parece indicar que la adaptación a los factores eco-geográficos sería una de las principales causas de la estructura de población observada en las entradas de la colección nuclear española. Otras investigaciones llevadas a cabo con cebada han mostrado también una relación entre distribución de la diversidad genética y factores ambientales. En este sentido, Huang *et al.* (2002), analizando 54 genotipos de *H. spontaneum* originarios de 4 micro-hábitats, afirmaron que la frecuencia alélica fue significativamente diferente entre los 4 micro-hábitats. Estos autores comentaron que la distribución de los alelos aparecía estructurada según factores de estrés ecológico (climáticos y edáficos). Por otra parte, Russell *et al.* (2003) estudiaron la diversidad existente en una colección de *landraces* de cebada provenientes de diferentes regiones eco-geográficas de Siria y Jordania con microsatélites nucleares y cloroplásticos. Estos autores mostraron que los alelos frecuentes de los loci Bmac399, HvM020, Bmac134, Bmag125, Bmag013, HvOLE, HvM003, HvM067, Bmac113, Bmac316, Bmac018 y Bmac156 fueron distintos de una población a otra. Igualmente, al analizar 25 genotipos de *H. spontaneum* originarios de todo el rango geográfico de la cebada silvestre con 18 genes nucleares, Morrell *et al.* (2005) encontraron que 11 de los 18 loci demostraron una estructura de la diversidad haplotípica dependiente de la geografía.

La aparición de una estructura de población asociada a un gradiente eco-geográfico no implica que todos los marcadores que presentan correlaciones con variables climáticas o geográficas indiquen la presencia de QTLs de adaptación en las zonas vecinas. Cualquier marcador cuyos alelos tengan una distribución desigualada entre las poblaciones aparecerá asociado con las variables eco-geográficas. Este podría ser el caso de algunos de los loci que mostraron una asociación con las variables climáticas y geográficas en la colección nuclear. Este fenómeno ya ha sido observado por Zhang *et al.* (1990). Estos autores compararon 101 entradas de *H. spontaneum* procedentes de 10 colecciones representativas de cinco hábitats de la especie en Israel y observaron que las cinco regiones difieren considerablemente entre sí en composición y grado de variabilidad genética. Para explicar este fenómeno, los autores propusieron la hipótesis de que este grado de diferenciación entre regiones geográficas y taxones (*H. vulgare* y *H. vulgare* ssp. *spontaneum*) indicaría que la domesticación y la adaptación subsiguiente han ocurrido no como un proceso de selección de alelos en loci individuales, sino que esos procesos habrían reorganizado el genoma en haplotipos multilocus adaptados a los distintos ambientes. En este mismo sentido, Allard (1999)

indicó que es esperable que la autogamia sea especialmente eficaz en la reorganización de alelos favorables en haplotipos que permitan un comportamiento favorable en respuesta al ambiente, aunque los alelos se encuentren en cromosomas distintos. Los resultados encontrados en este estudio de la colección nuclear española concuerdan con estas hipótesis.

La hipótesis de la existencia de orígenes distintos en las cebadas españolas fue propuesta en el apartado 3-4-4 del capítulo 3, por la distribución que presentaban los haplotipos A y D del STS MWG699 en los grupos genéticos 3 y 4. La aparición de numerosas asociaciones entre los tipos A y D del STS MWG699 con las variables climáticas y geográficas parece apoyar esta hipótesis. Un posible doble origen de las dos principales poblaciones españolas de cebada explicaría en parte la aparición de diferentes loci asociados a las variables geográficas y climáticas consideradas, como consecuencia de un proceso de domesticación y adaptación divergente de los dos grupos 3 y 4. La distribución desequilibrada observada entre los tipos A y D del STS MWG699 (Figura 4.2A) y los alelos 2 y 4 del locus HvBM5A, gen candidato de respuesta a la vernalización (Figura 4.3B), parece apoyar esta idea. Una gran proporción de los genotipos de tipo D de MWG699, tienen el alelo 2 (4850 pb) del locus HvBM5A (72 entradas de 87), mientras que las entradas que tienen el tipo A de MWG699, presentan en su mayoría el alelo 4 (1200 pb) de HvBM5A (42 entradas de 63). Esto sugiere que durante el proceso de domesticación de la cebada, el requerimiento de vernalización que presentan los genotipos fue uno de los factores principales que influyó en su distribución en la Península Ibérica.

4-5. Referencias bibliográficas

- Allard RW. 1999.** History of plant population genetics. *Annual Review of Genetics*, 33: 1-27.
- Baek HJ, Beharav A, Nevo E. 2003.** Ecological-genomic diversity of microsatellites in wild barley, *Hordeum spontaneum*, populations in Jordan. *Theoretical & Applied Genetics*, 106: 397-410.
- Baum BR, Nevo E, Johnson DA, Beiles A. 1997.** Genetic diversity in wild barley, (*Hordeum spontaneum* C. Koch), in the Near East: A molecular analysis using random amplified polymorphic DNA (RAPD) markers. *Genetic Resources and Crop Evolution*, 44: 147-157.
- Bothmer R von, Jacobson N, Baden C, Jørgensen RBI, Linde-Laursen I. 1995.** An ecogeographical study of the genus *Hordeum*. *Systematic and Ecogeographic*

Studies on Crop Genepools, 7 International Plant Genetic Resources Institute, Rome, 2nd ed.

- Boyd WJR, Li CD, Grime CR, Cakir M, Potipibool S, Keveeta L, Men S, Jalal Kamali MR, Barr AR, Moody DB, Lance RCM, Logue SJ, Raman H, Read BJ. 2003.** Conventional and molecular genetic analysis of factors contributing to variation in the timing of heading among spring barley (*Hordeum vulgare* L.) genotypes grown over a mild winter growing season. *Australian Journal of Agricultural Research*, 54: 1277-1301.
- Casas AM, Yahiaoui S, Ciudad F, Igartua E. 2005.** Distribution of MWG699 polymorphism in Spanish European barleys. *Genome*, 48: 41-45.
- Ceccarelli S. 1994.** Specific adaptation and breeding for marginal conditions. *Euphytica*, 77: 205-219.
- Chalmers KJ, Waugh R, Waters J, Forster BP, Nevo E, Abbott RJ, Powell W. 1992.** Grain isozyme and ribosomal DNA variability in *Hordeum spontaneum* populations from Israel. *Theoretical & Applied Genetics*, 84: 313-322.
- Forster BP, Russell JR, Ellis RP, Handley LL, Robinson D, Hackett CA, Nevo E, Waugh R, Gordon DC, Keith R, Powell W. 1997.** Locating genotypes and genes for abiotic stress tolerance in barley: A strategy using maps, markers and wild species. *New Phytologist*, 137: 141-147.
- Fu D, Szűcs P, Yan L, Helguera M, Skinner JS, von Zitzewitz J, Hayes PM, Dubcovsky J. 2005.** Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics*, 273: 54-65.
- Hijmans RJ, Guarino L, Jarvis A, O'Brien R, Mathur P, Bussink C, Cruz M, Barrantes I, Rojas E. 2005.** DIVA-GIS, version 5.2. Manual. <http://www.diva-gis.org/>
- Huang Q, Beharav A, Li Y, Kirzhner V, Nevo E. 2002.** Mosaic microecological differential stress causes adaptive microsatellite divergence in wild barley, *Hordeum spontaneum*, at Neve Yaar, Israel. *Genome* 45: 1216-1229.
- Igartua E, Gracia MP, Lasa JM, Medina B, Molina-Cano JL, Montoya JL, Romagosa I. 1998.** The Spanish barley core collection. *Genetic Resources and Crop Evolution*, 45: 475-481.
- Ivandic V, Hackett CA, Nevo E, Keith R, Thomas WTB, Forster BP. 2002.** Analysis of simple sequence repeats (SSRs) in wild barley from the Fertile Crescent: Associations with ecology, geography and flowering time. *Plant Molecular Biology*, 48: 511-527.
- Ivandic V, Thomas WTB, Nevo E, Zhang Z, Forster BP. 2003.** Associations of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. *Plant Breeding*, 122: 300-304.
- Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH. 2000.** Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon

dynamics in response to sharp microclimatic divergence. Proceedings of the National Academy of Sciences, USA, 97: 6603-6607.

- Liviero L, Maestri E, Gulli M, Nevo E, Marmioli N. 2002.** Ecogeographic adaptation and genetic variation in wild barley, application of molecular markers targeted to environmentally regulated genes. *Genetic Resources and Crop Evolution*, 49: 133-144.
- Moralejo M, Swanaton JS, Muñoz P, Prada D, Elía M, Russell JR, Ramsay L, Cistué L, Codesal P, Casas AM, Romagosa I, Powell W, Molina-Cano JL. 2004.** Use of new EST markers to elucidate the genetic differences in grain protein content between European and North American two-rowed malting barleys. *Theoretical & Applied Genetics*, 110: 116-125.
- Morrell PT, Toleno DM, Lundy KE, Clegg MT. 2005.** Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite high rates of self-fertilization. Proceedings of the National Academy of Sciences, USA, 102: 2442-2447.
- Nevo E. 1992.** Origin, Evolution, Population genetics and resources of wild barley, *Hordeum spontaneum*, in the Fertile Crescent. In Ed. Shewry PA, Barley. Genetics, Biochemistry, Molecular Biology and Biotechnology. CAB International, Wallingford, UK. 19-43
- Nevo E, Zohary D, Brown AHD, Haber M. 1979.** Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum*, in Israel. *Evolution*, 33: 815-833.
- Nevo E, Beiles, A, Zohary D. 1986.** Genetic resources of wild barley in the Near East: Structure, evolution and application in breeding. *Biological Journal of the Linnean Society*, 27: 355-380.
- Nevo E, Apelbaum- Elkaher I, Garty J, Beiles A, Owuor EW, Fahima T, Korol A. 1997.** Natural selection causes microscale allozyme diversity in wild barley and a lichen at 'Evolution Canyon', Mt. Carmel, Israel. *Heredity*, 78: 373-382.
- Nevo E, Baum B, Beiles A, Jonson DA. 1998.** Ecological correlates of RAPD DNA diversity of wild barley, *Hordeum spontaneum*, in the Fertile Crescent. *Genetic Resources and Crop Evolution*, 45: 151-159.
- Nevo E, Beharav A, Meyer RC, Hackett CA, Forster BP, Russell JR, Powell W. 2005.** Genomic microsatellites adaptive divergence of wild barley by microclimatic stress in 'Evolution Canyon' Israel. *Biological Journal of the Linnean Society*, 84: 205-224.
- Owuor ED, Fahima T, Belies A, Korol AB, Nevo E. 1997.** Population genetics response to microsite ecological stress in wild barley, *Hordeum spontaneum*. *Molecular Ecology*, 6: 1177-1187.
- Owuor ED, Beharav A, Fahima T, Kirzhner VM, Korol AB, Nevo E. 2003.** Microscale ecological stress causes RAPD molecular selection in wild barley, Neve Yaar microsite, Israel. *Genetic Resources and Crop Evolution*, 50: 213-224.

- Papadakis J. 1975.** Climates of the world and their agricultural potentialities. Libro de Edición Argentina, Buenos Aires. 200p.
- Ramsay L, Macaulay M, degli Ivanissevich S, MacLean K, Cardle L, Fuller J, Edwards KJ, Tuveesson S, Morgante M, Massari A, Maestri E, Marmioli N, Sjakste T, Ganal M, Powell W, Waugh R. 2000.** A simple sequence repeat-based linkage map of barley. *Genetics* 156:1997-2005.
- Ruimy A, Dedieu G, Saugier B. 1996.** TURC: A diagnostic model of continental gross primary productivity and net primary productivity. *Global Biogeochemical Cycles*, 10: 269-285.
- Russell JR, Booth A, Fuller JD, Baum M, Ceccarelli S, Grando S, Powell W. 2003.** Patterns of polymorphism detected in the chloroplast and nuclear genomes of barley landraces sampled from Syria and Jordan. *Theoretical & Applied Genetics*, 107: 413-421.
- SAS Institute. 1988.** SAS/STATTM User's Guide, Release 6.03. Edition. Cary, NC: SAS Institute Inc., 1028p.
- Saghai Maroof MA, Allard RW, Zhang Q. 1990.** Genetic diversity and ecogeographical differentiation among ribosomal DNA alleles in wild and cultivated barley. *Proceedings of the National Academy of Sciences, USA*, 87: 8486-8490.
- Schneider S, Roessli D, Excoffier L. 2000.** Arlequin ver2.000. A software for population genetics data analysis. *Genetics and Biometry Laboratory, University of Geneva, Switzerland.*
- Supronova T, Krugman T, Fahima T, Chen G, Shams A, Korol A, Nevo E. 2004.** Differential expression of dehydrin genes in wild barley, *Hordeum spontaneum*, associated with resistance to water deficit. *Plant, Cell and Environment*, 27: 1297-1308.
- Turpeinen T, Tenhola T, Manninen O, Nevo E, Nissilä E. 2001.** Microsatellite diversity associated with ecological factors in *Hordeum spontaneum* populations in Israel. *Molecular Ecology*, 10:1577-1591.
- von Zitzewitz J, Szűcs P, Dubcovsky J, Yan L, Francia E, Pecchioni N, Casas A, Chen THH, Hayes PM, Skinner JS. 2005.** Molecular and structural characterization of barley vernalization genes. *Plant Molecular Biology*, 59: 449-467.
- Zhang Q, Saghai-Maroof MA, Allard RW. 1990.** Effects on adaptedness of variations in ribosomal DNA copy number in population of wild barley (*Hordeum vulgare* ssp. *spontaneum*). *Proceedings of the National Academy of Sciences, USA*, 87: 8741-8745.
- Zhang Q, Saghai-Maroof MA, Kleinhofs A. 1993.** Comparative diversity analysis of RFLPs and isozymes within and among populations of *Hordeum vulgare* ssp. *spontaneum*. *Genetics*, 134: 909-916.

Capítulo 5

Caracterización agronómica

5-1. Introducción

En Europa occidental se ha producido una sustitución varietal prácticamente total en la cebada. A lo largo del siglo XX, las variedades locales se fueron sustituyendo paulatinamente por variedades mejoradas. Las viejas variedades locales (*landraces*) se guardan casi exclusivamente en los bancos de germoplasma. Sin embargo, ha habido algunas excepciones a este hecho. En las zonas marginales, donde las condiciones climáticas son difíciles, algunas variedades tradicionales de cebada o cultivares primitivos siguen siendo cultivadas hoy en día. La inestabilidad de las variedades modernas en estas zonas es la principal razón del mantenimiento del cultivo de las variedades tradicionales. Aunque estas últimas no suelen proporcionar rendimientos altos, si suelen ser variedades más estables, asegurando un rendimiento mínimo cada año.

En los países de la región WANA (*West Asia, North Africa*), las variedades tradicionales ocupan el 85% de la superficie destinada al cultivo de cebada (Fischbeck, 2003).

España también sufre del problema de la inestabilidad de los rendimientos de las variedades modernas en las zonas menos favorables. El cultivar tradicional Albacete sigue siendo ampliamente cultivado, principalmente en las zonas semi-áridas como Aragón, donde ocupó el 69% de la superficie dedicada a la cebada de seis carreras en 2002 (DGA, 2002).

La mayor estabilidad de las variedades tradicionales locales, bajo condiciones ambientales desfavorables, es un fenómeno ampliamente descrito por diferentes autores (van Oosterom *et al.*, 1993, Voltas *et al.*, 1999; Lasa *et al.*, 2001). Sin embargo, sus rendimientos suelen ser bajos en condiciones óptimas, comparados con los rendimientos de las variedades modernas.

Sería conveniente diseñar nuevas estrategias de selección para explotar los recursos de estas variedades locales y conseguir nuevas variedades más productivas y adaptadas a condiciones difíciles. Ceccarelli *et al.* (1987) defendieron que la mejora de genotipos de cebada para condiciones extremas se debe realizar precisamente en estas condiciones desfavorables. La hipótesis de estos autores se completa con las afirmaciones de que no existe una adaptación amplia para diferentes ambientes macro-agro-ecológicos, y que la selección para potencial de rendimiento alto no incrementa el rendimiento en condiciones desfavorables (Ceccarelli, 1989). También Ceccarelli y

Grando (1989) explicaron que no es eficiente utilizar material élite convencional para la selección en condiciones de estrés hídrico, siendo conveniente el empleo de material adaptado.

La singularidad genética de las cebadas españolas fue descrita por Tolbert *et al.* (1979), Igartua *et al.* (1998) y Lasa *et al.* (2001). Esta singularidad puede estar relacionada con caracteres de adaptación. Por lo tanto, la colección nuclear de cebadas españolas, representativa de la diversidad de germoplasma español, podría ser una base material para la selección de nuevas variedades adaptadas a las condiciones difíciles de las regiones mediterráneas, o para la identificación de caracteres de adaptación favorables. La evaluación del potencial agronómico ha de realizarse necesariamente en varios ambientes, ya que la interacción genotipo por ambiente (GxE) en la zona mediterránea es un factor de importancia capital. Los objetivos principales de este estudio han sido: 1) evaluar el comportamiento de las entradas de la colección nuclear de cebadas españolas en diferentes localidades, en zonas representativas del cultivo de cebada en España; y 2) analizar las relaciones entre los demás caracteres fenotípicos y el rendimiento, así como el efecto de las condiciones climáticas sobre estos caracteres y su consecuencia sobre el rendimiento.

5-2. Material y Métodos

5-2-1. Material vegetal

Un total de ciento ochenta y dos genotipos fueron evaluados en los ensayos de campo. La distribución de estos genotipos fue la siguiente:

- 145 líneas puras, de espiga de 6 carreras, pertenecientes a la colección nuclear española. Inicialmente, la colección incluía 148 genotipos de este tipo de espiga, pero tres de ellos fueron eliminados, pues resultaron ser duplicados tras la caracterización molecular. Esos tres genotipos se evaluaron en los ensayos de campo, pero sus resultados no fueron tenidos en cuenta;
- 11 líneas puras, de espiga de 2 carreras, de la colección nuclear española;
- 16 cultivares antiguos de amplio uso en España, 8 de dos carreras (Alpha, Beka, Hassan, Kym, Pallas, Union, Wisa y Zaida) y 8 de seis carreras (Ager, Albacete, Almunia, Barbarrosa, Dobra, Hatif de Grignon, Monlon y Pané), incluidas en la colección nuclear;

- 10 cultivares modernos, empleados como testigos (Candela, Gaelic, Graphic, Nevada, Orria, Plaisant, Seira, Steptoe, Tipper y Volga).

La procedencia de las líneas puras de la colección nuclear está recogida en la Tabla 3.2 del capítulo 3. Asimismo, el pedigrí de las variedades está descrito en la Tabla 3.3 del capítulo 3.

A lo largo de este trabajo denominaremos *landraces* a las líneas puras procedentes de poblaciones autóctonas que componen la colección nuclear. Se ha preferido este término a la denominación alternativa de entradas autóctonas, pues ésta incluiría también a tres de los dieciséis cultivares antiguos incluidos en la colección nuclear, que son autóctonos.

5-2-2. Diseño experimental y localidades

La evaluación de los genotipos se llevó a cabo en tres comunidades autónomas: Aragón, Castilla y León y Cataluña, donde se estudiaron en diferentes ambientes. Se dispone de un total de 10 ensayos, clasificados como sigue:

- Cataluña: los ensayos se realizaron en 3 localidades de la provincia de Lleida, que son Artesa de Segre (2002, 2004), Bell-lloc d'Urgell (2002) y Foradada (2003), y se llevaron a cabo bajo la responsabilidad del IRTA- UdL.
- Castilla y León: los ensayos se llevaron a cabo en la localidad de Macotera (Salamanca), en los años 2002 y 2003, bajo la responsabilidad del ITA-JCyL de Valladolid.
- Aragón: se dispone de 4 ensayos llevados a cabo en las localidades de Zuera, en la finca del Vedado (2003) y Sádaba (2002, 2003 y 2004). Estos ensayos fueron realizados por parte de la EEAD-CSIC.

Además de estos ensayos, se sembraron también experimentos en el Vedado de Zuera (2002 y 2004), y Macotera (2004), pero no llegaron a ser cosechados por distintas contingencias climáticas o agronómicas.

Las coordenadas geográficas, así como los principales datos climáticos de estas localidades durante los años de los ensayos, se presentan en la Tabla 5.1.

Los datos climáticos se consiguieron del Instituto Nacional de Meteorología, a partir de estaciones climáticas presentes en la misma localidad del ensayo, como es el caso de Artesa de Segre, Zuera (El Vedado) y Sádaba, o de estaciones presentes en localidades próximas para los otros ensayos, siempre a una distancia inferior a los 20 km. En estos casos, a la hora de la elección de la estación meteorológica, se tuvo en

cuenta la similitud geográfica con la localidad del ensayo y la ausencia de barreras montañosas importantes entre las estaciones y las localidades de ensayo.

Las estaciones elegidas fueron:

- Artesa de Segre: estación Artesa de Segre, código 9650;
- Bell-lloc d'Urgell: estación Mollerusa "Granja", código 9729;
- Macotera: estación Pedraza de Alba, código 2863;
- Sádaba: estación de Sádaba, código 9330;
- Zuera (El Vedado): estación de Zuera (El Vedado), código 9495.

Se obtuvieron datos diarios de pluviometría y temperaturas (mínimas, máximas y medias), de cada localidad para los años de ensayos. A partir de estos datos, se calculó la pluviometría mensual y los promedios mensuales para temperaturas. De la misma forma, se calculó la pluviometría y los promedios de temperaturas para los periodos de otoño (octubre, noviembre, diciembre), invierno (enero, febrero, marzo) y primavera (abril, mayo, junio). La evapotranspiración potencial (ETP) diaria fue calculada, basándose sobre los datos diarios de pluviometría y de temperaturas máximas y mínimas, empleando el método de Hargreaves (Hargreaves *et al.*, 1985). La evapotranspiración mensual se calculó sumando los valores diarios. También se calculó la ETP de otoño, invierno y de primavera sumando los datos mensuales correspondientes a cada periodo. La *ratio* entre pluviometría y ETP de cada periodo se obtuvo dividiendo esas dos variables.

La parcela ocupada por cada ensayo tenía un tamaño considerable que, en las condiciones de la mayoría de las localidades, supondría unas condiciones de suelo muy heterogéneas. Para mitigar el efecto de esa heterogeneidad sobre el experimento, se decidió subdividirlo en cuatro sub-ensayos, en cada localidad. En cada sub-ensayo se sembraron 44 entradas de la colección nuclear y 10 testigos, resultando un total de 54 entradas por sub-ensayo. Como sólo había 175 entradas en la colección nuclear (y el diseño adoptado requería 176), en uno de los sub-ensayos fue necesario rellenar una parcela con un testigo adicional. Tanto las líneas puras de 2 carreras como las variedades antiguas de 2 y 6 carreras fueron distribuidas de forma igualada en los sub-ensayos. La división en sub-ensayos también se hizo atendiendo al origen geográfico de las entradas. Así, las entradas de una misma provincia se dispusieron en sub-ensayos distintos, excepto en los casos en que el número de entradas por provincia era mayor que 4, en los que necesariamente se repetían entradas de esas provincias en algún sub-ensayo. También se atendió a la distribución por comunidades autónomas, de modo que

cada sub-ensayo fuera una representación lo más fiel posible de la diversidad geográfica presente en la colección nuclear.

El diseño experimental adoptado en cada ambiente, para cada sub-ensayo, fue un alfa-látice con 3 repeticiones. Los bloques incompletos se dispusieron en las dos direcciones, fila ($k=9$) y columna ($k=6$). Cada genotipo se sembró en una parcela de 6 surcos de 6 m de longitud, espaciados 0,2 m, con una dosis de 350 semillas m^{-2} . Cada sub-ensayo tenía una anchura de 27 parcelas, sembradas en seis hileras, con un pasillo de 1 m entre hileras. Así cada sub-ensayo formaba un cuadrado casi exacto. La disposición de los sub-ensayos en el campo fue en cuadrantes, conformando una parcela cuadrada de aproximadamente 81 m de lado (Figura 5.1). Las repeticiones se dispusieron en perpendicular a las hileras. Al no corresponder las repeticiones entre sub-ensayos (la repetición I del sub-ensayo 1 no se corresponde con la del sub-ensayo 4, por ejemplo), este factor estaba jerarquizado a los sub-ensayos en los análisis. Las siembras se realizaron en otoño durante, los meses de noviembre o diciembre (Tabla 5.1), mientras que las cosechas se efectuaron entre los meses de junio y julio de cada campaña.

Tabla 5.1. Coordenadas geográficas y características climáticas de las localidades experimentales, y fechas de siembra de los diferentes ensayos.

Comunidad Autónoma	Localidad	Latitud	Longitud	Altitud (m)	Año del ensayo	Código ensayo	Temperatura media (°C)			Temperatura mínima (°C)			Temperatura máxima (°C)		
							otoño	invierno	primavera	otoño	invierno	primavera	otoño	invierno	primavera
Cataluña	Artesa de Segre	41°53'	01°02'	320	2002	ART02	7,10	7,13	16,13	1,40	0,70	8,63	12,77	13,57	23,57
					2004	ART04	9,30	5,67	16,00	4,97	0,20	9,30	13,67	11,77	22,67
	Bell-lloc d'Urgell	41°38'	00°47'	210	2002	BEL02	8,33	7,97	16,83	3,43	2,60	9,43	13,13	13,37	24,23
	Foradada	41°52'	1°00'	455	2003	FOR03	9,80	5,83	18,07	4,43	-0,63	10,30	15,13	12,33	25,80
Castilla y León	Macotera	40°49'	05°17'	892	2002	VLD02	8,97	9,07	15,97	3,07	3,67	8,73	15,10	14,47	23,17
					2003	VLD03	11,73	6,33	14,70	6,83	-3,40	5,17	16,57	16,50	23,60
Aragón	Sádaba	42°16'	01°16'	440	2002	SAD02	8,93	8,57	16,20	2,97	2,97	9,13	14,87	14,07	23,20
					2003	SAD03	11,17	6,73	17,37	6,43	1,63	10,37	15,87	11,83	24,33
					2004	SAD04	9,33	6,00	15,50	4,93	0,93	8,70	13,77	11,03	22,23
	Zuera (El Vedado)	41°52'	00°39'	298	2003	VED03	11,07	7,13	18,10	6,53	2,17	11,33	15,60	12,07	24,80

Fuente: elaboración propia a partir de datos del Instituto Nacional de Meteorología

Tabla 5.1. (Continuación.)

Comunidad Autónoma	Localidad	Año del ensayo	Fecha de siembra	Pluviometría (mm)				ETP ¹			Ratio ²		
				otoño	invierno	primavera	total	otoño	invierno	primavera	otoño	invierno	primavera
Cataluña	Artesa de Segre	2002	31/10/2001	100,00	67,00	202,00	369,00	129,50	153,80	434,80	0,77	0,44	0,46
		2004	18/12/2003	129,80	126,00	167,00	422,80	113,30	133,40	410,00	1,15	0,94	0,41
	Bell-lloc d'Urgell	2002	07/11/2001	82,20	54,40	152,30	288,90	127,20	147,10	442,20	0,65	0,37	0,34
	Foradada	2003	07/11/2002	165,00	178,00	34,80	377,80	133,40	147,40	467,20	1,24	1,21	0,07
Castilla y León	Macotera	2002	08/11/2001	111,10	106,60	136,70	354,40	140,20	154,40	430,70	0,79	0,69	0,32
		2003	11/11/2002	199,30	169,00	71,60	439,90	136,90	141,90	418,70	1,46	1,19	0,17
Aragón	Sádaba	2002	28/11/2001	66,10	58,90	162,20	287,20	135,80	149,30	421,50	0,49	0,39	0,38
		2003	27/11/2002	146,50	189,10	77,00	412,60	127,60	132,60	436,90	1,15	1,43	0,18
	Zuera (El Vedado)	2004	23/12/2003	158,90	115,60	129,50	404,00	113,90	126,40	409,10	1,40	0,91	0,32
		2003	08/11/2003	138,00	127,00	106,00	371,00	125,40	133,90	438,70	1,10	0,95	0,24

¹ ETP: evapotranspiración total calculada según el método de Hargreaves

² Ratio: pluviometría dividido por la ETP

Fuente: elaboración propia a partir de datos del Instituto Nacional de Meteorología

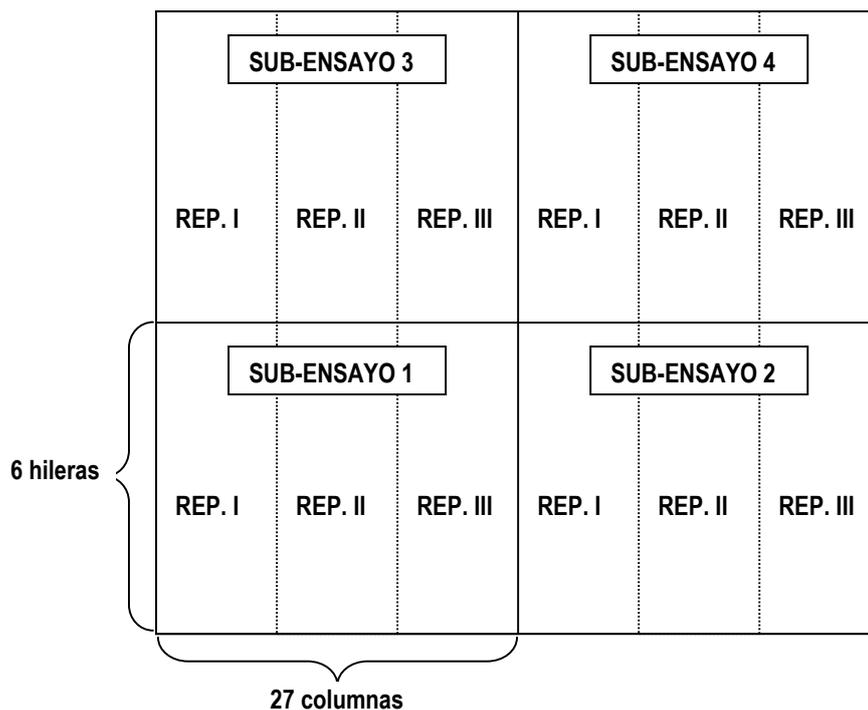


Figura 5.1. Esquema de los ensayos de campo, mostrando la disposición de los sub-ensayos y las repeticiones.

5-2-3. Caracteres estudiados

5-2-3-1. Espigado (ESP)

La fecha de espigado de cada genotipo se registró cuando el 50% de los tallos de la parcela presentaban las aristas emergidas al menos 2 cm de la vaina de la hoja bandera. El carácter se expresó en días transcurridos desde el 1 de enero.

5-2-3-2. Altura (ALT)

Se midió la altura de los genotipos en la fase de maduración. Eligiendo una planta representativa de la parcela, se midió su altura en cm a partir del nudo de ahijamiento hasta el punto de inserción de la hoja bandera.

5-2-3-3. Encamado (ENC)

Se evaluó en los ensayos en los que ocurrió de modo acusado. Para este carácter, se atribuyeron valores de forma subjetiva a las parcelas, basándose sobre una escala del IPGRI (1994), que varía de 1 a 9. Los niveles se definen como sigue:

- 1. Muy bajo
- 2. Muy bajo a bajo
- 3. Bajo
- 4. Bajo a intermedio
- 5. Intermedio
- 6. Intermedio a alto
- 7. Alto
- 8. Alto a muy alto
- 9. Muy alto

5-2-3-4. Rendimiento (RDT)

Tras la cosecha con cosechadora, se determinó el peso de cada parcela. A partir de estos datos se calculó el rendimiento de cada parcela en $q \text{ ha}^{-1}$.

Los resultados de rendimiento se han dividido en otras dos nuevas variables, según el rendimiento promedio en distintos ambientes. Así, se definió el rendimiento en ambientes de baja y alta productividad. Los valores del ensayo ART02 no se incluyeron en el análisis de este carácter, ante las irregularidades que se produjeron en la cosecha de ese experimento, inducidas por el alto encamado

5-2-3-5. Peso hectolítrico (PHL)

Es el peso de 100 litros de semillas, expresado en kilogramos, determinado mediante el empleo de un analizador de granos (modelo Dickey-John).

5-2-3-6. Peso de 1000 granos (PMG)

Tras la cosecha se contaron 1000 semillas de cada genotipo, mediante un contador de grano. Luego se determinó su peso expresado en gramos.

Cada carácter no fue estimado en todos los ensayos, aunque si se obtuvieron datos suficientemente representativos en la mayoría de ellos. La Tabla 5.2 resume los caracteres medidos en cada ensayo.

Tabla 5.2. Caracteres agronómicos registrados en cada ensayo.

Carácter Ensayo	ALT	ESP	ENC	RDT	PHL	PMG
ART02	x	x		x	x	x
BEL02		x	x	x	x	x
SAD02	x			x		
VLD02	x	x		x	x	x
FOR03		x	x	x	x	x
VLD03	x	x	x	x	x	x
SAD03	x			x	x	
VED03	x	x	x	x	x	x
ART04		x		x	x	x
SAD04				x	x	x

5-2-3-7. Componentes de rendimiento

Los componentes de rendimiento se evaluaron solamente en el ensayo del VED03, que presentó unas buenas condiciones de producción. Esta evaluación se hizo sólo en las dos repeticiones centrales de cada sub-ensayo, quitando las repeticiones de los lados. En cada parcela se eligió un surco, en el cual se cosecharon a mano las espigas de 1 m lineal. De estas plantas, se determinó:

- número de espigas m^{-2} (ESM);
- número de granos por espiga (GRE);
- número de granos m^{-2} (GRM);
- peso de 1000 granos (PMG).

5-2-3-8. Enfermedades

La evaluación de la incidencia de enfermedades foliares fue realizada en condiciones naturales, en los ensayos donde hubo ataques generalizados. Se atribuyeron

valores a las parcelas de forma subjetiva, basándose sobre una escala de susceptibilidad elaborada por el IPGRI (1994), cuyos niveles varían de 1 a 9, de forma siguiente:

- 1. Susceptibilidad no visible a muy baja
- 3. Baja
- 5. Intermedia
- 7. Elevada
- 9. Muy elevada

Las principales enfermedades observadas fueron el moteado de las hojas (MOT), provocado por el hongo *Pyrenophora teres*, manifestada principalmente en los ensayos ART02 y FOR03; y el oidio (OID), producido por el hongo *Blumeria graminis*, y que se reveló especialmente en los ensayos BEL02 y SAD02.

Anteriormente a estos ensayos, las líneas puras de la colección nuclear fueron evaluadas en varias localidades, en parcelas de menor tamaño. Dos de estos ensayos se llevaron a cabo en Gerona y Valladolid, en el año 2000. Cada genotipo fue sembrado en una parcela de 6 surcos, de 3 m de longitud, espaciados de 0,25 cm, con una única repetición (Lasa *et al.*, 2001). En este estudio se emplearon los datos de incidencia de enfermedades registradas en estas dos localidades. En Gerona, se observaron síntomas de oidio y roya (RYA) provocada por el hongo *Puccinia hordei*, mientras que en Valladolid sólo se notaron síntomas de oidio.

5-2-3-9. Calidad maltera

La evaluación de la calidad maltera de las entradas se llevó a cabo en la maltería DAMM de Bell-lloc en la provincia de Lleida, bajo la responsabilidad del Dr. Molina-Cano (IRTA-UdL).

Esta evaluación se realizó a partir de muestras cogidas del ensayo de FOR03 que presentó una calidad suficiente para ofrecer unos resultados válidos.

Los caracteres elegidos para la determinación de la calidad maltera se basaron sobre aspectos físicos y bioquímicos del grano:

- Calibre de los granos: este carácter se midió mediante dos variables, que fueron el porcentaje de granos que pasaban por una criba de 2,2 mm, y el porcentaje de granos que no pasaban por una criba de 2,5 mm).

- Porcentaje de materia seca en el extracto de malta: es un indicador esencial de la calidad maltera y se mide como el porcentaje de elementos solubles procedentes de la malta que se incorporan al mosto.
- Proteína total en malta: es el porcentaje de proteína total en harina de malta.

5-2-4. Análisis estadísticos

5-2-4-1. Análisis de cada ensayo

Como se ha descrito antes, cada ensayo de campo se componía de cuatro sub-ensayos. La combinación de los datos de estos sub-ensayos se realizó de la siguiente manera:

Para cada uno de los sub-ensayos, se hizo un análisis siguiendo el diseño experimental en alfa-látice con el procedimiento *proc mixed* (SAS, 1988); éste permitió estimar el valor medio de cada genotipo en cada sub-ensayo, para cada uno de los caracteres estudiados. A partir de los valores de los 10 testigos comunes incluidos en cada sub-ensayo (ver apartado 5-2-1), se verificó la homogeneidad de los cuatro sub-ensayos de cada localidad mediante un análisis de varianza (*proc glm*, SAS, 1988) seguido por una separación de medias de los 4 sub-ensayos (método Scheffe), siguiendo un diseño de bloques al azar, con tres repeticiones, con el factor repetición jerarquizado dentro de sub-ensayos. Cuando las medias de los testigos fueron significativamente diferentes para alguno de los caracteres estudiados, se procedió a corregir los valores estimados de los genotipos, para dicho carácter. El coeficiente de corrección se calculó dividiendo el promedio general de los genotipos (testigos no incluidos) por el promedio del sub-ensayo (testigos no incluidos). Los datos corregidos se obtuvieron multiplicando el correspondiente coeficiente por el valor de los genotipos incluidos en cada sub-ensayo. Esta corrección resultaba más conservadora que la calculada basándose en los 10 testigos comunes, que aparentemente tendía a sobre corregir los datos.

Una vez que se corrigieron los datos, se creó una nueva base de datos con un único valor para cada carácter por genotipo y ambiente, mediante el procedimiento *proc means* (SAS, 1988).

Los análisis que se describen en el siguiente apartado se basaron sobre esta nueva base de datos.

5-2-4-2. Análisis combinado de los ensayos

a. Estudio descriptivo de las variables

Se calcularon los parámetros estadísticos (mínimo, máximo, promedio y desviación estándar) para todas las variables analizadas, separando las 156 líneas españolas de las variedades comerciales. Estos cálculos se hicieron para cada uno de los ensayos donde fue medido cada carácter.

Basándose sobre la media general de cada genotipo, se representaron las distribuciones de valores (en porcentajes) de cada carácter para las variedades comerciales y las 156 *landraces*, mediante histogramas. Las clases de los histogramas presentados son de una amplitud de 0,5 desviaciones típicas para los caracteres cuantitativos continuos, y de una unidad para los caracteres discretos (enfermedades, encamado).

b. Análisis de varianza (interacción genotipo x ambiente)

Para cada carácter analizado, se ha realizado un análisis de varianza teniendo en cuenta el número de ensayos en los que se había medido. En una primera etapa, se analizaron los factores genotipo y ensayo así como la interacción genotipo x ensayo.

En una segunda etapa se dividió el factor genotipo en siete grupos, separando por un lado las líneas puras de la colección nuclear, de acuerdo a su clasificación en grupos genéticos, según la información molecular del apartado 3-3-3 del capítulo 3, y por otro lado las variedades.

La composición de estos siete grupos es como sigue:

- Grupo 1: 17 entradas de seis carreras de la colección nuclear.
- Grupo 2: 9 entradas de dos carreras de la colección nuclear.
- Grupo 3: 48 entradas de la colección nuclear, (47 de seis carreras), que proceden en su mayoría de la Meseta.
- Grupo 4: 82 entradas de la colección nuclear, procedentes del sur de la península, de la vertiente mediterránea y del valle del Ebro.
- Grupo 5: 8 variedades antiguas de dos carreras (Alpha, Beka, Hassan, Kym, Pallas, Union, Wisa y Zaida).
- Grupo 6: 8 variedades antiguas de seis carreras (Ager, Albacete, Almunia, Barbarrosa, Dobra, Hatif, Monlon y Pané).

- Grupo 7: los 10 cultivares, la mayoría modernos, introducidos como testigos en los ensayos de campo (Candela, Gaelic, Graphic, Nevada, Orria, Plaisant, Seira, Steptoe, Tipper y Volga).

Un análisis de varianza permitió estimar la variabilidad “entre” y “dentro de” estos grupos, así como la interacción genotipo por ensayo “entre grupos” y “dentro de” grupos.

El factor genotipo se ha tomado como aleatorio, pues las entradas suponen una muestra representativa de todas las cebadas españolas, y de las variedades cultivadas en España. El factor ambiente, sin embargo, por el número limitado de ensayos, se ha considerado como fijo. Este número es suficiente para una evaluación agronómica adecuada, pero insuficiente para considerarlo una muestra representativa de todos los ambientes posibles.

El cuadrado medio del error general de la tabla del análisis de varianza combinado fue calculado a partir de los análisis de cada ensayo, como el promedio de los errores de los diferentes sub-ensayos. Como el análisis combinado se realizó con las medias ajustadas de los genotipos, los valores de la suma de cuadrados se corrigieron multiplicándolos por el número de repeticiones.

Para las variables medidas en un único ensayo (componentes de rendimiento, datos de calidad maltera, y la enfermedad de la roya) se analizó solamente la varianza entre esos grupos, tomando como error del análisis la varianza dentro de los grupos.

Cuando el análisis de varianza detectaba diferencias significativas entre grupos, se procedió a una separación de medias de grupos con el método LSD, para una probabilidad de $\alpha=0,01$.

c. AMMI (*Additive Main effects and Multiplicative Interaction*)

Con objeto de describir mejor la interacción genotipo por ambiente observada para el carácter rendimiento, se realizó un análisis AMMI. Este análisis se basa sobre la utilización de componentes principales y un modelo que incorpora componentes aditivos y multiplicativos (Gauch, 1988, 1992).

La fórmula general del modelo AMMI para la estimación del valor del genotipo i en el ambiente j es:

$$E(Y_{ij}) = \mu + g_i + e_j + \sum_{k=1}^k c_{ik} d_{jk}$$

donde:

μ : es la media general;

g_i : es el efecto de genotipo;

e_i : el efecto del ambiente;

c_{ik} y d_{jk} son valores multiplicativos genotípicos y ambientales, respectivamente.

Este análisis se realizó con el programa SAS (SAS, 1988), basándose sobre un modelo de programa para el análisis AMMI del módulo de estadística del curso de mejora genética vegetal del IAMZ 2000-2001 (Romagosa, 2001).

El número (K) de términos AMMI necesarios para una descripción adecuada de la interacción genotipo por ambiente, se evaluó con el método de Gollob (1968). El cuadrado medio de cada término se obtiene a partir de los “*eigenvalues*” (suma de cuadrados) dividido por los grados de libertad, calculados según el procedimiento propuesto por Gollob (1968): $gl = G + E - 1 - 2n$ (G, número de genotipos; E: número de ambientes y n, número de ejes retenidos). Se calcula la *ratio* de la varianza (cuadrado medio / error general estimado), cuya significación se determina según una distribución de F.

Para una mejor interpretación del análisis AMMI, se procedió a una representación gráfica bidimensional descriptiva de los resultados. Las posiciones de los puntos de genotipos, así como las coordenadas de los ambientes resultan de los valores multiplicativos estimados respectivamente para los genotipos y los ambientes.

La representación gráfica del AMMI2 (modelo con dos factores multiplicativos) fue enriquecida con información de otras variables genotípicas y ambientales externas al análisis AMMI. Esto se llevó a cabo mediante las regresiones de las covariables genotípicas y ambientales sobre los valores genotípicos y ambientales (respectivamente) de los ejes 1 y 2 del AMMI. Los coeficientes de regresión (a_1 , a_2) con relación al origen (0,0), definen la dirección de las variaciones de la covariables en la gráfica, y permiten la proyección del vector de la covariable en el biplot. Solo se representaron en el AMMI, las covariables que presentaron coeficientes de determinación (R^2) elevados.

d. Regresión factorial

Las covariables genotípicas y ambientales pueden ser introducidas en el análisis de varianza, y especialmente de la interacción genotipo por ambiente (GxE), de un modo más directo, mediante regresión factorial. La fórmula general del modelo de la regresión factorial (van Eeuwijk, 1995) es :

$$E(Y_{ij}) = \mu + g_i + e_j + \sum_{k=1}^k x_{ik} \tau_{jk}$$

donde x_{ik} es el valor de la variable k genotípica o fenotípica del genotipo i , y τ_{jk} el valor de una variable ambiental k para el ensayo j .

En una primera fase, las regresiones factoriales fueron realizadas para descomponer al factor *genotipo* con las variables genotípicas. En una segunda fase se procedió a la descomposición de la interacción *GxE*, empleando las covariables genotípicas y ambientales, y la interacción entre ellas.

En los dos casos, este análisis se realizó de forma secuencial, introduciendo en una primera etapa una variable cada vez. En una segunda ronda, se conservaba en el modelo la variable más explicativa, según la suma de cuadrados de tipo I, y se introducía una segunda variable, etc. De esta forma se realizaron diferentes rondas de análisis, empleando todas las combinaciones entre las covariables, hasta llegar al modelo más parsimonioso.

5-3. Resultados

5-3-1. Análisis de cada ensayo

La verificación de la homogeneidad de los 4 sub-ensayos de cada ambiente se verificó mediante un análisis de varianza sobre los valores de los testigos, lo que permitió corregir los datos de los genotipos en los diferentes sub-ensayos, cuando esto fue necesario. En la Tabla 5.3 se presentan los promedios y desviaciones estándar de las variedades testigo así como los coeficientes de corrección (C) calculados para las variables agronómicas (rendimiento, espigado, altura, encamado, peso hectolítrico y peso de mil granos), en los casos en que fue necesario. Para algunas variables, no se pudo verificar la homogeneidad de la parcela por disponer solamente de datos de una repetición, como fue el caso del espigado y la altura de planta en el ensayo VLD03, del peso hectolítrico en los ensayos ART04, SAD03, SAD04 y VED03, y del peso de mil granos en los ensayos ART02, BEL02, FOR03 y SAD04. En los casos donde no hubo correcciones, fue debido a que no se observaron diferencias significativas entre sub-ensayos para las variedades testigo. Las correcciones se aplicaron a los promedios de los valores genotípicos calculados de acuerdo con el análisis de alfa-látice en dos direcciones seguido en cada sub-ensayo. Como se puede ver por el valor de los coeficientes C de las Tablas 5.3 y 5.4, las correcciones fueron muy moderadas.

Para las variables de componentes de rendimiento, evaluadas en el ensayo VED03, sólo se corrigieron los valores de la variable granos por espiga. En las otras variables no se observaron diferencias significativas en los promedios de los testigos de los 4 sub-ensayos (Tabla 5.4). Tampoco se observaron diferencias significativas en los promedios de los testigos para las enfermedades observadas en los ensayos ART02, BEL02, SAD02 y FOR03 (Tabla 5.5).

Tabla 5.3. Descripción de los 10 ensayos empleados en este estudio: promedios y desviaciones estándar de las variables agronómicas (rendimiento, espigado, altura, encamado, peso hectolítrico y peso de mil granos) de las variedades testigo por ensayo y sub-ensayo. C: coeficiente de corrección de los sub-ensayos.

Ensayo	Sub-ensayo	Variables																	
		Espigado (días)			Altura (cm)			Encamado			Rendimiento (q ha ⁻¹)			Peso hectolítrico (kg hl ⁻¹)			Peso 1000 granos (g)		
		Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C
ART02	1	120,4	4,4	—	68,1	11,3	0,9604	—	—	—	30,32	2,13	1,0122	68,9	2,7	—	37,4	4,0	—
	2	120,3	3,8	—	70,0	11,5	0,9643	—	—	—	29,71	1,88	0,9655	68,9	2,9	—	38,3	3,9	—
	3	120,5	3,9	—	65,3	9,0	1,0206	—	—	—	27,73	2,08	1,0189	69,6	2,5	—	39,0	3,0	—
	4	120,7	3,5	—	63,4	7,9	1,0618	—	—	—	26,62	2,66	1,0051	69,0	2,7	—	39,2	3,3	—
BEL02	1	107,6	4,1	—	—	—	—	1,1	0,3	—	38,13	3,90	—	69,5	2,7	0,9924	35,7	5,0	—
	2	108,4	3,8	—	—	—	—	1,1	0,3	—	36,39	4,12	—	66,9	3,1	1,0146	35,3	5,2	—
	3	107,8	4,0	—	—	—	—	1,0	0,1	—	38,75	4,38	—	69,7	3,6	0,9820	40,7	3,8	—
	4	106,8	5,3	—	—	—	—	1,1	0,3	—	37,76	3,01	—	68,0	3,8	1,0118	36,4	3,7	—
SAD02	1	—	—	—	46,0	6,5	0,9969	—	—	—	20,10	2,14	1,0018	—	—	—	—	—	—
	2	—	—	—	44,3	6,4	1,0287	—	—	—	19,18	1,71	1,0382	—	—	—	—	—	—
	3	—	—	—	48,5	5,9	0,9447	—	—	—	21,00	2,07	0,9564	—	—	—	—	—	—
	4	—	—	—	44,7	6,4	1,0351	—	—	—	19,80	2,19	1,0065	—	—	—	—	—	—
VLD02	1	119,6	2,9	—	63,8	11,8	—	—	—	—	43,57	2,81	0,9974	72,4	2,7	0,9929	45,1	2,9	—
	2	120,7	5,8	—	65,3	11,9	—	—	—	—	43,73	2,75	1,0000	70,6	2,6	1,0009	44,7	2,9	—
	3	119,4	2,7	—	64,4	11,7	—	—	—	—	44,18	3,35	0,9744	72,0	2,3	0,9883	44,9	3,2	—
	4	120,4	3,3	—	64,3	10,1	—	—	—	—	37,89	2,94	1,0305	70,0	2,5	1,0185	44,2	2,8	—
FOR03	1	114,7	2,5	—	—	—	—	1,0	0,8	—	20,18	2,55	1,0144	70,3	1,5	1,0115	35,6	3,9	1,0197
	2	114,5	2,4	—	—	—	—	1,4	1,2	—	24,12	3,64	0,9023	69,9	1,8	1,0083	35,6	3,5	0,9969
	3	115,3	3,1	—	—	—	—	1,3	0,8	—	16,70	2,63	1,0723	70,1	2,2	1,0031	36,3	3,1	1,0023
	4	114,4	3,5	—	—	—	—	1,1	0,1	—	17,37	1,40	1,0317	68,1	1,4	0,9779	38,1	2,2	0,9818

Tabla 5.3. (Continuación.).

Ensayo	Sub-ensayo	Variables																	
		Espigado (días)			Altura (cm)			Encamado			Rendimiento (q ha ⁻¹)			Peso hectolítrico (kg hl ⁻¹)			Peso 1000 granos (g)		
		Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C
VLD03	1	119,3	5,1	—	59,0	14,3	—	0,9	0,7	—	21,41	3,07	1,0237	68,3	2,7	1,0144	39,2	3,2	1,0125
	2	119,4	3,9	—	59,5	14,4	—	0,9	0,6	—	19,21	3,46	1,0987	69,5	2,4	0,9778	37,9	4,0	1,0213
	3	119,1	6,1	—	59,7	6,5	—	1,0	0,5	—	23,85	3,18	0,9849	67,5	2,2	1,0171	40,1	4,2	0,9820
	4	118,1	4,0	—	67,4	16,0	—	0,9	0,8	—	27,45	4,23	0,9065	69,9	2,3	0,9918	40,4	3,9	0,9853
SAD03	1	—	—	—	60,8	9,5	0,9558	—	—	—	19,37	3,05	0,9332	67,5	2,2	—	—	—	—
	2	—	—	—	58,0	8,3	1,0035	—	—	—	16,94	3,00	1,0475	67,8	2,2	—	—	—	—
	3	—	—	—	56,3	9,7	1,0040	—	—	—	16,74	3,53	0,9871	68,5	1,9	—	—	—	—
	4	—	—	—	56,9	8,4	1,0403	—	—	—	16,24	3,41	1,0409	68,1	2,1	—	—	—	—
VED03	1	109,2	3,9	—	64,1	12,6	—	1,4	0,6	—	30,12	3,75	—	69,3	3,0	—	40,1	3,2	—
	2	109,1	4,0	—	62,5	11,5	—	1,2	0,9	—	28,00	3,75	—	69,6	2,3	—	39,8	4,0	—
	3	110,9	2,9	—	63,1	14,5	—	1,3	0,4	—	30,20	3,41	—	69,4	2,9	—	39,0	4,6	—
	4	110,1	3,2	—	61,0	12,1	—	1,1	0,2	—	29,72	2,90	—	69,4	2,6	—	39,2	2,9	—
ART04	1	117,2	2,7	—	—	—	—	—	—	—	43,84	5,18	0,9960	71,4	2,6	—	41,4	4,6	—
	2	117,2	2,6	—	—	—	—	—	—	—	40,50	3,68	1,0236	69,7	4,3	—	39,6	5,1	—
	3	117,1	2,4	—	—	—	—	—	—	—	41,70	3,08	1,0004	71,4	2,3	—	39,6	4,3	—
	4	117,2	3,0	—	—	—	—	—	—	—	40,46	3,92	0,9803	68,2	3,4	—	37,3	4,1	—
SAD04	1	—	—	—	—	—	—	—	—	—	37,68	4,93	1,0104	67,3	3,6	—	—	—	—
	2	—	—	—	—	—	—	—	—	—	35,80	5,17	1,0240	69,5	2,5	—	—	—	—
	3	—	—	—	—	—	—	—	—	—	41,15	5,28	0,9510	69,0	2,3	—	—	—	—
	4	—	—	—	—	—	—	—	—	—	38,33	5,17	1,0187	69,4	3,0	—	—	—	—

Tabla 5.4. Promedio y desviación estándar de las variables de componentes de rendimiento (espigas m², número de granos m², granos por espiga y peso de mil granos), de las variedades testigo, por sub-ensayo; C: coeficiente de corrección de los sub-ensayos.

Sub-ensayo	Variables											
	Espigas m ²			Granos m ²			Granos por espiga			Peso 1000 granos (g)		
	Media	SD	C	Media	SD	C	Media	SD	C	Media	SD	C
1	322,3	70,0	—	7446,5	2096,3	—	23,4	5,8	1,0546	40,1	3,2	—
2	334,3	114,7	—	8080,3	3241,6	—	23,6	5,9	1,0434	39,8	4,0	—
3	334,3	120,4	—	7886,5	1957,3	—	25,4	8,6	0,9694	39,0	4,6	—
4	303,3	69,0	—	7624,0	1729,6	—	26,2	7,9	0,9416	39,2	2,9	—

Tabla 5.5. Promedios y desviación estándar de las enfermedades (moteado de las hojas y oídio) de los testigos observados en los sub-ensayos de los ensayos de Artesa2002, Bell-lloc 2002, Sádaba 2002 y Foradada 2003).

Ensayo	Sub-ensayo	Variables					
		Moteado de las hojas			Oídio		
		Media	SD	C	Media	SD	C
ART02	1	2,9	1,3	—	—	—	—
	2	3,0	1,5	—	—	—	—
	3	3,0	1,3	—	—	—	—
	4	3,2	1,1	—	—	—	—
BEL02	1	—	—	—	2,1	1,3	—
	2	—	—	—	1,9	0,8	—
	3	—	—	—	2,0	1,3	—
	4	—	—	—	2,2	2,0	—
SAD02	1	—	—	—	3,1	1,0	—
	2	—	—	—	2,8	0,9	—
	3	—	—	—	3,5	0,8	—
	4	—	—	—	2,4	1,1	—
FOR03	1	7,6	1,8	—	—	—	—
	2	7,5	1,6	—	—	—	—
	3	7,8	0,8	—	—	—	—
	4	7,7	1,1	—	—	—	—

5-3-2. Comparación entre las *landraces* y las variedades comerciales

Las *landraces* de la colección nuclear (156, tras la eliminación de los tres duplicados) y las variedades comerciales (26, de las cuales 16 están incluidas en la colección nuclear y los 10 testigos) fueron comparadas mediante un análisis estadístico descriptivo (Tabla 5.6). Esta comparación permite una valoración global de las características agronómicas de la colección nuclear en contraste con cultivares mejorados. A este respecto es conveniente resaltar las posibles desviaciones derivadas de la diferente incidencia de materiales de dos carreras en ambos grupos, el 54% de las comerciales frente al 7% de las *landraces*.

5-3-2-1. Espigado

Los ensayos de BEL02 y VED03 fueron los de espigado más precoz, mientras que los más retrasados fueron los de VLD02 y VLD03 (Tabla 5.6).

En todos estos ensayos, se observó que las *landraces* espigaron más tarde que las variedades comerciales (Tabla 5.6). Esta diferencia osciló entre los 2,3 días de ART02 y los 5,8 días de BEL02. Por otra parte, las diferencias entre los rangos de fechas de espigado en las variedades comerciales variaron entre 9,7 (ART04) y 19 días (VLD03). Para las *landraces*, sin embargo, el rango de variación del espigado fue muy superior, oscilando entre 17,5 (ART04) y 26 días (VLD03). En el histograma A de la Figura 5.2, se representa la distribución de los días de espigado (promedio de todos los ambientes) según lo explicado en el apartado 5-2-4-2.a, para los dos grupos considerados (*landraces* y variedades comerciales). Se observa que la mayoría de las variedades comerciales espigaron como promedio entre 114 y 116 días a partir del 1 de enero. Para las variedades autóctonas, la moda se situó hacia los 118 días desde el 1 de enero.

5-3-2-2. Altura de planta

La altura de planta fue máxima en los ensayos VLD02 y VLD03, y los valores menores se registraron en el ensayo SAD02. En todos los ensayos, las *landraces* mostraron unas alturas mayores que las variedades comerciales. Las diferencias en promedio entre estos dos grupos variaron de 9,1 cm (VLD02) a 25,3 cm (VLD03). Las variaciones entre los distintos ambientes para este carácter fueron muy grandes. Para las variedades comerciales, los promedios de altura variaron de 50,1 cm (SAD02) a 74,0 cm (ART02). Para las *landraces*, las alturas variaron de 61,8 cm (SAD02) a 92,3 cm

(VLD03). El histograma B de la Figura 5.2 ilustra la distribución de los individuos de cada uno de los dos grupos considerados, en función de las clases de altura calculadas. Se observa que más del 40% de las variedades no llegaron a 64 cm de promedio, mientras que la altura de la mayoría de las *landraces* estaba entre 79 y 83 cm.

5-3-2-3. Encamado

El encamado apareció en algún grado en la mayoría de los ensayos. Se tomaron valoraciones en cuatro ensayos en los que este carácter se manifestó de modo acusado. En todos los casos, las *landraces* fueron las más afectadas. Aunque se observaron algunos valores máximos de encamado en las variedades comerciales, estos fueron muy escasos. De hecho los promedios para este grupo no superaron la valoración de 2. Sin embargo, en las líneas puras de la colección nuclear, los promedios fueron más altos, con un máximo de 6,5 en el ensayo BEL02. Considerando el promedio de los genotipos en los cuatro ensayos, se observa que casi el 70% de las variedades comerciales tienen un valor de encamado de 1, pero que la mayoría de las *landraces* se sitúan entre los valores de 4 a 7 (histograma C, Figura 5.2).

5-3-2-4. Rendimiento

Considerando todos los genotipos analizados, los rendimientos promedio de los ensayos fluctuaron entre los 17,0 q ha⁻¹ de FOR03 y los 38,7 q ha⁻¹ de VLD02 .

En siete de los ensayos, las variedades comerciales mostraron unos rendimientos mayores que las *landraces*. Esto se notó especialmente en el ensayo ART04, donde las variedades comerciales superaron a las *landraces* en 11,98 q ha⁻¹ de promedio. En los otros 6 ensayos, estas diferencias variaron entre 1,93 (SAD02) y 6,85 q ha⁻¹ (ART02). En los ensayos VLD03, SAD03 y VED03, las líneas puras de la colección nuclear fueron, en promedio, más productivas que las variedades comerciales, con ventajas de 3,35, 2,49 y 1,19 q ha⁻¹, respectivamente. El histograma D de la Figura 5.2 muestra la distribución de rendimientos de los genotipos de los dos grupos considerados. Las dos distribuciones presentaron las mayores frecuencias en la misma clase de rendimiento (con el punto central en 27,2 q ha⁻¹), pero la distribución de las variedades comerciales está fuertemente sesgada hacia la izquierda, mientras que la de las *landraces* se acerca más a una distribución normal. O sea, que hay muchas más variedades comerciales entre los genotipos que presentaron mejores rendimientos medios. En cualquier caso, es

destacable que un número pequeño de *landraces* presentó rendimientos comparables a las variedades cultivadas (excepto a las tres o cuatro mejores variedades).

Con objeto de ver como se comportaron los genotipos de estos dos grupos en diferentes condiciones ambientales, se procedió a la división de la variable rendimiento en dos nuevas variables, según el potencial de rendimiento observado en los distintos ensayos. Este potencial de rendimiento se midió como el rendimiento promedio de los testigos exclusivamente. Se consideró que el rendimiento promedio de todos los genotipos del ensayo no era una buena estimación del potencial de rendimiento de un ambiente, pues estaba desvirtuado en aquellos ensayos donde ocurrió un encamado notable. Los testigos son mucho menos propensos al encamado y, por tanto, su promedio es una mejor aproximación al rendimiento potencial de un ambiente que el promedio general. Los ensayos SAD02, FOR03, SAD03, VED03 y VLD03 (rendimientos promedio de variedades testigo entre 16,12 y 27,71 q ha⁻¹) se definieron como de rendimiento en condiciones de baja productividad. De forma similar, los ensayos de BEL02, VLD02, ART04 y SAD04 (rendimientos promedio de variedades testigo entre 32,53 y 40,75 q ha⁻¹) constituyeron la muestra de rendimiento en condiciones de alta productividad.

En el histograma E de la Figura 5.2, donde se representa la distribución del rendimiento de los genotipos de los dos grupos en ambientes de alta productividad, se observó una clara superioridad de las variedades comerciales. De hecho, más del 80% de ellas se concentran en las tres clases de mayor rendimiento, por sólo un 10% de las *landraces*. Sin embargo, en los ambientes de baja productividad, la distribución de las *landraces* y de las variedades comerciales fue mucho más igualada, incluso con una ligera superioridad de las primeras (Figura 5.2, histograma F).

5-3-2-5. Peso hectolítrico

Los promedios de las variedades comerciales fueron de 68,0 kg hl⁻¹ (SAD03) a 71,7 kg hl⁻¹ (VLD02), mientras que los promedios de las *landraces* oscilaron entre 62,4 kg hl⁻¹ (ART02) y 65,6 kg hl⁻¹ (VLD02, VLD03). Las diferencias entre los dos grupos se destacaron más en los ensayos ART02 y ART04, en los cuales las variedades superaron a las *landraces* en 7,0 y 7,1 kg hl⁻¹, respectivamente. El histograma G de la Figura 5.2 muestra la distribución de los genotipos de estos dos grupos. Las dos distribuciones están muy separadas, con más del 70% de las variedades con valores superiores a 69 kg hl⁻¹, mientras que el 75% de las *landraces* presentaron un peso

hectolítrico inferior a 65 kg hl⁻¹. Sin embargo, la distribución de los valores de las *landraces* fue muy amplia, mostrando desde valores de peso hectolítrico muy bajos hasta valores elevados, similares a los de las mejores variedades comerciales.

5-3-2-6. Peso de mil granos

Los pesos de mil granos más bajos se observaron en el ensayo FOR03, mientras que los valores más altos se registraron en el ensayo VLD02. Las *landraces* presentaron consistentemente unos promedios de peso de 1000 granos más altos que las variedades comerciales. Esto se notó particularmente en el ensayo VED03, donde el peso de 1000 granos de las *landraces* fue en promedio 4,8 g más alto que el de las variedades comerciales. Este resultado fue inesperado, pues los genotipos de dos carreras suelen tener mayor peso individual de grano que los de seis carreras, y la proporción de genotipos de dos carreras es mucho mayor en las variedades comerciales que en las *landraces*.

Considerando todos los ensayos donde se evaluó este carácter, se observó que el 70% de las variedades comerciales tienen un peso de 1000 granos que osciló entre 35,2 y 42,6 g. Sin embargo el 80% de las *landraces* presentó valores superiores, entre 37,7 y 47,7 g (histograma A de la Figura 5.3).

Tabla 5.6. Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables agronómicas (rendimiento, espigado, altura, encamado, peso hectolítrico y peso de mil granos) observados en los dos grupos estudiados: *landraces* y variedades comerciales, en los diferentes ensayos.

Ensayo	Grupos	Variables																							
		Espigado (días)				Altura (cm)				Encamado				Rendimiento (q ha ⁻¹)				Peso hectolítrico (kg hl ⁻¹)				Peso de 1000 granos (g)			
		Media	Sd	Min	Max	Media	Sd	Min	Max	Media	Sd	Min	Max	Media	Sd	Min	Max	Media	Sd	Min	Max	Media	Sd	Min	Max
ART02	<i>Landraces</i>	121,4	5,2	111,0	132,5	84,7	6,5	66,9	98,9	-	-	-	-	20,93	3,17	13,65	29,48	62,4	2,9	56,8	73,7	39,8	5,1	26,8	52,4
	V. comerciales	119,1	4,0	110,5	126,0	74,0	10,1	56,8	96,5	-	-	-	-	27,78	2,74	18,35	32,11	69,4	3,5	61,9	74,6	40,0	4,9	30,1	50,3
BEL02	<i>Landraces</i>	112,1	4,7	97,4	121,7	-	-	-	-	6,5	1,3	1,0	8,7	28,50	4,95	16,36	43,97	63,5	3,0	56,8	72,9	39,8	5,0	27,2	55,9
	V. comerciales	106,3	4,6	97,9	114,1	-	-	-	-	1,7	1,8	1,0	8,0	35,23	4,36	26,65	43,17	68,7	3,5	60,5	76,2	37,9	4,9	30,0	51,8
SAD02	<i>Landraces</i>	-	-	-	-	61,8	5,0	46,9	81,7	-	-	-	-	17,01	2,81	6,08	24,54	-	-	-	-	-	-	-	-
	V. comerciales	-	-	-	-	50,1	7,3	38,4	64,1	-	-	-	-	18,94	2,44	13,59	23,78	-	-	-	-	-	-	-	-
VLD02	<i>Landraces</i>	124,0	5,1	113,6	137,4	89,5	6,7	66,4	119,0	-	-	-	-	38,36	5,91	20,35	51,02	65,6	3,0	56,3	73,4	46,2	3,7	35,7	53,9
	V. comerciales	119,2	3,2	112,7	128,6	70,8	12,4	51,8	94,2	-	-	-	-	40,75	4,81	28,14	47,51	71,1	3,5	60,3	75,3	44,4	3,3	37,6	50,5
FOR03	<i>Landraces</i>	116,8	3,5	106,7	130,0	-	-	-	-	5,2	2,7	1,0	9,0	16,69	2,33	9,75	21,75	63,6	2,6	55,4	72,2	37,9	3,6	28,4	45,4
	V. comerciales	114,1	2,6	108,4	118,6	-	-	-	-	2,0	1,8	1,0	8,0	19,18	2,52	14,05	24,28	69,3	3,3	62,9	76,3	36,9	3,7	29,9	45,2
VLD03	<i>Landraces</i>	123,4	6,2	109,0	135,0	92,3	11,4	56,0	116,0	5,2	2,5	1,0	9,0	25,82	4,00	9,97	35,38	65,6	2,3	58,0	71,6	44,0	5,0	30,9	54,3
	V. comerciales	117,6	4,4	108,0	127,0	67,0	14,1	48,0	95,0	1,3	1,2	1,0	7,0	22,47	3,13	15,72	28,25	69,3	2,4	63,7	73,1	40,3	5,3	32,7	53,5
SAD03	<i>Landraces</i>	-	-	-	-	76,2	6,1	58,5	95,8	-	-	-	-	18,61	3,31	6,92	26,57	62,9	2,5	55,0	70,9	-	-	-	-
	V. comerciales	-	-	-	-	61,9	8,3	49,8	78,5	-	-	-	-	16,12	3,28	8,77	21,29	68,0	2,8	61,0	73,5	-	-	-	-
VED03	<i>Landraces</i>	113,2	4,4	102,3	123,1	81,8	6,7	64,3	102,7	4,0	2,0	1,0	9,0	28,90	3,58	17,35	36,12	64,8	3,0	57,5	75,1	45,1	5,1	30,5	56,6
	V. comerciales	109,5	3,5	104,6	116,0	67,6	11,5	50,8	89,2	2,0	2,0	1,0	8,3	27,71	3,31	20,68	34,25	69,4	3,5	61,8	75,4	40,3	5,2	31,0	53,9
ART04	<i>Landraces</i>	120,0	4,0	112,0	129,5	-	-	-	-	-	-	-	-	26,45	3,13	16,22	35,16	63,5	3,2	55,5	72,5	39,6	4,3	30,2	53,2
	V. comerciales	116,6	2,5	112,3	122,0	-	-	-	-	-	-	-	-	38,43	5,97	23,43	47,46	70,6	4,4	62,4	79,7	38,4	4,2	26,8	48,1
SAD04	<i>Landraces</i>	-	-	-	-	-	-	-	-	-	-	-	-	34,94	4,48	18,51	46,53	63,0	3,4	54,2	73,5	40,6	4,9	27,8	53,6
	V. comerciales	-	-	-	-	-	-	-	-	-	-	-	-	37,05	4,84	24,55	44,31	69,3	3,5	58,1	74,2	38,3	5,2	30,6	49,6

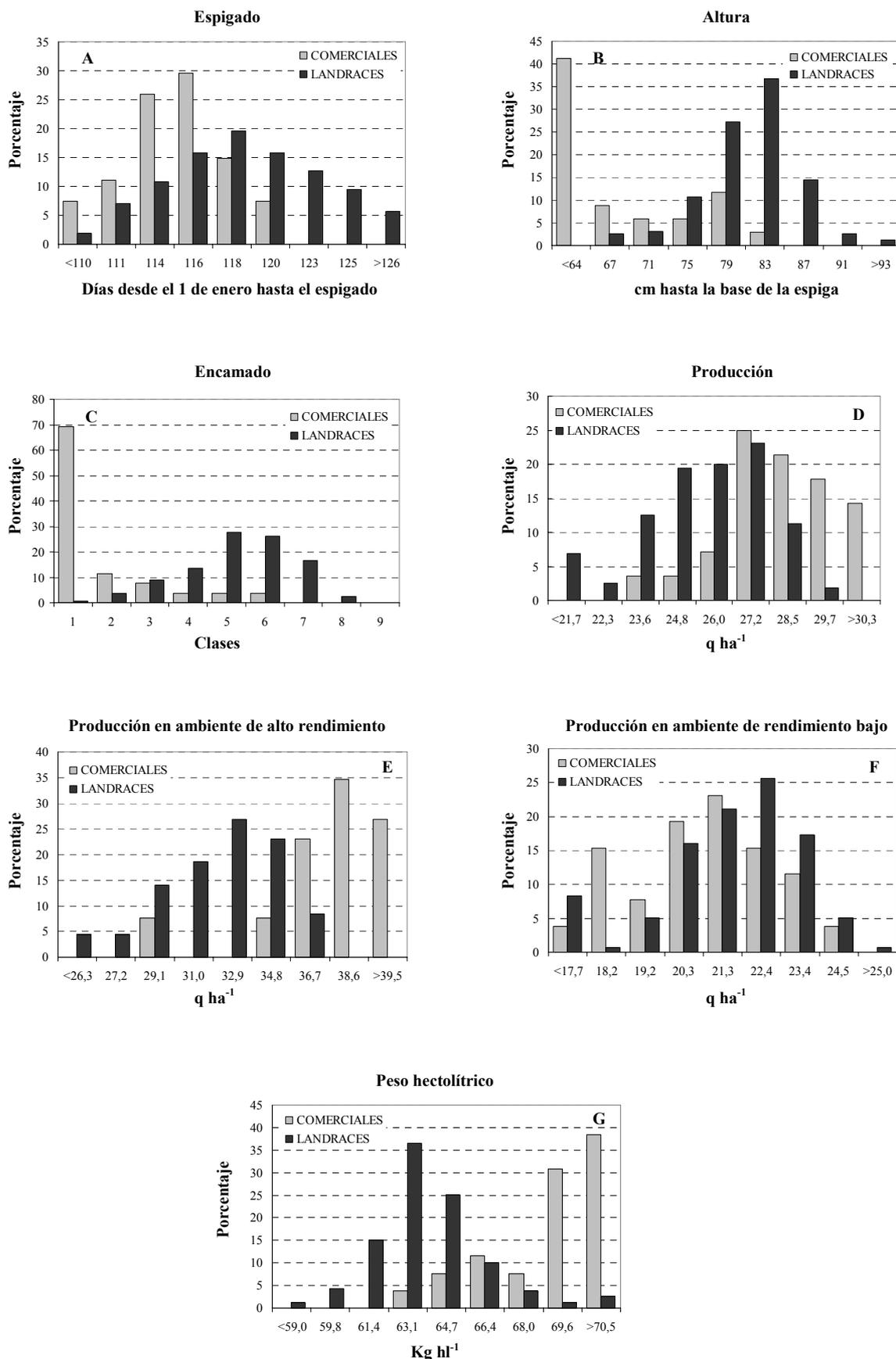


Figura 5.2. Distribución de los valores (en porcentajes) de los caracteres: espigado, altura, encamado, rendimiento, rendimiento en ambiente de alta productividad, rendimiento en ambiente de baja productividad y peso hectolítrico, para las variedades comerciales y *landraces*. Cada clase de los histogramas es de una amplitud de 0,5 desviaciones típicas a partir de la media para las variables cuantitativas y de 1 unidad para la variable encamado

5-3-2-7. Componentes de rendimiento

Espigas m⁻²: el número de espigas m⁻² fue más alto en las variedades comerciales en comparación con las *landraces*. Estas últimas presentaron un promedio de 255 espigas m⁻², mientras que el promedio de las variedades comerciales fue de 319 (Tabla 5.7). La representación en histograma de la distribución de los genotipos de los dos grupos en función del número de espigas confirma este hecho. El 70% de las *landraces* mostraron un número de espigas m⁻² entre 189 y 301. Sin embargo, la mayoría de las variedades comerciales mostraron un número de espigas m⁻² que osciló entre 301 y 396 (histograma B, Figura 5.3).

Granos m⁻²: según el histograma C de la Figura 5.3, la distribución de los genotipos de los dos grupos fue parecida. Aunque las *landraces* presentaron un promedio relativamente más bajo que las variedades comerciales (Tabla 5.7), debido a un cierto número de *landraces* con valores muy bajos (histograma C, Figura 5.3).

Granos por espiga: aunque el promedio de esta variable fue parecido en los dos grupos, las *landraces* superaban ligeramente a las variedades comerciales. De hecho, en las primeras el número de granos por espiga variaba de 15 a 65, mientras que las variedades comerciales presentaron valores entre 16 y 41 (Tabla 5.7). El histograma D de la Figura 5.3, ilustra la distribución de los genotipos de ambos grupos. Se observa que el 60% de las variedades presentan valores entre 17 y 24, Sin embargo la mayoría de las *landraces* se distribuyen entre 24 y 30. Este carácter está muy probablemente influido por la desproporción de genotipos de dos y seis carreras entre *landraces* y variedades comerciales.

Tabla 5.7. Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables de componentes de rendimiento (espigas m⁻², granos m⁻², granos por espiga y peso de mil granos) observados en los dos grupos estudiados: *landraces* y variedades comerciales, en el ensayo VED03.

Variables		<i>Landraces</i>	V. comerciales
Espigas m⁻²	Media	255,0	319,0
	Sd	69,0	87,0
	Min	113,0	180,0
	Max	483,0	490,0
Granos m⁻²	Media	6813,0	7536,0
	Sd	1861,0	2099,0
	Min	2973,0	4493,0
	Max	11610,0	12500,0
Granos por espiga	Media	27,0	24,0
	Sd	6,0	7,0
	Min	15,0	16,0
	Max	64,0	41,0
Peso 1000 granos (g)	Media	45,1	40,3
	Sd	5,1	5,2
	Min	30,5	31,0
	Max	56,6	53,9

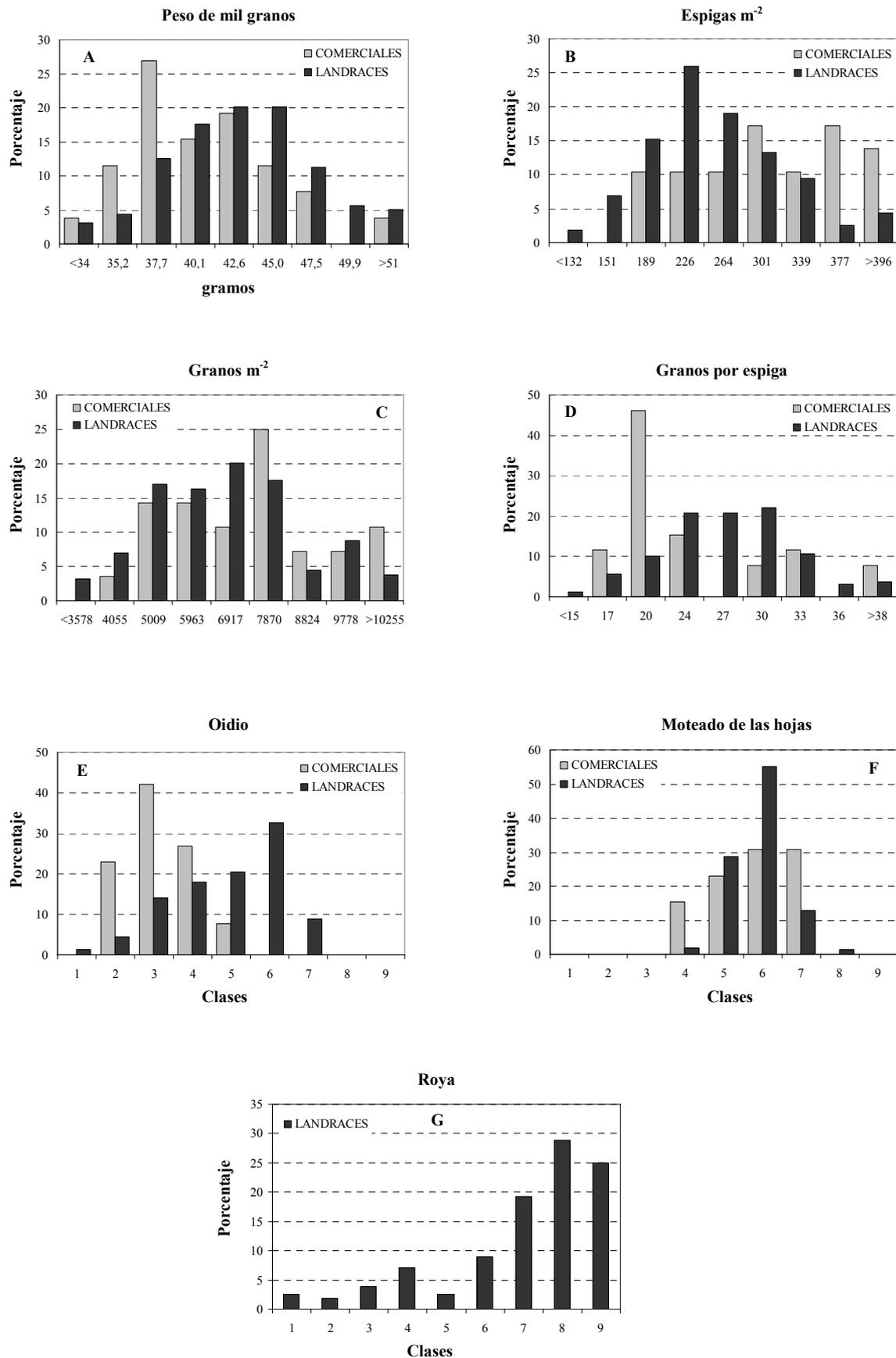


Figura 5.3. Distribución de los valores (en porcentajes) de los caracteres: peso de mil granos, espigas m², granos m², granos por espiga, susceptibilidad a oídio, moteado de las hojas y roya para las variedades comerciales y *landraces*. Cada clase de los histogramas es de una amplitud de 0,5 desviaciones típicas a partir de la media para las variables cuantitativas y de 1 unidad para las variables reacción a oídio, moteado de las hojas y roya.

5-3-2-8. Enfermedades

Oídio: los ataques de oídio fueron aparentemente más fuertes en los ensayos del 2000 de Gerona y Valladolid en comparación a los ensayos BEL02 y SAD02. En estos dos últimos, se observa que los promedios de los dos grupos comparados, son parecidos (Tabla 5.8). Sin embargo, el histograma E de la Figura 5.3 muestra que las *landraces* son en general más susceptibles que las variedades comerciales. Más del 90% de estas últimas, se encuentran en valores entre 2 y 4. Contrariamente, la mayoría de *landraces* se sitúan entre los valores 4 y 7 de la escala de evaluación (histograma E, Figura 5.3).

Tabla 5.8. Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables de enfermedades (oídio, moteado de las hojas y roya) observados en los dos grupos estudiados: *landraces* y variedades comerciales, en los diferentes ensayos.

Ensayo	Grupos	Oídio				Moteado de las hojas				Roya			
		Media	Sd	Min	Max	Media	Sd	Min	Max	Media	Sd	Min	Max
ART02	<i>Landraces</i>	-	-	-	-	3,5	0,9	1,3	6,4	-	-	-	-
	V. comerciales	-	-	-	-	3,3	1,2	1,3	5,7	-	-	-	-
BEL02	<i>Landraces</i>	3,9	1,6	0,3	7,5	-	-	-	-	-	-	-	-
	V. comerciales	3,1	1,4	0,8	5,4	-	-	-	-	-	-	-	-
SAD02	<i>Landraces</i>	3,1	0,9	1,0	5,0	-	-	-	-	-	-	-	-
	V. comerciales	3,0	0,8	1,8	5,0	-	-	-	-	-	-	-	-
FOR03	<i>Landraces</i>	-	-	-	-	8,1	0,8	4,3	9,2	-	-	-	-
	V. comerciales	-	-	-	-	7,9	1,1	5,1	9,2	-	-	-	-
GER00	<i>Landraces</i>	6,3	2,8	1,0	9,0	-	-	-	-	7,0	2,0	1,0	9,0
	V. comerciales	-	-	-	-	-	-	-	-	-	-	-	-
VLD00	<i>Landraces</i>	5,9	1,6	1,0	8,0	-	-	-	-	-	-	-	-
	V. comerciales	-	-	-	-	-	-	-	-	-	-	-	-

Moteado de las hojas: el hongo *Pyrenophora teres* provocó daños en los ensayos ART02 y FOR03. El ataque fue aparentemente más acentuado en éste último ambiente. Los promedios anotados para las *landraces* y variedades comerciales en el ensayo FOR03 fueron, respectivamente, de 8,1 y 7,9 (Tabla 5.8). La susceptibilidad de las *landraces* así como las variedades comerciales parece ser similar. En el histograma F de la Figura 5.3 no se aprecian diferencias en la distribución de los genotipos de los dos grupos.

Roya: el hongo *Puccinia hordei* se manifestó particularmente en el ensayo de Gerona de 2000, donde fueron evaluadas únicamente las *landraces*. Estas se revelaron muy susceptibles a esta enfermedad. El promedio observado fue de 7 (Tabla 5.8) y

aproximadamente el 70% de los genotipos presentaron unos valores muy altos, entre 7 y 9, según la escala de evaluación (histograma G, Figura 5.3).

5-3-2-9. Calidad maltera

Calibre de los granos: las variedades comerciales se destacan por el tamaño grande del grano contrariamente a las *landraces*, que presentaron granos más pequeños. La media del porcentaje de granos que no pasaron por una criba de 2,5 mm en las *landraces* fue menor que en las variedades comerciales (Tabla 5.9). También este carácter pudo verse afectado por el desequilibrio entre genotipos de dos y seis carreras.

Los histogramas A y B de la Figura 5.4 muestran como se distribuyeron los genotipos de ambos grupos, en función del porcentaje de granos que pasaron la criba de 2,2 mm y del porcentaje de granos que no pasaron la criba de 2,5 mm. En el primero, se observa que la mayoría de las variedades comerciales tienen un porcentaje de granos que pasan la criba de 2,2 mm, que se sitúa entre valores de 1,5 y 2,3, mientras que más del 40% de las *landraces* presentan valores superiores a 2,3. El histograma B confirma claramente la superioridad de las variedades comerciales para estos caracteres, pues su porcentaje de granos retenidos por la criba de 2,5 mm fue de 85% en más de la mitad de los genotipos, valor muy superior al de las *landraces*.

Porcentaje de materia seca en el extracto de malta: el porcentaje promedio de materia seca en el extracto de malta fue ligeramente más alto en las variedades comerciales (79,3%) en comparación con las *landraces* (73,4%). El 90% de estas últimas se situó entre 74 y 78%, mientras que la mayoría de las variedades comerciales presentaron un porcentaje de 78% (histograma C, Figura 5.4).

Tabla 5.9. Parámetros estadísticos (Media, Sd: desviación estándar, Min: mínimo, Max: máximo) de las variables de calidad maltera (calibre <2,2 mm, calibre >2,5mm, materia seca en el extracto de malta y proteína total en la malta) observados en los dos grupos estudiados: *landraces* y variedades comerciales, en el ensayo de Foradada 2003.

Variables		<i>Landraces</i>	Variedades comerciales
Calibre <2,2 mm (%)	Media	4,1	2,3
	Sd	3,3	1,3
	Min	0,8	0,4
	Max	20,5	5,7
Calibre >2,5 mm (%)	Media	69,3	79,3
	Sd	14,8	9,6
	Min	17,3	56,5
	Max	91,3	91,9
Materia seca (%)	Media	73,4	78,7
	Sd	9,5	2,1
	Min	35,5	74,8
	Max	81,5	82,2
Proteína (%)	Media	11,5	11,1
	Sd	1,1	1,4
	Min	9,4	9,3
	Max	14,9	12,8

Proteína total en malta: los porcentajes de proteína en la harina de malta fueron parecidos entre las *landraces* y las variedades comerciales. Los porcentajes promedio observados fueron respectivamente de 11,5 y 11,1. Sin embargo el 20% de las variedades presentaron unos porcentajes inferiores a 9,4 (Figura 5.4, histograma D).

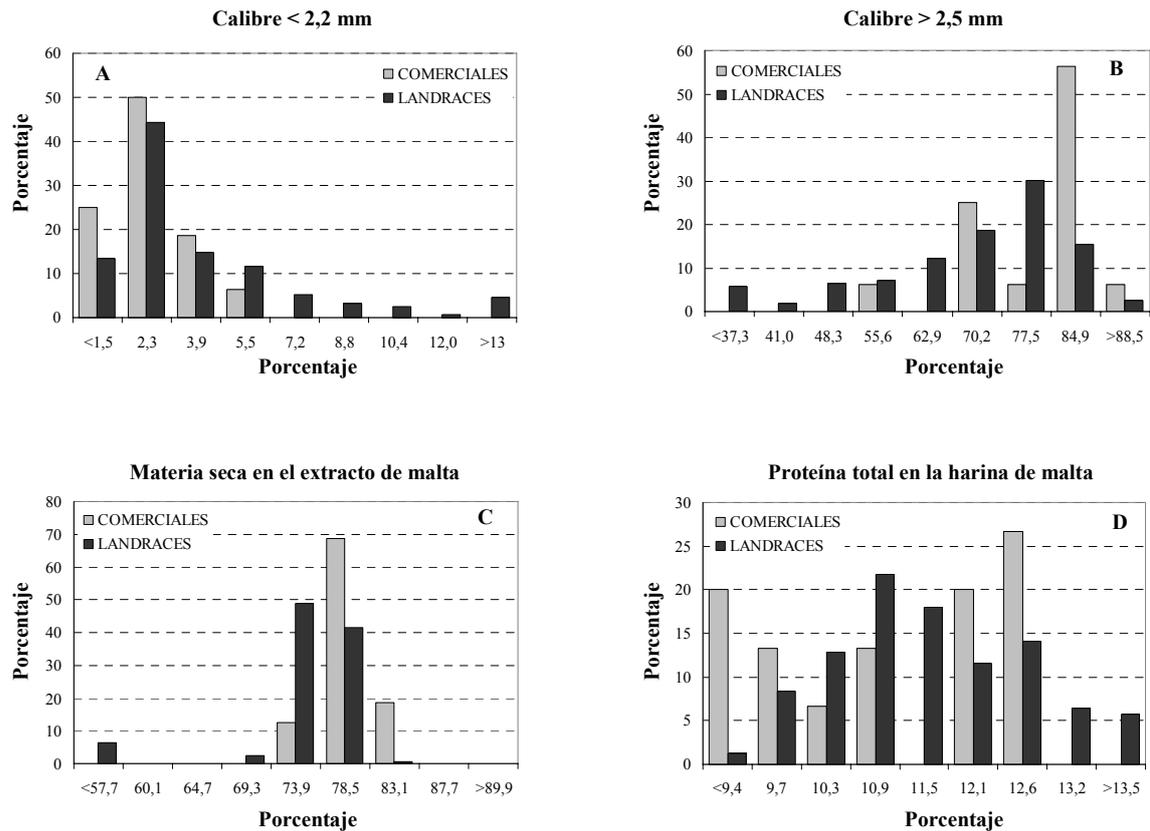


Figura 5.4. Distribución de los valores (en porcentajes) de los caracteres calibre 2,2 mm, calibre superior a 2,5 mm, materia seca en extracto y proteína total en la harina de malta, para las variedades comerciales y *landraces*. Cada clase de los histogramas es de una amplitud de 0,5 desviaciones típicas a partir de la media.

5-3-3. Análisis combinado de los ensayos

Este análisis se ha llevado a cabo para todos los caracteres estudiados en varios ambientes, con objeto de evaluar la importancia relativa de los factores genotipo y ensayo, así como la interacción entre ellos. También se evaluó la varianza dentro y entre los siete grupos de genotipos definidos en el apartado 5-2-4-2.b, de este capítulo, así como la interacción genotipo por ensayo entre y dentro de grupos.

5-3-3-1. Caracteres agronómicos

Los efectos ensayo y genotipo, así como la interacción genotipo x ensayo (GxE), fueron altamente significativos ($P < 0,01$) para las variables espigado, altura, encamado, rendimiento, peso hectolítrico y peso de mil granos (Tablas 5.10a y 5.11a).

La división del factor genotipo en grupos mostró que el valor de los cuadrados medios entre grupos para todos estos caracteres fue mucho mayor que el cuadrado medio dentro de grupos, aunque ambos componentes fueron significativos. Lo mismo se puede decir en cuanto a la interacción

La separación de medias de los grupos permitió identificar 4 a 6 valores distintos, según el carácter analizado. En general, las mayores diferencias se produjeron entre los grupos de variedades comerciales (números 5 a 7) y los de *landraces* (números 1 a 4). Pero también se observaron diferencias significativas en todos los caracteres entre los grupos de *landraces*, demostrando que los grupos definidos mediante marcadores moleculares muestran también respuestas agronómicas diferenciales.

Para la variable espigado, se distinguieron 4 promedios distintos entre los 7 grupos (Tabla 5.10b). El espigado fue más tardío en las *landraces*, particularmente en el grupo 1. Los genotipos más tempranos fueron las variedades comerciales antiguas de 2 carreras (grupo 5).

Para el carácter altura de planta, se diferenciaron 5 promedios distintos. La altura más baja (59,8 cm) se observó en el grupo 7 (variedades testigo) seguido por el grupo de variedades antiguas de 2 carreras (grupo 5). Con la excepción del grupo 2 (fundamentalmente, genotipos de 2 carreras), las *landraces* presentaron unos promedios de altura superiores a 80 cm (Tabla 5.10b).

Se detectaron 4 niveles de promedios en los grupos para la variable encamado. Las variedades testigo (grupo 7) así como las variedades antiguas de 2 carreras (grupo 5) prácticamente no se encamaron. Contrariamente, el encamado fue muy acusado en los dos grupos principales de cebadas de seis carreras autóctonas (3 y 4). Las variedades comerciales antiguas de 6 carreras (grupo 6), así como los grupos 1 y 2 de las *landraces* presentaron valores intermedios (Tabla 5.10b).

Tabla 5.10a. Cuadrados medios de los análisis de varianza para las variables espigado (días desde el 1 de enero), altura de planta (cm) y encamado (valoración visual) evaluados en los diferentes ensayos de campo.

Fuente	Espigado			Altura			Encamado		
	gl	CM	**	gl	CM	**	gl	CM	**
Ensayo (E)	6	10528,0	**	5	53550,9	**	3	414,5	**
Genotipo (G)	181	360,5	**	181	1047,2	**	181	41,2	**
entre grupos	6	2867,7	**	6	19626,3	**	6	676,7	**
dentro de grupos	175	274,5	**	175	410,2	**	175	19,4	**
GXE	1086	12,9	**	905	99,0	**	543	9,8	**
GXE entre grupos	36	52,4	**	30	638,7	**	18	27,7	**
GXE dentro de grupos	1050	11,5	**	875	80,3	**	525	9,2	**
Error	2002	2,8		1820	21,1		1456	2,6	

** Significativo al nivel P <0,01.

Tabla 5.10b. Separación de medias de los grupos analizados, para las variables espigado (días desde 1 de enero), altura de planta (cm) y encamado.

Grupos	Espigado		Altura		Encamado	
	Promedio*		Promedio		Promedio	
1	122,9	A	85,4	A	4,0	B
2	118,5	B	73,7	C	3,6	BC
3	118,8	B	81,4	B	5,5	A
4	116,9	C	80,8	B	5,5	A
5	112,5	E	63,0	D	1,4	D
6	114,2	D	74,1	C	3,0	C
7	114,7	D	59,8	E	1,1	D

*medias seguidas por la misma letra no son significativamente distintas para $P < 0,05$.

La separación de medias del rendimiento distinguió 4 niveles de rendimiento entre los grupos. Los testigos (grupo 7) mostraron el promedio más alto. El rendimiento más bajo fue observado para el grupo 1. El rendimiento promedio de las variedades comerciales antiguas de 6 carreras fue más alto que el de las de 2 carreras (grupo 6 vs. grupo 5). Lo mismo ocurrió entre los principales grupos de *landraces* de la colección nuclear (grupos 3 y 4 vs. grupo 2, Tabla 5.11b).

Para el peso hectolítrico se distinguieron 6 niveles de promedios distintos. El mejor peso hectolítrico fue observado en las variedades de 2 carreras (grupo 5), seguido por los testigos y el grupo 2 de las *landraces*, que contiene entradas de 2 carreras. Los grupos 3 y 4 presentaron unos pesos hectolítricos muy bajos (Tabla 5.11b). Las entradas de 6 carreras mostraron unos pesos de mil granos muy altos. Este fue todavía mayor en las *landraces* que en las variedades comerciales. Los grupos 2, 5 y 7 presentaron promedios de peso de mil granos menores y similares entre ellos. El grupo 1 presentó el peso de mil granos claramente más bajo (Tabla 5.11b).

Tabla 5.11a. Cuadrados medios de los análisis de varianza para las variables rendimiento ($q\ ha^{-1}$), peso hectolítrico ($kg\ hl^{-1}$) y peso de mil granos (g), evaluados en los diferentes ensayos de campo.

Fuente	Rendimiento			Peso hectolítrico			Peso mil granos		
	gl	CM	**	gl	CM	**	gl	CM	**
Ensayo (E)	8	33116,4	**	8	451,3	**	7	2870,2	**
Genotipo (G)	181	162,1	**	181	175,8	**	181	229,1	**
entre grupos	6	1868,0	**	6	3733,1	**	6	2293,9	**
dentro de grupos	175	103,6	**	175	53,8	**	175	158,3	**
GXE	1448	43,5	**	1448	8,3	**	1265	14,4	**
GXE entre grupos	48	451,4	**	48	38,7	**	42	62,0	**
GXE dentro de grupos	1400	29,5	**	1400	7,3	**	1223	12,8	**
Error	3276	8,1		912	3,6		1274	4,5	

** Significativo al nivel $P < 0,01$.

Tabla 5.11b. Separación de medias de los grupos analizados, para las variables rendimiento (q ha⁻¹), peso hectolítrico (kg hl⁻¹) y peso de mil granos (g).

Grupos	Rendimiento		Peso hectolítrico		Peso mil granos	
	Promedio*		Promedio		Promedio	
1	23,27	D	65,5	D	35,8	D
2	24,01	D	69,3	B	38,5	C
3	25,96	C	62,7	F	42,6	A
4	27,08	BC	63,6	E	42,6	A
5	26,75	C	72,1	A	38,7	C
6	28,27	B	67,0	C	41,0	B
7	29,90	A	69,3	B	39,1	C

*medias seguidas por la misma letra no son significativamente distintas para P <0,05.

El análisis de varianza de la variable rendimiento se realizó para los ambientes de productividad baja y alta, definidos en el apartado 5-3-2-4. Los efectos genotipo, ensayo y la interacción GxE fueron altamente significativos en los dos tipos de ambientes. Al analizar la varianza entre y dentro de los grupos de genotipos, se observó que ésta fue mayor entre grupos que dentro de grupos. Sin embargo esta varianza fue significativa en los dos casos (Tabla 5.12a).

También se analizó la varianza del efecto GxE entre y dentro de grupos. Esta interacción se reveló significativa entre y dentro de grupos, en los dos tipos de ambientes.

La separación de medias diferenció 5 niveles entre los grupos en los ambientes de alta productividad. Los testigos se destacaron con un rendimiento más alto, seguidos por las variedades de 2 y 6 carreras. En este tipo de ambientes, los promedios de las *landraces* fueron bajos. Dentro de las *landraces*, el grupo 4 fue el que presentó mayor rendimiento. En los ambientes de baja productividad, la separación de medias, diferenció solamente dos clases de grupos. Las *landraces* de los grupos 3 y 4 presentaron rendimientos similares a los testigos y a las variedades de 6 carreras (grupo 7 y 6, respectivamente). Las variedades de 2 carreras (grupo 5) mostraron rendimientos relativamente bajos en los dos tipos de ambientes. Sin embargo, mostraron una tendencia similar a los grupos de entradas de seis carreras. El grupo 5 (variedades de dos carreras de la colección nuclear) fue significativamente mejor que el 2 (compuesto sólo por *landraces* de dos carreras) en los ambientes de alta productividad, mientras que en los de baja productividad, estos dos grupos presentaron producciones similares (Tabla 5.12b).

Tabla 5.12a. Cuadrados medios de los análisis de varianza para la variable rendimiento ($q\text{ ha}^{-1}$), evaluada en los diferentes ensayos de los 2 ambientes de alta y baja productividad.

Fuente	Rendimiento					
	Ambientes de alta productividad			Ambientes de baja productividad		
	gl	CM		gl	CM	
Ensayo (E)	3	13341,3	**	4	15690,5	**
Genotipo (G)	181	171,9	**	181	66,2	**
entre grupos	6	2367,9	**	6	636,3	**
dentro de grupos	175	96,6	**	175	46,6	**
GXE	543	57,6	**	724	24,7	**
GXE entre grupos	18	561,2	**	24	197,8	**
GXE dentro de grupos	525	40,3	**	700	18,8	**
Error	1456	10,01		1820	6,6	

** Significativo al nivel $P < 0,01$.

Tabla 5.12b. Separación de medias de los grupos analizados, para la variable rendimiento ($q\text{ ha}^{-1}$), en ambientes de alta y baja productividad.

Grupos	Alta productividad		Baja productividad	
	Promedio*		Promedio	
1	28,8	E	18,86	B
2	30,4	DE	18,90	B
3	31,4	D	21,63	A
4	33,3	C	22,08	A
5	36,5	B	18,97	B
6	36,6	B	21,59	A
7	40,0	A	21,84	A

*medias seguidas por la misma letra no son significativamente distintas para $P < 0,05$.

En cuanto a los componentes de rendimiento, evaluados en el ensayo VED03, las diferencias entre grupos fueron altamente significativas para las variables espigas m^{-2} , granos por espiga, peso de mil granos y no significativa para la variable granos m^{-2} (Tabla 5.13a).

La separación de medias distinguió 3 grupos que se solapan para la variable espigas m^{-2} . Un factor que influye notablemente en las variables espigas m^{-2} y granos por espiga es el tipo de espiga de los genotipos. Esto se tratará con más detalle en un apartado posterior. El grupo con *landraces* de 2 carreras presentó un número más alto de espigas m^{-2} . En cuanto a los grupos de seis carreras, el grupo 1, es el que menor número de granos por espiga presentó, seguido por el grupo 3. En cuanto a la variable granos por espiga se diferenciaron 4 grupos. Los grupos 2 y 5 presentaron promedios igualados, con un bajo número de granos por espiga. El número más alto de granos por espiga correspondió al grupo 1. Las entradas de 6 carreras de los grupos 3 y 4 mostraron unos promedios altos, que no se diferenciaron de los testigos y de las variedades comerciales antiguas de 6 carreras (Tabla 5.13b). La separación de medias para la

variable peso de mil granos, evaluada separadamente para este ensayo, mostró 4 grupos que se solapan (Tabla 5.13b). Los pesos de mil granos más altos se registraron en los dos grupos de *landraces* 3 y 4, mientras que el grupo 1, junto con las variedades de 2 carreras (grupo 5) mostraron los promedios más bajos.

Tabla 5.13a. Cuadrado medios de los análisis de varianza para las variables espigas m⁻², granos m⁻², y granos por espiga y peso de mil granos evaluados en el ensayo del Vedado 2003.

Fuente	gl	Espigas m ⁻²	Granos m ⁻²	Granos por espiga	Peso mil granos†
		CM	CM	CM	CM
Grupos	6	81506 **	8054221 ns	584 **	636 **
Error	175	8865	7249772	64	37

ns: no significativo; **: significativo al nivel P <0,01.

† se analizó el peso de mil granos del ensayo VED03, por mostrar la variación de los componentes de rendimiento en el mismo ensayo.

Tabla 5.13b. Separación de medias de los grupos analizados, para las variables espigas m⁻², granos m⁻², y granos por espiga evaluados en el ensayo del Vedado 2003.

Grupos	Espigas m ⁻²		Granos m ⁻²		Granos por espiga		Peso mil granos	
	Promedio*		Promedio		Promedio		Promedio	
1	213	C	6844	A	32,4	A	38.5	D
2	340	A	6297	A	18,7	D	39.4	CD
3	246	C	6506	A	26,3	B	47.3	A
4	259	BC	7043	A	27,8	AB	45.9	AB
5	374	A	7236	A	19,1	CD	38.9	D
6	259	BC	7558	A	28,5	AB	42.6	BC
7	323	AB	7759	A	24,6	BC	39.5	CD

*medias seguidas por la misma letra no son significativamente distintas para P <0,05.

5-3-3-2. Enfermedades

El análisis de varianza para el oídio y el moteado de las hojas reveló la existencia de diferencias significativas entre ensayos y genotipos. La interacción GxE fue también significativa. Se observó que las diferencias entre genotipos están más concentradas entre grupos que dentro de grupos. Asimismo, la interacción genotipo por ambiente se debe más a la interacción entre grupos que dentro de grupos (Tabla 5.14a).

La separación de medias de los grupos analizados reveló que los genotipos del grupo 3 son los más susceptibles al oídio. Por otro lado, las variedades son más resistentes en general, particularmente el grupo de testigos y las variedades de 2 carreras. En cuanto al moteado de las hojas, las variedades de 6 carreras son relativamente las más afectadas, seguidas por el grupo 3 de la colección nuclear. En el resto de los genotipos la susceptibilidad es prácticamente la misma (Tabla 5.14b).

Para la roya, el análisis de varianza sugiere que existen diferencias significativas entre los 4 grupos de genotipos de la colección nuclear (Tabla 5.14a). La separación de medias mostró que el grupo 4 es relativamente menos susceptible que los otros 3 grupos (Tabla 5.14b).

Tabla 5.14a. Cuadrado medios de los análisis de varianza para las variables oídio, moteado de las hojas y roya evaluadas en los diferentes ensayos.

Fuente	Oídio			Moteado de las hojas			Roya	
	gl	CM		gl	CM		GI	CM
Ensayo (E)	3	619,3	**	1	5742,7	**		
Genotipo (G)	181	10,4	**	181	3,2	**		
entre grupos	6	127,3	**	6	13,2	**	3	44,0
dentro de grupos	175	6,4	**	175	2,8	**	152	3,4†
GXE	491	2,7	**	181	1,5	**		
GXE entre grupos	12	10,4	**	6	3,2	**		
GXE dentro de grupos	479	2,6	**	175	1,4	**		
Error	284	1,7		728	0,8			

** Significativo al nivel $P < 0,01$.

†ante la ausencia de repeticiones para esta variable, se empleó como error el cuadrado medio de la varianza entre grupos

Tabla 5.14b. Separación de medias de los grupos analizados, para las variables oídio, moteado de las hojas y roya.

Grupos	Oídio		Moteado de las hojas		Roya	
	Promedio*		Promedio		Promedio	
1	4,6	B	5,7	BCD	7,4	AB
2	4,5	B	4,9	E	8,1	A
3	5,9	A	5,9	AB	8,2	A
4	4,2	B	5,8	BC	6,2	B
5	2,9	C	5,2	DE		
6	3,9	B	6,3	A		
7	2,5	C	5,3	CDE		

*medias seguidas por la misma letra no son significativamente distintas para $P < 0,05$.

5-3-3-3. Calidad maltera

Estas variables fueron evaluadas para los genotipos de la colección nuclear, o sea, excluyendo los testigos (grupo 7). Para estimarlas, se juntó la semilla de las tres repeticiones del ensayo FOR03 en una única muestra. Por tanto, no disponemos de una verdadera estimación del error y no puede evaluarse la significación de las diferencias entre genotipos. Como alternativa, se evaluó la significación de la diferencia entre grupos, empleando como error el cuadrado medio de las diferencias dentro de grupos. La diferencia entre grupos fue altamente significativa para las variables calibre de granos (Calibre menor que 2,2 mm y Calibre mayor que 2,5 mm) y porcentaje de

proteína total en harina de malta. Sin embargo, no hubo diferencias significativas entre grupos para el porcentaje de materia seca en el extracto de malta (Tabla 5.15a).

Tabla 5.15a. Cuadrado medios de los análisis de varianza para las variables calibre <2,2 mm, calibre > 2,5 mm, materia seca en el extracto de malta, proteína total en harina de malta evaluados en el ensayo de FOR03.

Fuente	gl	Calibre <2,2	Calibre >2,5	MS	Proteína
		CM	CM	CM	
Grupos	5	67,5 **	1350,3 **	181,3 ns	7,9 **
Error	166	8,8	179,9	81,8	1,2

ns: no significativo; **: significativo al nivel P <0,01.

Tabla 5.15b. Separación de medias de los grupos analizados, para las variables calibre <2,2 mm, calibre >2,5 mm, materia seca en el extracto de malta, proteína total en harina de malta evaluados en el ensayo de Foradada 2003.

Grupos	Calibre <2,2		Calibre >2,5		Materia seca		Proteína	
	Promedio*		Promedio		Promedio		Promedio	
1	8,0	A	53,4	C	76,5	A	11,3	AB
2	3,9	B	66,4	BC	78,2	A	12,0	A
3	3,4	B	73,2	AB	73,2	A	10,9	AB
4	3,7	B	70,6	AB	72,3	A	11,9	A
5	2,0	B	81,5	A	79,9	A	11,8	A
6	2,6	B	77,1	AB	77,5	A	10,5	B

*medias seguidas por la misma letra no son significativamente distintas para P <0,05.

La separación de medias de los caracteres que se revelaron significativos mostró que las diferencias para la variable calibre menor que 2,2 mm ocurre entre el grupo 1 y el resto de los grupos. El grupo 1 es el que tiene más granos de tamaño pequeño. Sin embargo, para el carácter calibre mayor que 2,5 mm, el grupo 5 (variedades antiguas de dos carreras) se reveló como el grupo de genotipos que tienen mayores granos. Las variedades comerciales de 6 carreras y los grupos 3 y 4 de la colección nuclear presentaron porcentajes de granos grandes menores que el grupo 5, pero con diferencias no significativas. En cuanto al porcentaje de proteína total en harina de malta, la separación de medias diferenció dos grupos que se solapan. Las diferencias significativas ocurrieron entre el grupo 6 y los grupos 2, 4 y 5. Estos últimos presentaron valores relativamente altos en comparación a los demás (Tabla 5.15b).

5-3-4. Comparación entre genotipos de 2 y 6 carreras

Se realizó una comparación entre genotipos de 2 y 6 carreras mediante un análisis estadístico descriptivo, y separando las variedades comerciales de las *landraces*. Este

análisis se realizó para ayudar en la interpretación de algunos de los caracteres agronómicos, que podrían mostrar una expresión diferencial según el tipo de espiga. Por tanto, es probable que algunas de las diferencias observadas entre los siete grupos de genotipos definidos en el apartado 5-2-4-2.b, se debieran al hecho de que algunos de estos grupos contengan solamente entradas de 2 ó 6 carreras y otros una mezcla de los dos tipos, especialmente el grupo de las variedades testigo (grupo 7), que agrupa 6 variedades de dos carreras y 4 de seis carreras. Para esta comparación se consideraron un total de 4 grupos de germoplasma tal como están definidos en la Tabla 5.16.

Los caracteres considerados en esta comparación son el rendimiento, el peso hectolítrico, los componentes de rendimiento, y el calibre de los granos.

Rendimiento: los genotipos de 6 carreras presentaron unos rendimientos mayores que los de 2 carreras, tanto en la comparación de *landraces* como entre las variedades comerciales. Estas diferencias son más acusadas entre las *landraces*, donde la separación de medias distinguió dos grupos distintos, mientras que las variedades comerciales presentaron un solapamiento entre los dos grupos (2 y 6 carreras) (Tabla 5.16).

Peso hectolítrico: los genotipos de 2 carreras presentan unos promedios claramente mayores que los de 6 carreras, tanto para *landraces* como para variedades comerciales.

Peso de mil granos: los genotipos de 6 carreras presentaron pesos de 1000 granos mayores que los de 2 carreras. Las *landraces* de 2 carreras mostraron valores muy cercanos a las variedades comerciales de 6 carreras. (Tabla 5.16).

Espigas m⁻²: los genotipos de 2 carreras mostraron un número de espigas m⁻² mayor que los de 6 carreras. Las variedades comerciales de 2 carreras presentaron el mayor número de espigas m⁻², pero no se diferenciaron significativamente de las *landraces* de 2 carreras. Lo mismo se observó entre *landraces* y variedades comerciales de 6 carreras (Tabla 5.16).

Tabla 5.16. Parámetros estadísticos (Media, Sd: desviación estándar) de las variables rendimiento (q ha⁻¹), peso hectolítrico (kg hl⁻¹), peso de mil granos (g), espigas m⁻², granos m⁻², granos por espiga, calibre <2,2 mm y calibre >2,5 mm), en los grupos *landraces* y variedades comerciales, según el tipo de espiga 2 ó 6 carreras, considerando el número de ensayos en los que fue tomado cada carácter.

Tipo	Grupos	N. entradas	Variables								
			Rendimiento (q ha ⁻¹)			Peso hectolítrico (kg hl ⁻¹)			Peso de 1000 granos (g)		
			Media*	Sd		Media	Sd		Media	Sd	
2 carreras	<i>Landraces</i>	11	23,9	2,3	C	68,7	2,5	B	39,4	4,1	AB
	V. comerciales	14	27,9	2,1	AB	71,3	1,7	A	39,0	2,8	B
6 carreras	<i>Landraces</i>	145	26,3	2,3	B	63,5	1,7	D	41,8	3,8	A
	V. comerciales	12	29,1	2,1	A	67,2	2,4	C	40,2	5,3	AB
		N. ensayos	9			9			8		

Tipo	Grupos	N. entradas	Variables								
			Espigas m ⁻²			Granos m ⁻²			Granos por espiga		
			Media	Sd		Media	Sd		Media	Sd	
2 carreras	<i>Landraces</i>	11	336,0	94,0	A	6364,0	1894,0	B	19,0	2,0	B
	V. comerciales	14	367,0	75,0	A	7184,0	1542,0	AB	19,0	1,0	B
6 carreras	<i>Landraces</i>	145	249,0	63,0	B	6847,0	1861,0	AB	28,0	6,0	A
	V. comerciales	12	264,0	67,0	B	7948,0	2619,0	A	30,0	7,0	A
		N. ensayos	1			1			1		

Tipo	Grupos	N. entradas	Variables								
			Calibre <2,2 (%)			Calibre >2,5 (%)					
			Media	Sd		Media	Sd				
2 carreras	<i>Landraces</i>	11	4,2	2,6	A	65,0	17,2	B			
	V. comerciales	8	2,0	0,6	A	81,5	6,7	A			
6 carreras	<i>Landraces</i>	145	4,1	3,4	A	69,6	14,6	AB			
	V. comerciales	8	2,6	1,8	A	77,1	11,9	A			
		N. ensayos	1			1					

*medias seguidas por la misma letra no son significativamente distintas para P <0,05.

Granos m⁻²: las entradas de 6 carreras presentaron un número de granos m⁻² relativamente mayor que los de 2 carreras. Las diferencias se observaron principalmente entre las *landraces* de 2 carreras y las variedades comerciales de 6 carreras. (Tabla 5.13b).

Granos por espiga: el número de granos por espiga fue claramente mayor en las entradas de 6 carreras que en las de 2 carreras. Entre grupos de genotipos del mismo tipo de espiga, no se observaron diferencias entre *landraces* y variedades comerciales (Tabla 5.16).

Calibre del grano : no se observaron diferencias significativas entre grupos de 2 y 6 carreras para el porcentaje de granos que pasan la criba de 2,2 mm. Tampoco se notaron diferencias significativas entre genotipos de 2 y 6 carreras, ni entre *landraces* y variedades comerciales, para el porcentaje de granos que no pasan la criba de 2,5 mm. El porcentaje más bajo se observó en las *landraces* de 2 carreras, que se diferenciaron significativamente de las variedades comerciales de 6 carreras (Tabla 5.16).

5-3-5. Análisis de los factores G y GxE del rendimiento

5-3-5-1. AMMI

El análisis de la variable rendimiento mostró una interacción GxE notable. Esta interacción fue analizada mediante el modelo de efectos principales aditivos e interacción multiplicativa (AMMI). Esta interacción fue descompuesta en ejes principales, cuya significación fue comprobada con un análisis de varianza. Éste mostró que los 3 primeros ejes de componentes principales eran significativos, explicando un 48,0%, 28,0% y 15,5% de la suma de cuadrados de la interacción GxE, con unos grados de libertad respectivos de 188, 186 y 184 (Tabla 5.17). El residuo dejado por el modelo no es significativo. En realidad la mayor parte de la interacción (75,94%) está explicada por los dos primeros términos multiplicativos. Existe la posibilidad de que el modelo AMMI3 sobre ajuste (*over-fit*) los datos, ya que una parte del ruido total de la interacción está retenida por el modelo AMMI. Para comprobarlo, se calculó la proporción de ruido retenida en la interacción a partir del error, según el método propuesto por Gauch (1992), multiplicando el CM del error por los grados de libertad de la interacción. En este caso la

cantidad de ruido contenida en la interacción fue de $(8,1 \times 1448 = 11728,8)$, lo que representa el 19% de la suma de cuadrados de GxE. Esto implica que lo que queda realmente de la suma de cuadrados de la interacción, eliminado el error es el 81% de la suma de cuadrados de GxE. Este porcentaje es ligeramente mayor que el porcentaje explicado por los dos primeros ejes del AMMI. Lo que nos permite confirmar que el AMMI2 es más parsimonioso para describir adecuadamente la interacción excluyendo el ruido contenido en el modelo.

Tabla 5.17. Análisis de varianza del AMMI para rendimiento

Fuente	gl	SC	CM	F	
Ensayo (E)	8	264931	33116,4	4076,6	**
Genotipo (G)	181	29340	162,1	3,7	**
GxE	1448	62942	43,5	5,4	**
AMMI1	188	30205	160,7	19,8	**
AMMI2	186	17594	94,6	11,6	**
AMMI3	184	9751	53,0	6,5	**
Residuo	890	5392	6,1	0,7	ns
Error	3276		8,1		

ns: no significativo, **: significativo al nivel $P < 0,01$

Para interpretar correctamente la interacción GxE, se realizaron dos representaciones gráficas distintas. En el primer gráfico se presenta el modelo AMMI1, con los valores medios de la variable rendimiento de cada genotipo y ensayo en el eje X, frente a sus correspondientes valores en el primer eje principal (eje Y), que explicó el 48,0 % de la interacción (Figura 5.5). En la segunda gráfica se presentan los valores de los genotipos y los ensayos sobre los dos primeros componentes principales, que explicaron un 76,0 % de la interacción (Figura 5.6).

En la Figura 5.5, se observa que los ensayos ART04 y BEL02 tienen valores negativos elevados en el primer componente principal (eje Y), en contraste con unos elevados valores positivos de los dos ensayos de Macotera (VLD02, VLD03). Esto indica que las diferencias en las respuestas de los genotipos a las condiciones contrastantes encontradas en estas parejas de ensayos son probablemente las principales causas de la aparición de esa interacción. Por su ubicación con respecto al eje X, las variedades testigo (grupo 7) son las que mayor rendimiento presentan, seguidas por algunas variedades de 6 carreras del grupo 6 y algunas entradas de la colección nuclear del grupo 4. Todas estas entradas presentan una interacción de signo negativo en el primer componente principal. Estas entradas parecen estar mejor adaptadas relativamente a los ensayos ART04, BEL02

y SAD04, que también presentan signos negativos en ese primer componente principal. Una mayoría de las entradas del grupo 4 presentan signo negativo en el primer componente principal. Sin embargo, la mayoría de las entradas del grupo 3 presenta signo positivo en ese eje y tienden a ser relativamente más productivas en los ensayos de VLD02 y VLD03, que presentan los valores positivos más acusados en el primer componente principal.

La representación de los ensayos sobre los dos primeros ejes de componentes principales (modelo AMMI2, Figura 5.6), permitió ver mejor las agrupaciones de genotipos y ensayos. En cuanto a los ensayos, se pueden distinguir cinco grupos: SAD03 y VED03, ART04 y BEL02, VLD03, VLD02, todos ellos con una importante contribución a la interacción GxE explicada por los dos primeros componentes principales, y un quinto grupo formado por SAD02, FOR03 y SAD04, que contribuyeron poco a esta interacción. Los testigos, así como las variedades comerciales de 2 y 6 carreras incluidas en la colección nuclear, presentaron signos negativos en el eje Y (componente principal 2), al igual que BEL02 y ART04. De hecho, las variedades comerciales presentaban el mayor diferencial con las *landraces* justo en estos ensayos (Tabla 5.6). Se observa también que el grupo 3, compuesto sobre todo por entradas de la Meseta, tienden a tener una adaptación mejor en los ensayos de esa misma región geográfica (VLD02, VLD03).

Las regresiones de las variables genotípicas y climáticas sobre los valores correspondientes de los genotipos y ensayos en los ejes 1 y 2 permitieron determinar las variables posiblemente más relacionadas con ellos. Estas fueron incorporadas a la gráfica del AMMI2. El criterio de selección de estas variables fue el coeficiente de determinación R^2 de la regresión de cada variable sobre los dos componentes principales. El umbral arbitrario fijado para este coeficiente fue de 0,25 para las variables genotípicas y 0,40 para las variables climáticas.

El conjunto de las covariables ambientales fue analizado mediante un análisis de componentes principales para revelar las relaciones y redundancias que existen entre ellas (no mostrado). En el caso de covariables claramente redundantes, sólo se incluyó en el gráfico del AMMI, la covariable cuyo coeficiente de determinación fue más alto.

Las variables genotípicas que parecen tener una influencia sobre las variaciones del rendimiento son, principalmente: altura de planta ($R^2=0,56$), encamado ($R^2=0,36$), espigado ($R^2=0,26$), peso hectolítrico ($R^2=0,51$), oídio ($R^2=0,43$) y roya ($R^2=0,29$). La información proporcionada por estas variables coincide con lo que ya hemos visto en el

análisis descriptivo. Se observa que las variedades comerciales y testigos espigan relativamente más temprano, presentan unas alturas de planta bajas, se encaman poco, son más resistentes al oídio y presentan granos de peso hectolítrico alto, contrariamente a las *landraces*. Dentro de estas últimas, las entradas de la Meseta (grupo 3) parecen ser las más altas y tardías. También son las que presentan un mayor encamado y una susceptibilidad alta a las enfermedades (oídio y roya), mientras que algunas entradas del grupo 4 parecen presentar reacciones opuestas.

Las variables climáticas que más probablemente influyeron en el comportamiento diferencial de los genotipos en los ensayos fueron particularmente las temperaturas máximas de otoño ($R^2=0,83$), las temperaturas máximas de invierno ($R^2=0,50$), la evapotranspiración de otoño ($R^2=0,51$) y la *ratio* entre precipitación y evapotranspiración de primavera ($R^2=0,41$). De hecho los ensayos SAD03, VED03 y VLD03 se caracterizaron por unas temperaturas altas en otoño, y una *ratio* de primavera baja contrariamente a los ensayos BEL02 y ART04. Por otra parte las temperaturas máximas de invierno fueron relativamente altas en los ensayos VLD02 y VLD03, y bajas en SAD04.

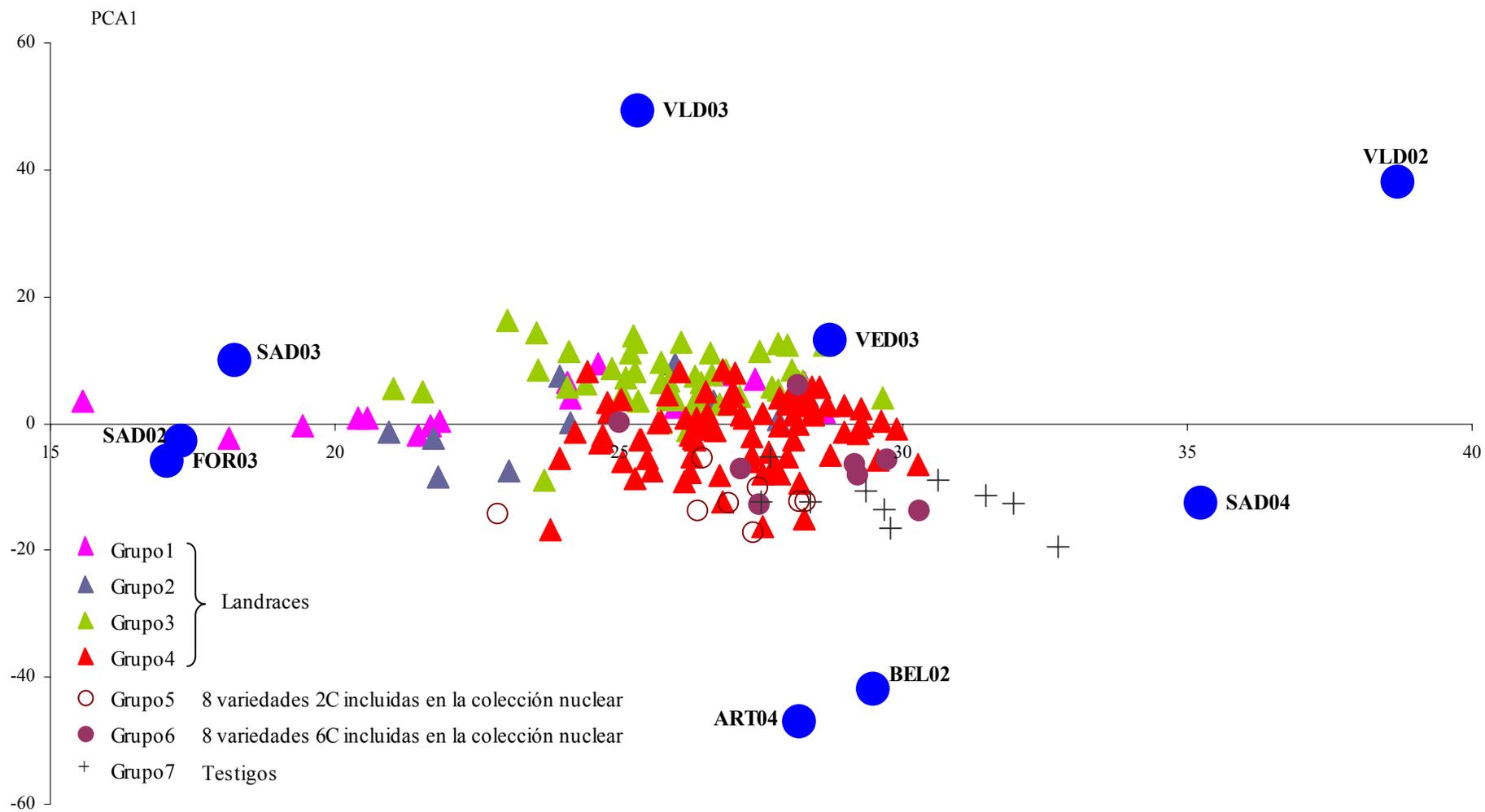


Figura 5.5. Representación del PCA1 del AMMI versus rendimiento medio ($q\ ha^{-1}$) de 182 genotipos divididos en 7 grupos, y evaluados en 9 ensayos.

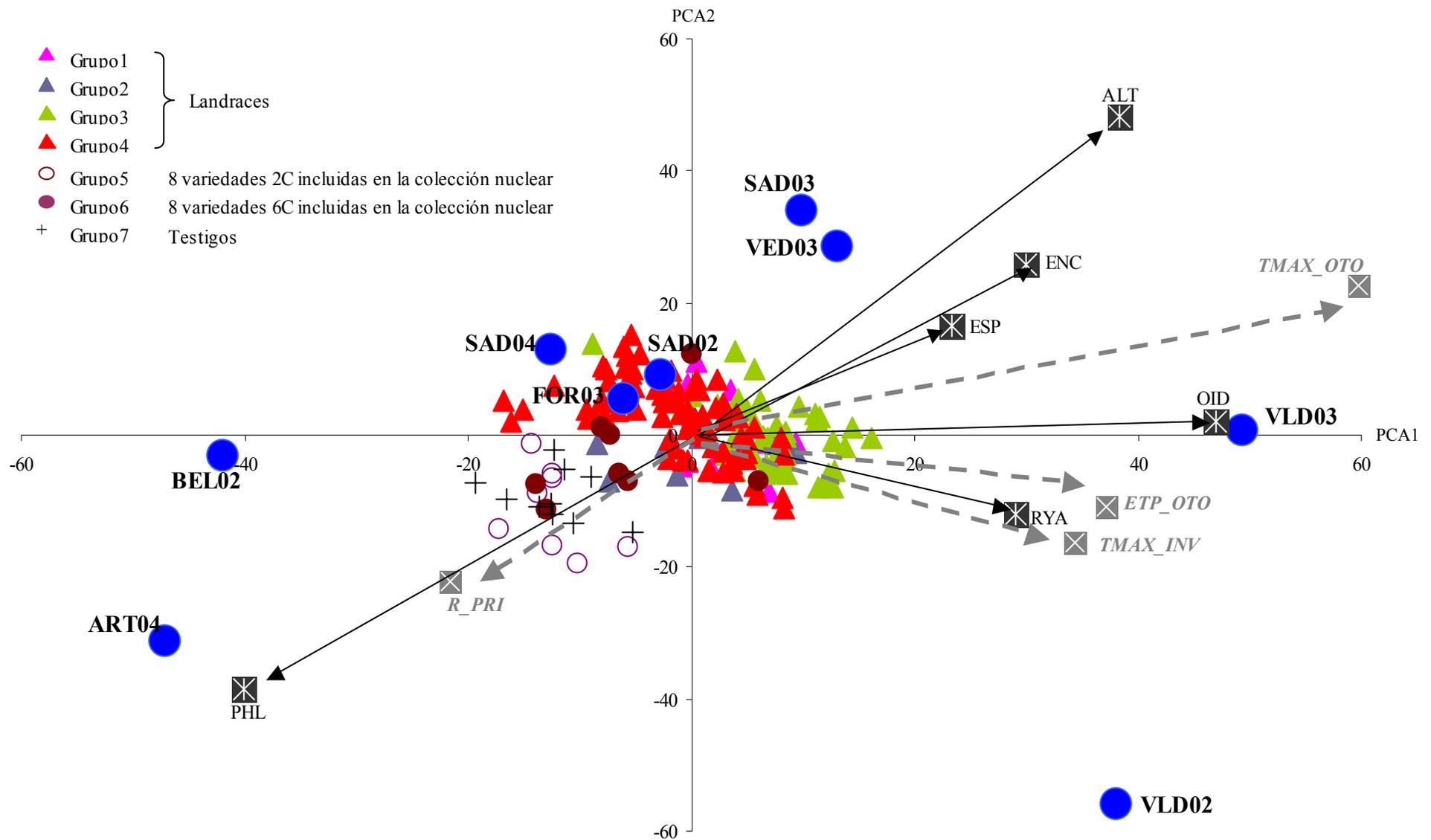


Figura 5.6. Representación de los dos componentes principales del modelo AMMI2, para 182 genotipos divididos en 7 grupos evaluados en 9 ensayos. Están incluidas las variables genotípicas y climáticas explicativas del Gx_E. Los vectores indican en que dirección se da el mayor cambio de las variables genotípicas (líneas enteras) y climáticas (líneas discontinuas). El tamaño de los vectores es proporcional al R^2 de cada variable explicado por los dos ejes principales.

5-3-5-2. Análisis de regresión factorial

Las variables genotípicas fueron utilizadas en un análisis de regresión factorial, para profundizar en la explicación del factor principal “Genotipo”. Este análisis permitió identificar las covariables más explicativas. En la Tabla 5.18, se exponen los resultados del modelo seleccionado. Este incluye 5 variables genotípicas (ESP, PMG, ALT, GRM, ESM) que explican el 34,9% del efecto del genotipo, con sólo el 2,8% de los grados de libertad. Sólo la covariable espigado explicó el 17% del efecto del genotipo con un único grado de libertad. Las desviaciones que corresponden al resto de la suma de cuadrados del genotipo, no explicadas por las covariables, fueron también significativas.

Tabla 5.18. Regresión factorial para el factor principal “Genotipo”, incluyendo covariables genotípicas.

Fuente	gl	SC	CM	F	
Ensayo	8	264931,1	33116,4	4076,6	**
Genotipo	181	29339,6	162,1	3,7	**
ESP	1	5098,5	5098,5	117,3	**
PMG	1	1041,0	1041,0	23,9	**
ALT	1	2128,2	2128,2	48,9	**
GRM	1	1247,2	1247,2	28,7	**
ESM	1	726,5	726,5	16,7	**
Desviaciones	176	19098,2	108,5	2,5	**
GxE	1448	62942,1	43,5	5,4	**
Error	3276		8,1		

** : Significativo al nivel $P < 0,01$.

Las mismas variables genotípicas y las variables climáticas definidas en el capítulo de Material y Métodos fueron empleadas en un análisis de regresión factorial, para evaluar su contribución a la suma de cuadrados de la interacción GxE. Este análisis se realizó en diferentes etapas. En una primera fase se analizaron las covariables genotípicas, luego en una segunda etapa se emplearon las covariables climáticas. Y como última etapa, se evaluó la interacción entre los dos tipos de covariables. El orden de introducción de las covariables se hizo de la siguiente manera: primero se corrieron los modelos con una covariable, y se eligió la covariable más explicativa (según la R^2) para guardarla en el modelo. A continuación se calculaban los modelos con dos covariables, la primera la elegida en el paso anterior, y la segunda cada una del resto de covariables. De la misma forma que antes, se escogía la segunda covariable más

explicativa para incluirla de forma permanente en el modelo, e introducir una tercera etc. De este modo se realizaron todas las rondas necesarias del análisis hasta incluir todas las covariables significativas. Este proceso se repitió con las covariables genotípicas y las climáticas y, en una última fase combinando los dos tipos de covariables, hasta conseguir el modelo más parsimonioso.

Los resultados del análisis de la regresión factorial de las covariables genotípicas x ensayo permitió determinar las principales covariables explicativas de la interacción GxE. Estas fueron, por orden de importancia: PHL, ESP, OID, ALT y ENC, las cuales explicaron el 45% de la interacción GxE, con el 2,8% de grados de libertad de esta última (Tabla 5.19). Las desviaciones no explicadas por el modelo fueron también significativas.

Tabla 5.19. Repartición de la interacción GxE, mediante análisis de regresión factorial con las covariables genotípicas.

Fuente	gl	SC	CM	F	
GxE	1448	62942,1	43,5	5,4	**
PHL*Ensayo	8	11307,8	1413,5	174,0	**
ESP*Ensayo	8	7043,7	880,5	108,4	**
OID*Ensayo	8	5758,5	719,8	88,6	**
ALT*Ensayo	8	2670,7	333,8	41,1	**
ENC*Ensayo	8	1422,8	177,8	21,9	**
Desviaciones	1408	34738,6	24,7	3,0	**
Error	3276		8.1		

** : Significativo al nivel P <0,01

En cuanto a las regresiones factoriales con las covariables climáticas, se detectaron principalmente dos covariables que contribuyen explicativamente en la interacción GxE: las temperaturas máximas de otoño e invierno (TMax_oto, TMax_inv). La TMax_oto parece ser la covariable climática que más influye en el comportamiento de los genotipos, como ya se mostró en la gráfica del modelo AMMI2. La interacción TMax_oto x genotipo se llevó el 33,8% de la suma de cuadrados del GxE (Tabla 5.20), con la variable TMax_inv en un segundo lugar. La interacción de estas dos covariables con el factor Genotipo, explicó el 45,7% de la interacción GxE con un 25% de sus grados de libertad (Tabla 5.20). Este modelo es menos parsimonioso que el anterior y que el AMMI1.

Tabla 5.20. Repartición de la interacción GxE, mediante análisis de regresión factorial con las covariables ambientales.

Fuente	gl	SC	CM	F	
GxE	1448	62942,1	43,5	5,4	**
Genotipo*TMax_oto	181	21253,2	117,4	14,5	**
Genotipo*TMax_inv	181	7485,4	41,4	5,1	**
Desviaciones	1086	34203,4	31,5	3,9	**
Error	3276		8,1		

** : Significativo al nivel P <0,01

Tabla 5.21. Repartición de la interacción GxE, mediante análisis de regresión factorial con las covariables genotípicas y ambientales.

Fuente	gl	SC	CM	F	
GxE	1448	62942,1	43,5	5,4	**
OID*ensayo	8	10977,6	1372,2	168,9	**
OID*TMax_oto	1	8049,2	8049,2	990,8	**
Desviaciones	7	2928,5	418,4	51,5	**
ESP*ensayo	8	6903,5	862,9	106,2	**
ESP*TMax_oto	1	5203,8	5203,8	640,6	**
Desviaciones	7	1699,7	242,8	29,9	**
ENC*ensayo	8	6539,3	817,4	100,6	**
Residuo	1424	38521,6	27,1	3,3	**
Error	3276		8,1		

** : Significativo al nivel P <0,01

La tercera etapa de las regresiones factoriales consistió en combinar las covariables genotípicas con las climáticas. El mejor modelo conseguido con este análisis, permitió explicar el 31,5% de la interacción GxE con solamente 10 grados de libertad, que representan el 0,7% de los grados de libertad del GxE. Este modelo incluye las interacciones entre oídio*TMax_oto, ESP*TMax_oto y ENC*ensayo (Tabla. 5.21).

Aunque este modelo es menos explicativo que los anteriores y que el AMMI, se considera como el modelo más parsimonioso, pues explica una cantidad de la suma de cuadrados de GxE por grado de libertad notablemente superior a los demás modelos.

5-3-6. Clasificación de los genotipos

Con objeto de describir los mejores genotipos, y de facilitar la búsqueda de posibles parentales para cruzamientos, se procedió a su clasificación de acuerdo con:

- rendimiento medio en todos los ensayos;
- rendimiento medio en ambiente de alta productividad;

- rendimiento medio en ambiente de baja productividad;
- índice agronómico, calculado basándose sobre las clasificaciones (o rankings) de las variables rendimiento, espigado, altura de planta, encamado y peso hectolítrico.

Para calcular éste índice, se asignaron coeficientes a las clasificaciones de las variables consideradas, según su importancia relativa en el programa nacional público de mejora de cebada y para este tipo de material, que fue como sigue:

Índice agronómico = (ranking rendimiento * 0,4) + (ranking espigado * 0,1) + (ranking altura * 0,1) + (ranking peso hectolítrico * 0,2) + (ranking encamado * 0,2).

Los rankings de los genotipos fueron asignados de forma descendente para las variables rendimiento y peso hectolítrico y de forma ascendente para las variables espigado, altura y encamado. En la Tabla 5.22, se presentan los 25 mejores genotipos según el carácter considerado.

En cuanto a los promedios de rendimiento en la totalidad de los ensayos, se observa la superioridad de algunas variedades como Orria, Graphic, Steptoe, Candela y Barbarrosa superiores a todas las *landraces*. Por otra parte, se observó la presencia de 14 entradas de la colección nuclear entre los 25 mejores genotipos. Doce de estas entradas pertenecen al grupo 4 (entradas fundamentalmente del Sur y del Valle del Ebro), lo que reafirma la superioridad general de este grupo, ya apuntada en los análisis combinados de ensayos (apartado 5-3-3-1). Ninguna entrada de los grupos 2 (*landraces* de 2 carreras), y sólo una entrada de los grupos 1 y 3, entraron entre estas primeras clasificadas. La mejor entrada del grupo 2 fue la 148, que se quedó en la posición 57 (en cuanto a rendimiento total). Las mejores *landraces* resultaron ser las entradas 91 y 42, que se situaron respectivamente en los puestos 6 y 7 (Tabla 5.22).

Tabla 5.22. Clasificación de los genotipos según el promedio de rendimiento en los 9 ensayos (sin ART02), promedio de rendimiento en ambiente de alta productividad, promedio de rendimiento en ambiente de baja productividad e índice agronómico. Los números corresponden a las entradas de la colección nuclear, entre paréntesis se indica a que grupo genético pertenecen. Las variedades comerciales aparecen en negro, y los testigos además están escritos en cursiva.

Clasificación	Rendimiento total	Rendimiento en ambientes de alta productividad	Rendimiento en ambientes de baja productividad	Índice agronómico
1	<i>Orria</i>	<i>Orria</i>	73 (G 4)	<i>Graphic</i>
2	<i>Graphic</i>	<i>Graphic</i>	81 (G 4)	Barbarrosa
3	<i>Steptoe</i>	<i>Steptoe</i>	16 (G 4)	<i>Gaelic</i>
4	<i>Candela</i>	<i>Gaelic</i>	27 (G 3)	<i>Orria</i>
5	Barbarrosa	Barbarrosa	91 (G 4)	<i>Candela</i>
6	91 (G 4)	<i>Candela</i>	38 (G 4)	<i>Volga</i>
7	42 (G 4)	<i>Nevada</i>	18 (G 4)	<i>Nevada</i>
8	<i>Gaelic</i>	<i>Volga</i>	75 (G 4)	Alpha
9	Monlón	Zaida	42 (G 4)	Hassan
10	<i>Nevada</i>	Kym	<i>Orria</i>	<i>Tipper</i>
11	27 (G 3)	Dobla	106 (G 1)	Monlón
12	81 (G 4)	Ager	76 (G 4)	<i>Steptoe</i>
13	18 (G 4)	Alpha	39 (G 4)	Ager
14	<i>Volga</i>	Monlón	31 (G 3)	Dobla
15	125 (G 4)	Hassan	79 (G 3)	Zaida
16	73 (G 4)	<i>Tipper</i>	51 (G 3)	<i>Plaisant</i>
17	98 (G 4)	91 (G 4)	49 (G 4)	Pané
18	Pané	Pané	14 (G 4)	Kym
19	16 (G 4)	145 (G 4)	92 (G 4)	Union
20	Ager	<i>Seira</i>	109 (G 4)	<i>Seira</i>
21	38 (G 4)	42 (G 4)	<i>Steptoe</i>	98 (G 4)
22	110 (G 4)	98 (G 4)	146 (G 3)	148 (G 2)
23	14 (G 4)	116 (G 4)	11 (G 3)	Beka
24	52 (G 4)	4 (G 1)	25 (G 4)	8 (G 3)
25	4 (G 1)	125 (G 4)	125 (G 4)	145 (G 4)

En ambientes de alta productividad, la superioridad de las variedades comerciales fue muy destacada, pues ocuparon los 16 primeros puestos. Las variedades *Orria*, *Graphic* y *Steptoe* ocuparon los primeros lugares. Sólo 7 *landraces* entraron en la lista de los 25 mejores genotipos, 6 de ellas del grupo 4. Las entradas 91, 145 y 42 resultaron ser las más productivas en estas condiciones, entrando respectivamente en los puestos 17, 19 y 21 de la clasificación. El único genotipo del grupo 1 (entrada 4), que se

encuentra en esta lista, se colocó en el puesto 24. Las entradas 148 y 27, que pertenecen respectivamente a los grupos 2 y 3, son las más productivas de su grupo, ocupando respectivamente los puestos 73 y 33 en la clasificación.

En condiciones de baja productividad la situación fue muy distinta. Sólo las variedades Orria y Steptoe entraron en la lista ocupando respectivamente las posiciones 10 y 21. Las entradas del grupo 4 fueron dominantes en esta lista (16 de las 25), del mismo modo que los cultivares dominaban la clasificación en los ambientes de alta productividad. Del grupo 3 se encontraron 6 genotipos, de los cuales el 27 fue el mejor, situándose en el puesto 4. La entrada 106 es la única del grupo 1, que se encuentra en la lista, en la posición 11. Se observa que ninguna entrada del grupo 2 está presente. La entrada 148, que pertenece a este grupo, resultó ser la más productiva en estas condiciones, al igual que en los ambientes de alta productividad, clasificándose en el puesto 40.

La clasificación de los genotipos según el índice agronómico calculado, confirmó la superioridad de las variedades comerciales, ocupando los 20 primeros puestos de la lista. Sólo 4 *landraces*, de los grupos 2, 3 y 4, entraron en la lista de los mejores genotipos, mientras que el mejor genotipo del grupo 1 (entrada 4), se clasificó en el puesto 47. Es también digno de destacar que las 10 variedades comerciales empleadas como testigos están entre las 25 primeras por el índice agronómico, mientras que de las 16 variedades más antiguas, que forman parte de la colección nuclear, están incluidas 11 de las 16. De este modo se pone de manifiesto la superioridad general de las variedades más modernas (las empleadas como testigos, en general) sobre las antiguas, aunque hay excepciones notables, como los cultivares veteranos Barbarrosa y Alpha, aún ampliamente empleados por los agricultores en la actualidad.

5-4. Discusión

Esta serie de ensayos de campo se componía inicialmente de 15 ensayos, de los cuales 5 no pudieron llevarse a término por distintas circunstancias, lo cual no es extraordinario en este tipo de ambientes mediterráneos, expuestos a numerosas vicisitudes climáticas.

El diseño experimental empleado, con la división en cuatro sub-ensayos por cada ensayo de campo, se justifica por la reconocida heterogeneidad de las parcelas de experimentación en al menos dos de las localidades empleadas. El sistema de corrección está basado en el “diseño enriquecido”, y se ha empleado una corrección leve, sólo en

los casos donde las diferencias entre los testigos en los sub-ensayos fueron significativas. Así, éstas sólo fueron necesarias en 17 de las 50 ocasiones posibles (Tablas 5.3 a 5.5). Como se esperaba por el conocimiento previo, los ensayos en la localidad de Sádaba precisaron el mayor número de correcciones (4, de un total de 5 variables medidas). Sin embargo, la localidad de El Vedado, cuyas parcelas son habitualmente muy heterogéneas, no precisó más que de una corrección sobre 9 variables medidas en el único año en que fue cosechado el experimento (2003). Aunque pudo influir el hecho de que este año fue particularmente favorable en esa zona, en la que los rendimientos raramente exceden los 20,00 q ha⁻¹.

Los parámetros estadísticos de las variables analizadas, calculados para los dos conjuntos de genotipos (*landraces* y variedades comerciales) resaltaron el comportamiento distinto de los dos grupos principales de genotipos. También, se observaron diferencias importantes entre los promedios de los distintos ensayos.

La distribución de clases de las variables reveló la amplia diversidad que existe entre los genotipos de las *landraces* de la colección nuclear en comparación con las variedades comerciales (Figuras 5.2 – 5.4), aunque hay que tener en cuenta el menor tamaño de muestra de estas últimas. También se notó la presencia de muchos caracteres indeseables, en porcentajes relativamente altos de entradas de la colección nuclear, tales como una excesiva altura del tallo que favorece el encamado, una mayor susceptibilidad a enfermedades y un bajo rendimiento. La mayor altura y encamado en las *landraces* era un resultado esperado, pues se trata de caracteres muy seleccionados en los procesos de mejora de cereales de invierno, por lo que es lógico que las variedades actuales presenten valores menores para ambos parámetros. No era esperable, sin embargo, la mayor longitud del ciclo de desarrollo de las *landraces*. La diferencia encontrada, de aproximadamente 4 días en promedio de fecha de espigado a favor de las variedades comerciales, es importante y es difícil que sea casual. En principio, las *landraces* deberían mostrar una adaptación óptima al medio, y especialmente en cuanto a longitud del ciclo, siendo un carácter fácilmente manipulable y de alta heredabilidad. Debería existir alguna explicación que justificara este desplazamiento en la preferencia de los agricultores hacia variedades más tempranas (recordemos que todas las variedades comerciales contempladas en este estudio son o han sido ampliamente cultivadas en España, o son productos recientes de programas de mejora locales). Una hipótesis posible es que las *landraces* proceden, prácticamente en su totalidad, de una época anterior a la mecanización del campo, que ocurrió relativamente tarde en España, con

posterioridad a la Guerra Civil. En esa época las cosechas se realizaban a mano o con tecnologías artesanales, a lo largo de periodos largos de tiempo. Por el contrario, la cosecha mecánica de las variedades actuales es extremadamente rápida. Por tanto, las variedades empleadas antes de la mecanización del campo debían ser necesariamente más tardías que las actuales, para evitar pérdidas excesivas de rendimiento por el “descabezamiento” (caída de las espigas al suelo) de las plantas que permanecieran demasiado tiempo en el campo tras la madurez fisiológica (JM Lasa, com. pers.). Por el contrario, el tiempo mínimo hasta el espigado en *landraces* y variedades comerciales es muy similar, indicando probablemente la longitud de ciclo mínima que debe tener una variedad para evitar las heladas primaverales y aprovechar las lluvias de primavera, típicas del clima mediterráneo español.

La diferencia en altura de planta entre *landraces* y variedades comerciales es muy sustancial, en torno a los 15 cm a favor de las primeras (Tabla 5.6). Esto sugiere una muy escasa presencia de alelos de enanismo en las *landraces*, cuya altura se sitúa, en general, muy por encima del formato actual de cultivar de cebada. Este hecho debería ser corregido convenientemente cuando estos materiales se empleen en programas de mejora. También hay que tener en cuenta que las *landraces* proceden de una época anterior a la generalización del uso de fertilizantes químicos por lo que, probablemente alcancen mayor desarrollo vegetativo y mayor altura en las condiciones experimentales actuales que en los ambientes agrícolas a los que estaban adaptadas. Por tanto, no es extraño que muestren un encamado acusado (Tabla 5.6), que podría estar motivado en parte por ese desarrollo extraordinario en respuesta a condiciones de fertilización modernas.

Las distribuciones de las *landraces* y de las variedades comerciales (Figura 5.2 - 5.4) ponen de manifiesto la mayor variabilidad existente en las primeras para muchos caracteres, exceptuando el rendimiento, el peso de mil granos, el número de granos m^{-2} , el número de granos por espiga y la resistencia al moteado de las hojas. Parece que la colección nuclear contiene una gran diversidad en caracteres agronómicos de todo tipo, en consonancia con la notable diversidad detectada a nivel molecular (capítulo 3 de este trabajo). En cuanto a los caracteres de calidad maltera, las *landraces* tienen en general unos valores bastante pobres para todos ellos (Tabla 5.9 y Figura 5.4). Sin embargo, algunos genotipos extremos llegan a mostrar valores interesantes de extracto o proteína, por lo que no es descartable que puedan existir fuentes de calidad maltera en la colección nuclear.

El análisis de varianza de las variables consideradas permitió ver la existencia de diferencias significativas entre los genotipos estudiados. Dividiendo éstos en 7 grupos, como está definido en el apartado 5-2-4-2.b, se detectó que las diferencias son mayores entre grupos que dentro de grupos. Estas diferencias entre grupos se deben en la mayoría de los casos, principalmente a las diferencias que existen entre los grupos definidos genéticamente en la colección nuclear (1 a 4) con los grupos de las variedades comerciales (5 a 7). La principal causa de estas diferencias se puede deber a la presencia de una combinación de caracteres agronómicos deseables en las variedades comerciales, como resultado del proceso de mejora. Un ejemplo son las variables como espigado, rendimiento, peso hectolítrico y resistencia al encamado. Las Tablas 5.10b y 5.11b, muestran claramente la superioridad de las variedades comerciales frente a las *landraces*. Los grupos 5 a 7, se caracterizaron por un espigado precoz, un rendimiento alto, una resistencia al encamado importante y un peso hectolítrico mayor. Contrariamente los grupos 1 a 4, que fueron más tardíos en espigar, presentaron rendimientos y pesos hectolítricos relativamente más bajos, y un fuerte encamado. También se observaron diferencias entre grupos del mismo lote (*landraces* o variedades comerciales) para algunos caracteres. En alguno de estos casos, las diferencias son consecuencia de la comparación de grupos de genotipos de diferente tipo de espiga (2 y 6 carreras). Este sería el caso de las variables peso hectolítrico, número de espigas m⁻² o número de granos por espiga. Las cebadas de 2 carreras se caracterizan por un grano redondeado y siempre presentan un peso hectolítrico mayor. También para los componentes de rendimiento, los genotipos de 2 carreras siempre muestran un número de espigas m⁻² más alto, que compensa el menor número de granos por espiga. Por otra parte, se observaron diferencias entre los grupos 3 y 4, que representan la mayoría de las *landraces* de 6 carreras, para algunos caracteres como espigado, rendimiento, resistencia a oídio y roya. Las entradas del grupo 4 espigaron antes, presentaron unos promedios de rendimiento más altos y mayor resistencia al oídio y la roya que las del grupo 3. El hecho de que la mayoría de las entradas del grupo 4 sean originarias de zonas caracterizadas por temperaturas relativamente altas y pluviometría escasa en primavera podría explicar su precocidad. Como adaptación a su ambiente, estas entradas aceleran su ritmo de desarrollo para escapar al estrés hídrico característico de las zonas mediterráneas en esta época. También la humedad alta característica de las zonas costeras, favorece el crecimiento de los hongos. Para adaptarse, algunas entradas del grupo 4 pudieron desarrollar una resistencia a las enfermedades en respuesta a la

presión ejercida por los patógenos, contrariamente a lo ocurrido en los ambientes más secos de la Meseta de los que son originarias las entradas del grupo 3.

Los ensayos fueron llevados a cabo en localidades de características geográficas y climáticas distintas, lo que se reflejó en el comportamiento de los genotipos. Esto se tradujo en una interacción GxE significativa para todas las variables medidas en estos ensayos. Esta interacción GxE fue mayor entre los 7 grupos definidos que dentro de grupos. La adaptación específica de los grupos a ambientes distintos parece ser una de las causas de la interacción GxE. Esto se ilustra en el análisis de la variable rendimiento con el AMMI1, donde destacó el mejor comportamiento de las variedades comerciales en los ambientes de alta productividad, SAD04 y VLD02, aunque muestren una interacción negativa con este último. Estas variedades parecen mejor adaptadas a los ensayos de BEL02 y ART04. Por el contrario, en los ensayos de VLD02 y VLD03, las entradas del grupo 3 (mayoritariamente de la Meseta) expresan mejor su potencial genético. Diferentes autores han encontrado que la comparación entre grupos de germoplasma es una de las explicaciones de la aparición de la interacción GxE, como por ejemplo Dofing *et al.* (1992), al evaluar el rendimiento de genotipos de cebada del sub-ártico y de ambientes templados en las dos zonas de origen. También, Ceccarelli y Grandó (1991) describieron una interacción similar al comparar genotipos locales de Oriente Medio con variedades modernas.

El modelo AMMI1, puso también de relieve la superioridad de los testigos en comparación con las variedades comerciales incluidas en la colección nuclear, lo que de nuevo refleja el efecto de la mejora en las variedades modernas. Por otra parte, se observó que las entradas del grupo 1 presentaron los peores promedios de rendimiento, debido probablemente a su espigado tardío. De hecho, se observó una correlación negativa general entre la fecha de espigado y el rendimiento ($r = -0.41$, $P < 0,0001$). El espigado tardío de estas entradas hace que sufran más las altas temperaturas y el estrés hídrico que suelen ocurrir en la fase final del cultivo, en condiciones mediterráneas. Villegas *et al.* (2000) comentaron que el factor más importante que confiere a un cultivo la aptitud de desarrollarse mejor bajo condiciones de estrés hídrico, es su ajuste fenológico. Sobre este mismo tema, Slafer *et al.* (2005) explican que los cambios en el desarrollo fenológico permiten escapar al cultivo del estrés, evitando estar expuesto a ello durante el ciclo del cultivo o evitando la coincidencia de las fases sensibles del cultivo con los periodos de más probable ocurrencia del estrés. Una de las fases más sensibles del cultivo de cebada es el periodo de llenado del grano. El espigado tardío

hace que los genotipos llenen sus granos en un periodo en el cual las condiciones climáticas son desfavorables. Por lo tanto, se obtienen unas espigas con granos pequeños. Esto es consistente con el bajo peso de mil granos observado en las entradas del grupo 1, muy tardías, lo que se traduce en un bajo rendimiento en este grupo. En apoyo de esta hipótesis se puede decir que se observó una correlación negativa general entre el espigado y el peso de mil granos ($r = -0,26$, $P = 0,003$). Stone y Nicolas (1994) comentaron que las temperaturas máximas altas en las fases post-floración, tienen un efecto adverso sobre el rendimiento, provocando una bajada en el peso del grano. También, Brancourt-Hulmel (1999) relacionó la caída de los rendimientos de variedades tardías de trigo, por haber estado sometidas a ambientes desfavorables durante el llenado de grano.

Viendo las fechas de espigado tardías de las entradas del grupo 1 y la alta biomasa que producen, se puede pensar que estos genotipos pudieran estar destinados preferentemente a un uso mixto o forrajero.

La información aportada por el modelo AMMI2, permitió una mejor separación de los grupos de la colección nuclear frente a las variedades comerciales, y la clara disociación de los dos principales grupos de esta colección. Las entradas del grupo 3 presentaron una mejor adaptación en los ensayos de Castilla y León, mientras que los genotipos del grupo 4 tienden a tener mejor comportamiento en los ensayos del Valle del Ebro. La superioridad de los testigos y de los grupos de variedades comerciales se destacó particularmente en los ensayos catalanes. La inclusión de la información aportada por los vectores de las variables genotípicas permitió adelantar hipótesis sobre causas de la interacción GxE. La altura excesiva del tallo, el fuerte encamado, junto con la susceptibilidad a enfermedades (oídio y roya) y el espigado tardío que caracterizan las entradas de la colección nuclear, son las principales variables genotípicas relacionadas con la interacción GxE entre grupos. Estos resultados coinciden con los de Brancourt-Hulmel (1999), que comentó que la precocidad del espigado, junto con la susceptibilidad al oídio y el encamado son las covariables genotípicas que tuvieron mayor efecto en la interacción GxE, en ensayos de trigo de invierno.

Por otra parte, la variable peso hectolítrico, que explica el 50% de la variación GxE y presenta una dirección contraria a los vectores de espigado, altura, encamado y oídio, indica fundamentalmente la presencia de granos de calibre mayor en las variedades comerciales. Esta variable no se puede considerar como responsable de los altos rendimientos de estas variedades, sino como una consecuencia resultante de las

características genotípicas deseables que poseen estas últimas. En efecto, por ser más tempranas, estas variedades se benefician de un periodo de llenado del grano más largo que las líneas de la colección nuclear. Además, la baja altura que caracteriza estas variedades les procura una resistencia al encamado, que muchas veces ocurre en el periodo de llenado de los granos.

La posición de los vectores de las variables oídio y roya con respecto a las entradas de los grupos 3 y 4, parece indicar que la susceptibilidad de los genotipos del grupo 3 (Meseta) es mayor que la de los del grupo 4 (Sur y Valle del Ebro). Asimismo, se aprecia la presencia de algunos genotipos resistentes en las entradas del grupo 4. Esto se podría explicar por el hecho de que las entradas provenientes de las zonas costeras (Sur y Valle del Ebro) estuvieron siempre expuestas a estas enfermedades, conduciendo a la selección a favor de la resistencia en algunas de ellas. Por el contrario, a mayor altitud, la incidencia de las enfermedades es menor. El efecto del clima sobre el comportamiento de los genotipos, en los diferentes ensayos se manifiesta principalmente para las variables R_pri, Tmax_oto, ETP_oto y Tmax_inv.

La dirección del vector de la R_pri, indica que los ensayos de BEL02 y ART04 tuvieron más agua disponible en las fases finales del ciclo de cultivo, lo que explicaría la mejor adaptación de las variedades comerciales en ambos ensayos. Al haber sido seleccionados esencialmente en ambientes favorables, estos cultivares pueden expresar mejor su rendimiento potencial cuando las condiciones son buenas. Por otra parte, el bajo rendimiento de las *landraces* en estos ensayos podría ser en parte consecuencia del encamado. En este tema, Brancourt-Hulmel (1999) comenta que el encamado tiene un efecto importante sobre el rendimiento. En los ensayos de SAD03, VED03 y VLD03, donde hubo un estrés hídrico en las fases finales del cultivo (Tabla 5.1), las *landraces* superaron relativamente a los rendimientos de las variedades comerciales. Garcia del Moral *et al.* (2003) comentaron que, en ambientes de baja productividad por falta de agua disponible al final del ciclo del cultivo, los genotipos precoces rinden mejor que los tardíos. Este no fue el caso en este estudio, ya que las variedades comerciales fueron más tempranas. Este fenómeno parece indicar que el material local pudiera ser más tolerante a la sequía. En este mismo sentido, Voltas *et al.* (1999) comentaron que variedades de seis carreras de alto rendimiento con amplia adaptación, como Barbarrosa, Orria y Plaisant, sufren en condiciones de estrés hídrico terminal, mientras que el cultivar español Albacete no se ve afectado en el mismo grado por estas condiciones. van Oosterom *et al.* (1993), indicaron que en ambientes de baja

productividad, causada por estrés hídrico y térmico, las variedades insensibles a las fluctuaciones ambientales aseguran una mejor estabilidad de rendimiento.

Por otra parte los vectores de las temperaturas máximas de otoño e invierno presentan respectivamente ángulos agudos con los vectores de las variables oídio y roya. Esto parece indicar que la incidencia de las enfermedades fue favorecida por las altas temperaturas de estas épocas. Manners (1986) comenta que el patógeno puede estar latente durante un periodo largo cuando la temperatura sea desfavorable y volver rápidamente a la actividad cuando vuelvan las condiciones óptimas. Si el patógeno se vuelve inactivo, por las temperaturas desfavorables, la epidemia se retrasa, pero si muere a causa de ellas, dicha epidemia no se produce. Esto, nos lleva a pensar que al tener temperaturas favorables durante todo el ciclo del cultivo, la actividad del patógeno no se ve afectada y su ataque se ve multiplicado.

Las regresiones factoriales sugieren que las principales diferencias de rendimiento entre genotipos se deben principalmente al espigado, la altura y los componentes de rendimiento (PMG, ESM, GRM), de entre las variables que fueron medidas. Sin embargo, el espigado parece tener más efecto en estas diferencias, ya que explica el 17,4% de la variación genotípica del rendimiento con solamente 1 grado de libertad, lo que corresponde a un 0.006% de los grados de libertad del genotipo. Esto indica la importancia de este carácter en la selección para rendimiento. Las desviaciones que corresponden al resto de la suma de cuadrados del genotipo, no explicadas por las covariables genotípicas, fueron también significativas. Esto significa que hay diferencias genotípicas importantes que no han sido explicadas por las covariables estudiadas. Pero no es sorprendente, pues las variables estudiadas no componen un panorama completo de las posibles causas que afectan al rendimiento.

Por otra parte, las regresiones factoriales indicaron que tanto las variables genotípicas como las climáticas fueron importantes en la explicación del GxE. Sin embargo, estas primeras fueron más explicativas (45% del GxE con 40 gl) que las climáticas (45,7% con 362 gl).

Tanto las regresiones de las variables genotípicas y climáticas sobre los ejes del AMMI como la regresión factorial apuntan a las mismas variables como relacionadas con la interacción GxE del rendimiento. Sin embargo, en la regresión factorial no aparecieron ni la roya, ni la evapotranspiración de otoño, ni la *ratio* de primavera, que si se revelaban en el grafico del AMMI2. Este se puede deber al hecho de la existencia de

multicolinealidad entre las variables, y a que la regresión factorial no incorpora variables que estén claramente correlacionadas.

El modelo más parsimonioso se alcanzó al integrar la interacción entre variables genotípicas y ambientales. Este modelo, sin ser una descripción completa de las causas de la aparición de la interacción GxE del rendimiento, sí que permite avanzar algunas hipótesis sobre estas causas. Este modelo, que incluye la interacción Tmax_oto x oídio, Tmax_oto x espigado, y encamado x ensayo, llegó a explicar 31,5% del GxE con menos del 1% de los grados de libertad de esta interacción, aunque deja un residuo significativo importante. Las interacciones Tmax_oto x oídio y encamado x ensayo, podrían tener una explicación lógica en este estudio, ya que la susceptibilidad al encamado y la incidencia del oídio ayudado por temperaturas óptimas, son los principales factores que provocan diferencias de rendimiento en los diferentes ensayos, para las líneas puras de la colección nuclear. Sin embargo, no encontramos una explicación a la interacción espigado x Tmax_oto.

La clasificación de los 182 genotipos según el rendimiento medio en la totalidad de los ensayos y en el rendimiento medio en ambientes de alta productividad mostró la superioridad agronómica general de las variedades comerciales frente a las *landraces*. Dentro de estas últimas, se distinguieron algunas entradas, particularmente del grupo 4, con unos rendimientos relativamente aceptables. Sin embargo en ambientes de baja productividad, las *landraces*, también especialmente algunas del grupo 4, se destacaron con unos rendimientos relativamente más altos que las variedades comerciales. Esto coincide con las afirmaciones de Voltas *et al.* (1999) y Lasa *et al.* (2001) en cuanto a la estabilidad de las variedades tradicionales locales bajo condiciones ambientales desfavorables y sus rendimientos bajos en condiciones óptimas, comparados con los rendimientos de las variedades modernas. Similarmente, Ceccarelli *et al.* (1991) afirmaron que *landraces* colectadas en regiones áridas (*Steppe*) de Siria mostraron un promedio de rendimiento dos veces más altos que las variedades modernas, cuando estas fueron cultivadas en una zona desfavorable. Ceccarelli *et al.* (1992) y Ceccarelli (1996) indicaron que variedades de alto potencial de rendimiento bajo condiciones óptimas no son rentables en malas condiciones y viceversa.

Las variedades Orria y Steptoe son las únicas que se clasificaron entre los 25 mejores genotipos en ambientes de baja productividad. Estas dos variedades se clasificaron como las mejores variedades de 6 carreras en condiciones óptimas. La estabilidad de estas variedades se debería probablemente al hecho que la variedad Orria

fue obtenida recientemente en España, y que Steptoe es una variedad americana, con parte de sus ancestros mediterráneos y que fue ampliamente cultivada en España durante el siglo pasado. Esto concuerda con las afirmaciones de Ceccarelli *et al.* (1987) que recomendaron llevar la mejora de genotipos de cebada para condiciones extremas, en las mismas condiciones, y Ceccarelli y Grando (1989) que defendieron el uso de un material adaptado para la selección en condiciones de estrés hídrico.

La superioridad de las variedades comerciales se confirmó con la clasificación según el índice agronómico. Este índice es totalmente subjetivo, pero mide de algún modo la aceptabilidad de las entradas desde el punto de vista agronómico. Aunque las *landraces* han mostrado características muy interesantes a lo largo de todo este estudio, este índice revela que, en su formato actual no pueden competir con los cultivares, por lo que su aprovechamiento debe realizarse a través de programas de mejora.

5.5 Referencias bibliográficas

- Brancourt-Hulmel M. 1999.** Crop diagnosis and probe genotypes for interpreting genotype environment interaction in winter wheat trials. *Theoretical & Applied Genetics*, 99: 1018-1030.
- Ceccarelli S. 1989.** Wide adaptation: How wide? *Euphytica*, 40: 197-205.
- Ceccarelli S. 1996.** Positive interpretation of genotype by environment interactions in relations to sustainability and biodiversity. *In Plant Adaptation and Crop Improvement*. Ed. Cooper M, Hammer GL, CAB International, Wallingford, England. 467-486.
- Ceccarelli S, Grando S. 1989.** Efficiency of empirical selection under stress conditions in barley. *Journal of Genetics & Breeding*, 43: 25-31.
- Ceccarelli S, Grando S. 1991.** Environment of selection and type of germplasm in barley breeding for low-yielding conditions. *Euphytica*, 57: 207-219.
- Ceccarelli S, Nachit MM, Ferrara GO, Mekni MS, Tahir M, van Leur JAG, Srivastava JP. 1987.** Breeding strategies for improving cereal yield and stability under drought. *In Drought tolerance in winter cereals*. Ed. Srivastava JP, Porceddu E, Acevedo E, Varma S. John Wiley and Sons, New York, 101-114.
- Ceccarelli S, Acevedo E, Grando S. 1991.** Breeding for yield stability in unpredictable environments: single traits, interaction between traits, and architecture of genotypes. *Euphytica*, 56: 169-185.
- Ceccarelli S, Grando S, Hamblin J. 1992.** Relationships between barley grain yield measured in low and high yielding environments. *Euphytica*, 64: 49-58.
- DGA. 2002.** Encuestas productivas de cultivos herbáceos de invierno. Campaña 1999-2000. Departamento de Agricultura del Gobierno de Aragón. <http://portal.aragob.es>

- Dofing SM, Berke TG, Baenziger PS, Knight CW. 1992.** Yield and yield component response of barley in subarctic and temperate environments. *Canadian Journal of Plant Science*, 72: 663-669.
- Fischbeck. G. 2003.** Diversification through breeding. *In Diversity in barley (*Hordeum vulgare*)*. Ed. Bothmer R von, Hintum T van, Knüpfner H, Sato K. Elsevier Science B.V. Amsterdam, The Netherlands, 29-52.
- García del Moral LF, Rharrabti Y, Villegas D, Royo C. 2003.** Evaluation of grain yield and its components in durum wheat under Mediterranean conditions: An ontogenic approach. *Agronomy Journal*, 95: 266-274.
- Gauch HG. 1988.** Model selection and validation for yield trials with interaction. *Biometrics*, 44: 705-715.
- Gauch HG. 1992.** Statistical analysis of regional yields trials. Ed. Gauch HG, Elsevier, Amsterdam. 278 p
- Gollob HF. 1968.** A statistical model that combines features of factor analysis and analysis of variance techniques. *Psychometrika*, 33: 73-115.
- Hargreaves GL, Hargreaves GH, Riley JP. 1985.** Irrigation water requirements for Senegal River Basin. *Journal of Irrigation and Drainage, Engineering.*, ASCE, III: 265-275.
- Igartua E, Gracia MP, Lasa JM, Medina B, Molina-Cano JL, Montoya JL, Romagosa I. 1998.** The Spanish barley core collection. *Genetic Resources and Crop Evolution*, 45: 475-481.
- IPGRI. 1994.** Descriptors for barley (*Hordeum vulgare* L.). International Plant Genetic Resources Institute, Rome, Italy, 45 p.
- Lasa JM, Igartua E, Ciudad FJ, Codesal P, García EV, Gracia MP, Medina B, Romagosa I, Molina-Cano JL, Montoya JL. 2001.** Morphological and agronomical diversity patterns in the Spanish barley core collection. *Hereditas*, 135: 217-225.
- Manners JG. 1986.** Introducción a la fitopatología. Ed. Limusa. México. 295 p.
- Romagosa I. 2001.** Statistics and experimental design. Plant Breeding Course. IAMZ. CIHEAM. (octubre 2000-junio 2001)
- SAS Institute. 1988.** SAS/STATTM User's Guide, Release 6.03. Edition. Cary, NC: SAS Institute Inc., 1028p
- Slafer GA, Araus JL, Royo C, García del Moral LF. 2005.** Promising eco-physiological traits for genetic improvement of cereal yields in Mediterranean environments. *Annals of Applied Biology*, 146: 61-70.
- Stone PJ, Nicolas ME. 1994.** Wheat cultivars vary widely in responses of grain yield and quality to short periods of post-anthesis heat stress. *Australian Journal of Plant Physiology*, 21: 887-900.
- Tolbert DM, Qualset CO, Jain SK, Craddock JC. 1979.** A diversity analysis of a world collection of barley. *Crop Science*, 19: 789-794.

- van Oosterom EJ, Ceccarelli S, Peacock JM. 1993.** Yield response of barley to rainfall and temperature in Mediterranean environments. *Journal of Agricultural Sciences, Cambridge*, 121: 307-313.
- van Eeuwijk FA. 1995.** Linear and bilinear models for the analysis of multi-environment Trials: I. An inventory of models. *Euphytica*, 84: 1-7.
- Voltas J, van Eeuwijk FA, Sombrero A, Lafarga A, Igartua E, Romagosa I. 1999.** Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley I. Individual grain weight. *Field Crops Research*, 62: 63-74.
- Villegas D, Aparicio N, Nachit MM, Araus JL, Royo C. 2000.** Photosynthetic and developmental traits associated with genotypic differences in durum wheat yield across the Mediterranean basin. *Australian Journal of Agricultural Research*, 51: 891-901.

Capítulo 6

Mapeo por Asociación

6-1. Introducción

El análisis de QTLs es un enfoque útil para descubrir y diseccionar caracteres complejos e identificar alelos favorables (Paterson *et al.*, 1988). La identificación de marcadores moleculares estrechamente ligados a QTLs de interés agronómico es el primer paso para poder desarrollar una selección asistida por marcadores.

Habitualmente, la búsqueda de QTLs en plantas se ha llevado a cabo en poblaciones con un pedigrí conocido, tales como generaciones F₂, BC₁, líneas doble haploides o RILs, procedentes de cruces de dos parentales. Sin embargo, este enfoque presenta varias limitaciones. Por un lado, estas poblaciones son el producto de unos pocos ciclos de recombinaciones meióticas, lo que limita la resolución de los mapas genéticos que proporcionan (Varshney *et al.*, 2005). Por otra parte, estas poblaciones permiten el estudio simultáneo de solamente 2 alelos por locus (Flint-Garcia *et al.*, 2003), lo que no es representativo del germoplasma de la especie. Por eso, los QTLs detectados en un cruzamiento no son directamente extrapolables a otros materiales. Por eso, se han buscado otras alternativas que soslayen estas limitaciones.

La búsqueda de asociación entre caracteres fenotípicos y marcadores moleculares en una población natural por el método de *linkage disequilibrium mapping* podría ser un procedimiento complementario al método del mapeo con poblaciones de origen biparental. Este método está basado en la asociación no aleatoria de los alelos de distintos loci, causada por la historia de mutación y recombinación que comparten (Buckler y Thornsberry, 2002; Rafalski, 2002). Los métodos de asociación evalúan si ciertos alelos dentro de una población aparecen en determinados fenotipos con una frecuencia mayor que la esperada.

El *linkage disequilibrium mapping*, también conocido como análisis de asociación, ha sido utilizado extensivamente en la disección del control genético de enfermedades en humanos y en la genética de animales. Recientemente ha sido extendido a plantas para la identificación de asociaciones entre caracteres y marcadores moleculares. En uno de los primeros estudios de este tipo, Virk *et al.* (1996) emplearon regresiones lineales múltiples para explicar 6 caracteres con marcadores moleculares en arroz. Beer *et al.* (1997) usaron el método de *linkage disequilibrium mapping* en 64 *landraces* y cultivares de avena, para los que consiguieron encontrar asociaciones entre 30 caracteres cuantitativos y marcadores moleculares. Thornsberry *et al.* (2001), encontraron polimorfismos en el locus *Dwarf8*, asociados con variación en el tiempo a

la floración en el maíz. También en esta especie, Wilson *et al.* (2004), encontraron asociaciones entre caracteres de composición del grano (proteína, almidón, etc.) y genes candidatos. Gebhardt *et al.* (2004) emplearon el *linkage disequilibrium mapping* para buscar asociaciones entre marcadores y la resistencia a enfermedades y duración del ciclo en la patata. En la cebada, Igartua *et al.* (1999), examinaron la asociación entre RFLPs, localizados en las zonas del genoma responsables de la respuesta a fotoperíodo y vernalización, y el carácter espigado. Kraakman *et al.* (2004), evaluaron asociaciones entre AFLPs y caracteres como rendimiento, estabilidad y adaptabilidad. Skøt *et al.* (2005) usaron este método, para identificar genes de floración en poblaciones naturales de *Lolium perenne* (L.).

Este método aumenta considerablemente la resolución del mapeo en comparación con las poblaciones estándar de mapeo. Además, muchas especies de plantas presentan altos niveles de diversidad, por los cuales el método de asociación sería más adecuado para evaluar los numerosos alelos existentes (Flint-Garcia *et al.*, 2003). En una completa revisión sobre este método, Gupta *et al.* (2005), también resaltaron que el mapeo por desequilibrio de ligamiento es más problemático metodológicamente, pero también ofrece una mayor recompensa, porque permite una localización más precisa de la posición de los QTL que controlan el carácter de interés. Comparando el análisis de ligamiento clásico en poblaciones biparentales (*linkage analysis*) y el *linkage disequilibrium mapping* para la detección de QTL, se ha demostrado que el *linkage analysis* es más útil para la búsqueda de QTLs a lo largo de todo el genoma, mientras que el *linkage disequilibrium mapping* proporciona una localización más precisa de un QTL. De hecho, la utilidad del mapeo por asociación en los cultivos vendrá determinada por la estructura genética de las poblaciones y, sobre todo, por la extensión del desequilibrio de ligamiento. Si éste decae rápidamente en distancias de unos miles de pares de bases, el mapeo de todo el genoma requeriría un número de marcadores irrealizable, y el método serviría para analizar asociaciones de polimorfismos en regiones concretas del genoma con genes candidato. Si el desequilibrio de ligamiento no decae hasta distancias considerables, incluso en el rango de centiMorgans, el mapeo al nivel de todo el genoma es posible pero, en cualquier caso, con un número suficiente de marcadores (Nordborg *et al.*, 2002, Flint-Garcia *et al.*, 2005, Varshney *et al.*, 2005).

El riesgo más obvio de este método es la identificación de falsos positivos (Turpeinen *et al.*, 2003). Sin embargo, Sackville Hamilton *et al.* (2002) proponen que una selección cuidadosa de la población a analizar podría reducir estos riesgos. Entre las

medidas que proponen para efectuar esa selección están el incluir poblaciones con el mismo fenotipo y provenientes de diferentes orígenes geográficos, junto con selección de poblaciones de fenotipos divergentes originarios de la misma zona geográfica, y repetir este proceso para diferentes localidades. Sköt *et al.* (2005) comentaron que estas etapas pueden ayudar en la reducción de falsos positivos, pero no puedan eliminar la estructura de población que es la causa más frecuente de la detección de asociaciones espurias. Se han desarrollado diversos programas estadísticos que permiten detectar la estructura de poblaciones y realizar los análisis de asociación eliminando el efecto de esa estructura (Devlin y Roeder, 1999; Pritchard *et al.* 2000), lo que podría solucionar este problema. Sin embargo, hace falta evaluar el comportamiento de estos métodos en situaciones reales.

En este sentido, Remington *et al.* (2001), comentaron que asociaciones encontradas entre caracteres fenotípicos y microsatélites en maíz, debidas el efecto de la estructura de población, desaparecieron al centrar los análisis en una sub-población. Pero ciertas asociaciones entre los microsatélites y el carácter floración mostraron altos niveles de significación dentro de las sub-poblaciones.

El objetivo de este estudio es el análisis de asociación o *linkage disequilibrium mapping* en las entradas de la colección nuclear española de cebada. Se analiza la asociación de marcadores moleculares con treinta caracteres agronómicos y morfológicos. Algunos de estos caracteres son de herencia simple con genes de localización conocida, por lo que podrán emplearse como testigos para comprobar la bondad del método. También se evalúa el efecto de la estructura de la población, descrita en el capítulo 3, sobre la detección de asociaciones.

6-2. Material y Métodos

6-2-1. Material vegetal

En este capítulo nos centramos sobre el estudio de ciento cincuenta y seis líneas puras de la colección nuclear (de un total de 159, excluidos los tres duplicados descritos en el capítulo 3). Estas se dividen en 11 entradas de 2 carreras y 145 entradas de 6 carreras. La procedencia de estos genotipos esta descrita en la Tabla 3.2 del capítulo 3. Se han excluido de este estudio los cultivares de la colección nuclear y otros cultivares de referencia. Los motivos son, por un lado que uno de los objetivos de este estudio es identificar lo mejor posible variabilidad genética presente en la colección nuclear, que

pueda ser empleada posteriormente en programas de mejora. Por otro lado, el conjunto total de genotipos analizados en el capítulo 3 es una representación desequilibrada del germoplasma de la cebada, y genéticamente muy heterogénea, como quedó de manifiesto en el capítulo 3. Por tanto, cabe esperar que los genotipos de la colección nuclear, muy mayoritarios, pesen más en los resultados de los análisis de asociación. El gran desequilibrio que existe entre *landraces* y cultivares afectaría fuertemente a los análisis. Por tanto, con objeto de conseguir unos resultados más claros y representativos, se decidió acometer esta parte del trabajo sólo con las *landraces* de la colección nuclear.

6-2-2. Marcadores moleculares

Un conjunto de setenta y tres marcadores moleculares fueron empleados en este análisis. Sesenta y cinco de estos marcadores están descritos en el apartado 3-1-2-2 del capítulo 3, mientras que los otros ocho están descritos en el apartado 4-2-2 del capítulo 4.

6-2-3. Caracteres estudiados

Las 156 *landraces* analizadas en este estudio fueron evaluadas en 10 ambientes. Los datos de los caracteres agronómicos, de incidencia a enfermedades y de calidad maltera, registrados en estos ensayos, fueron utilizados en este estudio. Las condiciones experimentales así como los detalles sobre estos caracteres están descritos en el capítulo 5.

Una serie de caracteres morfológicos descriptivos fueron también incluidos en este estudio. Estos caracteres, catorce en total, fueron tomados por personal del laboratorio de mejora de cebada de la Estación Experimental de Aula Dei, siguiendo las normas del IPGRI (1994), y se describen a continuación:

Pigmentación del tallo y de la aurícula

Estos dos caracteres fueron observados en un total de 5 plantas por entrada, en el campo. Se distinguieron dos clases de pigmentación:

- 1- Verde.
- 2- Púrpura.

Vellosidad de la vaina basal

Este carácter fue evaluado en 3 plantas por entrada, cultivadas en condiciones de invernadero. Este carácter presenta 2 clases que son:

- 1- Ausencia.
- 2- Presencia.

Glumas y aristas de la gluma

- 1- Longitud de la gluma y de la arista más corta que el grano.
- 2- Longitud de la gluma y de la arista igual que el grano.
- 3- Gluma y arista más larga que el grano.

Tipo de Lemma

- 1- Lemma lisa no dentada.
- 2- Lemma dentada.
- 3- Lemma vellosa.

Forma de las aristas de la lemma

- 1- Sin aristas.
- 2- Aristas cortas.
- 3- Aristas largas.
- 4- Arista en forma de capucha sécil.
- 5- Arista en forma de capucha alargada.

Barbas de las aristas de la lemma

- 3- Arista lisa (pocas barbas al extremo de la arista).
- 5- Arista intermedia (barbas cortas en la mitad superior de la arista).
- 7- Arista rugosa.

Color de la gluma

- 1- Blanca.
- 2- Amarilla.
- 3- Marrón.
- 4- Negra.

Color de las aristas

- 1- Blancas.
- 2- Amarrillas.
- 3- Marrones.
- 4- Rojas.
- 5- Negras.

Longitud de los pelos de la raquilla

- 1- Pelos cortos.
- 2- Pelos largos.

Longitud y densidad de espiga

Estos dos caracteres fueron evaluados eligiendo 5 espigas representativas de cada entrada. La longitud de espiga se clasificó según 5 categorías, de acuerdo con la distribución de valores observada:

- 1- Espigas inferiores a 4,90 cm.
- 2- Espigas entre 4,90 y 6,46 cm.
- 3- Espigas de 6,46 a 8,02 cm.
- 4- Espigas de 8,02 cm a 9,58 cm.
- 5- Espigas superiores a 9,58 cm.

Para la densidad de espiga se consideran 3 clases que son:

- 3- Laxa.
- 5- Semidensa.
- 7- Densa.

Espiguillas por espiga

El número de tripletes de espiguillas por espiga fue evaluado eligiendo 5 espigas representativas de cada entrada. Se determinaron 5 clases, según la variación del número de tripletes.

- 1- Menos de 14 tripletes.
- 2- Entre 14 y 16 tripletes.
- 3- Entre 17 y 19 tripletes.

4- Entre 20 y 22 tripletes.

5- Más de 22 tripletes.

Tipo de espiga

2- Dos carreras.

6- Seis carreras.

Los detalles de la distribución de estos caracteres morfológicos en las entradas de la colección nuclear, no se presentarán en los resultados, por no ser el objetivo de este estudio. La Tabla 6.1 resume los caracteres analizados en este trabajo.

Tabla 6.1 Caracteres analizados

Caracteres agronómicos	Rendimiento (RDT)
	Espigado (ESP)
	Encamado (ENC)
	Altura de planta (ALT)
	Peso hectolítrico (PHL)
	Peso de mil granos (PMG)
	Espigas m ⁻² (ESM)
	Granos m ⁻² (GRM)
	Granos por espiga (GRE)
Enfermedades	Roya (RYA)
	Moteado de las hojas (MOT)
	Oídio (OID)
Calidad maltera	Porcentaje de granos de calibre inferior a 2,2 mm (CA1)
	Porcentaje de granos de calibre superior a 2,5 mm (CA2)
	Porcentaje de proteína total en harina de malta (PRT)
	Porcentaje de materia seca en el extracto de malta (MLT)
Caracteres morfológicos	Pigmentación del tallo (PIT)
	Pigmentación de la aurícula (PIA)
	Vellosidad de la vaina basal (VVB)
	Glumas y aristas de la gluma (GAG)
	Tipo de lemma (TLM)
	Forma de las aristas de la lemma (FAL)
	Barbas de las aristas de la lemma (BAL)
	Color de la gluma (COG)
	Color de las aristas (COA)
	Longitud de los pelos de la raquilla (LPR)
	Longitud de Espiga (LES)
	Densidad de espiga (DES)
	Espiguillas por espiga (EES)
	Tipo de espiga (TES)

6-2-4. Análisis de asociación

La asociación entre los marcadores moleculares y los caracteres estudiados se realizó mediante análisis de regresión entre los caracteres y los marcadores para las *landraces* de la colección nuclear. Éste se llevó a cabo de dos maneras distintas, teniendo en cuenta o no la estructura poblacional. Además se efectuó otro análisis considerando sólo los individuos pertenecientes a una de las poblaciones (el grupo 4), en la que el desequilibrio de ligamiento fue mínimo. A partir de ahora se emplearán los términos “grupo” o “población” como sinónimos.

En primer lugar, se evaluó la asociación de cada uno de los marcadores individualmente con cada uno de los caracteres estudiados. Aquellos marcadores que resultaron asociados significativamente con cada carácter se analizaron posteriormente en modelos multilocus. El motivo de realizar estos últimos análisis es identificar sólo aquellas regiones que puedan estar más cerca de verdaderos QTLs, minimizando los falsos positivos que pudieran derivar de la presencia de desequilibrio de ligamiento entre marcadores no ligados, un problema similar al de la multicolinealidad en regresión múltiple. De este modo es esperable que se cometan más errores de tipo 2 (no detección de asociaciones verdaderas) pero, para el propósito de este estudio, son más tolerables los errores de tipo 2 que los de tipo 1 (detección de asociaciones no verdaderas).

A continuación se explican estos distintos modelos de regresión con más detalle.

6-2-4-1. Regresión marcador a marcador

Se evaluaron las asociaciones entre cada marcador y el carácter considerado mediante un análisis de regresión múltiple, en el cual el carácter fenotípico era la variable dependiente y todos los alelos de cada marcador eran las variables independientes. En cada análisis se excluyeron los alelos presentes con una frecuencia inferior al 5%, o sea, presentes en menos de 8 individuos en el conjunto de las 156 *landraces*, o en menos de 4 individuos en el caso del grupo 4, como recomendaron Wilson *et al.* (2004). La información de los marcadores se codificó creando una variable para cada alelo, expresando mediante un “1” la presencia de un alelo en un individuo, y con un “0” su ausencia en el resto de variables. Por ejemplo, un marcador con dos alelos quedaba plenamente descrito con dos variables.

Los distintos modelos fueron calculados a partir de diferentes análisis de regresión, siguiendo el proceso que se detalla a continuación:

- 1) Análisis sin tener en cuenta la estructura poblacional que, aplicado a las 156 *landraces*, proporciona directamente el modelo SG y aplicado a las entradas del grupo 4 da como resultado el modelo G4.

carácter = alelos del marcador

- 2) Análisis de regresión con el modelo completo

carácter = grupos + alelos del marcador

- 3) Análisis de regresión entre el carácter y los grupos, que permite estimar la variación del carácter debida a la estructura de población.

carácter = grupos

- 4) Calculo de la variación debida a exclusivamente al marcador (modelo CG), restando la suma de cuadrados del modelo (3) de la del modelo (2). Para ello, y para cada marcador, se calculó la suma de cuadrados explicada por los alelos del mismo, como diferencia de la suma de cuadrados del modelo completo y la explicada por los clusters. A partir de la suma de cuadrados que corresponde a los alelos, se calculó el cuadrado medio dividiendo por los grados de libertad correspondientes al número de alelos del marcador, menos uno. Posteriormente se calculó una F, como cociente entre el cuadrado medio de los alelos y el cuadrado medio del error del modelo completo. La significación de la F fue evaluada con el test de distribución de probabilidad F. Se eligieron dos niveles de significación de 0,0007 y 0,0014, que corresponden a una corrección de Bonferroni para un nivel α igual a 0,05 y 0,1 respectivamente, para una comparación múltiple de 73 marcadores.

En los modelos que tienen en cuenta la estructura poblacional, la probabilidad de pertenencia de cada individuo a cada población o grupo (matriz Q proporcionada por el análisis de STRUCTURE, ver capítulo 3) se incluyó en el análisis como cuatro variables independientes en el modelo, previamente a la introducción de la información de los marcadores.

Todos estos análisis se llevaron a cabo mediante el programa SAS con el procedimiento *proc reg* (SAS, 1988)

Los coeficientes de determinación ajustados de los loci, calculados para cada modelo (SG, CG, G4), fueron comparados mediante un análisis de varianza (para cada carácter), con el procedimiento *proc glm* (SAS, 1988).

6-2-4-2. Análisis multilocus

Los loci que presentaron asociaciones significativas para un carácter fueron analizados en modelos multilocus. Este análisis se realizó para ver si existen efectos aditivos entre los alelos de los diferentes loci, y por otra parte determinar cuál es la combinación de loci que explica mejor el carácter analizado.

Para ello, una vez identificados los loci que presentaban asociación en el análisis marcador a marcador, se analizaron todas las combinaciones posibles entre esos loci, introduciendo desde 2 a n loci en el modelo, siendo n el número total de loci significativos. La asociación entre los marcadores y el carácter se evaluó mediante el procedimiento *proc mixed*, método *ML* (*Maximum likelihood*) del programa SAS (SAS, 1988), considerando los loci como factor fijo y los grupos como factor aleatorio.

La elección del mejor modelo se llevó a cabo atendiendo a los siguientes criterios de información o ajuste (*information criteria*): AIC (Akaike, 1974), AICC (Hurvich y Tsai, 1989), HQIC (Hannan y Quinn, 1979), BIC (Schwarz, 1978) y CAIC (Bozdogan, 1987).

Para cada carácter se fueron evaluando de forma sucesiva distintos modelos, incluyendo combinaciones de todos los loci significativos en el análisis marcador a marcador. En cada ronda de análisis (de 2 a n loci en el modelo) se eligió aquel modelo en el que todos los marcadores incluidos fueran significativos y que presentaba los valores más bajos de la mayoría de los criterios de información.

Este análisis multilocus fue también ejecutado, considerando solamente las entradas de la población 4. En este caso, los únicos factores que se introducen en el modelo son los loci, que se consideraron como factores fijos. Se repitió el mismo proceso descrito anteriormente para la elección del modelo más explicativo del carácter.

Para el modelo más explicativo de cada carácter, se procedió a una separación de medias de los alelos de cada locus incluido en el modelo. Este análisis, se realizó mediante la opción *lsmeans / pdiff* del programa SAS (SAS, 1988).

Se calculó un coeficiente de determinación R^2 aproximado, para el modelo elegido de cada carácter. Éste análisis se llevó a cabo mediante el procedimiento *proc glm* del programa SAS (SAS, 1988). En el análisis de las 156 *landraces*, se dedujo el coeficiente de determinación de los loci del modelo, restando del modelo completo (grupos + loci), la parte correspondiente a los grupos.

6-3. Resultados

6-3-1. Análisis de regresión locus por locus

Este análisis permitió detectar un cierto número de asociaciones entre las variables estudiadas y los marcadores analizados. Este número varió según el carácter analizado y el modelo de análisis utilizado. Sobre 2190 comparaciones posibles (30 variables y 73 marcadores) se detectaron 311 asociaciones positivas, o sea el 14,2% con el modelo SG (sin tener en cuenta la estructura de población). El número de asociaciones positivas se redujo a 65 casos (2,9%) al considerar la estructura de población (modelo CG). Por otro lado, considerando únicamente la población 4 (modelo G4) el número de casos de asociaciones positivas descendió a 45, o sea el 2,1% del total de las comparaciones (Tabla 6.2). Comparando los loci con asociaciones positivas entre los distintos modelos de análisis, se observó que había 48 loci comunes entre los modelos SG y CG. El modelo G4 tuvo 22 y 20 loci en común con los modelos SG y CG, respectivamente. La comparación entre los 3 modelos mostró 15 asociaciones que fueron detectadas en todos los casos.

Para los caracteres donde se detectaron asociaciones con los loci, se calcularon los promedios de los coeficientes de determinación ajustados de los distintos modelos. El análisis de varianza llevado a cabo para comparar estos promedios de los diferentes modelos permitió ver que, en la mayoría de los casos, estos son mucho más altos en el modelo donde no se tiene en cuenta la estructura de población (SG), que en los modelos realizados teniendo en cuenta esta estructura (CG), o en el modelo del grupo 4 (G4), como se muestra en las Figuras 6.1 y 6.2. Este fenómeno fue más evidente para los caracteres agronómicos que los morfológicos. Por otra parte, el modelo G4 mostró un promedio relativamente más alto que el modelo CG, para todas las variables analizadas (Figuras 6.1 y 6.2)

Tabla 6.2. Número de loci que presentaron asociaciones positivas con las variables estudiadas, según el modelo de análisis empleado (sin grupos, con grupos o grupo 4); número de loci comunes entre los diferentes modelos.

Variables	SG	CG	G4	Comunes SG-CG	Comunes SG-G4	Comunes CG-G4	Comunes SG-CG-G4
Rendimiento	21	9	2	5	1	1	1
Espigado	17	12	16	8	9	10	6
Encamado	14	2	1	2	0	0	0
Altura	8	0	0	0	0	0	0
Peso hectolítrico	34	0	1	0	1	0	0
Peso de mil granos	19	2	2	2	2	2	2
Espigas m ²	3	0	0	0	0	0	0
Granos m ²	0	0	1	0	0	0	0
Granos por espiga	6	0	0	0	0	0	0
Roya	21	4	4	3	4	3	3
Moteado de la hojas	2	0	0	0	0	0	0
Oídio	14	0	0	0	0	0	0
Calibre <2,2 mm	9	2	5	2	2	1	1
Calibre >2,5 mm	9	0	0	0	0	0	0
Proteína	9	0	1	0	0	0	0
Extracto de malta	1	0	0	0	0	0	0
Pigmentación del tallo	3	3	0	2	0	0	0
Pigmentación de la aurícula	2	0	4	0	0	0	0
Vellosidad de la vaina basal	13	4	0	2	0	0	0
Glumas y aristas de la gluma	9	4	0	3	0	0	0
Tipo de lemma	18	3	0	2	0	0	0
Forma de las aristas de la lemma	0	0	0	0	0	0	0
Barbas de las aristas de la lemma	0	0	0	0	0	0	0
Color de la gluma	0	0	0	0	0	0	0
Color de las aristas	0	0	0	0	0	0	0
Longitud de los pelos de la raquilla	11	5	3	4	2	2	2
Longitud de espiga	32	5	3	3	1	1	0
Densidad de espiga	3	2	1	2	0	0	0
Espiguillas por espiga	1	0	1	0	0	0	0
Tipo de espiga	32	8	0	8	0	0	0
Total	311	65	45	48	22	20	15

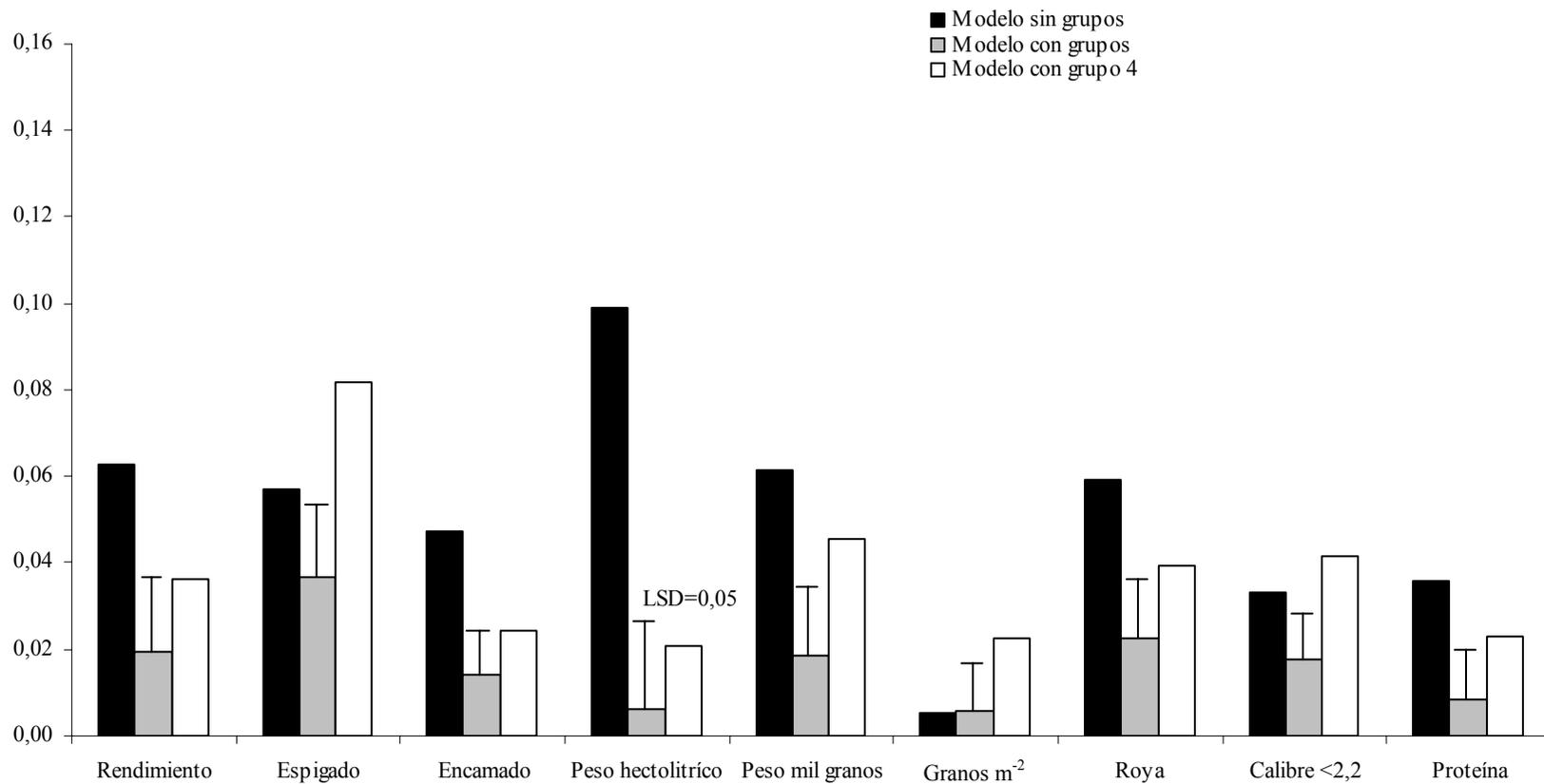


Figura 6.1. Promedios de los coeficientes de determinación ajustados resultantes de los análisis de regresión entre los 73 marcadores y las variables agronómicas, de tolerancia a enfermedades y calidad maltera según los 3 modelos SG, CG y G4. Solo se representaron los promedios de las variables que mostraron alguna asociación significativa en los modelos CG y G4.

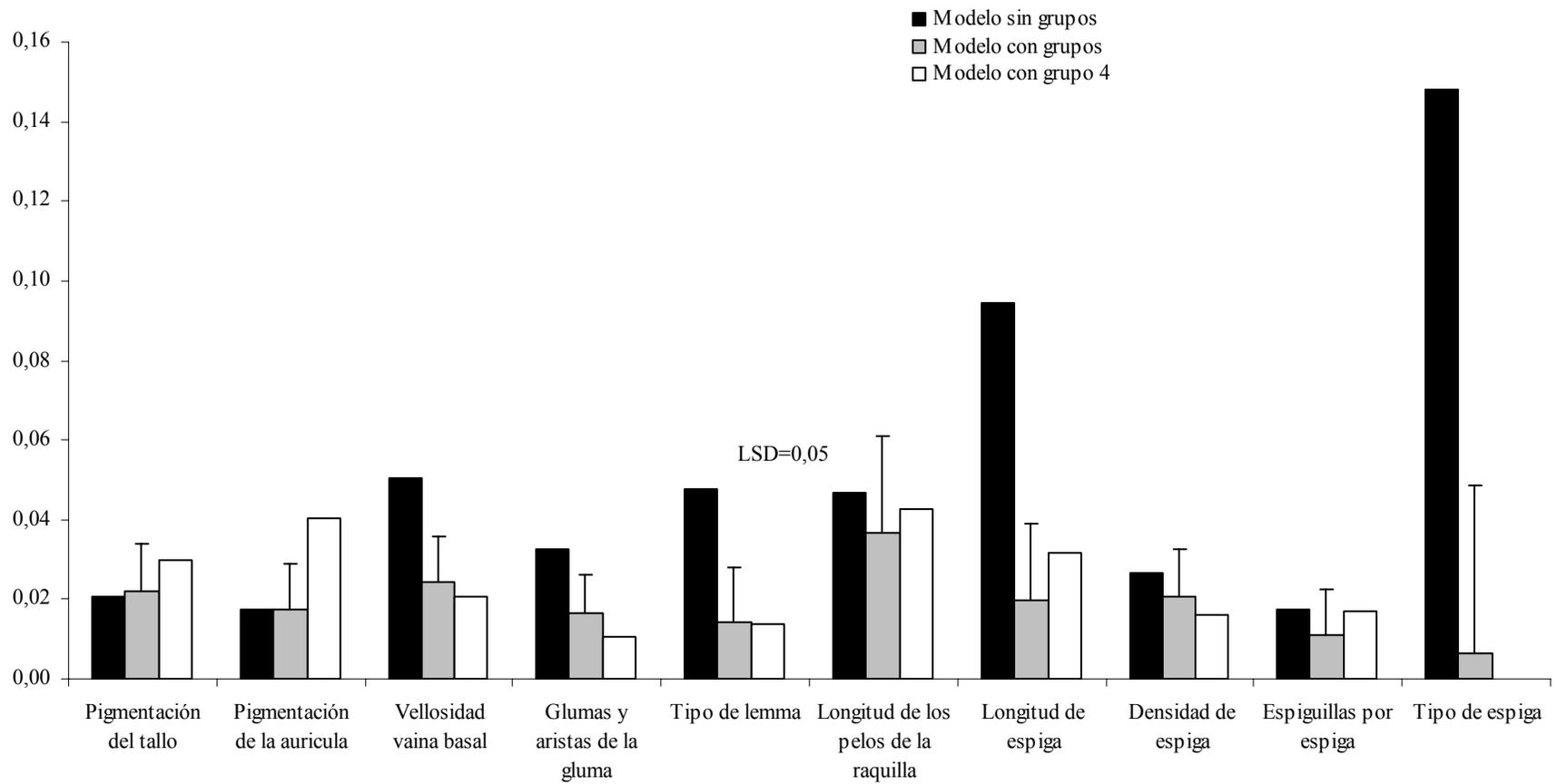


Figura 6.2. Promedios de los coeficientes de determinación ajustados resultantes de los análisis de regresión entre los 73 marcadores y las variables morfológicas según los 3 modelos SG, CG y G4. Solo se representaron los promedios de las variables que mostraron alguna asociación significativa en los modelos CG y G4.

6-3-1-1. Asociaciones con caracteres agronómicos

Se detectaron asociaciones con 6 variables agronómicas: rendimiento, espigado, encamado, peso hectolítrico, peso de mil granos y granos m^{-2} . Para las variables altura de planta, número de espigas m^{-2} y granos por espiga, se detectaron asociaciones sólo en el modelo SG, o sea, sin incluir la estructura de población. Estas desaparecieron al introducir en el análisis la estructura de población, o al considerar el modelo G4 (Tabla 6.3). A continuación se detallan las asociaciones encontradas para cada carácter.

Rendimiento

Veintiuna asociaciones positivas fueron detectadas entre esta variable y los 73 loci analizados con los modelos realizados sin tener en cuenta la estructura de población (SG). Estas se redujeron a 9, al tener en cuenta dicha estructura (CG). Cinco de ellas son comunes entre los dos modelos. Los 9 loci se distribuyen en 6 cromosomas de la forma siguiente: en los cromosomas 1H, 5H, 6H y 7H se detectó una asociación positiva en cada cromosoma, que corresponden respectivamente a los loci WMC1E8, HvBM5A, Bmag009 y HvCMA. En el cromosoma 2H, 2 loci, Bmac134 y scssr07759 se mostraron significativos y en el cromosoma 4H, 3 loci, Bmag353, EBmac701 y HvBAMY, revelaron asociaciones positivas. Sin embargo, ninguna asociación fue detectada en el cromosoma 3H (Tabla 6.3). Los loci Bmac134, EBmac701 y HvBM5A fueron los más explicativos para este carácter, con unos coeficientes de determinación de 0,129, 0,135 y 0,133, respectivamente.

Por otro lado, en el modelo G4 sólo se encontraron dos asociaciones que corresponden a los loci scssr07759 (2H) y Bmac018 (6H) (Tabla 6.3). Los coeficientes de determinación de estos dos loci fueron respectivamente de 0,233 y 0,190. El locus scssr07759, es el único marcador significativo común entre los diferentes modelos.

La comparación de los promedios de los coeficientes de determinación de los 73 loci mostró que el modelo SG explicó en general un porcentaje de varianza mayor que los otros dos (Figura 6.1). La diferencia encontrada entre los promedios de los 2 modelos sin y con grupos corresponde a la parte explicada por la estructura de población.

Espigado

El número de loci asociados a este carácter varió de 12 (CG) a 17 (SG). Ocho loci fueron comunes entre los dos modelos (Tabla 6.2). Los doce loci encontrados asociados al espigado en las 156 *landraces* se distribuyeron en los 7 cromosomas como sigue: Bmac032 (1H), HvM020 (1H), Bmac134 (2H), scssr07759 (2H), EBmac640 (2H), HvLTPPB (3H), HvM062 (3H), Bmac029 (3H), HvBAMY (4H), Bmag223 (5H), Bmac018 (6H) y Bmag369 (7H).

Los loci HvLTPPB, scssr07759, Bmac032 y Bmag369 resultaron ser los más explicativos, con unos coeficientes de determinación respectivos de 0,153, 0,160, 0,133 y 0,134.

Por otro lado, al considerar solamente las entradas del grupo 4, se detectaron 16 asociaciones que correspondieron a los loci siguientes: HvGLUEND (1H), Bmac134 (2H), MWG858 (2H), scssr07759 (2H), EBmac640 (2H), MWG699 (2H), HvLTPPB (3H), Hv13GEIII (3H), HvM062 (3H), Bmac029 (3H), EBmac701 (4H), HvBAMY (4H), Bmag337 (5H), Bmag223 (5H), Bmac018 (6H) y Bmag369 (7H). En este caso, los marcadores más explicativos fueron EBmac701, Bmac134 y scssr07759, con unos coeficientes de determinación notables (0,335, 0,319 y 0,287, respectivamente). El número de marcadores comunes entre este modelo y los modelos SG y CG fueron, respectivamente, 9 y 10. Mientras que sólo 6 loci significativos fueron comunes entre los 3 modelos.

El análisis de varianza de los promedios de los coeficientes de determinación ajustados de los loci, calculados para cada modelo, reveló diferencias significativas entre los promedios de los 3 modelos (Figura 6.1). Curiosamente, el promedio correspondiente al modelo G4 resultó ser el más alto.

Encamado

Al no incluir la estructura de población en los análisis, 14 loci mostraron asociaciones positivas con este carácter. Estas se redujeron a dos loci al quitar esta estructura. Estos dos loci son Bmac032 (1H) y Bmag006 (3H). Los coeficientes de determinación de estos dos loci fueron respectivamente de 0,126 y 0,154.

En el grupo 4, el locus Bmac064 fue el único que mostró una asociación significativa con este carácter, y no se había revelado significativo con el modelo sin grupos (Tabla 6.3). Su R^2 fue bastante elevado (0,193).

Los coeficientes de determinación de los modelos CG y G4 fueron similarmente bajos, y significativamente menores que los del modelo SG (Figura 6.1).

Peso hectolítrico

Para esta variable se detectaron 34 asociaciones al hacer los análisis sin tener en cuenta la estructura de población (SG), pero estas asociaciones desaparecieron por completo en el modelo CG. Sin embargo, los análisis llevados a cabo solamente con el grupo 4, revelaron la asociación del locus EBmac701 (4H) con este carácter, con un coeficiente de determinación de 0,321. Este locus también había resultado significativo con el modelo SG (Tabla 6.3).

Como en el carácter anterior, los promedios de los coeficientes de determinación de los modelos CG y G4 fueron similares, e inferiores al del modelo SG (Figura 6.1).

Peso de mil granos

Los análisis con el modelo SG mostraron 19 asociaciones, que se redujeron a 2 al incluir la estructura de población. Estos dos loci fueron MWG699 (2H) y HvBM5A (5H). Los dos se revelaron significativos en los modelos SG y G4 (Tabla 6.3). Los coeficientes de determinación de estos dos loci fueron respectivamente de 0,103 y 0,106 en las 156 *landraces* y de 0,172 y 0,187 en el modelo G4.

La comparación de los promedios de los coeficientes de determinación de los 3 modelos presentó diferencias significativas, siendo mayor el promedio del modelo SG (Figura 6.1).

Granos m⁻²

Únicamente el microsatélite HdAMYB (4H) presentó una asociación positiva con esta variable, en el análisis del modelo G4 (Tabla 6.3). De hecho, el promedio de los coeficientes de determinación para esta variable, fue más alto en este modelo que en los otros dos, que no se diferenciaron entre sí (Figura 6.1). El locus HdAMYB presentó un coeficiente de determinación de 0,229.

6-3-1-2. Asociaciones con tolerancia a enfermedades

Los análisis de regresión entre los marcadores y los datos de incidencia de enfermedades en campo revelaron asociaciones positivas, sobre todo en el modelo SG. El número de asociaciones con ese modelo fue de 2, 14 y 21 para moteado de las hojas, oídio y roya, respectivamente. Sin embargo, al incluir la estructura de población, estas asociaciones desaparecieron en el caso del oídio y del moteado de las hojas, mientras que en el caso de la roya, sólo 4 loci mostraron una asociación positiva. Lo mismo ocurrió en el modelo G4 (Tabla 6.3). Los loci que se revelaron significativos para la roya en el modelo CG fueron Bmac399 (1H), HvBTAI3 (4H), Bmag337 (5H) y HvBM5A (5H), con unos coeficientes de determinación respectivos de 0,165, 0,085, 0,115 y 0,170. El locus Bmag337 no mostró una asociación significativa en el modelo SG (Tabla 6.3). En el modelo G4 fueron significativos cuatro marcadores que ya aparecían en el modelo SG, y tres de ellos en el modelo CG. Estos fueron Bmac399 (1H), MWG699 (2H), HvBTAI3 (4H) y HvBM5A (5H). En este último modelo, estos 4 loci presentaron coeficientes de determinación de 0,361, 0,128, 0,160 y 0,200, respectivamente.

La comparación de los promedios de los coeficientes de determinación de los 3 modelos presentó diferencias significativas. El promedio del modelo G4 fue mayor que el promedio del modelo CG (Figura 6.1).

Tabla 6.3. Asociaciones positivas de los loci con las variables estudiadas, según el modelo de análisis empleado (A: sin estructura de población o SG; B: considerando la estructura de población o CG; C: grupo 4 ó G4); * y X corresponden respectivamente a un nivel de significación α de 0,05 ó de 0,1, empleando la corrección de Bonferroni para 73 tests independientes.

6-3-1-3. Asociaciones con calidad maltera

Los análisis de regresión entre marcadores y los datos de calidad maltera permitieron detectar asociaciones con las 4 variables estudiadas en el modelo SG. Sin embargo, al analizar estas variables, teniendo en cuenta la estructura de población, o analizando únicamente la población 4, solamente dos variables presentaron asociaciones positivas con los loci. Estas variables fueron el calibre inferior a 2,2 mm y el contenido de proteína del grano.

Calibre inferior a 2,2 mm

Nueve loci mostraron una asociación significativa con esta variable en el modelo SG, que se redujeron a 2 en el modelo CG. Estos 2 loci fueron Bmac134 (2H) y Bmag125 (2H), y presentaron respectivamente coeficientes de determinación bastante elevados, de 0,182 y 0,125. Cinco loci mostraron asociaciones significativas en el modelo G4. Tres de estos loci no se revelaron significativos en el modelo SG y únicamente uno de los 5 loci (Bmag125) fue común para los tres modelos (Tabla 6.3). Los 5 loci se distribuyen en sólo 2 cromosomas: scssr07759 (2H), Bmag125 (2H), HvM054 (2H), Bmag225 (3H) y Hv13GEIII (3H) y explicaron respectivamente 0,186, 0,255, 0,189, 0,240 y 0,256 de este carácter.

Los promedios de los coeficientes de determinación de los 3 modelos se diferenciaron significativamente. El modelo G4 presentó un promedio mayor que los otros dos modelos (Figura 6.1).

Proteína

Nueve loci mostraron asociaciones con esta variable en el modelo SG. Estas desaparecieron al incluir en el análisis la estructura de población (Tabla 6.2). En el modelo G4 se detectó la asociación del locus Bmac067 con esta variable ($R^2=0,232$). Este marcador no se reveló significativo en el modelo SG (Tabla 6.3).

Los promedios de los coeficientes de determinación de los 3 modelos presentaron diferencias significativas; el promedio más alto se registró en el modelo SG (Figura 6.1).

6-3-1-4. Asociaciones con variables morfológicas

Para facilitar la interpretación de los análisis de asociación con las variables morfológicas, se incluyen en la Tabla 6.4 los resultados del polimorfismo fenotípico observado en las 159 entradas de la colección nuclear, aunque solo se utilizaron los datos de 156 en los análisis.

Tabla 6.4. Distribución fenotípica de las variables morfológicas en las 159 *landraces*, de acuerdo con la escala aplicada en la descripción de cada carácter.

Carácter	Escala								
	1	2	3	4	5	6	7	8	9
Pigmentación del tallo	60	99							
Pigmentación de la aurícula	60	99							
Vellosidad vaina basal	79	80							
Glumas y aristas de la gluma	52	94	13						
Tipo de lemma	5	20	134						
Aristas de la lemma			159						
Barbas de las aristas de la lemma			4		27		128		
Color de la gluma	19		140						
Color de las aristas	19		140						
Longitud de los pelos de la raquilla	90	69							
Longitud de espiga	8	37	86	19	9				
Densidad de espiga			71		71		17		
Espiguillas por espiga	5	38	77	27	11				
Tipo de espiga		11				148			

Los análisis de regresión con las 14 variables morfológicas estudiadas, permitieron detectar asociaciones con 10 de ellas. Para el carácter aristas de la lemma no se observó polimorfismo en las 159 *landraces*, por lo cual no se puede detectar ninguna asociación. Los caracteres barbas de las aristas de la lemma, color de la gluma y color de las aristas presentaron polimorfismo pero no se detectó ninguna asociación con los marcadores, en ningún modelo (Tabla 6.3).

Pigmentación del tallo

Tres loci se revelaron significativos en los modelos SG y CG, siendo comunes dos de ellos entre los dos modelos. Sin embargo, en el modelo G4 no se detectó ninguna asociación (Tabla 6.3). Los loci con asociaciones positivas en el modelo CG fueron MWG699 (2H), Bmag013 (3H) y HdAMYB (4H), y presentaron coeficientes de determinación de 0,068, 0,188 y 0,160, respectivamente.

La comparación de los promedios de los coeficientes de determinación de los 3 modelos no mostró diferencias significativas (Figura 6.2).

Pigmentación de la aurícula

Los análisis de regresión sin la estructura de población detectaron dos loci con asociaciones positivas. En el modelo CG, sin embargo, ningún marcador reveló una asociación significativa (Tabla 6.2). El análisis en el G4 detectó 4 asociaciones: HvBKASI (2H), Hv13GEIII (3H), Bmac029 (3H) y Bmag223 (5H), todos ellos distintos de los detectados para el modelo SG. Los coeficientes de determinación de estos loci fueron respectivamente de 0,192, 0,265, 0,187 y 0,265.

Los promedios de los coeficientes de determinación de los modelos SG y CG no presentaron diferencias significativas, pero ambos presentaron promedios menores al modelo G4 (Figura 6.2).

Vellosidad de la vaina basal

Trece loci mostraron asociaciones positivas con este carácter, en el SG. Estas se redujeron a 4 al incluir la estructura de población. Los loci involucrados fueron Bmac399 (1H), Bmag337 (5H), Bmag500 (6H) y EBmac806 (6H), con coeficientes de determinación de 0,123, 0,107, 0,122 y 0,106, respectivamente. No hubo asociaciones significativas para el modelo G4.

El modelo SG presentó un promedio de los coeficientes de determinación significativamente más alto que los otros dos modelos, CG y G4, que no difieren entre ellos. (Figura 6.2).

Glumas y aristas de la gluma

Al no incluir la estructura de población en los análisis de regresión, se detectaron 9 asociaciones. Sin embargo, solamente 4 marcadores mostraron asociaciones positivas con la inclusión de los grupos en el análisis (modelo CG). Tres de ellos aparecieron en los dos modelos (Tabla 6.2). Los 4 loci que manifestaron asociación con este carácter fueron Bmac399 (1H), MWG699 (2H), Bmag500 (6H) y Bmag369 (7H), y mostraron respectivamente coeficientes de determinación de 0,120, 0,092, 0,122 y 0,075. Ninguna asociación fue detectada en el modelo G4 (Tabla 6.3).

El modelo SG presentó no sólo el que más asociaciones, sino que, en conjunto el promedio de los coeficientes de determinación de este modelo fue superior al de los otros dos modelos (Figura 6.2).

Tipo de lemma

Dieciocho loci mostraron asociaciones con este carácter en el modelo SG. Al incluir en el análisis la estructura de población, el número de asociaciones bajó drásticamente, a sólo tres, que fueron *scsr07759* (2H), *HvBAMY* (4H) y *HvBM5A* (5H). Los coeficientes de determinación de estos loci fueron respectivamente de 0,125, 0,096 y 0,144. Como en el carácter anterior, no se detectaron asociaciones para esta variable con el modelo G4 (Tabla 6.3).

El promedio de los coeficientes de determinación del modelo SG fue significativamente mayor que el de los otros dos modelos, que resultaron similares (Figura 6.2).

Longitud de los pelos de la raquilla

Los análisis mostraron 10 asociaciones con el modelo SG. Al considerar la estructura de población en los análisis, solamente 5 loci mostraron una asociación positiva. Estos fueron *MWG699* (2H), *HvM054* (2H), *Bmac067* (3H), *Bmac113* (5H) y *Bmag223* (5H). Los coeficientes de determinación correspondientes a estos loci fueron respectivamente de 0,086, 0,203, 0,161, 0,136 y 0,467. En el grupo 4, tres loci revelaron una asociación con este carácter: *HvM020* (1H), *HvM054* (2H) y *Bmag223* (5H), y presentaron coeficientes de determinación de 0,325, 0,307 y 0,671. Los loci *HvM054* y *Bmac113* no presentaron una asociación con la variable en el modelo SG.

La comparación de los promedios de los coeficientes de determinación de los 3 modelos no mostró diferencias significativas (Figura 6.2).

Longitud de espiga

Los análisis de regresión realizados sin tener en cuenta la estructura de población mostraron 32 asociaciones con la longitud de la espiga. Con la inclusión de los grupos en los análisis, solamente 5 loci presentaron una asociación positiva. Estos fueron: *Bmac399* (1H), *EBmac701* (4H), *Bmac040* (6H), *HvM004* (7H) y *Bmag120* (7H). Estos loci presentaron coeficientes de determinación de 0,143, 0,137, 0,134, 0,126 y

0,148, respectivamente. En el grupo 4, los loci HvGLUEND (1H), HvM036 (2H) y Bmac040 (6H) son los únicos que presentaron una asociación con este carácter, con coeficientes de determinación de 0,185, 0,187 y 0,307, respectivamente. Los coeficientes de determinación de los modelos CG y G4 no se diferenciaron significativamente, pero los dos presentaron promedios significativamente inferiores al promedio del modelo SG (Figura 6.2).

Densidad de espiga

Tres loci mostraron una asociación con este carácter en el modelo SG. Con la introducción de la estructura de población en el análisis, solamente dos de los tres loci siguieron revelando una asociación significativa. Estos fueron Bmac399 (1H) y Bmag120 (7H), con coeficientes de determinación de 0,176 y 0,200. En el modelo G4, el locus Bmac040 (6H) fue el único que presentó una asociación positiva, con un coeficiente R^2 de 0,289.

La comparación de los promedios de los coeficientes de determinación de los 3 modelos no reveló ninguna diferencia significativa (Figura 6.2).

Espiguillas por espiga

Un único locus presentó una asociación con este carácter en el modelo SG. Al incluir la estructura de población en los análisis, no se encontró ninguna (Tabla 6.3). Sin embargo, en el modelo G4, el locus Bmac018 (6H) mostró una asociación con este carácter, con un coeficiente R^2 de 0,184.

La comparación de los promedios de los coeficientes de determinación de los 3 modelos no reveló ninguna diferencia significativa (Figura 6.3).

Tipo de espiga

Los análisis de regresión llevados a cabo sin incluir la estructura de población revelaron la asociación de 32 loci con este carácter. Este número se redujo a 8 con la introducción de los grupos en los análisis. Estos loci fueron Bmac032 (1H), scssr10477 (1H), HvM036 (2H), MWG699 (2H), HvM054 (2H), Hv13GEIII (3H), HvZCCT (4H) y HvBM5A (5H), como se muestra en la Tabla 6.2. Los loci HvM036 y MWG699 son los que presentaron coeficientes de determinación relativamente más altos, de 0,060 y

0,052, respectivamente. No se realizó el análisis del modelo G4 pues la mayoría de las entradas de este grupo son de 6 carreras, sólo hay una de 2 carreras.

Los promedios de los coeficientes de determinación de los dos modelos se diferencian significativamente (Figura 6.2).

6-3-2. Análisis multilocus

Los loci que presentaron asociaciones significativas para cada carácter en el modelo CG fueron analizados en modelos multilocus, para examinar si sus efectos son aditivos. Este análisis permitió determinar la mejor combinación de loci que explica cada carácter, basándose sobre los criterios de información o de ajuste del *proc mixed* de SAS (SAS, 1988). Los resultados se resumen en las Tablas 6.5 y 6.6.

6-3-2-1. Análisis multilocus con caracteres agronómicos

Rendimiento

El análisis multilocus realizado con diferentes combinaciones de los nueve loci que fueron significativos para este carácter en las 156 *landraces*, reveló que bastan cuatro para encontrar un modelo tan explicativo como el modelo completo con nueve. El modelo multilocus agrupa a los loci *scssr07759*, *EBmac701*, *HvBM5A* y *Bmag009*, que conjuntamente explican un 32,9% de la variación del rendimiento (Tabla 6.5).

Una separación de medias de los alelos de estos loci permitió ver cuáles son los posibles efectos sobre el rendimiento de cada alelo en particular. Para el locus *scssr07759*, se distinguió el alelo 2 (213 pb) asociado a una producción alta (Figura 6.4 a). En el locus *EBmac701*, los alelos 1 (133pb) y 5 (149 pb) son los que presentaron un promedio de producción alto (Figura 6.6a). Para los loci *HvBM5A* y *Bmag009*, solamente se destacó de los demás un único alelo, que presentó un promedio de producción bajo. Fueron el alelo nulo para *HvBM5A* (Figura 6.7a) y el alelo de 176 pb para *Bmag009* (Figura 6.8a).

En el grupo 4, los dos loci *scssr07759* y *Bmac018* explicaron el 31,5% de la variación del rendimiento. Para el locus *scssr07759*, se observó una asociación con alta producción para el alelo de 213 pb (Figura 6.4b). Sin embargo en el locus *Bmac018* se distinguieron 2 alelos distintos (138 y 144 pb) asociados con promedios de rendimientos más elevados (Figura 6.8b).

Espigado

De los 12 loci que mostraban independientemente asociación con el espigado, 9 entraron en el modelo más explicativo. Estos fueron Bmac032, Bmac134, scssr07759, HvLTPPB, HvM062, HvBAMY, Bmag223, Bmac018 y Bmag369 (Tabla 6.5). Este modelo explicó el 57,6% de la variación en fecha de espigado. Los alelos de cada locus relacionados con el espigado temprano fueron el alelo de 235 pb de Bmac032 (Figura 6.3a); el alelo de 213 pb de scssr07759 (Figura 6.4 a); los alelos de 156 y 158 pb de Bmac134 (Figura 6.4a); el alelo de 221 pb de HvLTPPB (Figura 6.5a); el alelo de 247 pb de HvM062 (Figura 6.5a); los alelos de 129 y 131 pb de HvBAMY (Figura 6.6a); el alelo de 125 pb de Bmag223 (Figura 6.7a); los alelos de 138, 142 y 144 pb de Bmac018 (Figura 6.8a) y el alelo de 189 pb de Bmag369 (Figura 6.9a).

En el grupo 4, el modelo multilocus más explicativo agrupó sólo 5 loci de los 16 que mostraron una asociación con el carácter espigado de manera independiente. Estos fueron HvGLUEND, scssr07759, HvLTPPB, EBmac701 y Bmac018. Este modelo explicó el 75,7% de la variación de la fecha de espigado.

Los alelos asociados con el espigado precoz fueron los de 260 y 262 pb de HvGLUEND (Figura 6.3b); el alelo de 213 pb de scssr07759 (Figura 6.4b); el de 221 pb de HvLTPPB (Figura 6.5b); los de 133, 135 y 137 pb de EBmac701 (Figura 6.6b) y los de 138 y 144 pb de Bmac018 (Figura 6.8b).

El alelo de 142 pb del locus Bmac018, que se clasificó como temprano al considerar las 156 *landraces*, se encuentra clasificado como tardío al considerar solamente las entradas del grupo 4 (Figuras 6.8a, 6.8b).

Encamado

Los dos loci asociados con encamado en las 156 *landraces*, Bmac032 y Bmag006, explicaron en el modelo multilocus un 26,4% de la variación de este carácter. Los alelos asociados con un encamado acentuado son los de 233 y 235 pb de Bmac032 (Figura 6.3a) y los de 170, 182 y 196 pb de Bmag006 (Figura 6.5a).

En el grupo 4, donde se detectó únicamente el locus Bmac064 asociado con este carácter, el alelo de 157 pb, es el que menos encamado presenta (Figura 6.9b).

Peso hectolítrico

El locus EBmac701 fue el único que mostró una asociación con este carácter, en el grupo 4, explicando el 32,1% de su variación. Las entradas que tienen el alelo de 147 pb presentaron el peso hectolítrico más alto, contrariamente a las que tienen el alelo de 133 pb (Figura 6.6b).

Peso de mil granos

Un 15,4% de la variación del peso de mil granos fue explicada por el modelo multilocus con MWG699 y HvBM5A, en las 156 *landraces*. En el locus MWG699, el peso de mil granos más bajo fue observado para el tipo D (Figura 6.4a), mientras que para el locus HvBM5A, el alelo nulo y el alelo de 150 pb presentaron los promedios más bajos (Figura 6.7a).

En el grupo 4, el análisis multilocus de estos loci anuló la significación de HvBM5A. La separación de medias del locus MWG699 confirmó el bajo peso de mil granos de las entradas de tipo D (Figura 6.4b).

Granos m⁻²

Solamente el locus HdAMYB reveló una asociación con el número de granos m⁻², en la población 4. Este marcador explica el 22,9% de la variación de este carácter. El número más alto de granos m⁻² corresponde al alelo de 186 pb, seguido por el de 212 pb (Figura 6.6b).

6-3-2-2. Análisis multilocus de tolerancia a enfermedades

El análisis multilocus para la tolerancia a la roya en las 156 *landraces*, permitió ver que la combinación de los loci Bmac399, Bmag337 y HvBM5A era la más explicativa. Este modelo explicó un 30,7% de la variación de la incidencia de esta enfermedad. Los alelos menos susceptibles de cada locus fueron los de 131 y 137 pb para Bmac399 (Figura 6.3b), el de 143 pb para Bmag337 y el alelo de 1200 pb para el locus HvBM5A (Figura 6.7b).

En el grupo 4, el modelo multilocus más explicativo agrupó a los loci Bmac399, MWG699 y HvBTAI3, y explicó el 47,1% de la variación de la incidencia de la roya en esta población. Los alelos menos susceptibles para cada locus en esta población fueron

los de 131 y 137 pb para Bmac399 (Figura 6.3b), el tipo A de MWG699 (Figura 6.4b) y el alelo de 228 pb de HvBTAI3 (Figura 6.6b).

6-3-2-3. Análisis multilocus con las variables de calidad maltera

Calibre inferior a 2,2 mm

Los dos loci Bmac134 y Bmag125 que mostraron asociación con este carácter, juntos explican un 27,1% de la variación en el análisis multilocus. El alelo de 138 pb de Bmac134 corresponde al promedio de porcentaje más alto de granos que pasan la criba de 2,2 mm. Para el locus Bmag125, los alelos de 134 y 136 pb son los que presentan un porcentaje alto (Figura 6.4a).

En el grupo 4, el modelo más explicativo agrupó a los loci scssr07759 y Bmag225; los dos explican un 40,0% de la variación de la variable, y corresponden a regiones cromosómicas distintas de las detectadas en el modelo CG. Los alelos de 195 pb de scssr07759 (Figura 6.4b) y de 140 pb de Bmag225 son los que presentan un porcentaje alto de granos que pasan la criba de 2,2 mm (Figura 6.5b).

Proteína

El locus Bmac067 fue el único marcador que reveló una asociación con el porcentaje de proteína total en harina de malta en la población 4, explicando un 23,2% de la variación de este carácter. La separación de medias mostró que el menor contenido en proteína está relacionado con el alelo de 179 pb (Figura 6.5b).

6-3-2-4. Análisis multilocus con las variables morfológicas

Pigmentación del tallo

El modelo más explicativo del carácter pigmentación del tallo, en las 156 *landraces* combina los loci Bmag013 y HdAMYB, descartando MWG699. Este modelo explicó el 27,7% de la variación del carácter. Los alelos de 149, 153, 155 y 159 pb del locus Bmag013 y los de 190, 192 y 211 pb de HdAMYB parecen ser los relacionados con la presencia de antocianina (Figura 6.5a, 6.6a).

Pigmentación de la aurícula

Las diferentes combinaciones entre los 4 loci que mostraron asociaciones con la pigmentación de la aurícula en el grupo 4 permitieron ver que el modelo más explicativo es el que agrupa a los 3 loci HvBKASI, Hv13GEIII y Bmag223. Este modelo explicó el 50,1% de la variación del carácter. Los alelos que parecen estar relacionados con la presencia de antocianina para cada locus son los de 198 y 200 pb para HvBKASI (Figura 6.4b), el de 175 pb para Hv13GEIII (Figura 6.5b) y los de 155, 157 y 159 pb para Bmag223 (Figura 6.7b).

Vellosidad de la vaina basal

De los 4 loci que mostraron asociaciones con la vellosidad de la vaina basal, 3 entraron en el modelo multilocus más explicativo. Estos fueron Bmac399, Bmag337 y EBmac806. Los 3 explican combinadamente un 28,2% de la variación. Los alelos que parecen estar relacionados con la ausencia de vello son los de 109, 127 y 151 pb de Bmac399 (Figura 6.3a), el de 147 pb de Bmag337 (Figura 6.7a) y el de 172 pb de EBmac806 (Figura 6.8a).

Glumas y aristas de la gluma

El modelo multilocus que agrupa los loci MWG699, Bmag500 y Bmag369, resultó ser el más explicativo de este carácter, con un coeficiente de determinación de 22,3%, descartándose Bmac399. La separación de medias de los alelos permitió ver que las entradas que presentan los tipos D y K de MWG699 tienden a tener las glumas y aristas de la gluma más cortas que el grano (Figura 6.4a). Lo mismo ocurre con las entradas que presentan el alelo de 164 pb de Bmag500 (Figura 6.8a) y los alelos de 187 y 189 pb de Bmag369 (Figura 6.9a).

Tipo de lemma

El modelo multilocus más explicativo de este carácter agrupó a los loci scssr07759 y HvBM5A, quedando fuera del modelo HvBAMY. Estos dos loci explican juntos un 22,6% de la variación observada en el tipo de lemma. El alelo de 195 pb de scssr07759 y los alelos nulo y de 150 pb de HvBM5A parecen ser los relacionados con el tipo de lemma dentada (Figuras 6.4a, 6.7a).

Longitud de los pelos de la raquilla

Cuatro loci entraron en el modelo multilocus más explicativo para este carácter, en las 156 *landraces*. Estos fueron MWG699, HvM054, Bmac113 y Bmag223. Este modelo explicó un 59,3% de la variación observada en la longitud de los pelos de la raquilla, aunque la mayor parte de la variación está explicada por el locus Bmag223 (46,6%). La separación de medias correspondientes a los diferentes alelos permitió distinguir que el alelo de 157 pb de HvM054, el de 191 pb de Bmac113 y los alelos de 155, 157 y 159 pb de Bmag223 corresponden a entradas que tienden a tener los pelos de la raquilla cortos (Figuras 6.4a, 6.7a). Para el locus MWG699, las entradas de tipo K se diferencian de los otros dos, con la tendencia a tener unos pelos de raquilla largos (Figura 6.4a).

En el grupo 4, este carácter está más explicado por el modelo que combina los loci Bmag223 y HvM054. Este modelo explica el 73,6% de la variación en la longitud de los pelos de la raquilla en estas entradas. El locus Bmag223 explicó, él solo, el 67,1% de esta variación. Los alelos de 155, 157 y 159 pb de Bmag223 y el alelo de 157 pb de HvM054 son los que corresponden a entradas con pelos de la raquilla cortos (Figuras 6.4b, 6.7b).

Longitud de espiga

Los 5 loci (Bmac399, EBmac701, Bmac040, HvM004 y Bmag120) que mostraron asociación con el carácter longitud de espiga en las 156 *landraces*, entraron en el análisis multilocus, explicando conjuntamente el 34,5% de la variación de este carácter. El alelo de 137 pb de Bmac339, los de 135 y 149 pb de EBmac701, los de 176, 218, 226 y 232 pb de Bmac040, así como el alelo de 238 pb de Bmag120, parecen ser los más relacionados con una longitud de espiga corta (Figuras 6.3a, 6.6a, 6.8a, 6.9a).

En el grupo 4, solamente 2 loci (HvGLUEND y Bmac040) entraron en el modelo más explicativo de este carácter, quedando fuera HvM036. Explicaron un 46,8% de la variación en la longitud de espiga en esta población. Los alelos de 176 y 232 pb de Bmac040 y el alelo de 264 pb de HvGLUEND son los que más se destacan en esta población, con una longitud de espiga relativamente más corta (Figuras 6.3b, 6.8b).

Densidad de espiga

Los loci Bmac399 y Bmag120 explicaron juntos en el modelo de análisis multilocus un 27% de la variación de la densidad de espiga en las 156 *landraces*. Los alelos de 137 y 141 pb de Bmac399 así como los alelos de 238 y 244 pb de Bmag120 parecen tener espigas relativamente más densas (Figuras 6.3a, 6.9a).

En el grupo 4, el locus Bmac040 fue el único que presentó una asociación con este carácter, explicando el 28,9% de su variación. Las entradas que tienen el alelo de 176 pb se caracterizaron por unas espigas relativamente más densas (Figura 6.8 b).

Espiguillas / espiga

El locus Bmac018 fue el único que mostró una asociación con este carácter en la población 4, explicando el 18,4% de su variación. El alelo de 144 pb es el más destacable, con un número de espiguillas por espiga más alto (Figura 6.8b).

Tipo de espiga.

De los 8 loci que presentaron una asociación con este carácter, 7 (scsr10477, HvM036, MWG699, HvM054, Hv13GEIII, HvZCCT y HvBM5A) entraron en el modelo multilocus más explicativo, explicando conjuntamente sólo el 14,5% de la variación de este carácter.

Tabla 6.5. Coeficientes de determinación mostrados por los diferentes loci que revelaron asociaciones positivas con las variables analizadas en el modelo con grupos, en los análisis locus por locus y en el modelo multilocus.

Variables	Modelo	Coeficiente de determinación R^2	
		Grupos + locus	Locus
Rendimiento	Grupos	0,276	
	WMC1E8	0,306	0,030
	Bmac134	0,405	0,129
	scssr07759	0,324	0,048
	Bmag353	0,389	0,113
	EBmac701	0,410	0,135
	HvBAMY	0,329	0,053
	HvBM5A	0,409	0,133
	Bmag009	0,337	0,061
	Multilocus (scssr07759 + EBmac701 + HvBM5A + Bmag009)		0,329
Espigado	Grupos	0,193	
	Bmac032	0,326	0,133
	HvM020	0,275	0,082
	Bmac134	0,351	0,157
	MWG858	0,263	0,070
	scssr07759	0,353	0,160
	EBmac640	0,272	0,079
	HvLTPPB	0,346	0,153
	HvM062	0,321	0,128
	HvBAMY	0,272	0,079
	Bmag223	0,333	0,140
	Bmac018	0,309	0,116
	Bmag369	0,327	0,134
Multilocus (Bmac032 + Bmac134 + scssr07759 + HvLTPPB + HvM062 + HvBAMY + Bmag223 + Bmac018 + Bmag369)		0,576	
Encamado	Grupos	0,192	
	Bmac032	0,318	0,126
	Bmag006	0,351	0,159
	Multilocus (Bmac032 + Bmag006)		0,264
Peso mil granos	Grupos	0,352	
	MWG699	0,456	0,104
	HvBM5A	0,458	0,106
	Multilocus (MWG699 + HvBM5A)		0,154
Roya	Grupos	0,205	
	Bmac399	0,370	0,165
	HvBTAI3	0,290	0,085
	Bmag337	0,320	0,115
	HvBM5A	0,375	0,170
	Multilocus (Bmac399 + Bmag337 + HvBM5A)		0,307
Calibre <2,2 mm	Grupos	0,168	
	Bmac134	0,350	0,182
	Bmag125	0,293	0,125
	Multilocus (Bmac134 + Bmag125)		0,271

Tabla 6.5. (continuación)

Variables	Modelo	Coeficiente de determinación R^2	
		Grupos + locus	Locus
Pigmentación del tallo	Grupos	0,068	
	MWG699	0,136	0,068
	Bmag013	0,256	0,188
	HdAMYB	0,227	0,160
	Multilocus (Bmag013 + HdAMYB)		0,277
Vellosidad de la vaina basal	Grupos	0,184	
	Bmag399	0,307	0,123
	Bmag337	0,291	0,107
	Bmag500	0,307	0,122
	EBmac806	0,290	0,106
Multilocus (Bmac399 + Bmag337 + EBmac806)		0,282	
Glumas y aristas de la gluma	Grupos	0,169	
	Bmac399	0,290	0,120
	MWG699	0,261	0,092
	Bmag500	0,291	0,122
	Bmag369	0,244	0,075
Multilocus (MWG699 + Bmag500 + Bmag369)		0,223	
Tipo de lemma	Grupos	0,195	
	scssr07759	0,320	0,125
	HvBAMY	0,290	0,096
	HvBM5A	0,339	0,144
	Multilocus (scssr07759 + HvBM5A)		0,226
Longitud de los pelos de la raquilla	Grupos	0,138	
	MWG699	0,224	0,086
	HvM054	0,341	0,203
	Bmac067	0,298	0,161
	Bmac113	0,274	0,136
	Bmag223	0,605	0,466
	Multilocus (MWG699 + HvM054 + Bmac113 + Bmag223)		0,593
Longitud de espiga	Grupos	0,392	
	Bmac399	0,536	0,143
	EBmac701	0,530	0,137
	Bmac040	0,527	0,134
	HvM004	0,519	0,126
	Bmag120	0,540	0,148
	Multilocus (Bmac399 + EBmac701 + Bmac040 + HvM004 + Bmag120)		0,345
Densidad de espiga	Grupos	0,055	
	Bmac399	0,232	0,176
	Bmag120	0,255	0,200
	Multilocus (Bmac399 + Bmag120)		0,270
Tipo de espiga	Grupos	0,808	
	Bmac032	0,853	0,045
	scssr10477	0,821	0,013
	HvM036	0,868	0,060
	MWG699	0,860	0,052
	HvM054	0,830	0,023
	Hv13GEIII	0,821	0,013
	HvZCCT	0,824	0,017
	HvBM5A	0,844	0,036
	Multilocus (scssr10477 + HvM036 + MWG699 + HvM054 + Hv13GEIII + HvZCCT + HvBM5A)		0,145

Tabla 6.6. Coeficientes de determinación mostrados por los diferentes loci que revelaron asociaciones positivas con las variables analizadas en el modelo G4, en los análisis locus por locus y en el modelo multilocus.

Variables	Locus	Coefficiente de determinación R^2
Rendimiento	scssr07759	0,233
	Bmac018	0,190
	Multilocus (scssr07759 + Bmac018)	0,315
Espigado	HvGLUEND	0,272
	Bmac134	0,319
	MWG858	0,124
	scssr07759	0,287
	EBmac640	0,195
	MWG699	0,148
	HvLTPPB	0,253
	Hv13GEIII	0,210
	HvM062	0,235
	Bmac029	0,247
	EBmac701	0,335
	HvBAMY	0,217
	Bmag337	0,203
	Bmag223	0,254
	Bmac018	0,248
Bmag369	0,201	
Multilocus (HvGLUEND + scssr07759 + HvLTPPB + EBmac701 + Bmac018)	0,755	
Encamado	Bmac064	0,193
Peso de mil granos	MWG699	0,172
	HvBM5A	0,187
	Multilocus (MWG699)	0,172
Peso hectolítrico	EBmac701	0,321
Granos m⁻²	HdAMYB	0,229
Roya	Bmac399	0,361
	MWG699	0,128
	HvBTAI3	0,160
	HvBM5A	0,200
	Multilocus (Bmac399 + MWG699 + HvBTAI3)	0,471
Calibre < 2,2 mm	scssr07759	0,186
	Bmag125	0,255
	HvM054	0,189
	Bmag225	0,240
	Hv13GEIII	0,256
Multilocus (scssr07759 + Bmag225)	0,400	
Proteína	Bmac067	0,232
Pigmentación de la aurícula	HvBKASI	0,192
	Hv13GEIII	0,265
	Bmac029	0,187
	Bmag223	0,265
Multilocus (HvBKASI + Hv13GEII + Bmag223)	0,501	

Tabla 6.6. (continuación)

Variables	Locus	Coefficiente de determinación R^2
Longitud de los pelos de la raquilla	HvM020	0,325
	HvM054	0,307
	Bmag223	0,671
	Multilocus (Bmag223 + HvM054)	0,736
Longitud de espiga	HvGLUEND	0,185
	HvM036	0,187
	Bmac040	0,307
	Multilocus (HvGLUEND + Bmac040)	0,468
Densidad de espiga	Bmac040	0,289
Espiguillas / espiga	Bmac018	0,184

Los marcadores que se retuvieron en los modelos multilocus para todas las asociaciones significativas encontradas se han representado sobre un mapa revisado de la población Oregon Wolfe (Costa *et al.*, 2001), al que se ha añadido la información sobre la asignación de los marcadores a *bins*, según la distribución de estos últimos propuesta por Kleinhofs y Han (2002). Esta información está recogida en las Figuras 6.3 a 6.9. Cada figura se compone de tres partes. La columna vertebral de cada figura es un mapa de *bins* de cada cromosoma, con los marcadores que definen los *bins*, un esqueleto de marcadores de la población Oregon Wolfe, y los marcadores analizados en este estudio. En las dos primeras partes de cada figura (una para el modelo CG y otra para el modelo G4) se incluyen además en el lado derecho tablas con los alelos de los marcadores que han mostrado asociaciones significativas y que se han retenido en los análisis multilocus, el número de entradas en las que estaba presente cada alelo, y los valores de los caracteres fenotípicos correspondientes, con la separación de medias correspondiente al modelo multilocus. En la tercera parte de cada figura se presenta, en el lado izquierdo, una tabla con las asignaciones de posiciones de QTLs encontrados en otras poblaciones (Tabla 6.7), y de genes descritos (Tabla 6.8), solamente para los caracteres con asociaciones significativas en este estudio.

Figuras 6.3a - 6.9a. Representación de un mapa consenso de *bins* de los cromosomas 1H-7H. Los marcadores en negrita corresponden a los analizados en este estudio. Los loci subrayados marcan el inicio del *bin*. Las tablas corresponden a la separación de medias de los alelos de loci que mostraron asociaciones positivas con los caracteres analizados en el modelo multilocus de las 156 *landraces* con el modelo con grupos (CG).

Figuras 6.3b - 6.9b. Representación de un mapa consenso de *bins* de los cromosomas 1H-7H. Los marcadores en negrita corresponden a los analizados en este estudio. Los loci subrayados marcan el inicio del *bin*. Las tablas corresponden a la separación de medias de los alelos de loci que mostraron asociaciones positivas con los caracteres analizados en el modelo multilocus de las entradas del grupo 4 (G4).

Figuras 6.3c- 6.9c. Mapa consenso de *bins* de los cromosomas 1H-7H, Los marcadores en negrita corresponden a los analizados en este estudio. Los loci subrayados marcan el inicio del *bin*. Las tablas representan las asignaciones de posiciones de QTLs encontrados en otras poblaciones, solamente para los caracteres con asociaciones significativas en este estudio. Los números corresponden a las poblaciones donde se localizaron los QTLs, los números marcados en negrita representan el pico del QTL en dicha población, las poblaciones están descritas en la Tabla 6.7. Las letras corresponden a la referencia de descripción de los genes de los caracteres correspondientes, que se presentan en la Tabla 6.8

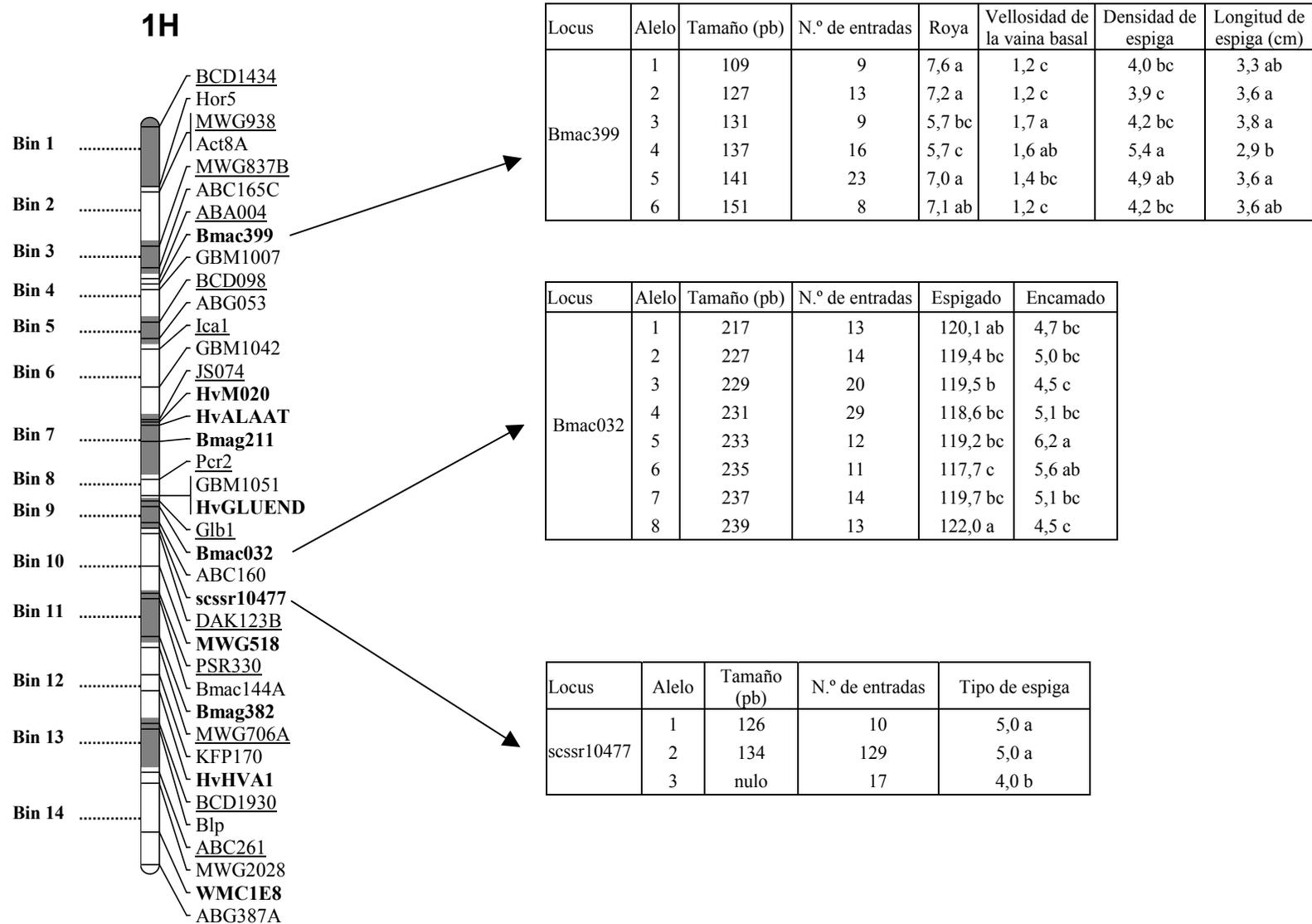


Figura 6.3a

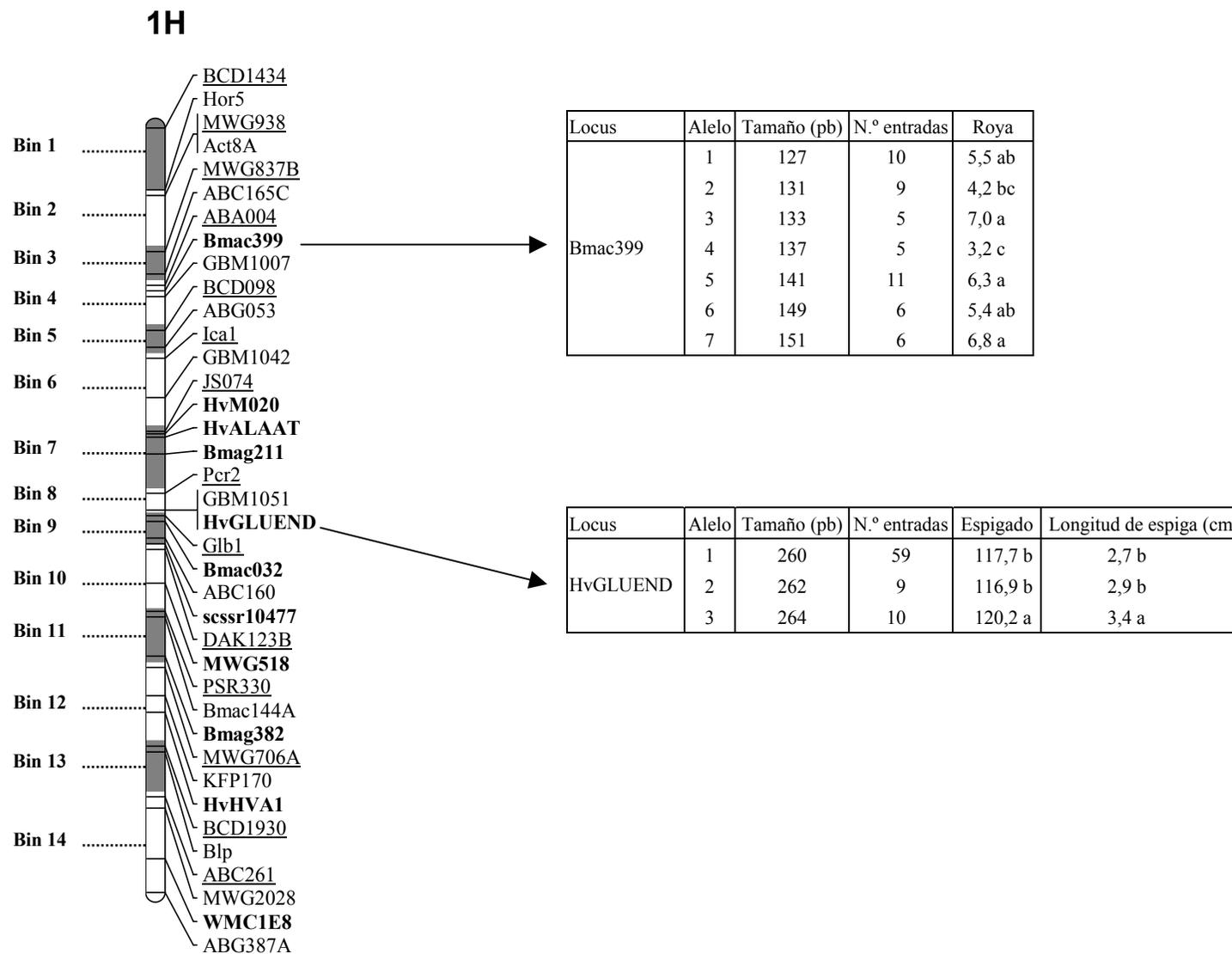


Figura 6.3b

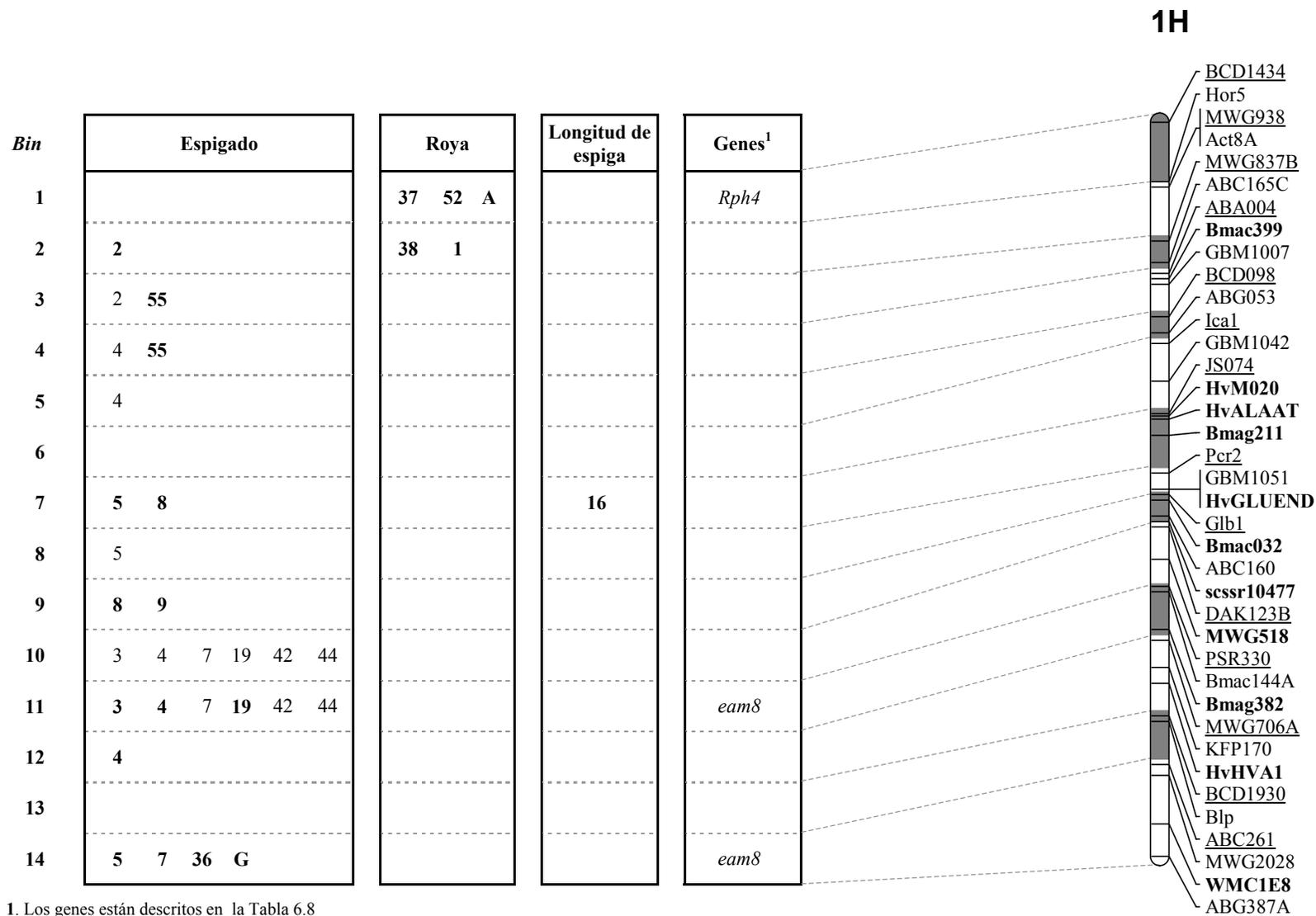


Figura 6.3c

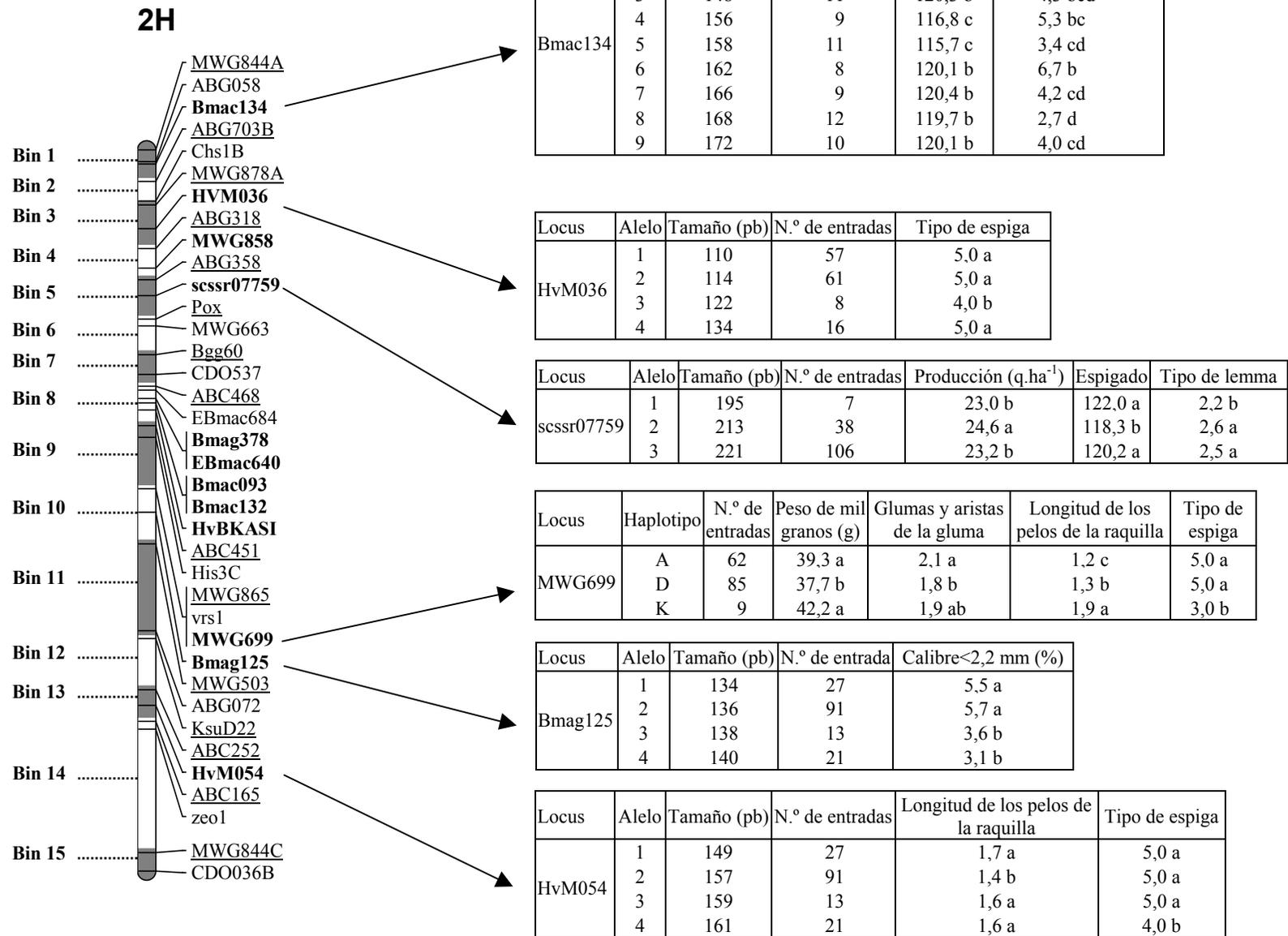


Figura 6.4a

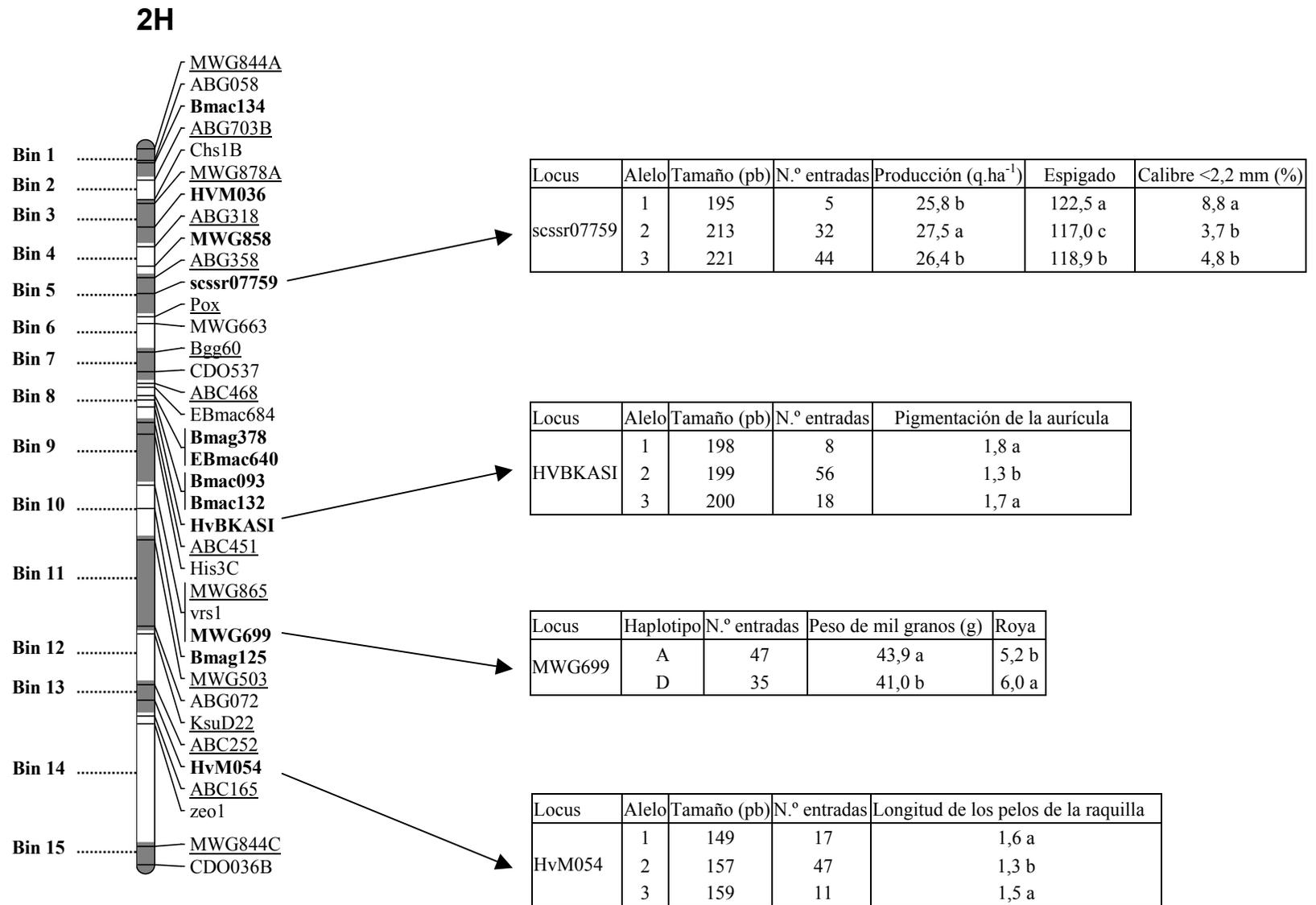


Figura 6.4b

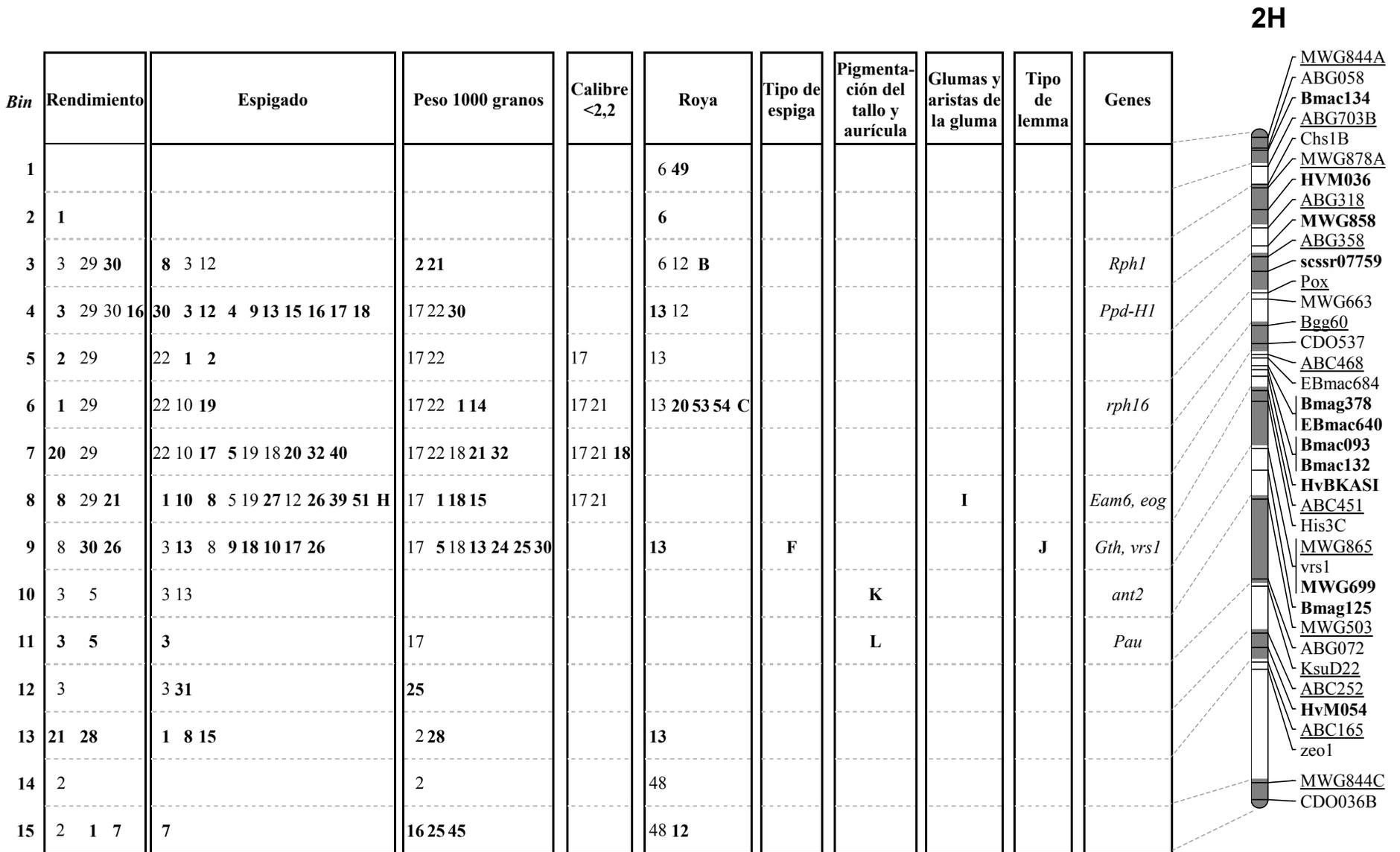


Figura 6.4c

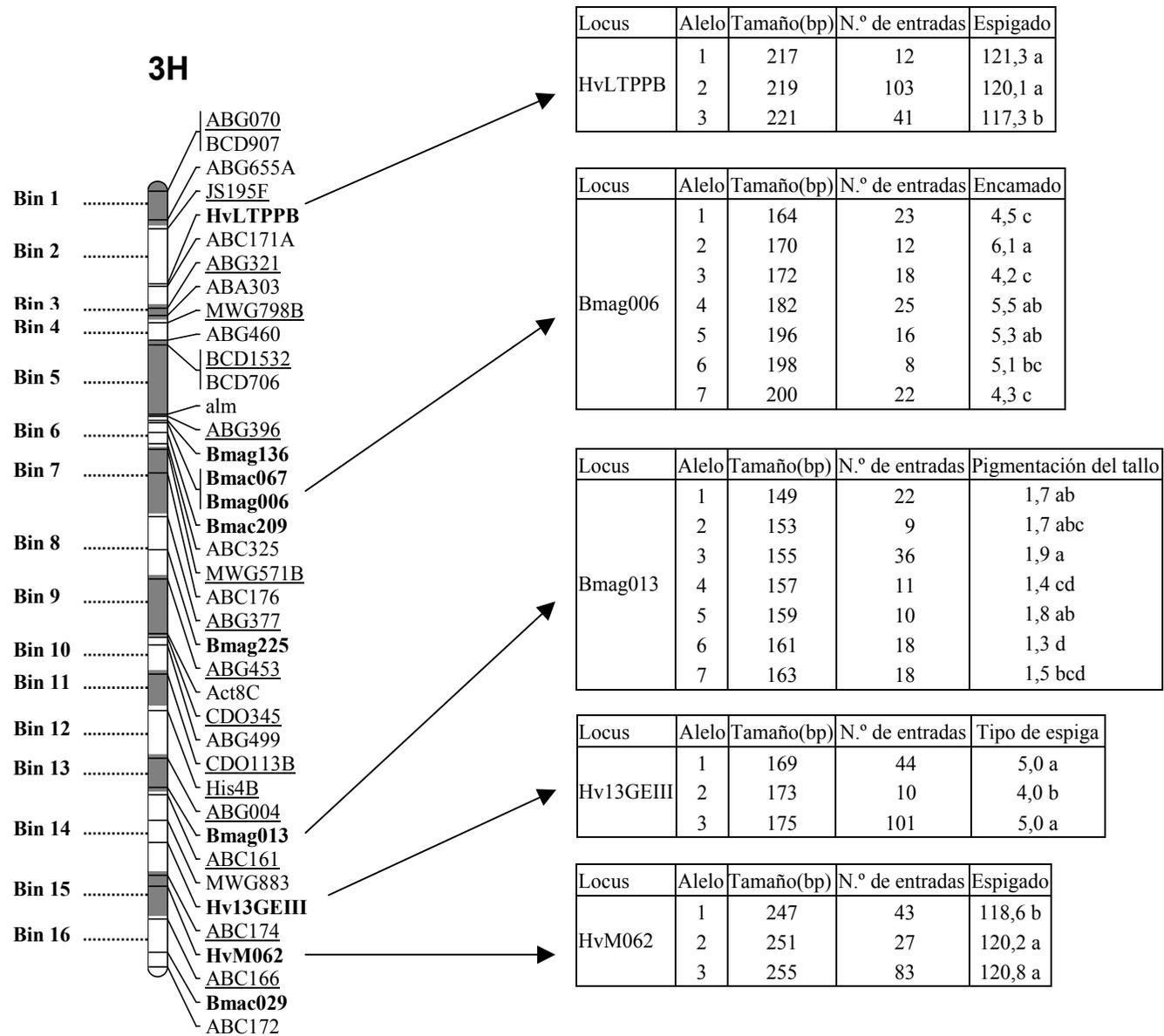


Figura 6.5a

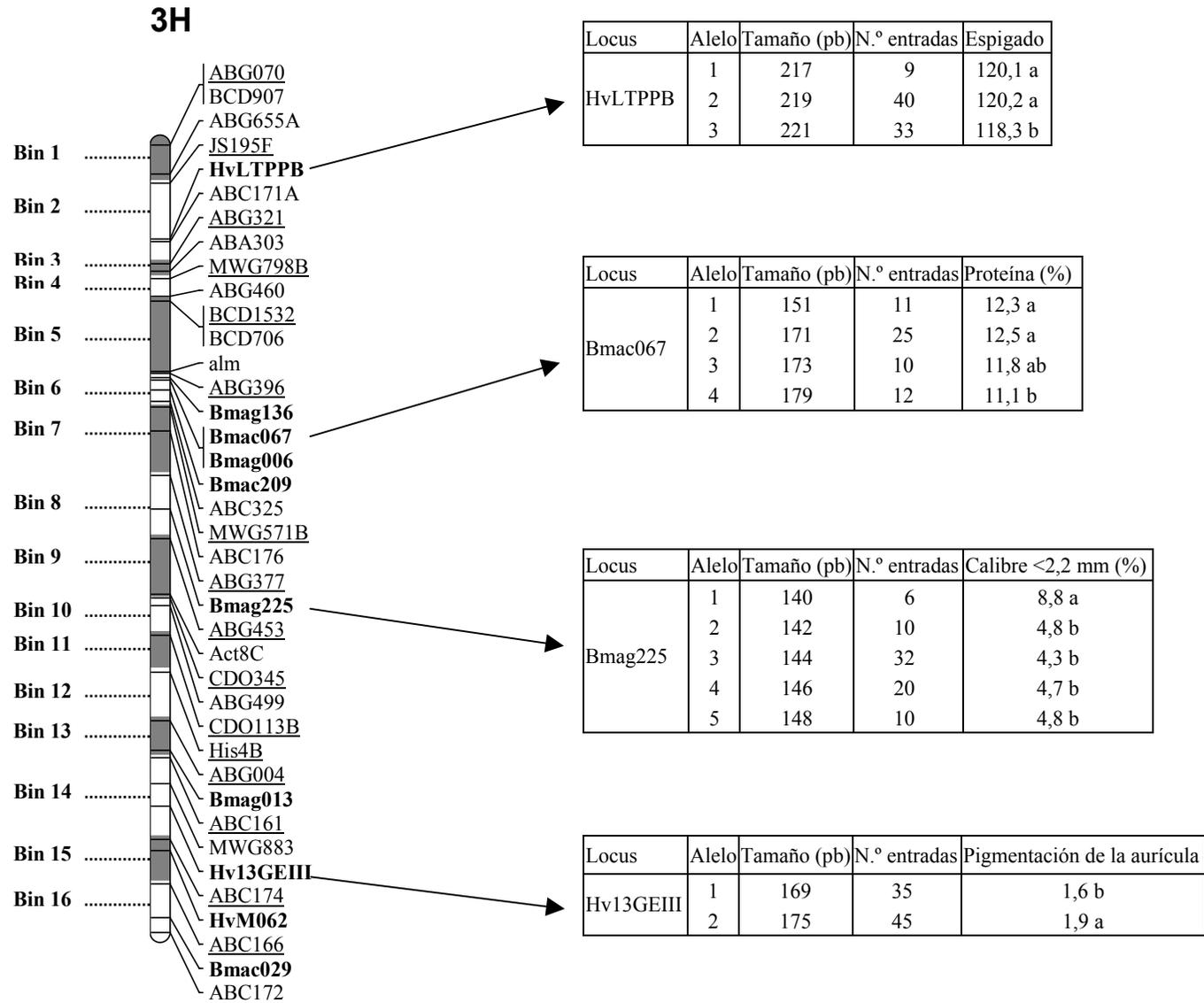


Figura 6.5b

3H

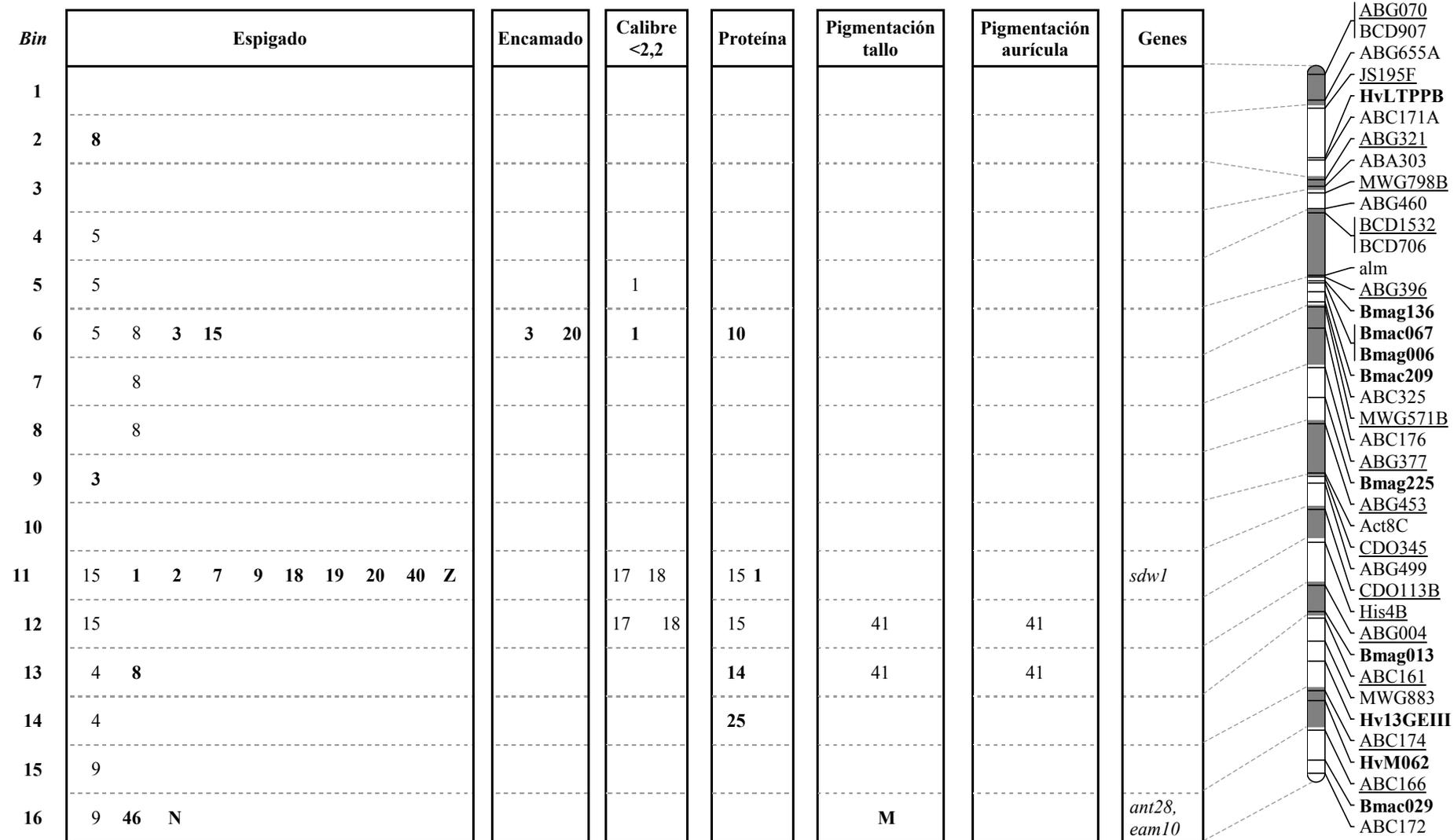


Figura 6.5c

4H

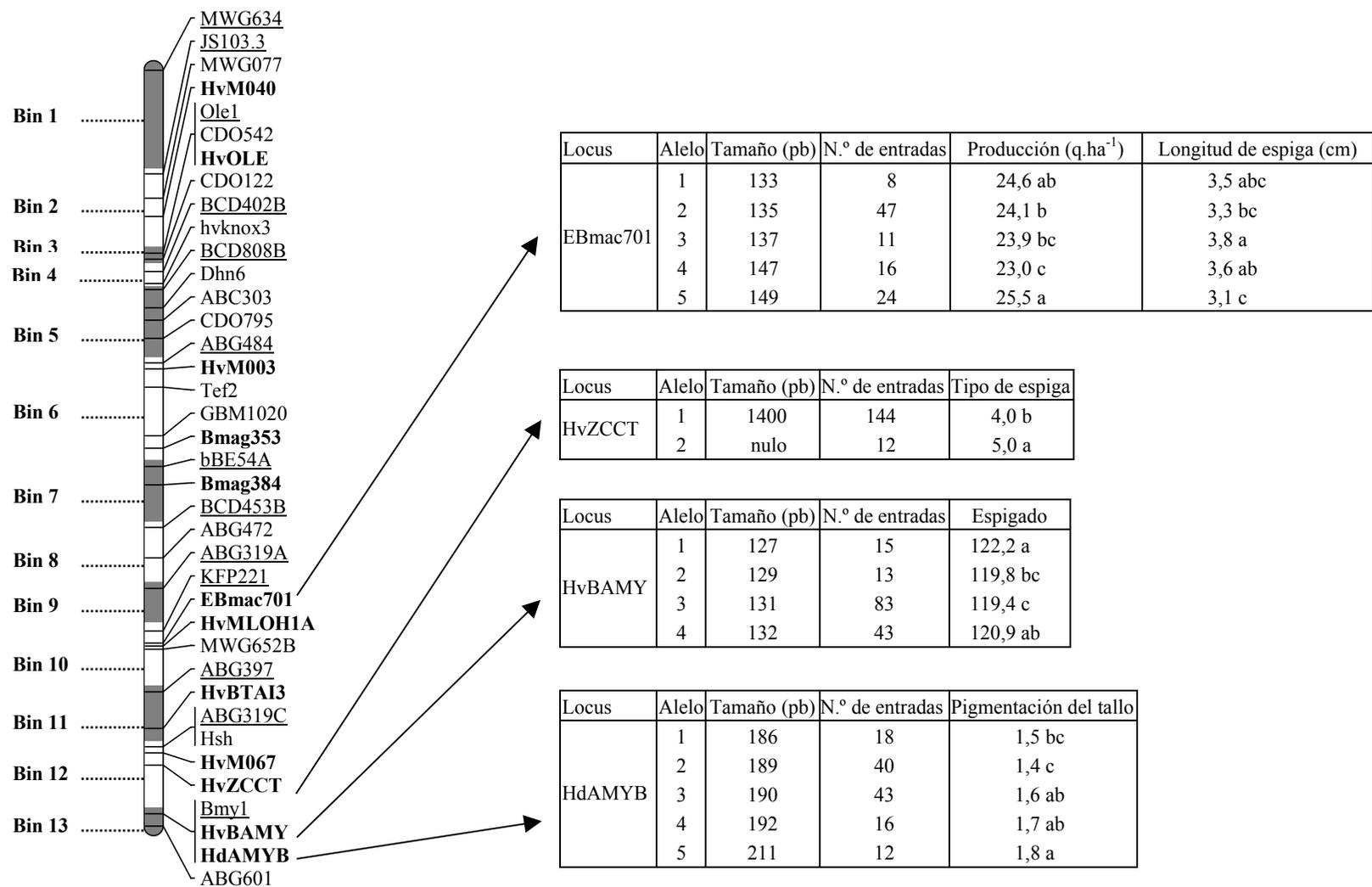
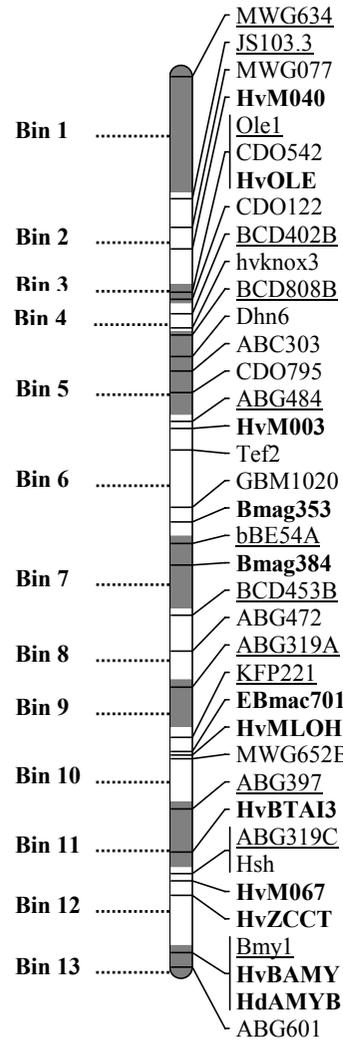


Figura 6.6a

4H



Locus	Alelo	Tamaño (pb)	N.º entradas	Espigado	Peso hectolítrico (kg hl ⁻¹)
EBmac701	1	115	6	121,5 a	63,9 b
	2	133	6	118,0 bc	61,8 c
	3	135	28	117,6 c	63,3 b
	4	137	7	117,7 bc	63,3 b
	5	147	7	120,1 ab	65,9 a
	6	149	6	122,5 a	63,5 b

Locus	Alelo	Tamaño (pb)	N.º entradas	Roya
HvBTAI3	1	228	17	5,0 b
	2	234	65	6,2 a

Locus	Alelo	Tamaño (pb)	N.º entradas	Granos m ⁻²
HdAMYB	1	186	11	9129 a
	2	189	25	6970 b
	3	190	17	6760 b
	4	192	9	6269 b
	5	212	5	7617 ab

Figura 6.6b

4H

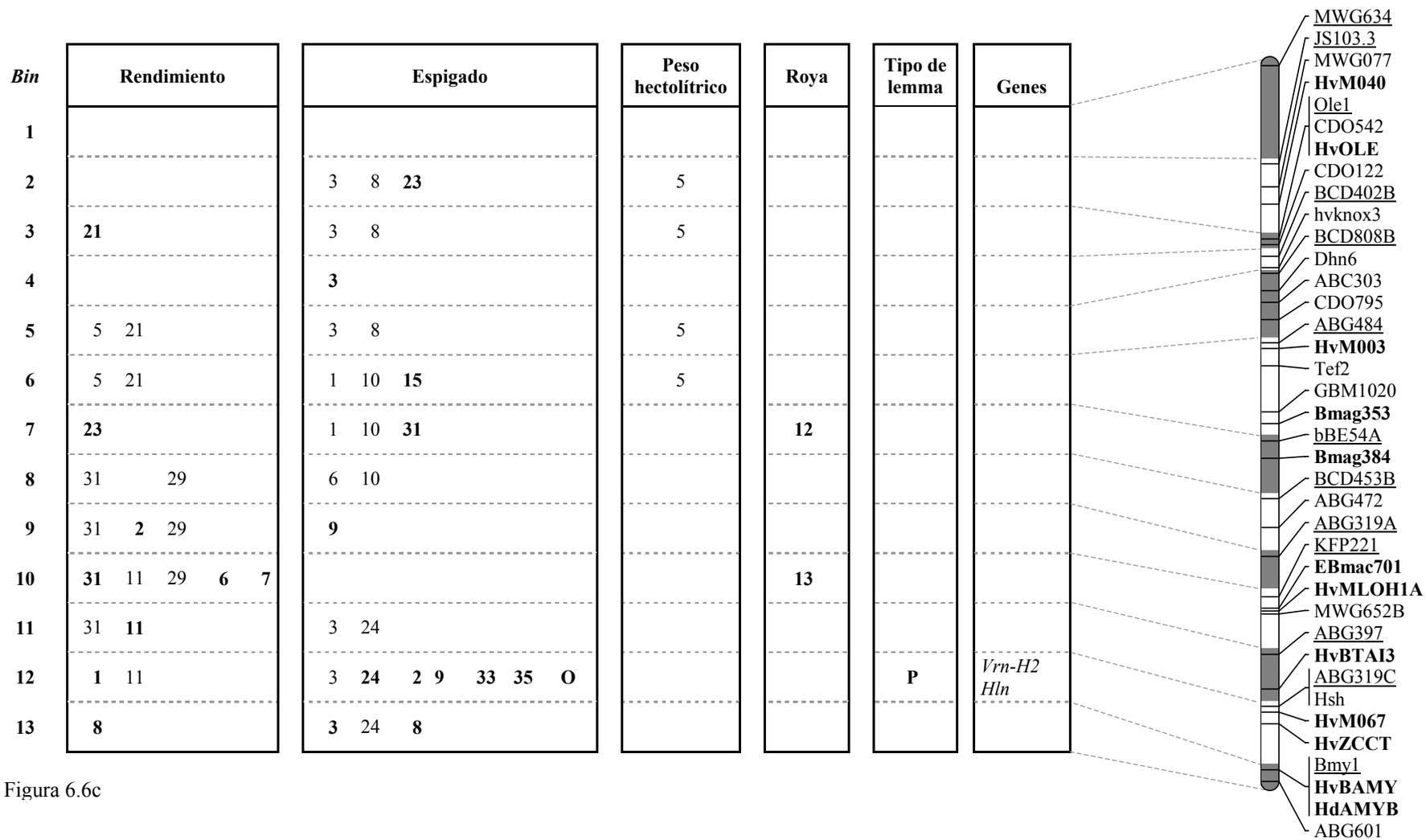
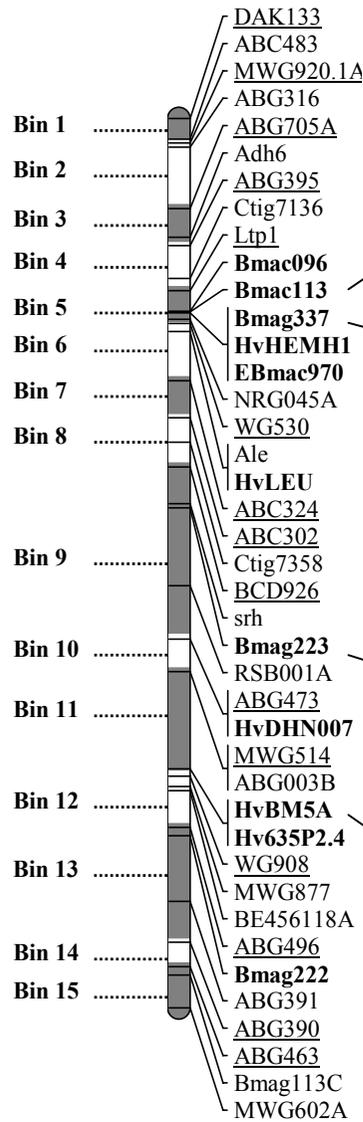


Figura 6.6c

5H



Locus	Alelo	Tamaño (pb)	N.º de entradas	Longitud de los pelos de la raquilla
Bmac113	1	181	65	1,7 a
	2	191	15	1,2 c
	3	193	22	1,5 b
	4	195	20	1,5 b
	5	197	10	1,5 ab

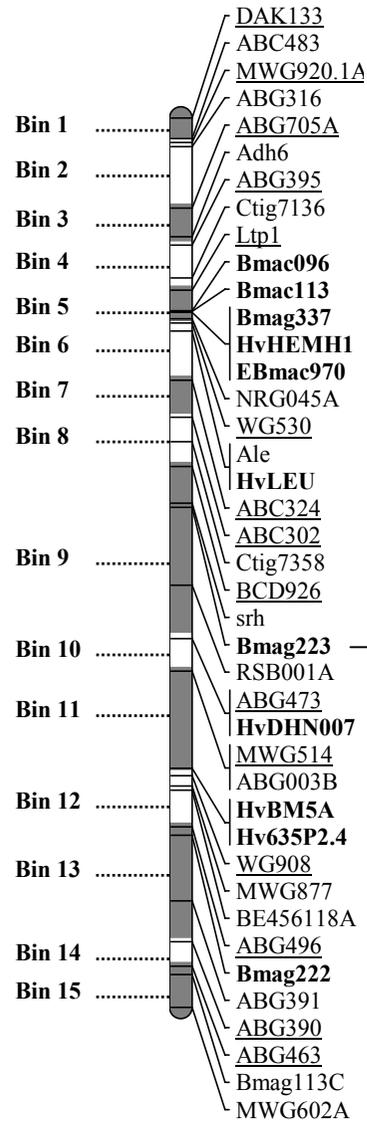
Locus	Alelo	Tamaño (pb)	N.º de entradas	Roya	Vellosidad de la vaina basal
Bmag337	1	127	19	7,5 ab	1,3 ab
	2	129	61	7,5 a	1,5 a
	3	143	13	5,7 c	1,4 a
	4	145	20	6,6 bc	1,4 a
	5	147	28	6,8 b	1,1 b
	6	149	8	6,4 abc	1,5 a

Locus	Alelo	Tamaño (pb)	N.º de entradas	Espigado	Longitud de los pelos de la raquilla
Bmag223	1	125	12	117,3 d	1,4 cd
	2	127	16	120,9 ab	1,9 a
	3	155	17	120,3 b	1,1 e
	4	157	35	119,4 bc	1,2 e
	5	159	10	119,6 abcd	1,2 de
	6	165	9	121,6 a	1,8 ab
	7	169	11	118,2 cd	1,7 bc

Locus	Alelo	Tamaño (pb)	N.º de entradas	Producción (q.ha ⁻¹)	Peso de mil granos (g)	Roya	Tipo de lemma	Tipo de espiga
HvBM5A	1	4850	92	25,9 a	41,4 a	7,2 a	2,7 a	5,0 a
	2	1100	50	25,7 a	42,6 a	5,7 b	2,6 a	5,0 a
	3	150	7	24,9 a	38,2 b	7,8 a	2,2 b	4,0 b
	4	nulo	7	20,2 b	36,8 b	6,5 ab	1,7 c	4,0 b

Figura 6.7a

5H



Locus	Alelo	Tamaño (pb)	N.º entradas	Pigmentación de la aurícula	Longitud de los pelos de la raquilla
Bmag223	1	125	12	1,4 b	1,6 b
	2	127	14	1,4 b	2,0 a
	3	155	9	1,8 a	1,1 c
	4	157	20	1,7 a	1,1 c
	5	159	8	1,9 a	1,1 c

Figura 6.7b

5H

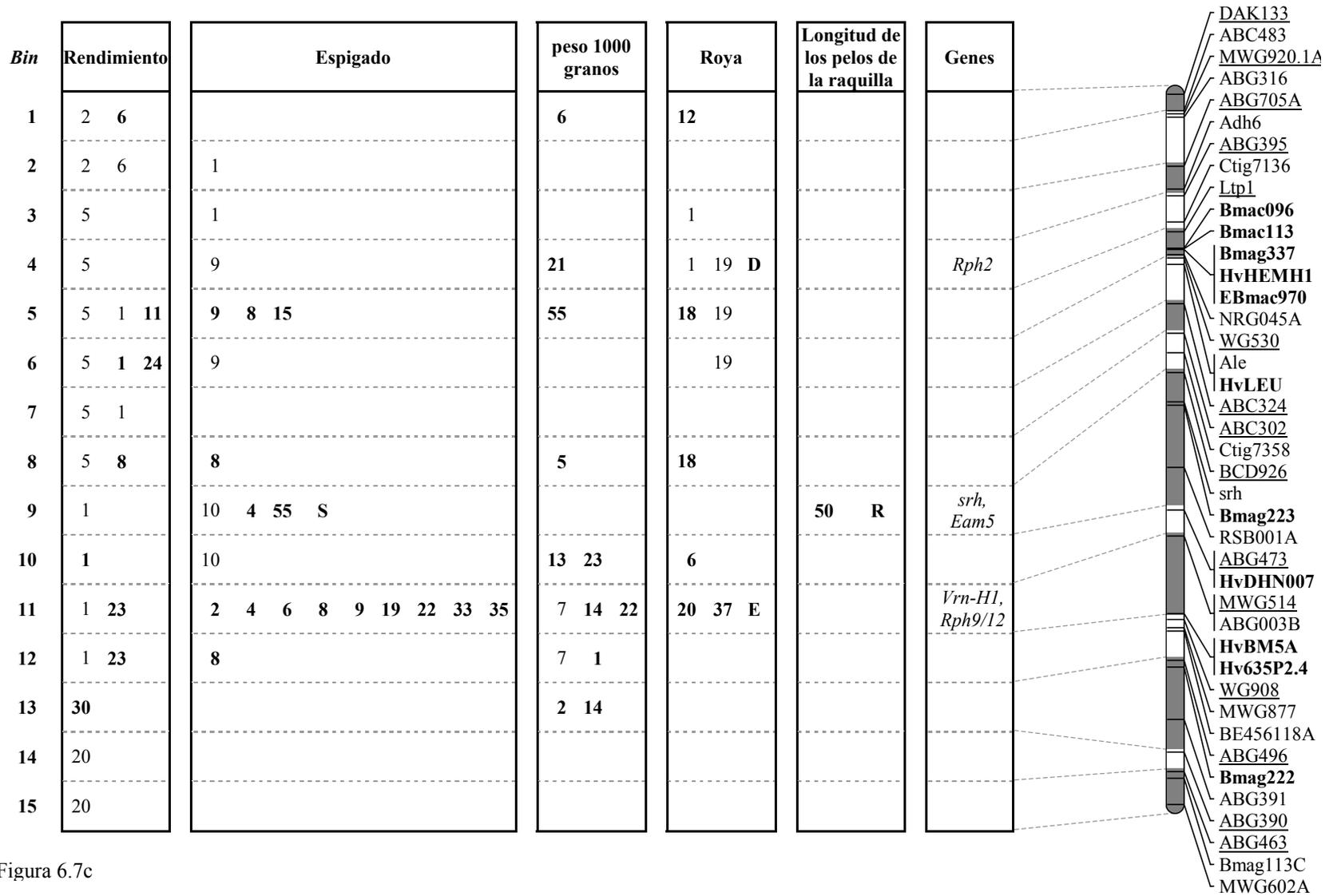


Figura 6.7c

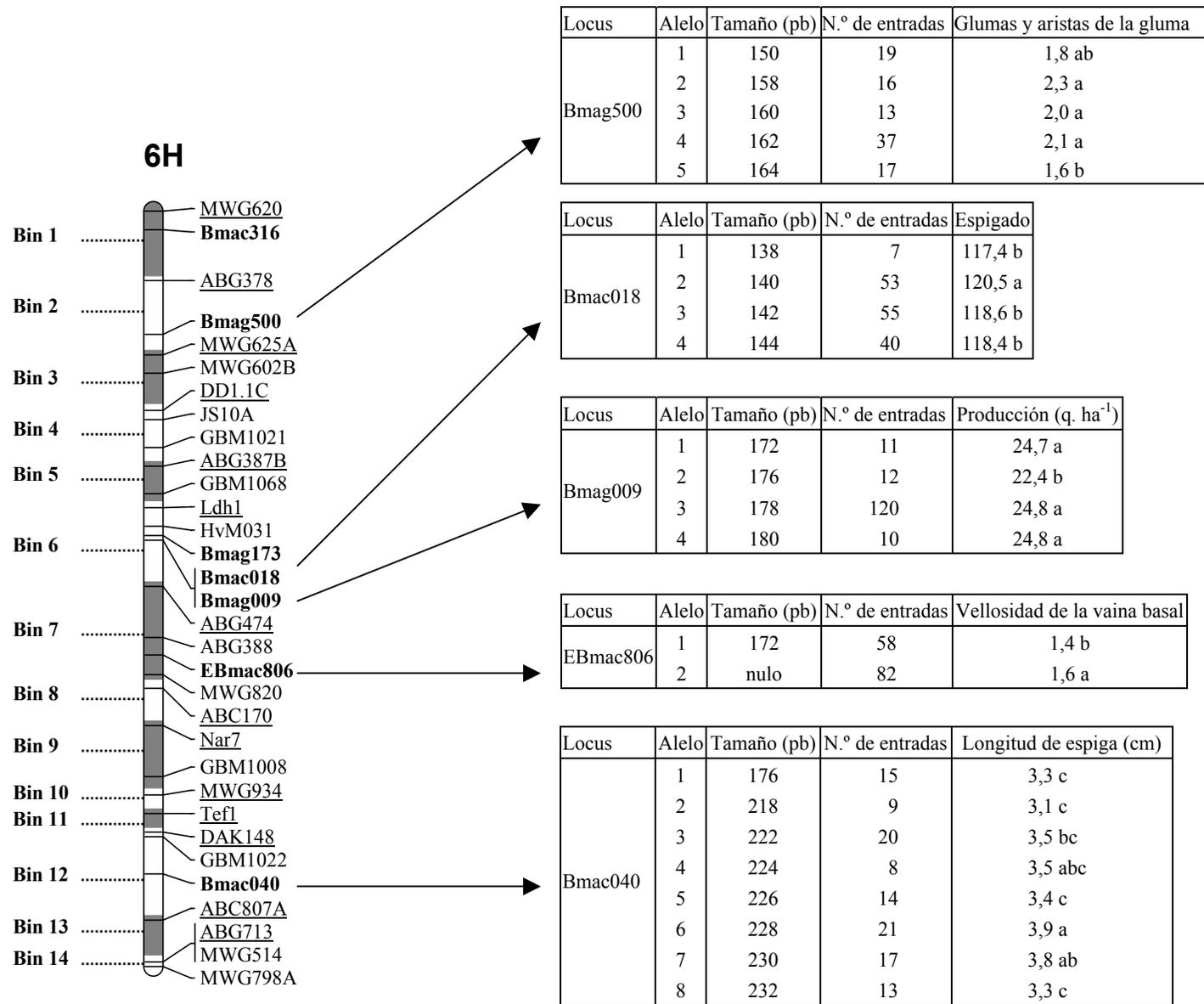


Figura 6.8a

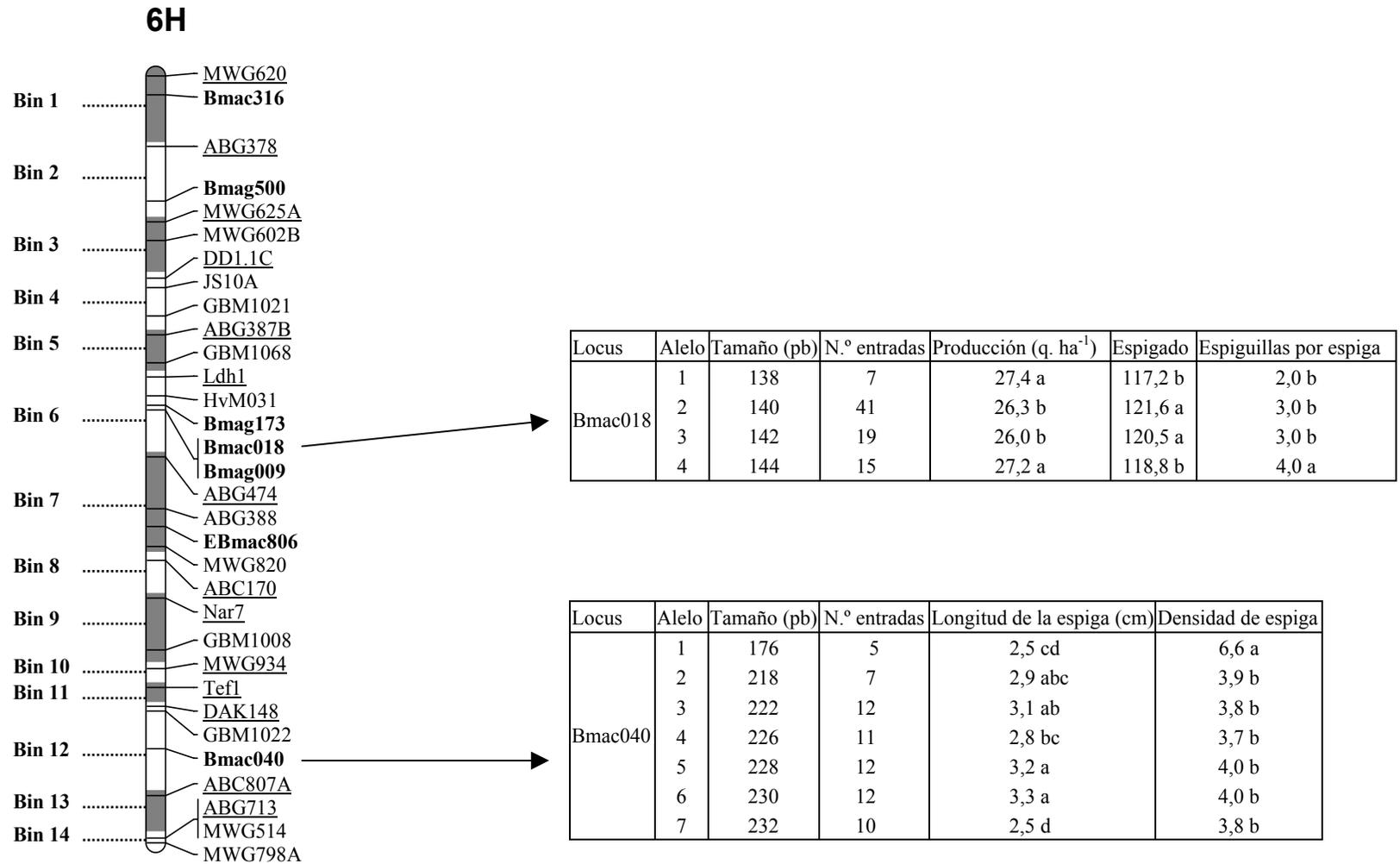


Figura 6.8b

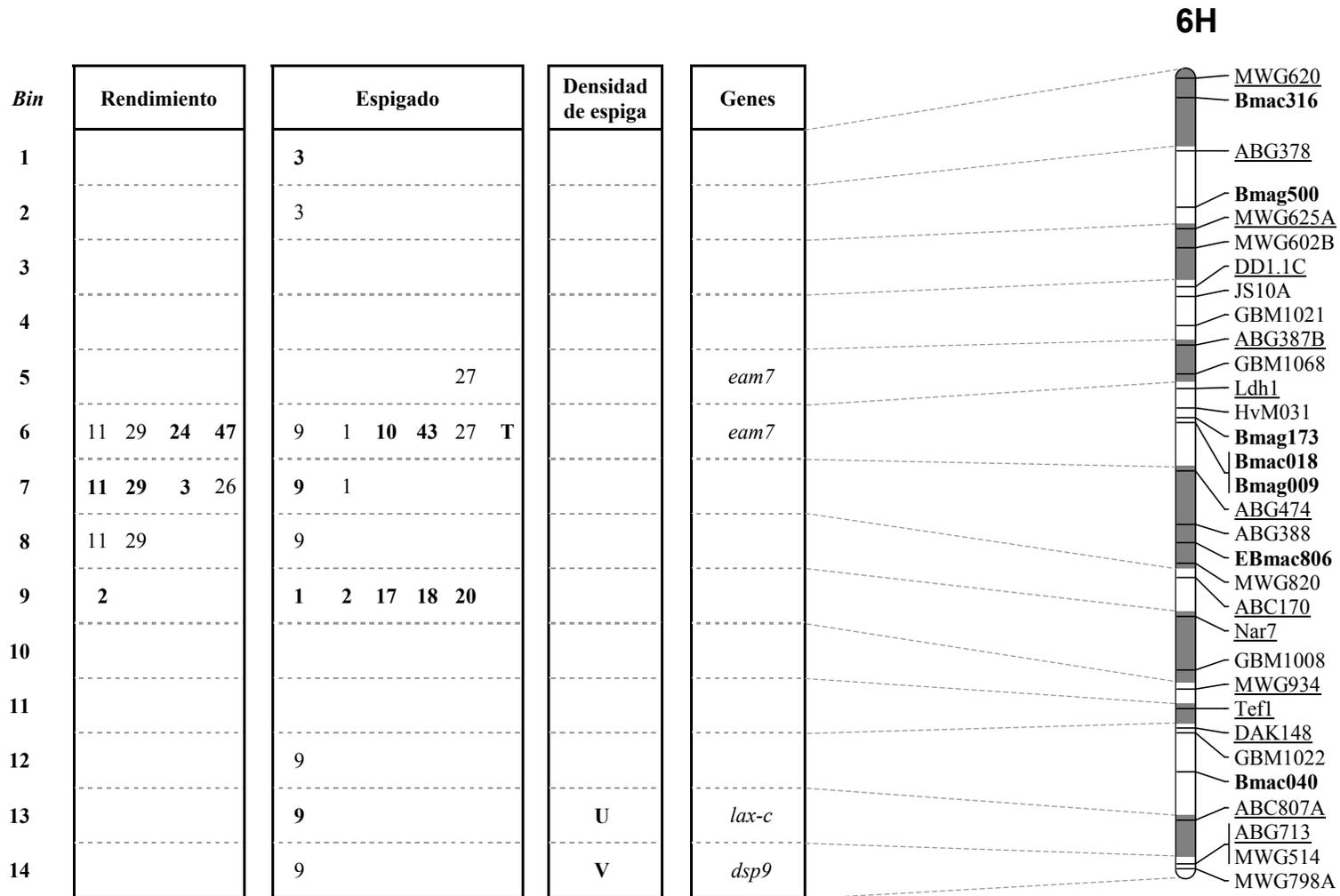
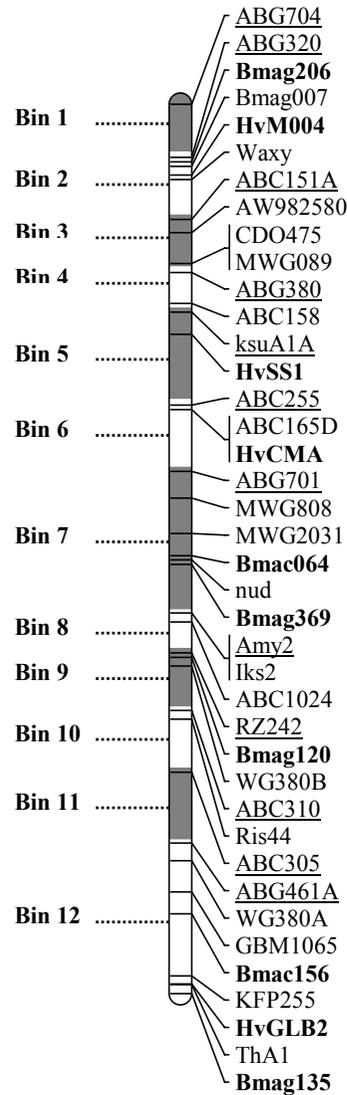


Figura 6.8c

7H



Locus	Alelo	Tamaño (pb)	N° entradas	Longitud de espiga (cm)
HvM004	1	194	18	3,3 bc
	2	198	13	3,5 ab
	3	200	15	3,1 c
	4	202	31	3,6 a
	5	248	10	3,6 ab
	6	252	12	3,7 a
	7	256	11	3,4 abc

Locus	Alelo	Tamaño (pb)	N.º de entradas	Espigado	Glumas y aristas de la gluma
Bmag369	1	185	31	118,6 a	2,2 a
	2	187	62	118,5 a	1,8 b
	3	189	53	117,0 b	1,8 b

Locus	Alelo	Tamaño (pb)	N.º de entradas	Densidad de espiga	Longitud de espiga (cm)
Bmag120	1	230	10	4,3 bc	3,5 ab
	2	232	8	4,6 b	3,8 ab
	3	234	21	4,3 bc	3,4 bc
	4	236	48	4,0 bc	3,6 ab
	5	238	8	4,7 ab	3,0 c
	6	244	9	5,8 a	3,2 bc
	7	262	9	4,5 b	3,4 abc
	8	264	21	4,3 bc	3,5 ab
	9	268	8	3,3 c	3,9 a

Figura 6.9a

7H

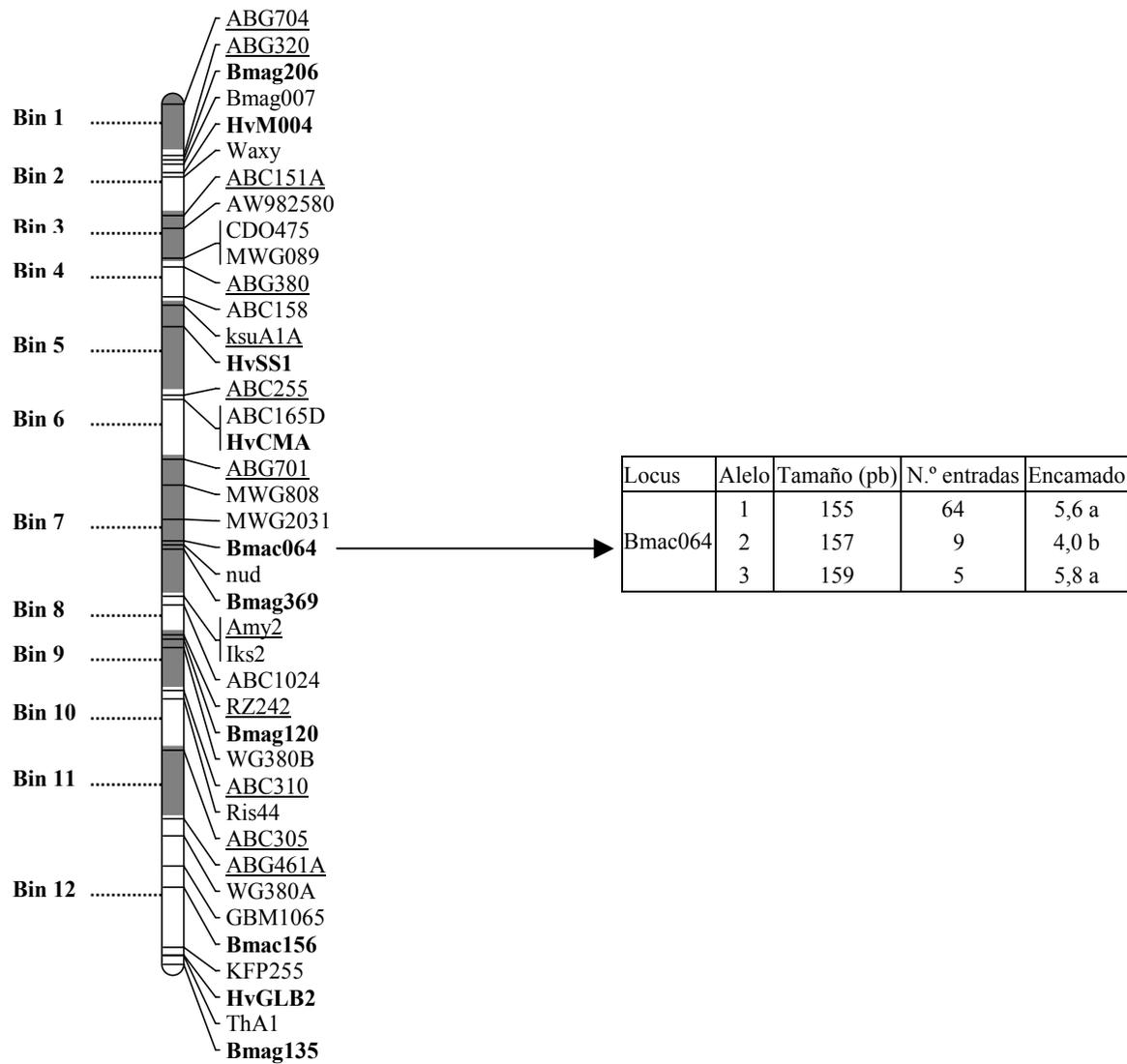


Figura 6.9 b

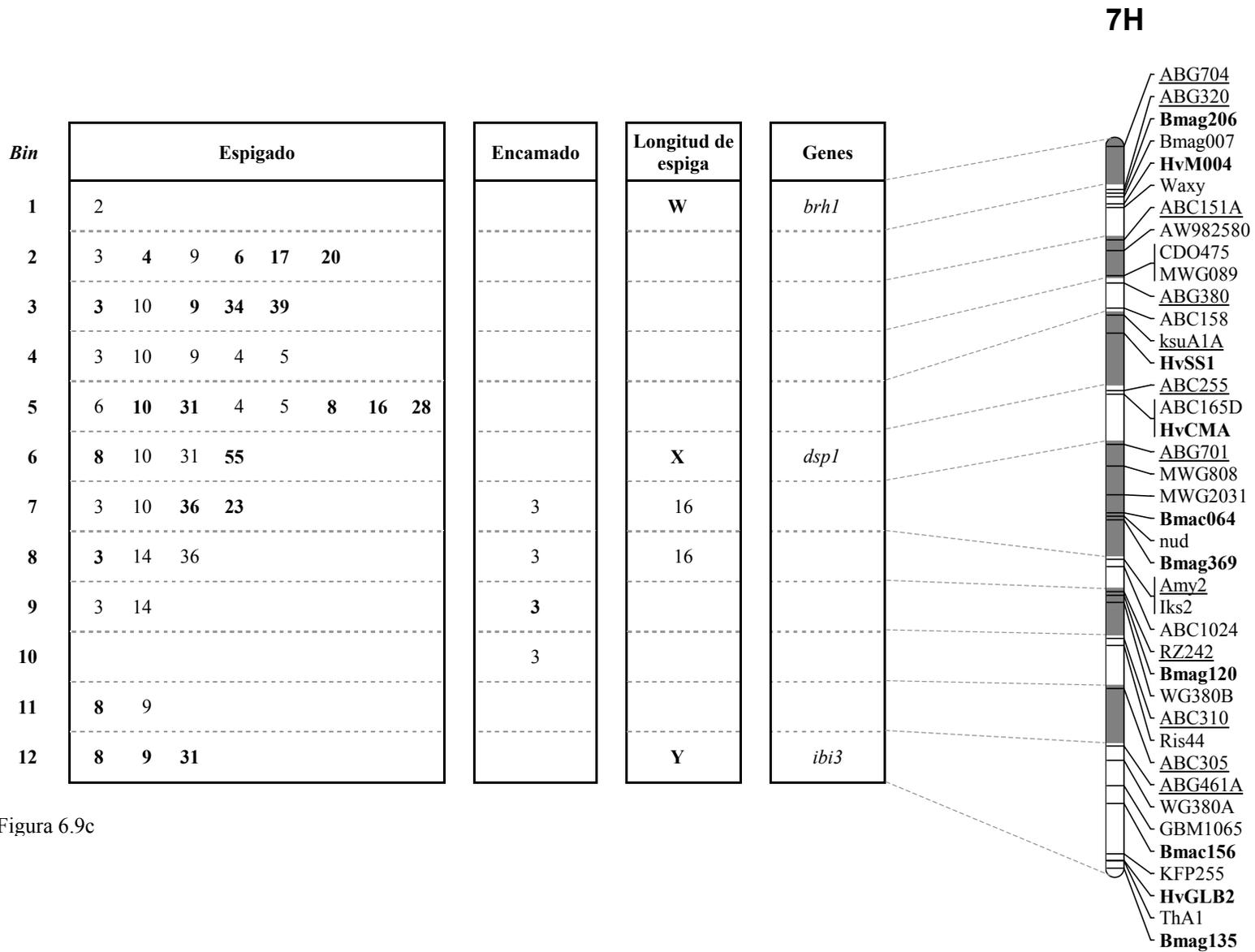


Figura 6.9c

Tabla 6.7. Descripción de las poblaciones numeradas en las figuras 6.3c-6.9c y sus correspondientes referencias.

Población	Número de la población	Autores
Blenheim x E224/3	1	Barua <i>et al.</i> , 1993; Thomas <i>et al.</i> , 1995; Powell <i>et al.</i> , 1997.
Blenheim x Kym	2	Bezant <i>et al.</i> , 1996, 1997.
Stephoe x Morex	3	Hayes <i>et al.</i> , 1993; Boyd <i>et al.</i> , 2003.
Dicktoo x Morex	4	Pan <i>et al.</i> , 1994; Boyd <i>et al.</i> , 2003; von Zitzewitz <i>et al.</i> , 2005.
Harrington x Morex	5	Marquez-Cedillo <i>et al.</i> , 2000, 2001; Ayoub <i>et al.</i> , 2002
Harrington x TR306	6	Tinker <i>et al.</i> , 1996; Mather <i>et al.</i> , 1997; Spaner <i>et al.</i> , 1998.
Prisma x Apex	7	Yin <i>et al.</i> , 1999, 2002, 2005.
Harry x ISR101-23	8	Pillen <i>et al.</i> , 2004.
Igri x Triumph	9	Laurie <i>et al.</i> , 1994, 1995.
Chevron x M69	10	de la Pena <i>et al.</i> , 1999; Canci <i>et al.</i> , 2003, 2004.
Derkado x B83/21/5	11	Thomas <i>et al.</i> , 1998; Ellis <i>et al.</i> , 2002.
L94 x Vada	12	Qi <i>et al.</i> , 1998.
Hor1063 x Krona	13	Kicherer <i>et al.</i> , 2000.
Tadmor x ER(Apm)	14	Teulat <i>et al.</i> , 2001.
Arta x <i>H. spontaneum</i>	15	Baum <i>et al.</i> , 2003.
Brenda x HS213	16	Li <i>et al.</i> , 2005.
Chebec x Harrington	17	Barr <i>et al.</i> 2003b; Boyd <i>et al.</i> 2003.
Sloop x Alexis	18	Barr <i>et al.</i> , 2003a; Boyd <i>et al.</i> , 2003; Karakousis <i>et al.</i> , 2003b.
Sloop x Halcyon	19	Boyd <i>et al.</i> , 2003; Read <i>et al.</i> , 2003; Park <i>et al.</i> , 2003.
Tallon x Kaputar	20	Cakir <i>et al.</i> , 2003; Boyd <i>et al.</i> , 2003; Park <i>et al.</i> , 2003.
Mundah x Keel	21	Long <i>et al.</i> 2003.
Amagi Nijo x WI2585	22	Pallotta <i>et al.</i> , 2003; Reinheimer <i>et al.</i> , 2004.
Apex x ISR101-23	23	Pillen <i>et al.</i> , 2003.
Calicuchima-sib x Bowman	24	Hayes <i>et al.</i> , 1996, 1997.
Caesarea 26-24 x Harrington	25	Matus <i>et al.</i> , 2003.
Karl x Lewis	26	See <i>et al.</i> , 2002.
M92-299 x M81	27	Canci <i>et al.</i> , 2004.
Post x Vixen	28	Scheurer <i>et al.</i> , 2001.
SM73 x SM145	29	Zhu <i>et al.</i> , 1999a.
Tystofte Prentice x Vogelsanger Gold	30	Kjaer <i>et al.</i> , 1995; Kjaer y Jensen, 1996.
VB9524 x ND11231*12	31	Emeberi <i>et al.</i> , 2003.
Beka x Logan	32	Moralejo <i>et al.</i> , 2004.
Dicktoo x Kompolti korai	33	Karsai <i>et al.</i> , 2005.
Gobernadora x CMB643	34	Zhu <i>et al.</i> , 1999b.
Nure x Tremois	35	Francia <i>et al.</i> , 2004; von Zitzewitz <i>et al.</i> , 2005.
Oregon Wolfe Barley	36	Börner <i>et al.</i> , 2002.
Patty x Tallon	37	Park <i>et al.</i> , 2003.
Shyri x Galena	38	Toojinda <i>et al.</i> , 2000.
Stander x MNS93	39	Canci <i>et al.</i> , 2004.
Arapiles x Franklin	40	Boyd <i>et al.</i> , 2003.
Arapiles x Caminant	41	Garvin <i>et al.</i> , 1998.
2571 x Betzes	42	Börner <i>et al.</i> , 2002.
Atsel x Betzes	43	Stracke y Börner, 1998.
mat-a11 x Betzes	44	Börner <i>et al.</i> , 2002.
Igri x Danilo	45	Backes <i>et al.</i> , 1995.
Super precoz x Betzes	46	Börner <i>et al.</i> , 2002.
Tankard x Livet	47	Rajasekaran <i>et al.</i> , 2004.
Vada x 1B-87	48	Backes <i>et al.</i> , 2003.
Vada x 81882	49	Pickering <i>et al.</i> , 1998.
Galleon x Haruna Nijo	50	Karakousis <i>et al.</i> , 2003a.
Fredrickson x Stander	51	Mesfin <i>et al.</i> , 2003.
Gold x Prior	52	Collins <i>et al.</i> , 2001.
L94 x HS078	53	Ivandic <i>et al.</i> , 1998.
L94 x HS084	54	Ivandic <i>et al.</i> , 1998.
<i>H. spontaneum</i>	55	Ivandic <i>et al.</i> , 2003.

Tabla 6.8. Descripción de los genes citados en las figuras 6.3c - 6.9c, con los códigos asignados a las referencias

Cr	Genes	Descripción del gen	Carácter relacionado con el gen	Fuente	Código
1H	<i>Rph4</i>	Resistencia a <i>Puccinia hordei</i>	Roya	McDaniel y Hathcock, 1969.	A
	<i>eam8</i>	<i>Early maturity</i>	Espigado	Franckowiak <i>et al.</i> , 1997.	G
2H	<i>Rph1</i>	Resistencia a <i>Puccinia hordei</i>	Roya	Roane y Starling, 1989.	B
	<i>Ppd-H1</i>	Respuesta a fotoperíodo	Espigado	Laurie <i>et al.</i> , 1994.	9
	<i>Rph16</i>	Resistencia a <i>Puccinia hordei</i>	Roya	Drescher <i>et al.</i> , 2000.	C
	<i>Eam6</i>	<i>Early maturity</i>	Espigado	Franckowiak y Konishi, 2002.	H
	<i>eog</i>	<i>Elongated outer glume</i>	Glumas y aristas de la gluma	Franckowiak y Lundqvist, 1997.	I
	<i>Gth</i>	<i>Toothed lemma</i>	Tipo de lemma	Franckowiak, 1997.	J
	<i>vrs1</i>	<i>Deficiens</i>	Tipo de espiga	Komatsuda <i>et al.</i> , 1999.	F
	<i>ant2</i>	<i>Anthocyanin-less (purple straw)</i>	Pigmentación del tallo	Franckowiak y Lundqvist, 1997.	K
3H	<i>Pau</i>	<i>Purple auricle</i>	Pigmentación de la aurícula	Nilan, 1964.	L
	<i>sdw1</i>	Denso	Altura, (Espigado)	Laurie <i>et al.</i> , 1993.	Z
	<i>ant28</i>	<i>Proanthocyanidin</i>	Pigmentación	Jende-Strid, 1998.	M
4H	<i>eam10</i>	<i>Early maturity</i>	Espigado	Börner <i>et al.</i> , 2002.	N
	<i>Vrn-H2</i>	Respuesta a vernalización	Espigado	Konishi y Franckowiak, 1997.	O
5H	<i>Hln</i>	<i>Hairs on lemma nerves</i>	Tipo de lemma	Franckowiak y Konishi, 1997.	P
	<i>Rph2</i>	Resistencia a <i>Puccinia hordei</i>	Roya	Borovkova <i>et al.</i> , 1997.	D
	<i>srh</i>	<i>Short rachilla hairs</i>	Longitud de los pelos de la raquilla	Franckowiak, 1997.	R
	<i>Eam5</i>	<i>Early maturity</i>	Espigado	Franckowiak, 2002.	S
	<i>Vrn-H1</i>	Respuesta a vernalización	Espigado	Laurie <i>et al.</i> , 1995.	9
	<i>Rph9/12</i>	Resistencia a <i>Puccinia hordei</i>	Roya	Franckowiak y Jin, 1997.	E
6H	<i>eam7</i>	<i>Early maturity</i>	Espigado	Franckowiak y Gallagher, 1997.	T
	<i>lax-c</i>	<i>Laxatum-c</i>	Densidad de espiga	Lundqvist y Franckowiak, 1997.	U
	<i>dsp9</i>	<i>Dense spike</i>	Densidad de espiga	Konishi y Franckowiak, 1997.	V
7H	<i>brh1</i>	<i>Brachytic</i>	Longitud de espiga	Franckowiak, 1997.	W
	<i>dsp1</i>	<i>Dense spike</i>	Densidad de espiga	Franckowiak y Konishi, 1997.	X
	<i>lbi3</i>	<i>Long basal rachis internode</i>	Longitud de espiga	Franckowiak, 1997.	Y

6-4. Discusión

Los análisis de asociación realizados en las 156 *landraces*, sin tener en cuenta la estructura de población, llevaron a la detección de un número elevado de asociaciones. De hecho, las 311 asociaciones detectadas en los análisis sin la inclusión de la estructura de población, que suponen un 14,0% de todos los análisis realizados, se redujeron a sólo la quinta parte (64, o sea, el 2,9%) al considerarla, por lo que cabe pensar que muchas de las asociaciones detectadas con el modelo SG fueran posiblemente falsos positivos.

Estos resultados concuerdan con los resultados de Thornsberry *et al.* (2001), quienes consiguieron disminuir el número de asociaciones positivas por un factor de 4,7 para algunos caracteres, en un estudio de asociación del carácter fecha de floración en maíz, al incluir el efecto de la estructura de la población. También Wilson *et al.* (2004) encontraron que la asociación entre el gen candidato *shrunken2* (*sh2*) y el contenido en almidón, previamente descrita por Prioul *et al.* (1999) en una muestra de líneas puras de maíz, no era tan evidente al analizar un conjunto mayor de líneas puras, teniendo en cuenta la estructura de población en el análisis. Estos autores concluyeron que la asociación encontrada en esta región pudo ser consecuencia de una interacción genotipo por ambiente, o de la influencia exclusiva de la estructura de población. Por eso, Flint-Garcia *et al.* (2005), recomiendan que, cuando se hacen estudios de asociación para cualquier carácter, hay que examinar previamente el efecto de la estructura de población. Aranzana *et al.* (2005) comentaron en este sentido que, aunque sea reconocido que la estructura de población causa correlaciones espurias, la predicción de los falsos positivos es difícil, pues las dimensiones de este problema no son conocidas, por la escasez de estudios de asociación a nivel de genomas completos. En el presente estudio, los caracteres que presentaron un número mayor de asociaciones en los análisis, sin tener en cuenta el efecto de la estructura de población, fueron el tipo de espiga (dos o seis carreras), peso de mil granos, peso hectolítrico, tipo de lemma y longitud de espiga (Tabla 6.2). El número de asociaciones detectadas con el modelo SG para estos caracteres fue muy elevado, entre un 26% y un 47% de las asociaciones analizadas. Estos porcentajes son exageradamente elevados y, sin duda, se deben a la presencia de asociaciones espurias. De hecho, la mayoría de estas asociaciones desaparecieron al incluir el efecto de la estructura de población en los análisis. Estos caracteres están relacionados con la morfología de la espiga y del grano, por lo que es posible que estén fuertemente influenciados por la presencia de entradas de dos y seis carreras en la

colección nuclear. De hecho, uno de los grupos de entradas definidos por el análisis de grupos está compuesto casi exclusivamente por entradas de dos carreras (el grupo 2). El grupo 2 ya mostró un peso hectolítrico significativamente superior a los otros tres grupos, compuestos casi exclusivamente por entradas de seis carreras (Tabla 5.11b del capítulo 5). Este hecho se repite para el carácter longitud de espiga (datos no mostrados). De esta manera, existe un polimorfismo cierto para estos caracteres fenotípicos en la colección nuclear, pero parece que la estructura de población detectada se hace cargo de una parte importante de ese polimorfismo fenotípico. Así, el número de asociaciones de marcadores con esos cuatro caracteres disminuye acusadamente en el modelo CG (Tabla 6.2), pues la variación dentro de grupos es relativamente menor que entre los grupos. El marcador MWG699, que se sitúa junto al gen principal que determina el número de carreras (*vrs1*), aparecía asociado a un gran número de caracteres que se distribuyen de forma diferente entre 2 y 6 carreras (p.ej. altura, peso hectolítrico, peso de mil granos, granos por espiga o longitud de la espiga), al no considerar la estructura de población. Sin embargo, la mayor parte de esas asociaciones dejaron de ser significativas al tener en cuenta dicha estructura.

Esto se confirmó también con la comparación de los promedios de los coeficientes de determinación de los modelos sin grupos y con grupos (Figura 6.1, 6.2), donde estos coeficientes fueron mucho más altos en los modelos sin grupos.

El espigado es otro carácter que mostró muchas asociaciones, tal vez debido a un efecto de selección para adaptación a las distintas zonas geográficas, lo que probablemente habrá influido en la formación de los grupos genéticos. En este sentido, Remington *et al.* (2001), comentaron que la estructura de población en maíz parece reflejar los efectos de la selección de caracteres adaptativos como la floración. Este papel adaptativo de la floración y su influencia en la estructuración de las poblaciones se puso también de manifiesto en el estudio de Aranzana *et al.* (2005), quienes, al realizar un estudio de mapeo por asociación, descubrieron que la floración está fuertemente correlacionada con el origen geográfico en *Arabidopsis thaliana*. Esta situación coincide con la que resultaría para un carácter bajo selección latitudinal.

En general, caracteres muy correlacionados con la estructura de población pueden mostrar una elevada tasa de falsos positivos. Sin embargo, puede darse el caso de que existan alelos funcionales cuya distribución coincida con la estructura de población y que no sean detectados en el análisis al tener en cuenta dicha estructura (falsos negativos). Por este motivo, Flint-Garcia *et al.* (2005), recomiendan que la manera más

directa para validar una asociación potencial es evaluar el polimorfismo candidato en una muestra poblacional distinta, con una estructura poblacional diferente. En este sentido, para validar las asociaciones encontradas entre microsatélites y caracteres fenotípicos en maíz, Remington *et al.* (2001), realizaron los análisis en una de las poblaciones generadas por el programa STRUCTURE. En el presente estudio, la búsqueda de asociaciones en una de las poblaciones, donde la estructura de población era despreciable, redujo los casos positivos a 45. De estos 45, solamente 19 habían sido encontrados en el análisis de la totalidad de los genotipos, teniendo en cuenta la estructura de población, y otros 22 se detectaron inicialmente en el análisis sin tener en cuenta dicha estructura. Estos últimos pueden ser casos en los que las frecuencias alélicas sean bastante diferentes entre grupos, por lo que las asociaciones que presenten desaparecerían en buena medida en un modelo que considere la estructura de población. Sin embargo, el hecho de que reaparezcan en el modelo G4, donde el desequilibrio de ligamiento es muy bajo, sugiere que se trata de asociaciones tan probables como las que puedan aparecer con el modelo CG.

La eliminación de los alelos presentes en una frecuencia inferior al 5% en la muestra de genotipos analizada, hizo que los alelos que entraron en el análisis de las 156 *landraces* y en el grupo 4 no fueran comunes en su totalidad para algunos loci. Eso también podría explicar en parte la aparición de nuevas asociaciones en la población 4, que no se detectaron en el análisis de las 156 *landraces*.

Para averiguar hasta qué punto las asociaciones encontradas son reales, se compararon los resultados obtenidos en este estudio con los de QTLs encontrados por otros autores en estudios de ligamiento en poblaciones segregantes biparentales clásicas, y con las posiciones de genes conocidos.

Rendimiento

Una cantidad apreciable de loci mostró asociación con el rendimiento en el análisis de las 156 *landraces*. La mayoría coincidieron con regiones donde se encontraron QTLs en poblaciones biparentales (Figura 6.4c, 6.5c, 6.6c, 6.7c, 6.8c). Los loci WMC1E8 y Bmac134 no coincidieron con ningún QTL de rendimiento conocido, aunque la asociación positiva que mostraron se puede deber a la correlación que existe entre el rendimiento y el espigado ($r=-0,41$, $P<0,0001$). De todos los loci que presentaron asociación al mismo tiempo para rendimiento y espigado, se comprobó que

los alelos que corresponden a una floración precoz, fueron siempre los que mayor rendimiento presentaron. El locus Bmac134 mostró asociación con espigado, pero no así el locus WMC1E8. La no-asociación de este locus con el espigado, se debe al nivel de significación estricto elegido. Este marcador mostró un nivel de significación cercano al umbral de significación elegido en este estudio, lo que podría ser un caso de falso negativo.

El análisis multilocus de este carácter mostró que los loci scssr07759, EBmac701, HvBM5A y Bmag009 fueron los que más explican la variación de rendimiento. Diferentes QTLs fueron detectados en estas regiones en varias poblaciones. En el *bin* 5, del cromosoma 2H donde se localiza el locus scssr07759, fue detectado un QTL de rendimiento en las poblaciones Blenheim x Kym (Bezant *et al.*, 1997) y SM73 x SM145 (Zhu *et al.*, 1999a) (Figura 6.4c). Otro QTL para rendimiento fue descrito dentro del *bin* 10 del cromosoma 4H, en el cual se encuentra el locus EBmac701, en las poblaciones Harrington x TR306 (Tinker *et al.*, 1996); Prisma x Apex (Yin *et al.*, 1999, 2002); VB9524 x ND1123*12 (Emebiri *et al.*, 2003); SM73 x SM145, entre los *bins* 8 a 10 (Zhu *et al.*, 1999a) y Derkado x B83-12/21/5, entre los *bins* 10 y 12 (Thomas *et al.*, 1998; Ellis *et al.*, 2002) (Figura 6.6c). En el cromosoma 5H, fueron encontrados QTLs de rendimiento en las poblaciones Blenheim x E224/3 (Barua *et al.*, 1993, Powell *et al.*, 1997) y Apex x ISR101-23 (Pillen *et al.*, 2003), en el *bin* 11 donde se localiza el locus HvBM5A (Figura 6.7c). En el *bin* 6 del cromosoma 6H, donde se sitúa el locus Bmag009, se detectaron QTLs en las poblaciones Calicuchima-sib x Bowman. (Hayes *et al.*, 1996) y Tandard x Livet (Rajasekaran *et al.*, 2004). También Zhu *et al.* (1999a) así como Ellis *et al.* (2002) detectaron QTLs respectivamente en las poblaciones SM73 x SM145 y Derkado x B83-12/21/5, entre los *bins* 6 y 8.

En la población 4, solo se detectaron dos loci, scssr07759 y Bmac018. Este último se localiza en una posición muy cercana al locus Bmag009, ya descrito en el párrafo anterior.

El locus scssr07759 es el único que reveló asociación con el rendimiento en la totalidad de las entradas y en la población 4. El alelo de 213 pb fue el que se caracterizó por un promedio de rendimiento alto en los dos análisis (156 *landraces* y población 4).

La distribución no aleatoria de los alelos en los grupos podría ser una de las principales causas que engendran asociaciones con distintos loci entre las dos muestras de entradas. De hecho para los loci HvBM5A y Bmag009, el alelo que presentaba

diferencias significativas en el análisis de las 156 *landraces*, no estaba presente o se trataba de un alelo raro, que no entraba en el análisis de la población 4.

Espigado

Doce y dieciséis loci mostraron asociaciones con este carácter, respectivamente en las 156 *landraces* y en la población 4. De todos los loci que mostraron asociación con el espigado en las 156 *landraces* o en la población 4, solamente 2 no coinciden con regiones donde se describieron QTLs para espigado en la literatura. Los numerosos estudios existentes en la literatura ponen de manifiesto la alta frecuencia de detección de QTLs para este carácter. Esto no es sorprendente por una parte, pues es un carácter de medida sencilla, por lo que se evalúa en casi todos los experimentos de búsqueda de QTLs. Es un carácter que difícilmente se podría calificar como cualitativo o cuantitativo, más bien es un carácter intermedio entre estas dos situaciones. Está controlado por parte de unos pocos genes mayores, tres de respuesta a la vernalización y dos de respuesta a fotoperiodo (Takahashi y Yasuda, 1970; Laurie *et al.*, 1994, 1995), que configuran una respuesta cualitativa, pero bastante variada y aún no plenamente descrita. Además existen otros genes de menor efecto que, en unión con los genes mayores, terminan por formar una respuesta aparentemente cuantitativa. Pero, por otra parte, su heredabilidad es alta (Manzjuk y Barsukov, 1974, Ma *et al.*, 2000), habitualmente gracias a una varianza del error reducida. Este hecho favorece la detección de QTLs, al menos por comparación con caracteres cuantitativos también medidos habitualmente, como el rendimiento. De los 9 loci que entraron en el modelo más explicativo de este carácter, 8 coincidieron con regiones donde se encontraron QTLs en poblaciones biparentales. El locus Bmac032 se localiza en el *bin* 9 del cromosoma 1H, en el cual se describieron QTLs en las poblaciones Harry x ISR101-23 (Pillen *et al.*, 2004) e Igri x Triumph (Laurie *et al.*, 1995). Entre los *bins* 10 y 11 de este cromosoma, se encontraron más QTLs para espigado en otras poblaciones (Figura 6.3c). Este marcador podría estar en desequilibrio de ligamiento con el gen *eam8*, asignado al *bin* 11 por Börner *et al.* (2002) al analizar las poblaciones 2571 x Betzes y mat-a11 x Betzes, aunque éste fue descrito por Franckowiak *et al.* (1997) en el *bin* 14, donde se encuentra el locus WMC1E8, que reveló asociación con rendimiento.

En el *bin* 1 del cromosoma 2H, donde se encuentra el locus Bmac134, no se ha descrito ningún QTL para espigado en la literatura, sin embargo en el *bin* 5, donde se

localiza scssr07759, fueron detectados QTLs en las poblaciones Blenheim x E224/3 (Powell *et al.*, 1997) y Blenheim x Kym (Bezant *et al.*, 1996). En esta última, el alelo de Kym fue relacionado con el espigado tardío. Este también es el caso del locus scssr07759, para el cual las *landraces* que presentan el alelo de Kym (221 pb) espigaron 1,9 días más tarde. Las mismas variaciones de los alelos fueron observadas en la población 4. Este locus podría estar en relación con el gen de fotoperíodo *Ppd-H1* descrito por Laurie *et al.* (1994) en el *bin* 4 de este cromosoma.

En el *bin* 2 del cromosoma 3H, donde se localiza el locus HvLTPPB, solo se describió un QTL en la población Harry x ISR101-23 (Pillen *et al.*, 2004). En este locus, el alelo de 221 pb, que correspondió al espigado precoz, es típico del material español, y se ha encontrado solamente en 2 variedades (Kym y Cameo) de las empleadas en el estudio de diversidad (capítulo 3). En el *bin* 15 de este mismo cromosoma, donde se encuentra el locus HvM062, se encontró un QTL en la población Igri x Triumph, donde el alelo de esta última variedad fue descrito como tardío (Laurie *et al.*, 1995). Estos resultados coinciden con los observados en el presente estudio para los alelos de HvM062. Las *landraces* que presentan el alelo de Igri (247 pb) fueron más precoces que las que tienen el alelo de Triumph (255 pb), como se muestra en la Figura 6.5c. Este locus podría tener relación con el gen *eam10*, descrito por Börner *et al.* (2002), aparentemente situado en el *bin* 16 de este cromosoma.

En el *bin* 13 del cromosoma 4H, donde se encuentra el locus HvBAMY, se describieron QTLs para espigado en las poblaciones Steptoe x Morex (Hayes *et al.*, 1993), Calicuchima-sib x Bowman (Hayes *et al.*, 1996) y Harry x ISR101-23 (Pillen *et al.*, 2004). En la población Steptoe x Morex, los genotipos que presentaron el alelo de Steptoe adelantaron sus espigados 1,5 días con respecto a los que presentaron el alelo de Morex. También en las *landraces*, las entradas que presentaron el alelo de tamaño similar al de Steptoe (129 pb) en el locus HvBAMY espigaron 1 día antes que las entradas que presentaron el alelo de Morex (132 pb) (Figura 6.6c). Este locus podría tener relación con el gen de vernalización (*Vrn-H2*) situado en el *bin* 12, donde se encontraron QTLs para espigado en más poblaciones (Figura 6.6c). Los resultados del cruzamiento Steptoe x Morex son particularmente interesantes para compararlos con los de este estudio, pues Steptoe es una variedad californiana de ancestros traídos desde España.

En el *bin* 9 del cromosoma 5H, donde se sitúa el locus Bmag223, se describieron QTLs de espigado en las poblaciones Dicktoo x Morex (Pan *et al.*, 1994) y Chevron x

M69 (de la Pena *et al.*, 1999). También Ivandic *et al.* (2003), encontraron una asociación entre el microsatélite Bmag223 y la fecha de espigado en *H. spontaneum*. Este marcador podría estar en relación con el gen *Eam5*, que se encuentra en el mismo *bin* (Franckowiak, 2002).

En el *bin* 6, del cromosoma 6H se encuentra el locus Bmac018. En este mismo *bin* se describieron QTLs para espigado en diferentes poblaciones: Igri x Triumph (Laurie *et al.*, 1995), Blenheim x E224/3 (Barua *et al.*, 1993), Chevron x M69 y M92-299 x M81 (Canci *et al.*, 2004) o Atsel x Betzes (Stracke y Börner, 1998). Estos últimos autores localizaron el gen *eam7* en este *bin*, aunque este gen había sido descrito en el *bin* 5 por Franckowiak y Gallagher (1997). El locus Bmac018 mostró asociación con el espigado en la población 4. Tanto en las 156 *landraces* como en la población 4, las entradas que tienen los alelos de 138 y 144 pb se caracterizaron por un espigado precoz, mientras que las que presentan el alelo de 140 pb, fueron tardías. Curiosamente, el alelo de 142 pb fue precoz en las 156 *landraces* y tardío al analizar solamente la población 4.

Bmag369 es el noveno marcador que entra en el modelo multilocus más explicativo del espigado. Este se sitúa en el *bin* 7 del cromosoma 7H, en el cual fueron descritos QTLs en varias poblaciones. Pillen *et al.* (2003) describieron un QTL en la población Apex x ISR101-23, en el marcador Bmac064, que se encuentra a pocos centiMorgans del locus Bmag369. Börner *et al.* (2002) encontraron un QTL en el mismo *bin*, en la región del gen *nud* que se sitúa entre los dos loci Bmac064 y Bmag369, en la población Oregon Wolfe Barley. Por otra parte Canci *et al.* (1994) describieron un QTL en la población Chevron x M69, entre los *bins* 6 y 7. También Hayes *et al.* (1993), encontraron un QTL entre los *bins* 7 y 9, en la población Steptoe x Morex.

La mayoría de los loci que mostraron asociación con el espigado en la población 4 coincidieron con los encontrados en la totalidad de las *landraces*. De los 6 loci que mostraron asociación únicamente en esta población, sólo el locus EBmac701 (*bin* 10) no coincidió con ningún QTL de espigado conocido. Aun así, este marcador es el más explicativo de la variación de espigado en los análisis locus por locus, y entra en el modelo más explicativo de la variación del espigado en esta población. El QTL más cercano a esta posición descrito en la literatura es un QTL en el *bin* 9 en la población Igri x Triumph, en ensayos de siembra de otoño (Laurie *et al.*, 1995). Los genotipos de tipo Triumph espigaron 2 días más tarde que los genotipos de tipo Igri. Sin embargo, en el locus EBmac701 se observó el efecto contrario con respecto a los alelos de esas dos

variedades. Las *landraces* que tienen el mismo alelo que Triumph (135 pb) espigaron 2,5 días antes que las *landraces* que presentan el alelo de Igri (147 pb) (Figura 6.6b), por lo que parece tratarse de un polimorfismo distinto.

Encamado

Para este carácter se detectaron asociaciones distintas en las 156 *landraces* y en la población 4. Esto podría tener relación con el hecho de que los alelos que entraron en el análisis en cada muestra no son exactamente los mismos.

Los loci Bmac032 y Bmag006 fueron los únicos que revelaron asociación con el encamado en las 156 *landraces*. El primer locus no coincide con ningún QTL conocido. Sin embargo para la posición de Bmag006, localizado en el *bin* 6 del cromosoma 3H, se han descrito QTLs en las poblaciones Steptoe x Morex (Hayes *et al.*, 1993) y Tallon x Kaputar (Cakir *et al.*, 2003).

Los dos loci presentaron un efecto aditivo en la explicación de la variación del encamado.

El locus Bmac064, situado en el *bin* 7 del cromosoma 7H, es el único que presentó asociación con el encamado en la población 4. Hayes *et al.* (1993) encontraron un QTL para este carácter entre los *bins* 7 y 10 en la población Steptoe x Morex.

Peso hectolítrico

EBmac701 fue el único locus que mostró una asociación con este carácter en la población 4. Este se sitúa en el *bin* 10 del cromosoma 4H, donde no se describió ningún QTL para este carácter en la literatura. La detección de esta asociación podría ser el resultado de la correlación que podría existir entre peso hectolítrico y rendimiento. De hecho, esta es la región del cromosoma donde más QTLs para rendimiento se han descrito en la literatura (Figura 6.6c).

Peso de mil granos

Los loci MWG699 y HvBM5A mostraron asociación con el peso de mil granos en los modelos CG y G4. Los dos loci fueron significativos en el modelo multilocus en el análisis de todas las *landraces*; sin embargo, la significación del locus HvBM5A se anuló en la población 4. Esto se debe principalmente a que estos dos marcadores se encuentran en una situación de desequilibrio de ligamiento en esta población ($r^2=0,32$,

$P < 0,0001$). La mayoría de las entradas de esta población que presentan el tipo A para el locus MWG699 tienen a su vez el alelo de 1200 pb en el locus HvBM5A, mientras que las de tipo D (en MWG699) corresponden en su mayoría al alelo de 4850 pb (en HvBM5A), por lo cual la información obtenida para los dos loci en esta población es redundante.

La región donde se sitúa el locus MWG669 en el cromosoma 2H (*bin* 9), coincide con la zona donde se han descrito más QTLs para el carácter peso de mil granos (Figura 6.4c). También en el *bin* 11 del cromosoma 5H, donde se localiza el locus HvBM5A, se encontraron QTLs en las poblaciones Prisma x Apex (Yin *et al.*, 2002), Tadmor x ER(Apm) (Teulat *et al.*, 2001) y Amaji Nijo x WI2585 (Coventry *et al.*, 2003; Pallotta *et al.*, 2003).

Granos m⁻²

El locus HdAMYB fue el único marcador que mostró asociación con este carácter en la población 4. Esta asociación no coincide con ningún QTL descrito en la literatura para este carácter. El alelo de 186 pb, que se diferenció con un mayor promedio de granos m⁻² que los otros alelos, es exclusivo de algunas entradas españolas. No se ha encontrado en ninguna de las variedades testigo analizadas en el estudio de diversidad (capítulo 3).

Roya

De los cinco loci que mostraron asociación con la roya, dos coincidieron en *bins* que albergan genes de resistencia conocidos. El locus Bmac399, en el *bin* 4 del cromosoma 1H, es el más explicativo de la variación de la incidencia de esta enfermedad en la población 4, y el segundo más explicativo en las 156 *landraces*. Ningún QTL fue descrito en esta región del cromosoma en la literatura. Sin embargo en el *bin* 1 de este cromosoma, se encuentra el gen de resistencia *Rph4* (McDaniel y Hathcock, 1969). Este marcador podría estar en desequilibrio de ligamiento con el gen.

En la población 4 se detectó una asociación con resistencia a roya para el locus MWG699. No se describió ningún gen de resistencia a esta enfermedad cercano a este locus en la literatura. Se han descrito dos genes de resistencia a roya en este cromosoma, *Rph1* (Roane y Starling, 1989) y *Rph16* (Drescher *et al.*, 2000), que se localizan respectivamente en los *bins* 3 y 6. Sería posible que existiera un desequilibrio

de ligamiento entre el locus MWG699 y uno de esos genes. También se ha descrito otro gen de resistencia a roya en este cromosoma, *Rph15* (Franckowiak y Chicaiza, 1998). Su posición relativa en el cromosoma es desconocida, aunque fue asignado al brazo largo de este cromosoma, por lo cual no se descartaría una posible relación entre este gen y el locus MWG699. Por otra parte, Kicherer *et al.* (2000) identificaron un QTL en esta región del cromosoma en la población HOR 1063 x Krona.

El locus HvBTAI3 mostró asociaciones con esta enfermedad en las dos muestras de entradas, pero no coincide con ningún QTL para resistencia a roya de hojas descrito en esta región. Tampoco se describieron genes de resistencia a roya en este cromosoma. Sin embargo, Hayes *et al.* (1996) encontraron un QTL para resistencia a la roya amarilla (*Puccinia striiformis*) en el mismo *bin* (11) donde se encuentra el locus HvBTAI3. Por otra parte Franckowiak y Jin (1997) comentaron la existencia de dos genes de resistencia a *Puccinia hordei*, *Rph6* y *Rph8*, cuya posición en el genoma es desconocida. El locus Bmag337, localizado en el *bin* 4 del cromosoma 5H, podría estar marcando la presencia de polimorfismo en el gen *Rph2*, descrito por Borovkova *et al.* (1997) en esta zona del cromosoma. En este mismo *bin*, se describieron QTLs para resistencia a la roya de las hojas, en dos poblaciones: Blenheim x E224/3 (Thomas *et al.*, 1995) y Sloop x Halcyon (Park *et al.*, 2003).

El locus Bmag337 no mostró asociación con resistencia a roya en la población 4, posiblemente porque en este grupo de entradas se analizaron menos alelos de este marcador. De hecho, el alelo menos susceptible (143 pb) no entró en el análisis de esta población.

El gen de resistencia a la roya *Rph9/12* se sitúa en el *bin* 11 del mismo cromosoma (Borovkova *et al.*, 1998), donde también se encuentra el locus HvBM5A, con quien podría estar en desequilibrio de ligamiento. Otras evidencias de la presencia de genes de tolerancia en esta región son los QTLs para resistencia a roya descritos en este mismo *bin* en las poblaciones Tallon x Kaputar y Patty x Tallon (Park *et al.*, 2003).

Aunque el locus HvBM5A es el segundo marcador más explicativo de la resistencia a roya en la población 4, éste no entró en el mejor modelo del análisis multilocus. Esto se debe probablemente al desequilibrio de ligamiento que existe entre este marcador y el locus MWG699, y una posible redundancia de la información entre ambos.

En el análisis de las 156 *landraces*, el locus HvBTAI3 fue el que se quedó excluido del modelo multilocus más explicativo. Esto es consistente con el hecho de

que fuera el locus que presentó el coeficiente más bajo en los análisis locus por locus de estas entradas.

Calibre inferior a 2,2 mm

Dos y cinco loci, respectivamente, presentaron una asociación aparente con este carácter en las 156 *landraces* y en la población 4. El locus Bmag125 es el único común entre los dos modelos. No se han descrito QTLs para este carácter en el *bin* 9 del cromosoma 2H, donde se localiza este locus. Sin embargo, en el *bin* 8, se describieron QTLs en dos poblaciones, Chebec x Harrington (Barr *et al.*, 2003b) y Mundah x Keel (Long *et al.*, 2003).

El locus Bmac134, situado en el *bin* 1 del cromosoma 2H, fue el más explicativo para este carácter, en los análisis locus por locus del modelo CG. De nuevo, no coincide con QTLs para calibre del grano conocidos en esta región del cromosoma. Este locus mostró también asociación con la fecha de espigado y con rendimiento. Se observó que los alelos que mostraron un espigado tardío presentaron rendimientos peores. Esta es una situación que se da con bastante frecuencia en ambientes mediterráneos (Voltas *et al.*, 1999). Cabe dentro de lo posible que un estrés hídrico progresivo afectara de modo más acentuado a las entradas más tardías, influyendo en el tamaño y peso de los granos. De hecho, las entradas que presentaron el alelo de 138 pb del locus Bmac134 son las que espigaron claramente más tarde, y también son las que presentaron un porcentaje más alto de granos de calibre inferior a 2,2 mm (Figura 6.4a). Sin embargo, no se detectó asociación para el carácter peso de mil granos en este locus, lo que habría apoyado aún más la hipótesis expuesta.

Aparte del locus Bmag125, los loci scssr07759, HvM054, Bmag225 y Hv13GEIII mostraron asociación con este carácter en el análisis de la población 4. Pero ninguno coincide con QTLs descritos en las regiones del cromosoma donde se localizan. En el caso del locus scssr07759 se observó lo mismo que en el caso del locus Bmac134. Las entradas tardías son las que mayor porcentaje de granos pequeños presentan, aunque tampoco en esta zona se detectó asociación con el peso de mil granos. Por lo cual la asociación de este carácter con estos dos loci puede deberse a la correlación que existe entre los dos caracteres ($r = -0,26$, $P = 0,003$).

Proteína

El locus Bmac067 fue el único que reveló asociación con el porcentaje de proteína en la harina de malta, en el análisis de la población 4. Este locus se sitúa en el *bin* 6 del cromosoma 3H, donde ya se localizó un QTL para este carácter en la población Chevron x M69 (Canci *et al.*, 2003). Este locus no reveló asociación en el análisis de las 156 *landraces*. La introducción de más alelos en el análisis de la población 4, que no fueron incluidos en las 156 *landraces*, podría ser la razón de la detección de la asociación de este locus con proteína en esta población.

Pigmentación del tallo y de la aurícula

Se encontraron asociaciones para el carácter pigmentación del tallo en las 156 *landraces*, mientras que la asociación con la pigmentación de la aurícula se detectó únicamente en la población 4.

En el *bin* 10 del cromosoma 2H, se localiza un gen de antocianina *ant2* (*anthocyanin less*). Está cerca del locus MWG699, que mostró asociación con el carácter pigmentación del tallo. Esta asociación parece estar relacionada con el tipo de espiga. Las entradas de 2 carreras que presentan el haplotipo K del locus MWG699 presentan pigmentación púrpura en el tallo. Como el marcador MWG699 está ligado al gen *vrs1*, que determina el tipo de espiga 2 ó 6 carreras, cualquier carácter que tenga una distribución relacionada con el tipo de espiga, mostrará una asociación con este marcador.

Los loci HdAMYB y Bmag013 son los más explicativos de la variación en la pigmentación del tallo. Ningún gen responsable de la pigmentación ha sido descrito en el *bin* 13 del cromosoma 4H, donde se encuentra el locus HdAMYB. Tampoco se describieron QTLs para este carácter en esta zona del genoma. El locus Bmag013 se sitúa en el *bin* 13 del cromosoma 3H, donde se detectó un QTL para pigmentación del grano en los *bins* 13 y 14, en la población Arapiles x Caminant (Garvin *et al.*, 1998). Estos autores relacionaron el QTL con el gen *ant28* (proantocianidinas), aunque poco después este gen fue asignado al *bin*16 (Jende-Strid, 1998).

También los loci Hv13GEIII y Bmac029, que mostraron asociación con la pigmentación de la aurícula en la población 4, podrían tener relación con este gen, por su localización respectivamente en los *bins* 14 y 16 del cromosoma 3H. HvBKASI es otro locus que reveló asociación con la pigmentación de la aurícula. Se localiza en el *bin*

8 del cromosoma 2H y podría tener relación con el gen *Pau* (*Purple auricle*) aunque su situación, en el *bin11*, parece bastante alejada. El locus Bmag223 es uno de los más explicativos de este carácter con un coeficiente de determinación de 26,5%, pero no coincide con ningún gen.

Vellosidad de la vaina basal

Franckowiak y Konishi, (1997) describieron un gen (*Hsh*) para este carácter, en el cromosoma 4H. En este estudio, ningún marcador de este cromosoma mostró asociación con este carácter. Sin embargo se encontraron asociaciones con cuatro loci localizados en los cromosomas 1H, 5H y 6H, en el análisis de las 156 *landraces*. Estos no coincidieron con ningún QTL o gen descrito en la literatura.

Glumas y aristas de la gluma

En el cromosoma 2H, se describió un gen para este carácter en el *bin 8* (*eog: elongated outer glume*) por Franckowiak y Lundqvist (1997). El locus MWG699 reveló una asociación con este carácter, y podría estar relacionado con este gen por su localización próxima. Otros tres loci mostraron asociación con este carácter, pero no coincidieron con genes conocidos. Tampoco se describieron en la literatura QTLs para este carácter.

Tipo de lemma

Los loci scssr07759, HvBAMY y HvBM5A mostraron asociación con el tipo de lemma. De estos 3 loci, solamente el locus HvBAMY está localizado cerca del gen *Hln*, (*hairs on lemma nerves*) descrito por Franckowiak y Konishi (1997), en el *bin 12* del cromosoma 4H.

En el *bin 9* del cromosoma 2H, se encuentra el gen *Gth* (*toothed lemma*) descrito por Franckowiak (1997). La relación entre este gen y el locus scssr07759 es improbable, pero tampoco se descartaría un posible desequilibrio de ligamiento entre el gen y el locus. Sin embargo, no se ha descrito ningún gen para este carácter en el cromosoma 5H. Curiosamente, los loci scssr07759 y HvBM5A son los marcadores más explicativos de este carácter.

Longitud de los pelos de la raquilla

Un único QTL, fue descrito en la literatura para este carácter, localizado en el *bin* 9 del cromosoma 5H, en la población Galleon x Haruna Nijo (Karakousis *et al.*, 2003a). La posición del QTL coincide con la del gen *Srh* (Franckowiak, 1997). En este mismo *bin* se localiza el locus Bmag223, que mostró una asociación con la longitud de los pelos de la raquilla en este estudio. De hecho, de los loci que revelaron asociación con este carácter tanto en las 156 *landraces* como en la población 4, el locus Bmag223 fue el más explicativo de todos. Este marcador está muy próximo al gen, por lo cual podría estar ligado a él, lo que explicaría los valores altos de los coeficientes de determinación que mostró. Kraakman *et al.* (2006) también encontraron una asociación entre el locus Bmag223 y la longitud de los pelos de la raquilla, en un estudio de ‘*association mapping*’ en cultivares de cebada de dos carreras de primavera.

El locus HvM054 se mostró asociado a este carácter en las dos muestras de genotipos, pero no coincide con ningún gen conocido. El locus HvM020 reveló una asociación con este carácter en la población 4, pero no entró en el modelo multilocus más explicativo. Este locus se encuentra en desequilibrio de ligamiento con el locus Bmag223 en este grupo de entradas. Por este motivo, y dada la fuerte asociación existente entre el carácter y Bmag223, su contribución al carácter desaparece en el modelo multilocus.

Los loci MWG699, Bmac067 y Bmac113 que también mostraron asociación con la longitud de los pelos de la raquilla, en las 156 *landraces*, no coinciden con ningún gen conocido.

Longitud de espiga

Cinco y tres loci, respectivamente, mostraron asociación con la longitud de espiga en las 156 *landraces* y en la población 4. El locus Bmac040 es el único marcador común que mostró asociación con este carácter en los dos análisis. No se conoce ningún gen de este carácter que coincida con los marcadores que mostraron asociación con la longitud de espiga. Por otra parte, de los QTLs descritos en la literatura para este carácter, solamente dos están cerca de los *bins* de dos loci que presentaron asociación. Uno se localiza en el *bin* 7 del cromosoma 1H, y el otro entre los *bins* 7 y 8 del cromosoma 7H. Fueron descritos por Li *et al.* (2005), en la población Brenda x HS213. El locus HvGLUEND, que manifestó una asociación con la longitud de la espiga en el

análisis de la población 4, se encuentra en el *bin* 8 del cromosoma 1H, mientras que el locus Bmag120 que reveló asociación con este carácter en el análisis de las 156 *landraces*, se sitúa en el *bin* 9 del cromosoma 7H.

Densidad de espiga

Los loci Bmac399 y Bmag120 mostraron asociación con la densidad de espiga en el análisis de las 156 *landraces*. En el análisis de la población 4, solamente el locus Bmac040 presentó asociación con este carácter. En el cromosoma 1H, Franckowiak (1999) describió dos genes para densidad de espiga que son *dsp.ab* y *dsp.ao*. La posición exacta de estos genes es desconocida, El gen *dsp.ao* fue asignado al brazo corto del cromosoma. El locus Bmac399, se encuentra en el *bin* 4 del cromosoma 1H, por lo que podría tener relación con este gen. En el *bin* 9 del cromosoma 7H, donde se localiza el locus Bmag120, no se ha descrito ningún gen para este carácter. Sin embargo, se conoce un gen de densidad de espiga en el *bin* 6 del cromosoma 7H. Los genes *lax-c* y *dsp9* se encuentran respectivamente en los *bins* 13 y 14 del cromosoma 6H. La asociación del locus Bmac040, situado en el *bin* 13, con este carácter, podría deberse a la proximidad con respecto al gen *lax-c*, ya que se localizan en el mismo *bin*.

Espiguillas / espiga

El locus Bmac018, fue el único marcador que presentó asociación con este carácter, y sólo en el análisis de la población 4. Esta posición no coincide con ningún gen conocido ni QTL descrito en la literatura.

Tipo de espiga.

Ocho loci mostraron asociación con el tipo de espiga (2 ó 6 carreras), aunque solamente el locus MWG699 tienen relación con este carácter, por estar ligado el gen de determinación de tipo de espiga 2 ó 6 carreras *vrs1* (Komatsuda *et al.*, 1999). Este gen está distribuido de una manera muy desigual entre los grupos (Tabla 3.12 del capítulo 3). Es posible que los demás loci mostraran asociación con este carácter solamente por el hecho de que los alelos frecuentes de estos loci en las entradas de 2 y 6 carreras son distintos. Los coeficientes de determinación que presentaron estos loci en el modelo CG, locus MWG699 incluido, son muy bajos, debido al hecho de que las entradas de 2 carreras se encuentran en su mayoría en un grupo, y las entradas de 6 carreras repartidas

en otros 3 grupos. Por lo cual la variación real del carácter esta muy correlacionada con la estructura de población. En este sentido, Flint-Garcia *et al.* (2005) comentan que, en el caso de que la estructura de población explique una parte importante de la variación del carácter, los análisis de asociación estructurados tendrán poca potencia para detectar los efectos de genes individuales, como sería este caso.

Al final de este estudio, se comprobó que muchas asociaciones detectadas pueden ser reales, ya que coinciden con numerosos QTLs hallados en poblaciones biparentales, o con posiciones de genes conocidos. Estas asociaciones son las que con mayor probabilidad pueden estar indicando la presencia de QTLs para los caracteres involucrados.

Por otra parte también se observaron casos de asociaciones que no coincidían ni con QTLs ni con genes. No se puede descartar que se trate de falsos positivos. El desequilibrio de ligamiento entre loci que realmente no están ligados (loci localizados en distintos cromosomas), sería una de las causas que induce a la detección de falsos positivos. Este pareció ser el caso entre los loci MWG699 y HvBM5A así como HvM020 y Bmag223, respectivamente para la roya y longitud de los pelos de la raquilla, en la población 4, detectados gracias a la redundancia puesta de manifiesto en los análisis multilocus. También es posible que algunos de los falsos positivos fueran efecto del azar, aunque se eligió conscientemente un nivel de significación bastante estricto. Es conocido que la corrección de Bonferroni produce tests excesivamente estrictos.

La correlación entre caracteres sería también una posible causa de la detección de falsos positivos. Gupta (2002) comentó que una de las limitaciones del método de ‘*association mapping*’ es la confusión del efecto de un QTL por otros que influyen en el carácter. En cualquier caso, este tipo de falsos positivos debidos a pleiotropía afecta en la misma medida a los análisis clásicos de QTLs en poblaciones biparentales. En este estudio, este fenómeno se observó principalmente entre el espigado y rendimiento, con marcadores asociados a ambos caracteres, siempre coincidiendo los alelos que marcaban fecha de espigado más temprana con mayores rendimientos. El carácter tipo de espiga influyó también sobre muchos caracteres, pero al estar correlacionado con la estructura de población, estas asociaciones fueron anuladas en su mayoría al incluir esta estructura en los análisis. Por otra parte, la correlación entre el carácter tipo de espiga y la estructura de población, complicó la detección de la asociación entre el marcador

MWG699 y el gen *vrs1*, que presentó un coeficiente de determinación muy bajo aunque este locus fue descrito por Tanno *et al.* (2002) como un marcador ligado al gen *vrs1*. Es un caso claro de pérdida de poder estadístico al tratarse de un carácter altamente correlacionado con la estructura de la población. En este tema, Wilson *et al.* (2004), y Flint-García *et al.* (2005), comentaron que la pérdida de significación debida a la inclusión de la estructura de la población en el análisis podría tener varias causas. Un motivo posible sería que el polimorfismo no fuera funcional, y la asociación aparente estuviera solamente causada por la estructura poblacional; otro motivo podría ser que el polimorfismo tuviera un origen funcional, pero éste coincidiera con la estructura de población. En este caso, el resultado observado será un falso negativo.

Por otra parte, la aplicación de la corrección de Bonferroni para reducir los falsos positivos, podría resultar en la eliminación de algunos positivos verdaderos. Esto sería probablemente el caso del locus WMC1E8 con el espigado, que presentó un nivel de significación ligeramente más alto del nivel escogido en este análisis.

Los resultados obtenidos en este estudio son en general prometedores en cuanto a la aplicabilidad de los estudios de asociación para la detección de QTLs en cebada. Sin embargo, esta no se puede llevar a cabo sin tener en cuenta los factores citados más arriba.

El desequilibrio de ligamiento variará muy probablemente entre distintas colecciones de materiales, dependiendo de su historia y relaciones filogenéticas. Una colección de variedades locales como la estudiada aquí presentará unos niveles de desequilibrio de ligamiento seguramente menores a los de una colección de variedades modernas procedentes de programas de mejora. En este último caso, las variedades tendrán ancestros comunes a un número relativamente bajo de meiosis de distancia, lo que habrá significado menores oportunidades de recombinación. Aunque Kraakman *et al.* (2004), analizando 146 cultivares modernos de cebada de 2 carreras de primavera, con 236 AFLPs, encontraron un nivel de desequilibrio de ligamiento similar al observado en este estudio. Estos autores llevaron a cabo un estudio de ‘*association mapping*’ para los caracteres de rendimiento y estabilidad de rendimiento en este grupo de germoplasma, y encontraron que la mayoría de las asociaciones detectadas coinciden con QTLs descubiertos anteriormente en poblaciones biparentales. Posteriormente Kraakman *et al.* (2006), analizando una muestra de 148 cultivares modernos de cebada de 2 carreras de primavera, con marcadores AFLPs y microsatélites, encontraron asociaciones entre estos marcadores y caracteres agronómicos (espigado y altura de

planta), resistencia a enfermedades (roya de hojas y virus del enanismo amarillo de la cebada) y caracteres morfológicos (longitud de los pelos de la raquilla y tamaño de la lodícula). Estos autores confirmaron que la mayoría de las asociaciones detectadas coinciden con QTLs encontrados en poblaciones biparentales o con zonas de genes conocidos. Basándose en los resultados obtenidos, estos autores concluyen que el *linkage disequilibrium mapping*, es una herramienta adicional útil para la búsqueda de asociaciones entre marcadores y genes o QTLs.

Otros trabajos, tanto en especies autógamas como alógamas, confirman la eficacia del método de *linkage disequilibrium mapping*, en la búsqueda de QTLs. Breseghello y Sorrells (2006), analizaron la asociación de microsatélites localizados en los cromosomas 2D, 5A y 5B con el tamaño de grano y la calidad harinera dentro de una colección de 96 cultivares modernos de trigo blando. Aunque el desequilibrio de ligamiento observado en los cromosomas 2D y 5A, decae respectivamente a unas distancias de 1 cM y 5 cM, las asociaciones halladas por estos autores concuerdan con QTLs detectados en estudios previos en poblaciones biparentales.

En un estudio de 95 ecotipos de *Arabidopsis thaliana*, se demostró que el desequilibrio de ligamiento decae más rápido, a una distancia de 50 kb (Nordborg *et al.*, 2005). Aun así, Aranzana *et al.* (2005) examinaron la posibilidad de llevar a cabo un estudio de asociación a gran escala (*genome-wide association*) en la misma muestra, buscando asociaciones entre polimorfismos de tipo SNP y los caracteres fecha de espigado y resistencia a enfermedades, de los que se conocía el control genético. Estos autores llegaron a identificar los genes responsables, a pesar de detectar un elevado número de falsos positivos.

El primer estudio de asociación en plantas se llevó a cabo en la especie alógama maíz. Remington *et al.* (2001) identificaron una asociación entre el locus *Dwarf8* y la fecha de floración mediante asociación a gran escala con microsatélites. A su vez, Thornsberry *et al.* (2001) describieron ciertos polimorfismos dentro de ese locus que son responsables de la asociación con fecha de floración empleando un análisis de gen candidato. Otros ejemplos de la utilidad de este tipo de análisis basado en genes candidato son los de Wilson *et al.* (2004) o Yamasaki *et al.* (2005).

Aunque se trata de una metodología de análisis recientemente empleada en plantas, los resultados de los distintos estudios son muy prometedores.

6-5. Referencias bibliográficas

- Akaike H. 1974.** A new look at statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716- 723.
- Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jacob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M. 2005.** Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *Plos Genetics*, 1: 531-539.
- Ayoub M, Symons SJ, Edney MJ, Mather DE. 2002.** QTLs affecting kernel size and shape in a two-rowed by six-rowed barley cross. *Theoretical & Applied Genetics*, 105: 237-247.
- Backes G, Graner A, Foroughi-Wehr B, Fischbeck G, Wenzel G, Jahoor A. 1995.** Localization of quantitative trait loci (QTL) for agronomic important characters by the use of a RFLP map in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 90: 294-302
- Backes G, Madsen LH, Jaiser H, Stougaard J, Herz M, Mohler V, Jahoor A. 2003.** Localisation of genes for resistance against *Blumeria graminis* f.sp.*hordei* and *Puccinia graminis* in a cross between a barley cultivar and a wild barley (*Hordeum vulgare* ssp. *spontaneum*) line. *Theoretical & Applied Genetics*, 106: 353-362.
- Barr AR, Jefferies SP, Broughton S, Chalmers KJ, Kretschmer JM, Boyd WJR, Collins HM, Roumeliotis S, Logue SJ, Coventry SJ, Moody DB, Read BJ, Poulsen D, Lance RCM, Platz GJ, Park RF, Pannozo JF, Karakousis A, Lim P, Verbyla AP, Eckermann PJ. 2003a.** Mapping and QTL analysis of the barley population Alexis x Sloop. *Australian Journal of Agricultural Research*, 54: 1117-1123.
- Barr AR, Karakousis A, Lance RCM, Logue SJ, Manning S, Chalmers KJ, Kretschmer JM, Boyd WJR, Collins HM, Roumeliotis S, Coventry SJ, Moody DB, Read BJ, Poulsen D, Li CD, Platz GJ, Inkerman PA, Pannozo JF, Cullis BR, Smith AB, Lim P, Langridge P. 2003b.** Mapping and QTL analysis of the barley population Chebec x Harrington. *Australian Journal of Agricultural Research*, 54: 1125-1130.
- Barua UM, Chalmers KJ, Thomas WTB, Hackett CA. 1993.** Molecular mapping of genes determining height, time to heading, and growth habit in barley (*Hordeum vulgare*). *Genome*, 36: 1080-1087.
- Baum M, Grando S, Backes G, Jahoor A, Sabbagh A, Ceccarelli S. 2003.** QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' × H. *spontaneum* 41-1. *Theoretical & Applied Genetics*, 107: 1215-1225.
- Beer SC, Siripoonwiwat W, O'Donoghue LS, Souza E, Matthews D, Sorrells ME. 1997.** Associations between molecular markers and quantitative traits in an oat germplasm pool: Can we infer linkages?. *Journal of Agricultural Genomics*, 3, paper 1. <http://www.cabi-publishing.org/jag/papers97/paper197/indexp197html>.

- Bezant J, Laurie D, Pratchett N, Chojecki J, Kearsey M. 1996.** Marker regression mapping of QTL controlling flowering time and plant height in spring barley (*Hordeum vulgare* L.) cross. *Heredity*, 77: 64-73.
- Bezant J, Laurie D, Pratchett N, Chojecki J, Kearsey M. 1997.** Mapping QTL controlling yield and yield components in a spring barley (*Hordeum vulgare* L.) cross using marker regression. *Molecular Breeding*, 3: 29-38.
- Börner A, Buck-Sorlin GH, Hayes PM, Malyshev S, Korzun V. 2002.** Molecular mapping of major genes and quantitative trait loci determining flowering time in response to photoperiod in barley. *Plant Breeding*, 121:129-132.
- Borovkova IG, Jin Y, Steffenson BJ, Kilian A, Blake TK, Kleinhofs A. 1997.** Identification and mapping of a leaf rust resistance gene in barley line Q211861. *Genome*, 40: 236-240.
- Borovkova IG, Jin Y, Steffenson BJ. 1998.** Chromosomal location and genetic relationship of leaf rust resistance genes *Rph9* and *Rph12* in barley. *Phytopathology*, 88: 76-80.
- Boyd WJR, Li CD, Grime CR, Cakir M, Potipibool S, Kaveeta L, Men S, Jalal Kamali MR, Barr AR, Moody DB, Lance RCM, Logue SJ, Raman H, Read BJ. 2003.** Conventional and molecular genetic analysis of factors contributing to variation in the timing of heading among spring barley (*Hordeum vulgare* L.) genotypes grown over a mild winter growing season. *Australian Journal of Agricultural Research*, 54:1277-1301.
- Bozdogan H. 1987.** Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika*, 52: 345-370.
- Breseghello F, Sorrells ME. 2006.** Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*, 172: 1165-1177.
- Buckler IV ES, Thornsberry JM. 2002.** Plant molecular diversity and applications to genomics. *Current Opinion in Plant Biology*, 5: 107-111.
- Cakir M, Poulsen D, Galwey NW, Ablett GA, Chalmers KJ, Platz GJ, Park RF, Lance RCM, Panozzo JF, Read BJ, Moody DB, Barr AR, Johnston P, Li CD, Boyd WJR, Grime CR, Appels R, Jones MGK, Langridge P. 2003.** Mapping and QTL analysis of the barley population Tallon x Kaputar. *Australian Journal of Agricultural Research*, 54: 1155-1162.
- Canci PC, Nduulu LM, Dill-Macky R, Muehlbauer GJ, Rasmusson DC, Smith KP. 2003.** Genetic relationship between kernel discoloration and grain protein concentration in barley. *Crop Science*, 43: 1671-1679.
- Canci PC, Nduulu LM, Muehlbauer GJ, Dill-Macky R, Rasmusson DC, Smith KP. 2004.** Validation of quantitative trait loci for Fusarium head blight and kernel discoloration in barley. *Molecular Breeding*, 14: 91-104.
- Collins N, Park R, Spielmeier W, Ellis J, Pryor AJ. 2001.** Resistance gene analogs in barley and their relationship to rust resistance genes. *Genome*, 44: 375-381.

- Costa JM, Corey A, Hayes PM, Jobet C, Kleinhofs A, Kopisch-Obusch A, Kramer SF, Kudrna D, Li M, Riera-Lizarazu O, Sato K, Szűcs P, Toojinda T, Vales MI, Wolfe RI. 2001.** Molecular Mapping of the Oregon Wolfe Barley: A phenotypically polymorphic double-haploid population. *Theoretical & Applied Genetics*, 103: 415-424.
- Coventry SJ, Barr AR, Eglinton JK, McDonald GK. 2003.** The determinants and genome locations influencing grain weight and size in barley (*Hordeum vulgare* L.). *Australian Journal of Agricultural Research*, 54: 1103-1115.
- de la Pena RC, Smith KP, Capettini F, Muehlbauer GJ, Gallo-Meagher M, Dill-Macky R, Somers DA, Rasmusson DC. 1999.** Quantitative trait loci associated with resistance to Fusarium head blight and kernel discoloration in barley. *Theoretical & Applied Genetics*, 99: 561-569.
- Devlin B, Roeder K. 1999.** Genomic control for association studies. *Biometrics*, 55: 997-1004.
- Drescher A, Ivandic V, Walther U, Graner A. 2000.** High resolution mapping of the *Rph16* locus in barley. *In Barley Genetics 8, 'Proceedings of the 8th International Barley Genetics Symposium'*. Ed Logue S. Adelaide University. Volume II: 95-97
- Ellis RP, Forster BP, Gordon DC, Handley LL, Keith RP, Lawrence P, Meyer R, Powell W, Robinson D, Scrimgeour CM, Young G, Thomas WTB. 2002.** Phenotype/genotype associations for yield and salt tolerance in a barley mapping population segregating for two dwarfing genes. *Journal of Experimental Botany*, 53: 1163-1176.
- Emebiri LC, Moody DB, Panozzo JF, Chalmers KJ, Kretschmer JM, Ablett GA. 2003.** Identification of QTLs associated with variations in grain protein concentration in two-row barley. *Australian Journal of Agricultural Research*, 54: 1211-1221.
- Francia E, Rizza F, Cattivelli L, Stanca AM, Galiba G, Toth B, Hayes PM, Skinner JS, Pecchioni N. 2004.** Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) x 'Tremois' (spring) barley map. *Theoretical & Applied Genetics*, 108: 670-680.
- Flint-Garcia SA, Thornsberry JM, Buckler IV ES. 2003.** Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, 54: 357-374.
- Flint-Garcia SA, Thuillet AC, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES. 2005.** Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant Journal*, 44: 1054-1064.
- Franckowiak JD. 1997.** New and revised descriptions of barley genes. *Barley Genetics Newsletter*, 26: 44, 73, 106, 269.
- Franckowiak JD. 1999.** Semidwarf genes. *Barley Genetics Newsletter*, 29: 74-79.
- Franckowiak JD. 2002.** New and revised Barley Genetic Stock descriptions. *Barley Genetics Newsletter*, 32:109.

- Franckowiak JD , Chicaiza O. 1998.** BGS 96. Barley Genetics Newsletter, 28: 96.
- Franckowiak JD, Gallager LW. 1997.** New and revised descriptions of barley genes. Barley Genetics Newsletter, 26: 233.
- Franckowiak JD, Jin Y. 1997.** New and revised descriptions of barley genes. Barley Genetics Newsletter, 26: 107.
- Franckowiak JD, Konishi T. 1997.** New and revised descriptions of barley genes. Barley Genetics Newsletter, 26: 53, 189, 202.
- Franckowiak JD, Konishi T. 2002.** New and revised Barley Genetic Stock descriptions. Barley Genetics Newsletter, 32: 86-87
- Franckowiak JD, Lundqvist U, 1997.** New and revised descriptions of barley genes. Barley Genetics Newsletter, 26: 93, 118.
- Franckowiak JD, Lundqvist U, Konishi T, Gallagher LW. 1997.** New and revised descriptions of barley genes. Barley Genetics Newsletter, 26: 213-215.
- Garvin DF, Miller-Garvin JE, Viccars EA, Jacobsen JV, Brown AHD. 1998.** Identification of molecular markers linked to *ant28-484*, a mutation that eliminates proanthocyanidin production in barley seeds. Crop Science, 38: 1250-1255
- Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schüler K. 2004.** Assessing genetic potential in germplasm collections of crop plants by marker-trait association: a case study for potatoes with quantitative variation of resistance to late blight and maturity type. Molecular Breeding, 13: 93-102.
- Gupta PK. 2002.** Molecular markers and QTL analysis in crop plants. Current Science, 83: 113-114.
- Gupta PK, Rustgi S, Kulwal PL. 2005.** Linkage disequilibrium and association studies in higher plants: Present status and future prospects. Plant Molecular Biology, 57: 461-485.
- Hannan EJ, Quinn BG. 1979.** The determination of the order of an autoregression. Journal of the Royal Statistical Society, Series B 41: 190-195.
- Hayes PM, Liu BH, Knapp F, Chen F, Jones B, Blake T, Franckowiak J, Rasmusson D, Sorrells M, Ullrich SE, Wesenberg D, Kleinhofs A. 1993.** Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. Theoretical & Applied Genetics, 87: 392-401.
- Hayes P, Prehn D, Vivar H, Blake T, Comeau A, Henry I, Johnston M, Jones B, Steffenson B, St Pierre CA, Chen F. 1996.** Multiple disease resistance loci and their relationship to agronomic and quality loci in a spring barley population. Journal of Agricultural Genomics, 2, paper 2. [http://www.cabi-publishing.org/jag/papers96/paper296/index 296.html](http://www.cabi-publishing.org/jag/papers96/paper296/index%20296.html).
- Hayes PM, Ceroni J, Witsenhoer H, Kuiper M, Zabeau M, Sato K, Kleinhofs A, Kudrna D, Kilian A, Saghai-Marouf M, Hoffman D, The North American Barley Genome Mapping Project. 1997.** Characterizing and exploiting genetic diversity and quantitative traits in barley (*Hordeum vulgare*) using AFLP markers.

Journal of Agricultural Genomics, 3, paper 2. [http://www.cabi-publishing.org/gateways/jag/papers97/paper297/index 297.html](http://www.cabi-publishing.org/gateways/jag/papers97/paper297/index%20297.html).

- Hurvich CM, Tsai CL. 1989.** Regression and time series model selection in small samples. *Biometrika*, 76: 297-307.
- Igartua E, Casas AM, Ciudad F, Montoya JL, Romagosa I. 1999.** RFLP markers associated with major genes controlling heading date evaluated in a barley germ plasm pool. *Heredity*, 83: 551-559.
- Ivandic V, Walther U, Graner A. 1998.** Molecular mapping of a new gene in wild barley conferring complete resistance to leaf rust (*Puccinia hordei* Otth). *Theoretical & Applied Genetics*, 97: 1235-1239.
- Ivandic V, Thomas WTB, Nevo E, Zhang Z, Forster BP. 2003.** Associations of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. *Plant Breeding*, 122: 300-304
- IPGRI. 1994.** Descriptors for barley (*Hordeum vulgare* L.). International Plant Genetic Resources Institute, Rome, Italy, 45 p.
- Jende-Strid B. 1998.** Proanthocyanidin-free 28, *ant28*. *Barley Genetics Newsletter*, 29: 97.
- Karsai I, Szűcs P, Mészáros K, Filichkina T, Hayes PM, Skinner JS, Láng L, Bedő Z. 2005.** The *Vrn-H2* locus is a major determinant of flowering time in a facultative x winter growth habit barley (*Hordeum vulgare* L.) mapping population. *Theoretical & Applied Genetics*, 110: 1458–1466.
- Karakousis A, Barr AR, Kretschmer JM, Manning S, Logue SJ, Roumeliotis S, Collins HM, Chalmers KJ, Li CD, Lance RCM, Langridge P. 2003a.** Mapping and QTL analysis of the barley population Galleon x Haruna Nijo. *Australian Journal of Agricultural Research*, 54: 1131-1135.
- Karakousis A, Barr AR, Chalmers KJ, Ablet GA, Holton TA, Henry RJ, Lim P, Langridge P. 2003b.** Potential of SSR markers of plant breeding and variety identification in Australian barley germplasm. *Australian Journal of Agricultural Research*, 54: 1197-1210.
- Kicherer S, Backes G, Walther U, Jahoor A. 2000.** Localising QTLs for leaf rust resistance and agronomic traits in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 100: 881-888.
- Kjær B, Jensen J. 1996.** Quantitative trait loci for grain yield and yield components in a cross between a six-rowed and a two-rowed barley. *Euphytica*, 90: 39-48.
- Kjær B, Jensen J, Giese H. 1995.** Quantitative trait loci for heading date and straw characters in barley. *Genome*, 38: 1098-1104.
- Kleinhofs A, Han F. 2002.** Molecular mapping of the barley genome. *In* *Barley Science. Recent Advances from Molecular Biology to Agronomy of Yield and Quality*. Ed. Slafer GA, Molina-Cano JL, Savin R, Araus JL, Romagosa I. Haworth Press, New York, 31-63.

- Komatsuda T, Li W, Takaiwa F, Oka S. 1999.** High resolution map around the *vrs1* locus controlling two- and six-rowed spike in barley, *Hordeum vulgare*. *Genome*, 42: 248-253.
- Konishi T, Franckowiak JD. 1997.** New and revised descriptions of barley genes. *Barley Genetics Newsletter*, 26: 188, 239.
- Kraakman ATW, Niks RE, van der Berg P, Stam P, van Eeuwijk FA. 2004.** Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, 168: 435-446.
- Kraakman ATW, Martínez F, Mussiraliev B, van Eeuwijk FA, Niks RE. 2006.** Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. *Molecular Breeding*, 17: 41-58.
- Laurie DA, Pratchett N, Romero C, Simpson E, Snape JW. 1993.** Assignment of the *denso* dwarfing gene to the long arm of chromosome 3 (3H) of barley by use of RFLP Markers. *Plant Breeding*, 111: 198-203.
- Laurie DA, Pratchett N, Bezant JH, Snape JW. 1994.** Genetic analysis of a photoperiod response gene on the short arm of chromosome 2 (2H) of *Hordeum vulgare* (barley). *Heredity*, 72: 619-627.
- Laurie DA, Pratchett N, Bezant JH, Snape JW. 1995.** RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter x spring barley (*Hordeum vulgare* L.). *Genome*, 38: 575-585.
- Li JZ, Huang XQ, Heinrichs F, Ganai MW, Röder MS. 2005.** Analysis of QTLs for yield, yield components and malting quality in a BC₃-DH population of spring barley. *Theoretical & Applied Genetics*, 110: 356-363.
- Long NR, Jefferies SP, Warner P, Karakousis A, Kretschmer JM, Hunt C, Lim P, Eckermann PJ, Barr AR. 2003.** Mapping and QTL analysis of the barley population Mundah x Keel. *Australian Journal of Agricultural Research*, 54: 1163-1171.
- Lundqvist U, Franckowiak JD. 1997.** New and revised descriptions of barley genes. *Barley Genetics Newsletter*, 26: 244.
- Ma Z, Steffenson BJ, Prom LK, Lapitan NLV. 2000.** Mapping of quantitative trait loci for *Fusarium* head blight resistance in barley. *Phytopathology*, 90: 1079-1088.
- Manzjuk VT, Barsukov PN. 1974.** Genetic studies of some quantitative characters of barley. *Barley Genetics Newsletter*, 4: 48-49.
- Marquez-Cedillo LA, Hayes PM, Jones BL, Kleinhofs A, Legge WG, Rossnagel BG, Sato K, Ullrich SE, Wesenberg DM, the North American Barley Genome Mapping Project. 2000.** QTL analysis of malting quality in barley based on the doubled-haploid progeny of two elite North American varieties representing different germplasm groups. *Theoretical & Applied Genetics*, 101: 173-184.
- Marquez-Cedillo LA, Hayes PM, Kleinhofs A, Legge WG, Rossnagel BG, Sato K, Ullrich SE, Wesenberg DM, the North American Barley Genome Mapping**

- Project. 2001.** QTL analysis of agronomic traits in barley based on the doubled haploid progeny of two elite North American varieties representing different germplasm groups. *Theoretical & Applied Genetics*, 103: 625-637.
- Mather DE, Tinker NA, LaBerge DE, Edney M, Jones BL, Rossnagel BG, Legge WG, Briggs KG, Irvine RB, Falk DE, Kasha KJ. 1997.** Regions of the genome that affect grain and malt in a North American two-row barley cross. *Crop Science*, 37: 544-554.
- Matus I, Corey A, Filichkin T, Hayes PM, Vales MI, Kling J, Riera-Lizarazu O, Sato K, Powell W, Waugh R. 2003.** Development and characterization of recombinant chromosome substitution lines (RCSLs) using *Hordeum vulgare* subsp. *spontaneum* as a source of donor alleles in a *Hordeum vulgare* subsp. *vulgare* background. *Genome*, 46: 1010–1023.
- McDaniel ME, Hathcock BR. 1969.** Linkage of Pa4 and Ml-a loci in barley. *Crop Science*, 9: 822.
- Mesfin A, Smith KP, Dill-Macky R, Evans CK, Waugh R, Gustus CD, Muehlbauer GJ. 2003.** Quantitative trait loci for Fusarium head blight resistance in barley detected in a two-rowed by six-rowed population. *Crop Science*, 43: 307-318.
- Moralejo M, Swanston JS, Muñoz P, Prada D, Elía M, Russell JR, Ramsay L, Cistué L, Codesal P, Casas AM, Romagosa I, Powell W, Molina-Cano JL. 2004.** Use of new EST markers to elucidate the genetic differences in grain protein content between European and North American two-rowed malting barleys. *Theoretical & Applied Genetics*, 110: 116-125.
- Nilan RA. 1964.** The cytology and genetics of barley 1951-1962. Monographic supplement, Washington State University, Volume 32: 93-176.
- Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D. 2002.** The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, 30: 190-193.
- Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R, Jakobsson M, Kim S, Morozov Y, Padhukasahasram B, Plagnol V, Rosenberg NA, Shah C, Wall JD, Wang J, Zhao K, Kalbfleisch T, Schulz V, Kreitman M, Bergelson J. 2005.** The pattern of polymorphism in *Arabidopsis thaliana*. *Plos Biology*, 3: 1289-1299.
- Pallotta MA, Asayama S, Reinheimer JM, Davies PA, Barr AR, Jefferies SP, Chalmers KJ, Lewis J, Collins HM, Roumeliotis S, Logue SJ, Coventry SJ, Lance RCM, Karakousis A, Lim P, Verbyla AP, Eckerman PJ. 2003.** Mapping and QTL analysis of the barley population Amagi Nijo x WI2585. *Australian Journal of Agricultural Research*, 54: 1141-1144.
- Pan A, Hayes PM, Chen F, Chen THH, Blake T, Wright S, Karsai I, Bedö Z. 1994.** Genetic analysis of the components of winterhardiness in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 89: 900-910.

- Park RF, Poulsen D, Barr AR, Cakir M, Moody DB, Raman H, Read BJ. 2003.** Mapping genes for resistance to *Puccinia hordei* in barley. *Australian Journal of Agricultural Research*, 54: 1323-1333.
- Paterson SH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD. 1988.** Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment polymorphism. *Nature*, 225: 721-726.
- Pickering RA, Steffenson BJ, Hill AM, Borovkova I. 1998.** Association of leaf rust and powdery mildew resistance in a recombinant derived from a *Hordeum vulgare* x *Hordeum bulbosum* hybrid. *Plant Breeding*, 117: 83-84.
- Pillen K, Zacharias A, Léon J. 2003.** Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 107: 340-352.
- Pillen K, Zacharias A, Léon J. 2004.** Comparative AB-QTL analysis in barley using a single exotic donor of *Hordeum vulgare* ssp. *spontaneum*. *Theoretical & Applied Genetics*, 108: 1591-1601.
- Powell W, Thomas WTB, Baird E, Lawrence P, Booth A, Harrower JW, McNicol JW, Waugh R. 1997.** Analysis of quantitative traits in barley by the use of amplified fragment length polymorphisms. *Heredity*, 79: 58-59.
- Prioul JL, Pelleschi S, Sene M, Thevenot C, Causse M, de Vienne D, Leonardi A. 1999.** From QTLs for enzyme activity to candidate genes in maize. *Journal of Experimental Botany*, 50: 1281-1288.
- Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959.
- Qi X, Niks RE, Stam P, Lindhout P. 1998.** Identification of QTLs for partial resistance to leaf rust (*Puccinia hordei*) in barley. *Theoretical & Applied Genetics*, 96: 1205-1215.
- Rafalski A. 2002.** Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology*, 5: 94-100.
- Rajasekaran P, Thomas WTB, Wilson A, Lawrence P, Young G, Ellis RP. 2004.** Genetic control over grain damage in a spring barley mapping population. *Plant Breeding*, 123: 17-23.
- Read BJ, Raman H, McMichael G, Chalmers KJ, Ablett GA, Platz GJ, Raman R, Genger RK, Boyd WJR, Li CD, Grime CR, Park RF, Wallwork H, Prangnell R, Lance RCM. 2003.** Mapping and QTL analysis of the barley population Sloop x Halcyon. *Australian Journal of Agricultural Research*, 54: 1145-1153.
- Reinheimer JL, Barr AR, Eglinton JK. 2004.** QTL mapping of chromosomal regions conferring reproductive frost tolerance in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 109: 1267-1274.
- Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler IV ES. 2001.** Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences, USA*, 98: 11479-11484.

- Roane CW, Starling TM. 1989.** Linkage studies with genes conditioning leaf rust reaction in barley. *Barley Newsletter*, 33:190-192
- Sackville Hamilton NR, Skøt L, Chorlton KH, Thomas ID, Mizen S. 2002.** Molecular genecology of temperature response in *Lolium perenne*: 1. preliminary analysis to reduce false positives. *Molecular Ecology*, 11: 1855-1863.
- SAS Institute. 1988.** SAS/STAT User's Guide, Release 6.03. Edition. Cary, NC; SAS Institute Inc., 1028p
- Scheurer KS, Friedt W, Huth W, Waugh R, Ordon F. 2001.** QTL analysis of tolerance to a German strain of BYDV-PAV in barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 103: 1074-1083.
- Schwarz G. 1978.** Estimating the dimension of a model. *Annals of Statistics*, 6: 461-464.
- See D, Kanazin V, Kephart K, Blake T. 2002.** Mapping genes controlling variation in barley grain protein concentration. *Crop Science*, 42: 680-685.
- Skøt L, Humphreys MO, Armstead I, Heywood S, Skøt KP, Sanderson R, Thomas ID, Chorlton KH, Sackville Hamilton NR. 2005.** An association mapping approach to identify flowering time genes in natural populations of *Lolium perenne* (L.). *Molecular Breeding*, 15: 233-245.
- Spaner D, Shugar LP, Choo TM, Falak I, Briggs KG, Legge WG, Falk DE, Ullrich SE, Tinker NA, Steffenson BJ, Mather DE. 1998.** Mapping of disease resistance loci in barley on the basis of visual assessment of naturally occurring symptoms. *Crop Science*, 38: 843-850.
- Stracke S, Börner A. 1998.** Molecular mapping of the photoperiod response gene *ea7* in barley. *Theoretical & Applied Genetics*, 97: 797-800.
- Takahashi R, Yasuda S. 1970.** Genetics of earliness and growth habit in barley. *Barley Genetics II*. 'Proceedings of the 2nd International Barley Genetics Symposium'. Ed. Nilan RA. Washington State University Press, Pullman, 388-408.
- Tanno K, Taketa S, Takeda K, Komatsuda T. 2002.** A DNA marker closely linked to the *vrs 1* locus (row type gene) indicates multiple origins of six rowed cultivated barley. (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 104: 54-60.
- Teulat B, Merah O, Souyris I, This D. 2001.** QTLs for agronomic traits from a Mediterranean barley progeny grown in several environments. *Theoretical & Applied Genetics*, 103: 774-787.
- Thomas WTB, Powell W, Waugh R, Chalmers KJ, Barua UM, Jack P, Lea V, Forster BP, Swanston JS, Ellis RP, Hanson PR, Lance RCM. 1995.** Detection of quantitative trait loci for agronomic, yield, grain and disease characters in spring barley (*Hordeum vulgare* L.). *Theoretical & Applied Genetics*, 91: 1037-1047.
- Thomas WTB, Baird E, Fuller JD, Lawrence P, Young GR, Russell J, Ramsay L, Waugh R, Powell W. 1998.** Identification of a QTL decreasing yield in barley linked *Mlo* powdery mildew resistance. *Molecular Breeding*, 4: 381-398.

- Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler IV ES. 2001.** *Dwarf8* polymorphisms associate with variation in flowering time. *Nature Genetics*, 28: 286-289.
- Tinker NA, Mather DE, Rossnagel BG, Kasha KJ, Kleinhofs A, Hayes PM, Falk DE, Ferguson T, Shugar LP, Legge WG, Irvine RB, Choo TM, Briggs KG, Ullrich SE, Franckowiak JD, Blake TK, Graf RJ, Dofing SM, Saghai-Marooof MA, Scoles GJ, Hoffman D, Dahleen LS, Kilian A, Chen F, Biyashev RM, Kudrna DA, Steffenson BJ. 1996.** Regions of the genome that affect agronomic performance in two-row barley. *Crop Science*, 36: 1053-1062.
- Toojinda T, Broers LH, Chen XM, Hayes PM, Kleinhofs A, Korte J, Kudrna D, Leung H, Line RF, Powell W, Ramsay L, Vivar H, Waugh R. 2000.** Mapping quantitative and qualitative disease resistance genes in a doubled haploid population of barley (*Hordeum vulgare*). *Theoretical & Applied Genetics*, 101: 580-589.
- Turpeinen T, Vanhala T, Nevo E, Nissilä E. 2003.** AFLP genetic polymorphism in wild barley (*Hordeum spontaneum*) populations in Israel. *Theoretical & Applied Genetics*, 106: 1333-1339.
- Varshney RK, Graner A, Sorrells ME. 2005.** Genomics-assisted breeding for crop improvement. *Trends in Plant Science*, 10: 621-630.
- Virk PS, Ford-Lloyd BV, Jackson MT, Pooni HS, Clemeno TP, Newbury HJ. 1996.** Predicting quantitative variation within rice germplasm using molecular markers. *Heredity*, 76: 296-304.
- Voltas J, van Eeuwijk FA, Sombrero A, Lafarga A, Igartua E, Romagosa I. 1999.** Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley. I. Individual grain weight. *Field Crops Research*, 62: 63-74.
- von Zitzewitz J, Szűcs P, Dubcovsky J, Yan L, Francia E, Pecchioni N, Casas A, Chen THH, Hayes PM, Skinner JS. 2005.** Molecular and structural characterization of barley vernalization genes. *Plant Molecular Biology*, 59: 449-467.
- Wilson LM, Whitt SR, Ibañez AM, Rocheford TR, Goodman MM, Buckler IV ES. 2004.** Dissection of maize kernel composition and starch production by candidate gene association. *The Plant Cell*, 16: 2719-2733.
- Yamasaki M, Tenailon MI, Bi IV, Schroeder SG, Sánchez-Villeda H, Doebley JF, Gaut BS, McMullen MD. 2005.** A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. *The Plant Cell*, 17: 2859-2872.
- Yin X, Stam P, Johan Dourleijn C, Kropff MJ. 1999.** AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. *Theoretical & Applied Genetics*, 99: 244-253.
- Yin X, Chasalow SD, Stam P, Kropff MJ, Dourleijn CJ, Bos I, Bindraban PS. 2002.** Use of component analysis in QTL mapping of complex traits: a case study on yield in barley. *Plant Breeding*, 121: 314-319.

- Yin X, Struik PC, van Eeuwijk FA, Stam P, Tang J. 2005.** QTL analysis and QTL-based prediction of flowering phenology in recombinant inbred lines of barley. *Journal of Experimental Botany*, 56: 967-976.
- Zhu H, Briceño G, Dovel R, Hayes PM, Liu BH, Liu CT, Ullrich SE. 1999a.** Molecular breeding for grain yield in barley: an evaluation of QTL effects in a spring barley cross. *Theoretical & Applied Genetics*, 98: 772-779.
- Zhu H, Gilchrist L, Hayes PM, Kleinhofs A, Kudrna D, Liu Z, Prom L, Steffenson B, Toojinda T, Vivar H. 1999b.** Does function follow form? Principal QTLs for *Fusarium* head blight (FHB) resistance are coincident with QTLs for inflorescence traits and plant height in a doubled-haploid population of barley. *Theoretical & Applied Genetics*, 99: 1221-1232.

Capítulo 7

Discusión general

7-1. Discusión general

7-1-1. Diversidad y adaptación

La colección nuclear española de cebada presenta una diversidad elevada, comparable a la encontrada por Struss y Plieske (1998) en una colección mundial. Los resultados de este estudio resaltaron la singularidad del germoplasma español en comparación con el europeo. Aunque se encontraron alelos específicos en los dos tipos de germoplasmas empleados, éstos fueron más numerosos en la colección nuclear española que en los genotipos de referencia. Los alelos específicos aparecieron especialmente (184 de 669 alelos detectados en la totalidad de los genotipos analizados) en las entradas de 6 carreras que son mayoritarias en esta colección. Este resultado reafirma la singularidad del germoplasma español, que ya fue comentada por Casas *et al.* (1998), al comparar mediante RFLPs un número mucho menor de variedades españolas antiguas de 6 carreras con otro germoplasma europeo. Las divergencias que existen entre el material español y el europeo apuntan a la posibilidad de que los dos grupos de germoplasma tengan orígenes distintos, o se hayan separado en un periodo relativamente lejano.

El germoplasma español no es homogéneo. En él parecen coexistir dos orígenes distintos, cuyos representantes no están totalmente mezclados, por lo que son aún distinguibles. Es especialmente destacable que el programa STRUCTURE separe el material español en dos poblaciones mayoritarias, antes de dividir los grandes grupos de germoplasma de cebada (2/6 carreras, invierno/primavera). Esto induce a pensar que esa divergencia genética entre los grupos españoles es de una magnitud notable. La división dentro del material español ocurre fundamentalmente entre las entradas de la Meseta y las del conjunto Sur-Levante-Valle del Ebro. Analizando estos grupos con el STS MWG699, propuesto por Tanno *et al.* (2002) como marcador de domesticación de la cebada, se confirmó que este marcador también tiene una distribución muy marcada por la separación de los dos grupos. La mayoría de las entradas de la Meseta son de haplotipo D para MWG699, contrariamente al grupo del Sur, Levante y Valle del Ebro, que son mayoritariamente de haplotipo A (Tabla 3.12, Figura 4.2A.). Es también destacable que, pese a no haber incluido MWG699 en los análisis de cluster para diferenciar los grupos genotípicos, este marcador presenta la distribución más desigual entre los dos principales grupos españoles (poblaciones 3 y 4). Tanno *et al.* (2002)

comentaron que el haplotipo A tiene una distribución geográfica amplia, mientras que el D se encuentra exclusivamente en la región Mediterránea. También indicaron que el haplotipo A se encuentra en genotipos de *H. spontaneum* originarios de Turkmenistán mientras que el haplotipo D se localiza también en algunas cebadas cultivadas de 2 carreras del norte de África y entradas de *H. spontaneum* de Marruecos. En nuestro caso, hemos detectado el haplotipo D en variedades de invierno de toda Europa, mientras que el A se presenta sobre todo en España y en variedades europeas de primavera, especialmente las escandinavas (Casas *et al.* 2005). Anteriormente Molina-Cano *et al.* (1999) y Zohary (1999) comentaron la posibilidad de que Marruecos pudiera ser un centro independiente de domesticación o un centro de diversificación de la cebada, además del centro reconocido del Creciente Fértil. Todo esto sugiere la posible existencia de orígenes distintos en las cebadas españolas. Luego su distribución en la Península Ibérica parece que siguió fuerzas de adaptación a las condiciones agroecológicas. Un análisis de asociación entre el polimorfismo observado en la colección nuclear y las variables eco-geográficas de las localidades de origen confirmó que la distribución de la diversidad genética observada en esta colección sigue los gradientes de estas variables. La mayoría de los loci que mostraron un número alto de asociaciones con las variables eco-geográficas coinciden con los loci que separan los dos principales grupos de genotipos de 6 carreras españolas (poblaciones 3 y 4, Tabla 4.2, capítulo 4). Un caso especialmente llamativo es el del locus HvBM5A, marcador candidato del gen de vernalización *Vrn-H1* (5H). La distribución desequilibrada que mostraron los alelos de este locus con los del STS MWG699 es muy acusada. Esto parece indicar, por un lado, que las cebadas de los dos orígenes distintos portaban alelos de vernalización diferentes en *Vrn-H1* y, por otro lado, que durante el proceso de domesticación de la cebada, el requerimiento de vernalización fue uno de los determinantes de la distribución de los genotipos en la Península Ibérica.

7-1-2. Relación entre la diversidad genética y comportamiento agronómico

La evaluación agronómica de la colección nuclear española mostró la existencia de una gran diversidad en caracteres agronómicos de todo tipo, en consonancia con la notable diversidad detectada a nivel molecular. También se observó que las entradas españolas se comportan, en general, de forma distinta a las variedades europeas. Los análisis de varianza mostraron claramente que las diferencias observadas entre los

genotipos ocurren principalmente entre los grupos 1 a 4 (entradas de la colección nuclear) y los grupos 5 a 7 (variedades de referencia, mayoritariamente europeas), resaltando en general la superioridad de las variedades comerciales. Lo que es lógico debido a la presencia de una combinación de caracteres agronómicos deseables en las variedades comerciales, como resultado del proceso de mejora.

El comportamiento de los genotipos no fue similar en los diferentes ensayos. Se observó la existencia de una interacción genotipo x ambiente notable para el carácter rendimiento. Esta interacción se debió principalmente al hecho de que estos genotipos son de diferentes grupos de germoplasma, lo que coincide con las conclusiones de Ceccarelli y Grando (1991) y Dofing *et al.* (1992), al comparar variedades locales con variedades modernas o comparar grupos de germoplasma provenientes de zonas climáticas divergentes. En este estudio se observó que las variedades comerciales son mejores en ambientes de alta productividad, tales como BEL02 y ART04, pero son inestables. La superioridad relativa de las *landraces* en ambientes de baja productividad (VED03, SAD03) fue notable (Tabla 5.6, capítulo 5). Este fenómeno fue también observado por Voltas *et al.* (1999), al comparar el cultivar autóctono español Albacete con variedades de seis carreras de alto rendimiento con amplia adaptación, tales como Barbarrosa o Plaisant. Estos fenómenos probablemente se deban a la presencia de caracteres de adaptación en las *landraces*. Más arriba se ha discutido la posible influencia del requerimiento de vernalización en la distribución de la cebada en España. Este factor podría tener también su reflejo en el rendimiento. Otro factor de adaptación que ha sido observado en estudios de este tipo es el estrés hídrico, al cual las *landraces* suelen ser más tolerantes. van Oosterom *et al.* (1993) explicaron que la estabilidad del rendimiento en ambientes de baja productividad, causada por estrés hídrico y térmico, la alcanzan aquellas variedades que presentan una insensibilidad a las fluctuaciones ambientales.

La divergencia genética observada entre los grupos 3 y 4 de la colección nuclear se reflejó también en su comportamiento agronómico. Estos dos grupos muestran una interacción de tipo *cross-over*. Las entradas del grupo 3 (mayoritariamente de la Meseta) rindieron relativamente mejor en los ensayos realizados en esta región, VLD02 y VLD03 (Figura 5.5 y 5.6, capítulo 5), y mejor en términos absolutos que las entradas del grupo 4. Por el contrario, las entradas del Sur, Levante y Valle del Ebro fueron más productivas que las del grupo 3 en los ensayos SAD02, SAD03, SAD04 y VED03, todos ellos en el Valle del Ebro. Esto indica que las entradas parecen estar mejor

adaptadas a sus regiones de origen. En este sentido, Zhang *et al.* (1993) propusieron la hipótesis de que durante la domesticación y la adaptación subsiguiente habría ocurrido una reorganización del genoma en haplotipos multilocus adaptados a distintos ambientes, hipótesis que fue posteriormente apoyada por Allard (1999). La situación que hemos descrito en este trabajo concuerda plenamente con esta hipótesis.

7-1-3. Utilidad de los recursos fitogenéticos de la colección nuclear española.

La diversidad genética observada en la colección nuclear española y las singularidades que presenta parecen indicar que ésta podría ser una nueva fuente de alelos para la mejora de la cebada. Las entradas del grupo 3 mostraron una adaptabilidad en su ambiente (VLD02, VLD03) que se tradujo en unos rendimientos superiores o iguales a los de las variedades comerciales. Efectivamente, en el ensayo de VLD02, que se caracterizó por unos promedios de rendimientos altos, las *landraces* de este grupo mostraron un promedio de rendimiento de 40,82 q. ha⁻¹, que no se distingue del promedio de las variedades comerciales en este ensayo (40,75 q. ha⁻¹). Sin embargo en el ensayo de VLD03, donde las condiciones de productividad fueron desfavorables, las *landraces* de este grupo superaron a las variedades comerciales (28,25 vs 22,47 q. ha⁻¹). Lo mismo ocurrió con las *landraces* del grupo 4 en los ensayos de SAD02, SAD03, SAD04 y VED03. En estos ensayos, el estrés hídrico parece ser la causa más probable de la caída de los rendimientos de las variedades comerciales, mientras que las *landraces* parecen tener más tolerancia a este estrés. Las adaptaciones específicas que mostraron estos dos grupos de *landraces* (3 y 4) concuerda con las afirmaciones de Ceccarelli *et al.* (1987), que comentaron que no existe una adaptación amplia para diferentes ambientes macro-agroecológicos. Por lo cual estas adaptaciones diferenciales se deben tener en cuenta a la hora de integrar este material en un programa de mejora. Ceccarelli (1996) comentó que cuando la interacción genotipo x ambiente es importante, la selección debe ser enfocada hacia las adaptaciones específicas. Las *landraces* del grupo 4 serían probablemente más apropiadas para la selección para condiciones desfavorables causadas por estrés hídrico. Ceccarelli *et al.* (1991) demostraron la estabilidad de rendimientos, en condiciones desfavorables, de líneas seleccionadas a partir de variedades tradicionales. Por otra parte, el empleo de las entradas del grupo 3 sería probablemente más adecuado para la obtención de nuevas variedades, destinadas a zonas de la Meseta.

Las *landraces* del grupo 1 son las que peor rendimiento mostraron, pero parecen más productivas en biomasa vegetativa. Por el aspecto que presentan se puede pensar que estos genotipos fueran de uso mixto o forrajero. Su aprovechamiento en este sentido parece interesante.

Algunas de las *landraces*, la mayoría del grupo 4, mostraron un cierto nivel de resistencia a las enfermedades (Figura 5.3). Los ambientes más cálidos, de inviernos más suaves y, en algunos casos, más húmedos que caracterizan a las regiones de origen de los genotipos del grupo 4 hacen sospechar que sean zonas más favorables a la actividad de los patógenos fúngicos. Parece que algunas de las entradas del grupo 4 desarrollaron una resistencia a las enfermedades provocadas por hongos. Por tanto, estas entradas podrían ser una fuente de alelos de resistencia que se podrán introducir en programas de mejora.

7-1-4. Desequilibrio de ligamiento y mapeo por asociación

En el presente estudio se ha observado que en general el desequilibrio de ligamiento en la cebada parece extenderse hasta una distancia de 10 cM, lo que coincide con los resultados de Kraakman *et al.* (2004). Este grado de desequilibrio de ligamiento es prometedor para llevar a cabo un análisis de mapeo por asociación a escala del genoma (*Genome-wide association mapping*). Estudios llevados a cabo en especies donde el desequilibrio de ligamiento no está tan extendido, como el maíz (Remington *et al.*, 2001, Wilson *et al.*, 2004) o el trigo Breseghello y Sorrells (2006), mostraron resultados interesantes en cuanto a la eficacia del método de mapeo por asociación. Por tanto, cabe esperar unos resultados también positivos en la cebada. El obstáculo mayor que puede complicar la aplicación de este método es la presencia de una estructura de población, que resultara en la detección de muchos falsos positivos. Este fue el caso en el presente estudio. Los resultados mostraron claramente que el número de posibles falsos positivos que se detectan al no incluir la estructura de población en los análisis es elevado (Tabla 6.2, capítulo 6). En este estudio se detectó un elevado desequilibrio de ligamiento entre loci de distintos cromosomas. Por tanto, a menos que se controle la estructura de población, en estudios como este aparecerán asociaciones entre QTLs y loci no ligados. Por otra parte la inclusión de la estructura de población en los análisis puede engendrar también casos de falsos negativos, cuando el polimorfismo observado en el carácter estudiado coincide con la estructura de población (Wilson *et al.*, 2004, Flint-Garcia *et al.*, 2005). En el presente estudio, se observó este fenómeno para el

carácter tipo de espiga, donde el locus MWG699 ligado al gen *vrs1* mostró un coeficiente de determinación muy bajo, por el hecho que las entradas de 2 y 6 carreras se distribuyen en diferentes poblaciones. Remington *et al.* (2001) recomendaron la evaluación de las asociaciones en un grupo de germoplasma que no presente estructura de población. Siguiendo estas recomendaciones, se evaluaron las asociaciones en la población 4, donde subsiste una estructura de sub-población poco importante. Estos resultados confirmaron muchas de las asociaciones detectadas en la totalidad de los genotipos analizados, pero también aparecieron nuevas asociaciones. Esto es debido al hecho de que las frecuencias alélicas son distintas en las diferentes poblaciones, lo que hace que alelos no considerados en la totalidad de los genotipos, por no llegar al nivel de frecuencia del 5% escogido, si entraran en el análisis de la población 4.

Empleando un número relativamente alto de marcadores en este tipo de análisis, cabe esperar que se detectan falsos positivos por el azar (error de tipo I). Por eso la aplicación de un análisis de permutación o una corrección de Bonferroni (la empleada en este estudio) es necesaria (Flint-Garcia *et al.*, 2005). Esta última corrección es muy estricta y puede inducir a una tasa de errores de tipo II (falsos negativos) elevada, pero estos errores son menos preocupantes en este tipo de análisis, en el que se trata fundamentalmente de demostrar la viabilidad del método. El análisis multilocus proporciona un nivel adicional de depuración de los resultados. Si aún quedaran asociaciones entre loci y QTL no ligados después de controlar la estructura de población, es decir, si hubiera varios loci, en falso desequilibrio de ligamiento, ligados a un carácter, el análisis multilocus permite elegir el más probable de entre todos ellos. Por lo cual se puede confiar en que buena parte de las asociaciones detectadas en este trabajo son reales, especialmente las detectadas en el grupo 4, y en los análisis multilocus. De hecho, la mayoría de las asociaciones coincidieron con QTLs detectados en poblaciones biparentales o con posiciones de genes conocidos. También se encontraron asociaciones que no coincidieron con QTLs o posición de genes conocidos. Esto se explicaría probablemente por el hecho de que en este estudio se analizaron más alelos que en poblaciones biparentales, y también que algunos caracteres no habían sido estudiados anteriormente.

Tanto los resultados encontrados en el presente estudio, como los observados en otros trabajos apuntan a la viabilidad del método de mapeo por asociación, aunque posiblemente los tests estadísticos deberán ser más estrictos que en una población

biparental, para eliminar con mayor seguridad los falsos positivos, que pueden influir en los resultados.

7-1-5. Referencias bibliográficas

- Allard RW. 1999.** History of plant population genetics. *Annual Review of Genetics*, 33: 1-27.
- Breseghele F, Sorrells ME. 2006.** Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*, 172: 1165-1177.
- Casas AM, Igartua E, Vallés MP, Molina-Cano JL. 1998.** Genetic diversity of barley cultivars grown in Spain, estimated by RFLP, similarity and coancestry coefficients. *Plant Breeding*, 117: 429-435.
- Casas AM, Yahiaoui S, Ciudad F, Igartua E. 2005.** Distribution of MWG699 polymorphism in Spanish European barleys. *Genome*, 48: 41-45.
- Ceccarelli S. 1996.** Positive interpretation of genotype by environment interactions in relations to sustainability and biodiversity. *In Plant Adaptation and Crop Improvement*. Ed. Cooper M, Hammer GL, CAB International, Wallingford, England. 467-486.
- Ceccarelli S, Grandó S. 1991.** Environment of selection and type of germplasm in barley breeding for low-yielding conditions. *Euphytica*, 57: 207-219.
- Ceccarelli S, Nachit MM, Ferrara GO, Mekni MS, Tahir M, van Leur JAG, Srivastava JP. 1987.** Breeding strategies for improving cereal yield and stability under drought. *In Drought tolerance in winter cereals*. Ed. Srivastava JP, Porceddu E, Acevedo E, Varma S. John Wiley and Sons, New York, 101-114.
- Ceccarelli S, Acevedo E, Grandó S. 1991.** Breeding for yield stability in unpredictable environments: single traits, interaction between traits, and architecture of genotypes. *Euphytica*, 56: 169-185.
- Dofing SM, Berke TG, Baenziger PS, Knight CW. 1992.** Yield and yield component response of barley in subarctic and temperate environments. *Canadian Journal of Plant Science*, 72: 663-669.
- Flint-García SA, Thuillet AC, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES. 2005.** Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant Journal*, 44: 1054-1064.
- Kraakman ATW, Niks RE, van der Berg P, Stam P, van Eeuwijk FA. 2004.** Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, 168: 435-446.
- Molina-Cano JL, Moralejo M, Igartua E, Romagosa I. 1999.** Further evidence supporting Morocco as centre of origin of barley *Theoretical & Applied Genetics*, 98: 913-918.

- Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler IV ES. 2001.** Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proceedings of the National Academy of Sciences, USA, 98: 11479-11484.
- Struss D, Plieske J. 1998.** The use of microsatellite for detection genetic diversity in barley populations. Theoretical & Applied Genetics, 97: 308-315.
- Tanno K, Taketa S, Takeda K, Komatsuda T. 2002.** A DNA marker closely linked to the *vrs 1* locus (row type gene) indicates multiple origins of six rowed cultivated barley. (*Hordeum vulgare L.*). Theoretical & Applied Genetics, 104: 54-60.
- van Oosterom EJ, Ceccarelli S, Peacock JM. 1993.** Yield response of barley to rainfall and temperature in Mediterranean environments. Journal of Agricultural Sciences, Cambridge, 121: 307-313.
- Voltas J, van Eeuwijk FA, Sombrero A, Lafarga A, Igartua E, Romagosa I. 1999.** Integrating statistical and ecophysiological analyses of genotype by environment interaction for grain filling of barley I. Individual grain weight. Field Crops Research, 62: 63-74.
- Wilson LM, Whitt SR, Ibañez A, Rocheford TR, Goodman MM, Buckler IV ES. 2004.** Dissection of maize kernel composition and starch production by candidate gene association. The Plant Cell, 16: 2719-2733.
- Zhang Q, Saghai-Marooif MA, Kleinhofs A. 1993.** Comparative diversity analysis of RFLPs and isozymes within and among populations of *Hordeum vulgare ssp. spontaneum*. Genetics, 134: 909-916.
- Zohary D. 1999.** Monophylitic vs polyphylitic origin of the crops on which agriculture was founded in the Near East. Genetic Resources and Crop Evolution, 46: 133-142.

Capítulo 8

Conclusiones

8. Conclusiones

1. La colección nuclear de cebadas españolas presenta una diversidad genética elevada y diferente del germoplasma de invierno del resto de Europa.
2. El análisis de la estructura de poblaciones subyacente en la muestra de cebadas estudiada identificó la presencia de cuatro grupos genéticos mayoritarios, de los cuáles dos son específicos de la colección nuclear, en la que aparecen sólo parcialmente mezclados. Estos dos grupos parecen tener orígenes distintos.
3. La asociación entre los marcadores y los datos climáticos y geográficos de los sitios de origen de las entradas de la colección nuclear, apunta a que la distribución de la diversidad de estas entradas no es aleatoria, sino que parece indicar adaptación a condiciones ambientales diferentes.
4. En general, las entradas de la colección nuclear de cebadas españolas presentan caracteres agronómicos más desfavorables que las variedades modernas. Sin embargo, algunas muestran una estabilidad notable en condiciones desfavorables, lo que sugiere la presencia de caracteres de adaptación en estas entradas.
5. Los dos principales grupos españoles definidos genéticamente han presentado comportamientos agronómicos también diferenciados. Estos grupos de entradas, en promedio, tienden a comportarse mejor en los ensayos realizados en ambientes más similares a sus regiones de origen, lo que revelaría la posible existencia de adaptaciones específicas en estos grupos, probablemente ligadas a las respuestas a la temperatura.
6. La búsqueda de QTLs mediante mapeo por asociación en la cebada parece viable, siempre que se tenga en cuenta la posible estructura de población subyacente en los análisis. Los resultados encontrados en este estudio mostraron que un gran número de las asociaciones halladas coincidieron con QTLs encontrados en poblaciones biparentales o con posiciones de genes conocidos.
7. Finalmente, considerando conjuntamente los análisis de diversidad genética, la distribución eco-geográfica de la misma y la evaluación agronómica de la colección nuclear de cebadas españolas, se concluye que esta colección constituye un recurso fitogenético de primer orden para la búsqueda de nuevos genes y alelos para la mejora de la cebada, especialmente para ambientes mediterráneos.

Anexos

Anexo1

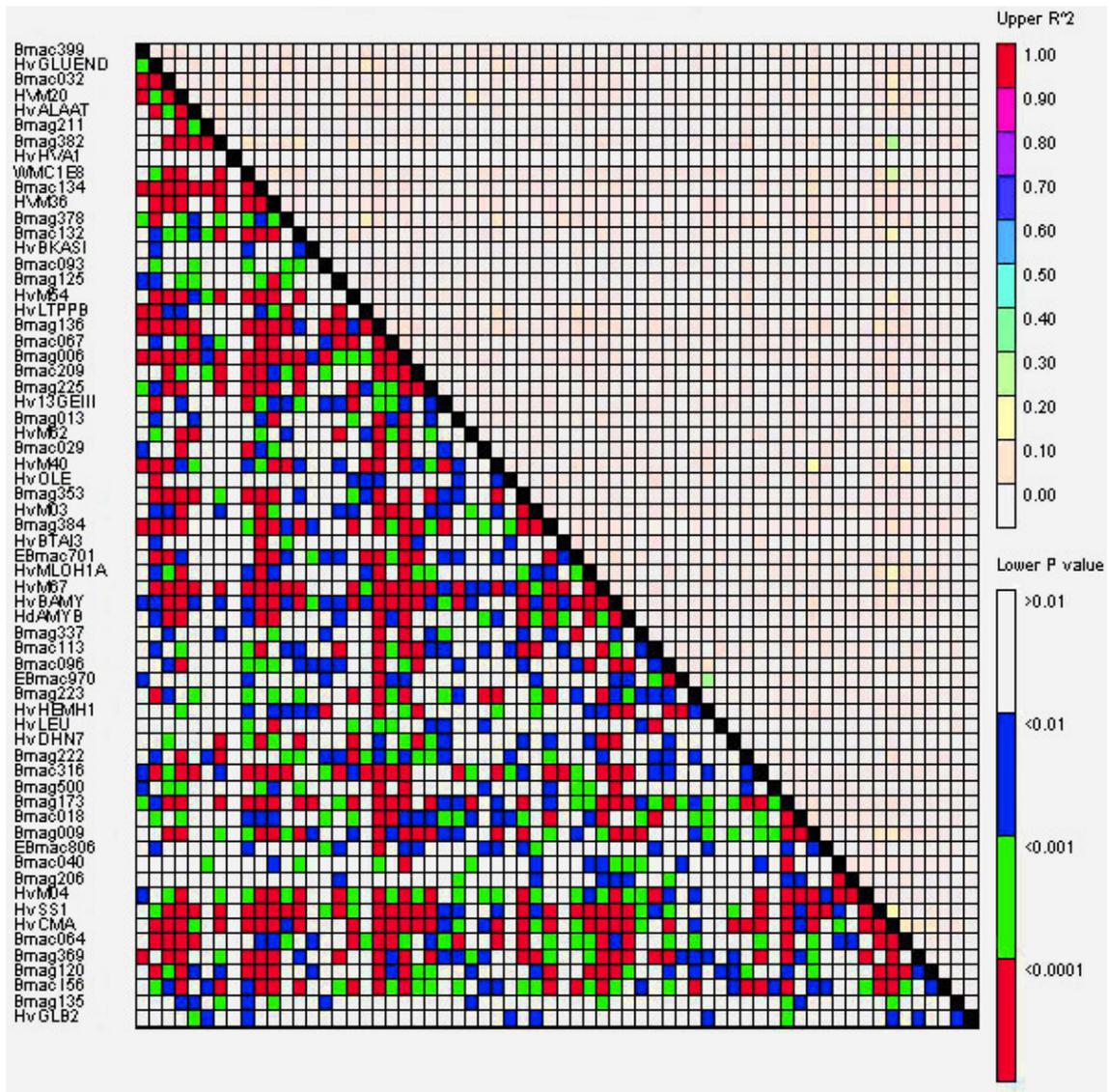


Figura 1. Desequilibrio de ligamiento entre 64 loci, evaluado en los 225 genotipos analizados. A la derecha se representan los coeficientes r^2 y a la izquierda las probabilidades de significación correspondientes, según los colores de la leyenda

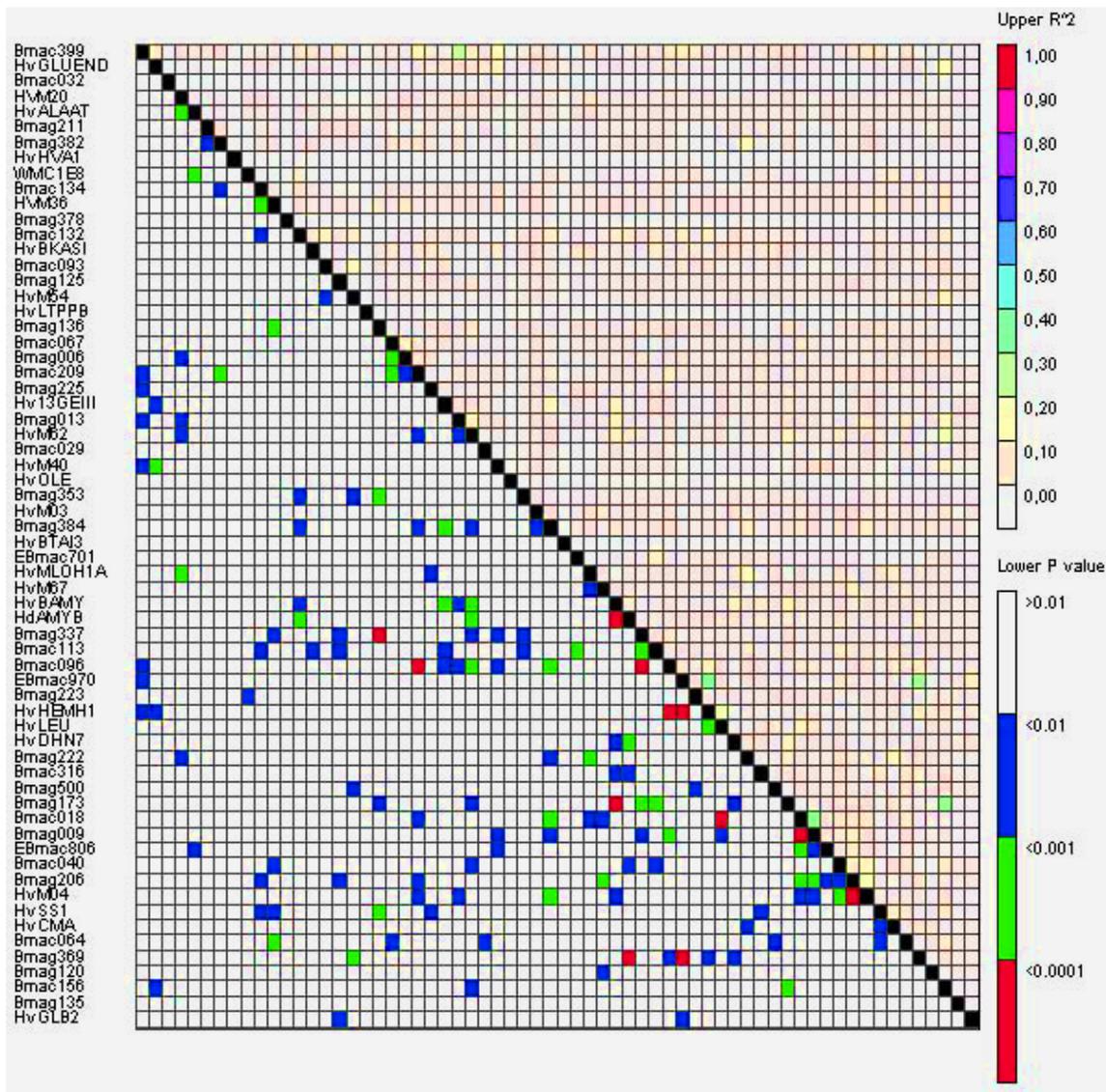


Figura 2. Desequilibrio de ligamiento entre 64 loci, evaluado en las entradas de la población 1. A la derecha se representan los coeficientes r^2 y a la izquierda las probabilidades de significación correspondientes, según los colores de la leyenda

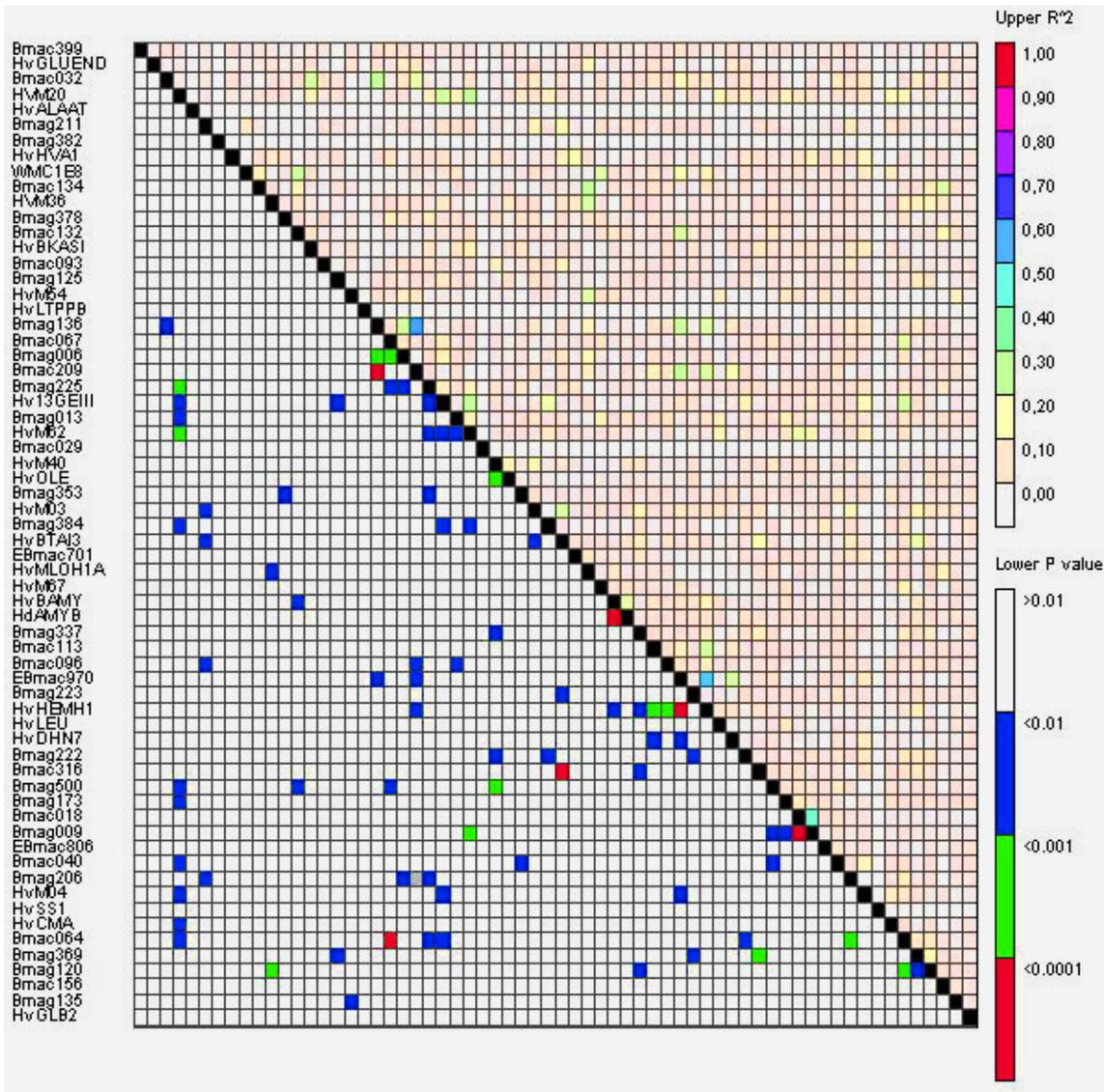


Figura 3. Desequilibrio de ligamiento entre 64 loci, evaluado en las entradas de la población 2. A la derecha se representan los coeficientes r^2 y a la izquierda las probabilidades de significación correspondientes, según los colores de la leyenda

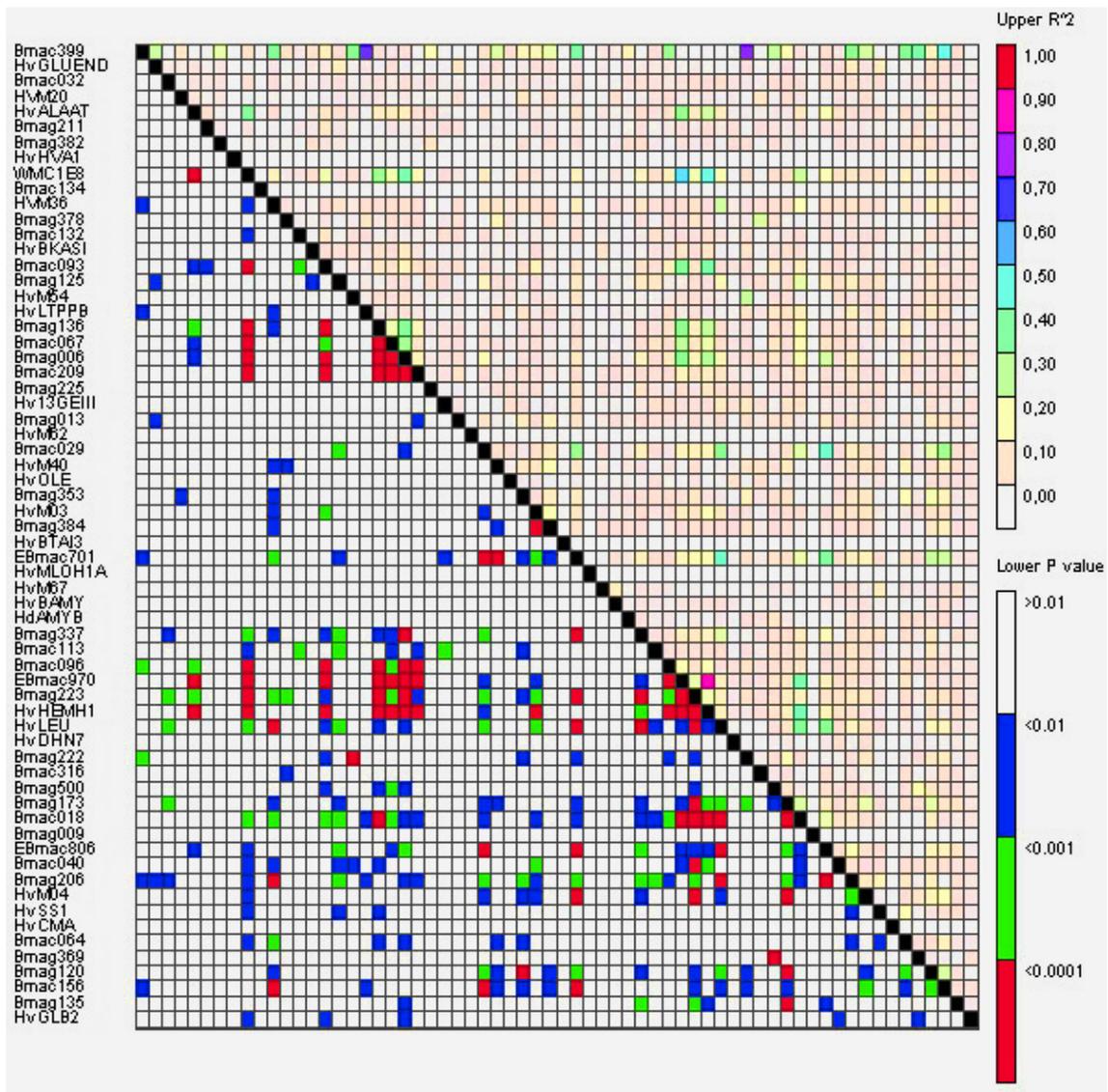


Figura 4. Desequilibrio de ligamiento entre 64 loci, evaluado en las entradas de la población 3. A la derecha se representan los coeficientes r^2 y a la izquierda las probabilidades de significación correspondientes, según los colores de la leyenda

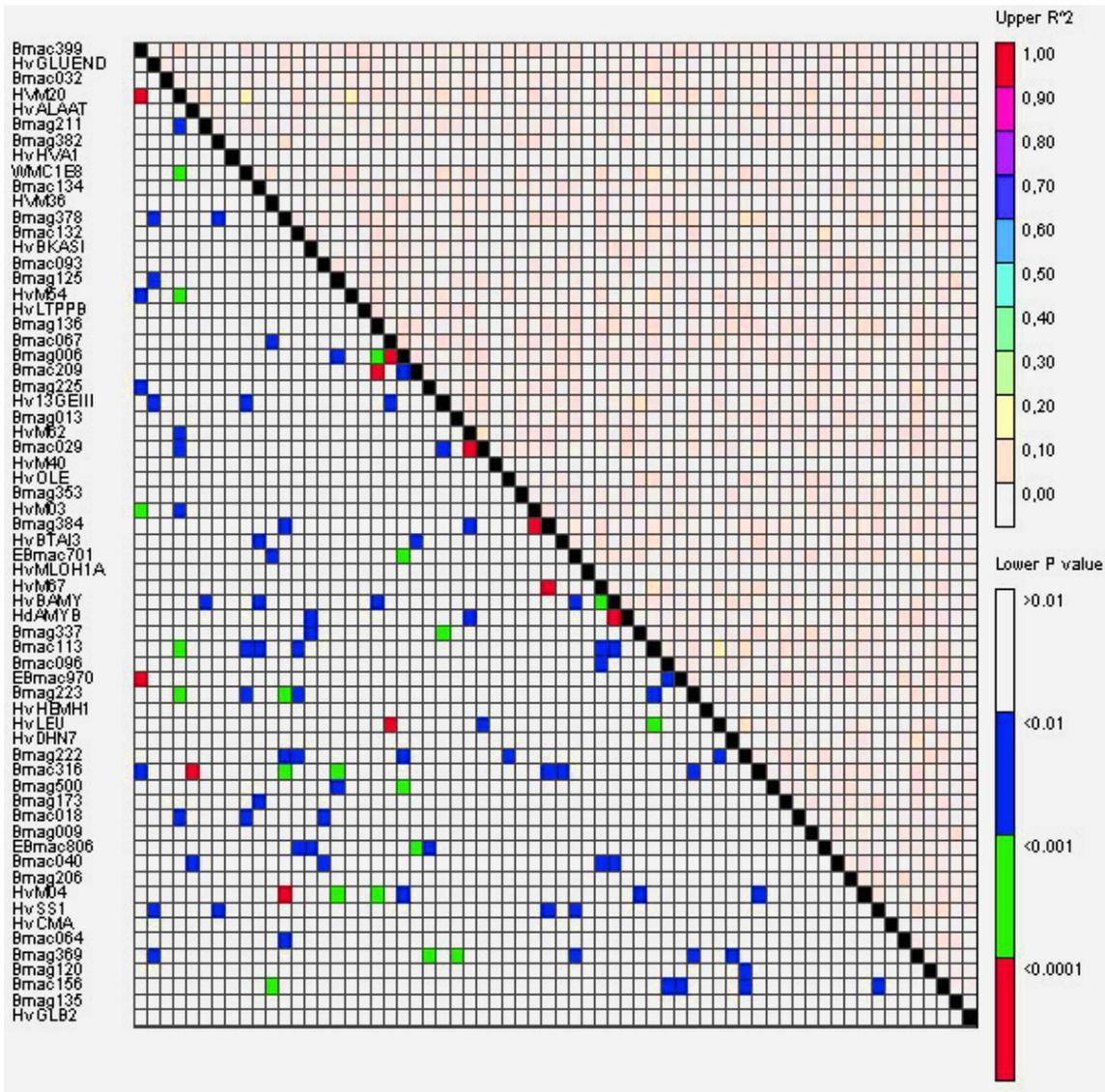


Figura 5. Desequilibrio de ligamiento entre 64 loci, evaluado en las entradas de la población 4. A la derecha se representan los coeficientes r^2 y a la izquierda las probabilidades de significación correspondientes, según los colores de la leyenda

Anexo 2

Tabla 1. Número total de alelos detectados en los 225 genotipos descritos en el capítulo 3, rango de tamaño de los alelos de cada locus; número de alelos comunes y alelos específicos de los 8 marcadores introducidos en el capítulo 4 y 6, en cada uno de los 4 grupos de germoplasma establecidos a priori.

Cr	Locus	Total	Número de alelos				Tamaño (pb)	Número de alelos específicos			
			6 C Esp	6 C Ref	2 C Esp	2 C Ref		6 C Esp	6 C Ref	2 C Esp	2 C Ref
1H	MWG518	2	2	2	2	2	-	0	0	0	0
	scssr10477	6	3	5	2	3	Nulo-158	1	2	0	0
2H	MWG858	2	2	2	2	2	-	0	0	0	0
	scssr07759	6	4	4	2	2	181-223	2	1	0	0
	EBmac640	5	4	3	3	3	174-190	1	0	1	0
4H	HvZCCT	2	2	2	2	2	Nulo-1400	0	0	0	0
5H	Hv635P2.4	10	8	9	6	6	188-199	1	2	0	0
	HvBM5A	7	3	7	4	5	Nulo-5250	0	1	0	0

Tabla 2. Distribución de los alelos frecuentes en las 159 *landraces* en los 4 grupos generados por STRUCTURE

Alelo	Tamaño (pb)	Número de genotipos			
		Grupo1	Grupo2	Grupo3	Grupo4
MWG518-1		5	9	3	19
MWG518-2		13	0	46	64
scssr10477-1	126	7	1	48	75
scssr10477-2	134	3	0	1	6
scssr10477-5	nulo	8	8	1	1
scssr07759-3	195	0	0	2	5
scssr07759-4	213	0	0	7	32
scssr07759-5	221	18	5	41	44
MWG858-1		12	9	13	17
MWG858-2		6	0	37	65
EBmac640-1	174	12	2	41	70
EBmac640-3	178	6	2	8	11
Hv635P24-3	192	7	1	15	26
Hv635P24-4	193	4	3	16	19
Hv635P24-5	194	3	1	2	18
Hv635P24-6	195	2	1	5	12
Hv635P24-7	196	1	3	4	5
Hv635P24-8	197	0	0	7	2
HvZCCT-1	1400	14	2	49	82
HvZCCT-2	nulo	4	7	1	0
HvBM5-2	4850	8	0	49	37
HvBM5-4	1200	5	1	1	44
HvBM5-6	150	0	6	0	1
HvBM5-7	nulo	5	2	0	0