Genetic Resources and Crop Evolution **45:** 475–481, 1998.

The Spanish barley core collection

E. Igartua[1,*], M.P. Gracia[1], J.M. Lasa[1], B. Medina[1], J.L. Molina-Cano[2], J.L. Montoya[3] & I. Romagosa[2]
[1] *Department of Genetics and Plant Production, Estación Experimental de Aula Dei, CSIC, P.O. Box 202, 50080-Zaragoza, Spain;* [2] *Universitat de Lleida-Institut de Recerca i Tecnologia Agroalimentàries, Av. Rovira Roure 177, 25196-Lleida, Spain;* [3] *Servicio de Investigación Agraria, Junta de Castilla y León, P.O. Box 172, 47080-Valladolid, Spain (* Author for correspondence: e-mail: igartua@eead.csic.es)*

Received 4 December 1997; accepted in revised form 30 April 1998

Key words: barley, core collection, Spain

Abstract

Spanish barleys constitute a germplasm group of particular interest for breeding purposes, as Spain has been proposed as a possible centre of origin of the crop. The Spanish National Germplasm Bank (Banco Nacional de Germoplasma, BNG), holds a collection of about 2000 barley accessions, mostly landraces collected in Spain prior to extensive introduction of modern varieties. The objective of this work is to create a core collection of barleys representative of old barley genotypes grown in Spain. The core collection will be constituted by three groups of germplasm: successful old varieties (15); entries in common with previously existing barley core collections (15); and 2-row (8) and 6-row (122) entries from the BNG, for a total of 160 entries. Entries were allocated by stratified sampling in agro-ecological uniform zones of barley cultivation in Spain. Classification of agro-ecological regions for barley was based on historical yield records for Spanish provinces. The number of entries for each region was determined in proportion to the logarithm of historical barley acreage. Final choice of accessions within provinces tried to maximize the diversity and avoid duplications by looking at passport data, and to agronomic evaluation data available for a group of about 900 accessions.

Abbreviations: BNG – Banco Nacional de Germoplasma.

Introduction

Core collections were devised as a tool for researchers and plant breeders to explore and utilise the genetic diversity present in germplasm collections (Frankel, 1984). Barley is a good example of a crop which would benefit from more extensive use of germplasm collections, as non-conventional germplasm has evidenced potential to contribute to the improvement of current cultivars (Ceccarelli & Grando, 1989; Veteläinen, 1994; Hadjichristodoulou, 1995). Recently, a large scale project has led to the creation of the International Barley Core Collection (Knüpffer & van Hintum, 1995), which intends to represent genetic diversity present in barley germplasm collections worldwide. Also, the USDA-ARS National Small Grains Collection is assembling a core subset from their barley collection, with world-wide scope as well (Bockelman, 1996). The study and evaluation of these large core collections will definitely contribute to increase our knowledge of the genetic diversity of the crop and its distribution. It will also help to identify geographical regions as potential donors of desirable traits for breeding purposes. In some cases, however, there exists sufficient prior knowledge about the potential usefulness of local germplasm resources as to justify a more detailed investigation. Spanish barley constitutes a group of particular interest, as its cultivation dates back to 5000 years BC, according to archaeological records (Moralejo et al., 1994). Thus, Spanish landraces evolved under Mediterranean conditions for several thousand years, and can be a likely source of alleles conferring adaptation to these conditions. This local germplasm has been so far under-exploited in breeding programmes, though its merit is demonstrated by the prolonged success of some old varieties (the best example would be cultivar Albacete, obtained from a local population 40 years ago, and still widely grown), and the persistence of some landraces still under cultivation, despite the general availability of modern varieties. This germplasm is also interesting because some results suggest that the West-Mediterranean region was a possible centre of origin for barley (Molina-Cano et al., 1987; Moralejo et al., 1994). The scarce evidence available (Moralejo et al., 1994; Casas et al., 1997) indicates the uniqueness of protein and RFLP patterns in Spanish landraces or old cultivars (2- and 6-row) compared to other European germplasm. Thus, Spanish landraces may be a source of new genetic variants for mainstream barley breeding programmes in Europe.

The Spanish National Germplasm Bank (BNG), located at the Centro de Recursos Fitogenéticos - Instituto Nacional de Investigaciones Agrarias in Alcalá de Henares, holds a collection of 2155 barley accessions, mostly (91%) landraces collected in Spain prior to extensive introduction of modern varieties. The main collection effort was actually carried out in 1942, well before significant introduction of foreign varieties, and covering the whole geographic range of barley cultivation in the country. The objective of this work was to create a core collection of Spanish barleys which could glean most of the genetic diversity contained in the whole collection, and that may constitute a research tool for breeders and plant scientists.

Materials and methods

The procedure to assemble the core collection essentially followed the recommendations given by van Hintum (1996), which can be summarized in the following steps: 1) determination of domain and size of the core collection; 2) division into genetically distinct types; 3) allocation of entries over groups; and 4) choice of entries.

For step 3), the number of entries to constitute each group was based on previous knowledge on the relative importance and diversity of the groups. An special procedure was followed for the large number of 6-row accessions present at the BNG collection. They were classified into groups of potentially distinct types, according to an agro-ecological classification of barley growing areas, based on a cluster analysis of historical yield records for the Spanish provinces. For this purpose, records of average provincial barley yields for the last 47 years (Spain has 50 provinces, with an average area of 10,000 km^2), were extracted from the Spanish Yearbooks of Agricultural Statistics (Anuario de Estadística Agraria, MAPA, 1948–1994). Clustering of provincial yields was carried out following the incremental sum of squares or Ward method, in SAS/STAT proc cluster (SAS Institute, 1989).

The number of entries to represent each agro-ecological region defined by the cluster analysis was established as proportional to the area devoted to barley in each agro-ecological region. For this purpose, however, the same data set as for defining the agro-ecological regions could not be used, since distribution of barley acreage in Spain changed markedly during the last three decades. Thus, the data set including the last 47 years used for classification of agro-ecological zones would not be representative of barley cultivation during the period prior to the introduction of modern varieties, which we propose as the domain of the core collection. For this reason, we gathered barley acreage data available for the period prior to 1960, when it can be assumed that most varieties grown were either primitive cultivars or local landraces. Data from only fifteen years, scattered from 1940 till 1959 were found. This was considered representative enough for an assessment of area devoted to barley per province and agro-ecological region, as barley acreage was very stable among years during that period (though it increased sharply from the early 1960s onwards).

Allocation of entries over agro-ecological regions was done according to stratified logarithmic sampling on the agro-ecological regions, based on the area devoted to barley calculated, as mentioned in the previous paragraph. Logarithmic proportionality increases the probability of capturing alleles conferring wide or local adaptation, better than a mere random sampling (Brown, 1989; van Hintum, 1996). The superiority of stratified sampling has been confirmed in several studies comparing a variety of sampling methods (Charmet & Balfourier, 1995; Diwan et al., 1995).

Choice of accessions within provinces, was made using information of passport data and of agronomic evaluations of a sub-group of accessions from the BNG, carried out at Zaragoza, Lleida and Valladolid (Spain), from 1986 to 1991. These evaluations were made in unreplicated trials with check plots, sown in Autumn, with plot size of two rows 120 cm long, 20 cm apart, for a variable number of accessions per year (from 38 to 244, for a total of 933). Several agronomic characters, such as days to flowering, days to maturity, length of grain filling period, plant height at maturity, grain yield, test weight, and 1000 kernel weight, were recorded at all trials.

Results and discussion

The steps suggested by van Hintum (1996) were completed as follows:

1. Domain and size of the core collection

The domain of the core collection was established as the *Hordeum vulgare* ssp. *vulgare* types grown in Spain prior to the extensive introduction of modern varieties. No indigenous ssp. *spontaneum* has been described in Spain. A set of about 160 entries was considered representative enough, and of manageable size.

2. Making of groups based on knowledge of germplasm

It was decided to include several groups of entries, for a variety of reasons. These groups were (Table 1):

I. Cultivars with a long history of successful cultivation in Spain. The inclusion of these entries ensures that materials presenting the best combination of loci conferring adaptation and good agronomic performance under Spanish conditions are included in the core collection.

II. Entries in common with other barley core collections, included as reference genotypes for comparisons across collections.
   - A. Entries in common with the European part of the International Barley Core Collection.
   - B. Entries in common with the barley core subset of the USDA-ARS National Small Grains Collection (NSGC).

III. Entries from the BNG collection, which make the basis of this core. They were further subdivided into the only genetically distinct types which could be identified inequivocally from passport data:
   - A. Two-row accessions, which constitute a small part of the collection, as most of the barley traditionally grown in Spain was six-row.
   - B. Six-row accessions from the BNG collection, other than in group I.

Table 1. Groups (and number) of entries for the Spanish Barley Core Collection

	Group		no. entries
I.	Successful old varieties		15
	Six – row	Two – row	
	Ager	Beka	
	Albacete	Hassan	
	Almunia	Kym	
	Barberousse	Pallas	
	Dobla	Union	
	Hatif de Grignon	Wisa	
	Monlon	Zaida	
	Pané		
II.A.	Spanish entries from European part of the International Barley Core Collection		2
II.B.	Spanish entries from Barley Core Subset of the USDA-ARS National Small Grains Collection (NSGC)		13
III.A.	Two-row entries from BNG		8
III.B.	Six-row entries from BNG		122
	TOTAL		160

3. Allocation of entries over the groups

About 10 per cent of the total number of entries was allocated to each of groups I and II. The remaining 80 per cent was allocated to the Spanish accessions from the BNG. Relative size of groups III.A and III.B was linearly proportional to their representation in the BNG collection (95% six-row, 5% two-row), which on the other hand reflected their agronomic relevance in Spain at the time of germplasm collection. Final allocation of entries over these groups is summarized in Table 1.

Allocation of entries within group III.B was more elaborate, due to the abundance of entries of this type. Several methods have been proposed to tackle the problem of choosing a representative subset of entries from a large germplasm collection. When there are not other descriptive data available, passport data should be used. Also, well-established agro-ecological regions can be used to define groups. Other approaches can be taken if evaluation data, such as agronomic, morphological, or molecular traits, are available. In such case, several multivariate analysis techniques have been proposed to identify groups of entries which account for most of the phenotypic diversity detected (Peeters & Martinelli, 1989; Spagnoletti Zeuli & Qualset, 1993; Noirot et al., 1996). Charmet & Balfourier (1995) attempted to benefit from both approaches by making use of both evaluation data and geographical origin of the entries by means of geostatistics.

Any of these approaches depends on the relative frequency of entries from a particular group in the whole collection, and assumes that the germplasm collection is an undistorted representation of crop distribution and its relative importance in the target area.

In many cases, however, it is not possible to test that assumption. Thus, some regions may be over- or under-represented in germplasm collections due to unknown factors which may have affected the process of germplasm collection. For this reason, we decided to use an independent source of information to determine agro-ecological zones with respect to barley cultivation, and sample the six-row entries according to their abundance in these areas. For this purpose, the historical records on barley cultivation in Spain were used, following a two-step process:

First, the classification into agro-ecological zones was done for the 34 provinces where there was significant barley cultivation during the last 47 years. Sixteen provinces where either barley cultivation was negligible, or where the six-row barley acreage decreased during the time period considered till almost disappear, were not taken into account for clustering pusposes.

In most of these provinces, we detected a significant increase in yield over time, but differing in magnitude across provinces. It was necessary to remove these trends prior to classification, as the reasons for their existence were most likely outside the subject of the analysis. Agro-ecological classification should be based upon environmental factors affecting historical crop adaptation over the whole period of cultivation. Increasing trends in average yields in the database, however, are a consequence of the adoption of better management practices, improved varieties, and increase of inputs in the last decades. The increase was more steep for some provinces, specially the highest yielding ones, where the higher probability of getting good yields and better profits justified bigger investments. Thus, to remove these undesired confounding effects, the yield time-series for each province was detrended as follows: first, a linear regression of year on yield was fitted for each province. Then, the residuals from the regression were added to the mean average yield across 47 years for each province. This new set

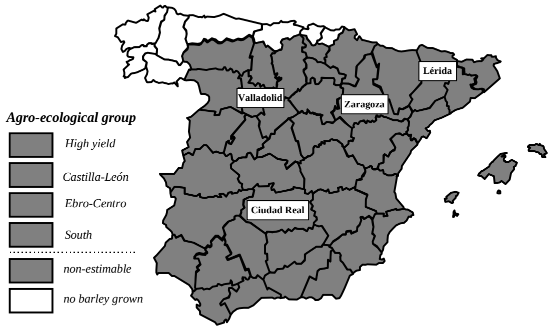


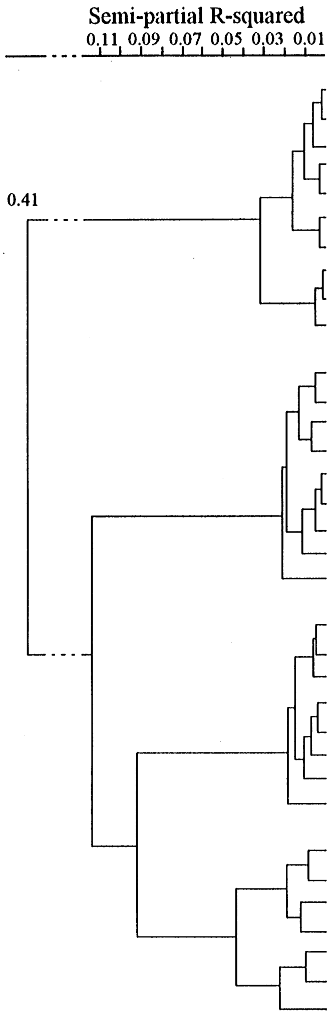
Fig. 1. Map of Spanish provinces showing the grouping of agro-ecological regions for barley cultivation, according to yield records of the last 47 years.

of data actually presents the same average provincial yields as the original one, but without time trends. Clustering of provinces according to detrended data produced groups of provinces with both similar average yield, and consistency in the fluctuation of yield across years. The analysis produced four clearly distinct groups, represented in Fig. 1. The truncation level was subjectively chosen as the appearance of a sudden increase in the semi-partial R-squared figure during the agglomeration process. Differences between the four clusters accounted for 62% of the total variance.

The first three clusters included contiguous provinces (Fig. 1), quite consistent regarding climatic characteristics, whereas the fourth included the four provinces from Cataluña and another three geographically more scattered high-yielding provinces. A last group was assembled with six provinces not included in the cluster analysis, in which 6-row barley cultivation was important in the early years of the 47-year period considered, but later declined till almost disappear. For the same reason, yield averages for these provinces are not provided in Fig. 1: their yield estimates for most of the period were calculated for very small acreages, which may not be representative of the larger acreages occupied by barley in these provinces in earlier years.

Once the agro-ecological zones were defined, the second step was the decision on how many entries should represent each zone in the core collection. As mentioned in the Materials and Methods section, we decided to sample in logarithmic proportion to the area devoted to barley in each region, prior to the introduction of modern varieties. These areas, in hectares and in percentage from the total, are summarized in Table 2. Within each region, the allocation of number of accessions per province was done according to linear

Table 2. Agro-ecological regions produced by clustering of Spanish provinces based on 47 years of yield averages from Agricultural Yearbooks. Also presented are the number of accessions from the BNG allocated to each province in the Spanish barley core collection, and historical records on barley cultivation area used for the stratified sampling over provinces



Agro-ecological group / Province	Historical (1930-60) 6-row area (%)	Mean yield (1948-94) (kg ha^{-1})	no. of entries for the core: 6-row	no. of entries for the core: 2-row
South	**42.4**	**1086**	**32**	**1**
Córdoba	4.1	1299	3	
Jaén	4.3	1224	3	
Valencia	0.8	1226	1	
Ciudad Real	8.5	1230	6	
Granada	4.2	1149	3	
Badajoz	8.5	1057	6	1
Cáceres	3.5	895	3	
Almería	2.2	757	2	
Murcia	4.7	863	4	
Baleares	1.6	1007	1	
Ebro-Centro	**22.5**	**1576**	**26**	**4**
Zaragoza	2.3	1537	3	
Teruel	1.4	1389	2	1
Albacete	6.1	1475	6	
Castellón	0.4	1445	1	
Madrid	1.9	1656	2	
Toledo	5.8	1601	7	1
Guadalajara	1.7	1672	2	1
Soria	1.7	1890	2	1
Huesca	1.1	1703	1	
Castilla-León	**14.9**	**1876**	**23**	**2**
Burgos	3.0	1885	5	
Palencia	1.4	1926	2	
Valladolid	2.8	1944	4	1
Avila	1.3	1848	2	
Zamora	1.5	1875	2	1
Salamanca	1.9	1639	3	
León	0.7	1777	1	
Segovia	2.4	2245	4	
High yield	**9.3**	**2056**	**20**	**1**
Barcelona	0.9	2212	2	
Gerona	0.2	2096	1	
Lérida	2.2	1994	5	1
Tarragona	0.9	1989	2	
Navarra	1.2	2195	2	
Rioja	0.9	2360	2	
Cuenca	3.0	1915	6	
Non- clustered (see text)	**10.1**	-	**21**	
Alava	0.3	-	2	
Alicante	0.9	-	3	
Cádiz	0.9	-	3	
Huelva	0.8	-	3	
Málaga	2.1	-	4	
Sevilla	4.7	-	5	
Tenerife	0.5	-	2	
TOTAL	**1490000 ha**	**1543**	**122**	**8**

proportionality to barley acreage, except for the group of provinces not classified in the cluster analysis, for whom logarithmic proportion was again used. Final allocation of entries for all provinces across groups is presented in Table 2.

Given the small number of two-row accessions in the BNG collection, no systematic criterion for allocation was used.

4. Choice of entries

Selection in groups I through III.A. was agreed among the authors, taking account of diversity and completeness of passport data, and of previous knowledge on the past importance and denominations of some entries. For group III.A, for instance, care was taken to include entries representing the *ladillas* and *pámulas,* two of the main traditional two-row types grown in Spain.

For groups III.A and III.B, the choice of entries within each province was made among accessions presenting complete passport data (1489) and, if possible, among these already evaluated for several agronomic traits (933). As growing conditions varied considerably among seasons, raw agronomic data could not be assembled in a unique database. Thus, all traits were transformed into discrete characters, with three levels each, either early-medium-late, short-medium-long, or low-medium-high, depending on the traits. Then, a new database was assembled, including all accessions with complete passport data (agronomically evaluated or not), together with the transformed evaluation data. Choice of entries for each province was done on this final database.

The selection criteria for germplasm inclusion aimed at maximizing the potential genetic diversity of each province's sample. For this purpose, we took into account both the evaluation data and also the altitude of the collection site as, given the mountainous nature of the landscape, altitude may be one of the main factors influencing crop adaptation in Spain. As yield and yield components present very large year by accession interaction under our conditions, producing dramatic changes in the ranking of accessions across seasons, less importance was given to these data compared to phenologic and height data. Thus, the accessions selected for each province were chosen to represent (mainly) the diversity of altitude of collection site, flowering date, and plant height.

Most two-row and six-row entries from the BNG barley germplasm collection finally selected to constitute the core collection were collected prior to 1950 (85%), and almost totally before extensive introduction of modern varieties (96% prior to 1960). Regarding geographical information of collection sites, the range of latitudes extended from 36°:32'N to 43°04'N (except two accessions from the Canary Islands, collected at 28°:39' and 28°:48'N); altitude of collection sites varied from 7m to 1587m above sea level. Regarding the evaluation data, there were ample ranges of variation for heading dates, as they fluctuated from 14 to 24 days across seasons (average 19), and for plant height, which varied between 21 cm and 40 cm (average 29 cm). Even these rough measures of variability already suggest the presence of potentially large genetic diversity for relevant agronomic traits.

Seed increases of all entries in the core collection (all groups in Table 1) will be carried out during the 1997/98 season, and evaluated for several agronomic and morphological traits at a minimum of three different locations. All the entries are reported to be inbred lines, and will be multiplied as such by harvesting bagged heads of a single plant in only one of the evaluation trials. Regarding the two- and six-row accessions from the BNG, a surplus of 20–25 entries will be sown, to prevent multiplication failures or potential duplications in the material, but their final number will be reduced to the figures presented in Tables 1 and 2. Seed will be available for distribution by 1999.

Acknowledgements

This work was funded by the INIA (Instituto Nacional de Investigaciones Agrarias). Seed and specific information of Spanish entries belonging to the USDA barley collection was kindly provided by Dr. Harold Bockelman, curator of the National Small Grains Collection. Seed from the entries shared with the International Barley Core Collection were kindly provided by its curator, Dr. H. Knüpffer.

References

Bockelman H.E., 1996. Developing a barley core subset in the USDA-ARS National Small Grains Collection. In: Skinlared A., G. Scoles & B. Rossnagel (Eds), Proceedings of the V International Oat Conference & VII International Barley Genetics Symposium. Poster Sessions, vol. 1, p. 154. University of Saskatchewan, Saskatoon.

Brown A.H.D., 1989. Core collections: a practical approach to genetic resources management. Genome, 31: 818–824.

Casas, A., E. Igartua, M.P. Vallés, F. Ciudad & J.L. Montoya, 1997. Genetic characterization of barley heading patterns in Spain. EUCARPIA, section Cereals Meeting, Tulln (Austria): 'Application of marker aided selection in cereal breeding programmes', September 21–23.

Ceccarelli S. & S. Grando, 1989. Efficiency of empirical selection under stress conditions. J. of Gen. and Breeding 43: 25–31.

Charmet G. & F. Balfourier, 1995. The use of geostatistics for sampling a core collection of perennial ryegrass populations. Genetic Resources and Crop Evolution 42: 303–309.

Diwan N., M.S. McIntosh & G.R. Bauchan, 1995. Methods of developing a core collection of annual *Medicago* species. Theor. Appl. Genet. 90: 755–761.

Frankel O.H., 1984. Genetic perspectives on germplasm conservation. In: Arber W., K. Llimensee, W.J. Peacock & P. Starlinger (Eds), Genetic Manipulation: Impact on Man and Society, pp. 161–170. Cambridge University Press, Cambridge UK.

Hadjichristodoulou A., 1995. Evaluation of barley landraces and selections from natural outcrosses of *H. vulgare* ssp. *spontaneum* with ssp. *vulgare* for breeding in semi-arid areas. Genetic Resources and Crop Evolution 42: 83–89.

Knüpffer, H. & Th. J.L.van Hintum, 1995. The Barley Core Collection: an international effort. In: Hodgkin T., A.H.D. Brown, Th. J. L. van Hintum & E. A. V. Morales (Eds), Core Collections of Plant Genetic Resources pp. 171–178, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore.

MAPA. 1949-1994. Anuario de Estadística Agraria. Publicaciones del Ministerio de Agricultura, Pesca y Alimentación (MAPA). Subdirección General de Estadística, Servicio de Estadísticas Agrarias.

Molina-Cano, J.L., P. Fra-Mon, G. Salcedo, C. Aragoncillo, F. Roca-de-Togores, 1987. Morocco as a possible domestication center for barley: biochemical and agromorphological evidence. Theor. Appl. Genet. 73: 531–536.

Moralejo M., I. Romagosa, G. Salcedo, R. Sánchez-Monge & J.L. Molina-Cano, 1994. On the origin of Spanish two–rowed barleys. Theor. Appl. Genet. 87: 829–836.

Noirot M., S. Hamon & F. Anthony, 1996. The principal component scoring: A new method of constituting a core collection using quantitative data. Genetic Resources and Crop Evolution 43: 1–6.

Peeters J.P. & J.A. Martinelli, 1989. Hierarchical cluster analysis as a tool to manage variation in germplasm collections. Theor. Appl. Genet. 78: 42–48.

SAS Institute. 1989. SAS/STAT Userïs Guide. Version 6, 4th ed. Vol 1 and 2. SAS Inst., Inc., Cary, NC.

Spagnoletti Zeuli P.L. & C.O. Qualset, 1993. Evaluation of five strategies for obtaining a core subset from a large genetic resource collection of durum wheat. Theor. Appl. Genet. 87: 295–304.

van Hintum, Th.J.L., 1996. Core collections in germplasm conservation: evaluation and use. In: Scoles G. & B. Rossnagel (Eds), Proceedings of the V International Oat Conference & VII International Barley Genetics Symposium. Invited Papers, pp. 113–119. University of Saskatchewan, Saskatoon.

Veteläinen M., 1994. Exotic barley germplasm: variation and effects on agronomic traits in complex crosses. Euphytica 79: 127–136.