Theor Appl Genet (2002) 105:91–99
DOI 10.1007/s00122-002-0903-7

P. Gauthier · B. Gouesnard · J. Dallard · R. Redaelli
C. Rebourg · A. Charcosset · A. Boyat

RFLP diversity and relationships among traditional European maize populations

Received: 14 July 2001 / Accepted: 22 November 2001 / Published online: 17 May 2002

Abstract Given the large extent of hybrid cultivation, the importance of conserving the diversity of crop genetic resources has given birth to numerous collections of old races. In the present paper, we conduct a molecular characterisation of a large collection of 488 European maize populations using the bulk RFLP analysis. The analysis of 23 RFLP loci showed a high allelic richness of 11.5 alleles per locus. Populations from eastern Europe (Poland, Austria, Germany, etc.) showed the lowest genetic diversity, a lower number of unique alleles and a higher percentage of fixed loci than populations from southern Europe. In fact, genetic diversity appeared higher in Southern regions where the first maize populations are thought to have been introduced. Molecular classification based on Rogers' distance (i.e. alleles frequencies) allowed us to distinguish three main clusters which were highly consistent with geographic origins. A Northeastern cluster grouped together early or intermediate populations from Northeastern countries and the Balkans, a southeastern cluster joined late and partially dent populations from Greece and Italy, and, a southwestern cluster was made up of early flint populations from northern Spain, Portugal and the Pyrenees. A correlation between allelic frequencies at some loci and latitude and/or longitude was observed. Such tendencies may reflect the direction of gene flow between different races of maize: for instance, North American (Northern flint) and Caribbean populations were introduced, respectively, to northern and southern Europe, in the past.

Communicated by H.C. Becker

P. Gauthier · B. Gouesnard (✉) · J. Dallard · A. Boyat
INRA,
Unité mixte de recherches Diversité et génome des plantes cultivées,
Domaine de Melgueil, 34130 Mauguio, France
e-mail: gouesnard@ensam.inra.fr
Fax: +33-4-67293990

R. Redaelli
ISC, Sezione di Bergamo, Via Stezzano, 24, 24126 Bergamo, Italy

C. Rebourg · A. Charcosset
INRA, Station de Génétique Végétale, Ferme du Moulon,
91190 Gif sur Yvette, France

Keywords *Zea mays* L. · Bulk RFLP · European landraces · Genetic diversity · Genetic structure

Introduction

In addition to an emphasis on production that has prevailed in the past for most crop varieties, present selection criteria have to take into account the necessity for diversification and environmental preservation. For instance, sources of pest tolerance and drought and cold tolerance have become important in order to reduce dependency on chemical pesticides and irrigation. In maize, open-pollinated varieties appear to be a major source of diversity with respect to these objectives. Since the introduction of maize to Europe five centuries ago, cultivated populations have evolved under the different selective pressures imposed in different regions and the needs of local farmers. The adaptation of landraces to many niches of European countries for many years explains the large variability which can be observed today in collections of populations (first stated by Brandolini 1969).

The necessity to preserve genetic resources appeared after the introduction of the first commercial hybrids 50 years ago (Edwards and Leng 1965) and led to the birth of many national maize collections. The necessity to characterise these collections in order to use their material properly in breeding programmes also appeared quickly. Morphological descriptions and classifications have been carried out on Spanish (Sanchez-Monge 1962), Italian (Brandolini and Mariani 1968), Yugoslavian and Romanian (Pavlicic and Trifunovic 1966), Portuguese (Costa-Rodrigues, 1971) and, more recently, French (Gouesnard et al. 1997) national collections. Using isozymes, Geric et al. (1989) analysed genetic diversity and relationships among 300 Yugoslavian accessions previously classified in 18 groups on the basis of morphological characters.

However, only a few studies focused on landraces at the European scale. Leng et al. (1962) described mor-

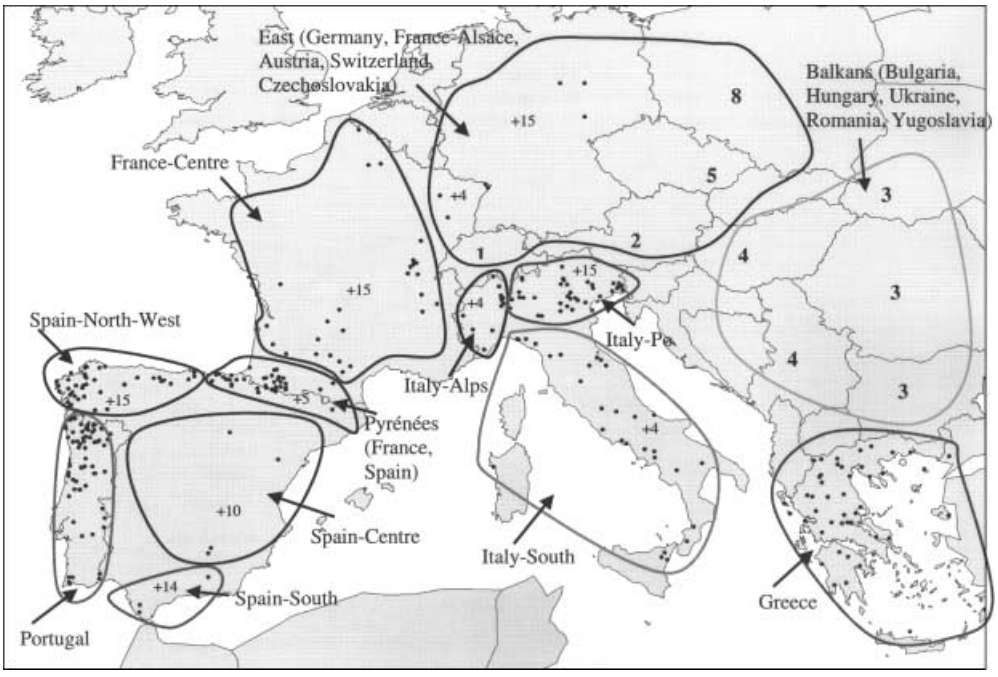

Fig. 1 Location of the European maize populations. *Numbers* indicate the number of populations when geographic coordinates are not available. Geographic groups or subgroups are also indicated

phological variation in populations from four countries of Southeastern Europe: Italy, Hungary, Yugoslavia and Romania. Comparing these populations from Southeastern Europe to populations from Spain, Edwards and Leng (1965) concluded that maize in these two regions derive from different American origins. Pavlicic (1971) also compared populations from Italy, Yugoslavia and Romania. Brandolini realised several syntheses on major European maize races (1969, 1970, 1971). More recently, Rebourg et al. (2001) studied the genetic structure of European populations by means of molecular analysis.

In 1996, seven countries, namely France, Germany, Greece, Italy, Portugal, Spain and the Netherlands, decided to set up a common programme of preservation, evaluation and use of maize landrace genetic resources. This programme, entitled RESGEN CT96-088, was partially funded by the European Union. Its first objective was to establish an exhaustive inventory of the genetic resources held by each country, to describe these resources using ecogeographical passport data and to characterise these resources using primary agromorphological descriptors. A total of 2900 accessions were placed in a European database accessible to the public. Following this, a series of agro-morphological characterisations enabled each country to define a representative national collection of its own populations.

In the study reported here, we analysed the resulting sub-sample of 394 populations listed by the different countries using restriction fragment length polymorphism (RFLP) markers. Compared with genetic variability of morphological traits, molecular polymorphism is generally considered to be independent of the environment. Moreover, RFLP markers have proved to be powerful tools for studying, maize population structures (Dubreuil and Charcosset 1998, 1999). In a separate study, the same populations have been characterised using 16 isozyme loci (Revilla, in preparation). From the results of both RFLP and isozyme analyses, a core collection (Brown 1989), i.e. a sub-collection of 100 accessions representative of the genetic diversity held in the total European maize collection, will be constituted.

The main objectives of this survey were (1) to investigate the genetic diversity and structure of European maize populations and (2) to clarify some historical hypotheses concerning their origin(s).

Materials and methods

Plant material

A total of 488 European maize populations were used in this study (Fig. 1), of which 394 were representative of the collections of six countries involved in the European project: Greece, Italy, Germany, France, Spain and Portugal. Different European institutes provided a representative sample of the diversity of their populations. Details concerning the origin (passport descriptors) and some primary descriptors of this European maize collection are available in the web database (http://www.ensam.inra.fr/gap/resgen88). The remaining 94 populations were previously analysed by Rebourg et al. (2001). These 94 populations, mostly from eastern Europe, were included in the present assay to complete our sample with eastern Europe accessions. The number of populations per country varied from one (Switzerland) to more than 100 (Italy or Spain). In order to compare genetic diversity according to geographic origins, populations were grouped into geographical groups and sub-groups (Fig. 1 and Table 2). A group may consist of populations from several countries (e.g. the Balkans), a single country (e.g. Greece) or a part of a country (e.g. France-Centre) or from several regions in different countries (e.g. Pyrénées). A subgroup may be represented by populations from a country (e.g. Ukraine) or a region (e.g. Spain-North-West).

The whole collection was sown in 1999, with the exception of the 94 populations studied by Rebourg et al. (2001). Two hundred populations were grown at the INRA maize Station in Mauguio

Table 1 Allelic richness, within-population diversity (H_w), total diversity (H_t) and between population diversity (G_{st}) estimated at 23 RFLP loci (sonde-enzyme combination)

Probe	Enzyme	Chromosomic location	Total number of alleles	Average number of alleles per population	H_w	H_t	G_{st}
BNL5.09	*Eco*RI	9	7	1.98	0.34	0.52	0.35
BNL5.10	*Eco*RI	9	18	3.36	0.55	0.77	0.29
BNL7.56	*Hin*dIII	5	4	1.77	0.20	0.32	0.37
BNL8.29	*Eco*RI	1	7	1.44	0.11	0.14	0.20
CSU81	*Hin*dIII	7	10	2.25	0.40	0.54	0.25
NPI270	*Eco*RI	4	18	4.02	0.53	0.78	0.32
NPI406	*Hin*dIII	1	8	1.72	0.17	0.27	0.37
SC322	*Eco*RI	5	25	4.17	0.58	0.81	0.28
UMC10	*Eco*RI	3	15	3.53	0.53	0.79	0.33
UMC19	*Hin*dIII	4	12	1.88	0.22	0.31	0.28
UMC47	*Eco*RI	4	7	1.78	0.18	0.25	0.26
UMC55	*Eco*RV	2	7	1.81	0.29	0.44	0.34
UMC89	*Eco*RV	8	8	2.16	0.30	0.52	0.41
UMC103	*Hin*dIII	8	9	1.88	0.25	0.35	0.30
UMC106	*Eco*RI	1	15	2.80	0.46	0.65	0.29
UMC107	*Hin*dIII	1	7	2.10	0.31	0.45	0.31
UMC132	*Eco*RV	6	3	2.43	0.42	0.65	0.36
UMC161	*Eco*RI	1	5	1.87	0.31	0.44	0.30
UMC168	*Eco*RV	7	15	3.31	0.49	0.74	0.33
BNL5.09	*Hin*dIII	9	16	3.20	–	–	–
BNL6.06	*Hin*dIII	3	18	2.85	–	–	–
BNL14.28	*Hin*dIII	9	14	2.33	–	–	–
UMC15	*Hin*dIII	4	16	2.63	–	–	–
		Mean	11.5	2.49	0.35	0.51	0.31

(near Montpellier, France), 114 at the INRA maize Station of Le Moulon (near Paris, France) and 80 at ISC station of Bergamo (Italy). Molecular analyses were performed in the laboratories of the same three stations. In each laboratory, all 394 populations were analysed using a subset of RFLP markers.

Molecular analyses

RFLP analyses were carried out using a pooled DNA sampling method described in Dubreuil et al. (1999) and Rebourg et al. (1999, 2001). Each population was represented by two DNA bulk samples, each extracted from leaf disks sampled on 15 individuals. DNA was extracted according to Tai and Tanksley (1990), and the samples were digested separately with three restriction enzymes (*Eco*RI, *Hin*dIII and *Eco*RV) and submitted to electrophoreses according to the Southern Blot procedure described by Sambrook et al. (1989). Separate DNA fragments were then vacuum-transferred from gels to nylon membranes.

We used 12 genomic probes UMC (University of Missouri, Colombia, Mo.), seven genomic probes BNL (Brookhaven National Laboratory, Upton, N.Y.), two genomic probes NPI (Native plants, Pioneer Hi-Bred Int) and two cDNA clones. Eight probes were assayed with *Eco*RI, four with *Eco*RV, nine with *Hin*dIII and one with both *Eco*RI and *Hin*dIII. The 23 probe-enzyme combinations and the chromosomic locations of the probes are indicated in Table 1. DNA probes were radiolabelled with [$_{32}$P]-dCTP by random priming synthesis (Feinberg and Vogestein 1983). Hybridisation was performed as described by Church and Gilbert (1984). After washing, nylon membranes were exposed to autoradiographic films.

The autoradiographic films were scanned. The ratio of the optic density of each band to total optic density of bands from the same lane was estimated using image analysis software (RFLP-SCAN, Scanalytics 1991). As all probes were chosen to be monolocus and to yield a single band pattern, the ratio estimated for a band could be interpreted as the allelic frequency of an allele. For each population, we estimated allelic frequencies by the average frequency of the two DNA pools representing the population. The accuracy of this method was established by Dubreuil et al. (1999) and fully implemented by Rebourg et al. (2001). In the present study, this method was applied when the quality of autoradiographic films was good enough to allow a proper use of the scanner, i.e. for 19 probes in all. The allelic nomenclature used was determined by C. Rebourg (C. Rebourg, personal communication).

Data analysis

The number of alleles per locus (further referred to as allelic richness) was determined for the entire collection and for various levels within the collection (population, sub-group, group). The existence of group- or sub-group-specific alleles was determined subsequently. For the 19 loci for which frequency data were estimated, we calculated total genetic diversity (H_T), genetic diversity within populations (H_w) and the proportion of diversity resulting from gene differentiation between populations (G_{st}) according to Nei (1987).

Two types of genetic distances between populations were calculated: (1) the modified Rogers' distance (Rogers 1972; Wright 1978) on frequency data and (2) Nei and Li's distance (Nei and Li 1979) on binary data (presence versus absence of alleles). The standardised Mantel coefficient, derived from the Z-statistics of Mantel (Mantel 1967) was computed to compare the two distances matrices using the Mantel option of GENETIX software (Belkhir 2000).

To investigate the relationships between populations, we carried out a Ward's hierarchical ascendant classification (Ward 1963) using the Cluster procedure from SAS (SAS institute 1989) with the two distances. For each allele, an analysis of variance (ANOVA procedure of SAS) was conducted on allelic frequencies using the main clusters as factor. *F* values were compared in order to determine the more structuring alleles.

In addition, two qualitative (kernel texture and ear conicity) and two quantitative morphological characters (ear row number and accumulative degree-days to female flowering (base 6)) were used to describe phenotypically the different clusters. They are available for each population in the database, given as primary descriptors.

Table 2 Partition of allelic richness among geographic groups and sub-groups

Group	Number of accessions	Total number of alleles	Mean number of alleles per locus	Mean number of alleles per locus and population	Mean percentage of fixed loci[a]	Number of unique alleles[b]
Sub-group						
Europe	488	264	11.48	2.48	19.1	
East	45	150	6.52	2.04	32.6	4
Poland	8	111	4.83	2.55	10.3	1
Czechoslovakia	5	94	4.09	2.10	30.4	1
Austria	2	48	2.09	1.54	56.5	
Germany	22	117	5.09	1.71	46.8	1
Switzerland	1	55	2.39	2.39	17.4	
France-Alsace	7	96	4.17	2.51	9.9	1
Balkans	17	132	5.74	2.36	18.9	4
Ukraine	3	59	2.57	1.71	37.7	
Romania	3	93	4.04	2.65	11.6	2
Bulgaria	3	72	3.13	2.20	23.2	
Yugoslavia	4	87	3.78	2.45	17.4	1
Hungary	4	107	4.65	2.67	8.7	
France-Centre	47	157	6.83	2.50	16.7	3
Italy	106	175	7.61	2.45	20.7	7
Italy-Alps	25	135	5.87	2.55	19.0	2
Italy-Po	46	159	6.91	2.63	15.7	2
Italy-South	35	102	5.37	2.14	28.4	1
Greece	50	152	6.61	2.17	27.2	3
Pyrénées	50	175	7.61	2.71	12.6	2
France-Pyrenees	26	137	5.96	2.64	14.0	
Spain-Pyrenees	24	160	6.96	2.80	11.1	2
Spain	102	185	8.04	2.55	16.9	9
Spain-Northwest	63	162	7.04	2.64	11.4	5
Spain-South	19	143	6.22	2.47	24.2	1
Spain-Centre	20	133	5.78	2.31	27.5	1
Portugal	71	172	7.48	2.75	11.7	9

[a] Percentage of fixed loci was estimated as the average percentage of homozygote loci per population
[b] Unique alleles are specific to a group or a sub-group as compared to the whole populations studied

Correlation between frequencies of structuring alleles and latitude or longitude of collection sites was tested for the 335 populations which had available geographic coordinates.

Results

Allelic richness

We found a total of 264 alleles for the 23 clone-enzyme combinations. The number of alleles per loci varied greatly from 3 (locus UMC132/*Eco*RV) to 25 (locus SC322/*Eco*RI), with an average value of 11.5 (Table 1). With a mean value of 2.49, within-population allelic richness accounted for 22% of the total allelic richness.

Populations from the Eastern and Balkan groups exhibited a lower allelic richness (6.52 and 5.74 alleles per locus, respectively) and a lower number of specific alleles (a total of eight) than populations from Southern groups (Table 2). The percentage of fixed loci per population was also highest (33%) in the Eastern group. In the France-Centre group and in the Southern groups (Italy, Greece, Pyrenees, Spain and Portugal), allelic richness ranged from 6.61 alleles per locus for the Greek populations to 8.04 for the Spanish populations, and the total number of specific alleles was 33.

Allelic diversity

Allelic frequencies could be determined for only 19 clone-enzyme combinations. We therefore calculated genetic diversity within and among populations only for these loci (Table 1). Total genetic diversity varied greatly among loci from 0.14 at locus BNL8.29/*Eco*RI to 0.81 at locus SC322/*Eco*RI, with an average value of 0.51. Within-population genetic diversity varied for the same loci from 0.11 to 0.58, with an average value of 0.35. The corresponding G_{st} value, which accounts for the proportion of between-population differentiation within total differentiation, was 31%.

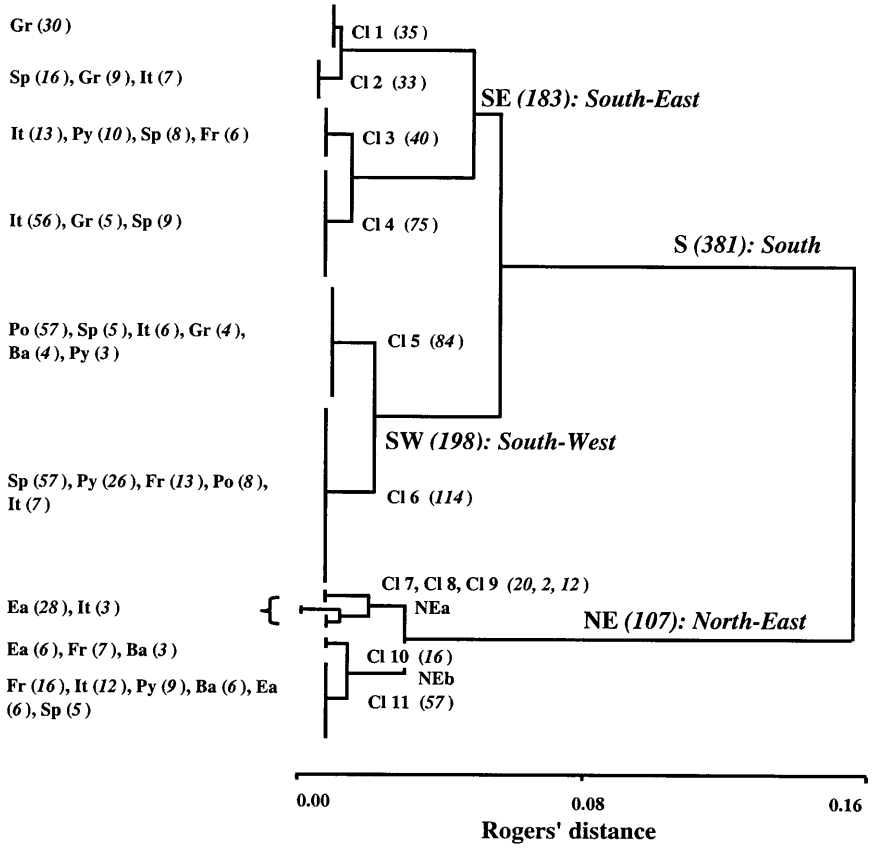

Fig. 2 Cluster analysis of the 488 European populations: Ward's classification based on Rogers' distances. The number of accessions per cluster are indicated in *italics*. *Cl* Cluster. The geographical origin of populations within each cluster is indicated. *Ea* East, *Ba* Balkans, *Fr* France-Centre, *Py* Pyrenees, *It* Italy, *Gr* Greece, *Sp* Spain, *Po* Portugal. For a description of the geographical groups, see Table 2. Only origins with three accessions or more within the cluster are presented

Relationships between populations

Roger's distance between populations ranged from 0.009 between two Italian populations to 0.655 between a Spanish (from Spain-Centre group) and a German population. Using Nei's index, the distance between populations ranged from 0.048 between two Portuguese populations to 0.737 between the Spanish population mentioned above and another German population.

A Mantel procedure was used to test the correlation between the genetic distance of Rogers and the genetic distance of Nei and Li. The two distances were highly correlated (correlation coefficient of 0.75, $P < 0.001$). Only results obtained using Rogers' distances will therefore be presented further. The cluster analysis (Fig. 2) first underlined a major differentiation between populations from the Northeastern (NE, 107 populations) and southern Europe (S, 381 populations) main clusters. The main cluster NE is principally made up of populations from north and northeast Europe: populations from the Balkans, East and France-Centre groups account for 70% of its size. It is divided into two clusters (NEa and NEb) and then into five sub-clusters (7, 8, 9, 10 and 11). In cluster NEa (sub-clusters 7, 8 and 9) 88% of populations originates from the north and the northeast of Europe, whereas these populations represent 60% of cluster NEb (sub-clusters 10 and 11). Main cluster S consists of 91% of populations from the south of Europe – i.e. populations from Spain, the Pyrenees, Portugal, Greece and Italy groups. It is separated into two clusters (SE and SW) and six sub-clusters (1, 2, 3, 4, 5 and 6). Cluster SE (sub-clusters 1, 2, 3 and 4) is made up of 67% of populations from the Greece and Italy groups i.e. population from southeastern Europe. Cluster SW (sub-clusters 5 and 6) consists of 79% of populations from Spain, Portugal and Pyrenees groups, i.e. populations from southwestern Europe.

Morphological variation in the different clusters

Morphological variation for four characters was observed in detail (Table 3) in the 11 sub-clusters described in Fig. 2. Sub-clusters 1, 2, 3 and 4, i.e. the South-east cluster (SE), are made up of late populations, part of which have dent kernels (Greek, Italian or Spanish populations). Some Greek populations of sub-cluster 1 present very conical ears.

Populations of sub-clusters 5 and 6, i.e. the Southwest cluster (SW), are rather early with flint kernels. Conical ears are observed in Portuguese populations of sub-cluster 5. In the North-east main cluster (NE), the earliness ranges from early in sub-clusters 7, 8 and 9 to late in sub-cluster 11, with intermediary populations in sub-cluster 10. Most populations are flint except some Italian dent populations of sub-cluster 11. Apart from some German populations of sub-cluster 9 showing conical ears, ears are cylindrical with a variable number of

Table 3 Morphological variations at two quantitative characters [degree-days to female flowering (base 6) and ear row number] and two qualitative characters (ear conicity and texture of kernel). *Bold types* indicate extreme values for quantitative characters. For conicity, a third intermediary class completes the percentages to 100; for texture, flint type, which is the more current for European maize populations, completes the percentages to 100

	Sub-cluster		Degree-days to female flowering		Ear row number		Conicity		Texture Dent (%)
	No	Number of accessions	Mean	Standard deviation[a]	Mean	Standard deviation[a]	Conical (%)	Cylindrical (%)	
SE	1	35	973.5	74.7	11.1	2.7	31	9	20
	2	33	**1039.4**	15.7	11.0	2.4	9	48	33
	3	40	978.9	152.7	12.6	3.1	5	67	15
	4	75	1011.6	120.4	13.3	2.9	8	53	7
SW	5	84	866.3	60.7	12.7	2.2	30	26	3
	6	114	867.1	145.7	12.3	3.6	5	53	5
NE	7	20	813.8	128.5	11.0	3.6	5	90	0
	8	2	**734.3**	127.3	**8.1**	0.0	0	100	0
	9	12	796.8	148.8	16.7	7.1	25	58	0
	10	16	890.5	166.5	**17.7**	6.2	6	81	6
	11	57	965.9	139.5	13.9	3.7	0	84	16
		Mean	925.3	150.4	12.8	3.6			

[a] Standard error deviation of mean estimates

Table 4 Alleles correlated with latitude and/or longitude coordinates of 335 of the 488 European populations studied

Alleles correlated with									
	Latitude		Longitude					Latitude and longitude	
	BNL7.56/ *Hin*dIII Allele 2	NPI270/ *Eco*RI Allele 5	UMC10/ *Eco*RI Allele 3	UMC161/ *Eco*RI Allele 2	UMC161/ *Eco*RI Allele 1	BNL8.29/ *Eco*RI Allele 3	SC322/ *Eco*RI Allele 4	SC322/ *Eco*RI Allele 1	BNL5.10/ *Eco*RI Allele 6
R^2	0.16	0.11	0.25	0.21	0.17	0.15	0.10	0.24/0.15	0.19/0.14
P	**	*	**	**	**	**	*	**/**	**/**
Frequency in main clusters									
NE	0.18	0.65	0.40	0.33	0.66	0.02	0.03	0.51	0.10
SE	0.03	0.20	0.18	0.15	0.82	0.03	0.04	0.22	0.46
SW	0.01	0.28	0.45	0.41	0.59	0.12	0.22	0.24	0.21

*P<0.05, **P<0.01

rows (from 8.1 to 17.6 in sub-clusters 8 and 10, respectively).

Geographical variation in allelic frequencies

For the whole set of European populations the classification shows a north-south (main cluster NE versus main cluster S) genetic structure and an east-west (cluster SE versus cluster SW) genetic structure for the Southern populations only. We thus decided to conduct an analysis of variance on allelic frequencies in order to determine which alleles discriminate these three major clusters.

Sixty alleles appeared as discriminating ($P < 0.001$, data not shown). Among these, the 20 alleles with the lowest P values were tested for their correlation with geographical coordinates. Two alleles were significantly correlated with latitude and five with longitude and two with both latitude and longitude (Table 4). Allelic frequencies for these alleles in the three major clusters are presented in Table 4.

Discussion

The study context

Most genetic research on maize has concerned inbred lines and their assignation to heterotic groups or pedigree relationships (e.g. Melchinger et al. 1991, 1992; Livini et al. 1992; Messmer et al. 1993; Mumm and Dudley 1994; Dubreuil et al. 1996). Only a few studies have been done on the genetic diversity and structure of maize collections. During the last decade a few isozyme studies describing relationships among restricted parts of maize populations have been made on national collections. Lefort-Buson et al. (1991) and Garnier (1992) characterised 115 and 65 populations, respectively, of the French

INRA-PROMAIS collection. Revilla et al. (1998) characterised 47 populations of an extensive Spanish collection. The present work uses a much larger sample than previous studies and uses a more powerful molecular technique. Comparing the efficiency of isozymes and RFLP to study genetic diversity within and among ten maize populations, Dubreuil and Charcosset (1998) showed the superior discriminative ability of RFLP data. The use of RFLP for large-scale molecular evaluations of genetic diversity in populations formerly required expensive and time-consuming effort. The use of bulk analysis (Michelmore et al. 1991) and its use for maize diversity analysis with RFLP (Dubreuil et al. 1999) enabled larger sized samples to be analysed. Using bulk RFLP method, Rebourg et al. (1999, 2001) described genetic relationships among the 65 populations previously studied by Garnier (1992) and among 131 other European populations. Finally, bulk RFLP analyses have been efficiently used to characterise genetic diversity among the 488 European maize populations reported in the present paper.

Molecular diversity of European populations

The present study showed a higher molecular allelic richness (11.5 alleles per locus) than have previous studies of European maize populations. Working on 131 European maize populations for the same 23 loci as described here, Rebourg et al. (2001) found an average number of 9.1 allele per locus; i.e., 2.4 alleles fewer per locus, on average. Our G_{st} value (31%) was lower than that found in Rebourg et al. (2001) (36%) but similar to that previous studies that used a lower number of populations based on RFLP (Dubreuil and Charcosset 1998; Rebourg et al. 1999) or isozymes (Garnier 1992). This difference with Rebourg et al. (2001) may result from a difference in sampling strategy. The latter sampled relatively fewer populations than we did in the Southern regions, which displayed the highest within-population diversity. This sample should have increased the relative genetic differentiation between populations.

Allelic richness was lower in northeastern Europe (6.52 allele per locus for the East group and 5.74 allele per locus in the Balkans group) than in southern Europe (8.04 allele per locus in Spain and 7.61 allele per locus in Italy). As previously stated by Rebourg et al. (1999, 2001), this result suggests that southern Europe was the location of most of the introductions of maize into Europe, or (and) that northern introductions had a lower genetic diversity than southern introductions. A high selection pressure for adaptation to a cold climate may also have contributed to the decrease of northern European maize diversity. In addition to this north versus south differentiation, we found a differentiation between southwestern and southeastern Europe, which suggests that the origins of maize are not the same in the two regions. This hypothesis is supported by the existence of numerous different unique alleles in each region (ten in Italy and Greece and 18 in Spain and Portugal).

Among the north-eastern populations, those from Germany and Austria display a particularly low within-population polymorphism: 1.7 and 1.5 alleles per locus and population, respectively. These countries are characterised by a high number of fixed loci (57% for German populations and 47% for Austrian populations). Among the southern populations, Greek populations show higher rates of gene fixation and lower scores for total number of alleles (6.6), number of alleles per locus and populations (2.17) and for number of unique alleles (3). As previously discussed by Sabounat and Pernès (1986) and Rebourg et al. (2001), it is possible that these populations were multiplied in a manner that increased inbreeding. Moreover, according to Brandolini (1970), maize was not introduced into Greece directly from America but more probably derived from Balkan populations. Such an historical trajectory of introduction may also explain the relatively low diversity of Greek populations. In contrast, allelic richness and number of unique alleles is maximum in Portugal, Spain and Italy, where several great navigators are known to have introduced maize populations from America during the 16th century (Revilla et al. 1998). Alternatively, the high allelic richness observed in Southern Europe may be due to the large extent of maize growing in this region for several centuries.

Genetic structure of European populations

Even if the hierarchical classification of the European populations of the present study (Fig. 2) was obtained with only 19 loci and 200 alleles, it is in agreement with the previous work by Rebourg et al. (2001) based on 29 loci and 278 alleles. The latter study suggested five major groups that could be considered to be races: German Flint, Northeastern European Flint, Southern European Flint, Italian Flint and Pyrenees Galice Flint. German Flint populations of Rebourg et al. (2001) are included in sub-clusters 7, 8 and 9 of the present study, which all belong to main cluster NE. Northeastern European Flint populations are mainly included in sub-cluster 11, which also belongs to main cluster NE. Southern European Flint populations are mostly included in sub-clusters 5 and 2, which belong, respectively, to main clusters SW and SE. This separation into two clusters is likely due to the larger number of southern populations in the present study (see discussion below). Italian Flint populations are included in sub-clusters 3, 4 of cluster SE. Pyrenees Galice Flint mostly belong to sub-cluster 6 of cluster SW.

The 19 loci appear to be sufficient to investigate the genetic structure in a large set of populations. Investigating the sampling variance of the RFLP data set in maize, Tivang et al. (1994) found that the number of bands required for a coefficient of variance of 10% was 388, 150 and 38 for closely, intermediately and distantly related inbred, respectively. The existence of 200 bands would thus be sufficient to characterise 488 populations – i.e. relatively distantly related entities.

In addition to the confirmation of a north-south structure of diversity, our results underline a clear east-west genetic structure of southern populations, that was not evident in the study by Rebourg et al. (2001). This difference could be due to the sampling strategy used in the present study, which permitted the analysis of representative samples of Greek and Portuguese populations [50 and 71 populations in the present study, respectively, when compared to 0 and 1 in Rebourg et al. (2001)]. The observed genetic structure agreed with the geographical grouping of populations into groups and sub-groups. Of the early and flint populations from the East and Balkans groups, 80% were joined into the NE main cluster. Half of the France-Centre group and 26% of the populations from the north of Italy (Italy-Po) also joined this cluster. The intermediary location of these French and Italian populations may explain their genetic and morphological affinities with northeastern populations. Ninety percent of Greek and 74% of Italian populations grouped together in cluster SE defining a late, flint-dent cluster. Eighty seven percent of the Spain-North-West (Galicia), 92% of Portuguese and 79% Spain-Pyrenean populations grouped together into cluster SW. These populations are on average flint and early (Table 3), which can be explained by the wet Atlantic climatic conditions of these regions. Populations from other Spanish sub-groups (Spain-South or Spain-Centre) preferentially joined (79%) cluster SE. As compared to Spain-North-West, these sub-groups were shown to be more fixed, with a lower allelic richness. Climatic conditions in these regions include a summer drought, which may explain their similarity with late populations from the southeast region. Populations from France–Pyrenees, France-Centre and Italy-Po tend to be distributed among the three main clusters, which suggests mixed origins and possible hybridisation phenomena in these regions. Moreover, the presence of dent populations in sub-clusters 1, 2, 3 and 11 indicates a probable introgression of recently introduced dent races into older races.

A high correlation was obtained when matrices based on Nei and Li distances and Rogers distances were compared (R=0.75 in Mantel's test), suggesting that both distances give similar estimates of genetic relationships among the accessions tested. Such a result means that populations or groups of populations are differentiated not only by contrasting allele frequencies but also by different alleles.

The origins of European maize and its present genetic structure

We found a correlation between the frequency of seven structuring alleles with latitude and/or longitude (Table 4). Such a tendency may indicate that these probes, supposed to be neutral, are submitted to selection effects or linked to selected genes on a chromosome. However, no referenced quantitative trait loci (QTL) were found to be associated with these structuring RFLP loci in the Maize Genome Database (Maize DB on http://www.agron.missouri.edu/). On the other hand, such gradual variation in frequency with latitude and/or longitude may reflect the direction of gene flow between different races of maize introduced into different European locations. We have thus compared the frequency distributions of these structuring alleles in the European populations and several representative American origins (Rebourg 2000). Indeed, three alleles: allele 5 of locus NPI270/*Eco*RI, allele 1 of locus SC322/*Eco*RI and allele 6 of locus BNL510/*Eco*R1 were correlated with latitude. The first two were highly frequent in the East and North-Balkan groups, moderately frequent in the France-Centre group and lowly frequent in the Southern groups. They were also present in high frequencies in American Northern-Flint. In contrast, allele 6 of locus BNL510/*Eco*R1, which shows a very low frequency in Northern-Flint and higher frequencies in South American, Andean and Caribbean materials was only found in high frequencies in populations from the South of Spain, Portugal, Italy and Greece. The contribution of American Northern-Flint material to the genetic basis of European was shown by Rebourg (2000). However, it would be necessary to increase the number of American populations per origin and the number of American origins themselves to investigate in more detail hypotheses concerning the origins of European maize.

Conclusion

A study of cytoplasmic markers like ribosome or chloroplast DNA, which are less polymorphic than nuclear DNA, appears to be necessary to study more precisely the different American origins of European maize. A long period of isolation followed by multiple events of hybridisation is likely to have complicated the genetic constitution of these European populations. Nevertheless, the present study confirms that European maize has multiple origins, south-eastern populations being clearly distinct from southwestern ones as previously suggested by the morphological studies of Leng et al. (1962). The conservation of these traditional populations appears to be of crucial importance as a reservoir of genetic resources. In this way, the agronomic evaluation of 100 of these populations selected as representative of the genetic variability of the whole collection ('Core collection') has been started in April 2001 in the different countries included in the original European maize project (RESGEN CT96-088). This more complete evaluation (associating genetic markers to agronomic evaluation) will probably lead to the use of the most interesting populations in breeding programmes.

Acknowledgements This research was partially supported by the European Union (contract: RESGEN CT96-088). We gratefully acknowledge J. Moreno-Gonzales, I. Brichette-Mieg, A. Ordas, R.A. Malvar, P. Revilla, A. Alvarez, M. Motto, E. Lazarou-Kallinaki, G. Evgenidis, A.M. Barata, R.M. Farias, P. Noël, P. Ruaud, T.J.L. Van Hintum and A. Börner, for providing plant material and for

helpful discussions concerning plant germplasm. We are deeply indebted to C. Tollon, M. Merlino, A. Chitto, S. Santoni, V. Combes and F. Dumas for expert technical support in the molecular analyses, to G. Lebfevre for cartographic help, to J. Ronfort for statistical help and to T. Bataillon and J. Thompson for their careful reading of the manuscript.

References

BelKhir K (2000) GENETIX 4.01. Belkhir Biosoft, Montpellier

Brandolini AG (1969) European races of maize. Annu Corn Sorghum Res Conf Proc 24:36–48

Brandolini AG (1970) Razze Europee di mais. Maydica 15:5–27

Brandolini AG (1971) Preliminary report on south European and Mediterranean maize germplasm. In: Kovacs I (ed) Proc 5th Meet Maize Sorghum Sect Eucarpia. Hungarian Academy of Sciences. Budapest, pp 108–116

Brandolini AG, Mariani G (1968) Il germosplama italiano nella fase attuale del migliorgamento ganetice del mais. Genet Agr 22:189–206

Brown AHD (1989) Core collections: a practical approach to genetic resources management. Genome 31:818–824

Church MG, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Costa-Rodrigues L (1971) Races of maize in Portugal. Agron Lusit 31:239–284

Dubreuil P, Charcosset A (1998) Genetic diversity within and among maize populations: a comparison between isozyme and nuclear RFLP loci. Theor Appl Genet 96:577–587

Dubreuil P, Charcosset A (1999) Relationships among maize inbred lines and populations from European and North-American origins as estimated using RFLP markers. Theor Appl Genet 99:473–480

Dubreuil P, Dufour P, Krejci E, Causse M, De Vienne D, Gallais A, Charcosset A (1996) Organization of RFLP diversity among inbred lines of maize representing the most significant heterotic groups. Crop Sci 36:790–799

Dubreuil P, Rebourg C, Merlino M, Charcosset A (1999) Evaluation of DNA pooled-sampling strategy for estimating the RFLP diversity of maize populations. Plant Mol Biol Rep 17:123–138

Edwards RJ, Leng ER (1965) Classification of some indigenous maize collections from Southern and South eastern Europe. Euphytica 12:161–169

Feinberg A, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Garnier P (1992) Contribution à l'étude de la variabilité génétique inter et intra population chez le maïs (*Zea mays* L.): valorisation d'informations agromorphologiques et enzymatiques. PhD thesis, INAPG

Geric I, Zlokolica C, Geric C, Stuber C (1989) Races and populations of maize in Yugoslavia. Isozyme variation and genetic diversity. IBPGR, Rome

Gouesnard B, Dallard J, Panouille A, Boyat A (1997) Classification of French maize populations based on morphological traits. Agronomie 17:491–498

Lefort-Buson M, Lavergne V, Daudin JJ, Charcosset A, Sampoux JP, Gallais A (1991) Genetic variability among populations of maize germplam. 2. enzymatic polymorphism and its relationship to quantitative trait diversity. Maydica 36:237–246

Leng ER, Tavcar A, Trifunovic V (1962) Maize of Southeastern Europe and its potential value in breeding programs elsewhere. Euphytica 11:263–272

Livini C, Ajmone-Marsan P, Melchinger AE, Messmer M, Motto M (1992) Genetic diversity of inbred lines within and among heterotic groups revealed by RFLPs. Theor Appl Genet 84:17–25

Mantel N (1967) The detection of disease clustering and a generalized approach. Cancer Res 27:209–220

Melchinger AE, Messmer MM, Lee M, Woodman WL, Lamkey KR (1991) Diversity and relationships among U.S. inbreds revealed by restriction fragments length polymorphisms. Crop Sci 31:669–678

Melchinger AE, Boppenmaier J, Dhillon BS, Pollmer WG, Herrmann RG (1992) Genetic diversity for RFLPs in European maize inbreds. II: Relation to performance of hybrids within versus between heterotic groups for forage traits. Theor Appl Genet 84:672–681

Messmer MM, Melchinger AE, Herrmann RG, Boppenmaier J (1993) Relationships among early European maize inbreds. 2. Comparison of pedigree and RFLP data. Crop Sci 5:944–950

Michelmore RW, Paran AE, Kesseli RV (1991) Identification of markers linked to desease resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Mumm RH, Dudley JW (1994) A classification of 148 U.S. maize inbreds. 1. Cluster analysis based on RFLPs. Crop Sci 4:842–851

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Pavlicic J (1971) Contribution to a preliminary classification of European open pollinated maize varieties. In: Kovacs I (ed) Proc 5th Meet Maize Sorghum Sect Eucarpra. Hungarian Academy of Sciences. Budapest, pp 93–107

Pavlicic J, Trifunovic V (1966) A study of some important ecologic corn types grown in Yugoslavia and their classification. J Sci Agric Res 19:44–62

Rebourg C (2000) Diversité génétique de populations européennes et américaines de maïs: Analyse moléculaire et morphologique. PhD thesis, Université Paris XI, Paris

Rebourg C, Dubreuil P, Charcosset A (1999) Genetic diversity among maize populations: bulk RFLP analysis of 65 accessions. Maydica 44:237–249

Rebourg C, Gouesnard B, Charcosset A (2001) Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. Heredity 86:574–587

Revilla P, Soengas P, Malvar RA, Cartea ME, Ordas A (1998) Isozyme variation and historical relationships among the maize races of Spain. Maydica 43:175–182

Rogers JS (1972) Measures of similarities and genetics distances. In Wheeler MR (ed) Studies in genetics VII. University of Texas, Austin, pp 145–153

Salanoubat M, Pernès J (1986) Enzyme polymorphims within and between European maize populations. Maydica 31:269–278

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn, Cold Spring Harbor Press, Cold Spring, Harbor, N.Y.

Sanchez-Monge E (1962) Razas de maiz en Espana. Ministerio de Agricultura, Madrid

SAS (1989) SAS, 6.03. SAS Istitute, N.C.

Scanalytics (1991) RFLPSCAN 2.01. Bionis, Fairfax, Va.

Tai TH, Tanksley SD (1990) A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. Plant Mol Biol Rep 8:297–303

Tivang JG, Nienhuis J, Smith OS (1994) Estimation of sampling variance of molecular-marker data using the bootstrap procedure. Theor Appl Genet 89:259–264

Ward JH (1963) Hierarchical grouping to optimize an objective function. Am Stat Assoc J 56:236–244

Wright S (1978) Evolution and genetics of populations. University Chicago Press, Chicago