Heredity **81** (1998) 100–110 *Received 7 November 1997, accepted 19 January 1998*

Genetic differentiation within and between natural populations of perennial and annual ryegrass (*Lolium perenne* and *L. rigidum*)

FRANÇOIS BALFOURIER*, GILLES CHARMET & CATHERINE RAVEL

INRA, Station d'Amélioration des Plantes, 234 avenue du Brézet, 63039 Clermont-Ferrand Cedex, France

Genetic structure of 120 wild populations of *Lolium perenne* and 50 populations of *L. rigidum* was studied using starch gel electrophoresis. Allelic frequencies were obtained from 12 polymorphic isozyme loci. Gene diversity indices (number of alleles (A), observed (H_o) and expected (H_e) heterozygosity) were significantly higher in *L. rigidum* ($A = 3.13$; $H_o = 0.369$; $H_e = 0.405$) than in *L. perenne* ($A = 2.72$; $H_o = 0.308$; $H_e = 0.322$). For both species, most of the diversity appeared to be within populations ($G_{ST} = 0.110$ and 0.170 for *L. perenne* and *L. rigidum*, respectively). Despite this weak genetic differentiation, significant patterns of geographical variation for diversity indices and allele frequencies were observed in *L. perenne* populations; the three genetic indices (A, H_o, H_e) showed the same trend of variation, with the lowest values in the north-west part of the distribution area (United Kingdom, Ireland) and the highest ones in the south-east (Turkey, Lebanon, Cyprus, Iraq, Iran). In the same way, as indicated by logistic regression analyses between allelic frequencies and geographical data of *L. perenne* populations, the latitudinal gradient of allelic frequencies appeared to be more pronounced, although significant relationships also existed with longitude. In contrast, no spatial organization of the diversity was detected in *L. rigidum*. Hypotheses concerning the taxonomic relationships and the genetic and geographical origins of the two species are discussed. *Lolium perenne* could be derived from a small bottleneck of *L. rigidum* populations in the Middle East, and its present distribution area in Europe could be explained either by the extension of primitive agriculture from the fertile crescent, or as a consequence of post-glacial recolonization from southern refugia.

Keywords: allozymes, clines, colonization, forage grasses, geographical structure, population genetics.

Introduction

Lolium perenne and *L. rigidum* are self-incompatible outbreeding species (Terrell, 1968): perennial ryegrass (*L. perenne*) is one of the most important forage grasses commonly cultivated in Europe for permanent pasture and amenity grassland, whereas *L. rigidum* is an annual species mostly used in Australia for winter grazing. These two species are still widespread as natural populations: *L. perenne* is found in most of Europe, and part of the Mediterranean and Middle East area, whereas *L. rigidum* is distributed all around the Mediterranean. They are interfertile species; consequently, gene flow may occur between the two species in the limited area where they are sympatric.

Allozyme diversity studies in *Lolium* have already been carried out on limited regional samples, e.g. 37 populations from the United Kingdom (Hayward, 1985), 60 populations from France (Charmet *et al*., 1993), 24 from Italy and Corsica (Balfourier & Charmet, 1994) and, more recently, a sample of 20 populations from Great Britain and Central Europe (Fernando *et al*., 1997). However, these studies only concerned *L. perenne* and thus they did not provide any insight about the origin of this species and its relationships with *L. rigidum*.

The aim of this paper is to analyse the genetic diversity of the two species, using isozymes, in a large sample of natural populations originating from

*Correspondence. E-mail: balfour@clermont.inra.fr

the whole distribution area, in order to compare the amount and the distribution of their genetic diversity. The results are used to explain the evolutionary relationships between these two *Lolium* species, and to provide some hypotheses concerning their genetic and geographical origins. A better knowledge of the spatial organization of genetic diversity is of interest for a more efficient collection and preservation of natural populations as genetic resources.

Materials and methods

Plant material

One hundred and twenty populations of *L. perenne* and 50 of *L. rigidum* were used in the present study; the geographical origin of *L. perenne* populations ranged from Ireland to Afghanistan and from Norway to Spain, whereas *L. rigidum* populations were sampled from the Canary Islands to Afghanistan and from all around the Mediterranean (Fig. 1). Most populations were collected as bulked seed from at least 50 plants taken from an ecologically homogeneous area of 100–1000 m^2. Geographical data (latitude, longitude) were recorded for each population and are available on request. For each species, populations were clustered into groups according to their geographical proximity: 12 groups for *L. perenne* populations and six groups for *L. rigidum* were chosen as a compromise between number of populations and geographical proximity.

Electrophoretic procedures

On average, 75 plants were studied from each of these 170 populations. Isozyme electrophoretic techniques followed those of Hayward *et al*. (1995). Slices of two starch gels and two different buffer systems were used and permitted the study of 10 enzyme systems, giving 12 readable loci: Histidine/citrate buffer for acid phosphatase (ACP, EC 3.1.3.2, two loci), isocitrate dehydrogenase (IDH, EC 1.1.1.42), malate dehydrogenase (MDH, EC 1.1.1.37), phosphoglucoisomerase (PGI, EC 5.3.1.9), phosphoglucomutase (PGM, EC 5.4.2.2), peroxidase (PRX, EC 1.11.1.7) and shikimate dehydrogenase (SKD, EC 1.1.1.25); Tris-citrate/lithium-borate buffer for glutamic-oxaloacetic transaminase (GOT, EC 2.6.1.1, two loci), diaphorase (DIA, EC 1.8.1.4) and superoxide dismutase (SOD, EC 1.15.1.1).

Fig. 1 Map of collection sites of *Lolium perenne* (▲) and *L. rigidum* (△) accessions.

Genetic diversity analysis

Allele frequencies were determined by direct allele counting. Standard statistics for characterizing genetic variability were computed for all accessions using the 's' programming environment (Becker *et al*., 1988). Population genetic statistics were computed for each species, on a population basis: mean number of alleles per locus (A), observed heterozygosity (H_o) and expected heterozygosity under panmixia (H_e).

The following Nei's genetic diversity statistics were also calculated for each species: the total gene diversity (H_T), subdivided into two components (H_S and D_{ST}, $H_T = H_S + D_{ST}$, Nei, 1973) where H_S is the within-population gene diversity and D_{ST} the between-population gene diversity, and the fixation index (F_{IS}) of individuals relative to each species' subpopulations ($F_{IS} = (H_T - H_S)/H_T$; Nei, 1977). Finally, the coefficient of differentiation among populations $G_{ST} = D_{ST}/H_T$ was also determined for each species. Confidence intervals, for mean values of each parameter, were calculated empirically by a resampling procedure over populations (1000 bootstrapping samples).

For each geographical group, gene diversity parameters (A, H_o, H_e) and the coefficient of differentiation within geographical groups were computed by using separately H_T and H_S within each group. Because of different numbers of populations (n) per group, we calculated absolute differentiation (D_m) (Nei, 1977): $D_m = (n/(n-1))\ (H_T - H_S)$. The coefficient of absolute gene differentiation (G'_{ST}) is then: $G'_{ST} = D_m/(H_T + D_m)$.

Geographical variation of gene diversity and allele frequencies

Overall geographical differences were evaluated by one-way analysis of variance of gene diversity indices corresponding to the different geographical groups of each species.

Linear correlation coefficients between gene diversity indices and geographical data were also calculated in order to identify geographical trends in variation. Then, in order to investigate relationships between allelic frequencies and geographical data, logistic regression analysis was carried out (Collett, 1991).

Modified Rogers's distances (Wright, 1978) were calculated between pairs of populations and among the 18 geographical groups of populations. The matrix of genetic distances between groups was then illustrated by means of a UPGMA dendrogram (Sneath & Sokal, 1973). All these analyses were carried out using the 's' programming environment (Becker *et al*., 1988).

Results

Genetic diversity indices

The mean values of genetic diversity indices at the level of each species are presented in Table 1: A, H_o, H_e and H_T are significantly higher in *L. rigidum* than in *L. perenne*, indicating greater diversity within the annual species.

The partition of the total gene diversity (H_T) within and between populations indicates that the

Table 1 Mean values of genetic diversity indices calculated over all populations of *Lolium perenne* and *L. rigidum*

	L. perenne	*L. rigidum*
Number of populations	120	50
Number of loci	12	12
A	2.723 (2.664–2.782)	3.129 (3.031–3.227)
H_o	0.308 (0.299–0.317)	0.369 (0.355–0.384)
H_e	0.322 (0.315–0.333)	0.405 (0.394–0.424)
H_T	0.362 (0.352–0.372)	0.488 (0.468–0.504)
G_{ST}	0.110 (0.075–0.145)	0.170 (0.127–0.213)
F_{IS}	0.043 (0.006–0.080)	0.089 (0.042–0.136)

A, number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; H_T, total gene diversity; G_{ST}, coefficient of gene differentiation over all populations; F_{IS}, within-population fixation index. Values in parentheses indicate confidence intervals obtained by resampling procedures (1000 bootstrapping samples).

overall population gene differentiation is smaller for *L. perenne* ($G_{ST} = 11\%$) than for *L. rigidum* ($G_{ST} = 17\%$). For both species, most of the diversity appears to be within populations.

Finally, the fixation indices (F_{IS}) are significantly positive for both species, more so for *L. rigidum*, corresponding to a deficit of heterozygotes.

Allele frequencies

For the 12 loci analysed, a total of 57 and 61 alleles were revealed in *L. perenne* and *L. rigidum* populations, respectively. Data are not presented here but are available on request. Allele nomenclature is that of Hayward *et al*. (1995).

Large differences were detected in the allele frequencies among populations at all loci analysed. With nine alleles, the *Pgi2* locus is the most polymorphic, whereas other loci showed three to six different alleles, with one or two very frequent, and one or two much rarer alleles.

For *L. perenne*, 23 alleles can be considered as common and widespread, according to the classification of Brown (1978), 10 as rare and widespread (mean frequency less than 5% but present in more than half of the populations) and 24 as rare and sporadic. For *L. rigidum*, we observed 34 common alleles, only two rare widespread and 25 rare and sporadic alleles. Four alleles, rare in *L. rigidum* (*Got3–10*, *Idh1–10*, *Dia1–50* and *Mdh1–05*) were not detected in *L. perenne* populations. Rare alleles in *L. rigidum* were generally also rare in *L. perenne*. In contrast, rare widespread or rare and sporadic alleles in *L. perenne* were generally more frequent in *L. rigidum*, with the exception of five alleles (*Pgi2–10*, *Pgi2–45*, *Acp1–60*, *Got3–50* and *Idh1–50*) which were rarer in *L. rigidum*.

Geographical distribution

Analyses of variance between the geographical groups in each species revealed that there are significant group differences in the three *L. perenne* gene diversity indices (A, H_o, H_e), but only allelic richness is significantly different in *L. rigidum* (Table 2). The correlations between the three gene diversity indices of *L. perenne* and geographical data of individual populations indicate that variation in these indices may be geographically distributed. The three parameters show a similar trend of variation: perennial ryegrass populations from the north-west part of the distribution area show the lowest values of allelic richness and lowest heterozygosity. In contrast, populations from the south-east have the highest values of gene diversity indices. No significant correlation with geographical data was observed for *L. rigidum*.

Genetic diversity and differentiation parameters for geographical groups of both species are given in Table 3. Comparisons among means can reveal spatial organization of diversity. For instance, mean H_o in *L. perenne* appears to show a clinal trend from the north to the south of the distribution area. Observed heterozygosity is lower in groups 11, 1 and 5 (northern Europe), increases in central and eastern countries of Europe, and is highest in Mediterranean areas (groups 2, 3 and 12). The spatial distributions of the other indices of *L. perenne* are quite similar and show the same trend of variation, indicating that the centre of diversity of this species is probably located in the south-east part of the Mediterranean area (group 12), as confirmed by the highest values of both G'_{ST} and H_T for this group. Spatial distribution of genetic indices is not evident for *L. rigidum*, and group differences are only significant for allelic richness.

Table 2 ANOVA computed over geographical groups and correlation between geographical data and gene diversity indices for *Lolium perenne* and *L. rigidum*

	L. perenne			*L. rigidum*		
	A	H_o	H_e	A	H_o	H_e
F-test	4.28	11.01	9.67	4.79	0.53	0.86
P-values	0.0001	0.0001	0.0001	0.0014	0.7514	0.5133
Longitude	0.099	0.371**	0.346**	0.154	0.112	0.009
Latitude	−0.250*	−0.500**	−0.481**	−0.089	0.007	0.047

A, number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity. Correlations significant at $^*P < 0.01$ and $^{**}P < 0.001$.

Table 3 Mean values of genetic diversity and differentiation parameters within geographical groups of *Lolium perenne* and *L. rigidum*

Species	N° g	Countries, regions of countries	*n*	*A*	H_o	H_S	H_T	G'_{ST}
L. perenne	1	Germany, Austria, Switzerland	9	2.435	0.268	0.290	0.306	0.065
	2	Italy (centre and south)	8	3.042	0.342	0.363	0.385	0.069
	3	Corsica, Sardinia	12	2.792	0.352	0.359	0.390	0.084
	4	Italy (north), France (south-east)	12	2.549	0.286	0.296	0.327	0.099
	5	United Kingdom, Ireland	7	2.560	0.269	0.280	0.304	0.094
	6	France (west)	8	2.990	0.307	0.334	0.345	0.044
	7	Holland, Belgium, France (east)	10	2.792	0.300	0.318	0.327	0.036
	8	Spain, Portugal, France (south-west)	11	2.720	0.277	0.297	0.322	0.087
	9	Greece, Yugoslavia, Bulgaria	9	2.750	0.339	0.356	0.386	0.085
	10	Slovakia, Hungary, Croatia, Romania	10	2.558	0.333	0.344	0.367	0.073
	11	Norway, Sweden, Denmark, Poland	13	2.603	0.266	0.280	0.298	0.068
	12	Turkey, Lebanon, Cyprus, Iraq, Iran	11	2.962	0.362	0.378	0.427	0.118
L. rigidum	13	France, Corsica, Sardinia, Italy	7	3.155	0.372	0.408	0.467	0.136
	14	Spain, Portugal	9	3.009	0.370	0.421	0.495	0.151
	15	Canary Islands, Morocco, Algeria	9	2.907	0.356	0.395	0.482	0.177
	16	Tunisia, Libya, Egypt	8	3.385	0.394	0.438	0.495	0.123
	17	Greece, Turkey, Cyprus, Israel, Jordan	9	3.389	0.365	0.401	0.461	0.133
	18	Iraq, Iran, Afghanistan, Kirghizia	8	2.938	0.361	0.393	0.454	0.139

N° g, geographical group number; *n*, number of populations; *A*, number of alleles; H_o, observed heterozygosity; H_S, gene diversity within population; H_T, total gene diversity; G'_{ST}, absolute coefficient of gene differentiation.

The logistic regression between allelic frequencies and geographical data for *L. perenne* populations reveals several significant relationships, whereas no consistent relationships are observed for *L. rigidum* populations. Considering one allele per locus (generally the most common one), the different combinations between allele frequency and latitude and/or longitude were tested in a Generalized Linear Model (McCullagh & Nelder, 1983). The effects of latitude and longitude always appear significant and relatively independent. Table 4 gives the deviance explained by logistic regression for *L. perenne* when using only one effect in the model: 10 alleles show significant relationships with latitude and only five with longitude, when these effects are used independently. Logistic regressions between these 10 allelic frequencies and the latitudes of the collection sites are presented in Fig. 2; all these alleles show a trend towards a latitudinal clinal variation. To illustrate this result for *L. perenne* populations, Fig. 3 shows the geographical frequency distribution of the *Dia1–30* allele. There is a clear trend from the northern populations of the United Kingdom and Ireland, characterized by the lowest frequencies, to the southern populations of Italy, Corsica and Middle Eastern countries with the highest allele frequencies.

The dendrogram based on the matrix of modified Rogers's distance among the 18 groups of *L. perenne* and *L. rigidum* populations (Fig. 4) clearly reveals the genetic divergence of the two species. *Lolium perenne* populations from the Middle East (group 12) appear to be the closest to the overall *L. rigidum* population. Among *L. perenne* populations, this

Table 4 Table of deviance explained by logistic regression in *Lolium perenne*

Allele	Longitude	Latitude
Pgi2–45	82.60**	147.95**
Acp1–40	NS	NS
Acp2–20	47.38**	274.83**
Got2–20	9.77*	89.99**
Got3–30	NS	NS
Sod1–20	NS	86.70**
Prx1–30	NS	89.48**
Idh1–30	25.09**	220.94**
Dia1–30	46.67**	229.53**
Skd1–20	NS	143.56**
Pgm1–30	NS	94.82**
Mdh1–20	NS	43.84**

*$P<0.01$, **$P<0.001$.

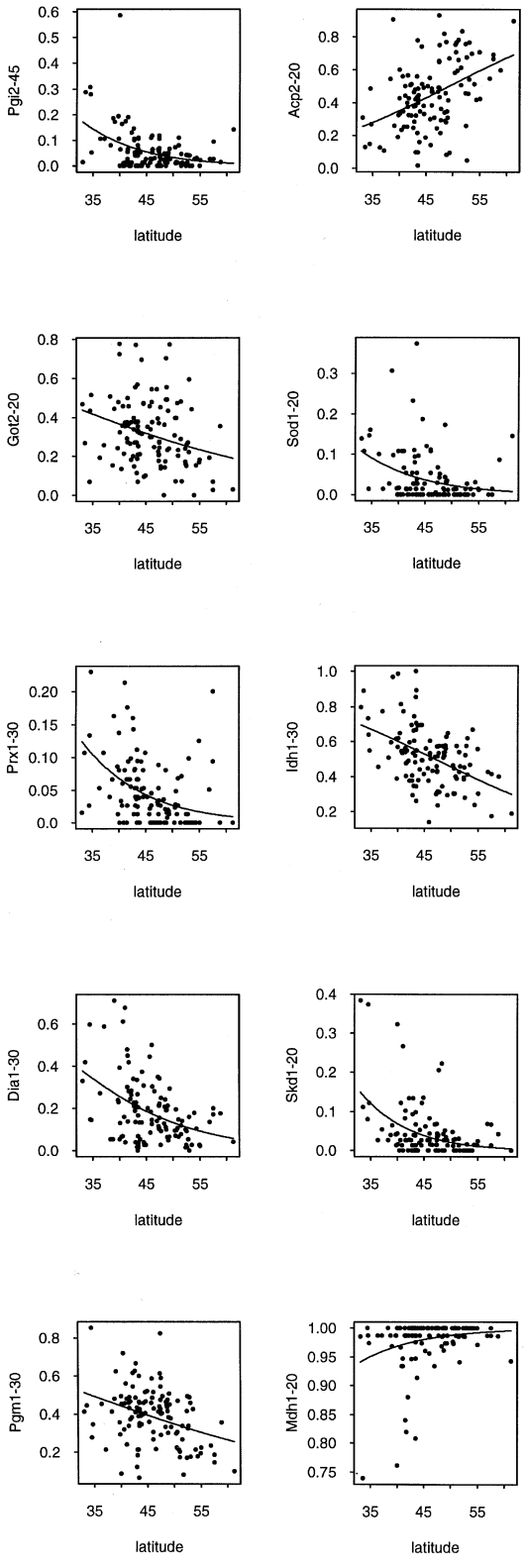

Fig. 2 Logistic regression between latitude of collection sites and allelic frequencies in *Lolium perenne* populations.

Middle East group is separate, and there is a smaller distance between southern French, Spanish and Italian populations (groups 4, 8, 2). The other groups of countries are divided in two distinct clusters: a first one with population groups 11, 1, 7, 5 and 6 which are located to the north-west of the Alps, and a second cluster with groups 3, 9 and 10 distributed to the south-east of the Alps. In *L. rigidum* accessions, geographical differences are not so clear: populations from Spain, Portugal, Morocco and Algeria (groups 14 and 15) seem to be nearest to *L. perenne* populations. The other groups are gathered in the same cluster. However, slight differentiation appears between populations of groups 17 and 18, from the Eastern part of the Mediterranean basin, and those of groups 13 and 16 from the Western part.

Discussion

Organization of genetic diversity at the species level

A primary goal of this research was to characterize the genetic diversity in these two species of *Lolium*. This study examined a substantially larger number of populations and loci than earlier studies and allows a thorough overview of isozyme diversity across the species' geographical ranges.

The mean values of genetic diversity indices (A, H_o, H_e) observed in *L. perenne* are consistent with results obtained previously (Hayward, 1985; Charmet *et al*., 1993). For *L. rigidum*, our values are higher than those reported by Charmet & Balfourier (1994), probably because of a larger sample of populations in the present study.

The total gene diversity at the species level is lower in *L. perenne* ($H_T = 0.362$) than in *L. rigidum* ($H_T = 0.488$). This is probably because effective population size is lower in *L. perenne* than in *L. rigidum*, which could be explained either by a bottleneck effect at the origin of *L. perenne* or by reproduction and life cycle characteristics. In *L. perenne*, only a small number of plants may participate in reproduction as it is a perennial species used for intensive grazing in natural and permanent pastures.

For the two species, most of the total gene diversity is accounted for by the within-populations variation component. This result agrees with those of Hamrick & Godt (1990) who reported a mean G_{ST} of 0.098 for 134 outcrossing wind-pollinated species. Our perennial ryegrass populations show a differentiation level of the same magnitude ($G_{ST} = 0.110$). However, the mean coefficient of gene differentia-

tion is clearly (but not significantly) higher in *L. rigidum* ($G_{ST} = 0.170$). This is probably because *L. rigidum* is an annual species with a wider ecological range than *L. perenne* (Hamrick *et al*., 1992).

The deficit of heterozygotes ($F_{IS} > 0$), greater in *L. rigidum*, may have two causes: partial selfing and a restricted neighbourhood size (or the Wahlund effect). Because the two species are known to be self-incompatible, the Wahlund effect is more likely to explain the deficit of heterozygotes. A smaller neighbourhood size, cause by the life cycle characteristics of *L. rigidum*, may explain the higher F_{IS} value for this species.

Allele frequencies

Sixty-one alleles were found among the 50 wild accessions of *L. rigidum*, whereas only 57 alleles were observed in the 120 populations of *L. perenne*. This relative loss of genetic variation in *L. perenne*, and the lower gene differentiation of this species may be the consequence of *L. perenne* being a recent derivative of *L. rigidum* populations after establishment of this annual species. Several arguments favour this theory.

First, *L. rigidum* possesses all the alleles of *L. perenne*, but the converse is not true; rare alleles may have been lost in *L. perenne* by drift during founding events. Moreover, the mean number of alleles observed in 50 populations of *L. rigidum* (3.13) is greater than that of *L. perenne* (2.72) from 120 populations. Furthermore, rare alleles of *L. rigidum* are also rare in *L. perenne*. The rare alleles that are more frequent in *L. perenne* than in *L. rigidum* (i.e. *Pgi2–10*, *Pgi2–45*, *Got3–50*, *Idh1–50* and *Acp1–60*) could be explained by a bottleneck effect because populations of *L. perenne* may have undergone severe restriction in their population size

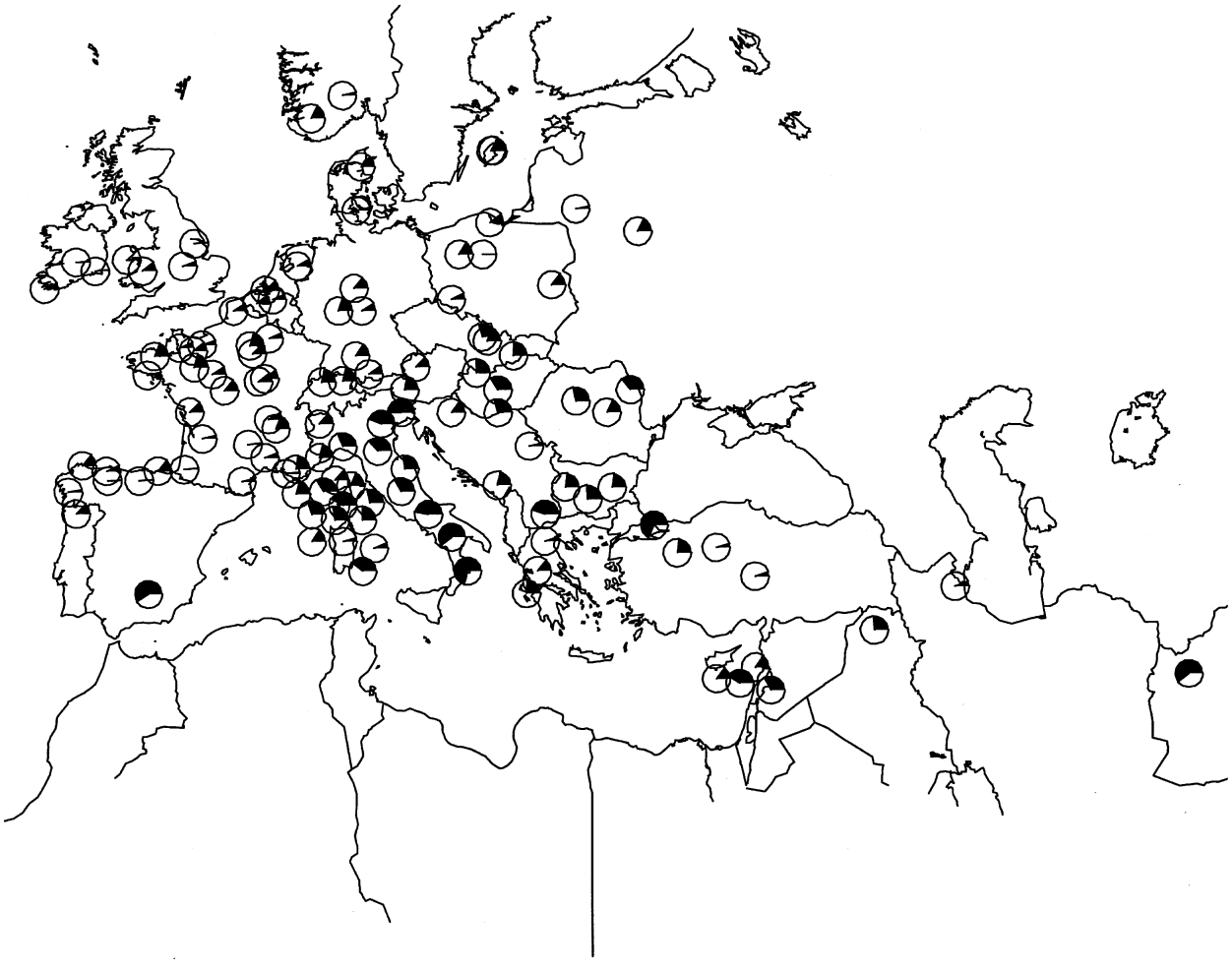

Fig. 3 Geographical pattern of variation for *Dia1–30* allele frequency in 120 populations of *Lolium perenne* (filled sectors are proportional to allele frequency).

during their history; the increase in the frequency of rare alleles after a founder effect is a well-known phenomenon (Maruyama & Fuerst, 1985).

Secondly, theory predicts that natural selection for adaptive traits, from a limited portion of wild gene pools, will produce a genetic bottleneck and decrease of diversity (Brown, 1978; Doebley, 1989). This apparently is the case between *L. perenne* and *L. rigidum*: we observe a lower diversity in *L. perenne* populations which may have been subject to natural selection pressure, caused by animal grazing, during different colonization stages.

Finally, phylogenetic analyses in the *Festuca–Lolium* complex, first using isozymes (Charmet & Balfourier, 1994) and then molecular markers and ITS rDNA (Charmet *et al*., 1997), indicate that the genus *Lolium* diverged from the broad-leaved fescues group around 2 Myr ago, and its differentiation into species is very recent (about 1 Myr ago). According to these studies, *L. rigidum* is the most likely common ancestor of *L. multiflorum* and *L. perenne* species: *L. multiflorum* accessions are very close to each other within a group including some *L. rigidum* accessions, whereas *L. perenne* accessions are in another cluster. Divergence between *L. rigidum* and *L. perenne* may have occurred very recently. As for *L. multiflorum*, Borrill (1976) suggests that this species, also named *L. italicum*, may have been derived by mutation from an *L. rigidum* Italian population, which would have been selected by farmers during the thirteenth century.

Geographical distribution within each species

The results obtained in this study for *L. rigidum* show that spatial distribution of this species is not simple. There is a between-population structure of this species, indicated by high values of G'_{ST}, but this

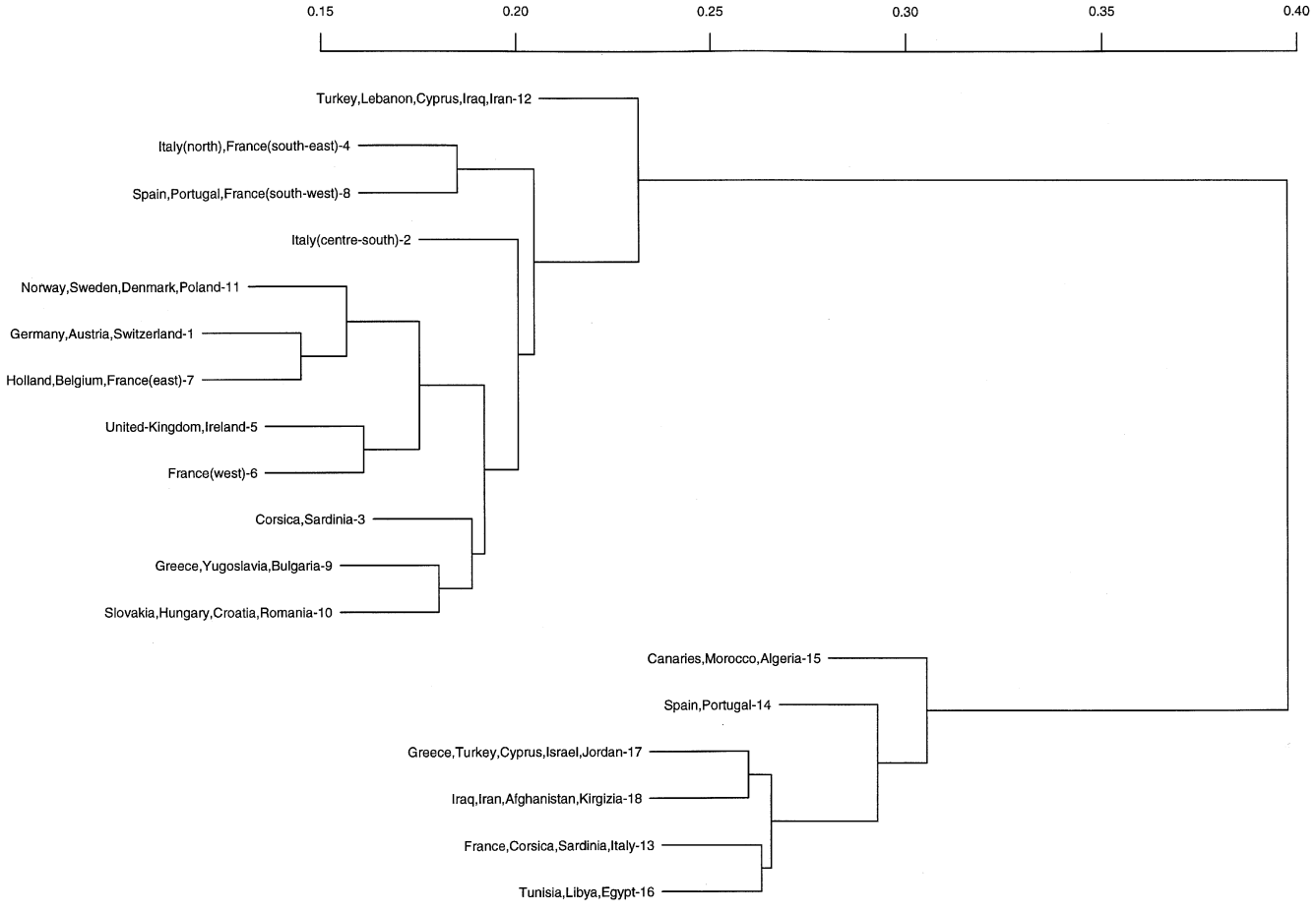

Fig. 4 Dendrogram of the 18 geographical groups of *Lolium perenne* (upper part), and *L. rigidum* accessions (lower part) based on the UPGMA method using the modified Rogers's distance matrix (numbers correspond to geographical group number).

structure is not correlated with distance. Different hypotheses can be proposed to explain this result: the distribution area studied could be too small, although it is very similar to that of *L. perenne* and we observed a spatial structure for the perennial species. Alternatively, gene flow may be reduced by fragmentation of the distribution area all around the Mediterranean. Furthermore, our sample of 50 populations was probably also too small, and local genetic drift and stochastic effects may be more important than eventual gene flow produced by migration processes on a large scale.

In contrast, the results obtained for *L. perenne* exhibit clear geographical patterns of variation for both diversity indices and individual allele frequencies, as indicated by strong and significant relationships with geographical data. A Mantel test indicates a significant correlation (at the 1% level) between G_{ST} and the geographical distance matrix of the populations. A latitudinal gradient appears to be more pronounced across loci and diversity indices, although significant relationships also exist with longitude. More specifically, populations from the most northern part of the distribution area (groups 11, 1, 5) exhibit the lowest levels of heterozygosity and allelic richness, whereas the highest values are observed in the southern part (groups 2, 3, 12).

The same pattern of clinal variation is observed for many allele frequencies from different loci. Clines in gene frequencies can be generated by selection for different alleles along a geographical or ecological gradient: this is the selection hypothesis (Endler, 1977). However, they can also be formed as a consequence of long-range dispersal and migration along a geographical axis (Wijsmann & Cavalli-Sforza, 1984). In fact, both processes can act separately or in combination (Felsenstein, 1976). According to Barbujani (1988), the selection hypothesis permits identification of different axes for different loci, unless they are tightly linked. In contrast, if the same axis is observed for all loci, the migration hypothesis must be considered. In the case of *L. perenne*, 10 loci out of 12 show significant relationships between allele frequencies and latitude. Therefore, the migration hypothesis (i.e. gene flow) is likely to be the main explanation of the south/north clinal variation in allele frequencies.

As before, a plausible explanation for such geographical differentiation in *L. perenne* is that it was derived from a small bottleneck of *L. rigidum* populations which occurred in the Middle East area a few thousand years ago. From that time onwards, two scenarios could explain the present distribution area of *L. perenne* in northern Europe.

The first scenario would fit historical processes such as the emergence of primitive agriculture of cereals crops in the Middle East, about 10,000 years ago, and its expansion towards Europe. According to historical studies of human populations, two different pathways of colonization from the fertile crescent are known: a south-western one (Mediterranean movement) towards Mediterranean countries (Greece, Italy and the Iberian Peninsula, then southern France, Corsica, etc.), and a north-western one (Danubian movement) towards Bulgaria, Romania, Slovakia, Poland, Germany, the north of France, then the United Kingdom, Ireland and Scandinavia. Most authors (Malik, 1967; Borrill, 1976) agree that the Mediterranean basin is the centre of origin of the genus *Lolium*, and that 'true self-pollinating' species of this genus (e.g. *L. temulentum*, *L. persicum and L. remotum*), ancestors of *L. perenne* and *L. rigidum*, are known only as weeds of cultivated crops. This possible scenario of colonization of Europe by *L. perenne* populations, carried as weeds of cereals crops by the first farmers, fits quite well with the dendrogram obtained in the present study (Fig. 4) and with a previous study using molecular markers (Charmet *et al*., 1997), and attesting to the monophyletic origin of *L. perenne*.

The second scenario would be colonization of Europe, from the Middle East area, followed by the last glaciation events (Würm period), just before the expansion of primitive agriculture. The present distribution area of *L. perenne* would be the result of recolonization from several southern postglacial populations. In this case, the possible recolonization pathway could have followed three different directions: (i) a northern movement from all refugia (Spain, south of France, Italy and the Middle East), then (ii) an eastern migration from the Iberian peninsula and France to those countries located north-west of the Alps (groups 11, 1, 7, 5 and 6), and (iii) another migration from Italy to countries of groups 3, 9 and 10 which are located to the south-east of the Alps. In this scenario, the Alps would have acted as a natural barrier to gene flow. This may explain the existing relationships between some allele frequencies and longitude. Collections of *L. perenne* populations in the Sierra Nevada, in Southern Spain, and in the Lebanon mountains confirm the existence of postglacial relic flora which now remain at altitude in the extreme south of the present *L. perenne* distribution area.

This last scenario of recolonization has already been reported for different European forest trees species such as beech (Leonardi & Menozzi, 1996) and oak (Michaud *et al*., 1995; Zanetto & Kremer, 1995).

For oak, these authors used analysis of fossil pollen maps to confirm their hypothesis of primary refugia areas in Spain, Italy and the Middle East. Unfortunately, such palynological data do not exist for *Lolium* species, and it is thus difficult to confirm this scenario.

Conclusions

Considering previous studies and the present results, we can assume that *L. perenne* is probably derived from a limited number of *L. rigidum* populations which occurred in the Middle East. Thus, the collection and preservation of natural populations from this presumed 'centre of origin' should be recommended. In the case of *in situ* conservation, suitable sites should be found both in the Middle East and in Western countries. Moreover, other sites, intermittently located along the 'migration routes' should also be preserved.

It is still quite difficult to choose between different scenarios to explain the present distribution area of *L. perenne*. Many arguments, based on molecular studies, favour the first hypothesis (extension with primitive agriculture), but, for a very short time-scale, the use of a molecular clock on our isozyme data is not sufficiently accurate to decide conclusively between the different scenarios. Geostatistical analyses are in progress on isozyme data, in order to confirm and describe in more detail the geographical patterns observed; the use of directional variograms should make it possible to detect the main direction of geographical variation and to choose between the scenario of extension with agriculture (mainly a longitudinal direction) and the postglacial recolonization hypothesis (mainly a latitudinal direction).

Finally, as the chloroplast genome is likely to reflect past colonization dynamics, a chloroplast DNA restriction polymorphism study is being undertaken in the same populations, in order to observe eventual divergence between chloroplast lineages. These complementary studies will be presented in further papers.

Acknowledgements

The authors wish to thank B. Boller (Switzerland), M. Bounejmate (Morocco), V. L. Bradley (U.S.A.), S. Bullita (Sardinia), V. Connolly (Ireland), A. Dovrat (Israel), M. Falcinelli (Italy), A. Hamadache (Algeria), M. Hayward (U.K.), L. Horvath (Hungary), K. Kolshus (Sweden), W. Majtkowski (Poland), C. Mousset (France), A. Oliveira (Spain), L. Ostrem (Norway), P. Perrino (Italy), D. Reheul (Belgium), D. Shamov (Bulgaria), I. D. Thomas (U.K.), T. Vaitsis (Greece), T. Van Hintum (Netherlands), E. Willner (Germany) and the French Association des Créateurs de Variétés Fourragères (ACVF) for providing seeds. The authors also gratefully acknowledge N. Pask and M. H. Tixier for their technical assistance in carrying out the electrophoretic analysis.

References

BALFOURIER, F. AND CHARMET, G. 1994. Geographical patterns of isozyme variation in Mediterranean populations of perennial ryegrass. *Heredity*, **72**, 55–63.

BARBUJANI, G. 1988. Detecting and comparing the direction of gene frequency gradients. *J. Genet.*, **67**, 129–140.

BECKER, R. A., CHAMBERS, J. M. AND WILKS, A. R. 1988. *The New S Language. A Programming Environment for Data Analysis and Graphics*. Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA.

BORRILL, M. 1976. Temperate grasses. In: Simmonds, N. (ed.) *Evolution of Crop Plants*, pp. 137–142. Longman, London.

BROWN, A. H. D. 1978. Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.*, **52**, 145–157.

CHARMET, G. AND BALFOURIER, F. 1994. Isozyme variation and species relationships in the genus *Lolium* L. (ryegrasses, Graminaceae). *Theor. Appl. Genet.*, **87**, 641–649.

CHARMET, G., BALFOURIER, F. AND RAVEL, C. 1993. Isozyme polymorphism and geographic differentiation in a collection of French perennial ryegrass populations. *Genet. Res. Crop Evol.*, **40**, 77–89.

CHARMET, G., RAVEL, C. AND BALFOURIER, F. 1997. Phylogenetic analysis in *Festuca-Lolium* complex using molecular markers and ITS rDNA. *Theor. Appl. Genet.*, **94**, 1038–1046.

COLLETT, D. 1991. *Modelling Binary Data*. Chapman and Hall, London.

DOEBLEY, J. F. 1989. Isozymic evidence and the evolution of crop plants. In: Soltis, D. E. and Soltis, P. S. (eds) *Isozymes in Plant Biology*, pp. 165–191. Dioscorides Press, Portland, OR.

ENDLER, J. A. 1977. *Geographic Variation, Speciation, and Clines*. Princeton University Press, Princeton, NJ.

FELSENSTEIN, J. 1976. The theoretical population genetics of variable selection and migration. *Ann. Rev. Genet.*, **10**, 253–280.

FERNANDO, W. M. U., HAYWARD, M. D. AND KEARSEY, M. J. 1997. Isozyme and quantitative traits polymorphisms in European provenances of perennial ryegrass (*Lolium perenne* L.). *Euphytica*, **93**, 263–269.

HAMRICK, J. L. AND GODT, M. J. W. 1990. Allozyme diversity in plant species. In: Brown, A. H. D., Clegg, M. T.,

Kahler, A. L. and Weir, B. S. (eds) *Plant Population Genetics, Breeding and Genetic Resources*, pp. 43–63. Sinauer, Sunderland, MA.

HAMRICK, J. L., GODT, M. J. W. AND SHERMAN-BROYLES, S. L. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests*, **6**, 95–124.

HAYWARD, M. D. 1985. Adaptation, differentiation and reproductive systems in *Lolium perenne.* In: Jacquard, P., Hein, G. and Antonovics, J. (eds) *Genetic Differentiation and Dispersal in Plants*. *NATO-ASI Series*, **G5**, 83–93. Springer, Berlin.

HAYWARD, M. D., DEGENNARS, G. H., BALFOURIER, F. AND EICKMEYER, F. 1995. Isozyme procedures for the characterisation of germplasm exemplified by the collection of *Lolium perenne* L. *Genet. Res. Crop Evol.*, **42**, 327–337.

LEONARDI, S. AND MENOZZI, P. 1996. Spatial structure of genetic variability in natural stands of *Fagus sylvatica* L. (beech) in Italy. *Heredity*, **77**, 359–368.

MALIK, C. P. 1967. Cytogenetic studies of the F_2 hybrid of *L. multiflorum* × *L. rigidum* and the species relationships in the genus *Lolium*. *Der Züchter*, **37**, 261–274.

MARUYAMA, T. AND FUERST, P. A. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics*, **111**, 675–689.

McCULLAGH, P. AND NELDER, J. A. 1983. *Generalized Linear Models.* Chapman and Hall, London.

MICHAUD, H., TOUMI, L., LUMARET, R., LI, T. X., ROMANE, F. AND DI GIUSTO, F. 1995. Effect of geographical discontinuity on genetic variation in *Quercus ilex* L. (holm oak). Evidence from enzyme polymorphism. *Heredity*, **74**, 590–606.

NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 3321–3323.

NEI, M. 1977. *F*-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.*, **41**, 225–233.

SNEATH, P. H. A. AND SOKAL, R. R. 1973. *Numerical Taxonomy.* W.H. Freeman, San Francisco, CA.

TERRELL, E. E. 1968. A taxonomic revision of the genus *Lolium*. *Tech. Bull. U.S. Dept. Agric.*, **1392**, 1–65.

WIJSMANN, E. M. AND CAVALLI-SFORZA, L. L. 1984. Migration and genetic population structure with special reference to humans. *Ann. Rev. Ecol. Syst.*, **15**, 279–301.

WRIGHT, S. 1978. *Evolution and the Genetics of Populations*, vol. 4, *Variability Within and Among Natural Populations.* University of Chicago Press, Chicago.

ZANETTO, A. AND KREMER, A. 1995. Geographical structure of gene diversity in *Quercus petraea* (Matt.) Liebl. l. Monolocus patterns of variation. *Heredity*, **75**, 506–517.