Agronomie 18 (1998) 659–667

Original article

Comparison between maize inbred lines: genetic distances in the expert's eye

Christine Dillmann*[a], Daniel Guérin[b]

[a] Station de génétique végétale, INA-PG, ferme du Moulon, 91190 Gif-sur-Yvette, France
[b] Inra, Geves, La Minière, 78285 Guyancourt cedex, France

(Received 15 June 1998; accepted 8 October 1998)

Abstract – To be registered on the national list for protection purposes, a new variety has to be distinct from all existing varieties, uniform and stable. In France, distinction relies on morphological traits, but the final decision is made by an expert committee. In the future, molecular markers could also be used, either for the management of reference collections or for the assessment of essential derivation. Our aims were to investigate the relationships between morphological or molecular distances and the opinion of experts in maize. Experts were asked to give a visual estimation of the relatedness between inbred lines. The results validate the French approach of distinction where a morphological distance index is first computed, and then used as a sieve to carefully observe only the closest varieties. Experts are also shown to be able to reconstruct the relatedness between two inbred lines from their phenotype. In that sense, they perform much better than classical morphological distance indices. (© Inra/Elsevier, Paris.)

maize / genetic distance / molecular markers / plant breeding protection

Résumé – Comparaisons entre lignées de mais : les distances génétiques et l'œil de l'expert. Pour être inscrite sur la liste nationale comme protégée, une nouvelle variété doit être distincte de toutes les variétés existantes, homogène et stable. En France, la distinction est fondée sur des caractères morphologiques, mais la décision finale est prise par un comité d'experts. Dans le futur, les marqueurs moléculaires pourraient aussi être utilisés pour gérer les collections de référence, ou pour instruire la dérivation essentielle. Dans le but d'étudier la relation entre ces différents modes d'appréciation de la distance entre variétés, nous avons demandé à des experts du maïs d'estimer visuellement l'apparentement entre des couples de lignées connues. Les résultats permettent de valider l'approche française de la distinction, qui consiste à utiliser comme tri un indice de distance morphologique pour n'observer plus attentivement que des lignées proches. On montre aussi que, contrairement aux distances morphologiques classiques, les experts sont capables de déduire du phénotype les relations d'apparentement entre deux lignées proches. (© Inra/Elsevier, Paris.)

maïs / distance génétique /marqueurs moléculaires /protection des obtentions végétales

Communicated by André Gallais (Gif-sur-Yvette, France)

---

* Correspondence and reprints
Tel: (33) (0)1 69 33 23 48; fax: (33) (0)1 69 33 23 40; e-mail: dillmann@moulon.inra.fr

1. INTRODUCTION

In Europe, plant breeder's rights depends on UPOV (*Union pour la protection des obtentions végétales*) guidelines. To be registered on the national list, a new variety has to be distinct from all previously registered varieties, uniform and stable (distinction, uniformity, stability, or DUS criteria). In France, distinctness testing is performed by GEVES (*groupe d'étude et de contrôle des variétés et des semences*) and relies upon a set of morphological traits, as well as upon the opinion of an expert committee in a two-stage procedure. The first year, all candidates and a set of reference varieties are grown and compared in the field on the basis of the standard UPOV criteria. The second year, each candidate is compared with its closest reference varieties. They are examined by experts, who make the final decision. As the number of new varieties in main culture crops is increasing, the whole procedure becomes more and more complex. In particular, it implies the management of huge reference collections, which could be assisted by the use of molecular markers. Molecular markers also appear as possible tools to assess essential derivation, which was defined in the 1991 UPOV convention. In that case, the problem is to accurately estimate the relatedness between two varieties. Hence, while molecular markers potentially appear as a choice method in the context of plant breeding protection [11], it seems necessary to study their relationships with the traits which are currently used to discriminate between varieties.

As the variability revealed by a given molecular marker concerns in general only a few base pairs, it is admitted that molecular markers could be used for distinction or essential derivation through the measurement of a synthetic distance index, pooling the information from a set of markers. Several studies have focused on the statistical properties of such molecular genetic distance, taking into account the sampling of markers throughout the genome [1, 7, 12, 13]. Regarding essential derivation, some methods have been proposed to improve the estimation of the relatedness between varieties and to remove the confusion between identity by descent and alikeness in state [3, 9]. The relationship between molecular and morphological distances have also been studied and appeared to be triangular, the molecular distance being the limiting factor of the morphological distance [2, 4, 6].

Knowing the specificity of DUS testing, which leaves the final decision to the experts, it seemed important to include the opinion of experts in such studies. Maize was chosen here as a model crop for two reasons. First, a large number of well characterised RFLP markers were available [6]. Secondly, the entities which are protected in maize are either the commercialised hybrids or their parental inbred lines, and they have an homogeneous genetic structure. The aims of the present study were: i) to validate the DUS protocol used in France; and ii) to compare the molecular distance between maize inbred lines to the opinion of maize experts.

2. MATERIALS AND METHODS

2.1. Distance indices

Three different genetic distances were computed, depending on the kind of information taken into account: maize experts, morphological and molecular distances.

Maize experts

They were chosen among public and private maize plant breeders on a volunteer basis. Some of them were members of the maize expert committee in France. They were asked to give a visual appreciation of the distance between pairs of inbred lines, based on the phenotype of the inbred lines at flowering time and at maturity. Two questions were asked:

(1) Taking into account the current DUS regulation, do you find the two inbred lines distinct? Please answer by yes or no.

(2) Give an estimation of the distance between the two inbred lines, taking into account your knowledge of maize and using the following scale: very close (1); close (3); intermediate (5); different (7); very different (9).

Actually, the second question aimed at visually estimating the relatedness between two inbred lines.

There were 31 experts at flowering time and 25 at maturity, with 18 experts having been present both times. It must be noted that the experts have only seen the plants twice and were not able to take into account all the morphological and physiological traits expressed during plant development.

Morphological distance

In maize, a hybrid variety can be protected as soon as the parental inbred lines satisfy the DUS criteria. The LCLM (*logiciel de comparaison de lignées de maïs*) distance index [8] is used in France in the distinction protocol for comparison of inbred lines. It is a phenotypic distance index, computed from 31 morphological traits measured in two locations with two repetitions per location, and 17 enzymatic loci. The traits and the corresponding evaluation protocols are described in the UPOV guidelines. When comparing two inbred lines, only traits exhibiting significant differences are taken into account. Each trait is given a weight of 2, 3 or 6, depending on the number of genes involved and on the heritability of the trait. The weight of 6 is given to polygenic traits which are mildly influenced by the environment, such as earliness, type of kernel (flint/dent) or ear height. More complex traits like ear length and ear diameter are given a weight of 3. A weight of 2 is given to the other traits, including most of the pigmentation traits. As for the electrophoretic patterns, the number of differences between a given pair of inbred lines is given a weight of 2, so that the weight of all electrophoretic patterns pooled together does not exceed the weight of one morphological trait. The resulting distance index may range from 0 to 108. A minimum distance of 6 was set empirically. Hence, depending on the traits, one to three significant differences are necessary for two given inbred lines to be declared distinct. The LCLM distance indices used in the present study have been elaborated at GEVES between 1989 and 1993.

Molecular distance

Results from 80 RFLP markers were used in this study. The enzyme/probe combinations were chosen (i) to be monolocus, (ii) to be distributed all over the genome, and (iii) to give an interpretable band pattern. Details on these RFLP markers are given in Dillmann et al. [7]. For each locus, each different band pattern was scored as a different profile or allelic form. The molecular distance was simply computed as the percentage of markers which differ between two inbred lines, regardless of their position on the maize genetic map.

2.2. Choice of inbred lines

In order to cover most of the possible critical distinction situations and to characterise essential derivation, 160 pairs of maize inbred lines were chosen.

In previous studies, a set of 145 inbred lines, representative of the parents of the varieties commercialised in France, have been analysed [1, 2, 6]. The LCLM distance, as well as the molecular distance have been computed for each of the 10 440 pairs of inbred lines. A subset of 160 pairs among the closest lines was chosen for the present study according to the following criteria: the inbred lines should originate from the same genetic background and the molecular distance should be lower than 30 %. Among all the pairs satisfying these criteria, a stratified sampling has been performed when possible in order to best represent all molecular distance classes.

As most of the material consisted of already protected and, therefore, distinct inbred lines, only three selected pairs of inbred lines have a molecular distance lower than 5 %. Pairs of very close inbred lines were under-represented in our sampling. The sample consisted of 14 pairs of European inbred lines, 62 pairs of SSS lines, 74 pairs of Iodent lines, 4 pairs of Wisconsin lines, 3 pairs of Canadian flint lines, and 3 pairs of lines of unknown origin. The distribution of the LCLM distance and of the molecular distance among the 160 pairs are given in *figure 1*. These distributions simply reflect the characteristics of the chosen set of pairs of lines.

2.3. Experimental design

All chosen pairs have been grown in 1995 at Le-Magneraud (westcentral France). They have been noted by maize experts at flowering time (middle of July) and at harvesting (end of August). There were two completely randomised repetitions for each pair. The experimental unit was two rows with 10 plants of the first inbred line in the first row, and 10 plants of the second inbred line in the second row. The distance between the rows was 80 cm.

3. RESULTS

Twenty-one pairs of inbred lines were discarded from the study, either because one of the inbred

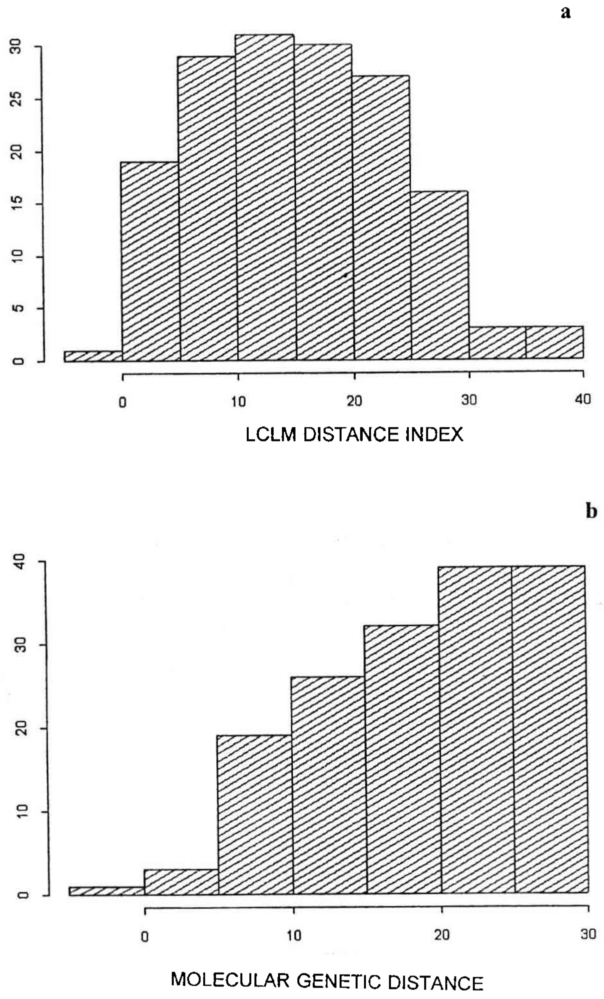

Figure 1. Distribution of the LCLM distance (a) and the molecular genetic distance (b) among the 160 pairs of maize inbred lines of the study.

lines did not germinate (one inbred line), or because they did not flower before the experts came (20 pairs).

3.1. Distinction analysis

First, the experts were asked if they found the inbred lines to be distinct or not. The repetition effect is not significant. Each of the 139 pairs was found to be distinct by at least one expert. A distance index has been computed for each date of notation as the proportion of experts who found the inbred lines to be distinct. The only pair with a distance index lower than 50 % is the one which included an inbred line which had failed the official distinction test. Among the 139 pairs, the distribution of the distance index is extremely skewed. Most of the experts have declared the inbred lines distinct. The median value is 90 % at flowering time, and 98 % at maturity. Because of the skewness of the distribution, there is no correlation between this distance index and other distances. Distances are generally higher at maturity than at flowering time.

3.2. Similarity study

The second question aimed at estimating the similarity between inbred lines. Some significant examples of the distribution of the expert's notation with regard to the 139 pairs of inbred lines are represented in *figure 2*. Generally, the experts have used all the notations possible between 1 and 9, and the distributions tend to be symmetrical. The mode of the distributions is between 4 and 5. The symmetry of these distributions, compared to the distributions presented in *figure 1* suggest that the experts have adjusted their notation to the variability among the chosen set of inbred lines.

An analysis of variance was performed for each date of observation to test the effect on notation of the inbred line pairs and of the expert. Both line and expert effects are significant. However, a comparison of means reveals that the expert effect came from a small number of experts who gave highly extreme notations. A principal component analysis has also been performed with the experts as variables. The first principal component explains 52 % of the variation at flowering time, and 41 % of the variation at maturity, and corresponds to the average expert notation for each couple of inbred lines. The percentages of explanation drop to 4.2 % and 6.2 %, respectively for the second principal compo-

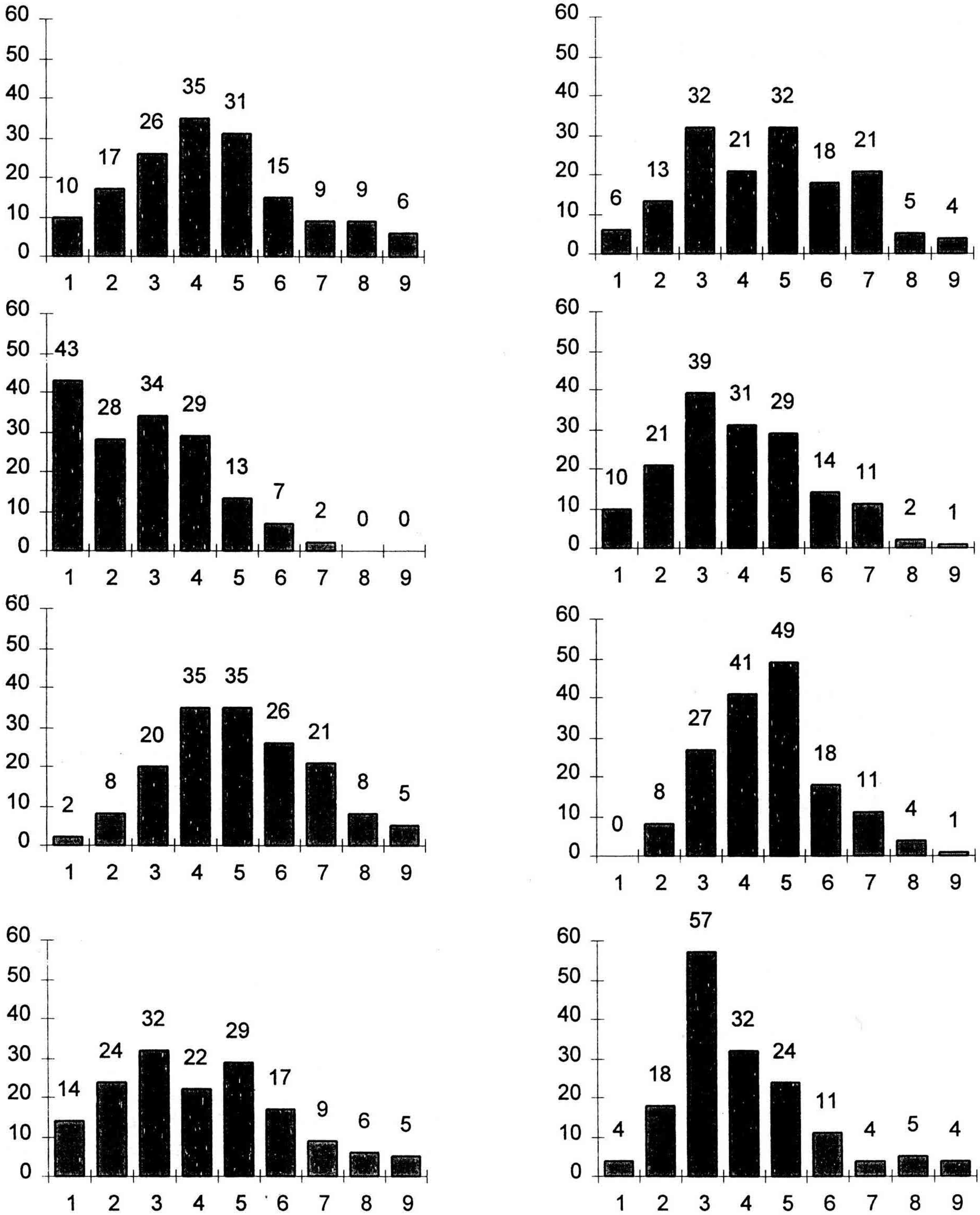

Figure 2. Distribution of eight different experts notation with regard to the pairs of inbred lines at flowering time. The notation estimates the relatedness between the inbred lines. It ranges from 1 (very related) to 9 (very different) and it is averaged over two repetitions.

nent. We have therefore chosen to compute the experts' distance index, for each date of observation and each pair of line, as the average notation across experts and across repetitions. At flowering time, the expert's distance ranges between 1.37 and 9, and the average distance across all pairs of lines is 4.57. At maturity, it ranges from 1.58 to 9, with an average distance of 4.63 across all pairs of lines. As previously, the average distance seems to be higher at maturity than at flowering time.

In order to take into account most of the characteristics expressed during plant development, a synthetic distance index has been computed as the maximum distance between flowering time and maturity.

MAXI = Max (expert's distance at flowering time; expert's distance at maturity)

The MAXI distance index ranges between 2.3 and 9, with an average value of $\mu_{\text{maxi}} = 5.1$ and a standard deviation of $\sigma_{\text{maxi}} = 1.3$. Its distribution is symmetrical around the average.

3.3. Relationship between MAXI and LCLM distance indices

As shown in *figure 3*, the relationship between MAXI distance and LCLM distance is not linear. LCLM distance is used by GEVES as a sieve to assess distinctness during the first year of submission, when it is possible, and to choose inbred lines in the reference collection which are to be implemented the second year for further comparisons with candidate lines. An empirical threshold has been set at 6 for distinction, which has been used for years as a safe cut-off point. So, each pair with a LCLM distance higher than 6 is considered as distinct. Similarly, it is possible in *figure 3* to define an arbitrary threshold for MAXI as the minimum MAXI distance for pairs of inbred lines already distinguished by LCLM. In our case, it is equal to $\mu_{\text{MAXI}} - 1.5\sigma_{MAXI}$ corresponding to a MAXI distance of 3.15.

The two thresholds are delimiting 4 sections in *figure 3*. The upper-left section corresponds to pairs of inbred lines that are distinguished by experts and not by LCLM. Apparently, experts are able to integrate some traits which distinguishe maize inbred lines which are not taken into account in the UPOV guidelines. The lower-right section in *figure 3* is empty. It corresponds to inbred lines that would have been declared distinct according to LCLM and found to be related by experts. This result validates the use of LCLM as a tool for a fast sorting of truly distinct inbred lines. However, three pairs of inbred lines have a MAXI value very close to the threshold (*figure 3*). For one of these pairs, the LCLM distance of 27 was obtained with traits such as silk pigmentation and height of ear insertion. Those traits had probably not been taken into account by the experts, either because they were not noticeable when the experts came, or because they were not expressed in 1995 at the experimental station. Notice that in that case, the LCLM distance has been computed from three years of experimentation in two locations.

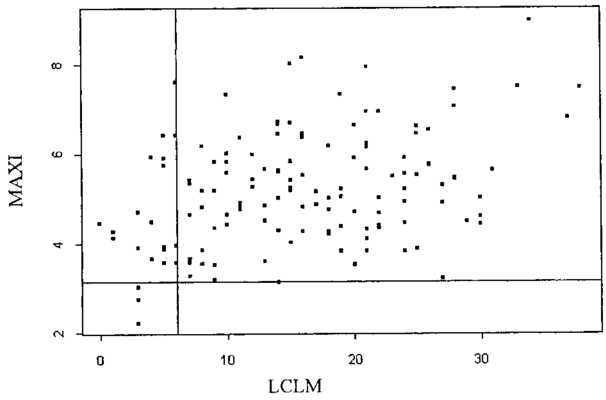

Figure 3. Relationship between the LCLM distance index and the expert's distance (MAXI). The vertical line represents the LCLM = 6 threshold. The horizontal line represents the arbitrary threshold of MAXI = 3.15.

3.4. Relationship between the expert's distance and the molecular distance

In our experiment, the experts were able to distinguish between related inbred lines more accurately than LCLM. *Figure 4a* shows the relationship between the molecular distance and the LCLM distance for the 139 pairs of inbred lines. As previously observed [6], this relationship is triangular, and low phenotypic distances may correspond to relatively high molecular distances. On the contrary, two distinct sections appear in the relationship between the molecular distance and the MAXI distance in *figure 4b*. In the first section on the left of *figure 4b*, the three pairs with a MAXI distance lower than the 3.15 threshold also exhibit low molecular distances. The second section on the right of *figure 4b* corresponds to inbred lines which are distinguished by the experts. The minimum molecular distance in this section is a 5 % difference, and the molecular distance appears as a limiting factor for MAXI, just as it was for the molecular distance in connection with LCLM. Taking the molecular distance as a measure of relatedness between inbred lines, the comparison of *figure 4a* and *4b* shows that, in our experiment, the experts were able to separate related and non-related inbred lines more accurately than LCLM.

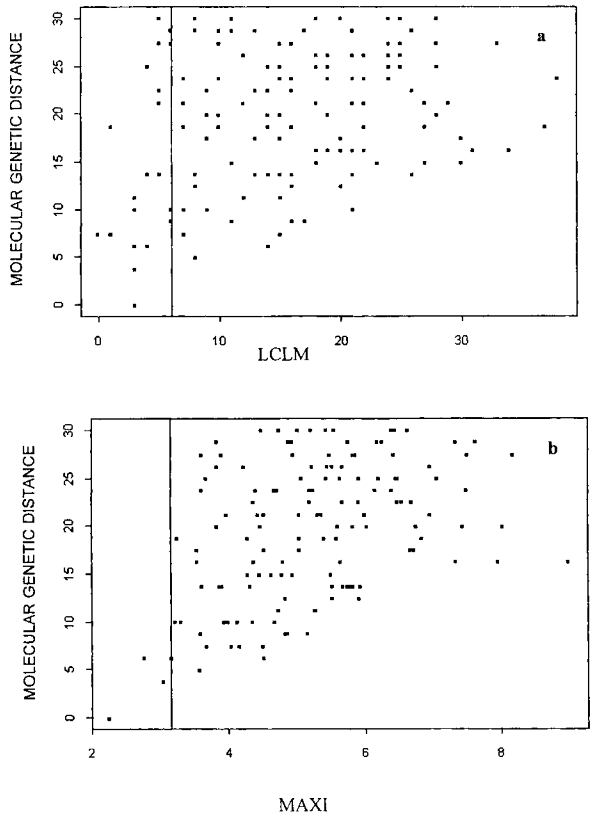

Figure 4. Relationship between the molecular genetic distance and the morphological distances (a) LCLM, and (b) MAXI. (a) The vertical line represents the LCLM = 6 threshold. (b) The vertical line represents the arbitrary threshold of MAXI = 3.15.

In order to fine tune this result, the molecular distance and the MAXI distance have been centred and reduced, and the ordering of the different pairs have been compared. We find a good agreement between the two classifications for pairs with molecular distances lower than 7 %. For pairs with molecular distances between 7 and 15 %, the experts are systematically overestimating the molecular distance: the standardised MAXI distance is always higher than the standardised molecular distance. For pairs with molecular distances higher than 15 %, there is no relationship between the expert distance and the molecular distance.

4. DISCUSSION

Our study was centred on the investigation of the relationship between several yardsticks which may be used to assess distinction or essential derivation in maize. In particular, given the important role played by maize experts, we were interested in the relationship between their opinion and morphological or molecular distances. First, we wanted to know the criteria on which the opinion of experts regarding distinction was based. Second, we wanted to know whether they were able to assess relatedness between two genotypes in the same way that molecular markers do.

Only one year of experimentation in one location has been performed with maize experts. As a result, some pairs of inbred lines have had to be removed from the study because of environmental effects (no germination or no flowering). Moreover, some discriminant traits have obviously not been taken into account by experts for some pairs of inbred lines. This emphasises the need to perform multilocational experiments over several years whenever phenotypical traits are concerned. However, additional years of experimentation would have probably resulted, in our case, in an enhancement of the discriminative power of the experts, as observed at GEVES with the LCLM distance index. Hence, the results obtained here concerning the visual appreciation of the genetic distance by maize experts seem to us sufficiently clear to set the basis for a reflection on the possible use of molecular markers in Plant Breeder's Rights.

The expert committee which makes the final decision concerning DUS in France is designated by CTPS (*Comité technique permanent pour la sélection*). It is the official body in charge of varietal registration, and consists of representatives of public and private maize breeding companies. They have a good knowledge of maize selection and of plant development. As the experts are the ones who are primarily interested in plant breeding protection, we have chosen to take their opinion as reference, and calibrated the other potentially discriminant traits according to this reference.

The present study clearly validates the French distinctness protocol. All the inbred lines studied here which are currently protected in France were also found distinct by the experts. Among the 139 pairs of inbred lines, none of them was distinguished by the morphological LCLM distance while found to be very close by the experts. It was, therefore, very easy to set an arbitrary minimum expert's distance for distinction. Moreover, this result emphasises the advantage of a two-stage practice which consists in screening all the new candidates along with a set of reference genotypes during the first year of observation, and then growing only relatively close pairs of candidate and reference varieties during the second year of observation. This practice facilitates the management of the reference collections, and allows to allocate more resources for the observation of the closest genotypes.

In general, expert's notations were more discriminant than LCLM morphological distances. Weights given by the experts to the different traits, integrating their knowledge of the crop, were probably different from LCLM weights. Experts may also take into account other traits than those of the UPOV maize guidelines. Our results suggest that the most discriminating traits in maize are the ones that are expressed at the end of the life-cycle of the plants.

Experts were also able to identify closely related inbred lines. They clearly found more similarity between inbred lines which are closely related, with a molecular distance less than 7 %, than between inbred lines which are alike on an UPOV morphological basis, but less related on a molecular basis. In that sense, experts perform better than morphological distances. As a matter of fact, they have a good knowledge of phylogenetic relationships within the species, and they know how the varieties they produce were obtained. It seems that experts are able to reconstruct the relatedness between two inbred lines from their phenotype, up to a limit situated around 90 % of genome shared.

Despite the very few pairs of closely related inbred lines present in this study, the present results can help design new experiments concerning essential derivation in maize. Further work should focus around 10 % of molecular divergence and include mostly closely related inbred lines. Within such a material, molecular markers are expected to be in linkage disequilibrium [5], and the relationships between molecular and morphological distance should be linear [4, 10]. To check whether experts are really able to estimate the relatedness between inbred lines, such work should also include pairs of unrelated inbred lines which are similar on a phenotypic basis.

Acknowledgements: This study was supported by GEVES and SEPROMA. The experiment has been conducted by Véronique Lacoste. We thank G.Y. Berthe and J. Guiard for useful discussions concerning the analysis and the interpretation of the results, and two anonymous reviewers for helpful comments on the manuscript.

REFERENCES

[1] Bar-Hen A., Charcosset A., Relationship between molecular and morphological distances in a maize inbred

lines collection. Application for breeder's rights protection, in: Van Oijen J.W., Jansen J. (Eds.), Biometrics in plant breeding: applications of molecular markers, Proc. Ninth Meeting Eucarpia Section Biometrics in Plant Breeding, 1995, pp. 57–66.

[2] Bar-Hen A., Charcosset A., Bourgoin M., Guiard J., Relationship between genetic markers and morphological traits in a maize inbred lines collection, Euphytica 84 (1995) 145–154.

[3] Bernardo R., Estimation of coefficient of coancestry using molecular markers in maize, Theor. Appl. Genet. 85 (1993) 1055–1062.

[4] Burstin J., Charcosset A., Relationship between phenotypic and marker distances: theoretical and experimental. Investigations, Heredity 79 (1997) 477–183.

[5] Charcosset A., Essioux L., The effect of population structure on the relationship between heterosis and heterozygosity at marker loci, Theor. Appl. Genet. 89 (1994) 336–343.

[6] Dillmann C., Bar-Hen A., Guérin D., Charcosset A., Murigneux A., Comparison of RFLP and morphological distances between inbred lines. Application to maize (*Zea mays* L.) inbred lines, Theor. Appl. Genet. 95 (1997) 92–102.

[7] Dillmann C., Charcosset A., Goffinet B., Smith J.S.C., Dattée Y., Best linear unbiased estimator of the molecular genetic distance between inbred lines, in: Krajewski P., Kaczmarek Z. (Eds.), Advances in biometrical genetics, Proc. Tenth Meeting Eucarpia Section Biometrics in Plant Breeding, 1997, pp. 105–110.

[8] Gruau, Logiciel de Comparaison de Lignées de Maïs, GEVES, 1989.

[9] Lynch M., Estimation of relatedness by DNA fingerprinting, Mol. Biol. Evol. 5 (1988) 584–599.

[10] Smith O.S., Smith J.S.C., Bowen S., Tenborg R.A., Wall S.J., Similarities among a group of maize inbred lines as measured by pedigree, F1 grain yield, grain yield, heterosis and RFLPs, Theor. Appl. Genet. 80 (1990) 833–840.

[11] Soller M., Beckmann J.S., Genetic polymorphism in varietal identification and genetic improvement, Theor. Appl. Genet. 47 (1983) 179–190.

[12] Tivang J.G., Nienhuis J.S., Smith O.S., Estimation of sampling variance of molecular marker data using booststrap procedure, Theor. Appl. Genet. 89 (1994) 259–264.

[13] Weir B.S., Gaut B.S., Variance for distances between plant sequences, New Zealand J. Bot. 31 (1993) 317–322.