TWC/15/16 Rev.
ORIGINAL: English
DATE: June 11, 1997

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Fifteenth Session
Budapest, June 3 to 5, 1997

REVISED DOCUMENT ON
IDENTIFICATION OF RYEGRASS (*LOLIUM SPP.)* CULTIVARS
BY MEANS OF AFLP MARKERS

Document prepared by experts from Belgium

IDENTIFICATION OF RYEGRASS (*Lolium spp*.) CULTIVARS BY MEANS OF AFLP MARKERS

Isabel Roldán-Ruiz[1], Kristiaan Van Laecke[1], Jan De Riek[1], Jochen Dendauw[1], Ann Depicker[2], Erik Van Bockstaele[1] & Marc De Loose[1].

[1] Department of Applied Plant Genetics and Breeding, Centre for Agricultural research, 9820-Merelbeke, Belgium.
[2] Laboratory of Genetics, University of Gent. K.L. Ledeganckstraat, 35, 9000-Gent, Belgium.

Introduction

It has been shown for different crops that DNA profiling techniques can help to distinguish among commercial cultivars (Zhang *et al.* 1993, Morrell *et al.* 1995). Most of the work has been done on cultivars produced by clonal propagation, inbreeding or selfing and little is known about the applicability of these methods to outcrossing crops. In outcrossing crops like *Lolium spp*. which are commercialised as synthetics, a cultivar is a population of related genotypes. In this case, genetic variation is expected to be found both within and among cultivars. This implies that when the purpose is to characterise a cultivar, a (representative) number of individuals should be analysed and appropriate statistical analyses should be used to test for significant differences between cultivars (Morrel *et al*. 1995). Not only the polymorphic markers but also the non polymorphic markers might be of interest when the purpose is to estimate genetic distances between cultivars.

AFLP (Vos *et al.* 1995) consists in the amplification of a subset of restriction fragments; by changing the number and the sequence of selective bases an almost infinite number of markers can be obtained per genotype. The advantages of AFLP over other molecular marker methods is that no sequence data are necessary and that large numbers of polymorphic markers can be scored in each reaction, an important characteristic when the information content of each individual marker is low, as it is the case with AFLP. AFLP allows a quick determination of 'allele frequencies' for many markers distributed over the genome and is one of the more powerful techniques for gene pool and population genetic analysis.

AFLP was originally developed as a radioactive technique. In this case ^{33}P or ^{32}P labelled primers are used and the visualisation of the amplification products takes place on a phosphorimager system or on X-ray films. An obvious step to increase the throughput of the AFLP technique is the implementation of gel electrophoresis and detection systems like the ABI Prism 377 DNA Sequencer (ABI377). The ABI377 automatically analyses DNA molecules labelled with fluorescent dyes. At the moment four different fluorescent 'colours' are available for labelling. The use of the ABI377 for AFLP analysis offers the advantage that, together with the sample, an internal lane standard (labelled with a different fluorescent dye) can be loaded in each lane, allowing a very accurate 'between lanes' and 'between gels' normalisation. The main disadvantage of AFLP markers is that, when scored using the ABI377, they behave in a dominant fashion.

In this work an evaluation is made of the AFLP marker technique for the identification of ryegrass cultivars. We present here the results for four diploid commercial cultivars of

Lolium perenne belonging to three European breeding companies. The influence of the number and identity of AFLP markers used on the final conclusions is discussed.

Material and methods

Plant material

Four diploid *Lolium perenne* cultivars were analysed. Between 48 and 51 plants from the third generation of each cultivar were individually fingerprinted (table 1).

Table 1. Characteristics of the plant material used.

Cultivar	Company	Number of fingerprinted plants
Merganda	RvP (BE)	51
Paddok	RvP (BE)	49
Morimba	Van der Have (NL)	49
Barylou	Barenbrug (NL)	48
	TOTAL	197

The plant material was obtained from 4-6 weeks-old seedlings that were nursed in the greenhouse. Approximately 100 mg fresh plant material was harvested per plant. After harvesting, the plant material was lyophilised, grounded in a mill and vacuum-packed. DNA was extracted using the CTAB procedure described by Weisig *et al.* (1991).

AFLP reactions

Individual DNA samples were used to generate AFLP markers. AFLP analysis was performed according to Vos *et al.* (1995) and using available kits (Gibco BRL and Perkin Elmer). The DNA was digested using *Eco*RI (hexa-cutter) and *Mse*I (tetra-cutter). Afterwards, double-stranded adapters were ligated to the restriction fragments. A pre-amplification step was performed using primers complementary to the *Eco*RI and *Mse*I adapters with an additional selective 3' nucleotide: *Eco*RI+A, *Mse*I+C. For the selective amplification both the *Eco*RI primer and the *Mse*I primer contained the same sequences as those used in the pre-amplification but with three selective nucleotides at the 3' end (in this way only 1/4096 fragments originated after the digestion were amplified): *Eco*RI+AGG, *Mse*I+CTT. Only the *Eco*RI primer was labelled with a fluoresceïne group 'JOE' (Perkin Elmer).

The samples were loaded on a 5% polyacrilamide gel and analysed with an ABI377. GS-500 Rox labelled size standard (Perkin Elmer) was loaded in each lane in order to facilitate the automatic analysis of the gel and the sizing of the fragments. The computer program GeneScan 2.0.2. was used for the analysis of the data and the generation of Sample Files.

Statistical analysis

The data were further analysed using Genotyper 2.0 (Perkin Elmer) for the analysis of the GeneScan files generated by the ABI377. Markers were selected which showed a relatively good amplification in at least one of the 197 plants analysed and that were easy to identify, if present, in the rest of the plants. In this way, a total of 72 markers were selected and automatically scored for presence/absence in all the 197 plants analysed. A scoring table (1/0) was generated that was further analysed using Microsoft Excel, Microsoft Access and the statistical package SPSS. The following calculations and analyses were performed:

1. Percentage of appearance of each marker in each cultivar;
2. Pairwise distances between plants belonging to the same and to different cultivars, using the Jaccard and the Hamman coefficients. The full data set (72 markers) was used for this calculation;

$$\text{Jaccard } (x,y) = \frac{a}{(a+b+c)}$$

$$\text{Hamman } (x,y) = \frac{(a+d)-(b+c)}{a+b+c+d}$$

where a = number of fragments present in x and y; b = number of fragments present in x and absent in y; c = number of fragments absent in x and present in y; d = number of fragments absent in x and y.

When using the Jaccard coefficient ‘d’ is not taken into account and it is assumed that the absence of a trait does not simply relate to the degree of similarity between individuals. It ranges from 0 to 1.

Hamman is a non-metric similarity measure that gives the probability that a characteristic has the same state in both items (present in both or absent from both) minus the probability that a characteristic has different states in the two items (present in one and absent from the other). It has a range of -1 to +1 and is monotonically related to the Simple Matching Coefficient.

3. Discriminant analysis. Only the polymorphic markers (69) were included in the analysis.

Afterwards, two subsets of markers were selected based on two different criteria. In the first subset a random selection of 39 markers was made; in the second subset only those markers which were present in at least 50% of the plants were included (37 markers fulfilled this criterion). The discriminant analysis was repeated in both cases and the results were compared with those of the whole data set.

Results

The differences found among cultivars were at the level of frequency of appearance of a number of markers and only a few could be considered ‘diagnostic’ markers, as they were

present (or absent) only in one of the cultivars studied. Only 3 markers of the total of 72 that were analysed, were not polymorphic and were present in all the 197 plants analysed.

Thirty seven markers showed appreciable differences among cultivars. Two of the 37 markers were exclusively present in one of the cultivars, three were present only in two cultivars and one was present in three cultivars and absent in one. The rest of the markers that showed differences (30) were present in all cultivars, but in different proportions.

In tables 2 and 3 are summarised the average similarities between couples of plants belonging to the same or to different cultivars using the Jaccard and the Hamman similarity measures respectively. As expected, the average similarity values between plants belonging to the same cultivar (diagonal in the tables) were in all cases significantly higher (ANOVA, $p<0.001$) than the average similarity values between plants belonging to different cultivars.

Table 2. Average similarity values calculated using the Jaccard's coefficient, between couples of plants belonging to the same and to different cultivars.

	Barylou	**Merganda**	**Morimba**	**Paddok**
Barylou	0.661			
Merganda	0.613	0.647		
Morimba	0.599	0.612	0.654	
Paddok	0.599	0.625	0.615	0.650

Table 3. Average similarity values calculated using the Hamman's coefficient, between couples of plants belonging to the same and to different cultivars.

	Barylou	**Merganda**	**Morimba**	**Paddok**
Barylou	0.499			
Merganda	0.435	0.519		
Morimba	0.397	0.406	0.506	
Paddok	0.400	0.474	0.443	0.509

A very interesting finding is that the highest average similarity between couples of plants belonging to different cultivars was that found between Paddok and Merganda (shadowed in the tables), which are two cultivars developed by the same breeding institute (RvP).

Performing discriminant analyses (Digby & Kempton 1987) we combined the information about all the markers to make the differences between cultivars clearer than is possible on the basis of the values of a single marker. It was chosen to work with this mutlivariate technique and not with the more known 'Correspondence Analysis (CA)' or 'Principal Component Analysis (PCA)' because by using discriminant analysis a differential weight is given to the markers; those markers which have an erratic appearance are unlikely to be reliable indicators and are given a lower weight than markers with more consistent abundance when constructing an ordination. The results obtained when all polymorphic markers (69) were included in the analysis are summarised in figure 1A. The four cultivars appear clearly differenciated and, based on the three discriminant functions obtained, only two plants were 'misclassified' (placed on a different cultivar).

The results regarding the second data set of 39 markers selected at random are shown in figure 1B. The capacity of discrimination of the analysis when based on this number of markers is lower than when based on the whole data set and the differentiation among cultivars is less clear. In this case 26 plants were 'misclassified'.

In figure 1C is represented the output of the discriminant analysis based on a subset of 37 markers, selected based on their abundance values. Thirty five plants were 'misclassified' in this case and the cultivars are not easily differenciated in the ordination diagram.

Discussion

The results show that AFLP markers are a powerful tool for identification purposes even for outcrossing crops. As expected in the case of outcrossers, the AFLP markers analysed were highly polymorphic among cultivars, but also within cultivars. Nevertheless, it was possible to differentiate clearly among cultivars and the differentiation was clearer as more markers were included in the analysis. Apparently, the number of markers included had a big influence on the capacity of discrimination of the analysis.

With respect to the definition of a 'minimum genetic distance' threshold for PBR, the figures presented in tables 2 and 3 demonstrate that although the distances between plants belonging to different cultivars are significantly higher than the distances between plants belonging to the same cultivar, the margin is quite small. Therefore, to determine a threshold for 'minimum genetic distance' it is necessary to perform a detailed analysis of the genetic distances between the cultivars that are currently accepted to be different based on morphological characteristics.

The striking result that the highest similarity was found between cultivars produced by the same institute (Paddok and Merganda) would strongly support the usefulness of this marker system in the genetic analysis of natural and commercial ryegrass populations. Nevertheless, these results should be taken with some reserve because they are based only on one primer combination and we still cannot predict the result of the inclusion of more markers (obtained from other primer combinations) in the analysis.

The Plant Breeder Rights Act 1994 places the onus on the breeder of a new cultivar to show that the cultivar meets the DUS criteria (Morrel *et al.* 1995). Following Stam (1994), for a fingerprinting method to be useful to a company it has to be easy to automate and consistent in results; both requisites are fulfilled by the AFLP technique, specially when a sequencer is used. Prerequisites for the use of DNA profiling techniques for cultivar identification are: (1) that the markers are reproducible; (2) that the inheritance of the marker is mendelian; (3) that the allele frequencies are stable. It has already been shown that AFLP markers are highly reproducible and that their inheritance is mendelian (Schondelmaier *et al.* 1996), what converts them in a promising tool for future applications.

References

Digby, P.G.N. & Kempton, R.A. (1987). *Multivariate analysis of ecological communities*. Chapman & Hall Ltd, New York.

Morrell, M.K.; Peakall, R.; Appels, R.; Preston, L.R. & Lloyd H.L. (1995). DNA profiling techniques for plant variety identification. Australian Journal of Experimental Agriculture 35: 807-819.

Schondelmaier, J.; Steinrücken, G. & Jung, C.(1996). Integration of AFLP markers into a linkage map of sugar beet (*Beta vulgaris* L.). *Plant Breeding* 115: 231-327.

Stam, P. (1994). Marker-assisted breeding. In: *Proceedings of the 9th meeting of the EUCARPIA Section Biometrics in Plant Breeding*, 6 - 8 July 1994, Wageningen (The Netherlands). J.W. Ooijen & J. Jansen (eds.). pp 32-44.

Vos, P., Hogers, R., Bleeker, M., Reijans, M.; van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J. Kuiper, M. & Zabeau, M. (1995). AFLP™: a new technique for DNA fingerprinting. *Nucleic Acids Research 23(21):* 4407-4414.

Weisig, K., Beyerman, B., Ramser, J. & Kahl, C. (1991). Plant DNA-fingerprinting with radioactive and dioxigenated oligonucleotide probes complementary to simple repetitive DNA sequences. *Electroforesis* 159-169.

Zhang, Q; Saghai Maroof, M.A.; Kleinhofs, A. (1993). Comparative diversity analysis of RFLPs and isozymes within and among populations of *Hordeum vulgare* spp. *spontaneum*. *Genetics* 134: 909-916.

Acknowledgements

The authors thank the companies Advanta-Van der Have & Barenbrug for providing the seeds of two of the cultivars. We also thank Ariane Staelens and Greet De Smet from the RvP staff, for excellent technical assistance.

This work was supported by a grant from the Flemish institute for the promotion of the scientific-technologic research in the industry (I.W.T.), and the Belgian ministry of middle class and agriculture.

[End of document]