Disclaimer: unless otherwise agreed by the Council of UPOV, only documents that have been adopted by the Council of UPOV and that have not been superseded can represent UPOV policies or guidance.

This document has been scanned from a paper copy and may have some discrepancies from the original document.

_____

Avertissement: sauf si le Conseil de l'UPOV en décide autrement, seuls les documents adoptés par le Conseil de l'UPOV n'ayant pas été remplacés peuvent représenter les principes ou les orientations de l'UPOV.

Ce document a été numérisé à partir d'une copie papier et peut contenir des différences avec le document original.

_____

Allgemeiner Haftungsausschluß: Sofern nicht anders vom Rat der UPOV vereinbart, geben nur Dokumente, die vom Rat der UPOV angenommen und nicht ersetzt wurden, Grundsätze oder eine Anleitung der UPOV wieder.

Dieses Dokument wurde von einer Papierkopie gescannt und könnte Abweichungen vom Originaldokument aufweisen.

_____

Descargo de responsabilidad: salvo que el Consejo de la UPOV decida de otro modo, solo se considerarán documentos de políticas u orientaciones de la UPOV los que hayan sido aprobados por el Consejo de la UPOV y no hayan sido reemplazados.

Este documento ha sido escaneado a partir de una copia en papel y puede que existan divergencias en relación con el documento original.

BMT/1/3
ORIGINAL : English/French
DATE : March 10, 1993

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

WORKING GROUP ON BIOCHEMICAL AND MOLECULAR TECHNIQUES AND DNA-PROFILING IN PARTICULAR

First Session
Geneva, April 19 and 20, 1993

THE ADVANTAGES AND LIMITS OF THE USE OF DNA POLYMORPHISM WITH RELATION TO PLANT VARIETY PROTECTION

Document prepared by experts from France

Introduction

Considerable progress has been made over the last ten years in the field of genetic marking as a result of developments in molecular biology techniques.

The investments made in mapping the human genome will lead, without doubt, to further diversification and simplification of these techniques, coupled with increased reliability and reproduceability.

Somewhat tardily, but nevertheless, significantly, the techniques for analysing DNA polymorphism are now being transferred to higher plants. The identification of genes coding for characteristics of agronomic value, or closely bound to the loci influencing those characteristics, is a main objective when seeking to use molecular biology in relation to the breeding of new varieties. Molecular techniques are also likely to be applied in assessing variety DUS and for establishing "essential derivation" between varieties.

Description of Different Methods of Molecular Biology

Although biochemical markers are very helpful as an identification tool for varieties, they evidence only a very small part of the variability existing in the genome. On the one hand it is rarely possible to have more than 20 to 30 genes per species studied (compared with the tens of thousands in one genome), while on the other hand, even the totality of the genes represents less than 10% of the genome so that the variability of 90% of the genome is not yet investigated.

Accordingly, molecular biology, which deals with the DNA itself, is becoming a tool for an increasing number of geneticists and population biologists. Also for the protection of plant varieties, some concern arises from the increasing number of varieties and the convergence of breeding objectives make the visible differences less and less obvious.

Organization of the DNA

Plant DNA is generally more complex than the DNA of animals. Whereas mammal cells contain 4 to 5 µg of DNA, plant cells may contain up to 100 µg. This DNA is located in the organelles (the chloroplast and the mitochondria) and the nucleus. The organelle DNA is relatively short (220 to 2,500 kb to be compared to 10^9 kb for the nuclear genome) and not very complex. The nuclear DNA on the other hand is highly redundant: only a little part of the DNA bears genetic information. This part, the genes, consists of "unique sequences" of nucleotides which are repeated once or a few times in the genome. For the greater part of the DNA, however, of which the function is not yet known, there appear a great deal of "repetitive sequences" of more or less important stretches of DNA which are repeated up to a few thousand times. During the evolution of a species, these repetitive sequences have accumulated variations like point mutations, deletions, duplications or amplifications of particular portions of the sequences. These repetitive sequences are called satellites and amount to more than 50% of the genome.

Polymorphism of the Organelle DNA

The variations occurring in the DNA of organelles can be evidenced by cutting the DNA (or digesting it) with restriction enzymes. These enzymes are able to recognize particular sequences of 4 to 6 nucleotides (restriction sites) and cut the DNA inside or near these particular sequences. Any mutation occurring at these restriction sites will cause the enzyme to fail to recognize and to cut the sequence. Thus, as different DNAs have their restriction sites located differently, they will yield restriction fragments (fragments of DNA obtained after the action of the restriction enzyme) with different sizes. Electrophoresis will separate these fragments according to their sizes and give patterns specific to each different DNA.

The polymorphism thus observed is useful at the interspecific level. It has been used to classify subspecies of rice (Dally and Second, 1990) and other species but the intravarietal polymorphism encountered is low, especially in chloroplasts where the DNA is very small and contains merely "useful" information since it cannot afford many mutations. However, variation in organelle DNA has been shown to be correlated with cytoplasmic male sterility in a number of species. This is the case with sorghum, beet and other crops (Lee et al., 1989; Mikami et al., 1986; De Courcel et al., 1989), so that the study of this DNA can be very useful in recognizing plants with different types of cytoplasmic sterility.

Nuclear DNA

The nuclear DNA is a great deal more complex, and more sophisticated methods are used to study its polymorphism. We shall describe techniques based on the RFLP method and on the PCR method.

RFLP (Restriction Fragment Length Polymorphism)

Unlike organelle DNA, nuclear DNA will not yield discrete bands after electrophoresis of its restriction fragments. On the contrary, since the hundreds of thousands of bands of every size obtained after digestion (due to the much larger size of the DNA), will give a smear on the gel where the eye will not be able to see the variability. Therefore, a property of the DNA, the fact that complementary DNA strands do spontaneously associate with each other, is exploited. Thus, a probe, consisting of a particular sequence of DNA, is added to the gel and left to associate (hybridize) with the matching sequence in the smear. If the probe has been radioactively or biochemically labelled prior to the hybridization, it will be possible to locate in the gel the particular sequence of DNA. A great number of probes exist already and are conserved in private and public banks. They can be of different types: genomic DNA (gDNA) or copy DNA (cDNA) or else they can be synthetic oligonucleotides. According to the number of loci thus evidenced they are called monolocus or multilocus probes. The variations observed with monolocus probes arise mostly from mutations in the restriction sites leading to differences in the length of the restriction fragments. Conversely, multilocus probes, which are generally satellite DNA, reveal another type of variability which is due to differences in the number of repetitions of the particular sequence of DNA investigated. This is named VNTR (Variation in the Number of Tandem Repeats) or fingerprinting. The best known of these microsatellites are Jeffreys probe (Jeffreys, 1985) and the probe M13 (Vassart et al., 1987) which have been used extensively in human and animal genetics and later to analyze plant genomes. A number of papers have been published on poplars, pine trees, apple trees, Arabidopsis, different Rosaceae and other plants species (Rogstad et al., 1988; Nybom et al., 1990; Zimmerman et al., 1989; ..).

Lately, the SRM probes (Simple Repeat Motives) have been found to be very variable in the human genome (Ali et al., 1985). These probes are synthetic oligonucleotides consisting in repeats of 2 to 5 nucleotides in the same order. They have been shown to be ubiquitous in the animal and vegetal kingdom (Studer and Epplen, 1990; Weising et al., 1991).

Compared to monolocus probes, multilocus probes yield more complex patterns whose genetic control is not easily established. But they give more information per gel and they are also codominant and inherited in a mendelian way. An analogy can be found in the comparison of enzymes and stock proteins. Most of these satellites are available anywhere and are not species specific so that it is not necessary to make new probes when changing species. For monolocus probes, when studying a new species it is possible to use probes from more or less related species (heterologous probes).

PCR (Polymerase Chain Reaction)

Although RFLP methods display a lot of advantages (almost unlimited number or possible markers, codominance, mendelian segregation) they also have some disadvantages which are:

- the length of experiments (about one week),
- the difficulty of the method; only trained technicians can perform the experiment;
- the use of radioactivity; at present, good results can be obtained with non radioactively labelled probes in the case of multilocus probes but it is not yet so with monolocus probes.

Therefore the techniques based on PCR have been extensibely studied over the last few years. They consist in the amplification of specific portions of the genome which can then be visualized on agarose gels. The conditions required for the amplification of a given gene are:

- to know the sequence of the two extremities of the gene,
- to possess the two complementary sequences (primers) so that the polymerase can amplify the gene.

Polymorphism in the amplification products may arise either from a mutation in the sequence hybridizing with the primer or from a mutation between the two primers. In this latter case, the mutation will be detected without further experiment if it changes the length of the amplified fragment (important insertions or deletions). Otherwise, it will be necessary to digest the amplified fragment (method called PCRFLP) to detect punctual mutations occuring at a restriction site. This method is much used in medecine but for plant identification the requirement of knowing the sequence is at present not very practical. However, in the near future, sequencing is likely to become easier and easier and the information contained in gene data banks is increasing exponentially. Moreover, as the technique is much easier than the RFLP (much quicker and not requiring good quality extracted DNA), it is expected that it will be used more and more.

Another PCR-based method much spoken about is RAPD (Random* Amplified Polymorphic DNA). For this method, no preliminary information about the genes is required. Two random sequences of 10 to 20 bp are used. If complementary sequences exist in the genome and if they are not too far from each other, the strech of DNA in between is amplified. Sometimes numerous complementary sequences are found so that electrophoresis of the amplified fragments yields a "fingerprint". As first shown for soybean by Williams et al. (1990), these fragments are highly polymorphic. Thus, the technique appears very promising. However, two disadvantages remain:

- the repeatability of the method is not yet perfect--the patterns obtained depend on the polymerase used, on the apparatus and many other factors,
- the RAPD markers are dominant and it is generaly not possible to distinguish heterozygotes from the dominant homozygotes.

PCR technology can also be very useful for investigating the polymorphism of the microsatellites. The microsatellites are short sequences of di or tri or tetra nucleotide repeats which occur in the genome. In humans, they are known to be numerous and very variable (Hazan et al., 1992). Two alleles have a difference of two or a few nucleotides, that is not great enough to be detected by electrophoresis in agarose gels. Therefore, sequence gels are used enabling minor difference in length to be evidenced. Microsatellites have recently been shown to occur also in plants (Beckmann and Soller, 1990) and their variability has been investigated in sunflower (Brunel, in press). Although not much is known yet on the microsatellites of plants, it seems that they could be interesting markers in view of their codominance.

---

* (Random means random association of the nucleotides used as primers. As far as this association has been defined, it becomes definite.)

The following table summarizes the advantages of the different methods.

	RFLP monolocus	RFLP multilocus	PCRFLP	RAPD	Microsa: tellites
Disadvantages					
Use of radioactivity	+	-	-	-	-
Difficulty	+	+	-	-	+/-
Advantages					
Repeatability	+	+	+	-	+
Codominance	+	+	+	-	+
Specificity	+	+	+	-	+
Polymorphism within species	+	+	?	+	?

DISCUSSION

Technical constraints

The limits on using molecular marking in relation to plant variety protection currently have two aspects:

- the speed with which techniques evolve creates uncertainty amongst users,
- consequently, the debate more frequently concerns the choice of technique than reflection on the way to use the results.

The rapid evolution of methods (RFLP, RAPD, AFLP, PCRFLP) constitutes a brake on their practical application, and that for a number of reasons.

Numerous publications describe the technical aspects exclusively and do not permit an assumption to be made on the polymorphism that could be evidenced.

Other results concern inter or intra-specific genetic diversity more directly, but rarely that of bred plant material. Thus, with the exception of maize, it still remains difficult to assess the amplitude of inter-varietal variability demonstrated by molecular marking.

Finally, certain methods are pushed too rapidly, long before their reliability and reproduceability have been proved, which may end in discouraging bodies such as UPOV whose vocation is not to develop methods but, on the contrary, to apply well established methods.

Today, as a result, there exists no experimental comparisons to validate one or other of the methods in relation to the aims of plant variety protection and, with the exception of maize, the attempts to apply molecular marking to variety descriptions are linked more to technical opportunity than to genuine methodological choice.

Currently, RFLP is the method in which the plant molecular biology laboratories are most skilled and which is most used in relation to plant breeding. However, this technique remains expensive and dependant, for some years yet, on the use of radioactivity, whereas the techniques derived from PCR, although less well-known at present, are more readily automatable and do not necessarily require the use of radioactive elements.

Despite these uncertainties, there are two main reasons for UPOV to avoid concentrating its discussions on technical choice alone.

UPOV must be positioned as a user of well-tried methods; its role is not to promote techniques but to assess their relevance to its aims.

Further, whatever the technique or techniques used, the information gathered presents a number of associated advantages:

- The observed polymorphism is independant of environmental conditions, while the method of sampling plant material is frequently simplified when compared with that required for biochemical analysis,
- the very high number of potential markers enables precise identification,
- the resemblance between two varieties, established on the basis of molecular markers, is probably the most reliable measurement of their genetic relationship and, consequently, an appropriate tool for establishing derivation between varieties,
- the power of analysis is such that the necessary investments in automating the main steps are sure to be made, particularly in the medical field.

Quality of the Information

From the point of view of plant variety protection, the major difference between current methods of analysis of DNA polymorphism concerns the extent to which they permit genectic interpretation.

Certain methods, particularly RFLP's, make it possible to associate the observed polymorphism with mendelian segregation and, consequently, to establish a genetic map in terms of chromosome loci or linkage groups. These methods are the most informative since the markers that are chosen can represent the whole set of chromosomes. Any new association between two markers reflects a genetic recombination at the chromosome site considered.

For some economically important species (maize, sunflower, tomato, soya, peas, etc...), genetic mapping has been put in hand. Harmonization of methods is already underway in public and private laboratories.

Other methods, (fingerprinting, use of multilocus probes, RAPD) do not have such an immediate effect on determining the genetic map. Whilst they give a genetic fingerprint, the interpretation of the differences observed between two such fingerprints cannot be directly linked to allelic variations observed at identified loci. This situation will doubtlessly remain the most frequent one for a long time yet for most plant species, partly because their genetics are not well-known or are highly complex and also because the mapping investment cannot be made for economic reasons.

In the case of plant variety protection, where the aim is to distinguish between varieties constituting a genetic innovation and, in the near future, to establish links of dependency between essentially derived varieties,

thoughts must be given at the level of the quality and relevance of the information provided by molecular markers.

The quality of the information provided by molecular markers varies depending on whether or not they have been mapped.

A comparison between two varieties on the basis of mapped markers makes it possible to forecast a degree of similarity of genomes looked at as a whole. The accuracy of the measurement depends on the number of markers used, but if the marks are well distributed on the chromosomes, the hypotheses of formal genetics lead to the conclusion that the observed similarity is in direct relationship with the relatedness of the varieties.

Any observed similarity that is greater than the mean similarity between two varieties taken at random from a reference population could thus be considered an indication of derivation and constitute one of the objective factors for assessing dependency.

The same reasoning may be followed if the markers are not mapped. However, in this case, whether the marks used are representative of the overall genome remains unknown and the direct relationship between the observed similarity and the relatedness of the varieties is subject to the unverifiable hypothesis that distribution of the marks is representative.

It is important to note that in both cases, whether the mapping is known or unknown, each marker taken individually only constitutes a minor piece of information and that conclusions cannot be drawn from a genetic point of view unless it is one of a set of markers distributed to the extent that this is technicaly possible over the whole genome.

Subject to having a sufficient number of markers revealed by reliable techniques, the quality of the genetic information obtained cannot be called into question. All the work done on human genetics or on animal genetics shows clearly the accuracy that can be achieved in detecting relationships and relatedness.

Additionally, it is possible, and even probable, that molecular markers reveal the residual heterogeneity within lines. These markers are in most cases inactive and are not observed during breeding and may therefore remain polymorphous within a homogeneous, uniform and stable line without that fact indicating a lack of homogeneity or stability due to residual heterozygosity. That type of heterogeneity, where it has no consequence for the phenotypical homogeneity of the variety nor its stability, should be tolerated since an overstringent regulation of markers of this type would not be acceptable to the breeders and, indeed, would have no genetic basis.

This probable existence of residual heterogeneity in the molecular markers within a variety must be taken into account in our thinking alongside the need to base distinctness on an overall set of observed differences.

Relevance of Molecular Markers for Plant Variety Protection

The relevance of molecular markers in relation to the aims of plant variety protection is a different question from that of the quality of the information. The relevance of such markers must be discussed on the basis of their complementary rather than their exclusive use.

In the great majority of situations, the polymorphism revealed by the molecular markers is an inactive polymorphism that cannot be linked to a phenotypical expression. It often reflects tiny variations at DNA level. It may also affect repeated or highly repeated DNA sequences whose role in plants is little known. Therefore, a difference concerning just one or several markers taken in isolation cannot be held to constitute a significant difference at the genetic level. In the same way as for enzymatic polymorphism, for the purposes of distinctness, DNA polymorphism does not signify the existence of genetic innovation unless it is observed at sufficient number of sites or in conjunction with other morpho-physiological differences.

Furthermore, if we disregard the mutations (often single gene) affecting characteristics such as flower coloration, the genetic control of characteristics of interest when growing and using varieties is frequently polygenic. Even if research is carried out in this area, genes of agronomic interest will not be identified nor located, since the generally accepted hypothesis is that they are distributed over the chromosomes.

This supports the suggestion that using molecular markers for the purposes of plant variety protection can only be justified if the differences between the markers reflect sufficient differences between the chromosomes, particularly by new associations, which means that a large number must be available. Whether the molecular markers reflect the differences at the level of expressed DNA or non-expressed DNA is not the main topic of the discussion since, in every case, the variations shown are very small and, as things currently stand, can but rarely be linked to a phenotypical expression.

Despite this reservation, the number of possible markers, their total independence from the environment, their capability to reveal variations over the whole genome, mean that they constitute a remarkably accurate tool for analysis.

UPOV normally uses distinctness criteria that are most often polygenic and are subject to variation with environment. The phenotypical variations make it difficult to establish minimum distances. These difficulties will be even more marked when one attemps to establish essential derivation between varieties that are close genetically.

One can foresee an overall approach to distinctness or to essential derivation, based on the complementarity of the information provided by precise genetic probes such as biochemical and molecular markers and the distinctness criteria observed on plants. Methodological research on this subject is underway at GEVES in France.

This approach has the advantage of enabling the observed differences to be associated and weighted according to their importance. It should also enable the genetic distances to be assessed by taking into account divergencies at both the genome level and the level of the expressed characteristics. It reduces the risk of establishing distinctness on the basis of characteristics unrelated to growing the variety and it gives an estimate of genetic convergence on which the concept of essentially derived varieties is based.

REFERENCES

ALI S., MULLER C.R., EPPLEN J.T. (1986). Hum. Genet., 74 : 239-243.

BECKMANN J.S., SOLLER S. (1990) Bio/Technology, 8 : 930-932

DALLY A.M., SECOND G. (1990) Theor. Appl. Genet., 80 : 209-222.

DE COURCEL A.G.L., VEDEL F., BOUSSAC J.M. (1989). Theor. Appl. Genet. 77 : 793-798.

DEVOS K.M., GALE M.D. (1992). Theor. Appl. Genet. 84 : 567-572.

HAZAN J., DUBAY C., PANKOWIAK M.P., BECUWE N., WEISSENBACH J. (1992). Genomics, 12 : 183-189.

JEFFREYS A.J., WILSON V., THEIN S.L. (1985). Nature, 314 : 67-73

KADOWAKI K., YAZAKI K., OSUMI T., HARADA K., KATSUTA M., LEE S.H., MUTHUSKRISHNAN, SORENSEN E.L. and LIANG G.H. (1988). Theor. Appl. Genet., 77 : 379-382.

MIKAMI T., HARADA T., HISHIMA Y., SUGIURA and KINOSHITA T. (1986). Gamma Field Symposium, 25 : 131-143

NAKAGAHRA M. (1988). Theor. Appl. Genet. 78 : 809-814.

NYBOM H., ROGSTAD S.H. and SCHALL B.A. (1990). Theor. Appl. Genet., 79 : 153-156.

RIEDEL G.E., SWANBERG S.L., KURANDA K.D., MARQUETTE K., LAPAN P., BLEDSOE P., KENNEDY A., LIN B.Y. (1990). Theor. Appl. Genet., 80 : 1-10.

ROGSTAD S.H., PATTON J.C. and SHAAL B.A. (1988). Proc. Natl. Acad. Sci., USA, 85 :9176-9178.

TAUTZ D., RENZ M. (1984). Nucleic Acids Res., 12 : 2147-4138.

VASSART G., GEORGES M., MONSIEUR R., BROCAS H., LEGUARRE A.S., CHRISTOPHE D. (1987). Science, 235 : 683-684.

WEISING K., BEYERMANN B., RAMER J., KAHL G. (1991). Electrophoresis, 12 : 159-169.

WILLIAMS J.G.K., KUBELIK A.R., LIVAK K.J., RAFALSKI J.A., TINGEY S.V. (1990). 18-22 : 6531-6535.

ZIMMERMAN P.A., LANG-UNNASCH N. and CULLIS C.A. (1989). Genome, 32 : 824-828.

[End of Document]