Chapter 14 in: **Which DNA Marker for Which Purpose?** Final Compendium of the Research Project **Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees** in the European Union DGXII Biotechnology FW IV Research Programme **Molecular Tools for Biodiversity**. Gillet, E.M. (ed.). 1999. URL http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm

Application of the AFLP® technique in marker assisted breeding

Johan Peleman

Keygene N.V., Agro Business Park 90, P.O. Box 216, 6700 AE Wageningen, The Netherlands
Email: johan.peleman@keygene.com

1. Introduction

Through the development of RFLP markers in the early eighties, indirect selection in plant breeding using DNA markers became technically feasible. However, the laborious nature of the RFLP technique prevented a broad application of RFLP markers for marker assisted breeding. In the late eighties and the early nineties, molecular diagnostic methods based on PCR technology have been developed, such as RAPD, AFLP and microsatellites. These methods allowed an efficient detection of DNA fragments starting from small amounts of DNA. The AFLP technique offers the additional advantage that it is a robust technology which, unlike microsatellites, does not require an initial investment as it can be applied instantaneously on any crop. In an appropriate laboratory set-up this technology can now be used as a tool for marker-assisted breeding in a cost-effective way.

2. The AFLP technique

AFLP's (amplified fragment length polymorphisms) are molecular markers obtained by selective PCR amplification of restriction fragments. The technique involves three steps: 1. Restriction enzyme digestion; 2. Ligation of adaptors; 3. Selective amplification of restriction fragments. The selective amplification is based on the recognition of unique nucleotides flanking the restriction site. The principle of selective amplification can be used to adjust the number of fragments that are amplified in a single PCR reaction. Optimal amplification of genomic DNA fragments is obtained by using a combination of a rare cutter and a frequent cutter for the template preparation, followed by detection of the amplified fragments on denaturing polyacrylamide gels. The technique offers great flexibility in the number of loci that can be co-amplified in one PCR. Typically, 50 to 100 restriction fragments are co-amplified in one fingerprint. This technique is therefore extremely powerful for the identification of restriction fragment polymorphisms. An additional advantage is that relatively few primer pairs are needed to visualize a large number of loci. The AFLP technique has been applied successfully on a wide variety of plant species.

Extensive research, including the generation of many genetic maps in different crops, has demonstrated that the majority of the amplified restriction fragments correspond to unique loci in the genome. In collaboration with laboratories throughout the world, reference AFLP linkage maps are being constructed using segregating populations that have been mapped before by using RFLP markers. This approach allows the integration of AFLP and RFLP linkage maps. Such reference RFLP/AFLP maps have been or are being constructed in maize, wheat, barley, rice, oilseed rape,

sunflower, potato, tomato, lettuce, pepper, cucumber, *Arabidopsis*.

The AFLP technique generates highly informative gel images, in some cases containing up to » 8000 informative data points per sequencing gel. This huge information content has triggered the need for an image analysis software package that can convert these complex gel images into data sets that only contain the scores of the informative marker bands. Such specific software has been developed in Keygene. The software recognizes the different lane boundaries, identifies all the bands in the gel, corrects for mobility differences that may have occurred within the gel, and quantifies the intensity of all the bands. Subsequently, appropriate statistical analysis of the relative intensity of each band throughout the gel allows the automated scoring of the band. Moreover, extensive analysis of the relative intensities of the marker bands in segregating populations has demonstrated that the AFLP reactions are quantitative, and that co-dominant scoring (distinguishing homo- from heterozygosity) of the AFLP markers using the above described software can be achieved with a high level of confidence.

3. Applications in plant breeding

The most straightforward applications of the AFLP technique in marker assisted breeding include genetic distance analysis, variety identification, isolation of markers tightly linked to specific genes, and marker assisted backcrossing. These applications are discussed into more detail.

3.1. Genetic distance analysis, variety identification and seed purity analysis

Genotyping using DNA markers can be considered as the most reliable method for the identification of lines and varieties. Therefore the DNA fingerprinting methods can be used to analyze the purity of seed lots. Genetic distance analysis can be a powerful tool for breeders to identify different heterotic groups and to increase the efficiency of finding crosses with good combinability. To determine the genetic distance between lines and groups of lines, the lines are fingerprinted and the marker-presence or absence is scored for each line. Based on the obtained score table, similarity indices can be calculated for all combinations of lines. Subsequently, the relatedness amongst the lines can be visualized using a dendrogram display or PCA plots.

3.2. Indirect selection

Indirect selection can be an advantageous method of selection in plant breeding. Especially for traits for which the phenotypic tests are unreliable or expensive, markers can offer a solution. Before indirect selection can be applied, the genetic basis of the trait of interest needs to be elucidated and markers linked to the gene(s) of interest have to be identified. Once linked markers have been identified, the AFLP markers can be converted into simple PCR assays, which allow screening of large numbers of plants for the trait of interest in a cost effective manner. A suitable linked DNA marker should allow the prediction of the phenotype in a broad range of the germplasm. The occurrence of multiple alleles in the germplasm for a desired locus may sometimes complicate the identification of markers with a good predictive value in the germplasm.

3.2.1. Monogenic traits

For the identification of markers linked with monogenic traits, different approaches can be followed. The preferable approaches are all based on screening a limited number of samples with a relatively large number of primer pairs. This way, many loci can be screened with a limited effort. The number of lanes per fingerprint can be limited by screening on set(s) of Near Isogenic Lines

(NIL's), if these are available. Candidate markers that are identified in this way are then screened on a panel of phenotypically well characterised lines to confirm their linkage and to determine the predictive value of the markers. Another efficient approach consists of the 'Bulked Segregant Analysis' (B.S.A.) method (Michelmore *et al*., 1991). For this type of screening, individuals from a segregating population are pooled on the basis of their phenotype, and the pools are then screened until a sufficient number of markers emerges. This method can be used for both dominant and recessive monogenic traits. For dominant genes, 'cis' markers (linked with the trait of interest) will emerge from the screening, whereas 'trans' markers (linked with the opposite allele) will be identified for recessive traits.

The B.S.A. approach may also be useful for the identification of linked markers for oligogenic traits. This has been demonstrated in an AFLP marker screening for one type of disease resistance in tomato. The B.S.A. screening yielded trans markers for a recessive gene, involved in the resistance. When screening the trans markers on individuals of the segregating population, the resistance could only be predicted in » 75% of individuals, and the remaining 25% of individuals could not be explained as recombinants. The results suggested that an additional (dominant) gene is needed for full resistance. By bulking the sensitive individuals in which the resistance cannot be predicted by the markers for the recessive gene, and creating a 'positive' bulk of resistant individuals on the other hand, a screening can then be performed for markers linked with the putative dominant gene. The screening for this additional gene is currently underway.

3.2.2. Polygenic (quantitative) traits

The classical approach for the identification of loci involved in complex polygenic traits consists in the screening of a large number of individuals from a segregating population with a set of markers that are evenly distributed throughout the genome. Subsequently, statistical analysis is performed to identify regions in the genome that are involved in the trait. The laborious nature of this approach makes it unrealistic to screen sufficiently large populations to precisely locate the quantitative trait loci (QTL). As a consequence, the QTL cannot be localized precisely on the map and closely linked markers cannot be obtained, thereby preventing the broad scale application of indirect selection for quantitative traits.

For this reason, a new approach for the identification of QTL markers, based on the B.S.A. principle, was investigated. With this goal, an oilseed rape F2 population of » 2500 individuals segregating for two quantitative traits, glucosinolate and erucic acid contents, has been used. Based on the phenotypic scores, bulks were composed and approximately 2000 loci were screened on those bulks using the AFLP fingerprinting technique. Candidate markers that were identified on the bulk screening were then analyzed on » 200 randomly chosen individuals of the F2 population. This screening demonstrated that the candidate markers identified using the B.S.A. screening approach were derived from three different loci involved in glucosinolate content and two QTL involved in erucic acid content. The results have been confirmed by independent studies in which the same map positions have been identified to be involved in the respective traits (Toroser *et al*., 1995; Jourdren *et al*., 1996). The results demonstrate that a B.S.A. strategy may be useful even for the identification of markers for quantitative traits.

3.3. Marker Assisted Backcross Breeding

With the cost reductions that can be achieved using AFLP technology, Marker Assisted Backcross breeding is now at the verge of becoming a standard application in modern plant breeding. Two different aspects can be distinguished in backcross breeding: (i) Selection for high recurrent parent genomic content: In this application, the DNA fingerprints are used to calculate the % recurrent

parent genome in each backcross individual, hereby taking the genome representation of the markers into account. (ii) Selection against linkage drag: When negative characteristics are linked with the trait that needs to be introgressed, molecular markers can be used to select for recombinants in the region. After phenotypic testing of these recombinants, individuals may be selected in which the region responsible for the linkage drag has been removed from the locus of interest.

4. Conclusion

The efficiency that can be obtained using the AFLP technique for the identification of linked markers and selection in backcross populations demonstrates the power of this technique for marker assisted breeding applications. The technique may therefore open the way for a broad application of molecular markers in plant breeding.

5. References

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

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

Toroser D, Thormann CE, Osborn TC, Mithen R (1995) RFLP mapping of quantitative trait loci controlling seed aliphatic-glucosinolate content in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* 91: 802-808.

Jourdren C, Barret P, Horvais R, Foisset N, Delourme R, Renard M (1996) Identification of RAPD markers linked to the loci controlling erucic acid level in rapeseed. *Molecular Breeding* 2: 61-71.