

AFLP: a new technique for DNA fingerprinting

Pieter Vos*, Rene Hogers, Marjo Bleeker, Martin Reijans, Theo van de Lee, Miranda Hornes, Adrie Frijters, Jerina Pot, Johan Peleman, Martin Kuiper and Marc Zabeau

Keygene N.V., PO Box 216, Wageningen, The Netherlands

Received July 14, 1995; Revised and Accepted October 5, 1995

ABSTRACT

A novel DNA fingerprinting technique called AFLP is described. The AFLP technique is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. The technique involves three steps: (i) restriction of the DNA and ligation of oligonucleotide adapters, (ii) selective amplification of sets of restriction fragments, and (iii) gel analysis of the amplified fragments. PCR amplification of restriction fragments is achieved by using the adapter and restriction site sequence as target sites for primer annealing. The selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites. Using this method, sets of restriction fragments may be visualized by PCR without knowledge of nucleotide sequence. The method allows the specific co-amplification of high numbers of restriction fragments. The number of fragments that can be analyzed simultaneously, however, is dependent on the resolution of the detection system. Typically 50–100 restriction fragments are amplified and detected on denaturing polyacrylamide gels. The AFLP technique provides a novel and very powerful DNA fingerprinting technique for DNAs of any origin or complexity.

INTRODUCTION

DNA fingerprinting involves the display of a set of DNA fragments from a specific DNA sample. A variety of DNA fingerprinting techniques is presently available (1–11), most of which use PCR for detection of fragments. The choice of which fingerprinting technique to use, is dependent on the application e.g. DNA typing, DNA marker mapping and the organism under investigation e.g. prokaryotes, plants, animals, humans. Ideally, a fingerprinting technique should require no prior investments in terms of sequence analysis, primer synthesis or characterization of DNA probes. A number of fingerprinting methods which meet these requirements have been developed over the past few years, including random amplified polymorphic DNA (RAPD; 8), DNA amplification fingerprinting (DAF; 9) and arbitrarily primed PCR (AP-PCR; 10,11). These methods are all based on the amplification of random genomic DNA fragments by arbitrarily selected

PCR primers. DNA fragment patterns may be generated of any DNA without prior sequence knowledge. The patterns generated depend on the sequence of the PCR primers and the nature of the template DNA. PCR is performed at low annealing temperatures to allow the primers to anneal to multiple loci on the DNA. DNA fragments are generated when primer binding sites are within a distance that allows amplification. In principle, a single primer is sufficient for generating band patterns. These new PCR based fingerprinting methods have the major disadvantage that they are very sensitive to the reaction conditions, DNA quality and PCR temperature profiles (12–16), which limits their application.

This paper describes a new technique for DNA fingerprinting, named AFLP. The AFLP technique is based on the detection of genomic restriction fragments by PCR amplification, and can be used for DNAs of any origin or complexity. Fingerprints are produced without prior sequence knowledge using a limited set of generic primers. The number of fragments detected in a single reaction can be 'tuned' by selection of specific primer sets. The AFLP technique is robust and reliable because stringent reaction conditions are used for primer annealing; the reliability of the RFLP technique (17,18) is combined with the power of the PCR technique (19–21). This paper describes several features of the AFLP technique and illustrates how the technique can best be used in fingerprinting of genomic DNAs.

MATERIALS AND METHODS

DNAs, enzymes and materials

Lambda DNA was purchased from Pharmacia (Pharmacia LKB Biotechnology AB, Uppsala, Sweden). *Autographa californica* Nuclear Polyhedrosis Virus DNA (AcNPV) was a kind gift from Dr Just Vlak, Department of Virology, Agricultural University of Wageningen, The Netherlands, and was isolated as described previously (22). *Acinetobacter* DNA was a kind gift from Dr Paul Jansen, Department of Microbiology, University of Gent, Belgium, and was isolated from strain LMG 10554 according to the procedure of Pitcher *et al.* (23). Yeast DNA was isolated from strain AB1380 as described by Green and Olson with minor modifications (24). Tomato DNA (culture variety [cv] Money-maker, obtained from Dr Maarten Koornneef, University of Wageningen, The Netherlands), Arabidopsis DNA (Recombinant Inbred Line 240, obtained from Dr Caroline Dean, John Innes Center, Norwich, UK), maize DNA (strain B73, obtained from Dr Mario Motto, Istituto Sperimentale per La, Bergamo, Italy), cucumber DNA (cv Primera, obtained from De Ruiter

* To whom correspondence should be addressed

Seeds C.V., Bleiswijk, The Netherlands), barley DNA (cv Ingrid, obtained from Dr Paul Schulze-Lefert, University of Aachen, Germany), lettuce DNA (cv Calmar, obtained from Dr Richard Michelmore, UC Davis, Davis, CA, USA) and brassica DNA (oil seed rape, cv Major, obtained from Dr Thomas Osborn, University of Wisconsin, Madison, WI, USA) were isolated using a modified CTAB procedure described by Stewart and Via (25). Human DNA was prepared as described by Miller *et al.* (26) from a 100 ml blood sample of Mrs Marjo Bleeker, one of the co-authors of this paper. All restriction enzymes were purchased from Pharmacia (Pharmacia LKB Biotechnology AB, Uppsala, Sweden), except for the restriction enzyme *MseI*, which was purchased from New England Biolabs Inc. (Beverly, MA, USA). T4 DNA ligase and T4 polynucleotide kinase were also obtained from Pharmacia (Pharmacia LKB Biotechnology AB, Uppsala, Sweden). All PCR reagents and consumables were obtained from Perkin Elmer Corp. (Norwalk, CT, USA). All radioactive reagents were purchased from Amersham (Amersham International plc, Little Chalfont, Buckinghamshire, UK) or Isotopchim (Isotopchim SA, Ganagobie, France).

AFLP primers and adapters

All oligonucleotides were made on a Biotronic Synostat D DNA-synthesizer (Eppendorf GmbH, Maintal, Germany) or Milligen Expedite 8909 DNA-synthesizer (Millipore Corp. Bedford, MA, USA). The quality of the crude oligonucleotides was determined by end-labeling with polynucleotide kinase and [γ - 32 P]ATP and subsequent electrophoresis on 18% denaturing polyacrylamide gels (27). Oligonucleotides were generally used as adapters and primers for AFLP analysis without further purification.

AFLP adapters consist of a core sequence and an enzyme-specific sequence (28). The structure of the *EcoRI*-adapter is:

5-CTCGTAGACTGCGTACC
CATCTGACGCATGGTTAA-5

The structure of the *MseI*-adapter is:

5-GACGATGAGTCCTGAG
TACTCAGGACTCAT-5

Adapters for other 'rare cutter' enzymes were identical to the *EcoRI*-adapter with the exception that cohesive ends were used, which are compatible with these other enzymes. The *TaqI*-adapter was identical to the *MseI*-adapter with the exception that a cohesive end was used compatible with *TaqI*.

AFLP primers consist of three parts, a core sequence, an enzyme specific sequence (ENZ) and a selective extension (EXT) (28). This is illustrated below for *EcoRI*- and *MseI*-primers with three selective nucleotides (selective nucleotides shown as NNN):

	CORE	ENZ	EXT
<i>EcoRI</i>	5-GACTGCGTACC	AATTC	NNN-3
<i>MseI</i>	5-GATGAGTCCTGAG	TAA	NNN-3

AFLP-primers for other 'rare cutter' enzymes were similar to the *EcoRI*-primers, and *TaqI*-primers were similar to the *MseI*-primers, but have enzyme-specific parts corresponding to the respective enzymes.

Modification of DNA and template preparation

The protocol below describes the generation of templates for AFLP reactions using the enzyme combination *EcoRI/MseI*. DNA templates with other restriction enzymes were prepared using essentially the same protocol, except for the use of different restriction enzymes and corresponding double-stranded adapters.

Genomic DNA (0.5 μ g) was incubated for 1 h at 37°C with 5 U *EcoRI* and 5 U *MseI* in 40 μ l 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/ μ l BSA. Next, 10 μ l of a solution containing 5 pMol *EcoRI*-adapters, 50 pMol *MseI*-adapters, 1 U T4 DNA-ligase, 1 mM ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/ μ l BSA was added, and the incubation was continued for 3 h at 37°C. Adapters were prepared by adding equimolar amounts of both strands; adapters were not phosphorylated. After ligation, the reaction mixture was diluted to 500 μ l with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, and stored at -20°C.

AFLP reactions

Amplification reactions are described using DNA templates for the enzyme combination *EcoRI/MseI*. AFLP fingerprints with other enzyme combinations were performed with appropriate primers.

AFLP reactions generally employed two oligonucleotide primers, one corresponding to the *EcoRI*-ends and one corresponding to the *MseI*-ends. One of two primers was radioactively labeled, preferably the *EcoRI*-primer. The primers were end-labeled using [γ - 33 P]ATP and T4 polynucleotide kinase. The labeling reactions were performed in 50 μ l 25 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 5 mM DTT, 0.5 mM spermidine-3HCl using 500 ng oligonucleotide primer, 100 μ Ci [γ - 33 P]ATP and 10 U T4 polynucleotide kinase. Twenty μ l PCRs were performed containing 5 ng labeled *EcoRI*-primer (0.5 μ l from the labeling reaction mixture), 30 ng *MseI*-primer, 5 μ l template-DNA, 0.4 U *Taq* polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM MgCl₂, 50 mM KCl, 0.2 mM of all four dNTPs.

The PCR conditions differed depending on the nature of the selective extensions of the AFLP primers used for amplification. AFLP reactions with primers having none or a single selective nucleotide were performed for 20 cycles with the following cycle profile: a 30 s DNA denaturation step at 94°C, a 1 min annealing step at 56°C, and a 1 min extension step at 72°C. AFLP reactions with primers having two or three selective nucleotides were performed for 36 cycles with the following cycle profile: a 30 s DNA denaturation step at 94°C, a 30 s annealing step (see below), and a 1 min extension step at 72°C. The annealing temperature in the first cycle was 65°C, was subsequently reduced each cycle by 0.7°C for the next 12 cycles, and was continued at 56°C for the remaining 23 cycles. All amplification reactions were performed in a PE-9600 thermocycler (Perkin Elmer Corp., Norwalk, CT, USA).

AFLP fingerprinting of complex genomes generally involved an amplification in two steps. The first step of this amplification procedure, named preamplification, was performed with two AFLP primers having a single selective nucleotide as described above, with the exception that 30 ng of both AFLP primers was used, and that these primers were not radioactively labeled. After this preamplification step, the reaction mixtures were diluted 10-fold with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0, and used

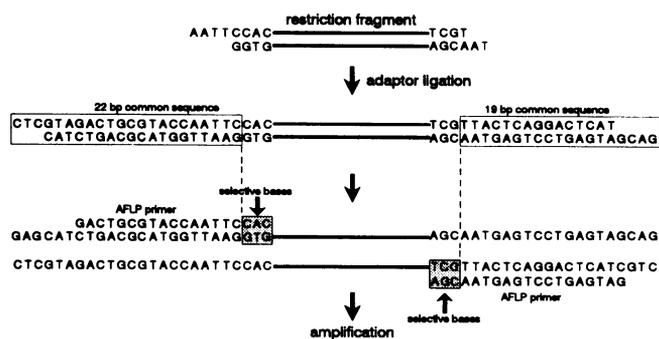


Figure 1. Schematic representation of the AFLP technique. Top: *EcoRI*-*MseI* restriction fragment with its 5' protuding ends. Center: the same fragment after ligation of the *EcoRI* and *MseI* adapters. Bottom: both strands of the fragment with their corresponding AFLP primers. The 3' end of the primers and their recognition sequence in the *EcoRI*-*MseI* fragment are highlighted.

as templates for the second amplification reaction. The second amplification reaction was performed as described above for AFLP reactions with primers having longer selective extensions.

Gel analysis

Following amplification reaction products were mixed with an equal volume (20 μ l) of formamide dye (98% formamide, 10 mM EDTA pH 8.0, and bromo phenol blue and xylene cyanol as tracking dyes). The resulting mixtures were heated for 3 min at 90°C, and then quickly cooled on ice. Each sample (2 μ l) was loaded on a 5% denaturing (sequencing) polyacrylamide gel (27). The gel matrix was prepared using 5% acrylamide, 0.25% methylene bisacryl, 7.5 M urea in 50 mM Tris/50 mM Boric acid/1 mM EDTA. To 100 ml of gel solution 500 μ l of 10% APS and 100 μ l TEMED was added and gels were cast using a SequiGen 38 \times 50 cm gel apparatus (BioRad Laboratories Inc., Hercules, CA, USA). 100 mM Tris/100 mM Boric acid/2 mM EDTA was used as running buffer. Electrophoresis was performed at constant power, 110 W, for ~2 h. After electrophoresis, gels were fixed for 30 min in 10% acetic acid dried on the glass plates and exposed to Fuji phosphoimage screens for 16 h. Fingerprint patterns were visualized using a Fuji BAS-2000 phosphoimage analysis system (Fuji Photo Film Company Ltd, Japan).

RESULTS AND DISCUSSION

Principle of the method

The AFLP technique is based on the amplification of subsets of genomic restriction fragments using PCR. DNA is cut with restriction enzymes, and double-stranded (ds) adapters are ligated to the ends of the DNA-fragments to generate template DNA for amplification. The sequence of the adapters and the adjacent restriction site serve as primer binding sites for subsequent amplification of the restriction fragments (Fig. 1). Selective nucleotides are included at the 3' ends of the PCR primers, which therefore can only prime DNA synthesis from a subset of the restriction sites. Only restriction fragments in which the nucleotides flanking the restriction site match the selective nucleotides will be amplified (Fig. 1).

The restriction fragments for amplification are generated by two restriction enzymes, a rare cutter and a frequent cutter. The AFLP

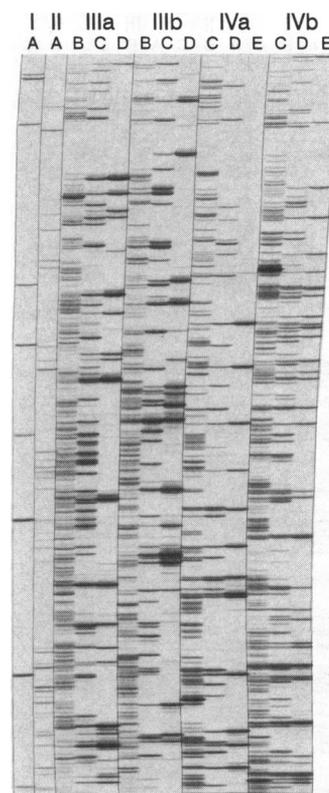


Figure 2. AFLP fingerprints of genomic DNAs of various complexities: λ DNA (panel I), AcNPV DNA (panel II), Acinetobacter DNA (panels IIIa and IIIb) and yeast DNA (panels IVa and IVb). Letters A, B, C, D and E refer to none, one, two, three and four selective bases in the AFLP primers respectively. The primer combinations used were from left to right: I. *EcoRI*+0/*MseI*+0, II. *EcoRI*+0/*MseI*+0, IIIa. *EcoRI*+0/*MseI*+A, *EcoRI*+C/*MseI*+A, *EcoRI*+C/*MseI*+AT, IIIb. *EcoRI*+0/*MseI*+T, *EcoRI*+C/*MseI*+T, *EcoRI*+C/*MseI*+TA, IVa. *EcoRI*+C/*MseI*+G, *EcoRI*+C/*MseI*+GC, *EcoRI*+CA/*MseI*+GC, IVb. *EcoRI*+C/*MseI*+T, *EcoRI*+C/*MseI*+TA, *EcoRI*+CA/*MseI*+TA. (+0 indicates no selective nucleotides, +A indicates selective nucleotide = A, etc). The molecular weight size range of the fingerprints is 45–500 nucleotides.

procedure results in predominant amplification of those restriction fragments, which have a rare cutter sequence on one end and a frequent cutter sequence on the other end (this will be explained below, see also Fig. 3). The rationale for using two restriction enzymes is the following. (i) The frequent cutter will generate small DNA fragments, which will amplify well and are in the optimal size range for separation on denaturing gels (sequence gels). (ii) The number of fragments to be amplified is reduced by using the rare cutter, since only the rare cutter/frequent cutter fragments are amplified. This limits the number of selective nucleotides needed for selective amplification. (iii) The use of two restriction enzymes makes it possible to label one strand of the ds PCR products, which prevents the occurrence of 'doublets' on the gels due to unequal mobility of the two strands of the amplified fragments. (iv) Using two different restriction enzymes gives the greatest flexibility in 'tuning' the number of fragments to be amplified. (v) Large numbers of different fingerprints can be generated by the various combinations of a low number of primers.

AFLP fingerprinting of simple genomes

The AFLP technique makes use of ds adapters ligated to the ends of the restriction fragments to create target sites for primer

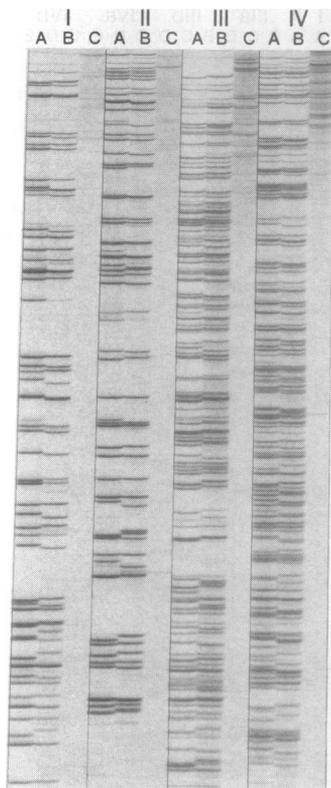


Figure 3. Illustration of the principle of preferential amplification of *EcoRI*–*MseI* fragments. AFLP fingerprints are shown of yeast DNA using primer combinations with a single selective base for the *EcoRI* primers and two selective bases for the *MseI* primer. Panels I, II, III and IV refer to the primer combinations *EcoRI*+C/*MseI*+AC, *EcoRI*+C/*MseI*+CA, *EcoRI*+T/*MseI*+AG and *EcoRI*+T/*MseI*+AT, respectively. Lanes A and B show standard AFLP fingerprints with either the *EcoRI* primers labeled (lanes A) or the *MseI* primers labeled (lanes B). In lanes C the *MseI* primers were labeled but the *EcoRI* primers were omitted from the AFLP reaction. The molecular weight size range of the fingerprints is from 80 to 500 nucleotides.

annealing in fragment amplification. A subset of the restriction fragments is specifically amplified by the use of selective nucleotides at the 3' ends of the AFLP primers. This principle was demonstrated by fingerprinting of four different simple genomic DNAs, varying in genome size from 48.5 to 16 000 kb. The following genomic DNAs were selected and fingerprinted using the enzyme combination *EcoRI*/*MseI*: phage λ DNA (48 451 bp) (29), AcNPV DNA (*Autographa californica* Nuclear Polyhedrosis Virus, 129 981 bp) (30), *Acinetobacter* DNA (estimated genome size 3000 kb) (31) and yeast DNA (estimated genome size 16 000 kb) (32) (Fig. 2).

For the DNAs of phage λ and AcNPV, the complete nucleotide sequence is known, and therefore all *EcoRI*/*MseI* fragments could be exactly predicted. Indeed, all predicted fragments were detected. Also for *Acinetobacter* and yeast DNA, the number of fragments corresponded well with the genome sizes of these organisms. Adding selective nucleotides to the AFLP primers reduced the number of bands ~4-fold with each additional selective base. Furthermore, the addition of an extra selective nucleotide always resulted in a fingerprint, which was a subset of the original fingerprint. This indicates that the selective nucleotides are an accurate and efficient way to select a specific set of restriction

fragments for amplification. The AFLP fingerprints showed that large numbers of restriction fragments were amplified simultaneously, and that in principle the number of bands detected is limited only by the power of the detection system, i.e. polyacrylamide gels. In general, the simultaneous amplification of DNA fragments using specific primer sets for each PCR fragment (multiplex PCR) appears to be rather troublesome (33–35). Our results suggest that multi-fragment amplification is efficient provided that all fragments use the same primer set for their amplification. This implies that the differences in amplification efficiency of DNA fragments in PCR (33–35), are mainly primer-associated and not fragment specific.

The ds adapters used for ligation to the restriction fragments are not phosphorylated, which causes only one strand to be ligated to the ends of the restriction fragments. Fragments were therefore not amplified if the template DNAs were denatured prior to PCR amplification (results not shown), except when *Taq* polymerase and dNTPs were present before denaturation. The filling-in of the 3' recessed ends by the *Taq* polymerase following the dissociation of the non-ligated strands during the heating step seems a matter of only seconds or less. Alternatively, the *Taq* polymerase may immediately displace the non-ligated strands at low temperatures in the process of assembling the reaction mixtures.

These results demonstrated that: (i) the AFLP technique provides an efficient way to amplify large numbers of fragments simultaneously, (ii) the amplified fragments are restriction fragments, (iii) the number of fragments obtained increased as the genome size increased and corresponded well with what was theoretically expected, (iv) the selective nucleotides at the ends of the AFLP primers reduced the number of bands precisely as would be expected.

Restriction of the genomic DNA with *EcoRI* and *MseI* will result in three classes of restriction fragments, *MseI*–*MseI* fragments, *EcoRI*–*MseI* fragments and *EcoRI*–*EcoRI* fragments. The vast majority (>90%) are expected to be *MseI*–*MseI* fragments, the *EcoRI*–*MseI* fragments will be about twice the number of *EcoRI* restriction sites and a small number of the fragments will be *EcoRI*–*EcoRI* fragments. In the previous experiments the *EcoRI* primer was labeled and, therefore, only the restriction fragments with an *EcoRI* site could be detected. To investigate the amplification of the *MseI*–*MseI* fragments a set of AFLP reactions on yeast DNA was performed with the *MseI* primer labeled instead of the *EcoRI* primer (Fig. 3): this is expected to show the *EcoRI*–*MseI* fragments as well as the *MseI*–*MseI* fragments.

Similar yeast fingerprints were obtained with either the *EcoRI* or *MseI* primer labeled (Fig. 3, compare lanes A and B). However, most of the bands showed a significant shift in mobility, due to the fact that the other strand of the fragments was detected. It was surprising to find that almost no additional fragments, i.e. *MseI*–*MseI* fragments, were detected upon labeling of the *MseI* primer in stead of the *EcoRI* primer. Therefore, additional reactions were performed to which only the *MseI* primer was added, which will theoretically only allow amplification of *MseI*–*MseI* fragments (Fig. 3, C lanes): indeed, these reactions showed amplification products not observed in the presence of the *EcoRI* primers. These observations imply that amplification of the *MseI*–*MseI* fragments is inefficient in the presence of the *EcoRI* primer, i.e. that there is preferential amplification of *EcoRI*–*MseI* fragments compared with the *MseI*–*MseI* fragments in the AFLP reaction. This may be explained in two ways. (i) The

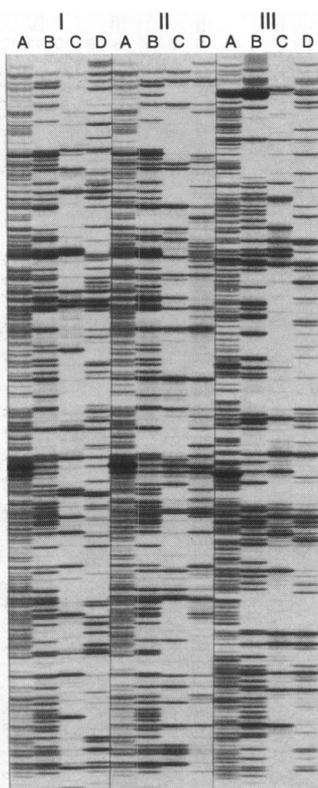


Figure 4. Illustration of the selectivity of AFLP primers. AFLP fingerprints are shown of yeast DNA using primer combinations with one selective base for the *EcoRI* primer and one (lanes A), two (lanes B), three (lanes C) and four (lanes D) selective nucleotides respectively for the *MseI* primer. The primer combinations used are *EcoRI*+A/*MseI*+CTCA (panel I), *EcoRI*+A/*MseI*+CTGC (panel II) and *EcoRI*+T/*MseI*+CTCA (panel III). The molecular weight size range of the fingerprints is from 40 to 370 nucleotides.

MseI primer has a lower annealing temperature than the *EcoRI* primer, making amplification of *MseI*–*MseI* fragments less efficient compared with *EcoRI*–*MseI* fragments under the conditions used. We have indeed observed stronger amplification of *MseI*–*MseI* fragments using longer or alternative *MseI* primers and adapters. (ii) The *MseI*–*MseI* fragments have an inverted repeat at the ends, due to the fact that they are amplified by a single primer. Therefore, a stem–loop structure may be formed by base–pairing of the ends of the fragments, which will compete with primer annealing. This is also confirmed by the observation that only larger fragments were amplified in AFLP reactions when only the *MseI* primer was added. Formation of a stem–loop structure will be more difficult in these larger fragments. We have indeed demonstrated that amplification of large fragments (>1 kb) with a single primer is quite efficient (results not shown).

Careful primer design is crucial for successful PCR amplification. AFLP primers consist of three parts: the 5' part corresponding to the adapter, the restriction site sequence and the 3' selective nucleotides. Therefore, the design of AFLP primers is mainly determined by the design of the adapters, which are ligated to the restriction fragments. Various adapter designs were tested (results not shown), all with good results provided that the general rules for 'good PCR primer design' were followed (36,37). A different adapter design will demand different PCR

conditions, and it is therefore important to select a specific design for the adapters, and to optimize the conditions for this design. An important feature of the AFLP primers is that all primers start with a 5' guanine (G) residue. We have found that a 5' G-residue in the unlabeled primer is crucial to prevent the phenomenon of double bands. The double bands appeared to result from incomplete addition of an extra nucleotide to the synthesized strands (results not shown). This terminal transferase activity of the *Taq* polymerase is quite strong, and has been frequently reported (38–40). We have also found that the 3' nucleotide additions were influenced by the concentration of dNTPs, and that double bands occur at low concentrations of dNTPs regardless of the 5'-residue in the PCR primers (results not shown). We conclude from our results that this terminal transferase activity is most strong (almost 100% of the synthesized strands) when the 3' nucleotide of the synthesized strand is a cytidine (C).

In the experiments described so far, AFLP primers were used with one or two selective 3' nucleotides. This low number of selective nucleotides was shown to provide an effective way to select the desired number of fragments for amplification. Next, primers with longer 3' extensions were tested to determine the maximum number of 3' selective nucleotides which would retain high selectivity in AFLP reactions. For this purpose, fingerprints were generated of yeast DNA using primers with up to four selective nucleotides. A single *EcoRI* primer was selected with one selective nucleotide and combined with four different *MseI* primers with one, two, three or four selective nucleotides, respectively. The *MseI* selective extensions were chosen in a way that with each additional selective nucleotide a fingerprint would be generated that is the subset of the preceding fingerprints, e.g. extensions +C, +CT, +CTC, +CTCA. The appearance of bands that do not occur in the preceding fingerprints is an indication that fragments are amplified which do not correspond to the sequence of the selective bases, and consequently that selectivity is incomplete (Fig. 4).

The fingerprints shown in Figure 4 and the fingerprints from previous experiments demonstrated that primer selectivity is good for primers with one or two selective nucleotides. Selectivity is still acceptable with primers having three selective nucleotides, but it is lost with the addition of the fourth nucleotide. The loss of selectivity with primers having four selective bases is illustrated by the amplification of numerous bands not detected in the corresponding fingerprints with primers having three selective bases (compare lanes D with lanes C). This indicates the tolerance of mismatches in the amplification of the fragments using AFLP primers with four-base extensions. It is most likely that mismatches will be tolerated at the first selective base, because this nucleotide is positioned most distant from the 3' end, and because the selectivity of three-base selective extensions was still adequate. Other researchers have also investigated the selectivity of the 3' nucleotides of PCR primers (41,42). They have found that mismatches were tolerated at both the 3' ultimate and penultimate nucleotides of PCR primers, which is in conflict with our findings. In other experiments, we have found that primer selectivity is relative and also depends on the number of fragments amplified in a single reaction, the PCR conditions and the primer design (results not shown). We feel that some level of mismatch amplification will always occur, but that the reaction conditions can prevent these mismatch bands to reach the

detection level. This, presumably, is the major difference with previously reported experiments (41,42).

AFLP fingerprinting of complex genomes

Initial experiments with AFLP fingerprinting of a number of plant and animal DNAs indicated that AFLP primers with at least three selective nucleotides at both the *EcoRI* and *MseI* primer were required to generate useful band patterns. Because primers with three selective bases tolerate a low level of mismatch amplification, a two-step amplification strategy was developed for AFLP fingerprinting of complex DNAs. In the first step, named preamplification, the genomic DNAs were amplified with AFLP primers both having a single selective nucleotide. Next the PCR products of the preamplification reaction were diluted and used as template for the second AFLP reaction using primers both having three selective nucleotides. We have compared this amplification strategy with a direct amplification of complex genomic DNAs without the use of the preamplification step. The two-step amplification strategy resulted in two important differences compared with the direct AFLP amplification: (i) background 'smears' in the fingerprint patterns were reduced, and (ii) fingerprints with particular primer combinations lacked one or more bands compared with fingerprints generated without preamplification. This is best explained assuming that the direct amplification with AFLP primers having three selective nucleotides resulted in a low level of mismatch amplification products, which caused the background smears and gave discrete amplified fragments corresponding to repeated restriction fragments. An additional advantage of the two-step amplification strategy is that it provides a virtually unlimited amount of template DNA for AFLP reactions. Figure 5 shows a number of typical AFLP fingerprints obtained with the two-step amplification strategy using DNAs of three plant species, *Arabidopsis thaliana*, tomato (*Lycopersicon esculentum*) and maize (*Zea mays*), and human DNA. For *Arabidopsis* DNA primer combinations with a total of five selective nucleotides were used, because of the small genome of this plant species (145 Mb, 43). For tomato, maize and human DNA, six selective nucleotides were used, three selective bases for both the *EcoRI* and *MseI*-primer.

DNA fingerprinting methods based on amplification of genomic DNA fragments by random primers have been found to be quite susceptible to the template DNA concentration (13,14). DNA quantities may vary considerably between individual samples isolated by standard DNA isolation procedures. Preferably, a DNA fingerprinting technique should be insensitive to variations in DNA template concentration. Therefore, the sensitivity of the AFLP technique for the template DNA concentrations was investigated.

AFLP fingerprints were performed using tomato DNA and the enzyme combination *EcoRI/MseI*. Template DNA quantities were varied from 2.5 pg to 25 ng. The standard two-step amplification protocol was used, with the exception of an extended preamplification step for the 2.5 and 25 pg DNA template samples. In tomato 2.5 pg of template DNA corresponds to approximately four molecules of each DNA fragment at the start of the AFLP reaction. AFLP fingerprint patterns were very similar using template quantities that ranged 1000-fold, i.e. from 25 ng to 25 pg (Fig. 6). Fingerprints generated with only 2.5 pg of template DNA were similar to the other fingerprints, although bands varied significantly in intensity and some bands were absent.

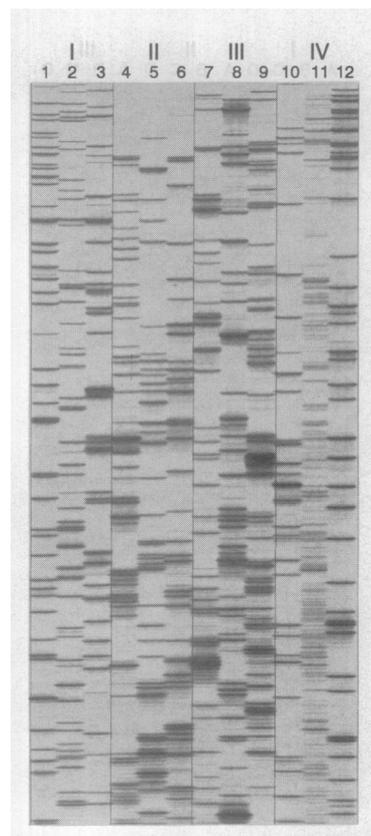


Figure 5. AFLP fingerprints of plant DNAs and human DNAs. Each panel shows three *EcoRI-MseI* fingerprints using three different primer combinations. Human DNA fingerprints are displayed in panel IV, plant fingerprints are displayed in panels I (Arabidopsis), II (tomato) and III (maize), respectively. Primer combinations are from left to right: 1. *EcoRI+CA/MseI+CTT*, 2. *EcoRI+CA/MseI+CAT*, 3. *EcoRI+CA/MseI+CTC*, 4. *EcoRI+ACC/MseI+CTT*, 5. *EcoRI+ACC/MseI+CTC*, 6. *EcoRI+ACC/MseI+CTA*, 7. *EcoRI+ACC/MseI+CAT*, 8. *EcoRI+AGG/MseI+CTT*, 9. *EcoRI+AGG/MseI+CAA*, 10. *EcoRI+CAC/MseI+CGA*, 11. *EcoRI+CAC/MseI+CAA*, 12. *EcoRI+CAG/MseI+CGA*. The molecular weight size range of the fingerprints is from 45 to 500 nucleotides.

These results demonstrate that the AFLP procedure is insensitive to the template DNA concentration, although aberrant fingerprints will be observed at very high template dilutions giving only a few template molecules at the start of the reaction. Most probably the individual restriction fragments are not randomly distributed at such low DNA concentrations explaining the observed differences in band intensities. This hypothesis is supported by our finding that comparison of a number of individual AFLP fingerprints obtained with only 2.5 pg of template DNA showed high variation in the intensity of the individual bands (results not shown).

A remarkable characteristic of the AFLP reaction is that generally the labeled primer is completely consumed (the unlabeled primer is in excess), and that therefore the amplification reaction stops when the labeled primer is exhausted. We have also found that further thermo cycling does not affect the band patterns once the labeled primer is consumed (results not shown). This characteristic is elegantly utilized in the AFLP protocol, which uses an excess of PCR cycles, which will result in fingerprints of equal intensity despite of variations in template concentration.

We have also noted that DNA fingerprints tend to be unreliable if the template concentration was below a certain absolute concentration (~1 pg), regardless of the complexity of the DNA

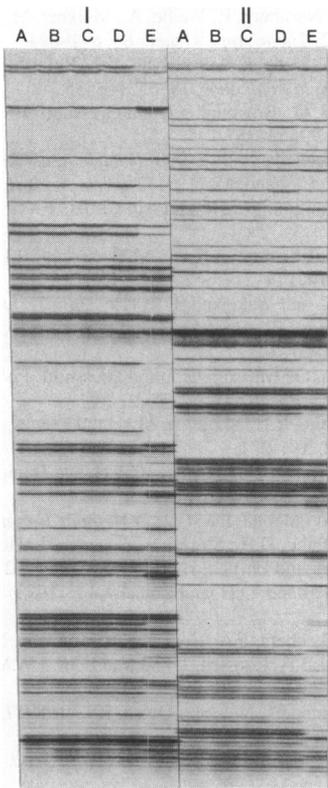


Figure 6. Illustration of the effect of template DNA concentration on AFLP fingerprinting. *EcoRI*–*MseI* fingerprints of tomato DNA are shown using two primer combinations *EcoRI*+*ACC/MseI*+*CAT* (panel I) and *EcoRI*+*ACC/MseI*+*CTT* (panel II). Five different DNA concentrations were used as AFLP templates: 25 ng (lanes A), 2.5 ng (lanes B), 250 pg (lanes C), 25 pg (lanes D) and 2.5 pg (lanes E). The molecular weight size range of the fingerprints is 40–370 nucleotides.

(results not shown). In the latter case, bands were observed which were template independent and which were also present if no DNA was added.

The use of other restriction enzyme combinations for AFLP fingerprinting of complex genomes was also investigated (results not shown). These enzyme combinations included *EcoRI*, *HindIII*, *PstI*, *BglIII*, *XbaI* and *Sse8387I* (eight base cutter) in combination with either *MseI* or *TaqI*. Fingerprints were generated on a variety of plant and animal DNAs. The use of *TaqI* as a four-base cutter in stead of *MseI* resulted in an unequal distribution of the amplified fragments, which were mainly present in the upper part of the gel. Most eukaryotic DNAs are AT-rich, and as a result *MseI* (recognition sequence TTAA) will generally produce much smaller restriction fragments compared with *TaqI* (recognition sequence TCGA). *MseI* is therefore preferred for AFLP fingerprinting because it cuts very frequently in most eukaryotic genomes yielding fragments that are in the optimal size range for both PCR amplification and separation on denaturing polyacrylamide gels. Other rare cutter enzymes generally performed equal to *EcoRI* in AFLP fingerprinting, and the number of bands obtained reflected the cleavage frequency of the various restriction enzymes. However, *EcoRI* is preferred because it is a reliable (low cost) six-cutter enzyme, which limits problems associated with partial restriction in AFLP fingerprinting (see below).

Incomplete restriction of the DNA will cause problems in AFLP fingerprinting, because partial fragments will be generated, which will be detected by the AFLP procedure. When various DNA samples are compared with AFLP fingerprinting, incomplete restriction will result in the detection of differences in band patterns, which do not reflect true DNA polymorphisms, i.e. when one sample is partially restricted and the others are not. Incomplete restriction will only become apparent when different DNA samples of the same organism are compared. It is characterised by the presence of additional bands in the lanes, predominantly of higher molecular weight.

CONCLUSIONS

AFLP is a DNA fingerprinting technique that detects genomic restriction fragments and resembles in that respect the RFLP (restriction fragment length polymorphism) technique, with the major difference that PCR amplification instead of Southern hybridisation is used for detection of restriction fragments. The resemblance with the RFLP technique was the basis to chose the name AFLP. The name AFLP, however, should not be used as an acronym, because the technique will display presence or absence of restriction fragments rather than length differences.

In our initial AFLP protocols an additional purification step was incorporated in the template preparation (28). Biotinylated *EcoRI*-adapters were used for ligation, and subsequently the biotinylated DNA fragments (*EcoRI*–*MseI* fragments and *EcoRI*–*EcoRI* fragments) were subtracted from the ligation mixture using streptavidin beads (28,44). This step reduced the complexity of the DNA template by removing all *MseI*–*MseI* fragments, which proved to be important for generating high quality fingerprints of complex DNAs. The protocol described in this paper omits this purification step, because amplification conditions are used which result in preferential amplification of *EcoRI*–*MseI* fragments with respect to the *MseI*–*MseI* fragments (Fig. 3). This is the result of careful adapter and primer design and inherent characteristics of the AFLP amplification reaction.

The AFLP technique will generate fingerprints of any DNA regardless of the origin or complexity, and in this paper we have presented AFLP fingerprints of DNAs differing in genome size as much as 100 000-fold. We have described that the number of amplified fragments may be controlled by the cleavage frequency of the rare cutter enzyme and the number of selective bases. In addition the number of amplified bands may be controlled by the nature of the selective bases; selective extensions with rare di- or trinucleotides will result in the reduction of amplified fragments. In general, there is an almost linear correlation between numbers of amplified fragments and genome size. This linear correlation is lost in the complex genomes of higher plants, which contain high numbers of repeated sequences and, hence, multicopy restriction fragments. Fingerprints of these complex DNAs consist predominantly of unique AFLP fragments, but are characterised by the presence of small numbers of more intense repeated fragments.

In complex genomes the number of restriction fragments that may be detected by the AFLP technique is virtually unlimited. A single enzyme combination (a combination of a specific six-base and four-base restriction enzyme) will already permit the amplification of 100 000s of unique AFLP fragments, of which generally 50–100 are selected for each AFLP reaction. Most AFLP fragments correspond to unique positions on the genome, and, hence, can be exploited as landmarks in genetic and physical

maps, each fragment being characterized by its size and its primers required for amplification. In addition, the AFLP technique permits detection of restriction fragments in any background or complexity, including pooled DNA samples and cloned (and pooled) DNA segments. Therefore, the AFLP technique is not simply a fingerprinting technique, it is an enabling technology in genome research, because it can bridge the gap between genetic and physical maps. First, the AFLP technique is a very effective tool to reveal restriction fragment polymorphisms. These fragment polymorphisms, i.e. AFLP markers, can be used to construct high density genetic maps of genomes or genome segments. In most organisms AFLP will prove to be the most effective way to construct genetic DNA marker maps compared to other existing marker technologies. Secondly, the AFLP markers can be used to detect corresponding genomic clones, e.g. yeast artificial chromosomes (YACs). This is most effectively achieved by working with libraries, which are pooled to allow for rapid PCR screening and subsequent clone identification. An AFLP marker will detect a single corresponding YAC clone in pools of as much as 100 YAC clones (M. Zabeau, M. Kuiper and P. Vos., manuscript in preparation). Finally, the AFLP technique may be used for fingerprinting of cloned DNA segments like cosmids, P1 clones, bacterial artificial chromosomes (BACs) or YACs (results not shown). By simply using no or few selective nucleotides, restriction fragment fingerprints will be produced, which subsequently can be used to line up individual clones and make contigs.

ACKNOWLEDGEMENTS

We would like to thank Richard Michelmore, Hanneke Witsenboer and Francesco Salamini for critically reading the manuscript. This work was financed by Keygene N.V. and six Dutch breeding companies: Royal Sluis, Cebeco Zaden, ENZA Zaden, De Ruiter Seeds, Rijkzwaan and Ropta ZPC. The work was subsidized by the Dutch ministry of economic affairs through a PBTS grant.

Note: The AFLP technique is covered by patents and patent applications owned by Keygene N.V. Information concerning licenses to practice the AFLP process for commercial purposes can be obtained from Keygene N.V. Research kits for AFLP fingerprinting of plant genomic DNA are available from Life Technologies (Gaithersburg, MD, USA) and Perkin Elmer (Applied Biosystems Division, Foster City, CA, USA).

REFERENCES

- 1 Jeffrey, A.J., Wilson, V. and Thein, S.L. (1985) *Nature*, **314**, 67–73.
- 2 Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. and White, R. (1987) *Science*, **235**, 1616–1622.
- 3 Jeffrey, A.J., Mcleod, A., Tamaki, K., Neil, D.L. and Monckton, D.G. (1991) *Nature*, **354**, 204–209.
- 4 Tautz, D. (1989) *Nucleic Acids Res.*, **17**, 6463–6472.
- 5 Weber, J.L. and May, P.E. (1989) *Am. J. Hum. Genet.*, **44**, 388–396.
- 6 Edwards, A., Civitello, A., Hammond, H.A. and Caskey, C.T. (1991) *Am. J. Hum. Genet.*, **49**, 746–756.
- 7 Beyermann, B., Numberg, P., Weihe, A., Meixner, M., Epplen, J.T. and Börner, T. (1992) *Theor. Appl. Genet.*, **83**, 691–694.
- 8 Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. (1990) *Nucleic Acids Res.*, **18**, 6531–6535.
- 9 Caetano-Anolles, G., Brassam, B.J. and Gresshoff, P.M. (1991) *BioTechnology*, **9**, 553–557.
- 10 Welsh, J. and McClelland, M. (1990) *Nucleic Acids Res.*, **18**, 7213–7218.
- 11 Welsh, J. and McClelland, M. (1991) *Nucleic Acids Res.*, **19**, 861–866.
- 12 Riedy, M.F., Hamilton, W.J.III. and Aquadro, C.F. (1992) *Nucleic Acids Res.*, **20**, 918.
- 13 Ellsworth, D.L., Rittenhouse, K.D. and Honeycutt, R.L. (1993) *BioTechniques*, **14**, 214–217.
- 14 Muralidharan, K. and Wakeland, E.K. (1993) *BioTechniques*, **14**, 362–364.
- 15 Micheli, M.R., Bova, R., Pascale, E. and D'Ambrosio, E. (1994) *Nucleic Acids Res.*, **22**, 1921–1922.
- 16 Caetano-Anolles, G., Brassam, B.J. and Gresshoff, P.M. (1992) *Mol. Gen. Genet.*, **235**, 157–165.
- 17 Botstein, D., White, R.L., Skolnick, M.H. and Davies, R.W. (1980) *Am. J. Hum. Genet.*, **32**, 314–331.
- 18 Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989) *BioTechnology*, **7**, 257–264.
- 19 Mullis, K.B. and Faloona, F.A. (1987) *Methods Enzymol.*, **155**, 335–350.
- 20 Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Ehrlich, H.A. (1988) *Science*, **239**, 487–491.
- 21 Ehrlich, H.A., Gelfand, D.H. and Sninsky, J.J. (1991) *Science*, **252**, 1643–1650.
- 22 Vlak, J.M., Klinkenberg, F.A., Zaal, K.J.M., Usmany, M., Klinge-Roode, E.C., Geervliet, J.B.F., Roosien, J. and Van Lent, J.W.M. (1988) *J. Gen. Virol.*, **69**, 756–776.
- 23 Pitcher, D.G., Saunders, N.A. and Owen, R.J. (1989) *Letts. Appl. Microbiol.*, **8**, 151–156.
- 24 Green, E.D. and Olson, M.V. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1213–1217.
- 25 Stewart, C.J., Jr and Via L.E. (1993) *BioTechniques*, **14**, 748–750.
- 26 Miller, S.A., Dykes, D.D. and Polesky, H.F. (1988) *Nucleic Acids Res.*, **16**, 1215.
- 27 Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
- 28 Zabeau, M. and Vos, P. (1993) *European Patent Application*, publication no: EP 0534858.
- 29 Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. and Petersen, G.B. (1982) *J. Mol. Biol.*, **162**, 729–773.
- 30 Ayres, M.D., Howard, S.C., Kuzio, J., Lopez-Ferber, M. and Possee, R.D. (1994) *Virology*, **202**, 585–605.
- 31 Bouvet, P.J.M. and Grimont, P.A.D. (1986) *Int. J. Syst. Bacteriol.*, **36**, 228–240.
- 32 Carle, G.F. and Olson, M.V. (1984) *Nucleic Acids Res.*, **12**, 5647–5664.
- 33 Chamberlain, J.S., Gibs, R.A., Ranier, J.E., Nguyen, P.N. and Caskey, C.T. (1988) *Nucleic Acids Res.*, **16**, 11 141–11 156.
- 34 Rossiter, B.J.F., Fuscoe, J.C., Muzny, D.M., Fox, M. and Caskey, C.T. (1991) *Genomics*, **9**, 247–256.
- 35 Ferrie, R.M., Schwarz, M.J., Robertson, N.H., Vaudin, S., Super, M., Malone, G. and Little, S. (1992) *Am. J. Hum. Genet.*, **51**, 251–262.
- 36 Dieffenbach, C.W., Lowe, T.M.J. and Dveksler, G.S. (1993) *PCR Methods Applic.*, **3**, S30–S37.
- 37 Breslauer, K.J., Ronald, F., Blocker, H. and Marky, L.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3746–3750.
- 38 Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfand, D.H. (1989) *J. Biol. Chem.*, **264**, 6247–6437.
- 39 Clark, J.M. (1988) *Nucleic Acids Res.*, **16**, 9677–9686.
- 40 Hu, G. (1993) *DNA Cell Biol.*, **12**, 763–770.
- 41 Kwok, S., Kellogg, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Sninsky, J.J. (1990) *Nucleic Acids Res.*, **18**, 999–1005.
- 42 Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C. and Markham, A.F. (1989) *Nucleic Acids Res.*, **17**, 2503–2516.
- 43 Arumuganathan, K. and Earle, E.D. (1991) *Plant Mol. Biol. Reporter*, **9**, 208–218.
- 44 Uhlen, M. (1989) *Nature*, **340**, 733–744.