TE-AFLP: combining rapidity and robustness in DNA fingerprinting

A. W. G. van der Wurff*, Y. L. Chan¹, N. M. van Straalen and J. Schouten¹

Institute of Ecological Science, Faculty of Biology, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands and ¹MRC Holland, Hudsonstraat 68, 1057 SN Amsterdam, The Netherlands

Received August 29, 2000; Revised and Accepted October 30, 2000

ABSTRACT

A new type of fingerprinting technique is presented, based on amplified fragment length polymorphism (AFLP). Rather than two endonucleases as in AFLP, we propose the use of three enzymes, hence the method is called three endonuclease (TE)-AFLP. Genomic DNA is digested and two sets of adapters are selectively ligated onto the restriction fragments in a single reaction volume. No adapters complementary to the ends generated by a frequent cutter are added. Due to the addition of a third endonuclease, the TE-AFLP method provides a high discriminatory power and a reduction in the number of bands. The latter makes it especially suitable for the analysis of complex genomes. TE-AFLP fingerprints are suitable for detection by automatic fluorescent sequencers and are obtained in less than half the time and at reduced costs compared to a typical AFLP. The reliability of this method was investigated by determining the influence of varying digestion, ligation and PCR components on the fingerprint. Moreover, cross-experiments to study inheritance of loci were performed with a primitive insect and with tomato strains. The features of TE-AFLP are discussed in comparison with conventional AFLP.

INTRODUCTION

Genomic fingerprints are increasingly used to study relationships at the intra- or even interspecific level. The fingerprints are obtained by visualising many parts of the genome. Differences in these fingerprints between individuals are interpreted as genetic distances. Obviously, the differences should reflect variations in DNA rather than artefacts due to a non-robust method. Furthermore, the method should provide the appropriate level of discriminatory power and it should be relatively rapid and cheap, especially in large-scale population genetics studies (see for example 1,2).

Up to now the most widely used PCR-based multilocus DNA fingerprint techniques are random amplified polymorphic DNA (RAPD) (3) and the more recently proposed amplified fragment length polymorphism (AFLP) (4). RAPD is generally regarded to be less time consuming, while the AFLP method is more robust (5,6). The technique proposed here combines both advantages.

Three endonuclease (TE)-AFLP is less time consuming compared to the authentic AFLP protocol because it allows a significant reduction in preparation steps. Furthermore, in contrast to RAPD (7), TE-AFLP is relatively insensitive to small fluctuations in reaction conditions.

The original AFLP technique involved digestion of the DNA by a pair of endonucleases, followed by ligation of short oligomers, i.e. adapters, to the ends of the fragments. A subset of the fragments is amplified by PCR, using primers that include the adapter sequence. The reaction products are separated by gel electrophoresis and visualised by a suitable detection method. To prevent amplification of too many fragments, the primers are generally extended at the 3′-terminus. The extension selects certain nucleotides in the restriction fragment immediately flanking the endonuclease recognition site. In theory, each selective base added to one of the primers reduces the number of amplified fragments 4-fold.

TE-AFLP differs from traditional AFLP by reducing the number of amplified fragments not only by primer extension, but also by selective ligation. Three endonucleases together with only two sets of adapters are added to a single reaction. As a consequence, the reduced number of potential amplifiable fragments diminishes competition during PCR, permitting stringent reaction conditions and thus eliminating the need for a two-step amplification in fingerprinting complex genomes.

Additionally, the third endonuclease provides discrimination at extra sites between genomes compared to typical AFLP, thereby increasing discriminatory power.

To test the reliability of the technique we studied: (i) the effect of different concentrations of digestion–ligation components and PCR reaction conditions on the fingerprint; (ii) the inheritance of fragments in crossing experiments with springtails and tomato strains. Finally, we discuss features of TE-AFLP in comparison with typical AFLP.

MATERIALS AND METHODS

DNA extraction

The technique was tested using the soil dwelling primitive insect Orchesella cincta (Collembola) and tomato plants (Lycopersicum sp.). Springtails (Collembola) were obtained from forests in The Netherlands (Roggebotszand, Lage Vuursche and Spanderswoud) and Italy (Siena).

*To whom correspondence should be addressed. Tel: +31 20 444 7073; Fax: +31 20 444 7123; Email: wurff@bio.vu.nl
 Approximately 250 ng DNA was extracted from each O. cincta individual using a modified CTAB protocol (8). The springtails were washed in 70% ethanol, ground with a sterilised glass rod and incubated in 500 µl of extraction buffer (100 mM Tris–HCl, pH 8.0, 1.4 M NaCl, 20 mM EDTA, pH 8.0, 2% w/v CTAB) and 2 µl of β-mercaptoethanol at 60°C for 1.5 h. After three extractions with 1 vol chloroform/isooamyl alcohol (24:1), DNA was precipitated at 4°C for 1 h by addition of 1/3 vol isopropanol. After centrifugation for 20 min at 4°C the precipitate was washed with 70% ethanol and vacuum dried. Finally, the precipitate was dissolved in 0.1× TE buffer and checked on a 1.2% (w/v) agarose gel containing ethidium bromide (0.5 µg/ml) using 0.5X TBE as electrophoresis buffer. DNA concentration was estimated by comparison with a molecular weight marker (white; MRC Holland).

Tomato strain DNA was obtained from the Department of Genetics, Vrije Universiteit, Amsterdam.

**TE-AFLP**

For fingerprints of *O. cincta*, *Lycopersicum esculentum* and *Lycopersicum penelli*, 20 ng DNA was added to a digestion–ligation mix in a final volume of 20 µl containing digestion–ligation buffer (10 mM Tris–HCl, pH 7.6, 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 0.5 mM ATP), 4 pmol of both *BamH*I and *BamHI* adapters, 0.50 Weiss units T4 DNA ligase and 6 U *BsaI*, 1.25 U *BamHI* and 1 U *RsaI* and incubated for 1.5 h at 30°C. Following digestion–ligation, 0.5 µl was used as template in a 50 µl PCR reaction containing 10 pmol of the labelled *Bsa*I-adapter primer and the unlabelled *BamHI*-adapter primer, PCR buffer (15 mM Tris–HCl, pH 8.5, 50 mM KCl, 1.5 mM MgCl₂, 0.01% Triton X-100), 200 µM each of dNTP (Boehringer Mannheim) and 2.5 U *Taq* polymerase. PCR was performed on a thermal cycler (PTC-100; MJ Research or Perkin Elmer Cetus) using an initial 2.5 min denaturation at 95°C followed by 10 cycles of 30 s denaturation at 95°C, 30 s annealing at 70°C, 60 s elongation at 72°C and 40 cycles of 30 s denaturation at 95°C, 30 s annealing at 60°C and 60 s elongation at 72°C. A final 20 min at 72°C allowed for maximal 3′ nucleotide addition owing to terminal transferase activity of the *Taq* polymerase (4). For analysis, 2.4 µl of the PCR product was mixed with 6.4 µl of formamide containing 5 mg/ml blue dextran and 2.4 µl of sterile double-distilled water and denatured for 5 min at 80°C. Of this sample, 8 µl was resolved on a polyacrylamide gel with 0.6X TBE electrophoresis buffer. For detection, an automatic AFLF sequencer with FITC-labelled primers or AFLF express II with Cy5-labelled primers (Amersham Pharmacia) was used according to the manufacturer’s instructions. ReadyMix or High Resolution (Amersham Pharmacia) gels were analysed with the aid of Fragment Manager 1.0, Sequence Analyses v.2.1 and Image 1D elite (Amersham Pharmacia). TE-AFLP fingerprints can be resolved on agarose gels. All enzymes were obtained from MRC Holland.

**Construction of oligonucleotides**

About 10 different combinations of endonucleases, primers and adapters were tested (unpublished results). The adapters and primers were selected based on number of bands, even distribution of bands along the gel, minimal background signal and highest level of polymorphism among five springtails from a single forest (see Fig. 2). The sequences of the *BamH*I adapters were 5′-ACgAAgTCcCCgCCAgCAa and 5′-gATcTTTg-CTggCcgg. The *BsaI* adapters were 5′-ACgTTgCggCcgg-CgAgA and 5′-CTAgTCTCgCcggCCgC. The unlabelled *BamHI*-C primer sequence was 5′-gTTTCgCcggCCAgCAAgACTCC and the labelled *BamHI*-CC primer was 5′-ggCgTCgAgACTAgACC (Fig. 1).

The annealing temperature in the PCR TE-AFLP protocol was optimised for the unlabelled primer. The labelled primer has a lower and the adapter oligomers have the lowest annealing temperature. Thus, amplification of fragments containing two labelled primers is severely hampered. To avoid double banding on the gel, the labelled primer was constructed with a 5′-terminal guanine to allow maximal 3′ nucleotide addition by the terminal transferase activity of the *Taq* polymerase (4). *Taq* polymerase was chosen because of its high discrimination ability in mismatching of 3′ primer termini (9) and low cost. All oligonucleotides were from Isogen or Amersham Pharmacia.

**RESULTS AND DISCUSSION**

To investigate the robustness of TE-AFLP, the influence of the digestion–ligation buffer components and reaction conditions on the banding pattern was studied.

No change in fingerprint was observed when the amount of tomato DNA in the digestion–ligation mix was varied from 1 to 1000 ng or when the reaction time was reduced from 90 to 30 min or increased to 4 h. Variations in incubation temperature between 25 and 37°C did not affect the fingerprint. Finally, the pattern remained identical when any of the reaction components, i.e. enzymes, DNA oligomers, ATP, Tris–HCl, NaCl, MgCl₂ and DTT, were increased or decreased by a factor of two.

Only an excess of magnesium and adapters in the digestion–ligation mix caused unreproducible fingerprints. The gel pattern was disrupted when >1.75 mM MgCl₂ and >8 pmol each adapter were present in the PCR reaction. In addition, disturbance of the fingerprint pattern occurred when adapters with an annealing temperature higher than those of the PCR primers were used. The intensity of the amplified fragments from 25 to 700 bp increased when the amount of the template digestion–ligation mix, *Taq* polymerase, PCR buffer and primers in the PCR was raised. Low intensity fragments were not detected when using exceptionally low concentrations of
labelled or unlabelled primer (<2 pmol). This is in contrast to typical AFLP, where an increase from 5 to 30 ng labelled primer results in a substantial increase in the number of amplified fragments detected by an automatic sequencer (10).

When the concentration of labelled primer was <20 pmol/50 µl reaction volume fragments >500 bp showed a decrease in signal intensity relative to the shorter fragments. Relatively low signal intensities of longer fragments are frequently reported (11, 12).

No change in pattern was revealed with an increase or decrease in annealing temperature by 3°C, with a change in the number of PCR cycles by 10, with magnesium concentrations ranging from 1.3 to 1.6 mM or by increasing KCl concentration in the PCR buffer by 20 mM. No hot start was necessary since there was no effect of immediate mixing of components on ice or at room temperature either immediately or 1 h before the start of the PCR.

The inheritance of parental markers by 16 *O.cincta* offspring and four tomato hybrids was complete, i.e. all markers present in the offspring were observed in at least one of the parents (Fig. 3). Furthermore, as suggested by Herbergs et al. (13), automatic sequencer detection allows the identification of signal intensities as co-dominant markers, i.e. heterozygotes and homozygotes.

In TE-AFLP the addition of more than three selective nucleotides at the 3’-terminus of the primer to reduce the number of amplified fragments is not necessary. This is crucial, since non-selective amplification owing to mismatches in the case of a >3 nt 3’ extension under standard reaction conditions has been reported (4). Although long primer extensions can give reproducible results, additional investigation is necessary so as to optimise reaction conditions (11).

In addition, the combination of digestion and respective ligation together with only one PCR program results in a substantial reduction in the time necessary to prepare fingerprints.

The use of a second rare cutter in TE-AFLP reduces the number of potentially amplifiable fragments compared to traditional AFLP. The expected number of fragments which are amplified with regard to the number of endonucleases is conveniently illustrated as follows. Suppose that a DNA genome is digested with three different restriction enzymes, i.e. A (a frequent cutter) and B and C (both rare cutters), having restriction sites a, b and c. For circular DNA with a random sequence order the total number of fragments obtained is equal to the number of restriction sites. Three endonucleases will generate fragments with six different restriction ends (AA, BB, CC, AB, AC, BB and BC). The expected frequencies of these fragments are, respectively: \(a^2/(a+b+c)^2\), \(b^2/(a+b+c)^2\), \(c^2/(a+b+c)^2\), \(2ab/(a+b+c)^2\), \(2ac/(a+b+c)^2\) and \(2bc/(a+b+c)^2\).

In TE-AFLP the fragments with restriction ends from enzyme A are not amplified. Assuming that only those fragments having two different adapters are efficiently amplified, the frequency of the dominant product is \(2bc/(a+b+c)^2\) and the total number of fragments which can be efficiently amplified is \(2bc\). Assuming that b and c are of the same order and a is 16 times larger, the expected frequency of the dominant

---

**Figure 2.** Example of an autoradiogram of TE-AFLP fingerprints of 37 *O.cincta* specimens (lanes 1–37). Individuals obtained from the Lage Vuursche and Spanderswoud forests (∼6 km apart, The Netherlands) are indicated by white and black squares, respectively. M is a 50 bp marker (numbers represent size in bp); C represents a negative control (i.e. no DNA added).
product is only $2b^2/(18b)^2 = 2/324 = 0.62\%$. In traditional AFLP no A enzyme is present and B usually has of the order of 16 times more restriction sites than C. The expected frequency of the fragments which are amplified efficiently is thus $32c^2/(17c)^2 = 32/289 = 11.1\%$.

Using the TE-AFLP technique the number of fragments which are potentially amplified can be easily reduced 20-fold compared to traditional AFLP. Furthermore, only $\sim 0.62\%$ of the fragments which have two different adapters have to compete during PCR with 0.62% of the fragments which have two identical adapters. This allows high annealing temperatures during PCR, which results in stringent amplification and a low background signal.

Both AFLP and TE-AFLP are designed such that the available amount of primers rather than the number of PCR cycles is the limiting factor in the amplification process. This allows an excessive number of amplification cycles and results in a minimal influence on the fingerprint by the amount of template DNA. Furthermore, the amplification-limiting amount of primer prevents competition and the appearance of artefactual bands in the final amplification cycles. In TE-AFLP artefacts arise in negative controls in which no DNA is added; however, these artefactual bands are not encountered in positive controls.

In fingerprinting honeybee DNA (*Apis mellifera* L.) (14) combined digestion and ligation with a single PCR profile was used with one endonuclease (*EcoRI*). However, care is needed with this technique, since small fragments with two complementary ends may form stem–loop structures preventing primer annealing during PCR (4). Furthermore, our results suggest that long fragments are less efficiently amplified. This restricts the use of AFLP with a single endonuclease to studies which include a careful method of optimisation concerning organisms with low genomic complexity. This is in contrast to TE-AFLP, where the embodiment of three restriction endonucleases and selective ligation offers great potential for fingerprinting, especially with large genome sizes.

The utilisation of three endonucleases under standard AFLP conditions was explored in 1997 (15). Three sets of adapters were added; however, only two PCR primers were used in amplifying the templates. Although the authors mentioned in their Materials and Methods that the modification gave improved results, the potential features of using three restriction enzymes were not discussed.

In general, the choice of a particular number of bands generated in a fingerprint should be a compromise between reliability, ease of analyses, statistical power and chance of revealing polymorphisms. In the case of a large number of fragments adjacent bands and non-homologues are hard to identify, making patterns difficult to score, whereas a low number of sparsely distributed fragments allows discrimination between bands, even on an agarose gel. On the other hand, the generation of a large number of bands increases the chance of revealing polymorphic loci and increases the power of statistical tests.

In conclusion, the choice of a particular technique is largely dependent on the hypothesis formulated (16,17) and the complexity or amount of genomic variation of the organism under scrutiny.

![Figure 3. Inheritance of TE-AFLP fragments in the springtail O.cincta. Each horizontal lane represents an individual fingerprint with peaks ranging from 50 to 300 bp. Lanes numbers 16 and 17 represent a male from Siena (Italy) and a female from Roggebotszand (The Netherlands), respectively. Lanes 1–15 represent offspring.](image-url)
ACKNOWLEDGEMENTS

G. Ernsting is acknowledged for commenting on the manuscript. The TE-AFLP technique is covered by patents and patent applications owned by Keygene N.V. Additional information is available on http://www.MRC-Holland.com.

REFERENCES