Research Journal of Applied Sciences 1 (1-4): 136-140, 2006


# Dna Fingerprinting-Review Paper

R. Blundell
Department of Physiology and Biochemistry, Biomedical Science Building,
Faculty of Medicine and Surgery, University of Malta, Msida MSD06, Malta

Before the Polymerase Chain Reaction (PCR) was established, DNA fingerprinting technology has relied for years on Restriction Fragment Length Polymorphism (RFLP) and Variable Number of Tandom Repeats (VNTR) analysis, a very efficient technique but quite laborious and not suitable for high throughput mapping. Since its, development, PCR has provided a new and powerful tool for DNA fingerprinting.

PCR is a amplification process based on a specialised polymerase enzymes, which can synthesise a complementary strand to a given DNA strand in a mixture the four DNA bases and two DNA oligonucleotides (primers) flanking the target sequence. The mixture is heated to separate the strands of double-stranded DNA containing the target sequence and then cooled to allow (1) the primers to find and bind to their complementary sequences on the separated strands and (2) the polymerase to extend the primers to new complementary strands. Repeated heating and cooling cycles multiply the target DNA exponentially. Thus in about 20 PCR cycles, the DNA fragment can be amplified by a millionfold[1].

According to the earlier PCR based approach, fingerprints are generated through the selective amplification of hypervariable loci such as mini- or microsatellites. More recently, a novel PCR-based strategy involving the use of arbitrary primers to amplify random genomic DNA fragments has been developed.

**Arbitrary primed PCR:** A significant advance in DNA fingerprinting technology was made in the early 1990s through the development of a novel Polymerase Chain Reaction (PCR)-based strategy which involves the use of single oligonucleotide primers of arbitrary sequence to amplify random genomic DNA fragments[2-4]. This strategy, often referred to as arbitrary primed PCR, allows detection of polymorphisms between individual (or strains) as difference between the patterns of DNA fragments amplified from the different DNAs using a given primer.

The random amplification strategy is shared by three methods called Arbitrarily Primed PCR (AP PCR)[2]; Random Amplified Polymorphic DNA RAPD[3] and DNA Amplification Fingerprinting DAF[4]. They differ from one another in the length of primers used, amplifcation conditions, separation and visualisation of amplified DNA fragments; consequently they generate markedly different fingerprint patterns varying from quite simple to complex.

The acronym MAAP (multiple arbitrary amplicon profiling) was proposed to collectively define these techniques: the multiple, arbitrary nature of targeted sites and the amplifcation of a range of characteristic DNA products. This acronym would also be appropriated in view of the fact that this methodology is often used to place markers in a genetic map[5].

The increasing popularity of random amplification technology is due to the substantial advantages it provides over previously available DNA fingerprinting techniques. Mainly the ability to easily and rapidly generate a number of polymorphic markers using a very small amounts of starting DNA, independently of any prior knowledge of the target DNA sequence. This feature makes random amplification technology a tool potentially useful in many areas of genetic research and particularly well suited to high throughput applications.

**Random amplification of polymorphic DNA:** While standard PCR reactions are based on the use of two specific oligonucleotides that selectively prime the amplification of the target DNA sequence, arbitrary primed PCR reactions involve the use of a single oligonucleotide of arbitrary sequence which primes the amplification of several discrete DNA products. Each of these anonymous but reproducible fragments is derived from a region of the genome that contains, on opposite DNA strands, two primer binding sites located within an amplifiable distance of each other.

It is generally assumed, that due to the low stringency condition ensured by an appropriate annealing temperature, the arbitrary primer binds to a number of sites randomly distributed in the genomic DNA template and primes DNA synthesis even in those studies in which the match between primer and template is imperfect. The amplification process is initiated on those genomic regions in distance not exceeding a few thousand

nucleotides. The outcome of the amplification reaction is determined by a competition in which those products representing the most efficient pairs of priming sites separated by the most easily amplifiable sequences (i.e., shorter sequences with little secondary structure) will prevail[6].

The assumption that not all amplifications are the result of perfect matching between primer and template is supported by the bacterial genomes. The number of fragments amplified from those genomes is higher than expected by statistics, if only perfect matching between primer and template is supposed to occur. This can only be explained through the assumption that imperfect matching is also involved[3].

A model to explain the amplification of DNA with arbitrary primers was proposed by Caetano-Anolles *et al.*[7]. The model is based on the competitive effects of primer-template and template-template interactions established predominantly during the first few cycles of the amplification process. In the first few temperature cycles a group of cognate amplicons is selected for amplification during a temperature screening phase driven by primer-template-enzyme interactions that can accommodate primer-template mismatching events. First-round amplification products are initially single-stranded, but have palindromic termini that can establish template-template interactions forming hairpin-loops and duplexes. The primer will have to recognise and displace these structures to allow enzyme anchoring and primer extension. In subsequent rounds of amplification, the different stages of the reaction tend to establish equilibrium, while the rare primer-template duplexes are enzymatically transformed into accumulating amplification products.

Reaction products are resolved by gel electrophoresis and patterns are generated from different genomic DNAs (e.g., DNAs from different individuals or strains) using the same primer are easily compared to identify possible differences (i.e. DNA fragments that are amplified form one genomic DNA but not from another). Any difference between the patterns of amplified fragments reveals a polymorphism in that it arises from a difference in template sequence that inhibits primer binding or otherwise interferes with amplification of the corresponding fragment.

Although DNA amplification using arbitrary primers, as originally developed, involves a single primer, it has been shown that pairwise combinations of primers can be used as well. Amplification performed using two arbitrary primers generate reproducible patterns that are different from those obtained with each single primer[8,9]. Welsh and McClelland[8] have also shown that more than 50% of the products generated by pairwise combinations of primers are different from those generated by either primer alone, that is the amount of products generated by both primers is higher than expected on a statistical basis.

An explanation for this observation could be the possibility that ssDNA products containing a primer at one end and its complement at the other end form a panhandle structure and are, therefore, at a disadvantage in competing for amplification compared to ssDNA products that do not form panhandle structures[10].

The use of pairwise combinations of primers provides an advantage over the amplification by single primers. It allows generation of a much higher number of fingerprints when a given number of primers is used both individually and in all possible pairwise combinations.

**Application of random amplified polymorphic DNA:** Random amplification technology is an efficient tool to quickly and easily screen a very large number of loci for possible DNA polymorphisms which are usually referred to as random amplified polymorphic DNAs (RAPDs). These polymorphisms have been proved able to discriminate between closely related individuals. This feature combined with their easy identification, makes RAPDs a valid type of marker, potentially useful in many areas of genetic research such as gene mapping, marker-assisted selection in breeding tasks and individual or strain identification. The dominant nature of RAPDs may be a drawback in some applications. Fingerprint patterns are scored for the presence or absence of a specific band where the presence of the band does not allow discrimination between homozygosis and heterozygosis at the corresponding locus[11].

Furthermore, RAPDs provide a powerful tool in population genetics as well as in phylogenetic analysis, though in the latter their utility is limited to closely related organisms[12].

Very interesting results have been obtained by applying random amplification to compare genomic DNAs from a tumor with that of normal tissues from the same individual[13-15]. These results demonstrate the ability of the random amplification assay to detect tumor specific genetic alterations. This technique has a particular promising future in experimental oncology.

**Alternative approaches:** Since the establishment of the basic methodology of DNA amplification by arbitrary primers, several variations aimed at further developing its potential have been proposed. These alternative approaches, by improving either primer design or amplification strategy, allow the researcher to tailor fingerprints. In this way peculiar requirements of specific

areas of application can be met, such as those pertaining to the complexity of amplification pattern, the level of polymorphisms detected and the nature of amplification products.

Strategies based on the improvement of primer design involve the use of the so-called mini-hairpin primers[16,17], which are primers containing hairpin-turn structure at their 5' end, or primers with some degenerate bases[18]. Furthermore, primers whose sequence has been biased to recognise particular sequence motifs, for example microsatellite repeats[8,10,19], promoter consensus sequences or gene family motifs[20], have also been used. Useful comments on the application of primers encoding particular sequence motifs can be found in McClelland and Welsh[10].

The improvement of amplification strategy is the basis for both Arbitrary Signatures for Amplification Profiles (ASAP) analysis[17] and Template Endonuclease Cleavage MAAP (tecMAAP)[21]. The former is a dual-step amplification strategy in which fingerprints are generated by re-amplifcation of a previous fingerprinting profile, while the latter involves endonuclease cleavage of the template prior to its amplification with arbitrary primers.

More recently further alternative strategies AFLP, RAHM, RAMPO and DS-PCR have been proposed. Vos *et al.*[22] have described a technique called AFLP which produces highly complex profiles by arbitrary amplification of subsets of restriction fragments ligated to adapter cassettes. Amplification is primed by tripartite oligonucleotides that consist of a sequence complementary to the adapter (at their 5' end), followed by another that is complementary to the restriction site and a very short arbitrary sequence (at their 3' end) which provides amplification selectively. Only restriction fragments in which the nucleotides flanking the restriction site match the 3' end of the primer will be amplified. The high reliability of this technique is due to the high stringency condition used for primer annealing.

A method which combines random DNA amplification with hybridisation to microsatellite-complementary oligonucleotide probes, named Random Amplified Hybridisation Microsatellite (RAHM)[23] and Random Amplified Microsatellite Polymorphisms (RAMPO) has been proposed[24]. The procedure initially involves genomic DNA amplification with an arbitrary primer. After electrophoretic separation and staining, the amplification products are transferred to a nylon membrane and hybridised to oligonucleotide probes carrying Simple Sequence Repeats (SSR). Fingerprint profiles not corresponding to the staining patterns and completely different for each probe are thus obtained. By enhancing the level of polymorphism detection, this approach can be usefully applied in genetic analysis of species where little or no intraspecific variation is detected by random amplifcation alone.

Double stringency PCR DS-PCR[25] is a technique that combines microsatellite-specific priming with arbitrary priming. During the first amplification cycle, a single microsatellite oligonucleotide primes the amplification of DNA between two microsatellite regions. In subsequent cycles, the annealing temperature (and the consequent stringency) is lowered in order to allow the arbitrary oligonucleotide to prime amplification within some of the previously amplified fragments. In this study, therefore, random amplification is targeted to highly polymorphic genomic regions.

**High throughput mapping systems:** The complete sequence of some genomes is known but to define the role of each gene in these genomes is a different and a more difficult task. The use of high throughput mapping systems is contributing to the knowledge of functional genomics. Recently devised high throughput-mapping systems are Serial Analysis of Gene Expression (SAGE), DNA microarrays, biological chips and PCR workstations.

**Serial Analysis of Gene Expression (SAGE):** The method of SAGE was first described by Velculescu[26] and the method is based on two principles. First, a short nucleotide sequence tag (9 to 10 base pairs bp) contains sufficient information to uniquely identify a transcript, provided it is isolated from a defined position within the transcript. Second, concatenation of short sequence tags allows the efficient analysis of transcripts in a serial manner by the sequencing of multiple tags within a single clone. Adaptation of this technique to an automated sequencer would allow the analysis of over 1000 transcripts in a single 3 h run.

**DNA microarrays:** A high-capacity system was developed to monitor the expression of many genes in parallel[27]. Microarrays prepared by high-speed robotics printing of complementary DNAs on glass were used for quantitative expression of the corresponding genes. Due to the small format and high density of the arrays, hybridisation volumes of 2 μL, was used to detect rare transcripts in probe mixtures derived from 2 μg of total cellular messenger RNA.

Schena[27] measured differential expression measurements of 45 *Arabidopsis* genes by means of simultaneous, two-colour fluorescent hybridisation. The procedure as a simple and economical way to explore gene expression patterns on a genomic scale. They amplified more than 6000 genes in about 4 months from the yeast

genome (*Saccharomyces cerevisiae*) and in only 2 days they printed a set of 110 microarrays of 6400 elements each.

**Biological chips:** Foder[28] described a method which uses light to direct the combinatorial chemical synthesis of biopolymers on a solid support. The identity and location of each biopolymer will be known and thus its interaction with a molecular binding agent can be measured.

An exciting application of oligonucleotide arrays is in DNA sequencing. Sequencing by Hybridisation (SBH) employs a chip containing a set of short oligonucleotides probes to generate hybridisation patterns from the complementary sequences that contain a longer target strand of DNA[29].

Foder[30] combined the technologies of light-directed combinatorial chemical synthesis with the laser confocal fluorescence scanning, thus obtained a multiplex assay format with an array resolution of $10^5$ synthesis sites per $cm^2$. An array of 65,536 8-mers is needed to sequence DNA up to 300 bases long and an array of 67,108,864 13-mers is needed for sequencing DNA of a few thousand bases long.

Parinov[31] suggested the use of short oligonucleotides (8-mers) and extending the length of only those that are hybridised to the DNA. Extension was carried out by the Contiguous Stacking Hybridisation (CSH) without any additional oligonucleotide synthesis. Computer simulations showed that hybridisation with 5-mers makes an 8-mer chip as efficient as a 13-mer chip and capable of sequencing a few thousand bases of DNA rather than 200 bases.

**PCR workstations:** Intelligent Automation Systems, incorporated, (IAS) designed a massive parallel factory-style automation system nicknamed the Genomatron[32]. The first station assembles PCR reactions in custom-fabrication 1536-well microtiter "cards" and seals the wells by welding a thin plastic film across the cards. The second station thermocycles the reactions by transporting the cards over three chambers that force temperature-controlled water to flow uniformly between the cards. The third station transfers the reactions from one microtiter card onto a hybridisation membrane affixed to the bottom of a second microtiter card by piercing the first card with a bed of 1536 hypodermic needles and sucking the reactions downward. These filter cards were then manually hybridised with a chemiluminescent probe and read by the Charge-Coupled Device (CCD) camera. The stations were computer controlled and the microtiter cards were assigned a bar code to facilitate sample tracking. Each station was designed to process 96 microtiter cards, providing a throughput of nearly 150,000 reactions per run.

## REFERENCES


1. Mullis, K., F. Faloona and S. Scharf, 1986. Specific enzymatic amplification of DNA *in vitro*: The Polymerase Chain Reaction. Cold Spring Harbour Symp. Quant. Biology, 51: 263-273.
2. Welsh, J. and M. McClelland, 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res., 24: 7213-7218.
3. Williams, J., G.K. Kubelik, R. Anne, Livak, J. Kenneth, Rafalski, J. Antoni, Tingey and V. Scott, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 22: 6531-6535.
4. Caetano-Anolles, G., G.J. Bassam and P.M. Greshoff, 1991. DNA amplification fingerprinting using very short arbitrary primers. Bio. Tech., 9: 553-557.
5. Caetano-Anolles, G., B .J. Bassam and P.M. Gresshoff, 1992b. DNA fingerprinting: MAAPing out a RAPD redefinition? Biotechnology, pp: 10-937.
6. McClelland, M., F. Mathieu-Daude and J. Welsh, 1995. RNA fingerprinting and differential display using arbitrary primed PCR. Trends Genet., 11: 242-246.
7. Caetano-Anolles, G., B.J. Bassam and P.M. Gresshoff, 1992a. Primer-template interactions during DNA amplification fingerprinting with single arbitrary oligonucleotide. Mol. Gen. Gent., 235: 157-165.
8. Welsh, J. and M. McClelland, 1991. Genetic fingerprinting using arbitrarily primed PCR and a matrix of pairwise combinations of primers. Nucleic Acids Res., 24: 7213-7218.
9. Micheli, M.R., R. Bova, P. Calissano and E. D'Ambrosio, 1993. Randomly amplified polymorphic DNA fingerprinting using combinations of oligonucleotide primers. Bio. Techiques, 15: 388-390.
10. McClelland, M., H. Arensdorf, R. Cheng and J. Welsh, 1994. Arbitrarily primed PCR fingerprints resolved on SSCP gels. Nucleic Acids Res., 22: 1770-1771.
11. Williams, J.G.K., M.K. Hanafey, J.A. Rafalski and S.V. Tingey, 1993. Genetic analysis using random amplified polymorphic DNA markers. Methods Enzymol., 218: 704-740.
12. Micheli, M.R. and R. Bova, 1997. Fingerprinting methods based on arbitrarily primed PCR. Springer Lab Manual.

13. Peinado, M.A., S. Malkhosyan and A. Velazques and M. Perucho, 1992. Isolation and characterisation of allelic losses and gains in colorectal tumours by arbitrarily primed polymerase chain reaction. Proc. Natl. Acad. Sci. USA, 89: 10065-10069.
14. Ionov, Y., M.A. Peinado, S. Malkhosyan, D. Shibata and M. Perucho, 1993. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature, 363: 558-561.
15. Kohno, T., K. Moroshita, H. Takano, D.N. Shapiro and J. Yokota, 1994. Homozygous deletion at chomosome 2q33 in human small-cell lung carcinoma identified by arbitrarily primed PCR genomic fingerprinting. Oncogene, 9: 103-108.
16. Caetano-Anolles, G. and P.M. Gresshoff, 1994. DNA amplification fingerprinting using arbitrary mini-hairpin oligonucleotide primers. Bio. Tech., 12: 619-623.
17. Caetano-Anolles, G. and P.M. Gresshoff, 1996. Generation of sequence signatures from DNA amplification fingerprints with mini-hairpin and microsatellite primers. Bio. Techniqes, 20: 6-1044.
18. Sakallah, S.A., R.W. Lanning and D.L. Cooper, 1995. DNA fingerprinting of crude bacterial lysates using degenerate RAPD primers. PCR Methods Applic., 4: 265-268.
19. Welsh, J., M. McClelland, 1992. PCR-amplified length polymorphisms in tRNA intergenic spacers for categorising staphlyococci. Mol. Microbio., 6: 1673-1680.
20. Birkenmeiier, E.H., U. Schneider and S.J. Thurston, 1992. Fingerprinting genomes by use of PCR primers that encode protein motifs or contain sequences that regulate gene expression. Mammalian Genome, 2: 537-545.
21. Caetano-Anolles, G., B.J. Bassam and P.M. Gresshoff, 1993. Enhanced detection of polymorphic DNA by multiple arbitrary amplicon profiling of endonuclease digested DNA: Identification of markers linked to the supernodulation locus of soybean. Mol. Gen. Genet., 241: 57-64.
22. Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. can de Lees, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper and M. Zabeau, 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acid Res., 23: 4407-4414.
23. Cifarelli, R.A., M. Gallitelli and F. Cellini, 1995. Random amplified hybridisation microsatellites (RAHM): Isolation of a new class of microsatellite-containing DNA clones. Nucleic acids Res., 23: 3802-3803.
24. Richardson, T., S. Cato, J. Ramser, G. Kahl and K. Weising, 1995. Hybridization of microsatellites to RAPD: A new source of polymorphic markers. Nucleic acid Res., 23: 18-3798.
25. Matioli, S.R. and R. de Brito, 1995. Obtaining genetic markers by using double-stringency PCR with microsatellites and arbitrary primers. Bio. Technique, 19: 752-756.
26. Velculescu, V.E., L. Zhang, B. Vogelstein and K.W. Kinzler, 1995. Serial analysis of gene expression. Sci., 270: 484-486.
27. Schena, M., D. Shalon, R.W. Davis and P.O. Brown, 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarrays. Sci., 270: 467-468.
28. Fodor, S.P.A., J.L. Read, M.C. Pirrung, L. Stryer, A.T. Lu and D. Solas, 1991. Light-directed, spatially addressable parallel chemical synthesis. Sci., 251: 767-773.
29. Southern, E.M., U. Maskos and J.K. Elder, 1992. Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: Evaluation using experimental models. Genomics, 13: 1008-1017.
30. Fodor, S.P.A., R.P. Rava, X.C. Huang, A.C. Pease, C.P. Holmes and C.L. Adams, 1993. Multiplexed biochemical assays with biological chips. Nature, 364: 555-556.
31. Parinov, S., V. Barsky, G. Yerskov, E. Kirillov, E. Timofeev, A. Belgovskiy and A. Mirzabekov, 1996. DNA sequencing by hybridization to microchip octa- and decanucleotides extends by stacked pentanucleotides. Nucleic aid Res., 15: 2998-3004.
32. Anderson, L., 1995. Entering the Postgenome Era. Sci., 270: 368-369.