

Different approaches to molecular scanning of point mutations in genetic diseases

Maurizio Ferrari, Paola Carrera and Laura Cremonesi

IRCCS. H.S. Raffaele, Clinical Molecular Biology Laboratory and DIBIT, Milan, Italy.

Abstract - Several advanced techniques to detect mutations have been developed during the last years. Scanning procedures can be classified as "physical" and "cleavage" methodologies. Physical methods include Denaturing Gradient Gel Electrophoresis, Single-Strand Conformation Polymorphism and Heteroduplex Analysis while cleavage methods include RNase A Cleavage, Chemical Mismatch Cleavage and Enzyme Mismatch Cleavage. A related technique, Protein Truncation Test, has been recently introduced.

INTRODUCTION

Sequence variations among individuals that may be associated with diseases or serve as genetic markers, can be investigated in a number of ways. Most techniques for mutation analysis exploit DNA heteroduplex formation. These make use of wild-type sequences to generate mismatched heteroduplexes with the putative mutated sequence. Mismatched heteroduplexes can be detected by either cleavage or physical methods. The major obstacle using conventional methodologies in clinical field is that they require handling of radioactivity, consequently many modifications of traditional techniques have been published to avoid use of radioactivity.

RNase A cleavage.

Among cleavage methods, RNase A (1,2) was the first described. Like many other scanning methods for mutation detection, this procedure relies on the principle of heteroduplex mismatch analysis. In the RNase A cleavage method, RNA-DNA heteroduplex between radiolabelled wild-type riboprobe and a mutant DNA, obtained by PCR amplification (3-6), is enzymatically cleaved by RNase A, by exploiting the ability of RNase A to cleave single-stranded RNA at the points of mismatches in RNA:DNA hybrids. This is followed by electrophoresis and autoradiography. The presence and location of a mutation are indicated by a cleavage product of a given size. This method can only detect approximately 50% of mismatches and for this reason it has been replaced by more efficient procedures.

Chemical Mismatch Cleavage (CMC)

In the Chemical Mismatch Cleavage (7-10) a radiolabeled DNA wild type sequence (probe) is hybridized to another amplified sequence containing the putative alteration to form a heteroduplex. A chemical modification, followed by piperidine cleavage, is used to remove the mismatch bubble in the heteroduplex. Gel electrophoresis of the denatured heteroduplex and autoradiography allow to visualize the cleavage product. Osmium tetroxide is used for the modification of mispaired thymidines and hydroxylamine for mismatched cytosines. Additionally, labelling the antisense strand of the probe DNA allows the detection of adenosine and guanosine mismatches.

The chemical cleavage of mismatch offers the unique advantage, over other mutation detection techniques, in that it can be used to detect almost 100% of mutations in long DNA fragments, being limited only by gel resolution and PCR efficiency, and that it is not very sequence context

dependent. Moreover, this method provides the precise characterization and the exact location of the mutation within the tested fragment. The major drawback of this technique is that it involves numerous manipulations using toxic and noxious chemicals. Recently, the method has been improved to make CMC more suitable for automation by using fluorescent primers also enabling multiplexing and thereby reducing the number of manipulations. Alternatively, fluorescently labelled dUTPs incorporated via PCR allow the internal labelling of both target and probe DNA strands and therefore labelling of each possible hybrid, doubling the chances of mutation detection and virtually guaranteeing 100% detection.

Emerging cleavage methods use enzymes which replace the toxic chemicals and result in a one step method. They use the T4 endonuclease VII resolvase and mismatch repair enzymes to cleave mismatched DNA.

Enzyme mismatch cleavage with T4 endonuclease VII resolvase.

The resolvases are enzymes that resolve branched DNA intermediates that form during genetic recombination. T4 endonuclease VII is the product of gene 49 of the bacteriophage T4 which catalyzes the resolution of branched DNA, Holliday structures, cruciforms and loops. It cleaves 3' and within 6 nt from the point of DNA distortion. Thus, when mutant and wild-type homoduplexes differing by a single nucleotide are melted and hybridized, they generate two heteroduplex species containing two pairs of mismatched bases. The mutation would be detected if any one strand containing one of the four mismatched bases were cleaved (11, 12).

Mismatch repair enzymes cleavage (MREC)

The mismatch repair enzymes cleavage assay (13) is an efficient method to identify single base substitutions. It relies on nicking enzyme systems specific for mismatch-containing DNA. The sequence of interest is amplified by PCR and homo- and heteroduplexes species may be generated at the end of the PCR, by denaturing and allowing to reanneal the amplified products. These hybrids are treated with mismatch repair enzymes and then analysed by denaturing gel electrophoresis. The MREC assay makes use of three mismatch repair enzymes. The MutY endonuclease removes adenines from the mismatches and is useful to detect both $A/T \leftrightarrow C/G$ transversions and $G/C \leftrightarrow A/T$ transitions. Mammalian thymine glycosylase removes thymines from T/G, T/C, and T/T mismatches and is useful to detect $G/C \leftrightarrow A/T$ transitions as well as $A/T \leftrightarrow C/G$ and $T/A \leftrightarrow A/T$ transversions. The all-type endonuclease or topoisomerase I from human or calf thymus can recognize all eight mismatches and can be used to scan any nucleotide substitution. The major advantages of this method are high sensitivity, reliability and precise localization of the mutation site. Moreover, MREC can use specific labels which can be incorporated into both DNA strands, thus allowing all four possible nucleotide substitutions in a give site to be identified. Due to its high sensitivity (it allows to detect a point mutation in a mutant DNA sample contaminated by 99% wild-type DNA), MREC could be an alternate to direct DNA sequencing.

Denaturing gradient gel electrophoresis (DGGE)

Denaturing gradient gel electrophoresis is a useful technique to separate and identify sequence variants. DGGE, either in presence of chemical or thermal gradients, represents one of the most advanced methods for detection of point mutations (14-16). DGGE is typically performed in constant-concentration polyacrylamide gel slabs, cast in presence of linearly increasing amounts of a denaturing agent (usually formamide and urea, cathode to anode). A variant of DGGE employs temperature gradients along the migration path and is known as TGGE. Separation by DGGE or TGGE is based on the fact that the electrophoretic mobility in a gel of a partially melted DNA molecule is greatly reduced as compared to an unmelted molecule. As the DNA fragment proceeds through the gradient gel, it will reach a position where the concentration of the denaturing agent equals the melting temperature (T_m) of its lowest melting domain causing denaturation, partial unwinding and consequent marked retardation due to frictional drag.

Visualization of the bands is achieved by ethidium bromide staining of the gel at the end of the run. In order to prevent the complete unwinding of the molecule, the duplex DNA is made to contain a high-melting domain, consisting in a stretch of 40 GC pairs, named GC-clamp (17), located at the 5' end of one of the two PCR primers. In order to identify all mutants, including those with a migration pattern almost indistinguishable from the wild-type, the wild-type and (presumptive) mutant chains are fully melted and re-annealed while mixed in solution, at the end of the PCR reaction. Four types of double helices will be formed: two homo-duplexes (Wt/Wt and M/M, for wild type and mutant, respectively) and two heteroduplexes (Wt/M and M/Wt). These four types of duplexes will have different melting temperatures (T_m) and thus branching and retardation of their mobility will occur at different positions along the gel. In case of individuals heterozygous for a point mutation, the expected four band pattern is in general obtained. DGGE using GC-clamped primers offers perhaps the most valid approach for identification of single base changes in PCR-amplified material.

The T_m and the entire melting profile of the DNA molecule of known sequence, together with calculations of the expected changes in electrophoretic mobility in gels under denaturing conditions, can be predicted with accuracy via the computational simulation program developed by Lerman and Silverstein.

Among the advantages of this methodology are: (a) a high detection sensitivity (ca. 99%); (b) improved heterozygote detection (a pattern of two homo- and two hetero-duplexes is obtained); (c) the possibility of optimizing the analysis by computer simulation (based on the Melt87 and SQHTX programs of Lerman) (18); (d) a non-radioactive protocol.

A recent modification of the TGGE technique is based on separation by capillary zone electrophoresis (CZE)(19). At present, this technique represents the most advanced electrophoretic technique, combining extremely high resolution with on-line detection and with sub-attomol sensitivity. Recently, a CZE method able to detect point mutations has been developed by running the PCR-amplified DNA fragments in temporal (as opposed to spatial) thermal gradient parallel to the migration direction. The thermal gradient is generated from "within" the capillary and controlled by a dedicated software previously developed for measuring precisely the temperature inside the separation chamber, by knowing the exact capillary diameter, the buffer conductivity, and the voltage applied. By selecting a buffer of proper conductivity and manipulating the applied voltage during the run, it is possible to create temporal temperature gradients (as opposed to gradients fixed in space, i.e. along the capillary length) with the correct slope and reaching the desired final value for optimum separation of homo- and hetero-duplexes. Producing denaturing temporal thermal gradients of any desired slope with laser-induced fluorescence detection, could become a formidable tool for large-scale screening of single nucleotide variations in genomic DNA.

Single-strand conformation polymorphism (SSCP)

Single-strand conformation polymorphism (20,21) is a method allowing the resolution of wild-type and mutated sequences based on a change in mobility of separated single-strand in non-denaturing polyacrylamide gel electrophoresis. Electrophoretic mobility depends on both size and shape of a particle, and single-stranded DNA molecules fold back on themselves and generate secondary structures which are determined by intramolecular interactions in a sequence dependent manner. A single nucleotide substitution can alter this secondary structure and, consequently, the electrophoretic mobility of the single strands, resulting in band shifts on autoradiographs.

The target sequences are amplified and labeled by a PCR reaction using either labeled primers or deoxynucleotides. The amplified fragments, after denaturation, are then resolved in polyacrylamide gel electrophoresis and mutations are detected as altered mobility in the autoradiogram with respect to a normal control run in parallel.

Most conformational changes caused by single nucleotide differences in various sequence contexts seem to be detected, although further accumulation of data may be necessary to estimate

accurately the rate of detection of this technique. The ability of a given variation to alter the conformation of the single strands is not predictable on the basis of an adequate theoretical model and base changes occurring in a loop or in a long stable stem of the secondary structure might not be detected by SSCP. Moreover, effects of sequence alterations on mobility in an SSCP gel may be very different between the two complementary strands. Moreover, many additional physical factors such as temperature and ionic strength, can affect the sensitivity of SSCP, and its optimization is highly empirical since sequence variations that have little or no effect on conformation in one set of conditions can dramatically affect mobility under different conditions. Standard SSCP reaches maximal reliability in detecting sequence alterations in fragments of 150-200 bp.

Due to its simplicity over other PCR-based scanning techniques, SSCP analysis has potential use in clinical molecular diagnosis provided that a nonradioactive detection method can be developed. Recently, visualisation of the bands of SSCP analysis has been obtained by silver staining. More advanced protocols, allowing the detection of mutations at a sensitivity equal to that of the radioactively-based SSCP analysis, have been developed. These last use fluorescence-labeled primers in the PCR and analyze the products with a fluorescence-based automated sequencing machine (22). Multi-color fluorescent SSCP also allows to include an internal standard in every lane, which can be used to compare data from each lane with respect to each other, thus effectively eliminating lane to lane variability.

Other variants to increase the detection rate includes a "dideoxy sequencing approach" based on dideoxy fingerprinting (ddF) (23) and restriction endonuclease fingerprinting (REF)(24), which are methods for detecting the presence of virtually all single-base substitutions. ddF is a combination between SSCP and Sanger dideoxy sequencing which involves nondenaturing gel electrophoresis of a Sanger sequencing reaction with one dideoxynucleotide. In this way a 250 bp fragment can be screened with 100% sensitivity. REF is a more complex modification of SSCP allowing the screening of more than 1 kb fragment with essentially 100% sensitivity. For REF, a target sequence is amplified with PCR, digested independently with five to six different restriction endonucleases, mixed, end-labeled, denatured, and analysed by SSCP on a nondenaturing gel. In case six restriction enzymes are used, a sequence variation will be present in six different restriction fragments, thus generating 12 different single-stranded segments. A mobility shift in any one of these fragments is sufficient to pinpoint the presence of a sequence variation. The restriction pattern obtained enables localization of an alteration in the region examined. The larger amount of DNA that can be analysed in each lane of the gel significantly increases the capacity of the method compared with traditional SSCP.

Heteroduplex analysis (HA)

Heteroduplex analysis (25) is a method for detecting single base substitutions in PCR products that is as fast and simple as SSCP. HA can be rapidly performed without radioisotopes or specialized equipment. This method takes advantage of the formation of heteroduplexes between wild-type and mutated sequences by heating and renaturing of PCR products. Due to a more open double-stranded configuration surrounding the mismatched bases, heteroduplexes migrate slower than their corresponding homoduplexes, and are then detected as bands of reduced mobility compared to normal and mutant homoduplexes on polyacrylamide gels.

The ability of a particular single base substitution to be detected by the HA method cannot be predicted merely by knowing the mismatched bases since the adjacent nucleotides have a substantial effect on the configuration of the mismatched region and length-based separation will clearly miss nucleotide substitutions.

Optimization of the temperature, gel crosslinking and percent acrylamide as well as glycerol and sucrose enhance the resolution of mutated samples. The HA method can be rapidly performed without radioisotopes or specialized equipment and, under optimized conditions, it should prove useful for rapidly screening large numbers of individuals for known mutations and

polymorphisms in sequenced genes. When HA is used in combination with SSCP, up to 100% of all alterations in a DNA fragment can be easily detected.

Protein Truncation Test (PTT)

The Protein Truncation Test (26, 27) is based on the *in vitro* coupled transcription and translation of PCR-amplified coding sequences. Genomic DNA or cDNA templates, obtained by reverse-transcription (RT)-PCR of RNA, are PCR amplified using a T7-modified primer containing a T7 promoter sequence and an eukaryotic translation initiation sequence. Translation and transcription of PCR products is subsequently performed in a commercially available rabbit reticulocyte lysate system, in the presence of radiolabeled amino acids. Translation products are analysed by SDS-PAGE electrophoresis and visualized by autoradiography. In the PTT, mutations that are predicted to cause termination of translation will result in the synthesis of truncated proteins. These shorter polypeptides will appear as anomalous bands on the autoradiogram. The size of the truncated protein points out the position of the stopcodon within the segment, thereby facilitating the further characterization of the mutation by DNA sequencing.

Direct Sequencing

DNA sequencing precisely defines the location and nature of the sequence alteration and thus represents the final step of any mutation detection method. The two classical DNA sequencing methods are the chemical Maxam-Gilbert (28) and Sangers enzymatic chain-terminating methods (29). Such protocols are subjected to cloning, maintenance and use of systems dependent on vectors and living host cells. More recently, the possibility to directly sequence PCR products (30-32), both from genomic and DNA-RNA templates, has greatly simplified sequencing without the need of library construction, screening and subcloning. This procedure also facilitates automation for large scale sequencing by using the new fluorescence detection technology (33), thus representing the method of choice for direct detection of mutation. For direct sequencing different approaches to generate single-stranded DNA templates have been developed, all employing the dideoxy termination sequencing protocol. The most used is based on asymmetric PCR to convert the double-stranded PCR product into a single-stranded sequencing template (34) using one primer in excess with respect to the other which is limiting or even absent. In alternative the PCR may be performed with a 5' biotin labelled primer, and streptavidine bound at a solid support is used to capture the biotinilated single stranded DNA (35).

CONCLUSIONS

There is an increasing need to improve current methodologies and develop new approaches for detecting disease gene mutations in human diseases and animal models.

Comparison of the rapidity, accuracy, and informativeness of the different methods for mutation detection indicates that each method has certain advantages but none is ideal. Methods accurately detecting the presence of a sequence change essentially 100% of the time while also providing precise information about the location of the sequence change are often labor intensive or technically challenging, and therefore are not practical for screening large numbers of individuals. Multiple variations of traditional methods have been developed. Many laboratories prefer to combine the use of two faster and easier methods, such as SSCP and HA, each identifying allelic variants on the basis of a different principle, which should complement one another in sensitivity, and can be simultaneously tested on the same PCR product.

Molecular diagnostic is now a clinical reality and the myriad of human mutations identified in the various genetic diseases can now be assayed in the clinical molecular biology diagnostics laboratory. This implies that new technologies suited for clinical use and dedicated computer software need to be implemented. Affordable automation ought to be the next step in raising molecular diagnostic services to such a high level of efficiency that a variety of rapid, accurate, cost effective tests reaching every field of medicine can be implemented.

Thus, the search for reliable and efficient approaches to the detection of known and unknown mutations will continue.

REFERENCES

1. R.M. Myers *et al.* *Science* **230**,1242-1246 (1985).
2. R.A. Gibbs and T. Caskey. *Science* **236**, 303-305 (1987).
3. R.K.Saiki *et al.* *Science* **230**, 1350-1354 (1985).
4. K. Mullis and F. Faloona. *Meth. Enzymol.* **55**, 335-350 (1987).
5. H.P. Fosberg. *Hum.Genet* **83**, 1-15 (1989).
6. K. Mullis *Sci. Am.* **262**, 56-65 (1990).
7. R.G.H. Cotton *et al.* *Proc Natl.Acad.Sci.USA.* **85**, 4397-4401 (1988).
8. M. Grompe *et al.* *Proc.Natl.Acad.Sci.USA.* **86**, 5888-5892 (1989).
9. S.M. Forrest *et al.* *Am.J.Hum.Genet.* **49**, 175-183 (1991).
10. J.A. Saleeba *et al.* *Hum.Mutat.* **1**, 63-69 (1992).
11. R. Youil *et al.* *Am.J.Hum.Genet.* 53 Abs 1257 (1993).
12. R. Youil *et al.* *Proc.Natl.Acad.Sci.USA* **92**,87-91 (1995).
13. A-L. Lu and I-C. Hsu. *Genomics* **14**, 249-255 (1991).
14. L.S. *et al.* *Ann.Rev.Biophys.Bioeng.* **13**, 399-423 (1984).
15. R.M. Myers *et al.* *Nucl.Acids Res.* **13**, 3131-3145 (1985).
16. R.M. Myers *et al.* *Meth.Enzymol.* **155**, 501-527 (1987).
17. V.C. Sheffield *et al.* *Proc.Natl.Acad.Sci.USA.* **86**, 232-236 (1989).
18. L.S.Lerman *et al.* *Meth.Enzymol.* **155**, 482-501 (1987).
19. C. Gelfi *et al.* *Electrophoresis* **15**, 1506-1511 (1994).
20. M. Orita *et al.* *Proc.Natl.Acad.Sci.USA.* **86**, 2766-2770 (1989).
21. M. Orita *et al.* *Genomics.* **5**, 874-879 (1989).
22. R. Makino *et al.* *PCR Methods and Applications* **2**, 10-13 (1992).
23. G. Sarkar *et al.* *Genomics* **13**, 441-443 (1992).
24. Q. Liu and S.S. Sommer. *Bio Techniques* **18**, 470-477 (1995).
25. C.M. Nagamine *et al.* *Am.J.Hum.Genet* **45**, 337-339 (1989).
26. P. Roest *et al.* *Hum.Mol.Genet.* **2**, 1719-1721 (1993).
27. R. Van der LuijT *et al.* *Genomics* **20**, 1-4 (1994).
28. A.M. Maxam and W. Gilbert. *Proc.Natl.Acad.Sci.USA.* **74**, 560-564 (1977).
29. F. Sanger *et al.* *Proc.Natl.Acad.Sci.USA.* **74**, 5463-5467 (1977).
30. M.A. Innis *et al.* *Proc.Natl.Acad.Sci.USA.* **85**, 9436-9440 (1988).
31. T.D. Kocher *PCR Methods and Appl.* **1**, 217-221 (1992).
32. I.S. Bevan *et al.* *PCR Methods and Appl.* **1**, 222-228 (1992).
33. M. Hattori *et al.* *Electrophoresis* **13**, 560-565 (1992).
34. U.B. Gyllensten and H.A. Erlich. *Proc.Natl.Acad.Sci.USA.* **85**, 7652-7656 (1988).
35. A.C. Syvanen *et al.* *FEBS Lett.* **258**, 71-74 (1989).