Chapter 12

Variants and Adaptations of the Standard PCR Protocol

The standard qualitative PCR protocol (as initially developed) involves the exponential amplification of a specific DNA sequence which is achieved by subjecting an optimized PCR mix to a predetermined number of thermocycles (cyclical heating and cooling steps). Each thermocycle within this standard PCR protocol actually comprises three distinct steps, namely: (i) a nucleic acid denaturation step (to disassociate the double-stranded DNA helix into single strands), followed by (ii) a primer annealing step (to allow the specific hybridization of a 15–25 base pair oligonucleotide primer), followed by (iii) a DNA extension/amplification step (to replicate and amplify the specific target nucleic acid sequence). After the given number of thermocycles are complete, and exponential DNA amplification has been performed, the presence or absence of specific target DNA sequences in the final PCR mix (and hence the presence or absence of specific target DNA sequences in the original sample) is established. Over the last few decades, many useful adaptations/variations of this standard qualitative PCR protocol have been developed. Some of the most important of these variants/adaptations are outlined below.

12.1 Generating Labelled PCR Amplimers for PCR Product Visualization, DNA Probes and Cloning

Several standard PCR protocol adaptations have now been developed which allow for the labelling of PCR amplimers as they are being created during PCR amplification. Some of the possible labels that may be incorporated are shown in Fig. 12.1, and include biotin, digoxygenin, many fluorescent dyes or fluorochromes (e.g. FAM, HEX, JOE, etc.; see also Table 5.4), and even (radioactive γ-^{32}P) phosphate groups. The advantage of attaching specific labels to PCR amplimers, is that the amplimers may be directly visualized or used in other "downstream" post-PCR applications such as solid phase hybridization or reverse line blots (see previous chapters).

Frequently, labelled amplimers are generated by adding specially modified nucleotides (to which labels have been chemically attached) to the PCR mix prior to PCR thermocycling. These modified nucleotides are added to the initial PCR mix alongside the normal, unmodified nucleotides (dATP, dCTP, dGTP and

E. Van Pelt-Verkuil et al., *Principles and Technical Aspects of PCR Amplification,*

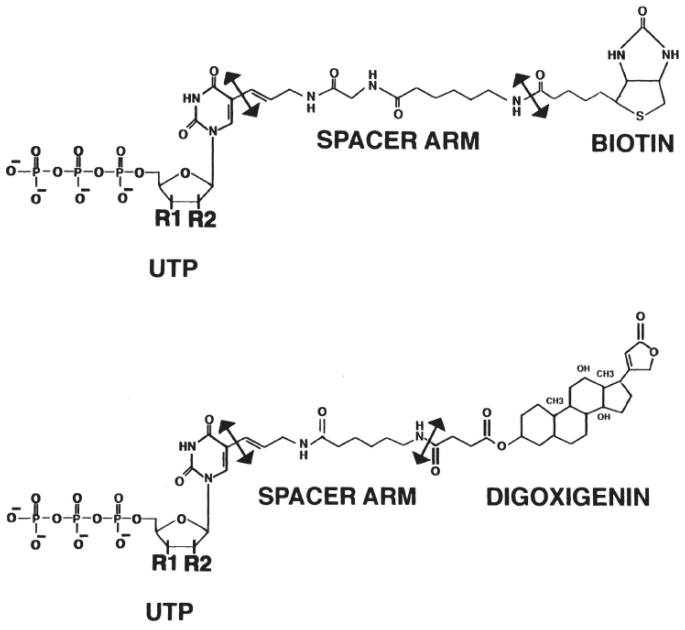

Fig. 12.1 Biotin-11-dUTP and digoxygenin-11-dUTP are two ligands which are frequently used to label PCR amplimers. When included in a PCR mix (along with unlabeled deoxynucleotides), they are incorporated into the growing DNA chain by the thermostable DNA polymerase. The 5′-end position of the pyrimidine nucleotide dUTP (deoxyuracil triphosphate) is most suited as substitution site for the addition of labels. The "11" identifies the number of carbon atoms present in the spacer arm. Radioactive compounds such as α-^{32}P dUTP are becoming increasingly less popular due to safety and disposal problems, though radioactive labelling still represents the most sensitive detection technique currently available. In uridine R1 = OH and R2 = OH, in deoxyuridine R1 = OH and R2 = H and in dideoxyuridine both R1 and R2 are H (Reproduced from Chevalier et al., 1997. With permission from Dennis G. Bashkin, Ph.D.)

dTTP), and become incorporated into the PCR amplimers via the action of DNA polymerase. When using PCR amplimer labelling methodologies, however, further optimization experiments, over and above those performed to optimize the standard PCR protocol, may need to be performed in order to achieve the correct ratio of modified/unmodified nucleotides within the PCR mix. Also, different thermostable DNA polymerases exhibit different incorporation efficiencies with different modified dNTPs, for example, the biotin labeled dNTP "Bio-16-dUTP" is incorporated more efficiently by Pwo (*Pyrococcus woesei*) thermostable DNA polymerase (Roche Diagnostics, Mannhein, Germany) than by Taq (*Thermus aquaticus*) or Tth (*Thermus thermophilus*) polymerases. However, modified compounds such as fluorescein-12-dUTP or digoxygenin-11-dUTP (DIG-11-dUTP) are incorporated equally efficiently by all three of these enzymes, (though the optimum PCR mix

ratio of modified/unmodified dNTP may nevertheless vary between 1:3 and 1:20 dependant on the enzyme used).

There are, however, important cost aspects for PCR protocols that use labelled deoxynucleotides in the initial PCR mix. Specifically, a relatively large fraction of the labelled deoxynucleotides will remain unused in the final PCR mix and both DNA strands will be labelled even if post-PCR experiments only require a single-stranded labelled DNA fragment (probe) to be generated. If cost is indeed an important factor, and single-stranded labelled DNA probes are required, then an asymmetric PCR protocol (where one of the PCR primers is present in great excess) using labelled deoxynucleotides may be a useful alternative [Jouquand et al., 1999]. However, asymmetric PCR amplification protocols amplify DNA in a mainly linear fashion, as opposed to the exponential amplification achieved using balanced concentrations of PCR primers, resulting in a significant reduction in amplimer yield. Finally, it should be noted that labelled deoxynucleotides will also be incorporated into any non-specific PCR amplification products generated during PCR amplification.

Alternatives to using *labelled dNTPs* for PCR amplimer visualisation/detection include the use of *end-labelled primers* as well as post-PCR labelling protocols involving either *nick translation*, *random primer labelling*, or *terminal deoxynucleotidyl transferase* (TdT) reactions. Many commercially available kits are currently advertised. The use of end-labelled primers in a PCR does not alter the PCR conditions compared to non-labelled PCR conditions. The use of such kits has advantages: DNA strand specificity can be accomplished and only primers that correctly anneal will be extended and labelled. Most end-labelling PCR mixes employ a single labelled primer, though multiple labelled primers may be useful in differentiating amplimer species in multiplex PCRs. The *nick translation* protocol uses the enzyme pancreatic exonuclease I to generate single nucleotide "nicks" (deletions on one of the strands of double-stranded DNA) in the PCR amplimer, which are then filled in by the 5′ to 3′-end repair mechanism of DNA polymerase I from *E. coli*. Moreover, by including labelled dNTPs in the DNA polymerase I reaction mix, the nicks will be filled in with labelled dNTPs, resulting in a labelled amplimer. *Random primer labelling* protocols work by annealing a random mix of (decanucleotide) primers to single-stranded PCR amplimers and allowing any "gaps" to be filled in using exonuclease-free Klenow fragment DNA polymerase. Figure 12.2 shows two amplimer labelling strategies currently available. The enzyme *terminal deoxynucleotidyl transferase* (TdT) has the ability to add 5–10 (labelled) deoxynucleotides to the 5′-ends of a double-stranded DNA molecule using manganese as a cofactor. If these nucleotides are labelled, then the PCR amplimer itself will also be labelled.

Different labels (e.g. dyes which fluoresce at different wavelengths upon excitation) may be attached to PCR amplimers arising from different PCRs. If closely related (with respect to their sequence identities), then these amplimers may be utilized in competitive hybridization assays, where differential (competitive) hybridization is indicated by the presence or absence of a particular fluorescent dye signal. This procedure is often used to detect differences in gene expression

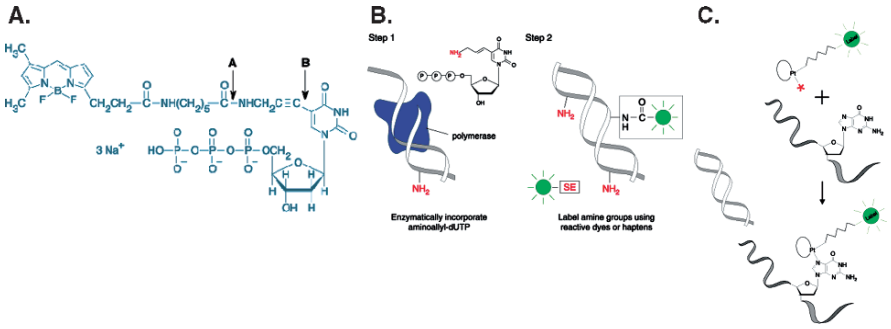

Fig. 12.2 Various important features of nucleic acid labelling technologies. A. The chromatide BODIPY FL-14-dUTP is representative for a variety of other aminoalkynyl dUTP nucleotides. Fluorophore labels are attached through a four-atom aminoalkynyl spacer (between arrows A and B). B. A two-step DNA labelling method in which aminoallyl dUTP is enzymatically incorporated (step 1) after which a reactive fluorophore is chemically coupled (step 2). C. Direct labelling of DNA using ULYSIS reagent. The reagent reacts directly with the N7 position in guanine and generates a stable coordination complex between nucleic acid and label (Copied from "The Handbook" at www.invitrogen.com. Copyright Molecular Probes, Inc.)

between different tissues or organisms, or between induced and non-induced tissues [Verhagen et al., 2000]. Labelled amplimers may be used as DNA probes in for example Southern hybridization reactions, however because single-stranded DNA is used in such protocols, one should ensure that the label is attached to the single-strand of the double-stranded amplimer that is to be used as a probe (see also other chapters).

Other means of labelling PCR amplimers includes the addition of phosphate groups to the 5′-end, allowing the amplimer to be used in cloning and ligation reactions (PCR amplimers are not phosphorylated at their 5′-ends). The enzyme T4 polynucleotide kinase may be used along with ATP (adenosine triphosphate) to phosphorylate the non-phosphorylated 5′-end of double-stranded PCR amplimers, thereby facilitating cloning of the amplimer into (plasmid) vectors using ligase enzymes. Alternatively, 5′-end phosphorylated primers may be utilized in the PCR mix *per se*. 3′-End phosphorylated primers will actually block PCR amplification (inert primer).

12.2 Two-Step PCR Protocol

Unlike the standard qualitative PCR protocol, a two step PCR protocol utilizes two distinct steps in each thermocycle instead of the usual three-steps (of denaturation, annealing and extension). The two-step PCR protocol uses a nucleic acid denaturation

step (whereby the reaction mix is heated to a temperature of between 92°C and 97°C), followed by a combined primer annealing and extension/amplification step whereby the reaction mix is heated to a temperature of between 50°C and 70°C. Taqman technology sometimes employs two step PCR protocols as well [Petitjean et al., 2005]. This contrasts with the more usual three-step PCR protocol where the annealing and extension temperatures of each thermocycle are separate. The advantage of using a two-step PCR protocol is that it allows the use of less stringent extension/amplification temperatures, a feature that may be particularly important when short amplimers with low annealing temperatures (e.g. random hexanucleotides or poly-dT primers of approximately 10–15 bp), or degenerate primers where PCR primer sequence mismatches may have to be tolerated, are to be used. During the heating and cooling periods of two-step PCR thermocycling, the ramp rate (i.e. the increase or decrease in temperature per unit time), may be varied in order to facilitate more efficient hybridization and elongation/amplification. A decreased ramp rate may increase the efficiency of a PCR, but unfortunately could also lead to the increased synthesis of non-specific PCR products. In some two-step PCR protocols dependant on the nature of the template (e.g. template molecules that are extremely rich in GC base pairs, it may be necessary to adjust the temperature ramp rate during the latter cycles of the PCR.

12.3 Booster PCR

The *booster* PCR protocol is designed to inhibit the accumulation of non-specific amplimers and primer dimer complexes, which may arise due to non-optimized PCR conditions and/or when target DNA concentrations are initially very low. Essentially, the first 15–20 cycles of booster PCR are performed using dNTP and primer concentrations which are 10–100 times less than in a standard qualitative PCR protocol so that the reaction equilibrium of the booster PCR is significantly shifted towards target-specific amplification, albeit at an initial cost to amplimer yield. After a set number of cycles, when sufficient quantities of specific PCR product have been amplified, thermocycling is paused and an extra quantity of primers and dNTPs (and thermostable DNA polymerase if required) are added to the reaction mix in order to achieve primer and dNTP concentrations similar to a standard PCR protocol. After addition of the extra quantities of reagents, the second *post-booster* phase of PCR amplification is performed using a further 30–50 PCR cycles with the same thermocycling parameters as the initial phase. Alternatives to the *booster PCR* are *nested PCR*, *hot-start PCR* (with or without and the use of thermally activated thermostable DNA polymerases) and *touchdown PCR*. It should be noted that all PCR protocols where the PCR vessel has to be opened after the initiation of thermocycling (including *booster PCR*, some *hot-start PCR*s and *nested PCR* protocols) carry an increased risk of being contaminated by amplified DNA (via aerosols, etc.) from neighbouring reaction tubes.

12.4 Hot-Start and Time-Release PCR Protocols

Non-specific primer hybridization and the formation of primer dimers may occur in all types of PCR protocols, especially if the initial target DNA: non-specific background DNA concentration is very low or multiple primer pairs are used within the same PCR mix (multiplex PCR). If a large concentration of non-specific background DNA is present in a non-optimized PCR mix, the vast majority of mis-priming events will occur during the initial pre-PCR heating step, a step required in order to melt the DNA duplex prior to the start of the PCR thermocycling program *per se* [Chou et al., 1992]. During this initial pre-PCR heating step when the temperature is sufficient to dissociate regions of the double stranded DNA template but not sufficient to ensure specific primer hybridization, primers will be able to hybridize to non-target regions of DNA, facilitating the amplification of non-specific PCR products, even at temperatures that are sub-optimal for amplification by the thermostable DNA polymerase. During subsequent PCR thermocycling, the temperature for primer hybridization does not drop below the pre-programmed hybridization temperature used in that particular PCR protocol, so that primer hybridization and amplification during these subsequent PCR cycles is specific. However, if primers have already hybridized to non-target DNA sequences and these non-specific sequences have been extended, then both non-specific DNA and target DNA will be co-amplified, resulting in a reduction in the yield of specific target amplimers and ongoing depletion of reaction components (and essentially a reduction in PCR sensitivity). The result is therefore a non-specific, insensitive PCR protocol, which is not ideal for either the clinical laboratory or research purposes. Several methods have now been developed which help to prevent non-target primer hybridization during the initial pre-PCR heating step. One alternative is the use of various single-stranded DNA "aptamers" that efficiently inhibit the polymerase activity of enzymes from the thermostable *Thermus* genus of bacteria (e.g. *Taq* and *Tth*), as long as the temperature remains below the melting temperature of the hairpin containing aptamer [Lin and Jayasena, 1997] (Fig. 12.3). Another, seldom used procedure involves the addition of "carrier DNA" to the PCR mix which shields all potential primer hybridization sites. Consequently, PCR primer hybridization only occurs at regions of DNA where the primers can successfully compete with the carrier DNA for binding sites, i.e. where the PCR primers bind in a highly specific manner to their specific DNA target sequence. Alternatively, some thermostable DNA polymerases show a much reduced activity at low temperatures, meaning that they are less capable of amplifying non-specific products during the initial PCR pre-heating step. However, neither of these methodologies is widely used in clinical or research laboratories. Instead, "hot-start" and "time release" PCR methodologies have become the most preferred methods used for inhibiting non-specific product amplification during the initial pre-PCR heating step (before the first cycle of thermocycling begins).

(i) Hot-start PCR: The hot-start PCR methodology, involves preparing PCR mixes with one of the essential ingredients (thermostable DNA polymerase, PCR primers,

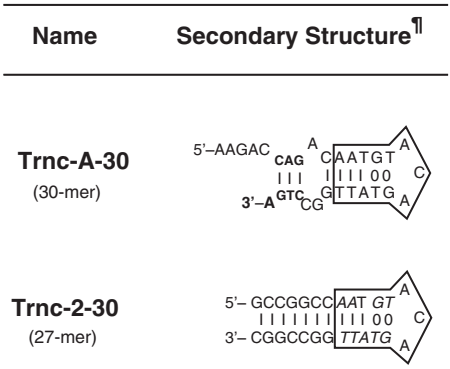

Fig. 12.3 Aptamers are small hairpin forming molecules that bind to Taq polymerase at temperatures below 40°C and exert a sequence specific inhibiting effect on polymerase activity. The complex dissociates upon heating and binding is reversible. Aptamers are used to prevent non-specific amplification (From Lin and Jayasena, 1997. Copyright Elsevier)

magnesium ions, dNTPs or rarely template DNA) missing from the mix. After heating to a temperature of between 92°C and 97°C (the pre-PCR heating step), the missing ingredient is then manually added to the reaction mix and thermocycling is allowed to proceed. Subsequent lowering of the temperature to the pre-programmed annealing temperature will then allow specific primer hybridization to occur and the amplification of specific target DNA during the first and subsequent PCR cycles [Mullis, 1991]. The hot-start PCR protocol helps to ensure that only specific PCR primer annealing occurs and allows the amplification of as few as 1–5 specific target DNA molecules within a background of up to 1 µg of non-specific DNA. A very serious disadvantage of the hot-start methodology involving the manual addition of PCR ingredients is the increased likelihood of carry-over contamination that arises when the PCR tube is opened (particularly if the PCR assay is being performed on multiple samples within the same PCR heating block, as all of the reaction tubes will have to be manually opened and closed, thereby increasing the chance of aerosol mediated contamination). Also, opening of reaction tubes at high temperature allows evaporation of the ingredients to occur, which may lead to small changes in the concentration of key PCR ingredients. To date, several ingenious alternative procedures have been developed to circumvent the contamination problems associated with the manual hot-start procedure. One option involves pipetting mineral oil onto the surface of each PCR mix (prior to any heating and thermocycling steps) and then pipetting the missing ingredient onto the surface of the mineral oil. Upon heating, the viscosity of the mineral oil will decrease and any ingredient pipetted onto the surface of the oil will readily diffuse into the PCR mix itself. N.B: Mineral oil was initially used to prevent evaporation of the PCR mix into the lid of the reaction tube during thermocycling, a necessity for older PCR machines which did not possess heated lids.

In a similar manner, solid DNase-free wax plugs may be used instead of mineral oil. The wax is added to the surface of the PCR mix as a semi-solid material and the reaction tube is then cooled on ice to allow solidification of the wax. The missing PCR reaction mix ingredient is then pipetted onto the surface of the solid wax layer. Subsequent heating of the PCR vessel results in the wax gradually melting, allowing diffusion of the missing ingredient into the PCR mix. In order to save time, many manufacturers now supply DNase-free wax "beads" which have been impregnated with magnesium chloride (an essential PCR ingredient). The wax bead is simply dropped onto the surface of the incomplete PCR mix and upon heating melts releasing the magnesium chloride. Finally, another ingenious hot-start PCR methodology involves the use of primers that contain internal self-complimentary sequences that form a hairpin-like stem-loop ("molecular beacon") structure at low temperatures. Upon heating the PCR mix, thermal energy disrupts the stem-loop structure generating a linear primer, allowing specific hybridization to occur [Kaboev et al., 2000].

(ii) Time Release PCR: Time release PCR has the same objectives as hot-start PCR, namely the reduction of non-specific PCR amplification products, though nowadays the term "hot-start PCR" is usually used for all hot-start and/or time release PCR methodologies. Time release PCR involves the use of novel thermostable DNA polymerases which require an extended period of heat treatment at relatively high temperatures (e.g. ±8 minutes at 95°C) in order to become activated (N.B. various alternative activation protocols have been described [Kebelmann-Betzing et al., 1998]). This means that PCR mixes can be prepared in their entirety prior to PCR thermocycling without the risk of non-specific primer hybridization (as the thermostable DNA polymerase will not be active until the temperature within the PCR mix is above the temperature required for specific primer annealing). The advantages of time release PCRs include the fact that they generate a greater yield of specific amplimer from a given concentration of reagents, are less sensitive to the action of PCR inhibitors, and generate better quality sequence data than non-hot-start PCR protocols [Moretti et al., 1998]. Examples of commercially available *time release PCR* enzymes include; AmpliTaq Gold (PE Biosystems), Thermostart (AbIgene), Hot Start Taq (Qiagen) and Proofstart (Qiagen) (Fig. 12.4). All of these particular enzymes have had their active sites artificially mutated so that their DNA polymerase activity is temporarily inactivated. However, upon heating, the three-dimensional conformation of the active site of the enzymes changes, restoring polymerase activity.

Other commercially available time release PCR-specific thermostable polymerase enzymes facilitate inhibition of amplification via binding of monoclonal antibodies to the enzyme, effectively blocking the active site of DNA synthesis as first described in the mid-1990s [Kellogg et al., 1994]). Upon heating, the monoclonal antibodies are denatured and passively dissociate from the enzyme, de-blocking the active site and restoring DNA polymerase activity. Commercially available examples of such enzymes include; Taq Platinum (Life Technologies), FastStartTaq (Roche), KOD polymerase [Mizuguchi et al., 1999], HotStarTaqDNAPol (Qiagen),

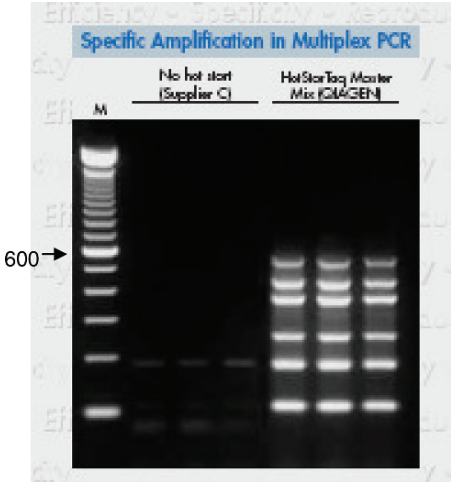

Fig. 12.4 Hotstart procedures involving activatable enzymes prevent the amplification of non-specific PCR procedures. At high DNA dilution there is an obvious clearing effect of the use of Hot Star Taq (five lanes on the right) as compared to non-hot start alternatives. Fragment of the p53 gene were amplified from genomic DNA in multiplex PCR. Parallel reactions were prepared using standard reaction conditions. On the left in the lane marked M, a 100 basepair molecular weight marker is shown with the 600 basepair fragment highlighted (From Critical Factors for Successful PCR at www.qiagen.com)

and RedHot DNA polymerase (ABIgene). It is also possible to purchase certain anti-DNA polymerase monoclonal antibodies from commercial suppliers, e.g. Jumpstart (Sigma) and TaqStart Antibody (ClonTech), which are added along with the non-time release thermostable DNA polymerase to the initial PCR mix. Newer monoclonal antibodies are currently under development which will be able to bind to different regions of specific polymerases. Another example of an antibody mediated time release enzyme is AccuPrime *Taq* DNA polymerase (Invitrogen), an enzyme mixed with accessory proteins of "unknown nature" (i.e. they are a commercial secret), which improve the specificity of primer annealing to target DNA. The protein may be similar to a "minor groove binding protein" previously described [Kutyavin et al., 2000]. The HotPrime polymerase (Qbiogene) is modified in yet another way, in that "heat labile blocking groups" have been coupled to some of the key amino acids of the enzyme. The application of heat (15 minutes at 95°C) dissociates the blocking groups restoring polymerase activity.

Certain classes of short double-stranded DNA fragments are able to inhibit non-specific amplification by *Thermophilus aquaticus* (*Taq*) and *Thermophilus flavus* (*Tfl*) enzymes at temperatures below 60°C. However, this inhibition effectively disappears at temperatures above 60°C [Kainz et al., 2000]. Ideally, the annealing temperature of these short double-stranded DNA fragments should lie between 60°C

and 70°C in order to exert their maximum effect. If the titre of these fragments is optimized with respect to a particular PCR protocol, then the quantity of specific amplimers generated in that particular PCR protocol will be greatly increased.

Reverse transcription (RT-PCR) protocols (where RNA is reverse transcribed into cDNA prior to PCR amplification) may also be transformed using either a hot-start or a time release RT-PCR protocol. The DuraScript HSRT PCR kit (Sigma) uses a mutant AMV RT (Durascript), which is robust enough to synthesize cDNA at elevated temperatures (up to 65°C) using random nine-mer primers equipped with 23 base pair oligo-dT extensions. The HSRT kit comprises the Durascript reverse transcriptase enzyme in conjunction with Sigma JumpStartAccu Taq LA thermostable DNA polymerase mix (bringing both exponential amplification and proofreading activity to the PCR). This (and other) adapted RT-PCR hot-start/time release protocols facilitate the reverse transcription of low copy numbers of mRNA and increase both the yield and specificity of RNA amplification. More of these commercially available RT-PCR enzyme combinations are now appearing on the diagnostic market.

12.5 Inverse PCR

PCR amplification proceeds in the 5′–3′ direction using primers that hybridize to known sequences of the target DNA molecule to be amplified. It is however possible to utilize PCR to amplify unknown "flanking" DNA regions that lie immediately upstream or downstream of the normal PCR primers used. This technique is called *inverse PCR* and can be used to amplify flanking regions of upto 3–4 kbp in length. The prerequisites for inverse PCR are that DNA sequences internal to the required flanking regions to be amplified are known, and that "upstream" and "downstream" restriction sites exists within these unknown flanking regions. The known DNA sequences are used to create "inverse" PCR primers, which point in the opposite direction to a standard PCR but still hybridize to complementary DNA. In effect, these inverse PCR primers amplify outwards and away from each other, rather than inwards towards each other (as is the case for a standard PCR).

The inverse PCR methodology *per se* involves; isolation of template DNA followed by restriction digestion using an appropriate enzyme(s), cleaning of the restriction fragments to remove any inhibitory enzyme and buffers, ligation and circularisation of the pooled fragments using DNA ligase, and then final PCR amplification using specific inverse PCR primers. The appropriate restriction enzyme(s) to be used are usually selected by "trial and error" by performing inverse PCR using a range of frequent cutting restriction enzymes, e.g. *Rsa I, EcoR I* etc or suitable frequent cutting isoschizomers (restriction enzymes that recognize the same DNA sequence), or by digesting the template DNA with a range of frequent cutting restriction enzymes and then hybridizing the fragments to PCR products generated from the known (internal) sequences of DNA to check that the restriction enzymes/combinations used do not

digest within the known sequence. After digestion of the template DNA, the many fragments are preferably ligated together in varying dilutions. One of these ligated fragment combinations will possibly be a circularized fragment containing the known DNA region of interest with extra upstream and downstream flanking DNA. Subsequent PCR amplification of the total ligation mix using inverse PCR primers, will allow PCR amplification outwards around the circularized DNA forming a linear PCR product. Thereafter, the unknown upstream and downstream flanking DNA sequences may be determined by sequencing using the inverse PCR primers as upstream and downstream sequencing primers (Fig. 12.5). If multiple amplimers are present after inverse PCR, non-contiguous DNA fragments (i.e. digested fragments from another part of the genome) may have erroneously ligated to the required DNA fragment. By electrophoresing, excising, purifying and sequencing each amplimer from an agarose gel, the sequences of all of the multiple inverse PCR amplimers may be determined. By comparing all of the sequences it will be possible to determine a conserved sequence, which will comprise contiguous DNA from the known and flanking DNA regions. The DNA sequences can also be searched for the presence of restriction digestion sequence sites recognized by the enzyme used in the initial inverse PCR digestion reaction. The required PCR product without contaminating DNA insertions should contain only one such restriction site as the ligation of non-contiguous fragments requires at least two restriction sites. Alternatively, restriction digestion of the correct circularized amplimer using the same enzyme as was used in the inverse PCR protocol, should yield only one digestion product. These results also provide information on the length of the upstream and downstream flanking sequences successfully amplified. However, logic dictates that the smallest amplimer will be the amplimer of interest, as the smallest amplimer is the one least likely to contain contaminating non-contiguous DNA.

When successful, inverse PCR can be used to circumvent time consuming and costly cloning and sub-cloning procedures, though the amount of sequence data provided for each flanking region may be somewhat limited (often in the region of 100–500 bp in length) due to restrictions on the size of linear DNA fragments that can circularize and self-ligate. The method is however well suited for improving the amount of sequence information available for short genomic regions, sequencing clones from genomic libraries, generating probes specific for uncharacterized flanking DNA sequences, and determining the presence of genomic translocations and gene fusion events by transposable or randomly integrating sequence elements. Recent advances in this field include PCR-based amplified RFLP [Liu et al., 2004] and long distance inverse PCR [Sonoki et al., 2004].

12.6 Asymmetric PCR

Asymmetric PCR is a PCR methodology where one of the PCR primers is present in a higher concentration than the other (usually at a concentration 1:50 to 1:100 higher) [Gyllenstein, 1989], allowing high concentrations of single-stranded

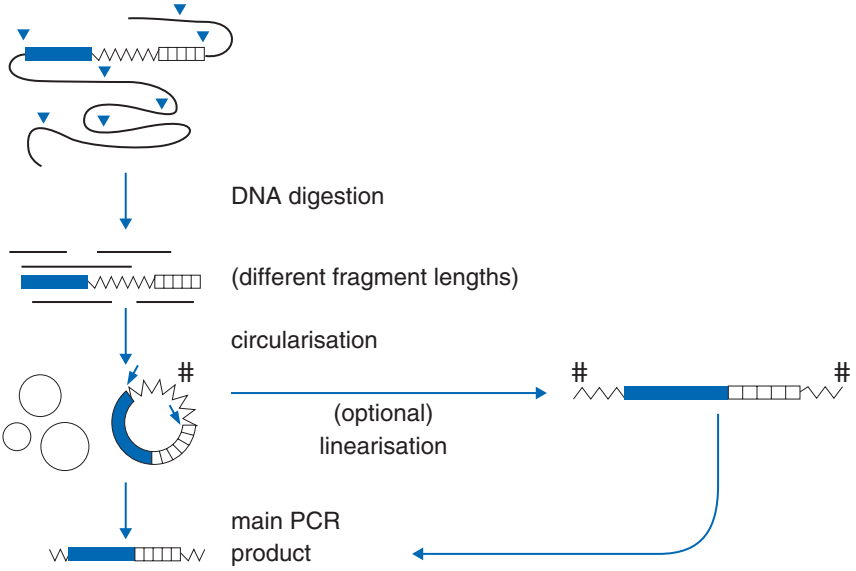

Fig. 12.5 Inverse PCR is a method for identifying unknown regions upstream and downstream (boxed) of a region of known sequence composition (zigzag). The flanking DNA regions are initially digested using an appropriate restriction enzyme (blue) and the resultant linear fragments are circularized by ligation. A PCR is then performed outwards from the known sequence using specifically designed inverse PCR primers. The final PCR product is sequenced to determine the composition of the flanking DNA sequence

DNA molecules up to 1,000 bp in length to be generated. During the first 15–20 cycles of an asymmetric PCR, the target DNA is amplified in a normal exponential fashion, however once the low concentration primer has been almost totally exhausted, exponential PCR amplification fails, as only the high concentration primer remains in the PCR mix. At this point and during the following 30–40 cycles, the vast majority of DNA molecules generated will be single-stranded, generated in a linear fashion. Asymmetric PCR therefore, is essentially a two-phase process, the first involving exponential amplification and the second phase linear amplification. If required for downstream applications, the single-stranded DNA molecules may be separated from any double-stranded amplimers by agarose gel electrophoresis and extracted from the agarose matrix using commercially available kits.

Asymmetric PCR may be utilized for generating single-stranded DNA probes for hybridization reactions, as well as in certain (rarely used) sequencing protocols which require for example the thermolabile *E. coli* DNA polymerase Klenow fragment (see Section 12.7.1) below). Asymmetric PCR also provides a convenient

alternative to classical cloning and sequencing strategies for phage λ clones (which are laborious to grow and purify).

12.7 PCR Mediated DNA Sequencing Strategies

Determining the nucleotide sequence of DNA amplimers, plasmids, genomic clones, etc. is an essential course of action for many research and clinical diagnostic laboratories. In order to obtain sequence data, the vast majority of laboratories use PCR mediated sequencing protocols, especially designed nucleotides called dideoxynucleotides and fluorescent labels. The most important advantage of PCR mediated DNA sequencing is the fact that the thermostable DNA polymerases used allow reactions to be performed at high temperatures under stringent conditions, as well as helping reduce any problems which may be encountered with template secondary structure. One particular advance over the past decade, is that PCR sequencing protocols have been extensively developed, simplified and adapted for high throughput applications that use specially designed automated sequencing machines, and (for the vast majority) fluorescent detection using fluorescent dye labels attached to either dideoxynucleotides or one of the PCR sequencing primers. Large advances have also been made with regard to sensitivity via the use of lasers and charge coupled device (CCD) cameras, gel visualization (using several different fluorescent labels in a single PCR sequencing mix), speed (via capillary electrophoresis formats), and sequence data analysis/sequence assembly (using specialized sequence alignment programs, e.g. Megalign, DNAstar; Sci Ed Central, Scientific and Educational Software etc) (see also Section 9.4).

Initially, there existed two major DNA sequencing methodologies, namely the Sanger method [Sanger et al., 1992], which used DNA chain terminating dideoxynucleotides, and the Maxam-Gilbert method [Maxam and Gilbert, 1977] which used chemical processes to sequentially remove nucleotides from a DNA chain. However, the Sanger method has now replaced the Maxam-Gilbert method in both research and clinical laboratories, not least because of the toxic and carcinogenic chemicals required to perform the Maxam-Gilbert method.

In the Sanger method, extension of the DNA strand by DNA dependent DNA polymerase is terminated by incorporation of a specific dideoxynucleotide (a deoxynucleotide analogue) that lacks the necessary 3′-hydroxyl group required for further DNA chain elongation. Moreover, by generating a whole series of prematurely terminated DNA fragments of different lengths whose termini end in one of the four dideoxynucleotide analogues, then gel electrophoresis (to separate the fragments) and detection (to identify which dideoxynucleotide terminates which individual fragment), will allow the sequence of the original template DNA to be determined. It should be noted that all DNA or PCR products for Sanger sequencing reactions should first be purified away from any contaminating DNA, primers, dNTPs, enzymes, etc. which may interfere with the sequencing reaction.

12.7.1 Generating Single-Stranded DNA for Sanger Sequencing Reactions

PCR amplification products and genomic DNA are double-stranded in nature, which makes them unsuitable for use in Sanger sequencing reactions (which require single stranded DNA). However, it is relatively easy to separate double stranded DNA into single strands by heating at 100°C for 5 minutes and then "*snap-cooling* or *snap-freezing*" the resultant single strands by immediately placing them in either ice, a dry-ice/ethanol bath or in liquid nitrogen. An alternative method involves the use of alkaline denaturation (often used for denaturing double-stranded extra-chromosomal DNA or plasmids into single strands), though this method tends to generate poor sequence data when applied to double-stranded PCR products.

One particular method for generating single-stranded DNA from PCR products, involves the use of 5′-end hapten labelled (e.g. biotin) PCR primers. This labelled primer becomes incorporated into subsequent PCR products and can be used to attach the double-stranded PCR products to a solid phase (e.g. streptavidin coated ELISA plate wells, streptavidin coated magnetic beads, etc.). The complementary DNA strand may then be removed (e.g. by alkaline treatment) ready for single-stranded DNA sequencing. Another PCR method for generating single strands involves asymmetric PCR (Section 12.6), which may be used to generate significant amounts (5–10 pmol) of single-stranded DNA up to 1,000 bp in length. However, the main disadvantage of this technique is that asymmetric PCR optimization has to be performed for each new target DNA molecule to be sequenced.

Another method for creating single-stranded DNA for sequencing reactions involves the use of the enzyme lambda exonuclease III [Higuchi and Ochman, 1989]. Basically, lambda exonuclease III is a 5′-3′ exonuclease which is specific for double-stranded DNA and which requires the presence of a phosphate group at the 5′-end position of the double-stranded DNA for its activity. The procedure involves the 5′-end labelling of double-stranded PCR products using T4 polynucleotide kinase and ATP (or alternatively the use of a 5′-end phosphorylated primer), followed by exonuclease degradation of the phosphorylated DNA strand. The remaining single-stranded DNA can then be used as a template in further sequencing reactions without the use of double-stranded denaturation protocols. By far the most and widely used method for generating single stranded DNA for use in Sanger sequencing reactions however, combines Sanger sequencing reactions with thermocycling (*cycle sequencing* methodologies).

12.7.2 Classical Sanger Sequencing of Single-Stranded PCR Products

Classical single-stranded DNA sequencing methodologies comprise two distinct steps and use a heat stable DNA polymerase. The start of the sequencing reaction is performed at 42°C for 5 minutes using a polynucleotide kinase enzyme or DNA

polymerase to label either specifically labelled sequencing primers or the nucleotide deoxyadenosinetriphosphate (dATP) with a (radioactive) label, e.g. ^{32}P. In the second step, four new reaction mixes are prepared each containing a normal mix of dNTPs and one of four dideoxynucleotide triphosphates (ddATP, ddCTP, ddGTP or ddTTP). DNA chain extension and chain termination is then performed at the optimum temperature for the polymerase enzyme (70°C) for a short period of time. Chain extension is terminated upon the incorporation of a ddNTP molecule. Products are separated by electrophoresis and visualized on the basis of the specific characteristics of the incorporated label. No amplification is involved in the classical Sanger sequencing methodology.

12.7.3 Direct PCR Sequencing

This sequencing protocol begins with a standard optimized PCR (15–30 cycles) that is used to generate millions of copies of the required DNA for sequencing. After PCR amplification is complete, a small aliquot of the PCR mix is added to a commercially available "sequencing mix" containing a special sequencing thermostable DNA polymerase that exhibits no 5′-end to 3′-end exonuclease (proofreading) activity, an activity which could remove any dideoxynucleotide terminators incorporated into the DNA chain (see Section 7.3). For the following sequencing reaction, one sequencing primer is used along with normal dNTPS, a labelled nucleotide (e.g. α–^{32}P dCTP), and one of the four dideoxynucleotide DNA terminators (either ddATP, ddCTP, ddGTP or ddTTP). The sequencing reaction in itself results in the linear accumulation of single-stranded DNA molecules (as only one primer is used) of different sizes, in which the detection label (e.g. α–^{32}P dCTP) has been incorporated. After PCR cycling, the DNA fragments generated in each reaction mix are analyzed using polyacrylamide gel electrophoresis and visualized by autoradiography or detection of fluorescence (Fig. 12.6). However, this somewhat old fashioned method is labour intensive, requires large quantities of template, and requires the use of very sensitive detection techniques, e.g. radioactive labels and autoradiography, which may have health and safety implications.

12.7.4 Four-Tube Cycle Sequencing

Cycle sequencing is performed using a standard PCR thermocycling machine and a sequencing reaction-specific thermocycling program. Both double-stranded and single-stranded DNA template may be used as starting material, as the temperature achieved during thermocycling is sufficient to thermally disassociate any double-stranded DNA into single strands. Cycle sequencing protocols utilize four standard PCR mixes, with each additionally containing one of four labelled chain terminating dideoxynucleotides (either ddATP, ddCTP, ddGTP or ddTTP). In contrast to a

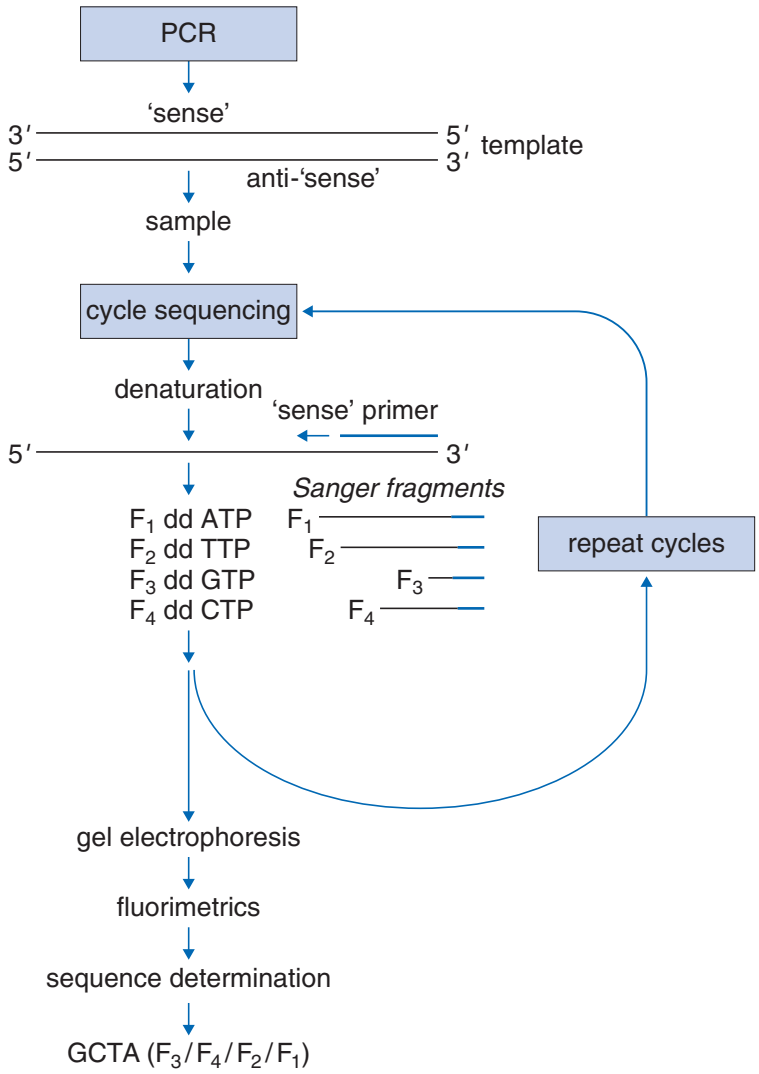

Fig. 12.6 Cycle sequencing using PCR. Cyclical denaturation, annealing and extension of target DNA in the presence of ddNTPs leads to the linear amplification of DNA molecules of different lengths

standard PCR methodology however, only a single (sequencing) primer is added to the sequencing reaction mix, which means that amplification is linear (as opposed to a standard PCR where two primers are used and the amplification is exponential). Various sequencing protocol options (including the use of either labelled dNTPs or labelled primers) are currently available. The terminated fragments in each of the four sequencing reactions are then separated by electrophoresis in a single polyacrylamide gel, and the labeled fragments then detected using (for

example) radioactive sensitive film (e.g. Hyperfilm, Amersham International). By examining the developed film, it is possible to associate a particular ddNTP with each terminated fragment length and hence determine the DNA sequence by reading down the gel.

Unfortunately, the fact that DNA synthesis during cycle sequencing is linear, means that this method is relatively insensitive, and requires the use of sensitive labelling/detection systems. Traditionally, radioactive labels have been used because of the high degree of sensitivity that these labels provide. However, developments in the field of fluorescent dye labels which are more able to efficiently transfer laser excitation energy into fluorescence energy (up to six times greater than previous earlier fluorescent dyes, Fig. 12.7), now means that fluorescent cycle sequencing is sensitive enough to generate readable sequence data using a relatively small amount of template and a reduced number of PCR cycles (10 ng for 15 cycles) from DNA fragments of up to 1,000 bp in length.

During the initial stages of the development of cycle-sequencing, it was observed that the available thermostable DNA dependent polymerases had problems in correctly recognizing the respective dideoxynucleotide terminators (ddNTPs), which were consequently incorporated into the growing DNA chain far less efficiently than the normal dNTPs, requiring extensive optimization of the ddNTP

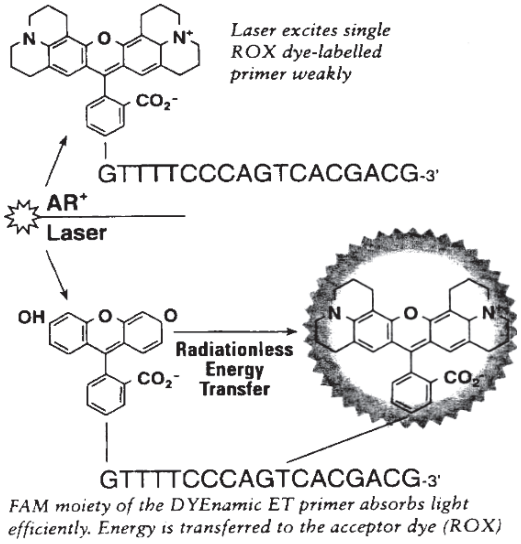

Fig. 12.7 ET fluorochromes are clever combinations of different fluorochromes which show increased fluorescence as compared to the single fluorochrome alone. For example, the excitation of the fluorescent dye ROX (structure on top) by an argon laser induces low levels of fluorescence; however if ROX is excited through transfer of energy from a FAM molecule (structure at bottom), then the resultant fluorescence is six times more intense

versus dNTP concentration. This effect also resulted in the poor sensitivity initially observed using the cycle sequencing method. However, the application of molecular techniques has now allowed the development and manufacture of newer, specifically engineered, thermostable DNA polymerases which show a much improved affinity for ddNTPs, produce uniform band intensities and exhibit a high degree of accuracy. Many companies which supply thermostable DNA polymerases for PCR now also supply variant polymerases specifically designed for sequencing purposes, e.g. Thermosequenase (Amersham International Ltd.), Sequencing Grade *Taq* DNA Polymerase (Promega Corporation) and Amplitaq DNA polymerase FS (Perkin Elmer/ABI).

12.7.5 One-Tube Cycle Sequencing

Sanger cycle sequencing has been completely revolutionized by the development of a broad spectrum of non-radioactive, fluorescent-based, dye labels which can be chemically attached to dideoxynucleotidetriphosphates and which fluoresce with their own specific spectral characteristics. The fact that these fluorescent labels exhibit their own specific fluorescent spectral characteristics (with a minimum of overlap of fluorescent wavelength between each individual dye), means that each of the four dideoxynucleotide terminators used in Sanger cycle sequencing (i.e. ddATP, ddCTP, ddGTP and ddTTP), may be individually tagged with its own specific dye label. Consequently, due to the dissimilar spectral characteristics of each of the four labels, all four tagged ddNTPS may be added to a single cycle sequencing reaction mix, allowing *one-tube cycle sequencing* to be performed. Examples of such fluorescent dye labels include FAM, HEX, TAMRA and ROX.

Basically, the methodology for *one-tube cycle sequencing* is the same as that used for *four-tube cycle sequencing*, except that only one reaction mix (containing all four differentially labelled ddNTPS) is used for each DNA template instead of four separate reaction mixes. After thermocycling, the whole range of terminated DNA fragments are separated by polyacrylamide gel electrophoresis, excited by a static laser, and the fluorescence measured as each fragment migrates down the gel. The quantity and particular wavelength of fluorescence emitted by each fragment is continuously measured and represented on a computer screen as a coloured peak (similar to a normal distribution curve). This allows the terminal ddNTP of each fragment to be determined by measurement of the fluorescent spectrum emitted by the fluorescent dye incorporated into that particular fragment. When collected together, the final result is presented as a wave-like line of different coloured peaks and troughs called a “chromatogram” (Fig. 12.8), with the colour black being associated with the nucleotide guanosine, the colour blue associated with the nucleotide cytosine, the colour red associated with the nucleotide thymidine and the colour green associated with the nucleotide adenosine as standard. Lately, special DNA sequencing machines have been developed which include laser and detection modules (e.g. the Perkin Elmer Cetus 3700 and 7600 series DNA sequencing machines,

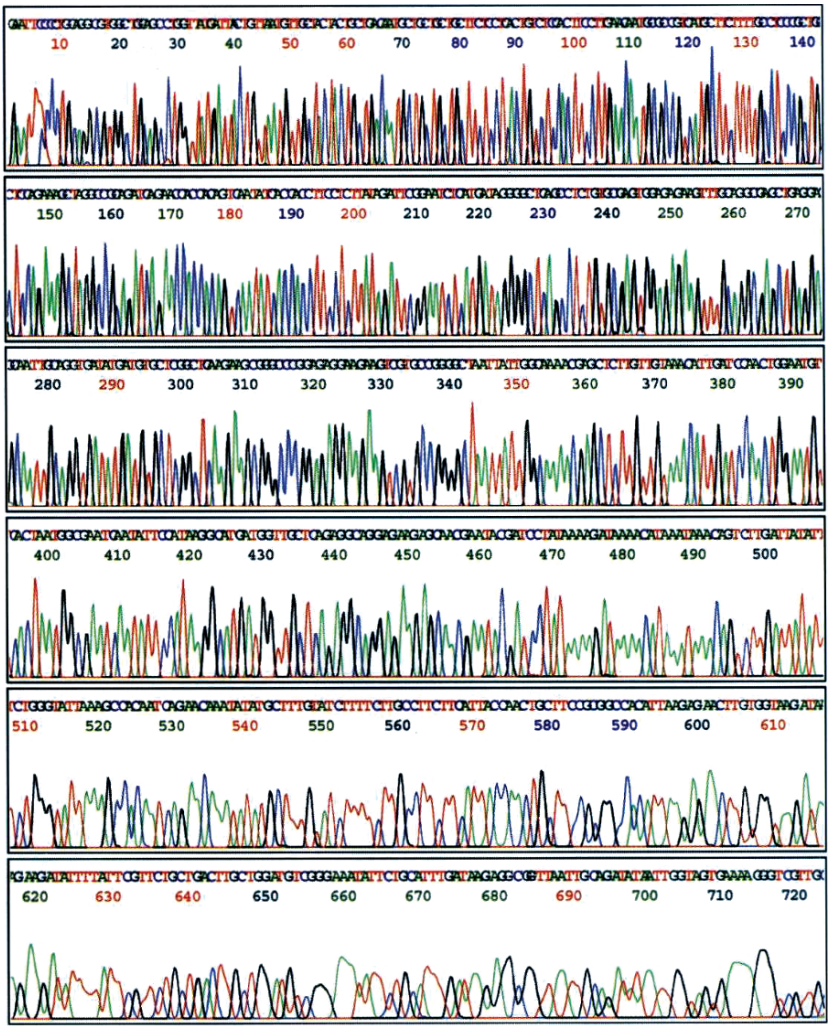

Fig. 12.8 DNA sequencing output: chromatogram showing the consecutive pattern of differentially labelled DNA products. Each peak represents the maximum fluorescence as obtained for each prematurely terminated DNA fragment. The fragments pass through a laser beam during polyacrylamide gel electrophoresis (PAGE)

capillary sequencer models), allowing 700–800 base pairs of sequence data to be generated from a single electrophoresis gel. Pre-prepared sequencing reaction mixes which contain the necessary labelled dideoxynucleotides, sequencing polymerase, dNTPs, magnesium ions etc are now also available for purchase from several commercial sources (e.g. Big Dye Terminator kits, Applied Biosystems).

12.7.6 Difficult to Sequence Templates

The presence of secondary structure in DNA molecules can interfere with Sanger sequencing reactions, as they hinder the progress of thermostable DNA polymerases along the DNA strand and may facilitate "base compressions", where incorrect sequence data is obtained due to "jumping" of the polymerase over any bases involved in secondary structure formation. Several novel options are available for helping to reduce the influence of such secondary structure on sequencing reactions. Options include sequencing the opposite DNA strand, altering the sequencing reaction conditions by increasing the melting temperature (e.g. to 98°C) or by including 40% formamide, increasing the extension temperature by 2–3°C, switching to Sequenase chemistry, utilizing 7-deaza-dGTP or dITP, adding betaine to a final concentration of 1.0 M, or by adding DMSO to a final concentration of 5% (vol/vol). Alternatively, the template may be shortened by restriction enzyme digestion outside of the region to be sequenced, thereby removing the secondary structure sequence entirely or generating shorter fragments that are less likely to form complex secondary structure. If the problem involves incomplete sequencing primer hybridization to a region of secondary structure, a new sequencing primer closer to the original sequencing primer may be designed to provide better hybridization properties. Another option is to sequence a set of smaller, overlapping fragments and then align all of the sequences into a single piece of data. Fidelity Systems Inc, offer ThermoFidelase I, a "hyperstable protein" which binds to DNA, and apparently helps DNA polymerases read through any secondary structure present (e.g. regions of high guanosine and cytosine content or repeat regions), as well as protecting DNA from thermal degradation. The protein is simply added to the sequencing reaction mix prior to thermocycling. Finally, it should be noted that problems with the sequencing of secondary structure may arise due to the type of dye-terminator chemistry used, with Big Dye terminators being more susceptible to problems than Rhodamine chemistry, though the introduction of a Big Dye dGTP kit (Applied Biosytems), specifically for sequencing templates with difficult secondary structure, has helped address the problem associated with Big Dye chemistry.

Homopolymeric regions (e.g. long stretches of adenosine nucleotides) can result in "slippage" of the polymerase (exact mechanism unknown) and short repetitive regions (200–500 bp in length) can also cause sequencing problems. It may be necessary in these cases to utilize transposon insertion as a method of introducing specific primer sequences into the regions followed by PCR sequencing in both directions, e.g. EZ-Tn5 and Hyper Mu (Epicentre).

"Next generation" sequencing strategies are currently reaching advanced stages of development and use. The reader is referred to "The year of sequencing", Nature Methods 2008; 5(1):11–21, for more information.

12.8 Touchdown and Touch-Up PCR

The touchdown PCR protocol is particularly suited for: (i) use on highly complex DNA samples containing a limited number of initial template molecules, (ii) for multiplex PCRs which contain sets of primers with very different annealing

temperatures, (iii) for the selective amplification of an intended target region where one or more primer sequence is very similar to other sequences present within the template DNA to be used, and perhaps most usefully, (iv) as a general thermocycling protocol for use with many different PCR primer pairs (without having to determine the exact annealing temperature required for each individual PCR primer pair and having to re-program the PCR thermocycling machine).

Touchdown PCR comprises a two-step methodology. In the first step (10–15 thermocycles), the annealing temperature of the very first thermocycle within this step is set at a temperature of approximately 10°C above the calculated T_m of the primer pair being used (e.g. 65–70°C). At every subsequent cycle in the first step the annealing temperature is decreased by a standard increment (usually either 0.5°C or 1°C) until the total number of cycles within the first step is complete and the final cycle annealing temperature is equivalent to the calculated T_M for the primer pair used. In the second step (15–30 thermocycles), thermocycling is performed at a single annealing temperature, the annealing temperature of the final cycle in the first step [Wu et al., 2005]. The advantage of touchdown PCR is that primer hybridization conditions are initially very stringent (and sometimes even too stringent to allow primer annealing to occur!). However, as first step thermocycling progresses, and the annealing temperature gradually decreases, the annealing temperature will begin to approach the T_m of maximum stringency for the primer pair(s) being used, resulting in the amplification of specific target DNA only. As the annealing temperature is gradually lowered, the conditions for primer hybridization will become less stringent and the efficiency of amplification will increase. Moreover, the target DNA specifically amplified in the first step of the touchdown PCR will tend to act as a preferential target for further DNA amplification, even if the annealing temperature falls a few degrees below the optimum T_m of the primer pair(s) being used. Further, if any non-specific PCR products are actually generated in the second step, they will be greatly outnumbered by the specifically amplified target DNA present in the final reaction products.

Optimized touchdown PCRs in combination with hot-start protocols using time release polymerases (see Section 12.4 above) help to ensure the maximum specificity of PCR amplification. If problems with the yield of PCR product are observed using the touchdown PCR protocol, it may be possible to decrease the annealing temperature used in the second step of the PCR protocol in order to improve PCR yield [Hecker and Roux, 1996]. It is also possible to simplify touchdown PCR protocols by using a steeper decrease in annealing temperature per first step cycle (so called *step-down* PCR protocols) whilst maintaining PCR specificity and efficiency. Touchdown PCR has also been successfully applied to amplify DNA of inferior quality from archival (archaeological or post-mortem) samples.

The reverse process or "touch-up" PCR has also been described as a method for increasing PCR specificity. Touch-up PCR protocols normally utilize specially designed "molecular beacon" primers that contain self-complementary sequences or "stem-loop" secondary structure (e.g. Touch-Up and Loop Incorporated Primers or "TULIPS" [Ailenberg and Silverman, 2000]). During touch-up PCR, the annealing temperature is gradually increased per cycle (e.g. from 60°C to 72°C over six

thermocycles), whereby the increasing thermal energy gradually facilitates the disassociation of the primer self-complementary "stem-loop" structure, allowing the primer sequences to specifically hybridize to the intended target sequence. A standard PCR at the calculated T_m of the primer (which will be lower than the annealing temperature previously used) is then performed for a further 30–40 cycles to increase the yield of specific amplimers. Designing these special touch-up primers requires specific computer software, as both the T_m of primer loop disassociation and T_m of primer-template hybridization need to be taken into consideration.

12.9 Multiplex PCR

Multiplex PCR protocols utilize multiple PCR primer pairs, all included in the same PCR mix, to amplify several different target regions at the same time. The upper limit on the number of primer pairs that can be included in a single multiplex PCR mix depends on the reaction conditions (annealing temperature, magnesium ion concentration, dNTP concentration, amplimer sizes, template concentration, etc.), and most importantly the absence of interactions between individual primers (cross-hybridization) within the reaction mix. Multiplex PCR primers are usually slightly longer (23–28 nucleotides) than those used in single-target "standard"PCR protocols, and are selected on the basis of similar T_m and similar (at least 40% of the total primer sequence) guanosine/cytosine nucleotide content. Multiplex PCRs require extensive optimization and titration of reaction components, and several different thermostable DNA polymerases may need to be tried [Moretti et al., 1998; Kebelmann-Betzing et al., 1998]. It should be noted that the concentrations of the different primer pairs used do not have to be equimolar for successful optimization, not even within individual PCR primer pairs [Berg et al., 2000]. In some more complex cases, the addition of certain additives, e.g. DMSO, glycerol or formamide, may help in obtaining specific amplification products due to the ability of these chemicals to interfere with DNA secondary structure formation which may influence primer-target hybridization accessibility and efficiency.

Multiplex PCR protocols are frequently used in clinical genetics and in forensic investigations. They facilitate the co-detection of several targets in a single PCR and hence the identification of complex disease genotypes, e.g. Duchenne's Muscular Dystrophy. For forensic purposes, various single nucleotide polymorphisms (SNPs) can be detected using a single test, and microsatellite regions (defined as regions within DNA where short sequences of two, three, or four nucleotides are repeated one immediately after the other) can be screened for variability, thereby saving time with respect to the generation of genetic fingerprint data. A more specialized multiplex PCR protocol for SNP analysis uses different fluorescent labelled ddNTPs, allowing SNPs to be identified using for example, a capillary sequencing machine [Hebert and Brazill, 2004]. Reverse transcription PCR (RT-PCR) protocols can also be multiplexed, allowing differences in expression between two different genes (e.g. a "target" and a "housekeeping control" gene) to

be assessed in a single multiplex RT-PCR. Of course, suitable DNA contamination controls have to be included in multiplex RT-PCR protocols similar to those required in standard single primer pair RT-PCR reactions. Recent advances in multiplex RT-PCR protocols also include the development of novel multiplex PCRs using self-quenched fluorescent primers which contain 5′-end hairpin loops, yielding a better "signal to noise" ratio [Nazarenko et al., 2002].

It should be noted that DNA contamination is a much more severe problem in multiplex PCR protocols as opposed to single primer pair PCR protocols, because there are more target sequences to be amplified. Therefore, as with all PCR protocols, anti-contamination control measures should be vigorously implemented.

12.10 PCR Using Degenerate Primers

Degenerate primers are mixtures of PCR primers designed to amplify the same genetic DNA target sequence where small numbers of nucleotide changes between different isolates or individuals is expected. These minor target nucleotide sequence changes may occur between different microbial isolates or species (the "quasi-species" concept), or within the same cell type of different individuals. Degenerate primers are usually a mix of primers where the main "core" nucleotide sequence of the primer is identical to all of the other primers, but where one (or more) of the primer sequences contain changes in one (or more) nucleotides compared to the other primers present within the degenerate primer mix. Under low stringency conditions mismatching and cross hybridization between non-perfectly matching sequences are allowed. This is especially relevant during the first few cycles of the PCR, where lower annealing temperatures, prolonged annealing times and slow temperature ramp rates should be used. However, it may be possible to increase the stringency of PCR thermocycling after the first few (5–10) cycles, without affecting the yield of PCR products, as the majority of amplimers will now contain the original degenerate primer sequences. Degenerate primer mixes may be ordered from commercial suppliers, in which case, standard letter codes are used to indicate which particular nucleotides should be added at a specific primer position. These letter codes are:

R = A or G (puRine) **Y** = C or T (pYrimidine) **K** = G or T (Keto) **M** = A or C (aMino)
S = G or C (Strong-3H bonds) **W** = A or T (Weak-2H bonds) **N** = any base
B = C, G or T **D** = A, G or T **H** = A, C or T **V** = A, C or G

A commonly used PCR protocol involving degenerate primers is the "degenerate oligonucleotide primer" (see Section 5.1.1. as well) or "DOP-PCR" protocol, outlined in Fig. 12.9. The DOP-PCR protocol is able to randomly amplify minute amounts (≤1 μg) of DNA, and has been successfully used in the amplification of purified chromosome preparations and the generation of probes for comparative chromosomal hybridization [Alcock et al., 2003].

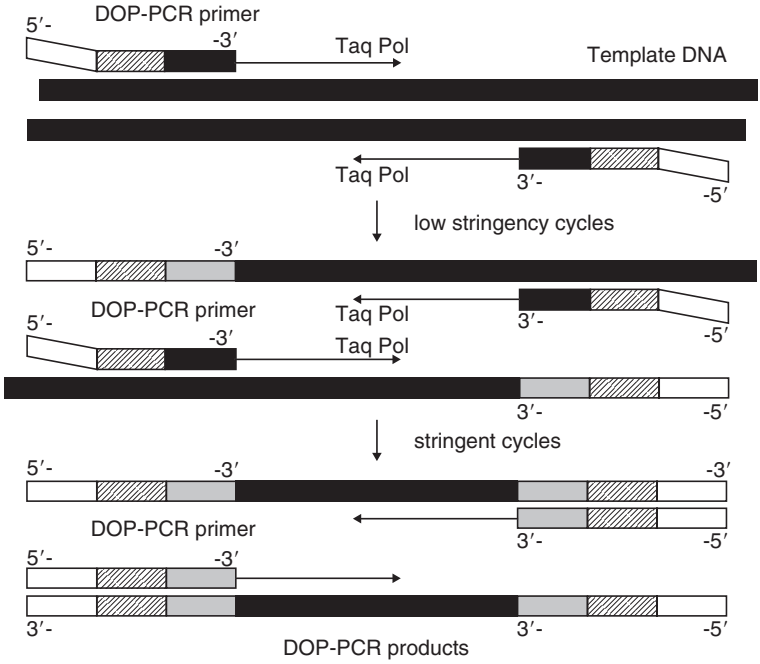

Fig. 12.9 DOP PCR. Degenerate oligomer primer PCR utilizes degenerate primers that contain an exact complementary six-nucleotide region at their 3′-end but contain six mixed nucleotides (where all four nucleotides have all been incorporated in equimolar amounts) in their central regions. At the 5′-end of the primer sits an additional region of nucleotides added to enhance the stability of hybridization interactions. Using DOP PCR primers and protocols generates a large variety of products depending on the complexity of the original template molecule used. Joos. S et al Roche product catalogue for ISH; Reproduced by permission of: 'Roche'

12.11 Repeat and Inter-repeat PCR

Strongly conserved repetitive sequences are found in nearly all genomic DNA molecules, even in organisms with relatively simple genomes, e.g. viruses. In bacteria, repeat DNA sequences or motifs, either in tandem or dispersed as single copies throughout the whole bacterial genome, have been successfully identified, characterized, and used in bacterial typing and epidemiological studies [Van Belkum, 1994]. With respect to the human genome, approximately 30% of all human DNA consists of repeat DNA sequences, which may vary from short (2–30 bp), to intermediate (300 bp), to large repeat units (≥300 bp), whilst mRNA is mostly devoid of these repeats. Highly repetitive sequences in the human genome are found in the telomeric and centromeric regions of chromosomes and short tandem repeat sequences or "STRs" are excellent genotyping targets and not only restricted to the human genome [Benecke, 1997; Singh et al., 2004; Leclair and Scholl, 2005].

Repeat sequences themselves are not usually 100% identical between different isolates or individuals, due to the fact that during DNA replication, DNA dependent DNA polymerases encounter problems attempting to copy repeat DNA sequences and may occasionally incorporate a mismatched nucleotide into the repeat sequence region. As the incorporation of mismatched nucleotides by DNA dependent DNA polymerase is a random process, the detection and characterization of repeat region variability provides an excellent means for identifying specific isolates or individuals. However, *in vivo*, the rate of such random mutations is held in check by several mismatch repair enzymes which act to prevent and correct mismatched nucleotide sequence errors, limiting such mutations to a rate of approximately 10^{-8} to 10^{-9} mutations per cell per generation [Boyer and Farber, 1998]. In fact, extensive repeat region sequence variation could point to a genetic deficiency in the DNA repair mechanism system of the individual presenting with such extensive errors, and microsatellite instability may be used to diagnose several hereditary diseases, including hereditary non-polyposis colorectal cancer (HNPCC) [Kruhoffer et al., 2005].

12.11.1 Repeat PCR

Repeat PCR protocols are designed to determine the length of genetic regions containing tandem repeat sequences [Van Belkum et al., 1995] (Fig. 12.10). Primers are designed that anneal specifically to either the 5′-end or 3′-end region of the repeat sequence to be investigated and are used as conserved PCR "initiating" regions. The length of the amplimer obtained using these protocols will be specific, and in the case of diploid organisms (where the presence of chromosome pairs means that heterozygosity can occur), it may be possible to PCR amplify two distinct repeat region fragments which will be distinguishable upon gel electrophoresis. Loss of one of these heterozygote amplimers (for example after malignant cell expansion during oncogenesis) can be easily detected. The subset of DNA repeat regions known as "short tandem repeats" are particularly suited for "repeat PCR" analysis, as they give rise to distinct and short PCR products, are easily amplified due to their relatively short size, and are easy to separate on standard agarose gels.

12.11.2 Inter-repeat PCR and Random Amplification of Polymorphic DNA (RAPD)

Inter-repeat PCR is based on the fact that repetitive DNA sequences may occur randomly within individual genomes, and that the intervening distances between these sequences in different isolates or individuals may also vary. In this case, PCR amplification from one particular repeat region to another repeat region gives rise

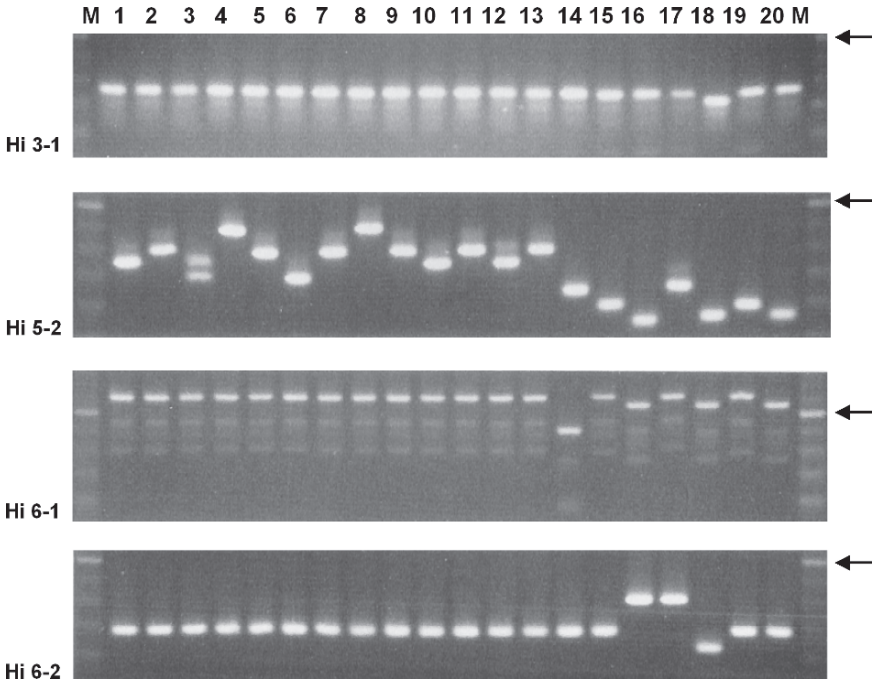

Fig. 12.10 Differences in repeat unit numbers at a particular repetitive DNA locus may be easily established using PCR primers designed to hybridise to neighbouring target sequence motifs [van Belkum et al., 1998]. Shown is an example of four different repeat detecting PCRs for a single set of *Haemophilus influenzae* isolates. Strains 1–13 were isolated during an outbreak of infection and it is obvious that in most (but not all) cases the repeat PCR is identical. Strains 14–20 are genetically unrelated and show a greater diversity. Lane M contains molecular weight markers and the arrow on the right identifies a 100 bp molecule

to variable fragment lengths and therefore variable genetic fingerprints within different isolates or individuals due to inter-repeat region length variation [Van Belkum, 1994] (Fig. 12.11). It should be noted that the presence and position of the repeat region does not have to be known prior to performing the PCR. In order to be successful however, the repeated motifs do need to be present in opposing orientations, i.e. facing in towards each other, so that a single PCR primer (a key feature of inter-repeat PCR protocols) can hybridize to both individual repeat sequences and generate a PCR amplimer. This process is generally referred to as random amplification of polymorphic DNA (RAPD) or arbitrary primed (AP-PCR). Such RAPD and AP-PCR protocols are usually performed with the help of relatively short primers (approximately ten nucleotides in length) which increases the chance of the primer "finding" and hybridizing to an opposing repeat sequence. In order to further increase the probability of primer hybridization, the annealing temperature used in these protocols is kept low (between 25°C and 42°C) so that partially identical stretches of DNA may also act as sites for primer hybridization. RAPD,

AP-PCR and indeed inter-repeat PCR protocols may generate complex mixtures of amplimers (and hence gel electrophoresis banding patterns), which need to be adequately separated for successful analysis. However, in most cases, simple agarose gel electrophoresis is sufficient to generate interpretable band pattern profiles (often referred to as "DNA fingerprints"). Densitometric scanning of the fingerprints may facilitate the automated interpretation of the sometimes dense arrays of DNA fragments. One particular problem encountered using RAPD and AP-PCR protocols is that the reproducibility of the PCR using a particular primer sequence and a specific DNA extract may be variable, and that selection of a suitable primer may be a complicated process involving multiple rounds of trial and error. However, these PCR protocols are convenient, rapid and flexible in the production of DNA fingerprints for genetic identification [Van Belkum, 2002]. PCR-mediated DNA fingerprinting

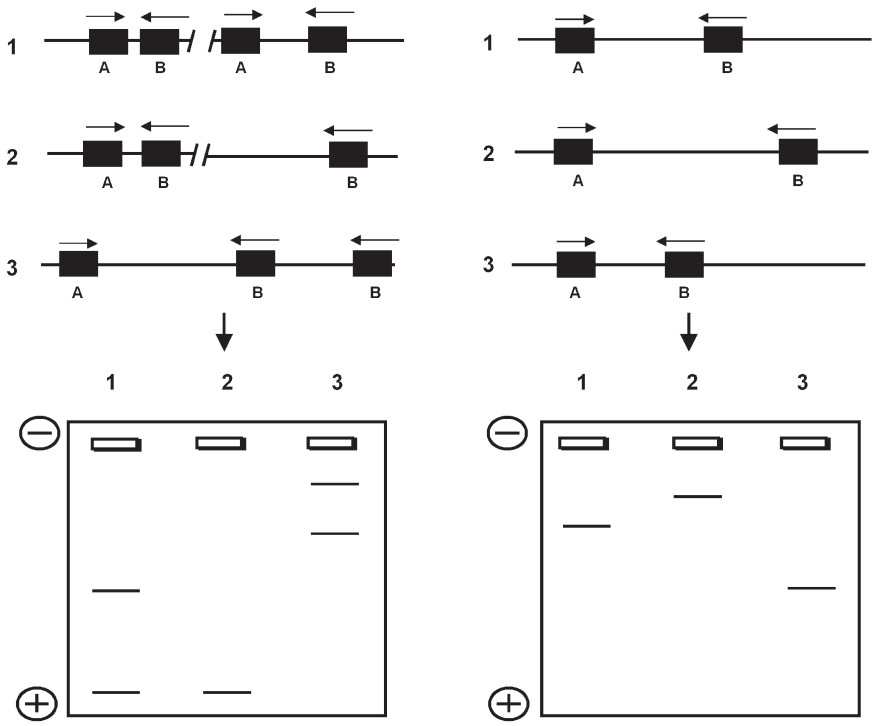

Fig. 12.11 Two variants of DNA fingerprinting mediated by PCR [Van Belkum, 1994]: (a) *Primer binding site heterogeneity*. DNAs 2 and 3 lack a site present in DNA 1. This results in the disappearance of a band in the electropherogram. In this example, multiple PCR primer pairs included in a single reaction mix may enhance the number of polymorphic sites that are detected; (b) *Primer binding site homogeneity*. Results of DNA amplification with primers that anneal to constant binding sites and span a variable segment of DNA. In this example, DNA segments 1 and 3 could contain deletions when compared to DNA segment 2. Alternatively, DNA segment 1 and 2 could harbour insertions, which are absent in DNA segment 3. The upper part of the panel provides the theoretical background for the electropherograms shown

protocols have been mainly used in epidemiological studies, to determine the evolutionary relatedness of different species (both in the eukaryotic and prokaryotic kingdoms), and in the determination of genetic polymorphisms between different individuals (paternity testing, forensic examinations, etc.) and may be readily adapted and applied to high-throughput screening strategies.

12.12 AFLP Fingerprinting

Amplification fragment length polymorphism (AFLP-PCR) analysis is a highly reproducible DNA typing and PCR amplification technique, which utilizes the presence/absence of restriction sites (restriction site polymorphisms) and PCR amplification to distinguish between different species or isolates [Vos et al., 1995]. The starting material for AFLP analysis comprises genomic DNA isolated from an individual organism, cell, tissue etc which is then digested to completion using one or more restriction enzymes (usually a "frequent" cutter and a "rare" cutter, e.g. *EcoRI* and *MseI*), so that a range of several thousands of short restriction fragments are generated of between 100 and 600 nucleotides in length. The best restriction enzymes to be used for a particular organism may be chosen by trial and error, or if the genome sequence is available, then computer software may be used. After complete restriction digestion, specific adapter molecules (short fragments of double-stranded DNA equipped with specific restriction sites and ligation linkers), are ligated to the "sticky" ends generated by the restriction endonuclease enzymes using DNA ligase. If two different restriction enzymes are used, then two different adaptors may be used together and ligated to the DNA fragments. A PCR is then performed using primers specifically designed to bind to the ligated adaptors. If two restriction enzymes have been used, then the PCR primers may be designed in such a way that only those fragments bordered by two dissimilar restriction sites (i.e. bordered by two different ligated adaptors) will be selectively amplified, so that fragments bordered by a single adapter type will not be amplified. After PCR amplification, the amplified fragments are separated by polyacrylamide gel electrophoresis so that the differences in banding patterns between isolates, cells, tissues, etc (due to amplification fragment length polymorphisms) may be detected. Individual fragments may be visualized using fluorescent or radioactively labelled primers or nucleotides, cut out of the gel and sequenced to reveal the genetic location of the restriction site polymorphism.

For many genomes and restriction enzyme combinations, the number of DNA fragments generated is so large that they are not separable, even by polyacrylamide gel electrophoresis. Hence, in many cases, minor modifications may have to be made to the AFLP primer sequences in order to limit the number of fragments PCR amplified. This is usually achieved by adding one or two "selective" nucleotides to the 3′-end of one or more of the PCR primers used. Hence, for a single primer with a single 3′-end extension, the number of fragments amplified will be reduced by one quarter, whilst for two primers each containing a single selective 3′-end nucleotide the number

of fragments amplified will be reduced by one sixteenth, under stringent thermocycling conditions (assuming a random and equal distribution of all four nucleotides within the genomic sequence to be tested by AFLP). In essence, AFLP is a high-throughput genotyping procedure, where multiple DNA samples may be analyzed and compared in parallel. Due to the versatile application of diverse combinations of restriction enzymes and PCR primers, the AFLP is, next to whole genome sequencing, the method of choice for high density mapping of genetic polymorphisms.

As well as being useful in determining the genetic relatedness of microbial isolates or individuals, AFLP banding patterns may also be used to examine the genetic basis of phenotypic differences between very similar organisms, tissues, etc. Genes or gene polymorphisms associated with a particular phenotypic characteristic may be putatively identified by: (i) generating AFLP DNA fingerprints between similar organisms, (ii) searching for band differences associated with that particular phenotype, (iii) cutting out and sequencing these bands from the polyacrylamide gel and finally, (iv) searching for sequence matches on Internet databases (e.g. the National Centre for Biotechnology Information, *http://www.ncbi.nlm.nih.gov/BLAST/* web site). AFLP exhibits several advantages or disadvantages over existing genotyping techniques [Torpdahl et al., 2005]

12.13 Base Excision Sequence Scanning (BESS-T-Scan) for Mutation Detection

BESS is a technique for determining T/A or A/T mutations and generating "genomic fingerprints" allowing differential comparisons between different isolates, species etc, to be made [Hawkins and Hofmann, 1999] (Fig. 12.12). The BESS PCR protocol is performed using a typical PCR mix but with one of the primers labelled at the 5′-end, and partial replacement of the dTTP (deoxythymidinetriphosphate) nucleotide by dUTP (deoxyuraciltriphosphate). This strategy overlaps with the one employed to prevent carry-over contamination. After PCR amplification, amplimers contain dUTP nucleotides, which when treated with the enzyme uracil N-glycosylase will be excised out of the amplimer (uracil N-glycosylase removes the organic base component from each incorporated dUTP nucleotide) creating single stranded "nick sites". These nick sites are subsequently available for digestion by the enzyme endonuclease IV, which cuts the DNA strand where the dUTP base has been removed. Separation of the single-stranded DNA fragments thus generated using denaturing polyacrylamide gel electrophoresis, followed by visualization using a label incorporated at the 5′-end of one of the primers, allows single nucleotide sequence differences between different amplimers to be detected. T/A or A/T mutations will result in the presence or absence of individual fragments respectively, and frame shift mutations, nucleotide deletions and differences in the number of DNA repeats can also be detected. The characteristic band patterns generated by this technique are dependant on the length and conformation of the individual single strands of DNA.

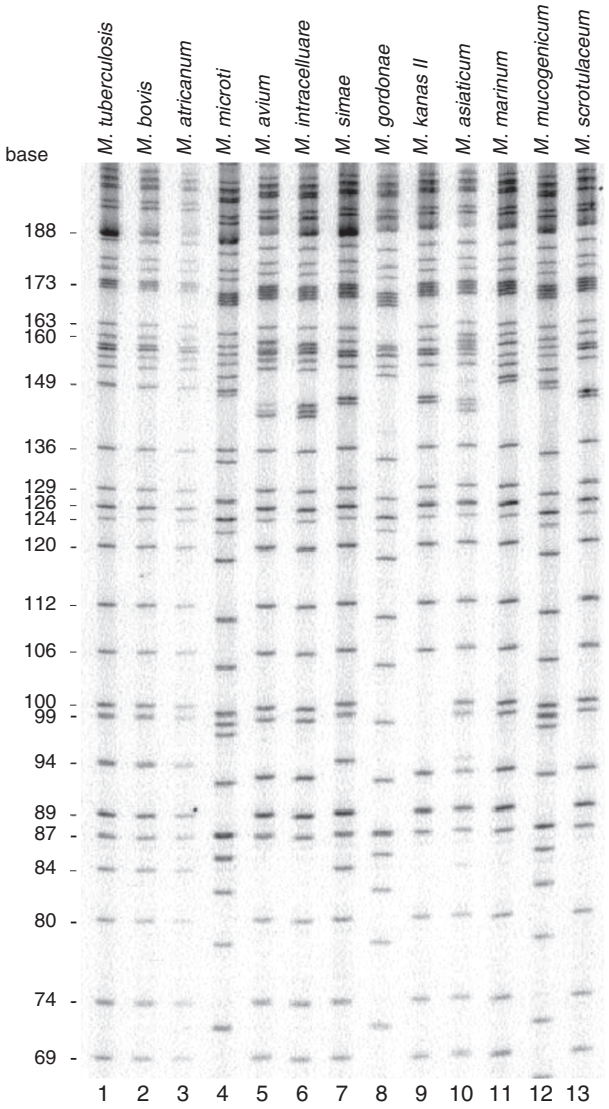

Fig. 12.12 Base excision sequence scanning. A standard PCR mix is prepared except that the nucleotide dTTP is partially replaced by the nucleotide dUTP. Upon PCR amplification, the amplimers contain dUTP, which when treated with the enzyme uracil N-glycosylase creates "nicks" by removing the organic base from each incorporated dUTP nucleotide. These nick sites are subsequently available for digestion by endonuclease IV. Separation of the single-stranded fragments generated using denaturing polyacrylamide gel electrophoresis, and visualisation using for example a label incorporated at the 5′-end of one of the primers, allows T/A and A/T mutations to be detected. Frame shift mutations, nucleotide deletions and differences in the number of DNA repeats, may also be detected. Shown is a BESS T-scan analysis of the 16S rRNA gene of 13 *Mycobacterium* strains. The samples were run on an 8% denaturing polyacrylamide gel (Reproduced from EPICENTRE Forum 5(2): "Single nucleotide typing of *Mycobacterium* species using the BESS T-scan mutation and localisation kit" by GA Hawkins. With permission from Epicentre Technologies Corp.)

12.14 Differential Display RT-PCR (DD-PCR)

Of the 10^5 different genes in a human cell, only 10–15% are actually expressed. Messenger RNA (mRNA) differential display techniques such as DD-PCR have been developed to visualize subtle differences in gene expression, both within different cell types and in the same cell type treated in different ways [Garcia and Castrillo, 2004; Castles et al., 1996]. The key steps in the DD-PCR protocol, begin with the extraction of mRNA transcripts (including DNase I treatment to remove any contaminating genomic DNA), and cDNA synthesis using a primer with a 5′-end comprising 10–12 oligo-dT nucleotides (eukaryotic mRNA molecules contain tracts of poly-adenosine nucleotides or poly-A tails and a 3′-end comprising either an A, C, or G nucleotide and one other nucleotide from A, C, G or T) so that the total length of the primer is approximately 12–14 nucleotides in length. The addition of the two extra nucleotides at the 3′-end, means that only a subset of all the many thousands of mRNA transcripts within the cells being compared are actually reverse transcribed into cDNA (thereby reducing the complexity of the final results to a manageable number of transcripts). After reverse transcription, an RT-PCR is performed using labelled dNTPs (e.g. radioactive ^{35}S-dATP), the 5′-end reverse transcription primer previously used in the reverse transcription reaction, and a 3′-end "arbitrary" primer (a random selection of ten nucleotides containing at least 50–70% GC content; AP-1 in Fig. 12.13). The DD-PCR products are then electrophoresed in a denaturing (results in better separation of bands) polyacrylamide gel and visualized (e.g. by autoradiography). Differences in banding patterns between different cell types or in the same cell type after different treatments may then be observed, and the identity of differentially expressed bands obtained by cutting-out the differentially expressed amplimer from the polyacrylamide gel, followed by cleaning and sequencing the unknown amplimer. If required, to further verify the results obtained using DD-PCR, "Northern blotting" may be performed by extracting the RNA from each of the original specimens tested and probing for the particular differentially expressed mRNA identified by DD-PCR. In this case, a probe is used based on the sequence obtained for the differentially expressed amplimer previously obtained using DD-PCR.

The selective nucleotides at the 3′-end of the reverse transcription primer, as well as the specificity of the arbitrary primer hybridization, help to ensure that the banding patterns obtained by DD-PCR are not too complex to interpret, though it is a question of *trial and error* finding the best combination of reverse transcription primer, arbitrary primer and correct DD-PCR thermocycling protocol to use that will yield an acceptable number of amplimers for investigation. Finding the best 5′-end reverse transcription and arbitrary primer combination may however be a time consuming and tedious process, and only RNA species harbouring an oligo-dA tail can be analyzed (a clear problem for researchers working with prokaryotic organisms). However, even with the "correct" reverse transcription primer, arbitrary primer, and thermocycling conditions, DD-PCR is not without its problems. If heterogeneous mixtures of cell types are used in mRNA isolation (e.g. liver versus kidney tissue) the pool of extracted mRNAs will usually be too diverse to be analyzed by DD-PCR. Therefore, it is advisable to use laboratory cultured (i.e. not

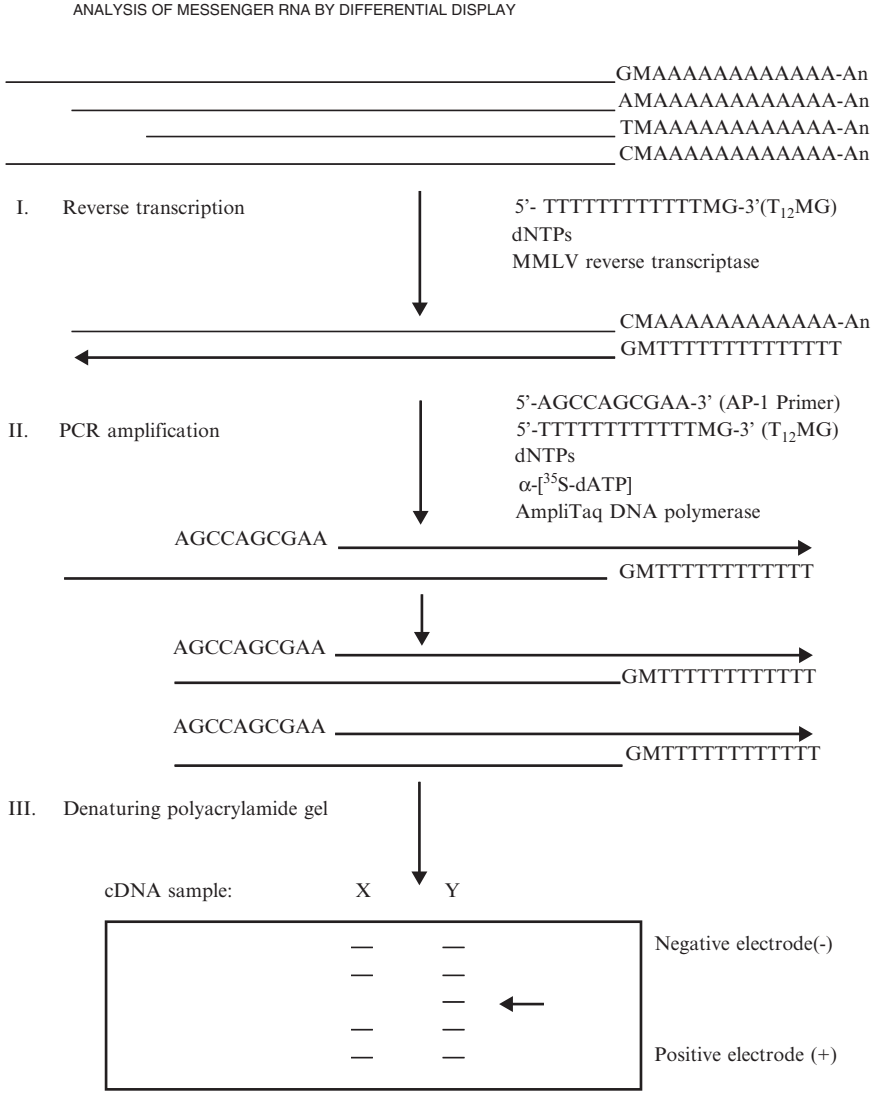

Fig. 12.13 Principle of differential display PCR. An mRNA pool is reverse transcribed into cDNA using a specially designed poly-dT primer which anneals to the 3′-end terminal poly-A tail of eukaryotic mRNA. The reverse transcribed cDNA pool is then PCR amplified using the reverse transcription primer, a randomly chosen (arbitrary) primer and labeled dNTPs. After polyacrylamide gel electrophoresis, differences in mRNA expression profiles may be observed as differences in banding patterns. The origin of these differences may then be determined by sequencing the differentially expressed band(s) (From Liang et al., 1996)

biopsy material) and growth synchronized cells for DD-PCR testing. Alternatively, it is possible to use specific cell types (e.g. tumour cells) which have been excised from paraffin embedded or air-dried coupes using micro-dissection. Finally, the DD-PCR fragments usually only encode the 3′-end non-translated region of the

mRNA molecule, rendering the identification of differentially expressed mRNA molecules by DD-PCR problematic without access to complete genomic DNA sequence databases, as translating differentially expressed DD-PCR amplimer sequences and searching in protein expression libraries may not yield the protein actually encoded by the differentially expressed mRNA transcript.

One alternative to DD-PCR is "RNA arbitrary primed-PCR" or "RAP-PCR". This technique is homologous to RAPD PCR, but uses reverse transcribed RNA/cDNA hybrids as PCR template instead of double-stranded genomic DNA. In contrast to DD-PCR, RAP-PCR utilizes a short primer of 10–20 randomly chosen nucleotides in the reverse transcription reaction, which means that problems associated with the poly-dT 3′-end mRNA orientated DD-PCR primer are avoided. In RAP-PCR, the reverse transcription primer may also be used for PCR thermocycling, with thermocycling being performed under low stringency conditions. Labelled RAP-PCR amplimers (labelled for example by incorporating radioactive dNTPs or fluorescently labelled primer in the RAP-PCR mix) may then be separated by denaturing polyacrylamide gel electrophoresis and the banding patterns visualized using for example autoradiography. An appropriate primer combination should be capable of generating at least 50 different RAP-PCR amplimers per reaction in order to adequately assess differences in gene expression.

12.15 The Protein Truncation Test (PTT)

Many inherited diseases (e.g. breast and colon cancer, muscular dystrophy, etc.) result from the accumulation of prematurely terminated proteins, which are often caused by the presence of mutations in genomic DNA that have led to the introduction of stop triplet codons (i.e. UGA, UAA or UAG) into transcribed mRNA molecules [Gite et al., 2003]. These shortened protein products may be detected using the *protein truncation test* or "PTT" [Hogervorst et al., 1995].

The principle of the PTT is shown in Fig. 12.14. The area of a gene in which the translation stop mutation is expected is PCR amplified from either reverse transcribed cDNA or genomic DNA to yield amplimers of 1–2 kb in length. PTT uses a specially designed "sense" primer (i.e. a primer which reads in the correct mRNA reading frame and can be translated directly into mRNA by DNA dependant RNA polymerases), and contains at its 5′-end: (i) a T7 RNA polymerase promoter sequence, (ii) a eukaryotic translation recognition (Kozak) sequence, and iii) an ATG "start" codon. The second PCR primer in the PTT test should be designed to contain a sequence complementary to the terminus of the gene sequence under investigation. Once PCR amplification has been performed, the resultant PTT amplimer is used as a template in a *coupled transcription and translation cell free system* (e.g. TNT Quick coupled Transcription/Translation System, Promega), which generates mRNA and protein via the 5′-end (T7 RNA polymerase recognition site) sequences included in the 5′-end PTT primers. For detection purposes, protein synthesis may be performed in the presence of radioactive amino acids or biotinylated lysine-tRNA and then detected (after SDS-PAGE gel electrophoresis)

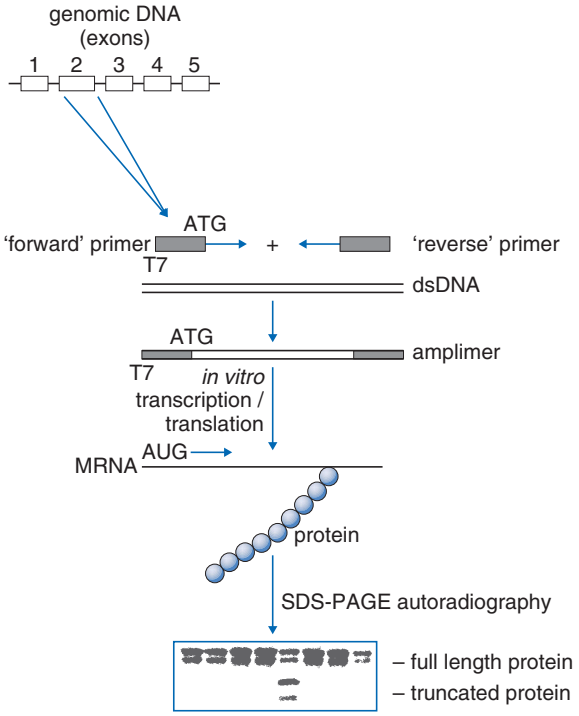

Fig. 12.14 Principle of the protein truncation test. Regions of genes to be analyzed are amplified from genomic, or reverse transcribed cDNA. The upstream primer is designed to include transcription sequences (e.g. for T7 RNA Polymerase) and eukaryotic translation signal sequences (Kozak sequences). Upon incubation of the resultant amplimer in a coupled transcription/translation cell free system (e.g. TNT Quick coupled Transcription/Translation System, Promega) the cDNA is translated back into RNA and then transcribed into protein. Full-length proteins and truncated proteins can then be detected by SDS-PAGE using autoradiography or Western blot techniques

by autoradiography or "Western" blotting using labelled anti-biotin antibodies [Roest et al., 1996]. Any truncated protein products will be visible as shorter polypeptide fragments, when compared to the expected protein fragment size after autoradiography or Western blotting.

Usually, PTT is performed on only a small part of the gene of interest (the part where most of the mutations are expected to occur), not least because eukaryotic genes are usually greater than 50–300 kb in length and the analysis of the entire 50–300 kb coding region is too complex to perform. Also, PTT does not detect genomic mutations that affect mRNA stability, mRNA processing, or mutations that do not introduce "stop" codons into the mRNA (i.e. mis-sense mutations, small insertions and short deletions).

However, PTT is a test that can potentially be adapted to high-throughput automated formats, allowing the processing of large numbers of samples in a single experiment. Also, PTT offers the ability to discriminate between clinically relevant

and clinically non-relevant mutations (as the presence of DNA sequence mutations may not actually be clinically relevant, whereas the presence of truncated proteins usually are) with good sensitivity and low false-positive rates.

12.16 Methylation Specific PCR and PCR in the Detection of Mutagens

Methylation of cytosine molecules in higher eukaryotes is a biologically important event, especially in epigenetic processes (i.e. in the developmental and organizational processes of organisms). Such methylation can bring about many effects in an organism including: (i) "gene silencing", (ii) the modulation of mRNA transcription, (iii) inactivation events on the X-sex chromosome, and (iv) occasionally the development of oncogenesis [Lehmann et al., 2003]. The precise mechanisms of methylation and/or protection of GC-rich regions against methylation are not yet fully understood, though it is known that methylation of genomic DNA can be both temporary or permanent in nature and only occurs if a cytosine nucleotide is immediately followed by guanosine (i.e. in the dinucleotide CpG). Also, methylation of cytosine bases in GC-rich sequences of promoter or enhancer regions of the genome (CpG islands) can be used to efficiently block mRNA transcription. Further, it has been shown that patterns of genomic methylation can be transferred from parents to offspring in a process referred to as "imprinting", with the dis-regulation of imprinting having been associated with several growth and behavioural defects, including Beckwith–Wiedemann syndrome, Prader–Willi syndrome, Angelman syndrome and cancer (Wilms tumour) [Reik et al., 2001].

Methylated DNA may be detected using several non-PCR techniques, including single-stranded conformational polymorphism (SSCP) analysis of native DNA to look for the presence of aberrant migration banding patterns upon gel electrophoresis. However, it is also possible to use novel PCR methodologies including methylation-specific PCR [Herman, 1996] or methylation specific AFLP.

In methylation-specific PCR, use is made of the fact that incubation of DNA containing non-methylated cytosine nucleotides for a short period of time at physiological temperature with sodium hydroxide (NaOH), renders the non-methylated cytosine nucleotides sensitive to overnight treatment at 55°C with acidic sodium bisulphite ($Na_2S_2O_3$) and hydroquinone, whereas C5-methylated cytosine nucleotides are not rendered susceptible by this treatment. For non-methylated (susceptible) cytosine nucleotides, such treatment facilitates cytosine deamidation, effectively changing the cytosine into a uracil nucleotide. The presence or absence of cytosine to uracil "metamorphoses" may then be detected using high stringency PCR conditions and a pair of sequence specific PCR primers, one of which contains either (i) a discriminatory adenosine nucleotide at its 3′-end (which will hybridize to a uracil nucleotide if present thereby indicating the original presence of non-methylated cytosine at that 3′-end position) or (ii) a discriminatory guanosine nucleotide at its 3′-end (which will hybridize to an undeaminated C5-methylated

cytosine nucleotide if present thereby indicating the original presence of methylated cytosine at that 3′-end position). Methylation specific PCR is complicated by the fact that the PCR primer that is not being used to probe the methylation site must comprise a sequence that is not affected by cytosine/uracil metamorphoses. In effect, this means that this particular primer should be designed from a region of DNA whose sequence contains no cytosine nucleotides (the primer sequence will contain no guanosine nucleotides). If multiple methylation sites are expected to be present within a short DNA sequence region (i.e. approximately 10–30 nucleotides), then guanosine or adenosine nucleotide substitutions may also be introduced into the "core" of the PCR primer as well as the 3′-end. However, using this methodology may require a large increase in the number of PCR primer combinations that have to be used. For example, to search for two methylation sites using the same primer sequence, may require primers containing 5′-nAnA–3′, 5′-nGnG–3′, 5′-nAnG–3′, and 5′-nGnA–3′ modifications (where "n" equals a given number of nucleotides), in order to probe all possible methylation combinations.

In the laboratory, methylation specific PCR may be used to track regions of inactivation on the chromosome or to make methylation charts for entire chromosomes, thereby providing clues as to the expression state of specific chromosomal regions. DNA from paraffin embedded tissues can also act as a template for methylation-sensitive PCR, though sensitivity may be improved by DNA precipitation, possibly supported by addition of carrier DNA or glycogen.

Methylation is a specific, biologically programmed nucleic acid modification mechanism. However, several chemical agents (e.g. antibiotics such as the anthracyclins) may act to non-specifically modify DNA or RNA, often resulting in nucleic acid damage or mutagenesis. The capability of such chemicals to facilitate DNA damage or mutagenesis can be assessed by observing a reduced PCR amplification yield in bacteriophage λ or P53 amplimer model systems [Jenkins and Parry, 2000]. High-throughput PCR testing of chemical mutagens is also an option, as many mutagens require co-factors (bleomycin requires Fe^{2+} ions, mitomycin C requires Cu^{2+} ions, etc.). By coupling the target DNA to 96-well microtitre plates, adding varying concentrations (or mixtures) of the suspected mutagenic agent(s) to each well, and subsequently performing high-throughput (nested) PCR and gel electrophoresis, either a decrease in PCR yield (mutagenic effect) or no decrease in yield (no mutagenic effect) may be observed. Novel generations of antibiotics may be (de-)selected on the basis of their mutagenicity using such PCR methods [Hotta, 1995].

12.17 Breakpoint PCR

An increasing number of diseases, and especially oncogenic malignancies, are associated with chromosomal translocations (reciprocal exchange of chromosomal regions). These translocations may be associated with the development and

initiation of oncogenesis and are scientifically and diagnostically interesting (e.g. the *Philadelphia chromosome* reciprocal translocation between chromosomes 9 and 22, t(9;222); chronic myeloid leukaemia, etc.). The breakpoint PCR is a simple PCR test designed to specifically amplify translocated regions of DNA. In essence, one primer of the breakpoint PCR pair is designed to anneal to a DNA sequence on the "recipient" chromosome and the other to a DNA sequence on the translocated chromosomal fragment. This means that an amplification product is obtained only when the (correct) translocation is present in the target DNA. A primer pair designed to amplify another (unaffected) region of DNA on the recipient chromosome may be used as an amplification control and included in the same breakpoint PCR mix resulting in a "multiplex breakpoint PCR" protocol. Since the regions where chromosomal breakage and translocations occur (the breakpoints) are not always exactly identical, breakpoint PCR primers may need to be spaced in an appropriate fashion so that the entire catalogue of already identified translocation sites associated with a particular disease are covered. However, this could mean that a "long and accurate" breakpoint PCR thermocycling methodology may need to be performed in order to generate long (5–30 kb) PCR amplimers. Breakpoint PCRs work well for mixtures of translocated cells in a background of healthy cells, since only the translocated cells will generate amplification products. If chromosomal translocations are accompanied by extensive chromosomal deletions, then the breakpoint PCR may yield a "false negative" result, in which case the use of comparative genomic hybridization (CGH), or array-CGH, may be a more appropriate diagnostic technique [Verhagen et al., 2000]. CGH may also be a better option if gene duplications during translocation are also usually observed.

12.18 Site Directed Mutagenesis by PCR

The selective introduction of nucleotide changes (mutations) within a known DNA sequence is a process called site-directed mutagenesis or SDM. The goal of such SDM may include (i) generating new restriction digestion sites, (ii) generating cloning sites, (iii) to change the sequence of a transcribed mRNA molecule so that a novel protein is translated, (iv) to introduce stop codons into DNA, (v) to help make fusion genes or vi) to study the effects of mutation events on the efficiency of promoters. SDM using PCR may be utilized to generate mutations by both nucleotide deletion and/or nucleotide substitution mechanisms.

Short nucleotide deletions may be introduced into DNA using a "degenerate" PCR primer, which at its 3′-end is complementary to the DNA sequence to be mutated, but lacks the nucleotides to be deleted near its 5′-end. After PCR amplification, the products will lack the deleted nucleotides. Due to the degenerate nature of the primer sequence (i.e. it will not perfectly hybridize to the intended target DNA sequence), the T_m of thermocycling must be lowered during the first few cycles of PCR in order to allow primer hybridization.

Degenerate primers may also be used to introduce nucleotide substitutions into specific DNA sequences, however the use of degenerate primers in a standard SDM deletion protocol (as mentioned above), would limit nucleotide substitutions to the outer limits (5′-ends) of the DNA to be amplified (dependent on the length of the degenerate primer to be used). Therefore, to facilitate nucleotide substitutions near the centre of a particular gene sequence, two separate PCRs need to be performed, with the two corresponding amplimers overlapping in the region where the mutation is to be introduced (see Fig. 12.15). Post-amplification, the two amplimers are purified away from excess dNTPs, enzyme, etc., and added to a single new PCR mix without PCR primers. The mix is then heated to denature the two amplimers and the temperature will be allowed to decrease so that cross-hybridization of the amplimers with each other may occur. Resultant hybrid molecules with single stranded ends extending beyond the common hybridization region will be filled in by the thermostable DNA polymerase present in the reaction mix, generating an elongated double-stranded template. The addition of 5′-end and 3′-end primers designed to hybridize to the termini of the new elongated DNA molecule, followed by PCR thermocycling, leads to amplification of the new elongated hybrid DNA construct, which may then for example, be ligated into a cloning vector for eventual protein expression studies or used to transform recipient cells, etc.

12.19 PCR Amplimers for Cloning and Expression

PCR is a powerful method for synthesizing large amounts of specific DNA fragments or complete genes (with or without their corresponding promoter regions). Further, the design of PCR primers with added 5′-end restriction enzyme recognition sequences followed by restriction enzyme digestion and ligation has allowed the cloning of such amplimers into vectors that contain identical restriction enzyme recognition sequences. However, PCR applications are constantly changing and improving with regard to speed, throughput and cost. For example, the introduction of "TA-cloning kits" (TA Cloning kit, Invitrogen) have been the greatest advance in PCR amplimer cloning and expression in recent years. Such kits include both special enzyme mixes and linear vectors containing 3′-end thymidine nucleotide "overhangs", and make use of a property of the majority of thermostable enzymes (including Taq but not the proofreading enzyme Pfu) to add an extra adenosine (or other) nucleotide to the 3′-end of PCR amplification products. Ligation of PCR amplimers to a special TA cloning plasmid than allows rapid insertion of the PCR product into a cloning/sequencing vector, without the need for PCR primers containing restriction enzyme recognition sites and multiple digestion and ligation experiments. Such advances in PCR cloning applications means that the construction of cumbersome random genomic or random cDNA libraries, along with labour intensive library screening, restriction mapping and sub-cloning, may now be avoided. The only requirement however, is that at least part of the sequence of the region to be amplified is known (in order to design suitable PCR primers). Once cloned into a relevant vector, PCR amplimers may be expressed in either prokaryotic or eukaryotic cells

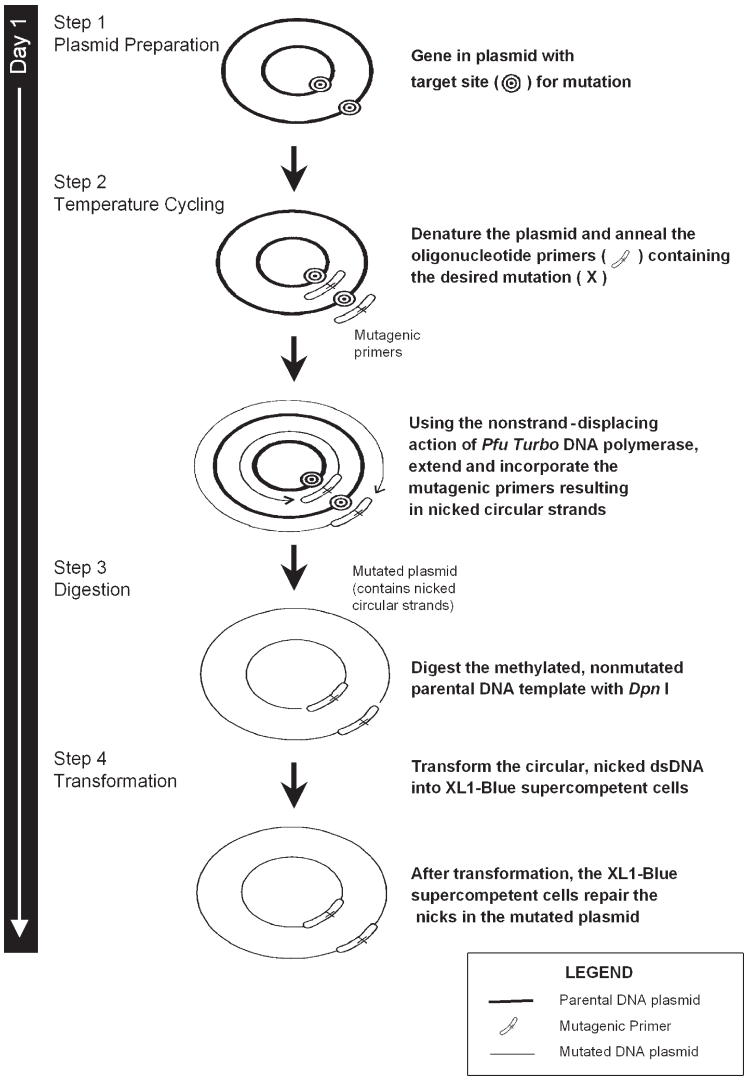

Fig. 12.15 Site directed mutagenesis by PCR. The target sequence is copied into two overlapping fragments containing the required nucleotide substitution(s) which are then co-amplified in a second PCR. In this way, nucleotide substitutions (mutagenesis) may be introduced into the target DNA sequence (www.stratagene.com/manuals)

(including bacterial, yeast, insect and mammalian cells). This depends on the expression system and promoter used. For this purpose, many commercially available expression vectors have been developed, e.g. the CheckMate Mammalian Two-Hybrid System and the the pALTER-MAX system for use in bacterial cells (Promega), and the BaculoDirect Expression System (Invitrogen) (Fig. 12.16). Alternatively, a

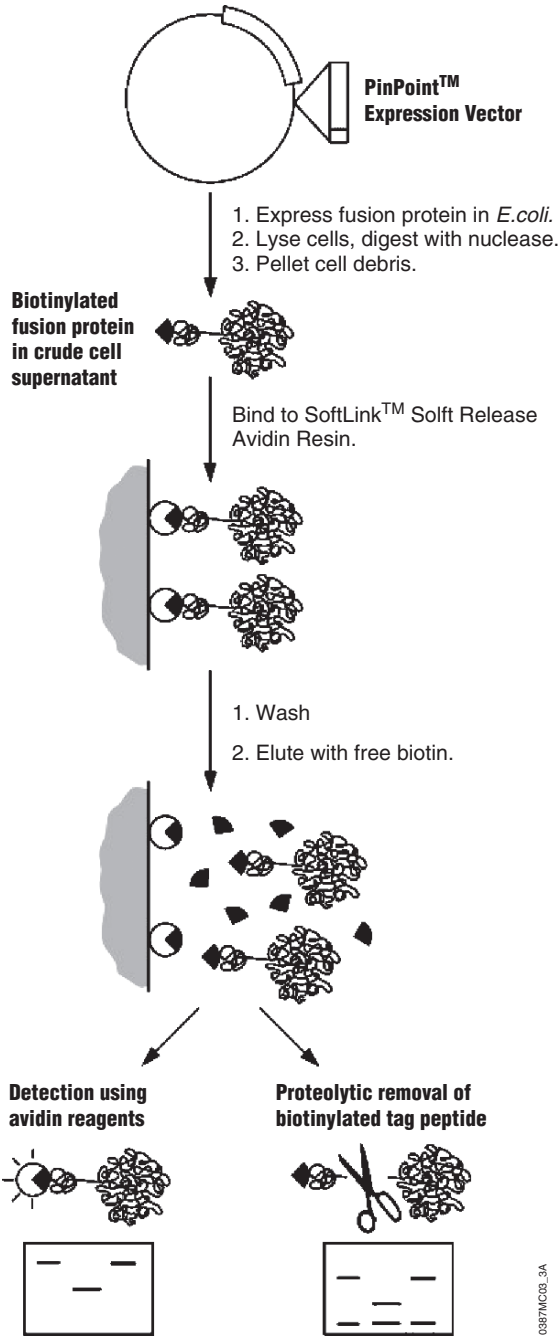

Fig. 12.16 The cloning of an amplified gene by PCR into a commercially available vector allows expression of the encoded protein. In this example, the vector has been designed to generate a fusion protein allowing easier isolation and manipulation of expressed protein (Promega Protocols)

gene may be PCR amplified and cloned into a vector which puts it under the control of a set of regulatory elements allowing for cell free lysate coupled transcription and translation, indeed some commercially available coupled transcription and translation systems are specifically designed for protein expression from the appropriate PCR products, e.g. TNT T7 quick for PCR system (Promega).

Large quantities of protein may also be generated from PCR products cloned into specific vectors by incorporating for example, a thioredoxin tag (to increase expressed protein solubility) in the vector, e.g. the His-Patch ThioFusion expression System (Invitrogen), whilst polyhistidine (a sequence that encodes 6–8 histidine residues) encoding vectors allow the purification of PCR amplimer generated proteins using nickel chelate resins in the form of columns or plates (Sigma). Alternatively, in the Pinpoint Vector and Expression Systems (Promega), the PCR product is cloned into a vector and synthesized as a fusion protein coupled to a 22 kDa polypeptide that is naturally biotinylated in *E. coli*. This means that the final fusion protein can be immediately coupled to streptavidin-coated beads for further purification or immediately used in downstream processes such as coating ELISA plates for sero-prevalence studies. For detection purposes, PCR amplimer cloning vectors are available which allow PCR amplimers to be expressed as fusion proteins containing such easily detectable "tags" as β-galactosidase, chloramphenicol acetyltransferase (CAT), enhanced green fluorescent protein (EGFP), luciferase and glutathione-s-reductase (GST).

12.20 SAGE

SAGE or "Serial Analysis of Gene Expression" (Fig. 12.17) is a method for mapping the expression of genes by constructing multiple repeats (of approximately 100 bp in length) of small (approximately 10 bp) cDNA "Tags" [Velculescu et al., 1995; Velculescu et al., 2000]. In the SAGE format, mRNA (representing all of the genes that were being expressed at the time of mRNA isolation) is extracted from cells and attached to a solid support. Next, cDNA synthesis is performed using oligo-dT or random hexamer primers, yielding double-stranded cDNA copies of the original mRNA transcripts available. These double-stranded cDNAs are then digested using the restriction enzyme *Nla*III generating a GTAC overhang. Two specific DNA adapters containing a recognition sequence for a second restriction enzyme (usually *BsmF*I), are then ligated to one half of the *Nla*III digested cDNA (one adapter per half), and further digested using *BsmF*I. As *BsmF*I cuts 10–14 bp downstream of the recognition sequence, small Tag/adapter products are liberated from the solid support. Next, the 5′-ends of the *BsmF*I digested products are "filled in" and products from each of the two adapters are pooled and ligated to form a "ditag". PCR amplification is then performed using the two adapter sequencers as primers and the PCR products again digested using *Nla*III. The "sticky ended" PCR amplimers are then ligated into long concatemers comprising multiple copies of the same ditag. These concatemers are then cloned

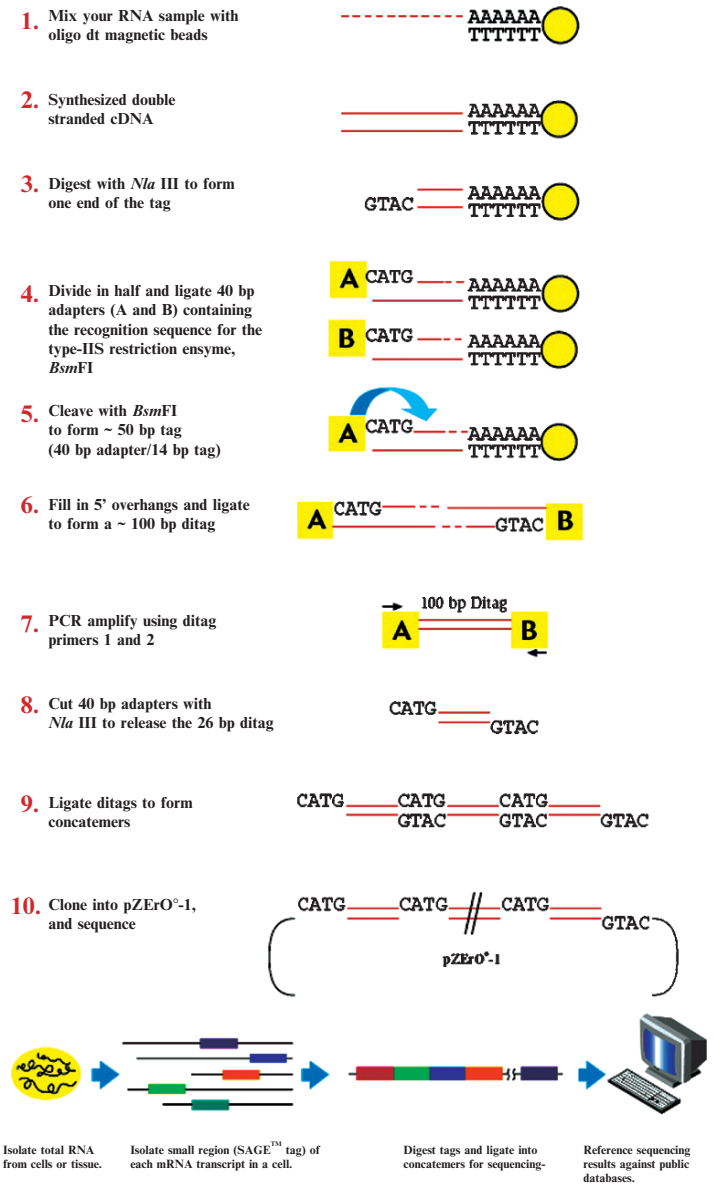

Fig. 12.17 Schematic presentation of the various steps of the Serial Analysis of Gene Expression (SAGE) protocol (see text in this chapter for more precise written explanation and www.invitrogen.com for more detail on the I-SAGE kit for Genome Wide Expression Analysis) (Copied from Invitrogen) (*see Color Plates*)

into a plasmid vector, sequenced, and the 10 bp tag compared to currently available databases (e.g. SAGEMAP at www.ncbi.nlm.nih.gov/projects/SAGE/) [Lash et al., 2000; Lal et al., 1999]. Such tags usually contain enough sequence information to reliably identify the source of the original complete mRNA transcript, and

the level of gene expression is indicated by the frequency of occurrence of a particular 10 bp tag. Commercial kits are currently available to perform this technique, e.g. I-SAGE Longkit (Invitrogen).

12.21 PCR Inhibition by DNA Specific Antibiotics and Mutagens

There exists a large multitude of agents capable of selectively damaging double stranded DNA. The degree of damage can be reflected in the decrease in the amount of the PCR product formed. The damaging agents can be found in food products and product formed during incineration of wood or charcoal. In addition, many chemotherapeutic agents do have DNA-mutagenic characteristics. This includes various antibiotics as well. Quality control of novel medications does include assessment of the compounds' mutagenic potential. DNA damage can be screened at high throughput using 384-wells ELISA plates. The wells contain a fixed amount of a biotinylated PCR product and these are subjected to a variety of (combinations of) chemicals to be explored. A subsequent quantitative nested PCR indicates the degree of DNA damage incurred by the chemicals. The same assay, of course, may also be used to identify compounds with anti-mutagenic effects. When testing antibiotics the effect of metal co-factors needs to be considered. Mitomycin, for instance, requires divalent copper ions for optimal activity.

References

Ailenberg M, Silverman M. 2000. Controlled hot start and improved specificity in carrying out PCR utilizing touch up and loop incorporated primers (TULIPS). Biotechniques 29:1018–1024.

Alcock HE, Stephenson TJ, Royds JA, Hammond DW. 2003. Analysis of colorectal tumour progression by micro-dissection and comparative genomic hybridization. Genes Chromosome Cancer 37:369–380.

Benecke M. 1997. DNA typing in forensic medicine and in criminal investigations: a current survey. Naturwissenschaften 84:181–188.

Berg H, Ziegler K, Piotukh K, Baier K, Lockau W, Volkmer-Engert R. 2000. Biosynthesis of the cyanobacterial reserve polymer multi-L-arginyl-poly-L-aspartic acid (cyanophycin): mechanism of the cyanophycin synthetase reaction studied with synthetic primers. Eur J Biochem 267:5561–5570.

Boyer JC, Farber RA. 1998. Mutation rate of a microsatellite sequence in normal human fibroblasts. Cancer Res 58:3946–3949.

Castles CG, Allred DC, Krieg SL, Benedix MG, Fuqua SA. 1996. RNA from air dried frozen sections for RT-PCR and differential display. Biotechniques 21:425–428.

Chevalier J, Yi J, Michel D, Tang XM. 1997. Biotin and digoxigenin as labels for light and electronmicroscopy in situ hybridization probes: where do we stand? J Histochem Cytochem 45:451–471.

Chou Q, Russell M. Birch DE, Raymond J, Bloch W. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy number amplifications. Nucleic Acids Res 20:1717–1723.

Garcia J, Castrillo JL. 2004. Differential display RT-PCR analysis of human choriocarcinoma cell lines and normal term trophoblast cells: identification of new genes expressed in placenta. Placenta 25:684–693.

Gite S, Lim M, Carlson R, Olejnik J, Zehnbauer B, Rothschild K. 2003. A high throughput nonisotopic protein truncation test. Nat Biotechnol 21:194–197.

Gyllenstein U. 1989. Direct sequencing of in vitro amplified DNA. In: Erlich H (ed) PCR technology. pp. 45–61.

Hawkins GA, Hoffman LM. 1999. Rapid DNA mutation identification and fingerprinting using base excission sequence scanning. Electrophoresis 20:1171–1176.

Hebert NE, Brazill SA. 2004. Microchip capillary gel electrophoresis with electrochemical detection for the analysis of known SNPs. Lab Chip 3:241–247.

Hecker KH, Roux KH. 1996. High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. Biotechniques 20:478–485.

Herman JG, Graf JR, Myohanen S, Nelkin BD, Baylin SB. 1996. Methylation specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci USA 93:9821–9826.

Higuchi RG, Ochman H. 1989. Production of single stranded DNA templates by exonuclease digestion following the PCR. Nucleic Acids Res 17:5865.

Hogervorst FB, Cornelis RS, Bout M, Van Vliet M, Oosterwijk JC, Olmer R, Bakker B, Klijn JG, Vasen HF, Meijers-Heijboer H. 1995. Rapid detection of BRCA1 mutations by the protein truncation test. Nat Genet 10:208–212.

Hotta K, Zhu CB, Phomsuwansiri P, Ishikawa J, Mizuno S, Hatsu M, Imai S. 1995. PCR inhibition assay for DNA-targeted antibiotics. J Antibiot 48:1267–1272.

Jenkins GJ, Parry JM. 2000. PCR-based methodology for the determination of the photoprotection afforded by sunscreen application. Biotechniques 29:1318–1326.

Joos S, Schutz B, Bentz M, Lichter P. Roche product manual "ISH to whole chromosomes". In: Procedures for In Situ Hybridization to Chromosomes, Cells, and Tissue Sections; Detection of Chromosomal Imbalances using DOP–PCR and Comparative Genomic Hybridization (CGH). p95–103. Also available at https://www.roche-applied-science.com/fst/publications.jsp? page=/PROD_INF/MANUALS/InSitu/InSi_toc.htm

Jouquand S, Cheron A, Galibert F. 1999. Microsatellite analysis using a two step procedure for fluorescence labeling of PCR products. Biotechniques 26:902–905.

Kaboev OK, Luchkina LA, Tretíakov AN, Bahrmand AR. 2000. PCR hot start using primers with the structure of molecular beacons (hairpin-like structures). Nucleic Acids Res 28:94.

Kainz P. Schmiedlechner A, Strack HB. 2000. Specificity-enhanced hot-start PCR: addition of double-stranded DNA fragments adapted to the annealing temperature. Biotechniques 28(2):278–282.

Kebelmann-Betzing C, Seeger K, Dragon S, Schmitt G, Moricke A, Schild TA, Henze G, Beyermann B. 1998. Advantages of new Taq DNA polymerase in multiplex PCR and time release PCR. Biotechniques 24:154–158.

Kellogg DE, Rybalkin I, Chen S, Mukhamedova N, Vlasik T, Siebert PD, Chenchik A. 1994. TaqStart antibody: hot start PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase. Biotechniques 16:1134–1137.

Kruhoffer M, Jensen JL, Laiho P, Dyrskjot L, Salovaara R, Arango D, Birkenkamp-Demtroder K, Sorensen FB, Christensen LL, Buhl L, Mecklin J, Jarvinen H, Thykjaer T, Wikman FP, Bech-Knudsen F, Juhola M, Nupponen NN, Laurberg S, Andersen CL, Aaltonen LA, Orntoft TF. 2005. Gene expression signatures for colorectal cancer microsatellite status and HNPCC. Br J Cancer 92:2240–2248.

Kutyavin IV, Afonina IA, Mills A, Gorn VV, Lukhtanova EA, Belousov ES, Singer MJ, Walburger DK, Lokhov SG, Gal AA, Dempcy R, Reed MW, Meyer RB, Hedgpeth J. 2000. 3′-minor groove binder DNA probes increase sequence specificity at PCR extension temperatures. Nucleic Acids Res 28:655–661.

Lal A, Lash AE, Altschul SF, Velculescu V, Zhang L, McLendron RE, Marra MA, Prange C, Morin PJ, Polyak K, Papadopoulos M, Vogelstein B, Kinzler KW, Strausberg RL, Riggins GJ. 1999. A public database for gene expression in human cancers. Cancer Res 59:5403–5407.

Lash AE, Tolstoshev CM, Wagner L, Schuler GD, Strausberg LR, Riggins GJ, Altschul SF. 2000. SAGEmap: a public gene expression resource. Genome Res 10:1051–1060.

Leclair B, Scholl T. 2005. Application of automation and information systems to forensic genetic specimen processing. Expert Rev Mol Diagn 5:241–250.

Lehmann K, Seemann P, Stricker S, Sammar M, Meyer B, Suring K, Majewski F, Tinschert S, Grzesik KH, Muller D, Knaus P, Nurnberg P, Mundlos S. 2003. Mutations in bone morphogenetic protei receptor 1B cause brachydactyly type A2. Proc Natl Acad Sci USA 100:12277–12282.

Liang P. 1996. Analysis of messenger RNA by differential display. In: Krieg PA (ed) A laboratory guide to RNA. Wiley, New York, pp. 224–235.

Lin Y, Jayasena SD. 1997. Inhibition of multiple thermostable DNA polymerases by a heterodimeric aptamer. J Mol Biol 27:100–111.

Liu WH, Kaur M, Wang G, Zhu P, Zhang Y, Makrigiorgos GM. 2004. Inverse PCR based RFLP scanning identifies low level mutation signatures in colon cells and tumours. Cancer Res 64:2544–2551.

Maxam AM, Gilbert W. 1977. A new method for sequencing DNA. Proc Natl Acad Sci USA 74:560–564.

Mizuguchi H, Nakatsuji M, Fujiwara S, Takagi M, Imanaka T. 1999. Characterization and application to hot start PCR of neutralizing monoclonal antibodies against KOD DNA polymerase. J Biochem 126:762–768.

Moretti T, Koons B, Budowle B. 1998. Enhancement of PCR amplification yield and specificity using Amplitaq Gold DNA polymerase. Biotechniques 25:716–722.

Mullis KB. 1991. The PCR in an anemic mode: how to avoid cold oligodeoxyribonuclear fusion. PCR Methods Appl 1:1–4.

Nazarenko I, Lowe B, Darfler M, Ikonomi P, Schuster D, Rashtchian A. 2002. Multiplex quantitative PCR using self quenched primers labeled with a single fluorophore. Nucleic Acids Res 30:37.

Petitjean J, Vabret A, Dina J, Gouarin S, Freymuth F. 2005. Development and evaluation of a real time RT-PCR assay on the LightCycler for the rapid detection of enterovirus in cerebrospinal fluid specimens. J Clin Virol 5:1271–1282.

Reik W, Dean W, Walter J. 2001. Epigenetic reprogramming in mammalian development. Science 293:1089–1093.

Roest PA, Bout M, van der Tuijn AC, Ginjaar IB, Bakker E, Hogervorst FB, van Ommen GJ, den Dunnen JT. 1996. Splicing mutations in DMD/BMD detected by RT-PCR/PTT: detection of a 19AA insertion in the cysteine rich domain of dystrophin compatible with BMD J Med Genet 33(11):9350939.

Sanger F, Nicklen S, Coulson AR. 1992. DNA sequencing with chain-terminating inhibitors. Biotechnology 24:104–108.

Singh A, Gaur A, Shailaja K, Satyare BB, Singh L. 2004. A novel microsatellite (STR) marker for forensic identification of big cats in India. Forensic Sci Int 141:143–147.

Sonoki T, Willis TG, Oscier DG, Karran EL, Siebert R, Dyer MJ. 2004. Rapid amplification of immunoglobulin heavy chain switch (IGHS) translocation breakpoints using long-distance PCR. Leukemia 18:2026–2031.

The year of sequencing. 2008. Nature Methods 5(1):11–21.

Torpdahl M, Skov MN, Sandvang D, Baggesen DL. 2005. Genotypic characterization of *Salmonella* by multilocus sequence typing, pulsed field gel electrophoresis and amplified fragment length polymorphism. J Microbiol Methods 63:173–184.

Van Belkum A. 1994. DNA fingerprinting of medically important microorganisms by use of PCR. Clin Microbiol Rev 7:174–184.

Van Belkum A. 2002. Molecular typing of microorganisms: at the centre of diagnostics, genomics and pathogenesis of infectious diseases? J Med Microbiol 51:7–10.

Van Belkum A, Scherer S, Van Alphen L, Verbrugh H. 1998. Short sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev 62(2):275–293.

Velculescu VE, Vogelstein B, Kinzler KW. 2000. Analysing uncharted transcriptomes with SAGE. Trends Genet 16:423–425.

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. 1995. Serial analysis of gene expression. Science 270:484–487.

Verhagen PC, Zhu XL, Rohr LR, Cannon-Albright LA, Tavtigian SV, Skolnick MH, Brothman AR. 2000. Microdissection, DOP PCR and comparative genomic hybridization of paraffin-embedded familial prostate cancers. Cancer Genet Cytogenet 122:43–48.

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Van Montagu M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414.

Wu WM, Tsai H J, Pang JH, Wang HS, Hong HS, Lee YS. 2005 Touchdown thermocycling program enables a robust single nucleotide polymorphism typing method based on allele-specific real-time polymerase chain reaction. Anal Biochem 339(2):290–296.