



Crop Protection Compendium

Analytical Methods

Text by CABI Bioscience UK Centre (Egham), UK

[DNA Extraction](#) . [RFLPs](#) . [DNA Fingerprinting](#) . [Repetitive DNA](#) . [PCR](#) . [PCR for Detection and Identification](#) . [PCR as a Taxonomic Tool](#) . [RAPDs](#) . [PCR Quality Control](#) . [Chromatography](#) . [Fatty Acid Analysis](#) . [PAGE and Protein Electrophoresis](#) . [PFGE](#) . [Ubiquinones](#)

DNA Extraction

A number of different procedures are available for the extraction of DNA from cells. In general these methods follow three basic steps or principles. The plethora of different techniques arises from the desire for a protocol that is quick, safe and efficient at producing DNA in a highly pure form. Problems can arise from a number of different sources: inability to break open the cells; degradation of the extracted DNA as a result of the co-extraction of nucleases; or co-isolation of proteins or high-molecular weight polysaccharides which can interfere with subsequent reactions. The three main stages generally found in DNA extraction protocols for bacteria and fungi are listed below:

1. Breakdown of cell wall and release of cell contents

Most Gram-negative bacteria can be lysed with 2% (w/v) sodium dodecylsulphate (SDS; also known as sodium laurylsulphate) treatment, whilst Gram-positive bacteria will require enzymic digestion using lysozyme, proteinase K and/or N-acetylmuramidase before adding SDS. Some Gram-positive bacteria appear to be resistant to lysozyme so the use of a lipase in combination with lysozyme may be more effective at removing the cell wall. Physical treatments such as sonication, shaking with glass beads, or freeze-thawing can also aid recovery. In general, most extractions from fungal biomass rely on physical treatments. Biomass can be freeze-dried then ground in a mortar and pestle, frequently with the addition of liquid nitrogen to disrupt the cell wall. Further refinements have been added to this step to aid the subsequent recovery of DNA: addition of 2-mercaptoethanol, which is frequently used to inactivate nucleases and is generally used in fungal procedures though it also appears to be effective with prokaryotes; the use of proteinase K to aid digestion of proteins; the use of cetyltrimethyl ammonium bromide (CTAB) for the removal of

polysaccharide-like contaminants. With this last point it should be noted that CTAB cannot be used in conjunction with SDS or phenol as these mixtures will form insoluble complexes.

2. Preparation of crude DNA extract

This is generally achieved by the addition of phenol or phenol-chloroform to the suspension containing the disrupted microbial biomass. Tubes containing the mixture are centrifuged, resulting in a biphasic separation. The upper aqueous phase contains the nucleic acids whilst the lower phase contains the phenol; there is generally a white semi-solid layer separating the two phases mainly made up of denatured proteinaceous materials. The upper aqueous phase is removed and the nucleic acids can be recovered by the subsequent precipitation using ice-cold ethanol. The process can often be enhanced by incubating at -20°C overnight.

3. Purification of DNA

In this step the nucleic acids are re-dissolved by the addition of a suitable buffer and gentle shaking, then the RNA is removed by the addition of RNase treatment. In general two enzymes such as RNase A and RNase T1 used together will eliminate all RNA from the material. Generally chloroform-isoamyl alcohol mixture is added, the solution shaken then centrifuged and the upper aqueous phase removed and the DNA precipitated again using ice-cold ethanol.

Standard molecular biology methods are well documented in a number of standard texts. Many different DNA extraction methods have been published for filamentous fungi and bacteria. Many of the chemicals and reagents used in DNA extraction are listed as toxic, harmful, irritant or corrosive. Appropriate safety data sheets should be consulted before undertaking any procedures.

It is not always necessary to obtain either large amounts of DNA or very pure DNA, particularly if it is required for amplification with specific primers (see **RAPDs**). In such cases it is often sufficient to obtain a small amount of partially purified DNA.

RFLPs

Restriction fragment length polymorphisms (RFLPs) are generated by digesting DNA with restriction enzymes, which 'cut' DNA at specific sites. These sites are generally determined by the sequence of particular bases: for example, the restriction enzyme EcoR 1 cuts DNA wherever the sequence GAATTC occurs. If the sequence occurs only once in the total DNA strand, the DNA will be cut into two fragments; however, if the sequence occurs many times, many fragments will be produced. Different enzymes are specific for different sites, therefore a large number of possible fragments may be derived from any given piece of DNA.

Chromosomal DNA generally yields many hundreds of fragments when digested with common restriction enzymes. Separating these fragments by electrophoresis results in a smear of DNA in the gel. Fragments that are present in multiple copies, such as genes for ribosomal RNA, may stand out as distinct bands, but in general little information can be

gained from restriction digestion on its own. It is usual to then hybridize probe DNA to the smear in order to obtain a smaller number of distinct bands that can give interpretable information.

A probe is a DNA fragment that can be expected to show homology to, and to hybridize with, the DNA of the organism being studied. It may have been generated from a closely related organism, or from an entirely different source. The probe DNA used in RFLPs is generally one of two types, either consisting of sequences from a particular gene or genes, or consisting of frequently repeated sequences. RFLPs with probes derived from particular genes will show differences in the sequence and arrangement of the regions around the conserved genes in the organisms studied, for example mitochondrial probes or probes for ribosomal RNA genes. These probes have been used in determining species and below-species relationships. Probes derived from commonly occurring repeated sequences may generate many bands in RFLPs. Examples of these include the mini-satellite repeat probes used in genetic fingerprinting, the repeat sequences from the phage M13 and simple repeated oligonucleotide sequences. In general these probes have been used to show lower levels of relatedness than probes derived from genes, and have been used in strain fingerprinting and population studies.

To generate RFLPs, it is necessary to label the probe so that it can be detected, and to transfer the smear of digested DNA to a medium suitable for the subsequent hybridization and detection. The general method for this transfer is by capillary action where the DNA is transferred to a nylon or nitrocellulose membrane by blotting; this membrane is then used in the hybridization reactions.

Radionucleotides have often been used to label probes: hybridization can be detected by exposure to X-ray film. This is a very sensitive system but can involve exposure times of some days for final detection. In recent years there has been considerable interest in the use of chemically labelled probes, and systems for labelling probes with biotin or dioxygenin are available commercially.

Probes are only required for RFLPs when a specific region of DNA is studied in a larger sample. RFLPs can be generated without probes where only a small number of bands are generated from the total sample. Three examples of this are polymerase chain reaction products ([PCR](#)), mitochondrial DNA and rare-cutting enzymes (see [Pulsed Field Gel Electrophoresis](#)).

RFLPs have been widely used at different taxonomic levels in fungal and bacterial systematics. A large selection of both mitochondrial and chromosomally derived probes have been used, including those derived from the ribosomal RNA gene complex, and these have yielded important information on relationships both within, and between species.

DNA Fingerprinting

Micro-organisms can be characterized using a variety of techniques that depend on the analysis of nucleic acids, generally referred to as molecular profiling or fingerprinting

techniques. These methodologies are reliant on the fact that closely related organisms will share common sequences, whereas more distantly related organisms will show wide differences. These techniques can exploit differences either by using restriction endonucleases to generate fragments that can then be separated by gel electrophoresis (restriction length polymorphic DNA analysis; also see [RFLP](#) and [PFGE](#)), or by the use of PCR primers to generate DNA fragments that can also be separated by gel electrophoresis to produce a characteristic fingerprint (e.g. [RAPDs](#)), ERICs, BOX, REP; see also [Repetitive DNA](#)), or a combination of both approaches. PCR-based fingerprinting methods offer some advantages over other nucleic-acid-based methods for strain characterization. These include speed, the small amount of material required and ease of performance.

Repetitive DNA

Characterization of bacterial genomes has led to the recognition of repeated DNA sequences that are conserved between diverse bacteria, particularly Gram-negative bacteria. Three families of repetitive DNA have been recognized that are unrelated at the DNA sequence level. These are referred to as: repetitive extragenic palindromic (REP) sequences; enterobacterial repetitive intergenic consensus (ERIC) sequences; and BOX element. Their functional significance is apparently related to the structural organization of the bacterial genome because REP, ERIC and BOX have the potential to form stem-loop structures.

Nucleotide sequence determination of these repetitive regions has enabled the design of PCR primers that are specific to each region. Accordingly, PCR with these primers gives rise to amplification products that reflect the number and distribution of repetitive sequences. This approach to genomic fingerprinting is referred to as 'rep-PCR' and offers a highly sensitive level of analysis, suitable for species, pathovar and, to a lesser extent, isolate identification.

PCR

The polymerase chain reaction (PCR), first described in the mid-1980s, has become a powerful technique for the selective amplification of DNA or RNA sequences. Numerous variations on the basic technique have now been developed, and a number of different applications are available.

PCR is a method for the production in vitro of large amounts of DNA from small amounts of a target DNA sequence. The process is analogous to the 'natural' replication of DNA as it occurs in living cells. The two strands of a parent DNA molecule are 'unzipped' to form single strands, which then act as templates for the synthesis of new daughter strands. These daughter strands are synthesized following the usual base-pairing rules: adenine (A) pairs with thymine (T), and cytosine (C) pairs with guanine (G). In this way two exact copies of double-stranded DNA are formed from one. This semi-conservative process always occurs in the 5' to 3' direction. In nature, DNA replication is reliant on the presence of a number of different enzymes and molecules, including the enzyme DNA polymerase, deoxyribonucleotides (dATP, dCTP, dGTP and dTTP) and RNA primer. The primer is required because DNA polymerase cannot start synthesis de novo and can only add nucleotides to a free hydroxyl group on the third carbon atom of deoxyribose or ribose. The primer anneals

with the parent template and provides a 3'-OH group, thus allowing polymerization to occur.

Amplification of DNA in vitro by PCR is based on the same principles of DNA replication. Target DNA, DNA polymerase, deoxyribonucleotides and primer DNA are mixed in an appropriate buffer. Primers consist of short single-stranded DNA sequences, commonly 10-20 bases long. The double-stranded target DNA is first unwound to produce single-stranded templates; this is usually achieved by heating. Primers then anneal to the templates and daughter strands are synthesized in the 5' to 3' direction. In essence, enzymatic amplification occurs of a DNA fragment that is flanked by two primers which anneal to opposite strands of the target sequence. The primers are orientated with their 3' ends pointing towards each other. Repeated cycles of heat denaturation of the template (to form single-stranded DNA), annealing the primers to their complementary sequences and extension of the annealed primers with DNA polymerase results in the amplification of the segment defined by the 5' ends of the PCR primers. The short, discrete, primer-terminated copies, which first appear in the second cycle, proceed to double with each subsequent cycle and rapidly become the predominant form of amplification product (e.g. one million copies of each target strand would be made in 20 cycles). The use of a thermostable polymerase such as Taq or Tth means that fresh polymerase does not have to be added after each heat denaturation step. One of the keys to successful PCR is primer design.

The exact conditions used for an amplification vary according to the specific application and usually require some optimization. Melting of the DNA is achieved by heating to ca 95°C. The temperature used for primer annealing is dictated by the base composition of the primers used (for each primer multiply the number of A/Ts by 2 and the C/Gs by 4 and add the total), however, this often requires optimization. Short primers, such as those used in RAPDs, have annealing temperatures around 35°C, whereas longer primers used for other applications typically have higher annealing temperatures of 50-70°C. Primers will anneal above and below their theoretical melting temperature but, in general, using a higher annealing temperature increases the strictness or 'stringency' of annealing and reduces the chances of artefactual PCR products. The extension phase is typically carried out at 72°C. The duration of each phase also requires optimization depending on the specific application and may vary from 30 seconds to several minutes. The number of cycles is usually 25-35. Amplifications usually end with a long extension phase (72°C) in order to ensure that the amplification products are fully extended.

PCR is a highly sensitive procedure and it has been shown that as little as one molecule of DNA can serve as a template for amplification. Because of the extreme sensitivity of the technique, contamination is one of the major problems of PCR use. If the sample to be amplified contains DNA from a foreign source, false positives or misleading results can easily be generated. This can only be completely avoided by spatially isolating every stage of the process to avoid possible sources of contamination. However, this is rarely feasible and a compromise must be arrived at. It is possible to avoid major reorganizations in laboratory practice by adhering to some simple routines: for example, good aseptic technique and laboratory hygiene; the use of molecular biology grade chemicals (including HPLC-grade water that has been filtered and autoclaved); a master-mix of PCR reagents can be prepared in advance and then the correct volume dispensed into each reaction tube to avoid pipetting errors in the preparation of reaction mixtures.

PCR for Detection and Identification

In the fields of plant pathology and plant quarantine it is necessary to identify the causal agent of a disease unambiguously, rapidly and at a level of infection that is not visually apparent. Application of the PCR reaction with primers that have been selected for specificity to the taxon of interest (species, genus etc.) can address these aspects.

Primers selected for identification are designed using DNA nucleotide sequence data that is unique to the taxon of interest and are typically greater than 20 bases in length. The amplified region may encode for a particular function: for example, a pathogenicity gene or a reporter gene, such as the *npt* gene common to many genetically modified organisms, or ribosomal genes.

Ribosomal genes are a particularly appropriate target for PCR-directed identification, as the genes occur in high copy numbers, are highly conserved and are flanked by spacer regions that contain comparatively variable sequences. DNA sequence data on ribosomal genes can be obtained by PCR with broad-range primers (universal primers) that anneal to the highly conserved ribosomal gene sequences and amplify across regions that contain nucleotide variation. Various sets of universal primers are available with application to bacteria and fungi. The nucleotide fragment generated is amenable to nucleotide sequencing from which, by comparison to a gene bank of analogous ribosomal gene sequences from other organisms, primers that are specific to the taxon of interest can be selected. Accordingly, the mix of conserved and variable regions facilitates the determination of specificity and the high copy number (60-200 copies) adds sensitivity to the level of detection. At some levels of identification it may not be necessary to sequence the rDNA fragment because sufficient variation may be apparent within the variable region for differentiation by frequent-cutting restriction endonucleases.

The use of specific primers also allows direct amplification and identification from plant or soil material, negating the requirement for culturing the organism. However, this approach is subject to inhibition of the PCR reaction by environmental compounds, notably phenolics. Such inhibition, linked to the exponential nature of PCR amplification, poses the greatest drawback to the use of PCR as a quantitative tool from environmental samples.

Random Amplified Polymorphic DNAs

Where information on nucleotide sequence data is not available, alternative approaches dependent on PCR may be used. Principally, these are reliant on primers of about 10 bases in length that, because of their short length, have homology to the organism of interest by chance association. These primers are called arbitrary primers and the amplification is referred to as random amplified polymorphic DNA (**RAPDs**). Typically, RAPDs result in complex fingerprints that lend themselves to taxonomic studies and population analysis rather than identification. In some cases fragments derived from RAPDs that appear unique to the taxon of interest have been isolated and used as DNA probes. These 'unique' probes can be used on environmental samples, hence circumventing the need for isolation of the pathogen and purification which is inherent to RAPDs due to their arbitrary nature. The

likelihood of obtaining a RAPD DNA fragment of identification value may be increased by a DNA procedure referred to as subtractive hybridization. Here, nucleic acid sequences specific to the taxon of interest are enriched by the hybridization and subsequent removal of sequences that are common to other closely related organisms. This DNA then undergoes RAPD amplification and specific probes are sought.

PCR as a Taxonomic Tool (Ribosomal RNA Gene Analysis)

Specifically targeted PCR primers can be used to study variation in a pre-determined part of the genome: the primers selectively amplify a section of DNA known to be of interest. One of the most widely studied of such cases is the eukaryotic chromosomal gene coding for ribosomal RNA, often referred to as the rDNA. It should be remembered, however, that the same principles apply to other genes.

Ribosomal DNA has been subjected to detailed study in a wide range of organisms. One reason for its popularity as a research topic is that the gene is present in multiple copies in the cell, making it an easier target to detect than single-copy genes. Also, because ribosomes are ubiquitous in living organisms, rDNA analysis theoretically allows comparisons to be made between widely different taxonomic groups, even between kingdoms, thus providing valuable information for phylogenetic studies examining relationships between higher taxa.

The multiple copies of the ribosomal RNA gene are not scattered throughout the genome, but typically occur in one or a few blocks, or tandem arrays, where the gene is repeated over and over again. The internal layout of the repeated gene unit has been studied in many organisms and occurs in several formats.

RNA polymerase transcribes a rRNA precursor which includes the 18S (small subunit), 5.8S and 28S (large subunit) genes, together with part of the following DNA sequence. Parts of this RNA molecule are subsequently cut out by endonuclease enzymes in the nucleolus and do not become part of ribosomes. These are termed 'internal transcribed spacer' regions (ITS1 and ITS2), because they are internal to the rDNA repeat unit, they are transcribed by RNA polymerase and they are non-coding but relatively consistent in size, suggesting that they have a role in the spatial organization of the DNA. After the 28S subunit DNA is a larger spacer region often termed the 'intergenic spacer' (IGS). In some taxa (e.g. in some Basidiomycota, in *Saccharomyces cerevisiae* and related yeasts) the IGS contains the ribosomal DNA 5S gene, which is separately transcribed and also becomes part of functional ribosomes. In many other taxa (e.g. *Schizosaccharomyces pombe*, *Neurospora crassa*, *Fusarium* spp.) the 5S gene occurs elsewhere in the genome and may be scattered, rather than forming an array.

Sequence conformity in the multiple copies of the rRNA gene in a repeat unit is maintained by a process termed 'concerted evolution'. Taxa with scattered 5S rRNA genes (such as *Neurospora crassa*) may contain several different sequences for this gene. The functional parts of the rRNA gene repeat unit (18S, 5.8S and 28S) have been found to have relatively conserved DNA sequences, although several slightly more variable domains are known to

occur within these. Theoretically, the non-coding regions (ITS, IGS) may accumulate base substitutions without impairing ribosome functionality, and so are more variable in base sequence. The ITS regions have a spacer role and tend not to vary significantly in length, being 550-600 base pairs in total, including the intervening 5.8S gene. The IGS, however, may accumulate both base substitutions and changes in length from insertions and deletions and so is the most variable part of the repeat unit.

DNA sequences have been published for the rRNA genes in a wide range of organisms. This has enabled primers to be designed to amplify selectively specific parts of the gene repeat unit: for example, to target conserved regions of the rDNA and hence to work with diverse organisms. By selecting pairs of primers facing in opposite directions, i.e. with their 3' ends facing each other, specific regions of the gene may be amplified.

Further work is required to extract useful information. At present there are three main approaches to the study of rDNA: sequence analysis; restriction fragment analysis; and the production of taxon-specific probes or primers. The method chosen depends upon the objectives of the particular study: classification, identification or detection.

Sequence analysis

DNA base sequences are determined for a specific gene or parts of a gene in a range of strains. These sequences are 'aligned', side by side, so that comparisons can be made of base substitutions, insertions, deletions and translocation events. Aligned base sequences are then analysed using cladistic techniques to infer points of branching and hence determine hypothetical phylogenetic relationships between the taxa studied. Sequence analysis is thus a powerful tool in the development of classification systems.

Sequence data for rDNA, especially for the small subunit (18S), have been published for many organisms and the information submitted to computer databases which act as data banks of freely available information, for example GenBank at NCBI, or the ribosomal database project based at the University of Illinois.

GenBank holds sequence data from a wide variety of sources and is freely accessible via the Internet.

1. Via email: Send a message to: info@ncbi.nlm.nih.gov
2. Via anonymous ftp: [ncbi.nlm.nih.gov](ftp://ncbi.nlm.nih.gov)
3. Via a World-Wide-Web server: <http://www.ncbi.nlm.nih.gov/>

GenBank databases may be searched using Blast or Retrieve email servers, or by network Entrez software, or by purchasing data on CD-ROM.

The ribosomal database project (rdp) database may be accessed and searched on the Internet as follows:

1. Via email. Send a message to: server@rdp.life.uiuc.edu with the single line:
INFORMATION

You will then be sent instructions on available commands.

2. Using anonymous ftp: [rdp.life.uiuc.edu](ftp://rdp.life.uiuc.edu)
3. Using gopher: [rdpgopher.life.uiuc.edu](gopher://rdpgopher.life.uiuc.edu)
4. Via a World-Wide-Web server: <http://rdp.life.uiuc.edu/>

A range of molecular biology software for various computer platforms (PC, Mac, VAX, Unix, etc.) is available on the Internet, for example:

from EMBL at <http://www.ebi.ac.uk>

from the Australian National University at <http://life.anu.edu.au/molecular/software/>

Although it may be possible to identify taxa using sequence analysis, with today's technology this would be prohibitively expensive. Sequencing may be used to help develop classification systems, but simple, cheaper methods are then required to enable routine identification. At present there are two main approaches: restriction-fragment analysis and production of taxon-specific primers or probes.

Restriction-fragment analysis

PCR amplification product/s may be digested using restriction endonucleases and the resultant fragments separated by electrophoresis. By using a suite of enzymes it may be possible to differentiate between species by comparison of their restriction fragment patterns. This approach is most suitable for PCR product/s obtained from the ITS or IGS regions, where base differences are more common than in the expressed regions of the rDNA. Restriction enzymes with short recognition sequences (4 or 5 bases) tend to give the most useful fragment patterns, e.g. Taq I, Hae III, Msp I, but as many enzymes as possible should be screened in order to find any with differential ability.

Taxon-specific primers or probes

DNA sequence data may be used to design primers which have sequences unique to specific taxa. Experimental data from strains sequenced locally can be supplemented by searching the ribosomal database gene banks for other taxa which might cross-react with potential primer sequences. Primers are usually at least 20 bases long in order to increase their specificity. The level of taxon specificity can be tailored as desired - species, genus, kingdom - providing that suitable, correlating DNA sequences can be found at the relevant rank. With such primers, an amplification product will only be obtained if the target taxon is present. This not only enables identification, but facilitates detection of the organism concerned, even in mixed or contaminated materials, soil or infected tissues. Although such primers need not target the ribosomal RNA gene, it does have the advantage of having a high copy number per cell, thus making it easier to detect at low population levels.

RAPDs

The basic RAPD (random amplified polymorphic DNA) technique is related to other PCR-based fingerprinting procedures that are known under different names: multiple arbitrary amplicon profiling (MAAP), arbitrarily primed PCR (AP-PCR) and DNA amplification fingerprinting (DAF). Subtle differences do exist between each procedure. In general, differences are related to the size of the PCR primers used: AP-PCR, 18 or more bases; RAPDs, generally 10 bases; DAF, typically 7-8 bases. These acronyms are often incorrectly used, thus creating some confusion; for the purposes of this short introduction they will be considered as one technique.

With RAPD analysis, sequences of related organisms are sampled in a way that reveals small sequence differences. Primers are chosen arbitrarily without the need for information on the gene sequences of the target organisms. The procedure requires the isolation of DNA from the target organisms, amplification from the DNA using one random primer (though primers have also been used in combination), separation and visualization using agarose or polyacrylamide gels and staining with ethidium bromide or silver staining, respectively.

Successful RAPD analysis depends on consideration of a number of factors. The temperature of primer annealing is usually low, typically between 30 and 40°C, which allows annealing even when the primer does not exactly match the template. In addition, relatively high concentrations of magnesium chloride are required, between 4 and 8 mM; this has multiple effects but will mainly decrease the stringency of primer annealing and enhance primer mismatching. Arbitrary priming can occur at as few as 5 matches out of 6 bases at the 3' end of the primer. Thus, even in relatively small genomes, such as found in bacteria, there are a huge number of potential sites for primer annealing. In general, primer annealing must occur on opposite strands of the target DNA with the 3' end of each primer facing each other. Suitable amplification occurs when annealing sites are separated by between 200 and 4000 bases.

Primer selection is a key step in successful RAPD analysis and is potentially the most time consuming. This process is largely empirical, although most random primers have a relatively high guanine and cytosine (G + C) content (in excess of 60%). A small cross-section of strains from the major study are selected and a wide selection of random primers is screened against them. Ideally, any primers producing 5-10 bands in all strains are selected for further study using the complete set of test organisms. Visualizing the product is also an important feature, generally agarose and ethidium bromide staining can be used to generate fingerprints. However, if resolution of bands of a similar size is required, polyacrylamide gels can be used and DNA fragments can be visualized by silver staining.

Typically, RAPDs result in complex fingerprints that lend themselves to taxonomic studies and population analysis rather than identification. In some cases fragments derived from RAPDs that appear unique to the taxon of interest have been isolated and used as DNA probes. These 'unique' probes can be used on environmental samples, hence circumventing the need for isolation of the pathogen and purification which is inherent to RAPDs due to their arbitrary nature. The likelihood of obtaining a RAPD DNA fragment of identification

value may be increased by a DNA procedure referred to as subtractive hybridization. Here, nucleic acid sequences specific to the taxon of interest are enriched by the hybridization and subsequent removal of sequences that are common to other closely related organisms. This DNA then undergoes RAPD amplification and specific probes are sought.

PCR Quality Control (The Development of Good Experimental Procedures)

PCR is a highly sensitive procedure and it has been shown that as little as one molecule of DNA can serve as a template for amplification. Because of the extreme sensitivity of the technique, contamination is one of the major problems of PCR use. If the sample to be amplified contains DNA from a foreign source, false positives or misleading results can easily be generated. This can only be completely avoided by spatially isolating every stage of the process to avoid possible sources of contamination. However, this is rarely feasible and a compromise must be arrived at. It is possible to avoid major reorganizations in laboratory practice and use if some simple routines are adhered to.

All plastic ware must be sterilized by autoclaving before use, used only once and then discarded. Gloves should be worn at all times and whenever possible, dispensing of reagents should take place in a laminar air-flow cabinet. Pipetting aids are potentially a major source of contamination, risks can be reduced by using positive displacement pipettes or by using aerosol resistant tips. In addition, pipetting aids should be regularly cleaned with dilute bleach to reduce the risk of contamination. Similarly, work surfaces should be regularly cleaned with a dilute bleach solution.

Wherever possible all reagents should be purchased from reputable companies and, if available, molecular biology grade chemicals should be used. All reagents should be handled using good aseptic technique. If mistakes are made in the preparation of reactions, these should be discarded and performed again. Water used in PCR reactions and in the preparation of reagents used in PCR must be of a high quality; HPLC-grade water that has been filtered through a 0.2 μm nitrocellulose filter, then autoclaved, is preferred. New reagents should be diluted to the correct concentration and aliquoted into small volumes, generally 50 μl . This ensures that stock reagents cannot become contaminated by poor pipetting technique. In addition, to avoid pipetting errors in the preparation of reaction mixtures, a master-mix of PCR reagents can be prepared in advance and then the correct volume dispensed into each reaction tube.

Chromatography

Chromatography is a technique for separating mixtures of molecules, or for characterization of individual molecules. A mobile phase containing the components to be separated is passed over a stationary phase to elicit separations. Mobile phases can be of a wide range of different solvents including gases and the stationary phase can also be of widely different types. There is an interplay between the three main elements: firstly, stationary phase from active to inactive; secondly, mixture to be separated (from non-polar (lipophilic) to polar (hydrophilic)); and thirdly, mobile phase (from non-polar to polar). Physical and chemical forces act between the solutes and the two phases to cause separations.

Two sub-classes exist within solid/liquid chromatography, these are normal-phase chromatography (NPC) and reversed-phase chromatography (RPC). The stationary phases in NPC are polar whereas RPC supports are essentially non-polar. Conceptualization of NPC can be of the retention of solutes being determined by the polar interactions of the stationary and mobile phases. These interactions are complex and involve such forces as dispersion forces, dipole moments, hydrogen bonding, and dielectric interactions between ions. In RPC the situation is reversed where the mobile phase is more polar than the stationary phase and hence attractive forces between the stationary and mobile phase are dominant. The stationary phase is usually an inert hydrocarbon and the interactions are hydrophobic, with selectivity being dominated by solvent effects.

Thin layer (both normal and reversed phase), preparative layer (normal), liquid (normal phase), and high-performance liquid chromatography (reversed phase) are commonly used.

TLC

In this procedure the stationary phase is layered over a support plate which can be of glass, aluminium or plastic. Samples in a suitable solvent are loaded onto the plate, and when dry it is placed into a sealed glass TLC tank containing a solvent. The solvent ascends the TLC plate by a form of capillary action.

HPLC

In the analytical system, liquid chromatography is scaled down and automated so that rapid, reproducible analysis of complex mixtures can be performed. The particle size of the stationary phase is very small, enabling small columns with high resolution power to be used which increase the speed of elution. The column is eluted with a mobile phase pumped through the column in a highly controlled manner. Minute sample (microlitre) volumes are injected onto the column through an injection port. The solutes are monitored by some chemical or physical property of the solute (e.g. absorbance, fluorescence, etc.) after the solvent has passed through the column detector systems.

Alkylphenone standards are a homologous series of compounds which are useful for the standardization of isocratic and gradient HPLC. They are stable, have a wide range of polarities, and are readily detectable. Retention indices from using alkylphenone standards are independent of column efficiency, variations of column efficiency, column flow rate, and temperature. Retention indices may be used between laboratories as a reliable guide to the tentative identification of, for example, mycotoxins because the use of alkylphenones reduces variability from chromatographic conditions.

Fatty Acid Analysis

All cellular organisms possess a cytoplasmic membrane as a component of their cell envelope, the composition of which is roughly 50% lipid. Membrane lipids are a diverse group of molecules which are frequently used as markers for the classification and

identification of micro-organisms. In particular, the amphipathic lipids (possessing hydrophilic and hydrophobic regions) have great relevance to microbial systematics, both as complete molecules and when broken down to their constituent parts. Amphipathic lipids such as glycolipids and phosphoglycerides consist of a polar 'head' and non-polar hydrocarbon 'tails' and are generally arranged to form a lipid-bilayer. The non-polar tails are generally esters of long-chain fatty acids.

The fatty acid composition of micro-organisms has been used extensively to aid microbial characterization. In general, the fatty acid profile of most bacteria ranges from n-C9 to n-C20. The fatty acids in most Gram-positive bacteria are located in the cytoplasmic membrane, though some, such as the mycobacteria and related genera, have long-chain lipids, known as mycolic acids (β -hydroxy fatty acids), which are found in an outer membrane-like structure. Gram-negative bacteria possess fatty acids in the cytoplasmic membrane though they can also be found in the cell-wall fraction, generally in the lipopolysaccharide. In common with most eukaryotic cells, fungal fatty acid composition is much simpler than the prokaryotes: n-C16 and n-C18 acids predominate. These are commonly found as components of more complex cytoplasmic membrane lipids such as acylglycerols and glycerophospholipids; they can also be found in the 'free' form, though the significance of this remains unclear.

Microbial fatty acids can be found in a variety of forms. The basic types of acids include straight-chain-saturated acids; mono- and polyunsaturated acids; methyl-branched acids; and those with cyclopropane rings or with hydroxy groups. Methyl-branched fatty acids of the iso- and anteiso- configurations occur predominantly in Gram-positive organisms, although there are Gram-negative exceptions such as members of the genus *Xanthomonas*. In general, cyclopropylated and hydroxylated fatty acids are more commonly found in Gram-negative bacteria. The hydroxylated acids are of particular importance as they are generally components of the lipopolysaccharides, and can be highly diagnostic in determining species identity, particularly among non-fluorescent pseudomonads. Polyunsaturated fatty acids generally do not occur in prokaryotes though they do have significance in fungal characterization.

The analysis of fatty acids can be broken down into five steps:

1. Cell culture under standard conditions.
2. Release of fatty acids from the cell surface (both cytoplasmic and outer membranes) by saponification.
3. Methylation of the fatty acids to increase volatility.
4. Analysis by gas chromatography, frequently using capillary columns to enhance sensitivity and resolution.
5. Exploitation of the fatty acid profile for classification or identification. Identification can occur by comparing the fatty acid profile of an unknown to those in the literature or by

computer library matching. Classification can involve statistical analysis using fatty acid profiles from related or similar organisms.

PAGE and Protein Electrophoresis

The movement of charged proteins on an inert support in an electric field enables their physical separation. This can be considered as a form of chromatography, where the stationary phase (the gel) is often assumed to be chemically inactive. However, chromatographic-like adsorption effects have been observed between gels and proteins, for example for proteins containing aromatic amino acids. In addition, separations by electrophoresis not only rely on charges on proteins but the results also depend upon other molecular properties. Thus electrophoretic mobility also depends on molecular size. Results are impossible to predict merely from knowledge of the charged groups on proteins. Although electro-osmosis effects are lacking with polyacrylamide gel electrophoresis (PAGE), molecular sieving phenomena are imposed, and in fact molecules of similar charge-to-size ratio may be separated on the basis of size alone.

Before undertaking the electrophoretic separation of proteins, it is desirable to standardize extracts to the same protein concentration. This is performed by a spectrophotometric assay for total proteins, and adjusting volumes of extracts to give between 4 and 8 mg protein in each millilitre of extract. There are a wide range of methods available for the spectrophotometric determination of total protein, including the Bradford reagent and Lowry and Biuret determinations.

Reliable separation of proteins according to their size can be obtained by binding the detergent sodium dodecyl sulphate (SDS) to the proteins. As protein and SDS bind in a constant ratio, and SDS is highly charged under standard electrophoretic conditions, the relative mobility of the protein-SDS complex is proportional to the size of the protein. This technique may be used with total protein methods, but is not suitable for isozyme analysis as the enzymes are inactivated by the binding procedure.

Electrophoresis is carried out in polyacrylamide gels with a discontinuous buffer system. The exact conditions vary depending on the enzyme system being studied. Electrophoresis results in a smear of proteins down a supporting gel, which can be visualized by staining the gel with protein-specific stains. The unstained gel can be incubated in specific enzyme-substrate reaction stains so that the protein bands corresponding to particular enzymes can be seen on the gel. Many enzymes are present in cells in more than one form and so a number of bands corresponding to different sizes, but showing the same enzymic activity may be seen. These different forms of the same enzyme are referred to as iso-enzymes or isozymes, and have been used extensively in microbial taxonomy to differentiate genera, species and populations, and have also been used extensively in microbial population genetics.

PFGE

Pulsed-field gel electrophoresis (PFGE) is appropriate to the separation of DNA fragments

larger than 20 kb, where the utility of standard electrophoresis declines. The technique is reliant on the alternation of electric fields, the frequency of which determines the size of the DNA fragments resolved. Simplistically, DNA under PFGE zigzags through the agarose matrix as a function of how quickly the DNA fragments re-orientate in response to the alternating electric field; small and large fragments are resolved by short and long pulse times, respectively.

A prerequisite to the analysis of large DNA fragments is that shearing forces acting upon the DNA must be minimized. Accordingly, DNA isolated by standard protocols is inappropriate for PFGE studies. Instead, DNA is isolated from cells fixed within an agarose matrix by treatments that remove the cell wall and disable enzymatic activity. The stringency of the cell-wall removal treatment depends on the composition of the wall. For example, Gram-negative bacterial cell walls are susceptible to detergents, such as sarkosyl; whereas Gram-positive bacterial cell walls require additional exposure to lytic enzymes, such as lysozyme. Fungal studies present a further aspect in that protoplasts have to be generated before cell fixing. In general terms this is achieved by the removal of the cell wall during culture by the action of a cocktail of enzymes. DNA-binding proteins and other proteins, and nuclease activity are removed in all cases by exposure to proteinase K and EDTA, respectively.

Applications

Knowledge of the genetic diversity within populations of the causal organism is fundamental to the development of a plant disease control strategy. Many methods are available that can type isolates comprising a population with varying degrees of differentiation. PFGE offers a relatively new approach that can be applied to bacterial ecology.

Bacterial isolates can be typed by a genomic fingerprint generated by a rare-cutting restriction endonuclease. This approach necessitates the use of PFGE for DNA fragments larger than 20 kb. The fingerprint is comparatively simple, making the data generated amenable to interpretation. Typically, a rare-cutting fingerprint contains in the range of 5-40 fragments.

The rare-cutting nature of a restriction endonuclease enzyme is an arbitrary parameter dependent on the DNA constitution of the organism under study. Accordingly, the identification of a restriction endonuclease that is rare-cutting to a particular organism necessitates a screening process. Towards this objective, the following rationale may be used as a guide: organisms of high G and C content are most likely to contain restriction recognition sites of rare frequency comprising A- and T-rich sequences, and vice versa. Having identified a rare-cutting restriction endonuclease, the conditions for PFGE need to be optimized for fragment resolution. The representative nature of the data generated is increased if each isolate is characterized by more than one restriction endonuclease.

The data generated by this approach are suitable for statistical analysis, and information regarding relatedness of isolates and genetic diversity of populations can be investigated. These statistical analyses rely on the assumption that co-migrating DNA fragments between isolates are homologous. This assumption holds between closely related isolates where the genomic fingerprints are ostensibly similar, but may lead to spurious similarities between

less related isolates: in such cases other approaches are more advisable. PFGE has been used to monitor bacterial phyllosphere dynamics, type pathogenic isolates from a single outbreak of disease, track the emergence of antibiotic-resistant isolates in hospitals, uncover subgroups within pathovars and to type the valuable production isolates in the manufacture of cheese and wine.

PFGE has also been applied to studies on filamentous fungi and yeast, as the technique can reliably separate pieces of DNA up to approximately 9 Mb. The chromosomes of many filamentous fungi and yeast are in this size range, and so a major application of PFGE is to obtain a karyotype for fungi, where the individual chromosomes from the fungal nucleus are arranged in a gel according to their size.

Chromosome numbers and sizes can vary considerably within species and genera of fungi, particularly in respect of length polymorphisms and the presence of small chromosomes. In some instances there is sufficient variability for karyotypes to be used as characters in strain typing. This is particularly the case for yeasts such as *Saccharomyces* and *Candida* species. Electrophoretic karyotypes have been determined for many fungi and similar variations in lengths and numbers have been observed in many genera. A few examples of these are *Fusarium oxysporum*, *Leptosphaeria maculans*, *Ustilago hordei* and *Colletotrichum gloeosporioides*. The significance of differences in karyotypes in filamentous fungi is not entirely clear. Some differences have been suggested to be correlated with differences in pathogenicity, but these correlations have not always been supported when large numbers of isolates have been studied, the lack of a sexual cycle in most of these isolates prevents studies on chromosomal assortment and replication.

Ubiquinones

Ubiquinones are being used increasingly in fungal taxonomy, and the methods of analysis for these compounds are similar to those employed for secondary metabolites.

Ubiquinones are isoprenoid quinones and a chemical class of terpenoid lipids. They are highly mobile electron carriers between the flavoproteins and the cytochromes of the electron transport chain, and consist of a benzoquinone ring with an isoprenoid chain, the latter component making ubiquinones highly non-polar. The length of the isoprenoid tail, and degree of saturation, can vary depending upon the biological system from which the ubiquinone was isolated.

Ubiquinones are being used increasingly in microbial chemotaxonomy, due to the inherent structural variation observed between some taxa. The isoprenoid tail makes the molecule highly non-polar, enabling rapid diffusion in hydrocarbon phases such as mitochondrial inner membranes. Ubiquinones are constituents of the eukaryotic mitochondrial plasma membranes and are important in electron transport, oxidative phosphorylation and, possibly, active transport. Transport of electrons is carried out by reduction and subsequent oxidation of the two quinone groups on the benzoquinone ring. Both partially saturated and unsaturated structures are found in fungi, whereas only unsaturated have been observed from bacteria. There is considerable interest in their use in fungi as taxonomic characters.

Many previous methods for ubiquinone analysis are time consuming, and specify high-performance liquid chromatography which is expensive, and/or rigorous extraction procedures, particularly saponification. In addition, many reports list only the most prominent ubiquinone as characters, and those at lower concentration, or which are not as readily detected, tend to be ignored. 'Direct' methods avoid these potential problems, and offer the advantages of not requiring expensive equipment, being experimentally undemanding, giving highly reproducible results, and using a mild extraction procedure. All the spots on the TLC plate are considered as characters (i.e. ubiquinone and non-ubiquinone) on the basis of all evidence being 'grist for the taxonomic mill'.

In the past ubiquinones were thought to be only of use at the genus level with fungi, for example, many members of the penicillia were grouped together as containing ubiquinone 9. However, by using the above method a range of ubiquinones can be detected. This indicates that the situation is more complex, and species identifications may now be possible.