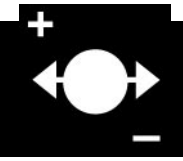
GENE TYPING: TWO-DIMENSIONAL ELECTROPHORESIS



N. J. van Orsouw, S. B. McGrath and J. Vijg, Institute for Drug Development, Cancer Therapy and Research Center, San Antonio, TX, USA
R. K. Dhanda, Mosaic Technologies, Boston, MA, USA
C. B. Scott, CBS Scientific Company, Del Mar, CA, USA

Introduction

With the human genome program drawing to a close, attention is now rapidly shifting from obtaining consensus sequences of all human genes to the detection of individual DNA sequence variations. Based on complete sequence information for all human genes, it is theoretically possible to generate a catalogue of all gene mutations and polymorphisms in the human genome and test them directly for association to relevant phenotypes, e.g. of health and disease. Unfortunately, current methods for detecting DNA sequence variants are not optimized for generating data on multiple genes in large numbers of individuals, e.g. in population-based studies or in the clinical setting. The most reliable system for comprehensive gene sequence analysis is still nucleotide sequencing itself, which is not compatible with cost-effective large scale population-based genetic screening.

Recently, various systems have been proposed to analyse gene-coding and regulatory sequences more effectively for all possible variations. Here we review the development and application of one such system, two-dimensional gene scanning (TDGS). This method is based on the two-dimensional separation of polymerase chain reaction (PCR)-amplified gene fragments on the basis of both size and base pair sequence in polyacrylamide gels. Attention will be focused on most recent developments in automation and miniaturization of the two-dimensional electrophoresis procedure. Future developments towards a dedicated fully automated high-throughput system for gene analysis will be discussed.

Two-Dimensional Gene Scanning: Background and Principles

Denaturing Gradient Gel Electrophoresis

TDGS is based on denaturing gradient gel electrophoresis (DGGE) as the mutation detection principle, in combination with PCR amplification to prepare the target sequences. In DGGE, DNA fragments are subjected to electrophoresis in a polyacrylamide gel against a gradient of ever higher temperature or chemical denaturants (i.e. a mixture of urea and formamide). Unlike nucleotide sequencing, DGGE detects mutations, including base pair substitutions and small insertions and deletions, on the basis of differences in the melting temperature of the target fragments. A given DNA fragment comprises one or more domains, each representing a stretch of between 50 and 300 base pairs with equal melting temperature (the temperature at which each base pair has a 50% probability of being in either the helical or the denatured state). Since the stability of each domain depends on its sequence, mutational differences among different fragments are revealed as migrational differences in the gel (**Figure 1**A).

In order to obtain virtually 100% accuracy in mutation detection, fragments to be subjected to DGGE can be clamped to a GC-rich sequence (a stretch of 30–50 G and C bases). A convenient way of attaching a GC-clamp to the target fragment is by making it part of one of the primers in a PCR. Without GC-clamping, a DNA fragment consisting of one melting domain will become completely single-stranded upon denaturation and run off the gel. By adding a GC-clamp, a single high-melting domain is artificially created at one end of the target fragment. As the GC-clamped target fragment migrates through the gradient of denaturants, melting of the target domain causes partial branching and halting of the fragment in the gel (Figure 1B). Thus, one function of the GC-clamp is to ensure branch formation after melting of the target fragment. However, when the target DNA fragment consists of multiple melting domains (Figure 1C), only mutations in the lowest melting domain are readily detected. To facilitate detection of all possible mutations, it is imperative that the target fragment represents only one melting domain. Fortunately, since the addition of a GC-clamp allows for stacking interactions with neighbouring bases, the entire fragment will often behave as one melting domain (Figure 1C). However, this is not always the case, and in practice, the target fragment needs to be designed, e.g. through the strategic positioning of PCR primers to achieve the ideal single

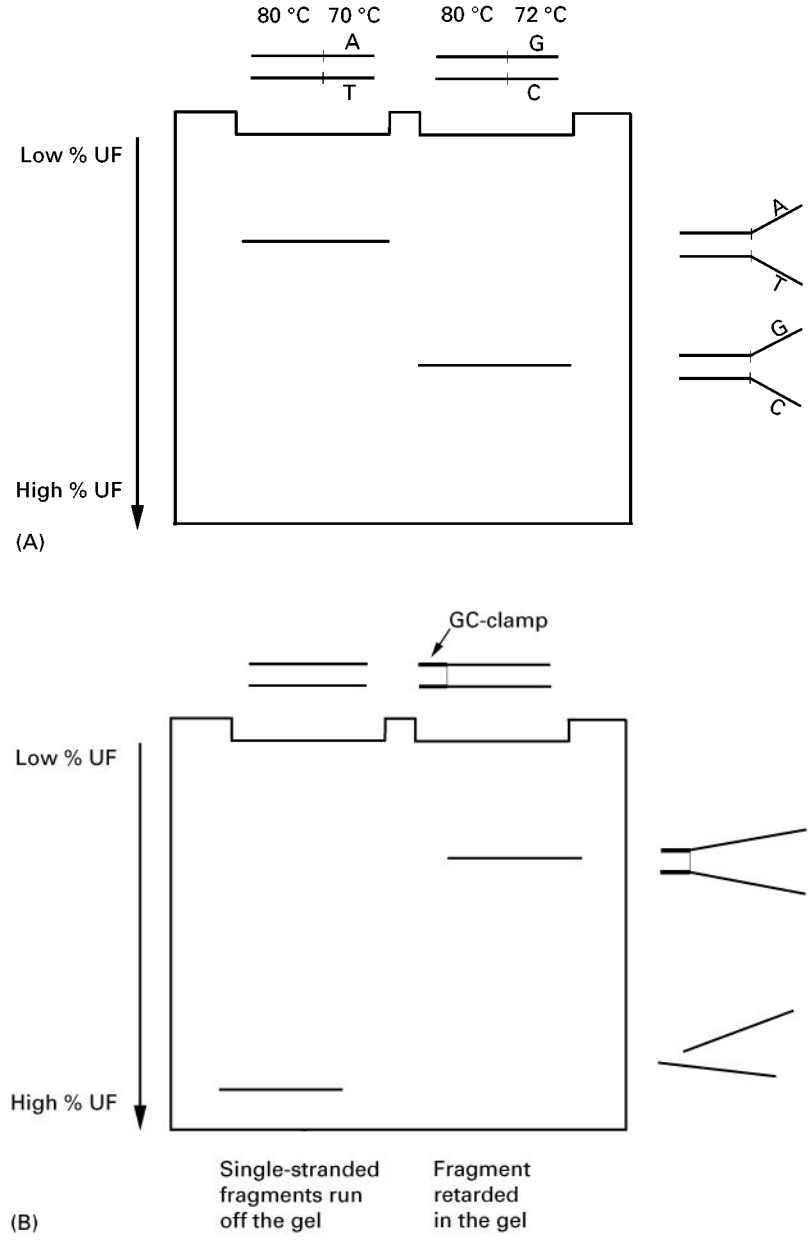
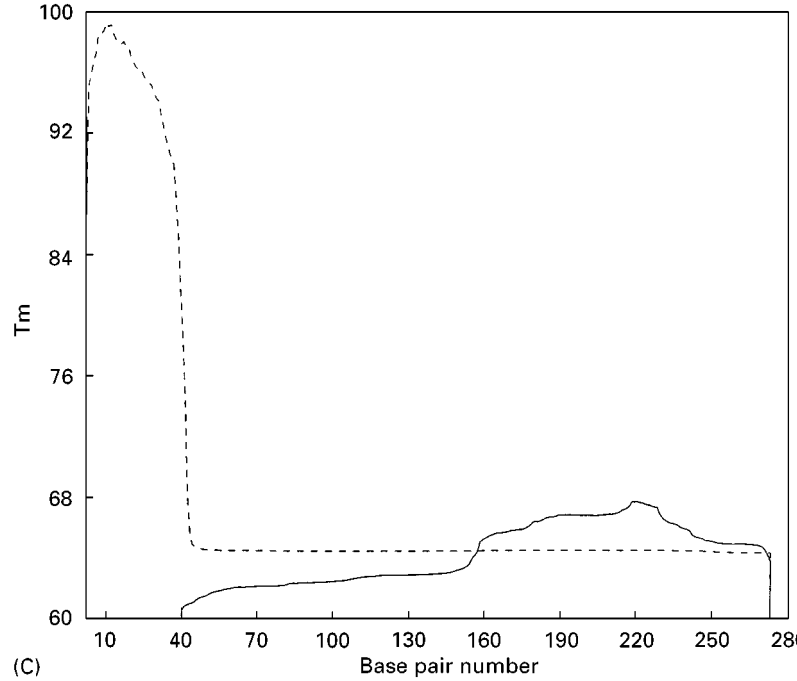


Figure 1 Principles of denaturing gradient gel electrophoresis. (A) Single base changes affect the melting temperature of a fragment, which results in a gel shift. (B) After complete denaturation, single-stranded fragments will run off the gel; the addition of a GC-clamp to the target fragment prevents complete denaturation and therefore fragments will be retarded in the gel. (C) The addition of a GC-clamp to a multiple-domain fragment can make the fragment behave as a single-domain fragment. Continuous line: target fragment without a GC-clamp. Dashed line: target fragment, including a 40 bp GC-clamp. (D) The introduction of a heteroduplex cycle at the end of PCR amplification of target fragments facilitates detection of heterozygous mutations as four molecules: two homoduplexes and two (early-melting) heteroduplexes.

melting domain. In general, target fragments in DGGE have an average size of 275 bp, including a GC-clamp and PCR primer sequences.

The sensitivity of DGGE for detecting variants is further enhanced by the introduction of a heteroduplexing step using one round of denaturation/renaturation, usually at the end of PCR amplification of the target fragment. In this manner, a heterozygous mutation is revealed as four different double-stranded fragments: two homoduplex molecules (one wild-type homoduplex and one mutant homoduplex) and two heteroduplex molecules (each comprising one wild-type and one mutant strand). Since the stability of heteroduplexes is so much lower, they always melt earlier than the homoduplex molecules (Figure 1D).



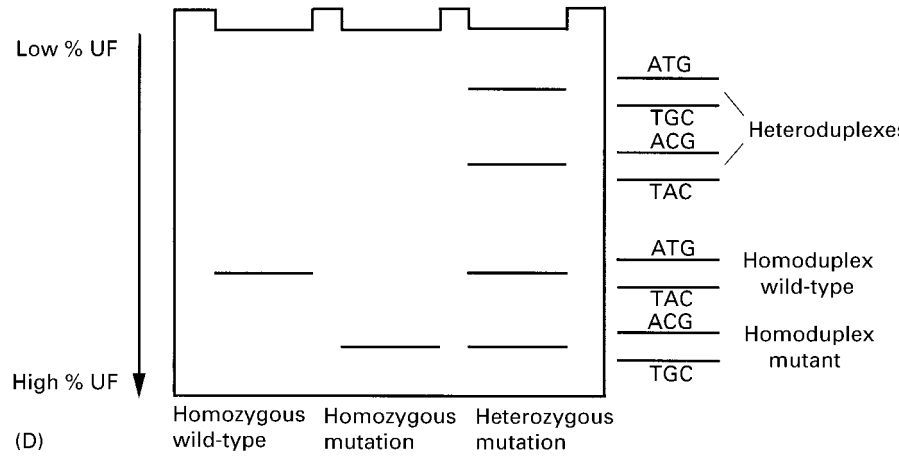


Figure 1 *Continued*

Although DGGE has the crucial advantage of having virtually 100% sensitivity in detecting mutations, it has typically been applied in a serial fashion, e.g. on a fragment-by-fragment basis. For analysing large genes or multiple genes this is not practical. A solution for this problem, which we adopted, is to apply the DGGE principle in the format as it was originally described, i.e. a two-dimensional system of separation by size followed by DGGE. Successful implementation of such a two-dimensional DNA electrophoresis system in mutation scanning of large genes requires an efficient multiplex PCR protocol. Indeed, without the possibility to PCR-amplify multiple target fragments (i.e. typically 10 or more) in one single reaction, the application of a parallel analysis system offers only a limited advantage. Multiplex PCR systems for genes and genetic markers are now becoming available and it has been demonstrated that as many as 26 fragments can be co-amplified in one single tube under the same reaction conditions.

Two-Dimensional DNA Electrophoresis

The major advantage of two-dimensional electrophoresis is that it provides a high resolution system to screen multiple fragments under the same conditions. It has been demonstrated that DGGE provides virtually 100% mutation detection sensitivity even when applied with a broad range gradient of

denaturants. This opens up the possibility to analyse multiple fragments for all possible mutations under the same set of experimental conditions. The total number of target fragments that can be analysed simultaneously depends on the resolution of the gel system used. Although high resolution can be obtained by using one-dimensional denaturing gradient gels, two-dimensional separation allows characterization of each fragment on the basis of two independent criteria, size and melting temperature. In practice, a fragment mixture corresponding to all exons of a gene is electrophoresed in a nondenaturing size gel. Fragments are further sorted out in a denaturing gradient gel as the second dimension (**Figure 2**). By using the two-dimensional system, it is possible to visualize completely all fragments corresponding to an entire gene for a particular DNA sample and immediately recognize each exon and variants therein. This has been demonstrated for several large human disease genes, including *CFTR*, *RB1*, *MLH1*, *TP53*, *TSC1*, *BRCA1*, as well as for a part of the mitochondrial genome.

Design Software and Instrumentation for TDGS Tests

Computer-Automated Design of Target Fragments for PCR and DGGE

A potential hindrance to the widespread application of TDGS to multiple novel genes involves the difficulties in the design of PCR primers generating

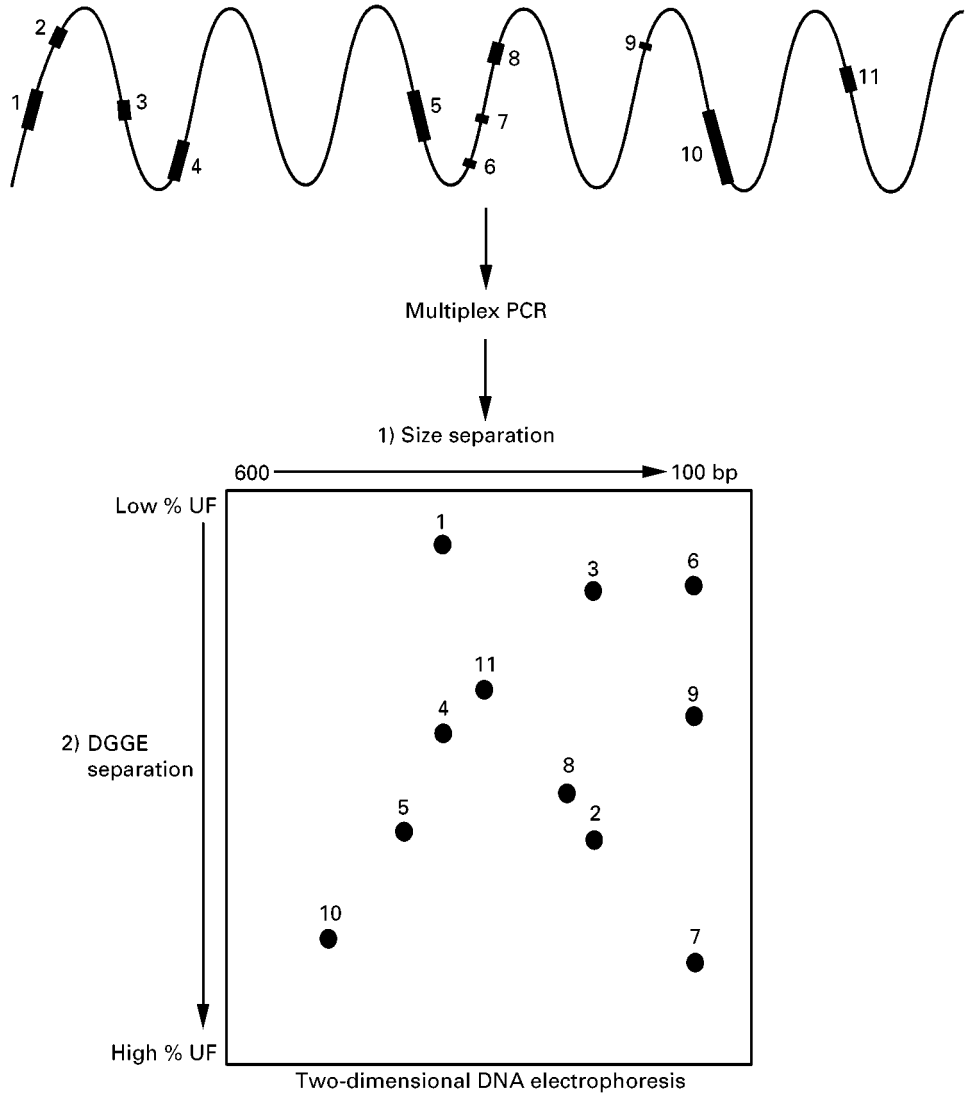


Figure 2 Schematic depiction of a TDGS test. All exons are amplified in an extensive multiplex reaction, and the fragments are resolved by size separation, followed by separation in a gradient of denaturants. Heterozygous mutations would show up as four spots instead of one.

single-domain fragments which can be resolved under one set of electrophoretic conditions. To design complete gene tests for mutational analysis by TDGS, an automated generally applicable computer program was developed, which was based on a commercially available primer design program (Primer Designer 3; Scientific and Educational Software, State Line, PA), the melting routine MELT87 and a newly generated spot distribution routine. After entering a gene's coding sequence as exons with their flanking intronic sequences, a rank of suitable PCR primers for each exon is designed by the PCR design subroutine. Next, the best primer pair is used in the melt subroutine to check for a one-domain target fragment. The program uses different GC-clamps at either the 5′ or 3′ end of the target fragment and, if necessary, additional small GC- or AT-clamps at either side of the target fragment. If it is impossible to design a one-domain fragment, the next optimal primer pair is tested, and so on. If a primer pair suitable to create a one-domain fragment cannot be found, the exon is split.

As soon as primers fulfil PCR and melting criteria, the fragment is positioned according to its size (x) and melting (y) coordinates. The spot distribution routine then checks for possible overlap. The output file of the program is a complete list of primers to be used in TDGS. (Questions regarding the use of the TDGS software should be directed to Accelerated Genomics, Concord, NH, http://www.accelerated genomics.com Tel.: (210) 616-5910; fax: (210) 692-7502.)

Electrophoresis

For two-dimensional DNA electrophoresis, originally two different gels were used for the first-dimension (separation according to size) and the subsequent second-dimension separation of these fragments by DGGE on the basis of their melting temperature. The first-dimension separation was carried out in polyacrylamide slab gels, which required staining of the gel to visualize the one-dimension separation pattern before this could be excised and transferred to the second-dimension denaturing gradient gel. Alternatively, tube gels have been employed for size separation, which obviated the need for gel staining and lane excision. However, routine application of TDGS requires standardization and automation, which is incompatible with the labour-intensive step of manual interference between the first- and second-dimension separation.

Recently, we developed a simple automated two-dimension instrument, which is based on an existing vertical electrophoresis system with an isolated horizontal unit on top (**Figure 3**). This top unit consists of two outer chambers and one middle chamber. The necessary contacts between the outer buffer chambers and the gel are provided by two strategically located openings in the inner glass plate.

In this system only one gel (a denaturing gradient gel) is used with the top part nondenaturing. This nondenaturing part functions as the lane for the first-dimension size separation. A slot former is placed in the top left part (**Figure 4**). In the current configuration, a gel is attached to each side of a gel holder, which can be placed in a buffer tank. Buffer tanks can hold multiple units so that multiple gels can be run simultaneously (**Figure 5**). The sample is electrophoresed on the basis of size horizontally in the nondenaturing top gel, and the second-dimension

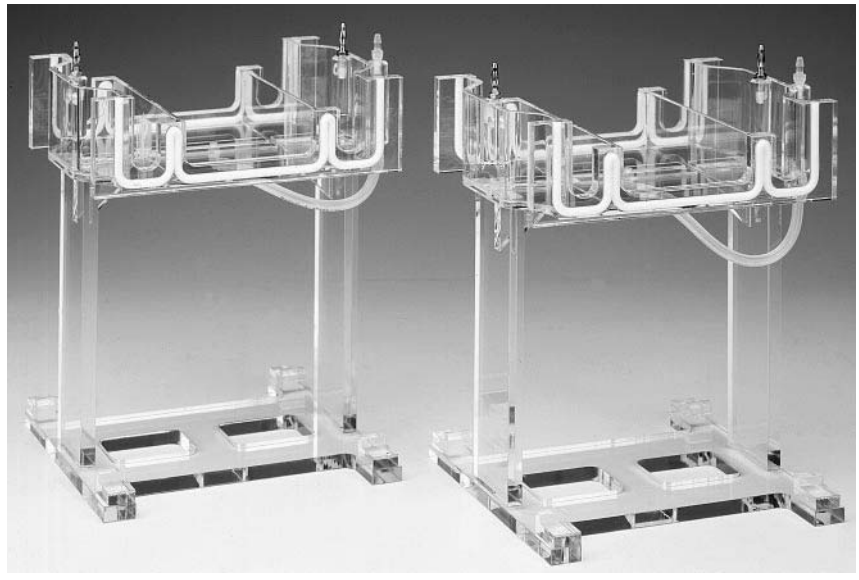


Figure 3 Two automatic dual-gel TDGS systems.

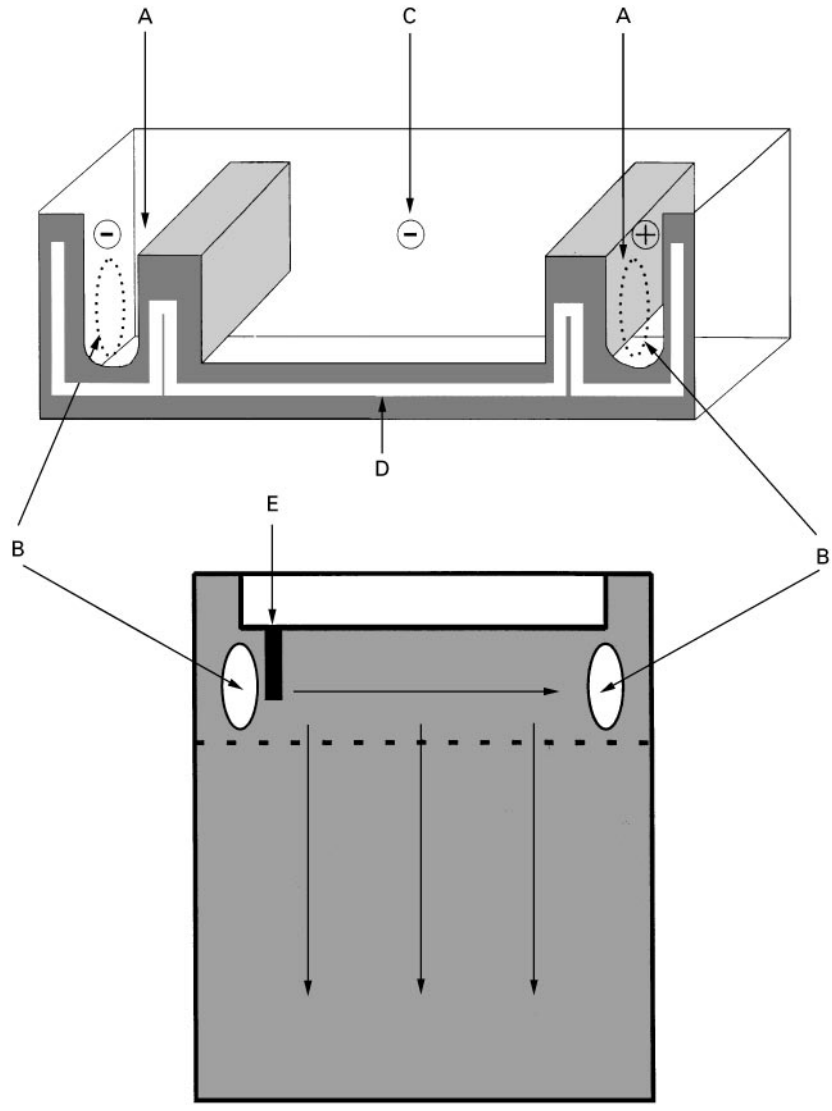


Figure 4 Schematic depiction of the automated two-dimensional electrophoresis unit. Buffer chambers for the first-dimension separation (A) are connected with the gel through openings in the (inner) glass plate (B). During the first-dimension electrophoresis, the middle buffer chamber (C) is isolated from the outer chambers. For the second-dimension run, buffer chamber C is flooded with buffer and the upper electrode is turned on in conjunction with a positive electrode in the lower reservoir (not shown in this figure). The gel cassette is sealed to the top unit with a serpentine silicone gasket (D), and sample is loaded in the single slot (E). The dashed line indicates the beginning of the gradient of urea/formamide.

electrophoresis is carried out vertically in the denaturing gradient gel. All components of the automatic TDGS electrophoresis system are depicted in **Figure 6**.

Gradient gels can be poured, up to nine at a time, using a simple linear gradient maker in combination with a multiple gel caster. The exact gradient that is to be applied is dependent on the GC-content of the gene(s) of interest and is determined by the TDGS primer designer software.

Miniaturization

Miniaturization of gene analysis systems, such as the TDGS system described here, offers two major advantages: increased speed and lower cost.

Speed The duration of electrophoresis depends on the voltage applied. For example, the optimal electrophoresis conditions for the retinoblastoma susceptibility gene *RB1* using standard 1.0 mm thick gels,

Figure 5 The entire automatic TDGS system. In this version of the system, four gels can be run simultaneously submerged in a buffer tank, which is equipped with a heater/stirrer to provide for a constant temperature. For more information, see http://www.cbssci.com

are 100 V, 5 h for the size separation and 100 V, 16 h for the second-dimension separation. Increasing the voltage increases the heat production, which negatively affects gel resolution. An obvious strategy is the use of thinner gels, which facilitate rapid heat dissipation into the surrounding buffer and thereby allow increasing the voltage while maintaining a good resolution. Currently, gels as thin as 0.35 mm are now run at 500 V, 0.8 h for the size separation and 500 V, 3.5 h for the second-dimension separation.

Cost The cost factor is of major importance for the large scale implementation of genetic testing. Since this is determined to a major extent by reagent and material cost, as well as space, miniaturization of analytical systems is of crucial importance. Miniaturization of TDGS results in thinner and smaller gels, which require less sample (smaller PCR volumes can now be applied) and lower gel and buffer amounts. Moreover, they take up less space. Instead of the current 17×22 cm format, two-dimensional patterns have already been produced on 10×10 cm mini-gels, and it is not unreasonable to expect that ultimately electrophoresis will be carried out on glass slides.

Detection of TDGS Patterns

After electrophoresis, the two-dimensional DNA fragment patterns can be visualized by incubating the gels with DNA staining dyes. Examples are ethidium bromide or the more sensitive dye Sybr-green. Patterns are photographed under UV light and evaluated for the occurrence of variations (in the form of four spot patterns; see Figure 1D). An example of a TDGS pattern is shown in **Figure 7**, depicting the *RB1* gene, containing a mutation in exon 2.

However, for large scale application of TDGS, dye primer technology for the in-gel detection of two-dimensional spot patterns is an obvious strategy. Test results indicate similar two-dimensional patterns and sensitivity for fluorescein-labelled primers compared to Sybr-green-stained gels.

Introduction of fluorescent detection offers two advantages over gel staining. First, the reduction in labour is considerable and loss of gels due to breakage is prevented. Second, since there is no need to release the gel from between the glass plates it has become possible to use thinner gels, which will allow shorter electrophoresis times (see above). To increase the efficiency even further it is possible to label different samples with different fluorophores. Current fluorescence imagers have the option to analyse multiples fluorophores in the same gel.

Future Developments

Routine application of TDGS requires standardization and further streamlining of the procedure. Ultimately, one could envisage a fully automated system of PCR amplification, sample loading, electrophoresis, scanning of gels by a fluorescent imager, followed by online interpretation of gels by image analysis systems.

Much of the labour that is involved in PCR amplification, as well as the error rate, can be greatly

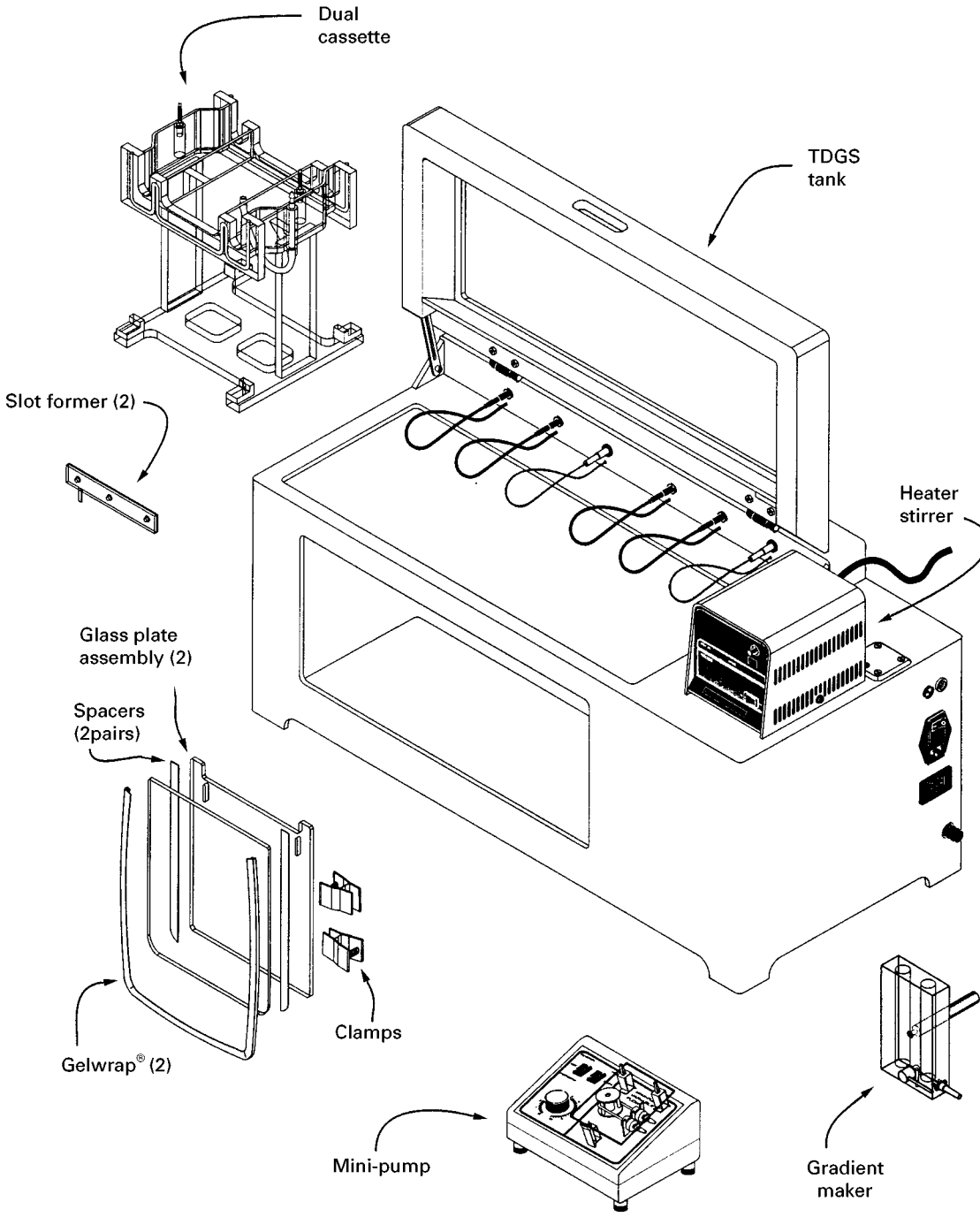


Figure 6 Line drawing of all the components of a 4-gel automatic TDGS system.

diminished by PCR robotics. Such instruments have now become widely available and, in combination with an ongoing effort to increase multiplex groups, are expected to increase greatly the front-end throughput of genetic testing. Multiple two-dimensional gels can be stacked for simultaneous electrophoresis of manifold samples. A simple robot arm could load the gel sandwiches into the fluorescent imager for quick gel scanning. Finally, while the actual interpretation of spot patterns is currently most conveniently done by eye, automated image analysis software is commercially available. The use of such software may, for example, facilitate the detection of subtle positional changes in the context of other spot variations. In this respect, one could envisage a programme with information on all possible spot positional variants to identify quickly recurrent mutations and polymorphisms on the basis of their unique

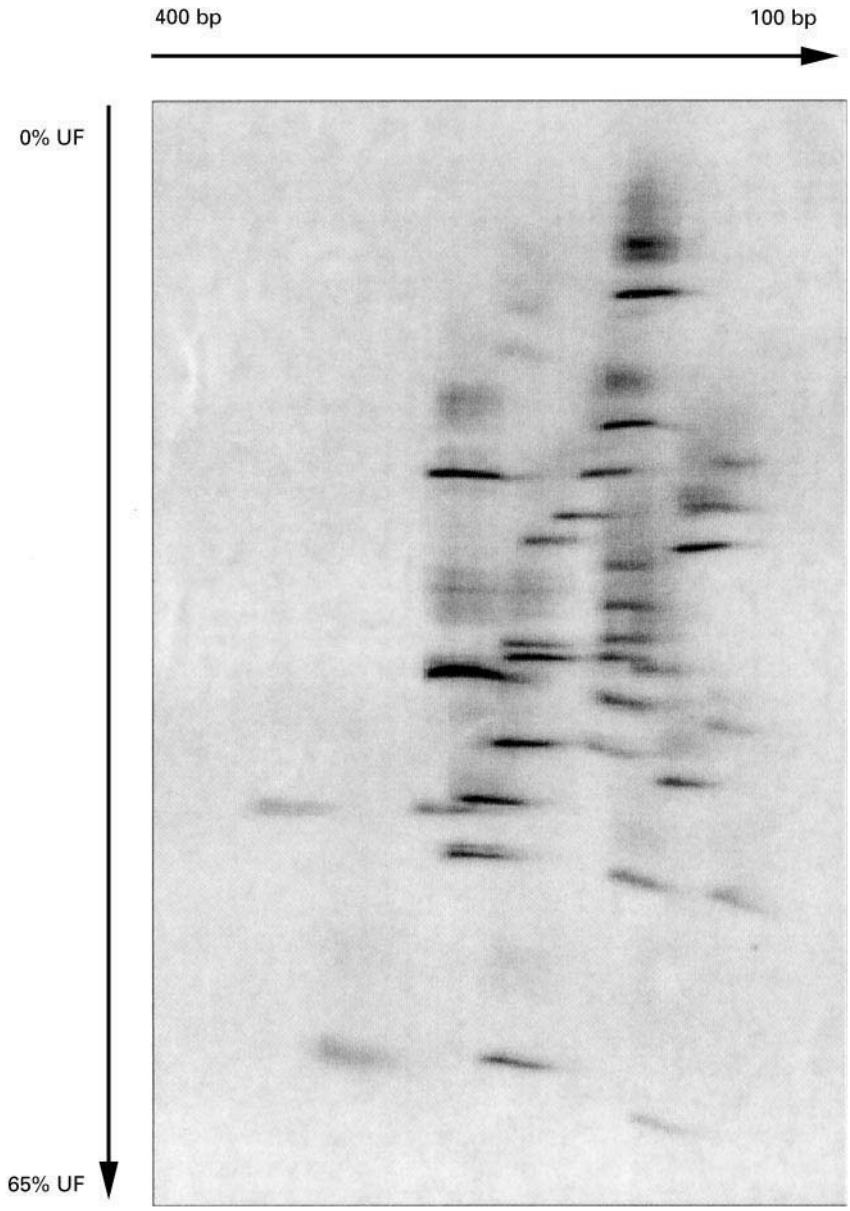


Figure 7 Empirical TDGS pattern of the retinoblastoma susceptibility gene *RB1*, containing a mutation in exon 2.

configuration. Such software should also be capable of storing two-dimensional patterns and link subsets of them in particular experiments requiring comparisons of large numbers of individuals. It could also provide for a sample tracking system.

Further Reading

Cotton RGH (1997) *Mutation Detection*. Oxford: Oxford University Press.

Dhanda RK, Smith WM, Scott CB *et al.* (1998) A simple system for automated two-dimensional electrophoresis: applications to genetic testing. *Genetic Testing* 2: 67–70.

Eng C and Vijg J (1997) Genetic testing: the problems and the promise. *Nature Biotechnology* 15: 422–426.

Fischer SG and Lerman LS (1979) Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. *Cell* 16: 191–200.

Lerman LS and Silverstein K (1987) Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. *Methods in Enzymology* 155: 501–527.

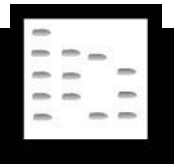
Sheffield VC, Cox DR, Lerman LS and Myers RM (1989) Attachment of a 40-base pair G + C rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proceedings of the National Academy of Science of the USA* 86: 232–236.

van Orsouw NJ, Li D, van der Vlies P *et al.* (1996) Mutational scanning of large genes by extensive PCR multiplexing and two-dimensional electrophoresis: application to the RB1 gene. *Human Molecular Genetics* 5: 755–761.

van Orsouw NJ, Dhanda RK, Rines DR *et al.* (1998) Rapid design of denaturing gradient-based two-dimensional electrophoretic gene mutational scanning tests. *Nucleic Acids Research* 10: 2398–2406.

Vijg J and van Orsouw NJ (1999) Two-dimensional gene scanning: exploring human genetic variability. *Electrophoresis* 20: 1239–1249.

GEOCHEMICAL ANALYSIS: GAS CHROMATOGRAPHY AND GC-MS



R. P. Philp, University of Oklahoma, Norman, OK, USA

Introduction

Geochemical analysis, and more specifically chromatography, is concerned with samples derived from two different sources: those of relatively recent origin, related to environmental problems; and those of a much greater geological age, related to fossil fuel exploration and exploitation. The chromatographic techniques utilized to analyse and characterize such samples are virtually identical regardless of the age and origin of the sample. The extracts from geochemical samples, whether they are rocks, soils, crude oil spills, contaminated wildlife or spills of refined products, are very complex mixtures of a wide variety of organic compounds. Compounds derived from fossil fuels typically will be complex mixtures of hydrocarbons, and the environmental samples from more recent sediments probably will contain a variety of other compounds such as chlorinated compounds, pesticides or herbicides. In view of the similarities of the techniques used for analysing the samples from these different sources, the majority of examples used in this article to illustrate the techniques will be based on the characterization of fossil fuel samples.

The major goal of any geochemical analysis is to take a sample and, through a variety of fractionations and analytical techniques, reach a point where either the presence or absence of specific target compounds can be determined, or fingerprints for specific classes of compounds can be obtained and used for correlation purposes. Applications related to petroleum exploration might use such fingerprints for oil–source rock or oil–oil correlation studies, whereas in environmental studies one is more concerned with correlating a spilled product with its original source material or trying to evaluate the extent of removal during clean-up procedures.

Geochemical samples are extremely complex mixtures of a wide variety of compound classes. The analytical techniques commonly used to characterize such mixtures involve some form of chromatography, such as gas chromatography (GC), gas chromatography–mass spectrometry (GC-MS), gas chromatography–mass spectrometry/mass spectrometry (GC-MS/MS), and more recently gas chromatography–isotope ratio mass spectrometry (GC-IRMS). Liquid chromatography (LC) and combined liquid chromatography–mass spectrometry (LC-MS) are also used in certain applications, but not to the same extent as GC and GC-MS. In addition to the analytical chromatographic separations, most geochemical analyses require some sort of fractionation into compounds classes prior to the actual analysis. There are certain cases where total sediment extracts or whole crude oils are analysed directly but generally the mixtures are so complex that an initial fractionation(s) is required to simplify the extracts for subsequent analyses. For example gas chromatograms of many crude oils (**Figure 1**) are dominated by *n*-alkanes but, for the most part, compounds that are of much greater geochemical importance are not readily observable in these chromatograms but are hidden in the baseline of the chromatogram. It should be noted that there are also many naphthenic crudes not dominated by *n*-alkanes, e.g. Venezuelan and Russian crudes. Most of these naphthenic crudes are either severely biodegraded or have been generated at relatively low levels of maturity from sulfur-rich kerogens. A fractionation step involving thin-layer chromatography, column chromatography or liquid chromatography, all of which involve partitioning of components between a liquid and solid phase, leads to the separation of