Review

Igor V. Kourkine
Christa N. Hestekin
Annelise E. Barron

Northwestern University,
Department of Chemical
Engineering,
Evanston, IL, USA

Technical challenges in applying capillary electrophoresis-single strand conformation polymorphism for routine genetic analysis

Recent and future advances in population genetics will have a significant impact on health care practices and the economics of health care provision only if a spectrum of patient-tailored, effective methods of DNA screening for sequence alterations has been developed. Genetic screening by capillary electrophoresis-single strand conformation polymorphism (CE-SSCP), which is based upon the differences in electrophoretic mobilities of wild-type and mutant DNA species, offers an important complement to other presently available techniques such as Sanger sequencing and DNA hybridization arrays due to its simplicity, versatility, and low cost of analysis. A two-part review of CE-SSCP that discusses its advantages and limitations is presented. Emphasis is placed on technological aspects of CE-SSCP (including such rarely addressed issues as sample preparation protocols and the nature of the polymeric DNA separation matrix) as well as on the potential of CE-SSCP for routine genetic analysis. An attempt is made to organize and present the information in sufficient detail to allow the use of SSCP for routine genetic screening even by those inexperienced in CE. Some discussion of CE-based heteroduplex analysis (HA) is also presented.

Keywords: Capillary electrophoresis / Heteroduplex analysis / Mutation detection / Review / Single-strand conformation polymorphism EL 4949

Contents

1 Introduction 1375
2 CE-SSCP parameter optimization 1376
2.1 PCR amplification 1376
2.2 Additives and thermal treatment of samples 1377
2.3 DNA size and sequence 1379
2.4 Temperature, buffer composition and pH . . . 1379
2.5 DNA separation matrices 1379
2.6 Capillary coating 1380
3 The outlook for routine use of CE-SSCP for genetic screening 1381
3.1 Sensitivity and specificity 1381
3.2 Throughput 1383
3.3 Direct mutation detection 1384
4 Conclusions 1384
5 References 1384

Correspondence: Professor Annelise E. Barron, Northwestern University, Department of Chemical Engineering, 2145 Sheridan Road, Evanston, IL 60208, USA
E-mail: a-barron@northwestern.edu
Fax: + 847-491-3728

Abbreviations: CAE, capillary array electrophoresis; **HA**, heteroduplex analysis; **LPA**, linear polyacrylamide

1 Introduction

The development of low-cost, high-throughput genetic screening technologies promises to enable a systematic search within populations for individuals possessing particular genotypes that are either associated with, predispose to, or directly lead to the development of disease [1]. The goals of genetic screening include medical diagnosis, management or treatment of illness, the provision of reproductive information, genetic mutation enumeration (estimation of prevalence of mutant alleles, their distribution and biological significance), and genetic research. Routine genotyping and genetic testing of large numbers of individuals is not without troubling implications and potential roadblocks, which can be broadly classified as either psychosocial or technical in nature. From the psychosocial viewpoint, it is important to differentiate between the screening for inherited mutations (germline mutations) and for genetic changes accumulated throughout the lifetime of an individual as a result of environmental exposures or other unexplained events (somatic mutations). Testing for somatic mutations is not troubled with the same "worrisome implications" as screening for inherited genetic mutations might be. For example, the detection of such "acquired" mutations hold no major ethical considerations such as the issue of impact on other family members or possible psychological trauma.

CE and CEC

© WILEY-VCH Verlag GmbH, 69451 Weinheim, 2002
0173-0835/02/1005–1375 $17.50+.50/0

From the technological perspective, there are real issues with the sensitivity of mutation detection. Routine genetic testing technology may fail to identify existent mutations in anywhere from 25 to 75% of the cases [2].

Single-strand conformation polymorphism (SSCP) is one of the most common, versatile, and inexpensive indirect methods for genetic screening, as evidenced by thousands of publications since its invention in 1989 by Orita *et al.* [3]. Unlike direct methods for genetic screening such as Sanger DNA sequencing or DNA hybridization arrays (*a.k.a.* GeneChips™), SSCP can determine the mutation status of a gene through a comparison to wild-type, but does not provide information on the exact nature of the mutation, *i.e.* the specific change in DNA sequence. SSCP is based on the principal that an alteration in DNA nucleotide sequence caused by a mutation will affect single-stranded DNA (ssDNA) folding and, hence, the electrophoretic mobility of a DNA fragment analyzed under native (nondenaturing) conditions. Traditionally, SSCP has involved time-consuming (up to 14 h), labor-intensive, and cumbersome electrophoresis of radioactively labeled DNA samples in a highly resolving separation medium such as a large-format, cross-linked 20% polyacrylamide gel. DNA bands typically were detected by long exposure of the gel to X-ray film (Fig. 1). These practical difficulties associated with slab-gel SSCP, which are ultimately technique-related, have resulted in selective and limited application of SSCP-based techniques in clinical medicine. The recent advent of capillary electrophoresis (CE), in which electrophoretic separation is carried out within arrays of microbore fused-silica capillaries of 50–75 μm inner diameter, not only provides a viable alternative to slab-gel SSCP, but also offers much higher throughput and sensitivity as well as greater reproducibility and full method automation [4, 5].

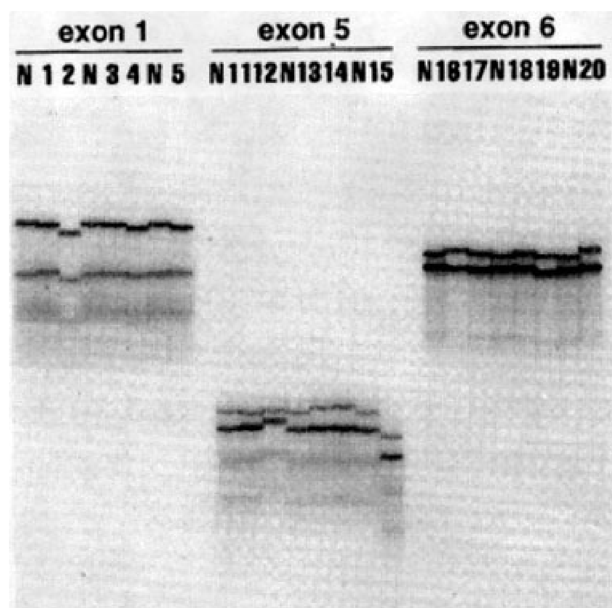

Figure 1. An example of genetic screening of the Factor IX exons 1 (424 bp), 5 (338 bp) and 6 (389 bp) by slab-gel SSCP, electrophoresed for 5 h. Lanes labeled N are loaded with wild-type controls, while those labeled 1–20 are loaded with samples of DNA obtained from hemophilia patients. Reprinted from Kukita *et al., Hum. Mutat*. 1997, *10*, 400–407, with permission.

While slab-gel SSCP protocols have been optimized extensively for sensitive mutation detection, precise, rational, and comprehensive information on the best methods for CE-SSCP is lacking. A few important CE-SSCP variables such as temperature, buffer composition and pH, and DNA size and sequence have been investigated with relative thoroughness. Others, such as the nature of the DNA separation matrix, capillary wall coating, and DNA sample preparation procedures, have received little or no attention. Moreover, information on CE-SSCP is presently fragmented and scattered throughout the literature, making it practically difficult for a novice in the field to start applying CE-SSCP to their needs.

In the first part of this review, we analyze the impact of DNA sample preparation methods and CE parameters on the sensitivity of mutation detection by CE-SSCP. In the second part, we outline some considerations related to the potential of CE-SSCP as an effective method for genetic screening in a clinical setting. The goal of this paper is to highlight the technical details of CE-SSCP protocols that are essential for reproducible and sensitive mutation detection, in order to facilitate the development of rational CE-SSCP protocols that will be useful in a clinical medicine, such as in the diagnosis and treatment of cancer.

2 CE-SSCP parameter optimization

2.1 PCR amplification

As indicated in many research papers and summarized in a recent review by Ren [6], two modes of detection can be used for CE-SSCP: fluorescence detection and UV (260 nm) detection. The use of UV detection dominated the field in the early 90's, but is now almost completely replaced by fluorescence detection due to the much higher sensitivity of the latter [7, 8]. Fluorescent labels may be introduced into dsDNA by PCR amplification with fluorescent primers, or by post-PCR fluorescent labeling using alkaline phosphatase and Klenow fragment enzymes. The latter method was pioneered by Inazuka *et al.* [9] and yields much cleaner and more homogeneous DNA samples, as well as easier-to-interpret CE-SSCP electropherograms. However, post-PCR labeling has yet

to gain widespread popularity, presumably because it is more involved than direct PCR primer fluorescent labeling, especially for the analysis of a small number of DNA samples (*i.e.,* without the aid of sample preparation robots). An alternative approach, recently described by Nishimura and Tsuhako [10], involves use of the intercalating dye YO-PRO-1, which was added to the DNA separation matrix and used for fluorescence detection of unlabeled SSCP conformers of the *N-ras* oncogene. The authors tested a variety of other intercalating dyes, none of which resulted in successful mutation detection. They also optimized the YO-PRO-1 concentration (optimum 0.2 μM) and observed that high concentrations of YO-PRO-1 led to decreased resolution, which was attributed to "disruption of higher-order structure of the ssDNAs [10]." Similarly, Zhang *et al.* [11] successfully utilized thiazole orange, another intercalating dye, for SSCP mutation detection in the *K-ras* gene.

The need for post-PCR purification is not well documented for CE-SSCP. Crude PCR samples contain a multitude of charged species including salt, unincorporated fluorescently labeled primers, excess dNTPs, and enzymes that reduce the efficiency of electrokinetic injection, which is the preferred way to introduce DNA samples into the capillary. More importantly, however, unincorporated primers can anneal to SSCP conformers to form primer-ssDNA constructs during denaturing and cooling prior to CE. The electrophoretic mobilities of these ssDNA-primer constructs can be very similar to those of SSCP conformers, and can complicate peak assignments [12].

Partial purification of crude PCR products by use of a Microcon-PCR (Millipore, Bedford, MA, USA) filtration device or QIAquick PCR Purification Kit (Qiagen, Valencia, CA, USA) reduces the amount of primer-ssDNA constructs [6, 13], while complete purification of crude PCR samples by preparative slab-gel electrophoresis can eliminate them entirely [14]. If additional PCR product purification is undesired or impossible, a similar but less pronounced beneficial reduction in primer-ssDNA annealing can be achieved by a mere decrease in initial primer concentration during PCR amplification [14]. Debernardi *et al.* [15] suggested an unconventional approach for pre-electrophoretic sample purification, in which they used biotinylated primers to perform PCR amplifications. Following the PCR amplifications, the resulting biotinylated DNA amplicons were purified from the buffer, salts, dNTPs, and primers with Dynabeads (M-280 Streptavidin; Dynal Biotech, Lake Success, NY, USA) and used for CE-SSCP with UV detection.

Fluorescent labeling of both the forward and reverse DNA strands offers several advantages over labeling of just one DNA strand, and requires little extra effort and moderate additional cost [16]. Notably, it increases the chances of mutation detection, as evidenced by many studies [17–19]. Two-color fluorescent labeling, in which the forward DNA strands are labeled with one type of dye and the reverse DNA strands with another type of dye, is superior to one-color fluorescent labeling in which only one dye is used. Not only does it allow for a straightforward identification of the peaks representing the forward and reverse ssDNA strands, but it also provides the means to differentiate between ssDNA and dsDNA species, including primer-ssDNA constructs, as described below [14, 18, 20].

We have observed that in two-color detection mode (forward strand – FAM, reverse strand – JOE), dsDNA exhibits two perfectly overlapping signals of two different colors in CE-SSCP electropherograms, while ssDNA shows only one major peak, of one color (often accompanied by a minor peak of the second color, Fig. 2) [14]. The minor peak, which can be attributed to minor overlap of the emission spectra of the two dyes, is centered in the same region as the major peak, but is of much lower intensity. We have used this feature to precisely and unambiguously assign all the peaks in electropherograms of 32 mutant DNA samples of the *p53* gene analyzed by combined CE-SSCP/heteroduplex analysis (HA) [20]. On the other hand, when a one-color detection mode (both strands – FAM) is employed, differentiation between SSCP conformers and ssDNA-primer constructs is virtually impossible, because the peaks look very similar in the two cases and have similar electrophoretic mobilities.

Attaching fluorescent dyes to DNA can lead to certain complications such as changes in DNA electrophoretic mobility due to the charge, hydrophobicity, or steric bulk of the dye [21]. If wild-type and mutant DNA are mixed and analyzed simultaneously, this complication can be addressed in a practical manner by labeling of the two corresponding ssDNA strands with the same dye (*i.e.,* both forward strands – FAM; both reverse strands – JOE), to cancel out or at least minimize these effects. Little is known about whether the nature of the dye affects folding of ssDNA during the formation of SSCP conformers, but it is very likely that it may have some effect. For the reasons of differences in mobility based on differences in dye structure, we do not recommend the use of four different dyes (*i.e.,* the use of different labels on each of the mutant and wild-type forward and reverse strands).

2.2 Additives and thermal treatment of samples

Denaturing agents, such as formamide, are commonly used to dilute CE-SSCP DNA samples prior to thermal denaturation. To our knowledge, there are no published

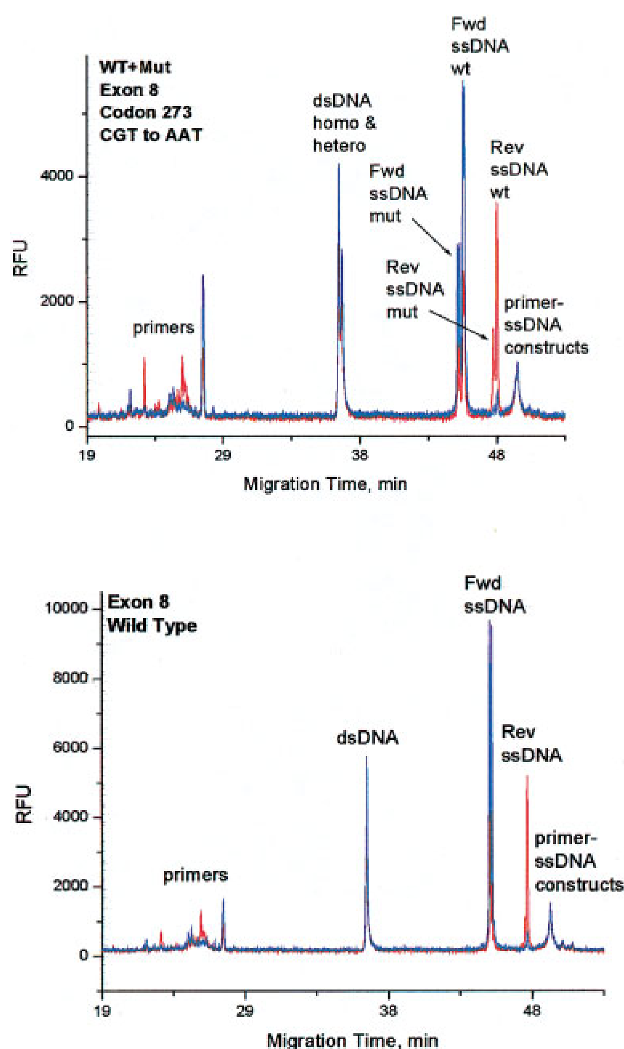

Figure 2. An example of genetic screening of a mutant DNA sample derived from the *p53* exon 8 (200 bp) by tandem capillary electrophoresis SSCP/HA using DNA samples prepared according to optimized protocols [20]. MegaBACE™ CAE instrument; total capillary array length 64 cm; effective capillary array length 40 cm; capillary inner diameter 75 μm; bare fused-silica dynamically coated with poly-*N*-hydroxyethyl acrylamide (polyDuramide™), LPA DNA separation matrices (6% w/v M_w 600 kDa); pre-electrophoresis 5 min at 140 V/cm; injection 230 V/cm for 25 s; electrophoresis at 180 V/cm at 27°C with an associated current of 10–12 μA.

reports that explain the reason for the adoption of this approach. However, in a related study, Casey and Davidson [22] investigated the melting of dsDNA and found that formamide simultaneously decreased the melting temperature (T_m) of the samples and improved the chances for complete denaturation. Elsewhere it was reported that high formamide concentrations (*i.e.,* 95%) depressed the T_m of dsDNA by as much as 60°C [23]. Therefore, we suppose that formamide has been added to SSCP samples prior to denaturation in order to ensure the completeness of initial DNA strand dissociation and unfolding.

Interestingly, Yip *et al.* [24] have added denaturants (formamide or urea) to the running buffer in slab-gel SSCP to “sharpen the bands of ssDNA fragments and clear the background that might otherwise obscure the bands that are diagnostic of a polymorphism,” and cited other studies that have used these and other additives (*i.e.,* SDS or PEG) for the same purpose. To explain this approach, the authors suggested that a mildly denaturing environment in the gel created by formamide reduced the number of different SSCP conformers and thus cleared up the smearing of DNA bands. Another potent denaturant, methylmercuric hydroxide (CH_3HgOH), was used by Weghorst and Buzard [25] to improve the crispness of DNA bands by facilitating sample denaturing and preventing the formation of partially annealed conformations that might be confused with true SSCPs. Sodium hydroxide is another chemical that is commonly added in a small amount to the DNA solution prior to denaturation, presumably to assist DNA denaturation by increasing pH. One report suggested that NaOH did not improve DNA denaturation, but did affect the mobility of SSCP conformers [18]. Clearly, the pH and ionic strength of the sample solution are both strongly affected by the addition of NaOH. However, another report has shown that high pH (11–12) can be used to fully denature DNA for sequencing [26].

Salt concentration in the sample is another variable that can have a significant impact on the outcome of CE-SSCP analysis. Prior to electrophoresis, PCR products are usually diluted 10–100 X in aqueous buffer or formamide to prevent fluorescence detector saturation and to conserve the sample, a process that affects the final salt concentration present in the sample. Although concentrations of > 50 mM NaCl (or KCl) are generally used for the determination of T_m of dsDNA to ensure correct nucleotide base pairing, no strict guidelines exist for salt content in samples to be analyzed by CE-SSCP, as required to effect partial nucleotide base pairing. An increase in salt concentration from 20–50 to 100 mM is reported to raise the T_m of dsDNA by 5–10°C, showing an opposite trend to formamide effects on the T_m and potentially complicating dsDNA denaturation [23]. At the same time, high salt concentrations will definitely decrease the amount of sample injected into a capillary electrokinetically, and thus can hamper mutation detection by decreasing the signal-to-noise ratio in the analysis.

We have recently optimized sample preparation methods for combined CE-SSCP/HA for PCR products amplified from the *p53* gene, and determined that no formamide or

extra salt are required if a PCR product is diluted 20–40 X with 1 X Tris-HCl buffer (pH 8.5) [14]. We also observed that snap-cooling the sample on ice reduced the percentage of primer-ssDNA constructs in the final mixture, in comparison to the distribution obtained when one allows the denatured dsDNA sample to cool at ambient temperature. Although addition of formamide may be helpful for the accomplishment of full, initial denaturation of DNA samples with high GC content, at this point, we consider its use in routine CE-SSCP to be unnecessary.

2.3 DNA size and sequence

About a dozen different genes have been analyzed by CE-SSCP, of which *p53*, ras, BRCA1, and BRCA2 have received the most attention. It is definitely easier to detect mutations in some gene regions than others, perhaps reflecting the impact of DNA sequence on chain folding. Many CE-SSCP researchers have concluded that as DNA fragment length increases, the sensitivity of mutation detection by CE-SSCP decreases [6]. While the optimal DNA size for SSCP analysis is believed to be in the range of 175–400 bp, even longer DNA fragments (503–650 bp [27]; 740 bp [9]; 1223 bp [28]) have been analyzed for mutations with reported success. At the other extreme, Arakawa *et al.* [8] have reported CE-SSCP analysis of DNA fragments as short as 71 bp.

The GC content of the DNA sequence also seems to have some impact on how easy it is to detect mutations by CE-SSCP. Nataraj *et al.* [29] reported that they had better mutation detection sensitivity for fragments with high GC content ($\sim$60%) than for fragments with lower GC content ($\sim$40%), under the same conditions of the analysis. Presumably, the greater number of G-C base pairs provides SSCP conformers of greater average stability. Although one might expect that the sensitivity of CE-SSCP mutation detection should be lower for sequence alterations located near the ends of DNA fragments (as in the case for HA), in at least one study no such correlation was observed [28].

2.4 Temperature, buffer composition and pH

These three variables have been discussed thoroughly by Ren [6] in his recent review. Herein, we would like to reiterate that the optimum temperature and pH conditions for sensitive mutation detection seem to be highly sequence-specific, and probably cannot be determined *a priori* because of the absence of a first-principles understanding and prediction of three-dimensional SSCP structures due to the idiosyncratic nature of DNA folding. Ellis *et al.* [21] attempted to model SSCP folding using the program Mfold, but the results were not consistent enough to be applied. While analysis temperatures between 20 and 30°C typically are recommended as a starting point in temperature optimization for attempts to analyze novel DNA fragments by SSCP, one needs to keep in mind that lower or higher temperatures may in fact provide more sensitive analysis. Tris-borate-EDTA (TBE) running buffer that contains 10% glycerol remains the top choice in most CE-SSCP protocols, but other buffers such as Tris-HEPES and Tris-MES-EDTA have also been used successfully [13]. Gelfi *et al.* [30] evaluated the combined effects of temperature, sieving matrix, applied voltage, capillary length, capillary inner diameter, and pH on CE-SSCP, and noted a significant improvement in the resolution and sensitivity of CE-SSCP when a low-pH buffer (Tris-MES-EDTA, pH 6.8) was used; at the same time, no beneficial changes were recorded when the temperature of analysis was lowered.

2.5 DNA separation matrices

In CE, the "sieving" of DNA is accomplished with fluid solutions of highly entangled, water-soluble polymers. These entangled polymers form a dynamic mesh of obstacles through which the SSCP conformers must negotiate or force a path. To engender differences in electrophoretic mobility for these small, folded DNA strands, which differ only in their shape, requires a relatively dense polymer solution. Normally, for the high-resolution separation of any distinct class of DNA samples (oligonucleotides, dsDNA fragments, DNA sequencing fragments), the polymer matrix must be optimized with respect to the chemical nature of the polymer, physical nature of the polymer (polymer molar mass distribution), and polymer concentration in the buffer.

The same is, of course, true for SSCP conformers; it is not yet clear what the true optimum matrix for CE-SSCP will be, as significant work remains to be done in this area. As described in several studies and summarized in the recent review by Ren [6], a number of different polymers such as hydroxyethylcellulose (HEC), polyvinylpyrrolidone (PVP), methylcellulose (MC), POP™ from ABI and linear polyacrylamide (LPA) have been applied to CE-SSCP analysis, with varying degrees of success. A related novel copolymer, LPA-polydimethylacrylamide (PDMA) was recently shown by Ren and Fang [31] to have good potential for use in CE-SSCP, but has yet to be optimized with respect to polymer molar mass and concentration. The stated advantage of LPA-PDMA is that it combines the excellent sieving ability of LPA with the intrinsic capillary-coating properties of PDMA, and hence allows for SSCP analysis in uncoated capillaries. These studies shed some light on the chemical compositions suitable

for polymers employed in CE-SSCP. However, there is a relative paucity of studies aimed at investigating the influence of the physical nature of the polymers (average molar mass (M_w), persistence length in aqueous solution, and polydispersity) on CE-SSCP performance. Previously, Barron *et al.* [32] showed for hydroxypropylcellulose (HPC) polymers that a minimum M_w of 300 kDa is required for the decent separation of dsDNA larger that 1 kbp, and also determined that "stiff" polymers such as HEC can be used in more dilute solutions than HPC or LPA to afford similar DNA separation ability. SSCP conformers are much more compact, and hence more challenging to separate than dsDNA fragments. Therefore, it is likely that adjusting the physical nature of the polymer that acts as a CE-SSCP matrix will also have significant impact on its separating ability and optimum useful concentration. To our knowledge, there is only one study that has touched upon this subject. Ren *et al.* [33] demonstrated that homemade, short-chain LPA could be used for SSCP detection of single point mutations in a number of genes, and yielded DNA separation matrices that also were relatively easy to load into the capillary because of their low viscosity. No exact information on the LPA molecular weight, or the effects of M_w on CE resolution, were provided.

This present lack of published information on the impact of polymer physical characteristics on CE-SSCP separation is most likely due to the general absence of the means and/or the desire among the majority of CE-SSCP researchers to synthesize and precisely characterize polymeric materials suitable for CE. Most of the aforementioned polymers used in CE-SSCP (HEC, HPC, MC) are modified natural products, and are only available in certain molecular weight ranges, which also vary substantially from batch to batch. Therefore, these natural polymers are unsuitable for precise and reliable studies of the effects of M_w on the sensitivity of mutation detection by CE methods.

In the absence of a large choice of polymers, some efforts have been directed at optimizing polymer concentrations, primarily POP™, in the DNA separation matrix. POP™ is a polymer of undisclosed chemical nature and molecular weight, which has been commercially available as a DNA separation matrix (dissolved form) only from Applied Biosystems for a number of years. Not surprisingly, most of these studies found that higher POP™ polymer concentrations (up to 5%, the maximum (stock) concentration), which were increasingly difficult to use because of the increased viscosity of the polymer solutions, yielded better results [9, 17, 18]. In addition to POP™, a few other polymers have also been briefly investigated for CE-SSCP, and followed the same trend of the dependency of their separation ability on concentration. For instance, Nishimura and Tsuhako [10] reported that an optimal MC concentration was 1%, while matrices with MC concentration $< 0.7\%$ afforded no resolution and those with MC concentration $> 1.2\%$ were too viscous to load into the capillary. Ren *et al.* [34] studied 2–6% LPA solutions as SSCP separation matrices and reported that 6% provided the best resolution.

Recently, we have conducted what we believe is the first systematic study of the impact of the molecular weight of LPA on CE-SSCP analysis using a high-throughput MegaBACE™ (Molecular Dynamics, Sunnyvale, CA, USA) capillary array electrophoresis (CAE) instrument [20]. In our study, we simultaneously varied the concentration (2, 4, 6% w/v) and M_w (200, 600, 1000 kDa, and 2400 kDa) of LPA, and applied a combined CE-SSCP/HA analysis to more than a dozen different mutant DNA samples derived from exons 7 and 8 of the *p53* gene (size 139–300 bp). The optimum DNA separation matrix for these samples was determined to be 6% w/v of 600 kDa LPA. LPA of lower molar mass (200 kDa) did not provide adequate SSCP resolution at any concentration, while the use of a higher molar mass (2400 kDa) polymer did not offer any advantage over the optimum M_w, and furthermore complicated capillary loading due to high viscosity of its solutions (Fig. 3). Use of the M_w 2400 kDa polymer matrix also led to significantly longer analysis times (1 h *vs.* 20–30 min for lower-M_w polymers). We plan to extend this study to longer DNA fragments and to other polymer compositions (*i.e.,* more hydrophobic PDMA) in an attempt to establish the generality of the results cited above, and to formulate a universal SSCP separation matrix (*i.e.,* mixtures of high and low M_w versions of the same polymer or, perhaps, polymers of different chemical compositions) that will allow for a high sensitivity of SSCP detection without case-by-case fine-tuning of other variables, most importantly, the temperature of analysis.

2.6 Capillary coating

Another greatly understudied CE-SSCP variable is the nature of the polymer coating applied to the inner walls of a capillary, which is aimed at greatly reducing the velocity of electroosmotic flow and minimizing the adsorption of the analytes to the capillary wall. While a detailed description of capillary wall coatings is beyond the scope of this review, we would like to stress that the presence of a stable coating with the proper chemistry has the utmost impact on the sensitivity of CE-SSCP mutation detection, in our experience. One particularly relevant study came from Tian *et al.*, [35], in which the authors emphasized that the efficacy of CE-HA, a genetic screening method closely related to CE-SSCP, was dependent on both the

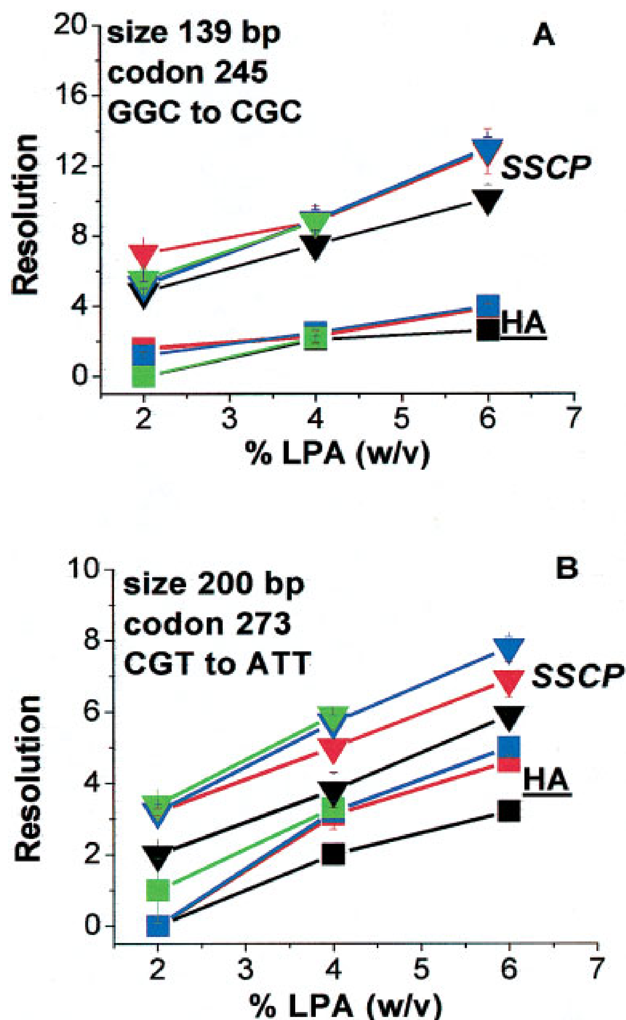

Figure 3. Combined impact of LPA M_w and matrix concentration on the resolution of capillary CE-SSCP and HA samples derived from (A) exon 7 and (B) 8 of the *p53* gene, performed in (--▼---■--) LPA M_w 200 kDA; (--▼---■--) LPA M_w 600 kDa; (--▼---■--) LPA M_w 1 kDa; (--▼---■--) LPA M_w 2.4 MDa. ▼, SSCP samples; ■, HA samples [20]. Conditions as in Fig. 2.

effective passivation of the capillary surface and on the choice of the correct polymer matrix for sieving. The authors tested a variety of silanizing reagents, polymer coatings, and DNA separation matrices for the analysis of fragments of the BRCA1 gene (DNA size *ca.* 200 bp) and found that the optimum performance for CE-HA was achieved when chlorodimethyloctasilane, PVP, and HEC were used in combination as the silanizing reagent, protective wall coating, and separation matrix, respectively (Fig. 4).

In a recent study, we have noticed that, under otherwise identical conditions, bare fused-silica capillaries dynamically coated with polyDuramide™, a novel adsorptive polymer jointly developed by BioWhittaker Molecular Applications (Walkersville, MD, USA) and our laboratory, provides much better CE-SSCP resolution than commercial capillary arrays covalently coated with LPA [20]. Interestingly, the same LPA-coated capillaries continued to provide high-quality DNA sequencing data while failing for SSCP, suggesting that the nature of folded ssDNA conformers and/or the CE conditions for SSCP place more stringent demands on the quality of the protective capillary wall coating. In particular, the absence of urea during SSCP separations may increase the tendency for analyte adsorption on the wall, which can lead to band-broadening and hence loss of peak resolution.

It is noteworthy that the pH of the buffer has a pronounced impact on the stability of some capillary wall coatings. For example, Ren *et al.* [34] showed that at pH 8.3 a dynamically applied dimethylacrylamide coating deteriorated faster than at pH 7.8. Similarly, we have noted that an LPA coating covalently bound to silica degrades rapidly during CE-SSCP analysis (*vide supra*) and compromises the sensitivity of mutation detection. This suggests a possibility that the apparent improvements in CE-SSCP sensitivity of mutation detection in low pH buffers may be at least partially due to the higher stability of protective capillary wall coatings at low pH.

3 The outlook for routine use of CE-SSCP for genetic screening

3.1 Sensitivity and specificity

For clinical genetic screening, it is essential to know the frequency of the incidents of false positives (specificity = 100 – frequency of false positives) and of false negatives (sensitivity = 100 – frequency of false negatives) of mutation detection in CE-SSCP. Both the specificity and the sensitivity of CE-SSCP mutation detection need to be at least 97%; lower sensitivity would impair confident mutation detection, while a lower specificity would unnecessarily disturb too great a number of patients.

The sensitivity of CE-SSCP has been anecdotally evaluated in several different studies, and is reputed to approach 100% if more than one experimental condition is employed [36–39]. However, the specificity of CE-SSCP has been much less investigated. According to some studies, SSCP analysis has produced no false positives [21, 37]. In another study, O'Connell *et al.* [36] analyzed 15 samples and identified 3 false positives. Elsewhere, Liechti-Gallati *et al.* [39] confirmed that some false positives resulted from the annealing of primers to SSCP conformers, and the consequent production of spurious peaks. The presence of these constructs led to the identi-

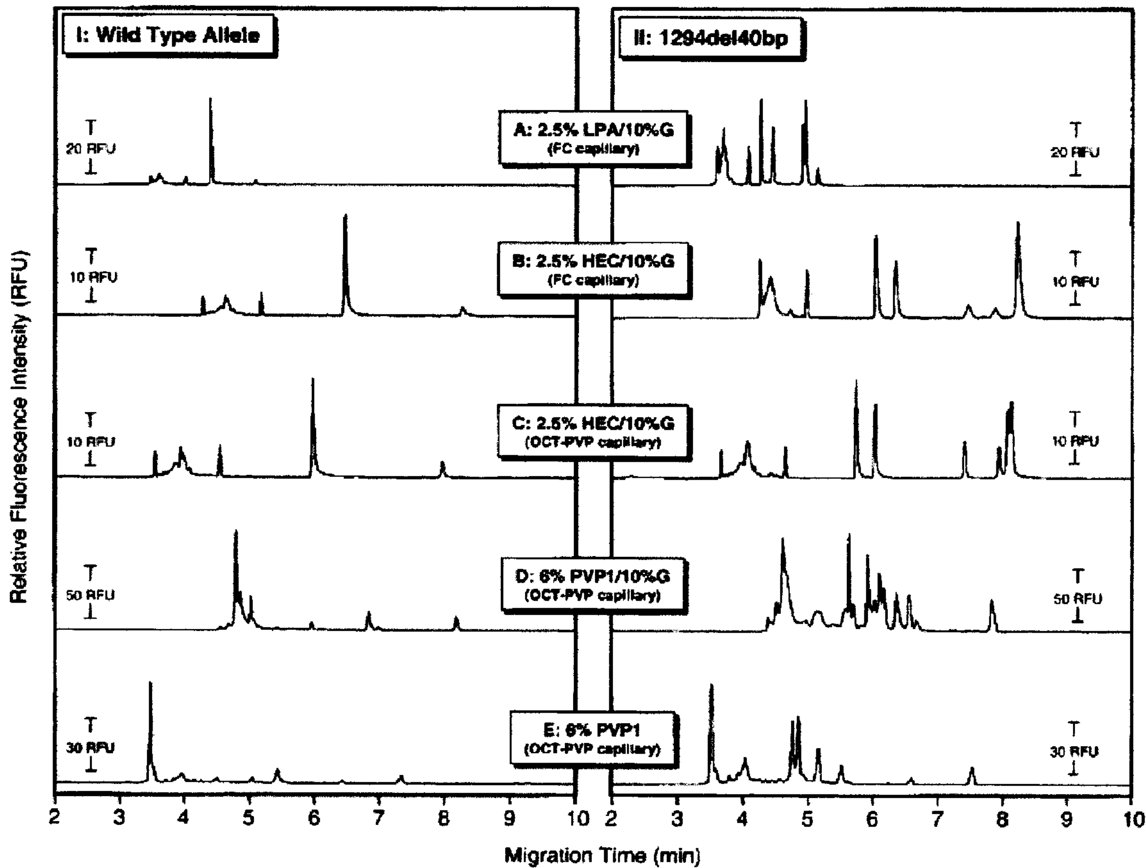

Figure 4. Comparison of different capillary wall coatings and separation matrices for CE based heteroduplex analysis. Reprinted from [35], with permission.

fication of some of the wild-type samples as mutants. While Ru *et al.* [37] concluded from their results that false positives should be more common in slab-gel SSCP than in CE-SSCP, Hayashi [12] stated in a recent review that false positives are unavoidable in genetic mutation detection in general since PCR products (which can be contaminated and/or imperfectly amplified by the enzymes) are used as starting materials for analysis.

Very few blinded studies, which can statistically assess the frequencies of false positives and false negatives, have been done in CE-SSCP. Notably, Wenz *et al.* [40] conducted two blind CE-SSCP studies. In the first study, 20 DNA samples containing two common mutations implicated in hereditary hemochromatosis were screened, and the results were in accord with those generated by restriction fragment length polymorphism (with 100% sensitivity). In the second study, ten *p53* gene exon 7 samples were analyzed by CE-SSCP against four reference samples, and again CE-SSCP correctly assigned the mutation status with 100% sensitivity [17]. Inazuka *et al.* [9] reasoned that the use of sophisticated computer software to assign the DNA mutation status should eliminate the need for blind studies. However, as long as a human factor is involved, blind studies are required for clinical assay validations regardless of the tools used for the data analysis.

The sensitivity of mutation detection by CE-SSCP and, as a matter of fact, most other mutation screening methodologies, is evaluated against and therefore cannot be better than the sensitivity of direct Sanger dideoxynucleotide sequencing (considered to be the "gold standard") in detecting sequence alterations. However, the latter method is not infallible either. For instance, Ahrendt *et al.* [41] used direct DNA sequencing to analyze a conserved region of the *p53* gene by slab-gel with automated fluorescence detection and detected only 76% of the mutations present (no information on number of repeat analyses is given). On the other hand, using the *p53* GeneChip™ hybridization assay, the authors detected 46 of 53 missense mutations (88% sensitivity) but none of the 5 frameshift mutations analyzed (0% sensitivity). Since even direct Sanger DNA sequencing or hybridization analysis does not give 100% sensitivity from a single analysis, it is important to ensure that sequencing of both the

forward and reverse DNA strands is done, with a sufficient number of repeats (3–5), to provide an accurate assessment of the genetic mutation methodologies to be tested.

Perhaps the best way to improve the sensitivity of an indirect mutation detection method such as CE-SSCP is to combine it with HA [39]. Similarly to SSCP, HA is based on the premise that dsDNA consisting of perfectly complementary strands (homoduplexes), will oftentimes exhibit different (usually higher) electrophoretic mobilities than "imperfect" DNA duplexes that contain nucleotide base mismatches (heteroduplexes). HA is reputed to be more useful for the detection of deletion and insertion mutations than for single-base substitutions, most of which typically remain undetected, especially in a slab-gel electrophoresis format [29, 42–44]. Typically, DNA samples being prepared for SSCP analysis are denatured at 95°C and cooled rapidly, while those destined for HA are denatured and cooled slowly over a period of 30–60 min to ensure complete dsDNA re-annealing. Combined SSCP/HA can be conducted by separately preparing SSCP and HA samples and analyzing them either simultaneously, or in a sequential fashion [39, 45, 46]. Recently, Ravnick-Glavač *et al.* [47] observed in their slab-gel electrophoresis experiments that SSCP/HA conformers suitable for tandem analysis can be generated simultaneously in the same vessel by utilizing the SSCP cooling protocols. Previously, doubts remained about whether the kinetics of annealing during fast cooling might yield incompletely re-annealed duplexes, that could result in false positives [48]. In the more recent study, Kozlowski and Krzyzosiak [27] for the first time demonstrated the possibility and utility of combined SSCP/HA in CE. In the latest development, we have investigated the effects of DNA sample preparation protocols and the physical nature of the DNA separation matrix on the resolution and sensitivity of combined CE-SSCP/HA performed in high-throughput CAE mode [14, 20].

In terms of sensitivity, Axton *et al.* [49] have reported that combined slab-gel SSCP/HA was able to detect 12 out of 12 (100%) mutations in the PAX6 gene, five of which were deletions and insertions and the rest single-base substitutions. Slab-gel SSCP alone correctly identified 10 out of 12 mutations; the remaining two undetected mutations (one deletion, one insertion) were identified by HA. It was unclear from the report what were the sizes of the analyzed DNA fragments. Ravnick-Glavač *et al.* [47] applied combined slab-gel SSCP/HA to the analysis of 27 mutated specimens of the CFTR gene (size 309–579) and were also able to achieve 100% sensitivity. Two mutations (one deletion, one insertion), which were indistinguishable by SSCP, showed different HA patterns. The authors again emphasized the need for the use of multiple running conditions for the most sensitive mutation detection, but identified a single gel condition (10% acrylamide/1.3% C; 10% glycerol) that produced a reasonably high sensitivity.

Recently, Kozlowski and Krzyzosiak [27] analyzed 31 mutant DNA samples of BRCA1 and BRCA2 (size 200–500 bp), most of which were single-base substitutions, by combined CE-SSCP/HA, and reported 100% sensitivity. For comparison, CE-SSCP alone detected 28 out of 31 mutations ($\sim$90%), and CE-HA alone detected 25 out of 31 ($\sim$80%) mutations. As expected, HA in CE mode was much more capable of detecting single-base substitutions than in a slab-gel electrophoresis mode. Interestingly, the authors reported two mutations in the BRCA1 gene (4427T/C and 3667 A/C) that actually showed two peaks in HA, one of which was assigned to two overlapping heteroduplexes and one homoduplex and the other of which was attributed to the second homoduplex. Thus, in these cases the presence of the mutation could be detected by the difference in electrophoretic mobilities of the two homoduplexes. Based on this observation, the authors argued that the correct name for the analysis should be SSCP/duplex analysis (DA) and not SSCP/heteroduplex analysis.

A potential problem with the use of SSCP for long-range genetic screening is the high occurrence of single nucleotide polymorphisms (SNPs) in the human genome (*ca.* 1 in 1000 bases in coding regions), which increases the probability of an SNP occurring in a large fragment of DNA that might be analyzed. Orban *et al.* [50] observed that SNPs masked detection of some mutants of BRCA1 and BRCA2 in slab-gel electrophoresis SSCP. However, Larsen *et al.* [45] have suggested that the high repeatability of SSCP, when run under the proper conditions, could be used to generate mutation and SNP "fingerprint" databases against which samples could be referenced. More studies need to be conducted to determine the extent of SNP effects on mutation detection. Notably, CE-SSCP has also been used for SNP detection [51, 52].

3.2 Throughput

Multiplexing, which involves the screening of a large number of DNA samples simultaneously, is desirable for clinical applications because of its higher throughput. Using multicapillary CE instruments equipped with multicolor laser-induced fluorescence detectors (*i.e.* the 4-color Applied Biosystems 3100™ or 3700™, or the MegaBACE™ CAE instrument), and labeling DNA with four different dyes instead of just two can further increase the throughput. In practice, Walz *et al.* [53] reported the ability to detect mutations in "seven pairs of exons" at the same time by multiplexing, without any loss of their SSCP-detection ability in

CE. However, Larsen *et al.* [45] in their study of the LQTS gene by combined CE-SSCP/HA cautioned that multiplexing might lead to decreased sensitivity of the analysis, due to the large number of DNA bands present, some of which may mask abnormal conformers. Optimization of the size of the DNA fragments analyzed by SSCP, as well as the use of high-resolution CE methods, can at least partially alleviate this problem [38, 54]. Another problem in multiplexing is that various rates of PCR amplification for different templates may result in unequal amounts of PCR end-products, which can lead to selective "drop-out" or overloading of some SSCP species relative to others [38].

3.3 Direct mutation detection

Generally, CE-SSCP is classified as an indirect mutation detection method that is capable of documenting the absence or presence of mutations only, but that cannot determine their position or nature. However, several reports have suggested that SSCP profiles are highly mutant-specific. For instance, years ago, Ravnik-Glavač *et al.* [47] observed that their samples gave distinct patterns in slab-gel electrophoresis that were "directly recognizable". In the previously mentioned study on the *p53* gene, Wenz *et al.* [17] screened ten unknown mutants against a panel of four reference mutants. Based on electrophoretic peak patterns, six samples were identified as having reference mutations, two were identified as having wild-type status, and two were shown to possess nonreference mutations. Characteristic SSCP/HA peak patterns could probably be determined for all of the "hotspot" mutants of the *p53* gene, to facilitate the efficiency of analysis. Larsen *et al.* [28] reported that according to their CE-SSCP analyses, each of 34 mutant DNA samples of the KVLQT1, HERG, and MYH7 genes possessed its own distinct SSCP pattern, and that each allele could be identified without subsequent DNA sequencing [28].

Several studies have addressed the issue of the high reproducibility of CE-SSCP that would be required for reliable DNA fingerprinting. For instance, Ellis *et al.* [21] noted that they had less than 0.02% standard deviation of peak migration times during multiple runs of their control [21]. Nishimura *et al.* [10] observed a relative standard deviation of 0.5% for the migration time and mobility of each peak for 9 repeat injections. Wenz *et al.* [17] found that their relative standard deviation was 0.01–0.05% for multiple injections of the same sample. Atha *et al*. [55] reported that standard deviations of the mobility of the sense and antisense DNA strands of the *p53* gene samples were 2.1 scans (0.47 s) and 0.79 scans (0.18 s). Hence, even though SSCP may not have a general reputation for high precision and reproducibility, these recent capillary-based studies would suggest that it deserves one.

4 Conclusions

In conclusion, we would like to cite B. Vogelstein [56], a well-known cancer researcher at John Hopkins University, who pointed out in a recent publication in *Science* that "Genetic testing in the future will no doubt involve a combination of methods designed to fit the mutation spectrum of specific patients and genes. . . . The best diagnostic medical tests have sensitivities and specificities approaching 100%, and it is not overly optimistic to expect that genetic testing will meet such exacting standards in the future". CE-SSCP, either alone or in combination with HA, may provide a cost-effective alternative for cases in which the sole purpose of analysis is to determine the presence or absence of a mutation to reduce the total frequency and cost of using direct mutation detection methodologies, which provide information on the precise nature of the mutation. In the latter case, DNA specimens can be efficiently prescreened by CE-SSCP/HA to identify mutant DNA specimens, which may then be analyzed by DNA sequencing, DNA hybridization arrays, or any other suitable direct technique.

Given the generally much higher cost of methods for direct mutation detection in comparison to CE-SSCP/HA, and the fact that only a relatively small percentage of patient samples may possess mutations (*i.e., ca.* 25% for the *p53* gene in breast cancer patients) CE-SSCP/HA can be used as a complementary method that will provide cancer researchers with significant cost and efficiency benefits. Further progress in the development of optimal materials and methods for CE-SSCP/HA analysis, as well as the implementation of these technologies on large sets of real patient samples to quantitate sensitivity and specificity for different genes of medical importance, will facilitate the transition of this technology into the clinical environment. Such routine application of low-cost genotyping technologies promises to make a substantial positive impact on human health in the new millennium.

Received December 17, 2001

5 References

[1] Gitzelman, R., *Hum. Genet. B: Med. Aspects* 1982, *84*, 425–436.

[2] van Ommen, G. J., Bakker, E., den Dunnen, J. T., *Lancet* 1999, *354*, SI5.

[3] Orita, M., Suzuki, Y., Sekiya, T., Hayashi, K., *Genomics* 1989, *5*, 874–879.

[4] Beale, S. C., *Anal. Chem.* 1998, *70*, 279R-300R.

[5] Haleem, I. J., *Electrophoresis* 2000, *21*, 1921–1939.

[6] Ren, J., *J. Chromatogr. B* 2000, *741*, 115–128.

[7] Oto, M., *Methods Mol. Med.* 1999, *27*, 127–138.

[8] Arakawa, H., Tsuji, A., Maeda, M., Kamahori, M., Kambara, H., *J. Pharm. Biomed. Anal.* 1997, *15*, 1537–1544.

[9] Inazuka, M., Wenz, H. M., Sakabe, M., Tahira, T., Hayashi, K., *Genome Methods* 1997, *7*, 1094–1103.

[10] Nishimura, A., Tsuhako, M., *Chem. Pharm. Bull.* 2000, *48*, 774–778.

[11] Zhang, Z., Wu, Y., Cheng, W., Wu, R., *Clin. Chim. Acta* 2000, *301*, 205–211.

[12] Hayashi, K., *Genet. Anal.* 1999, *14*, 193–196.

[13] Ren, J., Ueland, P. M., *Hum. Mutat.* 1999, *13*, 458–463.

[14] Kourkine, I. V., Hestekin, C. N., Magnusdottir, S. O., Barron, A. E., *BioTechniques* 2002, in press.

[15] Debernardi, S., Luzzana, M., De Bellis, G., *Frontiers Biosci.* 1996, *1*, c1–3.

[16] Merkelbach-Bruse, S., Kose, S., Losen, I., Bosserhoff, A. K., Buettner, R., *Combin. Chem. High Throughput Screen.* 2000, *3*, 519–524.

[17] Wenz, H. M., Ramachandra, S., O'Connell, C. D., Atha, D., *Mutat. Res. Genom.* 1998, *382*, 121–132.

[18] Ghozzi, R., Morand, P., Ferroni, A., Beretti, J. L., Bingen, E., Segonds, C., Husson, M.-D., Izard, D., Berche, P., Gaillard, J., *J. Clin. Microbiol.* 1999, *37*, 3374–3379.

[19] Cooksey, R. C., Morlock, G. P., Holloway, B. P., Mazurek, G. H., Abaddi, S., Jackson, L. K., Buzard, G. S., Crawford, J. T., *Mol. Diagn.* 1998, *3*, 73–80.

[20] Kourkine, I. V., Hestekin, C. N., Buchholz, B., Barron, A. E., *Anal. Chem.* 2002, in press.

[21] Ellis, L. A., Taylor, C. F., Taylor, G., *Hum. Mutat.* 2000, *15*, 556–564.

[22] Casey, J., Davidson, N., *Nucleic Acids Res.* 1977, *4*, 1539–1552.

[23] Bonner, G., Klibanov, A. M., *Biotech. Bioeng.* 2000, *68*, 339–344.

[24] Yip, S. P., Hopkinson, D. A., Whitehouse, D., *BioTechniques* 1999, *27*, 20–24.

[25] Weghorst, C. M., Buzard, G. S., *BioTechniques* 1993, *15*, 395–400.

[26] Mala, Z., Kleparnik, K., Boček, P., *J. Chromatogr. A* 1999, *853*, 371–379.

[27] Kozlowski, P., Krzyzosiak, W. J., *Nucleic Acids Res.* 2001, *29*, e71.

[28] Larsen, L. A., Christiansen, M., Vuust, J., Andersen, P. S., *Hum. Mutat.* 1999, *13*, 318–327.

[29] Nataraj, A. J., Olivos-Glander, I., Kusukawa, N., Highsmith, W. E., Jr., *Electrophoresis* 1999, *20*, 1177–1185.

[30] Gelfi, C., Vigano, A., Curcio, M., Righetti, P. G., Righetti, S. C., Corna, E., Zunino, F., *Electrophoresis* 2000, *21*, 785–791.

[31] Ren, J., Fang, Z., *Chin. Chem. Lett.* 2000, *11*, 1015–1018.

[32] Barron, A. E., Sunada, W. M., Blanch, H. W., *Electrophoresis* 1996, *17*, 744–757.

[33] Ren, J., Ulvik, A., Refsum, H., Ueland, P. M., *Anal. Biochem.* 1999, *276*, 188–194.

[34] Ren, J., Ulvik, A., Ueland, P. M., Refsum, H., *Anal. Biochem.* 1997, *245*, 79–84.

[35] Tian, H., Brody, L. C., Mao, D., Landers, J. P., *Anal. Chem.* 2000, *72*, 5483–5492.

[36] O'Connell, C. D., Atha, D. H., Oldenburg, M., Tian, J., Siebert, M., Handrow, R., Grooms, K., Heisler, L., de Arruda, M., *Electrophoresis* 1999, *20*, 1211–1223.

[37] Ru, Q. H., Jing, H., Luo, G.-A., Huang, Q., *J. Chromatogr. A* 2000, *894*, 171–177.

[38] Larsen, L. A., Christiansen, M., Vuust, J., Andersen, P., *Combin. Chem. High Throughput Screen.* 2000, *3*, 393–409.

[39] Liechti-Gallati, S., Schneider, V., Neeser, D., Kraemer, R., *Eur. J. Hum. Genet.* 1999, *7*, 590–598.

[40] Wenz, H. M., Baumhueter, S., Ramachandra, S., Worwood, M., *Hum. Genet.* 1999, *104*, 29–35.

[41] Ahrendt, S. A., Halachmi, S., Chow, J. T., Wu, L., Halachmi, N., Yang, S. C., Wehage, S., Jen, J., Sidransky, D., *Proc. Natl. Acad. Sci. USA* 1999, *96*, 7382–7387.

[42] Tian, H., Brody, L., Landers, J. P., *Genome Res.* 2000, *10*, 1403–1413.

[43] Wikman, F. P., Katballe, N., Christensen, M., Laurberg, S., Ørntoft, T. F., *Genet. Testing* 2000, *4*, 15–21.

[44] Jackson, H. A., Bowen, D. J., Worwood, M., *Brit. J. Haematol.* 1997, *98*, 856–859.

[45] Larsen, L. A., Andersen, P. S., Kanters, J. K., Jacobsen, J., Vuust, J., Christiansen, M., *Clin. Chim. Acta* 1999, *280*, 113–125.

[46] Pursall, M. C., Clay, T. M., Bidwell, J. L., *Eur. J. Immunogenet.* 1996, *26*, 41–53.

[47] Ravnik-Glavac, M., Glavač, D., Dean, M., *Hum. Mol. Genet.* 1994, *3*, 801–807.

[48] Hongyo, T., Buzard, G. S., Calvert, R. J., Weghorst, C. M., *Nucleic Acids Res.* 1993, *21*, 3637–3642.

[49] Axton, R. A., Hanson, I. M., Love, J., Seawright, A., Prosser, J., van Heyningen, V., *Mol. Cell. Probes* 1997, *11*, 287–292.

[50] Orban, T. I., Csokay, B., Olah, E., *BioTechniques* 2000, *29*, 94–98.

[51] Gonen, D., Veenstra-VanderWeele, J., Yang, Z., Leventhal, B., Cook, E. H. Jr., *Mol. Psychiat.* 1999, *4*, 339–343.

[52] Sasaki, T. I., Tahira, T., Suzuki, A., Higasa, K., Kukita, Y., Baba, S., Hayashi, K., *Am. J. Hum. Genet.* 2001, *68*, 214–218.

[53] Walz, T., Geisel, J., Bodis, M., Knapp, J.-P., Herrmann, W., *Electrophoresis* 2000, *21*, 375–379.

[54] Geisel, J., Walz, T., Bodis, M., Nauck, M., Oette, K., Herrmann, W., *J. Chromatogr. B* 1999, *724*, 239–247.

[55] Atha, D. H., Wenz, H. M., Morehead, H., Tian, J., O'Connell, C. D., *Electrophoresis* 1998, *19*, 172–179.

[56] Yan, H., Kinzler, K. W., Vogelstein, B., *Science* 2000, *289*, 1890–1892.