Interlaboratory Diagnostic Validation of Conformation-Sensitive Capillary Electrophoresis for Mutation Scanning

Christopher J. Mattocks,1* Gemma Watkins,1 Daniel Ward,2 Tom Janssens,3 Ermanno A.J. Bosgoed,4 Kim van der Donk,4 Marjolijn J. Ligtenberg,4 Bruno Pot,5 Joop Theelen,6 Nicholas C.P. Cross,1 Hans Scheffer,4 and Gert Matthijs3

BACKGROUND: Indirect alternatives to sequencing as a method for mutation scanning are of interest to diagnostic laboratories because they have the potential for considerable savings in both time and costs. Ideally, such methods should be simple, rapid, and highly sensitive, and they should be validated formally to a very high standard. Currently, most reported methods lack one or more of these characteristics. We describe the optimization and validation of conformation-sensitive capillary electrophoresis (CSCE) for diagnostic mutation scanning.

METHODS: We initially optimized the performance of CSCE with a systematic panel of plasmid-based controls. We then compared manual analysis by visual inspection with automated analysis by BioNumerics software (Applied Maths) in a blinded interlaboratory validation with 402 BRCA1 (breast cancer 1, early onset) and BRCA2 (breast cancer 1, early onset) variants previously characterized by Sanger sequencing.

RESULTS: With automated analysis, we demonstrated a sensitivity of >99% (95% CI), which is indistinguishable from the sensitivity for conventional sequencing by capillary electrophoresis. The 95% CI for specificity was 90%–93%; thus, CSCE greatly reduces the number of fragments that need to be sequenced to fully characterize variants. By manual analysis, the 95% CIs for sensitivity and specificity were 98.3%–99.4% and 93.1%–95.5%, respectively.

CONCLUSIONS: CSCE is amenable to a high degree of automation, and analyses can be multiplexed to increase both capacity and throughput. We conclude that once it is optimized, CSCE combined with analysis with BioNumerics software is a highly sensitive and cost-effective mutation-scanning technique suitable for routine genetic diagnostic analysis of heterozygous nucleotide substitutions, small insertions, and deletions.

© 2010 American Association for Clinical Chemistry

Mutation scanning involves comparing a defined region of interest in a DNA sequence with a reference sequence to determine what variants, if any, are present. These variants can be located anywhere in the region of interest and be of any type. This process is distinct from and substantially more demanding than genotyping, which determines the presence or absence of specific variants at specific locations. Direct capillary sequence analysis is often considered the gold standard for mutation scanning, particularly for diagnostic analysis, but this approach can be time-consuming and relatively expensive.

Prescreening strategies aimed at identifying variant-carrying PCR products without fully characterizing them are of considerable interest because they greatly reduce the requirement for confirmatory sequencing (1, 2). In recent years, the investigation of hereditary breast cancer has presented a substantial problem for genetic-diagnostic laboratories. Currently, a full screen consists of mutation scanning of the coding regions and flanking intronic sequences of the 2 hereditary breast cancer genes: BRCA17 (breast cancer 1, early onset; MIM 113705), which is located at 17q21 and consists of 5589 bp of coding sequence distributed over 22 exons (3), and BRCA2 (breast cancer 2, early onset; MIM 600185), which is at 13q12.3 and consists of 10257 bp of coding sequence distributed over 26 exons (4). More than 3400 different BRCA1 and BRCA2 variants have been identified (5), with the overwhelming majority being restricted to only a few
families. Apart from a few common mutations in specific populations, such as Ashkenazi Jews, variants are reasonably evenly distributed throughout the 2 genes, a fact that rules out the possibility of a focused screening approach.

Many different prescreening methodologies have been applied to breast cancer genes, including denaturing gradient gel electrophoresis \((6, 7)\), single-strand conformation polymorphism analysis \((8, 9)\), heteroduplex analysis \((10)\), conformation-sensitive gel electrophoresis \((11, 12)\), denaturing HPLC \((13)\), and mismatch-cleavage assays \((14)\). The throughput of these techniques is limited, however, and their diagnostic accuracies have generally been poorly reported in the literature \((15)\).

More recently, high-resolution melting analysis \((16–18)\) has been used. This technique allows a substantially higher throughput but requires intensive fragment-specific optimization \((17)\). In the absence of appropriate validation, the detection of certain insertions and deletions of single base pairs in homopolymer stretches may be questionable (see Supplemental Data I in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol56/issue4).

Conformation-sensitive capillary electrophoresis (CSCE)\(^8\) is a simple method for indirect mutation scanning that is based on the principle that homoduplex and heteroduplex DNA have different electrophoretic mobilities under partially denaturing conditions \((19–21)\). Fluorescently labeled PCR products are denatured and reannealed, allowing heteroduplexes to form. Wild-type samples can form only a single homoduplex species, whereas heterozygous samples can form 4 possible duplex species: 2 heteroduplexes and 2 homoduplexes. Theoretically, these 4 species have different conformations under partially denaturing conditions and migrate at different rates during electrophoresis. In practice, however, one or more of the different species often remain unresolved.

In common with most mutation-scanning methods, including Sanger sequencing, CSCE typically will not detect large deletions or duplications. A recently described approach that combines fluorescence multiplex quantitative PCR with CSCE does facilitate detection of these mutations \((22)\); however, a diagnostic strategy for genes such as \(BRCA1\) and \(BRCA2\), for which such mutations are important, usually requires the inclusion of mutation-detection techniques such as multiplex ligation-dependent probe amplification or quantitative PCR.

We report optimized conditions for CSCE analysis, together with an extensive formal diagnostic validation with the breast cancer genes as a model.

Materials and Methods

STUDY ORGANIZATION

The study was carried out in 3 laboratories (laboratories 1, 2, and 3). Laboratories 1 and 2 had previous experience with CSCE and performed the electrophoresis experiments on Applied Biosystems 3730 DNA Analyzer instruments (48 capillaries), whereas laboratory 3 had no prior experience and used an Applied Biosystems 3130xl Genetic Analyzer (16 capillaries).

CSCE ANALYSIS

CSCE conditions were based on the method developed at the Sanger Institute \((21)\). The separation matrix consisted of 50 g/L Conformation Analysis Polymer (Applied Biosystems) in 1/4 H\(_{11003}\) TTE Glycerol Tolerant Buffer (which contains Tris, taurine, and EDTA; National Diagnostics), with varying concentrations of urea and sucrose. Applied Biosystems supplied special run modules that allowed the run temperature to be reduced to 10°C.

MANUAL ANALYSIS

Manual analysis by visual inspection was carried out in GeneMapper software (version 3.7; Applied Biosystems). To determine the quantitative affect of varying technical parameters, we scored all analyses (except those indistinguishable from the wild-type control, which were scored 0) by assigning them a resolution score, which was equal to the number of peaks present plus shoulders, which counted as 0.5 (Fig. 1). Traces with peaks outside the analysis window for the instrument [500–6000 relative fluorescent units (RFU)] for

---

\(^8\) Nonstandard abbreviations: CSCE, conformation-sensitive capillary electrophoresis; RFU, relative fluorescence units; SRMS, single-sided root mean square; MAXDIFF, largest vertical difference between 2 compared curves; \(H_0\), null hypothesis.
the 3130xl instrument and 1000–30 000 RFU for the 3730 instruments) were scored as failures.

**AUTOMATED ANALYSIS**

The automated analysis used the heteroduplex analysis plug-in for BioNumerics software (version 5.10; Applied Maths) (23). Initially, traces with peaks below a defined threshold are excluded. Then, 5 parameters are used to compare the remaining traces with a specified wild-type control. The parameters used were SRMS (single-sided root mean square), which measures the percent peak broadening and compares it with that of the wild-type control, and MAXDIFF, which is largest vertical difference between 2 compared curves for the test and wild-type traces. The sum (SRMS + MAXDIFF) was used as a single measure of the efficacy of automated analysis. The other 3 parameters (SECPK, DFH3, and DFH4) were not used.

**OPTIMIZATION PHASE**

We used a systematic panel of 52 plasmid-based controls for the optimization phase (24). The physical parameters tested were the G + C content of the fragment, the position of the variant within the fragment, and the specific nature of the heteroduplex mismatch (single-base variants only). The process parameters tested were PCR dilution, electrophoresis loading parameters, urea content of the polymer, run temperature, and run voltage. In addition, we performed a series of experiments to determine the CSCE limit of detection with respect to the concentration of the variant allele in the sample. Full details of optimization experiments are described in Supplemental Data A in the online Data Supplement.

**FORMAL BLINDED VALIDATION**

The validation was performed with 79 PCR amplicons (235–509 bp in length) that covered the entire coding regions of the BRCA1 and BRCA2 genes (see Supplemental Data B in the online Data Supplement). The total analyzable region of interest was 23 769 bp. For validation we used blinded, anonymized patient samples that had previously tested positive for one or more BRCA1 or BRCA2 variants in one of the 3 laboratories. We selected 359 patient samples on the basis of (a) availability of DNA for analysis, (b) presence of non-redundant variants, and (c) exclusion of variants originally discovered by CSCE or a similar methodology.

We arrayed the DNA samples over 10 plates in 79 blocks that represented the 79 PCR amplicons analyzed (see Supplemental Data C in the online Data Supplement). Except for the 2 wild-type controls identified in each block, analyses within blocks were randomized and blinded. The final analysis panel consisted of 365 wild-type, 445 mutant, and 150 control analyses (n = 960). Because of the replication of variants, the mutant analyses consisted of 402 different variants: 267 point mutations, 24 insertions or duplications of up to 11 bp, 93 deletions of up to 62 bp, 2 insertion–deletions, and 16 compound variants (>1 variant in a fragment) (see Supplemental Data D in the online Data Supplement). The sample panel was confirmed by sequencing on a 3130xl Genetic Analyzer.

To facilitate fluorescence labeling, we used a 3-primer system in each laboratory to carry out the PCRs (25). For the purposes of manual scoring, 2 operators independently inspected all analyses. When the analysts’ results differed, the more conservative result (i.e., with the lower variant score) was taken.

**Results**

**LOSS OF RESOLUTION**

During the pilot analysis, we observed a cumulative loss in peak resolution with sequentially loaded plates in laboratory 1, but not in laboratory 2 or 3. This issue is discussed fully in Supplemental Data E in the online Data Supplement.

**OPTIMIZATION OF CSCE RUN CONDITIONS**

Full details of the optimization phase are described in Supplemental Data A in the online Data Supplement. Optimal conditions for the analyses of fragments with 20%–60% G + C content centered around 12.5 kV at 14 °C with 46 g/L urea, whereas the optimal conditions for fragments with 80% G + C content were 12.5 kV at 18 °C with 60 g/L urea (Fig. 2). Optimal cutoffs for the automated analysis parameters SRMS and MAXDIFF were 8.0 and 6.0, respectively.

**LIMIT OF DETECTION**

All but one of the mutations were detected at relative concentrations of ≥22%. Below this percentage, however, the detection rate dropped to 0% for all variants at 3% representation (see Supplemental Data A in the online Data Supplement).

**FORMAL BLINDED VALIDATION**

The panel of BRCA1 and BRCA2 samples was independently tested in each laboratory with the conditions determined in the optimization phase (validation raw data are given in Supplemental Data F in the online Data Supplement). The overall failure rate in the automated analysis was 5% (143 of 2880 analyses), ranging from 2.5% (24 of 960 analyses) in laboratory 1 to 7% (71 of 960 analyses) in laboratory 2. Both laboratory 1 and laboratory 2 detected all variants, but laboratory 3 failed to detect one of the variants. The overall false-positive rate was 8.3% (84 of 1011 analyses), ranging from 2% (7 of 358 analyses) in laboratory 1 to 13.4% (44 of 328 analyses) in laboratory 3.
In the manual analysis, the overall failure rate was 5% (142 of 2880 analyses), ranging from 1% (11 of 960 analyses) in laboratory 1 to 8% (77 of 960 analyses) in laboratory 3. The overall false-positive rate was 5.6% (57 of 1015 analyses), ranging from 1.6% (5 of 314 analyses) in laboratory 3 to 10.6% (36 of 339 analyses) in laboratory 2. All variants were detected in laboratory 1, but laboratories 2 and 3 missed 1 variant and 11 variants, respectively. Blinded reanalysis of the data for laboratory 3 showed that 5 of the missed variants were clearly detectable, with at least a shouldered peak to indicate their presence. The trace indicating a variant’s presence was subtle for 1 variant. Three of these variants were indistinguishable from the control peak, but

![Fig. 2. Pairwise response surface analyses.](image)

Plots show resolution score contours with respect to 2 variables. The key physical variables were position (P1, P2, or P3) of the variant and G + C content (A). P1 is the position of mutations closest to the end of the fragment in the generic mutation-detection (GMD) controls (62–84 bp), P3 is the position of mutations furthest from the end of the fragment in the GMD controls (215–223 bp), and P2 is the position of mutations approximately halfway between P1 and P3 in the GMD controls (142–152 bp). The effect of process variables (run voltage, urea concentration, and temperature) on resolution score was analyzed separately for fragments with 20%–60% G + C content (B) and fragments with 80% G + C content (C).
2 of these variants should have been detected but were not because of poor amplification (see Supplemental Data G in the online Data Supplement).

The final variant not detected by manual analysis in laboratory 3 was $BRCA1$:c.4676−34G>A, which is upstream of exon 16 and is located 66 bp from the end of the amplicon (G + C content, 45.5%). This variant also evaded detection by automated analysis in laboratory 3 and was the single variant not detected by manual analysis in laboratory 2. It represents the subtlest electrophoretic shift seen in this validation (Fig. 3A).

All of the insertion and deletion variants and 93% of the point variants produced at least 2 distinct peaks (95% overall). Only 5 of the point variants (<2%) consistently yielded a resolution score of 1.

A marginal loss of detection efficacy was seen in fragments with a higher G + C content and with variants at the ends of the fragments, but there were no statistically significant correlations (Fig. 4).

INTERLABORATORY COMPARISONS
The reproducibility of the variant peak shapes was measured by Spearman rank correlation analysis of resolution scores (for manual analysis) and SRMS + MAXDIFF analysis (for automated analysis). The distributions of variant analysis scores between laboratories were well correlated in terms of monotonic ordinality for both the manual and automated analyses (i.e., if the variants were placed in order of their analysis score within each laboratory, they would, in general, appear in the same order [$P < 0.0001$ in each case, Spearman rank correlation; null hypothesis (H$_0$): Lab $x$ ≠ Lab $y$]).

The distributions of variant analysis scores were not correlated with respect to location or dispersal ($P < 0.0001$, Friedman test; H$_0$: Lab $x$ = Lab $y$). For automated analysis, differences were caused by constant bias. With the results of laboratory 2 used as the baseline, laboratories 1 and 3 had a mean bias of −6.7 and −4.9, respectively; however, although laboratory 1

---

Fig. 3. Examples of CSCE analysis profiles.
In all panels but one (I), a broken line represents the wild-type trace, and a solid line represents the test trace. The most subtle variant (A), different variants at one position (B and C), compound variants (D, E, and F), 1 variant in 2 different PCR amplicons (G and H), and 2 different variants in 1 amplicon with virtually identical peak profiles (I).
yielded the lowest scores for the variants in the automated analysis, manual analysis produced the highest mean resolution scores (3.1 for laboratory 1; 3.0 for laboratory 2; 2.9 for laboratory 3). In addition, the data for laboratory 1 were notably less widely distributed than the data for the other laboratories (SRMS + MAXDIFF SDs: 10.9 for laboratory 1; 16.9 for laboratory 2; 19.4 for laboratory 3).

Discussion

SENSITIVITY AND SPECIFICITY

The accuracy of sequencing is such that any alternative must be validated to a very high standard if it is to be used for diagnostic purposes. In this study, we have shown that optimal performance for the analysis of fragments with 20%–60% G/C content can be achieved with a single set of run conditions (see Supplemental Data A in the online Data Supplement). Because only 23% of the exons in the genome (NCBI Build 36.3) have a G/C content >60%, the conditions we have validated should be applicable to most diagnostic situations.

Our extensive blinded validation shows that the sensitivity of CSCE for detecting heterozygous variants in fragments <500 bp with a G + C content 20%–60% is 98.3%–99.4% (95% CI) for manual analysis and 99.6% (95% CI) for automated analysis (Table 1). At these high levels, it is likely that interlaboratory variation would have a more important effect on sensitivity than any inaccuracy in estimating sensitivity. Therefore, for practical purposes we have shown that the sensitivity of CSCE for detecting heterozygous variants is indistinguishable from that for sequencing.

The overall specificity is dependent on the method of analysis used. With manual analysis, the estimated specificity is 93.1%–95.5% (95% CI), whereas the estimated specificity with automated analysis is 90.1%–93.1% (95% CI). The lower specificity achieved with automated analysis is largely due to the low-level background noise, which is often distant from the actual sample peak. Such noise can easily be excluded by visual inspection but cannot be identified as such by the analysis program (see Supplemental Data H in the online Data Supplement). Although an increase in the

Fig. 4. The effect of position and G + C content on detection in the validation panel.

(A and B), x Axis represents the distance of the respective variants from the nearest end of the fragment in which they occur. Data points are averaged analysis scores within 10-bp windows. (C and D), x Axis represents the percent G + C content of the fragments. The data points represent averaged analysis scores within 1% G + C windows.
false-positive rate with automated analysis increases the number of fragments requiring sequencing, the difference is marginal. The key benefit of automated analysis is its objectivity and consistency, which underpin a reliable sensitivity. Automated analysis also enables the development of schemes for objectively monitoring resolution and facilitates a substantial reduction in the time required for analysis.

AMPLICON DESIGN
A 50-bp buffer was included in the PCR designs to allow for decreased sensitivity in detecting variants close to the ends of fragments, an observation that has previously been reported (21). Seventeen different variants included in this study fell within the stipulated 50-bp buffer, with distances as close as 36 bp from the end of the amplified fragment. Although we detected all of these variants, we found a slight bias toward lower resolution scores in this group compared with the validation as a whole (Fig. 4). Therefore, a minimum 50-bp buffer is a prudent measure to maintain performance.

HOMOZYGOTE DETECTION
The detection of homozygous variants was not covered specifically in this validation, but it has been proved in principle by others (26). To detect homozygotes, one must mix samples in equal proportion with the wild-type control to allow the formation of heteroduplexes. Such mixing means that any heterozygous variant would be represented by only 25% of the sample. We have shown that the sensitivity of CSCE remains optimal only to approximately 22% representation of the mutant allele; sensitivity falls sharply below this value. It may therefore be prudent to use a lower proportion of the wild-type spike to keep the analysis of heterozygotes within the optimal range (e.g., 35% wild-type spike and 65% test sample).

LIMIT OF DETECTION
The findings that sensitivity is maintained down to 22% variant representation and that some variants are still detectable at 6% indicate that CSCE may be useful for the analysis of mitochondrial DNA, tumor samples, and mosaic mutations, for which mutant representation is expected to be <50%.

CONTROLLING RESOLUTION
CSCE is critically reliant on the discrimination of peak morphology. Therefore, robust monitoring of the resolution, ideally in all capillaries in every run, is critical for maintaining diagnostic standards. This goal can be accomplished most simply by reserving one of the color channels (typically LIZ) for running a resolution control. This control should consist of an amplicon

### Table 1. Summary results for methodology validation.

<table>
<thead>
<tr>
<th></th>
<th>Lab 1</th>
<th>Lab 2</th>
<th>Lab 3</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automated analysis, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True positive</td>
<td>385</td>
<td>371</td>
<td>390</td>
<td>1146</td>
</tr>
<tr>
<td>False negative</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>True negative</td>
<td>351</td>
<td>292</td>
<td>284</td>
<td>927</td>
</tr>
<tr>
<td>False positive</td>
<td>7</td>
<td>33</td>
<td>44</td>
<td>84</td>
</tr>
<tr>
<td>Analysis fails</td>
<td>24</td>
<td>71</td>
<td>48</td>
<td>143</td>
</tr>
<tr>
<td>Controls (excluded)a</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>450</td>
</tr>
<tr>
<td>Replicates (excluded)b</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>129</td>
</tr>
<tr>
<td>Total analyses</td>
<td>960</td>
<td>960</td>
<td>960</td>
<td>2880</td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>100</td>
<td>100</td>
<td>99.7</td>
<td>99.9</td>
</tr>
<tr>
<td>Lower 95% confidence limitc</td>
<td>99.2</td>
<td>99.2</td>
<td>98.8</td>
<td>99.6</td>
</tr>
<tr>
<td>Upper 95% confidence limitc</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>98</td>
<td>89.8</td>
<td>86.6</td>
<td>91.7</td>
</tr>
<tr>
<td>Lower 95% confidence limitc</td>
<td>96.4</td>
<td>86.7</td>
<td>83.1</td>
<td>90.1</td>
</tr>
<tr>
<td>Upper 95% confidence limitc</td>
<td>99.1</td>
<td>92.3</td>
<td>89.6</td>
<td>93.1</td>
</tr>
<tr>
<td>Manual analysis, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True positive</td>
<td>394</td>
<td>373</td>
<td>366</td>
<td>1133</td>
</tr>
<tr>
<td>False negative</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>True negative</td>
<td>346</td>
<td>303</td>
<td>310</td>
<td>958</td>
</tr>
<tr>
<td>False positive</td>
<td>16</td>
<td>36</td>
<td>5</td>
<td>57</td>
</tr>
<tr>
<td>Analysis fails</td>
<td>11</td>
<td>54</td>
<td>75</td>
<td>140</td>
</tr>
<tr>
<td>Controls (excluded)a</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td>450</td>
</tr>
<tr>
<td>Replicates (excluded)b</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>129</td>
</tr>
<tr>
<td>Total analyses</td>
<td>960</td>
<td>960</td>
<td>960</td>
<td>2880</td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>100</td>
<td>99.7</td>
<td>97.1</td>
<td>99</td>
</tr>
<tr>
<td>Lower 95% confidence limitc</td>
<td>99.2</td>
<td>98.7</td>
<td>95.2</td>
<td>98.3</td>
</tr>
<tr>
<td>Upper 95% confidence limitc</td>
<td>100</td>
<td>100</td>
<td>98.4</td>
<td>99.4</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>95.6</td>
<td>89.4</td>
<td>98.4</td>
<td>94.4</td>
</tr>
<tr>
<td>Lower 95% confidence limitc</td>
<td>93.4</td>
<td>86.2</td>
<td>96.7</td>
<td>93.1</td>
</tr>
<tr>
<td>Upper 95% confidence limitc</td>
<td>97.2</td>
<td>92</td>
<td>99.4</td>
<td>95.5</td>
</tr>
</tbody>
</table>

*a Defined controls for each fragment are excluded from the calculation of specificity.
*b Some variants were present in >1 analysis. Such replicates were excluded from the sensitivity calculations (i.e., only 1 copy of each variant analysis was considered). Replicates for exclusion were selected randomly and were the same for all analyses.
*c Confidence limits were calculated by the exact method for small numbers of failures (see http://www.itl.nist.gov/div898/handbook/prc/section2/prc241.htm).
containing a variant (preferably with a low resolution score) that produces a peak of known shape. This control can be monitored for appreciable changes in shape between capillaries and from run to run. Capillary analyses with results that fall outside the predefined limits should be considered as having failed.

RESOLUTION SCORES AND SPECIFIC EXAMPLES

Fig. 5A shows the distributions of mean resolution scores for manual analyses across the 3 laboratories, and Fig. 5B shows the distributions of the corresponding automated analysis scores. The majority (81%) of the insertion and deletion variants showed evidence of all 4 possible duplex species. Cases in which <4 peaks were detectable can be explained by lowered resolution due to proximity to the end of a fragment, chance equality of electrophoretic mobility of duplex species, or simply that peaks migrated beyond the electrophoresis data-collection window. Cases with >4 detectable species cannot be explained easily, because only 4 species are possible (2 homoduplexes and 2 heteroduplexes). It is notable that 22 of the 402 different variants in this study showed evidence of more than the predicted 4 duplex species (maximum, 8 species). This phenomenon occurred more frequently with insertion and deletion variants (9%) and compound variants (25%) than with single-base substitutions (2%). This finding suggests that for some variants, particularly when >1 base is affected, one or more of the duplexes formed are unstable and capable of assuming >1 conformation, thereby yielding >1 peak.

Aside from these observations, we found no specific correlations between the nature of the variation and detection efficacy; however, it is worth noting that point variations producing very subtle shifts, which are the most likely candidates for false-negative calls, are very rare. Only 5 variants (1.2%) showed no detectable shouldering; these variants were detectable only by a change in peak shape. Of these 5 variants, only 2 (<0.5%) appreciably challenged the automated analysis (i.e., they were not detected by both SRMS and MAXDIFF in all laboratories: BC1_16:c.4676-34G>A and BC1_20:c.5194-312G>A).

Various special cases were investigated during the workup for the validation (Fig. 3). To evaluate the suggestion that the presence of a polymorphism in the same fragment as a pathogenic mutation could confound its detection, we included 16 different compound variants in the study. Seven of these compound variants consisted of a point variant juxtaposed with a small insertion or deletion; the remaining 9 variants were pairs of point variants. In all cases, the compound variant showed numbers of peaks equivalent to or greater than the numbers produced by the individual variants, and the profiles could easily be distinguished from the individual variants where they were available for analysis. Although the numbers are not large enough to draw any statistically significant conclusions, the presence of a polymorphism appears, if anything, to render true variants more detectable.

In cases in which the same variant was present in >1 fragment (n = 34), the paired analyses yielded substantially different profiles. Similarly, in a single case in which 2 different base changes were located at the same nucleotide position, the 2 variants were clearly distinguishable (Fig. 3).

One variant, BRCA1:c.5369_5385del17 in exon 22, was given a resolution score of 2 by manual analysis in laboratory 3, but it actually had a 4-peak profile. Owing to

---

**Fig. 5. The effect of variation type on analysis score.**

Plots represent the distribution of analysis scores for the wild-type analyses and the different types of variants [wild type (WT), n = 365; point mutations (Point), n = 267; insertion/deletions (Insdel), n = 119; compound variants, n = 16]. Boxes, whiskers, and crosses represent the interquartile range, the upper and lower quartiles, and outliers, respectively. (A), Average manual analysis resolution score across the 3 laboratories. (B), Average automated analysis score above the most relevant analysis threshold [calculated by independently determining the mean of SRMS and MAXDIFF, deducting the appropriate analysis cutoff value (SRMS = 8; MAXDIFF = 6), and taking the higher of the 2 values].

---

600 Clinical Chemistry 56:4 (2010)
the separation of these peaks, only 2 peaks fell within the analysis window used. A similar error may have occurred in the analysis of BRCA1:c.2646-2648delTGC in exon 11, which was one of the variants missed in laboratory 3. It is therefore important that manual analysis be performed in 2 stages: an overall view to detect large shifts, followed by detailed examination of the primary peak in an analysis window 200–500 scans wide to detect any subtle shifts.

INSTRUMENT COMPARISON
Because of the width of the analysis window on the 3730 capillary arrays, it is necessary to include sucrose in the polymer to modify its refractive index so that all capillaries yield a suitable signal. This measure is not necessary on the 3130xl, although inclusion of sucrose did not detrimentally affect the analysis. The only important difference in the performance of the 2 instruments was optimizing the dilution and loading conditions so that all analyses fell within a suitable analysis window (500–600 RFU for the 3130xl and 1000–3 000 RFU for the 3730). This optimization requires substantially more effort for the 3130xl.

DEALING WITH NONPATHOGENIC POLYMORPHISMS
A number of variants in this study with low resolution scores gave very similar profiles that could not be reliably distinguished (Fig. 3). Therefore, because CSCE is not capable of unequivocally characterizing the variants detected, for diagnostic purposes all positive results, including those representing nonpathogenic polymorphisms, need to be confirmed by sequencing. This requirement can have a substantial effect on the cost and processing-time advantages gained by the use of CSCE. Alternative methods involving high-throughput or multiplex genotyping can be used to exclude known polymorphisms. For example, the Salisbury high-throughput screening facility uses a bespoke SNPlex assay (Applied Biosystems) to screen for 22 known polymorphisms in the breast cancer genes and has achieved an overall sequencing rate consistently <10% of the total fragments analyzed (unpublished data).

MATERIALS AND METHODOLOGY COMPARISON
For high-throughput diagnostic application, both CSCE and high-resolution melting analysis would be suitable candidates for a prescreening screen. The 2 methods are likely to have similar absolute sensitivities, although there may be differences with respect to the type of variants most likely to be called as false negatives (Supplementary Data I in the online Data Supplement). With respect to other aspects, there is little difference between the 2 methodologies. A key consideration in terms of methodology choice is likely to be the availability of instrumentation and the level of existing experience with the technologies.

Conclusion
In this study, we have defined optimal running conditions for applying CSCE to fragments with 20%–60% G + C content, which cover approximately 77% of all exons in the genome. We have derived a set of analysis guidelines for both manual and automated analyses. Although manual analysis can provide an effective mutation-scanning solution when the analysis guidelines are carefully applied, automated analysis is much faster and reduces the chances of operator error. With the use of multiplexing and automated analysis in the BioNumerics software, CSCE is substantially faster and cheaper than sequencing. This methodology is robust and relatively simple to implement for diagnostic applications. Furthermore, under the conditions defined and within the limitations described, CSCE with automated analysis provides a mutation-scanning technique for heterozygous nucleotide substitutions and small insertions and deletions that is, for practical purposes, equivalent to capillary-based Sanger sequencing in terms of sensitivity.

Additional Note
At the time of writing, Applied Biosystems was preparing a technical-application note regarding CSCE. This note will include a “troubleshooting” section. For any queries regarding CSCE or requests for specialist firmware and analysis modules mentioned in this report, please contact J. Theelen at Joot.Theelen@lifetech.com.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures of Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: B. Pot, Applied Maths.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: The project was jointly funded and supported by the UK Department of Health (via NGRI Wessex), EuroGentest, Applied Biosystems (Europe), and Applied Maths (Gent, Belgium).
Expert Testimony: None declared.
Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: We thank N. Arts and H. Ouchene for technical assistance.
References


