Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1051 (2004) 155–160
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

DNA profiling by capillary array electrophoresis with non-covalent fluorescent labeling

Nels A. Olson[a], Julia Khandurina[b], András Guttman[a,*]

[a] *Illumina Inc., San Diego, CA 92121, USA*
[b] *Diversa Corporation, 4955 Directors Place, B-1-131, San Diego, CA 92121, USA*

Abstract

Increasing need for large-scale DNA profiling necessitated the development of automated electrophoresis based methods enabling rapid, high performance analysis of nucleic acids in a wide molecular-mass range. In this paper, we report on the adaptation of a commercial 96-capillary array electrophoresis (CAE) instrument for high-throughput DNA fragment analysis and the evaluation of the effects of different non-covalent DNA staining dyes on separation efficiency. The applicability of different color internal fluorescent standards is shown with mathematical spectral overlap correction algorithms. Large-scale quality control assessment of oligonucleotide probes using non-covalent fluorophore labeling is also demonstrated. The method requires small sample amounts, offers automation and quantification capabilities to accommodate modern biotechnology industry needs.

Keywords: DNA profiling; Electrophoresis; Non-covalent fluorescent labeling

1. Introduction

Large-scale analysis and profiling of DNA molecules, such as polymerase chain reaction (PCR) products, restriction digest fragments and oligonucleotide probes is conventionally accomplished by means of agarose or polyacrylamide slab gel electrophoresis [1]. Electric field mediated separation methods have significantly advanced in the last two decades, primarily due to the development of microseparation technologies, such as capillary [2] and microchip electrophoresis [3]. These techniques have recently been applied to high-throughput DNA sequencing, featuring long read lengths [4]. On the other hand, development of automated devices for rapid, large-scale DNA fragment analysis, mapping and expression profiling from small quantities of samples is still in great demand. Our previous work reported successful adaptation of a slab gel based DNA sequencing instrument for double-stranded (ds) DNA fragment analysis up to 64 lanes, using agarose sieving medium [5]. Others applied capillary array electrophoresis (CAE), that combined advantages of the multilane slab gel format with the high separation performance of capillary electrophoresis [6–8]. At the early 1990's, Mathies and coworkers [9] introduced confocal fluorescent scanning for simultaneous interrogation of multiple capillaries. Later the same group have thoroughly investigated the possibility of multiplexing DNA fragment sizing by CAE using dimeric intercalator dyes, also evaluating ionic effects of buffer components and dye/DNA ratio [6]. Capillary array electrophoresis was also applied to high sensitivity mutation detection by single-strand conformation polymorphism/heteroduplex analysis [8].

Attempts to develop replaceable linear polymer sieving mediums for capillary electrophoresis have started more than a decade ago by the application of low viscosity non-cross-linked polyacrylamide gels for DNA analysis [10]. Since then, various other hydrophilic linear polymer solutions have been reported to accommodate separation of single and double-stranded DNA molecules in microbore capillary columns, including polyethylene oxides (PEO) [11], polyvinyl pyrrolidone [12], derivatized celluloses [13],

* Corresponding author. Tel.: +1 8585265421; fax: +1 8585265921.
E-mail address: aguttman@diversa.com (A. Guttman).

0021-9673/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2004.08.067

etc. Linear polymer matrices proved to be less sensitive to changes in temperature, and offered the possibility of pressure injection [14].

Visualization of DNA fragments by covalent labeling of nucleic acids with fluorescent tags, such as fluorescein and rhodamine derivatives, has been practiced for decades [15]. However, for large-scale DNA profiling, complexation based non-covalent labeling provides a universal, easy to use and more cost effective approach. In this case the staining dye is either directly added to the sieving medium (*in migratio* labeling), or mixed with DNA samples prior to electrophoresis. In both instances, the complexation phenomenon usually increases the separation window resulting in higher resolution [16].

In this paper, we report on the implementation of a commercial 96-capillary array DNA sequencer for large-scale dsDNA profiling. High separation efficiency was achieved using non-covalent staining procedure with fluorophore DNA binding dyes in conjunction with sensitive laser-induced fluorescence (LIF) detection. CAE with non-covalent labeling has also proved useful in high-throughput quality control analysis of oligonucleotide probes.

2. Experimental

2.1. Chemicals

Tris, boric acid, EDTA·Na_2, ethidium bromide, polyethylene oxides M_w 1.0×10^6 and 1.0×10^6, 8-aminopyrene-1,3,6-trisulfonic acid trisodium salt (APTS), sodium cyanoborohydrate (1 M in tetrahydrofuran), maltoheptaose and all other chemicals were from Sigma–Aldrich (St. Louis, MO, USA) in electrophoresis grade. SYBR Gold, SYBR Green and YO-PRO-1 staining dyes were from Molecular Probes (Eugene, OR, USA). The GelStar fluorescent dye was from Cambrex Bioscience (Rockland, ME, USA). The 100 base pairs (bp) and 1 k base pairs (kbp) DNA ladders (New England BioLabs, Beverly, MA, USA) were diluted with double deionized water (18 MΩcm) to the working concentration of 1.0–2.5 ng/μL and stored at −20 °C until use. All buffer solutions were filtered through a 0.2 μm nylon membrane syringe filter (Fisher Scientific, Pittsburgh, PA, USA) prior to use.

2.2. Capillary array electrophoresis system

A 96-capillary MegaBACE 1000 (Amersham Biosciences, Piscataway, NJ, USA) system was used either with coated (Amersham Biosciences) or uncoated bare fused silica (The Gel Company, San Francisco, CA, USA) capillary arrays. The effective separation distance was 40 cm with a total capillary length of 62 cm (75 μm I.D.). Samples were loaded into the individual wells of a 96-well plate and injected into the corresponding capillaries. Electrokinetic injection and separation were performed from the cathode side (catstage) towards the anode (anostage), since the negatively charged DNA fragments, as well as the APTS labeled internal standard mixture (see below) migrate towards the anode under applied electric field. In most experiments the labeling dye was added to the separation matrix for *in migratio* labeling. Six 2 mL tubes served as anode buffer reservoirs (one per each 16-capillary bundle), while a 96-well plate filled with the running buffer was positioned at the catstage. A plastic tray at the catstage was used during washing and conditioning steps. The instrument was equipped with a four color confocal scanning detection system. In the course of this work, the excitation and emission optics used were the same as in DNA sequencing applications (488 nm excitation wavelength and 520DF20, 555DF20, 585DF20 and 610LP emission filters).

2.3. Protocols

For dsDNA fragment analysis 1× TBE buffer (89 mM Tris, 89 mM boric acid and 2 mM EDTA·Na_2, pH 8.3) was used as running buffer containing 0.1% (M_w 1.0×10^6) and 0.5% (M_w 1.0×10^6) PEO. Samples were injected at 2–4 kV for 5–10 s and run at 10 kV (160 V/cm). To ensure good migration time reproducibility, the capillaries were sequentially washed between runs with 1 M HCl, followed by water and running buffer. Washing and buffer plate filling operations were performed using Instrument Control Studio commands (ICS, MegaBACE 1000) which allowed flexible control of time and pressure for solutions of different viscosity. The sample volume in the 96-well skirted plates (Gel Company) was kept in the 5–15 μL range. The separation temperature in the electrophoresis compartment was maintained at 30 °C.

APTS labeled internal standards were made as follows: 1 μL 5 mM maltoheptaose was labeled through reductive amination by the addition of 2 μL 0.2 M APTS in 15% acetic acid and 2 μL 1 M $NaBH_3CN$ in tetrahydrofuran [17]. The labeling reaction was incubated for one hour at 75 °C, followed by the addition of 100 μL water to stop the reaction. Prior to CE analysis, the reaction mixture was further diluted and mixed with samples resulting in approximately 5000-fold total dilution.

For single stranded oligonucleotide probe analysis, 2× MegaBACE buffer (Amersham Biosciences) and 0.5 mM EDTA (pH 8.3) were used as running buffer and sample loading buffer stock solution, respectively. Eighty and hundred μM oligonucleotide probe solutions were diluted 15-fold in water followed by further 10-fold dilution in the sample loading buffer and transferred into a microwell plate (5–15 μL per well). The sieving matrix in this instance was 4% linear polyacrylamide (LPA) in 1× MegaBACE running buffer containing 7 M urea and 20000-fold diluted GelStar fluorescent dye. LPA was polymerized in-house following the method of Heiger et al [18]. To minimize air bubble formation, the separation medium filled buffer tubes were spun at 8000 rpm for 2 min prior to introduction into the

capillaries. During separation the buffer tubes (anode side) and plate (cathode side) contained 1 mL and 150 μL/well running buffer, respectively. Separation temperature was 27 °C.

2.4. Data processing

Raw data sets from each run (96 binary *.rsd files) were translated into a single comma separated *.cvs file format using standard binary conversion software, followed by processing in Microsoft Excel to visualize the individual electropherograms. In cases were there was spectral overlap between the samples and internal fluorescent standards, a spectral correction was performed. To define the spectral correction matrix, the signal maxima and background were determined in two corresponding spectral channels for the internal standard (P1) and analyte (P2) peaks. This was followed by background subtraction, resulting in four numbers of the net signals in the two channels used, S1 and S2, respectively. Then the following 2 × 2 matrix was constructed:

$$\begin{bmatrix} 1 & \mathbf{B} \\ \mathbf{A} & 1 \end{bmatrix}$$

where $\mathbf{A} = S1_{P1}/S2_{P1}$ and $\mathbf{B} = S1_{P2}/S2_{P2}$. $S1_{P1}$ is the signal of P1 and $S1_{P2}$ is the signal of P2 in the first channel. $S2_{P1}$ is the signal of P1 and $S2_{P2}$ is the signal of P2 in the second channel.

An analytical inversion factor (AIF) was calculated as AIF = $(1 \times 1) - (\mathbf{A} \times \mathbf{B})$. Using this factor an algebraic solution was calculated for matrix inversion:

$$\begin{bmatrix} \mathbf{C} = 1/\mathrm{AIF} & \mathbf{E} = -\mathbf{B}/\mathrm{AIF} \\ \mathbf{D} = -\mathbf{A}/\mathrm{AIF} & \mathbf{F} = 1/\mathrm{AIF} \end{bmatrix}$$

To get spectrally separated traces the data points in the two channels were then adjusted for spectral separation based on these matrix values as follows:

$$\text{First channel}: \quad S1 \times \mathbf{C} + \mathbf{E} \times S2$$

$$\text{Second channel}: \quad S1 \times \mathbf{D} + \mathbf{F} \times S2 + \text{trace offset value}$$

3. Results and discussion

The first goal of this study was to evaluate the applicability of a commercial capillary array electrophoresis DNA sequencing instrument for dsDNA profiling in the size range of 100–10 000 bp, using low viscosity sieving matrix in conjunction with non-covalent fluorophore labeling. While automated DNA sequencers were originally designed for the analysis of covalently tagged single stranded DNA molecules at high temperatures, we adapted this system for the separation of unlabeled dsDNA fragments, either by adding ethidium bromide to the separation matrix or other fluorophore staining dyes to the samples prior to electrophoresis.

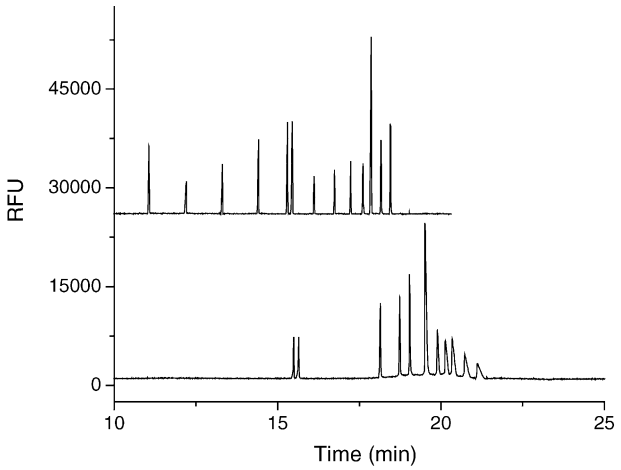

Fig. 1. Representative capillary array electrophoresis (CAE) separation traces of the 100 bp (upper trace) and the 1 kb (lower trace) dsDNA ladders. Peaks in the order of migration: 100, 200, 300, 400, 500, 517, 600, 700, 800, 900, 1000, 1200 and 1517 bp (upper trace); 0.5, 0.517, 1, 1.5, 2, 3, 4, 5, 6, 8 and 10 kbp (lower trace). Conditions: sieving matrix: 0.5% M_w 1×10^6 and 0.1% M_w 8×10^6 PEO in 1× TBE buffer (pH 8.3), containing 500 nM ethidium bromide. Injection: 5 s at 2 kV. Separation: 10 kV (160 V/cm). Temperature: 30 °C.

Since, in most instances, dsDNA analysis does not require sequencing grade resolution (i.e., single base) a dilute mixture of high-molecular-mass PEO polymers (M_w 1×10^6 and M_r 8×10^6) was employed. The low viscosity of this separation medium enabled easy introduction and replacement of the sieving matrix within the capillaries. Low conductivity TBE buffer allowed the application of high separation voltages without the generation of high electric current (<10 μA per capillary) and concomitantly negligible Joule heat production.

Fig. 1 delineate typical profiles of a 100 bp (upper trace) and 1 kb (lower trace) dsDNA ladder. The sieving matrix composition of 0.5% 1 MDa and 0.1% 8 MDa PEO provided good separation performance in a wide size range of 100–10 000 bp. Please note the excellent resolution between 500 and 517 bp fragments in both traces, as these usually co-migrate in conventional submarine agarose gel electrophoresis. Out of 96-capillaries, typically 4–8 capillaries per run exhibited poor performance (data not shown), presumably, either due to insufficient preconditioning (i.e., poor capillary surface quality), coating degradation (in case of permanently coated arrays) and injection failure, or bubble formation. Therefore, to ensure that all 96 samples are analyzed from a microtiter plate, samples should be present in replicates, or the same plate should be run repeatedly. Both options are still significantly more time efficient than sequential processing in single capillary settings.

As one can observe in Fig. 1, the shorter DNA fragments of 100–400 bp exhibited slightly fronting characteristics, while the peaks corresponding to the larger fragment range (over 5000 bp) are increasingly tailed. This suggests that additional fine tuning of the sieving matrix composition might be nec-

essary to optimize the separation performance in these size ranges. On the other hand, these fronting and tailing characteristics could be possibly related to electromigration dispersion phenomena, i.e., electrophoretic mobility mismatch between DNA molecules and borate co-ions in the buffer [19]. Separation efficiency in this case can be improved by either changing the running buffer pH, i.e., altering the mobility of the borate co-ion, or affecting the mobility of the DNA fragments by changing the polymer concentration and/or composition. However, enhanced separation performance of the larger DNA molecules is typically associated with performance loss for the shorter fragments (and vice versa), due to either increased mobility mismatch or inadequate sieving.

In the course of our study, we found that uncoated fused silica capillary arrays in combination with dynamic coating rendered more consistent performance and extended column life time compared to permanently coated arrays. Uncoated arrays were also less expensive. In addition, dynamic coating was easily regenerated prior to each run by sequential washing with HCl, water, and the separation matrix. When poor separation performance prevailed, an extended capillary cleaning procedure was applied using 1.0 M NaOH rinse followed by water and the above described regular regeneration sequence. On the other hand, permanently coated capillaries did not require extra washing steps, reducing in this way the total analysis time. However, we found some batch-to-batch variations in coating quality, and degradation of the separation performance usually became pronounced after 150–200 runs. It is important to emphasize that CAE based dsDNA analysis, utilizing low viscosity separation media, is quite sensitive to capillary surface uniformity and to any residual electroosmotic flow.

To address any possible capillary-to-capillary migration time inconsistency, internal standards were introduced to support accurate DNA fragment sizing. We attempted to apply a bracketing standard concept using a slow (maltoheptaose-APTS conjugate) and a rapidly migrating (APTS) fluorescent compounds with emission characteristics different than that of the DNA binding dyes. In this way, the standards and analyte molecules are visualized separately in two different spectral channels of the four color detection system of the CAE instrument. Fig. 2 compares simultaneous multicolor detection of the 100 bp DNA ladder *in migratio* labeled with ethidium bromide (upper red channel traces in Fig. 2A and B) and the APTS/maltoheptaose-APTS conjugate bracketing standard mixture (lower green channel traces), added to the samples for data normalization and size evaluation purposes. The first, high intensity peak (∼9 min) in the upper traces of Fig. 2A and B is the un-reacted APTS dye remained in the labeling reaction mixture. The maltoheptaose-APTS conjugate migrated at approximately 17.5 min. The extra peak at 12.5 min was an impurity in the APTS. The APTS and maltoheptaose-APTS conjugate internal standard mixture provided a good bracketing standard choice for the most usual PCR amplicon size range

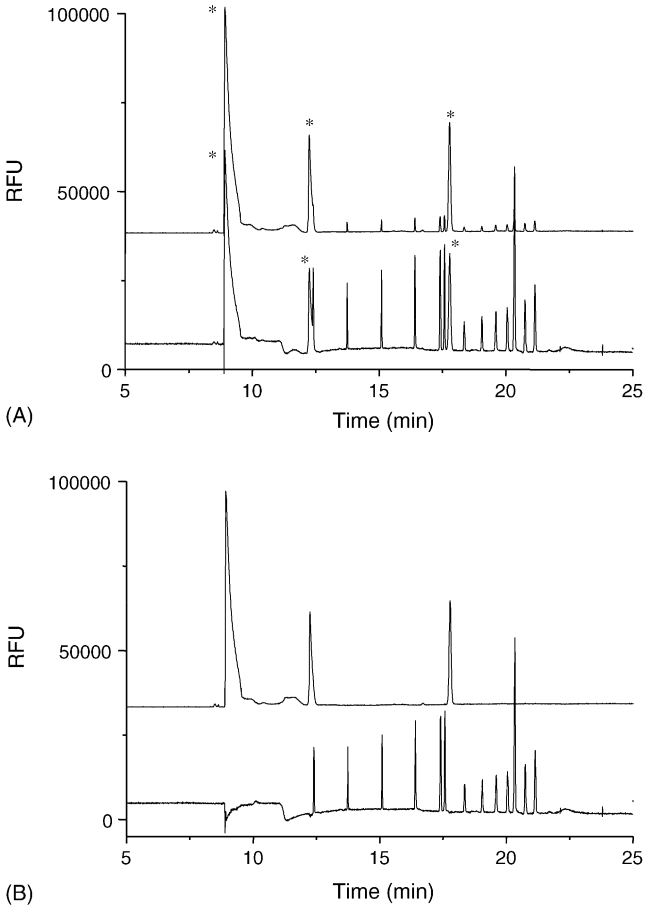

Fig. 2. Multicolor detection of *in migratio* labeled 100 bp dsDNA ladder (ethidium bromide, upper channel trace in each panel) and the APTS/maltoheptaose-APTS conjugate internal standards (lower channel trace in each panel) in CAE before (A) and after (B) a mathematical algorithm based spectral overlap correction. Peaks corresponding to the components of the internal standard mixture are marked with asterisks in panel A. Conditions as in Fig. 1.

of 100–500 bp, enabling unambiguous fragment size assessment. Higher degree of polymerization sugar molecules can extend the sizing range. Due to the broad fluorescent emission spectra of ethidium bromide and APTS, both asymmetrically extended towards the longer wavelength range, spectral overlap from the green channel to the red channel was particularly pronounced (Fig. 2A, upper channel trace). Fig. 2B depicts the same electropherograms after application of a spectral separation procedure based on the mathematical algorithm described in Section 2.4. As one can see, here the separation profiles of the DNA sample (red channel) and the internal standards (green channel) are clearly divided in the two spectral channels with no apparent signal overlap in between.

An alternative approach to mathematical algorithm based spectral separation is to use a fluorescent dye that emits in a higher (red) wavelength range, and a DNA binding dye that emits at the lower wavelength blue or green region, as

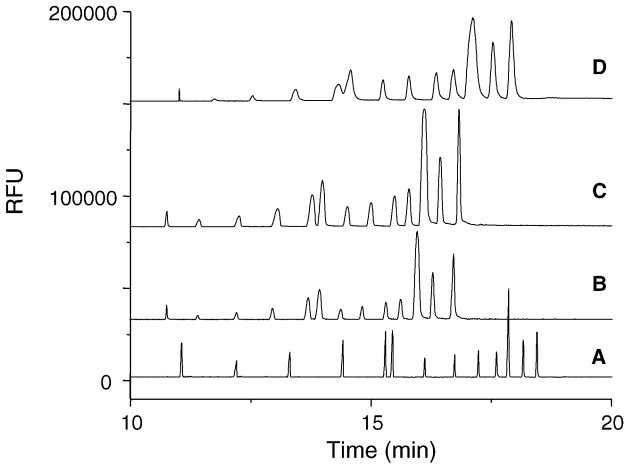

Fig. 3. Comparison of the effect of various non-covalent staining dyes on the separation efficiency of the 100 bp DNA ladder in CAE. (A) 500 nM ethidium bromide in the separation matrix; (B) 0.1× SYBR Green in the sample, (C) 0.1× SYBR Gold in the sample and (D) 1 μM YO-PRO-1 in the sample. Conditions as in Fig. 1.

this combination would significantly decrease spectral overlap. Fig. 3 compares the performance of several non-covalent DNA staining dyes, emitting in the low 500 nm range, to ethidium bromide. Trace A in Fig. 3 depicts the separation profile of the 100 bp DNA ladder with 500 nM ethidium bromide in the separation sieving buffer. Traces B, C and D delineate the electropherograms of the same ladder standard, pre-separation labeled with SYBR Green (0.1×), SYBR Gold (0.1×), and YO-PRO-1 (1 μM), respectively. Albeit, the analysis time in the three latter instances was approximately 20% shorter than that of with ethidium bromide containing sieving buffer, separation performance was apparently inferior, manifested in broader peaks and concomitantly lower resolution. Based on these results we have concluded that ethidium bromide containing separation matrix provided the highest separation performance with good detection sensitivity in CAE analysis of unlabeled dsDNA fragments.

A similar *in migratio* non-covalent labeling approach has been successfully applied to large-scale CAE analysis of single stranded oligonucleotide probes. In this instance, appropriate amount of GelStar staining dye (20000× diluted) was added to low concentration (4%) denaturing linear polyacrylamide gel matrix. Fig. 4A shows an overlay of 96 raw data traces obtained during regular quality control screening of a batch of in-house synthesized oligonucleotide probes in the size range of 50–70 bases. Fig. 4B and C depict two characteristic electropherograms illustrating better and worse synthesis performance, respectively. Please note that this low concentration sieving matrix in conjunction with non-covalent fluorophore labeling enabled rapid and adequate monitoring of all shorter oligonucleotide byproducts generated during the synthesis process.

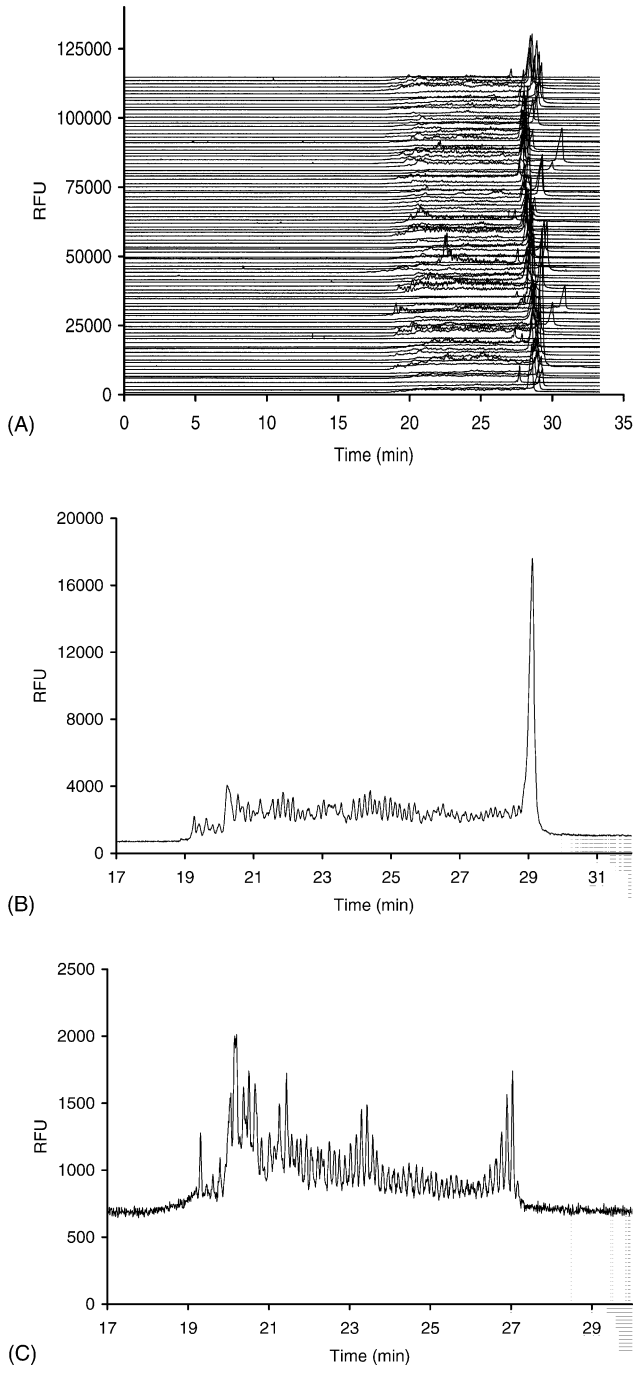

Fig. 4. Large-scale quality control screening of 96 in-house synthesized oligonucleotide probes using an automated CAE sequencing instrument: (A) overlay of 96 electropherograms. Selected separation traces in (B) and (C) depict better (63-mer) and worse (59-mer) quality synthesis products, respectively. Separation medium: 4% LPA, 7 M urea in 1× MegaBACE running buffer containing GelStar fluorescent staining dye (20,000 fold diluted). Conditions: 4 kV 10 s injection, 10 kV (160 V/cm) separation, temperature 27 °C.

4. Conclusion

A simple, rapid and economical DNA profiling method was developed using low concentration linear polymer based

separation matrices in a commercial 96-capillary array DNA sequencing system. The approach described in this paper greatly enhanced the productivity of DNA fragment analysis by automating the conventional manual procedure from sample injection to data analysis also reducing separation time and human intervention. Visualization of the dsDNA fragments was accomplished by non-covalent *in migratio* labeling with fluorescent dyes either dissolved in the separation matrix or added to the sample just prior to analysis. APTS and maltoheptaose-APTS conjugate were evaluated as internal standards detected in the green channel of the system. A mathematical algorithm was derived for adequate spectral overlap correction. Dynamic coating of the bare fused silica capillary arrays by the sieving matrix resulted in improved migration time reproducibility. Certain limitations of the CAE sequencing instrument, compared to a single capillary electrophoresis system, necessitated modifications of the operation protocols to optimize for DNA profiling. These limitations include the higher minimum separation temperature caused extra band broadening, possible siphoning effects due to unsynchronized movement of the cathode and anode stages, larger capillary diameter, and electric current constrain. The non-covalent labeling approach proved to be useful in large-scale quality assessment of in-house synthesized oligonucleotides probes.

Acknowledgment

The authors gratefully acknowledge the support of Illumina Corporation, especially Dr. Hongji Ren, as well as Drs. Mark Burk and Peter Domaille of Diversa Corporation.

References

[1] A. Chrambach, Practice of Quantitative Gel Electrophoresis, VCH, Deerfield Beach, FL, 1985.
[2] B.L. Karger, A.S. Cohen, A. Guttman, J. Chromatogr. 492 (1989) 585.
[3] J. Khandurina, A. Guttman, J. Chromatogr. A 943 (2002) 159.
[4] B.L. Karger, A. Guttman, Genet. Proteom. Technol. 3 (2003) 12.
[5] J. Khandurina, E. Legg, X. Wang, A. Guttman, Biotechniques 33 (2002) 1008.
[6] S.M. Clark, R.A. Mathies, Anal. Chem. 69 (1997) 1355.
[7] R.S. Madabhushi, M. Vainer, V. Dolnik, S. Enad, D.L. Barker, D.W. Harris, E.S. Mansfield, Electrophoresis 18 (1997) 104.
[8] I.V. Kourkine, C.N. Hestekin, B.A. Buchholz, A.E. Barron, Anal. Chem. 74 (2002) 2565.
[9] X.C. Huang, M.A. Quesada, R.A. Mathies, Anal. Chem. 64 (1992) 2149.
[10] A. Guttman, US Pat. 5332481 (1994).
[11] Y. Kim, E.S. Yeung, J. Chromatogr. A 781 (1997) 315.
[12] J.M. Song, E.S. Yeung, Electrophoresis 22 (2001) 748.
[13] H.E. Schwartz, K. Ulfelder, F.J. Sunzeri, M.P. Busch, R.G. Brownlee, J. Chromatogr. 559 (1991) 267.
[14] P.G. Righetti, C. Gelfi, J. Chromatogr. B 699 (1997) 63.
[15] L.M. Smith, J.Z. Sanders, R.J. Kaiser, P. Hughes, C. Dodd, C.R. Connell, C. Heiner, S.B. Kent, L.E. Hood, Nature 321 (1986) 674.
[16] A. Guttman, N. Cooke, Anal. Chem. 63 (1991) 2038.
[17] A. Guttman, F.T. Chen, R.A. Evangelista, N. Cooke, Anal. Biochem. 233 (1996) 234.
[18] D.N. Heiger, A.S. Cohen, B.L. Karger, J. Chromatogr. 516 (1990) 33.
[19] F.E.P. Mikkers, F.M. Everaerts, T.P.E.M. Verheggen, J. Chromatogr. 169 (1979) 1.