Evaluation of STR allele detection and data analysis using microcapillary array electrophoresis

Karin Crenshaw[1], Stephanie H.I. Yeung[2], Susan Greenspoon[3], Amy McGuckian[1], Cecelia Crouse[1] Jeff Ban[3], Richard A. Mathies[4]

[1]Palm Beach County Sheriff's Office Crime Laboratory, West Palm Beach, Fla; [2]UCSF/UCB Joint Graduate Group in Bioengineering, University of California Berkeley, Berkeley, Calif; [3]Virginia Department of Forensic Science, Forensic Biology, Richmond, Va; and [4]Department of Chemistry, University of California, Berkeley, Calif, USA

Abstract

The Palm Beach County Sheriff's Office (PBSO) Serology/DNA Section has conducted DNA analysis for over twelve years and has experienced the pressures associated with increased casework load. Since 2000, there has been a precipitous gap between the number of cases submitted and the number analyzed (46%). In order to address the issues of a burdened DNA program, the section identified and eliminated ineffective, laborious and time-consuming tasks while preparing for future technologies. Improvement in efficiency over the past few years includes using federally funded grants, internal support, and an external Process Mapping Team. The PBSO collaborated with forensic vendors, universities, and other forensic laboratories to enhance DNA testing procedures, including validation of the DNA IQ™ magnetic bead extraction system, robotic DNA extraction using the BioMek2000 and the ABI7000 Sequence Detection System. However, there is still a need to improve and streamline allele detection and analysis. In order to increase DNA profile output, the PBSO is currently engaged in a collaborative effort with the University of California, Berkeley and the Virginia Department of Forensic Science (VDFS) to evaluate a 96-channel microfabricated capillary array electrophoresis (μCAE) device for rapid separation of STR fragments. The μCAE instrument offers advantages over current multi-capillary STR detection platforms as it allows small sample volumes thus conserving evidence, increased sensitivity, and significantly greater throughput capabilities. Initial testing shows that 96 samples may be electrophoresed and analyzed in approximately 25 minutes. The system has the same capability of discerning minor alleles in mixture samples as current commercial CE instrument with a sensitivity of detecting from sample DNA down to 0.17 ng. Non-probative study has demonstrated the ability of the μCAE to analyze real-world casework samples.

Introduction

The forensic backlog crisis in more than 350 publicly funded crime laboratories of the nation was summarized in the United States Bureau of Justice Statistics 2002 Census of Publicly Funded Forensic Crime Laboratories report (1). The demands on public forensic DNA laboratories have increased dramatically over the past five years, causing a tremendous strain on laboratory space, scientists, quality-control/quality-assurance administrators, training programs, and information-system operators (1). Currently, the majority of forensic laboratories utilize the STR multiplex systems manufactured by Promega Corporation (Madison WI, PowerPlex®16 and 16BIO, PowerPlex 1.1, 2.1and 1.2) and Applied BioSystems (Foster City, CA AmpFℓSTR Profiler Plus, CoFiler and Identifiler). The most commonly used platforms for data capture and analysis include the ABI PRISM™ Genetic Analyzers 310, 3100, 3130 and 3100-Avant Capillary DNA Sequencer instruments and the MiraiBio II, IIe and III plus fluorescent image analysis systems using denaturing polyacrylamide gel electrophoresis. It takes several hours to nearly a day to conduct a complete analysis of amplified STR fragments. Regardless of the technologies and/or methodologies used in a forensic laboratory, strict quality-assurance and quality-control programs must be in place as well as adherence to national validation standards (2, 3). Therefore, evaluation of new technologies and developmental studies are imperative to their credibility and efficacy.

In order to increase the efficiency of allele detection and analysis, the forensic community is currently investigating the use of multi-capillary electrophoresis (CE) devices in order to increase the throughput of forensic DNA samples. The μCAE device coupled to a rotary confocal fluorescence scanning system developed at the University of California, Berkeley is currently being evaluated by both the PBSO and the VDFS for its feasibility and throughput for forensic use (4). This work will identify the optimal sample preparation and separation conditions needed to reliably perform these STR analyses. These studies will quantitatively benchmark the performance of our system compared to established commercial systems as well as more

advanced devices to be developed below. This paper introduces a portion of the evaluation studies that have been conducted using the Berkeley μCAE device to date. In addition, future studies will incorporate the use of energy-transfer cassettes-labeled primers for increased sensitivity (5).

Material and Methods

Samples: Biological samples used for the evaluation and comparison of the instruments and reagents presented herein are either samples previously tested at the PBSO including non-probative, semen, and population samples (28 Hispanic, 3 Asian and 17 African American), as well as sensitivity and mixed ratio samples generously provided by the National Institute of Standards and Technology (NIST). Sensitivity studies included serial dilution of GT37778 stock DNA (22, 11, 5.5, 2.75, 1.38, 0.69, 0.34, 0.17, 0.08, 0.043, 0.021, 0.011, and 0.0054 ng/μl) in TE^{-4} (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) buffer. Mixture and single-source study included rehydration of lyophilized female:male mixture samples at the ratios of 10:0, 9:1, 3:1, 3:2, 2:3, 1:3, 1:9 and 0:10 were reconstituted to a total concentration of 2 ng/uL with sterile water. Previously characterized population DNA extracts were re-quantified from lyophilized samples to an average of 1 ng/μL (0.75 – 1.5 ng/uL) for single source studies. Non-probative studies utilized PBSO DNA extracts from adjudicated forensic casework previously analyzed using the Hitachi FMBIO II allele detection instrument.

Extraction: A reagent negative (N) and positive control (P) were analyzed with all extraction runs. DNA samples were extracted and purified using a single step organic extraction (6). In addition, samples were extracted using the Promega DNA IQ™ (7) extraction kit (Promega, Madison, WI, USA). All procedure follows the manufacturer's instruction with modifications of the incubation temperatures (at 57 °C, 68 °C, or 95 °C) depending on the substrate composition for at least 30 minutes (if heat sensitive fabrics were used, e.g. polyester and nylon, samples were extracted without heating or at a lower temperature). For semen stains, differential extractions were performed as follows: a portion of the semen stain was incubated with 400 μL Tris/EDTA/NaCl, 25 μL 20% Sarkosyl, 75 μL H_2O, and 1 μL Proteinase K at 37 °C for 2 hours. After centrifugaion for 5 minutes, the supernatant removed (the Female fraction) and the sperm pellet washed 3 times in sperm wash buffer. The samples were then loaded onto the 96-well plate. The Beckman BioMek2000 was aligned and prepared following manufacturer recommendation (7,8).

Quantification: Quantification of DNA samples was conducted using QuantiBlot (Applied BioSystems, Foster City, CA, USA) and detected by the Hitachi CCDBIO (Hitachi, Alameda, CA, USA) as per protocol (9). In some cases, Millipore Microcon 100's (Fischer Scientific, Hanover, IL) were used to concentrate DNA samples.

PCR Preparation: DNA samples were amplified at the PBSO using the Promega Corporation GenePrint PowerPlex® 16 and ABI Profiler Plus STR multiplex systems according to the manufacturers' recommendations (10). 0.5 to 1 ng of DNA was used in the amplification reactions, unless otherwise specified. PCR fragments were resolved using an Embitec 3% agarose gel to assess the extent of amplified DNA (EmbiTec, San Diego, CA, USA) prior to polyacrylamide gel electrophoresis. Aliquots of the amplified products were shipped overnight to University of California, Berkeley laboratory and the VDFS for electrophoresis on the μCAE device and the ABI 310, respectively.

ABI Prism 310 Genetic Analyzer Sample Preparation and Operation: Amplification reactions, prepared in duplicate or halved for PowerPlex® 16 and AmpFℓSTR Profiler Plus, were analyzed on both the ABI 310 Genetic Analyzer and the μCAE. Sample preparations for the ABI Prism 310, capillary electrophoresis run parameters and data analysis using the GeneScan Data Collection v.2.1 and Analysis v.3.1 software programs and the PowerTyper™ 16 Macro, were as described in Technical Manual No. D012 (Promega Corp.).

μCAE Sample Preparation: The amplified PowerPlex® 16 samples (1 μL) was combined with 1 μL of ILS 600 in 6 μL of 50% formamide (1:1 deionized formamide solution in distilled deionized water). The PowerPlex® 16 allelic ladder (1.5 μL) was mixed with 1 μL of the ILS 600 in 4.5 μL of 50% formamide. ABI Profiler Plus® amplified samples were prepared as follows: 1 μL of sample was mixed with 0.8 μL of GeneScan 500 ROX and 6.2 μL of 50% formamide. The allelic ladder (4.5 μL) was mixed with 1.2 μL of GeneScan 500 Rox and 3.3 μL of 50% formamide. Prepared samples were denatured at 95 ºC for 3 minutes and quick cooled on ice prior to loading. 1.8 μL of each sample was manually loaded into 2 or 3 lanes per run. A minimum of 2 allelic ladder samples were run in parallel lanes.

Microdevice Design and Fabrication: The design and fabrication of the device are discussed in previous publications (11,12). All fabrication was performed at the Berkeley Microfabrication Laboratory, University of California. The microdevices were fabricated on 150-mm diameter borofloat glass wafers (Schott, Yonkers, NY). After photolithography, all features were isotropically etched to a depth of 25 μm with hydrofluoric acid (HF). Reservoirs were diamond-drilled into the etched wafers using a CNC mill. The patterned glass wafer was thermally bonded to a blank glass substrate to react a sandwich structure. After etch, the main separation channels are 200-μm wide by 25 μm deep capillaries stretching 16 cm from the twin-T 250-μm injector to the detection point (Fig. 1). Polyacrylamide is used to coat the channels using a modified Hjerten procedure (13), which prevents electro-osmotic flow.

μCAE Operation. The μCAE device is first filled with Long Read linear polyacrylamide (Amersham Biosciences, Piscataway, NJ) using a high-pressure filling station. Voltage to the samples was applied using an electrode-array ring placed in the sample wells. Continuous buffer reservoirs created using a polydimethlysiloxane (PDMS) elastomer ring was secured on top of the cathode and waste wells and filled with 5X TTE (250 mM Tris/ 250 mM TAPS / 5 mM EDTA, pH 8.3). The assembly was placed on the Berkeley rotary confocal fluorescence scanner, the stage heated to 67 ºC and equilibrated for 2 minutes. The PowerPlex® 16 samples were injected for 65 s and the Profiler Plus® samples for 45 s. Separation was achieved by grounding the cathode, applying 2500 V to the anode, 180 V to the sample and 200 V to the waste reservoirs. After each run, the sieving matrix was cleared from the microchannels and washed with deionized water using the same high-pressure apparatus previously used to load it.

Fluorescence detection, Data acquisition and Analysis. The Berkeley four-color rotary confocal scanner has previously been described in detail (14). A schematic of the system is presented in Figure 2. The four-color fluorescence data was first converted to binary format and appended with proper header information using a custom LabView program (National Instrument, Austin, TX, USA). The preprocessed data were then analyzed using the MegaBACE Fragment Profiler v1.2 and Genetic Profiler v2.2 software programs. Color separation matrices were created using either BaseFinder (15) or Fragment Profiler. To generate a color separation matrix using Fragment Profiler, the Matrix FL-JOE-TMR-CXR (Promega Corp.), was utilized. This matrix contains separate dye standards, each of which was electrophoresed in its own capillary on the μCAE. This is required by Fragment Profiler since each color is assigned a run file when generating the matrix table. Peak filters for the PowerPlex® 16 and AmpFℓSTR Profiler Plus STR multiplexes were created *de novo.* First, marker panels were defined in Genetic Profiler based on the size ranges for each locus provided by the manufacturers for both multiplexes using the Marker Panel Set Editor. These were imported into the Peak Filter Editor in Fragment Profiler. Minimum peak heights were defined by examining baseline noise for each of the dye channels and testing for false allele marking. For mixed samples, the maximum number of peaks allowed within a defined range (locus) was raised from 2 to 50 in order to accommodate mixed samples. The minimum height ratio for each range was set at 0.2 and the left stutter veto threshold was set at 0.66 with the pattern type defined as a simple multinucleotide right filter pattern. Allele bin sets were also created for both PowerPlex® 16 and AmpFℓSTR Profiler Plus multiplexes from an existing bin set, using Fragment Profiler. Dynamic binning was employed initially to assist in better defining the bin sets. Multiple independent μCAE runs were analyzed and numerous modifications made to the working bin sets. Once bin sets were defined to accurately labeled peaks, static binning was employed, however, slight adjustments to the positioning of the bin sets may occasionally be needed to accommodate run to run variation. The dye color and sizes of the in-lane-standards (ILS600 for PowerPlex® 16 and GeneScan ROX500 for Profiler Plus) were defined using the Manage Dye Standards feature in Fragment Profiler. Independent data analysis using Fragment Profiler and Genetic Profiler was performed at the VDFS and PBSO in order to evaluate μCAE genotyping data.

Results

The goal of the collaborative effort among the Department of Chemistry at the University of California, Berkeley, the VDFS and the PBSO is to evaluate the feasibility of using a 96-channel radial μCAE-based platform to perform DNA profiling on casework evidence. The performance evaluation of the μCAE instrument is being conducted using the FBI Standards (2,3) as a model and includes precision, sensitivity, reproducibility, mixture, and non-probative testing. Detailed results of these studies will be presented elsewhere (4, manuscript submitted).Additional testing will be conducted to determine if there is any carryover or contamination issue associated with the μCAE process.

Efficiency: Initial evaluation results have shown that the μCAE device separates STR fragments in a fraction of the time needed for flatbed or conventional capillary devices such the CE310 instrument. Figure 3

demonstrates the ability of the Berkeley μCAE device to separate the PowerPlex 16 allelic ladder and Internal Lane Standard in < 22 minutes and the ProfilerPlus ™ allelic ladder and Internal Lane Standard in <18 minutes (Figure 4). A typical amplified product separation occurs in < 25 minutes for up to 96 samples.

Resolution and Sensitivity: The ability of the μCAE to resolve single base pair differences between alleles is evident in Figure 5A with the 9.3 and 10 THO1 alleles easily resolved. The ability for an instrument to detect low level amplified products is important for forensic DNA casework. PowerPlex® 16 and Profiler Plus™ samples were amplified with DNA serially diluted including 22ng, 11, 5.5, 2.75, 1.38, 0.69, 0.34, 0.17, 0.08, 0.043, 0.021, 0.011 and 0.0054ng. Using both multiplex systems, all amplified STR alleles were detected at down to 0.17ng of DNA and with complete concordance and similar sensitivity as observed with the same samples analyzed using the ABI Prism 310. Figure 5B demonstrates the profile obtained for one of the PowerPlex® 16 channels at 0.56 ng of template DNA.

Non-Probative Samples: Typing of seventeen non-probative DNA samples from case evidence previously processed by PBSO using both the PowerPlex 16® and Profiler Plus™ systems in which the DNA samples were extracted from evidentiary samples including blood, sex assault, paternity and mixed blood stains. PBSO has previously analyzed these probative casework samples using the Hitachi FMBIO II allele detection system. All standards evaluated generated complete DNA profiles using the μCAE system and ABI Prism 310 analysis. All results obtained from the μCAE device were consistent with the ABI Prism 310 results and previously reported results.

Mixtures: Biological evidentiary samples in which two or more individuals have contributed to the evidentiary stain is common in a forensic setting. Resolution of DNA STR mixtures is essential in forensic casework. Comparison of the reliability of the μCAE device with the ABI 310 with PowerPlex® 16 amplified products was conducted. The samples consisted of female and male DNA at the ratios of 10:0, 9:1, 3:1, 3:2, 2:3, 1:3, 1:9 and 0:10. All minor components were successfully detected and typed at ratios of 3:1 and 1:3 samples (data not shown). Figure 6 demonstrates the results from the Berkeley μCAE device when a mixture from a non-probative case was tested. The results of the PowerPlex 16 profile show both the major and minor contributors to this blood mixture. More extensive studies consisting of non-probative casework mixtures will be conducted to further evaluate if the capacity to detect minor alleles is demonstrated with the μCAE instrument.

Software Analysis: The Berkeley μCAE device software for data analysis is the MegaBACE Fragment Profiler and Genetic Profiler software programs. These programs were selected because they allowed the user to define parameters such as spectral overlap matrix, size standards, dye sets, and bin sets (Fig. 7). Analyzed data can then be viewed on one screen and the user can choose to view multiple panes at one time. The sample pane (Fig. 8, A) shows the sample(s) that are currently selected for viewing. The filter pane (Fig. 8, B) shows the markers that are available to view and indicates the ones that have been selected. The histogram (Fig. 8, C) is available to view and shows the bin sets for each of the size standards associated with a particular genetic marker. The trace for either a single sample or multiple samples can be viewed at the same time as well (Fig. 8, D). The buttons at the top of the screen (8E) allow the user to make changes to the information that is viewed, including adding additional samples to the trace view, enlarging or minimizing the number of markers that are viewed in the trace view, the channels that are viewed, and the bin sets within the histogram can also be modified. Some minor adjustments were required with this software. Trouble-shooting a DNA profile in which no alleles were called for one of the amplified samples is shown in Figure 9. The Internal Lane Standard contains two peaks that were marked for analysis (9A) that should not have been marked and two peaks that were not marked (9B, C) that should have been marked for analysis. This caused a shift for all of the alleles for the amplified samples. This shift resulted in no allele calls for the entire sample; therefore, each allele was labeled with a question mark ("?"). This was corrected by first re-analyzing the ladder to take out the two lower molecular weight standards and adding in the higher molecular weight markers followed by clicking on the re-analyze tab (Fig. 10, A). The result is a profile in which all of the alleles have been identified and called correctly.

Conclusion

Many high-throughput forensic and medical DNA laboratories have implemented protocols using robotics and microchip devices. The evaluation of the next-generation 96-channel μCAE device by the University of Berkeley, VDFS and the PBSO is on-going. To date, it has demonstrated that there will be a remarkable increase in the efficiency of the forensic laboratory to conduct DNA analysis on an ever-increasing caseload in a timely manner at a lower cost without compromising the quality. It is expected that in conjunction with

extraction robotics and qPCR protocols, the use of the μCAE instrument will greatly enhance the throughput of forensic DNA laboratories. An additional important advantage is that future microfabricated chips will contain integrated sample preparations such as PCR, which will further enhance the speed and automation, improve reliability, and reduce cost and labor for analysis of forensic DNA cases (15).

Acknowledgements

These studies are funded through a National Institute of Justice grant. Technical help from Margaret Kline at the National Institute of Standards and Technology is greatly appreciated. The authors think Dennis Meyers at GE Healthcare for his valuable assistance in using the Fragment Profiler software.

References

1. United States Department of Justice, Bureau of Justice Statistics Bulletin. Census of publicly funded forensic crime laboratories, February 2005 NCJ 207205. Available at: *http://www.ojp.usdoj.gov/bjs/whtsnw2.htm*. Accessed: April 20, 2005.
2. Forensic Science Communications. Quality assurance standards for forensic DNA testing laboratories. Vol 2. 2000.
3 Federal Bureau of Investigation. Quality assurance standards for forensic DNA testing laboratories. October, 1998.
4. Yeung, S. H. I., Greenspoon, S.A., McGuckian, A.B. Crouse, C.A., Emrich, C.A., Ban, J. and Mathies, R.A. Rapid and high-throughput forensic short tandem repeat typing using a 96-lane microfabricated capillary array electrophoresis microdevice. *J. Forensic. Sci.* (2005). (Manuscript Submitted)
5. Berti L, Medintz IL, Tom J, Mathies RA. Energy-transfer cassette labeling for capillary array electrophoresis short tandem repeat DNA fragment sizing. Bioconjug Chem. 2001;12:493-500.
6. Sambrook J, Fritsch EF, Maniatis T. Bacteriophage lambda growth, purification and DNA extraction. In: Sambrook J, Fritsch EF, Maniatis T, editors. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1989. vol. 1 p. 2.60 – 2.81
7. Promega's DNA IQ* DNA isolation system automated on the Beckman Coulter Biomek® 2000 Laboratory Workstation. Available at: *http://www.beckman.com/resourcecenter/eLabNotebook/ForensicID/dna-iq-bmk2000/index.htm*. Accessed: April 20, 2005.
8. Promega Corporation: DNAIQTM System-Small Sample casework Protocol; http://www.promega.com
9. Crouse C, Stephanie Yeung, Susan Greenspoon, Amy McGuckian, Julie Sikorsky , Jeff Ban and Richard Mathies, "Improving the Efficiency of a Small Forensic DNA Laboratory: Evidence Tracking, Robotic Assays and Future Microcapillary Array Devices" (Croatian Medical Journal, August, 2005 Vol46, 563-577
10. PowerPlex® 16 BIO system. Technical Manual. Madison (WI): Promega Corporation; 2001.
11. Paegel BM, Emrich CA, Weyemayer GJ, Scherer JR and Mathies RA. High throughput DNA sequencing with a microfabricated 96-lane capillary array electrophoresis bioprocessor. Proc. Natl. Acad. Sci. U. S. A. 2002; 99:574-9.
12. Simpson PC, Wolley AT and Mathies RA. Microfabrication technology for the reduction of capillary array electrophoresis chip. Biomed. Microdevices 1998; 1:7-26.
13. Hjerten S. High-performance electrophoresis - elimination of electroendosmosis and solute adsorption. J. Chromatogr. 1985; 347:191-8.
14. Shi YN, Simpson PC, Scherer JR, Wexler D, Skibola C, Smith MT, et al. Radial capillary array electrophoresis microplate and scanner for high-performance nucleic acid analysis. Anal. Chem. 1999; 71:5354-61.
15. Paegel BM, Blazej RG, Mathies RA. Microfluidic devices for DNA sequencing: sample preparation and electrophoretic analysis. Curr Opin Biotechnol. 2003;14:42-50.

Figure 1. Microfabricated capillary array electrophoresis chip design. There are 48 doublet structures etched on a 150-mm diameter glass wafer forming 96 separation channels (A). A common cathode and waste well are shared by each doublet (B). The sample is first injected from the sample well towards the waste well. By switching the electric field between the cathode and the anode, the sample plug is electrophoresed down the length of the channel towards the central anode.

Figure 2. The rotary confocal fluorescence scanning system. An argon ion laser beam (488-nm) is passed through a dichroic beam splitter and travels up the hollow shaft of a stepper motor. The beam is deflected 1 cm off the axis of rotation by a rhomb prism and focused by a 60X objective on the μCAE channels. The fluorescence travels back along the same path, passes through the dichroic beam splitter and enters a 4-color confocal detector consisting of a series of beamsplitters, filters and four PMT detectors.

Figure 3. Four-color traces of a PowerPlex 16 Allelic Ladder sample analyzed on the μCAE device.

Figure 4. Four-color traces of a ABI Profiler Plus Allelic Ladder sample analyzed on the μCAE device.

Figure 5. μCAE traces of single-source DNA sample amplified with PowerPlex16 BIO and data interpreted using the MegaBACE program. (A) Resolution of the 9.3/10 THO1 alleles in a non-probative DNA sample. (B) Example of the TAMRA fluorophore channel for a 0.56-ng sensitivity sample.

Figure 6. Four-color μCAE traces of a PowerPlex 16® non-probative DNA mixture sample from case evidence previously processed by PBSO. All results obtained from the μCAE device were consistent with those previously reported.

Figure 7. Screen capture of "Add Run" window of the MegaBACE Fragment Profiler, showing settings and parameter options for analysis.

Figure 8. Screen capture of Fragment Profiler part of MegaBACE program, showing data analysis and review windows available. Circles indicate the control panel for the Sample Pane showing which samples are currently shown (A), the Filter Pane showing the genetic markers for the panel and the corresponding dye (B); the Histogram window which allows the user to view, add, delete or adjust bin sets (C); the Trace window showing the current sample and genetic marker(s) chosen (D) and adding lane traces, widening, or shrinking lane trace views, dye sets to be viewed, and histogram control panel (E).

Figure 9. Screen capture showing the adjustment needed when the software cannot determine any of the alleles in a sample. The internal lane standards for that sample show that the lowest two peaks (A) are being used, while two larger size standards (B and C) are not being used. This causes a shift in the allelic ladder.

Figure 10. Screen capture showing the correct settings for the internal lane standard. The lowest two peaks are no longer marked, while the larger two peaks are now marked to be included. The re-analyze button (A) will re-calculate the allele calls for this particular sample after these adjustments have been made.