Genotyping Energy-Transfer-Cassette-labeled Short-Tandem-Repeat Amplicons with Capillary Array Electrophoresis Microchannel Plates

IGOR L. MEDINTZ, LORENZO BERTI, CHARLES A. EMRICH, JENNIFER TOM, JAMES R. SCHERER, and RICHARD A. MATHIES* 

Background: Genetic analysis of microsatellite DNA is a powerful tool used in linkage analysis, gene mapping, and clinical diagnosis. To address the expanding needs of studies of short tandem repeats (STRs), we demonstrated high-performance STR analysis on a high-throughput microchannel plate-based platform.

Methods: Energy-transfer-cassette-labeled STR amplicons were separated and typed on a microfabricated 96-channel radial capillary array electrophoresis (CAE) microchannel plate system. Four-color detection was accomplished with a laser-excited confocal fluorescence rotary scanner.

Results: Multiplex STR analysis with single base-pair resolution was demonstrated on denaturing polyacrylamide gel media. The high-throughput multiplex capabilities of this genetic analysis platform were demonstrated by the simultaneous separation of STR amplicons representing 122 samples in ninety-six 5.5-cm-long channels in <8 min. Sizing values obtained for these amplicons on the CAE microchannel plate were comparable to those measured on a conventional commercial CAE instrument and exhibit <1% sizing variance.

Conclusions: Energy-transfer-cassette labeling and microfabricated CAE microchannel plates allow high-performance multiplex STR analyses.

© 2001 American Association for Clinical Chemistry

Microsatellite DNA or short tandem repeat (STR) sequences are important in many types of genetic analysis, such as loss of heterozygosity testing, linkage analysis, gene mapping and discovery, paternity and forensic identity testing, evolutionary studies, and clinical diagnostics (1–10). STR analysis is highly informative because of the large number (>5200) of mapped loci available (11) and because these loci are highly polymorphic and widely distributed across the genome (6). STR analysis usually consists of locus-specific PCR amplification followed by electrophoretic separation of the products (3, 10, 12).

Many STR-based genetic studies, such as linkage analysis, gene mapping, and loss of heterozygosity testing, require the screening of large numbers of microsatellite loci in thousands of individuals (1, 3–5, 10). To address the needs of these and future STR-based studies, new generations of genetic analysis devices and methods that can rapidly perform sensitive, specific, and high-throughput STR analyses are required (13, 14).

Microfabricated capillary electrophoresis (CE) devices offer numerous advantages that can help meet the expanding need for genetic analysis. These advantages include the benefits of smaller sample volumes, higher assay speed and sensitivity, and the ability to densely pack separation channels into monolithic arrays on chips or wafers (15). High-speed sizing of DNA restriction fragments, PCR products, and STRs, as well as multiplex STR typing have already been demonstrated on prototype microfabricated CE systems (16–20).

The rapid analysis capabilities of these devices have been used for the diagnosis of lymphoproliferative disorders, herpes simplex encephalitis, and genotyping of a hereditary hemochromatosis-associated mutation (21–23). Single-channel microchip devices have been used for the analysis of simple sequence length polymorphisms in mice and for...
the determination of single-nucleotide polymorphism sites in the P53 tumor suppressor gene (24, 25). Finally, high-performance single-nucleotide polymorphism genotyping of 96 MTHFR alleles in <90 s has been demonstrated on a radial capillary array electrophoresis (CAE) microchannel plate (MCP) system with laser-excited rotary scanning confocal fluorescence detection (20).

Energy-transfer (ET) labeling of DNA sequencing samples offers the advantages of excitation at a single common laser wavelength with distinctive and intense acceptor dye emission and matched electrophoretic mobilities of the labeled fragments (26–29). Covalent labeling with ET tags can also be used to increase the throughput of genotyping analysis with CAE. For example, ET-labeled primers were used to generate allele-specific PCR products for the C282Y (845G→A), H63D (187C→G), and S65C (193A→T) hereditary hemochromatosis diagnostic mutations at a reference population of >100 samples (30). The mixed allele-specific PCR products were rapidly separated and genotyped in <10 min on a 96-channel radial MCP with single base-pair resolution. Recently, Berti et al. (31) developed a new ET cassette technology that can be used to label any PCR primer, sequencing primer, or other target of interest. The utility of this ET-cassette-labeling strategy has already been demonstrated for sequencing, PCR fragment sizing, and analysis of STR loci on commercial CAE instrumentation (31, 32).

In this report we examine the utility of a 96-channel radial CAE microplate coupled to a laser-excited rotary confocal fluorescence scanner for high-performance STR analysis. ET-cassette-labeled STR amplicons were separated in parallel and genotyped in <8 min with single base-pair resolution, and multiplex analysis formats were easily achieved. The sizing values obtained on the microfabricated CAE device were comparable to those obtained for the same amplicons on the MegaBACE-1000 CAE system. This study demonstrates the powerful combination of CAE MCPs and ET-cassette labeling for high-performance STR analysis.

Materials and Methods

DNA

Purified DNA from the K562 cell line was purchased from Promega Corp. Universal Centre d’Etude du Polymorphisme Humain (CEPH) donor DNA (cat. no. NA10859) was purchased from the Coriell Cell Repository. A genomic DNA sample containing the previously characterized THO1 9.3/10 alleles was generously provided by G. Sensabaugh (School of Public Health, University of California, Berkeley, CA) (32, 33).

ET-CASSETTE-Labeled PRIMERS

The synthesis of the ET cassettes and their characterization have been discussed extensively by Berti and coworkers (31, 32). Briefly, an ET-cassette-labeled primer consists of three parts: the ET cassette itself, the primer, and the disulfide linker joining the two at their 5' ends (see Fig. 1). The cassette moiety consists of a sugar-phosphate spacer with a 6-carboxyfluorescein (6-FAM; emission maximum at ~497 nm) donor at the 3' end, an acceptor dye linked to a modified T-base at the 5' end of the spacer, and a mixed disulfide group for coupling to the 5' end of a thiol-modified primer (31, 32). Acceptor dyes include 6-carboxyfluorescein-110 (R110); emission maximum at ~525 nm), 6-carboxyfluorescin-6G (R6G; emission maximum at ~555 nm), carboxytetramethylrhodamine (TMRA; emission maximum at ~572 nm), and 6-carboxy-X-rhodamine (ROX; emission maximum at ~620 nm). A particular ET cassette is described by the abbreviation D-Sn-A, where D is the donor, A is the acceptor, and Sn indicates the number of sugar phosphate monomers constituting the spacer (27). The ET cassettes with R110 and TAMRA as the acceptor dyes used a spacer with eight sugar phosphates (S8), whereas cassettes with R6G and ROX as the acceptors used spacers with seven sugar phosphates (S7). In this cassette format, the modified T to which the acceptor dye was attached functioned as part of the spacer. Therefore, the spacings for the cassettes were functionally nine units for the R110- and TAMRA-containing cassettes and eight units for the R6G and ROX cassettes (31).

The primer sequences are presented in Table 1. Modified primers for ET-cassette conjugation were purchased from Operon Technologies with a 5'-protected thiol (C6 S-S modification). Additional primers were obtained from Gibco BRL. The coupling reaction is a two-step process consisting of primer deprotection followed by conjugation with the ET cassette. The ET-cassette-labeled primers were purified by HPLC and quantified by ultraviolet-visible spectroscopy (31, 32).

PCR AMPLIFICATION

Genomic DNA (50 ng) was used for each PCR reaction along with 15 pmol of both the ET-cassette-labeled primer and the appropriate reverse primer. Reactions used Qiagen PCR Master Mix. Reaction conditions used were as described for specific STR loci (34)2 on a MJ Research PTC-100 Programmable Thermal Cycler. Successful PCR was verified by agarose gel electrophoresis.

2 For more information on the instrument used in Ref. (34), see http://www.mdyn.com.
**Table 1. Primer sequences.**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Chromosomal location</th>
<th>Locus definition</th>
<th>Repeat sequence</th>
<th>Labeled sequence, 5'-3'</th>
<th>PCR reverse primer sequence, 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF1PO-forward</td>
<td>5q33.3-q34</td>
<td>Human c-fms protooncogene for CSF-1 receptor gene</td>
<td>AGAT</td>
<td>AACCTGAGTCTGCCAAGGACTAGC</td>
<td>TCCACACACACTGCGCATCTTC</td>
</tr>
<tr>
<td>CSF1PO-reverse</td>
<td>5q33.3-q34</td>
<td>Human c-fms protooncogene for CSF-1 receptor gene</td>
<td>AGAT</td>
<td>TTCCACACACACTGCGCATCTTC</td>
<td>AACCTGAGTCTGCCAAGGACTAGC</td>
</tr>
<tr>
<td>D7S820</td>
<td>7q11.21-q22</td>
<td>NA*</td>
<td>AGAT</td>
<td>GTGCATAGTTTAAGAACGAACTAAC</td>
<td>CTGAGATTATCAAAAAACTCGAGG</td>
</tr>
<tr>
<td>D13S317</td>
<td>13q22.2-q31</td>
<td>NA</td>
<td>AGAT</td>
<td>ACAGAAGATAGGATGTGGA</td>
<td>GCCCAAAAGACAGACAGAA</td>
</tr>
<tr>
<td>TH01</td>
<td>11p15.5</td>
<td>Human tyrosine hydroxylase gene</td>
<td>AATG</td>
<td>GTGGGCTGAAAAAGCTCCCAGATT</td>
<td>ATTCAGAGGATATCGGGCTGAGG</td>
</tr>
<tr>
<td>TPOX</td>
<td>2p25.1-pter</td>
<td>Human thyroid peroxidase gene</td>
<td>AATG</td>
<td>ACTGCCACAGAAACCGCATTTAG</td>
<td>GGAGGAACGAGGAAACACAGGT</td>
</tr>
<tr>
<td>vWA-reverse</td>
<td>12p12-pter</td>
<td>Human von Willebrand factor gene</td>
<td>AGAT</td>
<td>GGACAGATGATAAATACATAGGGATGAGG</td>
<td>GAAAGCCTTAGGTTAGTGAAGAATAAT</td>
</tr>
</tbody>
</table>

* Data from Promega technical manual (39), as cited at www.cstl.nist.gov.

**SAMPLE PREPARATION AND SIZING STANDARDS**

PCR amplicons were desalted using the Qiagen PCR Purification Kit. Samples were resolubilized in 50 μL of 0.5× Tris-EDTA buffer. Approximately 0.5 μL of sample was then mixed with 5 μL of 750 mL/L deionized formamide containing dye-labeled sizing markers at a concentration of 10 fmol/L. The FAM and TAMRA Map-marker Sizing Standards consist of 20 fragments (60, 90, 100, 120, 140, 160, 180, 190, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, and 400 bp) and were obtained from Bio Ventures Inc. The ET-400-R sizing standard obtained from Ventana Inc. The ET-400-R sizing standard contains 96 capillaries, each with a 40-cm effective separation length of 150 mm in diameter (20, 30).

**MegaBACE-1000 CAE**

The analysis of these samples on the MegaBACE-1000 capillary array electrophoresis system (Molecular Dynamics) is discussed extensively in Berti et al. (32). This CAE system contains 96 capillaries, each with a 40-cm effective separation length (10, 14, 34). Samples were separated in Long Read linear polyacrylamide matrix (LPA) containing 7 mol/L urea in a Tris-TAPS buffer (Amersham/Pharmacia Biotech). Denatured samples were introduced into the capillaries by electrokinetic injection for 45 s at 3 kV and then electrophoresed at 10 kV for 75 min.

**MICROFABRICATION AND MICROPLATE DESIGN**

CAE MCPs were fabricated at the University of California-Berkeley Microfabrication Laboratory as described previously (20, 35). The design of the chip shown in Fig. 2A is a modification of one presented previously in that the substrate is now 150 mm in diameter (20, 30). Isotropy etching with HF formed ~110-μm wide by 50-μm deep channels. The distance along the separation capillary from the 250-μm twin-T injector to the detection point was 55 mm. The microchannels were coated with polyacrylamide, as described by Hjerten (36), to prevent electroendosmotic flow. For electrophoresis, the microplates were filled with Long Read LPA by use of a microplate gel loader/pressure washer device (30).

**ELECTROPHORESIS AND MICROPLATE SCANNING**

Sample loading, injection, and separation were performed as described by Shi et al. (20) and Medintz and coworkers (23, 30). Briefly, the loaded microplates were placed on the microplate holder of the rotary confocal scanner and heated at 40 °C, and a circular electrode array was placed on top of the microplate, making electrical contact with all the reservoirs. Samples underwent electrokinetic injection for 100 s by the application +5 V to the sample reservoir, +425 V to the waste reservoir, +50 to the cathode reservoir, and +200 V to the anode reservoir. Separation was carried out immediately after injection by the application of +1350 V at the anode reservoir, +200 V at the cathode, and +325 V to the sample and waste reservoirs.

During separation, samples were detected within the microplate by the laser-excited, rotary confocal fluorescence scanner (see Fig. 2C). The design and function of the scanner have been discussed extensively by Shi et al. (20). Briefly, the detection system consists of a rotating objective head coupled to a confocal detection unit, allowing four-color analysis (20, 37). The present scanner has been updated by the addition of a stepper motor with a hollow shaft and a diode-activated trigger for scan initiation. A 488-nm beam from an Ar+ laser is used for excitation, and the fluorescence is gathered by the microscope objective and sent to the four fluorescence detection channels. The “blue” channel detects at 505–530 nm (R110 maximum emission, ~525 nm); the “green” channel detects at 530–
560 nm (R6G maximum emission, ~550 nm); the “black” channel detects at 560–590 nm (TAMRA maximum, ~572 nm); and the “red” channel detects at >590 nm (ROX maximum, ~620 nm). The detection limit of this system has been estimated to be ~3 pmol/L fluorescein (signal-to-noise ratio = 1).

DATA ANALYSIS

Data for each 96-channel run were collected and stored as a data-appended text (DAT) files written to a specific run folder. Raw data files were converted to electrophoretic signal data (ESD) file formats by a TEXT-to-ESD conversion program (38). ESD data files were analyzed with Genetic ProfilerTM Software, Ver. 1.1 (Molecular Dynamics), where they underwent background subtraction by a spectral separation or cross-talk matrix. Unknown samples were sized against the DNA sizing ladders included in each capillary by a third-order local algorithm (10, 38).

Results

CAE MCP STR ANALYSIS

Using the primers described in Table 1, we successfully generated all four-color ET-cassette-labeled amplicons for each primer from the K562 and CEPH DNA, as described by Berti et al. (32). Amplicons were also generated from a THO1 9.3/10 allele-containing sample. The THO1 9.3 allele is a common variant that differs from the 10 allele by only 1 bp, instead of the usual 4-bp repeat (33, 39). This pool of samples was then analyzed on the MCP. Fig. 3A shows the processed electropherogram of
the vWA-specific R110-labeled amplicons generated from the CEPH donor DNA and separated against the TAMRA Mapmarker ladder. Fig. 3B shows the D7S820-specific R6G-labeled amplicons generated from the same CEPH DNA and sized with the ET-400-R ladder. Fig. 3C shows the separation of the TAMRA-labeled THO1 9.3/10 alleles against the FAM Mapmarker ladder. Note that single base-pair resolution is achieved in this separation. Fig. 3D shows the separation of the TPOX-specific ROX-labeled amplicons generated from the CEPH DNA. Sample volumes analyzed were only 0.5 μL.

SIZING OF PCR AMPICONS

Shown in Table 2 are the K562 allele size ranges for all four of the ET-labeled amplicons at the seven loci, measured on the MegaBACE-1000 CAE system (32), compared with the values obtained on the MCP for the same amplicons. The largest difference in mean sizing value (2.2 bp) was observed at the D7S820 locus for the 11 allele. This difference was <1% of the 229.0–233.8 bp amplicon size range. At the other extreme, no difference in mean sizing value was obtained for the TPOX 8 allele. The difference in mean allele size range obtained when the ET cassette label was alternated between the forward or reverse CSF1PO primers (see Table 1) for amplicon generation was almost negligible (0.2–0.4 bp, or <0.1% for a >300-bp size product). The SD among all four colors for each locus obtained on the MCP was also small, ranging from 0.2 bp for both the THO1 9.3 allele and the TPOX 8 allele to 1.5 bp for the CSF1PO (rev) 10 allele.

MULTIPLEX ANALYSIS

Coanalyzing multiple differentially labeled amplicons within a single channel can greatly increase the throughput of STR analysis. To explore the feasibility of this multiplex STR approach on MCPs, we genotyped five different amplicons in a single separation. The K562 DNA vWA-ROX and THO1-R6G amplicons as well as the CEPH donor DNA THO1-R6G, TPOX-TAM, and CSF (fwd)-ROX amplicons were mixed together. Fig. 4 shows the four-color separation and genotyping of this fiveplex against the FAM Mapmarker sizing ladder. All of the amplicons were well separated, and their sizing values did not differ from the values presented in Table 2 or from their individual analyses (data not shown).

HIGH-THROUGHPUT CAE ANALYSIS

The essence of high-performance STR analysis is the ability to rapidly, efficiently, and precisely analyze multiple samples in parallel. To this end, we performed 96 simultaneous separations on the radial CAE MCP (Fig. 5). Each electropherogram in this image was individually aligned to match the peaks of the 60-, 70-, and 400-bp

---

### Table 2. Analysis of observed K562 allele sizes using ET-labeled STR amplicons.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Alleles</th>
<th>Determined size range of all four ET-cassette-labeled amplicons, bp</th>
<th>Difference between determined sizes, bp</th>
<th>SD of four-color MCP sizing values</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSF1PO (fwd)</td>
<td>9/10</td>
<td>315.6–317.5/320.5–321.5</td>
<td>0.4/0.3</td>
<td>0.9/0.9</td>
</tr>
<tr>
<td>CSF1PO (rvr)</td>
<td>9/10</td>
<td>317.3–318.3/321.3–322.4</td>
<td>0.4/0.2</td>
<td>1.1/1.5</td>
</tr>
<tr>
<td>D7S820</td>
<td>9/11</td>
<td>222.7–224.6/231.1–233.8</td>
<td>2.0/2.2</td>
<td>1.4/1.1</td>
</tr>
<tr>
<td>D13S317</td>
<td>8/8</td>
<td>187.6–188.5</td>
<td>1.6</td>
<td>0.5</td>
</tr>
<tr>
<td>THO1</td>
<td>9.3/9.3</td>
<td>209.7–210.3</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>TPOX</td>
<td>8/9</td>
<td>243.8–244.9/247.9–248.9</td>
<td>0.0/0.4</td>
<td>0.2/0.6</td>
</tr>
<tr>
<td>vWA</td>
<td>16/16</td>
<td>163.1–163.8</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

---

*a From Promega technical manual (39), as cited at www.cstl.nist.gov.

*b Size of each allele averaged from two capillaries (MegaBACE) or two microchannels (MCP).

*c Mean size of all four MegaBACE values – mean size of all four MCP values.

*d Fwd and rev indicate the primer position of the ET cassette.
standards. Channels 1–48 used the ET-400-R sizing standard with FAM-, R6G-, and TAMRA-labeled STR samples. Channels 49–96 used the FAM Mapmarker Sizing Standard with R6G-, TAMRA-, and ROX-labeled samples. Each channel contained one to three STR amplicons with a total of 122 samples present. Note that all separations were complete in <8 min and that all of the amplicons and standards were well resolved. The small amount of channel-to-channel variation in migration times and patterns observed was attributable to the nonuniform positioning of the electrodes in the injector reservoirs as well as microchannel variation and did not affect each channel’s internally referenced mobility (20, 30).

Discussion
This study was performed to explore the feasibility of using radial CAE MCPs with four-color fluorescence detection for high-performance STR analysis. We used a pool of ET-cassette-labeled amplicons derived from characterized DNA samples that had been amplified with seven different STR-specific primers, each in the four ET-cassette colors. These samples were analyzed on our MCP platform and on the MegaBACE-1000 CAE instrument as a control. The spacings between determined allele sizes were strictly maintained and corresponded to the 4-bp increments predicted for the CSF1PO, D7S820, TPOX, and vWA loci (see Table 2). The largest difference between the
mean sizes obtained with the MegaBACE-1000 and the MCP platform was 2.2 bp at the D7S820 locus for the 11 allele, which represents a <1% difference for an amplicon of this size. Eight of the other 11 mean sizing differences between the two systems were <1 bp (<0.2%). This negligible variability was obtained despite the fact that the microplate separation was performed in <8 min on a separation capillary only 5.5 cm long. The time for the MCP electrophoresis was 10-fold less than the 75 min required for the MegaBACE system, and the effective separation distance on the CAE microplate was ~8-fold less than the 40-cm separation distance in the MegaBACE-1000 instrument. This comparison demonstrates the improved STR analysis capabilities available on CAE MCPs.

The ability to analyze multiple differentially labeled samples in a single channel is important for high-throughput analysis. To demonstrate multiplex analysis with the microchannel platform, we premixed five different amplicons and analyzed this fiveplex on the CAE microplate. All five amplicons were clearly resolved and genotyped although there was an eightfold difference in relative intensity between the utilized amplicons. To verify the high-performance capabilities of the MCP format, we simultaneously separated and genotyped 122 STR samples in 96 channels in <8 min. Each sample contained one to three STR amplicons. All amplicons were clearly resolved and genotyped in this four-color analysis. This work demonstrates the feasibility of collecting large amounts of data on a very rapid time scale with radial CAE MCP systems.

The CAE MCP format together with ET-cassette labeling has numerous inherent advantages for STR analysis. With >5200 loci available for analysis in the human genome (11), the ability to rapidly and efficiently tag any STR primer with any ET-cassette label will be particularly useful in designing and implementing new assays. Importantly, primers to be labeled in this format can be ordered commercially for facile conjugation with ET cassettes (31, 32). ET-cassette technology will thus facilitate expanded multiplex analysis formats. Furthermore, the volumes of final product analyzed are very small, representing only 1/100th of the total amount of purified PCR amplicon generated. This should make it possible to scale down reactant volumes and realize substantial cost savings. Even with this small analyte volume, each of the amplicons is clearly identified and genotyped. Indeed, the results presented here show that this ET-labeling format coupled with high-performance radial CAE MCP analysis will make the rapid screening of large numbers of individuals at large numbers of polymorphic sites possible. The implementation of MCP-based STR analysis may facilitate many large-scale studies of genetic variation in the human genome and in other genomes of interest, including mice, cattle, parasites, and so forth (24, 40, 41).

We thank Melanie Mahtani, Kristin Pirkola, and David Shen at Molecular Dynamics Inc. (Sunnyvale, CA) for assistance in using the MegaBACE-1000 CAE system and for providing the Genetic Profiler software. We also thank G. Sensabaugh for providing genotyping samples. I. M. is supported by NIH Program Project Grant P01 CA77664 in collaboration with Johns Hopkins University. Additional support was also provided by Amersham Pharmacia Biotech, by NIH Grant HG01399, and the Director, Office of Science, Office of Biological and Environmental Research of the US Department of Energy, under Contract DE FG91ER6125.

References

16. Woolley AT, Mathies, RA. Ultra-high-speed DNA fragment separa-
tion using microfabricated capillary array electrophoresis chips.


