ELSEVIER

Journal of Chromatography A, 781 (1997) 295–305

JOURNAL OF CHROMATOGRAPHY A

Rapid sizing of polymorphic microsatellite markers by capillary array electrophoresis

Elaine S. Mansfield[a,*], Marina Vainer[a], Dennis W. Harris[a], Paolo Gasparini[b], Xavier Estivill[c], Saul Surrey[d], Paolo Fortina[e]

[a]*Molecular Dynamics, Sunnyvale, CA, USA*

[b]*Servizio di Genetica Medica, I.R.C.C.S. Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Foggia, Italy*

[c]*Departament de Genetica Molecular, L'Hospitalet de Llobregat and Genetics Services, Barcelona, Spain*

[d]*Department of Pediatrics, Jefferson Medical College, Philadelphia, PA and the du Pont Hospital for Children, Wilmington, DE, USA*

[e]*Department of Pediatrics, The Children's Hospital of Philadelphia, University of Pennsylvania School of Medicine, Philadelphia, PA, USA*

Abstract

Genetic mapping and DNA sequencing projects could potentially be completed more rapidly by using capillary array electrophoresis (CAE) systems running 48–96 capillaries simultaneously. Currently, multiplex polymerase chain reaction (PCR) and multicolor fluorescent dye-labeling strategies are used to generate DNA profiles containing 18–24 genotypes per sample. By using 4-color fluorescence detection and these multiplex PCR strategies, a CAE system has the capacity to generate up to 5.5 million genotypes per year. CAE offers extremely fast, high-resolution separation of DNA and more automated sample processing than conventional systems because the labor-intensive slab-gel pouring and sample-loading steps are eliminated. We used a prototype CAE system in an ongoing linkage analysis study of inherited deafness in Mediterranean families. CA-repeat markers linked to deafness susceptibility genes on chromosomes 7, 11 and 13 were analyzed and DNA profiles generated which contain 6 markers per color. Fragment sizes of over 28 000 short tandem repeat alleles and 3200 CA-repeat alleles have been determined by CAE. An average sizing precision of ±0.12 base pairs (bp) for fragments up to 350 bp was realized in 1-h runs. In addition, a versatile non-denaturing matrix was used to separate DNA sizing standards, restriction digests, and multiplex PCR samples. Application of this matrix to Duchenne muscular dystrophy exon deletion screening is also described. These CAE approaches should facilitate rapid genotyping of microsatellite markers and subsequent identification of disease-causing mutations. © 1997 Elsevier Science B.V.

Keywords: Capillary array electrophoresis; Gene mapping; DNA typing; Microsatellite typing; DNA

1. Introduction

The Human Genome Initiative has significantly increased the rate at which disease-causing genes are being mapped and sequenced. Short, tandemly repeated and highly polymorphic genetic markers have rapidly found extensive application in mapping of disease-causing genes [1,2]. The most popular of these markers are dinucleotide repeat (microsatellite) markers commonly referred to as CA-repeats [1]. New cost-effective methods using these genetic markers to locate genes and characterize disease-causing mutations require robust, reproducible and accurate measurements of DNA fragment lengths. Capillary array electrophoresis (CAE) offers rapid, high-resolution separations, high-throughput and

*Corresponding author.

0021-9673/97/$17.00
PII S0021-9673(97)00542-6

sensitive detection of fluorescently-labeled DNA fragments.

We previously applied CAE to analysis of simple sequence repeat markers and demonstrated that highly reproducible sizing measurements can be obtained using a denaturing hydroxyethylcellulose (HEC) matrix and a 48-capillary system [3]. Highly accurate sizing results were also obtained using CAE data when DNA reference standards, such as CEPH control DNA [4] are used to normalize genetic typing results. In the analysis of over 8000 genotypes using the CAE prototype system, all observations were within 1 bp of actual when results were normalized to known allele sizes [3]. We therefore used the CAE prototype system in an ongoing, population-based linkage study to characterize the genetics of inherited deafness in Mediterranean families.

Inherited deafness affects approximately 1:1000 live births. Limited clinical differentiation and marked genetic heterogeneity make molecular diagnostic identification very challenging. To date, twelve genetically distinct human candidate genes (DFNB) responsible for autosomal recessive deafness have been mapped, although none have been identified by sequence analysis. The first two deafness genes, DFNB1 and DFNB2, were located by homozygosity mapping within consanguinous Tunisian families to chromosomes 13 and 11, respectively [5,6]. An additional deafness locus, DFNB4, described in Middle Eastern families, was linked to chromosome 7 by Baldwin et al. [7]. The other DFNB genes identified to date were mapped in families from India, Bali, Pakistan, Lebanon, Palestine, and Syria (see Gasparini et al. [8] for historical description). DFNB1, DFNB2 and DFNB4 were chosen as the initial set of candidate genes by Gasparini et al. [8] and Fortina et al. [9] to analyze a group of 48 unrelated Mediterranean families. We compared genotyping results at three DFNB1-linked CA-repeat markers using conventional methods and CAE to test the performance of the new electrophoresis approach. Results profiling CA-repeat markers closely linked to DFNB1, DFNB2 and DFNB4 using capillary array electrophoresis are reported here.

Non-denaturing DNA separations using agarose gels also have extensive utility in molecular genetic diagnosis. For example, Chamberlain et al. [10] and Beggs et al. [11] described multiplex PCR assays which together detect 98% of the deletions leading to the X-linked neuromuscular disorder Duchenne muscular dystrophy (DMD). Traditionally, the PCR products are analyzed by agarose gel electrophoresis and ethidium bromide staining. More recently, Gelfi et al. [12] described a capillary electrophoresis (CE) method to analyze multiplex PCR profiles diagnostic for DMD which contain exon fragments amplified from within the dystrophin gene. However, only a single sample could be analyzed at a time using CE and the UV absorbance detection method used is much less sensitive than laser-induced fluorescence detection of DNA stained with intercalating dyes during electrophoresis.

The increased sensitivity of fluorescence detection simplifies the process of preparing samples for CAE. PCR samples can be diluted prior to electrokinetic injection without further sample manipulation steps. We previously demonstrated a limit of detection of 25 fg DNA using CAE and the highly sensitive DNA-binding dye, Vistra Green (Amersham International, Chicago, IL, USA) [13]. Here, we report using similar separation protocols to detect unlabeled-DNA fragments employing the PCR-based DMD exon scanning assay of Beggs et al. [11]. PCR samples were diluted 1:500, directly injected into the capillaries, and detected using Vistra Green staining. Capillaries used in these experiments were compatible with both denaturing and non-denaturing DNA separations, increasing the utility of CAE systems in genetic research. These results suggest that CAE should facilitate both rapid genotyping of microsatellite markers in genetic linkage analysis and the subsequent detection of disease-causing mutations.

2. Experimental

2.1. Reagents

DNA quantitation standards (200 bp, 500 bp and 1000 bp fragments supplied at 4 ng ml^{-1}) and DNA sizing standards (20 bp ladder (20 bp–1000 bp) as well as a 100 bp ladder (100 bp–2.2 kb)) were purchased from GenSura (Del Mar, CA, USA). Polyethyleneoxide (PEO) ($M_n = 10^5$, $4–10^5$, 10^6, $8–10^6$) was obtained from Union Carbide (Houston,

TX, USA). HEC (M_n=9–10^4, 1.5–10^5) was purchased from Polysciences (Wilmington, DE, USA). Vistra Green was provided by Amersham International (Chicago, IL, USA). Buffers and 9-aminoacridine were purchased from Sigma (St. Louis, MO, USA). PCR reagents and DNA synthesis reagents including fluorescent amidite reagent 6-FAM were purchased from Perkin–Elmer Applied Biosystems Division (Foster City, CA, USA).

2.2. Apparatus

We used a 48-capillary array prototype system previously described by Bashkin et al. [14] to separate fluorescently-labeled DNA fragments. DNA samples were introduced into an array of 48 capillaries from a cathode manifold with the footprint of a half-microtiter plate. Capillaries are threaded through a detection area beneath a microscope objective. The instrument optical design is based on a system described by Huang et al. [15]. Light from a 20 mW Argon-ion laser passes through the microscope objective and a lens sequentially focuses on each capillary in the array as it moves across the capillary holding stage at a scan rate of 1 cycle/s (1 Hz). A dichroic beamsplitter and longpass or bandpass optical filters can be inserted into the fluorescence emission path to enable spectral separation of two detection channels.

Capillaries (200 μm O.D., 75 μm I.D.; Polymicro Technologies, Phoenix, AZ), coated on the external surface with a protective layer of polyimide, were coated on the internal surface with polyacrylamide to prevent electroosmotic flow during electrophoresis using a modified procedure of Cobb et al. [16]. Coated capillaries were cut to 65 cm length and assembled into arrays of 16 capillaries with detection windows burned at 40 cm path length. We used a modified hydroxyethylcellulose (HEC) DNA sequencing matrix of Bashkin et al. [17] to separate CA-repeat markers. Matrix was introduced into the capillary at 2.8 MPa (400 p.s.i) for 2 min, allowed to equilibrate with ambient pressure for 5 min, and pre-electrophoresed for 10 min at 185 v cm^{-1}.

2.3. Subjects and materials

Thirty Italian and eighteen Spanish families with autosomal recessive deafness were included in a population-based study as previously described [8]. Three Italian and two Spanish consanguinous families were included in the group. An example pedigree of one of the consanguinous families included in the study is reported in Fig. 2. There were 382 individuals available for linkage analysis. The 30 Italian and 18 Spanish families included 76 and 45 affected individuals, respectively, leaving 261 unaffected relatives in the study. Informed consent was obtained from all subjects and from parents of patients under legal age. In the initial phase of this study, Gasparini et al. [8] reported identifying linkage to chromosome 13 in 79% of the families. In this report, we describe genotyping 115 individuals in the remaining families at six CA-repeat polymorphisms linked to DFNB2 and DFNB4, on chromosomes 11 and 7 respectively.

2.4. Procedures

Analysis of the DFNB1-linked markers on chromosome 13 was done as previously described [8]. PCR primers were synthesized on a 394 DNA Synthesizer (Perkin–Elmer Applied Biosystems Division) and fluorochrome-labeled primers prepared by chemically attaching the fluorescent amidite reagent 6-FAM to each of the forward primers. Deprotection and purification was done using the recommended protocol for dye-labeled oligonucleotides (Applied Biosystems, 1991).

High-molecular-mass DNA was extracted from peripheral blood leukocytes using previously described protocols [18] or with a nucleic acid extractor (Perkin–Elmer Applied Biosystems Division). DNA concentration was adjusted to 40 ng ml^{-1} and PCR amplification was performed on an automated thermal cycler (Perkin–Elmer, Norwalk, CT, USA) using conditions previously described [19]. Following PCR amplification, samples were multiplexed at the sample loading step. In order to pool the chromosome 7 and 11 CA-repeat markers for multiplexing in the CAE runs, we redesigned the primer sets as outlined in Table 1, used them to amplify each marker, and combined PCR products using reaction volumes as described in Table 2. All samples were combined with a new tetramethyl rhodamine-labeled sizing standard, TMR70-350 (Molecular Dynamics,

Table 1
Redesigned primers for CA-repeat markers linked to DFNB2 and DNFB4

Locus	Size range	Primer sequences
D7S496	104–115 bp	Forward 5′-CATAAAATACATTTTCTATAT-3′
		Reverse 5′-GCTATAACCTCATAANAAACCAAAA-3′
D11S527	128–152 bp	Forward 5′-CAAGTTCTTATCCAGCCTGTCC-3′
		Reverse 5′-GCCCCTCTACTTGTCTGGAG-3′
D11S937	160–184 bp	Forward 5′-CTAATAAACAAATCCCTCTACCT-3′
		Reverse 5′-GCAATAATCATCTCAGTGATTG-3′
D7S501	194–210 bp	Forward 5′-CTAAGACAGGAATTGAATCCTG-3′
		Reverse 5′-ATTTCTTACCAGGCAGACTGCT-3′
D7S523	224–240 bp	Forward 5′-CTGATTCATAGCAGCACTTG-3′
		Reverse 5′-AAAACATTTCCATTACCACTG-3′
D11S911	282–306 bp	Forward 5′-ACAACCTTTCTAACTCCAAAGAC-3′
		Reverse 5′-CCTTCTGAACAATTGCCACATAC-3′

Sunnyvale, CA, USA), specifically formulated for CAE separations [3]. TMR70-350 was formulated to minimize sequence-specific migration anomalies observed in capillary electrophoresis [20] and contains eight compositionally-balanced fragments between 70 and 350 bp in length. Samples were heat denatured at 90°C for 5 min. in a total volume of 4 μl 50% (v/v) formamide prior to injection. DNA was introduced into the capillaries by electrokinetic injection for 15 s at 14 kV, and DNA fragments were separated for 50 min at 185 v cm^{-1} field strength.

Data were analyzed using prototype fragment analysis programs developed at Molecular Dynamics, ArrayQuant and CapillArray and using Excel (Microsoft, Seattle, WA, USA). Typing results determined on the CAE system were compared to those obtained using the Applied Biosystems Model 373A semi-automated fluorescent sequencer (Perkin-Elmer Applied Biosystems Division) using protocols previously described [8]. Pedigrees were entered into computer-readable form using Cyrillic v2.01 (Cherwell Scientific, Oxford, UK). The genotypes and haplotype for each linkage group were checked both manually and using Cyrillic prior to exporting the data for multipoint linkage analysis.

3. Results and discussion

3.1. Genotyping CA-repeat microsatellite markers

The first phase of this population-based study of

Table 2
PCR pooling strategy and sizing results obtained in Phase II deafness linkage study

Marker	Vol. PCR in pool (μl)[a]	Average size difference		Number of differences[b]	Genotype concordance[c]
		CAE	373A		
D7S496	4	−1.57	−0.33	3	98.68
D11S527	2.5	0.55	0.85	4	98.26
D11S937	7.5	−0.22	−0.67	4	98.26
D7S501	5	0.14	0.65	7	96.95
D7S523	4	1.25	0.31	2	99.13
D11S911	4	0.90	0.49	4	98.26

[a] A 2-μl aliquot of the PCR pool was combined with TMR50–530 sizing standard and loading buffer for the CAE analysis.
[b] The number alleles in the dataset where a different call was made using CAE or the reference slab-gel system (N=230 alleles at each locus).
[c] Genotype concordance is expressed as the percentage of alleles called the same on both systems.

the genetics of deafness in Mediterranean families focused on testing linkage to DFNB1 on chromosome 13. In order to better understand the frequency with which DFNB1 lead to deafness in the patient population, linkage analysis in 48 large pedigrees segregating autosomal recessive deafness was determined [8]. The highest linkage observed was for the CA-repeat marker D13S115 with a maximum two-point LOD score of 7.28 within 79% of families [8,9]. A strikingly similar proportion of familes in the two populations demonstrated linkage to DFNB1: 80% of the Italian families and 78% of the Spanish families showed linkage to this locus. This demonstrated for the first time that DFNB1 was an important major locus leading to deafness in Mediterranean populations.

Genotyping of the chromosome 13-linked microsatellite markers was done using a semi-automated slab-gel electrophoresis system (ABI Model 373A, Perkin–Elmer Applied Biosystems Division). To test the feasibility of using CAE for genotyping, all samples at D13S143, D13S115 and D13S175 were reanalyzed using our 48-capillary prototype CAE system in seven runs. Sizing results and called genotypes were then compared for the two systems. As previously reported, we found a reproducible but systematic bias in the estimated sizes at each of the three CA-repeat markers when analyzed by CAE [3]. However, because these migration differences are reproducible and consistent for each locus, the published sizes of known alleles can be used to normalize the CAE size estimates.

We determined the estimated size of the most common allele for each locus on each run to determine a ratio of computed relative to actual size (Table 3). These data were very consistent from one run to the next and the standard deviations averaged ±0.13 bp and precision of all runs were ±0.31 bp or less (data not shown). In the raw size estimates, we found a 1.9% overestimate at D13S143, 2.3% at D13S115 and a 1.1% at D13S175. This corresponds to a deviation from actual of 2.49 bp at D13S143, 3.69 bp at D13S115 and 2.83 bp at D13S175.

CAE sizing results can be normalized to obtain accurate size estimates [3]. The ratio of computed (estimated) to actual size was then used to normalize all of the CAE measurements (N=764 alleles at each locus). After normalizing the CAE results, all sizes were within 1 bp of actual and interpreted genotypes were in agreement with results obtained using a semi-automatic model 373A fluorescent DNA sequencer (Perkin–Elmer Applied Biosystems Division). Further, we found that the process of normalizing estimates with the sizes of known control alleles does not impart any systematic bias in the processed results. In the normalized data, the average deviation

Table 3
Normalizing CA-repeat estimates by CAE improves sizing accuracy

	D13S143	D13S551	D13S175
Most common allele (bp)	130	162	251
CAE size estimates (bp)			
Run 1	132.51	165.67	253.96
Run 2	132.24	165.73	253.93
Run 3	132.42	165.72	253.89
Run 4	132.64	165.65	254.01
Run 5	132.46	165.75	254.01
Run 6	132.17	165.79	254.01
Run 7	132.36	165.53	253.83
Average	132.49	165.69	253.93
Av. difference from actual (bp)	2.49	3.69	2.83
Average/actual (normalizing ratio)	1.019	1.023	1.011
Normalized average estimate	130.02	161.996	251.01
Av. normalized diff. from actual (bp)[a]	0.018	−0.004	0.008

[a] This represents the average difference in size from the actual value after normalizing all alleles tested at the locus (N=785 alleles).

from actual of all alleles in the patient population was 0.018 bp at D13S143, −0.004 bp at D13S115, and 0.008 bp at D13S175.

3.2. Genotyping microsatellites linked to DFNB2 and DFNB4

CA-repeat markers tightly linked to DFNB2 and DFNB4 were used to ascertain the contribution of these deafness-susceptibility loci within families which did not show linkage to DFNB1. The closest CA-repeat markers to DFNB4 were D7S496, D7S501, and D7S523 on chromosome 7 while D11S527, D11S937 and D11S911 were the closest markers to DFNB2 on chromosome 11. Unfortunately, published primer sequences for these six markers would yield overlapping PCR product sizes. Primers for the six loci were therefore redesigned using published genomic sequences (Table 1) to minimize the number of runs required and to permit multiplexing of the markers into single, 6-marker profiles for each individual.

Primer sets listed in Table 1 were synthesized including the fluorescent dye 6-FAM on each of the forward primers. These were used to amplify each of the six CA-repeat markers from 115 remaining family members in the study set. Following PCR amplification, the average signal intensity at each locus was determined by analyzing a minimum of five samples at each locus. Relative signal intensity was then used to devise a PCR-pooling strategy. Aliquots of the six CA-repeat loci were combined as indicated in Table 2. Residual PCR buffer in the pooled samples were removed by float dialysis for 15 min as previously described [21], and an aliquot of the pool (2 μl) was combined with sizing standard. The resulting mixture of PCR samples yielded peaks sufficiently close in intensity at each of the loci to generate clear, multi-locus CA-repeat profiles for genotyping. The profiles from six family members (Fig. 1) are each unique due to the high heterozygosities of the microsatellite markers.

Although the synthesized PCR primers were specifically designed to keep the alleles at each locus separated in length by a minimum of 8 bp, new alleles at D11S527 and D11S937 were observed. Consequently, the observed difference between the longest allele at D11S527 and the shortest at D11S937 was only 3 bp (see the second and third loci in the top profile, Fig. 1). Mendelian inheritance and haplotype segregation in all father, mother and child groups were verified using Cyrillic 2.0 (Cherwell Software, Oxford, UK) prior to exporting the genotypes to MLINK or other linkage analysis programs. Approximately 5% of the loci were reanalyzed using dilutions of the original PCR samples when peaks were offscale or signal intensities were less than 100 relative fluorescence units (rfu).

3.3. Data analysis

Entry of family data and sample information, Mendelian inheritance verification, and haplotype checking is most easily performed using a graphical pedigree-editing program. Several pedigree editors have been written for the WindowsNT operating system (Microsoft) used by the CAE system. We have found it convenient to enter family data using Cyrillic (Cherwell Scientific, Oxford, UK). An example pedigree from one of the consanguinous families in our data set is illustrated in Fig. 2. Haplotypes of chromosome 11-linked markers are shown for informative family members in this pedigree. Identity by descent, resulting from the first-cousin marriage within the family, at all three chromosome 11 markers is clearly evident in the proband.

Genotypes at each locus were compared after independently reviewing the sizing data from each system. An excellent correlation of called genotypes was obtained when comparing CAE results with those from the reference slab-gel system (ABI 373A). Different internal sizing standards were used in the samples; TMR70-350 which was specifically developed for CE separations [3,21] was used for CAE and GeneScan350 (Perkin-Elmer, Foster City, CA) was used for the slab-gel system. For the CAE data, we computed a normalized size based on the method described in section 3.1. Both CAE and 373A data show the characteristic biases in calculated sizing results, although this effect is generally more subtle in the 373A data (see Table 2). Whereas the maximum difference between estimated and actual size by CAE was −1.57 bp, we observed a maximum difference of +0.85 bp in the 373A data.

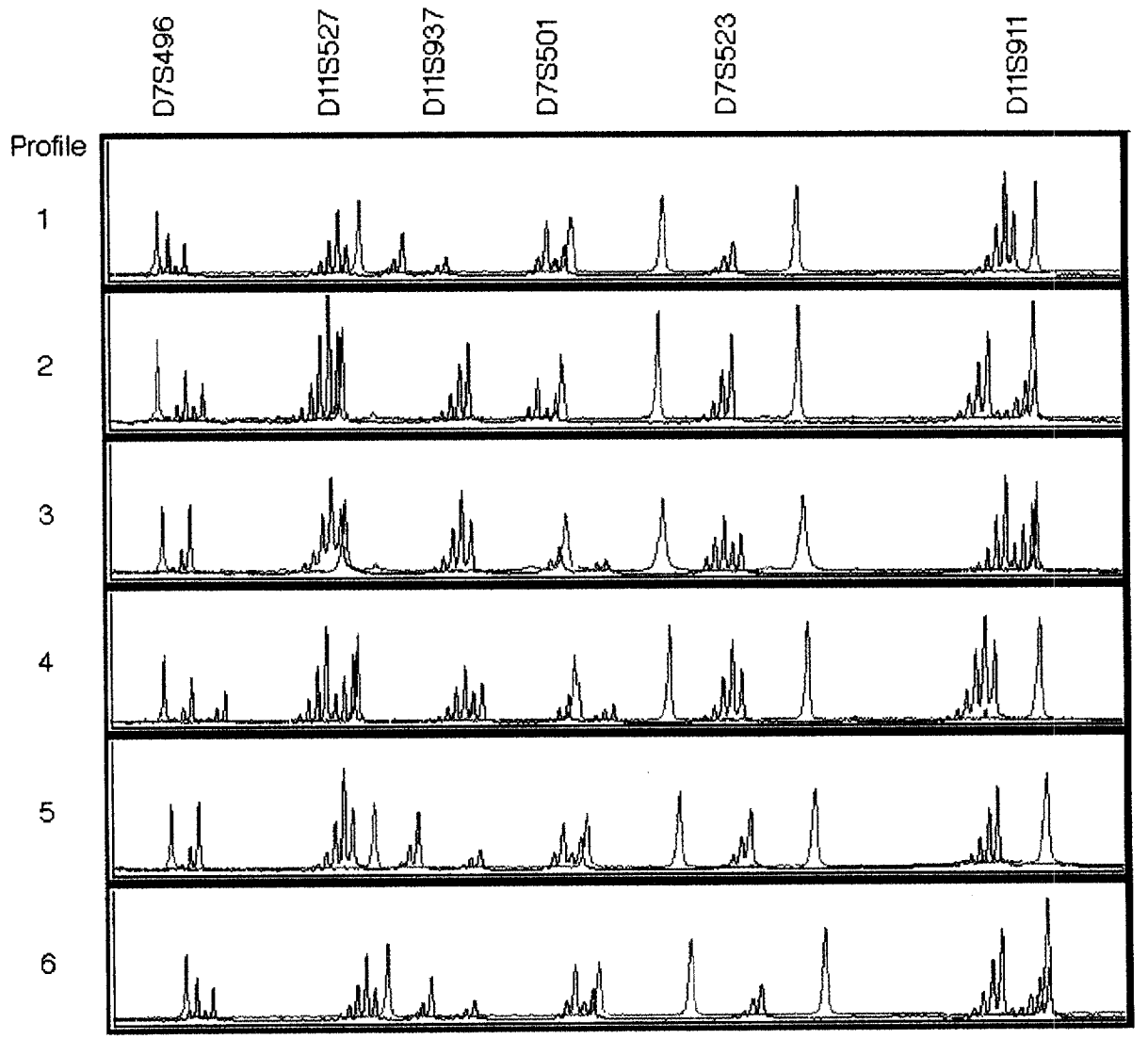

Fig. 1. CA-repeat profiles determined by CAE. Six CA-repeat polymorphic microsatellites were PCR amplified for 30 cycles, and then pooled and desalted by float dialysis. A 1.0-μl aliquot of the PCR pool was combined with 0.5 μl of TMR70-350. Following heat denaturation in 50% (v/v) formamide, the DNA was electrophoresed by CAE as described in section Section 2.4. The FAM-labeled PCR products (blue) were sized relative to the migration time of fragments in the TMR-labeled sizing standard (red). The microsatellite profiles of six family members are illustrated.

On each instrument system, the size of alleles at D7S496 and D11S937 were underestimated while alleles at the other four loci were overestimated. The magnitude of deviation from actual was greater at three of the loci using CAE system and at the other three loci using the slab-gel system.

A very low number of allele call differences was observed comparing data generated from CAE and the slab-gel system after verifying Mendelian inheritance and haplotype consistencies within the families. For example 2 differences at D7S523, 3 at D7S496, 4 at D11S527, D11S937 and D11S911, and 7 at D7S501 (N=230 alleles at each locus) were observed. Overall, this indicated an average of

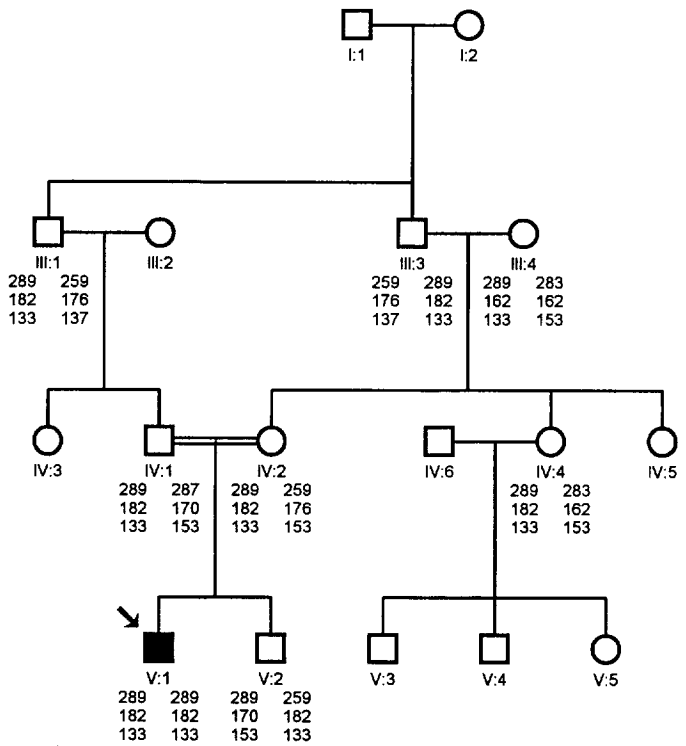

Fig. 2. Pedigree of a family with autosomal recessive deafness. The Cyrillic 2.1 pedigree drawing program (Cherwell) was used to enter pedigree S56 in a linkage study of Mediterranean families segregating non-syndronic autosomal recessive deafness. The haplotypes of informative family members for chromosome 11-linked CA-repeat markers are indicated. Homozygosity at each of the three markers in the proband (V:1; indicated by the '↓') is indicative of identity by descent at DFNB2, resulting from the consanguinous union of the affected child's parents.

>98% concordance in genotype results comparing the two separation methods. Thus, fewer than 2% of the samples would have to be reprocessed to reconcile the data despite the fact that different instruments, different gels and different sizing standards were used in the analysis. A major source of typing discrepancy could be attributed to sample processing errors or misinterpretation of peaks with low signal intensities. Ten genotypes in the reprocessed data were in a agreement with the original 373A calls, eight were in agreement with CAE calls and six yielded ambiguous profiles. The ambiguous profiles required reamplification and analysis of fresh PCR product to make an accurate genotype call.

Analysis of genotypes revealed 7 families with linkage to DFNB4 (chromosome 7) and two families with linkage to DFNB2 (chromosome 11). The 2-point LOD score for the chromosome 7 markers approached statistical significance and ranged from 2.8 to 3.2. The small number of informative families in this phase of the study contributed to the relatively low scores and LOD scores could increase by extending the study to include additional families or additional informative members in the identified families.

3.4. Duchenne muscular dystrophy screening

Non-denaturing agarose gel separations are frequently used in fragment analysis applications which do not require the high level of resolution needed for CA-repeat genotyping. Previously, we demonstrated that fragments between 80 bp and 40 kb could be completely resolved using a non-denaturing poly-

ethyleneoxide (PEO) matrix [13] and a mixture of four sizing ladders (20 bp, 100 bp, 1 kb and 5 kb). Here, we report using this matrix to analyze the nine exon screening panel of Beggs et al. [11] in multiplex PCR samples (Fig. 3). This assay is used to identify deletions in the dystrophin gene that lead to Duchenne (DMD) or Becker (BMD) muscular dystrophies. PCR reactions were performed using the procedure of Beggs et al. [11] and diluted 1:500 in distilled water prior to analysis by CAE.

Flanking standards of 100 bp and 700 bp were included in the CAE profiles to serve as both sizing

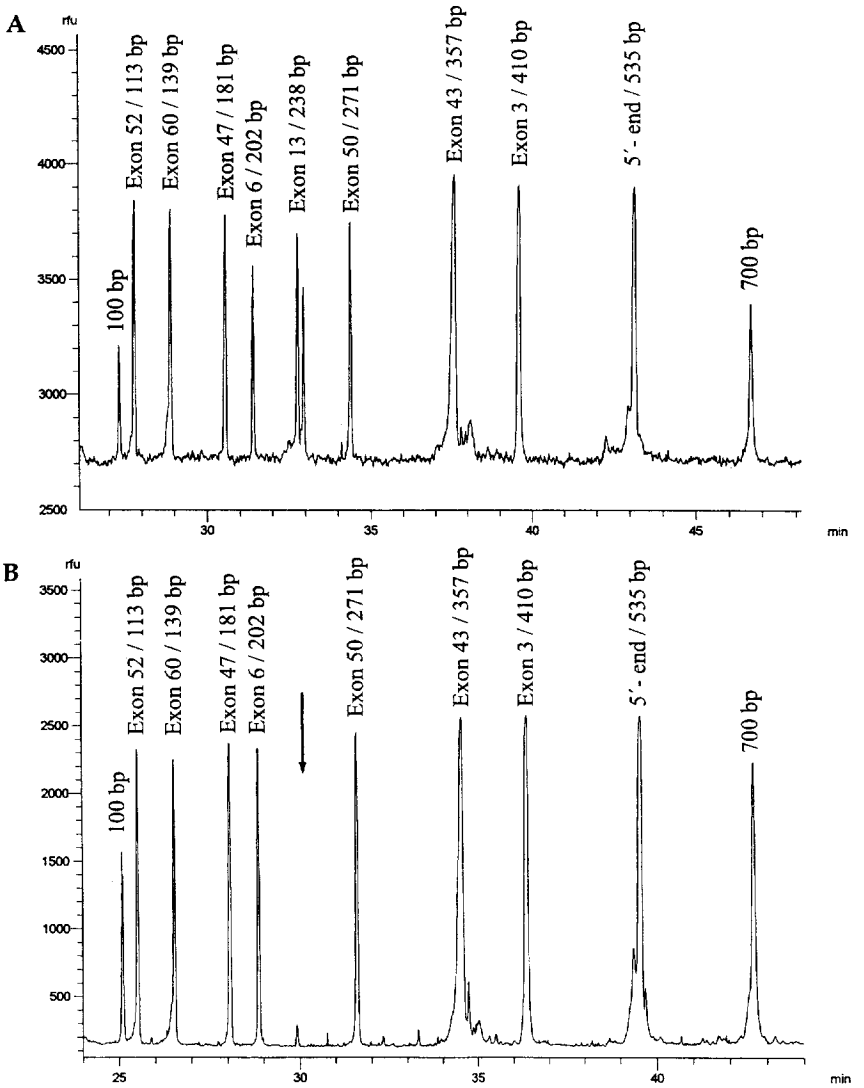

Fig. 3. Multiplex PCR profile of exons in DMD/BMD gene. A 1:500 dilution of PCR product amplified using the multiplex PCR assay of Beggs et al. [11] was combined with 0.5 pg of two flanking DNA standards (100 bp, 700 bp; BioVenture, Murfreesboro, TN, USA). DNA was electrophoresed in a non-denaturing matrix containing 0.5% (w/v) PEO ($M_n = 10^6$), 1X TBE, and 10 μM of 9-aminoacridine (Sigma). Conditions used were as follows: separation voltage: 108 v cm^{-1}; detection: Vistra Green (2 nM) (Amersham International) added to the matrix; excitation: 20 mW 488 nm argon-ion laser, 550 V PMT; and, emissions collected using a 530DF30 filter.

and sample injection controls. A normal female control profile (Fig. 3A) contains PCR products amplified from nine deletion-prone exons. The profile of a male affected with DMD is missing the 238 bp fragment from exon 13 (Fig. 3B). CAE separations are complete in under 50 min, less than half the time required using conventional procedures [11]. In addition, the CAE procedure using the new highly sensitive DNA stain, Vistra Green, offers higher resolution and up to 500 times greater sensitivity than ethidium bromide staining of DNA in agarose gels [13].

3.5. Conclusions

We demonstrated the utility of a 48-capillary CAE prototype system in two phases of a population-based linkage analysis project studying the genetics of deafness. In Phase I, we found that CAE genotyping results correlated with a slab-gel system (ABI 373A) using CA-repeat markers closely linked to the DFNB1 locus on chromosome 13. During Phase II, genotyping of 115 remaining family members were independently conducted using data generated on the two instrument systems. Over 98% of the called alleles agreed comparing the two systems although slight variation in absolute size measurements was observed in data collected by either instrument. In the CAE study, 5% of the processed samples were reanalyzed due to offscale or low-intensity peaks, 1% to resolve haplotyping inconsistencies, and 2% due to sample mix-ups, user differences in profile interpretation. These modest rates of sample reprocessing might be further improved by the use of robotic liquid handling systems for sample processing.

The CA-repeat profiles generated in Phase II contained six loci per run multiplexed by size. Multicolor fluorescent analysis of CA-repeat profiles would permit an even higher throughput of genotyping because up to three different profiles can be co-electrophoresed with the sizing standard. Using a 4-color multiplex approach, up to 18 genotypes could be analyzed in each sample per 50 min run. With continuous operation, 1920 samples could be processed per day by CAE. By using the 4-color multiplex strategy that has been applied to slab-gel electrophoresis systems, 34 560 genotypes could be determined per day with the system (1920 samples$\times$3 colors of labels$\times$6 loci/color). Due to the time and labor required for gel preparation and loading, and the longer electrophoresis time on current slab-gel systems, a maximum of 8 runs can be completed within 24 h (2 h runs, 1 h pre-electrophoresis and sample loading). This translates to processing up to 5184 genotypes per day (288 samples$\times$3 colors of labels$\times$6 loci/color). Thus, CAE offers approximately 7-times higher sample throughput than conventional slab-gel analysis.

In addition to higher throughput for genetic analysis, CAE offers several key advantages over current systems. The labor-intensive steps of fragment separations are eliminated by CAE. Slab-gel preparation is replaced by automated capillary filling using a low-viscosity gel matrix, refillable capillaries and a pressurized gel-filling station [14]. The process is complete in under 10 min. Gel loading, another tedious, labor-intensive step, is replaced with electrokinetic injection of up to 96 samples at a time into CAE instrument systems. Capillaries are then purged and refilled using the matrix-filling station. Thus, the increase in sample throughput is not limited to the electrophoresis time, but all steps of gel preparation and sample loading. In addition, both denaturing and non-denaturing separation conditions have now been reliably demonstrated using CAE. The CAE approach should facilitate rapid genotyping of microsatellite markers and the subsequent identification of disease-causing mutations.

Acknowledgments

This work was supported in part by Molecular Dynamics, a small business research innovation (SBIR) grant #1R43NS/MH34589-01 awarded by the N.I.H. to E.S.M. and N.I.H. grants DK16691 and P60HL38632. Additional support was provided by an advanced technology program grant #70NANB5H1031 awarded by the N.I.S.T to Molecular Dynamics (David L. Barker, Principal Investigator) and by grants from the Foerderer Fund for Excellence from the Children's Hospital of Philadelphia to P.F. and the Telethon to P.G.

References

[1] J.L. Weber, P.E. May, Am. J. Hum. Genet. 44 (1989) 388.
[2] C.M. Hearne, S. Ghosh, J.A. Todd, Trends Genet. 8 (1992) 288.
[3] E.S. Mansfield, M. Vainer, S. Enad, D.L. Barker, D. Harris, E. Rappaport, P. Fortina, Genome Res. 6 (1996) 893.
[4] C. Dibs, S. Faure, C. Fizames, D. Samson, A. Drout, A. Vignal, P. Millasseau, S. Marc, J. Hazan, E. Seboun, M. Lanthrop, G. Gyapay, J. Morissette, J. Weissenbach, Nature 380 (1996) A4.
[5] P. Guilford, S.B. Arab, S. Blanchard, J. Levilliers, J. Weissenbach, A. Belkahia, C. Petit, Nature Genet. 6 (1994) 24.
[6] P. Guilford, H. Ayadi, S. Blanchard, H. Chaib, D. Le Paslier, J. Weissenbach, M. Drira, et al., Hum, Mol. Genet. 3 (1994) 989.
[7] T. Baldwin, L.A. Farrer, S. Weiss, A.L. De Stefano, R. Adair, B. Franklyn, K.K. Kidd, et al., Hum, Mol. Genet. 4 (1995) 1637.
[8] P. Gasparini, X. Estivill, V. Volpini, A. Totaro, S. Castellvi-Bel, N. Govea, M. Mila, M. Della Monica, V. Ventruto, M. De Benedetto, P. Stanziale, L. Zelante, E.S. Mansfield, S. Surrey, P. Fortina, Eur. J. Hum. Genet. (1997), in press.
[9] P. Fortina, S. Surrey, X. Estivill, V. Volpini, A. Totaro, S. Melchiondia, S. Castellvi-Bel, N. Govea, M. Mila, M. Della Monica, V. Ventruto, M. De Benedetto, P. Stanziale, L. Zelante, E.S. Mansfield, P. Gasparini, Am. J. Hum. Genet. 59 (1996) A218.
[10] J.S. Chamberlain, J.E. Rainer, P.N. Nguyen, C.T. Caskey, Nucleic Acids Res. 16 (1989) 11141.
[11] A.H. Beggs, M. Koenig, F.M. Boyce, L.M. Kunkel, Hum. Genet. 86 (1990) 45.
[12] C. Gelfi, A. Orsi, F. Feocini, P.G. Righetti, I. Spiega, P. Carrera, M. Ferrari, Biotechniques 19 (1995) 254.
[13] R.M. Madabhushi, M. Vainer, V. Dolnik, S. Enad, D.L. Barker, D. Harris, E.S. Mansfield, Electrophoresis 18 (1997) 104.
[14] J. Bashkin, D. Roach, J. Leong, D. Barker, R. Johnston, J. Cap. Electrophoresis 3 (1996) 60.
[15] X.C. Huang, M.A. Quesada, R.A. Mathies, Anal. Chem. 64 (1992) 2149.
[16] K.A. Cobb, V. Dolnik, M. Novotny, Anal. Chem. 62 (1990) 2478.
[17] J. Bashkin, M. Marsh, D. Barker, R. Johnston, Appl. Theor. Electrophoresis 6 (1996) 23.
[18] M. Poncz, D. Solowiejzcyk, B. Harpel, Y. Mory, E. Schwartz, S. Surrey, Hemoglobin 6 (1982) 27.
[19] T.J. Hudson, M. Engelstein, M.K. Lee, E.C. Ho, M.J. Rubenfield, C.P. Adams et al., Genomics 13 (1992) 622.
[20] M. Wenz, Nucleic Acids Res. 22 (1994) 4002.
[21] M. Vainer, S. Enad, V. Dolnik, D. Xu, J. Bashkin, M. Marsh, O. Tu, D.W. Harris, D.L. Barker, E.S. Mansfield, Genomics 41 (1997) 1.