Research Report

Melting Curve Analysis of SNPs (McSNP®): A Gel-Free and Inexpensive Approach for SNP Genotyping

BioTechniques 30:358-367 (February 2001)

J.M. Akey, D. Sosnoski[1], E. Parra[1], S. Dios[1], K. Hiester[1], B. Su, C. Bonilla[1], L. Jin, and M.D. Shriver[1]
University of Texas Houston Health Science Center, Houston, TX and [1]Penn State University, University Park, PA, USA

ABSTRACT

High-throughput methods for assaying DNA variation require two important steps: (i) *discriminating the variation and* (ii) *detecting the signal. In this report, we describe a novel SNP genotyping method that we refer to as melting curve analysis of SNPs (McSNP®). McSNP combines a classic approach for discriminating alleles, restriction enzyme digestion, with a more recent method for detecting DNA fragments, melting curve analysis. Melting curve analysis is performed by slowly heating DNA fragments in the presence of the dsDNA-specific fluorescent dye SYBR® Green I. As the sample is heated, fluorescence rapidly decreases when the melting temperature of a particular fragment is reached. We show that it is possible to determine the composition of simple mixtures of DNA fragments, such as those that result from restriction enzyme digestions of short PCR products. McSNP is well suited for high-throughput genotyping because 96 samples can be analyzed and automatically scored in 20 min. Our results clearly demonstrate that McSNP is a simple, inexpensive, and accurate means of genotyping SNP variation.*

INTRODUCTION

The third-generation genetic map comprised of single nucleotide polymorphisms (SNPs) is being expeditiously developed (24). SNPs are the most abundant type of sequence variation in the human genome and will be useful tools in many diverse applications that include disease gene mapping, evolution, pharmacogenetics, and forensics (3,11). An impressive SNP resource already exists as nearly 300 000 have been deposited into publicly accessible databases (see http://www.ncbi.nlm.nih.gov/SNP/, http://snp.cshl.org/, and http://hgbase.cgr.ki.se/). However, without parallel progress in SNP genotyping technology, their true power and inherent benefits will not come to fruition.

Novel genotyping methods amenable to high-throughput analysis should ideally be gel-free, robust, inexpensive, and simple to perform. To this end, these requirements have inspired the development of a variety of genotyping assays, including the oligonucleotide ligation assay (OLA) (12), genetic bit analysis (GBA) (15), mass spectroscopy (8), "chip" technology (24), TaqMan® (14), and dynamic allele specific hybridization (DASH) (10). Although many SNP genotyping methods have been developed, no single technology has emerged as being clearly superior because of limitations such as cost, complexity, and accuracy.

Recently, the principles underlying kinetic PCR (9), which is predicated on monitoring the fluorescence of a diagnostic probe once per PCR cycle (25), have been applied to SNP genotyping methods (6,10,14). The dsDNA-specific dye SYBR® Green I (Moleclar Probes, Eugene, OR, USA) has been used to analyze the melting curves of PCR products, which are characterized by a rapid loss of fluorescence as the temperature is raised through the samples' melting temperature (T_m) (17). Since the T_m is a function of product length, sequence composition, and GC content, it should be possible to distinguish DNA fragments that differ with respect to these parameters by melting curve analysis (MCA). The general principles of MCA were used to develop T_m-shift genotyping, an SNP genotyping assay based on allele-specific PCR (6). While this represents an important advance in MCA applied to SNP genotyping, it is subject to the inherent limitations of allele-specific PCR, such as difficulty in reaction optimization.

In the burgeoning field of SNP genotyping technology, it is easy to forget the elegantly simple and accurate approach of assaying SNP variation by digesting PCR products with restriction enzymes [i.e., PCR-restriction fragment length polymorphism (RFLP) typing]. In this report, we describe how MCA can be applied to classical PCR-RFLP methods. Specifically, a simple, inexpensive, and high-throughput assay for SNP genotyping that we refer to as MCA of SNPs (McSNP®) is presented in which the melting profiles of restriction enzyme-digested PCR products are analyzed. In total, we present data on seven

SNPs, clearly demonstrating that this method is a practical, robust, and sensitive approach to assaying SNP variation.

MATERIALS AND METHODS

PCR

SNPs were amplified by PCR, and most reactions were conducted under a standard set of conditions. PCR was performed in either a Hybaid MultiBlock System (Hybaid, Franklin, MA, USA) or a GeneAmp® 9700 (Applied Biosystems, Foster City, CA, USA) and consisted of the following thermocycles: 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, and a final extension at 72°C for 5 min. PCRs were performed in 25-µL reactions consisting of 25 ng genomic DNA, 1.5 mM $MgCl_2$, 2.5 µL 10× PCR buffer (Life Technologies, Rockville, MD, USA), 0.1 mM dNTPs, 0.04 µM each primer, and 2.5 U *Taq* DNA polymerase (Life Technologies). The exception to these conditions was marker Fy-null, in which the final $MgCl_2$ was 2.0 mM. Primer sequences were as follows: CKMM (5′-GCAGGCGCCTACTTCTGG-3′ and 5′-AGCTCATGGTGGAAATGGAG-3′), 4019 (5′-CAGGCCAAGAGCGTCCTA-3′ and 5′-TGCCACTCTGTGAACAGCAA-3’), 14319 (5′-CATCTGAGTGCAAGATAAAAAGGA-3′ and 5′-CCCACCCCCAAATCATCTAT-3′), FY-null (5′-GCCCAGAACCTGATGGCCCTCATTAGTGCT-3′ and 5′-CTGTCAGCGCCTGTGCTT-3′), 14867(5′-GGCAGGACATTCCAAGGCTCTC-3′ and 5′-CACCCTGGGTTAACACATTCA-3′), LPL (5′-TGCAAGGGTTTTGCTTAATTCT-3′ and 5′-CAACAACAAAACCCCACAGC-3′), and M89 (5′-ACAGAAGGATGCTGCTCAGCTT-3′ and 5′-GCAACTCAGGCAAAGTGAGA**C**AT-3′), in which the bold denotes the mismatched base to create the putative *Nla*III restriction site (see Results section). SNP markers WI-4019 (dbSNP ID: 2420), WI-14319, and WI-14867 were identified by querying the Whitehead Institute Genome Center SNP database (http://www-genome.wi.mit.edu/). SNP markers LPL, CKMM (dbSNP ID: 5318), and FY-null were obtained from References 7, 5, and 22, respectively. Marker M89 was obtained from Reference 21. The samples used in this study represent a subset of samples that we are using in other studies and are comprised of African-American, European-American, and Hispanic populations. More specific details can be found elsewhere (16,20). Informed consent from each subject was obtained for participation in this study.

Restriction Enzyme Digestion

Restriction enzyme digestions were performed in a final volume of 25 µL, which consisted of 10 µL PCR product, 1 U appropriate restriction enzyme, 1× reaction buffer supplied with the restriction enzyme, and, when recommended by the supplier, 1× bovine serum albumin (BSA). In particular, markers 4019, M89, and 14867 were digested by the restriction enzyme *Nla*III, while markers CKMM, 14319, FY-null, and LPL were digested by the restriction enzymes *Taq*I, *Rsa*I, *Sty*I, and *Pvu*II, respectively. We did not find it necessary to purify the PCR products before restriction enzyme digestions, which were performed as recommended by the suppliers for a period of 4–24 h. All restriction enzymes were purchased from New England Biolabs (Beverly, MA, USA).

MCA

Melting profile reactions were analyzed in a final volume of 50 µL. All reactions contained 5 µL restriction enzyme product, 5 µL 10× SYBR Green I (which comes as a 10 000× stock supplied by Molecular Probes), and were brought to a final volume of 50 µL in double-distilled water. Various destabilizing agents were added to these reactions including formamide, dimethyl sulfoxide (DMSO), and urea (all from Fisher Biotech, Fair Lawn, NJ, USA). Here, we provide the details of only those additives that had an enhancing effect on MCA. The optimal final concentrations of these additives were either 10% formamide or 10% DMSO. All reactions were performed in polycarbonate 96-well plates (Hybaid). Real-time fluorescent monitoring was performed with the DASH machine (Hybaid). Other instruments that may be suitable for MCA include the GeneAmp® 5700 and ABI Prism® 7700 Sequence Detection Systems (Applied Biosystems), the LightCycler® (Roche Molecular Biochemicals, Indianapolis, IN, USA), the iCycler™ (Bio-Rad Laboratories, Hercules, CA, USA), the Smart Cycler® (Cephied, Sunnyvale, CA, USA), and the Sentinel™ (Stratagene, La Jolla, CA, USA). However, we have not explored the general accuracy, robustness, and scoring software associated with these machines, which may differ substantially. Melting curves are acquired by ramping the temperature from 35°C to 90°C at a rate of 0.04°C/s and monitoring the change in fluorescence of SYBR Green I at 520 nm. The total run time is approximately 22 min, which can be generally cut in half once the T_m is determined and the thermal range adjusted accordingly. It is during the slow ramp to 90°C that the fluorescence is captured. The raw data are first converted to relative fluorescence by dividing each point by the initial fluorescence (i.e., 35°C) (6). Final melting curves are reported as the three-point smoothed negative first derivative of fluorescence, with respect to temperature versus temperature with a baseline subtraction. The baseline correction for each data point is calculated by subtracting the slope from a linear regression line of the first nine data points. All genotypes were confirmed by agarose gel electrophoresis. Specifically, samples were amplified by a second set of PCR primers to produce larger products more suitable for scoring by agarose gel electrophoresis. These samples were then independently digested and scored by agarose gels.

RESULTS

Restriction Enzyme MCA Profiles

The experimental protocol of McSNP is very simple and is outlined in Figure 1. Figure 2 illustrates typical McSNP genotype profiles for six autosomal SNP loci. The peak of a melting curve denotes the T_m of a particular DNA fragment, which is defined as the temperature at which half of the fragments in solution have denatured. As expected,

the observed T_m values of undigested SNP alleles are higher compared to the digested alleles. Note the clear distinction between homozygous and heterozygous individuals, in which heterozygous melting curves are composites of the two homozygous melting curves.

More specifically, Figure 2, panel A, consists of homozygous undigested genotypes, and, thus, only one T_m peak is observed. Figure 2, panel B, contains examples of heterozygous genotypes in which one allele is digested and the other is undigested. In general, we detected two distinct types of heterozygous melting curves, in which heterozygotes formed either two (i.e., marker CKMM) or three (i.e., marker 4019, FY-null, and 14319) resolvable peaks. In three-peak heterozygotes, one high-temperature peak corresponds to the undigested allele, while the other two lower-temperature peaks result from the two fragments produced by endonuclease cleavage of the digested allele (Figure 1). Conversely, in two-peak heterozygotes, the two resulting DNA fragments from endonuclease cleavage of the digested allele have similar T_m values and melt simultaneously. Thus, the difference between two- and three-peak heterozygotes is whether the digested allele gives rise to DNA fragments with the same or different T_m. Note that LPL and 14867 are actually three-peak heterozygotes, but, for presentation purposes, the lowest T_m peak was not included in the figure (Table 1). Finally, Figure 2, panel C, demonstrates homozygous digested genotypes. Two possible patterns of homozygous digested melting curves are possible, depending on whether the digested allele results in two fragments with the same or different T_m.

Engineering a Restriction Site

In many instances, an SNP will not cause a restriction enzyme site difference. In these cases, a digestion site for one of the SNP alleles can be introduced by designing one of the PCR primers with a mismatch near the 3′ end (13). For instance, the Y-chromosome SNP, M89, was genotyped with McSNP by using a primer with a mismatched C 3 bp upstream of the 3′ end, thus creating a *Nla*III restriction site in one of the alleles. Individuals with the M89*T allele will be undigested and produce an 87-bp fragment, whereas those with the M89*C allele will be digested resulting in fragments of 20 and 67 bp, as seen in Table 1. Although a mismatched primer theoretically reduces the efficiency of PCR, these primers are not difficult to work with in practice and often do not require optimizations over that of a standard PCR. For more specific details, see References 1, 4, and 13.

Major Parameters Affecting McSNP Profiles

Critical factors in McSNP include the size and sequence of PCR products as well as additives, which destabilize the DNA duplex and lower the T_m. In particular, there is a greater difference in the T_m between smaller DNA fragments than there is for larger DNA fragments. Theoretical and empirical studies of DNA denaturation have shown that, as the size of DNA fragments increases, the difference in T_m between them decreases. Our laboratory has had success using fragments in the range of 50–150 bp.

Moreover, it is necessary to add a destabilizing agent to ensure that the T_m of a fragment is reached during the experiment (i.e., T_m<100°C). We have tested formamide, DMSO, and urea for their effects on melting curve profiles. The overall effects of these additives are similar in that they destabilize dsDNA to varying degrees and thus lower the T_m. However, they have different influences on the shape of melting curves. For example, the addition of urea resulted in broadening of peaks and decreasing the fluorescence of the sample (data not shown). The effects of DMSO and formamide were found to be superior because they resulted in more sharply defined peaks. Thus, DMSO and formamide have relatively equivalent effects on melting curve characteristics. Although we chose to use formamide, one could also use DMSO to lower the T_m.

T_m Estimation

Theoretical prediction of the T_m of a DNA molecule is complex and has been the subject of numerous studies (reviewed in Reference 23). We have investigated the empirical relationship between the observed and predicted T_m because it is important to have some *a priori* information regarding the expected resolution of digested and undigested DNA fragments. The two methods used for predicting T_m were a simple salt-adjusted formula (http://www.biotechlab.nwu.edu/OligoCalc.html) (18,19) and a nearest-neighbor algorithm (http://alces.med.umn.edu/rawtm.html) (2).

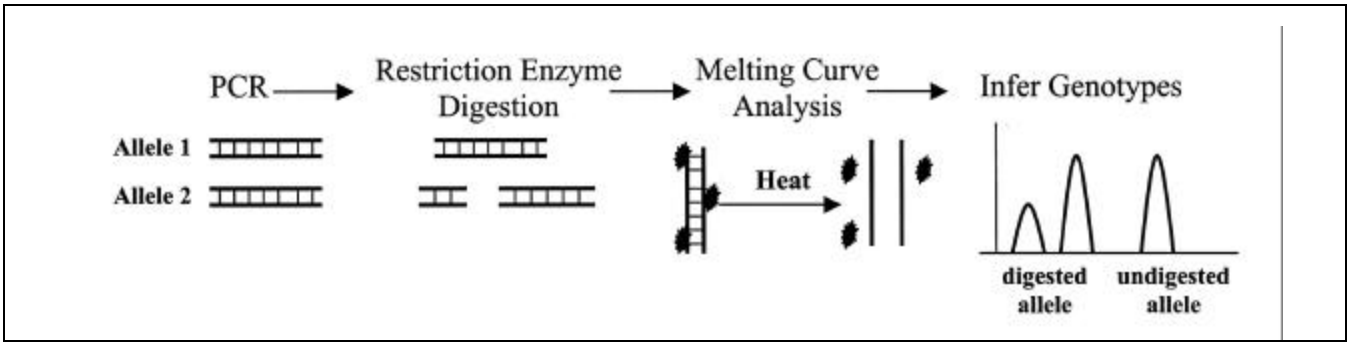

Figure 1. The experimental protocol of McSNP. The SNP is amplified by PCR and then subjected to restriction enzyme digestion. The identity of each DNA fragment is determined by the unique pattern of melting curve peaks acquired by analyzing the change in fluorescence of the dsDNA-specific dye SYBR Green I as the sample is heated. In this example, the individual is scored as a heterozygote having one undigested and one digested allele (alleles 1 and 2, respectively).

Table 1. Comparison of Observed and Predicted T_m

Locus	Fragment sizes (bp)[a]	Observed T_m(°C)	Predicted T_m (°C)[b]	Observed |1-2a| (°C)[c]	Predicted |1-2a| (°C)	Observed |1-2b| (°C)[d]	Predicted |1-2b| (°C)	Observed |2a-2b| (°C)[e]	Predicted |2a-2b| (°C)
CKMM	1: 82	1: 82.0	1: 84.3						
	2a: 48	2a: 74.1	2a: 73.5	7.9	10.7	7.9	10	0	0.2
	2b: 34	2b: 74.1	2b: 73.7						
4019	1: 62	1: 77.4	1: 72.2						
	2a: 37	2a: 67.4	2a: 63.8	10	13.3	15	15	5	4.1
	2b: 25	2b: 62.4	2b: 59.7						
FY-null	1: 61	1: 80	1: 78.4						
	2a: 38	2a: 75.5	2a: 68.1	6	10.1	10	14	7.9	11.1
	2b: 23	2b: 68.2	2b: 57.2						
14319	1: 84	1: 76.1	1: 76.5						
	2a: 52	2a: 73.4	2a: 68.1	4.7	8.4	11.8	21	5.3	6.1
	2b: 32	2b: 66.1	2b: 73.6						
14867	1: 81	1: 81.4	1: 80.4						
	2a: 52	2a: 77.7	2a: 74.6	3.7	5.8	15.7	19	12	5.8
	2b: 29	2b: 65.7	2b: 57.1						
LPL	1: 70	1: 74.4	1: 71.0						
	2a: 51	2a: 70.4	2a: 65.5	4	5.5	18.3	28	14.3	22
	2b: 19	2b: 56.1	2b: 43.5						
M89	1: 77	1: 75.5	1: 75.3						
	2a: 57	2a: 72.1	2a: 70.9	3.4	4.4	29.4	29.9	26	25.5
	2b: 20	2b: 45.4	2b: 46.1						

[a]Fragment size refers to the size of DNA fragments that result from restriction enzyme digestion. Fragments labeled 1 denote the fragment that results for undigested alleles, while fragments 2a and 2b are those that result for digested alleles

[b]Predicted T_m is based on a salt adjusted nearest-neighbor algorithm described in the text (2).

[c,d,e]Denotes the absolute value of the difference in observed T_m for fragments 1 and 2a, 1 and 2b, and 2a and 2b, respectively.

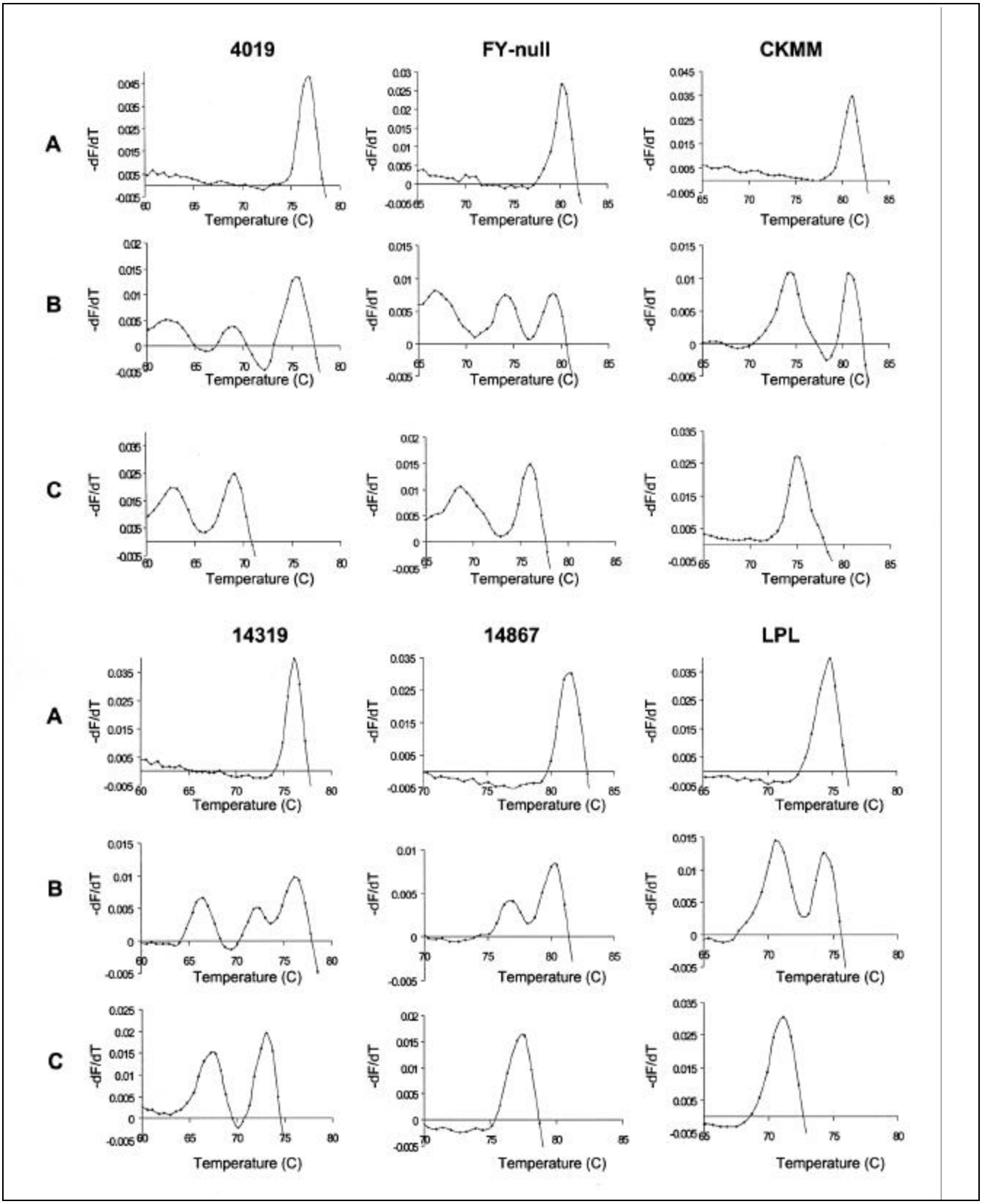

Figure 2. Typical McSNP profiles of the six autosomal SNPs investigated (4019, FY-null, CKMM, 14319, 14867, and LPL). Panels A, B, and C denote homozygous undigested, heterozygous, and homozygous digested genotypes, respectively. Note the clear resolution of homozygous and heterozygous genotypes. No template negative controls showed below background levels of fluorescence and are therefore not included in the figure.

Table 1 presents the average observed and predicted T_m for all DNA fragments of the seven SNPs investigated based on the nearest-neighbor algorithm. Similar results were obtained for the alternative formula described above (data not shown). The observed T_m is the average over one or two 96-well sample plates. There is a high concordance between the observed and predicted T_m levels for the DNA fragments ($r^2 = 0.94$) shown in Table 1 (Figure 3). Note the CKMM digested fragments (labeled 2a and 2b), which, despite a 14-bp difference in size, have predicted and observed T_m points that differ by only 0.2°C and 0°C, respectively. As indicated in Figure 2, these two fragments melt together in a single peak. These observations have important implications for the design of McSNP experiments because they suggest that the various T_m prediction formula can be used as an accurate tool to estimate *a priori* whether or not there will be a resolvable difference in T_m between two DNA fragments. We suggest that, as a rough approximation, two DNA fragments can be resolved if their predicted T_m difference is at least 3°C.

Automated Scoring

To investigate the utility and robustness of McSNP on a larger scale, we genotyped 853 individuals for marker CKMM. We used the automatic scoring feature of the DASH software (version 2.14), which allows genotypes to be "called" automatically after an experiment is complete. Of the 853 samples, 37 failed to amplify (i.e., PCR failures). Of the remaining 816 samples, 807 were correctly genotyped and scored (98.9%). Thus, only 1.1% (nine individuals) of the reactions that were amplified and automatically scored with the DASH software were repeated because of inconsistencies with the observed melting curve pattern. These problems can be classified into two categories. First, unusual patterns in the relative size of the melting curves (0.7%) are likely due to partial digestion of the PCR products, and after subsequent amplification and digestion with increased amounts of enzyme and/or fewer PCR cycles, the samples show the expected melting profiles. Second, unusual peaks in the scoring range (0.4%) disappeared when the same samples were amplified, digested, and scored again with the DASH software. These problems are attributable to artifacts in the amplification or monitoring of the T_m profile. Overall, the automatic scoring feature of the DASH software allows a fast and accurate classification of genotypes. However, a careful inspection of the automatic scores is recommended to detect and correct possible errors caused by PCR artifacts or partial digestion.

DISCUSSION

In this report, we have investigated how the general approach of MCA, which is based on the natural phenomenon of T_m differences between DNA fragments that differ in sequence and size, can be used for genotyping. Specifically, we have developed McSNP, a simple, high-throughput, and inexpensive SNP genotyping assay. Furthermore, we have established the general robustness of McSNP by genotyping larger numbers of individuals and using the genotyping software of the DASH machine to automatically score genotypes.

While restriction enzymes may seem archaic compared to some of the recent technologies developed for SNP genotyping (8,24), it should be noted that they are very accurate, as they rely on tools that have evolved in nature to recognize SNPs. The two main disadvantages of "classic" PCR-RFLP typing are *(i)* limited throughput and *(ii)* partial digestion of PCR products, which decreases accuracy. However, both of these concerns are addressed by analyzing digestion products by McSNP rather than agarose gel electrophoresis.

First, throughput of gel-based methods is limited because it requires substantial time and labor to pour, load, stain, and interpret gels. On the other hand, McSNP reactions are self-contained because it is possible to conduct PCR, restriction enzyme digestion, and MCA in the same 96-well plate, with the analysis step requiring less than 20 min.

Second, partial digestion is a potential problem in any method using restriction enzymes and originates primarily from either too many PCR cycles or decreased efficiency of the restriction enzyme digestion because of the PCR components (e.g., Triton® X-100). Simple solutions to these problems include performing fewer PCR cycles and diluting the PCR product in a larger volume before restriction enzyme digestion. Both methods are more feasible for McSNP typing because the sensitivity of MCA is considerably higher than gel-based detection (data not shown). So far, we have seen little evidence for partial digestion (<1%; see Automated Scoring section of Results) and do not anticipate this to be an obstacle for most McSNP experiments.

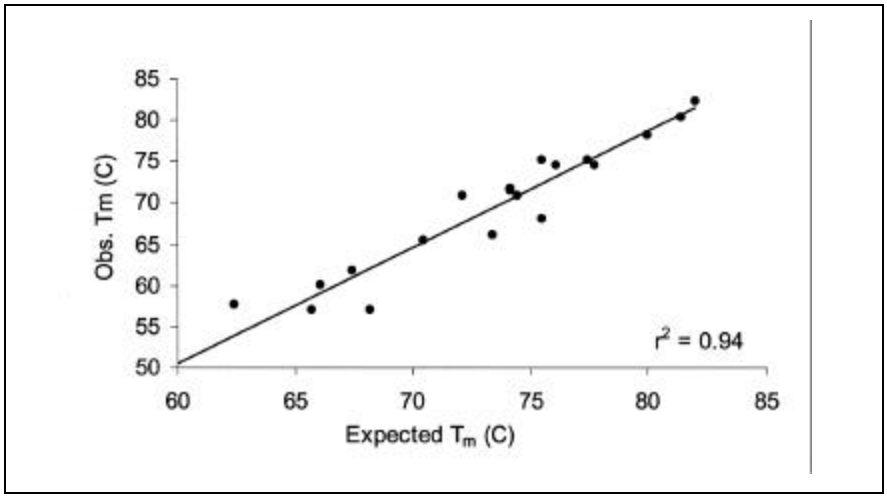

Figure 3. Relationship between observed and expected T_m points. We found a high correlation between observed and predicted T_m points. The predicted T_m is based on a salt adjusted nearest-neighbor algorithm described in the text (2), and the plotted data are from Table 1.

An important requirement of the McSNP assay is the presence of a restriction site generated by the SNP. However, it has been estimated that SNPs only occur in natural restriction sites 50% of the time (12). Thus, the ability to create artificial restriction sites during PCR is critical to the general applicability of McSNP. To this end, computer simulations have been performed, and it is estimated that 98% of all polymorphisms occur in either natural or artificial restriction sites for the commercially available four-base cutters (13). Hence, McSNP will be able to genotype the preponderance of identified SNPs.

It is also important to consider the cost of a genotyping assay. The reagent and consumable costs of a McSNP experiment are very low. The post-PCR cost per genotype can be broken down as follows: $0.06 for the restriction enzyme (an average of 15 commonly used enzymes), $0.001 for SYBR Green I, $0.0003 for formamide, and $0.06 for the polycarbonate plates (which can be used in the PCR step, further reducing costs). Thus, the cost of a McSNP experiment is approximately $0.12 beyond that of a PCR, making it one of the least expensive SNP genotyping methods available.

Finally, a realistic estimate of sample throughput is 2400 genotypes/day/instrument (96 samples × 25 runs/day, assuming each run requires 15 min), which translates into 576 000 genotypes/year (assuming 48 working weeks). The primary bottleneck of the McSNP protocol is the PCR step. Therefore, multiplexing McSNP experiments would be advantageous. In principle, it should be straightforward to multiplex two loci. However, solutions containing three or more loci would likely result in complex melting curve patterns that may substantially overlap. We are exploring the potential of multiplex McSNP experiments and are adapting McSNP to a microarray-based format. In conclusion, McSNP allows SNPs to be genotyped in a gel-free high-throughput manner and is a useful addition to the geneticist's genotyping toolbox.

ACKNOWLEDGMENTS

This work was supported in part by grant National Institutes of Health grant nos. DK53958 and HG02154 to M.D.S. There are several patents pending related to the contents of this manuscript.

REFERENCES

1. **Athma, P., N. Fidahusein, and M. Swift.** 1995. Single base polymorphism linked to the ataxia-telangiectasia locus is detected by mismatch PCR. Biochem. Biophys. Res. Commun. *210*:982-986.
2. **Breslauer, K.J., R. Frank, H. Blocke, and L.A. Marky.** 1986. Predicting DNA duplex stability from the base sequence. Proc Natl Acad Sci. USA *83*:3746-3750.
3. **Brookes, A.J.** 1999. The essence of SNPs. Gene *234*:177-186.
4. **Cohen, J.B. and A.D. Levinson.** 1988. A point mutation in the last intron responsible for increased expression and transforming activity of the c-Ha-ras oncogene. Nature *334*:119-124.
5. **Gennarelli, M., G. Novelli, A. Cobo, B. Baiget, and B. Dallapiccola.** 1991. 3′ creatine kinase (M-type) polymorphisms linked to myotonic dystrophy in Italian and Spanish populations. Hum. Genet. *87*:654-656.
6. **Germer, S. and R. Higuchi.** 1999. Single-tube genotyping without oligonucleotide probes. Genome Res. *9*:72-78.
7. **Gotoda, T., N. Yamada, T. Murase, H. Shimano, M. Shimada, K. Harada, M. Kawamura, K. Kozaki, and Y. Yazaki.** 1992. Detection of three separate DNA polymorphisms in the human lipoprotein lipase gene by gene amplification and restriction endonuclease digestion. J. Lipid Res. *33*:1067-1072.
8. **Griffin, T.J., J.G. Hall, J.R. Prudent, and L.M. Smith.** 1999. Direct genetic analysis by matrix-assisted laser desorption/ionization mass spectrometry. Proc. Natl. Acad. Sci. USA *25*:6301-6306.
9. **Higuchi, R., C. Fockler, G. Dollinger, and R. Watson.** 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology (NY) *11*:1026-1130.
10. **Howell, W.M., M. Jobs, U. Gyllensten, and A.J. Brookes.** 1999 Dynamic allele-specific hybridization. A new method for scoring single nucleotide polymorphisms. Nat. Biotechnol. *17*:87-88.
11. **Kwok, P.Y. and Z. Gu.** 1999. Single nucleotide polymorphism libraries: why and how are we building them? Mol. Med. Today *5*:538-543.
12. **Landegren, U., R. Kaiser, J. Sanders, and L. Hood.** 1988. A ligase-mediated gene detection technique. Science *241*:1077-1080.
13. **Li, H. and L. Hood.** 1995. Multiplex genotype determination at a DNA sequence polymorphism cluster in the human immunoglobulin heavy-chain region. Genomics *26*: 199-206.
14. **Livak, K.J., S.J. Flood, J. Marmaro, W. Giusti, and K. Deetz.** 1995. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl. *4*:357-362.
15. **Nikiforov, T.T., R.B. Rendle, P. Goelet, Y.H. Rogers, M.L. Kotewicz, S. Anderson, G.L. Trainor, and M.R. Knapp.** 1994. Genetic Bit Analysis: a solid phase method for typing single nucleotide polymorphisms. Nucleic Acids Res. *11*:4167-4175.
16. **Parra, E.J., A. Marcini, J. Akey, J. Martinson, M.A. Batzer, R. Cooper, T. Forrester, D.B. Allison, R. Deka, R.E. Ferrell, and M.D. Shriver.** 1998. Estimating African American admixture proportions by use of population-specific alleles. Am. J. Hum. Genet. *63*:1839-1851.
17. **Ririe, K.M., R.P. Rasmussen, and C.T. Wittwer.** 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal. Biochem. *245*:145-160.
18. **Rychlik, W. and R.E. Rhoads.** 1989. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Res. *17*:8543-8551.
19. **Sambrook, J., E.F. Fritsch, and T. Maniatis.** 1989. Molecular Cloning: A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, NY.
20. **Shriver, M.D., M.W. Smith, L. Jin, A. Marcini, J.M. Akey, R. Deka, and R.E. Ferrell.** 1997. Ethnic-affiliation estimation by use of population-specific DNA markers. Am. J. Hum. Genet. *60*:957-964.
21. **Su, B., J. Xiao, P. Underhill, R. Deka, W. Zhang, J. Akey, W. Huang, D. Shen et al.** 1999. Y-chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. Am. J. Hum. Genet. *65*:1718-1724.
22. **Tournamille, C., Y. Colin, J.P. Cartron, and C. Le Van Kim.** 1995. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. Nat. Genet. *10*:224-228.
23. **Turner, D.H.** 1996. Thermodynamics of base pairing. Curr. Opin. Struct. Biol. *6*:299-304.
24. **Wang, D.G., J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins et al.** 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science *280*:1077-1082.
25. **Wittwer, C.T., M.G. Herrmann, A.A. Moss, and R.P. Rasmussen.** 1997. Continuous fluorescence monitoring of rapid cycle DNA amplification. BioTechniques *22*:130-138.

Received 10 August 2000; accepted 16 October 2000.

Address correspondence to:

Joshua M. Akey
Human Genetics Center
Graduate School of Public Health
University of Texas Houston Health Science Center
6901 Bertner Avenue
Houston, TX 77030, USA
e-mail: jakey@gsbs3.gs.uth.tmc.edu