Parallel Single Nucleotide Polymorphism Genotyping by Surface Invasive Cleavage with Universal Detection

ARTICLE *in* ANALYTICAL CHEMISTRY · MAY 2005

Impact Factor: 5.64 · DOI: 10.1021/ac0483825 · Source: PubMed

4 AUTHORS, INCLUDING:

Michael Olivier
Texas Biomedical Research Institute
78 PUBLICATIONS **3,863** CITATIONS

SEE PROFILE

Anal. Chem. **2005,** *77,* 2400−2405

Parallel Single Nucleotide Polymorphism Genotyping by Surface Invasive Cleavage with Universal Detection

Yan Chen,[†] Michael R. Shortreed,[†] Michael Olivier,[‡] and Lloyd M. Smith*,[†]

Department of Chemistry, University of Wisconsin—Madison, 1101 University Avenue, Madison, Wisconsin 53706, and Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, Wisconsin 53226

Large-scale investigations of sequence variation within the human species will provide information about the basis of heritable variation in disease susceptibility and human migration. The surface invader assay (an adaptation of the invasive cleavage reaction to an array format) is capable of exquisitely sensitive and specific detection of genetic variation. It is shown here that this genotyping technology can be multiplexed in a DNA array format, permitting the parallel analysis of a panel of single nucleotide polymorphisms (SNPs) directly from an unamplified genomic DNA target. In addition, a "universal" mode of detection was developed that makes use of a mixture of degenerate templates for DNA ligation to the surface-bound cleaved oligonucleotides and thereby makes this strategy amenable to any desired SNP site or combination of SNP sites, without regard to their particular DNA sequences. This approach was demonstrated on a proof-of-principle scale using small DNA arrays to genotype 6 SNP markers in the PTPN1 gene and 10 mutations in the cystic fibrosis transmembrane conductance regulator gene. This ability to analyze many different genetic variations in parallel, directly from unamplified human genomic DNA samples, lays the groundwork for the development of high-density arrays able to analyze hundreds of thousands or even millions of SNPs.

The successful sequencing of the human genome has ushered in a new era for biological research, providing a framework and a foundation to unravel the complex sets of gene functions, interactions, and pathways that act in concert to comprise the human life form.[1] A critical aspect of this quest is the ability to understand, in as much detail as possible, the ways in which the individual parts of the system act and interact, i.e., to understand the function of the gene products encoded within the genome.[2] A key tool in analyzing gene function is determination of the relationship between genotype and phenotype. Hence, a major focus of current efforts in genomics is to elucidate the genetic variations extant within the human population, and to study the effects of these variations upon the human system.[3] Over 90% of genetic variations are single nucleotide polymorphisms (SNPs), each of which occurs every 500−1000 base pairs in the genome.[4] To date, over 8 million of these SNPs have been placed in a public database,[5] and 4.3 million SNPs have been validated in multiple studies. It has been suggested that large-scale pharmacogenetics and complex disease association studies will require the analysis of hundreds of thousands of SNPs from thousands of individuals (The SNP Consortium, http://snp.cshl.org). Recent advances in SNP genotyping technology[5−8] have enabled the simultaneous genotyping of as many as 100 000 SNPs. However, there remains much room for improvement.

Arguably the most significant single issue faced by many large-scale genotyping studies is the need for synthesis and handling of very large numbers of oligonucleotides. For example, if each SNP site is individually PCR amplified, then the analysis of 300 000 SNPs would require the synthesis and handling of 600 000 PCR primer oligonucleotides. Here, an approach is described to such large-scale analyses of genetic variation which addresses this problem by combining the specificity of invasive cleavage reactions[9] with the parallelism offered by DNA microarray technology.[10] By using well-developed photolithographic or other DNA array fabrication technologies, this approach renders obsolete the need for synthesis and handling of individual oligonucleotides for genetic analysis. Instead, all of the chemical complexity is built into the array, and the user of the technology has merely to add a genomic DNA sample and execute a few sequential steps.

* To whom correspondence should be addressed. Fax: (608) 265-6780. E-mail: smith@chem.wisc.edu.

† University of Wisconsin—Madison.

‡ Medical College of Wisconsin.

(1) Pennisi, E. *Science* **2003**, *300*, 409.

(2) Collins, F. S.; Green, E. D.; Guttmacher, A. E.; Guyer, M. S. *Nature* **2003**, *422*, 835−847.

(3) White, K. P. *Nat. Rev. Genet.* **2001**, *2*, 528−537.

(4) Kruglyak, L. *Nat. Genet.* **1997**, *17*, 21−24.

(5) Murray, S. S.; Oliphant, A.; Shen, R.; McBride, C.; Steeke, R. J.; Shannon, S. G.; Rubano, T.; Kermani, B. G.; Fan, J.-B.; Chee, M. S.; Hansen, M. S. T. *Nat. Methods* **2004**, *1*, 113−117.

(6) Hardenbol, P.; Baner, J.; Jain, M.; Nilsson, M.; Namsaraev, E. A.; Karlin-Neumann, G. A.; Fakhrai-Rad, H.; Ronaghi, M.; Willis, T. D.; Landegren, U.; Davis, R. W. *Nat. Biotechnol.* **2003**, *21*, 673−678.

(7) Oliphant, A.; Barker, D. L.; Stuelpnagel, J. R.; Chee, M. S. *Biotechniques* **2002**, Suppl., 56−58, 60−51.

(8) Matsuzaki, H.; Dong, S.; Loi, H.; Di, X.; Liu, G.; Hubbell, E.; Law, J.; Berntsen, T.; Chadha, M.; Hui, H.; Yang, G.; Kennedy, G. C.; Webster, T. A.; Cawley, S.; Walsh, P. S.; Jones, K. W.; Fodor, S. P. A.; Mei, R. *Nat. Methods* **2004**, *1*, 109−111.

(9) Lyamichev, V.; Mast, A. L.; Hall, J. G.; Prudent, J. R.; Kaiser, M. W.; Takova, T.; Kwiatkowski, R. W.; Sander, T. J.; de Arruda, M.; Arco, D. A.; Neri, B. P.; Brow, M. A. *Nat. Biotechnol.* **1999**, *17*, 292−296.

(10) Fodor, S. P.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767−773.

10.1021/ac0483825 CCC: $30.25
Published on Web 03/09/2005

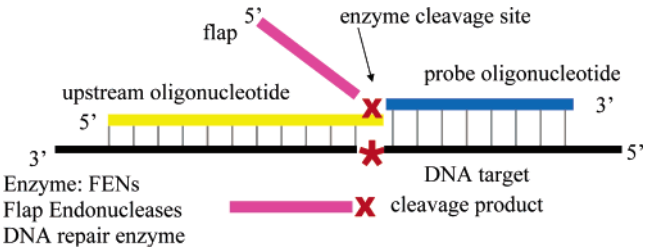

Figure 1. Schematic diagram of the invasive cleavage assay. The nucleotide at the 3′ end of the upstream oligonucleotide is designed to overlap at least one base into the downstream duplex formed by the probe and the target strand. The unpaired region on the 5′ end of the probe, or "flap", along with an immediate downstream paired nucleotide can then be removed by a class of structure-specific 5′-exonucleases. Absolute complementarity between the probe and the target sequence at the position of overlap is required for efficient enzymatic cleavage.

In invasive cleavage reactions, an upstream oligonucleotide and a downstream "probe" oligonucleotide hybridize in tandem to the target DNA strand and create a three-dimensional structure with a one-base-pair overlap (the invasive structure, Figure 1). The 5′ end of the probe oligonucleotide does not complement the target, but forms a 5′ arm or "flap" upon hybridization. This invasive structure is the substrate for a flap endonuclease (FEN), which cleaves the 5′ flap off the probe oligonucleotide after its first paired base. Absolute complementarity between the probe and the target sequence at the position of overlap is required for efficient enzymatic cleavage, which provides a cleavage rate ~300 times higher than for a noncomplementary substitute.[11] Reactions are performed near the melting temperature (T_m) of the hybridization region between the probe and target strand so that the cleaved and uncleaved probe oligonucleotides cycle during the course of the reaction to produce about 2000 cleavage events per hour, providing signal amplification.[9] The combination of the specificity afforded by the requirement that the upstream and downstream oligonucleotides hybridize adjacent to one another on the genomic target DNA, with the specificity conferred by the enzymatic recognition of the invasive cleavage complex, yields very high overall specificity for the cleavage reaction, suitable for the targeted recognition of a single nucleotide in a complex human genomic DNA sample.

It has been shown previously that the invasive cleavage assay can be adapted to an array format while maintaining sufficient sensitivity to detect polymorphisms in unamplified genomic DNA.[12] It is shown here that this reaction can be multiplexed in a DNA array format, permitting the parallel analysis of a panel of SNPs. Parallelization of the surface invader assay has no perceptable effect on the overall sensitivity, and thus, the assay is capable of detecting polymorphisms directly from unamplified genomic DNA. A universal mode of detection was developed to facilitate parallelization of the invader assay in the array format (Figure 2). This detection involves ligation of a primer for rolling circle amplification (RCA) directly to the cleaved probe. One primer can be ligated to all possible probe sequences because of the use of a degenerate template DNA. Once the primer is ligated to the probe, it is extended by RCA and labeled with a double-strand-specific fluorescent dye. The surface is imaged with an array scanner, and the polymorphisms are scored.

EXPERIMENTAL SECTION

Reaction Format. A three-dimensional structure (Figure 1) is formed by hybridization of two partially overlapping DNA oligonucleotides (probe and upstream oligonucleotides) to a target DNA strand.[9] An FEN specifically recognizes this structure and cleaves the probe oligonucleotide at the position of the overlap.[9] The probe oligonucleotides are provided in excess. Probe cycling is made possible by conducting the reactions at temperatures near the melting temperature of the probe−target duplex, yielding many cleavage events per minute per target (linear signal amplification). Uncleaved probes are terminated with a dabcyl blocking group which prevents nonspecific ligation. Other blocking groups such as trityl and biotin were also tested (data not shown). The results indicated that the dabcyl group gave better blocking efficiency, and accordingly, it was employed for this study.

Upstream and Probe Oligonucleotides. The surface-bound upstream and probe oligonucleotides were synthesized at the University of Wisconsin Biotechnology Center, Madison, WI. A series of 10 18-atom spacer moieties were also introduced in these oligonucleotides, providing a total spacer length of 240 Å. The use of such a spacer region between an oligonucleotide and a surface is often critical to obtaining good performance in surface hybridization.[13,14] Nucleotide sequences of upstream and probe oligonucleotides used in genotyping the 6 SNP markers in the PTP gene and in the detection of 10 cystic fibrosis mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene are given in the Supporting Information. Prior to purification, both 3′- and 5′-thiol-modified oligonucleotides were deprotected using standard procedures, purified by reversed-phase binary gradient elution HPLC (Shimadzu SCL-6A), and stored under an

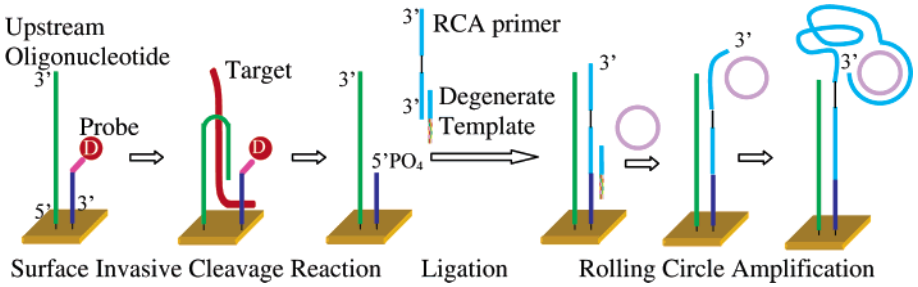

Figure 2. Schematic of the surface invasive cleavage reaction with rolling circle amplification (RCA). Probe and upstream oligonucleotides are coimmobilized onto the patterned surface. The degenerate template used in the ligation step has a sequence of 5′-NNNNNNGCATTCCG, where N = G, A, T, or C. The RCA primer is a special 3′−3′ oligonucleotide containing a primer sequence complementary to a sequence contained within the small single-stranded circular DNA template employed for RCA. The linear RCA product comprises repeating units of a sequence complementary to the circle sequence.

inert atmosphere. Oligonucleotide concentrations were determined by measuring absorption at 260 nm with an HP8453 UV−vis spectrophotometer.

Surface Preparation. The thiol-modified oligonucleotides were immobilized on photopatterned gold-coated glass substrates (Evaporated Metal Films, New York) via a multistep chemical modification.[15] A self-assembled monolayer (SAM) of $HS(CH_2)_{18}NH_2$[12] was formed on a 5 × 5 mm gold-coated glass substrate. Exposure of this SAM to UV light through a mask patterned with 500 μm square features resulted in selective oxidation of exposed regions of the surfaces. This surface was then rinsed to remove the oxidized material. Exposure of it to a 1 mM solution of $HS(CH_2)_{18}CH_3$ (Aldrich) resulted in the formation of a patterned surface with well-defined surface chemistry.[15] Free amine groups of the SAM were then reacted with the heterobifunctional linker sulfosuccinimidyl 4-(*N*-maleimidomethyl)cyclohexane-1-carboxylate (SSMCC) (Pierce), creating a thiol-reactive, maleimide-terminated surface that could covalently react with thiol-modified DNA strands.[15] Thiol-modified probe oligonucleotide (1 mM) and thiol-modified upstream oligonucleotide (1 mM) were mixed in equal volumes before deposition on the surface. The surface attachment reaction was permitted to occur for approximately 20 h in a humid chamber to prevent evaporation. The chips were then rinsed briefly with water, soaked sequentially in 2× SSPE for 2 h at 37 °C and in 8 M urea for 15 min (room temperature), and then rinsed with water again before the onset of the invasive cleavage reaction.

Human Genomic DNA Samples. Human genomic DNA samples (obtained from the Coriell Institute) were ethanol-precipitated, air-dried, dissolved in distilled water, and used directly in the cleavage reaction without further treatment.

Invasive Cleavage Reaction Conditions. In the invasive cleavage reaction,[12] 10 μL of a solution of human genomic DNA (1 pM, denatured at 95 °C and then snap-chilled on ice prior to use as a target for the surface invasive cleavage reaction), buffer (10 mM MOPS, pH 7.5, and 7.5 mM $MgCl_2$), and Afu FEN 1 enzyme (cleavase VIII, 50 ng, provided by Third Wave Technologies) were added to the photopatterned surface followed by incubation at 58.5 °C in a humid environment for 5 h. The chips were then rinsed with water, soaked in 8 M urea for 15 min (room temperature), and then rinsed again with water.

Ligation with Degenerate Templates. The mixture of degenerate ligation templates (obtained from the University of Wisconsin Biotechnology Center) with a sequence of 5′-NNN-NNNGCATTCCG (where N = G, A, T, or C) was synthesized by employing a mixture of nucleoside phosphoramidites at six positions at the 5′ end during chemical DNA synthesis. A specially modified 3′−3′ RCA primer (purchased from and PAGE purified by the Yale University Critical Technologies facility), 3′-AGCG-TACAGGATAGGAGTCGACACTAGTATTTT-(C_{18})-TTTTCGGAAT-GC -3′, was obtained by reversal of backbone polarity during chemical synthesis[16] immediately after the C_{18} spacer and was designed to have a priming sequence of 29 bases that is complementary to the circular DNA. In the primer shown above, C_{18} refers to spacer phosphoramidite 18 (Glen Research Corp., Sterling, VA), which is a hexaethylene glycol linker. The primer oligonucleotide (3 μM) for RCA was ligated specifically to the cleaved probes on the photopatterned surface at 0 °C overnight using degenerate ligation templates at an estimated total concentration of 300 μM. The reaction solution (20 μL) also contained 1 U/μL T4 DNA ligase in a ligation buffer supplied with the enzyme (all from New England Biolabs).

Determination of Ligation Efficiency. The approximate efficiency of the surface ligation step was determined by measuring the ratio of fluorescent signal that was obtained from the surface-ligated 3′−3′ RCA primer, which was fluorescently tagged, to that obtained from hybridization with fluorescent complements to the immobilized probe. Prior to the onset of the ligation reaction, surface-immobilized probes were hybridized with their fluorescein-tagged complements (2 μM in 2× SSPE) for half an hour at room temperature. The image of the chip was captured with a scanner (GeneTAC UC4×4, Genomic Solutions, Ann Arbor, MI), and the average fluorescence intensity of each probe spot was recorded for comparison later. Next, the hybridized probes were removed from the surface by treating the chip by soaking it in 8 M urea for 15 min followed by copious rinsing with water. The surface ligation reaction was then carried out as described above under various conditions of temperature and template length, with 2 μM ligation template and 2 μM 5′ fluorescein-tagged RCA primer. The chips were soaked again in 8 M urea for 15 min, rinsed with water, and scanned. The fluorescence signal from the surface-ligated 3′−3′ RCA primer was then divided by the corresponding probe hybridization signal to yield the ligation efficiency.

The approximate efficiency of the solution-phase ligation reaction was determined by measuring the ratio of fluorescent signal that was obtained from a fluorescently tagged ligation product to the sum of the same product and the residual fluorescently tagged probe. In the solution-phase ligation reaction, the probe oligonucleotide was tagged with a fluorescein moiety at its 3′ end. The reaction was carried out overnight as described above under various conditions of temperature and template length, with the ligation template, probe, and 3′−3′ RCA primer at a concentration of 2 μM. Following the ligation reaction, the product mixture was separated on a 15% denaturing PAGE gel. The unligated probe and ligation product appeared as two distinct bands due to the difference in their lengths. The fluorescence signals from each of these two bands was integrated and used in the calculation of ligation efficiency.

Rolling Circle Amplification. A linear precursor for the circular DNA 5′-PO_4-CGCATGTCCTATCCTCAGCTGTGATCA-TCAGAACTCACCTGTTAGACGCCACCAGCTCCAACTGTGA-AGATCGCTTAT-3′, guide DNA 5′-AGGACATGCGATAAGCGATC-3, and RCA detection probe 5′-FAM-TCAGAACTCACCTGTTAG-3′ were all obtained in PAGE-purified form from Integrated DNA

(11) Hall, J. G.; Eis, P. S.; Law, S. M.; Reynaldo, L. P.; Prudent, J. R.; Marshall, D. J.; Allawi, H. T.; Mast, A. L.; Dahlberg, J. E.; Kwiatkowski, R. W.; de Arruda, M.; Neri, B. P.; Lyamichev, V. I. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8272−8277.

(12) Chen, Y.; Shortreed, M. R.; Peelen, D.; Lu, M.; Smith, L. M. *J. Am. Chem. Soc.* **2004**, *126*, 3016−3017.

(13) Guo, Z.; Guilfoyle, R. A.; Thiel, A. J.; Wang, R.; Smith, L. M. *Nucleic Acids Res.* **1994**, *22*, 5456−5465.

(14) Shchepinov, M. S.; Case-Green, S. C.; Southern, E. M. *Nucleic Acids Res.* **1997**, *25*, 1155−1161.

(15) Brockman, J. M.; Frutos, A. G.; Corn, R. M. *J. Am. Chem. Soc.* **1999**, *121*, 8044−8051.

(16) Lizardi, P. M.; Huang, X.; Zhu, Z.; Bray-Ward, P.; Thomas, D. C.; Ward, D. C. *Nat. Genet.* **1998**, *19*, 225−232.

Technologies (Coralville, IA). To make single-stranded circles for the RCA step, the linear precursors (1 μM) were ligated by 1 U/μL T4 DNA ligase using guide sequences (5 μM) as templates at 16 °C overnight.[16] The product was then purified with the QIAquick Nucleotide Removal Kit (QIAGEN, Inc.). The concentration of the circular DNA was determined using UV−vis absorption (260 nm). Afterward, the ssDNA circle (0.1 μM in 10 μL of 50 mM Tris−HCl, pH 7.5) was applied to the surface, which had previously been ligated to the 3′−3′ RCA primers, for 30 min at 37 °C. Then, the RCA reaction was performed at 30 °C for up to 17 h upon the addition of 0.5 U/μL phage Ø29 DNA polymerase, 0.2 μg/μL BSA, and 3 mM dNTP in 50 μL of the reaction buffer (all from New England Biolabs).[16] Chip(s) were then incubated with the RCA detection probe (2 μM in 1× SSPE buffer, pH 7.5) for half an hour at room temperature and stained with SYBR Green I dye (Molecular Probes) in TE (10 mM Tris−HCl, 1 mM EDTA, pH 8.0) buffer at room temperature for 40 min. The fluorescein group on the RCA detection probe does not provide much additional detection sensitivity, but was useful for monitoring RCA probe hybridization during development of this system (e.g., for detecting larger amounts of target). The surface was then washed with 1× TBE buffer, and the surface fluorescence was imaged in 1× SSPE buffer (pH 7.5) using a GeneTAC UC4×4 microarray scanner.

RESULTS AND DISCUSSION

Universal Detection with Degenerate Ligation. Rolling circle amplification provides a mechanism to augment the amount of detectable signal from 5′-phosphate groups that are produced from flap cleavage during the invasive cleavage reaction.[12,16] Probes with 5′-phosphate groups are ligated to a special oligonucleotide (the 3′−3′ RCA primer described above) that contains a primer sequence for the rolling circle amplification. Noncleaved probes are unable to participate in ligation reactions because they remain blocked by dabcyl moieties.[17] The nature of the ligation template, which is required in the reaction, is the key issue that governs the generality of this detection scheme for parallel SNP analysis. The template DNA must possess sequence complementary to both the surface-bound probe oligonucleotide, corresponding to the SNP target, and the RCA circle. If each SNP targeted for analysis required a unique complementary template for ligation, the utility of the overall scheme would be severely compromised, as it would be necessary to synthesize a distinct ligation template for each SNP analyzed. This would obviate the point of the approach, which is to eliminate the need for individual syntheses of large numbers of oligonucleotides.

To address this generality issue, the use of degenerate templates for ligation was explored. It is well-known that degenerate mixtures of oligonucleotides are readily synthesized by employing a mixture of nucleoside phosphoramidites at a given position during chemical DNA synthesis.[18] It was reasoned that such a mixture of oligonucleotides might serve as a "universal" template for DNA ligation to the surface-bound cleaved oligonucleotides and thereby make the strategy shown in Figure 2 applicable to any desired SNP site or combination of SNP sites, without regard to their particular DNA sequences.

The two parameters of the template design that are particularly important are the overall length and the length of the degenerate portion. The overall length of the template has a large effect upon the stability of the duplex formed under the salt and temperature conditions employed in the ligation reaction. The length of the region in which degenerate bases are employed is important because the larger the number of different sequences that are present within the population (for the case in which all four nucleotides are combined at each of n positions, the mixture complexity is given by 4^n) the lower the concentration of any particular sequence represented in the mixture (given by 1 part in 4^n). The ligation reaction was optimized by varying both these parameters as well as the reaction temperature. T4 DNA ligase was employed because it yields high ligation efficiency and is active over a broad range of temperatures even though the optimal incubation temperature is 16 °C. Our studies showed that, at 16 °C, a minimum of six degenerate bases is required to yield a surface ligation efficiency of 25%, while in the solution the same condition gives an efficiency above 95% (data not shown). The reason for this is not clear. However, we found empirically that lowering the reaction temperature to 0 °C increased the surface ligation efficiencies to 40% when four degenerate bases were used in the template and to 80% when six degenerate bases were used in the template. Ligation results obtained in this manner are shown in Figure 3. Subsequent surface ligations were performed using six degenerate base templates.

Parallelization of the Surface Invader Assay. Two model systems were chosen on which to develop and demonstrate our approach to parallel genotyping. Although the eventual large-scale implementation of this approach will utilize a high-density array created by photolithographic or other means, the development of the basic chemistry and enzymology is readily done on a smaller scale using hand-spotted probes. The results of surface invasive cleavage reactions performed using genomic DNA as a target on a panel of six SNPs (dbSNP reference numbers rs6012953, rs932420, rs941798, rs1885176, rs6020605, and rs1885177) located in the PTPN1 gene16 on human chromosome 20 encoding protein tyrosine phosphatase 1 β (PTP-1β) are displayed in Figure 4. These SNPs are distributed across the entire gene region spanning over 68 kb, have allele frequencies ranging from 38.8% to 49.4%, and are in strong linkage disequilibrium. For each SNP, allele-specific probe and upstream oligonucleotide pairs were spotted in duplicate according to the pattern shown in the leftmost panel of Figure 4. The substrate was incubated for 5 h at 58.5 °C in a solution of denatured human genomic DNA from either of two individuals, along with FEN enzyme and buffer. DNA samples (20 μg or 10 amol) were from two individuals samples (nos. PD0003 and PD0004) of the Coriell Polymorphism Discovery Resource panel (PD24, C Coriell Cell Repository, Camden, NJ).[19] After the washing steps, the 3′−3′ RCA primer oligonucleotide was ligated to the surface at 0 °C using the degenerate template described above, followed by performance of the linear RCA reaction and subsequent hybridization, fluorescence labeling, and fluorescence imaging steps (see the Experimental Section for details). Inspection of the figure clearly reveals the genotype of both individuals. Individual no. PD0004 is heterozygous for five/

(17) Tyagi, S.; Bratu, D. P.; Kramer, F. R. *Nat. Biotechnol.* **1998**, *16*, 49−53.
(18) Oliphant, A. R.; Nussbaum, A. L.; Struhl, K. *Gene* **1986**, *44*, 177−183.
(19) Collins, F. S.; Brooks, L. D.; Chakravarti, A. *Genome Res.* **1998**, *8*, 1229−1231.

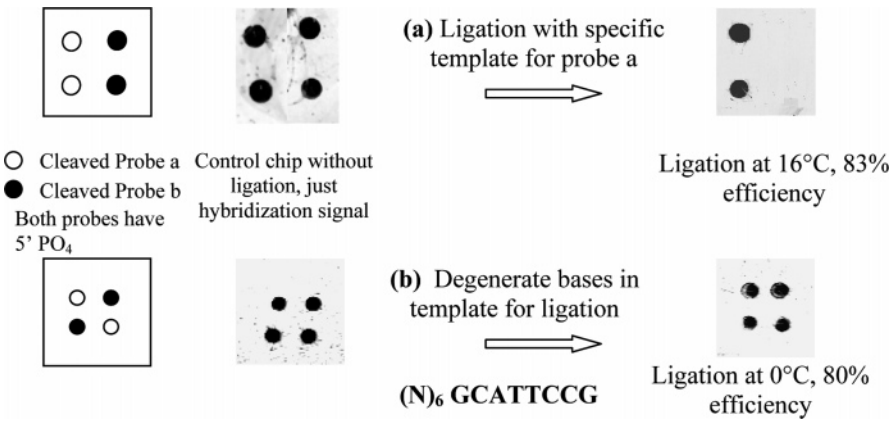

Figure 3. Surface ligation studies with specific and degenerate templates. The ligated oligo is fluorescein-tagged at the 5′ end with a sequence of 5′-FAM-CGGAATGC. For the degenerate template sequence [5′-NNNNNNGCATTCCG, N = G, A, T, or C], surface ligation efficiencies up to 80% could be obtained for ligation to a wide variety of immobilized probe sequences at 0 °C **(b)**, comparable to the case with a specific template **(a)**. Probe a = 5′-(PO_4)-CTTTTGCAGGTCATCGG-$(S_{18})_{10}$-SH. Probe b = 5′-(PO_4)-GAACGAACCCAACGCAA-$(S_{18})_{10}$-SH. Ligation probe = 5′-(6-FAM)-CGGAATGC.

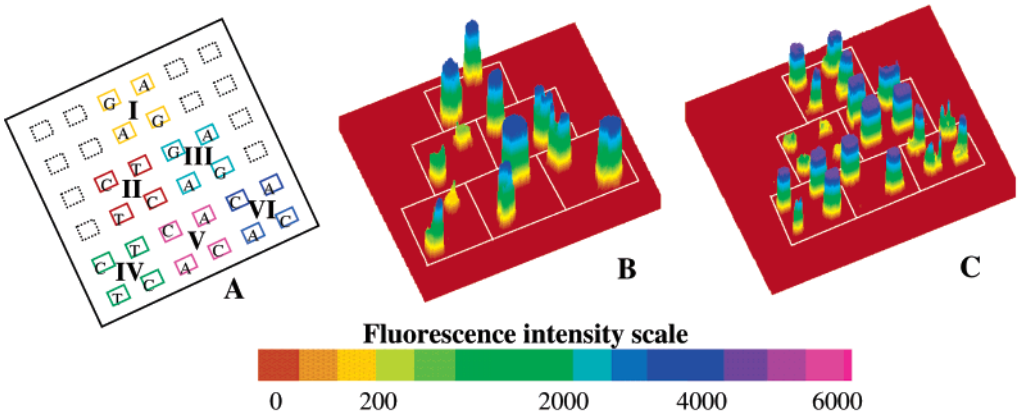

Figure 4. Genotyping six SNP markers in the PTP gene: (I) rs6012953, (II) rs932420, (III) rs941798, (IV) rs1885176, (V) rs6020605, and (VI) rs1885177. **(**A**)** Layout of the invasive cleavage array, showing positions of elements corresponding to the six SNPs being tested. Each allele is tested in duplicate. **(**B, C) Fluorescence images of genotyping results from two individuals (nos. PD0003 and PD0004 of the Coriell Polymorphism Discovery Resource panel (PD1−24), respectively). The presence of four peaks within a white square indicates a heterozygous genotype; two peaks indicate a homozygous genotype.

six SNPs, as indicated by the fluorescence signal obtained for both alleles in five cases, and for only one allele in the sixth case (PTP1159_291). In contrast, individual no. PD0003 is homozygous for all six SNPs, as indicated by the presence of a fluorescence signal for only one allele in each case. These genotypes are identical to the genotypes for these individuals as determined in standard solution invasive cleavage assays. It may be noted that the fluorescence signal obtained in the surface array format is not uniform for the two probe oligonucleotides for the same SNP and/or between probes for different SNPs under the same reaction conditions. This difference is likely attributable to the difference in the melting temperatures of the different probe oligonucleotides. This difference can be minimized, if desired, by varying the length and/or composition of the probe oligonucleotides to yield similar melting temperatures.[20]

The above study demonstrated the parallel analysis of a panel of common SNPs directly from unamplified human genomic DNA. A second system was chosen to show the ability to directly analyze a panel of causative mutations associated with a genetic disease; to this end, a panel of 10 mutations in the CFTR gene on chromosome 7 was selected.[21] Since the discovery of this gene in 1989, almost 1000 different mutations have been identified, only a few of which are of high frequency.[22] Recent recommendations by both the American College of Medical Genetics (ACMG) and the American College of Obstetricians and Gynecologists (ACOG) have led to the adoption of a standard CF gene mutation panel. It consists of the 25 most common mutations (those with a frequency of greater than 0.1% in the general U.S. population) as additional mutations are rare, are not well understood, and have minimal impact on the disease severity of an affected child. The majority of CF in the U.S. is caused by a three-base-pair deletion (ΔF508), accounting for approximately 66% of all CF chromosomes worldwide. This mutation causes a deletion of phenylalanine at position

(20) Kwiatkowski, R. W.; Lyamichev, V.; de Arruda, M.; Neri, B. *Mol. Diagn.* **1999**, *4*, 353−364.
(21) Drumm, M.; Collins, F. In *Molecular Genetic Medicine*; Friedmann, T., Ed.; Academic Press: San Diego, CA, 1993; Vol. 3, pp 33−64.
(22) Grody, W. W.; Cutting, G. R.; Klinger, K. W.; Richards, C. S.; Watson, M. S.; Desnick, R. J. *Genet. Med.* **2001**, *3*, 149−154.

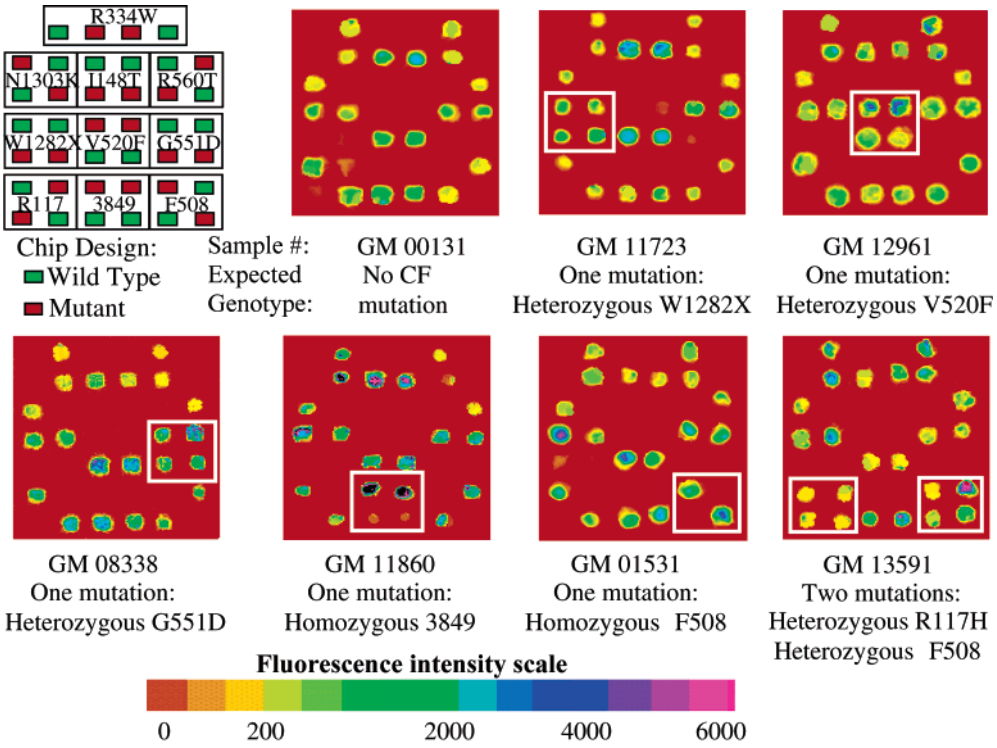

Figure 5. Detection of 10 cystic fibrosis mutations in the CFTR gene. A locator for the probe and upstream oligonucleotides specific for each of the 10 interrogated mutations is shown in the upper left corner of the figure. Genotype data from human genomic DNA samples including six known mutation carriers and one CF-free individual as a control are shown. The white boxes highlight genotyping of expected mutations. The genotypes obtained were in 100% concordance with the known genotypes as provided by the Coriell Institute.

508 of the protein product.[23] The CF mutations chosen in this study, ΔF508, G551D, W1282X, N1303K, R117H, R560T, 3849+10kbCT, V520F, R334W, and I148T, are a subset of the standard panel. Human genomic DNA samples with characterized genotypes including six unrelated individuals carrying cystic fibrosis mutations and one CF-free individual were obtained from the Coriell Cell Repository. The samples were denatured and genotyped on a surface invasive cleavage array as described above for the PTPN1 SNPs. The fluorescence images obtained for each individual are displayed in Figure 5. Among the six unrelated CF carriers, one was found to be homozygous for 3849+10kbCT, one homozygous for ΔF508, one heterozygous for both R117H and ΔF508, one heterozygous for W1282X/WT, one heterozygous for V520F/WT, and one heterozygous for G551D/WT. The genotypes obtained were in 100% concordance with the known genotypes provided by the Coriell Institute (http://locus.umdnj.edu/nigms/charmut/cysfib.html).

CONCLUSION

The present work provides a proof-of-principle demonstration of the ability to combine the specificity of the invasive cleavage reaction with the parallelism offered by the DNA array. This ability to analyze many different genetic variations in parallel, directly from unamplified human genomic DNA samples, lays the groundwork for the development of high-density arrays able to analyze hundreds of thousands or even millions of SNPs. A near-term application of the technology is in the analysis of panels of causative mutations. For example, it would be fairly straightforward to develop a CF chip that was able to analyze all of the ~1000 known mutations in the CFTR gene, rather than the small subset of mutations which can be cost-effectively addressed using current approaches.

ACKNOWLEDGMENT

We thank Third Wave Technologies (TWT) for providing the FEN enzyme and the designs of the invasive cleavage assays for the CF mutations. L.M.S. has a financial interest in TWT. This work was supported by NIH Grant No. R01HG02298.

SUPPORTING INFORMATION AVAILABLE

Nucleotide sequences of upstream and probe oligonucleotides used in genotyping the 6 SNP markers in the PTP gene and in the detection of 10 cystic fibrosis mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review November 1, 2004. Accepted January 27, 2005.

AC0483825

(23) Riordan, J. R.; Rommens, J. M.; Kerem, B.; Alon, N.; Rozmahel, R.; Grzelczak, Z.; Zielenski, J.; Lok, S.; Plavsic, N.; Chou, J. L.; et al. *Science* **1989**, *245*, 1066−1073.