Surface Amplification of Invasive Cleavage Products

Yan Chen,[†] Michael R. Shortreed,[†] Dora Peelen,[†] Manchun Lu,[†] and Lloyd M. Smith*,[†,‡,§]

Department of Chemistry, University of Wisconsin−Madison, 1101 University Avenue, Madison, Wisconsin 53706-1396, and Third Wave Technologies, 502 South Rosa Road, Madison, Wisconsin 53719-1256

Received October 31, 2003; E-mail: smith@chem.wisc.edu

Genetic variations that change the properties of the organism in a discernible fashion provide essential clues to the function of the altered gene. 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.[1] The most common type of genetic variations are the single nucleotide polymorphisms (SNPs), which occur every 500−1000 nt in the genome.[2] A critical issue in studies which seek to link such genetic variations to phenotypic variations is the scale upon which they must be performed. There are about 3 million SNPs within the human population.[3] So far, around 1.8 million SNPs have been mapped and placed in a public database by the not-for-profit SNP Consortium.[4] Large-scale population association studies to study the biological or medical significance of such variations may require the analysis of hundreds of thousands of SNPs on thousands of individuals. Thus, well-designed studies seeking to elucidate the genetic basis of common phenotypes such as heart disease, diabetes, or cancer would ideally be able to analyze billions of individual genetic differences.[5]

We are pursuing development of an approach to such large-scale analyses of genetic variation which combines the specificity of invasive cleavage reactions with the parallelism offered by DNA microarray technology.[6] In invasive cleavage reactions, a three-dimensional structure is formed by hybridization of two partially overlapping DNA oligonucleotides to a target DNA strand.[7] A flap endonuclease (FEN) specifically recognizes this structure and cleaves a flap from the 3′ probe oligonucleotide. The cleavage rate is approximately 300 times greater if the nucleotide at the cleavage site of the flap is complementary to the opposing nucleotide on the target strand, conferring single nucleotide specificity to the cleavage reaction.[8] Cycling of the probe oligonucleotide on the target strand results in about 2000 cleavage events per hour, providing signal amplification. Further amplification can be achieved in solution reactions using a second stage of invasive cleavage.[8] Detection of the cleavage products, which are specifically produced in a target-dependent manner with single-nucleotide specificity, signals the presence of the target DNA sequence in the sample. As described previously,[6] this approach may be parallelized by fabrication of a DNA array in which the two target-specific probe oligonucleotides are immobilized within each element in the array. Addition of target DNA to the solution permits formation of the specific invasive cleavage complex at each site of the array, followed by specific cleavage. A very significant advantage of this approach to large-scale SNP analysis is that all of the oligonucleotides needed may be synthesized in parallel on the support using array fabrication technologies. This is a tremendous advantage, as it eliminates the need for the synthesis and handling of millions of
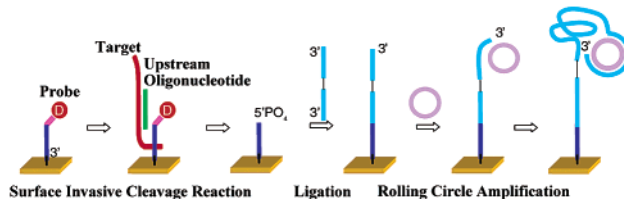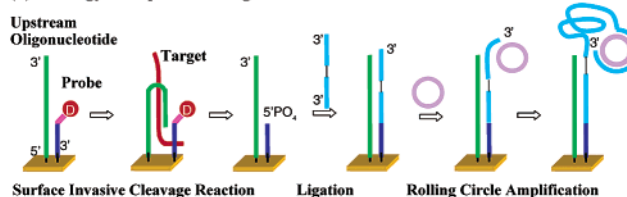

Figure 1. Schematic of the surface invasive cleavage reaction with rolling circle amplification. The linear RCA product comprises repeating units of a sequence complementary to the circle sequence.

individual probe oligonucleotides, a gargantuan and very expensive task of dubious practicality. In this approach, the simple addition of target DNA, enzyme, and buffer to the array, followed by incubation, washing, and detection steps, would reveal the genotype of the target DNA for every SNP site included in the array.

A major issue with this approach to the highly parallel and specific SNP analysis on a DNA array has been detection sensitivity. The concentrations of genomic target DNA which may practically be obtained from patient samples are in the picomolar range, which gives rise to cleavage at only a small fraction of the accessible surface sites. In previous work,[6] this limitation restricted the DNA targets which could be employed successfully to synthetic oligonucleotides, present at relatively high concentrations, or PCR amplified DNAs, which, due to the requirement for PCR primer synthesis for each SNP site, would obviate much of the advantage of the array approach for whole-genome analysis. Better detection sensitivity was reported with a microsphere-based[9] invasive cleavage assay. However, this approach does not land itself to parallelization. We describe here a highly sensitive approach to the detection of surface invasive cleavage reactions, utilizing rolling circle amplification (RCA) on the surface to produce large amounts of signal from relatively few cleavage sites. We use this approach to demonstrate the feasibility of SNP analysis on the surface directly from unamplified human genomic DNA, using as a model system a polymorphic site in codon 158 of the human ApoE gene, which plays a key role in the transport and metabolism of plasma cholesterol and triglycerides.[10]

Figure 1 shows a schematic of the strategies employed. A self-assembled monolayer (SAM) of the 18-mercapto octadecylamine, $HS(CH_2)_{18}NH_2$, was formed on a 5×5 mm gold-coated glass substrate. Exposure of this SAM to UV light through a mask

---

† University of Wisconsin−Madison.
‡ Third Wave Technologies.
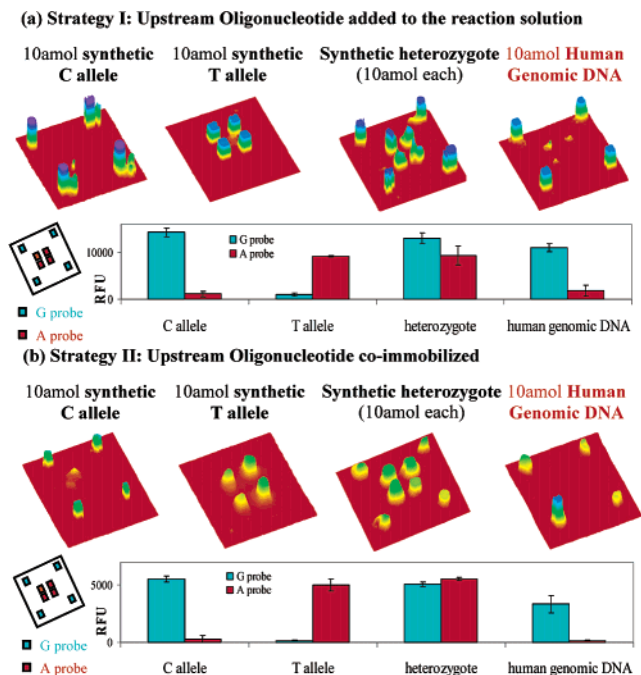§ Lloyd M. Smith has a financial interest in Third Wave Technologies, Inc.

Figure 2. Fluorescence images of results from the surface invasive cleavage reaction with rolling circle amplification.

patterned with 500 μm square features resulted in the selective oxidation of exposed regions of the surfaces. This surface was then rinsed to remove the oxidized material, and exposure of it to a solution of $HS(CH_2)_{18}CH_3$ resulted in the formation of a patterned surface with well-defined surface chemistry.[11] Free amine groups on $HS(CH_2)_{18}NH_2$ were then reacted with the heterobifunctional linker sulfosuccinimidyl 4-(*N*-maleimidomethyl) cyclohexane-1-carboxylate (SSMCC), creating a thiol-reactive, maleimide terminated surface that could covalently react with thiol-modified DNA strands.[11] Probe oligonucleotides specific for either allele of the target sequence and modified at their 5′ termini with a Dabcyl moiety, which serves in the present work as a 5′ blocking group, and at their 3′ termini with a thiol group, were aliquoted onto elements of the patterned surface for attachment. In strategy I, only the probe oligonucleotide was immobilized on the surface, while the upstream oligonucleotide was added in solution with a target strand. In strategy II, both the probe and the upstream oligonucleotides were immobilized on the surface in a 1:1 molar ratio, and only the target strand was added in solution. A small fraction of the probe oligonucleotides were then specifically cleaved in a target-dependent fashion by surface invasive cleavage, producing 5′ phosphate groups on the surface. These 5′ phosphates were ligated to 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, followed by performance of the linear RCA reaction to produce long single-stranded DNAs at the cleavage site.[12] After the RCA reaction, the surface was incubated with the RCA detection probe and stained with SYBR Green I, a double-strand-specific fluorescent dye, followed by washing and fluorescence scanning to yield a fluorescence image of the surface.

Figure 2 shows the experimental results obtained using both strategies described above. Control experiments were performed with synthetic targets corresponding to the c-allele, the t-allele, or a 1:1 mixture of both, giving the appropriate results in each case. The unamplified human genomic DNA sample at a concentration

of 1 pM gave an increased surface fluorescence only for the c-allele probe, indicating a homozygous c-allele genotype with both strategies. This result is consistent with the results obtained using a standard two-step solution invasive cleavage reaction. The lower fluorescence signal intensity observed on the co-immobilized surface in strategy II is due to the lower amount of probe available on the surface as compared to that on the probe-only immobilized surface.[6] Twenty micrograms (10 amol) of human genomic DNA was employed for this experiment, which can be extracted from an approximately 2 mL sample of whole blood. An approximate concentration detection limit of 10 fM target DNA was determined (data not shown). The results demonstrate the high degree of specificity characteristic of the invasive cleavage reaction and also show that this specificity is retained in the subsequent ligation and RCA steps. The concentration of target DNA which was detected here is approximately 3 orders of magnitude lower than that which was detectable in our previous work which employed detection by a fluorescence resonance energy transfer (FRET) mechanism.[6]

The present work provides a proof-of-principle demonstration of the direct analysis of single nucleotide polymorphisms from human genomic DNA in a surface array format. Two strategies were studied; although the co-immobilization strategy gives a lower fluorescence signal, it is a more practical format for large-scale genotyping on DNA arrays. Two stages of surface amplification were employed to provide the requisite sensitivity, the first provided by the cycling of the target DNA in the surface invasive cleavage reaction, and the second provided by the RCA amplification. In future work, this chemistry will be adapted for use on complex DNA arrays containing many hundreds of thousands of elements, each encoding a single SNP allele. The ability to rapidly and inexpensively analyze human genetic variation on such a very large scale will open manifold new opportunities for biological and medical research.

Supporting Information Available: Oligonucleotide sequences and experimental details (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) White, K. P. *Nature Rev. Genet.* **2001**, *2*, 528−537.
(2) Kruglyak, L. *Nat. Genet.* **1997**, *17*, 21−24.
(3) Abbott, A. *Nature* **2000**, *406*, 340−342.
(4) The SNP Consortium http://snp.cshl.org.
(5) Melton, L. *Nature* **2003**, *422*, 917−923.
(6) Lu, M. C.; Hall, J. G.; Shortreed, M. R.; Wang, L. M.; Berggren, W. T.; Stevens, P. W.; Kelso, D, M.; Lyamichev, V.; Neri, B.; Skinner, J. L.; Smith, L. M. *J. Am. Chem. Soc.* **2002**, *124*, 7924−7931.
(7) Lyamichev, V.; Mast, A. L.; Hall J. G.; Prudent, J. R.; Kaiser, M. W.; Takova, T.; Kwiatkowski, R. W.; Sander, T. J.; Arruda, M. D.; Arco, D. A.; Neri, B. P.; Brow, M. A. D. *Nat. Biotechnol.* **1999**, *17*, 292−296.
(8) 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.
(9) Rao, K. V. N.; Stevens, P. W.; Hall, J. G.; Lyamichev, V.; Neri, B. P.; Kelso, D. M. *Nucleic Acids Res.* **2003**, *31*, e66.
(10) Nickerson, D.; Taylor, S.; Fullerton, S.; Weiss, K.; Clark, A.; Stengard, J.; Salomaa, V.; Boerwinkle, E.; Sing, C. *Genome Res.* **2000**, *10*, 1532−1545.
(11) Brockman, J. M.; Frutos, A. G.; Corn, R. M. *J. Am. Chem. Soc.* **1999**, *121*, 8044−8051.
(12) Lizardi, P.; Huang, X.; Zhu, Z.; Bray-Ward, P.; Thomas, D.; Ward, D. *Nat. Genet.* **1998**, *19*, 225−232.

JA039355J