

Sabanayagam, C. R., Berkey, C., Lavi, U., Cantor, C. R., & Smith, C. L. (1999) in *Micro- and Nanofabricated Structures and Devices for Biomedical and Environmental Applications II*, vol. **3606**, ed. Ferrari, M. (Proceedings of The International Society for Optical Engineering, Bellingham, WA), pp. 90-98.

Molecular DNA switches and DNA chips

Chandran R. Sabanayagam^{*a}, Cristin Berkey^a, Uri Lavi^b, Charles R. Cantor^{a,c}, and Cassandra L. Smith^a

^aCenter for Advanced Biotechnology, Department of Engineering, Boston University,
Boston, MA 02215

^bAgricultural Research Organization, The Volcani Center, P.O.B 6 Bet-Dagan, 50250, Israel

^cSequenom, Incorporated, 11555 Sorrento Valley Road, San Diego, CA 92121

ABSTRACT

We present an assay to detect single-nucleotide polymorphisms on a chip using molecular DNA switches and isothermal rolling-circle amplification. The basic principle behind the switch is an allele-specific oligonucleotide circularization, mediated by DNA ligase. A DNA switch is closed when perfect hybridization between the probe oligonucleotide and target DNA allows ligase to covalently circularize the probe. Mismatches around the ligation site prevent probe circularization, resulting in an open switch. DNA polymerase is then used to preferentially amplify the closed switches, via rolling-circle amplification. The stringency of the molecular switches yields 10^2 - 10^3 fold discrimination between matched and mismatched sequences.

Keywords: DNA chip, DNA array, genome sensor, point mutation, mismatch, SNP, genotyping, rolling-circle amplification, isothermal amplification, DNA ligase

1. INTRODUCTION

The relatively new fields of molecular biology and biotechnology have focused our knowledge of human diseases at the molecular and genetic levels. It is estimated that the human genome encodes anywhere from 60,000-100,000 different genes, and certain mutations lead to dysfunctional proteins, giving rise to diseases. Therefore, a major goal of molecular biology and biotechnology is the fast evaluation of DNA sequences to ascertain genetic polymorphisms. In particular, mutation detection is usually focused on very specific sequencing, or minisequencing, of a particular gene locus. Genetic mutations can have several forms, including deletions, insertions, inversions, translocations and point mutations. A mutation assay which can identify all these types of variations, simultaneously, is strongly desirable.

Recently, much effort has been applied to point mutation detection, because it is the most wide spread genetic variation. By definition, a single-nucleotide polymorphism (SNP) is a nucleotide position in which an alternate base occurs with a frequency of at least 1%. An estimated 10^7 SNPs are currently present in the human genome. SNP frequency in humans is estimated to be 1:500-1,000 bp. They are excellent candidates for genetic markers and are likely to be associated with genetic diseases. High-density oligonucleotide arrays have been recently used to detect single-base substitutions in target DNA¹⁻⁵. Typically, the oligonucleotide probes are between 8-25 bases long, and targets consist of restriction enzyme digests of genomic DNA, PCR amplicons, or cDNAs. These arrays are based on the sequence-specific hybridization between two complementary strands of DNA. Mismatches between the probes and targets prevent efficient hybridization; and the genotype of the target DNA is elucidated from the hybridization patterns on the chip. However, these hybridization arrays require a redundant set of overlapping probes to successfully detect SNPs, because single-base mismatches between a target-probe complex can yield significant signals only 2-8 fold lower than those seen with perfect duplexes. The use of ligase for point mutation analysis on DNA arrays has considerable advantages over competitive hybridization methods alone. Because the efficiency of ligase is dramatically reduced in the context of mismatches, ligation allows for sensitive post-hybridization

* Correspondence: Telephone: 617 353 8505; Fax: 617 353 8501; Email: chandran@bu.edu

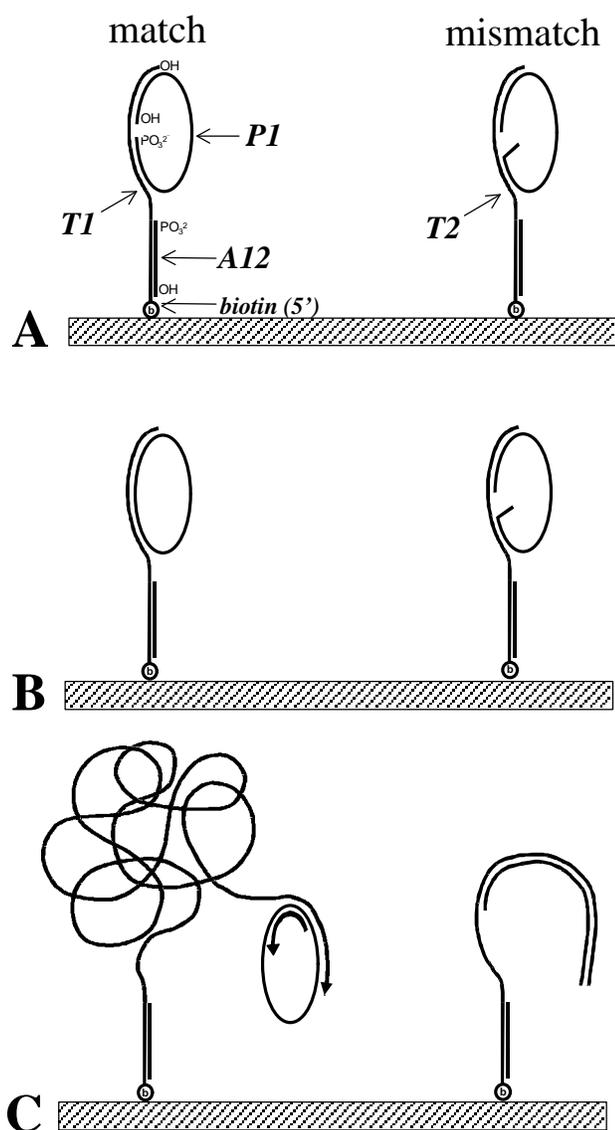


Fig. 1 Scheme for SNP analysis on surface-bound target DNA. (A) Targets oligonucleotides, T1 and T2, differ by a single nucleotide at position 38, a G or a T, respectively (see Fig. 2 for sequences). T1 and T2 are 5'-biotinylated so that they can be immobilized on streptavidin supports. A12 is an 18-base oligonucleotide which provides structural support for the targets. P1 is an oligonucleotide probe which can circularize on the targets. The P1/T1 complex forms 30 contiguous base pairs (left); a C:T mismatch is present at the 5'-terminus of P1 and nucleotide position 38 of T2 (right). (B) Proofreading by T4 DNA ligase. P1 contains a 5'-phosphate, incorporated during synthesis. Ligation is successful only if the P1 termini are perfectly hybridized to the template. (C) Signal amplification by DNA polymerase. Circularized P1 probes are used as templates for RCA (left), whereas only partial extension can occur for non-ligated probes (right).

proofreading⁶. Ligation products are typically 15-100 bases in length and may be linear or circular DNA. Circularized DNA has the advantage that it can be amplified specifically by rolling-circle amplification (RCA)⁷⁻¹¹, generating many tandemly-repeated copies of the original template, bound to the target DNA.

In this paper we describe a scheme for SNP analysis based on RCA on DNA arrays. Successful SNP detection can be performed by three coupled steps: sequence recognition (hybridization), proofreading and signal amplification. In our model system, an oligonucleotide probe, P1, was designed to circularize via hybridization on two immobilized templates, T1 and T2 (Fig 1A). T1 and T2 differ by only a single nucleotide such that the P1/T1 complex forms 30 contiguous base pairs, while the P1/T2 complex contains a C:T mismatch at the 5'-terminus of P1. This sequence recognition step is similar to hybridization on arrays, and it is difficult to distinguish between the P1/T1 and P1/T2 complexes because, in general, 3'- and 5'-end mismatches do not greatly affect duplex stability. The proofreading step is mediated by DNA ligase, because this enzyme is sensitive to single-base mismatches. In the context of perfect hybridization around the ligation point, the 3'-hydroxyl and 5'-phosphate of P1 can be successfully joined to form a covalently-closed circular probe (Fig. 1B). Thus, DNA ligase can be used as a proofreading enzyme to change the topological structure of the probe in an all-or-nothing type reaction. Finally, signal amplification is achieved by DNA polymerase in an extension reaction primed by the 3'-hydroxyl of the templates. In the case of circular probes, extension occurs via RCA. In contrast, only partial extension is possible with linear probes (Fig. 1C). Using isotopically labeled nucleotides during the extension reaction, we typically observe 10^2 – 10^3 fold more nucleotide incorporation with P1/T1 than with 5'- or 3'-mismatched probe/target complexes.

2. MATERIALS AND METHODS

2.1. Probe Circularization

The sequences and base pairing of a P1/T1 complex are shown in Fig. 2. P1 circularization was performed in solution by combining equimolar ratios of P1 and T1 to a final concentration of 5 μ M in TE (10 mM Tris-Cl, 1 mM EDTA) supplemented with 0.1M NaCl, heating the mixture to 95°C, followed by slowly cooling to room temperature. Ten pmol of the P1/T1 complex and 400 U T4 DNA ligase (New England Biolabs) were suspended in 50 μ l of T4 DNA ligase buffer supplied by the manufacturer. The cocktail was incubated at 16°C for 30 min in a PCR thermocycler, and the reaction was terminated by heating the samples to 75°C for 15 min. Non-ligated controls were made by omitting ligase from the above reaction.

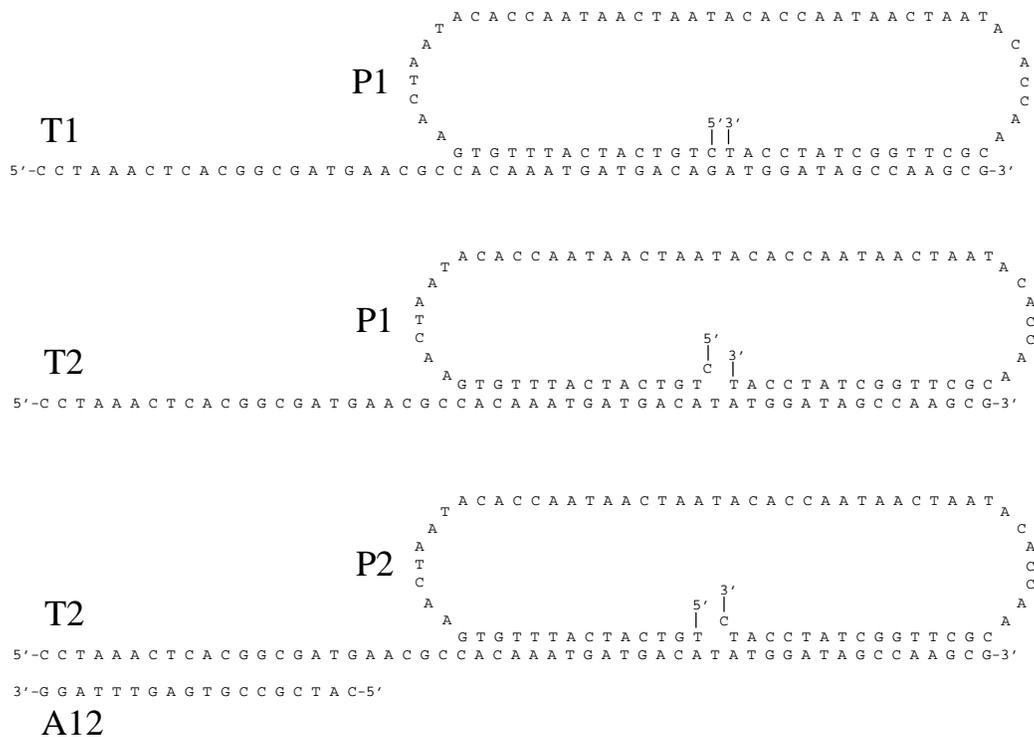


Fig. 2 Oligonucleotide sequences used for molecular DNA switches. The 74-base probe oligonucleotides, P1 and P2 were purchased with a 5'-phosphate, PAGE purified from Bio-Synthesis, Inc. The 53-base targets, T1 and T2 were purchased with a 5'-biotin modification, HPLC purified from Operon Technologies. The 5'-15 bases, and 3'-15 bases of P1 are complementary to positions 23-38, and 39-44 of T1, respectively. The P1/T2 complex contained a C:T mismatch at the 5'-terminus of P1; the P2/T2 complex is similar in sequence to P1/T2, except that the C:T mismatch occurs at the 3'-terminus of P2. The oligonucleotide, A12 (Operon Technologies), is complementary to the 5'-end of T1 and T2, and is used for structural support in chip-based experiments.

2.2. Rolling Circle Amplification in Solution

The RCA cocktail contained 1 mM dCTP, 1 mM dGTP, 1mM dATP, 0.1 mM dTTP, 0.2 $\mu\text{Ci}/\mu\text{l}$ [$\alpha\text{-}^{32}\text{P}$] dTTP (3,000 Ci/mmol; Amersham), and either 8.5 ng/ μl Φ29 DNA polymerase (a gift from John Cunniff, Amersham) or 0.25 U/ μl *E. coli* DNA polymerase I (Boehringer Mannheim) suspended in DNA polymerase I buffer (New England Biolabs). RCA was carried out with 3.3 pmol of ligated or non-ligated P1/T1 complexes suspended in 20 μl RCA cocktail. The mixture was incubated for the indicated times from 15 min to 24 hr at 37°C, pending on the experiment. RCA was terminated by heating to 95°C for 15 min.

2.3. Fabrication of Streptavidin-coated Microwells

Microwells were etched in <100> silicon according to standard practices¹². A 3" diameter silicon wafer contained approximately 250 microwells, with dimensions 2 mm X 2 mm X 200 μm deep, and spaced 2 mm apart. Biotin was discretely patterned inside the microwells using photolithography and has been previously described in detail¹³. Briefly, the micromachined wafers were silanized and an ethanolic solution of photoactivatable biotin (Pierce) was deposited on the wafers and allowed to evaporate. Photoactivatable biotin forms covalent bonds with nearby organic moieties upon exposure to UV light. A photomask was placed over the wafer such that only the microwells were exposed to UV light. The irradiated wafer was washed and then incubated with streptavidin, which only binds inside the biotinylated microwells. Because of streptavidin's tetrameric structure, two biotin-binding sites are used to immobilize the protein, leaving the remaining sites available to bind biotinylated oligonucleotides. Each well has a 500 nl volume.

2.4. Target Immobilization and RCA on DNA Chips

The 18-base oligonucleotide, A12, was complementary to the first 5'-18 bases of T1 and T2, and served as a structural support to lift the targets from the surface. Equimolar ratios of T1/A12 and T2/A12 were suspended at a final concentration of 5 μ M in TE supplemented with 0.1M NaCl, heated to 95°C, followed by slowly cooling to room temperature. Five hundred nl (1250 fmol) of both DNA complexes were spotted inside streptavidin-coated microwells. The chip was saturated with 50 μ M biotin in SPE buffer (0.1M Na-phosphate, pH 6.6, 1M NaCl), then washed 4 X 15 min in SPE buffer at 37°C, rinsed briefly with deionized H₂O, and air dried. A hybridization slide chamber (CoverWell PC50, Grace Bio-labs) was placed over the chip and secured with small paper binding clips. Two small holes were punctured in the hybridization chamber and served as fluid inlet and outlet ports. Probe hybridization was performed by injecting 85 μ l of the indicated probe solution suspended in TE supplemented with 0.1M NaCl. The injection ports were sealed, and the chip was incubated for 12 hr at 37°C. The chip was washed 4 X 15 min in SPE buffer at 37°C, followed by a rinse in deionized H₂O. A new hybridization chamber was placed on the chip and 85 μ l of ligation cocktail (see Fig. 4) was injected onto the chip and incubated for 30 min. The chip was washed 4 X 15 min in SPE buffer at 37°C, rinsed in deionized H₂O, and DNA polymerase cocktail was added as indicated in Fig. 4. The chip was washed 4 X 15 min in SPE buffer before exposing to phosphorimaging plates. A Molecular Dynamics STORM imager was used to scan the plates, and analysis was performed using software provided by the manufacturer.

3. RESULTS AND DISCUSSION

3.1. Probe Circularization and RCA in Solution

Both circular and linear P1 probes were used as templates for RCA. P1 was circularized on T1 templates using T4 DNA ligase. Ligation products were analyzed on 12% denaturing polyacrylamide gels (data not shown); circularized P1 product was observed as a supershifted band from the linear form. Approximately 60% of the P1 probes were circularized when equimolar ratios of P1 and T1 were ligated for 30 min at 16°C. Aliquots of either linear or circular P1 were used as templates for RCA reactions. The kinetics of RCA was explored by different incubation times of the P1/T1 complexes with *E. coli* DNA polymerase I or Φ 29 DNA polymerase. Figure 3A shows the resulting amplicons resolved on a 0.8% agarose gel. The gel was stained with SYBR Green I (FMC), a double-strand specific nucleic acid dye. The RCA amplicons generated from DNA polymerase I and Φ 29 DNA polymerase showed remarkable differences. Note that circularized P1 RCA products increase over time, for incubation periods up to 24 hr, consistent with the observation of others^{8,9}. Surprisingly, DNA polymerase I also generated amplicons with linear P1/T1 complexes after long incubation periods (between 6 and 24 hr).

3.1.1. Difference in RCA amplicons: single-stranded and double-stranded forms

Three observations led us to conclude that RCA mediated by Φ 29 DNA polymerase produced single-stranded (ss) amplicons while DNA polymerase I produces double-stranded (ds) amplicons. First, the amplicons generated by the different polymerases showed strong differences in their electrophoretic mobilities. Second, amplicons generated by DNA polymerase I showed a much greater increase in SYBR Green I fluorescence than the Φ 29 products. These results can be interpreted as differences in the fluorescence yield when the dye is bound to dsDNA versus ssDNA. To test the ss versus ds hypothesis, we used isotopically labeled dNTPs to assess the amount of amplicons produced by the different polymerases, and found that DNA polymerase I produced fewer polymers than did Φ 29 DNA polymerase (data not shown). Lastly, a 14-base oligonucleotide, with sequence similarity to P1 and containing a 5'-fluorescein modification, was used to probe RCA amplicons which had been UV immobilized on nitrocellulose membranes (Fig. 3B). The results show an increase in probe hybridization with Φ 29- versus DNA polymerase I-generated amplicons, further supporting the hypothesis that RCA mediated by Φ 29 polymerase-generated ssDNA, while DNA polymerase I produced dsDNA. It should also be noted that earlier work on RCA by Fire and Xu demonstrated that DNA polymerase I produced dsDNA amplicons, with only one primer and circularized template¹⁰. RCA amplicons generated by Φ 29 DNA polymerase are of particular interest in DNA array applications, because of their ability to hybridize to short oligonucleotide tags.

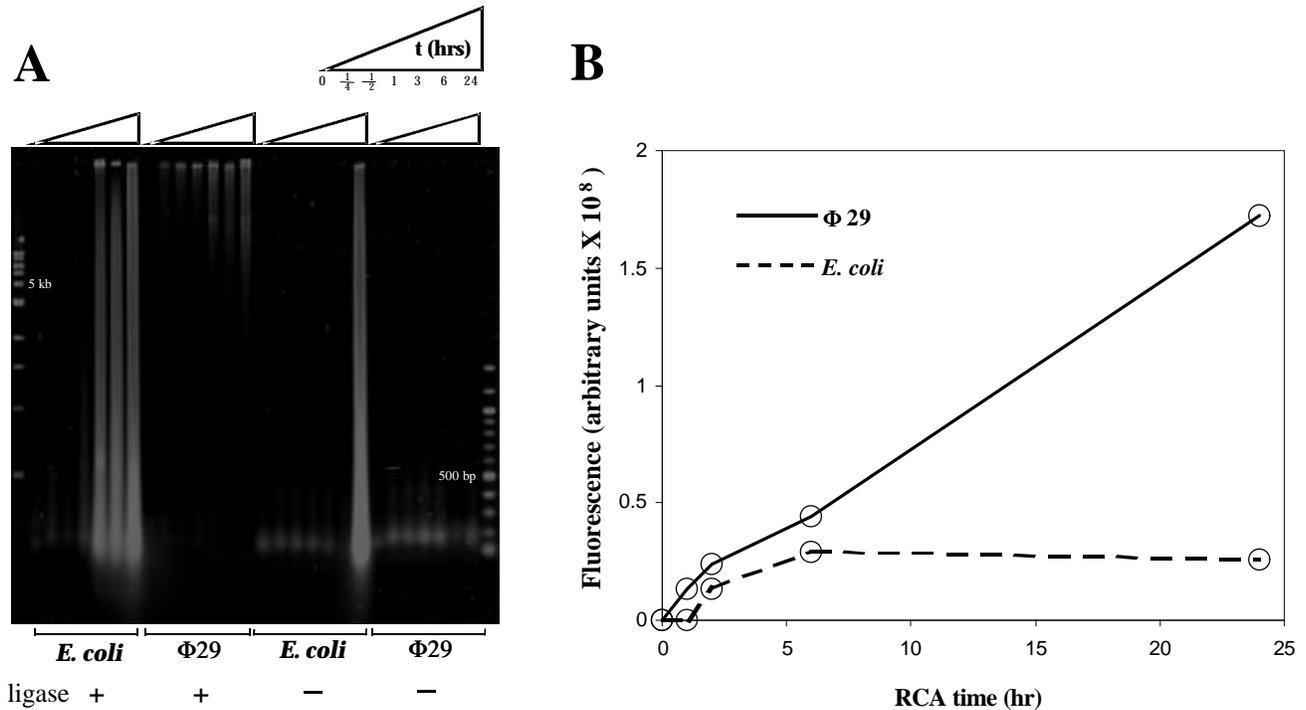


Fig. 3 Kinetics of RCA produced by $\Phi 29$ or *E. coli* DNA polymerases. (A) RCA amplicons resolved on a 0.8% agarose gel, stained with SYBR Green I. SYBR Green I is a DNA intercalating dye which can stain both double-stranded and single-stranded DNAs, but shows enhanced fluorescence when bound to duplexes. The incubation times were 0, $\frac{1}{4}$, $\frac{1}{2}$, 1, 3, 6 and 24 hr. One kb and 100 bp makers were loaded in the first and last lanes, respectively. RCA was performed using ligated (+) or non-ligated (-) P1 probe templates as indicated at the bottom of the image, and T1 primers. (B) Hybridization of 5'-fluorescein-labeled (F) oligonucleotide (5'-FAACTAATACACCAA) to RCA amplicons immobilized on nitrocellulose membranes by UV crosslinking. Hybridization was carried out overnight at 25°C in 6 X SSC buffer, followed by a 15 min wash in 6 X SSC, and a 15 min wash in 2 X SSC. Fluorescence intensity was determined by scanning the membranes with a Molecular Dynamics STORM imager. The x-axis represents RCA incubation times, and the y-axis is fluorescence.

3.2. SNP Detection on DNA Chips

The ability of $\Phi 29$ to generate significant amounts of RCA products, specifically on circularized templates, prompted us to explore its ability to amplify surface-bound target DNAs. Two biotinylated target oligonucleotides, T1 and T2, differing only at position 38, were immobilized inside streptavidin-coated microwells. The 5'-terminus of the P1 probe aligns with target positions 38, forming a C:G basepair and a C:T mismatch with T1 and T2, respectively. The P2/T2 complex contained a C:T mismatch at the 3'-terminus of P2 and position 38 of T2. Chip-based SNP detection was performed in three steps. The chips were incubated overnight with the probes in a solution of TE supplemented with 0.1M NaCl, and washed 4 X 15 min with SPE buffer to remove non-hybridized probes. Ligation was performed with T4 DNA ligase for 30 min at 37°C, followed by a wash. Chips were then incubated with DNA polymerase in the presence of dNTPs and [α -³²P] dTTP for 12 hr at 31°C. The chip was washed to remove unincorporated nucleotides, before radioactive imaging. Line profiles were determined using ImageQuant software (Molecular Dynamics). In general, even the mismatched probe/target complexes gave significant signals, and optimal conditions for SNP discrimination were determined by using different polymerases, varying the probe concentrations, and adjusting the ligation conditions. Figures 4A-C show autoradiograms of the chips with line profiles above the images. The specific reaction conditions are given in the figure legends.

As shown in Fig. 4A, $\Phi 29$ DNA polymerase was more stringent and efficient in amplifying closed switches (i.e., circularized probes) than *E. coli* DNA polymerase I. $\Phi 29$ DNA polymerase was thus our choice polymerase for RCA. Probe hybridization concentrations also affected circularization and RCA efficiencies (Fig. 4B). When probe concentrations were lowered from 2 μ M to 0.2 μ M, more amplicons were produced. The higher probe concentrations most likely resulted in

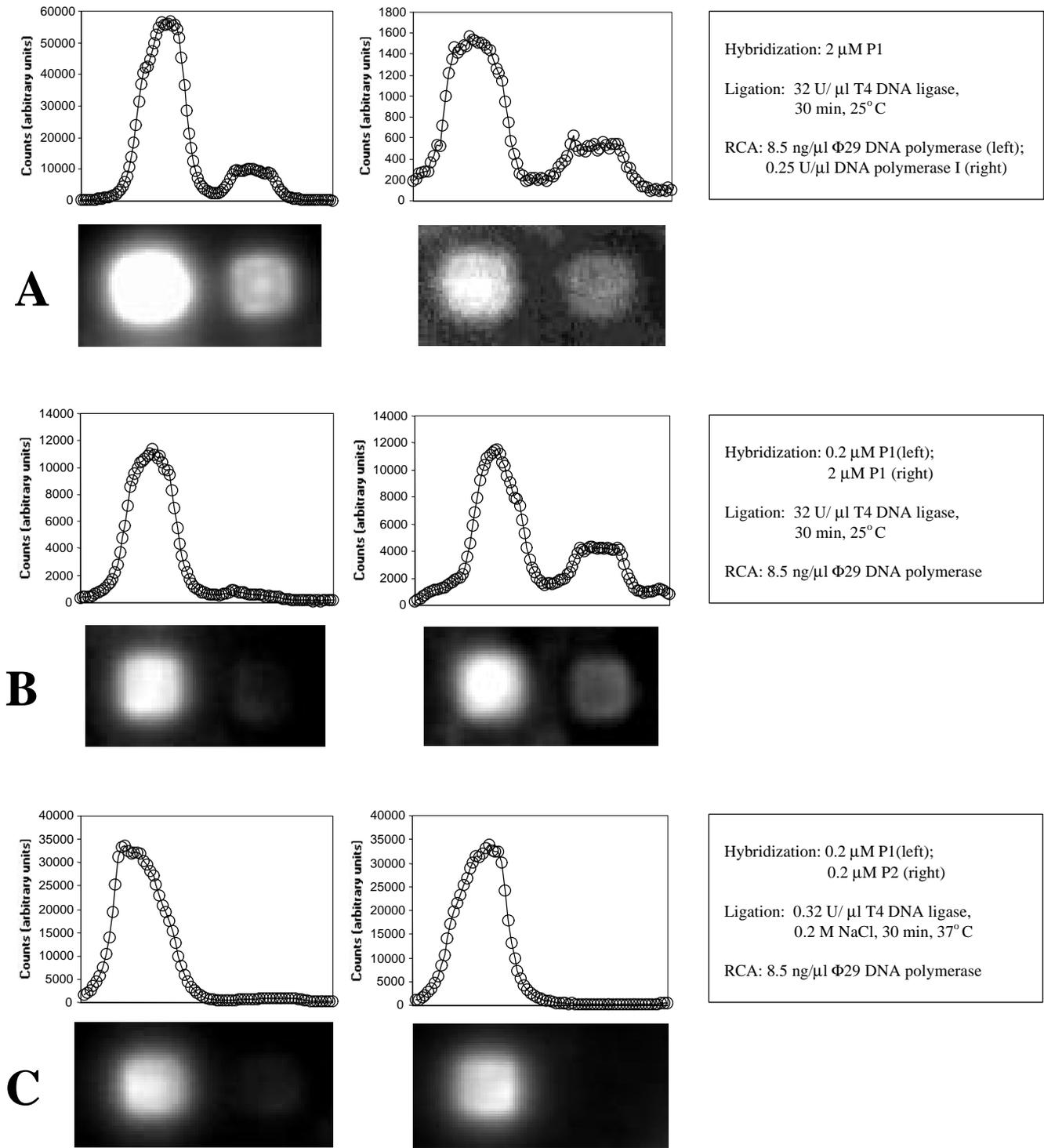


Fig. 4 Chip-based SNP analysis. Each image shows an autoradiogram of two wells containing immobilized T1 (left) and T2 targets (right). A line profile is given above the images. Reaction conditions are shown to the right of the panels. (A) RCA mediated by Φ 29 (left panel) or *E. coli* (right panel) DNA polymerase. (B) Effects of probe concentration during hybridization. (C) Optimized conditions for SNP discrimination.

undesired ligation products, namely probe dimers and trimers, in their linear forms. Optimization of ligation conditions led to a decrease in mismatch signals to background levels (Fig. 4C). It has been shown that the presence of NaCl or spermidine increases mismatch discrimination by T4 DNA ligase^{6,7,14}. The ligation conditions which produced the best discrimination between matched and mismatched probes consisted of 4U/ μ l T4 DNA ligase, 0.2 M NaCl, and incubation at 37° C for 30 min. Also, moving the mismatch from the 5'- to the 3'-terminus improved ligase discrimination by 5- to 8-fold, as documented by others^{7,14}. Under our most stringent conditions, the two target oligonucleotides could be distinguished with a signal-to-noise ratio of $\sim 10^3$.

4. CONCLUSIONS

The use of enzymatic reactions on DNA chips can greatly enhance sequence identification, with single-nucleotide resolution. We have presented a chip model for SNP analysis based on molecular DNA switches. These switches couple DNA hybridization, ligation and isothermal RCA to yield signal-to-noise ratios of 10^2 - 10^3 . This high ratio allows SNP analysis using a single switch, in contrast to the redundant set of overlapping probes required on DNA arrays based solely on hybridization. Thus, the system is particularly useful in screening numerous target DNAs at a polymorphic site, and should have utility in a variety of applications when sequence information is available. The system can be applied to the screening of known mutations (such as known genetic diseases and polymorphisms in agriculturally interested genes). In addition, the system can be used to genotype SNPs for forensic identification and linkage analysis. Although we presented a model for SNP analysis, DNA switches can also identify sequence deletions, insertions and inversions.

Another potential application for RCA of immobilized oligonucleotides is during the manufacturing process for DNA arrays. Format 2 hybridization arrays¹⁵ thus far consist of densely packed, monolayered probes. Concatemers of the immobilized probes can be synthesized directly on the solid support via RCA, thereby increasing the number of probe hybridization sites by at least two orders of magnitude. Extension of the probes from the support should lessen the problems associated with steric hindrance of hybridization near the surface-solution interface¹⁶.

ACKNOWLEDGEMENTS

We are grateful to John Cunniff of Amersham Life Technologies for gifts of Φ 29 DNA polymerase. Drs. Weichun Yuan and Jan Smits kindly fabricated the silicon microwells. This work was supported in part by a fellowship to C. Sabanayagam by the U.S. National Human Genome Research Institute (T32 HG00041-03), and by Sequemon, Incorporated. C. Berkey was funded by the National Science Foundation, Research Experience for Undergraduates program. U. Lavi was on sabbatical leave at the Center for Advanced Biotechnology at Boston University.

REFERENCES

1. D. G. Wang, J. Fan, C. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, R. Lipshutz, M. Chee, and E. S. Lander, "Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome," *Science* **280**, pp. 1077-1082, 1998.
2. J. G. Hacia, L. C. Brody, M. S. Chee, S. P. A. Fodor, and F. S. Collins, "Detection of heterozygous mutations in *BRCA1* using high density oligonucleotide arrays and two-colour fluorescence analysis," *Nature Genetics* **14**, pp. 441-447, 1996.
3. M. J. Kozal, N. Shah, N. Shen, R. Yang, R. Fucini, T. C. Merigan, D. D. Richman, D. Morris, E. Hubbell, M. Chee, and T. R. Gingeras, "Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays," *Nature Medicine* **2**, pp. 753-759, 1996.
4. Z. Guo, R. A. Guilfoyle, A. J. Thiel, R. Wang, and L. M. Smith, "Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports," *Nucleic Acids Res.* **22**, pp. 5456-5465, 1994.
5. A. C. Pease, D. Solas, E. J. Sullivan, M. T. Cronin, C. P. Holmes, and S. P. Fodor, "Light-directed oligonucleotide arrays for rapid DNA sequence analysis," *Proc. Natl. Acad. Sci. USA* **91**, pp.5022-5026, 1994.
6. N. E. Broude, T. Sano, C. L. Smith, and C. R. Cantor, "Enhanced DNA sequencing by hybridization," *Proc. Natl. Acad. Sci. USA* **91**, pp. 3072-3076, 1994.

7. U. Landegren, R. Kaiser, J. Sanders, and L. Hood, "A ligase-mediated gene detection technique," *Science* **241**, pp. 1077-1080, 1988.
8. J. Baner, M. Nilsson, M. Mendel-Hartvig, and U. Landegren, "Signal amplification of padlock probes by rolling circle replication," *Nucleic Acids Res.* **26**, pp. 5073-5078, 1998.
9. P. M. Lizardi, X. Huang, Z. Zhu, P. Bray-Ward, D. C. Thomas, and D. C. Ward, "Mutation detection and single-molecule counting using isothermal rolling-circle amplification," *Nature Genetics* **19**, pp. 225-232, 1998.
10. A. Fire and S. Xu, "Rolling replication of short DNA circles," *Proc. Natl. Acad. Sci. USA* **92**, pp. 4641-4645, 1995.
11. A. Hatch, T. Sano, J. Misasi, and C. L. Smith, "Rolling circle amplification of DNA immobilized on solid surfaces and its application to multiplex detection," *Genetic Analysis* **15**, pp. 35-40, 1999.
12. W. C. Choi and J. G. Smits, "A method to etch undoped silicon cantilever beams," *Journal of Microelectromechanical Systems* **2**, pp. 82-86, 1993.
13. C. R. Sabanayagam and C. R. Cantor, "Micropatterned streptavidin surfaces," submitted to *Nucleic Acids Res.*, Apr., 1999.
14. D. Y. Wu and R. B. Wallace, "Specificity of the nick-closing activity of bacteriophage T4 DNA ligase," *Gene* **76**, pp. 245-254, 1989.
15. C. R. Cantor, A. Mirzabekov, and E. Southern, "Report on the sequencing by hybridization workshop," *Genomics* **13**, pp. 1378-1383, 1992.
16. M. S. Shchepinov, S. C. Case-Green, and E. M. Southern, "Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays," *Nucleic Acids Res.* **25**, pp.1155-1161, 1997.