Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2269–2278

BIOORGANIC & MEDICINAL CHEMISTRY

APEX Disease Gene Resequencing: Mutations in Exon 7 of the p53 Tumor Suppressor Gene

John M. Shumaker,[a] Jeffrey J. Tollet,[a] Kellie J. Filbin,[a] Michael P. Montague-Smith[b] and Michael C. Pirrung[b,*]

[a]*Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA*
[b]*Department of Chemistry, Levine Science Research Center, Box 90317, Duke University, Durham, NC 27708-0317, USA*

Received 1 February 2001; accepted 21 March 2001

This paper is dedicated to Professor Peter B. Dervan in honor of his receipt of the 2000 Tetrahedron Prize and in appreciation of his pioneering efforts in integrating organic and nucleic acid chemistry.

Abstract—Detection of mutations in disease genes will be a significant application of genomic research. Methods for detecting mutations at the single nucleotide level are required in highly mutated genes such as the tumor suppressor p53. Resequencing of an individual patient's DNA by conventional Sanger methods is impractical, calling for novel methods for sequence analysis. Toward this end, an arrayed primer extension (APEX) method for identifying sequence alterations in primary DNA structure was developed. A two-dimensional array of immobilized primers (DNA chip) was fabricated to scan p53 exon 7 by single bases. Primers were immobilized with 200 μm spacing on a glass support. Oligonucleotide templates of length 72 were used to study individual APEX resequencing reactions. A template-dependent DNA polymerase extension was performed on the chip using fluorescein-labeled dideoxynucleotides (ddNTPs). Labeled primers were evanescently excited and the induced fluorescence was imaged by CCD. The average signal-to-noise ratio (S/N) observed was 30:1. Software was developed to analyze high-density DNA chips for sequence alterations. Deletion, insertion, and substitution mutations were detected. APEX can be used to scan for any mutation (up to two-base insertions) in a known region of DNA by fabricating a DNA chip comprising complementary primers addressing each nucleotide in the wild-type sequence. Since APEX is a parallel method for determining DNA sequence, the time required to assay a region is independent of its length. APEX has a high level of accuracy, is sequence-based, and can be miniaturized to analyze a large DNA region with minimal reagents.

Introduction

The increasing awareness that many human diseases have a genetic basis[1–3] has illuminated a need for rapid and accurate methodologies for detection and diagnosis of the informational errors that underlie such diseases. Consequently, many novel techniques for mutation detection have been developed. Current mutation detection methods have limitations in the efficient detection of *all* mutations. Single strand conformation polymorphism (SSCP),[4] denaturing gradient gel electrophoresis (DGGE),[5] and direct sequencing[6] all require a gel separation step for mutation identification. Furthermore, SSCP and DGGE require gel sequencing for specific mutation characterization. Hybridization-based analyses[7,8] eliminate the need for gel separation, but are limited to the identification of characterized mutations. Mutation analysis in the twenty-first century will require scanning an entire gene and identifying any type of mutation within the region of interest.

The completion of the sequencing of a human genome only augments interest in mutation detection in clinical diagnostic samples from individuals. Diseases in which the gene has a wide range of known mutations (and perhaps more lying undiscovered in latent populations) are a particular challenge. Whereas design of a genetic diagnostic test for a prevalent mutation might be readily accomplished using a number of different modalities, with highly mutated disease genes, it is essential that complete gene sequence information be obtained from each patient for comparison with the wild-type sequence. Posing an even greater challenge, significantly exceeding that of genomic sequencing, the source DNA may be derived from heterogeneous cell populations as well as being heterozygotic, meaning that high sensitivity, high dynamic range sequence detection technologies are required. Despite the success of gel-based Sanger chain-termination sequencing in genomic analysis, it is

*Corresponding author. Fax: +1-919-660-1591; e-mail: pirrung@chem.duke.edu

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00104-3

unlikely to offer a practical solution for disease gene comparison sequencing in the clinical setting. Nevertheless, an automated method to detect and score single nucleotide polymorphisms based on conventional sequencing has recently been offered.[9]

The power of the DNA microchip in genetic analysis has become abundantly apparent. Much of this research entails hybridization of labeled target nucleic acid to a chip followed by fluorescence detection. One application which has been made of DNA chips is in resequencing.[10] DNA chips can be generated with probes tailored to loci of interest and used in screening of populations by hybridization.[11,12] Although promising, the fidelity of hybridization-based assays is compromised by the inability to discriminate against end mismatches. The central portion of the oligonucleotide must therefore be designed to probe the DNA locus of interest to maximize the destabilizing effects of a mismatch. Consequently, additional mutations within a region surrounding the mutation site cannot be identified. More importantly, hybridization-based methods are limited to the identification of characterized mutations. Another drawback to hybridization-based analysis is the relatively narrow dynamic range of the signal for perfectly matched and mismatched hybrids. In non-optimal cases, resequencing may be ambiguous and detection of heterozygotes impossible. This shortcoming has called for higher fidelity, higher dynamic range assays for resequencing applications. To reduce background in hybridization assays, both enzymatic[13] and kinetic[14] approaches have been proposed.

Primer extension methods, based on template-dependent extension of DNA primers with a labeled dideoxynucleotide terminator, followed by detection of the label, have been earlier developed for the analysis of single-nucleotide polymorphisms (SNPs, genotyping).[15] These methods do not require gel separation and have been multiplexed in a 96-well microtiter format,[16–18] which is still limited in density and parallelism.

We have been developing,[19–22] as have other groups,[23–26] a novel technique for DNA chip analysis termed Arrayed Primer EXtension (APEX) that offers the potential to meet the foregoing need.[27–29] APEX is characterized by the attachment of DNA primers to a solid surface through their 5′-ends, presenting their 3′-ends for enzymatic processing. Primers bind to their complementary site within an analyte nucleic acid, and a DNA polymerase catalyzes the addition of a labeled dideoxynucleotide terminator. This fluorescence-based APEX method can directly determine sequence alterations in a DNA region of interest. APEX is essentially an array of one-base Sanger sequencing reactions, wherein the product is detected based on its position in the array rather than its mobility in gel electrophoresis. The advantages of APEX, which is marked by a pre-equilibrium step followed by an irreversible, polymerase-catalyzed terminator incorporation (Fig. 1), are worth emphasis. Both steps are dependent on the match between primer and template. Because of the very low error rate and the requirement of the polymerase for a perfect hybrid near the incorporation site (3′ end of the primer),[30] only mutations distal to the 3′-end can contribute to background in APEX. The covalent attachment of the fluorescent tag to priming sites in APEX permits very stringent washing steps. No labeling of analyte nucleic acid is required, and because each template molecule can, in principle, incorporate terminators at multiple priming sites, amplification is an inherent property of the APEX method. APEX also allows information to be gained without labeling or other modification of the analyte DNA.

The design of an APEX array for comparison sequencing is straightforward based on the complement of the gene of interest (Fig. 2). Primer lengths used in this study range from 14 to 18 nt, with the extension of each primer toward its 5′-end adjustable to level the hybridization of probes of different GC content using conventional primer design software to predict T_m. Arrays are produced with oligonucleotides in ca. 100 μm spots on a 200 μm spacing, creating a chip of density 2500 elements/cm^2.

The simplest format in which to consider the results of an APEX experiment is that used here, where one chip is used for each terminator, analogous to four different lanes in Sanger sequencing reactions. A simultaneous four-color detection system can also be used.[31,32] A

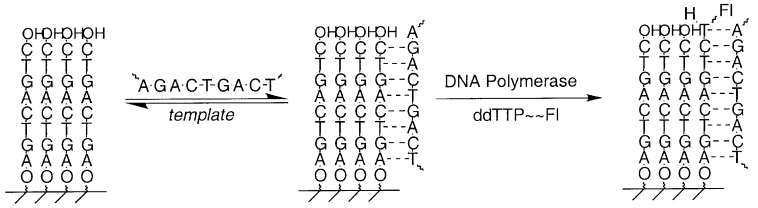

Figure 1. Reaction scheme for APEX. Equilibrium hybridization is followed by DNA polymerase-catalyzed addition of a dye-labeled dideoxynucleotide terminator to perfectly matched primer–template complexes.

```
Template: 3'←CACATTGTCAAGGACGTACCCGCCGTA→5'
Primer 3:  5'-GTGTAACAGTTCCTGCAT-3'          Tm (calc) = 46 ℃
Primer 9:            5'-TCCTGCATGGGCGG-3'    Tm (calc) = 46 ℃
```

Figure 2. Primer design. Two sites in the p53 tumor suppressor gene exon 7 are shown. A 15 nt primer for APEX interrogation of position 2557 would have a much lower T_m than the average for the array, so it was extended by 3 nt to match. A 15 nt primer for APEX interrogation of position 2563 would have a higher T_m than the average for the array, so it was reduced by 1 nt.

mock experiment to assist in demonstrating how APEX provides sequence information is shown in Figure 3. A significant advantage of APEX is that it can detect unknown mutations based upon a known sequence, whereas any hybridization-based assay must design specific probes for each mutation it detects. Deletions, substitutions, and insertions are readily detected, and are confirmed by the effect they have on primer extensions at nearby sequence loci. That is, mutations detected at one locus by an APEX signal different from wild-type are confirmed by the loss of APEX signals from primers that place this mutated locus close to the 3′-end of the primer–template complex. Experience has shown that mutations in the last 3–8 base pairs inhibit extension. This is likely related to specific recognition of the last five nucleotides of the duplex region by the polymerase, as recently demonstrated by crystallography.[33] This 'signal dropout' can cause double mutations spaced closer than ~6 bases to be missed if only one strand is sequenced. However, if both strands are sequenced, up

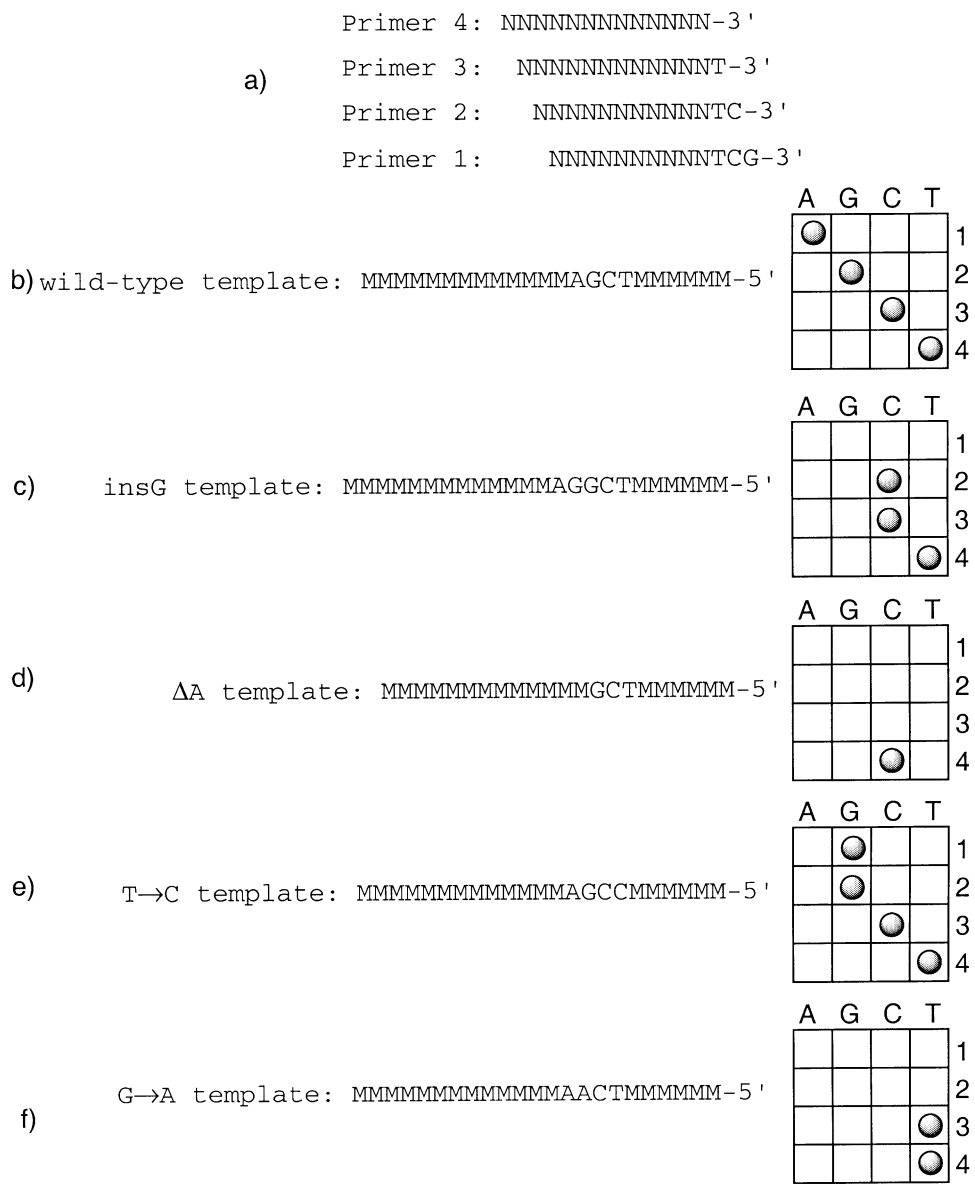

Figure 3. Mock APEX experiments. The effect of mutations observed in this work are shown. (a) Primers 1–4 are designed to address the four explicit bases in the wild-type template. Ns are nucleotides in the primers complementary to Ms in the template. (b) The primers are shown deposited vertically on the chips, and four chips, one for each ddNTP terminator, are used in each APEX. Primer 1 anneals crossing the explicit bases in the wild-type template, which leads to the addition of ddA opposite its T, hence the labeled spot for primer 1 on the A chip. No other primers incorporate A, thus they are unlabeled on the A chip. Likewise, ddG addition occurs only on primer 2, ddC addition only on primer 3, and dT addition only on primer 4. (c) For the insG mutant template, primers 3 and 4 are extended as for the wild-type. Primer 2 is extended on the C chip instead of the G chip because of the additional G. Primer 1 is not extended because of an G/G mismatch at its 3′-end. (d) For the delA mutant template, primer 4 leads to extension on the C chip instead of the T chip because of the G that has been moved opposite the extension site. The deletion causes mismatches at the 3′-ends of primers 1, 2, and 3, so they are not extended on any chip. (e) The T→C mutant template is identical to wild-type at primers 2, 3, and 4. The point mutation is reflected in extension on the G chip and a no extension on the A chip for primer 1. Primers addressing loci further to the 5′-end of the template (the next M region) would not be extended because of an A/C mismatch. (f) The G→A mutant template is identical to wild-type at primer 4. The point mutation is reflected in extension on the T chip and a no extension on the C chip for primer 3. Primers 1 and 2 are not extended because of the A/C mismatch.

to two base insertion mutations (reading one from each strand) can be found with APEX.

This work was undertaken to demonstrate the ability of APEX to detect a variety of sequence variants of the p53 tumor suppressor gene, a prototype of highly mutated disease genes. It focuses on a 37 nt sequence from exon 7 of the p53 gene, which encodes a 383 amino acid protein with a *trans*-activation region at the N-terminus and an oligomerization region at the C-terminus. A point mutation at the splice donor site at the 3′ end of exon 7 of the human p53 gene results in the retention of the intron 7 sequence in the mRNA, thereby inactivating the p53 protein.[34] Mutant p53s are found in inherited and spontaneous cancers and over 50 tumor types. It is the most commonly mutated gene in human tumors; half of patients diagnosed with cancer demonstrate p53 mutations. The p53 status is readily correlated with prognosis in colon, lung, cervical, bladder, prostate, breast, and skin cancers. The majority of known mutations are missense mutations in the center of the coding region. Interestingly, different mutagens result in different mutations in p53, raising the possibility of molecular epidemiology for cancer. Other methods for assessing p53 mutations have been reported.[35]

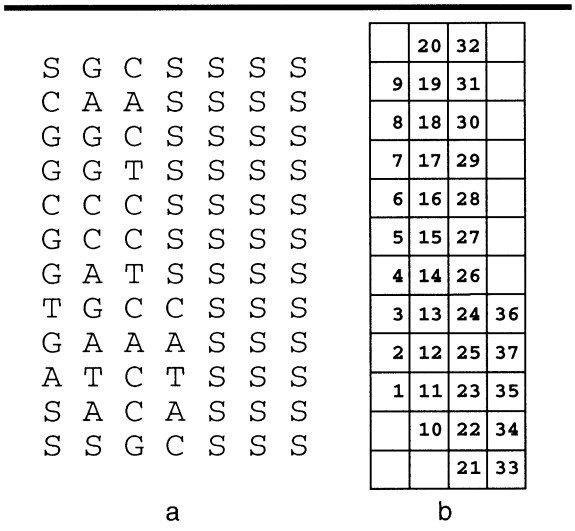

Figure 4. Spotting pattern. (a) The base that extends each primer with wild-type sequence template for the p53 exon 7 chip (7 columns, 12 rows) is given. S is a standard oligonucleotide used for grid alignment and control APEX. (b) Position of oligonucleotide primers listed in Table 1 in the section of the array used for p53.

Results

Chip fabrication and use

Oligonucleotide primers were obtained commercially bearing a 5′-aminolink group. They were spotted in 0.1 M NaOH onto epoxysilanated glass slides by delivery of nL volumes with a syringe robot or a stylus-based spotting system. This protocol produces 100 μm spots with 200 μm spacing, a density of 2500 probes/cm^2, which is sufficient to place all of the primers addressing both sense and anti-sense strands of p53 in one square centimeter. In this case, the primer set was limited to 37 probes covering a region of p53 exon 7. The sequence of the chip is complementary to that of the wild-type DNA template, and corresponds to the predicted sequence in Figure 4a. The templates for the data shown in Figures 5–9 were 72-mer synthetic oligonucleotides, though comparable results have been obtained with PCR products from human DNA samples (J. M. Shumaker, unpublished data). The very simple primer extension protocol uses Sequenase in a standard extension mix spread onto the chip with a cover slip, heated at 42 °C for 5 min with a hot plate. A single fluorescein-labeled ddNTP was used on each chip. The extension mix was removed with water and chips were imaged in fluorescence with evanescent excitation and cooled CCD detection.

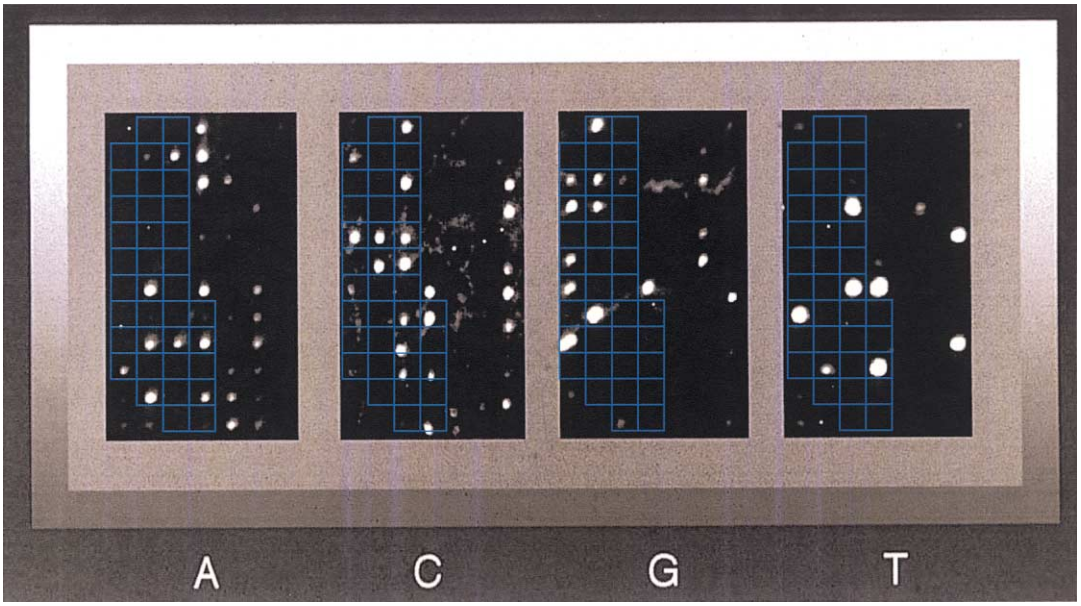

Figure 5. Wild-type. APEX with a 72 nt wild-type p53 exon 7 template. Grids corresponding to the spotting pattern in Figure 4 have been superimposed on the image.

Fluorescence image acquisition

The beam from an argon ion laser was propagated through the glass substrate by total internal reflection, creating an evanescent optical field near the surface of the glass/solution interface.[36–39] The immobilized and fluorescently labeled primers absorb the 488 nm laser light and fluoresce in a band centered at 520 nm. The background scatter of the laser is reduced in total internal reflection fluorescence (TIRF), as only those surface-bound particles in the evanescent field are excited, and a bandpass filter rejects the laser light. A cooled CCD camera was used to image the chip (Fig. 10).

Wild-type DNA

This experiment serves as a baseline for comparison to all other chips with mutant templates. The fluorescence image for the wild-type DNA is shown in Figure 5. The dynamic range of the data is 750 (primer #3) to 12. Calculated noise levels were generally < 25. Most primers give *visibly* readable sequence; only four primers produced a signal less than 100 (8, 19, 22, 25). Primer 19 produced a low signal (ca. 12) on the (correct) A chip, but this value was still much larger than the primer 19 signals on the remaining three chips. The calculated noise levels of three primers (2, 18, 29) were

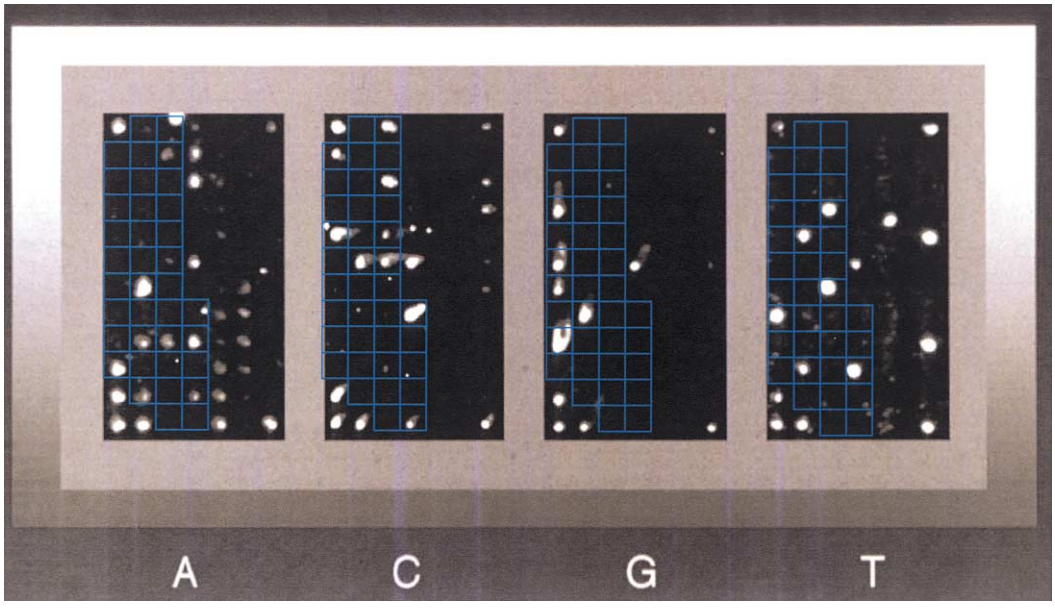

Figure 6. Substitution mutation. APEX with a 72 nt p53 exon 7 template with a G→A mutation at position 2570. A T signal is present in column, 2 row 5 (a change from a C). Following this locus, the chip reads no sequence for five bases due to that A/C mismatch. The first callable base following the mutation is the C in the column 3, row 10. In these arrays, template-specific marker extending oligonucleotides are located at the lower left corner.

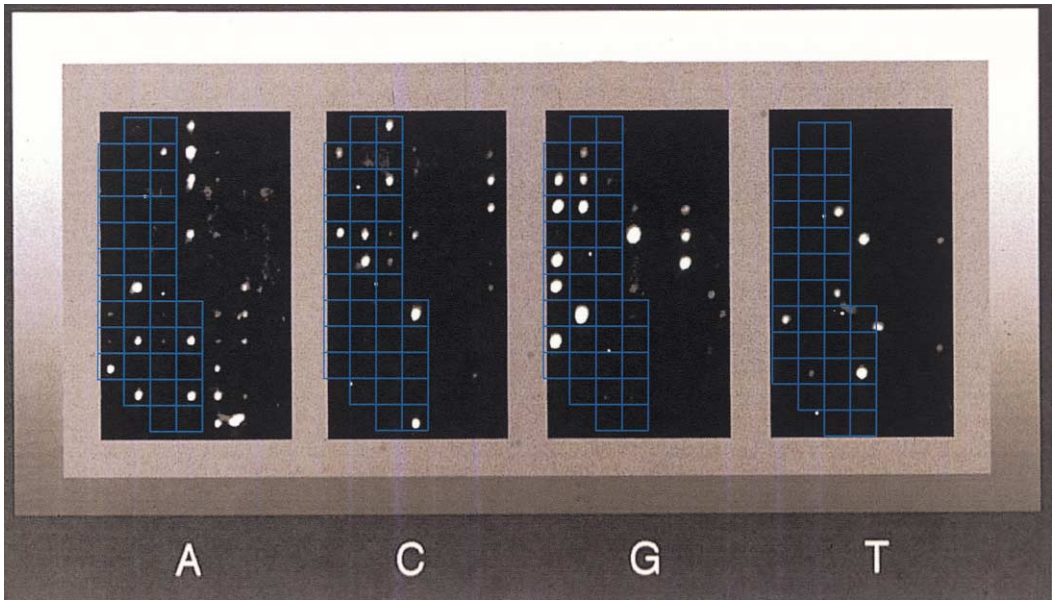

Figure 7. Substitution mutation. APEX with a 72 nt p53 exon 7 template with a T→C mutation at position 2573. A G signal is present in column 2, row 2 (a change from an A).

elevated due to 'fluorescent smearing' from adjacent primers. For example, the fluorescence from the G signal at primer 12 bleeds into the region occupied by primer 2.

Substitution mutation

The fluorescence image for the G→A substitution mutation is shown in Figure 6. The chip-read sequence deviates from wild-type at primer 16 where the wild-type C signal is replaced by a mutant T signal. Because a mutant DNA is probing a wild-type chip, thus producing a mismatch in primers 16–31, the next seven bases following the mutation yield no extension and no sequence information. The template-dependent DNA polymerase discriminates against mismatches at the 3′-end of an annealed duplex and produces a 'drop-out' region. The wild-type sequence read is recovered at the eighth base following the mutation site; in this case, T7 DNA polymerase can extend an annealed complex that has a mismatch eight bases from the 3′ end. Similar results are seen for the T→C mutation (Fig. 7).

Deletion mutation

The chip images for the ΔA mutation are shown in Figure 8. In this case, a mutant C signal replaces the wild-type T signal at primer 29. This is consistent with the sequence given in Table 1 of the ΔA template: a G follows a G (instead of an A following a G), producing

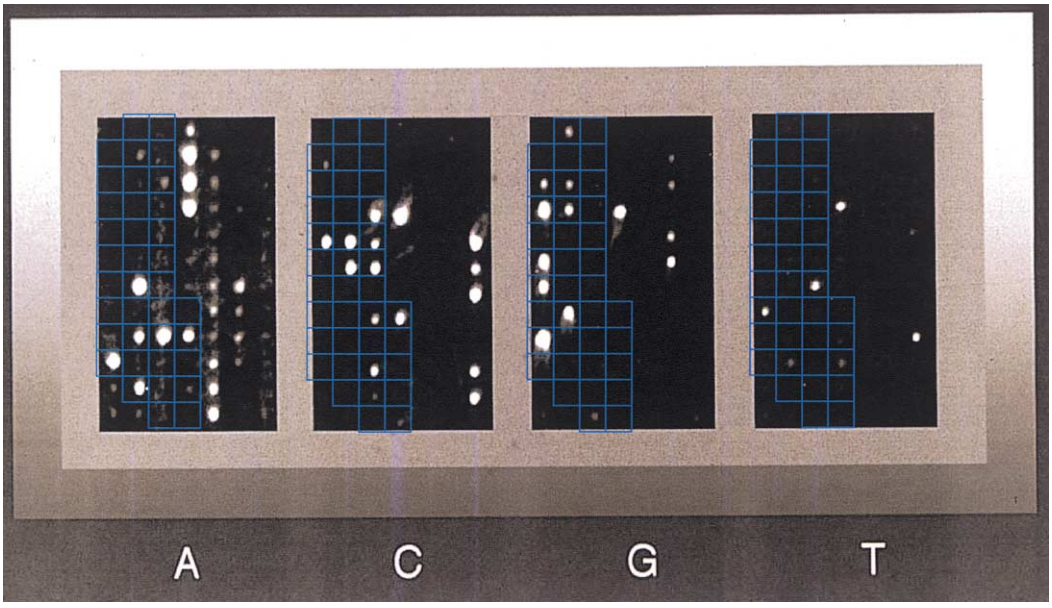

Figure 8. Deletion mutation. APEX with a 72 nt p53 exon 7 template with an A deletion mutation at position 2585. The T signal in column 3, row 4 of wild-type is replaced by a C signal, and sequence is not read following the mutation due to a G/T mismatch.

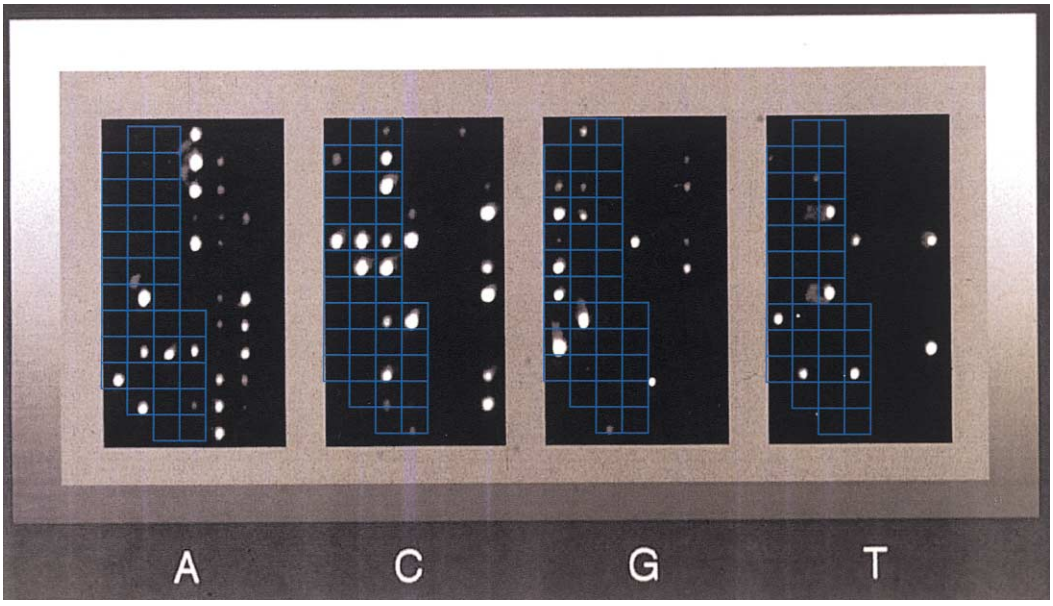

Figure 9. Insertion mutation. APEX with a 72 nt p53 exon 7 template with a G insertion mutation after position 2586. The wild-type A signal in column 3, row 2 is replaced by a C, and sequence is not read following the mutation due to a G/A mismatch.

a C signal on the grid. The 'drop-out' region is six bases in length, and the signal is recovered at primer 36.

Insertion mutation

The chip read for the insG mutation is shown in Figure 9. In this case, the wild-type A signal is replaced by a mutant C signal at primer 31. This image is consistent with the sequence of the insG template given in Table 1, where a G follows a G (instead of a T following a G). The 'drop-out' region for insG is only three bases in length and is terminated by the identification of T by primer 35.

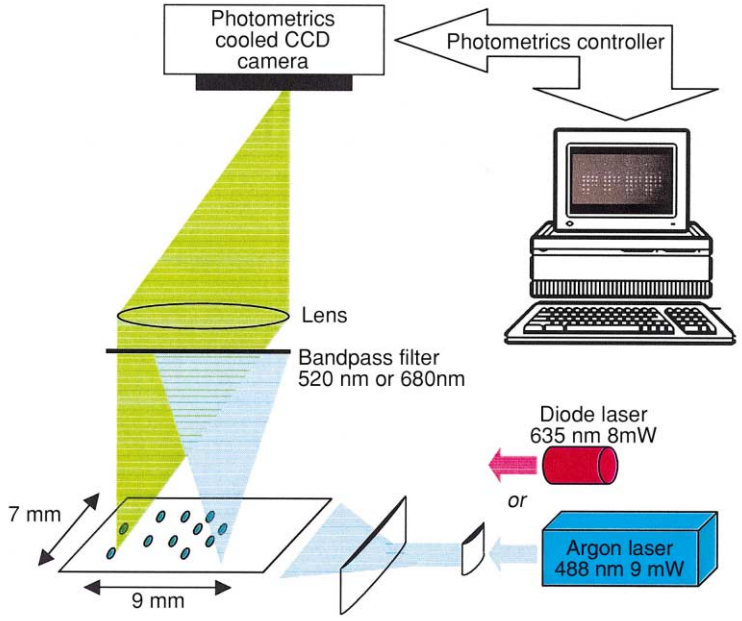

Figure 10. Instrument design. Evanescent excitation by total internal reflectance and detection of fluorescence on DNA chips.

Wild-type		insG		ΔA	
-**GGAGT**GGTAGT- -CC	+	-**GGAGGT**GGTAGT- -CC	+	-**GGGT**GGTAGT- -CC	+
-**GGAGT**GGTAGT- -CCT	+	-**GGAGGT**GGTAGT- -CCT	+	-**GGGT**GGTAGT- -CCT	-
-**GGAGT**GGTAGT- -CCTC	+	-**GGAGGT**GGTAGT- -CCTC	+	-**GG GT**GGTAGT- -CC$_T$C	-
-**GGAGT**GGTAGT- -CCTCA	+	-**GGAG^{G}T**GGTAGT- -CCTC A	-	-**GG GT**GGTAGT- -CC$_T$CA	-
-**GGAGT**GGTAGT- -CCTCAC	+	-**GGAG^{G}T**GGTAGT- -CCTC AC	-	-**GG GT**GGTAGT- -CC$_T$CAC	-
-**GGAGT**GGTAGT- -CCTCACC	+	-**GGAG^{G}T**GGTAGT- -CCTC ACC	-	-**GG GT**GGTAGT- -CC$_T$CACC	-
-**GGAGT**GGTAGT- -CCTCACCA	+	-**GGAG^{G}T**GGTAGT- -CCTC ACCA	+	-**GG GT**GGTAGT- -CC$_T$CACCA	-
-**GGAGT**GGTAGT- -CCTCACCAT	+	-**GGAG^{G}T**GGTAGT- -CCTC ACCAT	+	-**GG GT**GGTAGT- -CC$_T$CACCAT	+

Figure 11. Analysis of insertion/deletion results. Primer extension outcomes with wild-type and mutant templates are shown as + or −. For insertions and deletions, primer–template complexes can be formed with a bulged base on either the primer (deletion) or template (insertion) strand. The bulged bases are shown as a superscripts or subscripts.

Table 1. Sequences of the oligonucleotides used in this study[a]

p53-ex7 wt	
3′-ATGTACACATTGTCAAGGACGTACCCGCCGTACTTGGCCTCCGGGTAGGAGTGGTAGTAGTGT	
p53-ex7 G→A	
3′-ATGTACACATTGTCAAGGACGTACCCGCCGTACTTGACCTCCGGGTAGGAGTGGTAGTAGTGT	
p53-ex7 T→C	
3′-ATGTACACATTGTCAAGGACGTACCCGCCGTACTTGGCCCCCGGGTAGGAGTGGTAGTAGTGT	
p53-ex7 ΔA	
3′-ATGTACACATTGTCAAGGACGTACCCGCCGTACTTGGCCTCCGGGTAGGGTGGTAGTAGTGT	
p53-ex7 insG	
3′-ATGTACACATTGTCAAGGACGTACCCGCCGTACTTGGCCTCCGGGTAGGAGGTGGTAGTAGTGT	
1	TGTGTAACAGTTCCTGC
2	GTGTAACAGTTCCTGCA
3	GTGTAACAGTTCCTGCAT
4	GTAACAGTTCCTGCATG
5	TAACAGTTCCTGCATGG
6	ACAGTTCCTGCATGGG
7	CAGTTCCTGCATGGGC
8	GTTCCTGCATGGGCG
9	TCCTGCATGGGCGG
10	CCTGCATGGGCGGC
11	CCTGCATGGGCGGCA
12	CTGCATGGGCGGCAT
13	TGCATGGGCGGCATG
14	GCATGGGCGGCATGA
15	CATGGGCGGCATGAA
16	ATGGGCGGCATGAAC
17	TGGGCGGCATGAACC
18	GGGCGGCATGAACCG
19	GGCGGCATGAACCGG
20	GCGGCATGAACCGGA
21	CGGCATGAACCGGAG
22	GGCATGAACCGGAGG
23	GCATGAACCGGAGGC
24	CATGAACCGGAGGCC
25	ATGAACCGGAGGCCC
26	TGAACCGGAGGCCCA
27	GAACCGGAGGCCCAT
28	AACCGGAGGCCCATC
29	ACCGGAGGCCCATCC
30	CCGGAGGCCCATCCT
31	CGGAGGCCCATCCTC
32	GGAGGCCCATCCTCA
33	GAGGCCCATCCTCAC
34	AGGCCCATCCTCACC
35	GGCCCATCCTCACCA
36	GCCCATCCTCACCAT
37	CCCATCCTCACCATC

[a]The wild-type template is denoted by p53-ex7 wt; the substitution mutations are denoted by p53-ex7 G→A and p53-ex7 T→C; the deletion mutation is denoted by p53-ex7 ΔA; and the insertion mutation is denoted by p53-ex7 insG. The sequence of the 37 primers immobilized on the chip are listed at the bottom of the table.

Summary

The image characteristics of Figures 5–9 are very similar. The average signal level is 300 and the average background is 10, producing an average S/N of 30:1. Most of the noise levels greater than 25 are produced by 'fluorescent smearing' from adjacent elements. Primer 19 gives the lowest signal in the array.

Discussion

A feasibility study of applying APEX to the identification of sequence alterations in a known DNA sequence has been presented. DNA chips with immobilized primers on 200 μm spacing (2500 primers/cm^2) identify substitution, deletion, and insertion mutations. Labeled primers are evanescently excited, and the induced fluorescence is imaged onto a CCD camera. The entire assay can be completed within 10 min. The average S/N in these studies was 30:1. Since the signals of the elements surrounding primer 19 are large (Fig. 5), the low signal of primer 19 is attributed to an inefficient oligonucleotide synthesis.

The high S/N of APEX is significant because S/N determines the minimum level of mutant sequence that can be detected in a wild-type background. The results show that the method enables the identification of a very low mutant level.

The APEX method is very reliable because of its high redundancy in sequence determination. Each base is directly read once, and it is indirectly read several other times through its effect on the extension of primers addressing loci from 3 to 8 nt distant. While in this study only one DNA strand was used for analysis, when primers corresponding to the sequence 5′ to the mutation site and analyzing the sequence of both sense and antisense DNA strands are used, two primers of *different* sequence will address each locus directly, and indirectly address many nearby loci. These independent assessments of each nucleotide increases the confidence level in the analysis.

The ability of T7 polymerase to extend an annealed duplex with an internal mismatch under these conditions is dependent upon the type of mismatch. For a substitution mutation, T7 DNA polymerase requires a seven-base duplex region at the 3′ end of the immobilized primer; a six-base duplex region is not extended. For deletion and insertion mutations, formation of bulged duplexes with mutant template and wild-type primers is possible (Fig. 11), which implies they are responsible for extension. For a deletion mutation, the polymerase requires a six-base duplex region following the mutation; a five-base duplex is not extended. The polymerase requires only a four-base duplex following an insertion. These data suggests that the ability of T7 DNA polymerase to discriminate against frameshift mismatches depends upon the placement of the bulge. The polymerase discriminates more strongly against mismatches on the primer strand than against mismatches on the template strand; as a result, the polymerase can accept mismatches on the template strand closer to the 3′ end of the primer. Thus, the 'drop-out' region in APEX is smaller for an insertion mutation than for a deletion or substitution mutation.

The completion of the human genome sequence[40] demands new technologies to exploit this wealth of information to fulfill its promise for improvement of human health in the post-genomic era. Analysis of the variations of genomic sequence within human populations will be key not only to the discovery of the genetic basis of disease susceptibility, particularly due to multiple genetic loci, but also to individually tailored therapies as pharmacogenomics moves to the fore. It may soon be possible to address essentially all of the polymorphic loci in the genome and thereby genotype human populations. Methodologies such as APEX will be essential in this effort.

Experimental

Array preparation

The surface of a #2 micro cover glass (CMS, Inc., Houston, TX, USA) was cleaned with acetone prior to chemical modification. The surface was activated via an epoxysilane vapor deposition reaction (4 h, 80 °C) using 6 mL of toluene (Aldrich Chemical, Milwaukee, WI, USA), 2 mL of glycidoxypropyltrimethoxysilane (Aldrich) and 10 μL of *N,N*-diisopropylethylamine (Aldrich).[41] The slides were washed with ethyl acetate and dried.

Oligonucleotides were synthesized commercially with a 5′-amino linker for attachment to the glass support. The sequences of the DNA probes and immobilized primers are given in Table 1. The oligonucleotides were suspended in a basic solution (50 μM oligonucleotide, 100 mM NaOH) and attached to the solid support by pipetting onto the glass surface. Following oligo drying, the arrays were washed in 70 °C water, dried, and used in the APEX assay. The arrays were fabricated by Synteni, Inc. (Palo Alto, CA, USA).[42] The primers were immobilized 200 μm apart. The arrays were synthesized with a format of four columns of 12 elements (Fig. 4). Each number in Figure 4b denotes the position in the linear sequence that is identified at the respective element of the chip, and refers to the primer notation in Table 1. Each letter in Figure 4a denotes the wild-type nucleotide identified at the respective element of the chip. The oligonucleotides were provided by Pharmacia Biotech.

The 110 nt sequence of exon 7, bases 2501 to 2610, is:

1 GTTGGCTCTGACTGTACCACCATCCAC-
TACAAC*TACATGTGTAA-*
CAGTTCCTGCATGGGC

61 *GGCATGAACCGGAGGCCCATCCTCAC-*
*CATCATCACA*CTGGAAGACTCCAG

which is complementary to the wild-type template used herein in the italic region.

APEX assay

Fluorescein labeled ddNTPs (DuPont/NEN, Wilmington, DE, USA) were used as substrates for polymerase extension. Four separate experiments were performed, one for each labeled base. A typical reaction consisted of 1 μM target, 9 mM Tris–HC1 (pH 7.5), 2.5 μM $MnCl_2$, 6.25 mM NaCl, 10 mM $MgC1_2$, 20 mM EDTA, 0.5 mg/mL BSA, 10% glycerol, 7.9 mM DTT, 1 μM fluorescein-labeled ddNTP (e.g., fluorescein-ddATP), and 1.25 μM unlabeled ddNTPs (e.g., ddCTP, ddGTP, ddTTP), 0.1 units T7 Sequenase Version 2.0 DNA polymerase (Amersham, Cleveland, OH, USA), and 0.5 units pyrophosphatase (Amersham). The 40 μL reaction mixture was placed on the oligonucleotide array, covered with a cover slip, and incubated at 42 °C for 5 min. The array was washed in 70 °C water, dried, covered with a thin layer of Slow-Fade Light (Molecular Probes), and imaged by the fluorescence detection system.

Fluorescence image recording

The output from a 10 mW air-cooled argon ion laser (NEC GLS503 1, Tokyo, Japan) was asymmetrically expanded with a cylindrical lens and launched into a slide through a glass prism. The fluorescence from the

DNA chips was imaged onto a Photometrics PXL CCD camera (Tucson, AZ, USA) following filtering with a bandpass filter centered at 535 nm (Corion XM535, Holliston, MA, USA). Four separate images (A, C, G, T) were obtained for each DNA template.

Fluorescence image analysis

The images were analyzed using the OPUVUS software package (Version 5.2, Edmonds, WA, USA). Periodic variations of the background level were removed by a Fourier transform filter. A 2-D grid was imposed on the images, and the average signal intensity of each element was determined. Four matrices were obtained corresponding to the fluorescent intensities for A, C, G, and T. These intensity matrices represent relative fluorescence signal levels. The maximal fluorescent intensity at a given element determined the nucleotide identified at the respective position in the linear sequence, and the second largest value determined the noise associated at that element.

Acknowledgements

This work was financially supported by NIHGM 46720.

References and Notes

1. Zielenski, J.; Rozmahel, R.; Bozon, D.; Kerem, B.; Grzelczak, Z.; Riordan, J. R.; Rommens, J.; Tsui, L. *Genomics* **1991**, *10*, 214.
2. Miki, Y., et al.; Skolnick, M. H. *Science* **1994**, *266*, 66.
3. Wooster, R., et al.; Micklem, G. *Nature* **1995**, *378*, 789.
4. Orita, M.; Iwahana, H.; Kanazawa, H.; Hayashi, K.; Sekiya, T. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2766.
5. Myers, R. M.; Lumelsky, N.; Lerman, L. S.; Maniatis, T. *Nature* **1985**, *313*, 495.
6. Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5463.
7. Wallace, R. B.; Johnson, M. J.; Hirose, T.; Miyake, T.; Kawashima, E. H.; Itakura, K. *Nucleic Acids Res.* **1981**, *9*, 879.
8. Conner, B. J.; Reyes, A. A.; Morin, C.; Itakura, K.; Teplitz, R. L.; Wallace, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 278.
9. Nickerson, D. A.; Tobe, V. O.; Taylor, S. L. *Nucleic Acids Res.* **1997**, *25*, 2745.
10. Hacia, J. G. *Nat. Genet.* **1999**, *21*, 42.
11. Cronin, M. T.; Fucini, R. V.; Kim, S. M.; Masino, R. S.; Wespi, R. M.; Miyada, C. G. *Hum. Mutat.* **1996**, *7*, 244.
12. Kozal, M. J.; Shah, N.; Shen, N.; Yang, R.; Fucini, R.; Merigan, T. C.; Richman, D. D.; Morris, D.; Hubbell, E.; Chee, M.; Gingeras, T. R. *Nat. Med.* **1996**, *2*, 753.
13. Broude, N. E.; Sano, T.; Smith, C. L.; Cantor, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3072.
14. Stimpson, D. I.; Hoijer, J. V.; Hsieh, W. T.; Jou, C.; Gordon, I.; Theriault, T.; Gamble, R.; Baldeschwieler, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6379.
15. Nikiforov, T. T.; Rendle, R. B.; Goelet, P.; Rogers, Y.-H.; Kotewicz, M. L.; Anderson, S.; Trainor, G. L.; Knapp, M. R. *Nucleic Acids Res.* **1994**, *22*, 4167.
16. Syvänen, A. C.; Ikonen, E.; Manninen, T.; Bengtstrom, M.; Söderlund, H.; Aula, P.; Peltonen, L. *Genomics* **1992**, *12*, 590.
17. Syvänen, A. C.; Aalto-Setälä, K.; Harju, L.; Kontula, K.; Söderlund, H. *Genomics* **1990**, *8*, 684.
18. Livak, K. J.; Hainer, J. W. *Hum. Mutat.* **1994**, *3*, 379.
19. Shumaker, J. M.; Metspalu, A.; Caskey, C. T. In *Lab. Protoc. Mutat. Detect.*; Landegren, U., Ed.; Oxford University Press: Oxford, 1996; pp 93–95.
20. Tollett, J. J.; Kurg, A.; Shah, A.; Roa, B. B.; Richards, C. S.; Nye, S. H.; Pirrung, M.; Metspalu, A.; Shumaker, J. M. *Am. J. Hum. Gen.* **1997**, *61*, 1322.
21. Metspalu, A.; Saulep, H.; Tonisson, N.; Soo, V.; Kurg, A.; Shumaker, J. M. *Am. J. Hum. Gen.* **1997**, *61* (suppl. S), 1301.
22. Kurg, A.; Tollett, J. J.; Shah, A.; Roa, B. B.; Richards, C. S.; Nye, S. H.; Pirrung, M.; Metspalu, A.; Shumaker, J. M. *Am. J. Hum. Gen.* **1997**, *61* (Suppl. S), 68.
23. Dubiley, S.; Kirillov, E.; Mirzabekov, A. *Nucleic Acids Res.* **1999**, *27*, e19.
24. Pastinen, T.; Perola, M.; Niini, P.; Terwilliger, J.; Salomaa, V.; Vartiainen, E.; Peltonen, L.; Syvänen, A. *Hum. Mol. Genet.* **1998**, *7*, 1453.
25. Syvänen, A. C.; Landegren, U. *Hum. Mutat.* **1994**, *3*, 172.
26. Pastinen, T.; Kurg, A.; Metspalu, A.; Peltonen, L.; Syvänen, A. C. *Genome Res.* **1997**, *7*, 606.
27. Pirrung, M. C.; Connors, R. V.; Montague-Smith, M. P. *FASEB J.* **1997**, *11* (Suppl. S), A1214.
28. Pirrung, M. C.; Connors, R. V.; Montague-Smith, M. P.; Odenbaugh, A. L.; Walcott, N. G.; Tollett, J. J. *J. Am. Chem. Soc.* **2000**, *122*, 1873.
29. Pirrung, M. C.; Odenbaugh, A. L.; Davis, J. D. *Langmuir* **2000**, *16*, 2185.
30. Tabor, S.; Richardson, C. C. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4767.
31. Kurg, A.; Tonisson, N.; Georgiou, I.; Shumaker, J.; Tollett, J.; Metspalu, A. *Genet. Test.* **2000**, *4*, 1.
32. Tonisson, N.; Kurg, A.; Kaasik, K.; Lohmussaar, E.; Metspalu, A. *Clin. Chem. Lab. Med.* **2000**, *38*, 165.
33. Kiefer, J. R.; Mao, C.; Braman, J. C.; Beese, L. S. *Nature* **1998**, *391*, 304.
34. Takahashi, T.; D'Amico, D.; Chiba, I.; Buchhagen, D. L.; Minna, J. D. *J. Clin. Invest.* **1990**, *86*, 363.
35. Thompson-Hehir, J.; Davies, M. P. A.; Green, J. A.; Halliwell, N.; Joyce, K. A.; Salisbury, J.; Sibson, D. R.; Vergote, I.; Walker, C. *Diagn. Mol. Pathol.* **2000**, *9*, 110.
36. Garland, P. B. *Q. Rev. Biophys.* **1996**, *29*, 91.
37. Hutchinson, A. M. *Mol. Biotechnol.* **1995**, *3*, 47.
38. Thompson, N. L.; Lagerholm, B. C. *Curr. Opin. Biotechnol.* **1997**, *8*, 58.
39. Axelrod, D.; Thompson, N. L.; Burghardt, T. P. *J. Microsc.* **1983**, *129*, 19.
40. Yager, T. D.; Zewert, T. E.; Hood, L. E. *Acc. Chem. Res.* **1994**, *27*, 94.
41. Southern, E.; Maskos, U.; Elder, J. K. *Genomics* **1995**, *13*, 1008.
42. Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science* **1995**, *270*, 467.