5456–5465 Nucleic Acids Research, 1994, Vol. 22, No. 24

Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports

Zhen Guo, Richard A.Guilfoyle, Andrew J.Thiel, Renfeng Wang and Lloyd M.Smith*
Department of Chemistry, University of Wisconsin, Madison, WI 53706-1396, USA

Received July 5, 1994; Revised and Accepted October 24, 1994 GenBank accession no. M63238

ABSTRACT

A simple and rapid method for the analysis of genetic polymorphisms has been developed using allele-specific oligonucleotide arrays bound to glass supports. Allele-specific oligonucleotides are covalently immobilized on glass slides in arrays of 3 mm spots. Genomic DNA is amplified by PCR using one fluorescently tagged primer oligonucleotide and one biotinylated primer oligonucleotide. The two complementary DNA strands are separated, the fluorescently tagged strand is hybridized to the support-bound oligonucleotide array, and the hybridization pattern is detected by fluorescence scanning. Multiple polymorphisms present in the PCR product may be detected in parallel. The effect of spacer length, surface density and hybridization conditions were evaluated, as was the relative efficacy of hybridization with single or double-stranded PCR products. The utility of the method was demonstrated in the parallel analysis of 5 point mutations from exon 4 of the human tyrosinase gene.

INTRODUCTION

Methods for the analysis of genetic polymorphism have found wide utility in basic research, clinical diagnostics, forensics and other areas (1–4). One of the most powerful and widely used approaches for rapidly identifying mutations has been differential hybridization with allele-specific oligonucleotide probes (ASOs) (5–11). This type of analysis relies on knowledge of the polymorphic DNA sequences of the target alleles. With appropriate selection of oligonucleotide probes and hybridization conditions, it is possible to discriminate between a target DNA sequence and a variant differing in sequence by only a single base. Although the advent of PCR (12) has made possible the rapid and simple amplification of a target sequence of interest, the need for manual and laborious gel-based methods for analyzing PCR results has compromised its utility for the routine detection of DNA polymorphisms.

A potential solution to this problem is implicit in recent work by several groups exploring Sequencing by Hybridization, or SBH (13–18). This strategy, conceived of as an approach to large-scale DNA sequencing, has two major variants, termed Format I and Format II, respectively. In Format I, DNA targets to be sequenced are arrayed on a support, and each of a large combinatorial panel of oligonucleotides (e.g., all of the 65,536 possible 8mers) is sequentially hybridized to the support; the sequence of each target may then, in principle, be determined from the hybridization patterns (19). In Format II this scenario is inverted, with each of the large panel of oligonucleotides being arrayed on the support, and the sequence determined from a single hybridization of a suitably labeled target to the combinatorial array of oligonucleotides (20).

A number of substantial technical hurdles must be overcome for either of these approaches to be implemented. Format I faces practical problems such as how to perform such large numbers of successive hybridizations and how to label and detect the probe oligonucleotides cost effectively, as well as more fundamental issues such as deciphering the information when repetitive DNA regions are present and handling secondary structures in the targets. The latter issues are also present in the Format II approach, as are practical problems such as the need to develop methods for the synthesis of very complex combinatorial oligonucleotide arrays on solid supports, and for labeling and detecting the target molecules. Several groups are actively engaged in the development of these approaches. However, in the near-term the basic concept of Format II, using arrays of support-bound ASOs, also appears to offer a simple and attractive solution to the problem of large-scale detection and analysis of DNA polymorphisms.

This application of the Format II approach is simplified in a number of ways: first, large combinatorial arrays of oligonucleotides are no longer needed. Rather, less complex arrays suitable for the targeted analysis of defined genetic variants can be prepared by manual or robotic deposition of the desired oligonucleotides upon an appropriately derivatized support. Second, the problem of labeling the target strand to be analyzed is simplified by PCR: use of a suitably labeled primer in the PCR yields a labeled target strand which may then be used directly in hybridization to the support-bound oligonucleotides. Finally, the relatively small number of support-bound oligonucleotides employed reduces the difficulty of finding uniformly applicable hybridization conditions and decreases the spatial resolution required in detection.

*To whom correspondence should be addressed

In this manuscript methods are described for a) the preparation of planar arrays of ASOs on glass supports, b) the synthesis and purification of fluorescein-labeled single-stranded PCR probes, c) hybridization of the labeled purified PCR product to the ASO array and d) fluorescence scanning of the array to identify genetic polymorphisms. Oligonucleotide linkage chemistry and hybridization conditions were optimized. This direct fluorescence detection method, which we refer to as DAFT (Direct Allele-Specific Fluorescence Targeting), was applied to the parallel analysis of 5 point mutations present in exon 4 of the human tyrosinase gene.

MATERIALS AND METHODS

Oligonucleotides

All oligonucleotides were synthesized on an Applied Biosystems (Applied Biosystems, Inc., Foster City, CA, [ABI]) 308B or 391 DNA synthesizer. ASOs used in the fabrication of support-bound arrays contained a 5′ amino group introduced using the reagent N-trifluoroacetyl-6-aminohexyl-2-cyanoethyl N′,N′-diisopropyl-phosphoramidite (ABI). One of the two PCR primers employed was biotinylated by coupling NHS-LC-Biotin (Pierce, Rockford, IL) to the 5′ amino group as described (21). For fluorescence detection, the other primer was fluorescein-labeled by using 6-FAM phosphoramidite (ABI). The oligonucleotides were purified by reversed-phase high performance liquid chromatography, and oligonucleotide concentrations determined by UV spectrophotometry at 260 nm (22).

Preparation of ASO arrays on glass supports

Pre-cleaned microscope slides were immersed in 1% 3-aminopropyltrimethoxysilane solution (Aldrich Chemical, Milwaukee, WI) in 95% acetone/water for 2 min. The slides were washed ten times with acetone, 5 min per wash, dried for 45 min at 110°C, treated for 2 h with a solution of 0.2% 1,4-phenylene diisothiocyanate (PDC) solution (Aldrich) in 10% pyridine/dimethyl formamide and washed with methanol and acetone. The activated glass slides may be stored indefinitely at 4°C in a vacuum dessicator containing anhydrous calcium chloride without discernible loss of activity. Amino-oligonucleotides purified by HPLC were dried down on a Savant rotary evaporator and redissolved at a concentration of 2 mM in 100 mM sodium carbonate/bicarbonate buffer (pH 9.0). 2 μl of the resultant solutions were applied directly to the PDC-derivatized microscope slide in the desired grid pattern. Accurate positioning of the spots was facilitated by following a spot pattern on a paper template underneath the microscope slide. Slides were incubated at 37°C in a covered Petri dish containing a small amount of water for 1−2 h, removed, washed once with 1% NH_4OH, three times with water and air dried at room temperature. The slides were now ready for use in hybridization experiments. It is not recommended that slides be employed multiple times, as increased background is observed.

Oligonucleotide immobilization studies

The 5′ amino-modified oligonucleotides were 3′ end labeled with [α-^{32}P]dATP (Amersham, 400 Ci/mmol, 10 mCi/ml). A 100 μl labeling reaction containing 200 nmole oligonucleotide, 10 μl [α-^{32}P]dATP, 8 μl terminal deoxynucleotidyl transferase (USB, Cleveland, OH, 5 u/μl) and 20 μl 5×reaction buffer (500 mM sodium cacodylate, 10 mM $CoCl_2$ and 1 mM 2-mercaptoethanol) was incubated at 37°C for 60 min. The labeling mixture was purified on a NICK™ Spin column (Pharmacia Biotech, Piscataway, NJ). 2 μl of oligonucleotide solutions at various concentrations were applied to 2.5 cm×0.75 cm PDC-derivatized glass pieces at 37°C for 1 h. The glass surface was washed with 1% NH_4OH followed by water. Bound oligonucleotide was quantified by liquid scintillation counting on a Beckman LS6000 SE scintillation counter (Beckman Instrument Co., Fullerton, CA). Results of experiments not shown demonstrated that measured count rates were reduced by 50% due to screening of the emitted β-particles by the glass microscope slide. Accordingly these measured values were doubled for use in subsequent calculations.

DNA amplification and strand separation

Plasmid cDNA clones and human genomic DNA samples employed in these studies were kind gifts from Dr Richard Spritz (23) and are described further in the legend to Figure 4b. The tyrosinase gene fragment corresponding to exon 4 was amplified by PCR using one of two different primer sets. One set of PCR primers, 5′-biotin-TATTTTTGAGCAGTGGCTCC-3′ and 5′-(F)CTGAATCTTGTAGATAGCTA-3′ (where F indicates the fluorescein label for experiments employing fluorescence detection), yields a 157 nucleotide (nt) fragment from both the true tyrosinase gene exon 4 and the pseudogene. Because of this lack of specificity in the PCR amplification, this primer pair was used only for amplification of cloned plasmid DNAs in model studies. The second set of PCR primers, 5′-biotin-TTAATA-TATGCCTTATTTTA-3′ and 5′-(F)GGAATTCTAAAGTT-TTGTGTTATCTCA-3′, specifically yields a 347 nt fragment from exon 4 and adjacent intronic sequences of the true tyrosinase gene (38). This specificity is achieved by means of a mismatched base at the 3′ end of fluorescence primer in the pseudogene intronic region. As this primer pair utilizes intronic sequences it was not suitable for amplification from the plasmid cDNA clones. The specificity of the amplification for the true exon 4 was confirmed by restriction with *Msp*I as described (38). The amplification reaction (100 μl) contained 50 mM KCl, 10 mM Tris−HCl, 1.5 mM $MgCl_2$, 10 mg of gelatin, 50 ng DNA, 2 μM of each primer, 200 μM each dATP, dCTP, dTTP and dGTP, and 2.5 units of *Thermus aquaticus* (*Taq*) DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT). For radiolabeling of PCR products, 2.5 mM [α-^{32}P]dATP (Amersham, 400 Ci/mmol, 10 mCi/ml) was added to the reactions. PCR was performed in a Perkin-Elmer Cetus thermal cycler model 9600 using 35 cycles. For amplification of the 157 nt fragment the cycle employed was: 94°C for 30 s, 50°C for 1 min and 72°C for 2 min; for amplification of the 347 nt fragment the cycle employed was: 93°C for 45 s, 52°C for 2 min and 72°C for 1 min and 45 s.

The PCR mixture was purified using a QIAGEN™ PCR purification kit (QIAGEN Inc., Chatsworth, CA). For radiolabeled PCR, the PCR product was first purified using a NICK™ Spin column (Pharmacia Biotech) to remove excess [α-^{32}P]dATP. Strand separation of PCR products was accomplished using streptavidin-coupled magnetic beads (Dynabeads M-280, Dynal, Inc. Great Neck, NY) (21). 1 mg beads were prewashed in 1×TES (10 mM Tris−HCl, 1 mM EDTA, 1 mM NaCl, pH 8.2). The biotinylated PCR product was bound to streptavidin beads by adding it to the bead suspension and keeping the mixture at room temperature for 15 min, with frequent shaking. The beads were then washed twice with 1×TES. 100 μl 0.1 N NaOH was added to separate the

two DNA strands. The reaction was kept at room temperature for 10 min. The supernatant containing the non-biotinylated DNA strand was collected and then neutralized to pH 7 with 0.1 M HCl.

Hybridization optimization

To study the effect of oligonucleotide surface density on hybridization efficiency, a PDC-derivatized glass slide was cut into 2.5 cm×0.75 cm pieces, and ASO couplings were performed as described above, using 0.3–20 mM ASO solutions. 10 μl hybridization solution (20–50 nM single-stranded ^{32}P-labeled PCR product, 5×SSPE, 0.5% SDS) was applied to each piece of glass, which was then incubated for 3 h at 30°C in a closed Petri dish over water. Two 15 min washes were performed in a shaking water bath with 20 ml washing buffer (2×SSPE, 0.1% SDS) each. Hybridization signals were quantified by liquid scintillation counting of the glass piece.

To compare the hybridization of single-stranded and double-stranded PCR products, ASO-coupled glass pieces at different surface densities were prepared as described above. The single-stranded PCR products were prepared as described above. To prepare hybridization solution containing double-stranded PCR product, 10 μl ^{32}P-labeled PCR product was denatured by addition of 10 μl 400 mM NaOH and combined with 100 μl hybridization buffer (5×SSPE, 0.5% SDS). 10 μl hybridization solution containing either single-stranded or double-stranded ^{32}P-labeled PCR product was hybridized with the glass pieces for 3 h at 30°C in a closed Petri dish over water. Finally, the glass pieces were washed as above and hybridization signals were determined by scintillation counting.

To study hybridization kinetics, 2 μl solutions of 5 mM 422 wild type ASO were applied to glass pieces as above. Control glass pieces were prepared in an identical manner, except without ASO in the coupling solution. 10 μl single-stranded hybridization solution as above was incubated with each piece of glass at 30°C for 1, 2, 3, 4, 5 and 6 h, respectively. After incubation, the glass pieces were washed and scintillation counted. The scintillation counts obtained from incubation with PDC-derivitized slides lacking coupled ASOs were taken as backgrounds.

Fluorescence detection

After hybridization and washing, 50 μl washing buffer (2×SSPE, 0.1% SDS) was immediately applied to the glass slide, and the slide was covered with a 22 mm×50 mm microscope cover glass. This provides an aqueous environment for the fluorescence scanning. The slide was scanned using a Molecular Dynamics FluorImager 575. This instrument generates fluorescence images by scanning a laser (Ar^+ ion, 488 nm, 5 mW, 50 μm spot size) beam over the sample surface in a raster pattern, and monitoring the resulting fluorescence (through a 530 ± 15 nm bandpass filter) in pixel-by-pixel fashion. The spatial resolution of the scan was 100 μm per pixel, resulting in an overall image size of 750×250 pixels for the entire 75 mm×25 mm microscope slide. The fluorescence intensity of each pixel (as measured by a PMT) was digitized to 16-bit precision, and the data saved to computer disk as a 16-bit TIFF format file. After scanning, the fluorescence image was reconstructed from the digitized pixel intensities using software (ImageQuaNT v4.0) provided with the FluorImager 575.

The fluorescence standard curve was prepared by spotting 2 μl of a dilution series of fluorescein labeled oligonucleotide solution in 3 mm spots on the PDC-derivatized glass surface, yielding spots containing 3, 1, 0.3, 0.1, 0.03, 0.01, 0.003 and 0.001 pmol fluorescein, respectively. The slide was scanned before and after drying at room temperature. The signal on each spot was integrated using the ImageQuaNT software. The integrated fluorescence intensities were plotted versus the amount of fluorescence oligomer, as shown in Figure 3f.

For signal-to-noise ratio calculations, the ImageQuaNT software was used to sum pixel intensities within each spot image. The average value and standard deviation of pixel intensities were calculated and the local background level was subtracted. The signal-to-noise ratio was calculated as the fluorescence intensity level over background divided by the standard deviation of the intensities.

For digital image filtering, the data was exported to NIH Image v1.44, where it was convolved with a 3×3 median filter (24). The filter operates upon the center pixel of a 3×3 square neighborhood, replacing that pixel's value with the median pixel value of the neighborhood. This operation is performed upon every pixel in the original image, with the resulting values forming the new filtered image. The overall effect of this filtering is to smooth out pixels in the image which differ sharply from their immediate surroundings. The data are visually very similar before and after filtering; the primary effect of the filtering is upon the measured signal-to-noise ratio as discussed below.

The 3-dimensional representation of the data was generated using the color surface plot feature of NIH v1.55 (25).

RESULTS AND DISCUSSION

Support chemistry

A variety of solid supports have been investigated in DNA diagnostic applications, including nitrocellulose and nylon membranes (26,27), controlled-pore glass beads (28), acrylamide gel (29), polystyrene matrices (30), activated dextran (31), avidin-coated polystrene beads (32) and glass (33). Several technical issues must be considered in choosing an appropriate support. These issues include the level of scattering and fluorescence background, the chemical stability and complexity, the amenability to chemical modification or derivatization, surface area and loading capacity and the degree of non-specific binding.

In the present study, glass slides were chosen as the support media. Such supports were employed by Fodor *et al.* in the fabrication and use of combinatorial peptide and oligonucleotide arrays (34,35). These workers also showed that the glass supports were well-suited to fluorescence detection, a property used to advantage in the present studies. Glass is a readily available and inexpensive support medium, possessing a relatively homogeneous chemical surface whose properties have been well-studied, and which is amenable to chemical modification using very versatile and well developed silanization chemistry.

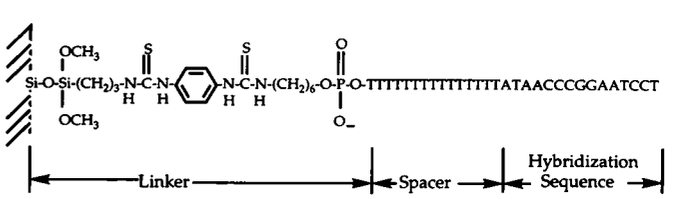


Figure 1. Diagram of linker chemistry employed for attachment of ASOs to glass supports at their 5′ termini.

Having chosen a support, the next aspect of the chemistry which must be developed is the coupling chemistry for linking the oligonucleotides to the support. This chemistry must meet several criteria (33): the linkage must be chemically stable, sufficiently long to eliminate undesired steric interference from the support, preferably hydrophilic enough to be freely soluble in aqueous solution, and not produce non-specific binding to the support. The linkage chemistry employed in the present studies is diagrammed in Figure 1, and is a variant of a linking chemistry developed previously for the preparation of derivatized glass membranes used in protein electroblotting (36). The surface derivatization is performed in three steps: a) reaction of the glass surface with aminopropyltrimethoxysilane to give an amino-derivatized surface; b) coupling of the amino groups with excess p-phenylenediisothiocyanate to convert the amino groups to amino-reactive phenylisothiocyanate groups; and c) coupling of 5′ amino-modified oligonucleotides to these amino-reactive groups to yield the surface bound oligonucleotide. This chemistry provides a 23 atom linker between the surface and the oligonucleotide. Additional distance between the hybridizing region of the oligonucleotide and the support may be obtained by including a non-specific poly dT spacer sequence at the 5′ end of the oligonucleotide, as shown in Figure 1. Experiments in which the length of this additional spacer sequence is varied are described below. In all experiments reported here the oligonucleotide probe is attached to the surface at the 5′ terminus: this choice was made primarily for convenience, as the chemistry for placing a primary amino group at the 5′ end is more widely available and straightforward than that for the 3′ end.

Tyrosinase model system

The model system employed in this study is exon 4 of the human tyrosinase gene. Tyrosinase is a copper-containing enzyme that catalyzes the first two reactions in the melanin biosynthetic pathway. A deficiency of human tyrosinase can cause the severe genetic disorder, type I oculocutaneous albinism. The human tyrosinase gene consists of five exons spanning more than 50 kb of genomic DNA on the long arm of human chromosome 11. A pseudogene containing a copy of the tyrosinase gene exons 4 and 5 has also been identified and localized to the short arm of chromosome 11 (38). At least 50 different pathologic mutations of the tyrosinase gene have been identified in patients with oculocutaneous albinism.

The fourth exon of the human tyrosinase gene is 182 nucleotides long and contains five known single-base mutations (GenBank accession number M63238). Figure 2 shows the complete sequence of exon 4 along with the flanking intron sequences (37–39). The figure also shows the nature and locations of the five single-base mutations examined. The five mutations are named P406L, G419R, R422Q, G446S and D448N, according to the codons in which they occur. The sequences utilized for ASOs are given in Table 1.

```
ttaat atatg cctta tttta cttta aaaat tttca aatgt ttctt ttata cacaa tatgt ttctt agtct gaata acctt ttcct

ctgca gT ATT TTT GAG CAG TGG CTC CGA AGG CAC CGT CCT CTT CAA GAA
      r  Ile Phe Glu Gln Trp Leu Arg Arg His Arg Pro Leu Gln Glu
                                                 CTT(Leu)

GTT TAT CCA GAA GCC AAT GCA CCC ATT GGA CAT AAC CGG GAA TCC TAC
Val Tyr Pro Glu Ala Asn Ala Pro Ile Gly His Asn Arg Glu Ser Tyr
                                    AGA(Arg)     CAG(Gln)

ATG GTT CCT TTT ATA CCA CTG TAC AGA AAT GGT GAT TTC TTT ATT TCA
Met Val Pro Phe Ile Pro Leu Tyr Arg Asn Gly Asp Phe Phe Ile Ser

TCC AAA GAT CTG GGC TAT GAC TAT AGC TAT CTA CAA GAT TCA Ggtaaa gttta
Ser Lys Asp Leu Gly Tyr Asp Tyr Ser Tyr Leu Gln Asp Ser A
                  AGC(Ser) AAC(Asn)

ctttc tttca gagga attgc tgaat ctagt gttac caatt tattt tgaga taaca caaaa cttta
```

Figure 2. DNA sequence of human tyrosinase gene exon 4 and adjacent intronic regions. Intron sequence is shown in lower case; exon sequence in capitals. The codons containing the 5 single-base mutations examined in this study are underlined. The wild-type sequences are shown with their allelic variants given beneath.

Oligonucleotide surface density, structure and hybridization efficiency

The surface density of the oligonucleotide probe is expected to be an important parameter of this system. A low surface coverage would presumably yield a correspondingly low hybridization signal, and decrease the hybridization rate. Conversely, high surface densities might result in steric interference between the covalently immobilized oligonucleotides, impeding access to the target DNA strand. To evaluate these parameters several experiments were performed. Unless otherwise indicated, all these experiments employed the 157 nt PCR product. In one set of experiments, different concentrations of a 5′-amino oligonucleotide labeled with ^{32}P at the 3′ end were reacted with the phenylenediisothiocyanate (PDC) modified glass surface. The resultant surface density was quantified by scintillation counting of the corresponding region of the glass slide, with the results shown in Figure 3a. These results show a linear relationship between the oligonucleotide concentration and the surface density over most of the range examined; thus it is possible to control the surface density of immobilized oligonucleotide by varying the oligonucleotide concentration. Knowledge of the specific activity of the radioisotope permits the surface density to be calculated: for the data of Figure 3a 1000 cpm corresponds to approximately 1250Å^2/molecule. The surface density of bound oligonucleotides was also influenced by the glass surface derivatization process. A higher concentration of 3-aminopropyl-trimethoxysilane (APTMS) during the preparation of the glass surface (see Materials and Methods) increases the surface density of bound oligonucleotide, but also causes an increase in non-specific binding to the activated surface (data not shown).

Table 1. Hybridization sequences employed for the allele-specific oligonucleotide(ASOs). Each ASO has a 15 nucleotide poly dT spacer sequence at the 5′ end and an aliphatic amino group at the 5′ terminus

	MUTATION		PROBE SEQUENCE
406	Mutant	CTT	5′ CGT CTT CTT CAA GAA 3′
	Wild Type	CCT	5′ CGT CCT CTT CAA GAA 3′
419	Mutant	AGA	5′ CCA TTA GAC ATA ACC 3′
	Wild Type	GGA	5′ CCA TTG GAC ATA ACC 3′
422	Mutant	CAG	5′ ATA ACC AGG AAT CCT 3′
	Wild Type	CGG	5′ ATA ACC GGG AAT CCT 3′
446	Mutant	AGC	5′ CAA AGA TCT GAG CTA 3′
	Wild Type	GGC	5′ CAA AGA TGT GGG CTA 3′
448	Mutant	AAC	5′ GCT ATA ACT ATA GCT 3′
	Wild Type	GAC	5′ GCT ATG ACT ATA GCT 3′

To determine the optimal surface density, hybridization experiments were performed using ^{32}P labeled single-stranded PCR product prepared as described in Materials and Methods, with the results shown in Figure 3b. As expected, there is an optimal surface coverage at which a maximum amount of the complementary PCR fragment is hybridized. This optimum occurs at about 5 mM oligonucleotide concentration, corresponding to a surface density of approximately 500Å^2/molecule. Under these conditions, approximately 20% of the surface bound ASOs are hybridized. Interestingly, the optimum surface density for hybridization to the longer 347 nt fragment was found to be approximately 30% lower (data not shown). This presumably reflects the greater steric interference of the longer fragment.

Another important parameter of this system is the nature of the PCR product hybridized to the surface. Although it is simpler to prepare double-stranded PCR products than single-stranded, hybridization of the double-stranded molecule to the support will necessarily suffer from competition of the complementary strand with the support bound oligonucleotide. To evaluate this, a comparison of hybridization efficiency using both single-stranded and double-stranded 157 nt PCR products was performed, with the results shown in Figure 3c. It is clear from these data that hybridization efficiency is much greater with the single-stranded than the double-stranded product. Similar results were obtained from longer PCR fragments up to 347 bases (data not shown). Accordingly, single-stranded PCR products were employed exclusively in subsequent hybridization experiments.

A third parameter examined was the length of the spacer between the support and the hybridization sequence. Initial work did not utilize a poly dT spacer (Figure 1), and no hybridization signal was observed. Hypothesizing that this was due to steric interference with the support, oligonucleotide probes with poly dT spacers of 0, 3, 6, 9, 12 or 15 nucleotides were synthesized and attached to supports. The results of hybridizations to these supports are shown in Figure 3d. The hybridization signal was virtually non-existent for poly dT spacers of six nucleotides or less, but then increased strongly with length up to 15 nucleotides. At this time, poly dT spacers longer than 15 have not been examined, although the data of Figure 6 suggest that further gains in hybridization signal strength may be attainable in this fashion, albeit requiring increased time and expense in oligonucleotide synthesis. Alternatively, longer chemical linkers could be employed as described by Maskos and Southern (33).

Hybridization kinetics

Hybridization kinetics were evaluated by varying hybridization time as shown in Figure 3e. In this figure plots are shown of both the non-specific background and the measured signal from the ASO-coupled region, as well as the difference of the two, corresponding to the net specific hybridization signal. Under the conditions employed the specific signal increases with increasing hybridization time up to 6 h; however, at about 3 h, a substantial non-specific background also begins to accrue. This background, whose origin is unclear, is patchy and non-uniform in nature. Efforts to reduce or eliminate it by use of non-specific blocking DNA and various detergents have thus far been unsuccessful. Therefore a hybridization period of 2–3 h was routinely employed for hybridization with the 157 nt fragment, yielding a consistent and reproducible hybridization signal without interfering background. In the case of the 347 nt fragment, it was found necessary to employ a longer hybridization time of 6–8 h to obtain good signal, in spite of the increase in background.

Design of ASO sequences

Discrimination in hybridization between perfectly matched and single-base mismatched DNA duplexes relies upon their stability differences, which can be estimated using semi-empirical relationships (40,41). In the present work, it is desirable (although not absolutely necessary) to find a single set of hybridization conditions for which a clear discrimination between matches and mismatches can be obtained for all the polymorphisms under analysis. This requires that the melting temperature of the chosen ASO sequences be comparable. It may be achieved by careful selection of the ASO length, base composition, and the mismatch position within the hybridization sequence.

The length of the hybridization sequence is the most significant factor affecting duplex stability. Generally, short hybridization sequences have the advantage of providing greater discrimination between matches and mismatches, but at a cost of lower overall stability; whereas longer hybridization sequences form more stable duplexes, but suffer from less discrimination due to the smaller fraction of the duplex which is mismatched. Another important factor affecting choice of oligonucleotide length is concerns about secondary structures forming in the single-stranded PCR product (26,42). Unlike double-stranded DNA, which is a relatively stiff complex, single-stranded DNA can be quite flexible, particularly in high salt conditions, and accordingly is prone to formation of internal structures. If the thermodynamic stability of such structures is greater than the stability of the duplex being formed between the PCR product and the ASO probe, the hybridizing regions of the single-stranded PCR product may not be accessible to the surface-bound ASO probes. This concern may be addressed in part by choosing longer ASO sequences, permitting higher temperatures to be employed in hybridization and therefore melting out internal structures in the single-stranded PCR product.

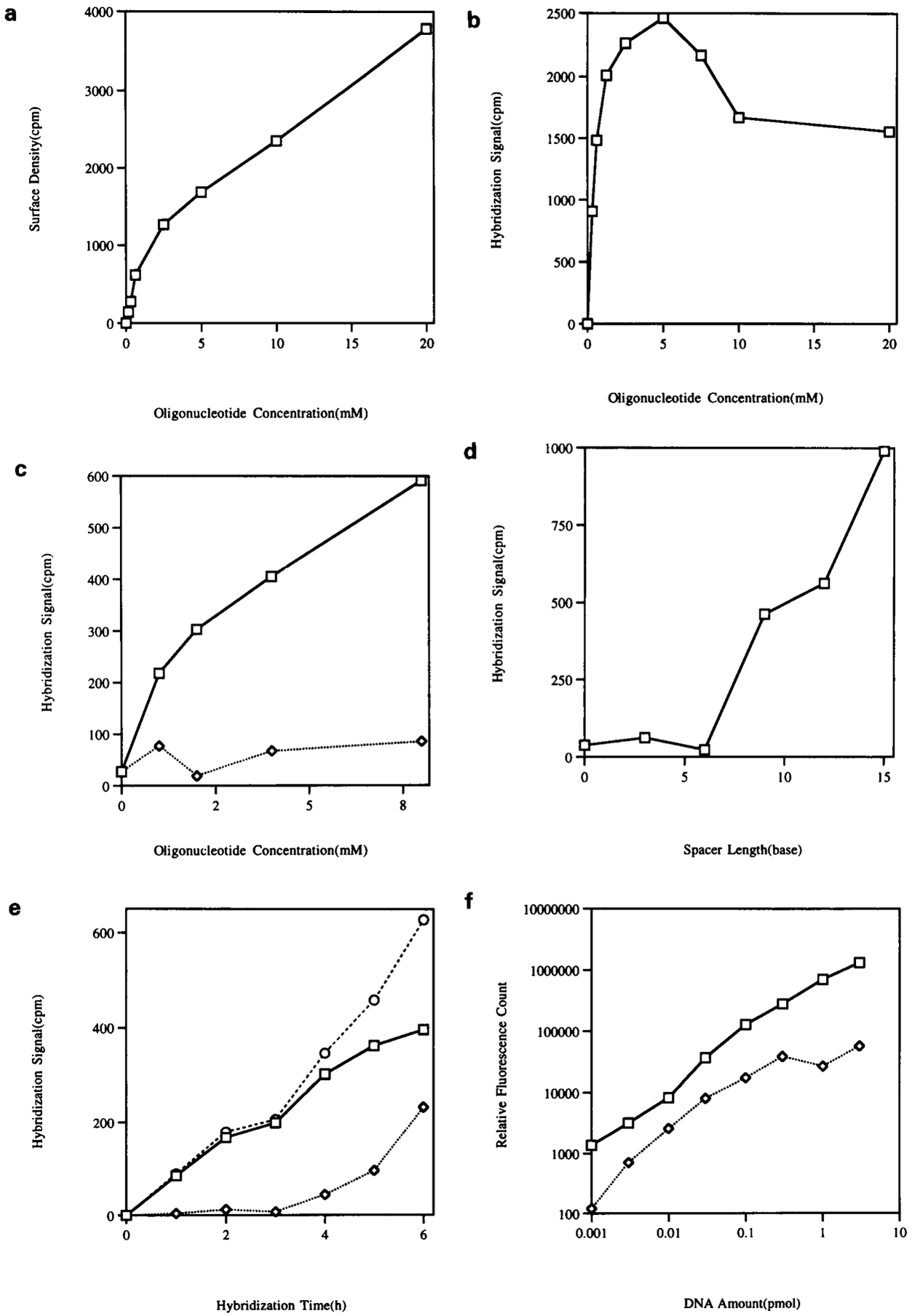
With these considerations in mind, ASO hybridization sequences of 12, 15 and 20 nucleotides in length were tested. Although all three lengths yielded approximately equal hybridization signals, single-base mismatch discrimination was best achieved with the 15-mer, whereas the 20-mer showed no reproducible discrimination and the 12 nucleotide sequence was difficult to use due to its low melting temperature (data not shown). An ASO hybridization sequence length of 15 nucleotides was therefore chosen for all subsequent work.

GC content also has a strong effect upon duplex stability. In the present work, all ASO sequences were chosen to have a GC content of 5–7 nucleotides out of 15. It was found that under these conditions a single set of hybridization conditions provided excellent discrimination for all 10 ASOs employed.

A third important factor is the location of the mismatched base in the hybridization sequence. Fodor and co-workers have shown that the discrimination between matches and mismatches is greatest when the mismatched base is placed near the center of the hybridization sequence (43). Accordingly, in this work all mismatches have been placed near the center of the ASO hybridization sequence.

Polymorphism detection in exon 4 of the human tyrosinase gene

The above system was applied to the analysis of 5 single-base mutations contained within exon 4 of the human tyrosinase gene.


a
Surface Density(cpm)
4000
3000
2000
1000
0
0
5
10
15
20
Oligonucleotide Concentration(mM)
b
Hybridization Signal(cpm)
2500
2000
1500
1000
500
0
0
5
10
15
20
Oligonucleotide Concentration(mM)
c
Hybridization Signal(cpm)
600
500
400
300
200
100
0
0
2
5
8
Oligonucleotide Concentration(mM)
d
Hybridization Signal(cpm)
1000
750
500
250
0
0
5
10
15
Spacer Length(base)
e
Hybridization Signal(cpm)
600
400
200
0
0
2
4
6
Hybridization Time(h)
f
Relative Fluorescence Count
10000000
1000000
100000
10000
1000
100
0.001
0.01
0.1
1
10
DNA Amount(pmol)

Ten 5′ amino-derivatized ASO probes were synthesized, each containing a 15 base hybridization sequence (Table 1) and a 15 base 5′ polythymidylate spacer. Five of the ASO probes correspond to the wild type sequences at the polymorphic site, and 5 to the mutant sequences. The ASO probes were applied in duplicate in 3 mm spots on derivatized glass slide surfaces. Due to the hydrophobic nature of the modified surface, the aqueous spots form well-defined circles rather than spreading as they would on a hydrophilic surface. The spots were arranged in a 4×5 array as shown in Figure 4a.

Fluorescently tagged single-stranded PCR products were prepared from either plasmid DNA clones of exon 4 or human genomic DNAs, and hybridized to the ASO arrays under conditions optimized for ^{32}P-labeled PCR products described above. After hybridization and washing steps, the microscope slide was scanned using a Molecular Dynamics FluorImager 575 fluorescence scanner. Figure 4b shows a composite of fluorescence data obtained from twelve such hybridization experiments: the left side shows data obtained from six slides hybridized individually with PCR products amplified from wild type or mutant cDNA clones; data on the right side was obtained with PCR products amplified from human wild type, homozygous or heterozygous genomic DNAs (see figure legend). It is apparent from these results that the discrimination of single-base mismatches is excellent. In addition, the ability to detect and discriminate both homozygotic and heterozygotic mutations in human genomic DNAs is demonstrated [compare the results shown on the right side from homozygous DNAs (panels 1, 2, 3) with those from heterozygotes (panels 4, 5, 6)].

Quantitative image analysis

Most ASO analyses to date have employed either direct radioactive detection or a variety of indirect enzyme-based systems generating a colored, fluorescent or chemiluminescent product. Radioactive detection, although extremely versatile and sensitive, is not well suited for clinical applications, and is increasingly expensive and difficult to use because of environmental and health concerns. Conversely, indirect non-isotopic enzyme systems suffer from complexity and do not generally permit accurate quantification. The fluorescence-based system described here permits direct quantification similar to that afforded by radioactive detection, but without the associated hazards. The simplicity and robustness of the chemistry employed is likely to increase the ease and reliability of the procedure substantially over that of indirect enzyme-based methods. The direct accessibility of the fluorescence data in digital form affords the advantage of facile quantification, comparison, image analysis and data archiving in this application. One example of this is the use of quantitative fluorescence data to provide an objective criterion for retaining or discarding the data in clinical diagnostic applications, and to assign a degree of certainty to the genotype based on the experimental results.

In quantitative analysis of the fluorescence image, it is assumed that the amount of fluorescent light emitted from each spot is representative of the amount of labeled DNA associated with that spot. To verify this assumption, various amounts of fluorescein-labeled oligonucleotide were spotted directly onto a microscope slide, and scanned twice: once while the droplets were still wet, and once after drying at room temperature. In the resulting image of fluorescent spots, the fluorescence intensity of each spot was integrated over the entire spot as described in Materials and Methods, with the results shown in Figure 3f. These standard curves reveal an important feature: the intensity and linearity of the measured fluorescence depends strongly on whether the sample is wet or dry. The upper trace corresponding to the wet sample shows a roughly linear relationship between the amount of fluorescent oligomer and the measured fluorescence intensity throughout the entire range examined (1 fmol to 3 pmol). The lower trace corresponding to the same slide scanned after drying at room temperature shows a markedly decreased fluorescence intensity which deviates from linearity for the higher amounts of DNA (1 and 3 pmoles). Rehydration regenerates the previous high level of fluorescence signal, demonstrating that photodestruction of the fluorophore is not responsible for the phenomenon (data not shown). A partial explanation for this behavior may be that the distance between fluorophores is decreased in the dry sample, relative to the wet sample, leading to increased Förster energy transfer and fluorescence 'quenching' (44,45). On the basis of these results, all scanning is performed on wet samples, as described in the Materials and Methods section, conferring greater sensitivity and linearity to the procedure. Under these conditions the instrumental detection limit corresponding to a signal to noise ratio of 2 is about 1 fmol of fluorescein per 3 mm spot, or one molecule per 1000Å×1000Å area.

An important consideration for quantitative analysis is the background arising from scattered light and non-specific fluorescence. Several factors can contribute to this background, including scattering of the excitation laser beam at the glass surfaces and dust or other contaminants on the slides. In the fluorescence image, this leads primarily to single dark pixels or small groups of pixels, whereas the true signal is composed of

Figure 3. Surface density and hybridization studies. Quantitation was performed by liquid scintillation counting of ^{32}P-labeled DNAs or fluorescence scanning as described in Materials and Methods. cpm = counts per minute. (**a**) Study of surface density. 0.3−20 mM solutions of 422 wild type ^{32}P-labeled oligonucleotide were applied to a PDC-derivatized glass surface. The coupling reaction was kept at 37°C for 1 h and then washed with 1% NH_4OH and water. (**b**) Effect of oligonucleotide surface density on hybridization efficiency. 10 μl single-stranded DNA hybridization solution was hybridized to surface bound 422 wild type oligonucleotide at various surface densities. (**c**) Comparison of hybridization signals of single-stranded DNA and of double-stranded DNA. Squares represent the hybridization signals obtained from single-stranded PCR product while diamonds represent the signals obtained from double-stranded PCR product. 422 wild type oligonucleotide was immobilized on PDC-derivatized glass surface, at various surface densities. 10 μl ^{32}P-labeled single-stranded or double-stranded DNA hybridization solution was incubated with glass pieces for 3 h. (**d**) Effect of spacer length on hybridization efficiency. Hybridizations were performed with an ASO surface density of 500Å^2/molecule obtained with a 5 mM ASO coupling solution, and all probes contain the 422 Wild type hybridization sequence with spacer lengths of 0, 3, 6, 9, 12 or 15 dTs. (**e**) Hybridization kinetics. Diamonds represent the background counts obtained in incubations with PDC-derivatized slides lacking ASOs. Circles represent the counts obtained from ASO-coupled regions under the same conditions. Squares represent the difference of the two, corresponding to the net specific hybridization signals. (**f**) Fluorescence standard curves in wet and dry environments. Squares represent the fluorescence intensities in wet environment while diamonds represent the signal intensities in dry environment. The amount of fluorophore ranges from 1−3 fmol. The fluorescence signals are decreased in the dry environment.

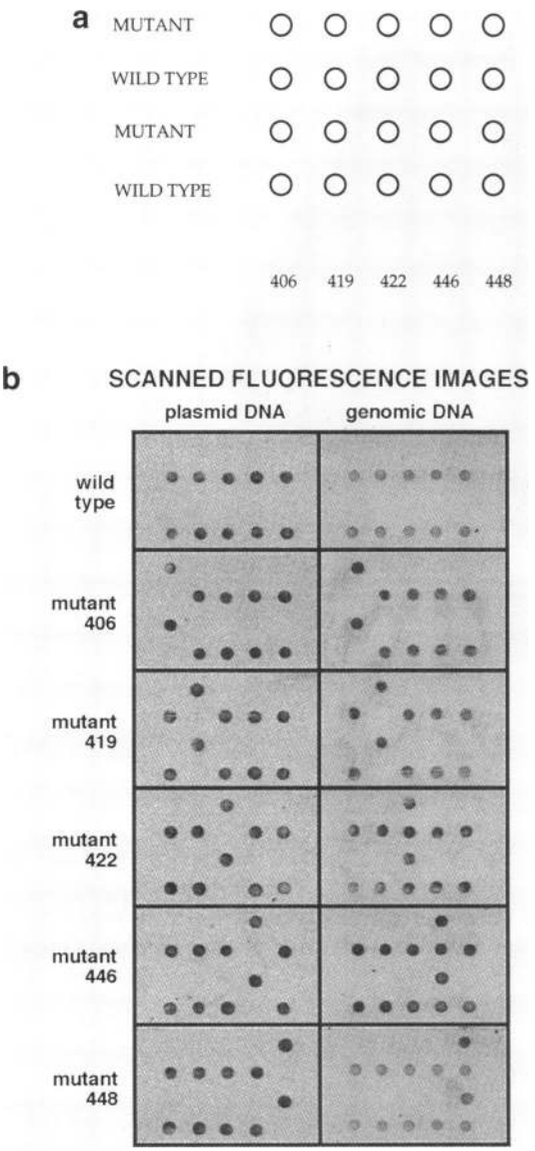


Figure 4. (**a**) Layout of the ASO array employed in the human tyrosinase gene exon 4 polymorphism analysis. The top and bottom two rows are duplicates. Column numbers 406, 419, 422, 446 and 448 refer to the ASOs employed as given in Table 1. 'Mutant' rows indicate surface-bound ASOs corresponding to the mutant allele; 'wild type' rows indicate surface-bound ASOs corresponding to the normal allele. (**b**) Fluorescence data obtained from slides hybridized with fluorescent single-stranded PCR products. The left column shows data from hybridization with 157 nt PCR products amplified from wild type or mutant plasmid cDNA clones. From top to bottom, the cDNA clones are pc*Tyr* (wild type) (46), pc*Tyr*P406L, pc*Tyr*G419R, pc*Tyr*R422Q, pc*Tyr*G446S, pc*Tyr*D448N (47). The right column shows data from hybridization with 347 nt PCR products amplified from human genomic DNA. From top to bottom, the genotypes are wt/wt, 406/406, 419/419, 422/wt, 446/373 and 448/wt. The precise nature of the mutations is shown in Table I and Figure 2. The data were processed by digital image filtering to reduce noise as described in Materials and Methods. Each spot contains 0.1−0.4 pmol of hybridized PCR product. The layout of the surface-bound ASOs is shown in Figure 4a.

large dark spots of many contiguous pixels. This difference in the nature of the true signal and background permits digital filtering to screen out the spurious background before quantification. This was accomplished by convolution of the raw data with a 3×3 median filter as described in Materials and Methods.

Table 2. Signal-to-noise ratios corresponding to data shown in Figure 4b, determined as described in Materials and Methods section

Source of DNA	Genotype	S/N (unfiltered)	S/N (filtered)
pcTyr	wild type	17.0	21.7
pcTyrP406L	Mutant 406	19.6	23.7
pcTyrG419R	Mutant 419	5.9	20.6
pcTyrR422Q	Mutant 422	14.5	19.9
pcTyrG446S	Mutant 446	8.4	23.7
pcTyrD448N	Mutant 448	16.1	23.6
Human Genomic DNA	wt/wt	9.6	15.7
	406/406	13.6	20.1
	419/419	17.4	20.1
	422/wt	18.2	26.8
	446/373	4.2	21.5
	448/wt	5.4	11.6

An advantage of this quantitative format is the ability to make objective statistical evaluations of data quality rather than relying upon subjective interpretations. To assess the quality of our experimental data, signal to noise ratios were calculated for each of the 12 frames shown in Figure 4b. This was first done on the original, unfiltered fluorescence images (images not shown), and then again on the filtered images. The results, shown in Table 2, clearly show that the digital filtering provides two advantages: a) higher signal to noise ratios due to the reduction of small spurious signals; and b) an increased uniformity of results presumably for the same reason. Based on these results, the 2:1 signal-to-noise detection limit of our hybridization experiments was calculated to be in the range of 10−40 fmoles of fluorescein-labeled PCR product per 3 mm spot. This is lower sensitivity than the instrumental detection limit determined above, reflecting the higher background levels in the hybridization experiment.

A final advantage of the quantitative nature of these data is that they may be easily stored on computer disk for archiving of the results, and later retrieved for further analysis and data manipulation. Various computer software programs can be used to add text labels to the images, re-scale and crop them, apply false colors and thresholds and even to display the data in alternative formats. Figure 5 shows a false colored, 3-dimensional representation of the data in which the height and color of the peaks along the Z-axis correspond to measured fluorescence intensity. This permits rapid visual identifications of positive hybridization signal.

DISCUSSION

We have described here a system for the direct fluorescence analysis of genetic polymorphisms by hybridization to glass support-bound arrays of allele specific oligonucleotides. Five single-base mutations contained within exon 4 of the human tyrosinase gene were studied as a test system; performing analyses in duplicate on each slide, and testing for the presence of both wild type and mutant sequences required construction of arrays of twenty oligonucleotides per slide. By limiting the hybridization sequence length to 15 nucleotides, and keeping GC content to between 5 and 7 nucleotides out of the fifteen total, a single set of hybridization and washing conditions were determined providing all-or-none discrimination between perfect matches and single-base mismatches.

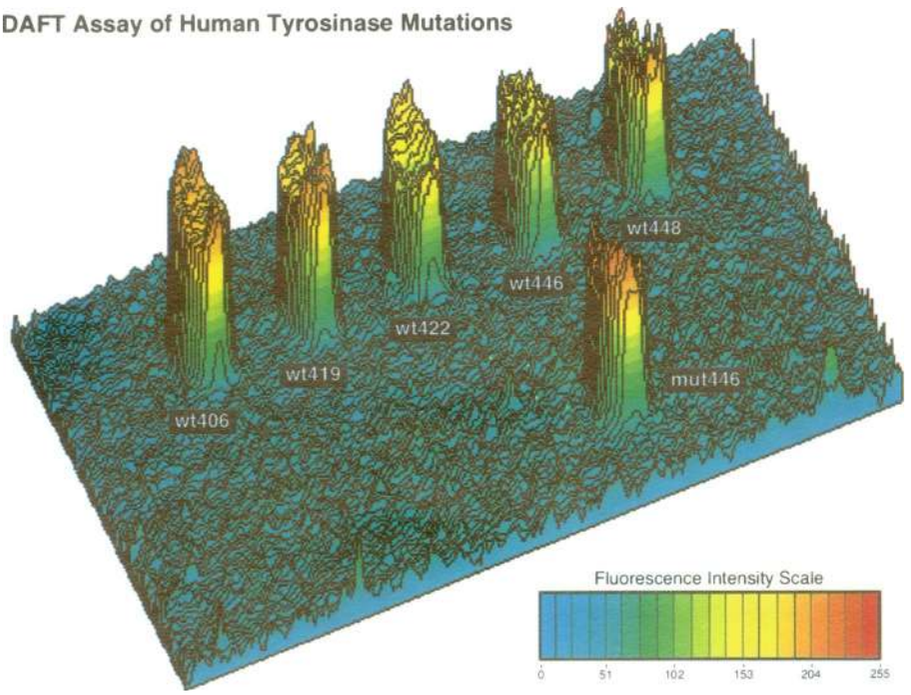


Figure 5. Three dimensional color representation of a portion of data from Figure 4b, right column, panel 5. The tyrosinase genotype analyzed is 446/373.

This system in its present form already may find substantial utility in a variety of genetic testing applications; further development of the system permitting the analysis of many more mutations in parallel is readily envisioned. This will probably require a) further increasing the complexity of the oligonucleotide array to include larger number of ASOs, b) using longer PCR products and c) simultaneous use of multiple PCR products corresponding to different regions of interest. Each of these improvements raises technical issues. The simple arrays used in the present study were prepared manually; the construction of substantially more complex arrays could be accomplished with robotic or photolithographic methodologies. Although longer PCR products will permit analyses of polymorphisms from larger regions, they are likely to possess higher degrees of secondary structure competing with ASO hybridization. The use of multiple PCR products at once will increase the burden of sample preparation. These and other issues will need to be addressed in the further development of this approach.

This having been said, the allure of these procedures is the possibility they present for the parallel analysis of hundreds or thousands of such polymorphisms. One can readily envision the commercial production of DNA 'chips' configured for tissue typing, cancer diagnosis, genetic identity testing, soil and environmental testing and many other applications. As shown here, such DNA 'chips' can be used to rapidly, quantitatively and non-isotopically yield the desired genetic information. It is our hope that the methods presented here will serve as a useful step towards this long-term goal.

ACKNOWLEDGEMENTS

We wish to recognize the generous support, advice and material provided by Dr Richard Spritz. Z.G. is a predoctoral trainee supported by NIH Molecular Biophysics training grant GM08293-05. A.T. is a predoctoral trainee supported by the Fannie and John Hertz Foundation. This work was supported by NSF grant DIR-8957582, DOE Human Genome Grant DE-FG02-91ER61122 and NIH Human Genome Grant 2R01HG00321.

REFERENCES

1. Cotton, R. G. H. and Malcolm, A. D. B. (1991) *Nature*, 353, 582–583.
2. Myers, R. M. and Maniatis, T. (1986) *Cold Spring Harbor Symp. Quant. Biol.*, 51, 275–284.
3. Landegen, U., Kaiser, R., Caskey, C. T. and Hood, L. (1988) *Science*, 242, 229–237.
4. Caskey, C. T. (1986) *Science*, 236, 1223–1229.
5. Wallace, B. R., Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T. and Itakura, K. (1979) *Nucl. Acids Res.*, 6, 3543–3557.
6. Wallace, B. R., Johnson, M. J., Hirose, T., Miyake, T., Kawashima, E. H. and Itakura, K. (1981) *Nucl. Acids Res.*, 9, 879–894.
7. Zhang, Y., Coyne, M. Y., Will, S. G., Levenson, C. H. and Kawasali, E. S. (1991) *Nucl. Acids Res.*, 19, 3929–3933.
8. Van Ness, J. and Chen, L. (1991) *Nucl. Acids Res.*, 19, 5143–5151.
9. Le Gall, I., Millasseau, P., Dausset, J. and Cohen, D. (1986) *Proc. Natl. Acad. Sci., USA*, 83, 7836–7840.
10. Lyonnet, S., Cailland, C., Rey, E., Berthelon, M., Frezal, J., Rey, J. and Munnich, A. (1988) *Lancet*, 507.
11. Cros, P., Allibert, P., Mandrand, B., Tiercy, J.-M. and Mach, B. (1992) *The Lancet*, 340, 879–874.
12. Saiki, R., Gelfand, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., Mullis, K. and Enrlich, H. (1988) *Science*, 239, 487–491.
13. Drmanac, R., Labat, I., Brukner, I. and Crkvenjakov, R. (1989) *Genomics*, 4, 114–128.
14. Bains, W. and Smith, G. C. (1988) *J. Theor. Biol.*, 135, 303–307.
15. Khrapko, K. R., Lysov, Y. P., Khorlyn, A. A., Shick, V. V., Florentiev, V. L. and Mirzabekov, A. D. (1989) *FEBS letters*, 256, 118–122.
16. Southern, E. M., Maskos, U. and Elder, J. K. (1992) *Genomics*, 13, 1008–1017.
17. Mirzabekov, A. D. (1994) *TIBTECH*, 12, 27–32.

18. Broude, N. E., Sano, T., Smith, C. L. and Cantor, C. R. (1994) *Proc. Natl. Acad. Sci., USA*, 91, 3072–3076.
19. Drmanac, R., Drmanac, S., Strezoska, Z., Paunesku, T., Labat, I., Zeremski, M., Snoddy, J., Funkhouser, W. K., Koop, B., Hood, L. and Crkvenjakov, R. (1993) *Science*, 260, 1649–1652.
20. Khrapko, K. R., Lysov, Y. P., Khorlin, A. A., Ivanov, I. B., Yershov, G. M., Vasilenko, S. K., Florentiev, V. L. and Mirzabekov, A. D. (1991) *DNA Sequence-J. DNA Sequencing and Mapping*, 1, 375–388.
21. (1989) Dynabeads M-280 Technical Handbook: Magnetic DNA Technology 6. Dynal Inc., Great Neck, New York.
22. Smith, L. M., Kaiser, R. J., Z., S. J. and Hoqd, L. E. (1987) *Methods in Enzymology*, 55, 260–301.
23. Sambrook, J., Fritsch, E. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
24. Lewis, R. (1990) Practical Digital Image Processing. Ellis Horwood Limited, West Sussex, England.
25. *NIH Image is a public domain software, written by Wagne Rasband at the U.S. National Institutes of Health, and available from the Internet by anonymous ftp from zippy.nimh.nih.gov or on floppy disk from NTIS, 5285 Port Royal Rd., Springfield, VA 22161, part number PB93-504868*, .
26. Saiki, R. K., Walsh, P. S., Levenson, C. H. and Erlich, H. A. (1989) *Proc. Natl. Acad. Sci., USA*, 86, 6230–6234.
27. Meinkoth, J. and Wahl, G. (1984) *Anal. Biochem.*, 138, 267–284.
28. Ghosh, S. S. and Musso, G. F. (1987) *Nucl. Acids Res.*, 15, 5353–5372.
29. Fahy, E., Davis, G. R., Dimichele, L. J. and Ghosh, S. S. (1993) *Nucl. Acids Res.*, 21, 1819–1826.
30. Rasmussen, S. R., Larsen, M. R. and Rasmussen, S. E. (1991) *Anal. Biochem.*, 198, 138–142.
31. Gingeras, T. R., Kwoh, D. Y. and Davis, G. R. (1987) *Nucl. Acids Res.*, 15, 5373–5390.
32. Lund, V., Schmid, R., Rickwood, D. and Hornes, E. (1988) *Nucl. Acids Res.*, 16, 10861–10880.
33. Maskos, U. and Southern, E. M. (1992) *Nucl. Acids Res.*, 20, 1679–1684.
34. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. and Solas, D. (1991) *Science*, 251, 767–773.
35. Kent, S., Hood, L., Aebersold, R., Teplow, D., Smith, L., Farnsworth, V., Cartier, P., Hines, W., Hughes, P. and Dodd, C. (1987) *BioTech.*, 5, 314–321.
36. Jacobs, J. W. and Fodor, S. P. A. (1994) *TIBTECH*, 12, 19–26.
37. Tripathi, R. K., Strunk, K. M., Giebel, L. B., Weleber, R. G. and Spritz, R. A. (1992) *Amer. J. Med. Genet.*, 43, 865–871.
38. Giebel, L. B., Strunk, K. M. and Spritz, R. A. (1991) *Genomics*, 9, 435–445.
39. Spritz, R. A., Strunk, K. M., Giebel, L. B. and King, R. A. (1990) *N. Engl. J. Med.*, 322, 1724–1728.
40. Wetmur, J. G. (1991) *Crit. Rev. Biochem. & Mol. Biol.*, 26, 227–259.
41. Breslauer, K. J., Frank, R., Blocker, H. and Marky, L. A. (1986) *Proc. Natl. Acad. Sci., USA*, 83, 3746–3750.
42. Southern, E. M., Case-Green, S. C., Elder, J. K., Johnson, M., Mir, K. U., Wang, L. and Williams, J. C. (1994) *Nucl. Acids Res.*, 22, 1368–1373.
43. Pease, A. C., Solas, D., Sullivan, E., Cronin, M. T., Holmes, C. P. and P.A., F. S. (1994) *Proc. Natl. Acad. Sci., USA*, 91, 5022–5026.
44. Cardullo, R., Agrawal, S., Flores, C., Zamecnik, P. C. and Wolf, D. E. (1988) *Proc. Natl. Acad. Sci., USA*, 85, 8790–8794.
45. Lee, L. G., Connell, C. R. and Bloch, W. (1993) *Nucl. Acids Res.*, 21, 3761–3766.
46. Bouchard, B., Fuller, B. B., Vijayasaradhi, S. and Houghton, A. N. (1989) *J. Exp. Med.*, 169, 2029–2042.
47. Tripathi, R. K., Hearing, V. J., Urabe, K., Aroca, P. and Spritz, R. A. (1992) *J. Biol. Chem.*, 267, 23707–23712.