ELSEVIER

Gene 188 (1997) 45–52

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Sequence analysis by hybridization with oligonucleotide microchip: identification of β-thalassemia mutations

Aleksei Drobyshev [a], Natalia Mologina [a], Valentine Shik [a], Diana Pobedimskaya [a], Gennadiy Yershov [a,b], Andrei Mirzabekov [a,b,*]

[a] *Joint Human Genome Program: Engelhardt Institute of Molecular Biology, 32 Vavilov Str., B-334 Moscow 117984, Russia*
[b] *Joint Human Genome Program, Center for Mechanistic Biology and Biotechnology, Argonne National Laboratory, 9700 S. Cass Ave., Argonne, IL 60439, USA*

Received 26 August 1996; revised 7 October 1996; accepted 16 October 1996; Received by D. Schlessinger

Abstract

Diagnostics for genetic diseases were run and sequence analysis of DNA was carried out by hybridization of RNA transcripts with oligonucleotide array microchips. Polyacrylamide gel pads ($100 \times 100 \times 20$ μm) were fixed on a glass slide of the microchip and contained allele-specific immobilized oligonucleotides (10-mers). The RNA transcripts of PCR-amplified genomic DNA were fluorescently labeled by enzymatic or chemical methods and hybridized with the microchips. The simultaneous measurement in real time of the hybridization and melting on the entire oligonucleotide array was carried out with a fluorescence microscope equipped with CCD camera. The monitoring of the hybridization specificity for duplexes with different stabilities and AT content was enhanced by its measurement at optimal, discrimination temperatures on melting curves. Microchip diagnostics were optimized by choosing the proper allele-specific oligonucleotides from among the set of overlapping oligomers. The accuracy of mutation detection can be increased by simultaneous hybridization of the microchip with two differently labeled samples and by parallel monitoring their hybridization with a multi-wavelength fluorescence microscope. The efficiency and reliability of the sequence analysis were demonstrated with diagnostics for β-thalassemia mutations.

Keywords: RNA; Diagnostics; Transcription; Discrimination temperature

1. Introduction

The number of discovered base changes in different genes is growing rapidly. These changes are associated with genetic diseases, with disease predispositions and cancers, with development of drug resistance in microorganisms, with genetic polymorphism, etc. The ability to simultaneously analyze many mutations in a gene in a simple, fast, and inexpensive way is essential in these fields and this need has stimulated the development of different methods for screening mutations.Hybridization of filter-immobilized DNA with allele-specific oligonucleotides (Conner et al., 1983) was suggested as an effective way for screening mutations. The development of an array of hundreds or thousands of immobilized oligonucleotides, the 'oligonucleotide chips', could extend this approach for simultaneous analysis of many mutations (for a review, see Mirzabekov, 1994). Such arrays can be manufactured by a parallel synthesis of oligonucleotides (Southern et al., 1992; Fodor et al., 1991; Pease et al., 1994; Matson et al., 1995) or by chemical immobilization of presynthesized oligonucleotides (Khrapko et al., 1989, 1991; Lamture et al., 1994; Ghu et al., 1994). Glass surfaces (Southern et al., 1992; Fodor et al., 1991; Ghu et al., 1994), glass pores (Beattie et al., 1995), polypropylene sheets (Matson et al., 1995), and gel pads (Khrapko et al., 1989, 1991) have been used as solid supports for oligonucleotide immobilization.Effective and precise sequence analysis by the hybridization of a probe with rather short microchip-immobilized oligonucleotides depends on many factors (Mirzabekov, 1994). The major factors are the reliability of the discrimination of perfect duplexes from duplexes containing mismatches, differences in stability of AT- and GC-rich duplexes, nucleic acid secondary

* Corresponding author at address b. Tel. +1 630 2523161; Fax +1 630 2523387; e-mail: amir@everest.bim.anl.gov

Abbreviations: bp, base pair(s); kb, kilobase(s) or 1000 bp; nt, nucleotide(s); TMR, tetramethylrhodamine.

PII S0378-1119(96)00775-5

A.

5′ $^{+1}$ACATTTGCTT CTGACACAAC TGTGTTCACT AGCAACCTCA AACAGACACC

$^{+51}$ATGGTGCACC TGACTCCTGA GGAGAAGTCT GCCGTTACTG CCCTGTGGGG

$^{+101}$CAAGGTGAAC GTGGATGAAG TTGGTGGT**G**^{26}A GGCCCTGGGC

$^{+141}$AG**G**^{1}**t**^{2}tg**g**^{5}**t**^{6}at caaggttaca agacaggttt aaggag 3′--

IVS1 ---------->

B.

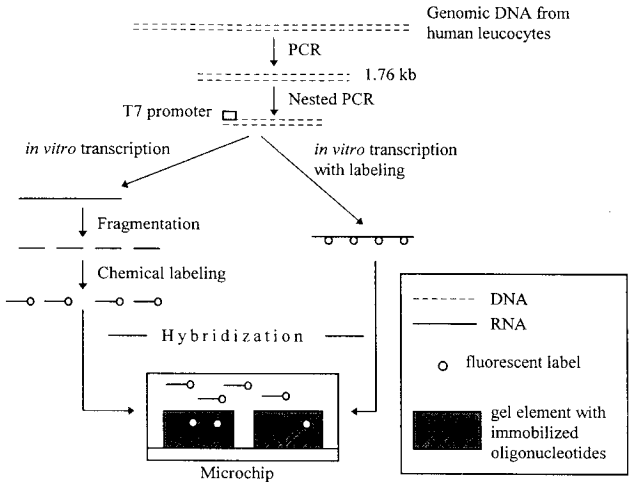

Fig. 1. Sequence analysis by hybridization of RNA transcripts with oligonucleotide microchips. (**A**) The sequence of the fragment of the β-globin gene from the first exon (Lawn et al., 1980). The bold letters and their superscripts indicate the point mutations and their positions in exon 1 (capital letters) and intron 1 (IVS1, lowercase letters after exon 1). The underlined sequences are complementary to the microchip-immobilized oligonucleotides (see Table 1). (**B**) PCR amplification of a 1.76 kb fragment of the human β-globin gene mapped from nucleotides −47 to +1714 (Lawn et al., 1980) was carried out with 1 μg genomic DNA (Poncz et al., 1982) and 50 pmol each of the forward primer 5′-GGAGCCAGGGCTGGGCATAAAAGT-3′ (−47→−23) and the reverse primer 5′-ATTTTCCCAAGGTTTGAACTAGCTC-3′ (+1689→+1714). The amplification was carried out in a DNA thermal cycler (GeneAmp PCR System 2400, Perkin Elmer Corporation) in 100 μl of a buffer containing 200 mM each of dATP, dCTP, dGTP, dTTP, 2.5 mM $MgCl_2$, 2 units of *Taq* DNA polymerase (BioMaster, Russia), 50 mM KCl, 10 mM Tris-HCl, pH 9.0, and 0.1% Triton X-100. The reaction conditions were 30 cycles, with 45 s at 95°C, 90 s at 66°C, and 120 s at 72°C. PCR product was purified from 2% low gel/melting temperature agarose gel (NuSieve agarose, FMC). The 159 bp and 102 bp DNA fragments were amplified with 10 ng of the 1.76 kb DNA with three nested primers, two containing T7 promoter sequence and a common reverse primer (see **A**). The nested primers were T7-V2L-45 (5′-GGAATTCCTAATACGACTCACTATAGGGA CACCATGGTGCACCTGACTCC-3′; +45→+67); T7-V2L-103 (5′-GGAATTCCTAATACGACTCACTATAGGGAGGTGAACGTGGATGAAGTTGG-3′; +103→+124); and 5′-TCTCCTTAAACCTGTCTTGTAACC-3′ (common reverse; +154→+177). The amplification was carried out in 25 cycles (15 s at 95°C, 30 s at 62°C, and 30 s at 72°C). PCR products were purified by QIAGEN QIAquick PCR Purification Kit. The PCR-amplified 159 or 102 bp DNA (4–5 μg) containing T7 promoter was transcribed with 400 units of T7 RNA polymerase (Promega) to produce 133 and 75 nt long RNA in 80 μl of buffer containing 300 mM Hepes, pH 7.6, 30 mM $MgCl_2$, 16 μg of BSA, 40 mM DTT, 30 units of RNasin (Promega) and 4 mM each of ATP, CTP, GTP, and UTP for 3 h at 38°C. Deproteinization of the reaction mixture was carried out in 20 mM EGTA, pH 8.0, 2% SDS, and Proteinase K (10 mg/ml) for 15 min at 37°C. The mixture was extracted first with equal volumes of phenol and then with equal volumes of chloroform, precipitated twice by 1 vol. of isopropyl alcohol, from 0.5 M $LiClO_4$ and dissolved on a Bio-Spin P6 column (BioRad). Fragmentation of 10–100 μg of RNA to an average length of 20–40-mers was carried out in 50 μl of 0.1 M KOH for 30 min at 40°C. Then 5 μl of 1 M Hepes, pH 7.6, and 15 μl of 1% $HClO_4$ were added at 4°C. The pellet of potassium perchlorate was removed by centrifugation and RNA was precipitated by 10 vol. of 2% $LiClO_4$ in acetone. The RNA was washed twice with acetone and dried for 20–30 min at room temperature. The fragmented RNA was dephosphorylated in 50 μl of 20 mM Tris-HCl, pH 8.0, 1 mM $MgCl_2$, 1 mM $ZnCl_2$, 10 units of RNasin, 5–7 units of calf intestine phosphatase (CIP) for 1 h at 37°C. RNA deproteinization and purification were carried out as described above. For chemical fluorescence labeling of RNA (Prudnikov and Mirzabekov, 1996), its 3′-terminal dephosphorylated nucleoside was oxidized in 20 μl of 10 mM sodium periodate for 20 min at room temperature. RNA was precipitated with acetone. A 10 molar excess of 10 mM TMR-hydrazine in 10% acetonitrile was added to oxidized RNA fragments in 20 μl of 20 mM sodium acetate, pH 4.0. The reaction mixture was incubated 30–40 min at 37°C, and the hydrazide bond between the RNA and dye was stabilized by reduction with freshly prepared 1.5 μl of 0.2 M $NaCNBH_3$ and incubation for 30 min at room temperature. Then mixture was extracted four times with water saturated *n*-butanol and precipitated with acetone. Alternatively, RNA was labeled by incorporation of fluorescein-UTP during the transcription with Ambion MEGAshortscript kit according to the manual.

structure, the efficiency of the hybridization, and simplicity in the preparation of the labeled samples for hybridization. Optimization of these factors is crucial for the application of microchips as diagnostics and for sequence analysis, and, in particular, for de novo DNA sequencing. Here we describe further developments that address these factors in the application of microchips containing gel-immobilized oligonucleotides (Khrapko et al., 1989, 1991; Yershov et al., 1996) for sequence analysis and for diagnostics. It was demonstrated that identification of base variations is significantly improved by parallel measuring of the melting curves of the duplexes formed on the entire oligonucleotide array, as well as by monitoring the simultaneous hybridization of two differently labeled samples at two wavelengths and by choosing proper allele-specific oligonucleotides.

2. Results and discussion

2.1. Oligonucleotide microchip

A number of the most commonly occurring β-thalassemia mutations within β-globin gene (Fig. 1A) have been tested as diagnostics using hybridization with oligonucleotide microchips. These mutations are splice-site mutations for the first, second, fifth, and sixth nt in the first intron (IVS I) of the β-globin gene: IVS I/1 G/A (G/A = substitution of G for A), IVS I/2 T/C, IVS I/5 G/T, IVS I/5 G/C, IVS I/6 T/C, and G/A substitution

Table 1
The sequence of the microchip allele-specific 10-mers

	#	Allele	Position of mutated base	Sequence	Location
	1	IVS (N)	-	5′-A TAC CAA CCT-gel	+141
M	2	IVS I/1 G/A	8	5′-A TAC CAA tCT-gel	+141
I	3	IVS I/1 G/T	8	5′-A TAC CAA aCT-gel	+141
C	4	IVS I/2 T/A	7	5′-A TAC Cat CCT-gel	+141
R	5	IVS I/2 T/C	7	5′-A TAC Cag CCT-gel	+141
O	6	IVS I/2 T/G	7	5′-A TAC Cac CCT-gel	+141
C	7	IVS I/5 G/A	4	5′-A TAt CAA CCT-gel	+141
H	8	IVS I/5 G/C	4	5′-A TAg CAA CCT-gel	+141
I	9	IVS I/5 G/T	4	5′-A TAa CAA CCT-gel	+141
P	10	IVS I/6 T/C	3	5′-A TgC CAA CCT-gel	+141
	11	CD 26 (N)	-	5′-G GCC TCA CCA-gel	+125
I	12	CD 26 G/A	6	5′-G GCC TtA CCA-gel	+125
	1	IVS (N)	9	5′-TGA TAC CAA C-gel	+143
	2	IVS I/2 T/G	9	5′-TGA TAC CAc C-gel	+143
M	3	IVS (N)	8	5′-GA TAC CAA CC-gel	+142
I	4	IVS I/2 T/G	8	5′-GA TAC CAc CC-gel	+142
C	5	IVS (N)	7	5′-A TAC CAA CCT-gel	+141
R	6	IVS I/2 T/G	7	5′-A TAC Cac CCT-gel	+141
O	7	IVS (N)	6	5′-TAC CAA CCT G-gel	+140
C	8	IVS I/2 T/G	6	5′-TAC CAc CCT G-gel	+140
H	9	IVS (N)	5	5′-AC CAA CCT GC-gel	+139
I	10	IVS I/2 T/G	5	5′-AC CAc CCT GC-gel	+139
P	11	IVS (N)	4	5′-C CAA CCT GCC-gel	+138
	12	IVS I/2 T/G	4	5′-C CAc CCT GCC-gel	+138
II	13	IVS (N)	3	5′-CAA CCT GCC C-gel	+137
	14	IVS I/2 T/G	3	5′-CAc CCT GCC C-gel	+137
	15	IVS (N)	2	5′-AA CCT GCC CA-gel	+136
	16	IVS I/2 T/G	2	5′-Ac CCT GCC CA-gel	+136

The oligonucleotides of microchip I are complementary to the coding strand of DNA of the β-globin gene of patients with β-thalassemia single-base mutations (G/A, substitution of A for G) in the first, second, fifth, or sixth nt of the first intron (IVS I/1, 2, 5, 6) of the β-globin gene and in the codon #26 (CD26) of the first exon. Oligonucleotides 1–16 of microchip II correspond to the normal and IVS I/2 G/T allele. The mutated and corresponding normal bases are placed from the second to the ninth positions of the 10-mers from their 3′ end. The mutated bases are shown in lowercase bold letters and corresponding oligonucleotide bases in the normal allele are underscored. The oligonucleotide synthesis and the microchip manufacturing were described by Yershov et al. (1996)

in the 26th codon (GAG) of the first exon (also known as abnormal hemoglobin E) (Diaz-Chico et al., 1988).A microchip with $100 \times 100 \times 20$ μm gel elements (Yershov et al., 1996) contains immobilized decadeoxyribonucleotides, 10-mers that correspond to normal and mutant β-thalassemia alleles. These 10-mers discriminate mismatches less reliably, but they are hybridized more efficiently than 8-mers, and 10-mers are thus used here. Some experiments were repeated, with the same results, on microchips with $60 \times 60 \times 20$ μm gel elements (not shown). Table 1 shows the sequence of the allele-specific oligonucleotides immobilized on the microchips. Mismatches within the duplexes have a much higher destabilization effect than mismatches at the terminal positions (Khrapko et al., 1991); therefore the mutated bases were placed inside of the immobilized oligonucleotides.

2.2. RNA samples

In the previous report (Yershov et al., 1996), single- and double-stranded PCR-amplified β-globin DNA fragments (of different length and after a random fragmentation) were tested for identification of some of these mutations. However, the hybridization of RNA offers some advantages over DNA hybridization. RNA fragments were derived from PCR-amplified genomic DNA by transcription with T7 RNA polymerase (Lipshutz et al., 1995). About 100 copies of unlabeled or fluorescently labeled RNA transcripts are synthesized per DNA molecule, providing a convenient way to prepare the sufficient amount of hybridization probes. RNA can be fragmented and one fluorescent dye molecule can be introduced per fragment (see Fig. 1).

The experimental scheme for preparation of fluorescently labeled RNA probes is shown in Fig. 1. A DNA fragment of the β-globin gene, 1.76 kb in length (from -47 to $+1714$), was amplified from genomic DNA isolated from human leukocytes. Then other sets of primers containing T7 RNA polymerase promoter were used for nested PCR amplification of 159 bp and 102 bp DNA fragments. Unlabeled or fluorescently labeled RNA transcripts of 133 and 75 nt long were synthesized using T7 RNA polymerase. The fluo-

rescently labeled RNA was directly used for hybridization with the microchips. Unlabeled RNA was fragmented with KOH, chemically labeled at the 3′-terminal dephosphorylated nucleoside of the fragments (see Fig. 1), and then hybridized with the microchips.

2.3. Hybridization

Polyacrylamide gel provides a stable three-dimensional support for immobilized oligonucleotides, shows negligible fluorescence background, and allows hybridized RNA to be easily washed off. Therefore, the oligonucleotide microchips can be dried, kept for months, and used repeatedly (15–50 times) for hybridization. The gel-immobilized oligonucleotides are well spaced and do not interfere with each other or with DNA upon hybridization. The hybridization within the gel looks more like a liquid-phase than a solid-phase reaction. These factors enhance the discrimination of perfect duplexes from mismatched ones.The hybridization of fluorescently labeled RNA with the microchips was carried out at 0°C for 18 h. Fragmented RNA and 19-mers can be hybridized for 30 min. In many cases, the intensities of the hybridization signals at 0°C were similar for perfect and mismatched duplexes. The perfect and mismatched duplexes as well as the duplexes having various GC and AT contents displayed different stabilities and therefore should be tested at different temperatures. The polyacrylamide gel provides more than 100-times higher capacity for three-dimensional immobilization of oligonucleotides than does a two-dimensional glass surface. The high concentration of immobilized oligonucleotides facilitates the discrimination of mismatched duplexes and enhances the sensitivity of measurements on the microchips. This allows the use of a CCD-camera-equipped fluorescence microscope (Yershov et al., 1996) although it is less sensitive than laser scanning systems (Lipshutz et al., 1995) but offers the advantage of monitoring the hybridization on a microchip at different temperatures in real time for measurement of the melting curves. The procedure, the software, and the hybridization microchamber (Yershov et al., 1996) have all been developed for recording melting curves at a wide range of temperatures simultaneously for perfect and mismatched duplexes formed upon hybridization of a probe with all microchip oligonucleotides.A significant amount of time is needed for the microchips hybridized with rather long RNA or DNA probes to achieve equilibrium. Therefore, non-equilibrium dissociation melting curves were measured. However, they are not far away from equilibrium and some difference in heating rate did not significantly affect the results. The melting curves for hybridization of synthetic 19-mers with the microchip oligonucleotides reached equilibrium under these conditions. The melting curves can also be measured after a few minutes, far away from equilibrium, if an internal standard is added to a tested sample. This standard can be differently labeled RNA of a normal allele. This speeds up significantly the identification of base changes (not shown). Fig. 2 shows an example of four melting curves for 75 nt long RNA fragments hybridized with the microchip oligonucleotides. The RNA was derived from a patient having the IVS I/2 T/A mutation. The curves were normalized to the initial hybridization signals. Melting curves 1 and 3 in Fig. 2 correspond to perfect duplexes; curves 2 and 4 correspond to duplexes containing internal T-T or G-T mismatches, respectively. The curves for the perfect and mismatched duplexes are shifted by about 10°C from each other.It appears that the greatest discrimination between perfect and mismatched duplexes can be achieved at a temperature at which the intensity of the hybridization signal from a perfect duplex drops to one-tenth of its initial value; at such a temperature, the hybridization intensities from mismatched duplexes usually approach the background level. We define the temperature at which the initial signal of hybridization drops by a factor of 10 as the discrimination temperature, T_d.The T_d values for perfect 40% and 70% GC-rich duplexes (curves 1 and 3 in Fig. 2A and B) are 52 and 64°C, respectively. One needs to consider that the dissociation curves for perfect and mismatched duplexes are parallel at the range of about 10°C (in the middle of the curves) when plotted on a semilogarithmic scale (Fig. 2). At this 10°C range, the ratios of the signals for perfect and mismatched duplexes remain rather constant. This makes the discrimination procedure robust to some inaccuracies in determining T_d. The discrimination temperature depends on experimental conditions (rate of heating, ionic strength, probe concentration, extent of fragmentation, etc.) which can vary from one experiment to another. However, these variations affect T_d and the relative intensities of the hybridization signals to a similar extent for all microchip elements and therefore do not significantly distort the discriminations. Therefore, to provide a reference T_d, the oligonucleotides CD26(N) and CD26 G/A, which form perfect and mismatched duplexes, respectively, with all RNAs tested (Table 2), are introduced into the microchip.The T_d can be determined by hybridization with an RNA sample if an allele DNA is available. If such DNA is unavailable, the T_d can be measured from the hybridization data with synthetic oligonucleotides corresponding to the mutated allele. Since T_d is robust to some inaccuracy in measurements, 19-mer oligodeoxynucleotides have been used in these experiments instead of more expensive oligoribonucleotides. One should, however, take into account differences in the stability of DNA-DNA homoduplexes relative to DNA-RNA heteroduplexes (Lesnik and Freier, 1995). The data in Table 2 show that the pattern of hybridization of the

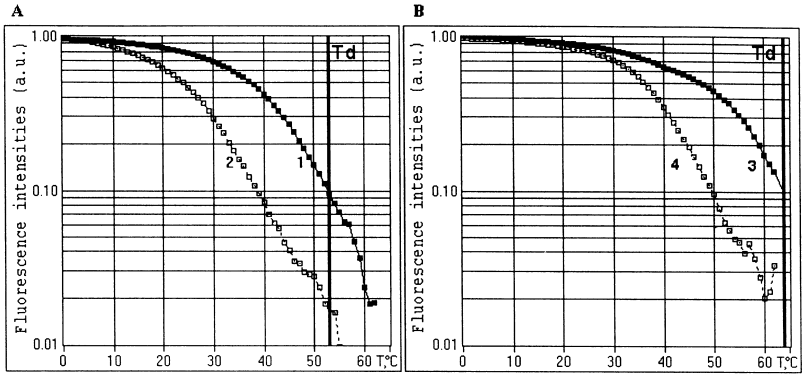

Fig. 2. Non-equilibrium melting curves of duplexes of RNA with microchip oligonucleotides. 75 nt long RNA enzymatically labeled with fluorescein was synthesized from the genomic DNA of a homozygote IVS I-2 T/A and CD26(N) β-thalassemia patient (Table 2, probe 2a). The RNA was hybridized with the following microchip-immobilized 10-mers: 1, IVS I/2 T/A; 2, IVS(N); 3, CD26(N); 4, CD26 G/A to produce perfect, 1 and 3, and mismatched duplexes, 2 and 4. The oligonucleotide microchip is an array of 12 polyacrylamide gel pads ($100 \times 100 \times 20$ μm) fixed on a hydrophobic glass surface and spaced by 200 μm (Yershov et al., 1996). Each gel pad contains chemically bonded 10-mers (see Table 1). The hybridization of RNA (1 pmol/μl) with the microchip was carried out in 5 μl of the hybridization buffer (1 M NaCl, 5 mM EDTA, 1% Tween-20, 10 mM $NaHPO_4$, pH 6.8) for 18 h at 0°C in a closed hybridization chamber. The melting curves were monitored simultaneously for all microchip elements by gradually increasing the temperature (1°C/min). The hybridization signals were measured at 1°C intervals for 1–5 s each time in parallel for all microchip elements using a fluorescence microscope equipped with a CCD camera and two sets of filters for fluorescein and TMR (Yershov et al., 1996). T_d, discrimination temperature; a.u., arbitrary units.

microchip with RNA derived from patients and with 19-mers was rather similar. Hybridization with corresponding synthetic oligonucleotides can be recommended as a control when a mutation is identified in an RNA sample by its hybridization with the diagnostics microchip.

2.4. Diagnostics

Table 2 summarizes the results of hybridization of the diagnostics microchips with (1) RNA probes derived from a number of homozygote and heterozygote β-thalassemia patients and (2) with corresponding 19-mers. The table shows the T_d for perfect duplexes formed on each microchip oligonucleotide. The relative intensities, R, of the hybridization signals for different microchip oligonucleotides in Table 2 are normalized to the signals for a perfect duplex at the T_d (estimated as 1.0). In most cases the ratios for mismatched duplexes are less than 0.1 and close to 0. These values are low enough to allow unambiguous identification of the homozygous and heterozygous mutations in patients at the T_d (when the hybridization signals only from perfect duplexes are observed). The hybridization of homozygote RNA (Table 2, probes 1a, 2a, 2b, and 3a) with the microchip shows the distinctive formation of a perfect duplex only with one immobilized oligonucleotide and mismatched duplexes with all others. Two perfect duplexes were unambiguously identified upon hybridization with a heterozygote RNA (Table 2, probe 4a).The noticeable exceptions are oligonucleotides corresponding to IVS 1/2 T/A and IVS 1/2 T/G mutations that show strong mismatched signals upon hybridization with non-corresponding samples of IVS(N) and IVS 1/2 T/A RNAs, respectively (Table 2, Exp. 1a, 2b, 4a, 6a and 6b). The relative intensities of these mismatched signals can be significantly decreased by choosing the proper oligonucleotides for immobilization (see Section 2.6). It appears that the diagnostics can be carried out with RNAs 75 nt long (Table 2, probes 1a, and 6a), and 133 nt long (probes 2b, and 6b), as well as with 133 nt long RNA fragmented to pieces 20–40 nt long (probes 3a, and 4a). However, the intensities of the hybridization signals after fragmentation are increased by about 5 times and the time of hybridization is decreased from several hours to tens of minutes (not shown).The longer RNA probes diffuse more slowly into the gel and can form stable secondary structures or aggregates. These factors interfere with their hybridization with rather short immobilized oligonucleotides. Thus, the fragmentation seems to be an essential step in sample preparation, since it enhances and speeds the hybridization.

2.5. Hybridization with two differently labeled probes

In addition to the measuring of the melting curves, the reliability of identification of mutations and base changes can be enhanced by the use of a multicolor fluorescence microscope (Yershov et al., 1996). For this purpose, the tested RNA is marked by one fluorescence label and hybridized with a microchip in the presence of a normal allele sample labeled with a different dye. The pattern and the ratio of hybridization measured with the two dyes will be similar for all microchip oligonucleotides except for those that correspond to different allele bases, i.e., mutations. Table 2 shows the results of such an experiment. Microchip I was hybrid-

Table 2
Identification of β-thalassemia mutations by hybridization with the microchip

21nHybridized probe				Immobilized 10-mer oligonucleotide											
No.		Allele	Size (nt)	IVS(N)	I/1 G/A	I/1 G/T	I/2 T/A	I/2 T/C	I/2 T/G	I/5 G/A	I/5 G/G	I/5 G/T	I/6 T/C	CD26(N)	CD26 G/A
				R at T_d =											
				42°C	39°C	38.5°C	42°C	48°C	45.5°C	37°C	44.5°C	40°C	50°C	54.5°C	49°C
1	a	IVS(N)	75	1.00	0.04	0	0.20	0.05	0.07	0	0	0	0.04	1.00	–
	b	IVS(N)	19[a]	1.00	0.09	0.07	0.02	0.03	0.01	0.03	0.03	0.07	ND	0	0
	a	IVS I/2 T/A	*Fl* 75	0.15	0	0	1.00	0.12	0.08	0	0	0	0	1.00	–
2	b	IVS I/2 T/A	*Fl* 133	0.03	0	0	1.00	0	0.30	0	0	0	0	1.00	0.19
	c	IVS I/2 T/A	19[b]	0.01	0	0	1.00	0.07	0.03	0	0	0	0	0.01	0
3	a	IVS I/1 G/A	133*fr*	0.03	1.00	0	0.01	0	0.02	0	0	0	0	1.00	–
	b	IVS I/1 G/A	9[c]	0.01	1.00	0.01	0.01	0	0	0.01	0	0	0	0	0
4	a	IVS I/1 G/A and IVS I/6 T/C	133*fr*	0.2	0.85	0	0.2	0	0.05	0	0	0	1.00	1.00	–
	b	IVS I/1 G/A	19[c]	0.01	1.00	0.01	0.01	0	0	0.01	0	0	0	0	0
	c	IVS I/6 T/C	19[d]	0.1	0	0	0	0	0	0	0	0	1.00	0	0
5	a	IVS I/5 G/T	19[e]	0	0	0	0	0	0	0.03	0.02	1.00	0	0	0
	b	CD26(N)	19[f]	0	0	0	0	0	0	0	0	0	0	1.00	0.03
	c	CD26 G/A	19[g]	0.03	0	0	0	0	0	0	0	0	0	0.04	1.00
6	a	IVS(N)	75	1.00	0.04	0	0.20	0.05	0.07	0	0	0	0.04	1.00	–
	b	IVS I/2 T/A	*Fl* 133	0.03	0	0	1.00	0	0.30	0	0	0	0	1.00	–

Microchip I (see Fig. 2) was successively hybridized with RNA 75 and 133 nt long without fragmentation or after fragmentation (133*fr*, probes 3a and 4a) and with six synthetic 19-mer oligodeoxyribonucleotides corresponding to β-thalassemia mutations. The RNA and 19-mers were labeled with TMR except for RNA probes 2a, 2b, and 6b, which were labeled with fluorescein (*Fl*). The melting curves (Fig. 2) were measured simultaneously for all microchip oligonucleotides at each hybridization. These curves provided values of hybridization intensities at the discrimination temperature, T_d. *R* is the ratio of the hybridization signal of a mismatched duplex (*Im*) to the signal of the perfect duplex (*Ip*) estimated at T_d in parallel for all microchip oligonucleotides. $R = Im/Ip$. 19[a–g]-synthetic 19-deoxymers were complementary to allele-specific 10-mers immobilized on the microchips.
[a]IVS(N): 5′-TMR-CCTGGGCAGGTTGGTATCA-3′; [b]IVS I/2 T/A: 5′-TMR-CCTGGGCAGG**a**TGGTATCA-3′; [c]IVS I/1 G/A: 5′-TMR-CCTGGGCAG**a**TTGGTATCA-3′; [d]IVS I/6 T/C: 5′-TMR-CCTGGGCAGGTTGG**c**ATCA-3′; [e]IVS I/5 G/T: 5′-TMR-CCTGGGCAGGTTG**t**TATCA-3′; [f]CD26(N): 5′-TMR-GTTGGTGGTGAGGCCCTGG-3′; [g]CD26 G/A: 5′-TMR-GTTGGTGGTaAGGCCCTGG-3′.

Table 3
The effect of the position of the allele base within 10-mers on mutation detection

Position of allele base	19-mer			RNA		
	$T_{0.1}$ of perfect duplex IVS(N)	$T_{0.1}$ of (G-A) mismatched duplex IVS I/2 T/G	$\Delta T_{0.1}$	$T_{0.1}$ of perfect duplex IVS(N)	$T_{0.1}$ of (G-A) mismatched duplex IVS I/2 T/G	$\Delta T_{0.1}$
9	40	32	8	35	37	−2
8	47	32	15	49	38	11
7	42	30	12	44	41	3
6	47	28	19	49	41	8
5	52	38	14	50	42	8
4	54	39	15	54	44	10
3	55	46	9	59	54	5
2	52	46	6	58	53	5

Microchip II contains two sets of 10-mers corresponding to the normal and IVS I/2 T/G alleles. The microchip was hybridized with the TMR-labeled normal allele 19-mer and to an RNA 75 nt long. $T_{0.1}$ is the temperature at which the hybridization signals for a microchip duplex drops to one-tenth of its initial value at 0°C. $\Delta T_{0.1} = T_{0.1}$ (a perfect duplex) minus $T_{0.1}$ (the corresponding mismatched duplex).

ized with a mixture of TMR-labeled, wild-type homozygote, 75 nt RNA (Table 2, probe 6a) and fluorescein-labeled, IVS I/2/ T/A homozygote, 133 nt RNA (Table 2, probe 6b). The patterns of hybridization detected at two wavelengths are very similar except for the normal and IVS I/2 T/A allele-specific oligonucleotides that indicate the presence of the corresponding mutation.

2.6. Optimal immobilized oligonucleotides

As shown in Table 2, the immobilized 10-mers matching mutations IVS I-2 T/G, IVS I-2 T/C, and IVS I-2 T/A are hybridized rather strongly with some RNA probes that correspond to other alleles. Different structural factors in RNA could cause this hybridization. One can minimize the effect of these factors by placing a variable IVS I-2 base into different positions of the 10-mers. The results of such experiments are shown in Table 3. Microchip II was successively hybridized with fragmented 75 nt long RNA or with a synthetic DNA 19-mer, both corresponding to the normal allele. Microchip II contained two similar sets of eight overlapped immobilized 10-mers that are complementary either to a normal allele or to IVS I-2 T/G allele (Table 1). The allele-specific bases A for the first set and C for the second set are located in these 10-mers in all internal positions from the second to the ninth. These bases form perfect A-T or mismatched A-G base pairs, respectively. As in Table 2, the stability of the perfect and mismatched duplexes formed on the microchip is determined as $T_{0.1}$, the temperature at which the initial hybridization signal of the duplex is decreased to one-tenth of the original intensity. $\Delta T_{0.1}$ corresponds to the difference in $T_{0.1}$ between the perfect and similar mismatched duplexes. A better discrimination of the perfect and mismatched duplexes is reflected in higher values of $\Delta T_{0.1}$. The discrimination efficiency (ΔT) was lower for hybridized RNAs than for the 19-mers. The discrimination was surprisingly low, $\Delta T = -2$°C and 3°C, when the allelic bases were placed at the ninth or seventh position, respectively, of the immobilized oligonucleotides. It appears that secondary structures and the presence of similar sequences in other regions of the RNA cause this lowering (not shown). These effects can be partly predicted from the sequence of the region that is searched for mutations. However, it is possible to reach a high discrimination ($\Delta T = 8$–11°C) when allele bases are placed in other positions, for example the 8, 6, 5, or 4 positions.

3. Conclusions

(1) The hybridization of RNA transcripts of PCR-amplified DNA with oligonucleotide microchips allows the reliable identification of base changes and homozygous and heterozygous β-thalassemia mutations in the genomic DNA of patients.
(2) The efficiency of the identification was significantly increased by: (1) measuring melting curves and comparing the hybridization signals for perfect and mismatched duplexes having different AT contents at a discrimination temperature specific for each duplex; (2) choosing the proper immobilized oligonucleotides from among the set of overlapping ones; (3) using a multi-wavelength fluorescence microscope and carrying out the hybridization of the microchips with a mixture of a mutated allele RNA and an internal standard of (a normal allele sample) stained with different dyes.
(3) The RNA transcription of PCR-amplified DNA provides an easier method for preparing a sufficient amount of labeled, single-stranded samples than the use of DNA PCR amplification.

Acknowledgement

We are grateful to Drs. Yu. Lysov, A. Perov, E. Timofeev, V. Barsky, E. Kreindlin, Mrs. V. Chupeeva, N. Kalganova, and J. Lebed for their contribution to this work. We are also grateful to Prof. T.H.J. Huisman for genomic DNA samples. This work was supported by grants 558 and 562 of the Russian Human Genome Program, by grant 96-04-49-858 of the Russian Basic Science Foundation, and by grant DE-FG02-93ER61538 of the US Department of Energy. We thank D.E. Nadziejka and L. Novikova for additional assistance.

References

Beattie, K.L., Beattie, W.G., Meng, L., Turner, S.L., Coral-Vazquez, R., Smith, D.D., McIntyre, P.M. and Dao, D.D. (1995) Advances in genosensor research. Clin. Chem. 41, 700–706.

Conner, B.J., Reyers, A.A., Morin, C., Itakura, K., Teplitz, R.L. and Wallace, R.B. (1983) Detection of sickle cell β-globin allele by hybridization with synthetic oligonucleotides. Proc. Natl. Acad. Sci. USA 80, 278–282.

Diaz-Chico, J.C., Yang, K.-G., Yang, K.-Y., Efremov, D.G., Stoming, T.A. and Huisman, T.H.J. (1988) Biochim. Biophys. Acta 949, 43–48.

Fodor, S.P.A., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T. and Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. Science 251, 767–773.

Ghu, Z., Guilfoyle, R.A., Thiel, A.J., Wang, R. and Smith, L.M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass support. Nucleic Acids Res. 22, 5456–5465.

Khrapko, K.A., Lysov, Yu., Khorlin, A., Shick, V., Florentiev, V. and Mirzabekov, A. (1989) Hybridization of oligonucleotides as a method of DNA sequencing. FEBS Letters 258, 118–122.

Khrapko, K., Lysov, Yu., Khorlin, A., Ivanov, I., Yershov, G., Vasilenko, S., Florentiev, V. and Mirzabekov, A. (1991) A method for DNA sequencing by hybridization with oligonucleotide matrix. J. DNA Sequencing 1, 375–388.

Lamture, J.B., Beattie, K.L., Burke, B.E., Eggers, M.D., Ehrlich, D.J., Fowle, R., Hollis, M.A., Kosicki, B.B., Reich, R.K., Smith, S.R., Varma, R.S. and Hogan, M.E. (1994) Direct detection of nucleic acid hybridization on the surface of a charge-coupled device. Nucleic Acids Res. 22, 2121–2125.

Lawn, R., Efstratiadis, A., O'Connel, C. and Maniatis, T. (1980) The nucleotide sequence of the human β-globin gene. Cell 21, 647–651.

Lesnik, E.A. and Freier, S. (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. Biochemistry 34, 10807–10815.

Lipshutz, R.J., Morris, D., Chee, M., Hubbell, E., Kozal, M.J., Shah, N., Shen, N., Yang, R. and Fodor, S.P.A. (1995) Using oligonucleotide probe arrays to access genetic diversity. BioTechniques 19, 442–447.

Matson, R.S., Rampal, J., Pentoney, S.L., Anderson, P.D. and Coassin, P. (1995) Biopolymer synthesis on polypropylene support: oligonucleotide arrays. Anal. Biochem. 224, 110–116.

Mirzabekov, A. (1994) DNA sequencing by hybridization: a megasequencing method and a diagnostic tool? Trends Biotechnol. 12, 27–32.

Pease, A.C., Solar, D., Sullivan, E.J., Cronin, M.T., Holmes, C.P. and Fodor, S.P.A. (1994) Light-directed oligonucleotide arrays for rapid DNA sequence analysis. Proc. Natl. Acad. Sci. USA 91, 5022–5026.

Poncz, M., Solowiejczyk, D. and Harpel, B. (1982) Construction of human gene libraries from small amount of peripheral blood: analysis of β-like globin genes. Hemoglobin 6, 27–40.

Prudnikov, D. and Mirzabekov, A. (1996) Chemical methods of DNA and RNA fluorescent labeling. Nucleic Acids Res. 24, 4535–4542.

Southern, E.M., Maskos, U. and Elder, J.K. (1992) Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. Genomics 13, 1008–1017.

Yershov, G., Barsky, V., Belgovskiy, A., Kirillov, E., Kreindlin, E., Ivanov, I., Parinov, S., Gushin, D., Drobyshev, A., Dubiley, S. and Mirzabekov, A. (1996) DNA analysis and diagnostics on oligonucleotide microchips. Proc. Natl. Acad. Sci. USA 93, 4913–4918.