

## ARTICLE

# High throughput parallel analysis of hundreds of patient samples for more than 100 mutations in multiple disease genes

Anthony P. Shuber<sup>†</sup>, Lesley A. Michalowsky, G. Scott Nass, Joel Skoletsky, Lisa M. Hire, Steve K. Kotsopoulos, Michael F. Phipps, Dana M. Barberio and Katherine W. Klinger<sup>\*</sup>

Department of Technology Development, Genzyme Genetics, Framingham, MA 01701-9322, USA

Received September 9, 1996; Revised and Accepted November 29, 1996

As more mutations are identified in genes of known sequence, there is a crucial need in the areas of medical genetics and genome analysis for rapid, accurate and cost-effective methods of mutation detection. We have developed a multiplex allele-specific diagnostic assay (MASDA) for analysis of large numbers of samples (>500) simultaneously for a large number of known mutations (>100) in a single assay. MASDA utilizes oligonucleotide hybridization to interrogate DNA sequences. Multiplex DNA samples are immobilized on a solid support and a single hybridization is performed with a pool of allele-specific oligonucleotide (ASO) probes. Any probes complementary to specific mutations present in a given sample are in effect affinity purified from the pool by the target DNA. Sequence-specific band patterns (fingerprints), generated by chemical or enzymatic sequencing of the bound ASO(s), easily identify the specific mutation(s). Using this design, in a single diagnostic assay, we tested samples for 66 cystic fibrosis (CF) mutations, 14  $\beta$ -thalassemia mutations, two sickle cell anemia (SCA) mutations, three Tay–Sachs mutations, eight Gaucher mutations, four mutations in Canavan disease, four mutations in Fanconi anemia, and five mutations in BRCA1. Each mutation was correctly identified. Finally, in a blinded study of 106 of these mutations in >500 patients, all mutations were properly identified. There were no false positives or false negatives. The MASDA assay is capable of detecting point mutations as well as small insertion or deletion mutations. This technology is amenable to automation and is suitable for immediate utilization for high-throughput genetic diagnostics in clinical and research laboratories.

## INTRODUCTION

Over the last several years, there has been a significant increase in the number of identified, cloned, and characterized genes responsible for inherited diseases in humans. As the number of disease-associated sequences has increased, the number of mutations identified within the genes has likewise increased. In some genes, only one or a few mutations are specifically responsible for the disease phenotype (e.g. sickle cell anemia) (1). However, in most disease genes, many different causative mutations exist with no single mutation present at a significant frequency within an affected patient population. Therefore, for both research and clinical diagnostic applications, improved methods of mutation analysis are required not only to confirm that a candidate gene truly represents the disease gene, but also to build mutation databases and provide clinical diagnostic assays. In addition, with the increasing number of cancer genes being

discovered (2–4), efficient, cost-effective, and highly informative mutation analysis procedures are necessary in order to better understand predisposition and polygenic diseases.

This has led to the development of two broad categories of mutation detection technologies (5,6). The first group, designed to scan for mutations within a gene, includes single-strand conformational polymorphism (SSCP) (7), denaturing gradient gel electrophoresis (DGGE) (8), heteroduplex analysis (HET) (9), chemical cleavage analysis (CCM) (10), ribonuclease cleavage (RNase) (11) and direct sequencing of the target (12). Although these procedures are highly informative, they can be tedious and are incompatible with high throughput and low cost. Given the need in the clinical diagnostic laboratory to be able to analyze large numbers of samples (>500 samples/analysis) cost-effectively, these scanning procedures are not used currently as routine methods of mutation detection (5). In the second group, more direct methods of mutation analysis have been developed

<sup>\*</sup>To whom correspondence should be addressed

<sup>†</sup>Present address: Exact Laboratories, 63 Great Road, Maynard, MA 01754, USA

such as allele-specific amplification (ASA) (13), oligonucleotide ligation assay (OLA) (14), primer extension (15), artificial introduction of restriction sites (AIRS) (16), allele-specific oligonucleotide (ASO) hybridization (17) and variations of these procedures. Together with robotics, these methods for direct mutation analysis have helped in reducing cost and increasing throughput when only a limited number of mutations need to be analyzed. However, given that many of the mutations identified in disease genes are rare, for most populations undergoing testing, large numbers of mutations must be analyzed in order to achieve significant detection frequencies.

Unfortunately, such comprehensive multiplex mutation analysis (>100 mutations) cannot be performed by any currently available diagnostic method while retaining the sample throughput and cost-effectiveness needed in a clinical diagnostic laboratory. A potential solution to this problem involves a miniaturized version of the reverse dot-blot procedure (18). The 'chip' technology, or sequencing by hybridization (SBH), involves the hybridization of a single labeled target to a dense panel of oligonucleotides arrayed on a solid support (19–23). Although this approach is an extremely appealing solution for large multiplex mutation analysis, a number of technical issues must be addressed before it can be applied routinely to clinical diagnostics (24). In addition, as with the reverse dot-blot procedure, each sample analysis is performed independently. Therefore, a significant effort is needed in order to develop a procedure that will allow large numbers of samples (100–500) to be analyzed in a single assay.

While our previous modified ASO approach allowed the simultaneous analysis of large numbers of patient samples for multiple cystic fibrosis (CF) mutations (25), relatively inefficient independent hybridizations were required to identify specific mutations. Herein, we present a methodology which combines high sample throughput and direct detection of a large number of sequence variants. The multiplex allele-specific diagnostic assay (MASDA) has the capacity to analyze large numbers of samples (>500) for a large number of mutations (>100) cost-effectively in a single assay. Like the more familiar 'chip' technologies (19–22), MASDA uses oligonucleotide hybridization to interrogate DNA sequences. However, in contrast to many oligonucleotide array approaches, in the MASDA technology, the target DNA is immobilized to the solid support, and interrogated in a combinatorial fashion with a pool of ASOs (i.e. a single solution of hundreds of different oligonucleotides with each oligonucleotide sequence specific for one mutation). By retaining the forward dot-blot format, we can analyze large numbers of samples (>500) for a large number of mutations (>100) simultaneously. A schematic diagram of the MASDA technology is shown in Figure 1. During the hybridization, the ASO(s) corresponding to a specific mutation(s) present in a given sample is hybrid-selected from the pool of probes by the target DNA. Following removal of unhybridized ASOs, sequence-specific band patterns associated with the bound ASOs are generated by chemical or enzymatic sequencing, and the mutation or mutations present in the sample are easily identified. Using the gene targets *CFTR* (26),  $\beta$ -globin (1), *HEXA* (27), *GCR* (28), *ASPA* (29), *BRCA1* (3) and *FACC* (30) as a model system, we demonstrate that MASDA not only allows different patient samples with different disease indications to be analyzed in a single assay, but allows the identification of multiple mutations in a single gene or multiple genes in a single patient's DNA sample.

## RESULTS

### Model system

Seven different gene targets, representing eight different diseases, were chosen as a model system for complex mutation detection (Table 1). A total of 106 different mutations were analyzed in a single hybridization and detection procedure, referred to as MASDA 106. Although large numbers of mutations have been identified within the majority of disease genes listed, for the purposes of this study a selected number of these mutations were used (Table 1 columns 3 and 4). The specific mutations chosen within each disease gene represented the most clinically relevant for diagnostic applications (except for CF, since the most common mutation  $\Delta 508$  was not included in this assay, and except for *BRCA1* where the development of mutation databases with genotype-phenotype correlations are ongoing). The largest number of mutations analyzed resided within the *CFTR* gene. In addition to the most frequently detected mutations within a CF patient population (Cystic Fibrosis Genetic Analysis Consortium, 1994; unpublished data), additional point mutations were included that lead to premature translation termination, and subsequently a truncated protein product. A total of 33 different amplification products were needed in order to interrogate for the presence or absence of the 106 different mutations (Table 1 column 5). It is important to note, however, that amplifications were performed in a disease-specific manner only, for example, if a patient was suspected to be a CF carrier, the DNA sample was amplified for the CF gene only.

The specific mutations examined within each disease gene are shown in Table 2. These tables also include the size (bp) of regions amplified, and the primers used for each amplification.

### Disease-specific target amplifications

Where applicable, multiplex PCR was performed to reduce the number of PCR reactions needed (Table 1, column 5). A total of nine reactions facilitated amplification of 33 different loci. All single or multiplex PCR reactions were performed in a disease-specific manner. In other words, individual DNA samples were amplified for the relevant disease gene only, and not for all loci examined in the assay. Examples of the various disease-specific amplification products are shown in Figure 2. In order to include 66 CF mutations within our study, 17 different regions within the *CFTR* gene were amplified using two multiplex PCR reactions (Fig. 2, lanes 1 and 2). Amplicon sizes ranged from 130 to 510 bp. A single amplification product of 1600 bp was sufficient to include 14  $\beta$ -thalassemia- and two sickle cell anemia-associated mutations within the  $\beta$ -globin gene (Fig. 2, lane 3). For Tay-Sachs assays, a 2-plex amplification reaction was designed to examine three mutations (Fig. 2, lane 4). To examine Canavan-associated mutations 3-plex amplification reactions were performed (Fig. 2, lane 6). A separate 4-plex amplification was performed for five breast cancer susceptibility-related mutations (Fig. 2, lane 7) and a single 3-plex amplification for Fanconi anemia-associated mutations (Fig. 2, lane 8). For Gaucher disease, the *GCR* pseudogene necessitated a 3-plex amplification and an independent, single amplicon amplification. For convenience of analysis, aliquots from both reactions were pooled and electrophoresed in the same lane of the analytical gel (Fig. 2, lane 5).

**Table 1.** Model system detailing the diseases and mutations examined using the MASDA technology

Disease	Gene	No. of known mutations	No. of ASO probes	No. of PCR reactions
Cystic fibrosis (CF)	<i>CFTR</i>	>500	66 CF1–31 CF33–68	2 (8-plex and 9-plex)
$\beta$ -Thalassemia (BT)	$\beta$ -globin	>90	14	1
Sickle cell (SCA)	$\beta$ -globin	2	2	Same amplicon as BT
Tay–Sachs(TS)	hexosaminidase A ( <i>HEXA</i> )	>28	3	1 (2-plex)
Gaucher (GCR)	glucocerebrosidase ( <i>GCR</i> )	>35	8	2 (3-plex + 1)
Canavan disease (CD)	aspartoacylase ( <i>ASPA</i> )	>4	4	1 (3-plex)
Breast cancer susceptibility (BRC)	<i>BRCA1</i>	>250	5	1 (4-plex)
Fanconi anemia (FA)	Fanconi anemia complementation C ( <i>FACC</i> )	> 8	4	1 (3-plex)
Total			106	9

**Table 2.** Cystic fibrosis mutations examined in *CFTR* 8-plex amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
12	510	16UPCFX12F 16UPCFX12R	CF26	1898+1
			CF36	1812–1
			CF37	Y563D
			CF38	P574H
19	450	UP3CFEX19F UP3CFEX19R	CF23	3849 + 4
			CF27	R1162X
			CF31	3659dC
			CF42	R1158X
			CF43	S1196X
			CF44	I1203V
			CF45	Q1238X
			CF46	3662dA
			CF47	3750dAG
			CF48	3791dC
CF49	3821dT			
9	375	15UPCFX9F 15UPCFX9R	CF28	A455E
13	335	UP3CFEX13F UP3CFEX13R	CF29	2183AA→G
			CF39	K710X
			CF40	2043dG
3	270	UP3CFEX3F UP3CFEX3R	CF30	G85E
			CF33	E60X
			CF34	405 + 1
5	172	UP3CFEX5F 15UPCFX5R	CF25	711 + 1
			CF35	G178R
14b	150	15UCFX14bF 15UCFX14bR	CF24	2789 + 5
16	130	16UPCFX16F 16UPCFX16R	CF41	3120G→A

**Table 3.** Cystic fibrosis mutations examined in *CFTR* 9-plex amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
Int 19	480	15UCFIN19F 15UCFIN19R	CF8	3849 + 10
21	421	REUPCFX21F 15UPCFX21R	CF5 CF67 CF68	N1303X W1310X W1316X
15	361	UP3CFEX15F UP3CFEX15R	CF60 CF61 CF62	Q890X 2869 + G 2909dT
4	307	15UPCFX4F 15UPCFX4R	CF6 CF10 CF21 CF50 CF51 CF52	R117H 621 + 1 Y122X 444dA 556dA 574dA
17b	285	L15UCF17BF L15UCF17BR	CF16 CF64 CF65	Y1092X W1089X 3358dAC
7	260	REUPCFX7F REUPCFX7R	CF12 CF17 CF18 CF20 CF53 CF54 CF55	1078dT R347H R347P R334W G330X R352Q S364P
11	240	15UPCFX11F 15UPCFX11R	CF1 CF2 CF9 CF11 CF15 CF19 CF22 CF59	G542X G551D R553X 1717-1 S549R R560T S549N A559T
10	215	15UPCFX10F 15UPCFX10R	CF7 CF13 CF14 CF56 CF57 CF58	DI507 Q493X V520F 508C C524X 1677dTA
20	195	15UPCFX20F 15UPCFX20R	CF3 CF4 CF66	W1282X 3905 + T S1255X

**Table 4.**  $\beta$ -Thalassemia and sickle cell anemia mutations examined in  $\beta$ -globin gene amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
1-3	1600	GH260 GH283	BT1 BT2 BT3 BT4 BT5 BT6 BT7 BT8 BT9 BT10 BT11 BT12 BT13 BT14	IVS1-1 IVS1-6 IVS1-5 IVS-110 NONS-39 IVS2-1 IVS-745 COD8/9 IVS-654 41/42 -29 71/72 COD24 -88
1-3	1600	GH260 GH283	SCA1 SCA2	HbS HbC

**Table 5.** Tay-Sachs mutations examined in *HEXA* gene amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
11/12	530	TSEX11F TSEX12R	TS2 TS3	Ex11 4 bp Ins Ex12 splice
7	190	TSEX7F TSEX7R	TS1	G269S

**Table 6.** Gaucher mutations examined in *GCR* gene amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
10/11	871	GCRDF GCRDR	GCR5 GCR6 GCR8	1448 1604
2	358	84IVSF 84IVSR	GCR3 GCR4	84GG IVS2+1
9	319	1226F 1226R	GCR1 GCR2 GCR7	1297 1226 1342

**Table 7.** Canavan mutations examined in *ASPA* gene amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
6	274	CD6F CD6R	CD3 CD4	E285A A305E
5	151	CD5F CD5R	CD2	Y231X
Int2/Ex3	147	CDInt2F CDEx3R	CD1	433-2

**Table 8.** Breast cancer susceptibility mutations examined in *BRCA1* amplifications

Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
20	450	BRCA20F BRCA20R	BRC4	5382+C
21	315	BRCA21F BRCA21R	BRC5	M1775R
2	290	BRCA2F BRCA2R	BRC1	185dAG
5	270	BRCA5F BRCA5R	BRC2 BRC3	C61G C64G

**Table 9.** Fanconi anemia mutations examined in *FACC* amplifications

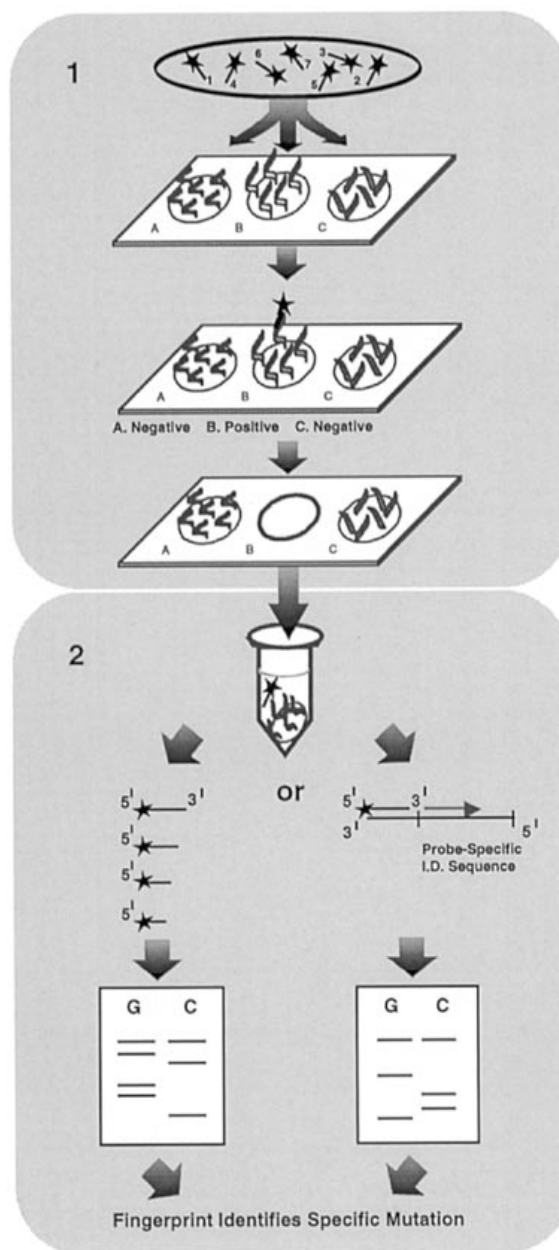
Exon	Amplicon size (bp)	Primers	Mutation number	Mutation name
1	366	FA1F FA1R	FA2	Q13X
6	329	FA6F FA6R	FA4 FA5	R185X D195V
4	274	FA4F FA4R	FA3	IVS4+4

### Mutation detection

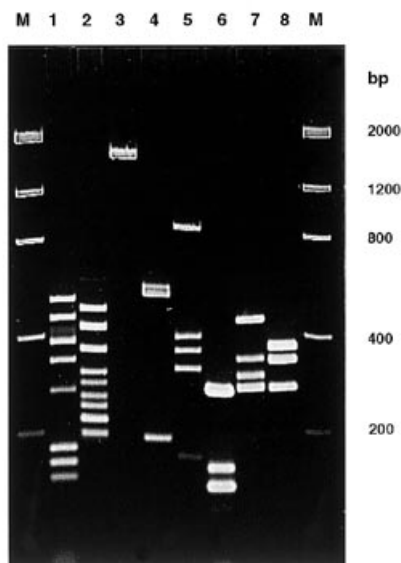
In order to analyze large numbers of samples simultaneously for the mutations listed in Table 1, we employed the standard forward dot-blot format and performed a single multiplex hybridization (Fig. 3). Although the percent G-C content of the 106 mutation-specific oligonucleotides ranged between 18 and 76%, the use of tetramethylammonium chloride (TMAC) (31) allowed all 106 mutation-specific oligonucleotides to be mixed together and hybridized in a single pool. Furthermore, the presence of TMAC in the hybridization and wash solutions allowed the hybridization and washes to be performed at the same temperature. As seen in Figure 3, only the 106 mutation-specific positive control samples generated signals upon autoradiography with no significant non-specific signal exhibited by the genotypically wild-type samples. Overall signal intensities and signal-to-noise ratios generated for the different mutation-specific positive control samples were optimized by adjusting the concentrations of each mutation-specific oligonucleotide in the hybridization.

### Mutation identification

The specific mutation present in a positive sample was identified by eluting the hybridized oligonucleotide from each individual dot and directly interrogating the oligonucleotide sequence. In one scheme, the eluted oligonucleotides were attached to a solid support, G and C base-specific chemical modification reactions were performed and the reaction products separated by polyacrylamide gel electrophoresis (32). Figure 4 represents an example of the CG-limited sequencing fingerprints produced from some of the oligonucleotides eluted from the mutation-specific positive control samples in Figure 3 (CF30, CF31, TS1, TS2, TS3, BT2,



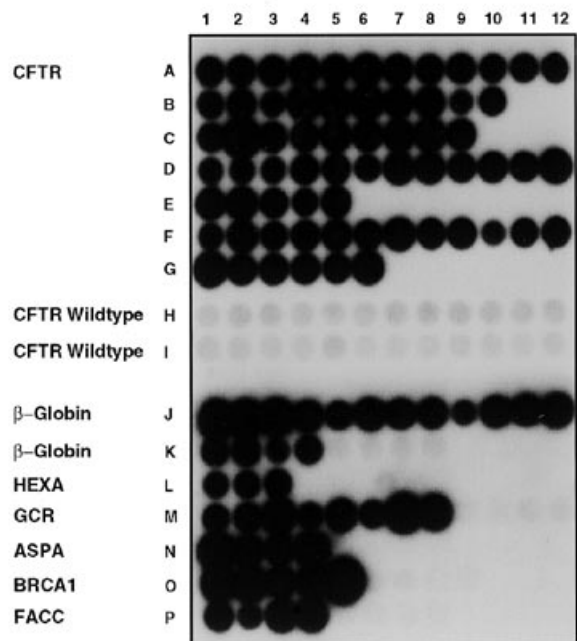
**Figure 1.** Schematic representation of MASDA procedure. In Phase 1 (panel 1), the probe corresponding to a specific mutation (one probe sequence per mutation interrogated) is hybrid-selected from the pool of probes by the target DNA on a dot-blot. In this diagram, seven different probes and three different patient samples (multiplexes A, B and C) are depicted whereas in the MASDA assay >100 probes are used routinely to interrogate >500 patient samples in each assay. Each dot represents the multiplex amplification performed on one patient DNA for one disease gene only. Unhybridized probes are washed away, and mutation-positive patient samples are located by the presence of the labeled probe. The second phase of the assay (panel 2) reveals the identity of the probe, and therefore the specific mutation present in the patient DNA. The dot from each positive sample is excised, and the probe eluted off the membrane disc. The identity of the probe is revealed by one of two methods: directly by chemical cleavage sequencing, or indirectly by using the probe as a primer in a cycle sequencing reaction. In the latter case, a pool of templates is used where only one template has a region complementary to the eluted probe and, therefore, downstream sequencing of the probe-specific identifier (I.D.) sequence reveals the identity of the probe. With limited sequencing of 'C' and 'G' residues only, the fingerprints obtained on sequencing gels are compared with a known database of sequences, the probe is identified and consequently unequivocal mutation identity is assigned.



**Figure 2.** Disease gene loci amplified by multiplex PCR for MASDA assay. DNA samples (2  $\mu$ g) were amplified in multiplexes specific for one of seven different genes (Table 1), and analyzed by gel electrophoresis. M =  $\phi$ X174/*Hae*III molecular weight marker; 1 = eight amplicon multiplex within the *CFTR* gene; 2 = nine amplicon multiplex within the *CFTR* gene; 3 = single amplicon within the  $\beta$ -globin gene; 4 = two amplicon multiplex within the *HEXA* gene; 5 = (3 + 1) multiplex (three amplicon multiplex + one independent amplicon) for the *GCR* gene; 6 = three amplicon multiplex within the *ASPA* gene; 7 = four amplicon multiplex within the *BRCA1* gene; 8 = three amplicon multiplex within the *FACC* gene.

BT3\*, BT6 and BT7). As shown in Figure 4, the oligonucleotide eluted from each dot generated a characteristic fingerprint which unambiguously identified the specific mutation present in the positive sample DNA. Unique band patterns were generated for each of 106 ASOs (data not shown). Sequence analysis of all 106 eluted oligonucleotides verified that a positive result from the single pooled hybridization represented specific oligonucleotide hybridization with no significant cross-hybridization between different probes. In addition, one sample containing two  $\beta$ -globin mutations (Fig. 4, lane BT3\*) generated a unique fingerprint made up of two superimposed oligonucleotide-specific band patterns. This demonstrated that a compound heterozygote genotype was readily identified using this technique.

In addition to the chemical modification and cleavage procedure, we devised an enzymatic protocol for eluted oligonucleotide identification. This procedure involved using the eluted mutation-specific oligonucleotide as a primer in a cycle sequencing reaction. The eluted oligonucleotide was added to a cycle sequencing reaction containing a mixture of synthetic (77mer) templates. Each synthetic template contained a different priming sequence complementary to only one of the mutation-specific oligonucleotides in the pooled hybridization, and a downstream specific identifier sequence to generate unique, mutation-specific fingerprints identifying the eluted ASO. Figure 5 is an example of the C and G band patterns generated from cycle sequencing reactions utilizing oligonucleotides eluted from positive (mutant genotype CF17, CF20, BT5 and BRC5) and negative (wild-type genotype CF, BT and BRC) samples. Each reaction performed with oligonucleotides eluted from positive control samples (Fig. 5, lanes 1–4) generated a common band pattern contained within

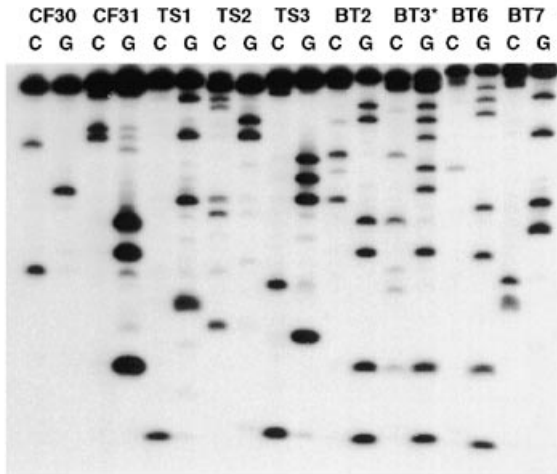


**Figure 3.** Simultaneous detection of 106 different mutations from seven different genes in a single hybridization assay. DNA samples, each positive for one of 106 different mutations, were amplified using disease-specific multiplexes. Amplified samples were denatured, spotted on membranes and hybridized under TMAC conditions to a mixture of 106 mutation-specific, labeled oligonucleotide probes. The blots were washed, dried and exposed to Kodak X-Omat X-ray film for 15 min at  $-80^{\circ}\text{C}$ . Rows A–G = detection of 66 *CFTR* mutations present in cystic fibrosis. Rows H–I = *CFTR* wild-type negative control samples for cystic fibrosis mutations. Rows J–K = detection of 14  $\beta$ -globin mutations in  $\beta$ -thalassemia and two  $\beta$ -globin mutations in sickle cell anemia. Row L = detection of three *HEXA* mutations present in Tay–Sachs. Row M = detection of eight *GCR* mutations present in Gaucher disease. Row N = detection of four *ASPA* mutations present in Canavan disease. Row O = detection of five *BRCA1* mutations present in breast cancer. Row P = detection of four *FACC* mutations present in Fanconi anemia. Disease-specific negative control (wild-type) samples follow the positive samples in rows K–P.

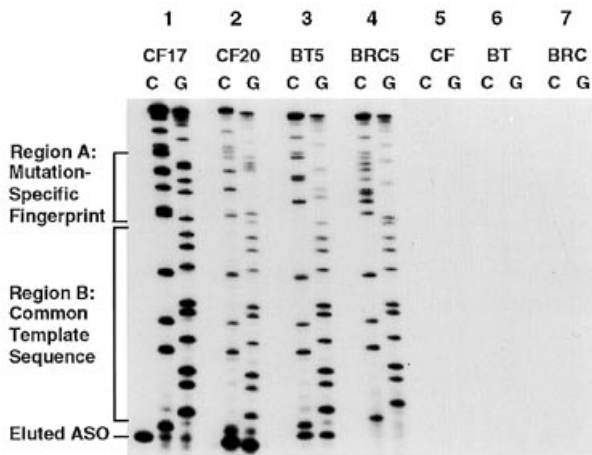
all synthetic templates (Region B) followed by a mutation-specific fingerprint (Region A). The pattern observed in the mutation-specific fingerprints allowed unequivocal identification of the corresponding ASO primer, and consequently the specific mutation present in the patient sample. No band patterns were observed when cycle sequencing reactions were performed with eluates from samples wild-type for the diseases interrogated (Fig. 5, lanes 5–7).

### Large-scale sample analysis

To validate the procedure, we performed a blinded analysis to assess the ability of the MASDA technique to identify mutations as envisaged in the diagnostic setting. More than 500 samples whose genotype had been established previously using other techniques were obtained from clinical collaborators. The genotypes remained unknown to the laboratory personnel. After the analysis, the MASDA results were compared with the known genotype. Figure 6 represents the hybridization results generated from analyzing >500 different DNA samples for the presence of 106 different mutations, in a single hybridization assay. All samples known to carry one of the 106 different mutations were identified correctly as positive in the hybridization (Fig. 6). The

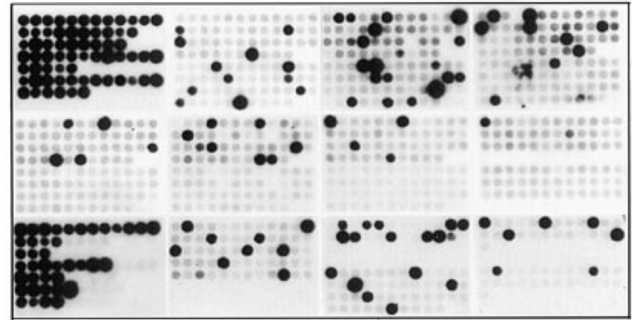


**Figure 4.** Band patterns generated by chemical cleavage of eluted ASOs reveal the identity of the mutation. Following hybridization, ASOs eluted from positive samples were subjected to chemical cleavage at C and G residues using a solid phase sequencing protocol (32), followed by gel electrophoresis. Unique band patterns were generated for all 106 mutations examined (data not shown) with representative data from 12 mutations shown in Figure 4. (See Tables 1–9 for definition of mutation abbreviations.) C = C cleavage reaction; G = G cleavage reaction.



**Figure 5.** Band patterns generated by an enzymatic sequencing procedure reveal the identity of the mutation. The ASO(s) eluted from mutation-positive samples were used to prime cycle sequencing reactions in a complex mixture of templates. Each template contained a region complementary to a specific ASO (the priming site), a common 'stuffer region' (Region B) and a downstream mutation-specific identifier sequence (Region A). With limited C and G sequencing (lanes C and G for each sample), the fingerprint generated from the mutation-specific identifier sequence unequivocally identified the specific ASO, and therefore the mutation present in the target DNA. CF = cystic fibrosis; BT =  $\beta$ -thalassaemia; BRC = breast cancer susceptibility *BRCA1* gene. Lane 1 = CF17 mutation; lane 2 = CF20 mutation; lane 3 = BT5 mutation; lane 4 = BRC5 mutation; lane 5 = CF wild-type control; lane 6 = BT wild-type control; lane 7 = BRC wild type control.

specific mutations were identified correctly in each positive sample by performing the chemical modification and cleavage procedure (data not shown). It is significant to note that no increase in non-specific background was observed when sample



**Figure 6.** Simultaneous detection of 106 different mutations in >500 different patient samples in a single ASO hybridization assay. DNA samples were amplified using disease-specific multiplexes, denatured, spotted on dot-blot and hybridized under TMAC conditions to a mixture of 106 different, labeled oligonucleotide probes. The blots were washed, dried and the autoradiographs were prepared by exposure to Kodak X-Omat X-ray film for 30 min at  $-80^{\circ}\text{C}$ .

throughput was increased to >500 samples in a single hybridization assay.

MASDA results are highly reproducible. These 500 samples subsequently were analyzed in a blinded manner in five separate assays. We did not observe any false positives from mutation-specific hybridization to a wild-type genotype, or false negative results from samples of known mutant genotype.

## DISCUSSION

We believe that the MASDA approach provides a major breakthrough in mutation detection. The combination of the forward dot-blot, complex simultaneous probe hybridization and direct mutation detection for the first time solves the dual problems of parallel, multiple sample analysis and parallel multiplex mutation detection. In contrast, a significant limitation in almost all of the currently available techniques is the inability to analyze a large number of samples simultaneously for a large number of mutations. The traditional dot-blot, wherein the PCR products are bound to a filter membrane and hybridized with allele-specific probes, is a very effective format for the analysis of large numbers of samples. However, in a standard forward dot-blot procedure, a separate hybridization is performed for each allele or mutation of interest. Because the number of probes required for genetic diagnosis is large, this procedure is very cumbersome. Previously we described a methodology for CF testing which pooled together multiple mutation-specific oligonucleotides into a single hybridization (25). This format retains the capacity for large sample throughput while reducing the number of hybridizations involved in performing multiple mutation analysis. Even with this approach, the limited number of reporters available to discriminate between probes necessitated secondary independent hybridizations on all positive samples to identify the specific mutation present, reducing the cost-effectiveness of the methodology. The ability of the reverse dot-blot and the miniaturization of reverse dot-blot (oligonucleotide arrays) to analyze multiple probes is an appealing solution to this problem. However, in this approach, a separate hybridization is performed for each sample, thus eliminating the capacity for large sample throughput.

In the MASDA approach, the disadvantages of both individual sample hybridizations and independent probe hybridizations are

avoided. By eluting and interrogating the sequence of the mutation-specific oligonucleotides hybridized to a patient's DNA sample, MASDA eliminates the need for secondary independent hybridizations. Therefore, in a single day, hundreds of different samples can be analyzed simultaneously in a single hybridization containing a complex mixture of hundreds of mutation-specific oligonucleotides. Also, by generating short and unique band patterns for the hybridized and eluted oligonucleotides, multiple samples can be analyzed by stagger loading samples across multiple lanes of a gel. Therefore, using currently available automated sequencers, specific mutation identification can be performed easily on hundreds of positive samples at a rate in the range of 150–900 samples/h.

When designing DNA diagnostic assays for widespread use in clinical laboratories, several parameters need to be considered in addition to high throughput sample capacity and multiplex mutation detection. These include ease of performance, overall assay economies of scale and ease and cost of modification [both in the assay design process and in implementation of the modification(s) in the clinical lab]. This is very important in a field such as genetic diagnostics, where both the number of relevant genes and mutations identified in each gene change rapidly, as do available instrumentation and detection systems. With these considerations in mind, MASDA has been designed to be a flexible, modular platform, allowing all of these challenges to be addressed.

There is minimal development necessary for the addition of new allele-specific oligonucleotides to a probe pool, and no incremental labor cost to perform the improved/expanded test. For example, optimization of the oligonucleotides in MASDA 106 involved nothing more than an independent hybridization of each oligonucleotide to determine the concentration necessary to yield a comparable signal to other mutation-specific probes in the pool. This allows new mutations easily to be added to any diagnostic assay. Because the hybridizations are performed with the oligonucleotide probes in solution, it is possible to mix and match probes on demand, therefore allowing clinical laboratories to customize diagnostic assays cost-effectively. Thus the assay can be readily modified without incurring the significant costs associated with the design and manufacture of new hardware or disposables. There is also flexibility in sample preparation and target detection. The sample nucleic acid can be either DNA or RNA. For the purpose of performing multiplex target amplifications, PCR was utilized as the amplification procedure for this work. However, MASDA is compatible with any target amplification technology, and does not require any processing of amplification products prior to mutation detection and identification. This becomes a very important issue when large numbers of samples need to be analyzed in a single assay. Since the sample nucleic acid does not need to be fragmented, long PCR products can be analyzed, as well as the multiplex amplicons demonstrated in this study.

A variety of different reporter groups and detection methods are compatible with MASDA. For example, while the data presented herein used an isotopic format, MASDA is equally powerful with non-isotopic detection methods such as chemiluminescence and fluorescence, and appropriate instrumentation for sequence analysis (data not shown). Similarly, we have presented both a base-specific chemical cleavage method, and an enzymatic method to identify the specific sequences of eluted oligonucleotides. Each approach meets different needs. The chemical cleavage method is

more direct and cost-effective and is very appropriate for high complexity laboratories, or ultimately, closed systems. The enzymatic method described avoids the use of organic sequencing reagents, but has higher reagent costs. Because MASDA is modular, as instrumentation changes, the method used to identify the eluted oligonucleotide can change, without requiring changes in the remainder of the assay. Thus, each of the individual components of MASDA can be modified individually, without altering the basic protocol. These features are important in the clinical laboratory, where continual protocol changes are disruptive, and frequent purchase of capital equipment may not be possible.

Finally, because all sequence variants are analyzed using the same reaction conditions, assays for rare disorders can be analyzed in the same 'run' as tests for common disorders. Thus, high volume economies of scale can be achieved for otherwise low volume tests. A more subtle benefit is the elimination of the internal support required to maintain multiple disease-specific protocols.

Currently, significant efforts are being made to develop informative databases on genotype-phenotype associations of existing and new mutations within known disease genes. In addition, there is an ever increasing interest in establishing the relationships between genotypes of patients involved in clinical trials and their response to various therapies. In addition to the diagnostic applications, MASDA will allow research laboratories to develop oligonucleotide libraries representative of previously identified expressed sequence tags or bi-allelic markers (polymorphisms) identified within the human genome. For this purpose, we currently are investigating the number of hybridized and eluted oligonucleotides that can be uniquely identified from any given positive sample. Given that the MASDA technology has the capacity to perform complex known mutation analysis on hundreds of patient samples with different disease indications in a single assay, we believe that it is suitable for immediate and future applications in both clinical and research laboratories.

## MATERIALS AND METHODS

### Genomic DNA samples positive (mutant) or negative (wild-type) for known mutations

Genomic DNA was extracted from whole blood as previously described (33).

### Cloned positive control DNA samples

When mutation-positive genomic DNA was not available, oligonucleotides representing 40 bp of endogenous gene sequence including the mutation were synthesized, cloned into pGEM®-3Zf(+) vectors (Promega Corporation, Madison, WI), and the presence of the mutation in each clone verified by sequencing (data not shown).

### DNA amplifications

As a model system for complex mutation detection, mutations were selected from 33 regions in seven different genes. The genes included the CF transmembrane conductance regulator gene (*CFTR*), the  $\beta$ -globin gene, the Tay-Sachs hexosaminidase gene (*HEXA*), the Gaucher gene (*GCR*), the Canavan aspartoacylase gene (*ASPA*), the breast cancer susceptibility gene (*BRCA1*) and the Fanconi anemia complementation group C gene (*FACC*).

PCR amplifications were performed using 1–2 µg of genomic DNA or 10 ng of plasmid DNA in 100 µl of reaction buffer containing 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 200 mM dNTPs and 0.05–0.1 U/ml of *Taq* polymerase (Perkin-Elmer, Norwalk, CT). For the different disease gene amplifications, the concentration of primers ranged from 0.2 to 1.6 µM.

For DNA amplifications involving simultaneous multiplexes of three or more amplicons (*CFTR* 8-plex and 9-plex, *ASPA* 3-plex, *BRCA1* 4-plex and *FACC* 3-plex), the primers were chimeras of a sequence-specific region with a common 'universal primer sequence' (UPS) as described by Shuber *et al.* (33). These primers facilitated rapid multiplex development and consistently robust amplifications. Primer sequences are not listed but may be furnished upon request.

DNA amplifications were performed using a Perkin-Elmer 9600 Thermal Cycler (Perkin-Elmer, Norwalk, CT). For *CFTR*, *HEXA*, *ASPA*, *BRCA1* and *FACC*, the amplifications were carried out for 28 cycles with ramping (94°C/10 s hold with 48 s ramp, 60°C/10 s hold with 36 s ramp, 72°C/10 s hold with 38 s ramp) and a final 74°C hold for 5 min before cooling. For β-globin and *GCR*, the amplification program consisted of 28 cycles with a 55°C anneal (94°C/10 s hold, 55°C/10 s hold, 74°C/10 s hold) and a final 74°C hold for 5 min before cooling.

Amplification products were analyzed by 2% agarose gel electrophoresis followed by ethidium bromide staining and visualization on a UV transilluminator (Fotodyne, New Berlin, WI).

### Specific mutations examined in the MASDA 106 hybridization assay

Mutations from seven different genes were selected as candidates for a complex mutation detection assay. The 106 mutations examined included point mutations, deletions and insertions. Details of the selected mutations and gene amplifications are listed in Table 1. References for any of the mutations will be provided upon request.

### Oligonucleotide pools

ASOs were 17mers synthesized and HPLC-purified by Operon Technologies (Alameda, CA). All oligonucleotides were quantitated by spectrophotometry and tested in independent hybridizations before being pooled. The sequences of the 106 different ASOs are not shown but may be provided upon request. Specified amounts of individual ASOs were combined into a pool of 106 ASOs so that the pool would contain the required amount of each specific ASO determined to be optimal for the pool hybridization. Aliquots of pooled ASOs were lyophilized and stored at –20°C.

### Dot-blots

Amplified products were denatured using 1.0 M NaOH, 2.0 M NaCl, 25 mM EDTA pH 8.0, containing bromophenol blue (30 µl of 0.1% bromophenol blue/10 ml denaturant) for 5 min at room temperature. Denatured products were blotted onto Biotrans membrane (ICN Biomedicals Inc., Aurora, OH) using a 96-well format dot-blot apparatus (Life Technologies, Gaithersburg, MD). Membranes were neutralized in 2× SSC (0.15 M NaCl, 0.015 M trisodium citrate) for 5 min at room temperature and baked in a vacuum oven at 80°C for 15 min. Immediately before

use, the membranes were rinsed in distilled water and placed in hybridization solution.

### Probe labeling

Aliquots of 106 pooled ASOs (one ASO sequence representing each mutation) were thawed, resuspended in distilled water and end-labeled in a single reaction containing 1× kinase buffer (New England Biolabs, Beverly, MA), 0.135 nmol of [ $\gamma$ -<sup>32</sup>P]ATP (DuPont, Boston, MA) and 35 U of T4 polynucleotide kinase (New England Biolabs, Beverly, MA). The labeling reactions were incubated at 37°C for 1 h. The efficiency of the kinase reaction was monitored by chromatography on cellulose polyethyleneimine (PEI) plates (J.T. Baker Inc., Phillipsburg, NJ) using 0.75 M NaH<sub>2</sub>PO<sub>4</sub> pH 3.5 buffer, followed by exposure of the plates to Kodak X-Omat X-Ray film (Eastman Kodak Company, Rochester, NY) at room temperature for 5 min.

### Hybridizations/ASO pooling

Hybridization and wash conditions, using TMAC, were the same as previously described (25). For this protocol, the 96-well array of spotted genomic samples was marked with a grid so that positives identified in the hybridization could be located easily for the subsequent elution and ASO sequencing. Signal intensities generated from the different mutation-positive samples were optimized by adjusting the concentrations of each mutation-specific oligonucleotide within the hybridization. In order to achieve uniform hybridization signals, the final concentration of each labeled mutant ASO in the pool hybridization ranged from 0.008 to 1.8 pmol/ml, with the concentration of cold normal ASOs ranging from 0- to 200-fold excess of the corresponding mutant ASO.

To ensure that all 106 ASOs within the pool were labeled, membranes containing a positive control sample for each mutation were included in each hybridization.

Once washed, the blots were wrapped in plastic wrap and exposed to Kodak X-Omat X-Ray film (Eastman Kodak Company, Rochester, NY) at –80°C for 15 min to 1 h.

### Specific mutation identification

*By chemical cleavage.* The ASO hybridized to each mutation-positive sample was identified by eluting and sequencing the ASO. For a pool of 106 ASOs, sequencing 'C' and 'G' bases only was sufficient to identify the ASO sequence unambiguously and therefore allowed unequivocal identification of the corresponding mutation in the DNA sample.

The region of membrane containing each mutation-positive sample identified in the pool hybridization was excised, and the disc of Biotrans membrane placed in 100 µl of distilled water and heated at 95°C for 10 min to elute the bound ASO. After cooling to room temperature, the membrane disc was discarded, and the eluted ASO was subjected to chemical sequencing.

Solid-phase chemical cleavage of the ASOs attached to a solid support was performed according to Rosenthal *et al.* (32) with minor changes. This method permitted simultaneous sequencing of all bound ASOs in a single reaction vessel. To attach the ASOs to a solid support prior to chemical cleavage, a small, labeled piece (6 mm × 3 mm) of CCS paper (32) was immersed in each tube containing eluted ASO, and incubated at 65°C for 1 h. All pieces of paper were then combined into a single 50 ml tube

containing 25 ml of distilled water. The papers were then washed at room temperature three times (30 s/wash) with distilled water (25 ml/wash) followed by three washes (30 s/wash) with 96% ethanol (25 ml/wash). Papers with attached ASOs could be batch washed without cross-contamination (data not shown).

Once washed, the papers were air-dried and each piece cut into two, with one-third assigned for the 'G' chemical cleavage reaction and two-thirds designated for the 'C' cleavage reaction. All 'C' reaction ASO solid supports were combined into one tube containing 1 ml of 4.0 M hydroxylamine HCl pH 6. For 'G' reaction modifications, the combined pieces of paper were placed in 1 ml of 50 mM ammonium formate pH 3.5 and 7 µl of dimethylsulfate added. Reactions were incubated at room temperature for 10 or 20 minutes for the 'G' and 'C' reaction respectively. Batch processing of >100 sequencing reactions was performed without cross-contamination of cleavage products (data not shown).

Washes were performed on the batch of 'C' reaction papers and the batch of 'G' reaction papers as described above for washes after attachment of the ASOs to the solid support. The papers were then air dried and each piece of paper placed in their designated location in a 96-well amplification tray (Perkin-Elmer, Norwalk, CT).

To cleave and elute the sequencing products off the solid support membrane, freshly prepared piperidine [50 µl of 10% (v/v) piperidine] was added to each well, the tray was covered with a rubber gasket, and incubated at 90°C for 30 min in a thermal cycler.

For each cleavage reaction, the piperidine solution containing the eluted cleavage products was transferred to a fresh 96-well amplification tray and the piperidine evaporated for 2–3 h at room temperature. The evaporation step was repeated twice with 50% ethanol (35 µl/well). ASO cleavage products were dissolved in 4 µl of loading dye (90% formamide, 1× TBE, 0.1% bromophenol blue, 0.1% xylene cyanol) before gel electrophoretic resolution. Sequencing gels (20% polyacrylamide/8 M urea/TBE) were pre-run for 1 h at 2000 V, the samples were loaded, and electrophoresis continued until the bromophenol blue dye had migrated 14 cm from the origin. Sequencing gels were exposed to Kodak X-Omat X-ray film for 24–48 h.

*By enzymatic sequencing.* Mutation-positive samples were identified, spots excised and ASOs eluted as described for the chemical cleavage method. The eluted samples were each placed in Microcon-10™ concentrators (Amicon Inc., Beverly, MA) and centrifuged at 14 000 r.p.m. for 15 min in a bench top microfuge. The eluates were then transferred to Microcon-3™ concentrators (Amicon Inc., Beverly, MA) and centrifuged at 14 000 r.p.m. for 30 min in a bench top microfuge. Both concentrators were washed twice with 100 µl of distilled water per wash. (The Microcon-3™ concentrators were washed with the eluate from the corresponding Microcon-10™ concentrators.) The eluted ASOs were recovered from the top portion of the Microcon-3™ concentrator by three serial rinses with 20 µl of distilled water/rinse, and the fractions pooled. The samples were lyophilized in a UniVapo™ concentrator (Integrated Separation Systems, Natick, MA), re-dissolved in 5 µl of distilled water, and used as sequencing primers in an enzymatic sequencing protocol designed to identify the eluted ASO. The sequencing reactions included a pool of oligonucleotide templates with each template (77mer) consisting of a 3' region (17 bp) as the primer-binding

site uniquely complementary to a specific ASO, and a second unique region (17 bp) consisting of an 'ASO-specific identifier sequence'. Sequencing products were only observed when an eluted ASO was bound to the complementary region of a unique template, acted as a primer and permitted cycle sequencing to reveal the identity of the downstream 'ASO-specific identifier sequence'.

The cycle sequencing reactions contained a pool of ASO-specific templates (5 fmol/template), 0.5 µl of Thermosequenase buffer concentrate (Amersham Life Science, Cleveland, OH), 0.125 µl of Thermosequenase (32 U/µl) and either 'G' termination mix (15 µM dATP, 15 µM dCTP, 15 µM dTTP, 15 µM 7-deaza-dGTP and 4 µM ddGTP) or 'C' termination mix (15 µM dATP, 15 µM dGTP, 15 µM dTTP, 15 µM 7-deaza-dGTP and 4 µM ddCTP) in a reaction volume of 8 µl. Cycle sequencing was performed between 95°C for 30 s and 70°C for 1 min for 30 cycles, followed by a 2 min incubation at 70°C. Sequencing products were resolved on a 15% acrylamide/7 M urea gel before being exposed to Kodak X-Omat X-ray film at –70°C for ~16 h.

## REFERENCES

- Wilson, J.T., Marotta, C.A., Forget, B.G. and Weissman, S.M. (1977) Structure of hemoglobin messenger RNA and its relation to hemoglobinopathies. *Trans. Assoc. Am. Physicians*, **90**, 117–126.
- Fearon, E.R. and Vogelstein, B. (1990) A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759–767.
- Miki, Y., Swenson, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P.K., Norris, F.H., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J.C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A. and Skolnick, M.H. (1994) Isolation of a strong candidate for the 17q-linked breast and ovarian cancer susceptibility gene, *BRCA1*. *Science*, **266**, 66–71.
- Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Patel, S., Rice, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., Gudmundsson, J., Ficenec, D., Kelsell, D., Ford, D., Tonin, P., Bishop, D.T., Spurr, N.K., Ponder, B.A.J., Eeles, R., Peto, J., Devilee, P., Cornelisse, C., Lynch, H., Narod, S., Lenoir, G., Egilsson, V., Barkadottir, R.B., Easton, D.F., Bentley, D.R., Futreal, P.A., Ashworth, A. and Stratton, M.R. (1995) Identification of the breast cancer susceptibility gene, *BRCA2*. *Nature*, **378**, 789–792.
- Cotton, R.G.H. (1993) Current methods of mutation detection. *Mutat. Res.*, **285**, 125–144.
- Forrest, S., Cotton, R., Landegren, U. and Southern, E. (1995) How to find all those mutations. *Nature Genet.*, **10**, 375–376.
- Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl Acad. Sci. USA*, **86**, 2766–2770.
- Myers, R.M., Lumelsky, N., Lerman, L.S. and Maniatis, T. (1985) Detection of single base substitutions in total genomic DNA. *Nature*, **313**, 495–498.
- Keen, J., Lester, D., Inglehearn, C., Curtis, A. and Bhattacharya, S. (1991) Rapid detection of single base mismatches as heteroduplexes on Hydrolink gels. *Trends Genet.*, **7**, 5.
- Cotton, R.G.H., Rodrigues, N.R. and Campbell, R.D. (1988) Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc. Natl Acad. Sci. USA*, **85**, 4397–4401.
- Myers, R.M., Larin, Z. and Maniatis, T. (1985) Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science*, **230**, 1242–1246.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

13. Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C. and Markham, A.F. (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.*, **17**, 2503–2516.
14. Landegren, U., Kaiser, R., Sanders, J. and Hood, L. (1988) A ligase-mediated gene detection technique. *Science*, **241**, 1077–1080.
15. Sokolov, B.P. (1990) Primer extension technique for the detection of single nucleotide in genomic DNA. *Nucleic Acids Res.*, **18**, 3671.
16. Cohen, J.B. and Levinson, A.D. (1988) A point mutation in the last intron responsible for increased expression and transforming activity of the c-Ha-ras oncogene. *Nature*, **334**, 119–124.
17. Wallace, B.R., Johnson, M.J., Hirose, T., Miyake, T., Kawashima, E.H. and Itakura, K. (1981) The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit  $\beta$ -globin DNA. *Nucleic Acids Res.*, **9**, 879–894.
18. Saiki, R.K., Walsh, P.S., Levenson, C.H. and Erlich, H.A. (1989) Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc. Natl Acad. Sci. USA*, **86**, 6230–6234.
19. Drmanac, R., Labat, I., Brukner, I. and Crkvenjakov, R. (1989) Sequencing of megabase plus DNA by hybridization: theory of the method. *Genomics*, **4**, 114–128.
20. Bains, W. and Smith, G.C. (1988) A novel method for nucleic acid sequence determination. *J. Theor. Biol.*, **135**, 303–307.
21. Khrapko, K.R., Lysov, Y., Khorlyn, A.A., Shick, V.V., Florentiev, V.L. and Mirzabekov, A.D. (1989) An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett.*, **256**, 118–122.
22. 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.
23. Pease, A.C., Solas, D., Sullivan, E.J., Cronin, M.T., Holmes, C.P. and Fodor, S.P.A. (1994) Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl Acad. Sci. USA*, **91**, 5022–5026.
24. Guo, Z., Guilfoyle, R.A., Theil, A.J., Wang, R. and Smith, L.M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.*, **22**, 5456–5465.
25. Shuber, A.P., Skoletsky, J., Stern, R. and Handelin, B.L. (1993) Efficient 12-mutation testing in the CFTR gene: a general model for complex mutation analysis. *Hum. Mol. Genet.*, **2**, 153–158.
26. Zielenski, J., Rozmahel, R., Bozon, D., Kerem, B.S., Grzelczak, Z., Riordan, J.R., Rommens, J. and Tsui, L.C. (1991) Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics*, **10**, 214–228.
27. Myerowitz, R. and Costigan, F.C. (1988) The major defect in Askenazi Jews with Tay–Sachs disease is an insertion in the gene for the alpha-chain of beta-hexosaminidase. *J. Biol. Chem.*, **263**, 18587–18589.
28. Horowitz, M., Wilder, S., Herowitz, Z., Reiner, O., Gelbart, T. and Beutler, E. (1989) The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics*, **4**, 87–96.
29. Kaul, R., Gao, G.P., Balamurugan, K. and Matalon, R. (1993) Cloning of the human aspartoacylase cDNA and a common missense mutation in Canavan disease. *Nature Genet.*, **5**, 118–123.
30. Whitney, M.A., Saito, H., Jakobs, P.M., Gibson, R.A., Moses, R.E. and Grompe, M. (1993) A common mutation in the FACC gene causes Fanconi anemia in Askenazi Jews. *Nature Genet.*, **4**, 202–205.
31. Wood, W.I., Gitschier, J., Lasky, L.A. and Lawn, R.M. (1985) Base composition-independent hybridization in tetramethylammonium chloride: a method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl Acad. Sci. USA*, **82**, 1585–1588.
32. Rosenthal, A., Schwertner, S., Hahn, V. and Hunger, H.D. (1985) Solid-phase methods for sequencing of nucleic acids I. Simultaneous sequencing of different oligodeoxyribonucleotides using a new, mechanically stable anion-exchange paper. *Nucleic Acids Res.*, **13**, 1173–1184.
33. Shuber, A.P., Grondin, V.J. and Klinger, K.W. (1995) A simplified procedure for developing multiplex PCRs. *Genome Res.*, **5**, 488–493.