Indian J Med Res 141, February 2015, pp 175-186

Development of low-density oligonucleotide microarrays for detecting mutations causing Wilson's disease

Manjula Mathur, Ekta Singh[*,†], T.B. Poduval[*] & Akkipeddi V.S.S.N. Rao

Molecular Biology Division, [*]Radiation Biology & Heath Sciences Division, Bhabha Atomic Research Centre, Mumbai, India

Received June 7, 2013

***Background & objectives*: Wilson's disease (WD) is an autosomal recessive disorder caused by mutations in *ATP7B,* a copper transporter gene, leading to hepatic and neuropsychiatric manifestations due to copper accumulation. If diagnosed early, WD patients can be managed by medicines reducing morbidity and mortality. Diagnosis of this disease requires a combination of tests and at times is inconclusive due to overlap of the symptoms with other disorders. Genetic testing is the preferred alternative in such cases particularly for individuals with a family history. Use of DNA microarray for detecting mutations in *ATP7B* gene is gaining popularity because of the advantages it offers in terms of throughput and sensitivity. This study attempts to establish the quality analysis procedures for microarray based diagnosis of Wilson's disease.**

***Methods*: A home-made microarrayer was used to print oligonucleotide based low-density microarrays for addressing 62 mutations causing Wilson's disease reported from Indian population. Inter- and intra-array comparisons were used to study quality of the arrays. The arrays were validated by using mutant samples generated by site directed mutagenesis.**

***Results*: The hybridization reaction were found to be consistent across the surface of a given microarray. Our results have shown that 52 °C post-hybridization wash yields better reproducibility across experiments compared to 42 °C. Our arrays have shown > 80 per cent sensitivity in detecting these 62 mutations.**

***Interpretation & conclusions*: The present results demonstrate the design and evaluation of a low-density microarray for the detection of 62 mutations in *ATP7B* gene, and show that a microarray based approach can be cost-effective for detecting a large number of mutations simultaneously. This study also provides information on some of the important parameters required for microarray based diagnosis of genetic disorders.**

Key words Discrimination score - hybridization probes - microarrays - mutations - Wilson's disease

Wilson's disease (WD) is a monogenic autosomal recessive disorder that clinically manifests at the average age of 12 yr (range five to 23) after copper accumulates in liver and brain, gradually leading to cirrhosis, lack of coordination, personality changes and early fatality. The complexities associated with its diagnosis, comprising clinical and biochemical investigations, in particular, parameters of copper metabolism[1] and the overlap of its symptoms with other disorders often lead to incorrect diagnosis. With early

[†]*Present address*: PA-II, EIRA Division, NEERI, Nehru Marg, Nagpur 440 020, Maharashtra, India.

and correct diagnosis, WD patients can be managed by medicines reducing morbidity and mortality.

The causal gene for WD is *ATP7B,* mapped to locus 13q14.3 and is coded by 21 exons spread over 80 kb of genomic DNA[2]. WD occurs at a frequency of 1 in 30,000 whereas its carrier frequency is estimated to be 1:90[3]. More than 500 different disease causing mutations in *ATP7B* have been reported[4] and are compiled as a database[5]. In addition to DNA sequencing, several assays have been developed for detection of point mutations[6]. These include amplification-refractory mutation system (ARMS)[7], combined chain reaction (CCR)[8], denaturing gradient gel electrophoresis, (DGGE)[9], and many variants of single-strand conformation polymorphism (SSCP) electrophoresis[10]. High resolution melting[11] (HRM) and quantitative real-time PCR[12] are some of the techniques that do not require a separation step for the detection of genetic differences. Methods based on nucleic acid hybridization[13] including microarrays of various types are gaining popularity due to the high throughput they offer[14].

For the diagnosis of Wilson's disease arrayed primer extension arrays (APEX)[15], hetero-duplex gel analysis[16] and DNA sequencing[17] have been used, while haplotype analysis has been used to identify the underlying common mutations among Indian patients[18]. Most of the patients with WD have been found to be compound heterozygotes (two different mutations) and the mutations seem to be population specific[1]. Screening for known mutations in a specific population can be an important first step in the genetic diagnosis of a disease. That one cannot rule out the disease in the absence of any of the tested mutations is a notable limitation of this approach. Here we report the development of an oligonucleotide microarray that addresses the 62 WD causing mutations reported from Indian population. We evaluated the quality of the arrays using inter- and intra-array comparisons and validated them with the help of site-direct mutagenesis derived mutant samples.

Material & Methods

This study was conducted in the Molecular Biology Division of Bhabha Atomic Research Centre (BARC), Trombay, Mumbai, Maharashtra, India, during 2009-2013. Table I lists the mutations tested in this study along with a reference to the article that identified / characterized the mutation(s). The overall strategy used in this study is depicted in Fig. 1.

PCR amplification of the 16 exons from gDNA samples: Primers for amplifying 250 to 700 base-pair fragments were designed using programme Primerquest (http://www.idtdna.com/Scitools/Applications/Primerquest). To minimize non-specific amplification, some of the amplicons were re-amplified with the use of nested primers. The sequences of the primers and the resulting fragment lengths are mentioned in Table II. Amplifications were carried out under standard conditions with 94°C melting, 65°C (or 55°C) annealing and 72°C, 30 sec extension and 30 cycles. Synthesized primers and reagents for PCR (Taq polymerase enzyme, buffer, dNTPs) were procured from Board of Radiation & Isotope Technologies (BRIT), Department of Atomic Energy (DAE), Mumbai. Amplification products were checked by agarose gel electrophoresis.

Generation of site-directed mutation derivatives: To evaluate the quality of the arrays with respect to detecting mutations and to simulate patient samples, site directed mutagenesis (SDM) products from five amplicons exon 4, 7, 8, 9 and 15 were generated. Overlapping and complementary internal primers carrying the mutations to be introduced in the exons were used in combination with the original end primers to generate two overlapping sub-fragments each from the desired exonic regions. All the primers used for SDM are listed in Table III. Annealing temperature was reduced to 40°C for the first 10 cycles of PCR and was kept 52.5°C, for the remaining 25 cycles. In the second round of PCR, the products of left and right mutated fragments from round I were mixed in 1:1 ratio and extension reactions were done by the addition of dNTPs and Taq polymerase at 72°C. Final products were obtained by re-amplification with the use of the respective end primers to arrive at the mutant products. These SDM products were verified by sequencing.

Design of probes and printing the microarray: DNA sequence of Homo sapiens *ATP7B* was downloaded from NCBI database (*http://ncbi.nlm.nih.gov*, Accession No. NG_008806). Information regarding Wilson's disease causing mutations specific to Indian population was extracted from Wilson Disease Mutation Database (http://www.wilsondisease.med.ualberta.ca/database.asp). Wild-type and mutant probes of length varying from 25 to 32 bases were designed spanning each mutation while keeping the GC content to about 50 per cent. The oligonucleotide probes were designed with the help of the software OligoArray (2.1)[21]. The WD arrays contained 204 probe triplets, 84 corresponding to WT probes, 93 mutant probes; 20

Table I. List of Wilson's disease causing variants used in the present study

Probe ID	Variant name (nucleotide)	Variant type	Amino acid change	Exon no.	No. of studies[$]	Reference*
P_01_2	c.174dupC	insertion	p.Thr59HisfsX19	2	1	75
P_02_2	c.448_452del	deletion	p.Glu150HisfsX11	2	1	75
P_03_2	c.561T>A	substitution	p.Tyr187Stop	2	1	75
P_04_2	c.813C>A	substitution	p.Cys271Stop	2	2	75,121
P_06_2	c.892delC	deletion	p.Gln298LysfsX2	2	1	75
P_07_2	c.997G>A	substitution	p.Gly333Arg	2	1	121
P_08_4@	c.1707+11dupGT	insertion	na	4	1	74
P_09_5	c.1708-1G>C	substitution	na	5	2	4,75
E5_fs_P	c.1747_1748insT	insertion	p.Asn581SerfsX232	5	1	4
P_12_5	c.1771G>A	substitution	p.Gly591Ser	5	1	120
P_13_5	c.1847G>A	substitution	p.Arg616Gln	5	1	121
M_13_5b	c.1849dupG	insertion	p.Asp617GlyfsX7	5	1	74
P_15_7	c.1963delC	deletion	p.Leu655CysfsX13	7	1	121
E7_fs_P@	c.2116_2117del	deletion	p.Val706ProfsX48	7	2	74,120
P_18_8	c.2128G>A	substitution	p.Gly710Ser	8	1	75
P_19_8	c.2145C>A	substitution	p.Tyr715Stop	8	1	121
P_20_8	c.2224insA	insertion	p.Val742AspfsX13	8	1	74
P_21_8@	c.2258dupC	insertion	p.Glu754Stop	8	1	74
E8_2fs_P	c.2292_2312del	deletion	p.Asp765_Phe771del	8	2	59,97
M_24_8a	c.2303C>T	substitution	p.Pro768Leu	8	1	121,75
M_24_8b	c.2304dupC	duplication	p.Met769HisfsX26	8	1	75,74
P_26_9@	c.2364delC	deletion	p.Ser789GlnfsX18	9	1	74,121
P_27_9	c.2383C>T	substitution	p.Leu795Phe	9	1	121
P_28_10(9)	c.2448-1G>A	substitution	na	10	1	74
M_29_10	c.2497dupG	insertion	p.Val833GlyfsX21	10	1	75
E11_fs_P	c.2582_2583insG	insertion	p.Met862HisfsX5	11	1	74
P_31_11	c.2623G>A/A>G	substitution	p.Gly875Arg	11	1	127
P_32_11	c.2728A>T	substitution	p.Lys910Stop	11	1	74
E12_trp_P	c.2815_2816insA	insertion	p.Trp939Stop	12	1	74
P_34_13	c.2906G>A	substitution	p.Arg969Gln	13	2	121,97
P_36_13	c.2930C>T	substitution	p.Thr977Met	13	1	120
P_37_13	c.2975C>A	substitution	p.Pro992His	13	1	74,73
P_37_13	c.2977dupA	insertion	p.Thr993AsnfsX35	13	2	74
P_40_13	c.3007G>A	substitution	p.Ala1003Thr	13	2	74
M_40_13b	c.3008C>T	substitution	p.Ala1003Val	13	2	106,121
M_43_13a	c.3026_3028del	deletion	p.Ile1009del	13	1	117
M_43_13b	c.3029A>G	substitution	p.Lys1010Arg	13	2	121,126
M_43_13c	c.3031_3032insC	insertion	p.Gly1011AlafsX17	13	1	74
P_47_14	c.3091A>G	substitution	p.Thr1031Ala	14	1	120

Contd...

Probe ID	Variant name (nucleotide)	Variant type	Amino acid change	Exon no.	No. of studies$	Reference*
P_48_14	c.3147delC	deletion	p.Thr1050HisfsX71	14	2	4,120
P_50_14	c.3182G>A	substitution	p.Gly1061Glu	14	2	121,126
E14_fs_P	c.3207C>A	substitution	p.His1069Gln	14	**	**
P_52_15	c.3282C>G	substitution	p.Phe1094Leu	14	2	121,115
P_54_15@	c.3301G>A	substitution	p.Gly1101Arg	15	1	4
P_55_15@	c.3305T>C	substitution	p.Ile1102Thr	15	4	4,74,106,120
P_59_15	c.3311G>A	substitution	p.Cys1104Tyr	15	1	74
P_60_15	c.3412+1G>A	substitution	na	15	1	120
P_61_16	c.3418delT	deletion	p.Val1140Ala-fs	16	1	74
P_62_16	c.3424dupC	insertion	p.Gln1142ProfsX11	16	1	74
P_63_16	c.3532A>G	substitution	p.Thr1178Ala	16	1	120
P_64_18	c.3722C>T	substitution	p.Ala1241Val	18	1	121
P_65_18	c.3767A>G	substitution	p.Gln1256Arg	18	2	74,106
P_67_18	c.3770_3771insG	insertion	p.Asn1257LysfsX2	18	1	74
P_68_18	c.3809A>G	substitution	p.Asn1270Ser	18	1	121
P_69_18	c.3839_3840insTAC	insertion	p.Met1280delinsIleThr	18	1	74
P_70_18	c.3890T>A	substitution	p.Val1297Asp	18	1	121
P_71_18	c.3895C>T	substitution	p.Leu1299Phe	18	1	121
P_71_18	c.3895delC	deletion	p.Ile1300SerfsX30	18	1	121
P_73_18#	c.3903+6T>C	substitution	na	18	1	121
P_74_19	c.4021G>A	substitution	p.Gly1341Ser	19	2	121,126
P_76_19	c.4021+3A>G	substitution	na	19	1	121
P_77_21	c.4310dupA	insertion	p.Pro1438AlafsX11	21	2	74,73

*The references numbered as per the listing in the Wilson disease Database (*http://www.wilsondisease.med.ualberta.ca/database.asp*)
**One of the most preponderant mutations across different countries and reported by many groups, but not in Indian patients
$ Number of independent studies reporting this variant
@ SDM mutations (Mutations tested by generating site directed mutagenesis products)
# This variant is also reported to be a highly polymorphic SNP (rs2282057).

exon specific control probes; two positive (digoxigenin labelled primer) controls and five blank (only buffer) spotting controls. For some of the mutations, the array contained probes for both strands, probes of altered length as well as a few probes printed at two locations on the slide. In all, the array addressed all the 62 WD causing mutations reported from Indian population.

Each probe was printed, in triplicate, on epoxy coated slides (Cat No: 40044, Corning, USA) in 3x SSC buffer at about 50 per cent relative humidity with the help of an in-house developed precision microarrayer. Post printing, the arrays were stored under desiccation at room temperature till further use.

Sample preparation: A commercial preparation of the human genomic DNA (Cat No. 11691112 001, Bangalore Genei, India) was used for amplifying the *ATP7B* regions. All the 16 amplicons that cover the 62 mutations were pooled in equal moles of each PCR product and were labelled with digoxigenin (DIG High Prime Labelling Kit, Roche Life Science, USA). The samples used for different experiments were derived from either WT chromosomal DNA or site-directed mutant samples (simulated Wilson disease samples). The SDM samples were prepared by replacing one of the wild-type exons with the corresponding mutant derivative. For example, SDM4 had the mutant

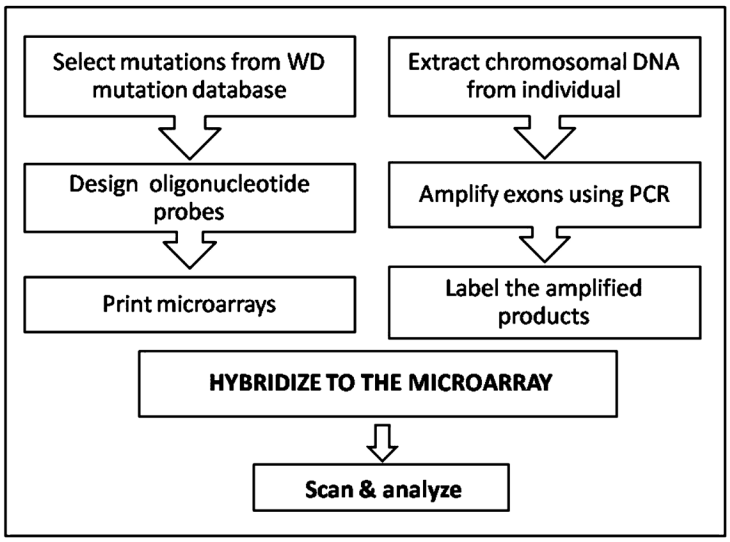

Fig. 1. Flowchart showing the overall strategy used in this study.

Table II. Primer sequences for the 16 amplicons used in this study

Sl. No	Exon no.	Amplicon size (bps)	Left forward primer	Right reverse primer
1	2a	692	TGCCAGAGAAGCTGGGATGTTGTA	ATATCAATTGGTCCCAGGCTTAAG
2	2b	723	CTTAAGCCTGGGACCAATTGATAT	CTCACCTATACCACCATCCAGGAG
3	4	291	TGACCTGATGGTTCCAGGTGTTCT	ACAAAAACCAGACACGTCCAAGATGG
4	5	495 455	AGTCCAGGGTCTTGAGAGCAGT TCTTGGCTGCCTGTTACCTAGACT	TTCACTGATATCCTCCCTCAGATTA
5	7	496 288	ATGTTTAGACCTCTAGATGCTCCCT ATCCAGGTGACAAGCAGCATCTGA	GAAAGCTGCAATAAAGTGCCAT ATATCTGAGGGCCACACACAGCAT
6	8	555 482	CAGTAGTCCTCTGAATGGGAAAGTA GCACAAAGCTAGAGGCTTTGCCAT	GCACCTTAATTATATGGAGGTTTCC
7	9	392 297	CATGTGTGGTGGATAGCAAGTAAC AGCTGTCTCTAACACCACGCTTGT	GAGTGGTGATCTTACTGTGTCTCTG TCTGCCCACACTCACAAGGTCTAT
8	10/ 11	916 546	TGATAAGTGGCGTTTGTTGCAGGG	CTACTCTGGCTTAGATTTTGCTGTC GGGATAATCTCCTTCATTTAACCAC
9	12	916 324	TGATAAGTGGCGTTTGTTGCAGGG AAGAGTTCTGGGAAATCAGACAGTT	CTACTCTGGCTTAGATTTTGCTGTC ACAACCACCATATAGCCCAAGGCA
10	13	500 431	TGTGGAATACCATCTGTTTCCG	GGCTACTCTGTTGCTACTGTTGTTA TGTCTTGAGTGGCTCTCAGGCTTT
11	14	419 370	TAGGAAGCTGTGCAGGTGTCTTGT AAGTTCTGCCTCAGGAGTGTGACT	TTTCCAGACCACACAGAGAAGGCT
12	15	497 393	TTAACCTTTCCTATCTGTTCCACCT	GAACATAAGAGAAACTTTCCTGGGT TCTGTGGTTTGACCCACCTCTACT
13	16	396	GGTGCTTACAAGGTTACAGTTTTTC	CCTGAAATTAAGAGAGGAAGGCT
14	18	479 458	AAGTCTTTGCAGAGGTGCTGCCTT GTTGACCAACATCACTGACTGG	AGGTTGATGCGTATCCTTCGGACA ACAGTCCTCTTGGAAAGGTGAAT
15	19	386	CTCACTGTGTGCTCGTCTCCATCA	ACAGCCAAGCATCTCCACTAGCTT
16	21	452	GAATGGCTCAGATGCTGTTGCGTT	CAGGGCAGGATGACTGGACATATC

All sequences are mentioned in the 5'->3' direction. Additional entries in a row indicate that a second round of PCR (nested) was done for that amplicon using the primer(s) mentioned in the second lane. Accordingly, the amplicon sizes are also listed. Sequence source: *http://ncbi.nlm.nih.gov*, Accession No. NG_008806

Table III. Primers used for generating site directed mutation (SDM) samples

Exon/bps	SDM fragments (bps)	Mutation type	Mutagenesis (Middle) primers (MP)
4 / 291	113, 207	Two base insertion	5'-CAGTAAGTACTGTGTGGGTGCGTTACG-3' 3'-GTCATTCATGACACACCCACGCAATGC-5'
7/ 288	238, 78	Two base deletion	5'-TTGTGTACCTTT--CCAGGTATATATG-3' 3'-AACACATGGAAA--GGTCCATATATAC-5'
8/ 482	253, 255	Single base insertion	5'-TGGTTGCTGTGGCCTGAGAAGGCGGA-3' 3'-ACCAACGACACCGGACTCTTCCGCCT-5'
9/ 297	116, 207	Single base deletion	5'-TAGAGCAAAA-CTCAGAAGCCCTGG-3' 3'-ATCTCCTTTT-GAGTCTTCGGGACC-5'
15/ 393	236, 182	G->A point mutation	5'-GTGCCAGGCTGTAGAATTGGGTGCA-3' 3'-CACGGTCCGACATCTTAACCCACGT-5'
15/393	236, 182	T->C point mutation	5'-GTGCCAGGCTGTGGAACTGGGTGCA-3' 3'-CACGGTCCGACACCTTGACCCACGT-5'

Exons and fragment length are given in column 1. Size of the two overlapping mutated fragments generated after the first round of PCR are listed in the 2nd column. The mutations were selected from Table I and the primers as given in Table II were used for final amplification of the mutant product.

amplicon for exon 4 and the wild-type amplicons from all the remaining amplicons. Similarly SDM7, SDM8, SDM9 and SDM15 were prepared.

WD microarray hybridizations: Hybridizations were carried out using reagents and protocols as per the manufacturer (Roche Life Science, USA). Briefly, prior to hybridization, the slides were treated in pre-hybridization buffer and were overlaid with labelled sample and were incubated overnight at 42°C or 52°C under a cover-slip. Post-hybridization washes were carried out as per the manufacturer's protocols. Hybridization signals were visualized using alkaline phosphatase conjugated anti-DIG antibody and NBT/BCIP (nitro blue tetrazolium and 5-bromo 4-chloro 3'-indolyl-phosphate) colorimetric substrate.

Image processing and quantification: The slides were scanned on a flatbed scanner at 600 dpi resolution and spot intensities were quantified with the help of ImageJ software and the plug-in 'Microarray Profile' (*http://imagej.nih.gov/ij*). This plug-in allows the user to define a grid of circles, fixing spot area for quantitative comparison of spots from the gray values obtained for each of the spot on hybridized microarray.

Statistical analysis & discrimination score: The intensity data from all the spots were analyzed by using Microsoft Excel (TM) and various statistical functions available therein. The mean intensity and the standard deviation were computed from the three spots for each probe and the spots that showed large standard deviation compared to the mean were physically examined to identify and remove outliers. Mean background was estimated by selecting six intra-spot regions (of the same size as was used for spot quantification) in each grid. The background subtracted intensity values were normalized with reference to the DIG labelled spots on the array to obtain the final intensity values for each probe.

Discrimination score (DS) is the most commonly used index for gene expression and genotyping microarrays[19]. The intensity values of the perfect match (PM) probe and the corresponding mismatch (MM) probe were used to compute the discrimination score for each of the probe pair (mutant and normal/wild type).

$$DS = (IP - IM) / (IP + IM)$$

Where IP is the average intensity value from the PM probe triplet and IM is the average intensity value from corresponding MM probes.

Ideally, the value of DS is always greater than 0 (IP > IM) for arrays hybridized with wild-type sample while DS becomes negative when the array is hybridized with a mutant sample (IP < IM). The greater the DS value, greater is the discrimination ability of that probe pair.

Results

Microarrays carrying probes to 62 mutations that were specific to Indian population were printed and

were evaluated for inter- and intra-array spot intensity variations, and were further assessed for their ability to detect mutation by using simulated patient samples, *i.e*. samples containing mutations at defined locations by site directed mutagenesis (Sigma-Aldrich, USA).

Microarray hybridizations: Image of one of the microarrays hybridized to wild-type sample is shown in Fig. 2. It should be noted that the picture depicted here is a magnified portion of a 13.5 mm x 13.5 mm region on the slide. The first and last triple spots (Fig. 2A) are the positive control spots (DIG labelled oligo). The spot intensities varied from probe to probe even for the wild-type probes. However, for most of the probe-pairs, the perfect-match (PM) probes yielded higher intensities compared to the corresponding mismatch (MM) probes (Fig. 2B).

Stringency of post-hybridization washes: The effect of altered post-hybridization wash temperatures (stringency) on the final spot intensities was assessed by comparing the spot intensities obtained with wild-type human chromosomal DNA samples treated at 52 and 42°C post-hybridization wash temperatures. The Pearson's product moment correlation coefficient, R and the R^2 values were computed from five independent samples at each temperature. The R^2 values were found to be in the range of 0.7 to 0.92 indicating a good correlation between the two data sets. One of the scatter plots is shown in Fig. 3 as an example. It was observed that the spot intensities obtained at 52°C were lower than those obtained at 42°C.

Inter-experiment comparison of spot intensities: Reproducibility of intensity data (of either PM or MM probes) across experiments is necessary for applying any further processing to the data. The reproducibility was assessed by comparing the intensity data from 61 perfect match probes from six independent experiments each carried out at 52 and at 42°C. The dispersion of intensities (standard deviation computed from inter-array experiments) were found to be lower for the hybridizations done at 52°C (Fig. 4) as compared to those at 42°C (data not shown). Inspite of the lower mean intensities obtained from post-hybridization washes at 52°C, better data reproducibility was observed at higher temperature. Hence the 52°C data were used for all further analysis.

Intra-experiment comparison of spot intensities: The design of arrays included a total of 20 probes that were spotted (in triplicates) at two different locations on the array. The spot intensities obtained from two different locations of the same array were found to be well correlated (Fig. 5, R^2=0.72). A similar correlation was observed for different repeats of the experiments

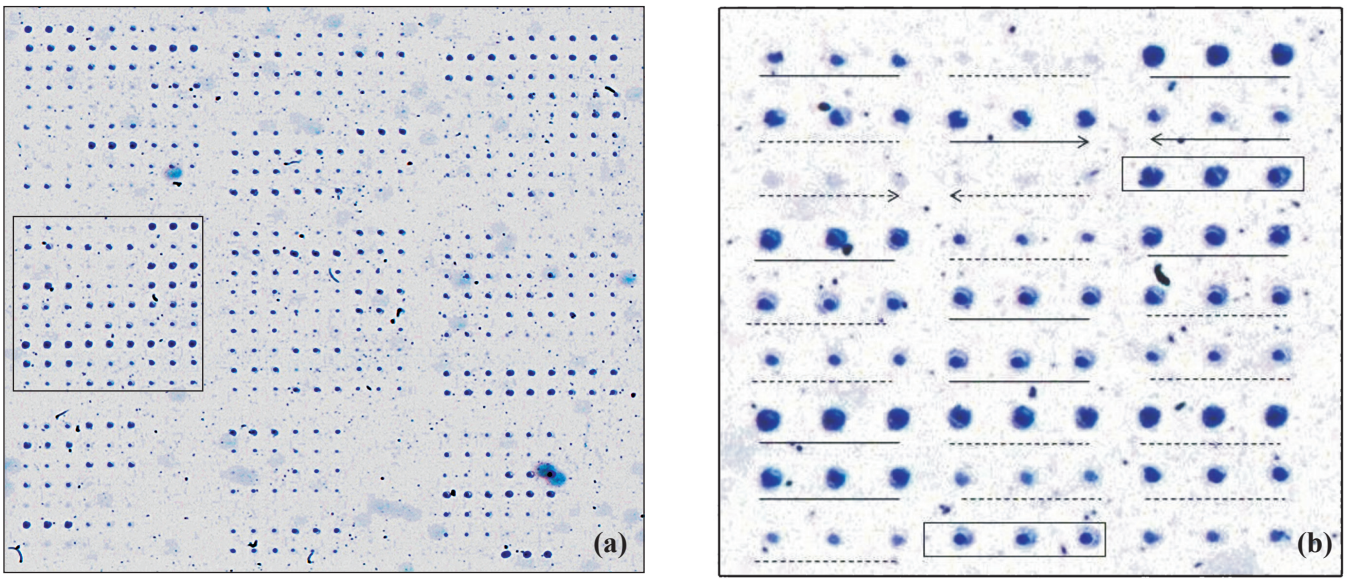

Fig. 2(a). Magnified view of a scanned microarray (13.5 x 13.5 mm) hybridized to wild-type human genomic DNA sample [inset is shown in Fig.2(b)]. **2(b).** A portion of the array wherein, the spots corresponding to wild-type and mutant and probes are shown as underlined with solid or dashed lines, respectively. Perfect match (PM) probes are underlined by a continuous line while the adjoining mismatch (MM) probes are underlined by dotted lines. The forward and reverse complement probes and their corresponding mutant probes are indicated by arrows. Spot triples shown in boxed region correspond to exon-specific positive control probes.

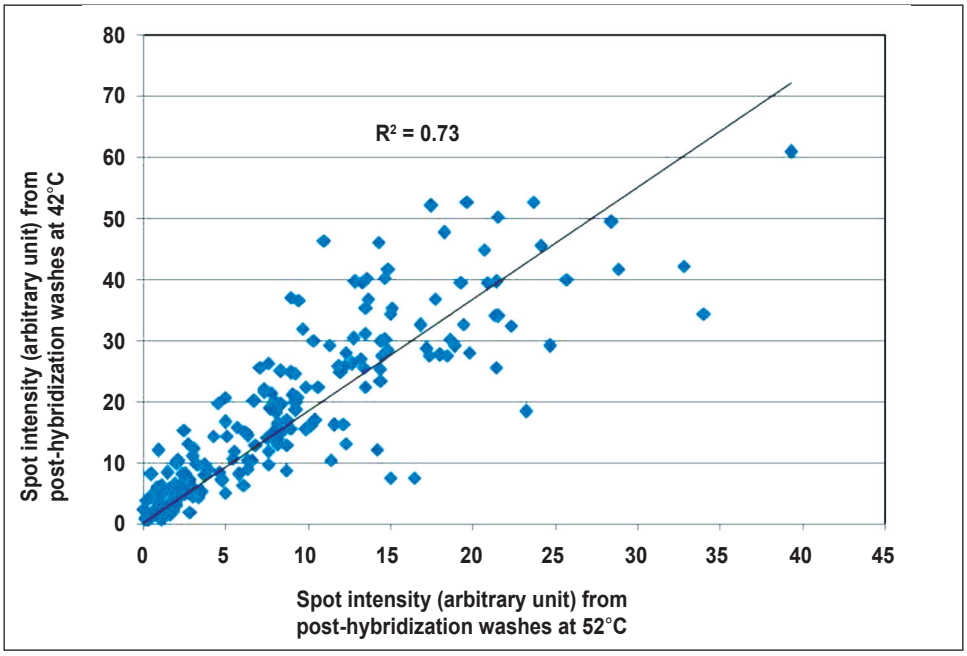

Fig. 3. A scatter plot of the spot intensities (arbitrary units) obtained from post-hybridization washes done at 52°C (X-axis) as compared to the spot intensities obtained from the corresponding spots at 42°C (Y-axis). The correlation coefficient R^2 for these data is indicated in the Figure.

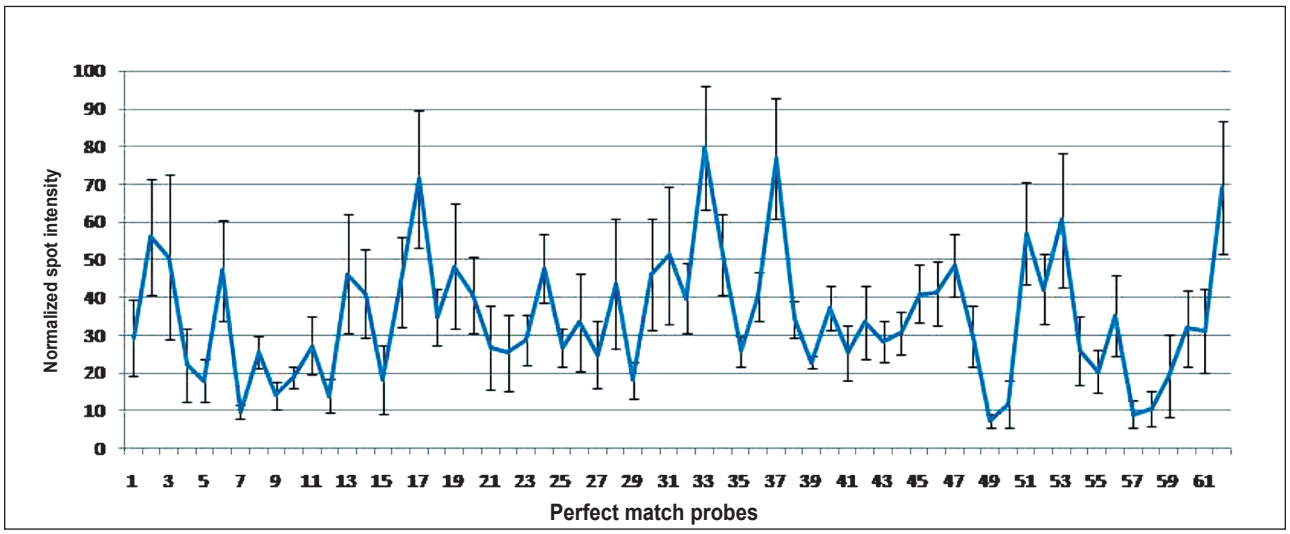

Fig. 4. Normalized spot intensities obtained for 61 perfect match probes from six independent hybridizations. The error bars indicate one sigma (Standard Deviation) from the mean.

indicating that the spot intensities (hybridization reactions) across a given slide were consistent.

Discrimination of perfect-match (PM) and mismatch (MM) probes: The ability of each PM and the corresponding MM probe pair to differentiate the mutant allele from the wild-type allele was assessed by computing the discrimination scores. Since the sample being labelled is double-stranded, any mutation can be examined by designing a probe for either of the strands. The possible effect of the strand (sequence composition) on detecting the mutation was addressed by comparing discrimination scores for probe pairs for four different mutations derived from forward (same as the mRNA sequence) and reverse complement strand

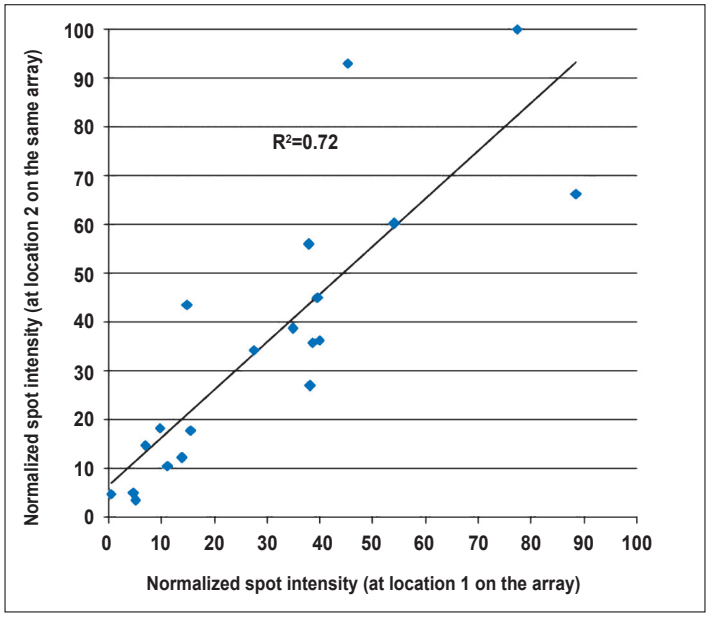

Fig. 5. Correlation observed for the mean intensity data for 20 different probe spots (triplets) spotted at two different locations on the same array. Hybridization was done with wild-type chromosomal DNA sample.

Table IV. Comparison of discriminations scores (DS) for complementary probe pairs

Probe ID	5'- Sequence of the WT probe -3' 5'- Sequence of the mutant probe -3'	DS Forward strand	DS Reverse complement
E5	GTGTCCACAACAT-AGAGTCCAAACT GTGTCCACAACATAAGAGTCCAAACT	0.40 ± 0.06	0.56 ± 0.12
E9	TAGAGCAAAACCTCAGAAGCCCTGG TAGAGCAAAAC-TCAGAAGCCCTGG	0.48 ± 0.16	0.45 ± 0.09
E9_a	TAGAGCAAAACCTCAGAAGCCCTGG TAGAGCAAAAC-TCAGAAGCCCTGG	0.51 ± 0.10	0.46 ± 0.10
E12	ACGTTGGTGGTAT-GGATTGTAATCG ACGTTGGTGGTATAGGATTGTAATCG	0.65 ± 0.10	0.48 ± 0.19

Values are mean ± SD of discrimination scores derived from five independent hybridizations (N=5). E9 and E9_a are the DS values obtained from the same probe pairs printed at two locations on the array.

Table V. Discrimination scores (DS) obtained from a subset of probes hybridized to wild-type sample and site directed mutated fragments (SDM)

Probe ID	Sequence (5'->3')	Nature of mutation	DS of WT sample (N=5)	DS of *SDM sample	Q-test probability
SDM4	CAGTAAGTACTGT--GGGTGCGTTACG	2BI	0.21 ±0.11	- 0.36	0.05
SDM7	TGTGTACCTTTGTCCAGGTATATATGAGAAAG	2BD	0.26 ±0.17	-0.76	0.01
SDM8	TGGTTGCTGTGGC-TGAGAAGGCGGA	SBI	0.16 ±0.05	- 0.02	0.10
SDM9	TAGAGCAAAAC**C**TCAGAAGCCCTGG	SBD	0.49 ±0.13	- 0.36	0.01
SDM15a	TGCCAGGCTGTGGAATTGGGTGCAAAG	G->A	0.41 ±0.17	0.08	0.05
SDM15b	CCAGGCTGTGGAATTGGGTGCAAAG	T->C	0.55 ±0.07	0.03	0.01

*Average of two probe-pairs each
Significance of deviation of the DS values for the mutant samples (6th column) from the wild-type sample is assessed by the Q-test[20]

Proble ID	Wild-type sample		SDM sample	
	PM probe	MM probe	PM probe	MM probe
SDM4				
SDM7				
SDM8				
SDM9				
SDM15-a				
SDM15-b				

Fig. 6. Images of the triplicate spots corresponding to the perfect match and mismatch (PM and MM) probes from one of the arrays hybridized with wild-type and site directed mutagenesis (SDM) samples, respectively.

(Table IV). The discrimination scores obtained from both the probe-pairs were found to be comparable to each other indicating that the strand specific variation of the hybridization signals is negligible.

Summary of probes' response: A total of 76 probe-pairs were spotted in the array including duplicate and reverse-complement probe pairs used for quality assessment. The DS values were computed for all these probe-pairs and the mean DS obtained from five independent hybridizations was used to assess the quality of each probe-pair. For a majority of the probe-pairs (51 of 76) the mean DS value was more than 0.2 while a total of 60 probes (>75%) showed a mean DS value greater than twice the standard deviation indicating that mutant alleles can be discriminated from the wild-type alleles for all these probe-pairs.

Of the 62 mutations that were considered, 43 could be detected with a DS value > 0.2, while an additional seven probe-pairs showed lower albeit reproducible DS values (DS > 2σ). This resulted in more than 80 per cent sensitivity towards detecting mutations. For ten probe-pairs that yielded poor discrimination, it was realized that the performance might be improved by changing the probe length. Most of the remaining probe-pairs showed low DS due to either high GC content (>60% GC) or runs of bases.

Discrimination scores for site directed mutation (SDM) samples: Six different mutations belonging to frameshift and missense categories *i.e.* one each of a two-base deletion, two-base insertion, single base deletion, single base insertion and two point mutations were introduced into specific exonic regions by site directed mutagenesis. For all the SDM samples, the MM probes were expected to yield higher spot intensity as compared to the corresponding PM probes at the site of mutation. A summary of hybridizations with these SDM samples is given in Table V. The significance of the deviation of the DS values for the mutant samples (6th column) compared to the wild-type sample is assessed by the Dixon's Q-test[20]. It was observed that all the simulated Wilson's disease mutations (except SDM8) could be detected with better than 95 % confidence. The spots corresponding to these hybridizations are shown in Fig. 6 and the appearance of the spot intensities are observed to be in line with the quantifications allowing to draw inferences with respect to their ability to detect mutations of different types.

Discussion

This study reports the design and evaluation of a low-density microarray for the detection of 62 mutations in *ATP7B* gene reported from Indian population. The arrays were printed using an indigenously developed

microarrayer and a total of 204 probes were incorporated into the present version of the array to assess the variability in DS due to probe length, mutation position within the probe, strandedness of the probe, position of the probes on the array, *etc*. The hybridization data obtained with 52°C post-hybridization wash showed better inter-array reproducibility compared to that with 42°C. The hybridization across the slide surface seemed to be uniform and the strandedness of the probe (forward versus reverse complement) did not seem to contribute to differences in DS. The DS varied from probe to probe, conceivably as a result of altered sequence composition and probe length. The sensitivity of our array was better than 80 per cent if all the probes that yielded reproducible DS values ($DS > 2\sigma$) were considered. It was observed that low discrimination scores were mainly due to probes that contained runs of bases (≥ 4) or probes with high GC per cent (> 60) or high melting temperature that could have high degree of secondary structures. These clues may be useful for improving the probe design which in turn will further improve the sensitivity of mutation detection. In view of the multiple parameters that affect the sensitivity of mutation detection, it is desirable that the mutations detected by microarray method, particularly the ones with low discrimination scores, are confirmed by an independent molecular biology method for diagnostic purposes.

Microarray technology has been applied for genotyping not only at genomic scale[22-24] but also at individual locus level. A microarray based approach was found to be cost-effective compared to sequence analysis[25], capillary based heteroduplex analysis[26] or re-sequencing[27]. Microarrays to screen for 301 disease-associated sequence variants in Leber Congenital Amaurosis (LCA) related genes[28], to study sporadic non-syndromic hearing loss in children[25] and a genotyping microarray for 298 Usher syndrome-associated sequence variants[27] are a few examples of the utility of this approach. This study also demonstrates that oligonucleotide based microarrays are useful for simultaneously examining a large number of mutations thus becoming an efficient first-pass screening tool in the diagnosis of genetic disorders. Even though fluorescence based methods are more sensitive, the use of colorimetric detection and a simple flat-bed scanner obviates the need for expensive laser based scanners making the assay cost-effective. Insights obtained from this study could be useful for custom designing microarrays for other genetic disorders in future.

Acknowledgment

The authors thank Dr Rita Mukhopadhyaya for sequencing the SDM derivatives and Ms Akanksha Agarwal for providing assistance in data analysis, and Drs Vinay Kumar, N Jawali, M Seshadri and MV Hosur for their valuable suggestions.

References

1. Huster D. Wilson disease. *Best Pract Res Clin Gastroenterol* 2010; *24* : 531-9.
2. Hung IH, Suzuki M, Yamaguchi Y, Yuan DS, Klausner RD, Gitlin JD. Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* 1997; *272 :* 21461-6.
3. Patil M, Sheth KA, Krishnamurthy AC, Devarbhavi H. A review and current perspective on Wilson disease. *J Clin Exp Hepatol* 2013; *3* : 321-36.
4. Luoma LM, Deeb TM, Macintyre G, Cox DW. Functional analysis of mutations in the ATP loop of the Wilson disease copper transporter, *ATP7B*. *Hum Mutat* 2010; *31* : 569-77.
5. Kenney SM, Cox DW. Sequence variation database for the Wilson disease copper transporter, *ATP7B*. *Hum Mutat* 2007; *28* : 1171-7.
6. Nollau P, Wagener C. Methods for detection of point mutations: performance and quality assessment. *Clin Chem* 1997; *43* : 1114-28.
7. Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, *et al.* Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* 1989; *17* : 2503-16.
8. Bi W, Stambrook PJ. CCR: a rapid and simple approach for mutation detection. *Nucleic Acids Res* 1997; *25* : 2949-51.
9. Fischer SG, Lerman LS. DNA fragments differing by a single base-pair substitution are separated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci USA* 1983; *80* : 1579-83.
10. Kakavas VK, Plageras P, Vlachos TA, Papaioannou A, Noulas VA. PCR-SSCP: a method for the molecular analysis of genetic diseases. *Mol Biotechnol* 2008; *38* : 155-63.
11. Wittwer CT. High-resolution DNA melting analysis: advancements and limitations. *Hum Mutat* 2009; *30* : 857-9.
12. Morlan J, Baker J, Sinicropi D. Mutation detection by real-time PCR: a simple, robust and highly selective method; *PLoS One* 2009; *4* : e4584.
13. Zhang Y, Coyne MY, Will SG, Levenson CH, Kawasaki ES. Single-base mutational analysis of cancer and genetic diseases using membrane bound modified oligonucleotides. *Nucleic Acids Res* 1991; *19* : 3929-33.
14. Yoo, SM, Choi JH, Lee SY, Yoo NC. Applications of DNA microarray in disease diagnostics. *J Microbiol Biotechnol* 2009; *19* : 635-46.
15. Gojova L, Jansova E, Kulm M, Pouchla S, Kozak L. Genotyping microarray as a novel approach for the detection of *ATP7B* gene mutations in patients with Wilson disease. *Clin Genet* 2008; *73* : 441-52.
16. Majumdar R, Al Jumah M, Al Rajeh S, Fraser M, Al Zaben A, Awada A, *et al*. A novel deletion mutation within the carboxyl

terminus of the copper-transporting *ATPase* gene causes Wilson disease. *J Neurol Sci* 2000; *179* : 140-3.

17. Gupta A, Chattopadhyay I, Dey S, Nasipuri P, Das SK, Gangopadhyay PK, *et al.* Molecular pathogenesis of Wilson disease among Indians: A perspective on mutation spectrum in *ATP7B* gene, prevalent defects, clinical heterogeneity and implication towards diagnosis. *Cell Mol Neurobiol* 2007; *27* : 1023-33.
18. Gupta A, Aikath D, Neogi R, Datta S, Basu K, Maity B, *et al.* Molecular pathogenesis of Wilson disease: haplotype analysis, detection of prevalent mutations and genotype-phenotype correlation in Indian patients. *Hum Genet* 2005; *118* : 49-57.
19. Zhou Y, Abagyan R. Algorithms for high-density oligonucleotide array. *Curr Opin Drug Discov Develop* 2003; *6* : 339-45.
20. Rorabacher DB. Statistical treatment for rejection of deviant values: critical values of Dixon's "Q" parameter and related subrange ratios at the 95 per cent confidence level. *Anal Chem* 1991; *63* : 139-46.
21. Rouillard JM, Zuker M, Gulari E. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res* 2003; *31* : 3057-62.
22. Kennedy GC, Matsuzaki H, Dong S, Liu WM, Huang J, Liu G, *et al.* Large-scale genotyping of complex DNA. *Nat Biotechnol* 2003; *21* : 1233-7.
23. Bejjani BA, Saleki R, Ballif BC, Rorem EA, Sundin K, Theisen A, *et al*. Use of targeted array-based CGH for the clinical diagnosis of chromosomal imbalance: is less more? *Am J Med Genet A* 2005; *134* : 259-67.
24. Syvänen AC. Toward genome-wide SNP genotyping. *Nat Genet* 2005; *37* (Suppl.): S5-10.
25. Hu X, Liang F, Zhao M, Gong A, Berry ER, Shi Y, *et al.* Mutational analysis of the *SLC26A4* gene in Chinese sporadic nonsyndromic hearing-impaired children. *Int J Pediatr Otorhinolaryngol* 2012; *76* : 1474-80.
26. Baaj Y, Magdelaine C, Ubertelli V, Valat C, Talini L, Soussaline F, *et al.* A highly specific microarray method for point mutation detection. *Biotechniques* 2008; *44* : 119-26.
27. Cremers FP, Kimberling WJ, Külm M, de Brouwer AP, van Wijk E, te Brinke H, *et al*. Development of a genotyping microarray for Usher syndrome. *J Med Genet* 2007; *44* : 153-60.
28. Yzer S, Leroy BP, De Baere E, de Ravel TJ, Zonneveld MN, Voesenek K, *et al*. Microarray-based mutation detection and phenotypic characterization of patients with Leber Congenital amaurosis. *Invest Ophthalmol Vis Sci* 2006; *47* : 1167-76.

Reprint requests: Dr Akkipeddi V.S.S.N. Rao, Molecular Markers & Gene Expression Studies, Molecular Biology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400 085, Maharashtra, India
e-mail: narayana@barc.gov.in, avssnr@gmail.com