

DNA DIAGNOSTICS FOR WILSON DISEASE

By

MANJULA MATHUR

LIFE01200804006

A thesis submitted to the

Board of Studies in Life Sciences

In partial fulfillment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

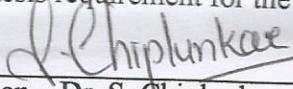


August 2014

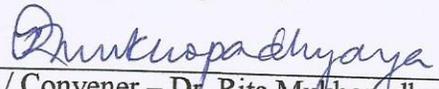
Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

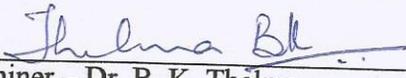
As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mrs Manjula Mathur entitled "DNA Diagnostics for Wilson Disease" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.


Chairman – Dr. S. Chiplunkar

Date: 10/04/2015


Guide / Convener – Dr. Rita Mukhopadhyaya

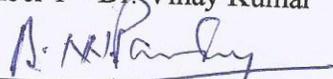
Date: 10/04/2015


Examiner – Dr. B. K. Thelma

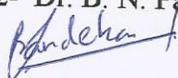
Date: 10/04/2015


Member 1 – Dr. Vinay Kumar

Date: 10/04/2015


Member 2- Dr. B. N. Pandey

Date: 10/04/2015


Member N – Dr. J. R. Bandekar

Date: 10/04/2015

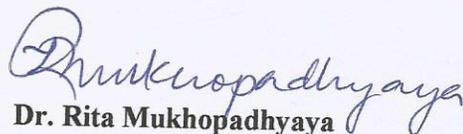
Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.

I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 10.04.2015

Place: BARC, TROMBAY, MUMBAI

Co-guide (if applicable)


Dr. Rita Mukhopadhyaya

Guide

¹ This page is to be included only for final submission after successful completion of viva voce.
Version approved during the meeting of Standing Committee of Deans held during 29-30 Nov 2013

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Manjula Mathur

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Manjula Mathur

DEDICATIONS

This thesis is dedicated to Wilson Disease patient *Mr. Ketan Ganatra* who inspired and encouraged me to take up this project for early and accurate diagnosis of WD.

My very high regards to my respected father *Late Shri I. D. Mathur*, who gave me training to think independently from very young age.

I dedicate my creative ideas to Almighty God to convert them into action for reducing human suffering, by development of Wilson Disease DNA Microarray (WDDM) for diagnostics of rare and complex Wilson Disease. The same method of DNA microarray can be applied for detection of other monogenic diseases in suspected patients.

ACKNOWLEDGEMENTS

I wish to acknowledge deep sense of gratitude with sincere thanks from the core of my heart to my guide Dr. T.B. Poduval, ex-Head, Immunology Section, Radiation Biology & Health Sciences Division (RBHSD), Bhabha Atomic Research Centre (BARC), Trombay. His support, valuable advise and constant encouragement for executing this research work was the constant source of uninstited inspiration for me.

My profound thanks to Dr. A.V.S.S. Narayan Rao, In-Charge, Microarray Facility, Molecular Biology Division (MBD), BARC, Trombay for sparing his valuable time from his very busy schedule. His illuminating guidance and scientific inputs for design and development of WDDM microarray helped me in executing my concept into actual experimental work. Without his painstaking contributions it would not have been possible to generate data for this thesis.

It is my great pleasure to thank Dr. Rita Mukhopadhyaya, In-Charge, DNA Sequencing Facility, MBD, BARC, Trombay who agreed to be my guide, consequent to retirement of Dr. T.B. Poduval. I am grateful for her valuable technical discussions and help in shaping the write up for this thesis.

I wish to express my gratitude to Dr. Vinay Kumar, High Pressure & Synchrotron Radiation Physics Division (HP&SRD), BARC, Trombay for his strong faith in me and in my project. His role as a doctoral committee member from inception of this work often demanded him to also be the soliciter among team members.

I shall be ever grateful to all members of my DRC who critically reviewed the research work. It is my great privilege to thank the present Chairman, Dr. J.R. Bandekar (Head, RBHSD, BARC, Trombay) and Dr. B.N. Pandey, RBHSD,

BARC, Trombay for spending their valuable time at the later stage of this work. My regards to Dr. M. Sheshdri, Dr. N. Jawali and Dr. M.V. Hosur ex-BARC scientists for their valuable guidance as members of my original DRC. Last but not the least, my sincere thanks to the members of BARC Medical Ethics Committee, for approving this project work that involved using human material.

Validation of this microarray would not have been possible without the support of all the patients and their relatives who agreed to give blood samples for research purpose. My sincere thanks to all of them and their families. My special thanks to Dr. Abha Naagral, Gastroenterologist, Jaslok Hospital, Mumbai, for clinical identification of Wilson Disease patients.

I am highly obliged to Dr. Pravin Potdar (Head, Molecular Medicine, Jaslok Hospital, Mumbai) whose talk (LIFE Science Symposium 2007 at BARC, Trombay) on WD patients planted the concept in my mind to take up this project of developing **“DNA Diagnostics for Wilson Disease”**.

I profusely thank my family, friends and well wishers for their continued support through difficult times and to overcome various technical and social challenges. I am indebted to my husband who stood by my side as the pillar providing me moral support during all ups and downs and in executing each of the subtle steps.

My acknowledgements would not be complete without the names of Project Trainees Ms. Ekta Singh Parihar and Ms Akansha Agrawal. Ms Ekta very meticulously repeated experiments providing confirmation of results on validation of WDDMs. Ms Akansha helped in analysis of scanned images.

CONCEPT ACKNOWLEDGEMENTS

WATSON CRICK MODEL OF DNA DOUBLE HELIX, KHORANA'S INVENTION OF CHEMICAL SYNTHESIS OF DNA AND DEVELOPMENT OF SOUTHERN HYBRIDISATION ARE THE TRIPOD OF SCIENTIFIC DEVELOPMENTS FROM THE LAST CENTURY FOR DNA-MICROARRAY DEVELOPMENTS IN THE 21ST CENTURY.

CONTENTS

Section No.	Topics	Page No.
	SYNOPSIS	1-13
	CHAPTER I – INTRODUCTION	
1.1	History & Current Scenario of Wilson Disease	14
1.2	Genetics, Epidemiology and Phenotype of WD	16
1.3	Regulation of Homeostasis of Transition Metals and Systems Approach to Understand Wilson Disease	17
1.4	Primary, Secondary and Tertiary Structure of <i>Atp7b</i> Protein and its function in normal Physiology	21
1.5	Clinical Diagnosis and Assay of WD	28
1.6	How is WD Treated	30
1.7	WD Causing Mutations in Gene <i>Atp7b</i>	30
1.8	Techniques for Mutation Detection	32
1.9	DNA-Micro arrays for Detection of Complex Genetic Disorders	33
1.10	DNA- Detection for WD Mutations	35
1.11	Rationale for the Present Study	37
1.12	Scope of the Present Study	38
	CHAPTER II - MATERIALS AND METHODS	
2.1	Molecular Sequence Data of the <i>Atp7b</i> For WDDM	39
2.2	Functional Domains And Corresponding Exons	40
2.3	Mutation Data Base for WDDM	42
2.4	Molecular Sequence Details of 16 Amplicons	45
2.5	Site Directed Mutagenesis	58
2.6	Printing of WDDM with Oligo Probes	59
2.7	Genomic DNA Isolation from Peripheral Blood	62
2.8	Softwares used	64
2.9	Amplification of DNA by PCR from GDNA Samples	64
2.10	Agraose Gel Electrophoresis of PCR Products	65
2.11	Re-PCR of 16 Exons	65

2.12	Clean up Protocol of PCR Products	66
2.13	Protocols for Site Directed Mutagenesis	67
2.14	Printing of Microarrays	70
2.15	Dig Labeling of Amplicons	71
2.16	Hybridization on WDDMS	72
2.17	Detection of Probe-Target Hybrids by Chromogenic Method	74
2.18	Quantification of Spot Intensities	75
2.19	Intra-Array and Inter-Array Correlation Study of Intensity Values	76
2.20	Differentiability of WT & its Corresponding Mutant Probes	77
2.21	Sequencing Protocol	79
2.22	Buffers Used	79
CHAPTER III - RESULTS AND DISCUSSION		
3.1	Genomic DNA Isolation	82
3.2.	PCR Amplification, Re-Amplification Including Nested PCR	82
3.3	PCR Amplifications from Patient and ST Samples for Testing on WDDM	85
3.4	Site Directed Mutagenesis of WT Amplicons And Kinetics of PCR Reactions in Mutagenesis	91
3.5	Sequencing of SDM Fragments	95
3.6	Prototype Microarrays for Mutation Detection	99
3.7	WD Detection Microarray Design Considerations	99
3.8	Sample Preparation	101
3.9	Hybridization and Post Hybridization Processing	102
3.10	Hybridizations Carried out for Evaluating WDDM	103
3.11	Effect of Post-Hybridization Wash Temperature on Spot Intensities	104
3.12	Statistical Analysis of Dispersion of Intensities	106
3.13	Intra array correlation of spot intensities	110
3.14	Discrimination Score (DS) For Each Probe Pair	111

3.15	DS Values Of Probes Tested By SDM	115
3.16	Hybridization Experiments with Patient Samples	116
3.17	Patient P-12 And Family Members	120
3.18	Cost Estimation of Detection Using Wddm	122
	CHAPTER IV	
4.1	Summary of Results	123
4.2	Conclusions	124
	References	126-133

LIST OF FIGURES

Figure No.	Description	Page No.
FIGURES IN CHAPTER I		
1.1	Dr. Samuel Alexander Kinnier Wilson (1878-1937)	15
1.2	A mammalian cell and copper trace element.	18
1.3	Proteins involved in homeostasis of Fe, Cu and Zn within a hepatocyte.	20
1.4A	Amino acid sequence of human ATP7B	23
1.4B	Conserved motifs in ATP7B sequence.	23
1.5	Schematic representation of the transmembrane organization of ATP7B.	24
1.6A	Hypothetical ribbon model of 3D structure of ATP7B	25
1.6B	Multiple species alignment of the conserved regions of ATP7A & ATP7B	26
1.7	A model to understand role of phosphorylation/dephosphorylation of ATP7B	27
1.8A	Algorithm for WD diagnosis as adopted by Mayo Clinic, USA	29
1.8B	A typical KF ring (brown) in the retina	29
1.9	Exon wise distribution of WD causing mutations on <i>ATP7B</i> gene	31
1.10	Exon wise distribution of WD mutations in Western Indian population.	32
1.11	Geographical map of places in India showing WD	32
FIGURES IN CHAPTER II		
2.1	Exonic structure of <i>ATP7B</i>	28
2.2	Exon - Intron junctions and splicing of RNA	29
2.3	Chromosome 13 alongwith 23 human chromosomes	41
2.4	Bond formation by epoxy coating on the surface of glass	61
2.5	Micro-arraying robotic machine used for preparation of WDDM	62
2.6	Format of printing WDDM on epoxy-coated glass slide	62
2.7	Brief schematics of methodology for WD mutation detection.	63
2.8	Schematic of site directed mutagenesis from WT amplicon	70
2.9	Schematics for Site Directed Mutagenesis Methodologies	71

2.10	Sixteen amplicons pooled for simulating SDM4 and SDM7.	72
2.11	Grid for microspots	79
	FIGURES IN CHAPTER III	
3.1	Agarose gel showing exon 5, 7, 8, 9, 11, 13,14,15 & 18	83
3.2A	Agarose gel showing Set A of amplicons (control DNA)	84
3.2B	Agarose gel showing Set B of amplicons (control DNA)	84
3.3	Agarose gel showing Set A of amplicons of patient P9, P9-1, P9-2	86
3.4	Agarose gel showing SetB of amplicons of patient P9, P9-1, P9-2	86
3.5	Agarose gel showing SetA of amplicons of P9-3, P12 and P12-1	86
3.6	Agarose gel showing SetB of amplicons of P9-3, (2) P12 and (3) P12-1	86
3.7	A set of 8 amplicons each of (1) P12-2, (2) P12-3 and (3) P5	87
3.8	B set of 8 amplicons each of (1) P12-2, (2) P12-3 and (3) P5	87
3.9	A set of 8 amplicons each of (1) P7, (2) P8 and (3) P10	87
3.10	B set of 8 amplicons each of (1) P7, (2) P8 and (3) P10	87
3.11	A set of 8 amplicons each of (1) C1, (2) C2 and (3) C3 gDNA templates	88
3.12	B set of 8 amplicons each of (1) C1, (2) C2 and (3) C3 gDNA templates	88
3.13	A set of 8 amplicons each of (1) C4, (2) C5 and (3) Ch	88
3.14	B set of 8 amplicons each of (1) C4, (2) C5 and (3) Ch	88
3.15	Aliquots from 1 st step of site directed mutagenesis PCR.	92
3.16A, B,C	SDM Fragments with kinetics	94
3.17	Snap shot of electrophoregram of forward strand of SDM4.	95
3.18	Snap shot of electrophoregram of forward strand of SDM7.	96
3.19	Snap shot of electrophoregram of forward strand of SDM8.	96
3.20	Snap shot of electrophoregram of forward strand of SDM9.	97
3.21	Snap shot of electrophoregram of reverse strand of SDM9.	97
3.22	Snap shot of electrophoregram of reverse strand of SDM15a.	97
3.23	Snap shot of electrophoregram of reverse strand of SDM15b.	98
3.24	Snap shot of electrophoregrams of reverse strands of 15a and 15b.	98

3.25	Hybridization on Prototype WDDM	99
3.26	Magnified view of Printed WDDM	100
3.27	Magnified view of WDDM hybridized with WT sample	102
3.28	Two arrays with different post-hybridization wash temperatures	104
3.29	Comparison of spot intensities across two locations on array	105
3.30 A & B	Spot Intensity Dispersion at A (42°C) and B (52°C) for all PM probes	107
3.31	Scan of HWDDM. Sample: SDM4, Hybridization temp: 42°C	108
3.32	Scan of HWDDM. Sample: SDM7, Hybridization temp: 42°C	108
3.33	Scan of HWDDM. Sample: SDM8, Hybridization temp: 42°C	109
3.34	Scan of HWDDM. Sample: SDM9, Hybridization temp: 42°C	109
3.35	Scan of HWDDM. Sample: SDM15, Hybridization temp: 42°C	110
3.36	Correlation between Spot intensities across two locations on the Array	111
3.37A, B,C	Scans of Hybridized arrays of P9, 9-1, 9-2	117
3.37D	Scans of Hybridized arrays of P9, 9-1, 9-3	118
3.38	Hybridization signals from Probe P_29_10 in Patient P-9 and relatives	118
3.39	Probable genotype of Patient Family	119
3.40	Snap shot of eletrophoregram of reverse strand of exon10 of P9	119
3.41A	Snap shot of eletrophoregram of forward strand of exon 13 of P12 covering probes P_34_13 & P_36_13	120
3.41B	Snap shot of eletrophoregram of forward strand of exon 13 of P12 covering probe P_37_13	120
3.42 A	Snap shot of eletrophoregram of forward strand of exon 18 of P12 covering probes P_65_18 & P_67_18	121
3.42 B	P-12 reverse strand. Snapshot covering probes P_65_18 & P_67_18	121
3.42 C	P-12 reverse strand. Snapshot covering probes P_68_18	121

LIST OF TABLES

Table No.	Topics	Page No.
I	List of Cuproenzymes and their function	27
II	Exons and their relation to functional domains of the ATP7B	41
III	List of mutations addressed in WDDM	43-45
IVA	List of PCR primers for 16 amplicons	55
IVB	Nested Primers for 2 nd round of PCR	55
V	List of WDDM Amplicons after 2 nd round of PCR	56
VI	Code numbers for WD indexed patients and their relatives	57
VII	Primers for Site Directed Mutagenesis (SDM)	58
VIIIA	Summary of SetA PCR products from all gDNA samples	89
VIIIB	Summary of SetB PCR products from all gDNA samples	90
IX	SDM fragments for simulating WD	91
X	Kinetics of SDM samples	93
XI	Sample description for hybridization on WDDM	101
XII	Summary of hybridizations carried out for the evaluation of arrays	103
XIII	Pearson Moment Correlation Coefficient	105
XIV	Computed DS scores for all the mutations	112 to 113
XV	One probe of exon 13 addressing three mutations	114
XVI	PM and MM probe pairs & DS scores from the simulated mutant samples (SDM) and WT	115
XVII	Dixon's Q-test on DS values	116



Homi Bhabha National Institute

Ph.D. PROGRAMME

Name of the student:	Manjula Mathur
Name of the Constituent Institution:	Bhabha Atomic Research Centre
Enrolment No. :	LIFE01200804006
Title of the Thesis:	DNA Diagnostics for Wilson Disease
Board of Studies:	Life Sciences

SYNOPSIS

Preamble:

In the post genomic era with availability of molecular sequences of human genome it is apt to use the DNA sequence information for clinical diagnostics. Thought of knowledge application for human welfare and availability of necessary tools at BARC, were put together to develop DNA Microarray based Diagnostics for Wilson Disease (WD). Phenotypic complexities of this inherited disorder make clinical diagnosis difficult. Whereas, if correctly diagnosed early it is a treatable condition and patients could be managed to lead normal life.

General Introduction and Literature Review

A paradigm shift in understanding genetic diseases and their complexities took place after completion of the Human Genome Project (HGP). Newer knowledge and development of new technologies helped to design new strategies for better prognosis, diagnosis and therapy for genetic diseases. Knowledge of human monogenic diseases that follow pattern of Mendelian inheritance (documented by Sir Victor A. McKusick) catalysed this post genome era of human molecular genetics research. Databases like OMIM (www.omim.org) which is Online version of Mendelian Inheritance in Man (MIM) at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, are being used for Knowledge Discovery Database (KDD)¹ by bioinformatics researchers for biological insights of genetic diseases. Wilson disease (WD) is one out of nearly 14000 known monogenic diseases documented in OMIM. This is an autosomal recessive disorder present from birth but clinically manifested at the median age of 12 to 23 years. WD is caused by accumulation of mutations in the gene *Atp7b* coding³ for WND (Wilson Disease) protein thereby affecting its normal function⁴. The carrier frequency is estimated to be 1:90² increasing chances of genetic burden in society.

The symptoms are recognizable after copper accumulates to toxic levels in liver, brain and kidney gradually leading to cirrhosis and early fatality. *Atp7b* is coded by 21 exons, covering 80kb of genomic DNA and maps to locus 13q14.3 on chromosome 13. ATP7B protein is evolutionarily conserved, membrane embedded P1-B type ATPase responsible for accurately regulating copper homeostasis for normal development and cellular metabolism. ATP7B has 6 copper binding domains at the N-terminus and eight trans-membrane

domains which have conserved phosphorylation, transduction, ATP-binding, and actuator sites executing intracellular copper trafficking in hepatocytes. Like any other protein WND is functional through its correct conformation by interacting with other proteins and specific cofactor binding domains. WD patients do not have an active WND, which is the cause of failure in hepatic copper excretion. As copper is required for numerous cellular processes, its homeostasis is highly regulated in healthy individuals. Nearly 600 different mutations in *Atp7b* gene that cause dysfunctional ATP7B have been correlated with WD⁵. Different mutations of *Atp7b* influence differently and hence the clinical manifestation of the disease is complex and varied. That is perhaps the reason that age of onset of symptoms is reported to vary from 3 to 70 years, making its diagnosis a miss. If diagnosed early for the presence of WD mutations, indexed patients can be managed by medicines and diet, reducing morbidity and mortality. Gupta et. al.⁶ have done extensive studies of WD in India using advanced molecular biology techniques. Fourteen new WD mutations were characterized in Western Indians by Annu et. al.⁷. Due to large number of causative mutations, a micro-array based detection approach becomes most desirable for carrier frequency screening as well as causative mutation in patients. The thesis describes the science and technical issues related to development of low density Wilson Disease DNA Microarrays (WDDMs) for correct and early identification of WD mutations among suspected individuals and carrier status among their family members. WDDMs designed in this project initially covered 62 mutations and 14 newly reported from Western India⁷ were added later in the updated design of 4th generation of WDDMs.

DNA micro-arrays, were first introduced by Pat Brown⁸ and his

colleagues at Stanford University in mid nineties. Originally cDNA microarrays were developed for high throughput measurements of expression levels of genes in cancerous cells. Oligonucleotide based microarrays evolved next, and these were designed for large scale screening of mutations in nucleotide sequences. The microarray technique takes advantage of the fundamental tool in molecular genetics i.e. nucleic acid hybridization which is DNA base pairing between individual single-stranded nucleic acid molecules to form double-stranded molecules. It involves mixing of nucleic acid molecules from two sources, a probe which consists of a homogenous population of probe molecules (chemically synthesized oligonucleotides) and a target which consists of heterogeneous population of nucleic acid molecules (amplicons of genes). Hundred percent hybridization commands high degree of base complementarity achieved by providing optimum annealing temperature, percent GC in sequence and concentration of salt in wash buffer. The rationale of the hybridization assay is to examine the presence or absence of *de novo* mutations in the sample within the region of known probe sequences. Digoxigenin (DIG) labeled target nucleic acid hybridizing to perfect match probes (labeled probe-target heteroduplex) are detected by anti-DIG antibody.

Bioinformatics and molecular analysis for Detection of WD Mutations:

Genomic DNA sequence (NG_008806) for *Atp7B* (human), its messenger RNA sequence (NM_000053) & amino acid sequence (P35670) were downloaded from NCBI website. Wilson Disease Mutation Database (<http://www.wilsondisease.med.ualberta.ca/database.asp>) was used to shortlist the

mutations covered on WDDMs. Based on the disease related functional mutations in *Atp7b* gene and their occurrence in Indian population^{6,7} strategy for WDDM design was undertaken. For designing PCR primers, computer program Primerquest (<http://www.idtdna.com/Scitools/Applications/Primerquest>) was used and suitable primers were selected in the intronic regions, so as to amplify complete exons of *Atp7b* gene. Sixteen DNA segments (amplicons) were shortlisted after excluding exons without any reported causal mutations. OligoArray software, was used for designing probes (25-32 single stranded DNA bases) keeping mutation in center and GC content to 50%.

Materials and Methods:

Approval of BARC Hospital Medical Ethics Committee was obtained vide Ref. No. BARCHMEC /10, dated 22nd December 2010 for using human material for the work. Peripheral blood sample was collected in EDTA from unaffected individuals, WD indexed patients and their relatives. Genomic DNA isolation was done following non-enzymatic method. gDNA preparations were checked for quality and quantity by optical density (OD) measurements at 260nm (for DNA) and 280nm (for protein) wavelengths and stored. Sixteen amplicons of *Atp7b* were amplified from each gDNA sample by 16 different Polymerase Chain Reactions (PCRs) using exon specific primers. Equi-molar proportions of 16 amplicons were labeled using DIG High Prime Labeling Kit (Roche, Germany) before hybridization on WDDMs. To assess sensitivity of WDDMs, six types of mutations were also generated by independent site directed mutagenesis experiments. These were used for training of WDDMs for detection of simulated mutations. Each mutated amplicon was pooled with remaining 15 wild type amplicons for hybridization on WDDMs and intensities at perfect match and mismatch positions were determined. The type of mutations, exon numbers and SDMs generated are listed in the Table 1.

Table 1. Simulated WD mutations

Exon	Size in Bps	Type of mutation	Name of mutated amplicon
4	291	<u>DBI</u>	SDM4
7	288	DBD	SDM7
8	482	<u>SBI</u>	SDM8
9	297	SBD	SDM9
15a	393	(G>A), SPM1	SDM15a
15b	393	(T>C), SPM2	SDM15b

WDDMs were prepared by printing each probe, in triplicate, on epoxy coated slides (Corning Cat No: 40044) in 3XSSC buffer at ~ 50% relative humidity with the help of an in-house developed precision micro-arrayer. Post printing, the arrays were stored under desiccation at room temperature till further use. Multi parameter hybridization experiments were conducted on WDDMs to arrive at reproducible spot intensities at perfect match and mismatch probes respectively. Hybridization signals were visualized by using Alkaline phosphatase conjugated Anti-DIG anti body and NBT/BCIP colorimetric substrate measurements. Slides were scanned on a flat bed scanner at 600 dpi resolution and spot intensities were quantified by Image J software (<http://rsb.info.nih.gov/ij/>). Intensities of spots on HWDDMs (Hybridised WDDMs) were quantified at perfect match (IP) and mismatch (IM) probes positions after background subtractions. Average **differentiation score** (DS) for all the probes were calculated from corrected IP and IM values, using the formula : $DS = (IP - IM) / (IP + IM)$. Ideally, in the absence of mutation, spot intensity at perfect match

should be higher than on the mismatched spot ($IP > IM$). In case of hybridization with mutant sample, intensity values of IP and IM at the corresponding probe position would be reversed i.e. ($IP < IM$) giving negative value of DS.

Results and Discussion

Sixty two mutations specific to Indian population were studied on the low density WD microarray. Figure 1 shows a magnified view of WDDMs with three arrays were printed per glass slide. Out of 204 probe triplets on each WDDM, 84 were WT probes, 93 were mutant probes, 20 exon specific control probes, 5 blanks (only buffer as negative control) and 2 were DIG labeled (as positive control).

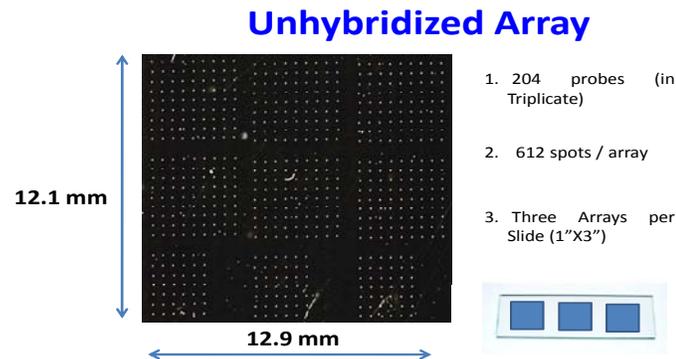


Figure 1. Magnified view of a typical WDDM

The Figure 2 shows an example of a scanned image after the array was hybridized to amplicons from gDNA of unaffected individuals. The picture depicted here is a magnified portion of 13.5 x 13.5 mm region on the slide. The first and last triple spots are the DIG labeled oligos used as positive controls for normalization of intensity values. Since nucleic acid hybridization efficiency is dependent on probe length and its GC% , spot intensities of different probe pairs varied

substantially between arrays. However, as expected most of the PM probes yielded higher intensity than MM probes for respective amplicons used for hybridization.

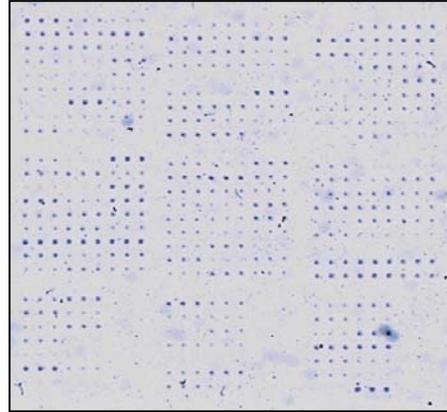


Figure 2. A magnified view of Hybridized WDDM (HWDDM)

For digitalization of data from HWDDMs all the spot intensity data were quantitatively acquired using Image J Software and saved in Microsoft Excel files after subtracting background signals. Inter and intra array analysis of data was done to statistically assess reproducibility of intensity values. For each probe pair Differentiability Score (DS) was calculated from average intensity values and tabulated for further analysis of probe pairs. Spot intensities within each array were compared and hybridization across the slide surface was found to be uniform. Specificity of mutation detection ability of probes was also studied by hybridizing same pool of amplicons on two HWDDMs with two different post hybridization washing temperatures, one at 42°C and other at 52°C temperatures. Pearson correlation coefficient was determined between respective spot intensities obtained on these two HWDDMs at PM positions of 84 probes for same pool of samples. Table 2 shows coefficient values (>0.7) calculated from five such different DNA samples showed good correlation between the two data sets obtained at two temperatures. WT refers to pool of all wild type amplicons, SDM refers to pool of

one representative of a mutant amplicon (numerical indicating the specific exon) with rest of fifteen WT amplicons.

Table 2. Intensity correlation coefficient between arrays at two temperatures

Sample	Correlation coefficient R^2 of 84 intensity values at 42°C and 52° C
WT	0.73
SDM4	0.92
SDM8	0.70
SDM9	0.82
SDM15	0.81

Probe wise intensity variation at PM positions was studied by normalizing IP values with respect to IP_{max} . Figure 3 is a graphical representation of probe wise dispersion of intensity values obtained from 5 hybridization experiments done at 52° C. The error bars represent one standard deviation from the mean intensity values.

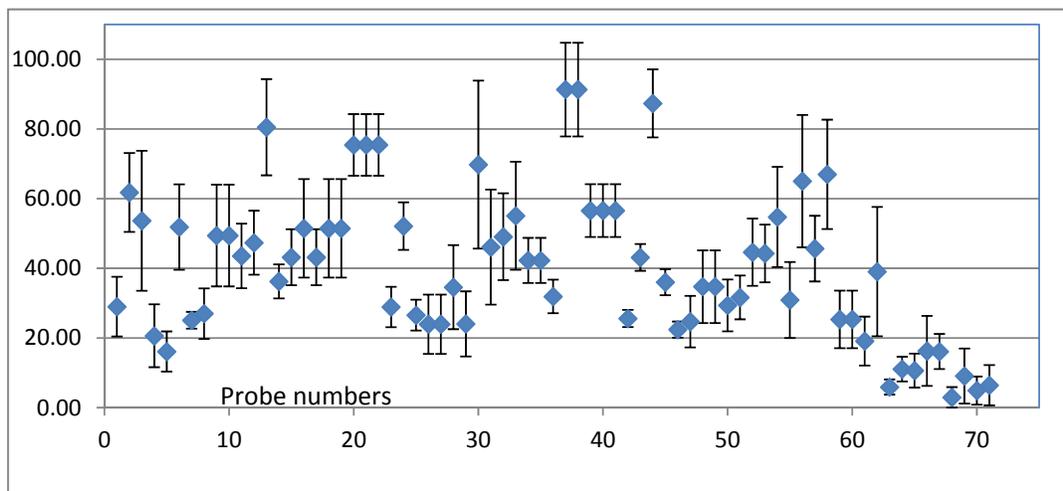


Figure 3: Spread of normalised intensities on Y-axis for 72 probes (X-axis)

These results indicated that intensity data from HWDDMs was reproducible although there existed a very wide variation in intensities across 72 probes studied on these WDDMs. DNA sequencing of SDM amplicons confirmed PCR generated mutation. Shown below in Table3, a HWDDM part depicting the spots obtained from hybridization of wild-type and one of the SDM sample. Average DS value for this particular probe pair is mentioned in last row. As conceptually required SDM yielded higher spot intensity at MM probe than on PM probe.

Table 3. Inverse intensity at PM and MM of WT and SDM probes

Hybridized with	Wild-type sample		SDM Sample	
Spot identity	PM probe	MM probe	PM probe	MM probe
Spot Image				
DS Score	0.49 ± 0.13		- 0.36 *	

Seven WD indexed patient and 6 family members of two patients were also added in the study. From visual inspection and spot intensity values a few HWDDMs of patient samples indicated presence of mutations which were also verified by DNA sequencing of the respective amplicons. After analyzing probe pairs with respect to their DS scores, those having very low values of DS were discarded and few probes were redesigned for better DS values. DS for each probe pair varied from probe to probe, conceivably as a result of altered sequence composition and probe length. Considering all the probes that yielded reproducible DS values (DS > 2 s.d.) the sensitivity of our WDDM array was better than 80%.

CONCLUSION AND FUTURE DIRECTION

Out of all the probe pairs spotted on WDDMs, nearly 80% were found to have average spot intensity differentiability (among perfect match and mismatch) required for minimal identification of a mutated allele from wild type alleles in the PCR amplified DNA samples. This study demonstrates that designed WDDMs and the protocols followed can be used for confirmation of the disease in suspected patients as a first pass screening tool. In comparison to widely available sequencing technology, low density microarray based technique is more efficient and cost-effective for mutation detection, in populations that may have large carrier frequency (1:90). Thus it is concluded that; human genome sequence data may be used for designing population specific microarrays for detection of prevalent mutations and prevalent genetic disorders by genetic counselors. This could prevent the burden of disease to future generations and thus reduce human suffering.

REFERENCES:

1. Hoan Nguyen, Tien-Dao Luu, Olivier Poch and Julie D. Thompson. Knowledge Discovery in Variant Database Using Inductive Logic Programming. *Bioinformatics and Biology Insights* 2013;7 119-131 doi: 10.4137/BBI.S11184
2. Schilsky M L . Wilson disease: genetic basis of copper toxicity and natural history; *Semin Liver Dis* (1996) 16 pp83–95.
3. Bull, P.C. et al. The WD gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nature Genetics* (1993) Vol. 5, pp 327-337

4. P de Bie, P Muller, C Wijmenga, L W J Klomp. Review: Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes. *Journal of Medical Genetics* (2007) Vol.44: 673-688
5. Kenny S.M., Cox D.W. Sequence variation database for the Wilson disease copper transporter ATP7B. *Hum. Mutat.* (2007) 28 (12) :1171-7
6. Arnab Gupta , Ishita Chattopadhyay, Sumit Dey, Poonam Nasipuri, Shyamal K. Das, Prasanta K. Gangopadhyay & Kunal Ray. 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 (8) : 1023-33; doi 10.1007/s10571-007-9192-7
7. Annu Aggarwal, Gursimran Chandhok, Theodor Todorov, Saloni Parekh, Sharada Tilve, Andree Zibert, Mohit Bhatt and Hartmut H.-J. Schmidt. Wilson Disease Mutation Pattern with Genotype-Phenotype-Phenotype Correlations from Western India : Confirmation of p.C271* as Common India Mutation and Identification of 14 Novel Mutations. *Annals of human genetics* (2013) Vol 77 (4) : 299-307; doi: 10.1111/ahg.1224.
8. Joseph DeRisi, Lolita Penlan, Patrick O. Brown, Michael L. Bittner, Paul S. Meltzer, Michael Ray, Yidong Chen, Yan A. Su and Jeffrey M. Trent. Use of cDNA microarray to analyse gene expression patterns in human cancer. *Nature Genetics*, Dec. 14, 1996, vol. 14, pp. 457-460.

“Only those with relaxed and loving mind find themselves in harmony with everyone.”

CHAPTER I

INTRODUCTION

1.1 HISTORY & CURRENT SCENARIO OF WILSON DISEASE

Dr. Samuel Alexander Kinnier Wilson (1878-1937) was a British neurologist who first described in 1912¹ that copper overload was the reason of the clinical condition of hepatolenticular degeneration. This autosomal recessive inheritable condition was later named after the discoverer as Wilson Disease (WD) and has been analysed extensively to understand the mechanism of the disease and its management. Year 1993 was golden year for WD when a breakthrough was achieved for molecular understanding of the disease. The *ATP7B* gene was cloned, sequenced and characterized^{2,3,4} to be coding for a putative copper transporter protein ATP7B similar to Menkes Disease protein ATP7A. The *ATP7B* gene is located on chromosome 13 at its long arm at locus 13q14.3 and is coded by 22 exons spread over 80kb of genomic DNA. A web site (<http://www.wilsondisease.med.ualberta.ca/references.asp>)^{wr1} dealing with DNA sequence mutations in WD gene *ATP7B* named “Wilson Disease Mutation Database” has 140 database references published between 1993 to the year 2009. This thesis is neither a study of copper toxicity due to environmental factors nor a study on nutritional deficiencies, but an attempt to develop a DNA based method to identify among us a few unfortunate individuals, who are born with an error of metabolism due to mutations in *ATP7B* gene. If these mutations can be detected in the individual he or she can be advised on his/her diet and medicines to live a normal healthy life like any other person without inborn error of metabolism. Research work described in this thesis resulted in the indigenous development of **Wilson Disease DNA Microarray (WDDM)**.



Figure 1.1 Dr. Samuel Alexander Kinnier Wilson (1878-1937)

Copper is a transition metal element, which has a natural property of getting easily converted between two redox states namely oxidized Cu (II) and reduced Cu (I). Because of this unique property, biological systems have made Cu metal to get manifested as an important catalytic co-factor for a variety of metabolic reactions in biological systems⁵. Copper, like vitamins, is required in tiny amounts for us to remain healthy and we get it from many foods. In healthy individuals the body gets rid of any excess copper while people with WD cannot get rid of excess copper and it builds up in the body, mainly in the liver, the brain, the cornea (the layer at the front of our eye) and kidneys. Abnormal levels of copper overload affects the individuals in varied ways depending upon the age, genetics and diet causing complex clinical hepato and/or neurotoxic manifestations. Too much copper in the liver cells (the hepatocytes) leads to liver damage while damage to brain tissue is known to occur in an area called the lenticular nucleus. Hence, Wilson's disease is sometimes also called hepatolenticular degeneration. There may be more physiological effects related to evolutionarily conserved ATPase proteins (Gupta and Lutsenko, 2012)⁶ that await conceptual discoveries by bioinformatics analysis (Chang & Mitchell, 2013)⁷ of metal transporter proteins using InterPro Database

<http://www.ebi.ac.uk/interpro>^{wr2}.

1.2 GENETICS, EPIDEMIOLOGY AND PHENOTYPE OF WD

Wilson disease (WD) is a treatable monogenic autosomal disorder present from birth but clinically manifested after copper overload at the median age of 6 to 23 years. Above certain levels of intracellular copper accumulation, it could even enter the nucleus and damage DNA causing unpredictable damage to the cellular functions and total chaos to the system. It is fatal unless detected early and treated before serious illness from copper poisoning develops. As of now, more than 30 million people worldwide are estimated to be affected by the disease. In affected patients, copper accumulates in liver and brain gradually leading to cirrhosis, lack of coordination, personality changes and early fatality (Huster, 2010)⁸. Aggarwal et al in 2009 developed a very novel Global Assessment Score (GAS) for WD⁹, which is a quantitative measure to monitor clinical parameters during treatment of WD patients. This system has scaling in two tiers that grades the multisystemic manifestations of the disease. Tier 1 scores the global disability in four domains: Liver, Cognition and Behavior, Motor, and Osseomuscular. Tier 2 is multidimensional scale for a fine grained evaluation of the neurological dysfunction. Scores of both tiers scale the disease burden on each system of the patient and the GAS for WD is also sensitive to clinical changes.

ATP7B, expresses in mitochondria of hepatocytes (Yamaguchi et al 1993)¹⁰. WD occurs at a frequency of 1 in 30,000 whereas its carrier frequency is estimated to be 1:90 (Schilsky, 1996)¹¹. The complexities associated with its diagnosis (Huster, 2010)⁸ and the overlap with other symptoms often lead to incorrect diagnosis for this treatable disorder. If diagnosed early and correctly, WD patients can be managed by medicines reducing morbidity and mortality. Because of

the complex phenotype of WD, it is highly advisable to do genotype diagnostics i.e. DNA diagnostics for WD, and hence it was decided to develop Wilson Disease DNA Microarray (WDDM). This thesis describes the scientific rationale and experiments conducted for development of WDDM.

1.3 REGULATION OF HOMEOSTASIS OF TRANSITION METALS AND SYSTEMS APPROACH TO UNDERSTAND WILSON DISEASE

Transition metals are frequently used as cofactors for enzymes and oxygen-carrying proteins that take advantage of their propensity to gain and lose single electrons (Wenjing et al 2003)¹². Metals are particularly important in mitochondria, where they play essential roles in the production of ATP and detoxification of reactive oxygen species. At the same time, transition metals (Ferrum/Iron (Fe), Cuprum/Copper (Cu) and Zinc(Zn)) can promote the formation of harmful radicals, necessitating meticulous control of metal concentration and subcellular compartmentalization. Mechanisms of fine regulation and homeostasis of trace elements is a subject of intense research all over the world and current understanding of Fe and Cu in mammalian mitochondrial biology and human diseases associated with aberrations in mitochondrial metal homeostasis are discussed in review article (Osredkar&Sustar,2011)¹³. Fig.1.2 is schematic depiction of copper distribution in a mammalian cell.

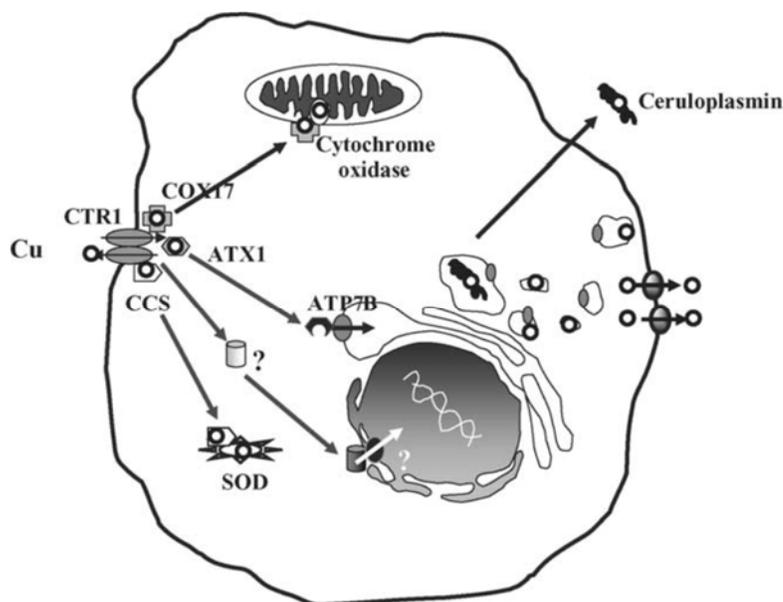


Figure 1.2 A mammalian cell and copper trace element. Adapted from (Safaei& Howell, 2005)¹⁴

Copper, taken up via hCTR1 is transferred to chaperones ATX1, CCS and COX17, which transfer the metal to ATP7A/ATP7B, Cu–Zn Superoxide Dismutase and Cytochrome C oxidase, respectively. Copper binding to ATP7A and ATP7B induces subcellular trafficking of vesicles that contain the two proteins from the trans-Golgi network (TGN) to more peripheral locations (Safaei& Howell, 2005)¹⁴. These proteins transfer Copper to cuproenzymes, such as tyrosinase and ceruloplasmin. Also the potential toxicity of Cu (I) is checked by a number of chaperones and buffering molecules that bind Cu (I). The central feature of the Copper homeostasis system is the presence of unique protein domains, rich in cysteine, methionine or histidine, called metal binding sequence (MBS). MBS bind Cu (I) in a protective pocket and hand it to the next protein through an intimate protein-protein interaction such that Cu atom is virtually never free in the cell. Copper binds to chaperones and transporters through labile electrostatic associations, and is easily trans-chelated between MBS-containing molecules. The

two Copper efflux proteins, ATP7A and ATP7B, are homologous both in structure (54% amino acid similarity) and in function (Pertrukhin et al, 1994)¹⁵. ATP7B is expressed in liver and kidney and to a lesser extent in brain of normal individuals, consistent with excessive Copper accumulation observed in these tissues when ATP7B function is lost in patients with Wilson's disease (Tanzi et al, 1993)⁴. ATP7A is expressed in the intestinal epithelium as well as most other tissues other than liver, and the pathology of Menkes Disease, caused by mutations in the gene *ATP7A*, reflects inadequate mobilization of Copper from a number of tissues. Both proteins have eight membrane spanning domains of which the sixth is conserved and forms the channel through which Cu atom moves (Bissig et al, 2001)¹⁶. Both proteins also have six MBS repeats of approximately 70 amino acids each at the N-terminal cytosolic side. Each MBS repeat has a core sequence of GMTCXXCIE that is similar to motifs found in mercury binding proteins and cadmium ATPase of bacteria (Silver et al, 1993)¹⁷. ATP7B and ATP7A are P1- (CPx)-type ATPases, and like other P-type ATPases, function as monomers (Soliozet al, 1996)¹⁸. An important feature of both ATP7B and ATP7A proteins is their ability to undergo constitutive and Cu-stimulated trafficking. Copper triggers the movement from the trans-Golgi network (TGN) to more peripheral membrane compartments. In case of ATP7A, the plasma membrane Cu-induced trafficking of the vesicles containing ATP7A and ATP7B proteins appears to be highly regulated by intramolecular phosphorylation (Lutensko & Petris, 2003)¹⁹ and trafficking molecules such as Rab5 and Rab7 GTPases (Pascale et al, 2003)²⁰.

Figure 1.3 is a model to represent proteins involved in intracellular trafficking of trace metals Fe, Cu and Zn focusing the mitochondria organelle of a hepatocyte.

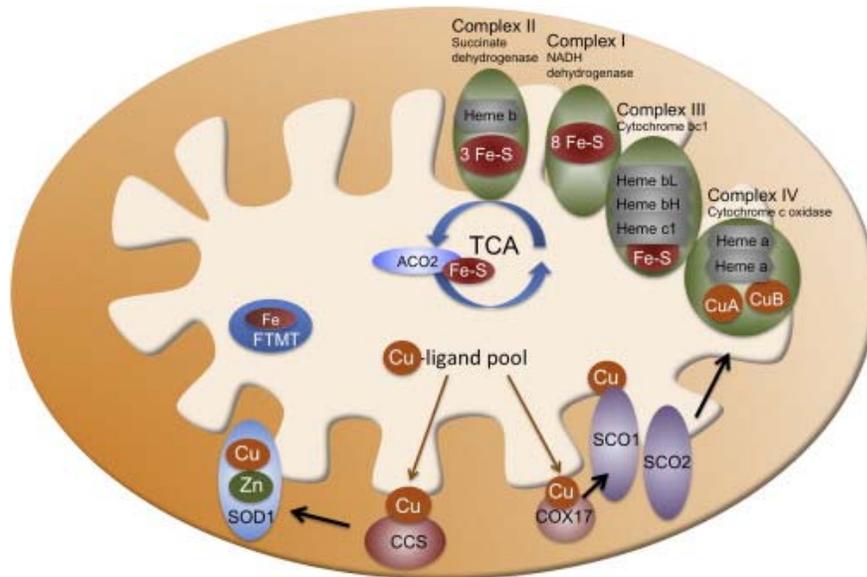


Figure 1.3 Proteins involved in homeostasis of Fe, Cu and Zn within a hepatocyte. Adapted from (Wenjing et al, 2013)¹²

Copper is critically important for developing mammals, but because of the potential toxicity of copper ions, the supply of copper to the fetus and neonate is carefully regulated by homeostatic mechanisms. Copper uptake into cells is mediated by Ctr1, and the essentiality of copper in development of the mouse is demonstrated by the early death of Ctr1 knock-out mice (Lee et al, 2001)²¹. There are many copper chaperones that function in the intracellular distribution of copper. The Cu-ATPases are important regulators of copper efflux from cells as well as supplying copper to secreted cuproenzymes. Menkes disease in humans and the *Mottled* mice mutants, are due to mutations of the Menkes gene, ATP7A. Wilson disease and the toxic milk mice are copper toxicosis disorders, due to mutations of the other Cu-ATPase gene, ATP7B. The proteins, ATP7A and ATP7B, play many roles in the physiological regulation of copper and are particularly important in the supply and distribution of copper to the fetus and developing neonate. Copper homeostasis involves changes in the intracellular location of the Cu-ATPases. Elevation in intracellular copper

induces the movement of these proteins from the transGolgi network to the plasma membrane (ATP7A) or endosomal-like vesicles (ATP7B). It was found that diverse missense mutations prevent the relocalization of the Cu-ATPases in response to copper, and other mutations alter the intracellular localization of the protein. The effects of mutations at a cellular level can explain many features of the genetic copper disorders very well reviewed by Prof. Mercer, J (2001)²². Professor Mercer's group at the School of Biological & Chemical Sciences, Deakin University, Australia is working on effective therapies for WD including gene therapy.

Within cells all proteins are functional through their correct conformation by interacting with other proteins and specific cofactor binding domains. ATP7B is a very large protein in the category of P-type ATPase and hence is being studied by reductionist approach by cloning in parts using molecular biology techniques and study Wild Type (WT) and mutated protein conformations. A recent multidisciplinary research by Braiterman et al (2014)²³ has focused on mutation S653 in the TM1 domain of ATP7B and established that mutation perturbs long-range inter domain interactions mediated by trans membrane segments TM1/TM2, suggesting a new functional role for this region. These studies at molecular sequence level of Wilson Disease Protein ATP7B also called WND may facilitate targeted therapy for patients.

1.4 PRIMARY, SECONDRY AND TERTIARY STRUCTURE OF ATP7B PROTEIN AND ITS FUNCTION IN NORMAL PHYSIOLOGY

Primary sequence of WND protein ATP7B, consisting of 1465 amino acids, is obtained by translating the cDNA sequence of *ATP7B*. By enzymatic activities classification, it is characterized as EC 3.6.3.4 with a computed molecular weight of

MW: 157334 (indicated in Figure 1.4A). The sequence data depicted in Figure 1.4A is taken from UniProtKB (Knowledge Base). The amino acid sequence motifs found in the amino acid sequence of ATP7B are presented as linear graphics in Figure 1.4B. PEPTIDE, PROSITE, TIGRFAM, PRINTS and Pfam indicated in Figure 1.4B are amino acid motif databases available for functional genomics research. Motifs for heavy metal transport are found in five metal binding domains at the N-terminal of ATP7B, as in the SUPERFAMILY and Pfam databases. Location of trans-membrane helices (of PEPTIDE database) on ATP7B shown in the first line of Figure 1.4B indicates that there are eight trans-membrane domains in this protein. The secondary structure of ATP7B derived from primary structure is depicted in Figures 1.5A and B (Ruslar Tsivkovskii et al, 2003)²⁴ highlighting different aspects of structure, function, and the alignment of its ATP binding domain with P1-B type ATPases of other organisms. Five MBS domains at the N-terminal of the protein are followed by trans-membrane regions and other conserved domains. Many of these domains are evolutionarily conserved in taxon specific manner in prokaryotes, eukaryotes and higher animals. It may be noted that amino acid residues AHSER...QT at position 1114-1143 (Figure 1.4A) in human ATP7B are not present in lower eukaryotes and bacteria. Splicing of mRNA, joining of exons and translation processes, sometimes introduce errors in amino acid sequence of the coded protein ATP7B that may also manifest as WD.

```

1 MPEQERQITA REGASRKILS KLSLPTRAW E PAMKKSFAFD NVGYEGGLDG LGPSSQVATS
61 TVRILGMTCC SCVKSIEDRI SNLKGII SMK VSLEQGSATV KYVPSVVLQ QVCHQIGDMG
121 FEASIAEGKA ASWPSRSLPA QEAVVKLRVE GMTCCSCVSS IEGKVRKLQGV VVRVKVSLSN
181 QEAVITYQPY LIQPEDLRDH VNDMGFEAAI KSKVAPLSLG PIDIERLQST NPKRPLSSAN
241 QNFNNSSETLG HQGSHVVTLQ LRIDGMHCKS CVLNIENENIG QLLGVQSIQV SLENKTAQVK
301 YDPSCSTSPVA LQRAIEALPP GNFKVSLPDG AEGSGTDHRS SSSHSPGSP RNQVQGTCTST
361 TLIAIAGMTC ASCVHSIEGM ISQLEGVQOI SVSLAEGTAT VLYNPSVISP EELRAAIEDM
421 GFEASVVSSES CSTNPLGNHS AGNSMVQTTD GTPTSVQEVA PHTGRLPANH APDILAKSPQ
481 STRAVAPQKC FLQIKGMTCA SCVSNIERNL QKEAGVLSVL VALMAGKAEI KYDPEVIQPL
541 EIAQFIQDLG FEAAVMEDYA GSDGNIELTI TGMTCASCVH NIESKLTRTN GITYASVALA
601 TSKALVKFDP EIIGPRDIK IIEEIGPHAS LAQRNPNAH LDHKMEIKQW KKSFLCLSLVF
661 GIPVMALMIY MLIPSNEPHQ SMVLDHNIIP GLSILNLIFF ILCTFVQLLG GWYFYVQAYK
721 SLRHRSANMD VLIVLATSIA YVYSLVILV AVAEKAERSP VTFDTPPML FVFIALGRWL
781 EHLAKSKTSE ALAKLSLQA TEATVVTLGE DNLIIREEQV PMELVQRGDI KVVPVGGKFP
841 VDGKVLGENT MADESLITGE AMPVTKKPGS TVIARSINAH GSVLIKATHV GNDTTIAQIV
901 KLVEEAQMSK APIQQLADRF SGYFVFPFIII MSTLTLLVWVI VIGFIDFGVV QKYFPNPNKH
961 ISQTEVIRF AFQTSITVLC IACPCSLGLA TPTAVMVG TG VAAQNGILIK GPKPLEMAHK
1021 IKTVMFDKTG TITHGVPRVM RVLLLGDVAT LPLRKVLAVV GTAEASSEHP LGVAVTKYCK
1081 EELGETETLGY CTDFQAVPGC GIGCKVSNVE GILAHSERPL SAPASHLINEA GSLPAEKDAV
1141 PQTFSVLIGN REWLRNGLT ISSDVS DAMT DHEMKGQTAI LVAIDGVLGC MIAIADAVKQ
1201 EAALAVHTLQ SMGVDVVLIT GDNRKRTARAI ATQVGINKVF AEVLPCHKVA KVQELQNKKG
1261 KVAMVGDGVN DSPALAQADM GVAIGTGTDV AIEAADVLLI RNDLLD VVAS IHLKRTVRR
1321 IRINLVLALI YNLVGIPIAA GVFMPIGIVL QPVMGSAAMA ASSVSVLSS LQLKCYKKPD
1381 LERYEAQAHG HMKPLTASQV SVHIGMDDRW RDSPRATPWD QVSYVSQVSL SSLTSDKPSR
1441 HSAADDDGD KWSLLLNGRD EEQYI // MW : 157334 Da

```

Figure 1.4A Amino acid sequence of human ATP7B (ATP7B_HUMAN in UniProtKB)

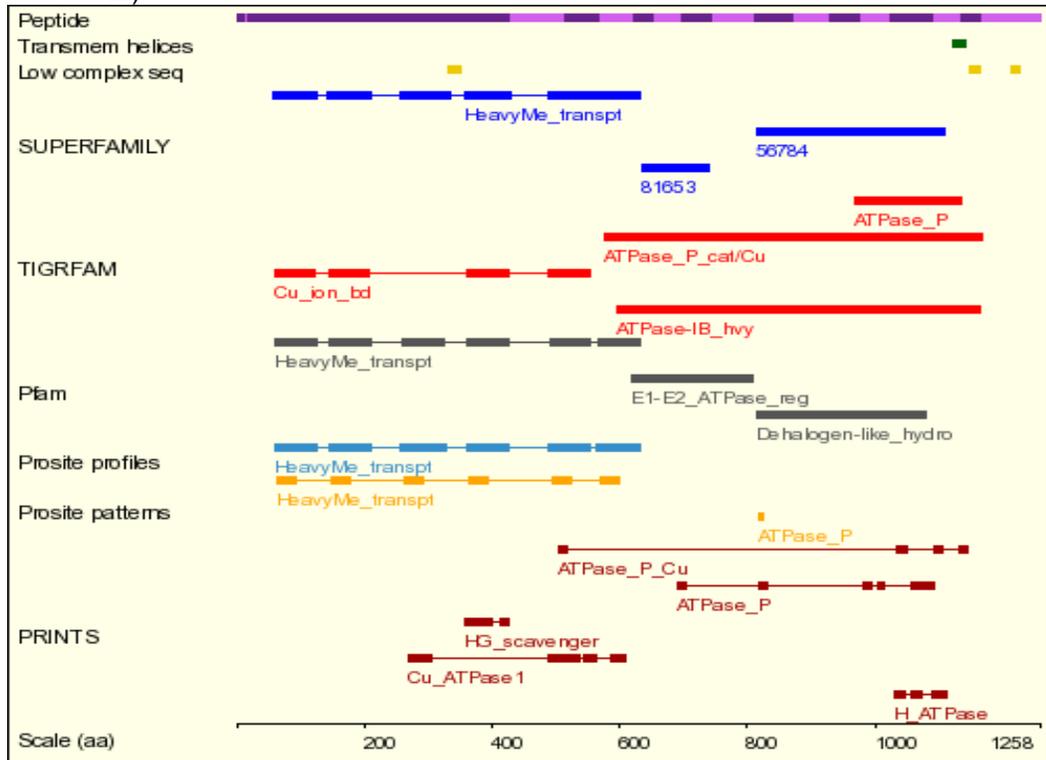


Figure 1.4B Conserved motifs in ATP7B sequence.

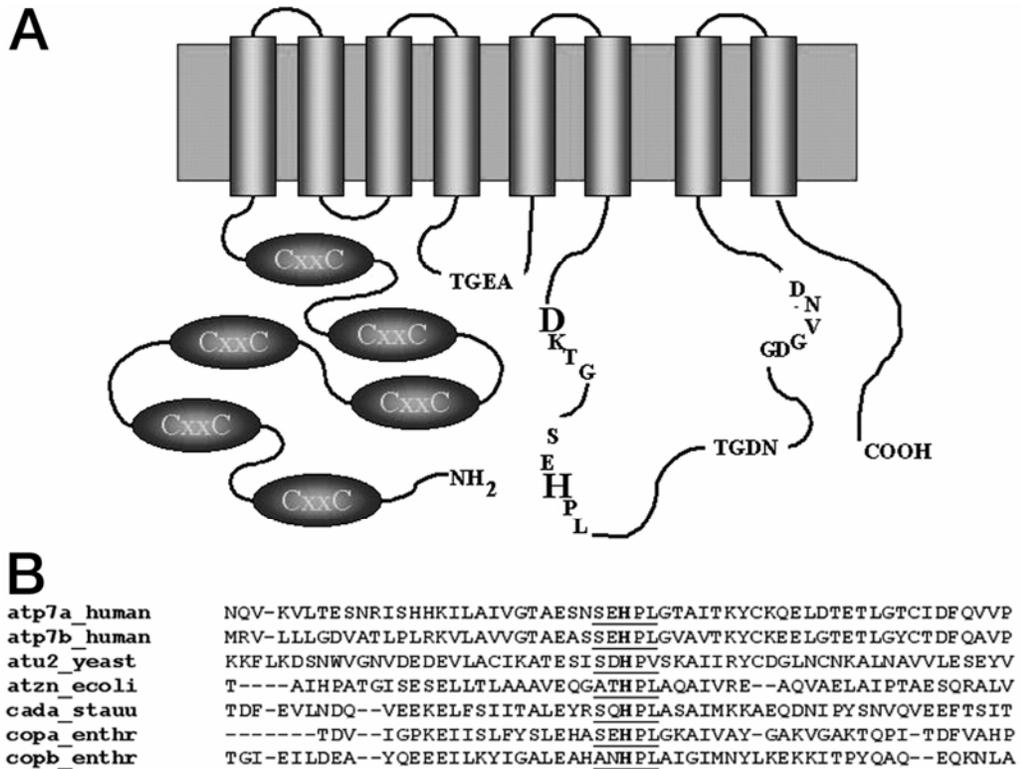


Figure 1.5 Schematic representation of the transmembrane organization of ATP7B²⁴.

A- the letters *TGEA*, *TGDN*, *DKTG*, and *GDGVND* indicate sequence motifs conserved in all P-type ATPases. The **bold letter D** in the *DKTG* motif marks the position of Asp-1027, an acceptor of P_i during catalysis. The *SEHPL* sequence is conserved in all P_1 -type ATPases and contains the invariant His-1069, marked by larger font. The *CXXC* motifs indicate copper-binding sites in the cytosolic copper-binding domain.

B- the alignment of the ATP-binding domain segments of several P_1 -type ATPase. The alignment was generated using ClustalW (www.ebi.ac.uk/clustalw/). The protein data base accession numbers are given in parentheses for the following: atp7a_human, MNKP (Q04656); atp7b_human, WNDP (P35670); atu2_yeast, yeast copper-transporting ATPase CCC2 (P38995); atzn_ecoli, lead-, cadmium-, and zinc-transporting ATPase (P37617); cada_stauu, cadmium-transporting ATPase (P20021); copa_enthr, copper-importing ATPase A from *Enterococcus hirae* (P32113); and copb_enthr, copper-exporting ATPase B from *E. hirae* (P05425). The invariant His is in **bold**; the *SEHPL*-like sequences are *underlined*.

Figure 6A is a hypothetical ribbon model of ATP7B derived from primary sequence. In the past 10 years, there has been significant progress in computational analysis of biological systems and study of regulation of Cu-ATPases has generated interest from clinical as well as fundamental view of research. Since this is a very large protein, it is cloned in part for study of atomic structure as in Protein Data Bank. PDB has four entries 2ARF, 2EW9, 2KMV & 2KIJ. Dr. Svetlana's group at Johns Hopkins University, Washington, USA are using structure and conservation to grade potential deleterious effects of mutations and for mechanistic understanding of WD (Schushan, 2012)²⁵. Effect of the mis-sense substitutions on the protein function can be checked using SIFT (Sorting Intolerant From Tolerant) (<http://sift-dna.org>)^{wr3} program that computes and depicts whether a substituted amino acid could affect protein function. A SIFT score <0.05 for an amino acid substitution is predicted to be deleterious (Ngak-Leng Sim et al, 2012)²⁶.

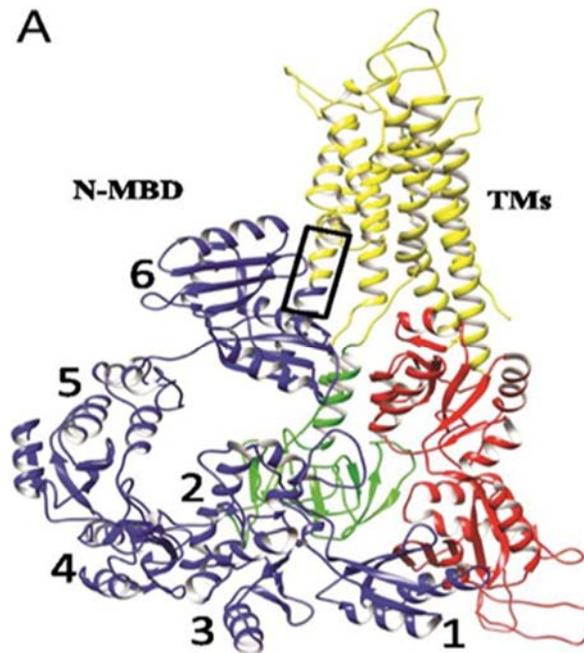


Figure 1.6A Hypothetical ribbon model of 3D structure of ATP7B generated by UCSF Chimera. Adapted from Braiterman et al (2014)²³.

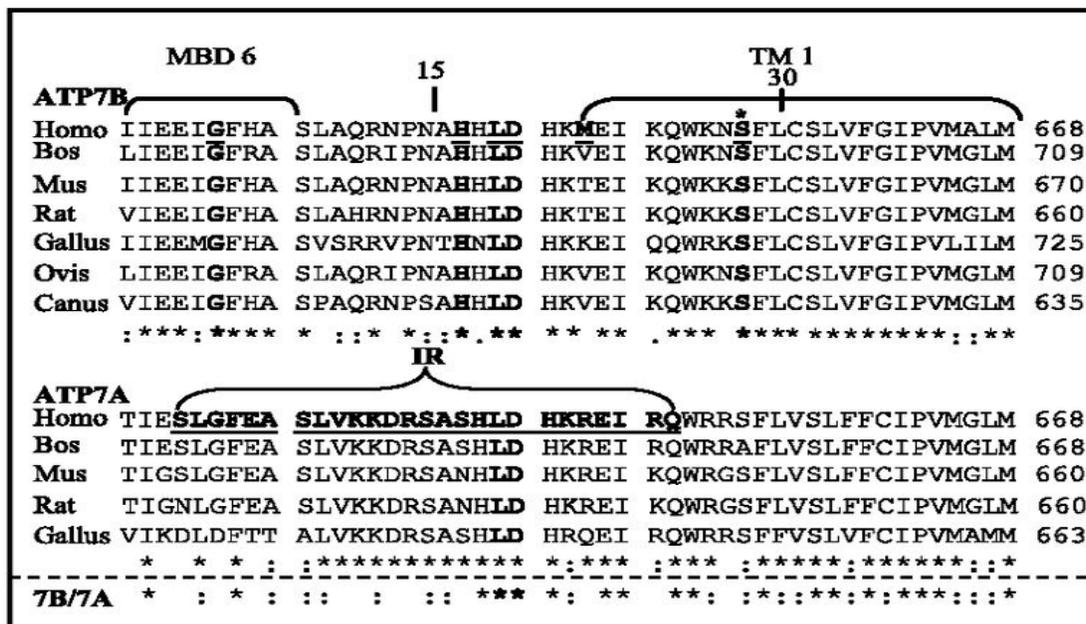


Figure1.6B. Multiple species alignment of the conserved regions, of ATP7A and ATP7B obtained by computer program ClustalW. Aligned residues are marked identical (*), conserved (:), and semiconserved (.)Braiterman et al (2014)²³.

Roles for conserved domains A (Dephosphorylation), P (Phosphorylation domain) and N (ATP binding) in regulation of copper binding by phosphorylation/dephosphorylation, in context of copper homeostasis by ATP7B using current understandings at the biochemical, molecular biology and protein structure level, are shown in Figure1.7 (P. de Bie et al, 2007)²⁷.

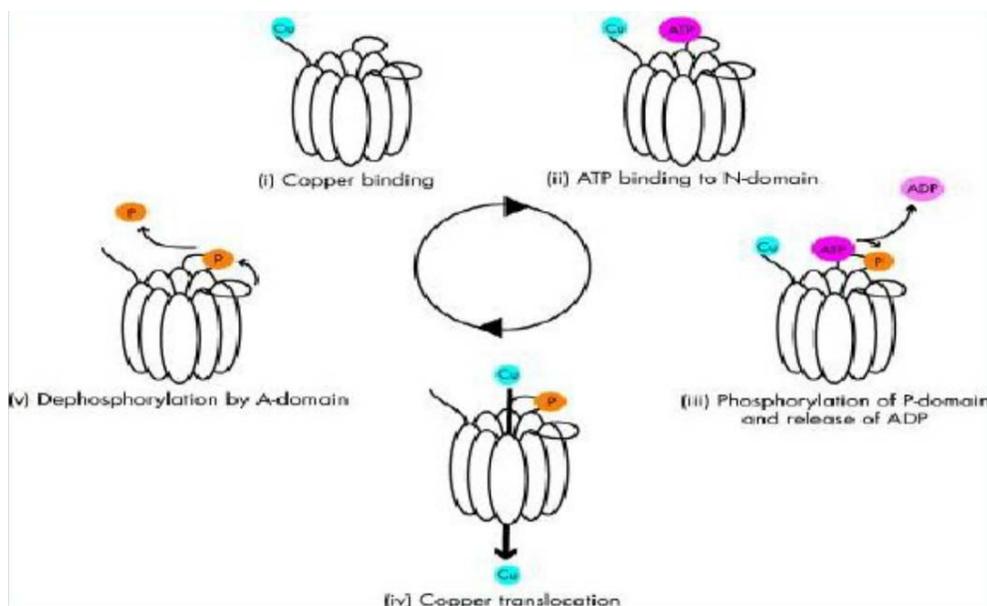


Figure 1.7 A model to understand role of phosphorylation/dephosphorylation of ATP7B in copper homeostasis. Figure adapted from (P de Bie et al, 2007)²⁷

A number of metabolic reactions, which are essential for normal function of brain and nervous system, are catalyzed by cuproenzymes as listed in Table I.

Table I. List of Cuproenzymes and their function

Cupro Enzyme Name	Enzyme Function
Cytochrome C Oxidase	Production of ATP, phospholipid synthesis.
Lysyl Oxidase	Connective tissue formation.
Ceruloplasmin (ferroxidase I)	RBC formation.
Dopamine-b-monoxygenase	Neurotransmitter synthesis.
Monoamine Oxidase	Neurotransmitter biosynthesis.
Tyrosinase	melanin formation.
Super Oxide Dismutase	Antioxidant functions.

Hence many biological functions are compromised due to defect in copper trafficking caused by mutations in *ATP7B*. This is perhaps the cause of complex

manifestation of WD.

1.5 CLINICAL DIAGNOSIS AND ASSAY of WD

In WD mutant individuals, copper begins to accumulate immediately after birth but the disease is manifested after copper overload exceeds its toxic threshold. Excess copper may attack the liver or brain, resulting in hepatitis, psychiatric, or neurologic symptoms. These symptoms usually appear in late adolescence. Patients may have jaundice, abdominal swelling, vomiting of blood, and abdominal pain. They may have tremors and difficulty with walking, talking and swallowing. They may develop all degrees of mental illness including homicidal or suicidal behaviors, depression and aggression. Women may have menstrual irregularities, absent periods, infertility or multiple miscarriages. No matter how the disease begins, it is always fatal if it is not diagnosed and treated. It is important to diagnose WD as early as possible, since severe liver damage can occur before there are any clear signs of the disease. Individuals with WD may falsely appear to be in excellent health (www.wilsondisease.org)^{wr4}. A very complex algorithm adopted by Mayo Medical Laboratories (<http://www.mayomedicallaboratories.com/test-catalog/Clinical+and+Interpretive/83697>)^{wr5} for managing WD diagnosis is given in Figure 1.8A. In our country clinical diagnosis of WD is made by relatively simple tests which can diagnose the disease in both symptomatic patients and people who show no signs of the disease. These tests include:

1. Ophthalmologic slit lamp examination for Kayser-Fleisher rings (Fig. 1.8B)
2. Serum Ceruloplasmin test (<200 mg/L) due to low copper excretion into the secretory pathway resulted by defective ATP7B .
3. 24-hour urine copper test (< 150µg normal)
4. Liver biopsy for histology and histochemistry and copper quantification

5. Genetic testing, haplotype analysis for siblings and mutation analysis

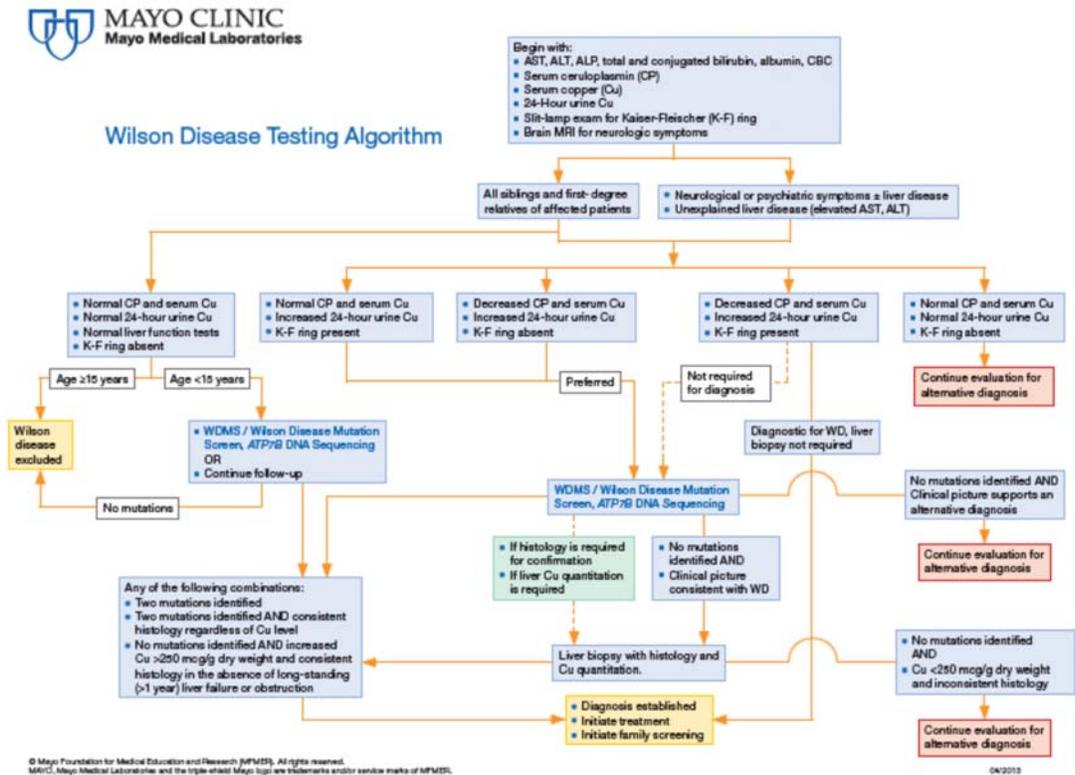


Figure 1.8A. Algorithm for WD diagnosis as adopted by Mayo Clinic, USA. (Website <http://www.mayomedicallaboratories.com/testcatalog/Clinical+and+Interpretive/83697>)

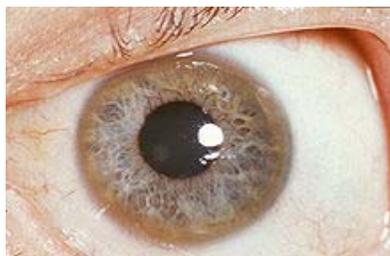


Figure 1.8B A typical KF ring (brown) in the retina

1.6 HOW IS WD TREATED

WD is a very treatable condition. With proper therapy, disease progress can be halted and very often symptoms can be improved. Treatment is aimed at removing excess accumulated copper and preventing its reaccumulation. Therapy must be lifelong. Patients may become progressively more sick from day to day so immediate treatment can be critical. Treatment delays may cause irreversible damage. Stopping treatment completely will result in death, sometimes as quickly as within three months. Decreasing dosage of medications also can result in unnecessary disease progression.

1.7 WD CAUSING MUTATIONS IN GENE *ATP7B*

Mutation screening by investigators has revealed defects that cover entire *ATP7B* gene. Prevalent mutations in *ATP7B* have been identified in different world populations. The spectrum of mutations in the *ATP7B* is available in the HGMD database (<http://www.hgmd.org>)^{wr6} and Locus Specific Universal Data Base hosted by European Council (<http://www.umd.be/ATP7B/>)^{wr7}. Nearly 600 disease causing mutations in *ATP7B* have been compiled as database (Kenney & Cox, 2007)²⁸. Following a clinical presentation, the diagnostic tests for Wilson's disease include mainly biochemical investigations, in particular, parameters of copper metabolism (Huster, 2010)⁸. Conducting genetic tests remains the preferred alternative in case of inconclusive biochemical tests or to identify affected but pre-symptomatic family members. This can include haplotype analysis, screening for known mutations and / or sequencing of the gene. Indigenous development of **WDDM** is an effort for simultaneous detection of known *ATP7B* mutations in suspected WD patients and their kin. Mutations covered in WDDM are given in Table II which includes 63 functional mutations in *ATP7B* reported in India and

world-over that cause WD. Exon wise distribution of WD causing mutations reported in Indians and worldover are given in Figure 1.9. It is noticeable that mutation pattern is not much different in Indian population than the world-over populations. This is because of the specificities of functional domains of this very large protein. As seen in Figure 1.9 exon 2, 8, 13, 14 and 18 have high percentage (>6% of the world patient population) of WD causing mutations in *ATP7B* gene.

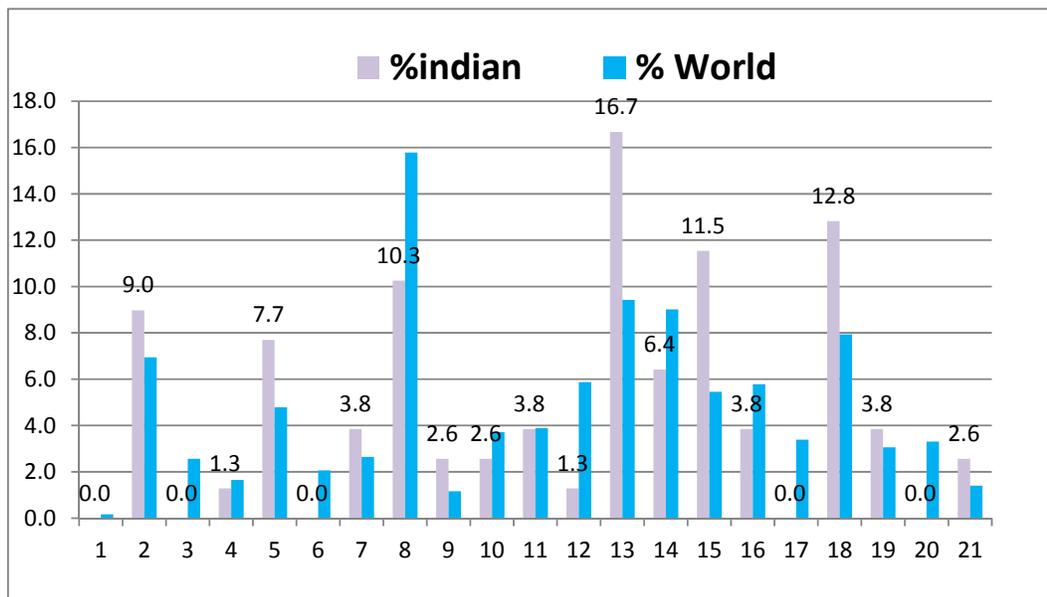


Figure 1.9 Exon wise distribution of WD causing mutations on Atp7b gene. X-axis represents 21 exons of *ATP7B* and Y-axis represent reported WD mutations (% of WD patients) in Indians and world over.

Annu et al (2014)²⁹ did extensive analysis of WD mutations in Western India in relation to their phenotype. Exon wise distribution of 36 mutations found in their study with German collaboration among a cohort of 52 patients is shown in Figure 1.10

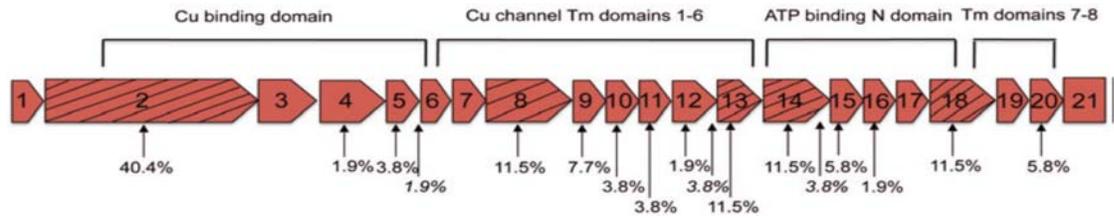


Figure 1.10. Exon wise distribution of WD mutations in western Indian population.

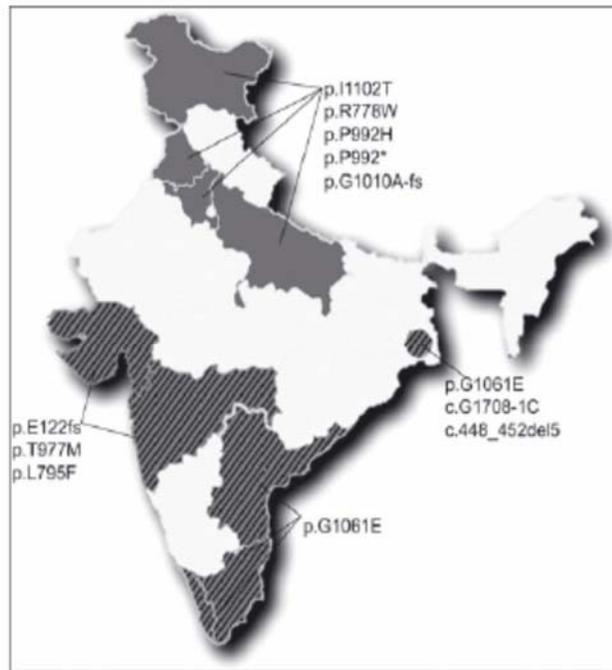


Figure 1.11: Grey areas indicate places, where WD studies have been taken

1.8 TECHNIQUES FOR MUTATION DETECTION

DNA sequencing is considered the gold standard and remains a definitive procedure for detection of mutations and nucleotide variability. A large number of molecular biology techniques that can detect point mutations, small insertions and deletions in genes have been developed (Nollau and Wagner, 1997)³⁰. These techniques involve the PCR amplification of the locus under consideration followed by examination of either the presence or altered mobility of the product. The different

techniques are ARMS, Amplification-Refractory Mutation System (Newton et al, 1989)³¹, CCR Combined Chain Reaction (Wanli & Stamboork,1997)³², DGGE Denaturing Gradient Gel Electrophoresis (Fischer and Lerman, 1983)³³. Altered restriction enzyme cleavage is commonly being used for detection of single point mutations as for Thalassaemia and other Haemoglobinopathies (Old, 1996;Book Series : Methods in Molecular Medicine)³⁴ and for ante-natal diagnosis (Old, 2003; Book series: Methods in Molecular Medicine)³⁵. High Resolution Melting (HRM) (Taylor, 2009)³⁶ and quantitative Real-Time PCR (Morlan et al, 2009)³⁷ are some of the techniques that do not require a separation step for the detection of genetic differences. (Wittwer, 2009)³⁸ has reviewed HRM Analysis application in relation to mitochondrial genome and BRCA1 mutations. Methods based on nucleic acid hybridization (Zhang et al, 1991)³⁹ including microarrays of various types are gaining popularity due to the higher throughput they offer (Yoo et al, 2009)⁴⁰. Gene expression microarrays have become very popular as a tool in cancer research and other omics studies for basic research in Life Science and drug response studies after they were first introduced in 1996 by Pat Brown and his group (DeRisi et al, 1996)⁴¹.

1.9 DNA-MICROARRAYS FOR DETECTION OF COMPLEX GENETIC DISORDERS

Similar to this study on disease specific DNA-diagnostics-Microarray, there are quite a few of the diagnostic genotype microarrays that have been developed elsewhere. Yzer et al (2006)⁴² developed a microarray for Leber Congenital Amaurosis (LCA) to screen for 301 previously identified disease-associated sequence variants in seven known LCA related genes. Cremers et al (2007)⁴³ developed a genotyping microarray that included 298 Usher syndrome-associated

sequence variants that were analyzed according to arrayed primer extension (APEX) technology. A group in China (Hu et al, 2012)⁴⁴ has developed DNA-microarray to study the role of SLC26A4 gene and four other genes in sporadic non-syndromic hearing loss in children. A highly specific microarray method for four point mutation detections causing a neurological disorder called Charcot-Marie-Tooth (CMT) disease was developed (Baaj et al, 2008)⁴⁵ by spotting stem-loop probes that yielded significantly higher discrimination ratios. The number of different loci that can be analysed at a time and the cost of analysis per locus determine the ultimate utility and the choice of DNA- microarray related detection methods. Stanford University, USA developed comprehensive efficient assay β -GMDA (β -Globin Mutation Detection Assay) which thereafter is in use for detection of *Hbb* mutations in Pan-ethnic population (Owen et al, 2010)⁴⁶. β -thalassemia mutations can still go undetected, especially when using detection strategies that identify only a specific subset of the various β -globin variants. Direct DNA sequencing enables more comprehensive detection of known and unknown mutations, if PCR primers are selected to cover whole of β -globin protein synthesis gene as in β -GMDA. The classical cytogenetics is being replaced by molecular cytogenetics using Chromosome Microarray Analysis (CMA). The Oligo-SNP CMA (Chromosome Microarray Analysis) method is in use by Quest Diagnostics (Nichols Institute, California, USA) and has demonstrated a viable alternative to the G-banding method in view of its advantages in detection of sub-microscopic genomic aberrations, shorter turnaround time due to elimination of time required for culture and a higher test success rate (Wang et al, 2014)⁴⁷. Electronic detection of DNA hybridization in microarray format has been reported by Blin et. al., (2014)⁴⁸ for genotyping GJB2 gene coding for connexin-26 protein involved in pre-lingual

non-syndromic deafness using Tas (Tagged allele specific) PCR.

1.10 DNA- DETECTION FOR WD MUTATIONS

Following are published articles in literature on WD Mutation detections:

1. For diagnosis of Wilson's Disease mutations, APEX (Arrayed Primer Extension) based microarray has been developed by (Gojova et al, 2008)⁴⁹.
2. MDE (Mutation Detection Enhancement) hetero-duplex gel analysis was used by Majumdar et al (2000)⁵⁰ to identify multiple sequence changes in the *ATP7B*.
3. Using dinucleotide repeat markers for haplotype analysis Gupta et al (2005)⁵¹ obtained distinct SNP based haplotypes to decipher the genotype-phenotype relationship in the patients.
4. Gupta et al (2007)⁵² identified eight mutations in Indians by DNA sequencing after amplification of exons using primers described by Waldenstrom et al (1996)⁵³. They studied Eastern Indian population for mutations in *ATP7B* and analysed their findings for genotype correlation with phenotype of the disease. From a cohort of 400 that included 109 WD patients they concluded that mutations in *Atp7B* could not be correlated with clinical manifestation of disease in terms of age of patients and hepato-lenticular or neurological presentation. Recently they have (Gupta et al, 2013)⁵⁴ characterized an *ATP7B* genetic variant (c.2623A/G) that is causal to the disease in Indians and is a non-disease causing SNP in Chinese.
5. Santosh et al (2006)⁵⁵ and Santosh et al (2008)⁵⁶ reported WD genotype in Tamil Nadu, Kerala and Andhra Pradesh states of south India. They reported 20 mutations, which include 12 missense, 2 nonsense, 2 deletions, 3 splice-site, and 1 putative mutation - eleven of these mutations are novel changes. The

mutations, c.2302 C > T (P768L), c.2906G > A (A969E), c.3008C > T (A1003V), and c.813C > A (C271Stop) account for 9, 9, 11, and 11% of the chromosomes respectively.

6. For North Indian population, WD mutation profiles have been reported by S. Khan et al, 2012⁵⁷. Entire coding region of the *ATP7B* was screened by them using PCR followed by SSCP. Within a group of 90 WD patients they observed exon 2,13 and 18 as hot spot exons with large number of variations. Kalita et al, 2010⁵⁸ reported absence of three mutations viz. H1069Q, I1102T and R778L in 26 neurologic WD patients in 25 families.
7. Annu et al (2013)²⁹ analyzed the WD mutational pattern in large region of western India. They identified 36 different disease-causing mutations (31 exonic and five intronic splice site variants). Fourteen novel mutations were identified and Exons 2, 8, 13, 14, and 18 accounted for the majority of mutations (86.4%). Mutation, p.C271Stop, and p.E122fs, were the most common mutations with allelic frequencies of 20.2% and 10.6%, respectively in their cohort. Frequent homozygous mutations (58.9%) and disease severity assessments allowed analysis of genotype–phenotype correlations. They concluded that including data from other parts of India p.C271Stop may be the most frequent mutation across India, and may harbor a moderate to severely disabling phenotype with limited variability.

Juan Gang et al. (2013)⁵⁹ reported a novel mutation in exon 7. In this study, 114 individuals of Chinese Han population living in north China were genotyped for WD. 36 mutations and 11 single-nucleotide polymorphisms (SNPs) were found out, of which 5 were firstly described in Chinese. Among these, p.R778L (21.5%),

p.A874V (7.5%) and p.P992L (6.1%) were the most frequent mutations. A genotype of p.L770L+p.R778L+p.P992L was the most frequent triple mutations and two pairs of mutations, p.L770L/p.R778L and p.A874V/p.I929V, were closely related. In addition, a database was established to summarize all *ATP7B* mutations, including those reported previously and those identified in this study (www.wjpch.com)^{wr8}. Popular algorithms were used to predict the functional effects of these mutations, and, by comparative genomics approaches, they predicted a group of mutation hot spots for *ATP7B*. This study will broaden our knowledge about *ATP7B* mutations in WD patients in north China, and be helpful for clinical genetic testing.

1.11 RATIONALE FOR THE PRESENT STUDY

Functional mutations of *ATP7B* protein influence the physiology in poorly understood manner. Also the manifestation of clinical symptoms of WD may be at any age, starting from 3 years to 70 years making the diagnosis for this treatable disorder difficult for physicians. The combinations of clinical tests remain inconclusive due to overlap of symptoms. In this study, an attempt has been made for the development of microarray that could detect any of 63 WD causing mutations (reported from Indian population) if present in the WD suspected patient. Inclusion of WDDM test in WD diagnosis could give confidence to the Doctor for required treatment of the patients and their asymptomatic kin if found positive on WDDM. During the process of natural selection, unlike most of the deleterious mutations, gene mutations in gene *ATP7B* are not lost from the population gene pool and also mutated alleles do not cause disease unless they are homozygous at that locus. Population wise study of *ATP7B* alleles could give insights regarding the inheritance pattern of the alleles. A website for the mathematical concepts of

population genetics has been developed at <http://www.dorak.info/genetics/popgen.html>⁹.

1.12 SCOPE OF THE PRESENT STUDY

Genetic testing is preferred alternative for suspected individuals with a family history. In this study a home-made microarrayer was used for the development of WDDM microarray that could detect 63 WD causing mutations reported from Indian population. Use of DNA microarray for detecting gene mutations is gaining popularity because of the advantages it offers in terms of throughput and sensitivity. During this study on development of WDDM quality analysis, procedures for microarray based diagnosis have been established. Inter and intra array comparisons were used for studying the quality of arrays with respect to their specificity and sensitivity. The arrays were validated by using six mutated amplicons (mutations generated by site directed mutagenesis) for hybridization on WDDM. Our arrays have shown better efficiency as sensitivity is more than 80% in detecting mutations thus giving confidence that a microarray based approach can be cost-effective for detecting large number of mutations simultaneously.

CHAPTER II

MATERIALS AND METHODS

2.1 MOLECULAR SEQUENCE DATA OF ATP7B FOR WDDM

Though the biological regulation is dynamic, we can search for mediating molecules. *ATP7B* mRNA is one such molecule relevant for WD and knowledge of its sequence became basis of the design of WDDM. Also the genomic sequence where *ATP7B* gene is mapped on human chromosome 13 is a 78,825 base pairs region, having genomic coordinates 52,506,806 to 52,585,630 on the web site of <http://genome.ucsc.edu>. The mRNA sequence derived and experimentally obtained has been made available by GeneBank, with sequence annotations and references. Figure 2.1 is functional depiction of *ATP7B* within 78,825 base pair genomic region. Horizontal curved lines are intronic regions that do not translate for the protein ATP7B. Vertical lines represent exonic regions, which are necessarily present in mRNA sequence.



Figure 2.1. Exonic structure of *ATP7B*

To map translating regions on gDNA, exons were numbered, starting from first translating region, followed by non-translating intervening sequence, intron and numbered one. Exon 2 follows after intron 1 and so on. Splice junctions are specific sequences present within the intron as shown in Figure 2.2.

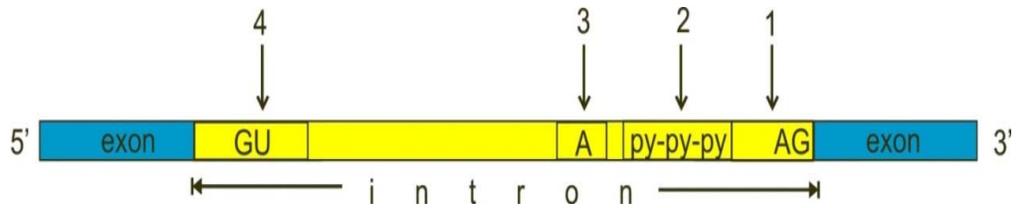


Figure 2.2 Intron junctions (http://en.wikipedia.org/wiki/RNA_splicing)^{wr10}

Point mutations in the underlying DNA or errors during transcription may activate a cryptic splice site in part of the transcript not usually spliced. This could result in a mature messenger RNA with a missing section of an exon. In this way, a point mutation, which usually only affects a single amino acid, can manifest as a deletion in the final protein. One such mutation in exon 8 of *ATP7B* is correlated with WD.

2.2 FUNCTIONAL DOMAINS AND CORRESPONDING EXONS

To cover most of the WD causing mutations including splice variants, genomic sequence and mRNA sequence were both consulted during designing of primers and probes. Nucleotide positions in mRNA sequence and corresponding functional domains across 21 exons are listed in Table II.

Note: DNA sequence of Homo sapiens ATP7B was downloaded from NCBI database (<http://ncbi.nlm.nih.gov>, Accession No. NG_008806). mRNA sequence of ATP7B with Annotations was downloaded from GenBank LOCUS – NM_000053.

Table II. Exons and their relation to functional domains of the ATP7 protein

Exon no.	Nucleotide position	Protein Domains	Exon no.	Nucleotide Positions	Protein Domains
1	158-208	Cu1	2	209-1442	Cu2/3/4
3	1443-1700	Cu4/5	4	1701-1864	Cu5/6
5	1865-2026	Cu6	6	2027-2103	Cu6/TM1
7	2104-2278	TM1/2	8	2279-2512	TM2/3/4
9	2513-2604	TM4	10	2605-2732	TM4/Td
11	2733-2887	Td/TM5	12	2888-3022	Ch/TM6
13	3023-3217	Phosphorylation domain	14	3218-3400	ATP loop
15	3401-3569	ATP loop	16	3570-3713	ATP binding
17	3714-3856	ATP hinge	18	3857-4060	TM7
19	4061-4178	TM7	20	4179-4281	TM7/8
21	4282-4555	TM8			

The WD causing mutations taken from mutation data base of Kenney & Cox 2007²⁸ , and other literature were studied for their exon wise distribution. A histogram prepared after detailed study of mutations world-wide and in India is given in Figure 1.9. It was observed that exon wise distribution of mutations in Indian population and world population was not much different. This would be because of the conservation of the molecular structure of human ATP7B

independent of the ethnicity of the human population and the geographical location. Most of the WD causing mutations identified so far were present on exon 8, 13, 14 and 18 of *ATP7B*.

2.3 MUTATION DATA BASE FOR WDDM

WD causing mutations covered in our study and covered on WDDM are listed in Table III. Mutation type, its mRNA coordinate, affected (mutated) amino acid, amino acid position, exon number and the protein domain of mutation are given in Table III for each of the mutation. Justification for selecting these mutations for WDDM is given in Introduction chapter. They cover 63 mutations present on 16 amplicons. Single stranded DNA probes were designed after extensive calculations using Oligoarray program. We had to design the left and right primers for PCR from genomic sequence to be able to amplify mutation containing DNA regions for testing them on WDDM. The amplified regions of size 291 to 723 bps (base pairs) were referred as ‘AMPLICONS’. Since our gene *ATP7B* has some introns/exons very short and some very long, the 16 amplicons have no direct correlation with exon numbers used in databases.

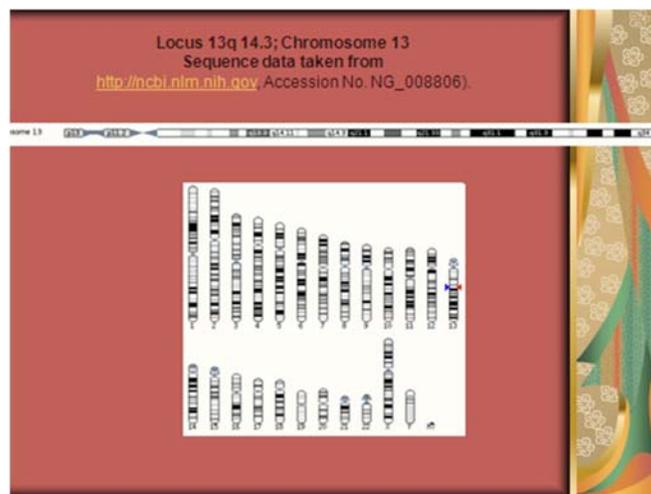


Figure 2.3. All human chromosomes in haploid, shown as 25 vertical shapes. Chromosome 13 is marked among these and is also shown in detail as horizontal line to show locus of Atp7B.

Table III. LIST OF MUTATIONS ADDRESSED IN WDDM

Mutation type	Nucleotide	Amino acid mutated	Exon /intron	Protein domain
insertion	c.174dupC	p.Thr59HisfsX19	2	Cu1
deletion	c.448_452del	p.Glu150HisfsX11	2	Cu2
substitution	c.561T>A	p.Tyr187Stop	2	Cu2
substitution	c.813C>A	p.Cys271Stop	2	Cu3
deletion	c.892delC	p.Gln298LysfsX2	2	Cu3
substitution	c.997G>A	p.Gly333Arg	2	Cu3
insertion	c.1707+11dupGT	na	4	Cu5
substitution	c.1708-1G>C	na	5	Cu5
insertion	c.1747_1748insT	p.Asn581SerfsX232	5	Cu5
substitution	c.1771G>A	p.Gly591Ser	5	Cu5
substitution	c.1847G>A	p.Arg616Gln	5	Cu6
insertion	c.1849dupG	p.Asp617GlyfsX7	5	Cu6
deletion	c.1963delC	p.Leu655CysfsX13	7	Cu6
deletion	c.2116_2117del	p.Val706ProfsX48	7	TM1
substitution	c.2128G>A	p.Gly710Ser	8	TM1
substitution	c.2145C>A	p.Tyr715Stop	8	TM1
insertion	c.2224insA	p.Val742AspfsX13	8	TM2
insertion	c.2258dupC	p.Glu754Stop	8	TM2
deletion	c.2292_2312del	p.Asp765_Phe771del	8	TM3
substitution	c.2303C>T	p.Pro768Leu	8	TM3
duplication	c.2304dupC	p.Met769HisfsX26	8	TM3
deletion	c.2364delC	p.Ser789GlnfsX18	9	TM3
substitution	c.2383C>T	p.Leu795Phe	9	TM4
substitution	c.2448-1G>A	na	10	TM4
insertion	c.2497dupG	p.Val833GlyfsX21	10	TM4
insertion	c.2582_2583insG	p.Met862HisfsX5	11	TM4
substitution	c.2623G>A/A>G	p.Gly875Arg	11	Td
substitution	c.2728A>T	p.Lys910Stop	11	Td

insertion	c.2815_2816insA	p.Trp939Stop	12	Td/TM5
substitution	c.2906G>A	p.Arg969Gln	13	Ch/TM6
substitution	c.2930C>T	p.Thr977Met	13	TM6
substitution	c.2975C>A	p.Pro992His	13	TM6
insertion	c.2977dupA	p.Thr993AsnfsX35	13	TM6
substitution	c.3007G>A	p.Ala1003Thr	13	TM6
substitution	c.3008C>T	p.Ala1003Val	13	TM6
deletion	c.3026_3028del	p.Ile1009del	13	Ph domain
substitution	c.3029A>G	p.Lys1010Arg	13	Ph Domain
insertion	c.3031_3032insC	p.Gly1011AlafsX17	13	Ph Domain
substitution	c.3091A>G	p.Thr1031Ala	14	Ph Domain
deletion	c.3147delC	p.Thr1050HisfsX71	14	Ph Domain
substitution	c.3182G>A	p.Gly1061Glu	15	Ph Domain
substitution	c.3207C>A	p.His1069Gln	14	Ph Domain
substitution	c.3282C>G	p.Phe1094Leu	14	ATP loop
substitution	c.3301G>A	p.Gly1101Arg	15	ATP loop
substitution	c.3305T>C	p.Ile1102Thr	15	ATP loop
substitution	c.3311G>A	p.Cys1104Tyr	15	ATP loop
substitution	c.3412+1G>A	na	15	ATP loop
deletion	c.3418delT	p.Val1140Ala-fs	16	ATP loop
insertion	c.3424dupC	p.Gln1142ProfsX11	16	ATP loop
substitution	c.3532A>G	p.Thr1178Ala	16	ATP loop
substitution	c.3722C>T	p.Ala1241Val	18	ATP binding
substitution	c.3767A>G	p.Gln1256Arg	18	ATP binding
insertion	c.3770_3771insG	p.Asn1257LysfsX2	18	ATP binding
substitution	c.3809A>G	p.Asn1270Ser	18	ATP binding
insertion	c.3839_3840insTAC	p.Met1280delinsIleThr	18	ATP binding
substitution	c.3890T>A	p.Val1297Asp	18	ATP hinge
substitution	c.3895C>T	p.Leu1299Phe	18	ATP hinge
deletion	c.3895delC	p.Ile1300SerfsX30	18	ATP hinge
substitution	c.3903+6T>C	na	18	ATP hinge

substitution	c.4021G>A	p.Gly1341Ser	19	ATP hinge
substitution	c.4021+3A>G	na	19	ATP hinge
insertion	c.4310dupA	p.Pro1438AlafsX11	21	TM7

2.4 MOLECULAR SEQUENCE DETAILS OF 16 AMPLICONS

Following pages give details of *ATP7B* genomic regions which are relevant for mutations covered on WDDM μ -arrays. The sequence of nucleotides A-adenine, C-cytosine, T-thymine, G-guanine, wherever appears in small letters represents intronic region while sequence in CAPITAL letters represents exonic regions numbered as in Genebank Locus [NM_000053](#). An intron always begins with ‘gt’ and ends with ‘ag’ (5’ and 3’ splice junctions). Within each of 16 genomic segments, left and right primers were shown in green color. These primers were designed by extensive sequence analysis of *ATP7B* using program (<http://www.idtdna.com/Scitools/Applications/Primerquest>)^{wr10} and were short listed such that the amplicon contains all probable mutations in that exon. Each 16 segments given below include translation of exonic regions with amino acid residues, in which WD correlated amino acids were shown in red color. Exon 2 spanning from 209 to 1442 in mRNA is the largest exon of size 1415 bases. There were two amplicons in this region which were named 2a and 2b. All the amplicon data is within **chr13:52,506,806 - 52,585,630; total 78,825 bp** referring to UCSC (University of California, Santa Cruz) genome browser at (<http://genome.ucsc.edu>)^{wr11} which is congruent to sequence at locus 13q14.3.

AMPLICON NUMBER 1

Exon2a: 209 to 901 in mRNA

tccatthttctcagtgccagagaagctgggatgttgtagaaaatatttggtttcaaggtta
aaaaatggtatthttctthttctthtttagATCTTATCTAAGCTTTCTTTGCCTACCCGTGCCT
GGGAACCAGCAATGAAGAAGAGTTTTGCTTTTGACAATGTTGGCTATG AAGGTGGTCTG
GATGGCCTGGGCCCTTCTTCTCAGGTGGCCACCAGCACAGTCAGGATCTTGGGCATGACT
TGCCAGTCATGTGTGAAGTCCATTGAGGACAGGATTTCCAATTTGAAAGGCATCATCAGC
ATGAAGGTTTTCCCTGGAACAAGGCAGTGCCACTGTGAAATATGTGCCATCGGTTGTGTGC
CTGCAACAGGTTTGCCATCAAATTTGGGGACATGGGCTTCGAGGCCAGCATTGCAGAAGGA
AAGGCAGCCTCCTGGCCCTCAAGGTCCTTGCTGCCAGGAGGCTGTGGTCAAGTCCGG
GTGGAGGGCATGACCTGCCAGTCCGTGTGCAGCTCCATTGAAGGCAAGGTCGGAAACTG
CAAGGAGTAGTGAGAGTCAAAGTCTCACTCAGCAACCAAGAGGCCGTCATCACTTATCAG
CCTTATCTCATTGACCCGAAGACCTCAGGGACCATGTAAATGACATGGGATTTGAAGCT
GCCATCAAGAGCAAAGTGGCTCCCTTAAGCCTGGGACCAATTGATATTGAGCGGTTACA

Expected size of amplicon: 692bps;

Amino acid translation:

atcttatctaagctthttctthtgcctaccctgccc
I L S K L S L P T R A
tgggaaccagcaatgaagaagagthtttgctthttgacaatgttggctatgaaggtggtctg
W E P A M K K S F A F D N V G Y E G G L
gatggcctgggccctthttctcaggtggccacagcacagtcaggatcttgggcatgact
D G L G P S S Q V A T S T V R I L G M T
tgccagtcatgtgtgaagtccattgaggacaggatthccaatthgaaagggcatcatcagc
C Q S C V K S I E D R I S N L K G I I S
atgaaggtthccctggaacaaggcagtgccactgtgaaatattgtgcatcggttggtgtgc
M K V S L E Q G S A T V K Y V P S V V C
ctgcaacaggtthgcatcaaattgggacatgggctthcgaggccagcattgcagaagga
L Q Q V C H Q I G D M G F E A S I A E G
aaggcagcctcctggccctcaaggtccttgctgcccaggaggctgtggtcaagctccgg
K A A S W P S R S L P A Q E A V V K L R
gtggagggcatgacctgcccagtcctgtgtcagctccattgaaggcaaggtccggaaactg
V E G M T C Q S C V S S I E G K V R K L
caaggagtgtgagagtcaaagtctcactcagcaaccaagaggccgtcatcacttatcag
Q G V V R V K V S L S N Q E A V I T Y Q
ccttatctcattcagcccgaagacctcaggaccatgtaaatgacatgggattthgaagct
P Y L I Q P E D L R D H V N D M G F E A
gccatcaagagcaaagtggtcctthtaagcctgggaccaattgatattgagcggttaca
A I K S K V A P L S L G P I D I E R L

Mutations covered in amplicon 1 = 3

AMPLICON NUMBER 2

Exon2b: 902 to 1442 in mRNA

TGCCATCAAGAGCAAAGTGGCTCCCTTAAGCCTGGGACCAATTGATATTGAGCGGTTACA
AAGCACTAACCCAAAGAGACCTTTATCTTCTGCTAACCAGAATTTAATAAATCTGAGAC
CTTGGGGCACCAAGGAAGCCATGTGGTCACCCCTCCAACCTGAGAATAGATGGAATGCATTG
TAAGTCTTGCGTCTTGAATATTGAAGAAAATATTGGCCAGCTCCTAGGGGTTCAAAGTAT
TCAAGTGTCTTGGAGAACA AAACTGCCCAAGTAAAGTATGACCCTTCTTGTACCAGCCC
AGTGGCTCTGCAGAGGGCTATCGAGGCACTTCCACCTGGGAATTTTAAAGTTTCTCTTCC
TGATGGAGCCGAAGGGAGTGGGACAGATCACAGGTCTTCCAGTTCTCATTCCCCTGGCTC
CCCACCGAGAAACCAGGTCCAGGGCACATGCAGTACCCTCTGATTGCCATTGCCGCAT

GACCTGTGCATCCTGTGTCCATTCCATTGAAGGCATGATCTCCCAACTGGAAGGGGTGCA
GCAAAATATCGGTGTCTTTGGCCGAAGGGACTGCAACAGTTCTTTATAATCCCTCTGTAAT
TAGCCCAGAAGAACTCAGAGCTGCTATAGAAGACATGGGATTTGAGGCTTCAGTCGTTTC
GCCAGAAGAACTCAGAGCTGCTATAGAAGACATGGGATTTGAGGCTTCAGTCGTTCTG
gtacgtagtgtgtttgaggcatgtcctgagcttgtctcctttctctttgtgtcttatag
ctcctggatggtggtataggtgagccctgctccctgctcccctgctcaacagtggaacaa

Expected size of amplicon: 723

tgccatcaagagcaaagtggctcccttaagcctgggaccaattgatattgagcggttacaa
A I K S K V A P L S L G P I D I E R L Q
agcactaaccctaaagagacctttatcttctgctaaccagaattttaataattctgagacc
S T N P K R P L S S A N Q N F N N S E T
ttggggcaccaaggaagccatgtggtcaccctccaactgagaatagatggaatgcattgt
L G H Q G S H V V T L Q L R I D G M H C
aagtcttgctcttgaatattgaagaaaatattggccagctcctaggggttcaaagtatt
K S C V L N I E E N I G Q L L G V Q S I
caagtgtccttgagaacaaaactgcccaagtaaagtatgacccttctgtaccagccca
Q V S L E N K T A Q V K Y D P S C T S P
gtggctctgcagagggctatcgaggcacttccacctgggaattttaagtttctcttctc
V A L Q R A I E A L P P G N F K V S L P
gatggagccgaagggagtgggacagatcacaggtcttccagttctcattcccctggctcc
D G A E G S G T D H R S S S S H S P G S
ccaccgagaaaccaggtccagggcacatgcagtaccactctgattgccattgcccggcatg
P P R N Q V Q G T C S T T L I A I A G M
acctgtgcatcctgtgtccattccattgaaggcatgatctcccaactggaaggggtgcag
T C A S C V H S I E G M I S Q L E G V Q
caaatatcgggtgtctttggccgaagggactgcaacagttctttataatccctctgtaatt
Q I S V S L A E G T A T V L Y N P S V I
agcccagaagaactcagagctgctatagaagacatgggatttgaggcttcagtcgtttgc
S P E E L R A A I E D M G F E A S V V S

Mutations covered in amplicon 2 = 3

AMPLICON NUMBER 3

EXON4: 1701 to 1864 in mRNA

caccagagtggttacagccatgacctgatggttccagGTGTTCTCTCCGT
GTTGGTTGCCTTGATGGCAGGAAAGGCAGAGATCAAGTATGACCCAGAGG
TCATCCAGCCCCTCGAGATAGCTCAGTTCATCCAGGACCTGGGTTTTGAG
GCAGCAGTCATGGAGGACTACGCAGGCTCCGATGGCAACATTGAGCTGAC
Agtaagtactgtgggtgcttacgggggttacaggcttctgacagtttgca
ttttggacacatccatctttgtgattagtaaatttcccatcttggacgt
gtctggttgggtggtggtggtatgtttggttcttcttcttcttgaag

Expected size of the amplicon = 291bp

Amino acid translation:

ggtgttctctccgtggttggcttgccttgatggcaggaagggcagagatcaagtatgacc
G V L S V L V A L M A G K A E I K Y D P
gaggtcatccagcccctcgagatagctcagttcatccaggacctgggttttgaggcagca
E V I Q P L E I A Q F I Q D L G F E A A
gtcatggaggactacgcaggctccgatggcaacattgagctgaca
V M E D Y A G S D G N I E L T

Mutations covered in amplicon 3 = 1 (splice variant) not shown in translated region.

AMPLICON NUMBER 4

Exon5: 1865 to 2026 in mRNA

agaatctcacatgccccgtttggaggtggagtcagggtcttgagagcag
tgctgaggagggaaaggctcttgctgcctgttacctagactccctggac
tggtcttcacaggctttccttgatcctgggtctgtgggattcttgccatc
ctgtgtgcagATCACAGGGATGACCTGCGCGTCCCTGTGTCCACAACATA
GAGTCCAACTCACGAGGACAAATGGCATCACTTATGCCTCCGTTGCCCT
TGCCACCAGCAAAGCCCTTGTAAAGTTTGACCCGAAATATCGGTCCAC
GGGATATTATCAAATTTATTGAGgtaagtaattcattaaaaaattgtagt
caccttttaaaaacagtaatatataatcagtgaagaaaaatgagaaaaat
atagctaagaaaaaatgaaaaataacaaaattcttcaacttttccatgg
aaataagctcagtcattttaaaataaactgcaggttaagtaataataaat
aatctgagggaggatatcagtgaaatgggtaagagaaagatgaggaaaaga
Expected size of amplicon = 495bp

Amino acid translation:

atcacagggatgacctgccccgtcctgtgtccacaacatagagtccaaactcacgaggaca
I T G M T C A S C V H N I E S K L T R T
aatggcatcacttatgcctccggtgcccttgccaccagcaaagcccttgtaagtttgac
N G I T Y A S V A L A T S K A L V K F D
ccggaaattatcggtccacgggatattatcaaaattattgag
P E I I G P R D I I K I I E

Mutations covered in amplicon 4 = 3

AMPLICON NUMBER 5

EXON7: 2104 to 2278 in mRNA

catctttggaacctccttatagtgatggttagacctctagatgctccctc
agatggccagtggtcagagagagagaggttcttactttcattttaaccctg
gtggtctgtcccagacatgtgacaaaggcagggtcttaactgtgtcctca
gaagggagtggtcttgtaatccaggtgacaagcagcatctgatatatctg
tggtgtgcatttgctttccagGTGGAAGAAGTCTTTCCTGTGACGCTG
GTGTTTGGCATCCCTGTGCATGGCCTTAATGATCTATATGCTGATACCCAG
CAACGAGCCCCACAGTCCATGGTCTTGGACCACAACATCATTCCAGGAC
TGTCATCTAAATCTCATCTTCTTTATCTTGTGTACCTTTGTCCAGgta
tatatgagaaagtgggcagacctctcccttccatgctgtgtgtggccctc
agatattctgcccgctaagcgcaaacatagtgcttttagcttggttaaat
ggcactttattgcagctttcctgccccaccaggcagttgctgctttttt

Expected size of amplicon = 496bp

Amino acid translation:

gtggaagaagtctttcctgtgcagcctgggtggtttggcatccctgtcatggccttaatgatc
W K K S F L C S L V F G I P V M A L M I
tatatgctgataccagcaacgagccccaccagtcctggtcctggaccacaacatcatt
Y M L I P S N E P H Q S M V L D H N I I

ccaggactgtccattctaaatctcatcttctttatcttgtgtacctttgtccag
P G L S I L N L I F F I L C T F V Q

Mutations covered in amplicon 5 = 2

AMPLICON NUMBER 6

Exon8 : 2279 to 2512 in mRNA

ctcttggatccattcctgtggacagtagtcctctgaatgggaaagtat
atctcataaacgcccacacagaggaagaagtactgtcacgactgtgcac
aaagctagaggctttgccatccccagggcccttggccctgtgtcgctcat
tgaactctcctccctacttgctggcagccttcaactgtccttgtctttcag
CTCCTCGGTGGGTGGTACTTCTACGTTTCAGGCCTACAAATCTCTGAGACA
CAGGTCAGCCAACATGGACGTGCTCATCGTCTGGCCACAAGCATTGCTT
ATGTTTATTCTCTGGTCATCCTGGTGGTGTGCTGTGGCTGAGAAGGCGGAG
AGGAGCCCTGTGACATTCTTCGACACGCCCCCATGCTCTTGTGTTTCAT
TGCCCTGGGCCGGTGGCTGGAACACTTGGCAAAGgtaacagcagcttcag
gttcagaaaagagctgtccttcagtaaacaatctcaacttccctctgaac
accatggtttagaattactaattatacacagcatagagacagacttaaaga
aataggaaactccatataattaagggtgtcttagtcactaatctccaaat

Expected size of amplicons = 555bp and 482 bps

Amino acid translation:

ctcctcgggtgggtggtacttctacgctcaggcctacaaatctctgagacacaggtcagcc
L L G G W Y F Y V Q A Y K S L R H R S A
aacatggacgtgctcatcgtcctggccacaagcattgcttattgtttattctctggtcatc
N M D V L I V L A T S I A Y V Y S L V I
ctgggtggttgtgtggctgagaaggcggagaggagccctgtgacattcttcgacacgccc
L V V A V A E K A E R S P V T F F D T P
cccattgctctttgtgttcattgccctgggcccgggtggctggaacacttggcaaag
P M L F V F I A L G R W L E H L A K

Mutations covered in amplicon 6 = 7

AMPLICON NUMBER 7

EXON9 : 2513 to 2604 in mRNA

ctgtggaagtgacatgtggccatgtgtggtggatagcaagtaagcccac
ctgcagagccttttatcgtgcccgtgcccgtgtttctctcgcaccagctgt
ctctaaccaccagccttgtgactctcaggctgggtttggacaggtctgctt
tcgatagctctcatttcacattctggttatttcctagAGCAAAACCTCAG
AAGCCCTGGCTAAACTCATGTCTCTCCAAGCCACAGAAGCCACCGTTGTG
ACCCCTGGTGAGGACAATTTAATCATCAGgtgagttatggttatcaaatg
tctttgtggttggtatctatcaatctgtgtgagctgcatcagatgcccat
gttgattgacattgcaatagaccttgtgagtggtggcagagacacagta
agatcaccactctcaatccagctacgaaagcaaggcattgaaactataaaa

Expected size of amplicon = 392bp

Amino acid translation:

agcaaaacctcagaagccctggctaaactcatgtctctccaagccacagaagccaccgtt
S K T S E A L A K L M S L Q A T E A T V
gtgacccttggtgaggacaatttaacatc
V T L G E D N L I I

Mutations covered in amplicon 7 = 2

AMPLICON NUMBER 8

Exon 10/11: 2605 to 2732 exon 10 & 2733 to 2887 exon 11 of mRNA

gaccgaatgagtgccatgtgagtgataaagtggcgtttgttcagGGAGG
AGCAAGTCCCATGGAGCTGGTGCAGCGGGCGATATCGTCAAGGTGGTC
CCTGGGGGAAAGTTTCCAGTGGATGGGAAAGTCTGGAAGGCAATACCAT
GGCTGATGAGTCCCTCATCACAGgtgagatggcttgtttcatgttccctc
aggaggatatcatagcagctgtcaggtcacatgagtgctggatggggctg
agcaagtgacagttgtctctttcctacgtctagGAGAAGCCATGCCAGTC
ACTAAGAAAACCCGAAGCACTGTAATTGCGGGGTCTATAAATGCACATGG
CTCTGTGCTCATTAAAGCTACCCACGTGGGCAATGACACCACTTTGGCTC
AGATTGTGAAACTGGTGAAGAGGCTCAGATGTCAAAGgtaatgaagaaa
tttttaaaactaacttcatctttctcgttttagaaattatgtgaagagtt
ctgggaaatcagacagttttattgagtagagattgattagtaaatgtggt
taaatgaaggagattatccc aatctttatccatgcttgtggtgttttatt

Expected size of amplicon = 547bp

Amino acid translation:

ggaggagcaagtccccatggagctgggtgcagcggggcgatatcgtcaagggtggccttggg
E E Q V P M E L V Q R G D I V K V V P G
ggaaagtttccagtgatgggaaagtccctggaaggcaataccatggctgatgagtcctc
G K F P V D G K V L E G N T M A D E S L
atcacaggagaagccatgccagtcactaagaaacccggaagcactgtaattgccccgtct
I T G E A M P V T K K P G S T V I A G S
ataaatgcacatggctctgtgctcattaaagctaccacgtgggcaatgacaccactttg
I N A H G S V L I K A T H V G N D T T L
gctcagattgtgaaactgggtggaagaggtcagatgtcaaag
A Q I V K L V E E A Q M S K

Mutations covered in amplicon 8 = 5

AMPLICON NUMBER 9

Exon 12: 2888 to 3022 in mRNA

tttttaaaactaacttcatctttctcgttttagaaattatgtgaagagtt
ctgggaaatcagacagtttattgagtagagattgattagtaaatgtggt
taaatgaaggagattatccc aatctttatccatgcttgtggtgttttatt
tcttcataggttgtaatttccatggcttgggtgttttattttcatagGC
ACCCATT CAGCAGCTGGCTGACCGTTT AGTGGATATTTGTCCATTTA
TCATCATCATGTCAACTTTGACGTTGGTGGTATGGATTGTAATCGGTTTT
ATCGATTTTGGTGTGTTTCAGAGATACTTTCTGtaagttgaatgccttg
ggctatatgggtggttgtgttttaataatctactgacattgatcctgttc

tttcatatcttagattcactgggctttaattattcattacatctttatttg
cttgcttctcttattgacagcaaaaatctaagccagagtaggataaacagt

Expected size of amplicon = 447bp

Amino acid translation:

gcaccattcagcagctggctgaccggtttagtggatattttgtcccatttatcatc
A P I Q Q L A D R F S G Y F V P F I I I
atgtcaactttgacgttgggtggatggattgtaacggttttatcgattttgggtgtgtt
M S T L T L V V W I V I G F I D F G V V
cagagatactttcct
Q R Y F P

Mutation covered in amplicon 9 = 1

AMPLICON NUMBER 10

Exon 13: 3023 to 3217 in mRNA

aacgtgttctctatgatggcagagcagtggtggaataccatctggttccgg
aacccaagttcgtcacgttgtgtccagtgccccctgaaatgtccttatg
tgattagagttctgggagcttcttattgaactctcaacctgcctctgac
tctgtcctgttttcagAACCCCAACAAGCACATCTCCAGACAGAGGTGA
TCATCCGGTTTGCTTTCCAGACGTCCATCACGGTGCTGTGCATTGCCTGC
CCCTGTCCCTGGGGCTGGCCACGCCACGGCTGTCATGGTGGGCACCGG
GGTGGCCGCGCAGAACGGCATCTCATCAAGGGAGGCAAGCCCCTGGAGA
TGGCGACAAGgtcagcctgtagcagggctttccccatcctgagagatga
aagtagtatctgtttactatctcacattgagagaaaagcctgagagccac
tcaagacagcagtgtaattacatagaataggaagtcaagtataactggg
aa**taacaacagtagcaacagagtagcc**accagtcataataatggaacgctt

Expected size of amplicon = 500bp

Amino acid translation:

aacccaacaagcacatctcccagacagaggtgatcatccggtttgctttccagacgtcc
N P N K H I S Q T E V I I R F A F Q T S
atcacggtgctgtgcattgcctgccccctgctccctggggctggccacgcccacggctgtc
I T V L C I A C P C S L G L A T P T A V
atgggtgggcaccggggtggccgcgagaaacggcatcctcatcaaggagggaagcccctg
M V G T G V A A Q N G I L I K G G K P L
gagatggcgcacaag
E M A H K

Mutations covered in amplicon 10 = 8

AMPLICON NUMBER 11

exon14 :3218 to 3400 in mRNA

ctgagattgaacgacagaggatcacgttaggaagctgtgcaggtgtcttg
tttctgtctgaggcaggttgggtgaagttctgcctcaggagtgtagta
tggaagccccctccatctgtattgtggtcagtgagttgtggttgtttttgg
cagATAAAGACTGTGATGTTTGACAAGACTGGCACCATTACCCATGGCGT

CCCCAGGGTCATGCGGGTGCCTCTGCTGGGGGATGTGGCCACACTGCCCC
TCAGGAAGGTTCTGGCTGTGGTGGGGACTGCGGAGGCCAGCAGTGAACAC
CCCTTGggcgtggcagtcaccaataactgtAAAGAGgtacgtggacttgg
Gcgtggccctgccctccccccaatgctctttttattcctcaccatgtcct
Tctctcctagctgccctcgaggagccttctctgtgtggtctggaaaacca

Expected size of amplicon = 419bp

Amino acid translation:

ataaagactgtgatgtttgacaagactggcaccattaccatggcgtccccagggatc
I K T V M F D K T G T I T H G V P R V M
cgggtgctcctgctgggggatgtggccacactgcccctcaggaaggctctggctgtggtg
R V L L L G D V A T L P L R K V L A V V
gggactgcggaggccagcagtgaaacccccttggcgtggcagtcaccaataactgtaaa
G T A E A S S E H P L G V A V T K Y C K
gag
E

Mutations covered in amplicon 11 = 4

AMPLICON NUMBER 12

exon 15 : 3401 to 3569 in mRNA

aattacaagttactagtcactatc**ttaacctttcctatctgttccacctc**
cctcccctcctttctctcagttcccgtttccgctgctctcttgccacct
tcaccctgtgtccctgtcctgctgcccctcccctttcacttcacccctct
tggettacagtttctctctctctctttccaccttcccagGAACTTGGAA
CAGAGACCTTGGGATACTGCACGGACTTCCAGGCAGTGCCAGGCTGTGGA
ATTGGGTGCAAAGTCAGCAACGTGGAAGGCATCCTGGCCACAGTGAGCG
CCCTTTGAGTGCACCGCCAGTCACCTGAATGAGGCTGGCAGCCTTCCCG
CAGAAAAAGgtattgctggcttttgtctctgcagctggttaaaagtagag
gtgggtcaaacacagagagcaccacgcccagcagtgattgcctctgctg
tgcgccagacggttcatggctaaggcaoccaagcctgcctccccac**acc**
aggaaagtttctcttatgttcttgggtgctgctaaattttgttgcttatg

Expected size of amplicon = 497bp

Amino acid translation:

gaacttggaaacagagaccttgggatactgcacggacttccaggcagtgccaggctgtgga
E L G T E T L G Y C T D F Q A V P G C G
attgggtgcaaagtcagcaacgtggaaggcatcctggcccacagtgagcgcctttgagt
I G C K V S N V E G I L A H S E R P L S
gcaccggccagtcacctgaatgaggctggcagccttcccgcagaaaaag
A P A S H L N E A G S L P A E K

Mutations covered in amplicon 12 = 4

AMPLICON NUMBER 13

Exon 16 : 3570 to 3713 in mRNA

attttactttttttttttttgtcctaaggatgctgtcacaagag**ggtgct**
tacaaggttacagtttttcagaatggttaaaaggatattttgctgttaaa

aggattgcatggTTTTtagttcacagtgaaattggaccatttagaaataa
ccacagcctctTTTTgaatagATGCAGTCCCCCAGACCTTCTCTGTGCTGA
TTGGAAACCGTGAGTGGCTGAGGCGCAACGGTTTAACCATTTCTAGCGAT
GTCAGTGACGCTATGACAGACCACGAGATGAAAGGACAGACAGCCATCCT
GGTGGCTATTGACGgtatcttctgcttctgccttccttccgctctctcag
Aaatacagtttctgcagatatcaggcaaaagagtcctcctttataaaaag
Aaaagaagacaacaaaagccttctctcttaatttcaggcctgttttcc

Expected size of amplicon exon 16=396bp;

Amino acid translation:

gatgcagtccccagaccttctctgtgctgattggaaaccgtgagtggctgagggcgaac
D A V P Q T F S V L I G N R E W L R R N
ggtttaaccatttctagcgtatgacgacgctatgacagaccacgagatgaaaggacag
G L T I S S D V S D A M T D H E M K G Q
acagccatcctggggctattgacg
T A I L V A I D

Mutations covered in amplicon 13 = 3

AMPLICON NUMBER 14

Exon 18 : 3857 to 4060 in mRNA

ccagggataaactggccctgtgacagcaaactgcaggggtgtggttgacc
aacatcactgactggaccagaggccgagggcagagggggcaagggtaac
ttgaggtttctgctgctatctgataccttttgccaacactaggcattgcc
ttccttttgtcttagGTTGGCATCAACAAAGTCTTTGCAGAGGTGCTGCC
TTCGCACAAGGTGGCCAAGGTCCAGGAGCTCCAGAATAAAGGGAAGAAAG
TCGCCATGGTGGGGGATGGGGTCAATGACTCCCCGGCCTTGGCCAGGCA
GACATGGGTGTGGCCATTGGCACCGGCACGGATGTGGCCATCGAGGCAGC
CGACGTCGTCCTTATCAGAgtgagcgtggctgcagccaggctgtgggtgc
tgggagggcaatgggcagacccttctcactgtgtgctcctctccatca
gAATGATTTGCTGGATGTGGTGGCTAGCATTACCTTTCCAAGAGGACTG
TCCGAAGGATACGCATCAACCTGGTCTGGCACTGATTTATAACCTGGTT
GGGATACCCATTGCAGCAGgtaggcagctcttaccactgtgctccagct

Expected size of amplicon = 479bp

Amino acid translation:

gttggcatcaacaaagtctttgcagaggtgctgccttcgcacaaggtggccaaggtccag
V G I N K V F A E V L P S H K V A K V Q
gagctccagaataaaggaagaaagtgcctatgggtgggggatggggatcaatgactccccg
E L Q N K G K K V A M V G D G V N D S P
gccttggcccaggcagacatgggtgtggccattggcaccggcagcggatgtggccatcgag
A L A Q A D M G V A I G T G T D V A I E
gcagccgacgtcgtccttatcaga
A A D V V L I R

Mutation covered in amplicon 14 = 7

AMPLICON NUMBER : 15

Exon 19 : 4061 to 4178 in mRNA

CGACGTCGTCCTTATCAGAgtagcgtggctgcagccaggctgtgggtgc
tgaggaggcaatgggcagacccttcctcactgtgtgctcctctccatca
gAATGATTTGCTGGATGTGGTGGCTAGCATTACCTTTCCAAGAGGACTG
TCCGAAGGATACGCATCAACCTGGTCTTGGCACTGATTATAACCTGGTT
GGGATACCCATTGCAGCAGgtaggcagctcttaccactgtgctccagct
GCGCCAGAAAGGCTTCTGTCTCCAGGTTCTGCTGGGGTTAGTGAGTG
GCTCACTCACTGGCTGGCTAGAGCGTTTTAGAAAGGCTGTTTTTTTTTT
TTTTTCATGTCCCCTGCTTTATATGTCTTTAAATGAGAACTGTGGAGA
GGACCTGAAACCTTCATTCACTCTGGTCTTTCTAAAAGCTAGTGGAGA
TGCTTGGCTGT

Expected size of amplicon = 386bp

Amino acid translation:

gaatgatttgcctggatgtgggtggctagcattcacctttccaagaggactgtccgaaggata
N D L L D V V A S I H L S K R T V R R I
cgcacacactggctcctggcactgattataacctgggtgggataccattgcagcaggt
R I N L V L A L I Y N L V G I P I A A G

Mutation covered in amplicon 15 = 1

AMPLICON NUMBER 16

Exon 21 : 4282 to 4235 in mRNA

tggatgagaggccttcaccaggcttagaaaaaaagccttgtttctagaatggctca
gatgctgttgcgcttccagCTATAAGAAGCCTGACCTGGAGAGGTATGAGGCA
CAGGCGCATGGCCACATGAAGCCCCTGACGGCATCCAGGTCAGTGTGCACATAGGCATG
GATGACAGGTGGCGGGACTCCCCAGGGCCACACCATGGGACCAGGTCAGCTATGTCAGC
CAGGTGTCTGCTGCTCCCTGACGTCCGACAAGCCATCTCGGCACAGCGCTGCAGCAGAC
GATGATGGGGACAAGTGGTCTCTGCTCCTGAATGGCAGGGATGAGGAGCAGTACATCtga
tgacttcaggcagggcgggcccggggcagggacttgcctccactcaccacaagctgagcagg
acagccagcagcaggatgggctgagctagcctocagcttggggacttccgctccctgga
tatgtccagtcactcctgcctgcagcagcggccttgtctgggtgcagctgggcttggcc
tggagaggacggcctgctgctcttggcctcacgggaccgtcagcatgggcttctgtct
tggactctagtccttggctggactgtagaaggtgagagggcagctcacctcctcacagac

Expected size of amplicon: 452bp

Amino acid translation:

tgctataagaagcctgacctggagaggtatgagggcacagggc
C Y K K P D L E R Y E A Q A
catggccacatgaagcccctgacggcatcccaggctcagtggtgcacataggcatggatgac
H G H M K P L T A S Q V S V H I G M D D
aggtggcgggactccccagggccacaccatgggaccaggctcagctatgtcagccagggt
R W R D S P R A T P W D Q V S Y V S Q V
tcgctgtcctcctgacgtccgacaagccatctcggcacagcgctgcagcagacgatgat
S L S S L T S D K P S R H S A A A D D D
ggggacaagtggtctctgctcctgaatggcagggatgaggagcagtacatc
G D K W S L L L N G R D E E Q Y I

Mutations covered in amplicon 16 = 1

Table IV A. List of PCR primers for 16 amplicons. 8 amplicons marked by * form set A and were amplified at elongation temp. 55 °C and set B of 8 amplicons were amplified at 60°C

Intron bps	Exon	ampli con bps	REVERSE (right) Primer-R	T _m	FORWARD (left) Primer-L	T _m
563	2a	692'	ATATCAATTGGTCCCAGGCTTAAG	59.9	TGCCAGAGAAGCTGGGATGTTGTA	59.9
---	2b	723'	CTCACCTATACCACCATCCAGGAG	59	CTTAAGCCTGGGACCAATTGATAT	59
82	4	291	ACAAAAACCAGACACGTCCAAGATGG	59.1	TGACCTGATGGTTCCAGGTGTTCT	59.7
8629	5	495'	TTCACTGATATCCTCCCTCAGATTA	60.3	AGTCCAGGGTCTTGAGAGCAGT	61.7
1161	7	496	GAAAGCTGCAATAAAGTGCCAT	59.8	ATGTTTAGACCTCTAGATGCTCCCT	61.0
1554	8	555	GCACCTTAATTATATGGAGGTTTCC	60.3	CAGTAGTCTCTGAATGGGAAAGTA	60.1
1992	9	392	GAGTGGTGATCTTACTGTGTCTCTG	59.4	CATGTGTGGTGGATAGCAAGTAAC	60.2
3183	10-12	916'	CTACTCTGGCTTAGATTTTGCTGTC	59.9	TGATAAGTGGCGTTTGTTCAGGG	62.2
7116	13	500'	GGCTACTCTGTTGCTACTGTTGTTA	59.1	TGTGGAATACCATCTGTTTCCG	62.0
703	14	419	TTTCCAGACCACACAGAGAAGGCT	60.2	TAGGAAGCTGTGCAGGTGCTTGT	59.9
1605	15	497	GAACATAAGAGAACTTTCCTGGGT	60.3	TTAACCTTTCCTATCTGTTCCACCT	60.6
1514	16	396'	CCTGAAATTAAGAGAGGAAGGCT	59.4	GGTGCTTACAAGGTTACAGTTTTTC	59.6
3410	18	479	AGGTTGATGCGTATCCTTCGGACA	59.8	GTTGACCAACATCACTGACTGG	60.5
1884	19	386'	ACAGCCAAGCATCTCCACTAGCTT	60.2	CTCACTGTGTGCTCGTCTCCATCA	59.8
36118	21	452'	CAGGGCAGGATGACTGGACATATC	58.7	GAATGGCTCAGATGCTGTTGCGTT	60

Table IV B. Nested Primers for 2nd round of PCR

NESTED (N) PRIMERS					
Exon	size (bps)	FORWARD (LEFT) PRIMER- LN	T _m	REVERSE (RIGHT) PRIMER- LR	T _m
5	208*	TCTTGGCTGCCTGTTACCTAGACT	59.3	ACTTAACAAGGGCTTTGCTGGTGG	59.7
7	288	ATCCAGGTGACAAGCAGCATCTGA	60.1	ATATCTGAGGGCCACACACAGCAT	60.2
8	312	GCACAAAGCTAGAGCTTTGCCAT	59.9	CCAGGGCAATGAACACAAAGAGCA	59.9
9	297	AGCTGTCTCTAACACCACGCTTGT	60.1	TCTGCCACACTCACAAGGTCTAT	59.5
11	209*	TGTCAGGTCACATGAGTGCTGGAT	59.9	TGAGCCTCTCCACCAGTTTCA	60.0
12	250*	CCCAATCTTTATCCATGCTTGTGGTG	58.0	ACAACCACCATATAGCCCAAGGCA	60.4

13	268*	ACAGAGGTGATCATCCGGTTTGCT	60.1	TGTCTTGAGTGGCTCTCAGGCTTT	60.0
14	256	AAGTTCTGCCTCAGGAGGTGACT	59.6	TACAGTATTGGTGACTGCCACGC	59.1
15	237	ACCTTCCCAGGAACTTGAACAGA	59.7	TCTGTGGTTTGACCCACCTCTACT	59.2
18	344	AAGTCTTGCAGAGGTGCTGCCTT	61.0	AGGTTGATGCGTATCCTTCGGACA	59.8

Table V. Size of WDDM Amplicons after 2nd round of PCR

S.N.	Round-I amplicon size in bps	Exon	amplicon size in bps	Primer pair	Forward Primer	Tm	Reverse Primer	Tm
1	692	2a	692	L+R	TGCCAGAGAAGCTGGGATGTTGTA	59.9	ATATCAATTGGTCCCAGGCTTAAG	59.9
2	72	2	723	L+R	CTTAAGCCTGGGACCAATTGATAT	59	CTCACATATACCACCATCCAGGAG	59
3	291	4	291	L+R	TGACCTGATGGTTCCAGGTGTTCT	59.7	ACAAAAACCAGACACGTCCAAGATGG	59.1
4	495	5	455	LN+R	TCTTGCTGCCTGTTACCTAGACT	59.3	TTCACTGATATCCTCCCTCAGATTA	60.5
5	496	7	288	LN+RN	ATCCAGGTGACAAGCAGCATCTGA	60.1	ATATCTGAGGGCCACACACAGCAT	60.2
6	555	8	482	LN+R	GCACAAAGCTAGAGGCTTTGCCAT	59.9	GCACCTTAATTATATGGAGGTTTCC	60.3
7	392	9	297	LN+RN	AGCTGTCTAACACCACGCTTGT	60.1	TCTGCCCACTCACAAGGTCTAT	59.5
8	900	10-11	547	L+R	TGATAAGTGGCGTTGTTGCAGG	62.2	GGGATAATCTCCTTCATTTAACCAC	60.3
9	900	12	334	L+RN	AAGAGTTCTGGGAATCAGACAGT	60.8	ACAACCACCATATAGCCCAAGGCA	60.4
10	500	13	431	L+RN	TGTGGAATACCATCTGTTCCG	62.0	TGTCTTGAGTGGCTCTCAGGCTTT	60.0
11	419	14	370	LN+R	AAGTTCTGCCTCAGGAGGTGACT	59.6	TTTCCAGACACACAGAGAAGGCT	60.2
12	497	15	393	L+RN	TTAACCTTCTCTATCTGTTCCACT	60.6	TCTGTGTTTGACCCACCTCTACT	59.2
13	396	16	396	L+R	GGTGCTTACAAGTTACAGTTTTTC	59.6	CCTGAAATTAAGAGAGGAAGGCT	59.4
14	479	18	458	L+R	GTTGACCAACATCACTGACTGG	60.5	ACAGTCTCTTGAAAGGTGAAT	60.4
15	386	19	386	LR	CTCACTGTGTCTCGTCTCCATCA	59.8	ACAGCCAAGCATCTCCACTAGCTT	60.2
16	452	21	452	L+R	GAATGGCTCAGATGCTGTTGCGTT	60	CAGGGCAGGATGACTGGACATATC	58.7

The Table VI gives details of WD indexed patients and relatives whose blood samples were processed and were taken up for hybridization on WDDMs. Blood samples from 3 females and 3 male healthy individuals were included in the study. Since presence of any WD mutation in patient sample is uncertain, site directed mutagenesis was done of five amplicons generated from commercial preparation of human chromosomal gDNA and six type of mutations were specifically generated

to simulate few frequently found WD mutations. These would act as standard for confirming differentiation ability of WDDM.

Table VI. Code numbers for WD indexed patients and their relatives

(B is for brother, S is for sister, in Sibs)

Code of patient	Sex/Age	Sib1 Sex/ Code	Sib2 Sex/Code	Mother	Father	Total
P1	F/32	B / P1-3		P1-2		3
P2	M/45	B / P2-3				2
P5	M/49	B / P5-3	S / P5-4			3
P6	M/24					1
P7	M/39					1
P8	F/15					1
P9	F/11	B / P9-3		P9-2	P9-1	4
P10	M/9					1
P11	F/22		S / P11-3 is WD patient	P11-2		3
P12	F/7		S / P12-3	P12-2	P12-1	4

2.5 SITE DIRECTED MUTAGENESIS

Following Six different types of WD causing mutations were selected for mutagenesis.

DBI – Double base insertion of GT in exon 4

DBD – Double base deletion of GT in exon 7

SBI – Single base insertion of C in exon 8

SBD – Single base deletion of C in exon 9

Two Point Mutations – SPM1 (G->A) 15a

and SPM2 (T->C) 15b in exon 15

Table VII. Primers for Site Directed Mutagenesis (SDM)

Exon / bps	SDM a&b Fragments	Type	Mutated Middle Primers (MP)
4 / 291	113,207 bp	DBI	CAGTAAGTACTGT <u>GT</u> GGGTGC GTTACG CGTAACGCACCCC <u>CA</u> ACAGTACTTACTG
7/ 288	238,78 bp	DBD	TTGTGTACCTTT– <u>CC</u> CAGGTATATATG CATATATACCTGG–AAAGGTACACAA
8/ 482	253,255 bp	SBI	TGGTTGCTGTGGC <u>C</u> TGAGAAGGCGGA TCCGCCTTCTCAG <u>G</u> CCACAGCAACCA
9/ 297	116,207 bp	SBD	TAGAGCAAAA–CTCAGAAGCCCTGG CCAGGGCTTCTGAG–TTTTGCTCTA
15a/ 393	236,182 bp	(G->A) SPM1	GTGCCAGGCTGT <u>A</u> GAATTGGGTGCA TGCACCCAATT <u>C</u> TACAGCCTGGCAC
15b/ 393	236,182 bp	(T->C) SPM2	GTGCCAGGCTGTGGAA <u>C</u> TGGGTGCA TGCACCCAGTTCCACA <u>G</u> CCTGGCAC

The details of protocols for Site Directed Mutagenesis are given in Section 2.14.

Sequencing of all mutagenised amplicons were done using the DNA sequencing facility of MBD.

2.6 PRINTING OF WDDM WITH OLIGO PROBES

Oligoarray 2.0 program was used for optimizing the length and sequence of the probes keeping the mutations in the centre of the probe sequence. Entropy, free energy, GC content, T_m, length and enthalpy for DNA duplex were optimized for each ssDNA probe sequence. Those ssDNA sequences were selected as probe sequences that could have mutation in the middle and total length between 24 to 32 bases. For each probe its wild type and mutated forms were synthesized in 50nMol scale with HPLC purified quality from commercial sources. Oligos, had available amine groups that would covalently attach to the epoxide ring at high pH. WDDM were printed on epoxy coated glass slides, for this purpose. Covalent linkage with the amine groups present on the synthesized oligonucleotides are shown in Figure 2.4. The printing was done using the in-house microarraying⁶² robot that was developed at BARC, Trombay during Xth Plan period. The micro-arrayer shown in Figure 2.5 has 3-axis (X,Y,Z) robotic system with positional accuracy of 1 micron. The spots formed were of diameter 100 to 150 μm and more than 6000 spots per slide can be printed with this microarrayer. Using this microarrayer, 20 to 30 slides were printed in one to two hours, maintaining 200 to 250 probe solutions in correct well plate format. The format of printing low density microarray WDDM for detection of WD mutations is shown in Figure 2.6. Hybridisation experiments were done on printed WDDMs with different samples of WD amplicons at different hybridizing temperatures. Hybridised WDDMs were scanned and referred as HWDDM for analysis purpose.

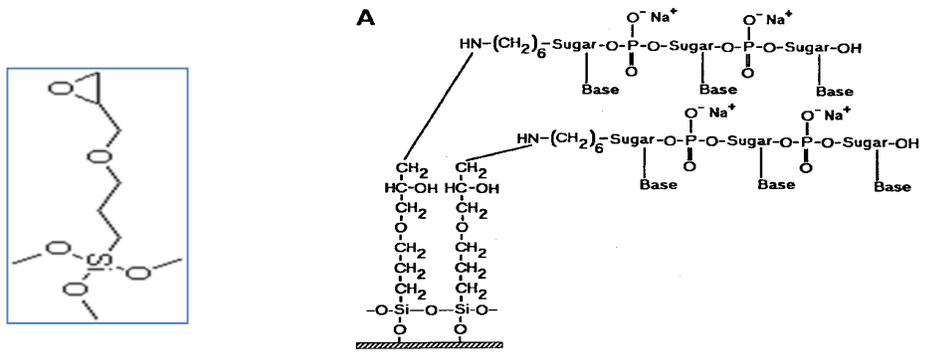


Figure 2.4 Bond formation by epoxy coating on the surface of glass.



Figure 2.5 Micro-arraying robotic machine used for preparation of WDDM

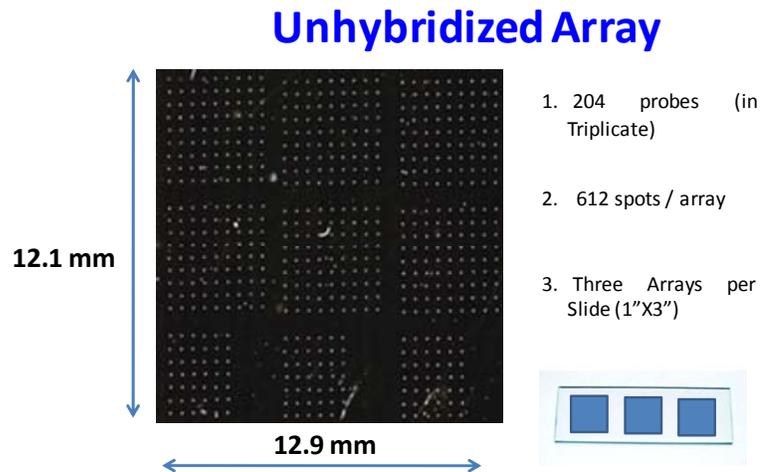


Figure 2.6. Format of printing WDDM on epoxy-coated glass slide.

The list of probes included in the WDDM and the nature of mutation covered by them are shown in Table III. These arrays were printed with total of 204 probe solutions, each in triplicate. Out of 204 probe triplets on each WDDM, 84 were WT probes (these include 63 Indian WD mutations, 4 SNP variants reported among these exons, repeated spots and reverse complements of probes) 93 were mutant probes, 20 exon specific control probes, 5 blanks (only buffer as negative control) and 2 were DIG labeled (as positive control). The printing was done asymmetrically to enable identification of probe pairs on HWDDMs by the pattern developed after hybridization. Dimension of one WDDM is 12.1 x 12.9 millimeter.

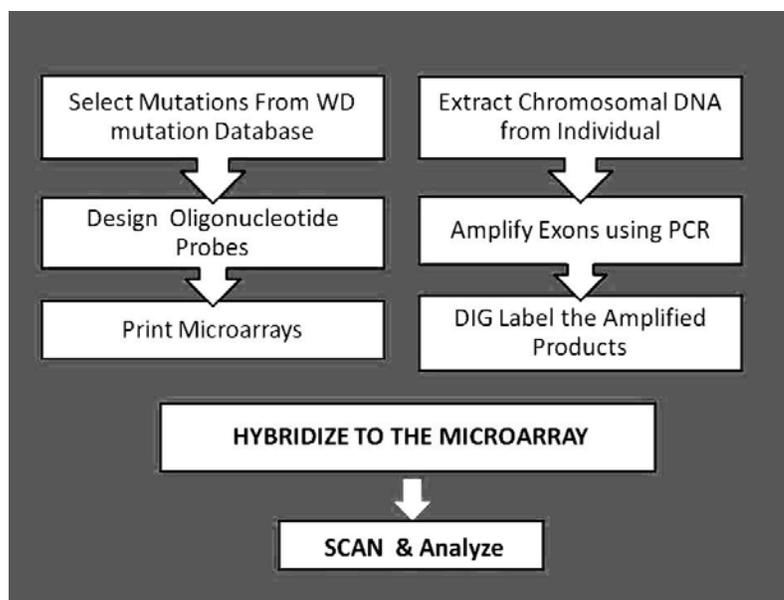


Figure 2.7. Brief schematics of methodology for WD mutation detection.

2.7 GENOMIC DNA ISOLATION FROM PERIPHERAL BLOOD

From practical and commercial point of view QIAGEN DNA isolation kit was used initially following the standard protocol provided by the QIAGEN. However, soon it was realized that for whole genomic DNA (gDNA) isolation from peripheral blood for PCR purpose, non enzymatic method of Lahiri and Numberger (1991)⁶⁰ is less expensive and straight forward.

PROCEDURE:

Peripheral Blood was collected (~10 ml) in vacutainer tubes containing 100µl of 15% EDTA. Equal volume of TKM1 buffer was added followed by addition of 125µl Nonidet P-40. Cell lysis was achieved by gently inverting the centrifuge tubes till RBCs were fully lysed. This mixture was centrifuged at 2200 rpm for 10 minutes at room temperature. The supernatant was slowly poured off saving the nuclear pellet. The pellet was resuspended in 5ml of TKM 1 and centrifuged as before. The pellet after washing should be devoid of red debris from RBCs. This

pellet containing cellular nuclei was gently suspended in 0.8 ml of TKM 2 buffer. Fifty microliter of 10% SDS was added to the nuclear suspension for releasing and denaturing nuclear proteins. A micro-pipetter was used for pipetting back and forth several times and this mix was incubated at 55°C for 10 minutes. 300 µl of 6M NaCl was added after 10 minutes and was mixed well by gentle shaking. This mix was centrifuged at 12000 rpm for five minutes in a micro-centrifuge. The cell debris pelleted at the bottom and genomic DNA remained in the supernatant. After transferring the supernatant in another tube, double volume of 100% ethanol at room temperature was added. This tube was inverted several times until the genomic DNA precipitated out of the suspension. This precipitated DNA was removed to another micro-centrifuge tube containing 1ml of ice-cold 70% ethanol. Micro-centrifugation at 12000 rpm at 4°C was done for 5 minutes to separate out DNA from remaining reagents and salts. This DNA pellet was air dried and re-suspended in 0.5 ml of Tris buffer by keeping at 65°C for 15 minutes. All DNA samples were stored at -20 °C after categorization.

Quality Check for gDNAs:

Optical density at 260 nm and 280 nm is a measure for assessing the protein contamination in isolated DNAs. A ratio in the range of 1.5 to 2.0 of A_{260}/A_{280} is acceptable for executing Polymerase Chain Reaction (PCR) on samples of isolated gDNAs. Optical density ratios were taken to ensure the quality of gDNAs. The concentration of DNA solution was estimated using the fact that OD (260nm) of 50ng/ µl is one. gDNA in eppendorf tubes were labeled in code numbers as in Table VI and stored at -20°C .

2.8 SOFTWARE USED

For designing of PCR primers Primerquest program^{wr10} was used. The 5' to 3' DNA sequence segments from the *Atp7B* gene were used as input sequence for calculation of T_m and GC content of the left and right primer sequences. The location of left and right primer sequences with similar T_m was most critical, such that PCR amplification of WD causing mutations could be facilitated. The sequence details of 16 amplicons selected for testing on WDDM are given in 16 different pages. The list of forward (Left) and reverse (Right) primers used for preparing 16 amplicons are given in the Table IV A. For Re-PCR nested primers were also designed for 10 amplicons. (Table IV B). For these 6 amplifications, the diluted amplicons from the first round were used as templates.. Remaining 6 amplicons were re-amplified with the same initial primers and using initial amplicons as templates. For 8 amplicons of set A elongation temperature was kept 55°C and 60°C for set B. In the second round of PCR for all the Gdna templates amplicons were amplified in 2 sets and in well plate format, while 1st set of PCR were done individually.

2.9 AMPLIFICATION OF DNA BY PCR FROM gDNA SAMPLES

Primers were synthesized at BRIT (Vashi, Mumbai) and were obtained from other commercial sources. Reagents for Polymerase Chain Reaction viz. Taq polymerase enzyme, buffer, dNTPs were obtained from JONAKI, Hyderabad. The protocol for the reaction is given below for 25µl reaction. It was ensured that in 25 µl PCR reaction mix approximately 200ng gDNA was present as template.

Water	8.3 µl
10X PCR Buffer	2.5 µl
2mM dNTPs	2.5 µl
0.1% BSA	2.5 µl
Forward Primer	2.0 µl

Reverse Primer	2.0 μ l	
Enzyme	0.20 μ l	
Template	5.0 μ l	(Reqd DNA is 200ng in 25 μ l rxn mix)

Total vol	25.0 μ l
Mineral oil	10.0 μ l

PCR Conditions:

Initial Denaturation	94 ^o C / 5 min.	} 35 cycles
Denaturation	94 ^o C / 30 sec.	
Annealing	65 ^o C / 30	
Extension	72 ^o C/1 min	
Final Extension	72 ^o C/7 min	
If required to hold	4 ^o C	

2.10 AGAROSE GEL ELECTROPHORESIS OF PCR PRODUCTS

All the PCR products were visualized by loading on ethidium bromide stained agarose gels (1 to 2% concentration). Electrophoresis was done in a field of approximately 10 V/Cm for nearly 1 hour along with a 100bp ladder DNA as a molecular weight marker. Five μ l to 10 μ l volume was loaded after mixing PCR product with commercial loading dye. The gels were viewed on trans-illuminator and documented using ‘Gel Doc System-Syngene’ with GeneTool software.

2.11 RE-PCR OF 16 EXONS INCLUDING NESTED PCR

The quality of mutation detection by microarrays depended upon the quality of labeled DNA. To minimize the carryover of non-specific DNA in the PCR products a re-PCR of all the 16 amplicons was done. In addition, this second round of PCR helped in increasing the yield of the amplicons. Ten of these reamplifications were conducted with primers nested within the previous amplicons. Nested Primers are given in Table IVB (Section 2.5). Remaining 6 amplicons were re-PCRred using the previously used primers.

The re-amplifications from 18 gDNA samples were carried out on microtitre plates using a liquid handling system EpMotion for quick pipetting of PCR reagents.

Amplification of 8 amplicons were conducted per 96 well microtitre plate at 550C and other set of 8 amplicons at 600C. All the amplification products were checked by agarose gel electrophoresis.

2.12 CLEAN-UP PROTOCOL OF PCR PRODUCTS

QIAGEN Spin column kit was used in which the spin column has silica gel membrane. DNA adsorbs to the silica membrane in the presence of high salt while all other materials pass through the column. Impurities are washed away and pure DNA is eluted with Tris buffer.

Material used-

- Binding Buffer PB –
- Wash buffer PE-
- Elution Buffer EB- 10mM Tris-Cl, pH-8.5

Procedure-

1. Add 5 volumes of Buffer PB to 1 volume of the PCR sample and mix
2. Place a QIAquick spin column in 2ml collection tube (provided with the kit)
3. Apply the sample to the QIAquick spin column and centrifuge for 1min
4. Discard flow through
5. Place the QIAquick spin column back into the same tube
6. Add 0.70ml buffer PE to the QIAquick spin column and centrifuge for 1min
7. Discard flow through and place the QIAquick spin column back in the same tube and centrifuge the column for additional 1 min
8. Place QIAquick spin column in a clean 1.5ml microcentrifuge tube

9. Add 30ul or 50ul elution buffer EB to the centre of the QIAquick membrane and centrifuge the column for 1 min . For increased DNA concentration, add 30 ul elution buffer to the centre of the QIAquick membrane and let the column stand for 1 min and then centrifuge for 1 min.
10. Store eluted DNA in -20 °C or 4 °C until needed

2.13 PROTOCOLS FOR SITE DIRECTED MUTAGENESIS

Commercially procured human gDNA (Ch DNA) was used for site directed mutagenesis. 1st round of PCR products of exons 4, 7, 8, 9 and 15 were generated in 50µl volumes of reaction. The primers used for site directed mutagenesis alongwith exons and type of mutation are given in Table VII . 2nd Round of PCR products were generated in 100 µl using 100 times diluted template from 1st round of amplification. These PCR products were cleaned by PCR clean-up kit described earlier and amplification products were checked by agarose gel electrophoresis.

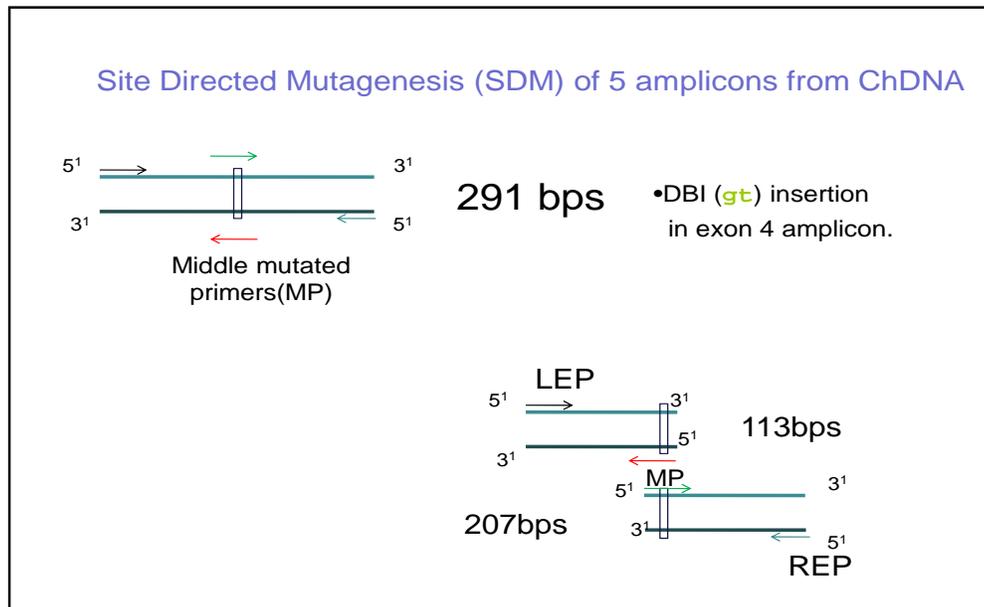


Figure 2.8 Schematic of site directed mutagenesis from WT amplicon.

The strategy used is depicted in this Figure 2.8 taking example of one of it i.e. DBI. The same strategy was used for all the products using their respective primers. To get good yield of SDM amplicons from WT amplicons, following procedures were adopted:

First step of experiments was to generate two short fragments using one mutated primer and one normal primer. Mutated primer concentration needed to be more because of reduced binding to template. After trial and error with PCR reactions the final methodology that gave good result was the following:

1) PCR protocol for **two short mutated fragments:**

PCR purified fragments of exon 4,7,8,9 and 15 were diluted three times 2:200 to generate final dilution of 1:100,00,00 (a million fold dilution). Primers were added in skewed ratio of 6:1 of mutated :WT. In the PCR mix mutated primer was 3 μ l while normal primer was 0.5 μ l. PCR program was modified from that of amplicon PCR program as follows:

10 cycles at Annealing temp: 40°C, Extension time: 60 sec.

+ 25 cycles at Annealing temp: 52.5°C, Extension time : 60 sec.

2) Extension PCR to Generate single Amplicon from two short fragments:

Ten microliter of each of the two short fragments from above PCR reaction was taken for extension PCR where 1 μ l dNTPs , 0.5 μ l enzyme and 2 μ l BSA were added for additional 10cycles of reaction. Complete schematics adopted is also given in Figure 2.9. From this, 1 μ l was saved to study dynamics of these SDM PCR reactions.

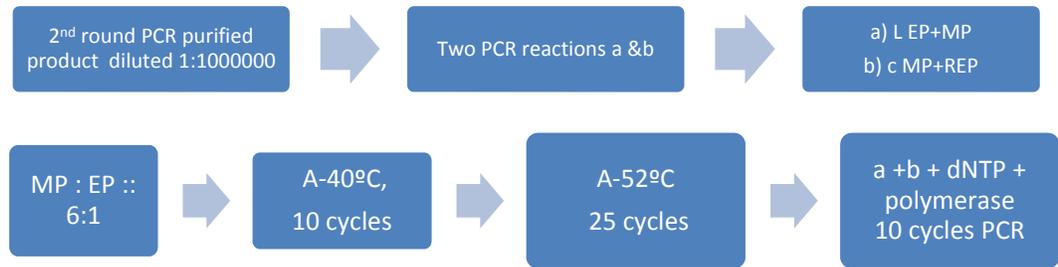


Figure 2.9 Methodology used for Site Directed Mutagenesis

3) Kinetics of PCR for Site Directed Mutagenesis:

-ve control was 1 µl from 2nd step before extension PCR.

+ve controls were after extension PCR.

These two were used as template after 1:1000 dilution.

In the 3rd step, end primers were used in 1:1 ratio.

Standard PCR program was followed.

Aliquots were taken out at 20th, 25th, 30th and 35th cycles during PCR for checking of products (referred W, X, Y, and Z from +ve control and referred W', X', Y', and Z' from -ve control) by electrophoresis. This step was to find kinetics of PCR such that PCR artifacts were not included and increased yield of mutagenised SDM fragment could be detected. A comparative study of yield of all these PCR products from +ve and -ve controls was made before selecting the final product for hybridization. Electrophoresis for these samples were done in 18 well gel apparatus in three sets of gel to accommodate 8 (W, W', X, X', Y, Y', Z, Z') aliquots from each of six SDMs alongwith 100bp DNA ladder as molecular weight marker.

Simulated WD mutation Samples

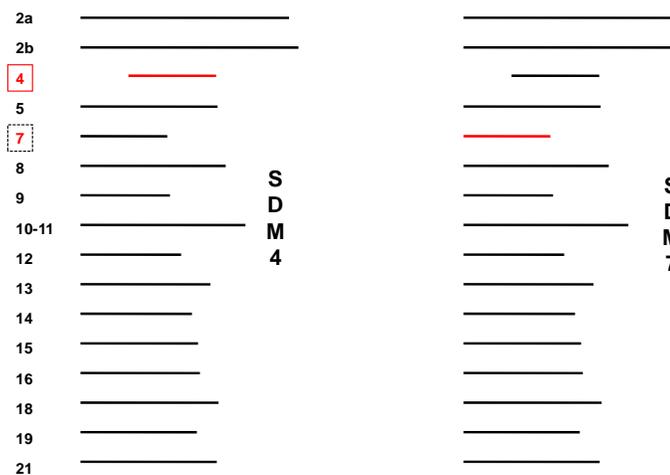


Figure 2.10. Sixteen amplicons pooled for simulating SDM4 and SDM7.

To test WDDM with a specific mutation, out of 16 WT amplicons one was replaced by its respective SDM as shown in Figure 2.10. Presence of mutated allele in SDM fragments was confirmed by sequencing (Sequencing electroporegram given in Results Section).

2.14 PRINTING OF MICROARRAYS

WDDM on epoxy coated slides were prepared by printing Oligo probes with the help of robotic system developed at BARC (Figure 2.5). The robot was programmed so that successive spots are spaced 0.8mm apart, a separation sufficient to avoid contact with the adjacent spots. Epoxy silane slides function by offering an epoxide ring that reacts with an amine group on the spotted material (ssDNA).

Spotting buffer SSC (saline-sodium citrate) is commonly used in hybridization as well as in post-hybridization washes. Hybridization is typically carried out in high ionic strength buffers to maximize the rate of annealing of the probe with its target. A decreased non-specific hybridization will occur in the presence of lower ionic strength buffer. The chemical and physical properties of the solvent greatly

influence DNA retention, spot morphology, and hybridization efficiency. A 384-microtiter plate was prepared for printing of the probes. Ten microliter of the 100 μ M probe was diluted using 32.5 μ l of St.Milli-Q water and 7.5 μ l of 20X SSC to obtain a 3X SSC concentration and a final volume of 50 μ l. While preparing the microtiter plate for the arrayer machine, it was ensured that air bubbles were not introduced in the well. The relative humidity was set to ~50% within the arrayer chamber. During the printing cycles, the pin heads picked up the solutions from the well plate, deposited them on all the epoxy slides, cleaned the pins in the ultrasonication cleaner, dried them by placing the pins on a vacuum chamber before going for next set of probe solutions in the well plate. The different time settings used during the printing cycle are shown below. It is to be noted that the time spent by the pin while picking-up the sample from well-plate is called (dwell-time).

- | | | |
|----|--------------------|----------|
| 1. | Ultrasonic washing | 15 (sec) |
| 2. | Vacuum dryer | 15 (sec) |
| 3. | Dwell time | 2 (sec) |

After printing the slides were kept in a high-humidity (about 75% RH) environment for 15 to 20 hours to ensure Immobilization of the spotted nucleic acids on the epoxy slides. This humidity was achieved by placing saturated ammonium sulphate solution inside an open petri plate kept in a desiccator. The printed slides were incubated overnight in one such desiccator and were later stored under vacuum until further use.

2.15 DIG LABELING OF AMPLICONS

Random primed labelling is the most widely used method for generating homogenously labeled DNA probes. In this labelling method the Klenow polymerase enzyme copies DNA template in presence of hexameric primers and alkali-labile DIG-dUTP. On average, the enzyme inserts one DIG moiety in every

stretch of 20-25 nucleotides. Pool of amplicons were prepared to have approximately 50 ng/ μ l DNA concentration and were DIG labeled as follows:

Protocol:

- Added Sterile Milli-Q water to the pool to make up 16 μ l volume.
- Heated this sample of DNA at 95°C for 5 minutes. This step was important for efficient labelling. Immediately kept on ice.
- After short spin added 4 μ l of DIG High Prime (Roche) Labelling mixture
- Incubated overnight at 37 °C.

2.16 HYBRIDISATION ON WDDMS

1) Pre-Hybridization treatment of printed array:

Before hybridization, the slide was incubated at 42°C with 20 ml of pre-warmed (to 42°C) DIG Easy-Hyb solution for 30 minutes with gentle shaking to remove excess/unbound probe. Six microliter of the labelled sample was heat treated at 95°C for 5 minutes and was chilled on ice for 10 min. After spinning the tube, 94 μ l of DIG Easy-Hyb solution was added to it and was used for setting up a hybridization.

2) Hybridization:

During hybridization, the labeled DNA would anneal to the complementary probe strand that is immobilized on the slide. After the pre-hybridization, the slide was removed from the solution and the excess liquid on it was removed by placing it on a clean lint-free tissue paper. Immediately, the hybridization mixture (about 100 μ l) containing the labeled DNA sample (see above) was carefully placed towards the edge on the printed side of the slide. A clean and dry cover-slip was carefully placed on the slide in such a way that it spreads the

hybridization solution evenly and without air bubbles across the area of interest. It must be ensured that the slide does not dry up during these steps. The slide was then placed inside a petri dish which in turn was placed in an airtight box containing wet tissues papers, for maintaining a humid environment during hybridization and to prevent the slides from drying up during hybridization. The box was then incubated overnight at 42°C in a hybridization oven.

3) Post-hybridization processing of the arrays

Even distribution of target and wash solutions over the array is the crucial element of successful hybridization. After hybridization, the non-specific binding sites (for the anti-DIG antibody) on the slide were blocked by the blocking buffer. Blocking methods provides the added advantage of washing away the unbound label from the surface that would otherwise compete with the labeled species, and helped in decreasing the background. Probe-target hybrids were detected with an enzyme-linked immunoassay.

4) Antibody solution:

Detecting the probe involved an anti-Digoxigenin-AP (alkaline phosphatase) conjugate antibody and the alkaline phosphatase chromogenic substrate. The antibody solution was centrifuged for 5minutes at 10,000 rpm in original vial prior to use. The Anti-Digoxigenin AP was diluted (1:250 to 1:500) in 1X blocking solution and used.

5) Post Hybridization Washes:

After allowing the hybridization reaction to proceed for a specified period of time, the following steps were carried out.

1. 2X SSC, 0.1% SDS was pre-warmed at 42°C

2. 10 ml of the above solution was poured over the slide and the cover-slip was allowed to float off
3. 10 ml of 2X SSC, 0.1% SDS was added and the slide was incubated with shaking for 5 minutes at 42°C
4. The slide was incubated in 10ml of pre-warmed 0.5 X SSC, 0.1% SDS for 5 minutes with shaking at 42°C
5. The slide was washed in 0.5X SSC, 0.1% SDS for 5 minutes with shaking at room temperature.
6. Washing buffer (Maleic acid buffer with 0.3% Tween20) was added for 5 minutes with shaking at room-temperature
7. The washing step (Step 6) was repeated.

2.17 DETECTION OF PROBE-TARGET HYBRIDS BY CHROMOGENIC METHOD

1. The slide was covered with approximately 1ml of 1X Blocking buffer and incubated at room temperature for 15 minutes.
2. The anti-body solution was prepared by spinning the anti-body conjugate vial for 5 minutes at 10,000 rpm
3. The anti-body (1:250) was prepared in 1 X blocking buffer.
4. Blocking solution was drained off from the slide and the slide was tapped on a lint- free tissue to remove the remaining blocking solution.
5. One ml. of diluted antibody solution was added to the slide and incubated for up to 30 minutes at room temperature.
6. The slide was washed with washing buffer for 10 minutes at room-temperature.
7. This step was repeated for 5 minutes at room-temperature.

8. One ml. of detection buffer was added to the slide and the slide was incubated for 5 minutes.
9. The slide was incubated with chromogenic (NBT/BCIP) substrate for 20minutes or longer (till color development)
10. The slide was washed in pre-warmed de-ionized water, dried immediately using a blower and stored at RT within petri dish, which is kept in a clean box for scanning.

Scanning the slide:-

After hybridization, the signal intensities of all the spots on WDDM were captured by flat-bed microarray scanner in the transparency mode. The arrays were scanned at 600 dpi resolution and 100 to 400 % scaling in a HP scan-jet model G4050.

2. 18 QUANTIFICATION OF SPOT INTENSITIES

Quantification of each of the spots was carried out by using the software Image J (version 1.40), a free general purpose image processing package from NIH (<http://rsb.info.nih.gov/ig/>). A plug-in, developed by Bob Dougherty and Wayne Rasband, that facilitates the analysis of microarray images (MicroArray Profiler) was downloaded and was used for carrying out the analysis of microspots on WDDM. The plug-in allows the user to define a grid of circles (Figure 2.11), which can be moved by dragging with the mouse.

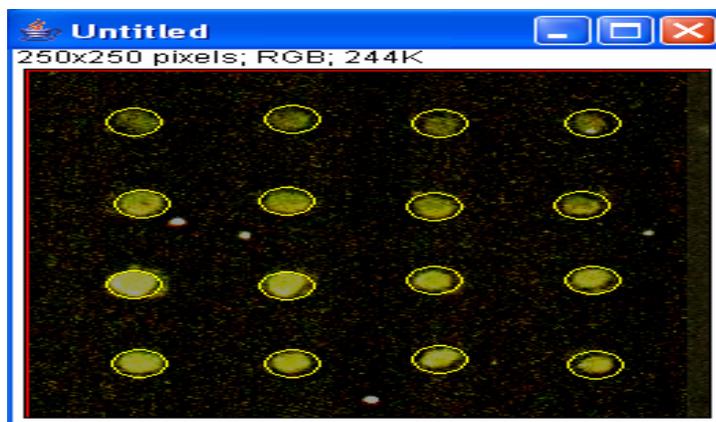


Figure 2.11 Grid for microspots

The final grid can be stored and recalled. The micro-spot intensity, which is mean grey scale value obtained using the grid, is saved in excel sheet columns having rows for each spot representing its coordinates on the HWDDM (Hybridized WDDM). Background correction of spot intensities was done using the statistical tools available within program Microsoft Excel and visual verification of scanned images of each of the HWDDM. Outliers were not considered as good data and were not included for next step of analysis described below to find probe differentiability.

2.19 INTRA-ARRAY AND INTER-ARRAY CORRELATION STUDY OF INTENSITY VALUES

HWDDM with SDM samples were generated for same Hybridisation temperature but different post hybridization washes (42 °C and 52 °C) to study the intensity variability of microspots on HWDDM. Pearson product moment correlation coefficients were calculated among intensity data for each SDM **using MS Excel tools.** PEARSON PRODUCT MOMENT CORRELATION COEFFICIENT developed by Karl Pearson also known as PPMCC is a measure of the linear correlation (dependence) between two variables X and Y , giving a value between +1 and -1 inclusive, where 1 is total positive correlation, 0 is no correlation, and -1 is

total negative correlation. It is widely used in the sciences as a measure of the degree of linear dependence between two variables. It is defined as given in the following formula:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Here X_i refers to intensity values from HWDDM at 42°C (post hybridization wash temperature) and Y_i refers to corresponding intensity values from HWDDM at 52°C (post hybridization wash temperature). Standard deviation was calculated and scatter of intensity data was quantitated using MS excel tools.

2.20 DIFFERENTIABILITY OF WT & ITS CORRESPONDING MUTANT PROBES

The mean intensity data of each spot for each HWDDM after background correction were stored in Microsoft Excel. For each HWDDM, these values were normalized with respect to average of intensities obtained at the spots where DIG labeled probes were spotted. All the probes were spotted in triplicate (WT as well as the mutant). The mean normalized intensity at the perfect match (WT probe) and mismatch (mutated probe) spots was used to compute DS (Differentiability Score). The **Discrimination Score Index (DS)** is independent of background correction and normalisation methods applied on the quantitative values of image intensities. This is most common index for Gene Expression Microarray Analysis (Zhou & Abagyan, 2003)⁶¹ but is also used for genotyping Microarrays. The detailed logic and statistics is given in Online Computation Biology Textbook at www.compbio.pbworks.com^{wr13}.

DS for each probe was calculated using the following formula:

$$\mathbf{DS} = (\mathbf{IP} - \mathbf{IM}) / (\mathbf{IP} + \mathbf{IM}) \quad ; \quad \mathbf{0} < \mathbf{DS} < \mathbf{1}$$

Where **IP** is average intensity value from triplet spots at WT probe (perfect- match).

IM is average intensity value from triplet at its mutant probe (mismatch).

IP is greater than **IM** (**IP > IM**), if hybridisation is done with WT amplicons.

Probe wise DS values were statistically analysed for their reproducibility by inter array and intra array comparisons. The DS value is dependent on the nature of hybridisation which is dependent on the probe sequence and post hybridisation washing temperature. The possible effect of the strand (sequence composition) on detecting the mutation was also addressed in our study by including 4 probes with their complementary probes for DS computations.

For HWDDM generated by hybridisation with SDM samples, (**IP < IM**) IP should be and is observed to be less than IM since the hybridisation sample has the mutant allele. Mathematically DS becomes negative if **IP < IM** (from definition itself).

DS values were classified and analysed with respect to the type of mutation.

2.21 SEQUENCING PROTOCOL

All the SDM amplicons and few of patient amplicons were sequenced. Sanger's dideoxy sequencing reaction, was carried out using Big Dye Terminator Ver 3.1 Cycle Sequencing Kit as per manufacturer's protocol. Products were loaded on ABI prism 377-18 DNA sequencer. Using forward and reverse primers two separate PCR reactions per DNA sample were generated using following program:

Program for cycle sequencing:

96⁰C for 10 sec. }
50⁰C for 5 sec. } 25 cycles
60⁰C for 4 min. }

4⁰C hold until ready to purify

Un-incorporated dye terminators were removed by ethanol precipitation, before loading on the sequencing machine to minimize noise on sequence data.

2.22 BUFFERES USED

All the chemicals used in this study were of Molecular Biology Grade supplied by MERCK (Germany) or SIGMA (USA).

BUFFERS FOR ISOLATION OF DNA FROM PERIPHERAL BLOOD

(Section 2.8)

1. TKM1 - 10mM Tris-HCl (pH 7.6); 10mM KCl, 10mM MgCl₂ and 2mM EDTA
2. TKM2 - 10mM Tris-HCl (pH 7.6), 10mM KCl, 10mM MgCl₂, 0.4 M NaCl and 2mM EDTA
3. TE - 10mM Tris-HCl and 1mM EDTA (pH 8.0)

20X SSC (SALINE & SODIUM CITRATE) BUFFER FOR SPOTTING

(Section 2.14)

NaCl - 175.3 gm (3.0 M)

Sodium citrate - 88.2 gm (0.3 M)

Dissolved in 800 ml of distilled water. The pH was adjusted to 7.0 with a few drops of 10 N NaOH and then the volume was adjusted to 1000 ml with distilled water. The solution was sterilized by autoclaving and stored at room temperature. Further, the solution that was used for printing the microarrays was filter sterilized with the help of 0.2 micron filter discs.

HYBRIDIZATION BUFFERS (Section 2.17)

1. SDS - 20%

20 g of SDS dissolved in 100 ml of sterile Milli-Q water. Store at RT.

2. Maleic acid buffer – 10X

Maleic acid (1.0 M), NaCl (1.5 M) Adjust pH to 7.5 with NaOH.

Autoclave and store at RT.

3. Washing buffer: Maleic acid Buffer -1X, Tween 20 - 0.3% (v/v)

Stored at room temperature.

4. Blocking solution : Freshly prepared by diluting 10X blocking stock solution (vial number 6 from the Roche kit) 1:10 with 1X Maleic acid buffer.

5. Detection buffer: Tris-HCl (0.1 M), NaCl (0.1 M) Adjust pH to 9.5 with NaOH. Stored at room temperature.

6. Substrate solution for colorimetric detection of DNA hybrids: One Tablet of NBT/BCIP dissolved completely in 10 ml sterile Milli-Q water. This is a colorless solution having 5-bromo-4-chloro-3-indoyl phosphate (BCIP) and nitro blue tetrazolium chloride (NBT).

“Peace comes from within. Do not seek it without it.”

CHAPTER III

RESULTS AND DISCUSSION

The microarray approach takes advantage of the fundamental tool in molecular genetics i.e. nucleic acid hybridization which is DNA base pairing between two complementary single-stranded nucleic acid molecules to form double-stranded molecules. It involves mixing of nucleic acid molecules from two sources, a probe which consists of a homogenous population of probe molecules (chemically synthesized oligonucleotides) and a target which consists of heterogeneous population of nucleic acid molecules (amplicons of genes). Hundred percent hybridization commands high degree of base complementarity achieved by providing optimum annealing temperature, percent GC in sequence and concentration of salt in wash buffer. The rationale of the hybridization assay is to examine the presence or absence of *de novo* mutations in the sample within the region of known probe sequences.

In the present study hybridization of Digoxigenin (DIG) labeled nucleic acid sample to immobilized Perfect Match or Mismatch probes (probe-target heteroduplex) and subsequent detection of hybrids by anti-DIG antibody was used. To develop a microarray based method for the detection of Wilson Disease causing mutations, the mutations that were reported in Indian population were extracted from literature. Details of all these mutations are listed in Table III (Section 2.4). As discussed earlier genomic regions surrounding these mutations were analysed with the help of the software OligoArray 2.0 to design perfect match (Probe_w) and its mutated pair (Probe_m) also called mismatch probe to address these mutations.

Designed probes were synthesized from a commercial source and were printed on epoxy coated slides with the help of an indigenously developed Microarrayer system. The quality of these Wilson Disease Diagnosis Microarrays (WDDM) was evaluated by examining the inter and intra array consistency of results after hybridization with DIG labeled amplicon samples. Their ability to detect / differentiate mutations was assessed by the use of experimental and statistical approaches described in this chapter. Few of the results of patient studies are given in section 3.14 and 3.15.

3.1 GENOMIC DNA ISOLATION

Human chromosomal DNA from about 5 ml of peripheral whole blood was isolated using a protocol mentioned in Section 2.10. Quality of the preparations were assessed by estimating OD₂₆₀/OD₂₈₀ on a NanoDrop (Thermus Fisher)®. Average 2 to 4 µg total DNA was obtained per sample. Approximately 100 nano gram (ng) DNA was used per PCR amplification reaction.

3.2 PCR AMPLIFICATION, Re-AMPLIFICATION INCLUDING NESTED PCR

Figure 3.1 is a representative agarose gel electrophoresis from a large number of PCR reactions done at the beginning of this project. Primers for PCR, temperatures for primer annealing, quantities of templates were suitably modified till desired specific PCR products were obtained. The primer combinations and annealing temperatures used were as in Table IV and Table IV A (nested PCR primer Table), protocol and PCR cycles used were as given in Section 2.10. Gel electrophoresis protocol was followed as in Section 2.11. This was very important to take the project further, since out of 3.4 billion base pairs of hgDNA, the probability of non specific amplifications is very high, inspite of best primers and best PCR conditions.

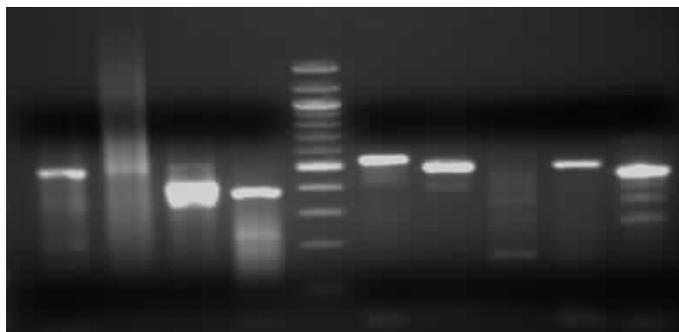
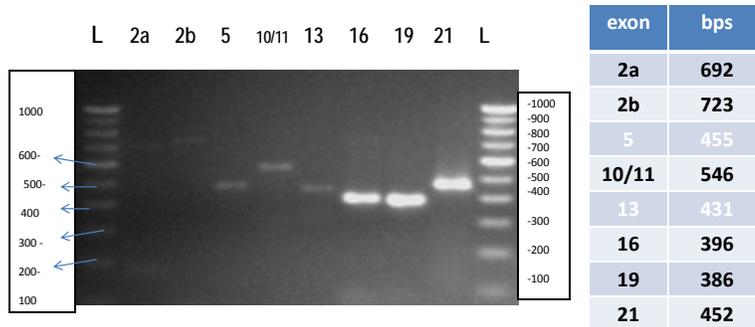


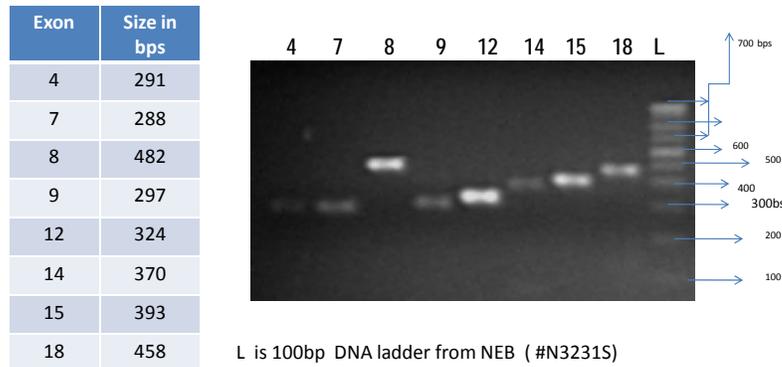
Figure 3.1 Agarose gel 1.8% with 100 bp DNA ladder. Gel stab of exon 5 (Lane1) , 7(Lane 2), 8(Lane3), 9(Lane4) &11(Lane6) and reamplified samples of exon 3(Lane7),14(Lane8),15(Lane9) & 18(Lane10). Lane 5 is 100bp DNA ladder.

To minimize the carry over contamination of the genomic DNA and the non-specific PCR products in the subsequent steps of mutation detection on WDDM, either nested PCR or re-amplification of all the 16 amplicons was carried out. Details of primers and amplicon sizes after the second round of amplification are given in Table V (Section 2.10). After optimization of all the PCR parameters, 16 amplicons were divided in two sets of eight amplicons each. The division of 16 amplicons in Set-A (55 °C) and Set-B (60 °C) was based on different annealing temperatures required for amplification of respective products. The gel pictures of electrophoresis of these two sets of samples are given in Figure 3.2A and 3.2B.



100bp DNA ladder from NEB (#N3231S) was used

Figure 3.2A. Picture of an Agarose gel (1.8%) of SetA of amplicons and 100bp DNA ladder



L is 100bp DNA ladder from NEB (#N3231S)

Figure 3.2B. Picture of an Agarose gel (1.8%) of SetB of amplicons and 100bp DNA ladder

As can be seen from these gel pictures, single bands present in each lane showed varying intensities. Though the same 5µl PCR product after the reaction is loaded on the gel, the intensities of bands indicated that the concentrations of amplicons

generated in each PCR reaction varies substantially. In spite of same PCR reaction mix and same cycles of PCR program, intensity differences in amplicon product thus could be due to poor quality or efficiency in primer-template binding. Each cycle of PCR, doubles the number of amplicon molecules and after n number of cycles, $2^n \times A$ number of molecules are generated. Here, A is the number of molecules produced in first cycles of PCR reaction. If A is low (because of poor primer-template binding and other reasons) the concentration at the end of the reaction would show as differences in band intensities. Careful examination of the intensities of 16 bands in two gel pictures indicated very poor to medium, good and very good primer-template binding efficiencies. Non specific DNA products (PCR artifacts) were absent and migration of bands in gel compared to migration of DNA ladder bands indicated size of amplicon that matched with the expected size and are tabulated alongside the gel picture for each lane.

3.3 PCR AMPLIFICATIONS FROM PATIENT AND WT SAMPLES FOR TESTING ON WDDM

Figures 3.3 to 3.14 are of electrophoresis gels of amplicons from respective gDNA samples that were used for hybridization on WDDMs. All the fragments sizes appeared as were expected. Intensity values from these gels for amplicons in two sets are tabulated in Table VIIIA & B. Amplicons were prepared from gDNA of five healthy individuals (C1 to C5), 6 WD indexed patients, 6 samples of patient families (P9, 9-1, 9-2, 9-3, P12, 12-1, 12-2, 12-3, P5, P7, P8, P10) and CH (commercial gDNA).

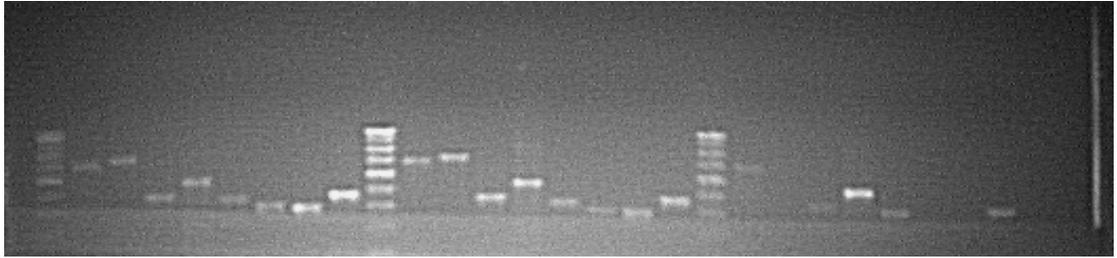


Figure 3.3 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder (L). Wells from left: A- set of 8 amplicons of P9, A- set of 8 amplicons of P9-1, A- set of 8 amplicons of P9-2.



Figure 3.4 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder. B set of 8 amplicons each of (1) P9, (2) P9-1 and (3) P9-2



Figure 3.5 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder A set of 8 amplicons each of (1) P9-3, (2) P12 and (3) P12-1

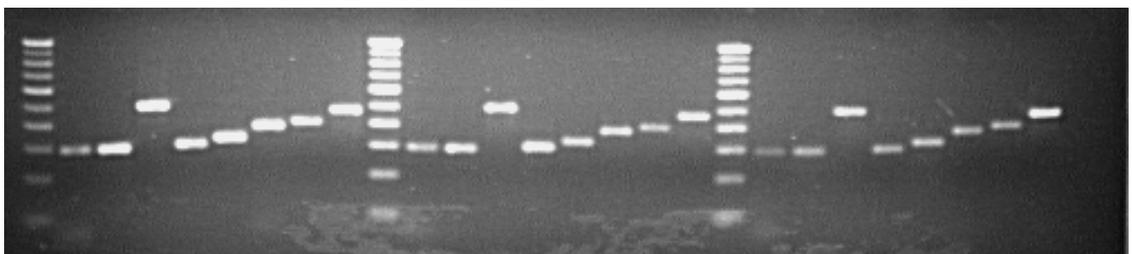


Figure 3.6 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder B set of 8 amplicons each of (1) P9-3, (2) P12 and (3) P12-1

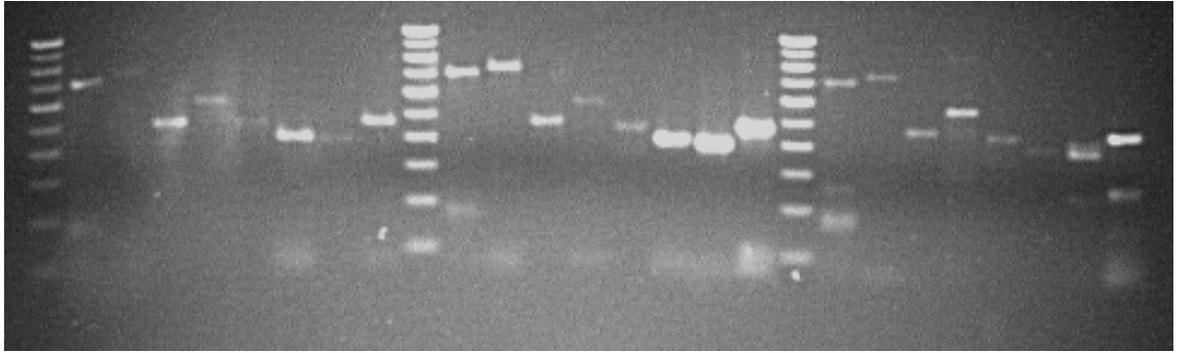


Figure 3.7 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
A set of 8 amplicons each of (1) P12-2, (2) P12-3 and (3) P5

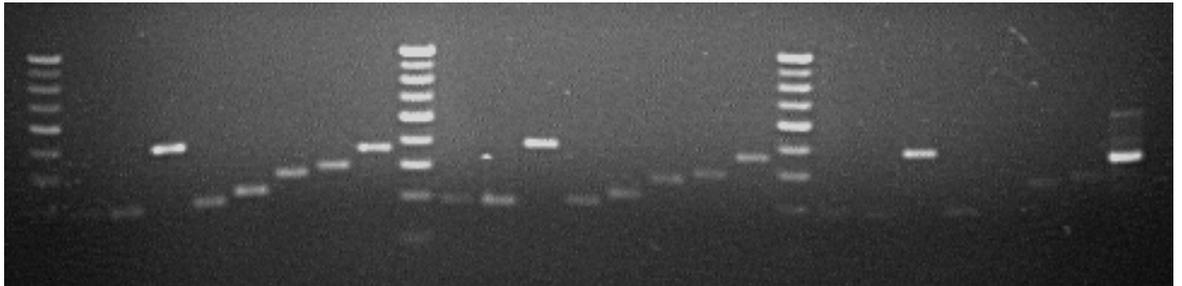


Figure 3.8 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
B set of 8 amplicons each of (1) P12-2, (2) P12-3 and (3) P5

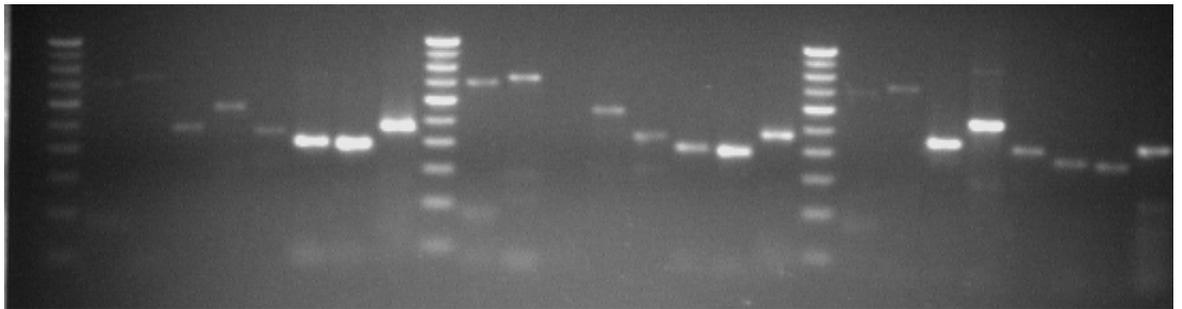


Figure 3.9 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
A set of 8 amplicons each of (1) P7, (2) P8 and (3) P10

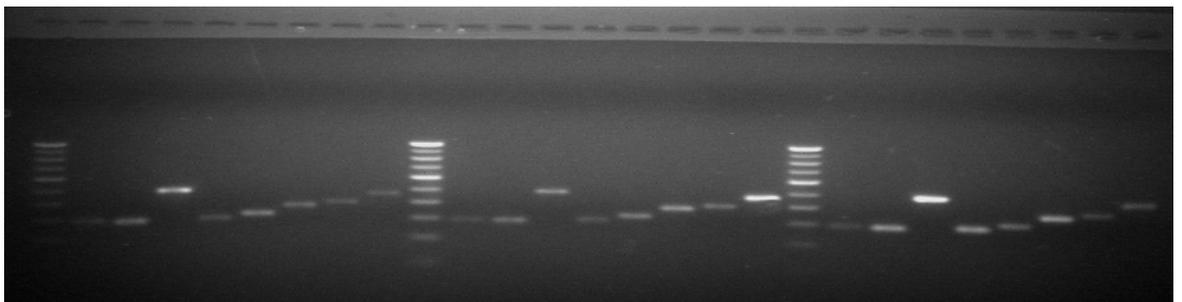


Figure 3.10 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
B set of 8 amplicons each of (1) P7, (2) P8 and (3) P10

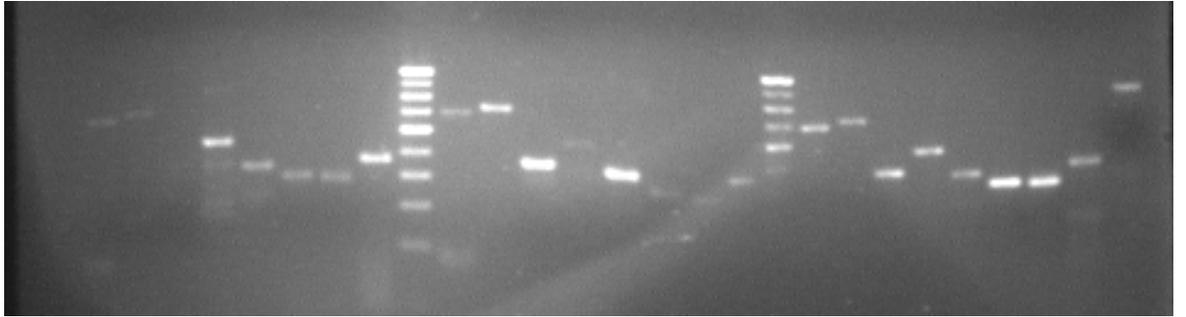


Figure 3.11 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
A set of 8 amplicons each of (1) C1, (2) C2 and (3) C3 gDNA templates

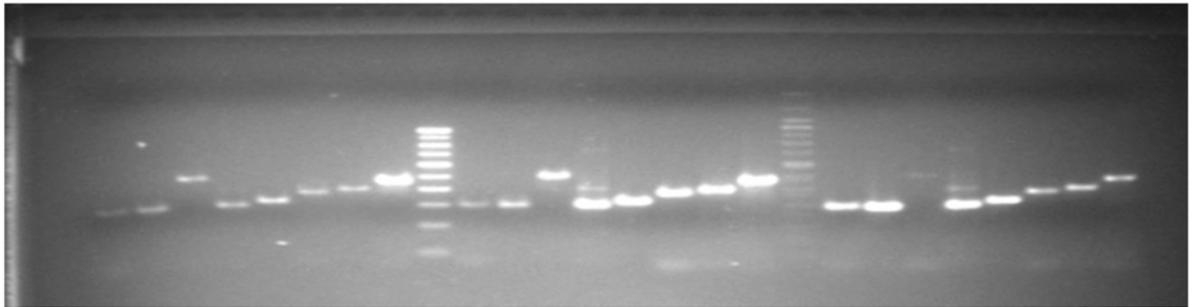


Figure 3.12 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
B set of 8 amplicons each of (1) C1, (2) C2 and (3) C3 gDNA templates

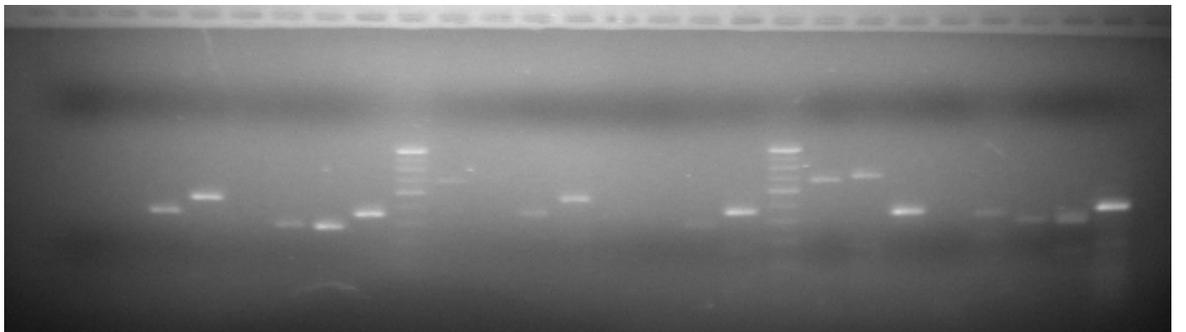


Figure 3.13 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
A set of 8 amplicons each of (1) C4, (2) C5 and (3) Ch

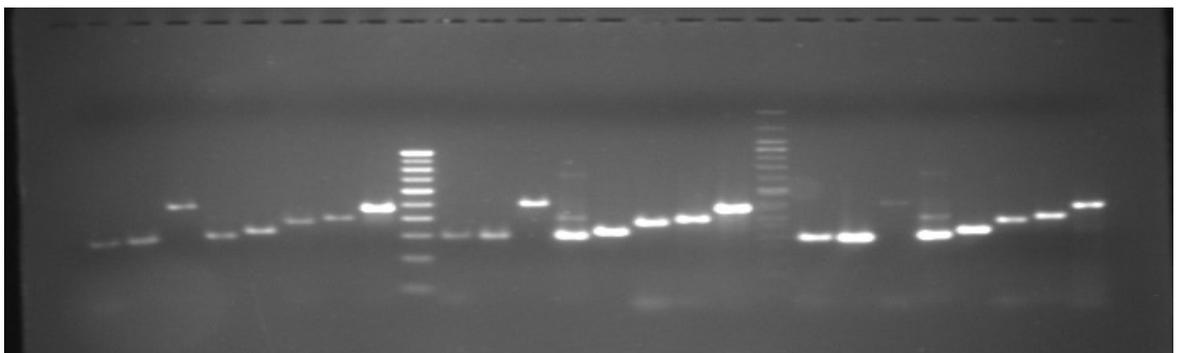


Figure 3.14 Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder
B set of 8 amplicons each of (1) C4, (2) C5 and (3) Ch gDNA templates

Table VIII A, comparative yields of Set A amplicons from 18 gDNA samples

SET-A of 8 Amplicons																		
Temp- late→ Exons	P9	9-1	9-2	9-3	P12	12-1	12-2	12-3	P5	P7	P8	P10	C1	C2	C3	C4	C5	CH
2a	++	+++	++-	++	-	+++	++	+++	++	++-	++	±		±	++	±	++	++
2b	++	+++	++	++	-	-	±	+++	++	±	++	++-		++	++	±	++	++
5	++	+++	++	+++	++	++	++	++	++	++	-	+++	++	±	+++	±	+++	+++
10/11	++	+++	+++	+++	+++	++	++	±	+++	++	++	+++	++	++	+++	++	±	+++
13	++	+++	++	++	+++	++	±	±	++	++	++	++	+++	+++	±	++	+++	+++
16	++	++	++-	++	+++	++	+++	+++	±	+++	++	++	++-	++	±	++	++-	+++
19	+++	+++	++-	++	++	++	±	+++	++	+++	+++	++	++	±	±	++	±	+++
21	+++	+++	++	++-	+++	+++	++	+++	+++	+++	+++	++	++	++	++	++++	++	++

++++ = Max intensity, +++- <++++, ++ = Medium intensity, + <++, - = no PCR and no template, Total amplicons = 16x18=288

Red marked were repeated in tubes after electrophoresis of 288 samples from PCR well plate.

(PCR were repeated in individual tubes wherever the PCR reaction failed in well plate.)

Table VIII B, comparative yields of Set B amplicons from 18 gDNA samples

SET-B of 8 Amplicons																		
<u>Temp- late→ Exons</u>	<u>P9</u>	<u>9-1</u>	<u>9-2</u>	<u>9-3</u>	<u>P12</u>	<u>12-1</u>	<u>12-2</u>	<u>12-3</u>	<u>P5</u>	<u>P7</u>	<u>P8</u>	<u>P10</u>	<u>C1</u>	<u>C2</u>	<u>C3</u>	<u>C4</u>	<u>C5</u>	<u>CH</u>
<u>4</u>	++	++	+++	++	++	++	±	±	±	++-	++-	++-	++	++	++	++-	++	+++
<u>7</u>	+++	+++	+++	+++	+++	+++	±	++	+++	++	++	++	++	++	+++	++	+++	++++
<u>8</u>	+++	+++	+++	+++	+++	+++	+++	+++	+++	++ ±	++	+++	+++	+++	+++	++	+++	+++
<u>9</u>	++	+++	++	+++	+++	++	++	++	±	++	++	++	++	+++	++	++	++++	+++
<u>12</u>	++	++	++	+++	+++	+++	++	++	+++	++	++	++	+++	+++	++	+++	++++	+++
<u>14</u>	+++	+++	++	+++	+++	++	++	++	±	++	++	+++	++	++	++	++	++++	+++
<u>15</u>	++	++	++	+++	++	++	++	++	±	++	++	++	+++	++	++	++	++++	+++
<u>18</u>	+++	+++	+++	+++	+++	+++	+++	++	+++	++	+++	++	+++	+++	++	++++	++++	+++

++++ = Max intensity, +++- <++++, ++ = Medium intensity, + < ++,

Red marked were done manually in tubes after electrophoresis of 288 samples from PCR well plate.

Re- PCR were repeated in individual tubes wherever the PCR reaction failed in well plate.

3.4 SITE DIRECTED MUTAGENESIS OF WT AMPLICONS AND KINETICS OF PCR REACTIONS IN MUTAGENESIS

To test the performance of the arrays for detection of mutation simulated mutant samples were prepared using site directed mutagenesis. The template for these reactions was human gDNA (hgDNA) from a commercial source. The different types of mutations generated included one point mutation each of G->A and T->C type, base deletion (one base and two base) and base insertion (single and double). Following Table lists the exons and the type of mutation in each one. The steps involved in generating these mutant products is detailed in **Section 2.14.**

Table IX: SDM fragments for simulating WD

Exon	WT amplicon / Mutated short amplicons (Size in Bps)	Type of mutation	Name of mutated amplicon
4	291 /205,112	<u>DBI</u>	SDM4
7	288/ 238,78	DBD	SDM7
8	482/ 253,255	<u>SBI</u>	SDM8
9	297/ 207,116	SBD	SDM9
15a	393/ 236,182	(G>A), SPM1	SDM15a
15b	393/ 236,182	(T>C), SPM2	SDM15b

The agarose gel picture in Figure 3.15 below shows the two small PCR amplification products (left and right) as generated with the use of mutated internal primers from each of the wild-type products. It must be noted that the sub-products were generated in separate reactions and were loaded on the gel together for the sake of comparison.

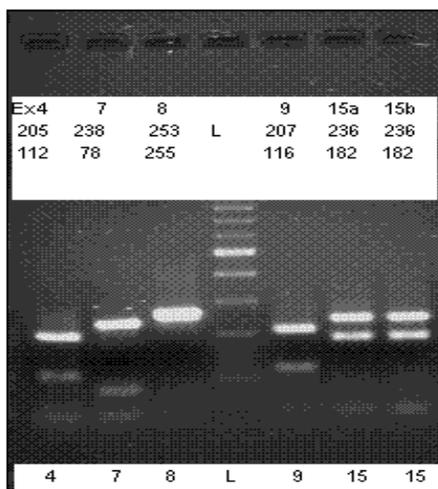


Figure 3.15. Gel picture of aliquots from 1st step (Before extension PCR) of site directed mutagenesis PCR. For all the exons viz 4,7,8,9&15 two fragments are visible, sizes of which are indicated in corresponding lanes.

The two smaller PCR products carrying the required mutation, were mixed in equi-molar concentrations and extension PCRs were done as detailed in Section 2.6. Amplification products from this step were diluted 1:1000 and used as templates in another round of amplification reactions to study kinetics of mutagenesis (Section 2.14). Three gel pictures have been put together in Figure 3.16 (A,B & C). Increase in band intensities from W to X to Y and further to Z was obvious in all 6 SDMs though band intensities are very different for each SDM. Visual observations from Figure 3.16 A,B,C for kinetics of each of SDM fragments in tabulated form are given in Table X. Examining band intensities of W',X',Y' and Z' indicated that till 25 cycles (W' and X') reference templates of each SDM were negligibly amplified (indicated by ' - ' in Table X). Y' aliquots which have undergone 30 cycles of PCR are ' + ' in intensity, indicating that carry over DNA amplification starts after 25 cycles of 3rd step of SDM PCR which gets further amplified in next 5 cycles (25 to 30) and are seen as intensity of Z' bands ('++' or '+') for each SDM. It was hence inferred that amplification of carryover contamination (of

WT allele of amplicon) takes over after 30 cycles of PCR. Hence 35 cycles PCR is not best for simulating WD mutations though yield of DNA fragment is high at this cycle. Hence Y aliquot was used for hybridization on WDDM.

Table X. Summary of visual observations from Fig 3.16(A,B,C)

Exon	Fragment size	3 rd Round Extension PCR products with and without extension PCR							
		No of PCR cycles / Reaction Type (WW'XX'YY'ZZ')							
		20 W	20 W'	25 X	25 X'	30 Y	30 Y'	35 Z	35 Z'
4	291	+	-	++	-	+++	+	++++	++
7	288	-	-	+	-	+++	+	+++++	++
8	482	+	-	++	-	+++	+-	++++	+
9	297	-	-	++	-	+++	+	+++++	++
15a	393	-	-	++	-	+++	-	++++	+
15b	393	+-	-	++	-	+++	+-	++++-	+

Plus + refers to the presence of DNA band which is barely visible while ‘++’ and ‘+++’ indicate increasing band intensities. Minus ‘-’ indicate the absence of a band. It was found that 30 cycles (Y aliquot) should have good yield of required SDM amplicons.

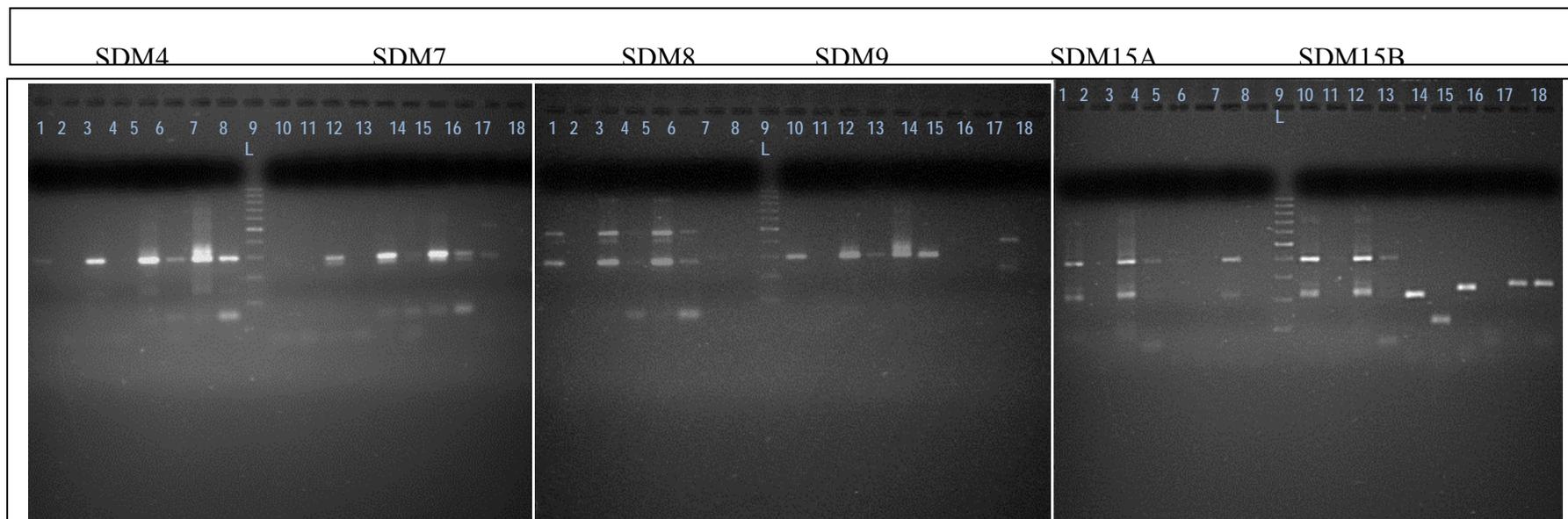


Figure 3.16 A

Figure 3.16 B

Figure 3.16 C

Figure 3.16 (A,B&C) Agarose gel electrophoresis with 1.8% gel and 100bp DNA ladder. W,W',X,X',Y,Y',Z,Z' of SDM4 (292 bps), SDM7 (288 bps), SDM8 (482 bps) and SDM15a and b (393bps). Remaining lanes of C gel were loaded with aliquots from 1st step of SDM-PCR.

3.5 SEQUENCING OF SDM FRAGMENTS

To confirm the designed changes in SDM amplicons all the six mutagenised fragments were sequenced using the protocol given in Section 2.22. The regions of electrophoregrams where the mutation was introduced are given in Figures 3.17 to 3.24 below. Sequence from database has been provided as reference. Any ambiguities in sequence data were resolved by examining the electro-phoregrams of reverse strand wherever needed.

I. SDM4 : Double Base Insertion-DBI

5' -CAGTAAGTACTGTGTGGGTGCGTTACG-3' (Mutated)
3' -GTCATTCATGACACACCCACGCAATGC-5'

5' -CAGTAAGTACTGTGGGTGCGTTACG -3' (WT)
3' -GTCATTCATGACACCCACGCAATGC -5'

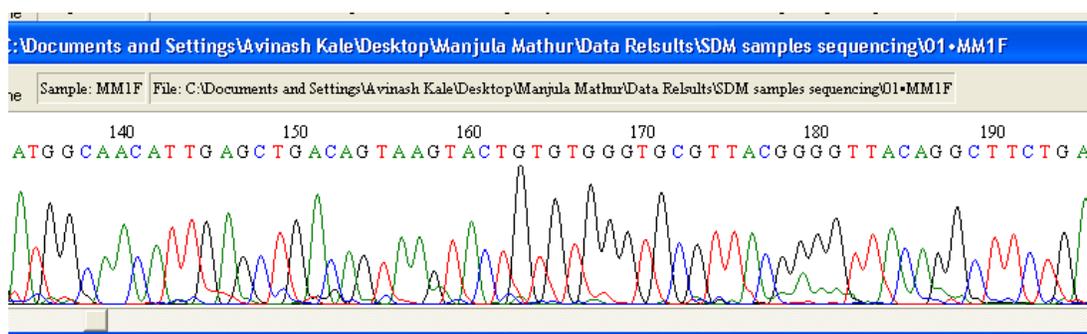


Figure 3.17 Snap shot of electrophoregram of forward strand of SDM4. Region between 160 -170 shows presence of inserted GT.

II. SDM-7: Double Base Deletion- DBD

5' -TTGTGTACCTTT--CCAGGTATATATG-3' (Mutated)
3' -AACACATGGAAA--GGTCCATATATAC-5'

5' -TTGTGTACCTTTGTCCAGGTATATATG-3' (WT)
3' -AACACATGGAAACTGGTCCATATATAC-5'

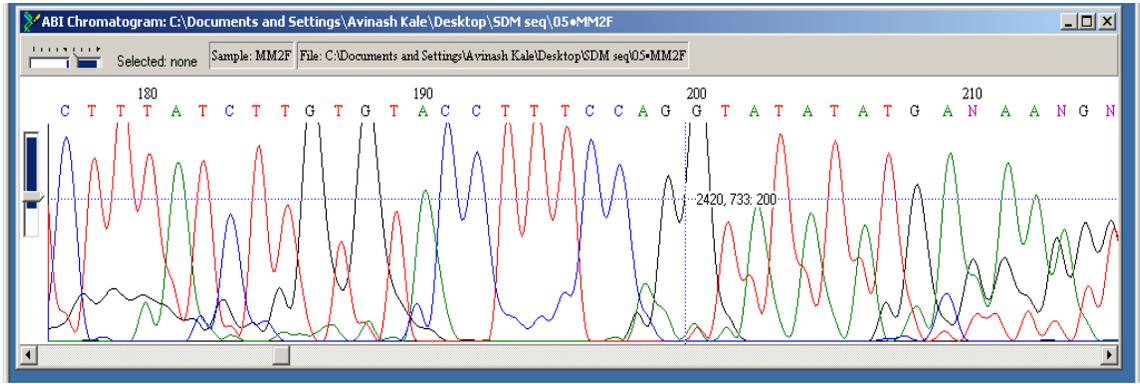


Figure 3.18 Snap shot of electrophoregram of forward strand of SDM7. Sequence from 192 reads as CTTTCCAG, while the WT sequence is CTTTGTCCAG

III. SDM-8: Single Base Insertion- SBI

5' -TGGTTGCTGTGGCCTGAGAAGGCCGA-3' (Mutated)
 3' -ACCAACGACACCGACTCTTCCGCCT-5'

5' -TGGTTGCTGTGGCTGAGAAGGCCGA -3' (WT)
 3' -ACCAACGACACCGACTCTTCCGCCT -5'

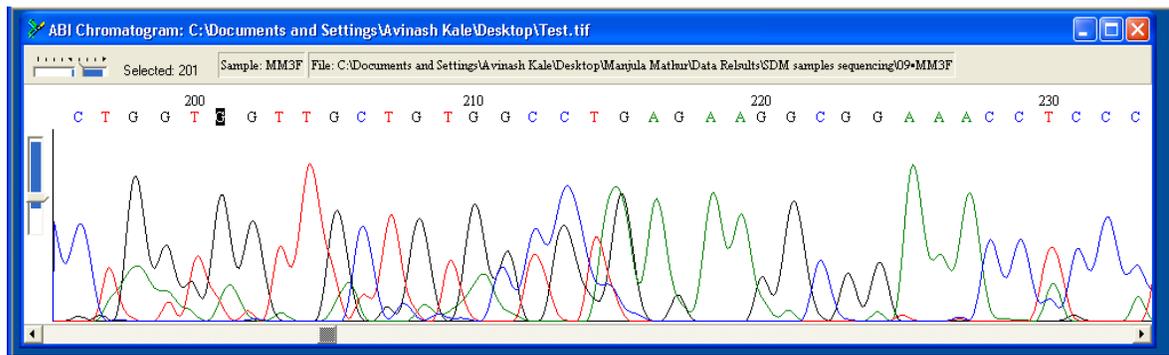


Figure 3.19 Snap shot of electrophoregram of forward strand of SDM8. Additional C seen in the region between 210 – 215 .

IV. SDM 9: Single Base Deletion – SBD

5' -TAGAGCAAAA-CTCAGAAGCCCTGG-3' (Mutated)
 3' -ATCTCGTTTT-GAGTCTTCGGGACC-5'

5' -TAGAGCAAAACCTCAGAAGCCCTGG-3' (WT)
 3' -ATCTCGTTTTGGAGTCTTCGGGACC-5'

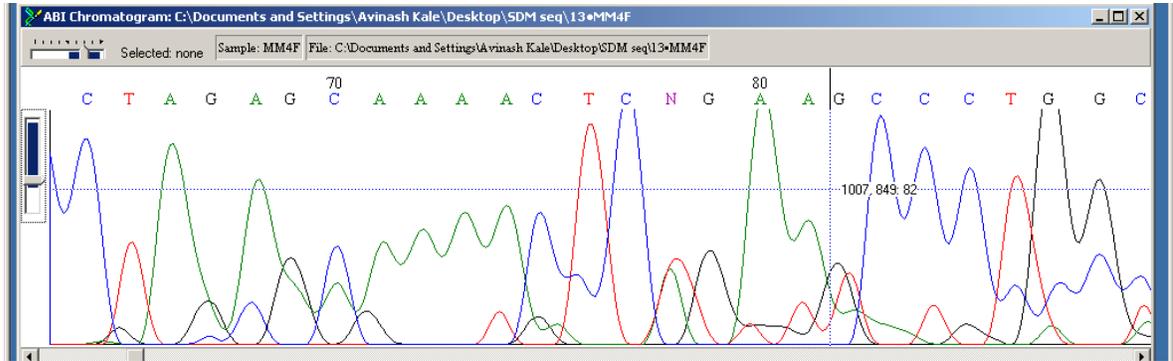


Figure3.20 Snap shot of electrophoregram of forward strand of SDM9. Sequence after Position 70 has CAAAACCT while the WT reads as CAAAACCT

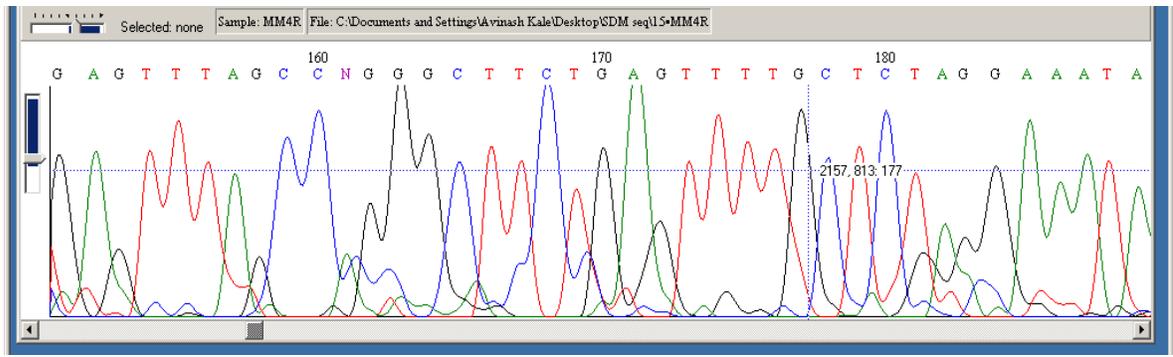


Figure3.21 Snap shot of electrophoregram of reverse strand of SDM9. Sequence after 170 has GAGTTTTT while in WT this reads GAGTTTTT

V. SDM 15a: Point Mutation- SPM1 (G>A)

5' -GTGCCAGGCTGTAGAATTGGGTGCA-3' (Mutated)
 3' -CACGGTCCGACATCTTAACCCACGT-5'

5' -GTGCCAGGCTGTGGAATTGGGTGCA-3' (WT)
 3' -CACGGTCCGACACCTTAACCCACGT-5'

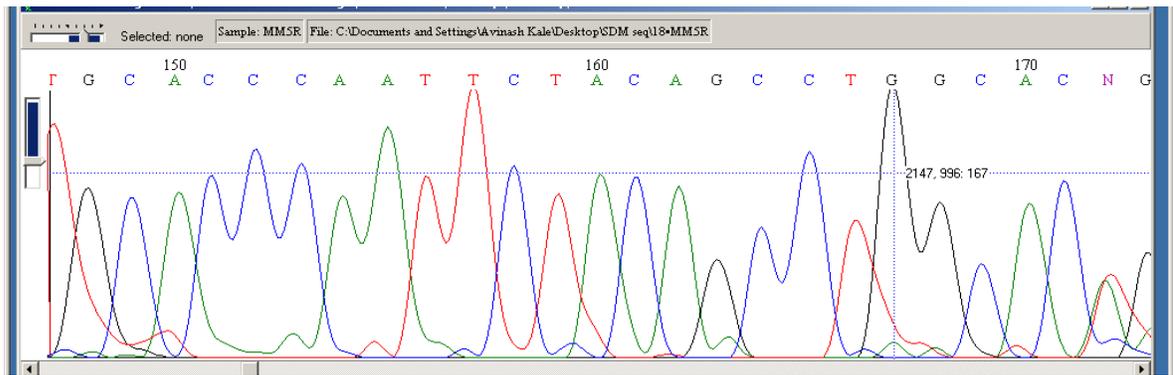


Figure3.22 Snap shot of electrophoregram of reverse strand of SDM15a. Region after 150 reads as TTCTACA while the WT reads TTCCACA (C>T)

VI. SDM 15b: Point Mutations- SPM2 (T>C)

5' -GTGCCAGGCTGTGGAAC TGGGTGCA-3' (Mutated)
 3' -CACGGTCCGACACCT T GACCCACGT-5'

5' -GTGCCAGGCTGTGGAAT TGGGTGCA-3' (WT)
 3' -CACGGTCCGACACCT TAACCCACGT-5'

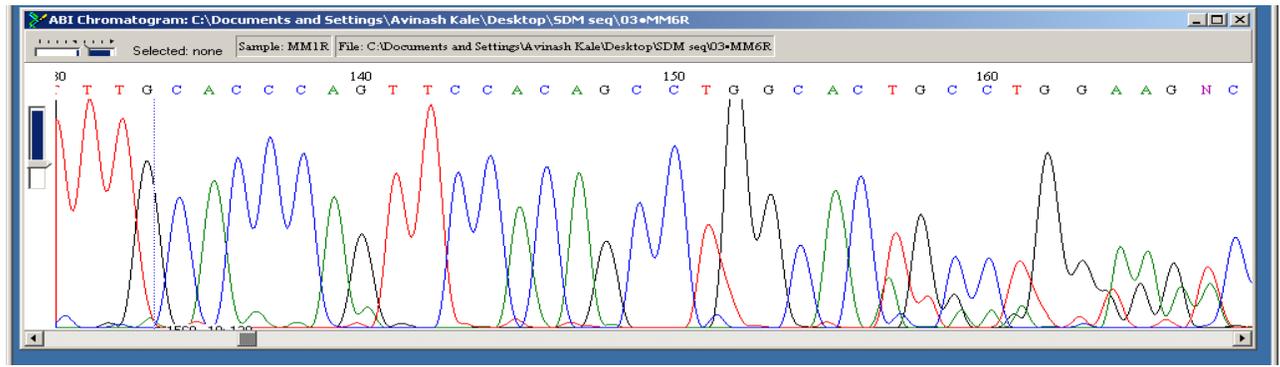


Figure 3.23 Snap shot of electrophoregram of reverse strand of SDM15b. Sequence region after 137reads as CCAGTTC while the WT is CCAATTC

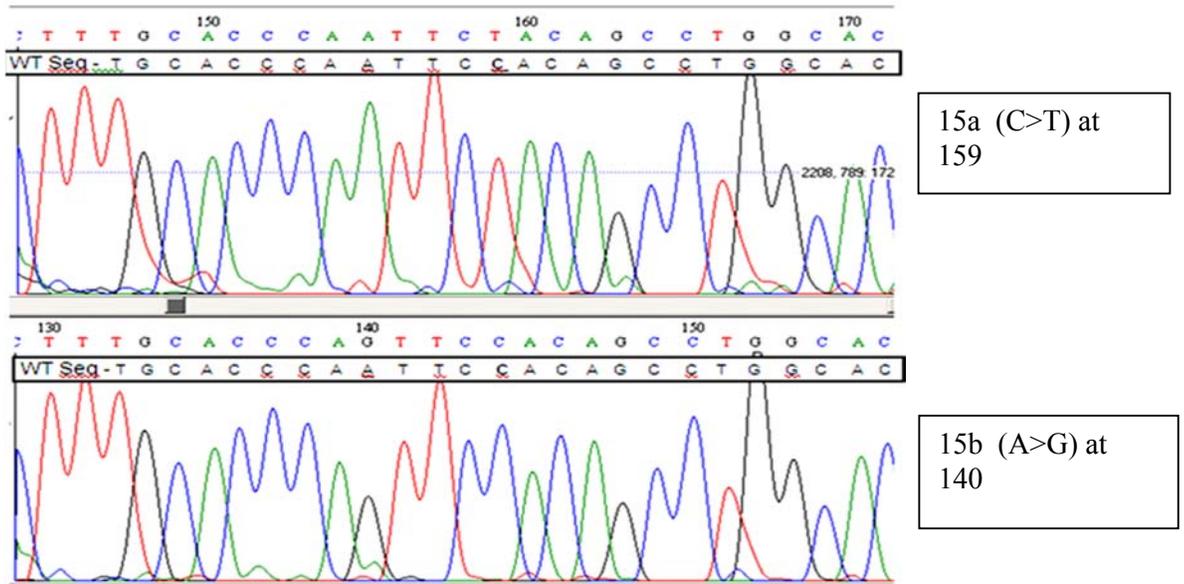


Figure 3.24 Snap shot of aligned electrophoregrams of reverse strands of SDM15a and SDM15b amplicons showing changes at respective positions in both

Thus, DNA sequencing of all the six SDM amplicons confirmed the generation of mutated allele to be used further.

3.6 PROTOTYPE MICROARRAYS FOR MUTATION DETECTION

Perfect match (PM) and mismatch probes (MM) for one mutation each in exon 5, 7, 8, 9, 11, 13, 14, and 15 and 18 were synthesized and were printed as Prototype WDDMs. Three DIG labeled amplicon samples (one WD patient and two normal) were subjected to hybridization experiments on prototype WDDM. Figure 3.25 is the scanned image of hybridized prototype WDDM. The text in upper boxes describes the respective sample of triplets of spots.

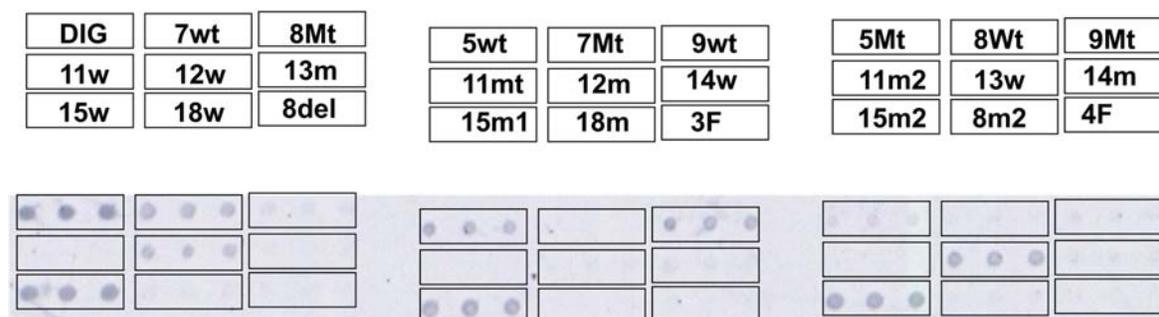


Figure 3.25 Hybridization on Prototype WDDM.

The spot intensity from the PM (Wild Type-wt/w) and MM (mutant –Mt/m) probes for the probes pair was compared with each other and following observations were made: 1) control DIG spots yielded highest intensity 2) probes from Exons 5, 7, 9 and 12 WT showed higher intensity spots compared to their mutant probes 3) PM and MM probes for exons 8, 11 and 18 did not yield any visible spots 4) for Exon 18 both the PM and MM probe spots yielded comparable intensities and 5) across arrays, the hybridization patterns remained comparable.

These observations suggested that immobilization of spotted oligonucleotides on epoxy coated slides could be achieved and such microarrays can be utilized to differentiate mutant alleles from their corresponding wild-type alleles using DIG based hybridizations.

3.7 WD DETECTION MICROARRAY DESIGN CONSIDERATIONS

Based on the results from the prototype arrays, a large number of probes corresponding to the PM and MM sequence of the region around the sites of interest were designed. These probes were evaluated for the occurrence of runs of bases and secondary structures using OligoArray software. Since the strandedness of the probe might determine the hybridization kinetics,

probes from both the strands were designed for some mutations. Additionally, the array layout included some probe solutions spotted at two different locations on each Array to evaluate the uniformity of hybridization across the slide surface. The list of probes included in the WDDM and the nature of mutation covered by them are shown in Table III. These arrays were printed with total of 204 probe solutions, each in triplicate. Out of 204 probe triplets on each WDDM, 84 were WT probes (these include 63 Indian WD mutations, 4 SNP variants reported among these exons, repeated spots and reverse complements of probes) 93 were mutant probes, 20 exon specific control probes, 5 blanks (only buffer as negative control) and 2 were DIG labeled (as positive control). The printing was done asymmetrically to enable identification of probe pairs on HWDDMs by the pattern developed after hybridization. Dimension of one WDDM was 12.1 x 12.9 millimeter. Figure 3.26 below shows a magnified view of WDDMs.

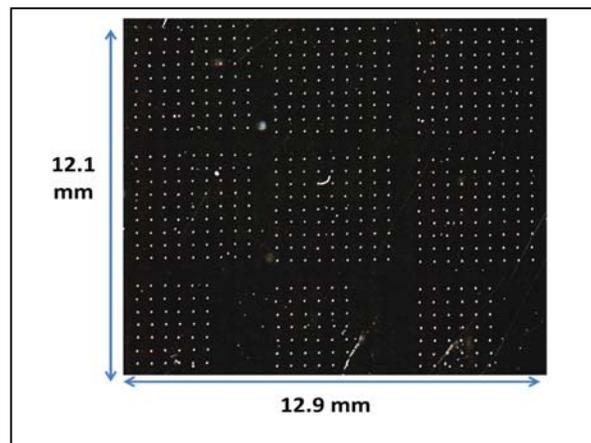


Figure 3.26 Magnified View of one of the Printed WDDM (magnified View)

3.8 SAMPLE PREPARATION

Sixteen amplicons of *ATP7b* spanning all the selected 63 mutations were amplified from each gDNA sample by 16 independent Polymerase Chain Reactions (PCRs) using exon specific primers. Equi-molar proportions of 16 amplicons were pooled to generate samples for DIG Labeling and hybridization on WDDMs. The following are the different types of samples prepared.

Table XI. Sample description for hybridization on WDDM

SI No	Sample type	Description
1	Wild-Type	All the 16 exons amplified from WT gDNA
2	SDM4	Amplicon corresponding to Exon 4 is mutated by SDM and all the remaining 15 amplicons were obtained from WT genomic DNA.
3	SDM7	Amplicon corresponding to Exon 7 is mutated by SDM and all the remaining 15 amplicons were obtained from WT genomic DNA.
4	SDM8	Amplicon corresponding to Exon 8 is mutated by SDM and all the remaining 15 amplicons were obtained from WT genomic DNA.
5	SDM9	Amplicon corresponding to Exon 9 is mutated by SDM and all the remaining 15 amplicons were obtained from WT genomic DNA.
6	SDM15a	Amplicon corresponding to Exon 15 is mutated by SDM (G->A) and all the remaining 15 amplicons were obtained from WT genomic DNA.
7	SDM15b	Amplicon corresponding to Exon 15 is mutated by SDM (T->C) and all the remaining 15 amplicons were obtained from WT genomic DNA.
	Patient / Control Samples	All exons amplified from a patient / control genomic DNA; Listed as a separate Table 6.

All the above samples were prepared in about 200 ng of total DNA and were taken for DIG (Digoxigenin) labeling using DIG High Prime Labelling Kit (Protocol described in Section 2.16). Among these, SDM4, SDM7, SDM8, SDM9, SDM15a and SDM15b simulate patient samples with known mutations in specific amplicons.

3.9 HYBRIDIZATION AND POST HYBRIDIZATION PROCESSING

The printed arrays (WDDMs) were hybridized with DIG labeled samples overnight under a cover slip, keeping them in humid and air tight boxes. The protocols for hybridization and post-hybridization are detailed in the **Section 2.17**. Hybridization detection reactions were done, the next day, using alkaline phosphatase conjugated Anti-DIG antibody and NBT/BCIP substrate. The hybridized WDDM (HWDDM) after drying were scanned on flatbed scanner at 600 dpi resolution. Scanned images were saved and for each array, intensities of all the spots were quantified with the help of ImageJ software (<http://rsb.info.nih.gov/ij/>)^{wr14}. When the hybridizations were carried out with wild-type sample at 42°C, as expected, most of the PM probes yielded higher intensities compared to the corresponding MM probes. The first and last triple spots were the DIG labeled oligos (positive controls) used for normalization of intensity values.

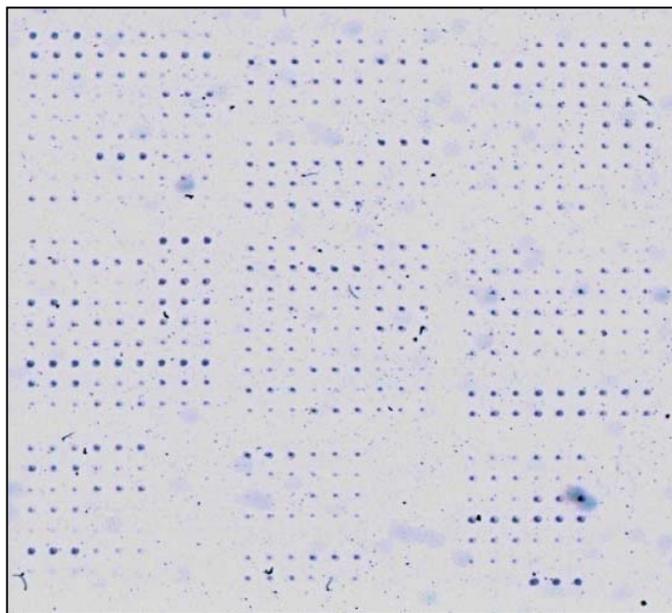


Figure 3.27 A magnified view of the WDDM hybridized with Wild-type sample

For standardization purpose, amplicons of WT chromosomal gDNA (Healthy normal volunteers) were taken for hybridization on WDDM. The Figure 3.27 shows an example of a scanned image after the array was hybridized to amplicons from gDNA of wild-type

individual. The picture depicted above is a magnified view of 13.5 x 13.5 mm region on the hybridized slide. The intensities from each spot were quantified by method given in Section 2.19 and were listed as a spread sheet.

3.10 HYBRIDIZATIONS CARRIED OUT FOR EVALUATING WDDM

A series of hybridizations were conducted to assess the quality of the printed WDDMs and also for determining the optimal experimental conditions of hybridizations on. Quality of all the microarray images of HWDDMs were qualitatively evaluated and summarized. Results obtained from all the HWDDMs are given in Table XII.

Table XII: Overall summary of hybridizations carried out for the evaluation of the arrays. #The quality of the microarray images after complete processing with each '+' representing a general improved quality of the image. Images that have atleast two “++” were used for quantification and further comparison.

Sample Name	Hybridization temp (°C)#	
	42	52
Wild-Type Chromosomal DNA	+++	++'
SDM4	+++'	+++'
SDM7	++'	+++'
SDM8	+++'	++'
SDM9	+++'	+++'
SDM15	+++'	+++'

The quality of scanned array images (final spot intensities) is a result of multiple factors including the quality of hybridization, wash conditions and image processing. Some images were negatively affected by the presence of high amount of back ground signal. The mean of the three spot intensities was calculated from the triplet spots of each probe and standard deviation (S.D.) was computed for identification of outlier spots. Spots having intensity value lower or greater than 2 x S.D. from mean value, were physically examined and if necessary, the intensity data from these outliers were not included in further analysis of

probes and array performance.

3.11 EFFECT OF POST-HYBRIDIZATION WASH TEMPERATURE ON SPOT INTENSITIES

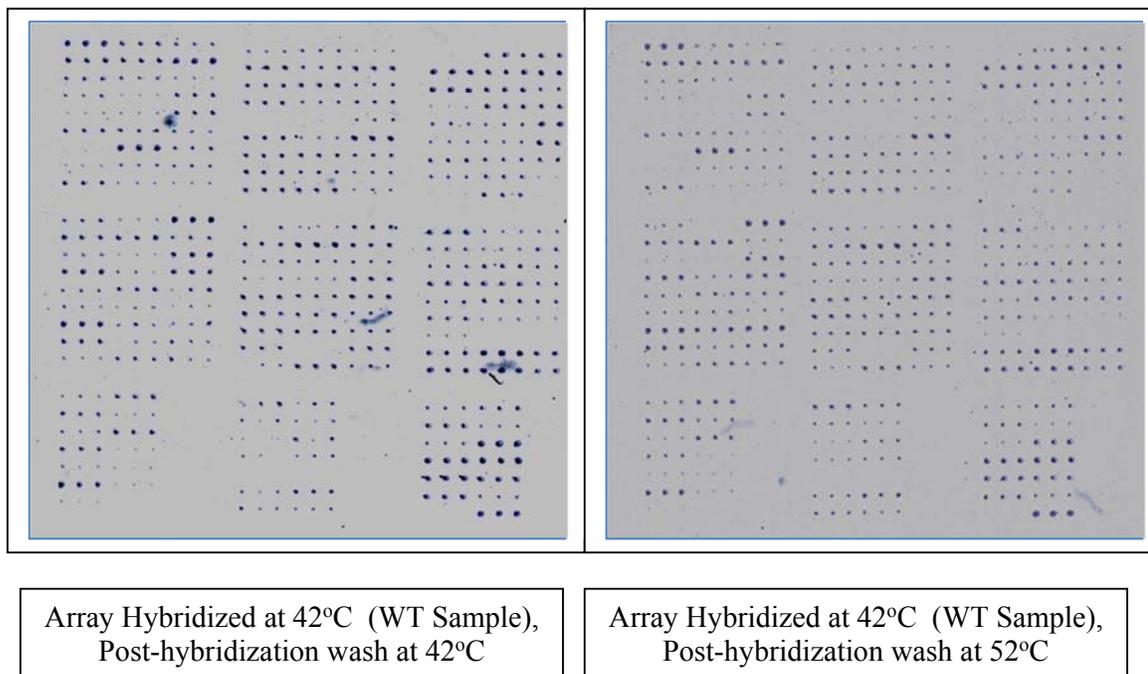


Figure 3.28 Two arrays with different post-hybridization wash temperature.

The Figure 3.28 represents two independent arrays hybridized with the same sample at 42°C and treated either at 42°C or at 52°C during the post-hybridization washes. The intensity of spots at these two temperatures were compared to assess the degree of correlation across temperatures. Figure below depicts the comparison of spot intensities as a function of post hybridization wash temperature. It can be observed that the spot intensities obtained from each spot correlated well ($R^2 = 0.73$) across experiments conducted at different post-hybridization temperatures. When the arrays were hybridized at 42°C and treated either at 42°C or at 52°C during the post-hybridization washes, it was observed that the spot intensities reduced at higher temperature. However, the discrimination of the PM and MM probes seemed to be lower at 42°C.

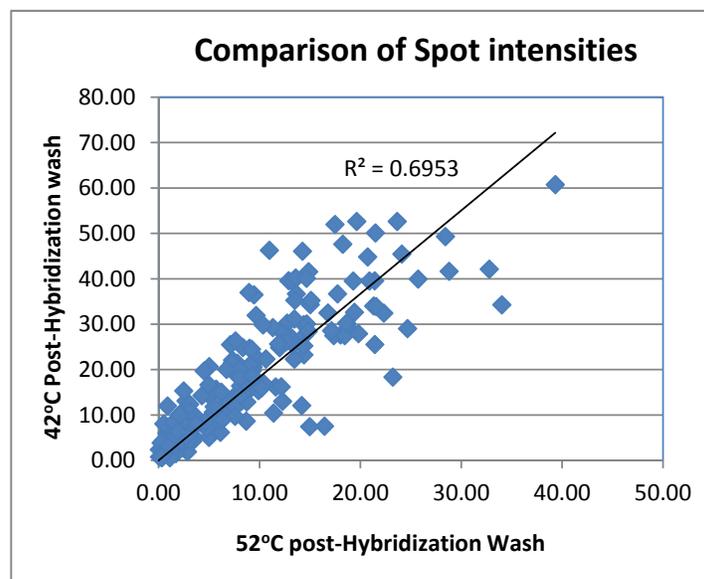


Figure 3.29 Comparison of Spot Intensities

Similarly correlation coefficients were computed from intensities obtained by differing post wash hybridization temperature from other hybridizing samples. The Table XIII below lists the R^2 values from all these pair-wise comparisons. The spot intensity data for these calculation of Correlation coefficient R^2 was taken from SDM hybridizations All these experiments and statistical analysis revealed a high degree of correlation when the arrays were processed at two temperatures after hybridization.

Table XIII. Pearson Moment Correlation Coefficient

Sample	R^2
WT (chromosome)	0.73
SDM4	0.92
SDM7	0.61
SDM8	0.71
SDM9	0.82
SDM15	0.80

3.12 STATISTICAL ANALYSIS OF DISPERSION OF INTENSITIES

Effect of temperature on spot intensities was extensively studied by statistical methods. Spot intensity data was taken from HWDDMs that used DNA amplicons from gDNA templates belonging to healthy individuals or commercial sources. Mean spot intensities from corresponding PM probe positions were taken and subjected to analysis using computation tools of MS Excel. The spot intensity data obtained from 5 independent repeats of each hybridization (42°C or 52°C post-hybridization wash) shown in Figure 3.30 revealed that the spread of observed spot intensities across repeat experiments were more pronounced at lower stringencies (42°C). The dispersion of intensities were found to be higher in spot intensity data in Figure 3.30A than data in Figure 3.30B. The error bars represent one standard deviation from the mean values. In spite of the lower mean intensities obtained from post-hybridization washes at 52°C, the dispersion of observed spot intensities across repeat experiments was less at higher stringency indicating better data reproducibility at higher temperature (52°C). Similarly, the ability to discriminate the PM and MM probes seem to be higher at 52°C. Hence, the 52°C data was used for all further analysis.

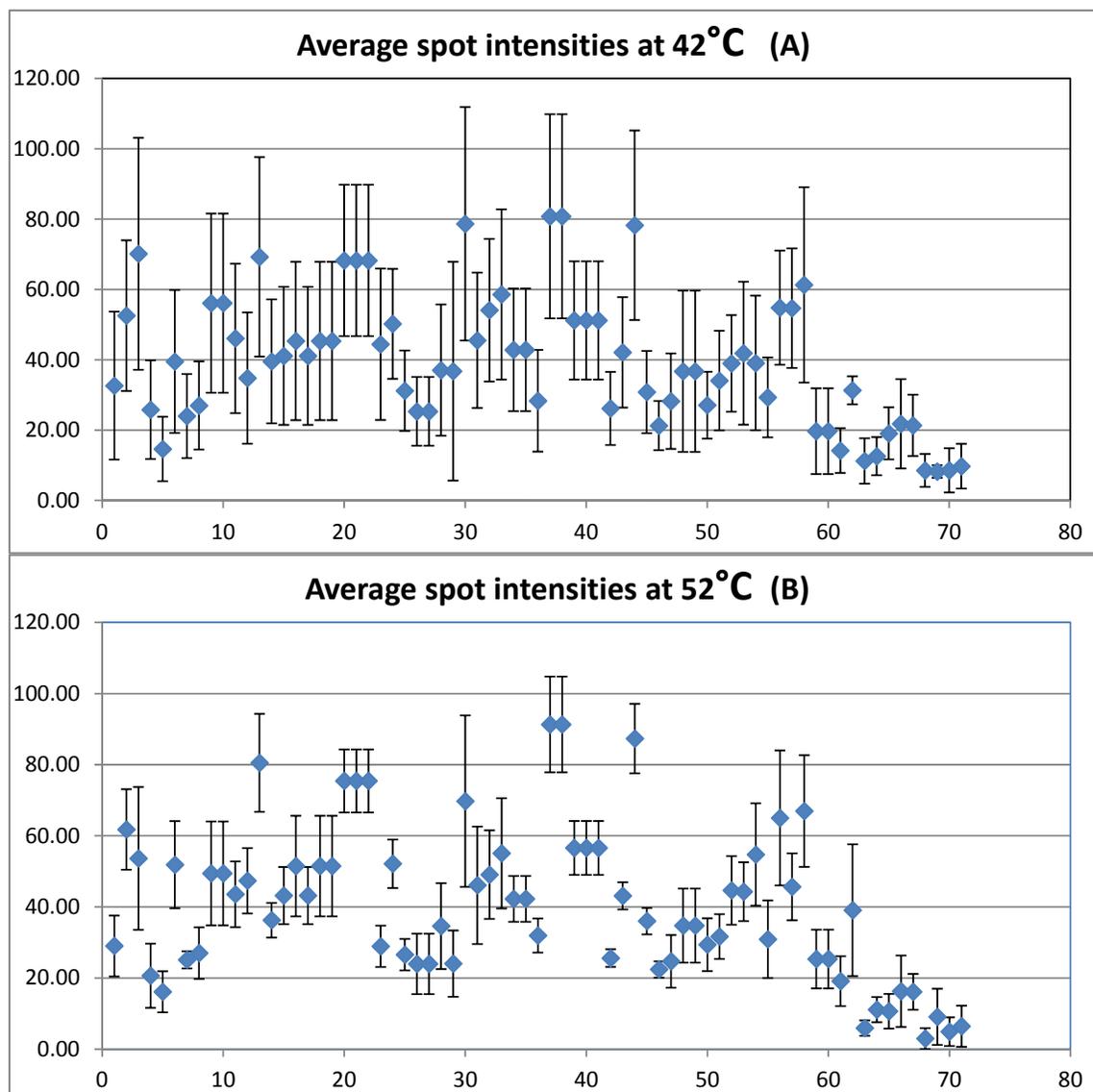
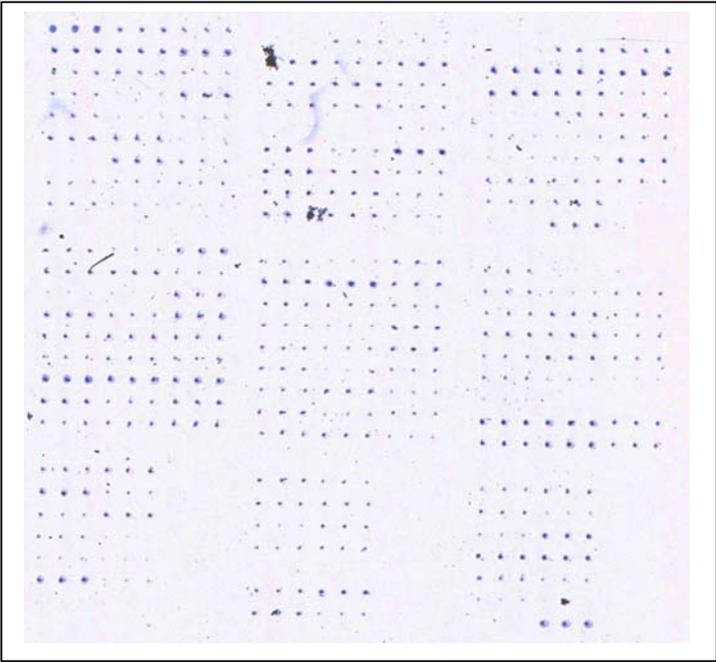


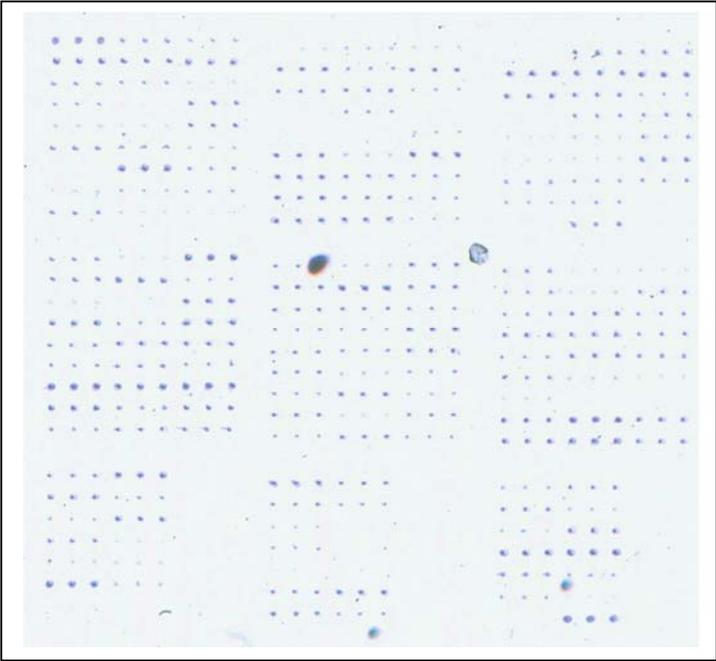
Figure 3.30 A and B are normalized spot intensities at perfect match position of 72 probes each from 6 HWDDM at 42°C (A) and at 52°C (B) respectively. The vertical bars represent one standard deviation about the mean values.

Shown below are the scanned pictures microarrays hybridized with SDM samples at 42°C and post-hybridization washes at 52°C (Figures 3.31 to 3.35). In general, the PM probes showed higher intensities compared to the corresponding MM probes. However, the intensity at PM (perfect match) spot were lower than its corresponding MM (mismatch) spot for that specific probe pair addressing the particular mutation present in SDM amplicon.

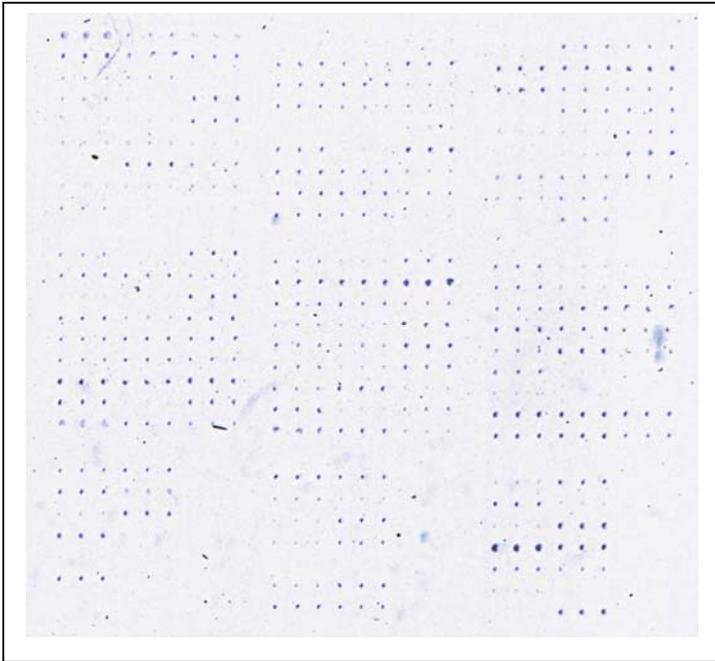
Sample: SDM4
Hybridization temp:
42°C
Post-Hyb Wash: 52°C
Figure 3.31



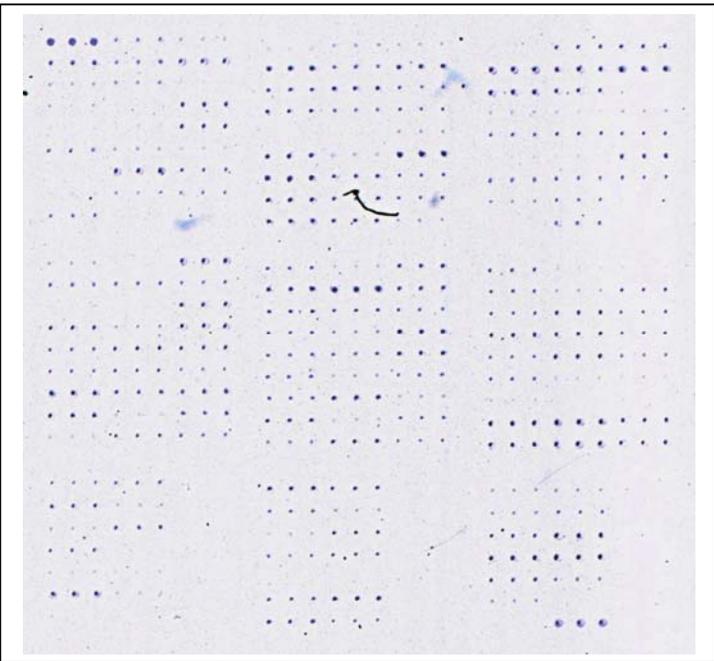
Sample: SDM7
Hybridization temp:
42°C
Post-Hyb Wash: 52°C
Figure 3.32



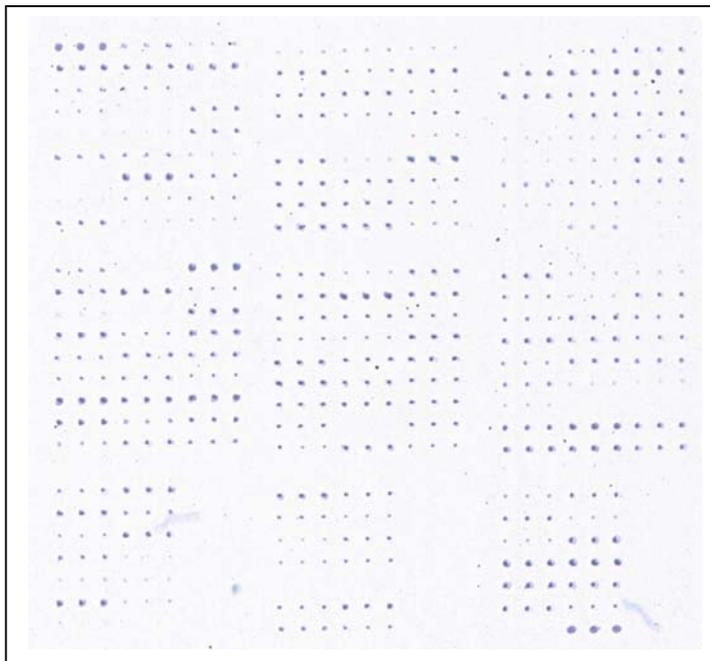
Sample: SDM8
Hybridization temp:
42°C
Post-Hyb Wash: 52°C
Figure 3.33



Sample: SDM9
Hybridization temp:
42°C
Post-Hyb Wash: 52°C
Figure 3.34



Sample: SDM15
Hybridization temp:
42°C
Post-Hyb Wash: 52°C
Figure 3.35



3.13 INTRA ARRAY CORRELATION OF SPOT INTENSITIES

A comparison of the intensities of repeat printing probes within the array would give assessment of the consistency of hybridization across the array (slide surface). Shown in the Figure 3.36 is a comparison of intensities obtained from such repeated spot pairs from array hybridization (52°C) obtained in SDM4 samples. The spot intensities obtained from two different locations of the same array were found to be well correlated (Figure 3.36, $R^2=0.72$). A similar correlation was observed for different repeats of the experiments indicating that the spot intensities (hybridization reactions) across a given slide were comparable.

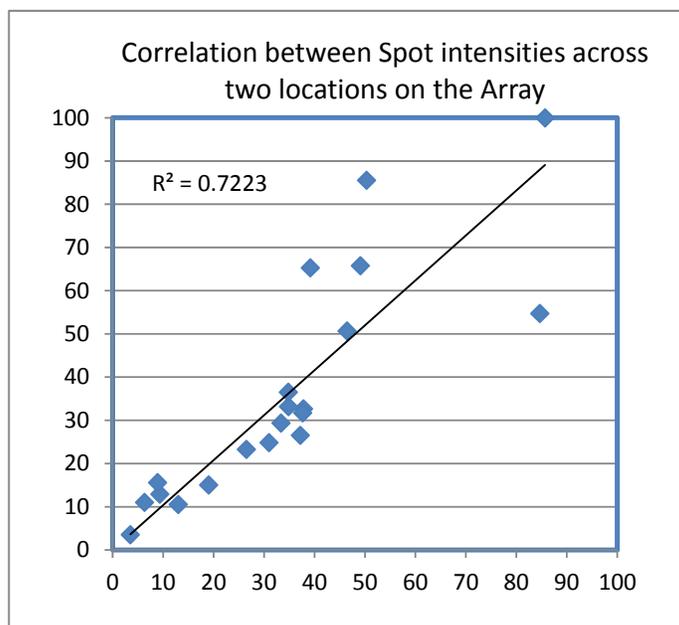


Figure 3.36 Correlation between Spot intensities across two locations on the Array. Both axes (X,Y) represent normalized intensity. The value of X is intensity at one position and Y is intensity at another position on array of the same probes.

These results confirmed that the kinetics of hybridization is uniform across the array.

3.14 DISCRIMINATION SCORE (DS) FOR EACH PROBE PAIR

The sensitivity of each probe with respect to its ability to detect the mutation is expected to be different. A quantitative assessment of each probe was done by computing the Discrimination Score (DS) based on the spot intensities at the PM and MM probe positions. The procedure for computing the DS is described in detail in the **Section 2.21**. Discrimination scores were computed for all the probe pairs from all the hybridization experiments as listed in Table XII (summary of hyb expts.). The DS values from the 52°C experiments have been listed here in Table XIV.

Table XIV: A Table of computed DS scores for all the mutations tested in the present study.

Probe_ID	Remark	DS_WTChr_3a_52	DS_SDM4_3b_52	DS_SDM7_3c_52	DS_SDM8_4a_52	DS_SDM9_4b_52	DS_SDM15_4c_52	(Avg)DS_52_Norm_Int	(SD) DS_52_Norm_int
P_01_2	Wt	-0.01	0.11	0.27	0.18	0.32	0.46	0.22	0.16
P_02_2	Wt	0.35	0.10	0.15	0.16	0.21	0.13	0.18	0.09
P_03_2	Wt	0.28	0.44	0.35	0.34	0.59	0.53	0.42	0.12
P_04_2	Wt	0.23	0.30	0.45	0.43	0.50	0.73	0.44	0.17
P_06_2	Wt	0.44	0.41	0.59	0.32	0.23	0.56	0.43	0.14
P_07_2	Wt	0.54	0.33	0.58	0.63	0.72	0.79	0.60	0.16
P_08_4	Wt-SDM	0.26	-0.55	0.02	0.35	0.20	0.10	0.06	0.32
S_4_Wt	Wt	0.35	0.23	0.55	0.16	0.27	0.41	0.33	0.14
P_09_5	Wt	0.34	0.17	0.49	0.31	0.52	0.70	0.42	0.19
E5_fs_P	Wt	0.36	0.38	0.42	0.32	0.49	0.43	0.40	0.06
P_12_5	Wt	0.02	0.30	0.19	-0.13	0.10	0.26	0.12	0.16
P_13_5	Wt	0.18	0.30	0.39	0.19	0.36	0.37	0.30	0.09
P_13_5	Wt	-0.14	0.07	0.01	-0.12	0.06	0.01	-0.02	0.09
P_15_7	Wt	0.37	0.42	0.41	0.32	0.32	0.30	0.36	0.05
E7_fs_P	Wt-SDM	0.35	0.62	-0.18	0.58	0.62	0.73	0.45	0.34
P_18_8	Wt	0.67	0.48	0.54	0.55	0.58	0.70	0.59	0.08
P_19_8	Wt	0.22	0.17	0.13	0.21	0.16	0.28	0.19	0.05
P_20_8	Wt	0.14	0.17	0.20	0.23	0.20	0.16	0.18	0.03
P_21_8	Wt-SDM	0.19	0.09	0.16	-0.07	0.15	0.10	0.11	0.09
E8_2fs_P	Wt	0.72	0.75	0.84	0.62	0.61	0.93	0.74	0.13
P_22_8	Wt	-0.03	0.14	0.24	-0.07	0.20	0.29	0.13	0.15
P_22_8	Wt	0.02	0.16	0.22	0.21	0.16	0.20	0.16	0.07
P_26_9	Wt-SDM	0.51	0.34	0.61	0.43	-0.65	0.78	0.34	0.51
P_27_9	Wt	0.34	0.46	0.65	0.33	0.41	0.80	0.50	0.19
P_29_10	Wt	0.28	0.14	0.34	0.21	0.15	0.34	0.24	0.09
P_30_11	Wt	0.13	0.01	0.33	0.11	0.09	0.54	0.20	0.20
P_31_11	Wt	0.66	0.79	0.71	0.81	0.83	0.62	0.74	0.09
P_32_11	Wt	0.63	0.45	0.61	0.44	0.36	0.71	0.53	0.14
E12_trp_P	Wt	0.66	0.69	0.74	0.56	0.52	0.75	0.65	0.10
P_34_13	Wt	0.37	0.35	0.49	0.39	0.51	0.48	0.43	0.07
P_36_13	Wt	0.19	0.28	0.38	0.13	0.30	0.41	0.28	0.11
P_37_13	Wt	0.11	0.06	0.05	-0.01	0.03	0.02	0.04	0.04
P_37_13	Wt	0.26	0.30	0.38	0.18	0.28	0.33	0.29	0.07
P_40_13	Wt	0.27	0.12	0.10	0.13	0.26	0.16	0.17	0.07
P_40_13	Wt	-0.17	0.16	0.02	0.09	0.18	0.03	0.05	0.13

P_43_13	Wt	0.73	0.38	0.36	0.62	0.64	0.42	0.53	0.16
P_43_13	Wt	-0.11	0.06	0.12	0.08	0.14	0.17	0.08	0.10
P_43_13	Wt	0.07	0.33	0.49	0.08	0.62	0.57	0.36	0.24
P_47_14	Wt	0.35	0.44	0.40	0.32	0.39	0.41	0.39	0.04
P_48_14	Wt	-0.31	-0.16	-0.08	-0.19	-0.13	0.01	-0.14	0.11
P_50_14	Wt	-0.26	0.36	0.30	-0.24	0.36	0.25	0.13	0.29
E14_fs_P	Wt	0.13	0.14	0.09	0.05	0.07	0.07	0.09	0.04
P_52_15	Wt	0.18	0.14	0.12	0.15	0.17	0.13	0.15	0.02
P_54_15	Wt-SDM	0.26	0.27	0.24	0.27	0.36	0.03	0.24	0.11
P_54_15	Wt-SDM	0.50	0.43	0.41	0.36	0.35	0.19	0.37	0.11
P_59_15	Wt	0.65	0.41	0.70	0.57	0.57	0.81	0.62	0.14
P_60_15	Wt	-0.40	-0.13	0.34	-0.36	0.15	0.45	0.01	0.36
S_15_Wt	Wt	0.53	0.70	0.59	0.44	0.58	0.60	0.57	0.08
P_61_16	Wt	0.45	0.36	0.37	0.34	0.37	0.45	0.39	0.05
P_62_16	Wt	0.19	-0.03	0.19	-0.05	-0.01	0.15	0.07	0.11
P_63_16	Wt	0.04	0.32	0.19	0.07	0.36	0.19	0.19	0.13
P_64_18	Wt	0.40	0.25	0.36	0.12	0.29	0.39	0.30	0.11
P_65_18	Wt	0.56	0.30	0.17	0.41	0.48	0.11	0.34	0.18
P_67_18	Wt	0.52	0.32	0.43	0.52	0.50	0.50	0.47	0.08
P_68_18	Wt	0.02	-0.08	-0.23	0.07	-0.12	-0.26	-0.10	0.13
P_69_18	Wt	-0.04	-0.14	-0.03	-0.06	-0.15	-0.08	-0.09	0.05
P_70_18	Wt	0.25	0.32	0.23	0.33	0.27	0.22	0.27	0.05
P_71_18	Wt	-0.11	-0.11	-0.36	-0.12	-0.16	-0.43	-0.21	0.14
P_71_18	Wt	-0.27	-0.25	-0.21	-0.20	-0.27	-0.40	-0.27	0.07
P_73_18	Wt	0.29	-0.03	0.21	0.34	0.12	0.33	0.21	0.14
P_74_19	Wt	0.84	0.50	0.66	0.70	0.61	0.63	0.66	0.11
P_76_19	Wt	0.04	0.21	-0.04	-0.01	0.08	0.03	0.05	0.09
P_77_21	Wt	-0.06	0.26	0.52	0.01	0.24	0.31	0.21	0.21

Notes:

- The data values displayed in red indicate outliers.
- The number of probes that display a DS score of > 0.2 are 40
- The number of probe pairs that display consistent DS values across hybridizations are 43.
- The mutation carrying probes in SDM samples are shown as yellow

From the above values, following conclusion about probe were made:

- DS for each probe pair varied from probe to probe, conceivably as a result of altered sequence composition and the length of the probe pair.
- The sensitivity of the array to detect a mutation (discriminate a mismatch probe from the corresponding perfect match probe) is about 65% (40 out of 63).

- A closer examination of the probe sequence and the DS values suggested that in about 10 cases the design could be improved by either increasing or decreasing the length of the probe. In fact in one case it was realized that a wrong probe sequence was used. The improved probe design would improve the sensitivity of the array to nearly 80% (50 out of 63).
- The probe P_37_13, addressed two mutations out of which, the insertion mutation had shown good discrimination (0.28) while the “C->A” mutation within the probe sequence could not be discriminated.

PM - CTGGCCACGCCCA-CGGCTGTCATGGT
 MM1 - CTGGCCACGCaCA-CGGCTGTCATGGT (C->a mutation)
 MM2 - CTGGCCACGCCCAACGGCTGTCATGGT (A insertion)

- Similarly, the probe P_43_13 also addressed three different mutations listed in Table XV below. However, only one of the mutation M_43_13a of 3 base deletion could be differentiated giving DS = 0.53. Other two mutations 13b (SPM) and 13c (SBD) gave poor results. Very low DS = 0.08 for 13b while 13c gave poor reproducibility, though DS= 0.36.
-

Table XV. Same probe in exon 13 addressing three mutations

P_43_13	ACGGCATCCTCATCA-AGGGAGGCAAG	DS
M_43_13a	ACGGCATCCTC---A-AGGGAGGCAAG	0.53
M_43_13b	ACGGCATCCTCATCA-gGGGAGGCAAG	0.08
M_43_13c	ACGGCATCCTCATCAGGGCGAGGCAAG	0.36*

* Showed poor reproducibility

- Remaining probe pairs (10 numbers) happen to reside in a high GC content region and hence could have poor discrimination score.

In summary, it can be concluded that the DS value of the probe depends on the nature of mutation, location of the mutation in the probe and the sequence context of the mutation.

3.15 DS VALUES OF PROBES TESTED BY SDM

The spots obtained from probe pairs that correspond to the mutant locus within each SDM sample are depicted in the following Table XVI. 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. It should be noted that the spots shown here were extracted from one microarray image each (Figures 3.31 to 3.35) and are a representation of the general pattern observed across multiple experiments.

Table XVI : PM and MM probe pairs & DS scores from the mutant loci from simulated samples (SDM) and WT

Mutation type	Average	Wild-type sample		SDM Sample		Average
	DS Score	PM Probe	MM Probe	PM Probe	MM Probe	DS Score
DBI	0.21±0.11					-0.36
DBD	0.26±0.17					-0.76
SBI	0.16±0.05					-0.02
SBD	0.49±0.13					-0.36
SPM1	0.41±0.17					0.08
SPM2b	0.55±0.07					0.03

It can be observed that the spot intensities from MM probes showed high intensities compared to the PM probes at most of the locations after hybridization. A summary of hybridizations with these SDM samples is given in Table XII. The significance of the deviation of the DS values for the mutant samples (6th column) compared to the wild-type (WT) sample was assessed by the Dixon's Q-test (Rorabacher, 1991)⁶³. Q-test probability of computed DS values from WT and SDM samples are given in Table XVII. It was observed that all the simulated WD mutations (except SDM8, that had 90% confidence)

could be detected with better than 95% confidence. The spots corresponding to these hybridizations are highlighted in boxes in scans given in Figures 3.27 to 3.31 and the appearance of the spot intensities are observed to be in line with the quantifications allowing us to draw inferences with respect to their ability to detect mutations of different types.

Table XVII. Dixon's Q-test on DS values

Probe ID	Sequence (5'->3')	Nature of Mutation	DS of WT Sample	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	TAGAGCAAAACCTCAGAAGCCCTGG	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					

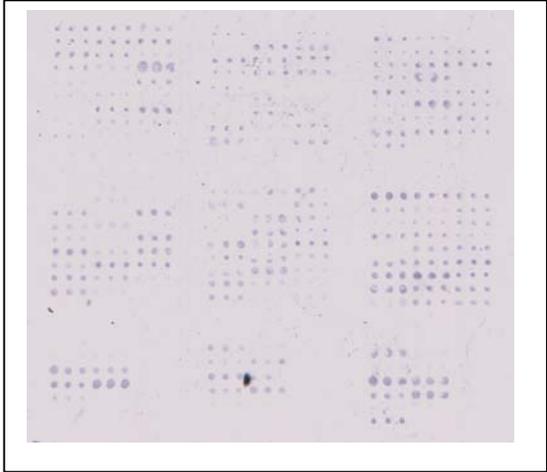
All these experiments demonstrated that the microarrays developed by us have the ability to detect Wilson Disease causing Indian mutations with a sensitivity of more than 65%.

3.16 HYBRIDIZATION EXPERIMENTS WITH PATIENT SAMPLES

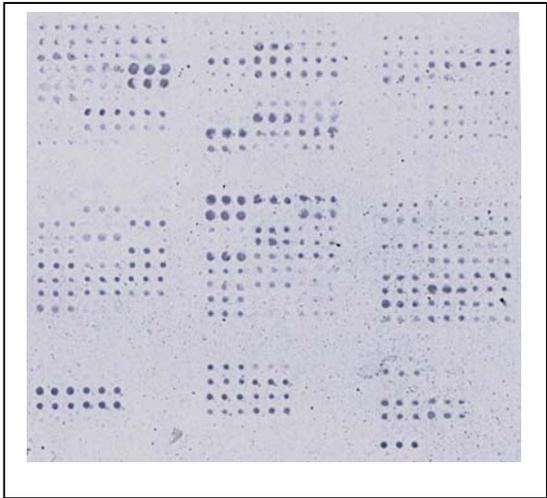
Hybridizations on second version of WDDMs were done for all the samples of gDNA (WD patients and five healthy individuals) as per details available in Table VIII of section 3.3. By this time newer version of Microarrays (Version 3) were printed by reducing the redundancy in the spotted probes. Thus, the layout of spots differed slightly in this version compared to the previous one. Amplicons from patient 9 and family members (P9, P9-1, P9-2 and P9-3) were hybridized to this new version 3 of

Microarrays. Scanned pictures of these arrays are given in Fig. 3.37 A, B, C & D.
below.

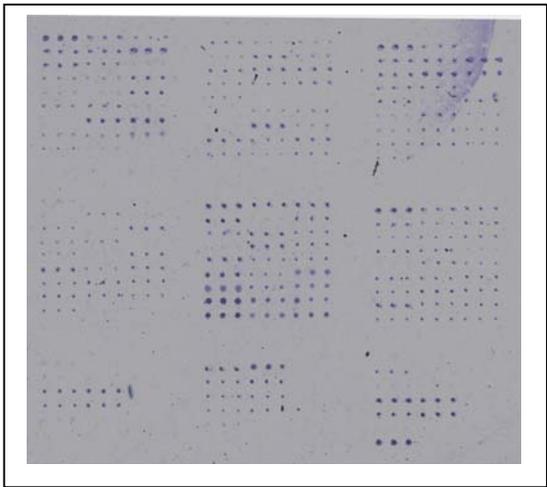
Sample: WD 9
Hybridization temp:
42°C
Post-Hyb Wash: 42°C
Figure 3.37A



Sample: WD 9-1
Hybridization temp:
42°C
Post-Hyb Wash: 42°C
Figure 3.37B



Sample: WD 9-2
Hybridization temp:
42°C
Post-Hyb Wash: 42°C
Figure 3.37C



Sample: WD 9-3
 Hybridization temp:
 42°C
 Post-Hyb Wash: 42°C
 Figure 3.37D

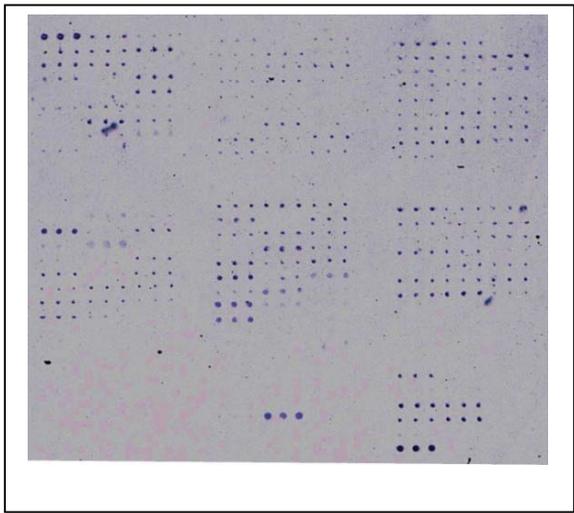


Figure 3.37 Hybridized scan of patient P9 and relatives 9-1, 9-2 & 9-3

In case of patient 9, a mutation like spot intensity pattern was observed for probe P_29_10. Details of spot intensities at this probe position from patient P9 and respective family members is shown in Figure 3.38.

Wild-Type		Father	
Patient-9		Mother	
Patient-9 (Independent Hybridization)		Brother	

Figure 3.38 Hybridization signals from Probe P_29_10 in Patient P-9 and relatives

In the WT sample, the intensity of the PM probe is much higher than that of the mutant probe, whereas, in the patient sample (scans of two arrays) given in Figure 3.40, the MM probe showed higher signal intensity than the PM implying that the patient could be carrying this particular mutation. Comparison of the signals obtained from the other family members suggest that all of them, i.e., father, mother and brother, are heterozygous for this locus and could be carriers of the mutation. A probable pedigree of this family is shown below (Fig. 3.39).

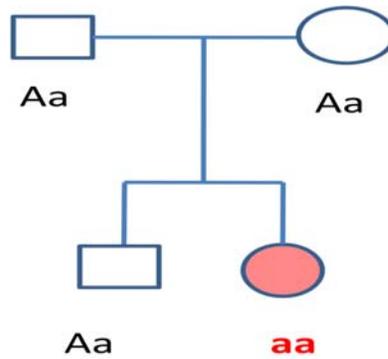


Figure 3.39 Probable genotype of patient family

Next, the DNA sequence from patient P-9 exon 10 was consulted for confirmation of hybridization results. Below are the snap shots from the specific region showing insertion mutation in this patient (Fig.3.40).

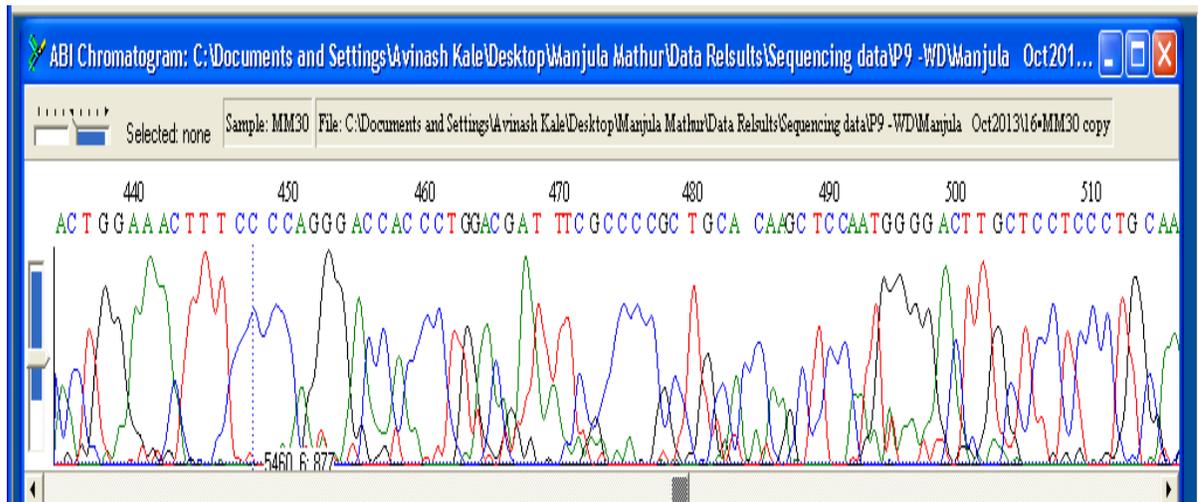


Figure 3.40. Reverse strand electropherogram of exon10 amplicon No.8 from P9 (WD patient). Sequence after position 460 is CCTGGACGATTCGCCCGC

P_29_10 PM – CGATATCGTCAAGG-TGGTCCCTGGG
P_29_10 MM - CGATATCGTCAAGGGTGGTCCCTGGG

The probe P_29_10 (sequence given above) corresponds to a ‘G’ insertion in exon 10 which results in “Val833GlyfsX21” frame-shift in the TM4 domain.

3.17 PATIENT P-12 & FAMILY MEMBERS

Another patient P12 and family members were screened for mutations by hybridizations but no significant changes in spot intensity patterns across any probe pairs were detected. It is possible that this patient might be carrying a hitherto unreported mutation or the mutation (though covered by WDDM) is not detected due to 65% sensitivity of its present version. However, it is also possible that the pathophysiological state of this patient is not due to dysfunctional WND protein. To resolve these issues, sequencing of all the amplicons of patient P12 and family members were taken up and analysed for the presence of mutations. Snap shots of sequence electrophoregrams covering mutation regions of exon 13 (Fig. 3.41 A and B) and exon 18 (Fig.3.42 A,B,C) are given below.

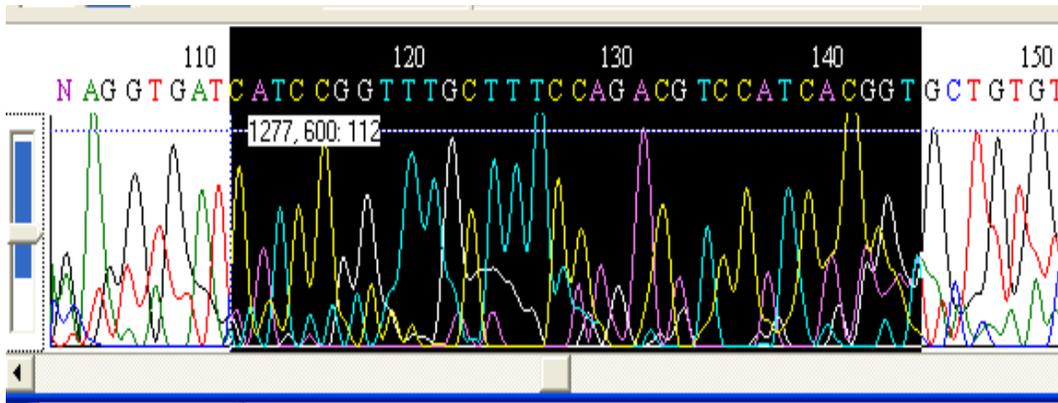


Figure 3.41 A. Snap shot of forward strand of Exon 13. Region of exon covering 2 mutations (Probe ID P_34_13 and P_36_13). Sequence after 112 reads as CATCCGGT referring to WT of probe P_34_13. Sequence after 137 is ATCACGGT referring to WT of probe P_36_13

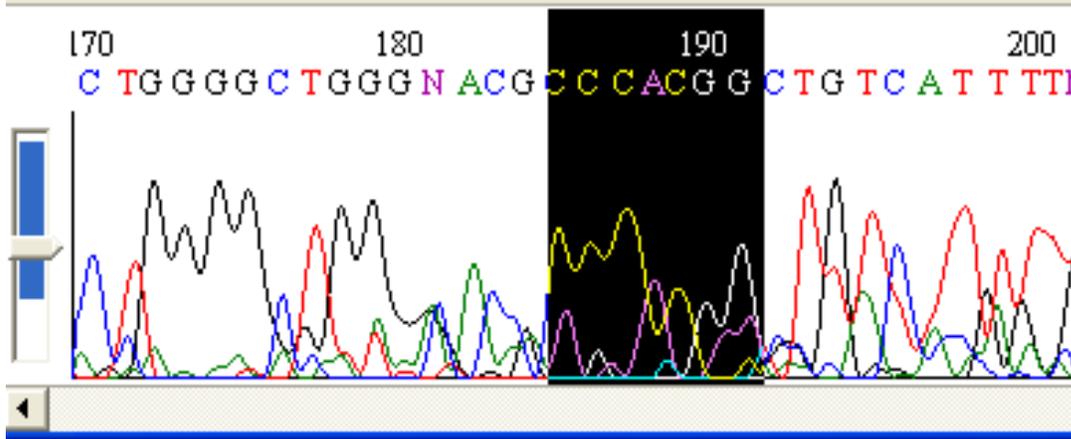


Figure 3.41 B. Snap shot of forward strand of Exon 13. Region of exon covering mutation probe ID P_37_13. Sequence after 184 reads as CCCACGG referring to WT

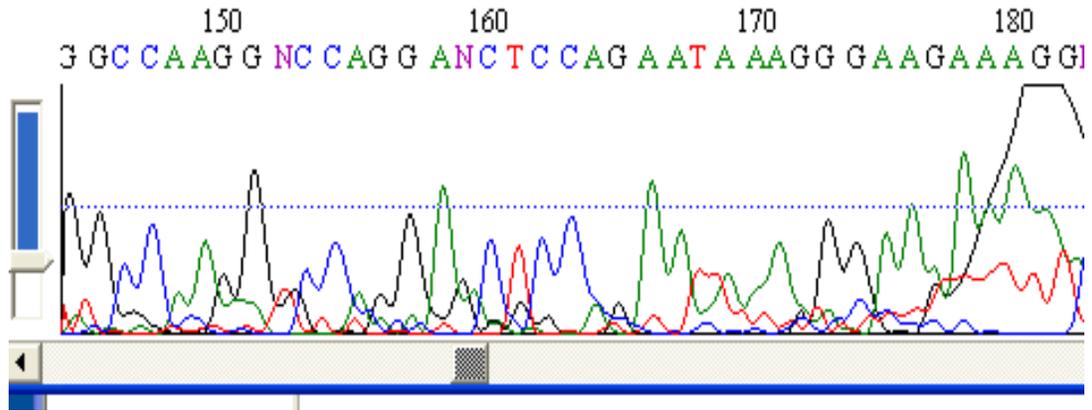


Figure 3.42 A Snap shot of forward strand of Exon 18. Region of exon covering 2 mutations (Probe ID P_65_18 and P_67_18). Sequence at 161 reads as TCCAGAAT referring to WT of probe P_65_18. Sequence at 164 is AGAATAA referring to WT of probe P_67_18

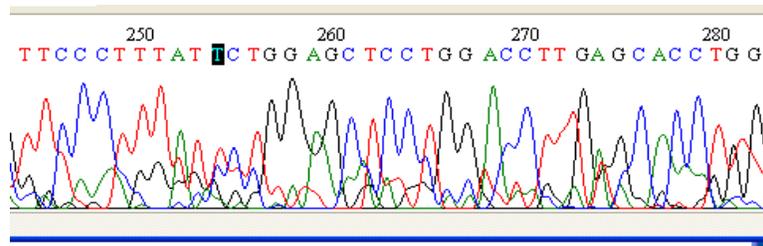


Figure 3.42 B Snap shot of reverse strand of Exon 18. Region of exon covering same two mutations (Probe ID P_65_18 and P_67_18). Sequence at 252 reads as ATTCTGGA referring to WT of probe P_65_18. Sequence at 250 is TTATTCTG referring to WT of probe P_67_18

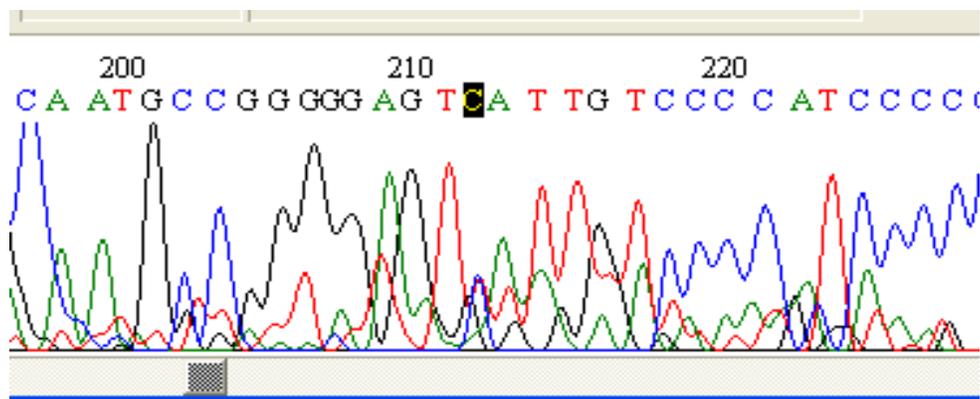


Figure 3.42 C Snap shot of reverse strand of Exon 18. Region of exon covering mutation (Probe ID P_68_18). Sequence at 209 reads as AGTCATTGT/A referring to WT.

3.18 COST ESTIMATION OF DETECTION USING WDDM

An integral part of any developmental model is cost effectiveness of the proposed procedure, its sustainability over a 5 -10 year period and ultimately utilization by the industry. The cost of detection of mutations calculated for the inhouse microarray developed by us is provided for reference below.

Cost of one epoxy slide ~ Rs 800/- to Rs 1000/-

Cost of synthesis of 200 DNA probes = 2 lacs

(The quantity of each DNA probe can print more than 500 slides using in house microarrayer).

Cost of one PCR reaction = ~Rs 10/-

Approximate cost of generating amplicons = Rs 400/-

Cost of DIG labelling pool of PCRed samples = Rs 1000/-

(Of the twenty microliter DIG labelled DNA obtained 5 μ l is required for one hybridization on one WDDM). Three WDDMs are printed per each glass slide. If three repeat hybridizations are done for each gDNA sample, cost of arrays = Rs 1000/-.

Major cost of labelling and of 3 WDDM (one slide) = Rs 1000+Rs1000=Rs2000/-

Cost of rest of the reagents & miscellaneous charges per slide = Rs1000/-

Thus maximum cost of consumables for processing one blood sample for WD detection ~ Rs3000/-

“Only those who see the invisible can do the impossible”

CHAPTER IV

SUMMARY AND CONCLUSIONS

4.1 SUMMARY OF RESULTS

- Prototype microarrays containing one mutation each from nine different exons were used to conclude that immobilization of spotted oligonucleotides on epoxy coated slides could be achieved and such microarrays can be utilized to differentiate mutant alleles from their corresponding wild-type alleles using DIG based hybridizations.
- The Microarray contained probes derived from 63 Indian WD mutations, 4 SNP variants, repeated spots and reverse complements of probes, exon specific controls, blanks (only buffer) and DIG labeled DNA as positive control. Samples were prepared by pooling equal proportions of specific set of wild-type / mutant amplicons (generated by site directed mutagenesis). Samples were labeled with DIG and were hybridized to the in-house printed Microarrays.
- Hybridizations were carried out at 42 and post-hybridization washes were done at two temperatures i.e., 42 °C and 52°C. The data reproducibility was observed to be better at higher temperature and hence, the 52°C data was used for all further analysis.
- The ability of each probe with respect to detecting the mutation was assessed by computing the Discrimination scores from the PM/MM spot intensities for each probe pair.
- All the simulated WD mutations (except SDM8 that had 90% confidence) could be detected with better than 95% confidence.
- The overall sensitivity of the array to detect a mutation is about 65% (40 out of 63) and if we consider all the probes that yielded reproducible DS values ($DS > 2$ s.d) the sensitivity of our WDDM array was better than 80%.

- Some of the probe pairs (10 numbers) happen to reside either in a high GC content region or regions of high secondary structure and have yielded poor discrimination scores.
- Using these arrays, we could detect a ‘G’ insertion mutation (exon 10) that resulted in “Val833GlyfsX21” frame-shift in the TM4 domain in one of the patient sample analyzed. This was further verified by DNA sequencing.

4.2 CONCLUSIONS

Molecular hybridization based DNA microarray methodology was conceptualized and executed to develop DNA Diagnosis for Wilson Disease. This thesis described the in-house preparation of low-density oligonucleotide microarrays (spotted on epoxy coated slides) for the simultaneous detection of 63 Wilson’s Disease causing mutations reported from Indian population. Authentication of the probe design and the sensitivity of the arrays towards mutation detection were assessed from the discrimination scores calculated from the perfect match and mismatch spots. Version 2 WDDM was subjected to extensive analysis of each of the probe pairs which showed close to 65% sensitivity. The ability to detect mutations by this technique depended on the nature and location of the mutation in the probe and the sequence around the mutation.

All of this analysis resulted in the design and development of WDDM version 4, now ready for detection of insertion, deletion and frame-shift mutations in the gene *ATP7B* and is expected to yield better than 80% sensitivity.

If the microarray format is integrated with solid state electronics it would improve the portability of the detection device. Label-free detection of DNA was demonstrated by Alexander Star et. al. in (2006)⁶⁴ using carbon nano-tube network field effect transistors. More recently a semiconductor device in microarray format has been used to demonstrate electronic hybridization by Blin et al. (2014)⁴⁸ for DNA genotyping.

This study demonstrated that designed WDDMs and the protocols followed can be used for confirmation of the disease in suspected patients as a first pass screening tool. Micro-array based diagnosis of WD is available from a few commercial / clinical R&D service providers internationally e.g. Asper Biotech, Ambry Genetics (www.gene-analysis-service.de)^{wr15}. In comparison, this is perhaps the first attempt in this direction in India for the detection of any heritable disorder. In comparison to the widely available sequencing technology, low-density microarray based technique is found to be more efficient and cost-effective. A rough estimate of the cost of testing each suspected patient is about Rs 3000/-. On similar lines arrays can be developed for detecting various haemoglobinopathies and other genetic disorders. Dr. Dipika Mohanty's editorial article in Indian Journal of Medical Research⁶⁵ brought out on the occasion of World Sickle Cell Day (June 19,2014) has discussed the need of screening sickle cell disease (SCD) causing mutations in India. This way the burden of disease to future generations could be reduced in the society.

The present work may be considered a Proof-Of-Concept for mutation detection by use of low-density microarrays. As a part of this study, a website (www.wilsonsdisease.com)^{wr16} was created which can be used to integrate the mutation data with other forms of data for obtaining insights into the patho - physiology of WD.

REFERENCES

1. Progressive lenticular degeneration: a familial disease associated with cirrhosis of liver, Wilson SAK,*Brain*, (1912)*Vol. 34*, 295-507.
2. The WD gene is a putative copper transporting P-type ATPase similar to the Menkes gene,Bull, P.C. et al,*Nature Genetics*, (1993),*Vol. 5*,327-337.
3. Mapping, cloning and genetic characterization of the region containing the WD gene, Petrukhin, K. et. Al,*Nature Genetics*, (1993), *Vol. 5*, 338-343.
4. WD gene is a copper transporting ATPase with homology to the Menkes disease gene. Tanzi, R.E. et al.,*Nature Genetics*, (1993), *Vol. 5*, 344-350.
5. Copper & Biological Health,Krupanidhi S, Sreekumar A and Sanjeevi CB,*Ind. Jou. Med. Res.*, (2008), *Vol.128*, 448-461.
6. Evolution of copper transporting ATPases in eukaryotic organisms,Gupta A and Lutsenko S.,*Curr Genomics.*, (2012), *Vol. 13 (2)*,124-133.
7. Wilson's disease and the copper ATPase Transporter. Chang H-Y and Alex Mitchel. *Inter Pro Protein Focus* (2013)
8. Wilson disease, *Best Practice & Research, Huster D*, (2010), *Clinical Gastroenterology*, *Vol. 24*, 531–539.
9. A Novel Global Assessment Scale for Wilson's Disease (GAS for WD) Movement Disorders, Aggarwal A, NitinAggarwal, AabhaNagraal, GovindjiJankharia, and Mohit Bhatt, (2009),*Vol. 24 (4)*, 509–518. © 2008 Movement Disorder Society
10. Isolation and characterization of a human liver cDNA as a candidate gene for Wilson disease, Yamaguchi Y, Heiny M E and Gitlin J D, (1993),*Biochem. Biophys. Res. Commun. Vol.197 (1)*,271-277.
11. Wilson disease: genetic basis of copper toxicity and natural history,Schilsky M L, (1996),*Semin Liver Dis.*, *Vol. 16*,83–95.
12. Iron and copper in mitochondrial disease, Wenjing Xu, Tomasa Barrientos, Nancy C.

- Andrews, (2013), *Cell Metabolism*, Vol. 17 (3), 319-328
13. Copper and Zinc, Biological Role and Significance of Copper/Zinc Imbalance, Osredkar J and Sustar N, (2011), *J Clin Toxicol*, S3:001. doi:10.4172/2161-0495.S3-001
 14. Copper transporters regulate the cellular pharmacology and sensitivity to Pt drugs, Safaei R and Howell S B, (2005), *Critical Reviews in Oncology/Hematology*, Vol.53, 13-23
 15. Characterization of the Wilson Disease gene encoding a P-type copper transporting ATPase: genomic organization, alternative, splicing and structure/ function predictions, Petrukhin K, Lutsenko S, Chernov I, et al., (1994), *Hum Mol Genet*, Vol. 3, 1647–56.
 16. Expression of the human Menkes ATPase in *Xenopus laevis* oocytes, Bissig KD, La Fontaine S, Mercer JF, Solioz M., (2001), *Biol Chem*, Vol. 382, 711–714.
 17. Human Menkes X-chromosome disease and the staphylococcal cadmium-resistance ATPase: a remarkable similarity in protein sequences. Silver S, Nucifora G, Phung LT., (1993), *Mol Microbiol*, Vol.10, 7–12.
 18. CPx-type. ATPases: a class of p-type ATPases that pump heavy metals, Solioz M, Vulpe C, (1996), *Trends Biochem Sci*, Vol. 21, 237–41.
 19. Function and regulation of the mammalian copper-transporting ATPases: insights from biochemical and cell biological approaches, Lutsenko S, Petris MJ., (2003), *J Membr Biol*, Vol. 191, 1–12.
 20. Endosomal trafficking of the Menkes copper ATPase ATP7A is mediated by vesicles containing the Rab7 and Rab5 GTPase proteins, Pascale MC, Franceschelli S, Moltedo O, et al, (2003), *Exp Cell Res*, Vol.291, 377–385.
 21. Essential role for mammalian copper transporter Ctr1 in copper homeostasis and embryonic development, Lee J, Prohaska JR and Thiele DJ., (2001), *Proc Natl Acad Sci U S A*. Vol. 98 (12), 6842-6847.
 22. The molecular basis of copper-transport diseases, Mercer, J. F. B., (2001), *Trends in Molec. Med.*, Vol. 7, 64–69.

23. Distinct phenotype of a Wilson disease mutation reveals a novel trafficking determinant in the copper transporter ATP7B, Lelita T. Braiterman, Amrutha Murthy, Samuel Jayakanthan, Lydia Nyasae, Eric Tzeng, Grazyna Gromadzka, Thomas B. Woolf, Svetlana Lutsenko, and Ann L. Hubbard, (2014), *PNAS Vol. 111 (14)*, E1364–E1373. *doi: 10.1073/pnas.1314161111*.
24. The Role of the Invariant His-1069 in Folding and Function of the Wilson's Disease Protein, the Human Copper-transporting ATPase ATP7B, Ruslan Tsivkovskii, Roman G. Efremov and Svetlana Lutsenko, 2003, *Journal of Biological Chemistry, Vol. 278, 13302-13308*. DOI: 10.1074/jbc.M300034200
25. A structural model of the copper ATPase ATP7B to facilitate analysis of Wilson disease-causing mutations and studies of the transport mechanism, Schushan M, Bhattacharjee A, Ben-Tal N & Lutsenko S, (2012), *Metallomics, Vol. 4, 669–678*. *doi: 10.1039/c2mt20025b*
26. SIFT web server: predicting effects of amino acid substitutions on proteins, Ngak-Leng Sim, Prateek Kumar, Jing Hu, Steven Henikoff, Georg Schneider, and Pauline C. Ng, (2012) *Nucleic Acids Res. Vol. 40 (Web Server issue)*, W452–W457. **DOI:** 10.1093/nar/gks539
27. Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes, P de Bie, P Muller, C Wijmenga, and L W J Klomp, (2007), *J Med Genet., Vol. 44 (11)*, 673-688. **DOI:** 10.1136/jmg.2007.052746
28. Sequence variation database for the Wilson disease copper transporter, ATP7B. Kenney S M and Cox D W., (2007), *Hum Mutat., Vol. 28 (12)* : 1171-1177.
29. Wilson Disease Mutation Pattern with Genotype-Phenotype Correlations from Western India: Confirmation of p.C271* as a Common Indian Mutation and Identification of 14 Novel Mutations, Anu Aggarwal, Gursimran Chandhok, Theodor Todorov, Saloni

- Parekh, SharadaTilve,AndreeZibert, Mohit Bhatt,and Hartmut H.-J. Schmidt, (2013),Annals of Human genetics, *Vol.77 (4)*, 299-307.**DOI: 10.1111/ahg.12024**
30. Methods for detection of point mutations: performance and quality assessment, Nollau P, and Wagener C. (1997) *Clinical Chemistry, Vol.43 (7)*,1114-1128.
 31. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Newton C R, Graham A, Heptinstall L E, Powell S J, Summers C, Kalsheker N, et al., (1989), *Nucleic Acids Res.,Vol. 17*, 2503-2516.
 32. CCR: a rapid and simple approach for mutation detection, Wanli Bi and Stambrook P J., (1997) *Nucleic Acids Research,Vol.25 (14)*, 2949–2951.
 33. DNA differing by a single base-pair substitution are separated in denaturing gradient gels. Correspondence with melting theory, Fischer, S. G., and L. S. Lerman., (1983), *Proc. Nad. Acad. Sci. USA.,Vol.80*, 1579-1583.
 34. **Book Title:** Molecular Diagnosis of Genetic Diseases, **BookSeries:**Methods in Molecular Medicine, Old John M, (1996), *Vol. 5*, 169-183, **DOI: 10.1385/0-89603-346-5:169**
 35. **Book Title:** Pediatric Hematology : Methods and Protocols **Series:** Methods in Molecular Medicine , Old John M, (2003) *Vol. 19*, 33-62, **DOI: 10.1385/1-59259-433-6:33**
 36. Mutation scanning using high-resolution melting,Taylor, C. F, (2009),*Biochemical Society Transactions, Vol.37*, 433–437.
 37. Mutation detection by Real-Time PCR: A simple and highly selective method, Morlan J, Baker J and Sinicropi D., (2009), *PLos ONE , Vol.4 (2)*,e4584.
 38. High-resolution DNA melting analysis: advancements and limitations,Wittwer CT., (2009), *Human Mutation, Vol.30 (6)*, 857–859.
 39. Single-base mutational analysis of cancer and genetic diseases using membrane bound modified oligonucleotides,Zhang, Y., Coyne, M.Y., Will, S.G., Levenson, C. H.,

- and Kawasaki, E.S., (1991), *Nucleic Acids Res.*, *Vol.19 (14)*, 3929–33.
40. Applications of DNA microarray in disease diagnostics, Yoo, Seung Min, et al., (2009), *J MicrobiolBiotechnol* , *Vol. 19 (7)*, 635-646.
41. Use of cDNA microarray to analyse gene expression patterns in human cancer, Joseph DeRisi, Lolita Penlan, Patrick O. Brown, Michael L. Bittner, Paul S. Meltzer, Michael Ray, Yidong Chen, Yan A. Su and Jeffrey M. Trent., (1996), *Nature Genetics*, *Vol. 14*, 457-460.
42. “Microarray-Based Mutation Detection and Phenotypic Characterization of Patients with Leber Congenital Amaurosis” Yzer et al (total 19 authors), (2006), *Investigative Ophthalmology & Visual Science*, *Vol.47 (3)*,1167-1176.
43. “Development of a genotyping microarray for Usher syndrome”, Cremers et al (total 30 authors), (2007),*J Med Genet* , *Vol.44*, 153–160.
44. Recognition of a C-C mismatch in a DNA duplex using a fluorescent small molecule with application for "off-on" discrimination of C/G mutation, Hu L, Wang Y, Wang W, Gao Q, Qi H, Zhang C., (2012), *ApplSpectrosc*, *Vol.66 (2)*, 170-174.
45. “A highly specific microarray method for point mutation detection”,Yasser Baaj, Corinne Magdelaine, VirginieUbertelli, Christophe Valat, Luc Talini, Françoise Soussaline, Elena Khomyakova, BenoîtFunalot, Jean-Michel Vallat, and Franck G. Sturtz, (2008), *BioTechniques*, *Vol. 44*,119-126
46. Comprehensive and Efficient HBB Mutation Analysis for Detection of - Hemoglobinopathies in a Pan-Ethnic Population,Owen T. M. Chan, Kenneth D. Westover, Lisa Dietz,James L. Zehnder, and Iris Schrijver, (2010),*Am J ClinPathol*, *Vol. 133*, 700-707. **DOI: 10.1309/AJCP7HQ2KWGHECIO**
47. Abnormalities in spontaneous abortions detected by G-banding and chromosomal microarray analysis (CMA) at a national reference laboratory, Wang Boris T, Thomas P Chong, Fatih Z Boyar, Kimberly A Kopita, Leslie P Ross, Mohamed M El-Naggar,

- Trilochan Sahoo, Jia-Chi Wang, Morteza Hemmat, Mary H Haddadin, Renius Owen, Arturo L Anguiano, (2014), *Mol Cytogenet*, *Vol.22*, 7:33.
48. Electronic hybridization detection in microarray format and DNA genotyping, Blin A, Cissé I & Bockelmann U, (2014), *Sci Rep*, *Vol.4*, 4194. **DOI: 10.1038/srep04194**
 49. Genotyping microarray as a novel approach for the detection of ATP7B gene mutations in patients with Wilson disease, Gojova L, Jansova E, Kulm M, Pouchla S, Kozak L., (2008), *Clin Genet*, *Vol.73*, 441-452. **DOI: 10.1111/j.1399-0004.2008.00989.x**
 50. Novel deletion mutation within the carboxyl terminus of the copper-transporting ATPase gene causes Wilson disease, Majumdar R, Al Jumah M, Al Rajeh S, Fraser M, Al Zaben A, Awada A, Al Traif I, Paterson M A., (2000), *J Neurol Sci.*, *Vol.179 (S 1-2)*, 140-143.
 51. Molecular pathogenesis of Wilson disease: haplotype analysis, Detection of prevalent mutations and genotype-phenotype correlation in Indian patients, Gupta A, Aikath D, Neogi R, Datta S, Basu K, Maity B, Trivedi R, Ray J, Das SK, Gangopadhyay PK, Ray K., (2005), *Human Genet*, *Vol. 118*, 49-57.
 52. Molecular Pathogenesis of Wilson Disease Among Indians: A Perspective on Mutation Spectrum in ATP7B gene, Prevalent Defects, Clinical Heterogeneity and Implication Towards Diagnosis, Gupta A, Ishita Chattopadhyay, Sumit Dey, Poonam Nasipuri, Shyamal K. Das, Prasanta K. Gangopadhyay, Kunal Ray, (2007), *Cell Mol Neurobiol*, *Vol.27 (8)*, 1023-33. **DOI 10.1007/s10571-007-9192-7**
 53. Efficient Detection of Mutations in Wilson Disease By Manifold Sequencing, Waldenstrom E, Lagerkvist A, Dahlman T, Westermark K, Lendegren U, (1996), *Genomics*, *Vol. 37*, 303-309.
 54. Genetic, metabolic and cellular factors influencing intracellular localization of the Wilson disease protein, ATP7B, Gupta A, Ashima Bhattacharjee, Nesrin Hasan, Lita Braiterman, Svetlana Lutsenko, Ann L Hubbard., (2014), *Molecular Cytogenetics* .

Vol.7 (Suppl 1),P68.

55. ATP7B mutations in families in a predominantly Southern Indian cohort of Wilson's disease patients, Santhosh, S., Shaji, R. V., Eapen, C. E., Jayanthi, V., Malathi, S., Chandy, M., Stanley, M., Selvi, S., Kurian, G., & Chandy, G. M., (2006), *Indian J Gastroenterol, Vol. 25, 277–282.*
56. Genotype phenotype correlation in Wilson's disease within families-a report on four south Indian families, S Santhosh, RV Shaji, CE Eapen, V Jayanthi, S Malathi, P Finny, N Thomas, M Chandy, G Kurian, GM Chandy, (2008), *World J Gastroenterol, Vol. 14 (29), 4672-4676.*
57. Mutation Profile in Wilson's Disease from North Indian Patients, S. Khan, M. Behari, S. Vivekanandhan, V. Goyal and B. K. Thelma, (2012), *Int J Hum Genet, Vol. 12 (2), 125-132*
58. R778L, H1069Q, and I1102T mutation study in neurologic Wilson disease, Jayantee Kalita, Bindu I Somarajan, Usha K Misra, Balraj Mittal, (2010), *Neurology India, Vol. 58 (4), 627-630*
59. Identification of one novel and nine recurrent mutations of the ATP7B gene in 11 children with Wilson disease, Juan Geng, Jian Wang, Ru-En Yao, Xiao-Qing Liu, Qi-Hua Fu, (2013), *World J Pediatr, Vol 9 (2), 158-162.*
60. A rapid non-enzymatic method for the preparation of the of HMW DNA from blood for RFLP studies, Debomoy K. Lahiri and John I. Numberger, Jr., (1991), *NAR, Vol.19 (19),5444.*
61. Algorithms for high-density oligonucleotide array, Yingyao Zhou & Ruben Abagyan, 2003, *Current Opinion in Drug Discovery & Development, Vol.6 (3),339-345*
62. Development of a 3-Axes Robotic System for Making DNA Microarrays, Sengar, R.S., K. D Lagoo, AVSS Narayana Rao, R. K. Puri and Manjit Singh, (2004), *International conference on Trends in Industrial Measurement and Automation, Chennai, India, 33-39.*

63. Statistical Treatment for Rejection of Deviant Values: Critical Values of Dixon's "Q" Parameter and Related Subrange Ratios at the 95% Confidence Level, Rorabacher, D.B., (1991) *Analytical Chemistry*, Vol. 63, 139-46.
64. Label-free detection of DNA hybridization using carbon nanotube network field-effect transistors, Alexander Star, Eugene Tu, Joseph Niemann, Jean-Christophe P. Gabriel, C. Steve Joiner, and Christian Valcke, 2006, *PNAS*, Vol. 103(4), 921-926.
65. A century after discovery of sickle cell disease: keeping hope alive, Deepika Mohanty, 2014, *Indian J Med Res.*, Vol (139), 793-795

WEBSITE REFERENCES:

1. <http://www.wilsondisease.med.ualberta.ca/references.asp>
2. <http://www.ebi.ac.uk/interpro>
3. <http://sift-dna.org>
4. www.wilsondisease.org
5. www.mayomedicallaboratories.com/testcatalog/Clinical+and+Interpretive/83697
6. <http://www.hgmd.org>
7. <http://www.umd.be/ATP7B/>
8. www.wjpch.com
9. <http://www.dorak.info/genetics/popgen.html>
10. http://en.wikipedia.org/wiki/RNA_splicing
11. <http://www.idtdna.com/Scitools/Applications/Primerquest>
12. <http://genome.ucsc.edu>
13. www.compbio.pbworks.com
14. <http://rsb.info.nih.gov/ij/>
15. www.gene-analysis-service.de
16. www.wilsonsdiseaseresearch.com

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 & Health 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 metabolism¹ 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 DNA². WD occurs at a frequency of 1 in 30,000 whereas its carrier frequency is estimated to be 1:90³. More than 500 different disease causing mutations in *ATP7B* have been reported⁴ and are compiled as a database⁵. In addition to DNA sequencing, several assays have been developed for detection of point mutations⁶. 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) electrophoresis¹⁰. High resolution melting¹¹ (HRM) and quantitative real-time PCR¹² are some of the techniques that do not require a separation step for the detection of genetic differences. Methods based on nucleic acid hybridization¹³ including microarrays of various types are gaining popularity due to the high throughput they offer¹⁴.

For the diagnosis of Wilson's disease arrayed primer extension arrays (APEX)¹⁵, hetero-duplex gel analysis¹⁶ and DNA sequencing¹⁷ have been used, while haplotype analysis has been used to identify the underlying common mutations among Indian patients¹⁸. Most of the patients with WD have been found to be compound heterozygotes (two different mutations) and the mutations seem to be population specific¹. 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 ^s	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

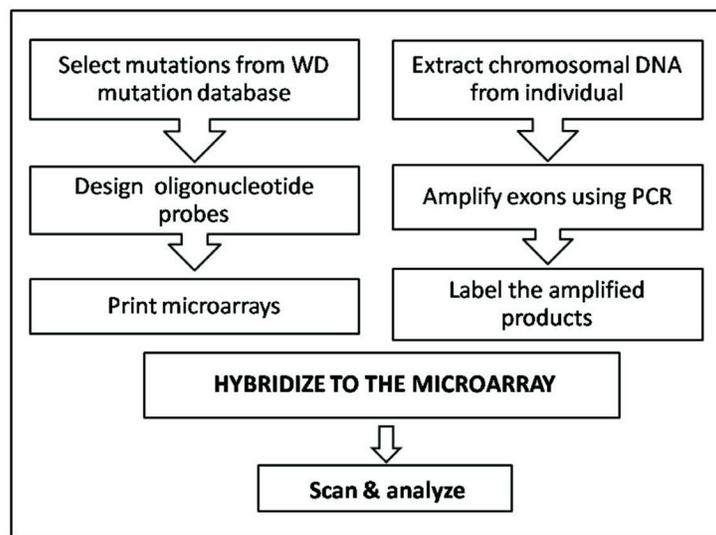


Fig. 1. Flowchart showing the overall strategy used in this study.

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	CTCACCTATAACCACCATCCAGGAG
3	4	291	TGACCTGATGGTTCCAGGTGTTCT	ACAAAAACCAGACACGTCCAAGATGG
4	5	495	AGTCCAGGGTCTTGAGAGCAGT	TTCACTGATATCCTCCCTCAGATTA
		455	TCTTGGCTGCCTGTTACCTAGACT	
5	7	496	ATGTTTAGACCTCTAGATGCTCCCT	GAAAGCTGCAATAAAGTGCCAT
		288	ATCCAGGTGACAAGCAGCATCTGA	ATATCTGAGGGCCACACACAGCAT
6	8	555	CAGTAGTCTCTGAATGGGAAAGTA	GCACCTTAATTATATGGAGGTTTCC
		482	GCACAAAGCTAGAGGCTTTGCCAT	
7	9	392	CATGTGTGGTGGATAGCAAGTAAC	GAGTGGTGATCTTACTGTGTCTCTG
		297	AGCTGTCTCTAACACCACGCTTGT	TCTGCCACACTCACAAGGTCTAT
8	10/ 11	916	TGATAAGTGGCGTTTGTTCAGGG	CTACTCTGGCTTAGATTTTGCTGTC
		546		GGGATAATCTCCTTCATTTAACAC
9	12	916	TGATAAGTGGCGTTTGTTCAGGG	CTACTCTGGCTTAGATTTTGCTGTC
		324	AAGAGTTCTGGGAAATCAGACAGTT	ACAACCACCATATAGCCCAAGGCA
10	13	500	TGTGGAATACCATCTGTTCCG	GGCTACTCTGTTGCTACTGTTGTTA
		431		TGTCTTGAGTGGCTCTCAGGCTTT
11	14	419	TAGGAAGCTGTGCAGGTGTCTTGT	TTCCAGACCACACAGAGAAGGCT
		370	AAGTTCTGCCTCAGGAGTGTGACT	
12	15	497	TTAACCTTTCCTATCTGTTCCACCT	GAACATAAGAGAACTTTCCTGGGT
		393		TCTGTGGTTTGACCCACCTCTACT
13	16	396	GGTGCTTACAAGGTTACAGTTTTTC	CCTGAAATTAAGAGAGGAAGGCT
14	18	479	AAGTCTTTCAGAGGTGCTGCCTT	AGGTTGATGCGTATCCTTCGGACA
		458	GTTGACCAACATCACTGACTGG	ACAGTCCTCTGGAAAGGTGAAT
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'-GTGCCAGGCTGTGGAAGTGGGTGCA-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 microarrays¹⁹. 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). In spite 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

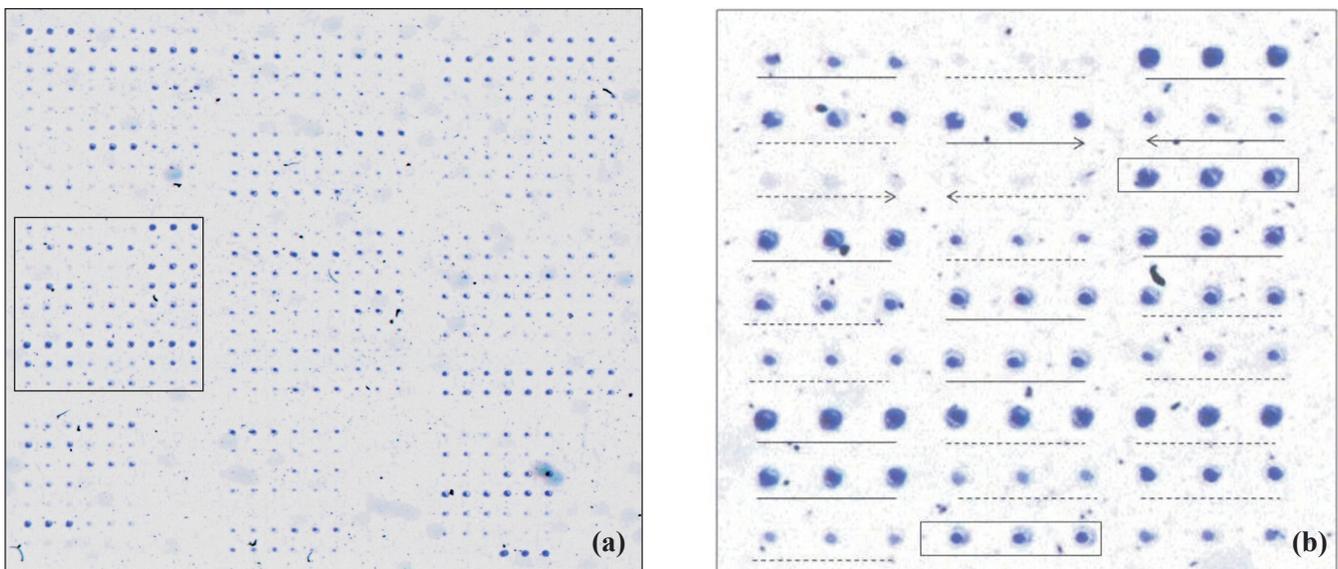


Fig. 2(a). Magnified view of a scanned microarray (13.5 x 13.5 mm) hybridized to wild-type human genomic DNA sample [inset is shown in Fig.2(b)]. **2(b).** A portion of the array wherein, the spots corresponding to wild-type and mutant probes are shown as underlined with solid or dashed lines, respectively. 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.

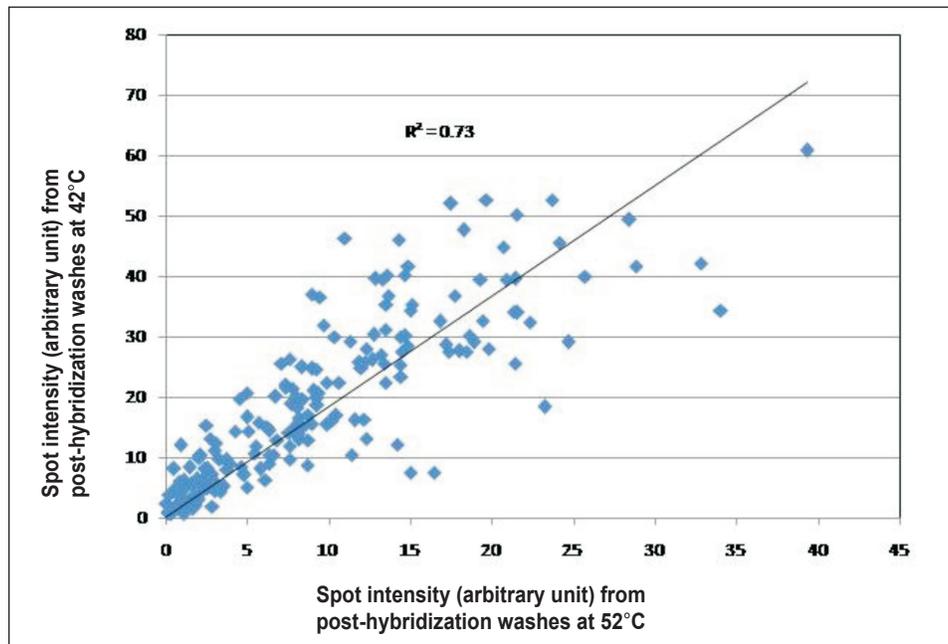


Fig. 3. A scatter plot of the spot intensities (arbitrary units) obtained from post-hybridization washes done at 52°C (X-axis) as compared to the spot intensities obtained from the corresponding spots at 42°C (Y-axis). The correlation coefficient R^2 for these data is indicated in the Figure.

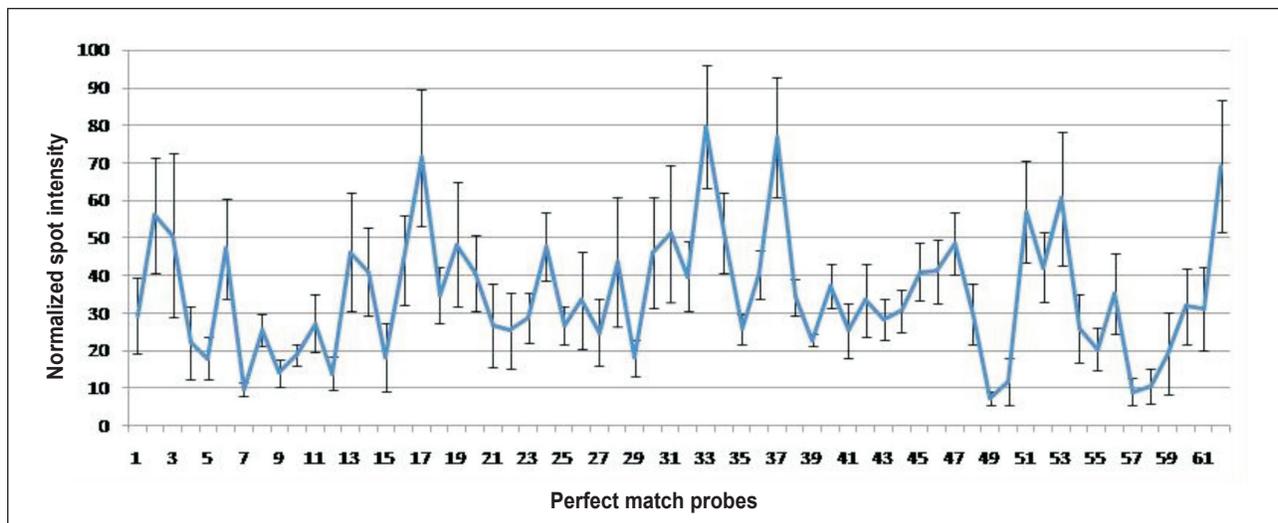


Fig. 4. Normalized spot intensities obtained for 61 perfect match probes from six independent hybridizations. The error bars indicate one sigma (Standard Deviation) from the mean.

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

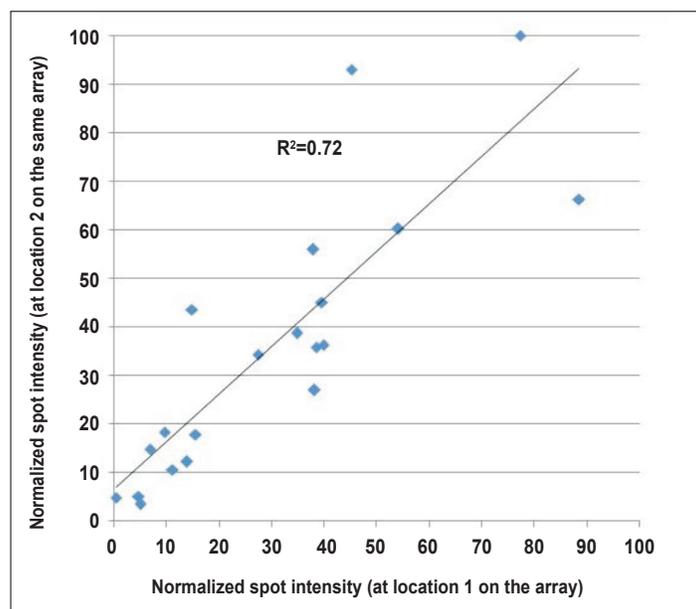


Fig. 5. Correlation observed for the mean intensity data for 20 different probe spots (triplets) spotted at two different locations on the same array. Hybridization was done with wild-type chromosomal DNA sample.

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	TAGAGCAAAACCTCAGAAGCCCTGG	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-test²⁰

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 ($DS > 2\sigma$). 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-test²⁰. 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 scale²²⁻²⁴ but also at individual locus level. A microarray based approach was found to be cost-effective compared to sequence analysis²⁵, capillary based heteroduplex analysis²⁶ or re-sequencing²⁷. Microarrays to screen for 301 disease-associated sequence variants in Leber Congenital Amaurosis (LCA) related genes²⁸, to study sporadic non-syndromic hearing loss in children²⁵ and a genotyping microarray for 298 Usher syndrome-associated sequence variants²⁷ 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

- Huster D. Wilson disease, best practice & research. *Clin Gastroenterol* 2010; 24 : 531-9.
- Hung I H, Suzuki M, Yamaguchi Y, Yuan D S, Klausner R D, Gitlin J D. Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* 1997; 272 : 21461-6.
- Patil M, Sheth KA, Krishnamurthy AC, Devarbhavi H. A review and current perspective on Wilson disease. *J Clin Exp Hepatol* 2013; 3 : 321-36.
- 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.
- Kenney SM, Cox DW. Sequence variation database for the Wilson disease copper transporter, *ATP7B*. *Hum Mutat* 2007; 28 : 1171-7.
- Nollau P, Wagener C. Methods for detection of point mutations: performance and quality assessment. *Clin Chem* 1997; 43 : 1114-28.
- Newton CR, Graham A, Heptinstall L E, Powell S J, 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.
- Bi W, Stambrook PJ. CCR: a rapid and simple approach for mutation detection. *Nucleic Acids Res* 1997; 25 : 2949-51.
- Fischer SG, Lerman LS. DNA 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.
- 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.
- Wittwer CT. High-resolution DNA melting analysis: advancements and limitations. *Hum Mutat* 2009; 30 : 857-9.
- Morlan J, Baker J, Sinicropi D. Mutation detection by real-time PCR: A simple and highly selective method; *PLoS One* 2009; 4 : e4584.
- 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.
- Yoo, SM, Choi JH, Lee SY, Yoo NC. Applications of DNA microarray in disease diagnostics. *J Microbiol Biotechnol* 2009; 19 : 635-46.
- 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.
- Majumdar R, Al Jumah M, Al Rajeh S, Fraser M, Al Zaben A, Awada A, *et al*. 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 S K, Gangopadhyay P K, *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 Devel* 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. 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 : 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, Voeselek K, *et al.* Microarray-based mutation detection and phenotypic characterization of patients with Leber Congenital Amaurosis. *Invest Ophth 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

Commentary

Low-density oligonucleotide microarrays - A major step in Wilson's disease diagnosis

Wilson's disease (WD) is a monogenic disorder of copper accumulation caused due to mutation in the copper transporting ATPase gene *ATP7B*^{1,2}. WD follows an autosomal recessive mode of inheritance with a world prevalence of 1 in 5000 to 1 in 30,000 live births and a carrier frequency of 1 in 90³. Though a single gene disorder, mutation detection in WD has been difficult, mainly due to two reasons: (i) size of the *ATP7B* gene is large (>6.5 kb) with 21 exons, and (ii) besides the prevalent mutations, extensive presence of rare mutations throughout the entire gene. To screen the entire gene one has to undertake amplification and sequencing strategy encompassing 24-26 exonic regions including intron-exon boundaries^{4,5}. This requires substantial time and effort. The second problem is heterogeneity of mutations *i.e.* the occurrence of hundreds of rare (non-prevalent) mutations in various world populations. Moreover, there also exists a huge variability in the prevalent mutations in these populations.

Indian population vs major world populations

Among the data available for a very few major world populations, H1069Q and E1064A are the major founder mutants in Caucasians⁶, R778L and P992L among the East Asians that includes Japanese, Korean and Chinese populations⁷⁻⁹. Among patients from Brazil, c.3402delC mutation had the highest frequency (30.8%), followed by the missense change, c.2123T>C (p.L708P) (16.7%)¹⁰. The Indian population is highly heterogeneous, comprising four major linguistic groups *i.e.* Indo-European, Dravidian, Tibeto-Burman and Austro-Asiatic, and 4693 communities with several thousand endogamous groups (The Indian Genome Variation Database)¹¹. Due to high genetic heterogeneity in the Indian population, prevalent mutations for a specific sub-population might not be present in others,

even if they are subgroups of the same larger linguistic group.

The data for WD mutations from India are quite divergent reflecting the ethno-genetic diversity of this large country. None of the common mutations detected has a prevalence over 25 per cent of all *ATP7B* mutations. In patients from southern India, c.G3182A (p.G1061E) in 16 per cent and c.C813A (p.C271X) in 12 per cent accounted for major *ATP7B* mutations¹², whereas p.E122fs (10.6%) and C271X detected in western Indian patients were the prevalent ones¹³. Patients from eastern and northern India harboured five prevalent mutations. Three mutations c.813C>A (Cys271Stop), c.1708-1G>C (IVS4-G>C) (splice site mutation) and c.3182 GGG>GAG (Gly1061Glu) were found to be associated with 16, 8.5 and 8 per cent of all the mutant chromosomes^{5,14}.

To summarize, multiple studies from different zones of India suggest that detecting WD by screening an individual for one major founder mutation would not be fruitful. A comprehensive and efficient strategy to screen a panel of multiple mutations (both common and rare) in a single multiplex reaction would address the issue.

Genetic diagnosis of WD has a tremendous advantage

Disease onset varies from as early as age of 2 years to as late as the 7th decade of life¹⁵. Early or presymptomatic diagnosis can alleviate disease progression and in some cases entirely prevent onset of disease symptoms. However, there is no confirmatory biochemical or clinical tests available for early disease detection. In the last 10 years major effort has been made in genetic diagnosis of the disease. Roberts and Schilsky laid guidelines for WD diagnosis approved by the American Association for the Study

of Liver Diseases (AASLD)¹⁶. Besides, typical clinical symptoms, such as presence of Keyser Fleischer ring, low serum ceruloplasmin (<20 mg/dl) and 24 h urinary copper excretion (>40 µg), molecular testing to determine the mutation has been considered crucial for disease diagnosis¹⁶. Also, genetic testing can be carried out to determine the mutation status in neonates, much before the onset of any clinical symptoms. Further, detection of carrier status of an individual can ensure proper genetic counselling. To date, major advances have been made to identify the disease status in siblings of WD patients. Gupta *et al*^{14,17} have utilized heterozygous microsatellite markers and informative single nucleotide polymorphisms (SNPs) to identify the mutant chromosome segregating within a WD family. The advantage of this study lies in rapid identification of presymptomatic individuals. However, in such studies the specific mutation often goes undetected. In some cases, high genetic homogeneity or consanguineous marriage results in non-informativeness of the SNP or microsatellite markers. The present study by Mathur *et al*¹⁸ further improves WD diagnosis by identifying specific mutations and not relying on informativeness or heterozygosity of markers.

Introducing low-density microarray for WD detection

In this issue the study by Mathur *et al*¹⁸ on microarray based diagnosis of WD is a major development not only in the area of genetic diagnosis of WD but also for monogenic disorders as a whole. The study attempts to establish detection and analysis procedures for microarray based diagnosis of WD. Presymptomatic identification by mutation detection would be now possible in neonates.

Mathur *et al*¹⁸ describes designing and utilizing an oligonucleotide microarray for detecting WD mutations. A subset of 62 mutations that were specific to Indian population was identified in the Wilson disease database and was used to design oligonucleotide probes for a low-density WD microarray. The entire process is simple but effective and can be split into two parts. Part I involves selection of mutations from the WD database, designing oligonucleotide probes and printing them as microarrays. The second part involved designing DNA amplicons to be tested. The amplicons were DIG labelled. As a proof of principle, the test amplicons were hybridized on the microarrays containing the probes, washed and then analyzed for binding. Around 60 per cent of the mutations showed

high detectability. On including weakly hybridized but reproducible amplicons, the detection rate rose to 80 per cent. Moreover, prevalent mutations in the Indian population, *i.e.* Tyr187Stop, Cys271Stop, Gly1061Glu, c.1708-1G>C were detected in the tests against controls. Authors also included ambivalent variants, *e.g.* c.2623G>A (Gly875Arg) that exhibit dual status in the WD database. Presence of this variant is linked to disease causation (DV) in southern Indian WD patients; however, is designated as a non-disease variant (NDV) in Han-Chinese population^{19,20}. *In vitro* studies carried out in cell lines have demonstrated that copper content of the cell dictates subcellular localization of the Arg⁸⁷⁵ variant and hence affecting intracellular copper content²¹. Early detection of these variants of ambiguous status, early in life of individuals will certainly help to take proper lifestyle and nutritional measures and possibly prevent the disease.

In a nation like India with great genetic diversity, inexpensive genetic testing of common as well as rare mutations would prove to be a blessing. Efforts should be made to improve accuracy and efficiency of this WD microarray platform to make it available for the healthcare system of the Indian subcontinent. Also, similar studies should be extended towards other monogenic diseases that show high mutational heterogeneity.

Arnab Gupta

Department of Cell Biology
Johns Hopkins University
School of Medicine, Baltimore
MD, USA 21205
agupta50@jhmi.edu

References

1. Kinnier Wilson SA. Progressive lenticular degeneration: a familial nervous disease associated with cirrhosis of the liver. *Brain* 1912; 34 : 295-509.
2. Bull PC, Thomas GR, Rommens JM, Forbes JR, Cox DW. The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat Genet* 1993; 5 : 327-37.
3. Figus A, Angius A, Loudianos G, Bertini C, Dessi V, Loi A, *et al*. Molecular pathology and haplotype analysis of Wilson disease in Mediterranean populations. *Am J Hum Genet* 1995; 57 : 1318-24.
4. Waldenstrom E, Lagerkvist A, Dahlman T, Westermarck K, Landegren U. Efficient detection of mutations in Wilson disease by manifold sequencing. *Genomics* 1996; 37 : 303-9.
5. 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.
6. Duc HH, Hefter H, Stremmel W, Castaneda-Guillot C, Hernandez Hernandez A, Cox DW, *et al.* Auburger G. His1069Gln and six novel Wilson disease mutations: analysis of relevance for early diagnosis and phenotype. *Eur J Hum Genet* 1998; *6* : 616-23.
 7. Kim EK, Yoo OJ, Song KY, Yoo HW, Choi SY, Cho SW, *et al.* Identification of three novel mutations and a high frequency of the Arg778Leu mutation in Korean patients with Wilson disease. *Hum Mutat* 1998; *11* : 275-8.
 8. Nanji MS, Nguyen VT, Kawasoe JH, Inui K, Endo F, Nakajima T, *et al.* Haplotype and mutation analysis in Japanese patients with Wilson disease. *Am J Hum Genet* 1997; *60* : 1423-9.
 9. Gu YH, Kodama H, Du SL, Gu QJ, Sun HJ, Ushijima H. Mutation spectrum and polymorphisms in *ATP7B* identified on direct sequencing of all exons in Chinese Han and Hui ethnic patients with Wilson's disease. *Clin Genet* 2003; *64* : 479-84.
 10. Deguti MM, Genschel J, Cancado EL, Barbosa ER, Bochow B, Mucenic M, *et al.* Wilson disease: novel mutations in the *ATP7B* gene and clinical correlation in Brazilian patients. *Hum Mutat* 2004; *23* : 398.
 11. Available from: www.igvdb.res.in, accessed on November 3, 2014.
 12. Santhosh S, Shaji RV, Eapen CE, Jayanthi V, Malathi S, Chandy M, *et al.* *ATP7B* mutations in families in a predominantly Southern Indian cohort of Wilson's disease patients. *Indian J Gastroenterol* 2006; *25* : 277-82.
 13. Aggarwal A, Chandhok G, Todorov T, Parekh S, Tilve S, Zibert A, *et al.* Wilson disease mutation pattern with genotype-phenotype correlations from Western India: confirmation of p.C271* as a common Indian mutation and identification of 14 novel mutations. *Ann Hum Genet* 2013; *77* : 299-307.
 14. Gupta A, Maulik M, Nasipuri P, Chattopadhyay I, Das SK, Gangopadhyay PK; Indian Genome Variation Consortium. Molecular diagnosis of Wilson disease using prevalent mutations and informative single-nucleotide polymorphism markers. *Clin Chem* 2007; *53* : 1601-8.
 15. Ala A, Borjigin J, Rochwarger A, Schilsky M. Wilson disease in septuagenarian siblings: raising the bar for diagnosis. *Hepatology* 2005; *41* : 668-70.
 16. Roberts EA, Schilsky ML; American Association for Study of Liver Diseases (AASLD). Diagnosis and treatment of Wilson disease: an update. *Hepatology* 2008; *47* : 2089-111.
 17. Gupta A, Neogi R, Mukherjee M, Mukhopadhyay A, Roychoudhury S, Senapati A, *et al.* DNA linkage based diagnosis of Wilson disease in asymptomatic siblings. *Indian J Med Res* 2003; *118* : 208-14.
 18. Mathur M, Singh E, Poduval TB, Rao AVSSN. Development of low-density oligonucleotide microarrays for detecting mutations causing Wilson's disease. *Indian J Med Res* 2015; *141* : 175-86.
 19. Wu ZY, Wang N, Lin MT, Fang L, Murong SX, Yu L. Mutation analysis and the correlation between genotype and phenotype of Arg778 Leu mutation in Chinese patients with Wilson disease. *Arch Neurol* 2001; *58* : 971-6.
 20. Santosh S, Shaji RV, Eapen CE, Jayanthi V, Malathi S, Finny P. Genotype phenotype correlation in Wilson's disease within families - a report on four south Indian families. *World J Gastroenterol* 2008; *14* : 4672-6.
 21. Gupta A, Bhattacharjee A, Dmitriev OY, Nokhrin S, Braiterman L, Hubbard A, *et al.* Cellular copper levels determine the phenotype of the Arg⁸⁷⁵ variant of *ATP7B*/Wilson disease protein. *Proc Natl Acad Sci USA* 2011; *108* : 5390-5.

EARLY DIAGNOSIS FOR WILSON DISEASE DUE TO MUTATIONS IN ATP7B CAUSING COPPER OVERLOAD

MANJULA MATHUR

Molecular Biology Division, Bhabha Atomic Research Centre,
Mumbai-400085, INDIA

Email: bioinf@barc.gov.in; mathur_manjula@yahoo.co.in

Abstract: - Wilson disease (WD) is a treatable monogenic autosomal disorder present from birth but clinically manifested after copper overload at the median age of 10 to 23 years. Copper is one of the nutritionally essential micronutrients as it forms cofactor to a large number of regulatory enzymes active in normal metabolism. If copper trafficking is defective, intracellular levels of copper start rising, which causes disturbances in structure and function of cuproenzymes and can generate superoxide radicals. Above certain levels of intracellular copper accumulation, it could even enter the nucleus and damage DNA causing unpredictable damage to the cellular functions. The gene associated with WD has been identified in last two decades as *atp7b*. Bioinformatics Applications are made to understand molecular genetic causes of the WD and to design a DNA chip for early and accurate detection of WD mutants in clinically complex patients.

Key-Words: - ATP7B, WD, *atp7b*, neuro-degeneration, Diagnosis, copper, liver

1. Introduction

The span of time between birth and death is called life. Most of us are born with normal health equipped with wonderful biological machinery that regulates our normal physiology and manages stresses generated during various life challenges. However, few of us are born with some genetic errors. Wilson Disease (WD) is one of such in-born errors of metabolism, affected by functionally defective ATP7B protein due to genetic mutations in *atp7b* (Fatemi and Sarkar 2002). Copper transporting ATPases ATP7A and ATP7B are evolutionarily conserved polytopic membrane proteins with essential roles in human physiology. Their transport activity is crucial for central nervous system development, liver function, connective tissue formation and many other physiological processes. WD is named after a British Physician Dr Samuel Alexander Kinnier Wilson who was first to document the disease as copper overload (Wilson 1912). WD may present as pediatric liver disease or may have neuropsychiatric presentation in adults. Thirty million people worldwide are affected by WD. It is fatal unless detected and treated before serious illness from copper poisoning develops. The WD gene *atp7b* was localized to the q14.3 band of chromosome 13 and cloned by two independent groups (Bull et. al. and Tanzi et. al) in 1993. With the successful completion of Human Genome Project, it is the natural duty of scientists to design sensitive specific DNA-Diagnostic chips for early but reliable disease diagnosis.

2. Physiology and pathophysiology of Wilson Disease (WD)

Small amounts of copper are as essential as vitamins. Copper is present in most foods, and most people have much more copper than they need. Healthy people excrete copper they don't need but WD patients cannot. Copper begins to accumulate immediately after birth in WD mutants. Excess copper attacks the liver or brain, resulting in hepatitis, psychiatric, or neurologic symptoms. The symptoms usually appear in late adolescence. Patients may have jaundice, abdominal swelling, vomiting of blood, and abdominal pain. They may have tremors and difficulty walking, talking and swallowing. They may develop all degrees of mental illness including homicidal or suicidal behaviors, depression and aggression. Women may have menstrual irregularities, absent periods, infertility or multiple miscarriages. No matter how the disease begins, it is always fatal if it is not diagnosed and treated. It is important to diagnose WD as early as possible, since severe liver damage can occur before there are any signs of the disease. If copper is not absorbed from the diet, it leads to Menkes disease which has been associated with functional defects caused by mutations in ATP7A (Bie et. al. 2007).

How is WD Diagnosed?

The diagnosis of WD is made by relatively simple tests which can diagnose the disease in both symptomatic and asymptomatic patients. These tests include:

- Ophthalmologic slit lamp examination for Kayser-Fleisher rings
- Serum Ceruloplasmin test
- 24-hour urine copper test
- Liver biopsy for histology and histochemistry and copper quantification
- Genetic testing, haplotype analysis for siblings and mutation analysis

However, individuals with WD may falsely appear to be in excellent health (www.wilsondisease.org).

How is WD Being Treated?

WD is a very treatable condition. With proper therapy, disease progress can be halted and oftentimes symptoms can be improved. Treatment is aimed at removing excess accumulated copper and preventing its reaccumulation. Therapy must be life long. Patients may become progressively sicker from day to day so immediate treatment can be critical. Treatment delays may cause irreversible damage. Stopping treatment completely will result in death, sometimes as quickly as within three months.

2.1 Molecular Mechanism of Copper Transport

Dietary copper is primarily absorbed through the intestinal mucosa and this process is regulated by copper intake. When the individual is copper deficient, uptake is more efficient and is reduced when copper stores are adequate (Bartinato and Abbe, 2004). In the serum, newly absorbed copper is transported to the liver bound to albumin and transcuprein. Most of the copper is taken up by the liver, the organ primarily responsible for regulating the copper status of the body.

2.1.1 A typical hepatocyte cell is shown in **Fig1** taken from Huster & Lutsenko 2007 depicting functions of different proteins in copper regulation. Copper uptake into cells is thought to be mediated by the plasma membrane protein CTR1 (Balamurugan, 2006). Small molecules such as glutathione (GSH) and metallothioneins bind copper for storage and/or detoxification and may provide an exchangeable pool of copper (Bertinato & Abbe 2004). Metallochaperones also bind copper and target

it to specific (Cobine et. al. 2006) destinations within the cell. CCS delivers copper to Cu/Zn-SOD1 (Ramasarma 2007), COX17 mediates copper transfer via HAH1 (Walker et.al. 2004) to the copper-ATPases, ATP7A and ATP7B for delivery to the secretory pathway and for efflux of excess copper from the cell. In hepatocytes, if copper levels remain within the normal range, most of the copper is incorporated into ceruloplasmin in the TransGolgi network (TGN), which is then secreted into the blood (Terada et. al. 1995). The ceruloplasmin holoenzyme contains six copper atoms and is a ferroxidase that has a role in iron mobilization. Ceruloplasmin contains the majority of serum copper and may have role in copper delivery to peripheral tissues. If copper supplies become excessive, then the biliary transport mechanisms are activated and the excess copper is excreted into the bile. Both the delivery of copper to ceruloplasmin and biliary excretion of copper are carried out by copper induced trafficking of ATP7B. At the cellular level, ion transport proteins must be targeted to the appropriate destination within the cell and retained there as long as needed and recycled back to its original location. Research over the last decade is leading to an integrated understanding of the role of copper transporting ATPases in cellular regulation of copper homeostasis (Fontaine and Mercer 2007).

2.2 Molecular sequence of Wilson protein ATP7B

Due to the very high cooperative nature of copper ATPase trafficking, clinical variability of WD is to the extent that the age of onset of symptoms of WD varies from 3 to 70 years. Also the clinical presentation of WD could be either hepatic or neurological symptoms. The genotype to phenotype (G2P) correlations are complicated because of compound heterozygosity of most of the WD patients. The *atp7b* is transcribed and spliced into 7500 bases of RNA in liver, brain, kidney and placental tissue. The mRNA is translated into the 1411 amino acids, 159kDa ATPase. It contains several functionally important domains. The molecular mechanism of copper transport in WD protein has been delineated by Fatemi and Sarkar 2002. **Fig. 2** is the pictorial representation of secondary structure of Wilson protein ATP7B. It contains several functionally important domains: a) for Cu binding MBDs 1-6 (Metal Binding Domain), b) for correct localization in the cell membrane TMs 1-7, c) for the passage of Cu through the membrane Transducing Domain (Td), d) for ATP dependent translocation of bound copper atoms- ATP binding, ATP hinge, ATP loop, Phosphatase and Channel domains. Kenney and Cox, 2007 have developed a database for reporting of mutations in *atp7b* that includes more than 518 variants (379 probable disease-causing and the remainder possible normal variants) from populations worldwide. Since not all of the missense mutations are disease causing, but are normal variants, we need to understand and assess the functional significance of these changes in terms of their affect on ATP7B catalytic activity and trafficking in relation to clinical heterogeneity of WD. To understand the disease causing effect of several ATP7B variants, experiments in *S. cerevisiae* were carried out by Forbes and Cox (2000). The website for *atp7b* is www.medicalgenetics.med.ualberta.ca/wilson/index.php.

3. Molecular Diagnostics and Genetics of WD

To design a WD DNA-diagnostic chip (microarray) of good efficacy, it is important to identify all the functional hot spots on ATP7B in relation to mutations database of *atp7b* (Kenny and Cox, 2007). An exhaustive set of oligonucleotides covering the mutation spots may be spotted on the microarray perspicaciously. The unknown DNA from human blood sample may be subjected to multiplex PCR (Polymerase Chain Reaction) for amplification of DNA segments that code for functional hotspots. PCR products may then be subjected to microarray along with fluorescently labeled dideoxynucleotides and polymerase enzyme for detection of mutations by primer extension. Let us try to understand the functions of different conserved domains from the published experimental results.

3.1 Metal Binding Domains: As depicted in Fig 2 and 3, ATP7B has six metal binding domains (MBD) in the N-terminal region. Each domain comprises approximately 70 amino acid residues, which share 20-60% sequence similarity and each has conserved metal binding site (MBS) GMxCxxC (x is any amino acid). MBDs 1-4 (Bunce, 2006) appear to be important for interaction with Atox1 and also have regulatory function. It is proposed that copper is transferred from MBD4 to MBD5-6 which then shuttle it to the intramembrane CPC site (Achila, 2006). In vitro and cell culture cell systems indicate that MBDs 1-4 are important for interaction with HAH1 (Human Atox1 Homologue) but are not required for ATPase trafficking.

3.2 Kinase mediated phosphorylation is a key regulatory mechanism that controls a vast array of signaling pathways. Vanderwerf et al (2001) demonstrated that in a variety of cell types (HepG2, HepG3 and primary epithelial cells) ATP7B was phosphorylated at a basal level, but becomes hyperphosphorylated in response to elevated copper levels. These authors identified separate sites for basal and copper dependent phosphorylation, the former occurring within the central region of the protein, and the latter mapped to a 63 amino acids domain in the N-terminal.

3.2.1 Catalytic phosphorylation: As part of their catalytic cycle the copper P-type ATPases undergo transient phosphorylation at the invariant aspartate residue (DKTG) to form an acyl-phosphate intermediate that catalyses transfer of Cu across the lipid bilayer. TGEA in the actuator domain are also shown to be phosphorylated.

3.3 Protein interactions: The dynactin subunit p62 was found to interact specifically with the ATP7B N terminus between MBD4 and 6 in a manner that was dependent on copper and the CxxC motifs (Lim et. al. 2006). Dynactin is a multi-subunit complex that binds both to microtubules and to the microtubule motor protein dynein, facilitating the interaction of dynein with other cellular constituents, and promoting dynein-mediated vesicle motility along microtubules. It was proposed that dynein/dynactin, through interaction between p62 and ATP7B may mediate microtubule-based copper-induced trafficking of ATP7B to the apical surface of hepatocytes. Microtubules are involved in the copper-induced trafficking of ATP7B to the apical surface of hepatocytes. COMMD1 (previously known as MURR1) is shown to interact directly with the N-terminal domain of hATP7B but not hATP7A (Lim et al 2006)

3.4 Molecular Genetics: Advances in molecular biology techniques and genetics have enabled genotyping of WD patients in hospitals. Gupta and Kunal have done genotypic studies on Indian WD patients (Gupta et. al. 2007a) on four SNPs and have contributed data to The Indian Genome Variation Consortium. In the Indian population prevalent WD mutations account for 41% of total mutations and there are numerous rare mutations which do not cluster to any specific region of the gene (Gupta et. al. 2007b).

4. Conclusion:

Due to complexity of manifestation of Wilson Disease, in addition to biochemical diagnosis, it is best to develop a database of DNA evaluation of the *atp7b* of WD suspected patients and their family. For this a microarray of *atp7b* should be designed judiciously covering the 80kb genomic DNA that should include mutation hotspots not only on 22 exons but including reported mutations on ncRNA. Accurate diagnosis of WD susceptibility is the answer for timely treatment of the condition to avoid unnecessary

deterioration of health of a *atp7b* mutant/variant. It needs to be diagnosed before the onset of organ involvement. Once WD predisposition is identified in individual, regular checkups, suitable nutrition, and preventive medication should be recommended to thwart disease progression and onset. Families affected with the disease should be encouraged to undergo carrier detection and appropriate genetic counseling could be given. Without appropriate molecular genetic diagnosis tool, predisposition for the disease would remain unnoticed in WD-mutant children until they manifest signs of the disease. Additional information is available at website www.bumc.bu.edu

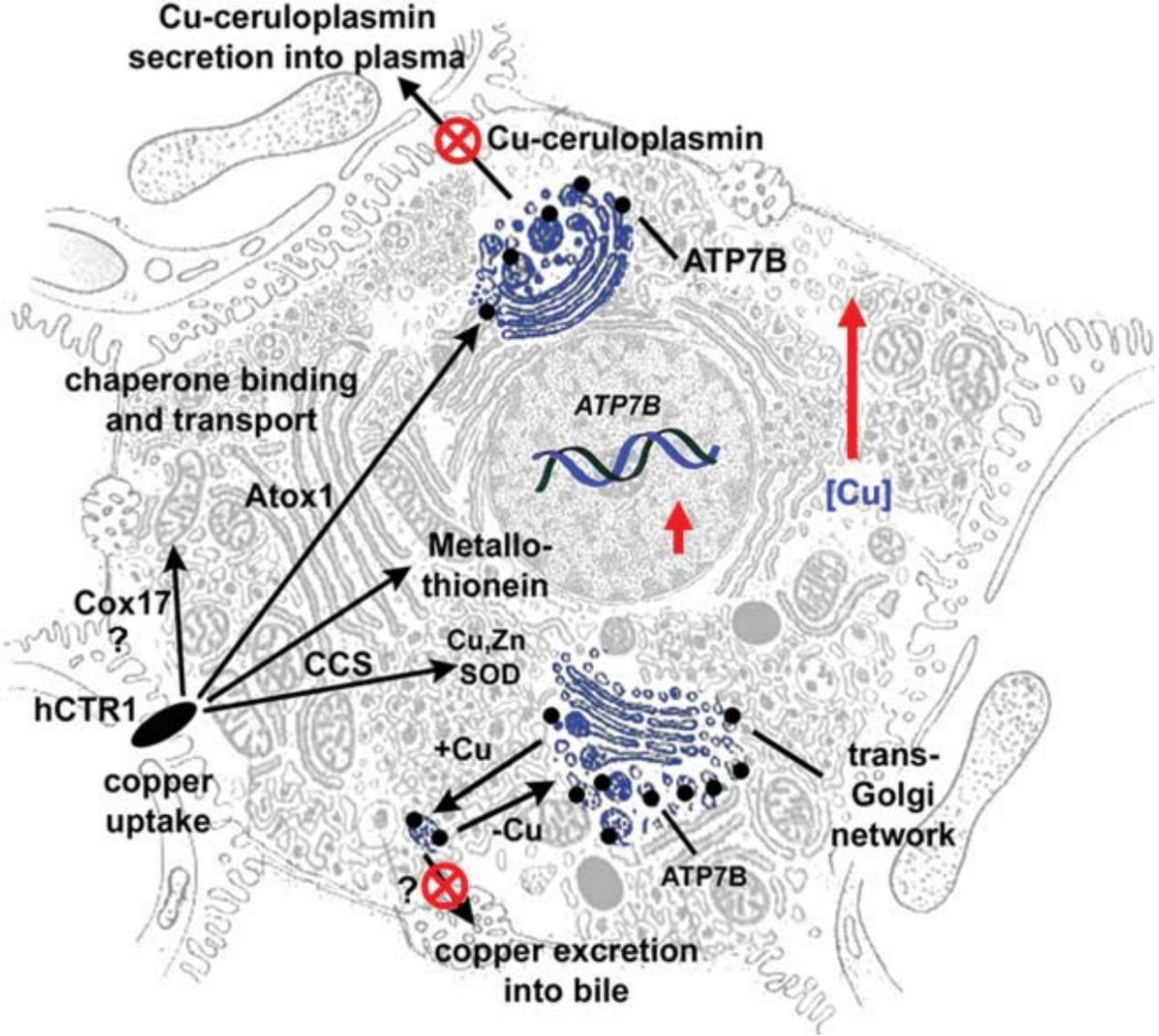


Fig 1. Cellular proteins in copper transport in a typical hepatocyte

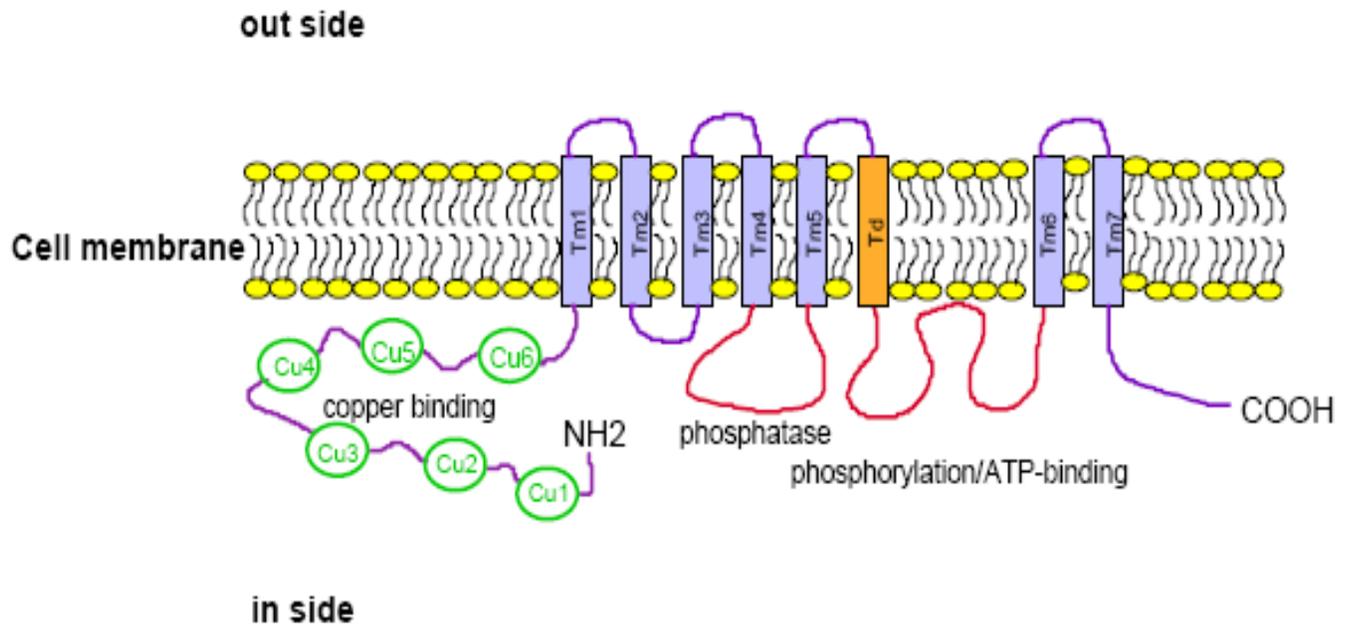


Fig 2. A model of WD protein ATP7B depicting conserved functional domains

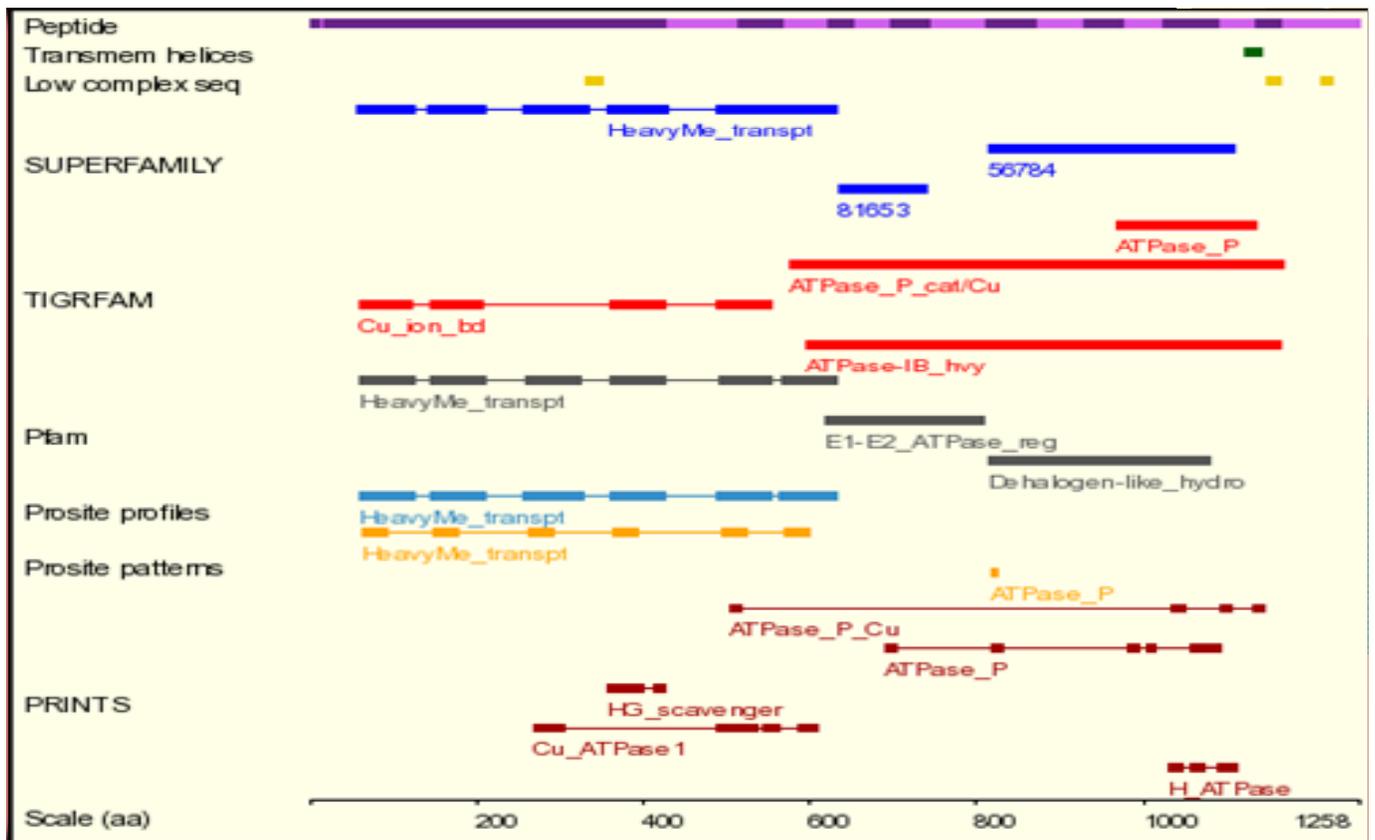


Fig 3. Bioinformatics Analysis of WD protein ATP7B

References:

- [1] Achila D, Banci L, Bertini I, Bunce J, Baffoni CS, Huffman DL, Structure of human Wilson protein domains 5 and 6 and their interplay with domain 4 and the copper chaperone HAH1 in copper uptake, *PNAS*, Vol. 103, 2006, pp. 29141-29147
- [2] Balamurugan K and Schaffner W, Copper homeostasis in eukaryotes: Teetering on a tightrope, *Biochim. Biophys. Acta* Vol. 1763, 2006, pp. 737-746
- [3] Bartinato J and Abbe MRL, Maintaining copper homeostasis: regulation of copper-trafficking proteins in response to copper deficiency or overload, *J. Nutr. Biochem.* Vol. 15, 2004, pp. 316-322
- [4] Bie PD, P. Muller, C. Wijmenga, L W.J. Klomp, Review: Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotype. *Journal of Medical Genetics* Vol. 44, 2007, pp. 673-688
- [5] Bunce J, David A. Chile, Evan Hetrick, Leighann Lesley and David L. Huffman. Copper transfer studies between the N-terminal copper binding domain one and four of human Wilson protein. *BBA*, Vol. 1760, No. 6, 2006, pp. 907-912.
- [6] Bull PC, Thomas GR, Rommens JM, Forbes JR, and Cox DW, The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. *Nat. Genet.* Vol. 5, 1993, pp. 327-337
- [7] Cobine PA, Pierrel F, Winge DR. Copper trafficking to the mitochondrion and assembly of copper metalloenzymes, *BBA*, Vol. 1763, 2006, pp. 759-772
- [8] Fatemi N and Sarkar B, Molecular Mechanism of Copper Transport in Wilson Disease. *Environmental Health Perspectives*, Vol. 110, No.5, 2002, pp. 695-698
- [9] Fontaine SL and Mercer JFB, Trafficking of the copper – ATPases, ATP7A and ATP7B: Role in copper homeostasis, *Archives of Biochemistry and Biophysics*, Vol. 463, 2007, pp. 149-167
- [10a] Gupta A, Maulik M, Nasipuri P, Chattopadhyay I, Das SK, Gangopadhyay PK, The Indian Genome Variation Consortium and Ray K, Molecular Diagnosis of Wilson Disease Using Prevalent Mutations and Informative Single-Nucleotide Polymorphism Markers. *Clinical Chemistry*, Vol. 53, No.9, 2007, pp. 1601-1608
- [10b] Gupta A, I. Chattopadhyay, S. Dey, P. Nasipuri, S. K. Das, P. K. Gangopadhyay and Kunal Ray. 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 DOI 10.1007/s10571-007-9192-7
- [11] Huster D and Lutsenko S, Wilson disease: not just a copper disorder. Analysis of a Wilson disease model demonstrates the link between copper and lipid metabolism. *Mol. BioSyst* Vol. 3, No. 12, 2007, pp. 816-824
- [12] Kenney SM and Cox CW, Sequence variation database for the Wilson disease copper transporter, ATP7B. *Human Mutation*, Vol.28, No.12, 2007, pp. 1171-1177.

- [13] Lim CM, Michael A. Cater, Julian F.B. Mercer and Sharon La Fontaine, Copper-dependent Interaction of Dynactin Subunit p62 with the N Terminus of ATP7B but not ATP7A, *J. Biol. Chem.*, Vol. 281, No. 20, 2006, pp. 14006-14014
- [14] Lutsenko S, Barnes NL, Bartee MY and Dmitriev OY, Function and Regulation of Human Copper – Transporting ATPases. *Physiological Reviews* Vol. 87, 2007, pp. 1011-1046
- [15] Ramasarma T, Many faces of superoxide dismutase, originally known as erythrocuprein, *Current Science*, Vol. 92, No. 2, 2007, pp. 184-191
- [16] Tanzi RE, et.al. The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. *Nature Genetics* Vol. 5, 1993, pp. 344-350.
- [17] Vanderwerf SM, Cooper MJ, Stetsenko IV & S. Lutsenko, Copper specifically regulates intracellular phosphorylation of the Wilson's disease protein, a human copper – transporting ATPase. *J. Biol. Chem.* Vol. 276, 2001, pp. 36289-36294.
- [18] Walker JM, Dominik Huster, Martina Ralle, Clinton T. Morgan, Ninian J. Blackburn, and Svetlana Lutsenko, The N-terminal Metal-binding Site 2 of the Wilson's Disease Protein Plays a Key Role in the Transfer of Copper from Atox1, *J. Biol. Chem.*, Vol. 279, Issue 15, 2004, pp. 15376-15384
- [19] Wilson SAK. Progressive lenticular degeneration: a familial disease associated with cirrhosis of liver. *Brain* Vol. 34, 1912, pp. 295-507