

**Evaluation of Denaturing  
Gradient Gel Electrophoresis (DGGE)  
as a method of mutation detection.**

Helene Cadiou

Department of Genetics and Biometry  
University College London  
University of London

submitted October 1994  
for the degree of Doctor of Philosophy

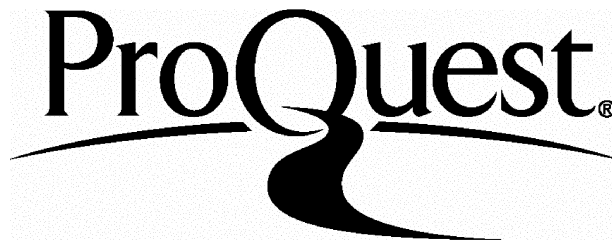
ProQuest Number: 10018526

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10018526

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.  
Microform Edition © ProQuest LLC.

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## **Abstract**

The success in the isolation and characterisation of genes responsible for genetic disease, the considerable molecular heterogeneity at these loci has placed a growing emphasis on the importance of mutation detection techniques both in the coding and non-coding regions of human DNA.

This thesis is principally concerned with the evaluation of one method of mutation analysis: Denaturing Gradient Gel Electrophoresis (DGGE), which has been used to identify sequence variation in three model systems: the well characterised alpha-1-antitrypsin (AAT) gene, with a range of known mutations; a segment of the Y chromosome where there is evidence for very little mutation and the Phenylalanine Hydroxylase (PAH) gene which mutated causes the autosomal recessive disorder Phenylketonuria (PKU), a disease characterised by severe mental retardation.

The AAT gene was used as a test locus to evaluate the general applicability of DGGE for the detection of single base pair polymorphisms. The sequence on the Y chromosome was used to set up the computer programs associated with the technique. Subsequently the technique was used to screen for mutation in 5 exons and their splice site sequences of the PAH gene in 50 unrelated Phenylketonuria (PKU) patients.

In the AAT gene, all six common point mutations were resolved. In contrast, no polymorphism was identified in the segment of Y chromosome investigated. In the PAH gene, the technique identified 10 different mutations, corresponding to 42% of affected chromosomes. Of these, three, K341R, L347F, and IVS10nt-3c->t were unique and not previously reported in

the international data base. Two further sequence variants which caused no obvious alteration in the function of the PAH gene were also identified. A comparison of the efficiency of DGGE, in relation to single strand conformation polymorphism (SSCP) as a mutation detection system was also performed. Only four out of the twelve mutations identified using DGGE were resolved using SSCP analysis. Furthermore SSCP analysis did not identify any other mutations not observed in the initial DGGE screen.

Overall, the study has evaluated the use of DGGE as a mutation detection system. It has resolved all the common point mutations in the AAT gene, confirmed the paucity of polymorphism in the Y chromosome and identified a number of mutations present in a South-East of England PKU population.

## **INDEX**

	<b>Pages</b>
<b>TITLE PAGE</b>	<b>1</b>
<b>ABSTRACT</b>	<b>2-3</b>
<b>CONTENTS</b>	<b>4-10</b>
<b>LIST OF FIGURES</b>	<b>11-17</b>
<b>LIST OF TABLES</b>	<b>18-19</b>
<b>ABBREVIATIONS</b>	<b>20-21</b>
<b>CONTRIBUTERS</b>	<b>22</b>
<b>ACKNOWLEDGMENTS</b>	<b>23</b>
<b>DEDICATION</b>	<b>24</b>
<b>CHAPTER 1: INTRODUCTION</b>	<b>25-57</b>
<b>CHAPTER 2: MATERIALS AND METHODS</b>	<b>58-78</b>
<b>CHAPTER 3: RESULTS</b>	<b>79-163</b>
<b>CHAPTER 4: DISCUSSION</b>	<b>164-188</b>
<b>CHAPTER 5: BIBLIOGRAPHY</b>	<b>189-206</b>

<b>CHAPTER 1: INTRODUCTION</b>	<b>25-57</b>
<b><u>1.1 Bakground</u></b>	<b>25</b>
<b><u>1.2 Mutation</u></b>	<b>25</b>
<b>1.2.1 Structural basis</b>	<b>25</b>
<b>1.2.2 Analysis of phenotypic variation</b>	<b>27</b>
<b>1.2.3 Analysis of genetic variation</b>	<b>31</b>
1.2.3.1 Utilisation of restriction enzymes	32
1.2.3.2 Minisatellites and microsatellites	34
<b><u>1.3 Mutation detection strategies</u></b>	<b>35</b>
<b>1.3.1 Denaturing Gradient Gel Electrophoresis (DGGE) and related techniques</b>	<b>36</b>
<b>1.3.2 Single Strand Conformation Polymorphism (SSCP)</b>	<b>44</b>
<b>1.3.3 Heteroduplex analysis</b>	<b>44</b>
<b>1.3.4 Chemical Mismatch Cleavage (CMC)</b>	<b>46</b>
<b>1.3.5 RNase protection Assay</b>	<b>48</b>
<b>1.3.6 Direct sequencing (DS)</b>	<b>48</b>
<b><u>1.4 Model systems investigated</u></b>	<b>49</b>

<b>1.4.1 Alpha-one-antitrypsin (AAT)</b>	<b>49</b>
<b>1.4.2 Y chromosome</b>	<b>50</b>
<b>1.4.3 Phenylketonuria (PKU)</b>	<b>51</b>
<b><u>1.5 Aims and objectives</u></b>	<b>56</b>
	<b>58-78</b>
<b>CHAPTER 2 : MATERIALS AND METHODS</b>	
<b><u>2.1 Materials and suppliers</u></b>	<b>58</b>
<b><u>2.2 Methods</u></b>	<b>62</b>
<b>2.2.1 Isolation of DNA</b>	<b>62</b>
2.2.1.1 DNA extracted from whole blood	62
2.2.1.2 plasmid DNA, extracted from bacterial cultures	62
<b>2.2.2 Concentration of DNA</b>	<b>63</b>
<b>2.2.3 PCR primer sequences and PCR reaction conditions</b>	<b>64</b>
2.2.3.1 Alpha-one-antitrypsin gene	64
2.2.3.2 Y sequence originated from the pY80 fragment	65
2.2.3.3 Phenylalanine hydroxylase gene	66

2.2.3.4 PCR protocol	67
<b>2.2.4 Agarose gel electrophoresis</b>	<b>67</b>
<b>2.2.5 Denaturing Gradient Gel Electrophoresis</b>	<b>67</b>
2.2.5.1 Gel preparation	67
2.2.5.2 Electrophoresis conditions	71
2.2.5.3 Visualisation of DGGE products	72
2.2.5.3.1 Ethidium bromide staining	71
2.2.5.3.2 Silver staining	71
<b>2.2.6 Constant Denaturant Gel Electrophoresis (CDGE)</b>	<b>73</b>
<b>2.2.7 Single Strand Conformation Polymorphism (SSCP)</b>	<b>73</b>
<b>2.2.8 Cloning of PCR product</b>	<b>74</b>
<b>2.2.9 Dideoxy sequencing</b>	<b>77</b>

## **CHAPTER 3: RESULTS**

<b><u>3.1 Mutation analysis of the human AAT gene.</u></b>	<b>79</b>
<b>3.1.1 Summary</b>	<b>79</b>

<b>3.1.2 Sample preparation</b>	<b>79</b>
<b>3.1.3 Generation of exon specific products</b>	<b>79</b>
<b>3.1.4 Screening of individual exons for mutation</b>	<b>81</b>
<b>3.1.5 Conciusion</b>	<b>91</b>
<b><u>3.2 Screening for mutation in the Y chromosome</u></b>	<b>92</b>
<b>3.2.1 Summary</b>	<b>92</b>
<b>3.2.2 DGGE analysis of pY-80</b>	<b>92</b>
<b><u>3.3 Mutation analysis of the Phenylalanine Hydroxylase (PAH) gene</u></b>	<b>105</b>
<b>3.3.1 Summary</b>	<b>105</b>
<b>3.3.2 Sample preparation</b>	<b>106</b>
<b>3.3.3 Generation of exon specific PCR products</b>	<b>106</b>
<b>3.3.4 Denaturing Gradient Gel Electrophoresis</b>	<b>107</b>
<b>3.3.5 Screening of individual exons for mutation</b>	<b>109</b>
<b>3.3.5.1 Exon 3</b>	<b>109</b>
<b>3.3.5.2 Exon 10</b>	<b>127</b>
<b>3.3.5.3 Exon 11</b>	<b>130</b>

3.3.5.4 Exon 12	146
3.3.5.5 Exon 5	154

### **3.3.6 Evaluation of SSCP as a method of mutation detection**

3.3.6.1 Evaluation of SSCP as a method of mutation detection for resolving the 12 different mutation types identified by DGGE. 154

3.3.6.2 SSCP as a technique to identify mutation other than those resolved by DGGE 160

## **CHAPTER 4: DISCUSSION**

<b><u>4.1 Summary</u></b>	164
<b><u>4.2 DGGE studies</u></b>	168
<b>4.2.1 Alpha-one-antitrypsin</b>	168
<b>4.2.2 Y chromosome</b>	169
<b>4.2.3 Studies on PKU</b>	171
4.2.3.1 Family studies: DGGE analysis	171
4.2.3.2 Population studies	174
4.2.3.3 Clinical aspects of PKU mutation analysis	176
<b><u>4.3 Mutation detection</u></b>	178

<b>4.3.1 Choosing a detection strategy</b>	<b>178</b>
<b>4.3.2 Applications of the different screening strategies</b>	<b>178</b>
<b>4.3.3 Studies evaluating different screening approaches</b>	<b>186</b>
<b><u>4.4 Conclusion</u></b>	<b>188</b>
<b>CHAPTER 5: BIBLIOGRAPHY</b>	<b>189</b>

## **LIST OF FIGURES**

**11-17**

### **CHAPTER 1**

1.1 Diagrammatical representation illustrating the principle of DGGE 37

1.2 Diagrammatical representation comparing the melting properties of two different DNA duplex forms by Parallel Denaturing Gradient Gel Electrophoresis 39

1.3 Diagrammatical representation of the generation of a high melting domain through the incorporation of a GC clamp at one end of the PCR product 40

1.4 Diagrammatical representation of a Perpendicular Denaturing Gradient Gel Electrophoresis 42

1.5 Diagrammatical representation of Single Strand Conformation Polymorphism (SSCP) 45

1.6 Diagrammatical representation of chemical mismatch cleavage (CMC) 47

1.7 Diagrammatical representation of the phenylalanine hydroxylation pathway in man as described by Kaufman in 1976 53

## **CHAPTER 2**

2.1 Diagrammatical representation of the pCR™ vector 76  
used to clone PCR reactions with EcoR1 sites flanking  
the cloning site

## **CHAPTER 3**

### **AAT**

3.1.1 An ethidium bromide stained perpendicular DGG 82  
of exon 3 from the AAT gene

3.1.2a An ethidium bromide stained parallel DGG of 84  
exon 3 from the AAT gene resolving the M and S alleles  
at codon position 265 using a 50-70% denaturing  
gradient

3.1.2b An ethidium bromide stained parallel DGG of 85  
exon 3 from the AAT gene resolving the M1Val and M1  
Ala alleles at codon position 213 using a 50-70%  
denaturing gradient

3.1.3a An ethidium bromide stained parallel DGG of 87  
exon 5 from the AAT gene resolving the M3 allele from  
the M1 allele at codon position 376 using a 55-70%  
denaturing gradient

3.1.3b An ethidium bromide stained parallel DGG of 88  
exon 5 from the AAT gene resolving the Z allele at  
codon position 342 using a 60-80% denaturing gradient

3.1.4 An ethidium bromide stained parallel DGG of exon 90 2 from the AAT gene resolving the M2 from the M allele at codon position 101 using a 56-76% denaturing gradient

## **pY-80**

3.2.1 Sequence of the pY-80 taken from Tsukahara *et al.*, 1990.

3.2.2 Melting map of the pY-80 sequence as determined 95 by the MELT87 program

3.2.3 a and b Melting map of the pY-80 sequence, 97 determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 5' and 3' end of the sequence respectively

3.2.4 Melting map of the pY-80 sequence, position 1- 98 270 bp determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 3' end of the sequence

3.2.5 Melting map of the pY-80 sequence, position 99 201-470 bp determined by the MELT87 program following the introduction of a 50 bp GC clamp at the 3' end of the sequence

3.2.6 PCR amplification of the Y1 and Y2 regions of pY- 101 80 corresponding to region positions 1-270 and 201-470 respectively

3.2.7 An ethidium bromide stained DGG of 9 male 103 samples and a positive control for the Y1 region, position 1-270bp, using a 60-80% denaturing gel

3.2.8 An ethidium bromide stained DGG of 2 male 104 samples and 2 positive controls for the Y2 region, position 201-470bp, using a 65-80% denaturing gel

## **PAH**

3.3.1 PCR amplification of exon/intron sequences of 108 exon 3, 5, 10, 11 and 12 of the PAH gene

3.3.2 Melting map of the sequence containing the exon 110 3 and splice site sequence of the PAH gene as determined by the MELT87 program

3.3.3 Melting map of the sequence containing the exon 111 3 and splice site sequence of the PAH gene determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 3' end of the sequence

3.3.4 Melting map of the sequence containing the exon 112 3 and splice site sequence of the PAH gene determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 5'end of the sequence

3.3.5 A silver stained CDG of the two mutation types 113 identified in exon 3 using a 52% constant denaturing gel

3.3.6 A silver stained CDG of the two mutation types 115  
identified in exon 3 using a 52% constant denaturing  
gel

3.3.7 EcoRI digestion of a cloned exon 3 PCR product of 122  
the PAH gene

3.3.8 A silver stained CDG of exon 3 mutation 1, 123  
identified using a 52% constant denaturant gel

3.3.9 Sequence comparison of mutated and normal 124  
alleles of the exon 3 region containing mutation 1

3.3.10 Comparison of the coding region containing the 125  
I65T mutation in the human, rat, mouse and drosophila  
PAH protein

3.3.11 Sequence comparison of mutated and normal 126  
alleles of the exon 3 region containing mutation 2

3.3.12 A silver stained DGG of the three mutation 128  
types identified in exon 10 using a 48-52% denaturing  
gradient

3.3.13 An ethidium bromide stained CDG of the three 129  
mutation types identified in exon 10 using a 49%  
constant denaturing gel

3.3.14 Sequence comparison of mutated and normal 131  
alleles of the exon 10 region containing mutation 1

3.3.15 Sequence comparison of mutated and normal 132  
alleles of the exon 10 region containing mutation 2

3.3.16 Sequence comparison of mutated and normal 133 alleles of the exon 10 region containing mutation 3

3.3.17 Comparison of the coding region containing the 134 K341R and L347F mutations in the human, rat, mouse and drosophila PAH protein

3.3.18a An ethidium bromide stained CDG of the five 136 mutation types identified in exon 11 using a 56% constant denaturant gel

3.3.18b A silver stained CDG of the five mutation 137 types identified in exon 11 using a 56% constant denaturing gel

3.3.19 Silver staining of a CDG of mutation 1 under 138 identical denaturing conditions using different percentages of acrylamide

3.3.20 A silver stained CDG of exon 11 mutation 2 and 140 3

3.3.21 Sequence comparison of mutated and normal 141 alleles of the exon 11 region containing mutation 1

3.3.22 Sequence comparison of mutated and normal 142 alleles of the intron 10 splice site region containing mutation 2

3.3.23 Sequence comparison of mutated and normal 143 alleles of the exon 11 region containing mutation 3

- 3.3.24 Sequence comparison of mutated and normal 144  
alleles of the exon 11 region containing mutation 4
- 3.3.25 Sequence comparison of mutated and normal 145  
alleles of the exon 11 region containing mutation 5
- 3.3.26 A silver stained DGG of the two mutation types 147  
identified in exon 12 using a 54-57% denaturing  
gradient
- 3.3.27 CDGE analysis of the PAH exon 12 in a PKU 150  
family showing mendelian inheritance of mutation1
- 3.3.28 Sequence comparison of mutated and normal 151  
alleles of the exon 12 region containing mutation 1
- 3.3.29 Sequence comparison of mutated and normal 152  
alleles of the intron 12 splice site region containing  
mutation 2
- 3.3.30 Comparison of the coding region containing the 153  
R408W mutation in the human, rat, mouse and  
drosophila PAH protein
- 3.3.31 A silver stained CDG of exon 5 using a 55% 155  
denaturing concentration
- 3.3.32 A silver stained SSCP gel corresponding to 157/158  
analysis of exon 3, 10, 11 and 12 of the PAH gene
- 3.3.33 A silver stained SSCP gel corresponding to 163  
analysis of 4 individual samples of exon 12 of the PAH  
gene

## **LIST OF TABLES**

18-19

### **CHAPTER 1**

1.1 Review of the historical approaches used to study 28  
genetic variation in man

### **CHAPTER 2**

2.1 Stock solutions for the preparation of 6.5% and 69  
10% acrylamide DGG

2.2 Volumes of stock solutions required to obtain the 70  
appropriate gradient of chemical denaturant in the DGG  
and temperatures as function of the denarurant  
concentration

2.3 AAT, Y and PKU DGGE conditions 72

2.4 SSCP silver staining protocol for the Phastsystem 75  
gels

## **CHAPTER 3**

3.3.1 The six common variant alleles of human AAT 80  
studied in this project.

3.3.2 Distribution of the identified PAH mutations in 116-8  
50 unrelated PKU patients and close relatives

3.3.3 Distribution of 50 unrelated PKU patient samples 119-  
and their relatives 120

3.3.4 Summary of the results of the SSCP analysis for 161  
the 12 different mutations identified by DGGE

## **CHAPTER 4**

4.1 DGGE applications in Human Molecular Genetics 165-7

4.2 Advantages and disadvantages of screening 179  
methods for detection of single base mutations.

4.3 Diagrammatical representation of the different 182  
types of DGGE banding patterns observed in the DGGE  
studies presented.

4.4 Genomic DNA versus mRNA as starting material for 185  
mutation analysis.

## Abbreviations

20-21

AAT	alpha-1-antitrypsin
bp	base pairs
°c	degrees centigrade
CDG	Constant Denaturing Gel
CDGE	Constant Denaturing Gel Electrophoresis
cDNA	complementary deoxyribonucleic acid
DGG	Denaturing Gradient Gel
DGGE	Denaturing Gradient Gel Electrophoresis
DNA	deoxyribonucleic acid
EDTA	ethylene-diamine-tetra acetic acid
FISH	Fluorescent in situ hybridisation
g	grams
g	gravity
Kb	kilobase
M	molar
Mb	Megabase
µg	microgram
ml	millilitre
µl	microlitre
Mins	minutes
mRNA	messenger ribonucleic acid
OD	optical density
PAH	Phenylalanine hydroxylase
PCR	polymerase chain reaction
PKU	Phenylketonuria
RFLP	Restriction Fragment Length Polymorphism
rpm	revolutions per minute
SSCP	Single strand conformation polymorphism
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
TEMED	N,N',N',N'-tetramethylethylenediamine

UV	ultraviolet light
VNTR	Variable number tandem repeat
XGal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside

## **Contributors**

### **Leonard Lerman**

provided the computer MELT87 and SQHTX algorithms

### **John Attwood**

set up the Lerman computer program

### **Sue Povey, David Whitehouse**

provided genomic DNA from a range of people with the most common AAT variants

### **Phillip Johnson**

with whom the AAT studies were carried out

### **Michel Vidaud**

offered the key to the understanding of the SQHTX algorithm

### **Dr D. Brenton and G. Gillett**

collected the PKU blood samples

### **Dr L. Tyfield**

offered informations on intron sequences of the PAH gene

## Acknowledgments

Professors E. Robson and D. Hopkinson, Dr Phillip Johnson who helped me throughout the course of my studies.

I am very grateful to Dr D. Brenton who was extremely supportive over the last year of my PhD, Margaret Lillburn for her dedication to the PKU patients in the Middlesex Hospital and her spontaneous help each time I needed some information

Thanks also to Isabelle Chantret, Joseph Nahmias, Wendy and Twee for their support.

I am grateful to PKU families for providing material to allow this study to be undertaken.

**To Jude and Anne**

## **Chapter 1 INTRODUCTION**

### **1.1 Background**

Techniques that allow the detection of genetic changes at the level of deoxyribonucleic acid (DNA) have had a major impact on our understanding of human genetic diseases, both by identifying specific mutations that result in disease, and by identifying DNA polymorphisms that are used as genetic markers in linkage studies.

This thesis is mainly concerned with the evaluation of a technique, Denaturing Gradient Gel Electrophoresis (DGGE), for the analysis of mutation in DNA, for the detection of polymorphisms which might be used as genetic markers for gene mapping and also to analyse the relationships between genotype and phenotype. This general introduction will present the background relevant to these studies, thereafter the different studies will be considered separately.

### **1.2. Mutation**

#### **1.2.1. Structural basis**

DNA mutations can be divided into three broad categories : those involving single base substitutions, the small deletion/insertion events involving just a few nucleotides and larger deletions/insertions which may affect part or all of a gene sequence. This omits the major chromosomal rearrangements and duplications which result in aberrant chromosomal constitution such as Down syndrome (trisomy 21), Klinefelter syndrome (XXY) and Turner syndrome (monosomy XO).

In coding sequence, single base pair changes (point mutations) can result in nonsense mutations (chain terminator mutations) and the generation of a stop codon in a translated sequence. They represent approximately 5% (Harris,1975) of all mutations. Synonymous mutations, where the nucleotide change does not result in a change in amino acid, due to degeneracy of the code, occur in approximately 25% of all mutations. Non synonymous mutations (missense mutations), where there is an amino acid substitution account for approximately 70% of all mutations in coding sequences. Point mutations can also occur in the region of splice site junctions or in the region of regulatory sequences of a gene.

Small deletions or insertions generally involve the loss or gain of a small number of nucleotides and may result in frameshift mutations or the generation of a mature protein containing or lacking a few amino acid residues if the mutation involves coding sequences. Such mutations can also affect the function of regulatory sequences and splicing mechanisms.

Large deletions/insertions or other rearrangments of greater than 1 megabase (Mb) of DNA may be visualised cytogenetically (Grompe *et al.*, 1993) while smaller submicroscopic changes are usually identified through the use of direct DNA analysis using restriction enzyme, gel electrophoresis, DNA transfer to filter support (Southern blot analysis, Southern,1975) and hybridisation to specific cloned DNA segments used as probes (Vetrie *et al.*, 1993). Alternatively, pulsed field gel electrophoresis (PFGE) studies on DNA can be used. In contrast to routine DNA analysis techniques which have an upper size limit of 40 Kb, PFGE can separate fragments in the size range of 0.5 to 10 Mb so that large areas of a gene can be screened for the presence of deletions in a single experiment. The introduction of fluorescently labelled DNA probes which are hybridized to chromosome spreads, termed

fluorescence *in situ* hybridization (FISH) have been used to detect microdeletions, duplications and rearrangements (Carter *et al.*,1992).

A further category of mutation, heritable unstable DNA, involves the amplification and expansion of a simple trinucleotide repeat sequence. This expansion of a trinucleotide repeat has been observed as a mechanism underlying the seven disorders, spinal and bulbar muscular atrophy (SBMA) also known as Kennedy disease (La Spada *et al.*,1991), fragile X (FX) (Yu *et al.*,1991), FRAXE fragile site (Knight *et al.*, 1993), myotonic dystrophy (DM) (Mahadevan *et al.*,1993), Huntington's disease (HD) (The Huntington's Disease Collaborative Research Group,1993), Spinocerebella ataxia (Sca1) (Orr *et al.*, 1993) and dentatorubropallidoluysian atrophy (DRPLA) (Nagafuchi *et al.*, 1994). Clinically these diseases exhibit a phenomenon called anticipation where increasing severity of the disease is observed in subsequent generations of an affected family. Molecular analysis of the trinucleotide repeat sequences in these pedigrees p(CGG)n in FX, p(GCC)n in FRAXE, p(CTG)n in DM and p(CAG)n in HD, SBMA, DRPLA and SCA1 suggest a correlation between the size of the amplification element in the patient and the severity of the phenotype observed. These trinucleotide repeats are located in the 5' (FX) and 3' (DM) untranslated regions and are also present in the coding sequence (SBMA, HD and SCA).

### **1.2.2 Analysis of phenotypic variation.**

The study of inherited variation in man began with the study of general phenotypic characteristics, whereas the most recent studies investigate variation at the molecular (DNA) level to establish precisely what determines the general phenotypic features (Table1.1).

APPROACHES	EXAMPLES		APPROXIMATE PERIOD	REFERENCES
<b>appearance</b>	facial features	Hapsburg lip	1910	Rubbrecht, 1910
	skeletal abnormalities	scoliosis	1934	Garland, 1934
	bleeding disorders	haemophilia A	1953	Alexander, 1953
<b>cells</b>	red cell	ABO blood groups	1900	Landsteiner, 1900, 1901
		Rhesus blood groups	1941	Landsteiner and Wiener, 1941
	other cells	histocompatibility systems (HLA)	1970	Teraski, 1970; Dausset, 1972
<b>metabolite</b>	alkaptonuria		1900	Garrod, 1902
	phenylalanine+related metabolites		1935	Penrose, 1935
<b>protein</b>	sickle cell anaemia		1949	Pauling, 1949
	haemoglobin variants			
	serum protein and isozyme polymorphism		1960	Harris and Hopkinson, 1972
<b>genotype</b>	DNA/ molecular genetics		1978	Kan and Dozy, 1978 Weatherall, 1985

**Table 1.1 Review of the historical approaches used to study genetic variation in man.**

Mendel was the first to study the inheritance of these phenotypic changes. By creating "artificial" crosses between different varieties of the common garden pea plant, he was able to follow the segregation of several traits in subsequent generations. From his observations, he subdivided these features into two distinct forms that he referred to as dominance and recessive. He defined dominance by those characters which are transmitted in their entirety and therefore constitute the character of the hybrid. Those which become latent in the process, he termed recessive. In essence a recessive allele will manifest its phenotypic effect only when present in the homozygous state, in contrast to a dominant allele which need only be present as a single copy to exert its effect on the phenotype.

In human beings, obvious morphological variations for example facial features, skeletal abnormalities and some genetic diseases can be traced in families in order to determine the recessive or dominant nature of a particular gene. Haemophilia A in the English royal family is probably one of the best examples. In the case of these royals, the pattern of segregation suggested that haemophilia A was an X-linked recessive disease. This ability to follow a particular characteristic through a pedigree is the basis for linkage analysis (section 1.1.1.3.) and has advanced considerably in recent years with the introduction of molecular techniques which can lead to the identification of a disease locus from a knowledge of its chromosomal position, an application known as positional cloning or reverse genetics (Collins, 1992; Orkin, 1986).

The first example of variation at the cellular level came from the discovery of different blood groups by Landsteiner at the turn of the century (Landsteiner 1900, 1901). He observed that sera taken from some individuals caused the agglutination of the red cells of others. By postulating the presence of two antigens, now

termed A and B, the red cells could be classified into four groups A, B, AB and O. Therefore, O individuals which have no antigens, have both antibody-A and antibody-B, whilst AB persons were without such antibodies. This has become the basis for blood transfusion. Since these initial studies, about 22 blood group systems (Daniels *et al.*,1993) have now been identified.

Landsteiner, suggested that the difficulties in matching for blood transfusion were similar to those encountered for transplantation. Early work in the mouse had established that genetic differences between recipient and donor tissue affected the outcome of a graft. This led to the discovery of the antigens important for transplantation called the major histocompatibility system by Snell (Snell *et al.*,1948). This initial work was carried out in the mouse and became an important model for studies of the Human Leucocyte Antigen (HLA) system.

Garrod in 1902 put forward the notion that there existed a relationship between the genetic constitution of an individual and his gene products, subsequently referred to as the well known one-gene-one enzyme (one polypeptide) theory. Through his work on the chemistry of the blood and other body fluids, a whole range of metabolic disorders have been identified. Among those, he described alkaptonuria which was the first example of a recessive inherited Mendelian disorder in man. This gene has only recently been localised to human chromosome 3q (Janocha *et al.*, 1994). In recognition of these achievements Garrod is referred to as the father of Biochemical Genetics.

The advent of electrophoresis (a technique that allowed the separation of proteins according to their physical properties) resulted in the discovery of the sickle-cell form of haemoglobin by Pauling in 1949. Very rapidly, this led on to the analysis of a vast spectrum of haemoglobin variation, as well as the analysis

of many plasma protein polymorphisms and enzyme polymorphisms in health and disease (Harris, 1975). The combination of electrophoresis and refinement strategies such as isoelectric focusing together with the development of methods for the specific detection of enzymes and other proteins allow the analysis of very large range of genes in families. The detailed studies on the variant haemoglobin polypeptides and the analysis of the structural variation underlying the haptoglobin polymorphisms layed the foundation of knowledge on the structural basis of mutation which introduced this chapter. It also set the scene toward the next major methodical developments in the analysis of genetic variation at the level of the gene itself.

It is only relatively recently, that studies have reverted from the investigation of altered proteins to the development of methods for the direct analysis of the DNA. Moreover, these techniques have facilitated the search for nucleotide changes in non coding region of the genome. These new so called molecular techniques enabled DNA to be visualised directly, for example by sequencing (Sanger *et al.*, 1977). As a result, more accurate diagnosis and better assessments of prognosis has become possible.

The next section considers the progress of the last ten years, in particular the utilisation of genetic variation within coding and non-coding region to allow the chromosomal assignment of disease loci and the range of approaches which have facilitated the identification of these genes and the characterisation of the mutations involved.

### **1.2.3 Analysis of genetic variation**

The following section looks at the approaches taken to analyse

mutations at the level of genomic DNA. At the present time changes responsible for genetic disease can be readily identified and the chromosomal position of different traits localised by linkage analyses. The most direct and unequivocal way of identifying the molecular changes underlying mutation is of course by DNA sequencing analysis. In the early days of molecular analysis, direct sequencing was employed to confirm the predictions from protein phenotypes in human genes such as the  $\alpha$  and  $\beta$  globin genes families. However this direct approach is clearly not feasible for the screening of large number of individuals and whole populations.

#### **1.2.3.1 Utilisation of restriction enzymes.**

The ability to manipulate DNA became possible with the discovery of a family of bacterial enzymes called restriction endonucleases which cleave DNA at specific nucleotide sequences (Arber *et al.*, 1974). Over a hundred such enzymes (Maniatis, 1991), have been isolated from various species of bacteria. Within bacterial cells these enzymes act to protect the cell from invading viruses by producing cuts in viral DNA. As a result, any foreign DNA can be fragmented and subjected to degradation by exonucleases. At the same time the chromosome of the cell is protected from destruction by the activities of additional enzymes, called methylases, which methylate these sites in the chromosomal DNA providing a means whereby "self" DNA can be distinguished from foreign DNA.

The DNA sequences recognized by most restriction endonucleases consist of four to six bases pairs symmetrically arranged as a palindrome. Some enzymes cleave both strands at the same point leaving blunt ends. Others make cuts which are offset, producing cohesive or sticky ends, so named because DNA cut in this way

tends to reassociate by base pairing. Initially this tendency of cohesive ends to anneal was used *in vitro* to join DNA from a foreign source with the DNA of a cloning vehicle, such as plasmid or phage DNA. For small viral genomes this provided a means to subdivide the DNA into smallish segments more amenable for extensive study.

Unfortunately the size of the human genome, approximately  $3 \times 10^9$  base pairs, prohibits its analysis in the same way. However it was possible to digest DNA into smaller fragments and to fractionate these depending on their size by gel electrophoresis. Through the introduction of Southern blotting (Southern, 1975) as a method of transferring DNA onto a solid support any cloned segment of DNA could be readily visualised by making this segment radioactive and hybridising it to the nylon support membrane containing the DNA. Restriction products showing homology to this region could be observed by autoradiography. Since restriction enzymes recognise specific DNA sequences, any changes in genomic DNA brought about by point mutations can lead to the loss or gain of a digestion site. Such variation in digestion patterns was first observed in the beta globin locus (Kan and Dozy, 1978) and this led to renewed efforts to identify genetic variation in the human genome.

These differences in the sizes of fragments resulting from the digestion of the corresponding region of DNA from homologous chromosomes have been termed restriction fragment length polymorphism (RFLP) (Botstein *et al.*, 1980). The use of more frequent cutters, for example 4 bp cutters, increase the likelihood of identifying RFLPs. In addition, enzymes, for example *Taq I* and *Msp I* containing the dinucleotide CpG in their recognition sequence (TCGA and CCGG respectively) also detect a far greater amount of polymorphic variation than other restriction enzymes. This reflects the methylation and

subsequent deamination of cytosine at a CpG dinucleotide in the genome which frequently results in the replacement of 5' methyl cytosine residue with thymidine leading to a transition mutation (pyrimidine/pyrimidine) (Cooper *et al.*,1988).

### **1.2.3.2 Minisatellites and microsatellites.**

A few RFLPs exhibit great variation in their length (Whyman and White, 1980; Jeffreys *et al.*,1985). These are termed minisatellites and consist of stretches of tandemly repeated units of between 11 and 60 base pairs in length (Nakamura *et al.*,1987). Jeffreys *et al.* (1985) made use of a minisatellite probe from the myoglobin gene which simultaneously revealed many polymorphic loci throughout the genome. The level of polymorphism was so high that it was possible to "fingerprint" an individual since the probability that two unrelated individuals share all fragments at cross-hybridising loci was negligible. Additional variable number tandem repeats (VNTR) were developed using the consensus nucleotides of some minisatellites to identify new VNTR loci for linkage studies. Their application however was limited by their tendency to cluster towards the telomeres of human chromosomes (Royle *et al.*,1990).

Other simple repeat sequences termed microsatellites, for example the poly(dC-dA)<sub>n</sub> could also display extensive length variation (Litt and Luty, 1989; Weber and May, 1989). These short simple repeat sequences are scattered throughout the genome with between 10-60 repeat units present in each block. The number of repeat units can be highly variable and through the application of the Polymerase Chain Reaction (PCR), an efficient system for their analysis has been developed. These repeats have the advantage that they occur once every 30-100 Kb in the genome (Stallings *et al.*,1991) therefore providing a vast new

resource of highly informative markers for the development of the human genome map. Fortunately, these microsatellites sequences are not restricted in their chromosomal position unlike minisatellites and have become the most commonly used marker in genetic studies (Weissenbach *et al.*,1992).

The PCR, introduced by Kary Mullis *et al.* in 1987 has revolutionised the whole field of molecular biology. It allows the enzymatic amplification of target sequences from small amount of genomic DNA. In the context of genetic diseases it has increased the efficiency of linkage analysis, but also it has allowed investigators to target regions of particular interest. Gene sequencing of these regions can then be used to evaluate and identify variation responsible for genetic diseases. Allele specific amplification, a simple and effective extension of the PCR reaction was developed by Wu *et al.*,1989. One of the pair of oligonucleotides is replaced by two oligonucleotides covering the mutated base. One of these would be perfectly matched on hybridisation and the other would create a mismatch at the 3' terminal with the same DNA. Conditions were established such that PCR product was produced in the former case and not in the latter case. This system was applied to the sickle allele of globin.

### **1.3 Mutation detection strategies**

Although sequencing is a standard procedure its application for the definitive analysis of whole genes from several different individuals is time consuming and laborious. Therefore additional screening methods which include Denaturing gradient gel electrophoresis (DGGE), Temperature gradient gel electrophoresis (TGGE), Single strand conformation polymorphisms (SSCP), chemical cleavage analysis (CCM), constant denaturing gel

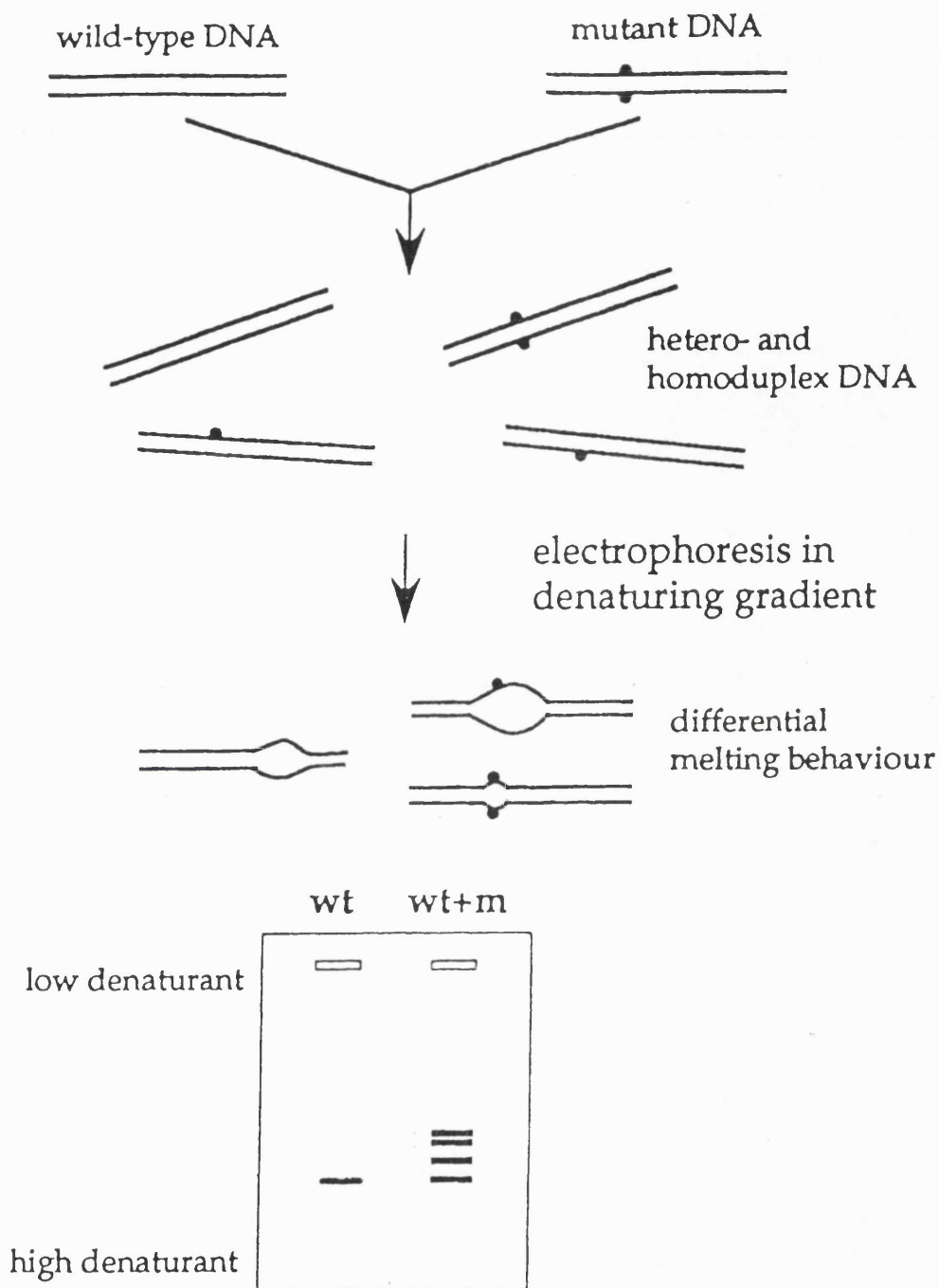
electrophoresis (CDGE), RNase-A mismatch cleavage method, DNA heteroduplexing have also been coupled with PCR to define the regions of the gene mutated in affected individuals, thus, directing sequencing approaches to small regions of particular interest. These strategies also allow efficient screening of a large number of individuals. Patients carrying the same mutation can be easily identified by comparing the position of aberrant DNA bands, avoiding the need for sequencing the corresponding region in every affected individual.

Naturally, the strategy used will vary on a number of parameters. In certain instances not only the sensitivity of a given approach but also its speed, simplicity and suitability for analysing both DNA and RNA templates in addition to the limitations of available sequence and structural information must be considered (see discussion 4.3).

I will now describe those techniques most widely used at the present time, in particular the DGGE which has been used predominantly for the work presented here. These have been reviewed most recently by Cotton (1993) and Grompe (1993).

### **1.3.1 Denaturing Gradient Gel Electrophoresis (DGGE) and related techniques**

DGGE, originally described by Fischer and Lerman in 1979, refined by Myers *et al.* in 1987 and Sheffield in 1989, involves the electrophoresis of double stranded DNA molecules in a polyacrylamide gel through increasing concentrations of chemical denaturing agents (formamide and urea) at a high ambient temperature (Fig.1.1). This technique adds another parameter to conventional electrophoresis which separates molecules according to size, charge and shape: the melting stability of



**Fig 1.1** Diagrammatical representation illustrating the principle of DGGE (Grompe, 1993).

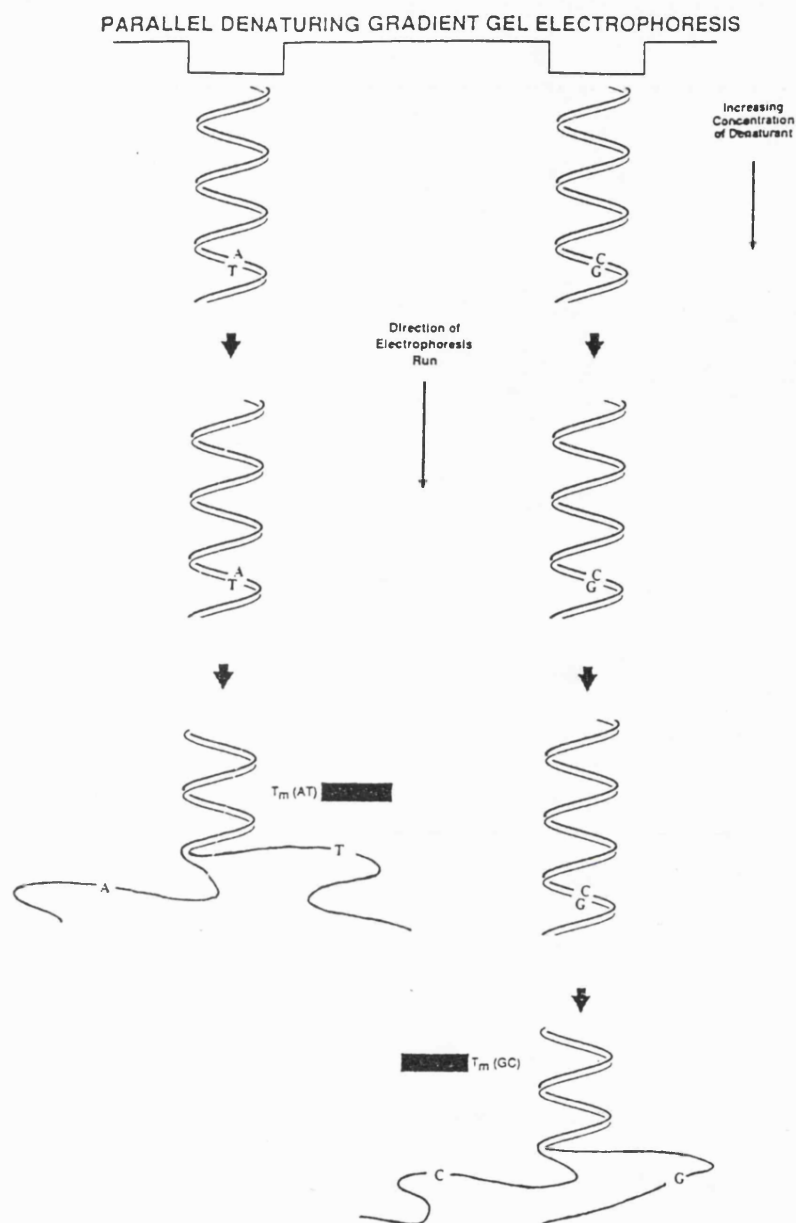
The figure takes as example a patient heterozygous for a particular mutation. PCR of the target DNA results in the generation of 4 different DNA duplexes, two homo and two heteroduplex forms. Electrophoresis, through a gel of increasing denaturant, results in the partial melting of the duplexes, in a sequence dependent fashion, leading to an altered pattern of migration. The four different bands in the wt + m lane correspond to the four different DNA duplexes.

biomolecules. DNA melts in domains at a temperature  $T_m$  which is directly related to the nucleotide sequence of the fragment of interest. A melting domain is a block of sequence within the fragment which melts cooperatively under the pressure of both high temperature and denaturing reagents.

Melting of DNA involves a transition from a tightly-packed double stranded, helical conformation to a single stranded random coil conformation and occurs in distinct regions called domains. Since the melting temperature ( $T_m$ ) of a melting domain, is directly related to the primary nucleotide sequence, a single base substitution within the region will result in a change in the melting temperature. This change can be utilised by the DGGE technique to resolve otherwise identical fragments of DNA. The coupling of PCR to DGGE allows small fragments of interest to be analysed independently.

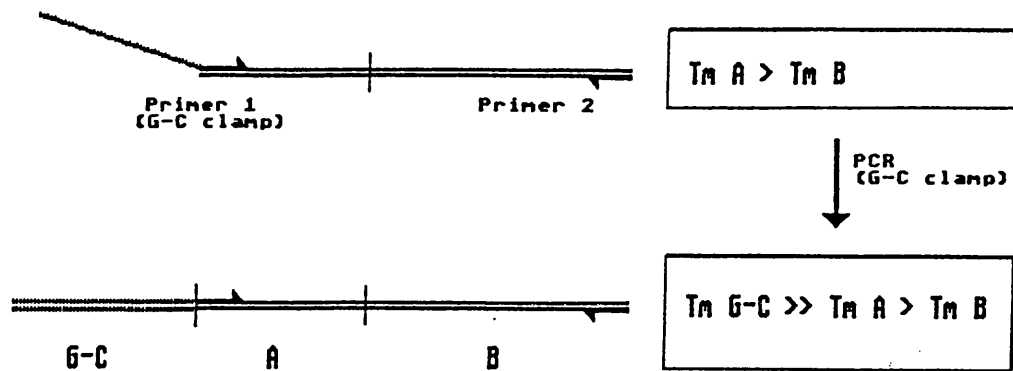
At the beginning of a denaturing gradient run, the double stranded (ds) DNA fragments run according to charge, size and shape into an increasing denaturing gradient until a point in the gradient is reached (corresponding to the  $T_m$ ) at which the lowest melting domain unwinds and electrophoretic mobility is therefore reduced (Fig. 1.2). A single base pair substitution in this melting domain will alter the  $T_m$  such that the partially melted DNA focuses at a higher or lower point within the gradient.

The melting of the highest melting domain results in the creation of single strands of complete melted DNA so that any variation within this domain will not be resolvable. Sheffield *et al.* (1989) increased the efficiency of the PCR based system by creating an artificial highest melting domain through the introduction of a 40 bp GC clamp at the 5' end of one of the PCR primers (Fig 1.3). This GC clamp now became an artificial highest melting domain increasing the number of mutations detectable by DGGE from 40%



**Fig 1.2 Diagrammatical representation comparing the melting properties of two different DNA duplex forms by Parallel Denaturing Gradient Gel electrophoresis.**

Two homoduplexes which differ by a single base substitution are electrophoresed through a gel of increasing denaturant. The change in base composition will result in a change in the stability of the DNA duplex. As the association between A-T is less stable than C-G the duplex containing A-T will denature earlier in the gradient. The partial melting of the duplex results in an abrupt decrease in mobility allowing its resolution as a distinct band.



**Fig 1.3 Diagrammatical representation of the generation of a high melting domain through the incorporation of a GC clamp at one end of the PCR product (Fodde *et al.*, 1994).**

Primers 1 and 2 generate a PCR product containing two separate melting domains, A and B. The melting temperature,  $T_m$ , for domain A is greater than the  $T_m$  of domain B. Therefore mutations in the B domain cannot be resolved as the DNA duplex will have completely dissociated. Through the addition of a GC clamp to either primer, in this instance primer 1, an artificially high domain is generated,  $T_m$  G-C is much greater than the  $T_m$  of domain A or B. Therefore all potential mutations in either domain can be resolved.

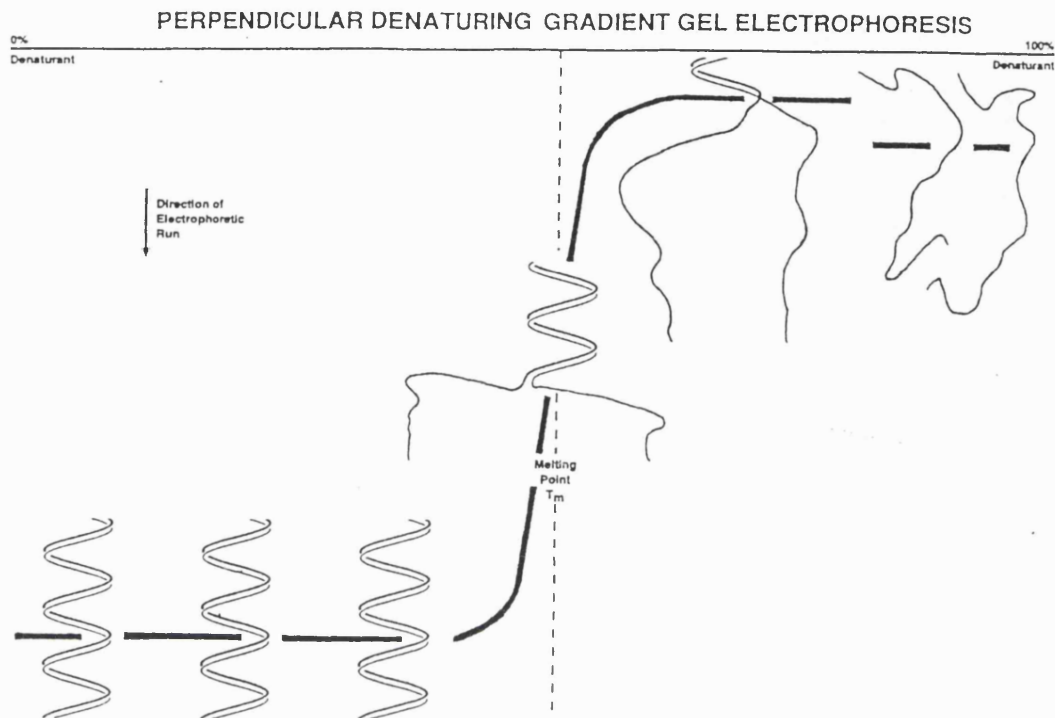
of all possible single base changes to close to 100% making it possible to resolve base changes anywhere within the amplified region.

DGGE resolves partially melted molecules. Strand dissociation leads to a loss of resolution. The irreversible transition due to strand separation can be shifted to higher temperatures through the use of GC clamps. The sequence of highest stability determines the temperature at which the strand dissociate completely. This results in a loss in continuity of the transition curve of a perpendicular denaturing gradient gel (Fig 1.4).

Computer analysis of the sequence of interest using two programs MELT87 and SQHTX (Lerman and Silverstein, 1987) enable the researcher to optimise the conditions for DGGE analysis. MELT87 calculates the melting map of a known sequence of DNA according to its nucleotide sequence. It plots the mid-melting point temperature at which each base pair is at 50% equilibrium between the helical and melted states as a function of its position along the DNA fragment. The electrophoretic conditions expected to provide optimal resolution and heteroduplex formation are determined by displacement calculation using the SQHTX program.

Among the first genes analysed using DGGE were the mouse and human  $\beta$  globin genes (Sheffield *et al.*, 1989). Since 1989, DGGE has found wide application in mutation detection including the genes for Cystic Fibrosis (Vidaud *et al.*, 1990; Fanen *et al.*, 1992; Mercier *et al.*, 1993), hemophilia A (Kogan and Gitschier, 1990; Higuchi *et al.*, 1991) and APC (Fodde, 1992). During the course of my studies, the DGGE technique has been applied as a mutation detection system, to a growing number of disease genes

Constant denaturant gel electrophoresis (CDGE) is a modification



**Fig 1.4 Diagrammatical representation of perpendicular gradient gel electrophoresis.**

The PCR product is electrophoresed perpendicular to a gradient of increasing denaturant. At the lower levels of denaturant the product remains double stranded until a concentration of denaturant is reached where the fragment starts to melt. This percentage denaturant corresponds to the melting of the lowest melting domain. In the example shown here the fragment contains a single melting domain. At higher denaturing conditions the product completely dissociates into single strands and can no longer be resolved.

of DGGE, initially described by Hovig *et al.* in 1991. It has mainly been applied to the p53 gene in human breast carcinoma and to the hypoxanthine phosphoribosyltransferase (HPRT) mutants. As with DGGE, the separation principle of CDGE is based on the melting behaviour of the DNA double helix of a given fragment. The gels used contain the same chemicals as for DGGE, but instead of a denaturant gradient, the gel is constituted of a uniform denaturant concentration. The configuration of the molecule is constant through the gel and the resolution is a function of the distance travelled. A specific denaturant concentration at which maximal separation between the wild-type and mutant fragments can be deduced from computer analysis using the algorithms developed by Lerman and by direct experimentation using a set of constant denaturant gels.

The denaturation or melting of the double stranded DNA can also be achieved by temperature and this method is termed temperature gradient gel electrophoresis (TGGE). As opposed to denaturing chemical gradients, TGGE induces no chemical interactions with the macromolecules and creates a stable non diffusable gradient. The TGGE approach is based on the same theoretical background however as DGGE, (Riesner *et al.*, 1989).

Another variant of DGGE is referred to as genomic DGGE (gDGGE). Genomic DNA is digested with a restriction enzyme, electrophoresed through a denaturing gradient gel, transferred to a nylon filter and finally hybridised to a unique DNA probe (Borrensens *et al.*, 1988). No PCR amplification is involved prior to the electrophoresis. The main difference relative to the "classic" DGGE is that it is not limited to a specific target (any available unique probe of any length can be used). Furthermore, no sequence information is required. The obvious drawback is that no clamps can be introduced and the approach relies entirely on the presence of "natural" clamps (sequences which have a higher

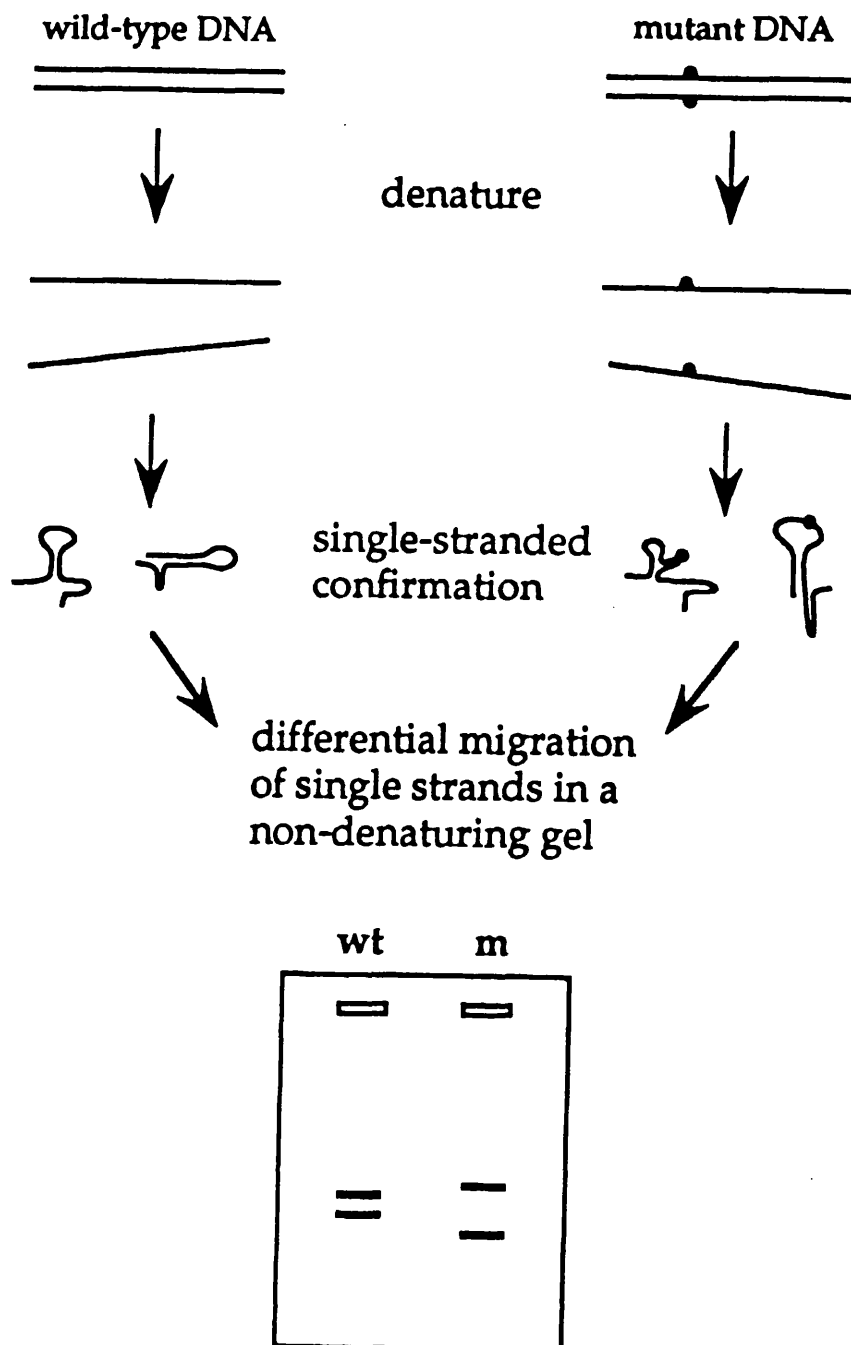
melting domain than the sequence containing the mutation). The mutation detection efficiency is therefore dramatically reduced to 20-60%. This might explain why this procedure has not been widely used so far. The use of different restriction enzymes in combination can overcome the poor efficiency to some extent and this modified system has been successfully used to screen for mutations (Gray *et al.*, 1992; Burmeister *et al.*, 1991, Krolewski *et al.*, 1992).

### **1.3.2 Single strand conformation polymorphism (SSCP).**

SSCP looks specifically at variation in single stranded molecules (Orita *et al.*, 1989). In non denaturing conditions, single-stranded DNA has a folded (tertiary) structure that is determined by intramolecular interactions and these in turn depend upon primary nucleotide sequence. A mutated sequence is therefore detected as a change in mobility caused by its altered folded structure (Fig 1.5). The high resolving power of polyacrylamide gel electrophoresis makes it possible to distinguish most conformational changes caused by subtle sequence differences such as a single base substitution in relatively small DNA fragment.

### **1.3.3 Heteroduplex analysis (HA).**

Heteroduplex DNA molecules, which differ in sequence may show different mobility in regular polyacrylamide gels from homoduplex DNA. This property is brought about by sequence dependent conformational changes in the dsDNA. The introduction of new gel matrices (such as Hydrolink) have enhanced the ability to detect mobility shifts in heteroduplex molecules. This technique was initially applied to the detection of mutations in



**Fig 1.5 Diagrammatical representation of Single Strand Conformation Polymorphism (SSCP) (Grompe, 1993).**

A single stranded molecule adopts a particular conformation as determined by its intramolecular interactions. As the sequence of a product varies so will these interactions result in different conformations and therefore altered mobility of the strand. SSCP is based on the observation that these changes in conformation can have a profound effect on electrophoretic mobility.

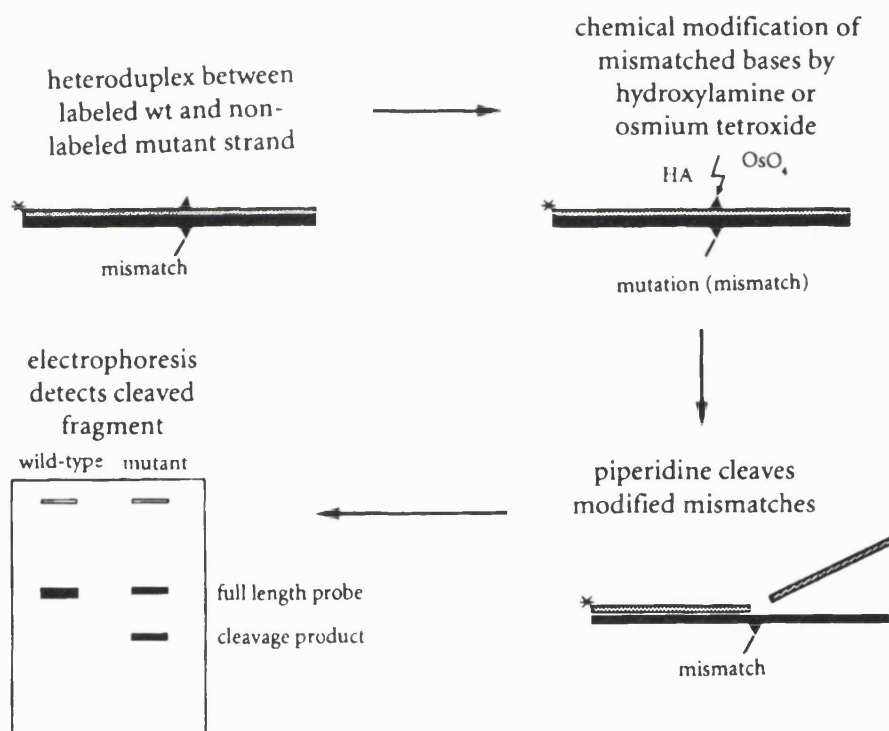
the rhodopsin gene of autosomal dominant Retinitis Pigmentosa patients (Keen *et al.*,1991) and has since been used to study a large number of additional diseases (Tassabehji *et al.*,1992; Artlich *et al.*,1992).

#### **1.3.4 Chemical Cleavage of Nucleic Acid Heteroduplexes, also called Chemical Mismatch Cleavage (CMC).**

The hydroxylamine and osmium tetroxide chemical cleavage method was first described by Cotton *et al.* in 1988. As in most other mutation detection systems the technique has been coupled with the PCR to generate target sequence (Fig 1.6).

A heteroduplex between a wildtype DNA molecule and mutant DNA (or RNA) is created by boiling and reannealing. Maxam-Gilbert sequencing chemistry is then used to chemically modify mismatched bases at the sites of mutations in the DNA-DNA or DNA-RNA heteroduplexes. Osmium tetroxide is used for the modification of mispaired thymidines and hydroxylamine for mismatched cytosines. Adenosine and guanosine mismatches of the wild-type sense strand are detected by also labelling the anti-sense strand of wild-type DNA in the heteroduplex. The labelled DNA is cleaved by piperidine at the site of the modification, and this is followed by denaturing polyacrylamide gel electrophoresis (PAGE) and autoradiography.

The wild type probe is either uniformly labeled or end-labeled and is hybridized to the unlabeled template DNA; a point mismatch is depicted as a bulge in Fig 1.6. After modification and cleavage of mismatched regions, the products are visualised by autoradiography. The wild type DNA remains uncut, and mutant DNAs display a mixture of cut and uncut products. If the probe is uniformly labeled two cleavage products are seen; if the probe is



**Fig 1.6 Diagrammatical representation of chemical mismatch cleavage (CMC) (Grompe, 1993).**

A heteroduplex between labelled (\*) wild-type DNA and unlabelled mutant DNA is shown. The mismatched base at the site of the mutation is chemically modified and subsequently cleaved by piperidine.

end-labeled only the fragment retaining the labeled end is observed.

### **1.3.5 RNase protection Assay: A mismatch cleavage method**

This method uses the ability of bovine pancreatic RNase to recognise and to cleave a significant percentage of single-base mismatches in RNA:RNA (Winter *et al.*,1985) or RNA:DNA (Myers *et al.*,1985) duplexes. A homogeneously labeled RNA probe, complementary to the gene being studied, is hybridised to cellular RNA or DNA, and the hybrids are digested with RNase A at the site of the mismatch. The resistant products are analysed by electrophoresis in denaturing polyacrylamide gels and autoradiography. Point mutations in the gene or gene transcripts are detected by the presence of mismatch-specific sub-bands. The size of the sub-bands is dependent on the position of the mismatches, therefore the mutations can be localised within the gene with an error of only a few nucleotides. This approach has been used in the identification of mutations in only a small number of genes which include the ras oncogene (Winter *et al.*,1985), the HPRT gene (Gibbs and Caskey, 1987) and the retinoblastoma gene (Dunn *et al.*,1989). However, since the introduction of PCR to generate target templates the technique in the main has been made obsolete.

### **1.3.6 Direct sequencing (DS)**

DNA sequencing defines the location and nature of the change and therefore is a final step to the mutation detection methods described previously. DS refers to the direct sequence analysis of PCR products. Several methods have been described to achieve

this. The asymmetric PCR method involves the reamplification of PCR product in a second reaction in which one of the oligonucleotide primers is in >100:1 excess of the other. This creates single-stranded PCR product which can be sequenced. The biotin method involves the biotinylation of one of the primers used for the PCR reaction. After the reaction, the double-stranded PCR product is captured on an avidin-coated magnetic bead. The non-biotinylated strand is melted away with NaOH and the sequencing reactions are carried out on the immobilized single-stranded template. The introduction of automated sequencing using commercially available machines from Pharmacia and Applied Biosystems has increased the speed and efficiency of sequence analysis over conventional radioactive sequencing.

#### **1.4 Model systems investigated.**

The following section briefly outlines the three model systems investigated in this study, the alpha 1 antitrypsin gene (AAT), a segment of the Y chromosome and the phenylalanine hydroxylase gene. These three model systems were used to evaluate the general applicability of DGGE for mutation detection analysis in the context of pathology and diagnosis of genetic disease and more generally in the detection of polymorphisms.

##### **1.4.1 Alpha 1 antitrypsin (AAT).**

AAT is a major proteinase inhibitor present in high concentrations in the plasma. It is a glycoprotein with a molecular weight 52,000 which is synthesised mainly in the liver. It inhibits a wide range of serine proteinases, although its principal target is thought to be neutrophil elastase (Heidtman *et al.*, 1986).

It is associated with early onset pulmonary emphysema and a form of neonatal liver disease, both of which are associated with abnormally low levels of plasma AAT (Fagerhol *et al.*,1981). The gene locus encoding AAT, called PI, is located on chromosome 14q31-32.3 (Darlington *et al.*,1982) and displays high levels of genetic polymorphisms at the protein level (Whitehouse *et al.*,1989).

In this regard the existence of six common and well characterised point mutations (M1 Ala, M1 Val, M2, M3, S and Z), associated with variant protein phenotypes, made this an ideal system to evaluate the general applicability of DGGE for detection of single base pair polymorphism.

#### **1.4.2 Y chromosome.**

The Y chromosome covers 1.7% of the whole genome comprising about  $5 \times 10^7$  bp. The long arm (almost 70% of the Y chromosome) is composed of highly and moderately repeated sequences and constitutes the large heterochromatic region. In contrast, the short arm contains the pseudoautosomal region (Yp11.2), which shows a high level of homology with the X chromosome. Y specific DNA probes from these regions are useful for various purposes including prenatal diagnosis of sex-linked disorders, following up the host-versus graft cells in patients with bone marrow transplants from the opposite sex and analyses of sex chromosomal disorders such as a Y/autosome translocation (Malaspina *et al.*, 1990).

In comparison to the extensive polymorphism observed in the genome (section 1.2.3), the Y chromosome contains very few polymorphic sites. Malaspina *et al.*, in 1990 failed to detect any RFLPs using 12 probes. In their study, although 2215 nucleotides were screened per subject, corresponding to a grand total of

46,515 nucleotides no polymorphisms were detected. Jakubiczka (1989) studied an additional 10 probes and only one polymorphism was detected. To date, similar RFLP studies have revealed polymorphisms at only 8 Y-specific loci (Malaspina *et al.*, 1990).

At the time of my study RFLP analysis had been the only approach used to screen for mutations on the Y chromosome. In this study, DGGE was used to look for variation in a 500 bp stretch of DNA from the Yp11.2-Ypter region, designated as pY-80 which had been completely characterised and had been demonstrated to be male specific (Tsukahara *et al.*, 1990).

Furthermore the MELT87 and SQHTX algorithms, devised by Lerman in 1987, to optimise the conditions for mutation detection using DGGE, were used prior to the practical investigations of this second model system. This also allowed me to evaluate the use of artificially mutated nucleotide primers as controls.

### **1.4.3 Phenylketonuria (PKU).**

Classical phenylketonuria (PKU) was first described by Asbjorn Folling in 1934. He established a relationship between the level of phenylpyruvic acid and the presence of mental retardation in a subset of 430 retarded children. As phenylpyruvic acid and phenylalanine (Phe) show a similar structure Folling suspected that phenylalanine might be the most probable source of phenylpyruvic acid. He called it "Imbecilitas phenylpyrouvica", reflecting its main clinical manifestations of irreversible mental retardation, behavioural disturbances and convulsive seizures and the importance of phenylpyruvic acid. The presence of 4 pairs of affected siblings in his affected group suggested a genetic component in the aetiology of the disease. A year later, Penrose

confirmed an autosomal recessive transmission for the disease renaming it PKU to reflect the high Phe level in affected individuals (Penrose,1935).

It was more than ten years later, in 1947, that the underlying biochemical pathway for PKU started to be deciphered by Jervis. He noticed that increased levels of tyrosine (Tyr) followed the ingestion of Phe in the blood of normal individuals but not in PKU patients. He confirmed these initial observations by comparing the transformation of Phe into Tyr in post mortem liver tissue from normal and PKU affected individuals and found that normal individuals could transform Phe into Tyr, whereas this reaction was absent in PKU patients.

Udenfriend and Cooper later identified this enzymatic system as a hydroxylation step which converted Phe to Tyr (Udenfriend and Cooper, 1952). It soon became apparent, from later work by Mitoma *et al.* that two protein products were required to sustain Phe hydroxylation (Mitoma, 1956), a phenylalanine hydroxylase (PAH) and a pteridine reductase (BH<sub>4</sub>) (Kaufman, 1959, 1963). A diagrammatic representation of the Phe hydroxylation pathway in man, as described by Kaufman in 1976, is presented in Fig 1.7. Although a defect in either protein could be responsible for the PKU phenotype most patients lacked functioning PAH rather than pteridine reductase (Kaufman, 1976). It is now known that only about 2% of PKU cases result from the BH<sub>4</sub> deficient form of PKU.

As a consequence it was conceivable that by limiting the importance of the hydroxylation reaction, ie by reducing the intake of Phe, the clinical implications of the disease could be controlled. The introduction by Woolf and Vulliamy of low Phe dietary therapy was found to greatly reduce clinical symptoms of the disease if started early in life. The severity of the disorder and the importance of early treatment, led to the implementation

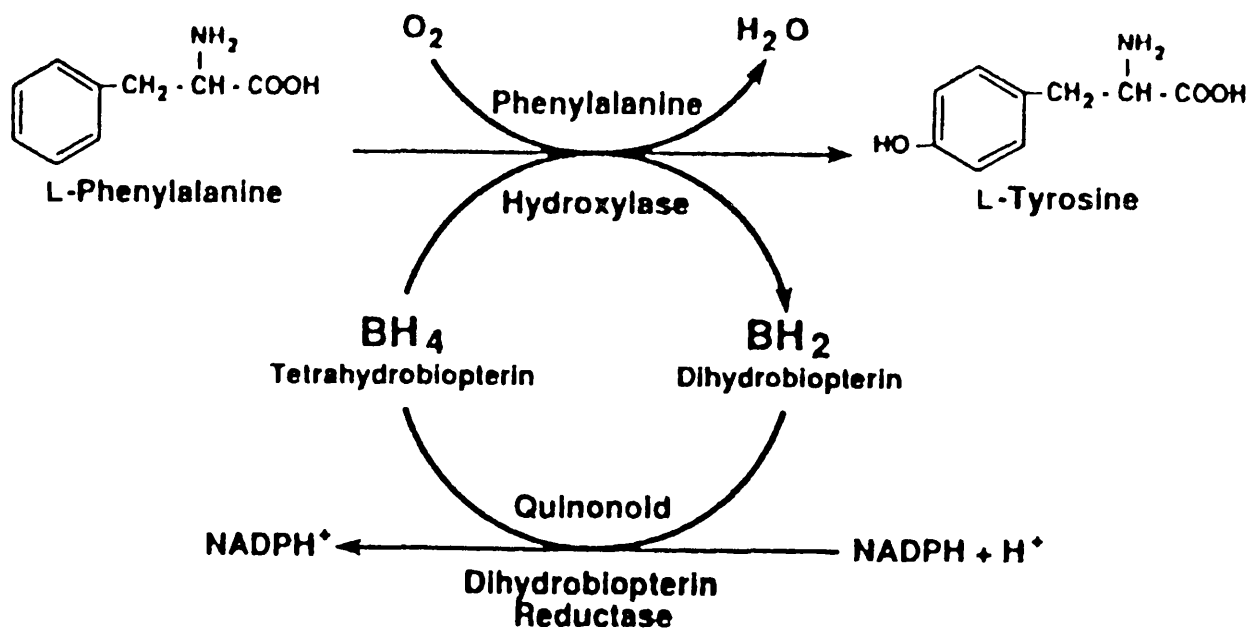


Fig 1.7 Diagrammatical representation of the phenylalanine hydroxylation pathway in man as described by Kaufman in 1976.

of a neonatal screening program in 1959 in Great Britain (Gibbs and Woolf, 1959; Boyd, 1961). The first assay used was the urinary ferric chloride test originally developed by Folling when he first described the disease. However, the discovery by Armstrong and Binkley in 1956 that excretion of phenylpyruvate is preceded by an elevation of Phe in serum led to the introduction of the Guthrie test (Guthrie, 1961; Guthrie and Susi, 1963), a bacterial inhibition assay for the detection of serum Phe levels. This gives an estimate of the frequency of PKU among Caucasians at approximately 1/10,000 live births corresponding to a carrier frequency of 1 in 50.

In addition to the neurological features of PKU, patients also exhibit hypopigmentation, other dermatological conditions and impaired postnatal growth among others (Scriver *et al.*, 1989). Patients can be subdivided into 3 distinct groups; classical PKU, mild PKU and benign hyperphenylalaninemia (Guttler, 1980). Classical PKU refers to the most severely affected class of hyperphenylalaninemic patients who would be profoundly retarded unless strict dietary therapy was implemented. Mild PKU refers to patients less severely affected than the previous group. They require dietary therapy but have a relaxed tolerance for Phe, relative to the classical group. Finally, benign hyperphenylalaninemia describes hyperphenylalaninemic individuals clinically normal without dietary therapy.

Following the cloning of the human PAH gene by Kwok *et al.* in 1985 it became clear that this clinical heterogeneity reflected parallel variation at the molecular level. Initial haplotype analysis suggested that PKU was caused by a small number of founder mutations, for example the first mutation identified, a splice site mutation in intron 12, accounted for almost 40% of mutations in the Danish population (DiLella *et al.*, 1986). Similarly, in the Bulgarian PKU population the R408W mutation in

exon 12 accounted for 34% of PKU alleles on a single haplotype background (Kalaydjieva *et al.*, 1990).

Subsequently 8 RFLPs have been identified in or near the human PAH locus. Since these RFLPs are tightly linked to the PAH gene, they have been used to follow the transmission of normal or mutant chromosomes in PKU families (Lidsky *et al.*, 1985). Although over 70 haplotype patterns have been identified, most PKU chromosomes are associated with haplotypes 1,2,3,4 in Caucasians populations. The application of haplotype analysis permitted the development of prenatal diagnosis of PKU (Riess *et al.*, 1987). This approach however has only limited application, for example prenatal diagnosis using RFLP analysis can only be provided to those families with a prior incidence of PKU. Unfortunately, over 95% of new cases of PKU are the result of random mating between previously undiagnosed carriers without a family history and thus are undetectable by this kind of haplotype analysis.

To overcome these limitations, the specific molecular lesions responsible for PKU must be identified and characterised in each major population group, to allow more rapid screening of carrier individuals. The introduction of PCR coupled mutation detection systems (section 1.1.1.3 ) allowed more detailed analysis of the PAH gene using chemical cleavage mismatch (Cotton *et al.*, 1990) and single strand conformation polymorphism (Dockhorn-Dworniczak *et al.*, 1991). By 1991, thirty one PAH mutations had been identified in PKU patients, spread over all 13 exons of the gene (Konecki *et al.*, 1991). The study also confirmed previous haplotype analyses which suggested that certain mutations predominate in a given population.

The identification of a large spectrum of PKU mutations at the PAH locus suggests a need to choose a technique which offers a

high degree of accuracy to screen an affected population. The objective of this study was to evaluate the DGGE technique as a rapid and reproducible method to identify and estimate the frequency of PKU mutations in a South East of England affected population. It was anticipated that this would confirm previous observations that the phenotypic heterogeneity observed in PKU reflected an underlying heterogeneity of the PAH genotype and would help define the correlation between a particular PKU mutation and its phenotype. In addition it was hoped that through a knowledge of the genotype of an individual the various levels of response of individuals to similar diets could also be explained. Moreover screening services could be made available to families for prenatal diagnosis or carrier detection of the disease.

### **1.5 Aims and objectives.**

Success in the isolation and characterisation of human genes responsible for genetic diseases coupled with the considerable molecular heterogeneity at these loci, has placed growing emphasis on the importance of mutation detection techniques both in coding and non-coding regions.

This thesis is concerned with the evaluation of one of these techniques : Denaturing Gradient Gel Electrophoresis (DGGE). This approach is tested in three model systems : 1. the well characterised human alpha -1 - antitrypsin (AAT) gene, with a range of known mutations, 2. a segment of the Y chromosome where there was no evidence for mutations and 3. the phenylalanine hydroxylase (PAH) locus where considerable heterogeneity has been encountered in families affected with phenylketonuria (PKU).

The AAT gene was used as the first test locus to evaluate the general applicability of DGGE for the detection of well defined

point mutations associated with variant protein phenotypes. The approach was largely experimental and involved a combination of perpendicular and parallel DGGE.

The computer algorithms designed by Lerman *et al.*, 1987 were used to optimise the conditions for DGGE screening using the second model system, a segment from the Y chromosome designated as pY-80. The approach was used to screen for mutations across this interval.

The most extensive study was carried out on the PAH gene where samples from fifty unrelated PKU patients and their relatives living in the South-East of England were investigated. The objective of this study was to evaluate the DGGE technique as an efficient method to identify and estimate the frequency of PKU mutations. It was anticipated that this would confirm other PKU population studies showing that the phenotypic heterogeneity observed in PKU reflected an underlying heterogeneity of the PAH genotype. In addition, it was hoped there would be a correlation between the genotype and the quality of response of patients to diet. Moreover screening services could be offered to families for prenatal diagnosis or carrier detection of the disease.

## **Chapter 2 MATERIALS AND METHODS**

### **2.1 Materials and Suppliers**

Unless otherwise stated, solutions, plastic or glassware were sterilised by autoclaving at 15psi, 121°C for 20-25 min.

#### **Glassware**

Standard laboratory glassware was washed in detergent and rinsed in double distilled water prior to autoclaving.

DGGE glass plates were purchased from Hoefer Scientific Instruments.

#### **Plasticware**

Disposable 50ml sterile screwtop tubes were purchased from Falcon (referred to as Falcon tubes). 1.5 ml microcentrifuge tubes (MCC) and 0.5 ml microcentrifuge tubes were purchased from Elkay and blue (1ml) and yellow (200µl) pipette tips from Algen.

DGGE plastic clamps and cams were purchased from Hoefer Scientific Instruments.

#### **General Reagents**

General laboratory reagents were of analytical grade and were purchased from BDH unless specified below

Formamide

Sigma Chemical Co. Ltd

Ficoll

Sigma Chemical Co. Ltd

Xgal	Bethesda Research Lab
Agarose	Sigma Chemical Co. Ltd
3MM paper	Whatman
Detergent	
(RBS 35 special glassware)	Pierce
Analytical Grade	
Mixed Bed Resin	Biorad

### Radionucleotides

[ <sup>35</sup> S] dATPaS	Amersham
---------------------------	----------

### Sequencing Reagents

Acrylamide	Kochlight Ltd.
Ammonium persulphate	Kochlight Ltd.
Boric acid	Sigma Chemical Co. Ltd
N,N',N',N'-tetra- methylethylenediamine (TEMED)	Sigma Chemical Co. Ltd
Urea	Sigma Chemical Co. Ltd

### Enzymes

Restriction enzymes (EcoRI)	BRL
--------------------------------	-----

### Kits

TA cloning kit	Invitrogen
Sequencing version 2.0 kit	United States Biochemical Corporation

### PCR reagents

Taq polymerase	Bioexcellence, Anglian, Applied Biosystems, Promega
10 X buffer	Applied Biosystems, Promega
Nucleotides	Pharmacia
Mineral oil	BDH
Glycerol	BDH
Oligonucleotides	Oswells, or in house Applied Biosystems synthesiser

### Media Reagents

Agar	Difco
Ampicillin	Sigma Chemical Co. Ltd
Kanamycin	Sigma Chemical Co. Ltd
Tryptone	Oxoid Ltd.
Yeast Extract	Oxoid Ltd.
Glucose	BDH

### General Solutions

TE buffer: 10mM Tris-HCl pH7.5, 0.1mM sodium EDTA pH8.0

10 X loading buffer: 25% (w/v) ficoll (Mr 400kD), 0.25% (w/v) Orange G, 0.25M EDTA in 1 X TAE buffer

10 X TAE: 0.4M Tris-acetate, 10mM sodium EDTA, pH 8.0

10 X TBE: 90mM Tris borate, 10mM sodium EDTA, pH 8.3

### DGGE Reagents

Acrylamide/Bis 37.5: 1	Biorad
Ammonium persulphate (APS)	Kochlight Ltd.
N,N',N',N'-tetra-	

methylethylenediamine (TEMED)	Sigma Chemical Co. Ltd
Urea	Sigma Chemical Co. Ltd
Formamide	Sigma Chemical Co. Ltd
Analytical grade mixed bed resin	Biorad

### **Solutions used for DGGE and gel analysis**

20 X TAE : 800mM Tris-HCl, 400mM Na Acetate, 20mM Na<sub>2</sub> EDTA, pH 7.4 adjust with Acetic Acid

DGGE loading buffer: 20% ficoll, 10mM Tris/pH 7.8, 1 mM EDTA, 0.5% orange G, 0.5% xylene cyanol, 0.5% bromophenol blue

### **Media and solutions for microbiological manipulation**

LB broth: (per 100ml) Bactotryptone 1g, Bacto-yeast extract 0.5g, NaCl 1g, Glucose 0.1g

LB agar: (per 100ml) Bactotryptone 1g, Bacto-yeast extract 0.5g, NaCl 1g, 1.5g Bacto-agar

Cell resuspension solution : 50mM Tris-HCl pH 7.5, 10mM EDTA, 100µg/ml RNase A

Cell lysis solution 0.2M NaOH, 1% (w/v) SDS made fresh

Neutralisation solution : 2.55M, potassium acetate pH 4.8

Column Wash solution: 200mM NaCl, 20mM Tris-HCl, pH7.5, 5mM EDTA (diluted 1:1 with 95% EtOH)

TE buffer: 10mM Tris-HCl, pH7, 1mM EDTA

### **Films**

positive/negative	Polaroid 665+/- Instant
	pack film
Hyper-film	Amersham

## **2.2 Methods**

### **2.2.1 Isolation of DNA**

#### **2.2.1.1 DNA extracted from whole blood**

DNA was extracted from blood essentially as described by Lahiri and Nurnberger Jr, 1991. An equal volume of low salt buffer (TKM1:10mM Tris-HCl pH 7.6, 10mM KCl, 10mM MgCl<sub>2</sub> and 2mM EDTA) was added followed by 125µl of Nonidet P-40 to lyse the cells. The cells were mixed well and pelleted by centrifugation at 6000 g. The supernatant was discarded and the pellet washed in TKM2 buffer (TKM1 and 0.4M NaCl) and pelleted as before. The pellet was resuspended in 0.8ml of high salt TKM2 buffer. 50µl of 10% SDS was added and mixed thoroughly before incubating for 10 min at 55°C, 0.3ml of 6M NaCl was added and the suspension was centrifuged for 5 min in a table microcentrifuge. The supernatant was retained and 2 x vol of 100% ethanol added to precipitate the DNA. The DNA was removed and washed in 70% ethanol and precipitated as before. The pellet was dried in a speed-vac (Howe, at 30°C for 15min) and resuspended in 0.5ml of TE. The concentration of the DNA was measured as described in section 2.2.2.

#### **2.2.1.2 Plasmid DNA extracted from bacterial cultures**

DNA probes were prepared from 1-3 ml bacterial cultures, grown overnight at 37°C, using Magic Minipreps DNA Purification system (Promega). Specifically the pellet was resuspended in 200 µl of resuspension buffer and transferred to a fresh tube before the addition of 200 µl of cell lysis solution. After inverting a few times, 200 µl of neutralising solution was added and the resulting suspension was pelleted by centrifugation in a bench centrifuge. 1 ml of purification resin was added to the supernatant and the DNA suspension was sucked under vacuum through a column where the DNA was retained. The DNA was washed by sucking through a further 2 ml of column buffer prior to being eluted by the addition of a small amount of preheated distilled water, typically 30-60 µl, followed by a brief 20 sec centrifugation in a bench centrifuge. DNA was quantified as described in section 2.2.2.

### 2.2.2 Concentration of DNA

The concentration of the DNA was measured by spectrophotometry in a CIBA CORNING 2800 spectrascan. Optical density (OD) readings were taken at 260nm and 280nm. An OD of 1 at 260nm corresponds to a approximately 50 µg/µl for double-stranded DNA and 40 µg/µl for single-stranded DNA. The ratio between the readings at 260nm and 280nm ( $OD_{260}/OD_{280}$ ) provided an estimate of the purity of the sample, a pure DNA preparation, without protein or phenol contamination having a value of 1.8. Alternatively the concentration of DNA was estimated by direct comparison of a small amount of DNA with a known amount of standard 1 kb ladder after gel electrophoresis.

### 2.2.3 PCR primer sequences and PCR reaction conditions

This section provides the primer sequences and the PCR conditions for the three different subjects studied.

#### 2.2.3.1 Alpha-one-antitrypsin gene

##### Primer sequences

##### Exon primer

##### 2

5' end \*GCA GGA CAA TGC CGT CTT CTG TCT C

3' end \*CCA CTA GCT TCA GGC CCT CGC TGA G

##### 3

5'end \*CTG GTG ATG CCC ACC TTC CCC TCT C

3' end \*GTC ACC CTC AGG TTG GGG AAT CAC C

##### 5

5' end GGG CCG TCC AGG GGG AGT GAG CGC T

3' end \*GGG TTT GTT GAA CTT GAC CTC GGG

40 bp GC clamp designated as: \* above

cgc ccg ccg cgc ccc gcg ccc gtc ccg ccg ccc ccg ccg c

##### PCR conditions

Exon	temperature			number of cycles
	denaturing	annealing	elongation	
2	92 <sup>0</sup> c	55 <sup>0</sup> c	72 <sup>0</sup> c	35
3	93	54	72	35
5	92	52	72	35

Taq polymerase used : Bioexcellence

### 2.2.3.2 Y sequences originated from the pY-80 fragment

Primers sequences used to amplify two fragments of the Y chromosome (originated from pY-80).

<b>Y1</b>	<b>primer</b>
5' end	AGG CCT GTC AAC CCC CAG CCC TGG G
3' end	*ACT GTG AGG GCC CAG GTT ACG GCA C
5' end mutant	AGG CCT GTC AAC CCC CAG CCC TGG GCT
(G->A)	GCT TCC CTG <u>A</u> CC TCT TCT CTG

<b>Y2</b>	<b>primer</b>
5' end	AGG TCC TTT GCC ATG ATG CCC CAT AG
3' end	*GGA GTA TTG CGT TGG CA
5' end mutant	AGG TCC TTT GCC ATG <u>A</u> GG CCC CAT AG
(T->G)	

40 bp GC clamp sequence designated as:\* above  
 cgc ccg ccg cgc ccc gcg ccc gtc ccg ccg ccc ccg ccg c

PCR conditions

fragment	temperature			number of cycles
	denaturing	annealing	elongation	
Y1	92 <sup>0</sup> c	52 <sup>0</sup> c	72 <sup>0</sup> c	35
Y2	91	50	72	35

Taq polymerase used: Anglian

## 2.2.3.3 PAH gene

Primer sequences used to amplify 5 exon and splice site junction regions of the PAH gene. The primer pairs for exon 10 and 11 were taken from Guldberg *et al.*, 1993.

Exon	Primer
<b>3</b>	
5' end	TAG GTT TTC CTG TTC TGG TTC TG
3'end	*CTT ATG TTG CAA AAT TCC TCT AAT TC
<b>5</b>	
5' end	*GCA CTG TCA TGG CTT TAG AGC CCC C
3' end	CAT CCT CAA CTG GAT GAG GGC AAG G
<b>10</b>	
5' end	*TTA ACG ATC ATA GAG TGT GC
3' end	ACA AAT AGG GTT TCA ACA AT
<b>11</b>	
5' end	TGA GAG AAG GGG CAC AAA TG
3'end	*GCC AAC CAC CCA CAG ATG AG
<b>12</b>	
5' end'	CGC CAA ATG GTG CCC TTC ACT CAA GC
3' end	*GTC TTC GAT TAC TGA GAA ACC GAG

40bp GC clamp designated as \* above

cgc ccg ccg cgc ccc gcg ccc gtc ccg ccg ccc ccg ccg c

PCR conditions

Exon	temperature			number of cycles
	denaturing	annealing	elongation	
3	94 <sup>0</sup> c	52 <sup>0</sup> c	72 <sup>0</sup> c	35
5	94	54	72	35
10	94	54	72	35
11	94	56	72	35
12	94	54	72	35

Taq polymerase used : Promega and Applied Biosystem

#### **2.2.3.4 Polymerase chain reaction (PCR) protocol.**

The PCR were carried out in a Hybaid automated heating block. The gilsons used to set up the reaction, gilson tips and centrifuge tubes were kept separate to those used in other experiments. Reactions were set up in a separate room to the one used for gel electrophoresis of the PCR products.

The reaction mix contained 1 X PCR buffer, 200 $\mu$ M dNTPs, 1 $\mu$ g genomic DNA, 50pMol of each primer and 2 units of Taq Polymerase in a final volume of either 50 or 100 $\mu$ l made up to volume with double distilled water. The reaction was given an initial 5 mins denaturation step and layered with paraffin to prevent evaporation of the mix. The conditions of the PCR were primer dependent as described previously (2.2.3.1 to 2.2.3.3).

After completion of the amplification reaction, the products were analysed by agarose gel electrophoresis.

#### **2.2.4 Agarose gel electrophoresis**

Agarose gel electrophoresis was used to assess the specificity and the efficiency of polymerase chain reaction. PCR products were resolved in 1.5-2.0% agarose gels prepared and run in 1 X TBE buffer. Samples were loaded with the addition of 0.1 vol of 10 X loading buffer. The DNA was stained with ethidium bromide at 0.5 mg/ml for 10 min and visualised under ultra-violet transillumination.

#### **2.2.5 Denaturing gradient gel electrophoresis.**

##### **2.2.5.1 Gel preparation.**

Denaturing gradient gel electrophoresis was carried out as described by Myers *et al.*, 1987 using a Hoeffer SE600 vertical electrophoresis apparatus placed in a purpose built tank fitted with a Grant TA circulating immersion thermostat containing TAE buffer (Tris-acetate 0.04M, EDTA 0.001M) pH 7.4 at 61°C.

The gels were made in glass moulds, 18 cm long and 14 cm wide with spacers of 0.75mm. Samples were run in acrylamide gels of varying concentration, depending on the size of the fragment being studied (generally, fragments of about 200bp in length were run in 10% acrylamide, while fragments of approximately 400 bp were run in 6.5% acrylamide).

The glass plates (14 cm x 18 cm x 0.75 mm) were cleaned in a solution of the special glassware detergent (RBS 35) followed by alcohol absolute and then acetone. The plates were finally rinsed in deionised water. The glass plates were dried and the spacers placed between the glass plates and clamped into position. This glass mould was then placed in a gel pouring stand.

6 ml solution of 10% acrylamide 0% denaturant was prepared, polymerised by the addition of 3µl of TEMED and 35µl of a 25% solution of ammoniumperoxydipersulphate (APS) and used to seal the bottom of the gel mould. The unpolymerised surface liquid layer was poured off.

A linear gradient of denaturant up to a maximum concentration of 40% formamide and 7M urea which is equivalent to 100% denaturant was prepared using standard stock solutions (Table 2.1). 9ml aliquots of the solutions corresponding to the highest and the lowest concentration of denaturant needed were prepared using either the 6.5% acrylamide stock solution

	6.5% acrylamide		10% acrylamide	
<b>% denaturant</b>	0 %	100 %	0 %	100 %
<b>TAE(20 x)</b>	5 ml	5 ml	5 ml	5 ml
<b>acrylamide</b> (50% stock)	13 ml	13 ml	20 ml	20 ml
<b>formamide</b>	0	40 ml	0	40 ml
<b>urea (7M)</b>	0	42 g	0	42 g
<b>deionised H<sub>2</sub>O</b>	up to 100 ml	up to 100 ml	up to 100 ml	up to 100 ml

**Table 2.1 Stock solutions for preparation of 6.5% and 10% acrylamide DGG.**

final denaturant concentration	stock solution	approximate temperature	
	0%	100%	
10%	0.90ml	8.10 ml	
15	1.35	7.65	
20	1.80	7.20	
25	2.25	6.75	
30	2.70	6.30	
35	3.15	5.85	
40	3.60	5.40	
45	4.05	4.95	
50	4.50	4.50	71
55	4.95	4.05	
60	5.40	3.60	74
65	5.85	3.15	
70	6.30	2.70	77
75	6.75	2.25	
80	7.20	1.80	80
85	7.65	1.35	
90	8.10	0.90	83

**Table 2.2 Volumes of stock solutions** required to obtain the appropriate gradient of chemical denaturant in the DGG and temperatures as a function of the denaturant concentration. The formula used to give an indication of the appropriate gradient corresponding to the  $T_m$  was empirically established  $T = \% + 185$

3.3

or the 10% acrylamide stock solution (Table 2.2).

The clean and dry gradient-maker was connected to the glass plates. The connecting tube between the two compartments was closed. The stirrer under the gradient maker was switched on. The polymerisation reagents, 4.5  $\mu$ l of TEMED and 50  $\mu$ l of 25% APS were added and mixed to each 9 ml solution. The 9ml high and low denaturant solutions were then placed in the inner and outer compartments respectively. The connecting tube was then opened and the pump switched on. The gel was left to polymerise for approximately 30 mins. The unpolymerised surface liquid layer was poured off and about 5 ml of the 0% denaturant stock solution containing suitable amount of TEMED and APS was poured in and the comb positioned. Once this had set, the comb was taken out and the gel was submerged in its holder in the electrophoresis tank.

#### **2.2.5.2 Electrophoresis conditions**

About 1 $\mu$ l or 25 $\mu$ l crude PCR products were loaded for subsequent silver or ethidium bromide staining respectively and electrophoresis was carried out at 61 $^{\circ}$ c, at about 35-40V, 70mA for 18 hrs, details are given in table 2.4.

#### **2.2.5.3 Visualisation of DGGE products**

##### **2.2.5.3.1 Ethidium bromide staining**

Polyacrylamide gels are stained as described in section 2.2.4

##### **2.2.5.3.2 Silver staining**

**AAT DGGE Conditions**

<b>Exon</b>	<b>PCR product length</b>	<b>Denaturant</b>	<b>Electrophoretic time</b>
<b>2</b>	502bp	56-76 %	40v for 20h
<b>3</b>	325	50-70	35v for 17h
<b>5</b>	402	55-70	35v for 20h
<b>3</b>	218	60-80	55v for 16h

**Y fragments DGGE Conditions**

<b>fragment</b>	<b>PCR product length</b>	<b>Melting temperature</b>	<b>Denaturant</b>	<b>Electrophoretic time</b>
<b>Y1</b>	315bp	75°C	60-80%	40v for 20h
<b>Y2</b>	334	77°C	65-80	40v for 19h

**PKU DGGE Conditions**

<b>Exon</b>	<b>PCR product length (exon+adjacent intron)</b>	<b>Melting temperature</b>	<b>Denaturant</b>	<b>Electrophoretic time</b>
<b>3</b>	338 bp	70°C	45-52 %	40v for 18h
<b>5</b>	313	74	55-65	40v for 18h
<b>10</b>	272	71	49	40v for 18h
<b>11</b>	341	72	56	40v for 18h
<b>12</b>	251	72.5	54 -57	40v for 18h

**Table 2.3 AAT, Y and PKU DGGE conditions**

All washings and incubations are carried out in 100ml volumes. After electrophoresis the gel was washed in 10% ethanol and 0.5% acetic acid for 2 X 3 min at room temperature with gentle shaking followed by incubation for 10 min in 0.1%  $\text{AgNO}_3$  solution. The gel was then washed in distilled water and stained in freshly prepared staining solution (1.5% NaOH, 0.01%  $\text{NaBH}_4$ , 0.4ml of 37% formaldehyde) for 20 min. The gel was fixed in 0.75%  $\text{Na}_2\text{CO}_3$  solution for 10 min. A permanent record was obtained by drying the gel onto a 3MM piece of filter paper.

#### **2.2.6 Constant denaturant gel electrophoresis (CDGE).**

The gels used in CDGE contained the same chemicals as for DGGE, but with a uniform denaturant concentration through the gel. Running conditions and staining were as for DGGE.

#### **2.2.7 Single strand conformation polymorphism (SSCP)**

SSCP analysis was carried out using the Phastsystem machine (Pharmacia). After PCR amplification the mineral oil layer was removed and 1.5  $\mu\text{l}$  of the PCR mix removed and an equal volume of formamide was added. The mix was incubated at 90°C for 10 mins and quenched on ice. The Phastsystem was set up to prerun, according to the manufacturers guidelines (400V, 10mA, 2.0W, 15°C, 10Vh) for the given acrylamide concentration of the commercially prepared phastsystem gels. When ready, approximately 2 $\mu\text{l}$  of each sample were loaded on the gel automatically (400V, 5mA, 2.0W, 2Vh). Samples were run for a period of 40 mins (400V, 10mA, 2.0W, 175Vh) at temperatures of 5°, 10° or 15°C. Some experiments at 5° were allowed to run for 80 Mins.

After electrophoresis the SSCP gels were stained automatically following the protocol shown in Table 2.4.

### **2.2.8 Cloning of PCR products**

PCR products were cloned using the TA cloning kit (Invitrogen). This system utilises the presence of 3' A overhangs on the PCR products after amplification. Ligations were carried out overnight at 12°C with approximately 1:3 molar ratio of vector : PCR product using 50 ng of pCR<sup>TM</sup> vector (Fig 2.1) in the 1X ligation buffer provided and 1 unit T4 DNA ligase in a final volume of 10µl.

Transformation of competent cells, Oneshot<sup>TM</sup> (Invitrogen), with ligated DNA was carried out as instructed by the manufacturer. Essentially, 1µl of the ligation mix and 2µl of 0.5M β-mercaptoethanol was added to 50µl of competent cells, incubated on ice for 30 min, heat shocked at 42°C for 30 sec and reincubated on ice for 2 min. 0.5 ml of medium was added and the cells incubated with shaking for 1 hr at 37°C in a rotary shaking incubator, prior to plating, to allow pre-expression of the antibiotic resistance.

Cells were plated onto kanamycin LB agar plates (at 50 µg/ml) and incubated overnight at 37°C. Recombinants were selected, blue/white selection, for inactivation of the X-gal gene by growing on X-gal (25µl of 40mg/ml X-gal stock spread on a 90mm petrie dish). White colonies were picked, grown overnight at 37°C and DNA extracted using the Magic Mini-prep Purification system (promega), for details see section 2.2.1.2. Insert DNA was liberated by digestion with EcoR1 which restricts the vector DNA flanking the insert DNA. Alternatively recombinants were identified by PCR of

Step number	Solution	Time (mins)	Temp (°C)	Remarks
1	20% TCA	5	20	Fixing
2	50% EtOH, 10% HAc	2	50	Wash
3	10% EtOH, 5% HAc	2	50	Wash
4	10% EtOH, 5% HAc	4	50	Wash
5	5% gluteraldehyde	6	50	Protein Sensitisation
6	10% EtOH, 5% HAc	3	50	Wash
7	10% EtOH, 5% HAc	5	50	Wash
8	Reagent grade water	2	50	Wash
9	Reagent grade water	2	50	Wash
10	0.4% silver nitrate	10	40	Stain
11	Reagent grade water	0.5	30	Wash
12	Reagent grade water	0.5	30	Wash
13	Developer	1	30	Develop
14	Developer	5	30	Develop
15	Background reducer	2	30	clear
16	10.0% glycerol	5	50	Stop

EtOH= Ethanol; TCA= Trichloroacetic acid; HAc=acetic acid.

**Table 2.4 SSCP silver staining protocol for the Phastsystem gels.**

**Comments for pCR<sup>TM</sup>II**  
3932 base pairs

Lac Z gene: bases 1-571  
 Sp6 promoter: bases 239-255  
 Multiple Cloning Site: bases 269-381  
 T7 promoter: bases 388-407  
 F1 origin: bases 572-986  
 Kanamycin resistance: bases 987-2114  
 Ampicillin resistance: bases 2133-2992  
 ColE1 origin: bases 3182-3765

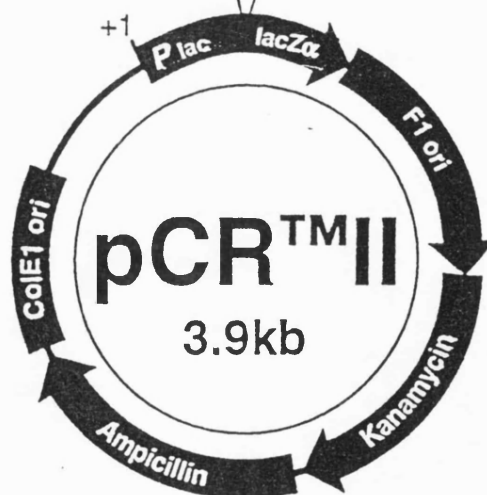
M13 Reverse Primer	So6 Promoter
CAG GAA ACA GCT ATG AC	AT TTA GGT GAC ACT ATA
GTC CTT TGT CGA TAC TG	TA AAT CCA CTG TGA TAT

NsiI	HindIII	KpnI	SacI	BamHI	SpeI
TAC TCA AGC TAT GCA TCA AGC TTG GTA CCG AGC TCG GAT CCA CTA GTA ACG GCC					
ATG AGT TCG ATA CGT AGT TCG AAC CAT GGC TCG AGC CTA GGT GAT CAT TGC CGG					

BstXI	EcoRI	PCR Product	EcoRI	EcoRV
GCC AGT GTG CTG GAA TTC GGC TT	AA	AA	GCC GAA TTC TGC AGA TAT	
CCG TCA CAC GAC CTT AAG CCG	TT	TT	CGG CTT AAG ACG TCT ATA	

Aval	PaeR7I
BstXI	NotI
CCA TCA CAC TGG CCG CCG CTC GAG CAT GCA TCT AGA GGG CCC AAT TCG	CCC TAT
GGT AGT GTG ACC GCC GGC GAG CTC GTA CGT AGA TCT CCC GGG TTA AGC	GGG ATA

T7 Promoter	M13 (-20) Forward Primer	M13 (-40) Forward Primer
AGT GAG TCG TAT TA	CTG GCC GTC GTT TTA C	T CGT GAC TGG GAA AAC
TCA CTC AGC ATA AT	GAC CCG CAG CAA AAT G	A GCA CTG ACC CTT TTG



The sequence detailed above represents the pCR<sup>TM</sup>II vector sequence with a PCR<sup>®</sup> product inserted by TA Cloning<sup>™</sup>. Note that the pCR<sup>TM</sup>II vector sequence listed in the following pages is modified at the unique *EcoR*I site during preparation for TA Cloning<sup>™</sup> so that the inserted PCR product is flanked on each side by *EcoR*I sites, as shown above.

V2.8-131027sa

**Fig 2.1** Diagrammatical representation of the pCR<sup>TM</sup>II vector used to clone PCR reactions with *EcoR*I sites flanking the cloning site.

individual bacterial cells using the appropriate primer pair, followed by growth and DNA extraction as described above.

### **2.2.9 Dideoxy sequencing**

Double stranded plasmid DNA (approximately 1 $\mu$ g) was alkali denatured in sodium hydroxide at room temperature for 10 min (40 $\mu$ l final volume). The DNA was precipitated in 2 X vol of 100 % absolute alcohol and 1/20 vol of 3M sodium acetate at -70°C for 5 min and recovered by centrifugation at 14 000 X g for 15 min.

The resultant pellet was washed in 70% alcohol, dried and resuspended in 10  $\mu$ l of water. 2 $\mu$ l of annealing buffer and 2 $\mu$ l primer (SP6 or T7) were added to the template, a final volume of 14 $\mu$ l. The tube was vortexed and spun briefly, incubated at 65°C for 5 min and at 37°C for 10min.

### **Sequencing reaction**

Reactions were carried out according to manufacturers recommendations (Pharmacia). Prior to loading, the samples were denatured by heating to 95°C for 5 min followed by quenching on ice.

### **Sequencing gels**

Sequencing was carried out using BIORAD apparatus. Samples were run at 1800V/60 watts for approximately 2-5 hrs depending on the size of the product being sequenced and the

position of the mutation within the fragment. After electrophoresis the gel apparatus was dismantled and the plates prised apart leaving the gel attached to the front plate. The gel was then submerged in fixer solution (10% acetic acid, 10% methanol) for 20 mins. The gel was transferred to a piece of 3MM whatman paper, covered with cling film and dried under vacuum for 2 hrs at 80°C. Once dry the cling-film was peeled off and the gel was autoradiographed overnight, without intensifying screens, at room temperature. The sequences were read using a light box.

## **Chapter 3 RESULTS**

### **3.1 Mutation analysis of the human AAT gene by Denaturing Gradient Gel Electrophoresis.**

#### **3.1.1 Summary**

The six major human AAT alleles (M1Ala, M1Val, M2, M3, S and Z), present in individual samples previously classified (Table 3.1.1) using isoelectric focusing, were used as a control system to evaluate the general applicability of DGGE as a mutation detection system. All of the mutants types could be distinguished by DGGE analysis of a wide range of phenotypes. As the approach used in this instance was entirely empirical, perpendicular DGG were run as a preliminary step in the analysis in order to optimise gradient and electrophoretic conditions for parallel DGGE.

#### **3.1.2 Sample preparation**

DNA, extracted from AAT individuals whose mutations had previously been characterised by isoelectric focusing (Whitehouse *et al.*, 1989), was kindly made available by Dr Susan Povey and Dr. David Whitehouse. The phenotypes studied included M1Val/Ala, Z, M1S, SZ, M1M3, M2.

#### **3.1.3 Generation of exon specific PCR products**

The polymerase chain reaction was used to selectively amplify exon 2, 3 and 5 of the AAT gene from all sample DNA. Primers were designed from intron sequences flanking each exon or within exon sequences. An artificially high melting domain, in

Allele	Frequency	Exon	Codon	Base change
M2	0.15	2	101	CGT-CAT
M1Val	0.45	3	213	GTG-GCG
M1Ala	0.25	3	213	GTG-GCG
S	0.03	3	265	GAA-GTA
Z	0.02	5	342	GAG-AAG
M3	0.09	5	376	GAA-GAC

**Table 3.1.1 The six common variant alleles of human alpha-1-antitrypsin studied in this project.**

the form of a GC clamp, was introduced at the 5' end of one primer of each of the primer pairs. Two sets of primer pairs were generated for each exon in order to investigate the effect of a GC clamp at either the 5' or 3' end of the target sequence. Amplification conditions for each exon were optimised to give a specific product of the appropriate size, corresponding to the region under investigation (section 2.2.3.1).

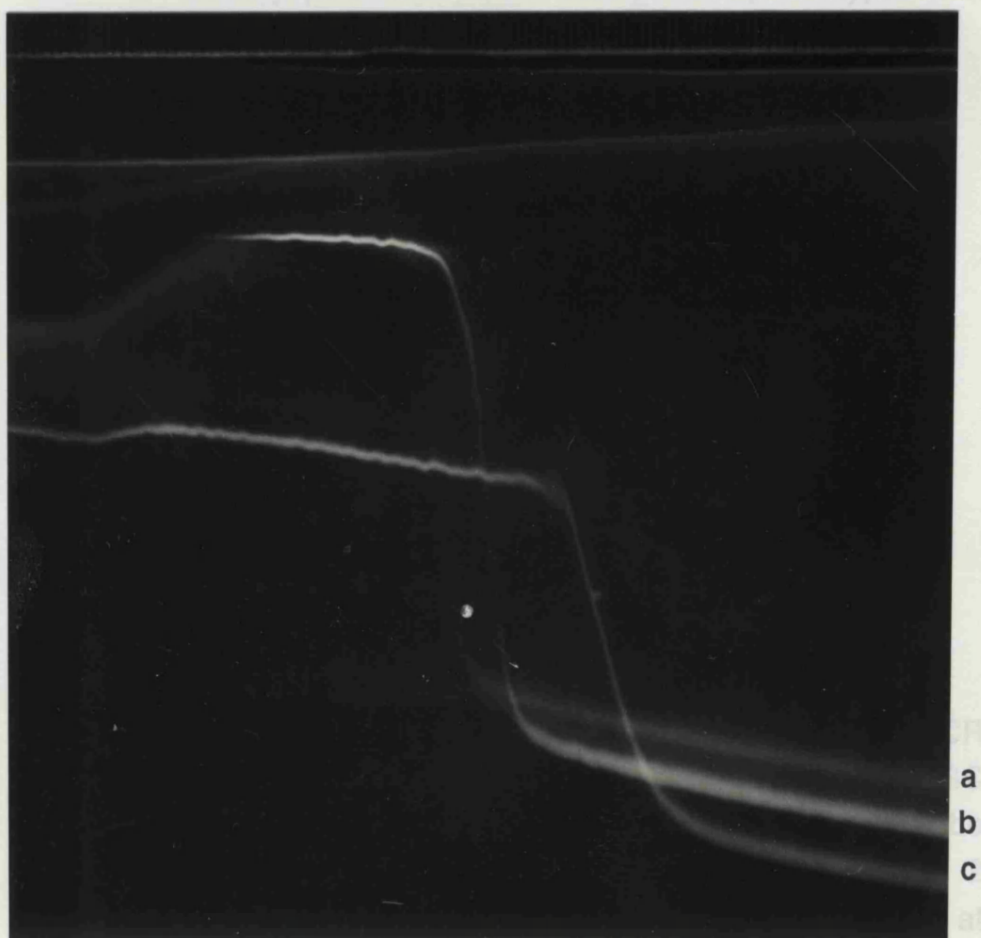
### 3.1.4 Screening of Individual Exons for mutation

To optimise the denaturing and electrophoretic conditions for mutation detection in multiple samples, perpendicular gel electrophoresis (in which the amplified DNA is layered across the top of the gel and subjected to electrophoresis at right angles to the direction of the denaturant gradient), was carried out for exons 2, 3 and 5. In general, the perpendicular gels were run with a 20-100% denaturant gradient in a 6.5% acrylamide gel at 61°C, 100 V for 4 hours.

#### Exon 3

Fig 3.1.1 shows a perpendicular gel separation carried out on exon 3 DNA fragments. These were generated using a combination of primer pairs designed to give PCR products with a GC clamp at both the 5' and 3' ends (a), a 3' GC clamp (b) and no GC clamp (c). The presence of the GC clamps increased the concentration at which melting occurred from 50 to 60% and greatly sharpened the melting transition. Therefore a 50 to 70% denaturant gradient was considered to be appropriate for the analysis of the exon 3 mutation by parallel DGGE. In the figure shown, the denaturing gradient runs on the horizontal axis from right 20% to left 100%. The vertical axis is the direction of electrophoresis.

Subsequently, 5 samples of individuals heterozygous or homozygous for exon 3 mutations at codons 213 and 205 were examined by parallel DGGE. The protein phenotypes and



**Fig 3.1.1** An ethidium bromide stained perpendicular DGGE of exon 3 from the AAT gene

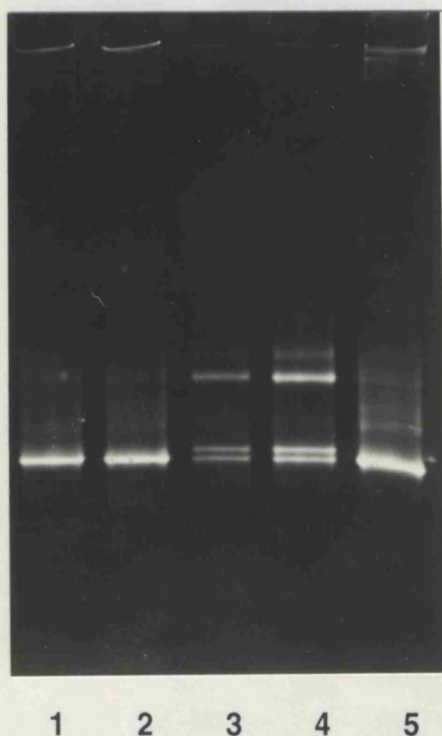
DGGE of a mixture of exon 3 DNA fragments amplified from genomic DNA of an AAT homozygote M1Val using the primers generating GC clamps at both ends (a), at the 3' end (b) and at neither end of the amplified fragment (c). The samples were run perpendicular to a 20-100% denaturant gradient in a 6.5% acrylamide gel at 61°C.

Subsequently, 5 samples of individuals heterozygous or homozygous for exon 3 mutations at codons 213 and/or 265 were examined by parallel DGGE. The protein phenotypes and DNA genotypes at both these positions are shown below.

Sample nº	1	2	3	4	5
<b>Protein phenotype</b>	M1	Z	M1S	SZ	M1M3
<b>codon 213</b>	M1Val M1Ala	M1Ala M1Ala	M1Val M1Val	M1Val M1Ala	M1Val M1Val
<b>codon 265</b>	M M	M M	M S	M S	M M

Fig 3.1.2a shows a DGGE gel of exon 3 for samples 1 to 5. PCR amplification was used to generate a product with a 5' GC clamp. The M1Val/Ala polymorphism at codon 213 could not be resolved under these conditions, samples 1 (M1Val/M1Ala), 2 (M1Ala/M1Ala) and 5 (M1Val/M1Val) which are polymorphic at this position give the same band mobility. The S variant at codon 265 present in sample 3 and 4 whose genotypes are M/S could be resolved from individuals carrying the homozygous common allele MM. These samples gave three resolvable bands corresponding to the two homo-duplex forms and a single band presumably corresponding to the two heteroduplex forms.

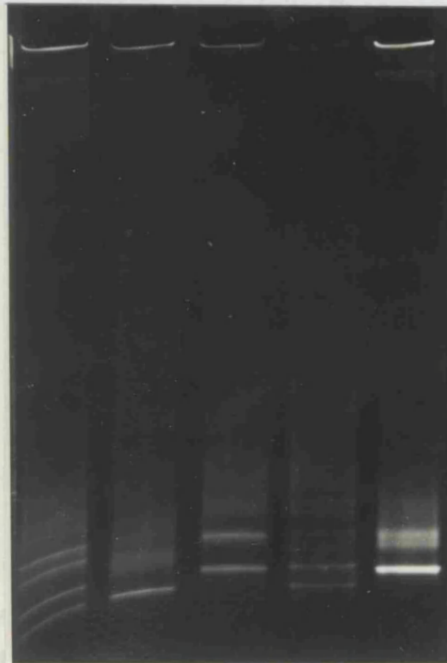
Fig 3.1.2b shows a DGG of exon 3 for samples 1 to 5. The M1Val/Ala (codon 213) alleles were resolved by PCR amplification using a primer pair which generated a PCR product with a 3' GC clamp.



**Fig 3.1.2a** An ethidium bromide stained parallel DGG of exon 3 from the AAT gene resolving the M and S alleles at codon position 265 using a 50-70% denaturing gradient

An ethidium bromide stained DGG of individuals 1-5 having genotypes MM, MM, MS, MS and MM respectively at codon position 265. PCR amplification was carried out so as to generate a product with a GC clamp at the 5' end. The band pattern is described in the text.

Individuals 1 and 4, both heterozygotes for the M1 polymorphism gave four resolvable bands, corresponding to the two homoduplexes and heteroduplexes forms. Individual 2 homozygous for M1Ala and individual 3 and 5 homozygous for M1Val generate a single band of fast and slow mobility respectively. The S mutation present in samples 3 and 4 could not be clearly resolved under these conditions. Some additional bands were observed in lanes which may correspond to the S mutation which has not



### Exon 5

DGGE analysis of exon 5 for patients samples 1 to 5 having the 2 alleles: 2 and M3 at codon position 213 respectively.

Fig 3.1.3a shows the DGG of exon 5 for patients samples 1 to 5 having the genotypes M1Val/Ala, M1Ala/Ala, M1Val/Val, M1Val/Ala and M1Val/Val respectively. The bands are resolved from each other under these conditions. Individuals 1 and 3 give the same single resolvable band. In contrast, individuals 2 and 4 and individual 5, heterozygous and homozygous respectively

**Fig 3.1.2b An ethidium bromide stained parallel DGG of exon 3 from the AAT gene resolving the M1Val and M1Ala alleles at codon position 213 using a 50-70% denaturing gradient**

An ethidium bromide stained DGG of individuals 1-5 having genotypes M1Val/Ala, M1Ala/Ala, M1Val/Val, M1Val/Ala and M1Val/Val respectively at codon position 213. PCR amplification was carried out so as to generate a product with a GC clamp at the 3' end. The band pattern is described in the text.

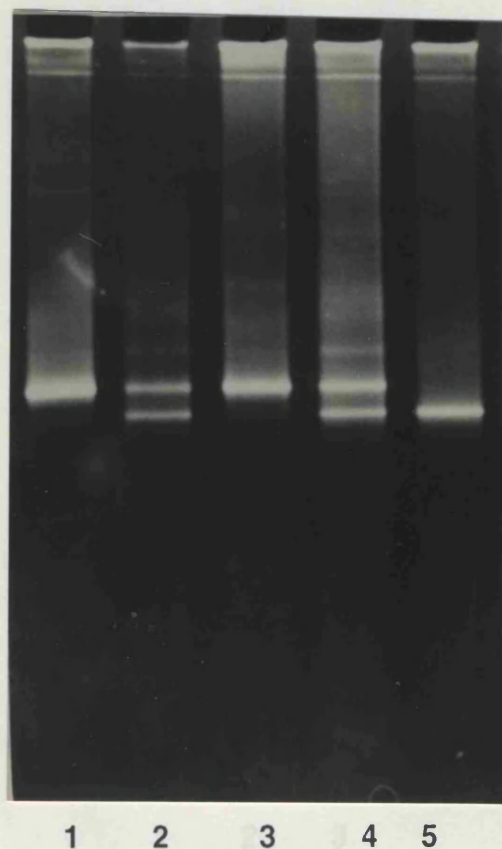
Individuals 1 and 4, both heterozygotes for the M1 polymorphism gave four resolvable bands, corresponding to the two homoduplexes and heteroduplexes forms. Individual 2, homozygous for M1Ala and individuals 3 and 5 homozygous for M1Val generate a single DGGE band of fast and slow mobility respectively. The S mutation present in samples 3 and 4 could not be clearly resolved under these conditions. Some additional bands are present in these lanes which may correspond to the melting domain containing the S mutation which has not been completely resolved.

## Exon 5

DGGE analysis of exon 5 succeeded in resolving the 2 alleles: Z and M3 at codon positions 342 and 376 respectively.

Fig 3.1.3a shows a DGGE gel of exon 5 with samples 1 to 5 having the genotypes M1M1, M1M3, ZZ, M3Z and M3M3 respectively. PCR amplification generated a product with a 5' GC clamp. The Z and M1 alleles could not be resolved from each other under these conditions. Individuals 1 and 3 give the same single resolvable band. In contrast, individuals 2 and 4 and individual 5, heterozygous and homozygous respectively for the M3 allele, display a faster resolvable band characteristic of M3 which distinguishes it from M1 and Z.

Fig 3.1.3b shows a DGGE gel of exon 5 for patients samples 1 to 3. The M1 and Z alleles were resolved from each other by DGGE using PCR amplification products from exon 5 which have a 3' GC clamp. Samples 1 and 3 homozygous for the M1 and Z alleles respectively both give a single resolvable band which can be distinguished from each other. Patient 2 who is



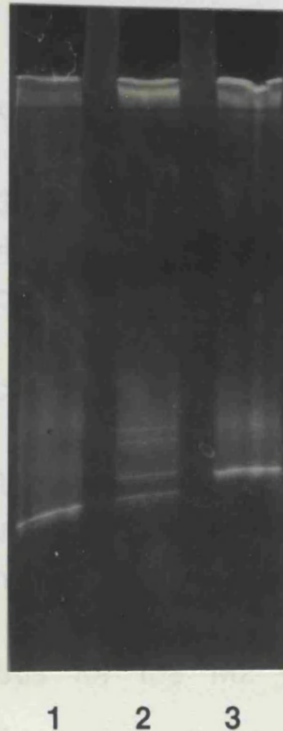
**Fig 3.1.3a** An ethidium bromide stained parallel DGG of exon 5 from the AAT gene resolving the M3 allele from the M1 allele at codon position 376 using a 55-70% denaturing gradient

An ethidium bromide stained DGG of individuals 1-5 having genotypes M1M1, M1M3, ZZ, M3Z and M3M3 respectively. Mutations at codon position 376 give rise to the M3 allele. The M1 and Z alleles result from sequence changes at other codon positions. PCR amplification was carried out so as to generate a product with a GC clamp at the 5' end. The band pattern is described in the text.

heterozygous for both these alleles generated four resolvable bands, two homoduplex forms corresponding to the bands generated by individuals 1 and 3 and two heteroduplex bands.

### Exon 2

DGGE analysis of exon 2 of the AAT gene was carried out using primers M2 and M (the mutation site is at codon position 101). The presence of a 5' or 3' GC clamp on the 2 PCR product provided a means of separating the two alleles. The different M/M2 phenotypes were resolved by reducing the overall GC content of the PCR amplification products prior to DGGE analysis. The exon 2 PCR product was digested with the restriction enzyme *SfiI* prior to DGGE analysis. This enzyme cleaves the exon into two fragments, one containing a single GC clamp as a result of the cleavage. The other fragment contained a GC clamp as a result of the cleavage. The band pattern for samples 1 to 3 is shown in Fig 3.1.4. Samples 1 and 2 represent individuals homozygous for the M allele, an individual heterozygous for the M and M2 alleles and an individual heterozygous for M2.



**Fig 3.1.3b** An ethidium bromide stained parallel DGG of exon 5 from the AAT gene resolving the Z allele at codon position 342 using a 60-80% denaturing gradient

An ethidium bromide stained DGG of individuals 1-3 having genotypes M1M1, M1Z and ZZ respectively. Mutations at codon position 342 give rise to the Z allele. The M1 allele results from sequence changes at a different codon position. PCR amplification was carried out so as to generate a product with a GC clamp at the 3' end. The band pattern is described in the text.

heterozygous for both these alleles generates four resolvable bands, two homoduplex forms corresponding to the bands generated by individuals 1 and 3 and two heteroduplex bands.

## Exon 2

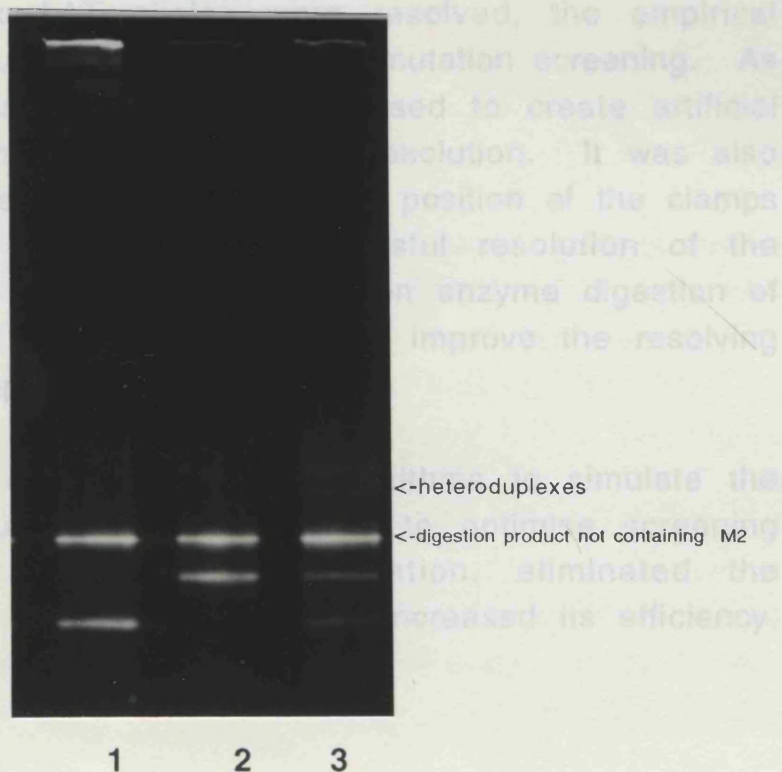
DGGE analysis of exon 2 succeeded in resolving the two alleles: M2 and M (the more common allele) at codon position 101. The presence of a 5' or 3' GC clamp at either end of the entire exon 2 PCR product provided poor resolution between individuals of different M/M2 phenotypes. The polymorphism was resolved by reducing the overall fragment size by digestion of the exon 2 amplification product with the restriction enzyme *StyI* prior to DGGE analysis. This enzyme digests the exon into two separate fragments both of which retain a single GC clamp as each primer pair used in its amplification contained a GC clamp. Fig 3.1.4 shows a DGG of exon 2 for samples 1 to 3 which represent an individual homozygous for the M allele, an individual homozygous for the M2 allele and an individual heterozygous for M2.

All three samples exhibit a constant band representing the digestion product not containing the M2 polymorphic site. Sample 1 in addition gives a lower band which represents the M allele. The M2 homozygous and heterozygous individuals, samples 2 and 3, give a different melting product representing the M2 allele. Individual 2 does not generate the lower band. This is consistent with the absence of the normal allele and its genotyping by isoelectric focusing. In contrast individual 3 heterozygous for M2 gives both homoduplex products in addition to two slower heteroduplex bands (arrowed).

### 3.1.5 Conclusion

Although all six mutations were resolved, the empirical nature of this study was a limitation. In similar studies, high melting denaturation is used to create artificial heteroduplexes. It was also interesting to observe the position of the clamps had a marked effect on the resolution of the heteroduplexes. The product provided a useful resolution of the heteroduplexes. An enzyme digestion of the product provided a useful resolution of the heteroduplexes. Improve the resolving capacity of the gel.

The availability of the cloning behaviour protocols, prior to the empirical nature of the study. These were used to generate a product with a GC clamp at both ends. The product was digested with the restriction enzyme *StyI* prior to DGGE analysis. The band pattern is described in the text.



**Fig 3.1.4** An ethidium bromide stained parallel DGG of exon 2 from the AAT gene resolving the M2 allele from the M allele at codon position 101 using a 56-76% denaturing gradient

An ethidium bromide stained DGG of individuals 1-3 having genotypes, MM, M2M2 and MM2 at codon position 101. PCR amplification was carried out so as to generate a product with a GC clamp at both ends. The product was digested with the restriction enzyme *StyI* prior to DGGE analysis. The band pattern is described in the text.

### 3.1.5 Conclusion

Although all six AAT alleles were resolved, the empirical nature of this study prevented optimal mutation screening. As in similar studies, GC clamps were used to create artificial high melting domains to allow band resolution. It was also interesting to observe that the 5' or 3' position of the clamps had a marked effect on the successful resolution of the mutations. The application of restriction enzyme digestion of product provided another alternative to improve the resolving capacity of the approach.

The availability of two computer algorithms to simulate the melting behaviour of a fragment and to optimise screening protocols, prior to any experimentation, eliminated the empirical nature of the approach and increased its efficiency. These were used in subsequent studies.

### **3.2 Screening for mutation in a 470 bp region of the Y chromosome, termed pY-80, by Denaturing Gradient Gel Electrophoresis. Utilisation of the Lerman algorithms to define optimal screening conditions.**

#### **3.2.1 Summary**

A 470 bp segment of the Y chromosome was screened using DGGE to look for possible variation in a sample of 20 unrelated male individuals. This DNA sequence was used as a test system to set up the Lerman algorithms, MELT87 and SQHTX, for optimisation of screening conditions in the DGGE technique. This process is described in detail. The 470 bp segment was subdivided into two separate parts, region Y1 and Y2 for analysis. An artificial "positive" control for the presence of mutation was generated through the introduction of a single base pair substitution in one amplification primer of each pair. No mutations were observed in this segment of DNA in the samples analysed.

#### **3.2.2 DGGE analysis of pY-80**

The sequence of a 470 bp segment of the Y chromosome, designated pY-80 (section 1.1.2.2), was analysed using the Lerman algorithms (section 1.1.1.3) to optimise conditions for mutation detection. These studies suggested that the entire pY-80 sequence could not be examined by DGGE in a single experiment. The sequence was consequently divided into 2 overlapping regions, Y1 (position 1-270 bp) and Y2 (positions 201-470) for screening purposes and each region was analysed independently. Fig 3.2.1 shows the complete sequence of this region and the primer pairs used for the PCR amplifications

**SEQUENCE pY-80**

**15'AGGCCTGTCAACCCCCAGCCCTGGG**CTGCTTCCCTGGCCTCTT  
 CTCTGTTCCCTCTCTGAGGGCCTAACTCCCTTGGGTAGTGCTGCAGAA  
 ATAGAGCCACAGGCCCTGGCTGATGATCTGGTGGACTGGGCAAATTGG  
 TCGTGACAGGTCAGGTTCTGGTTCAAAGCCAATTCCTCCGATGCCAAG  
 GAATGTCGAAGA**AGGTCCTTTGCCATGATGCCCCATAG**CTGCCC  
 ACCTCAGCAATC**GTGCCGTAACCTGGGCCCCTCACAGT**CAGACACC  
 AGCTGAAGAAGCTCAGGCAGTGACCTGCGGGAAACTCGGGCTTTACCC  
 TGCATGACCCTAGAACCACTGGACTGCAGTGGAGCCAGTCGCCCTGTA  
 TCCTGGAGGGAGACGAGTCAGGAAGGCGCACGCAGGCCAGCTCCCGA  
 GGTACTACCCCCTCTACTCCTCAGGGAGGA**TGCCAACGCAATACTC**  
**C** -3'<sup>470</sup>

**Fig 3.2.1 Sequence of pY-80 taken from Tsukahara *et al.*, 1990.**

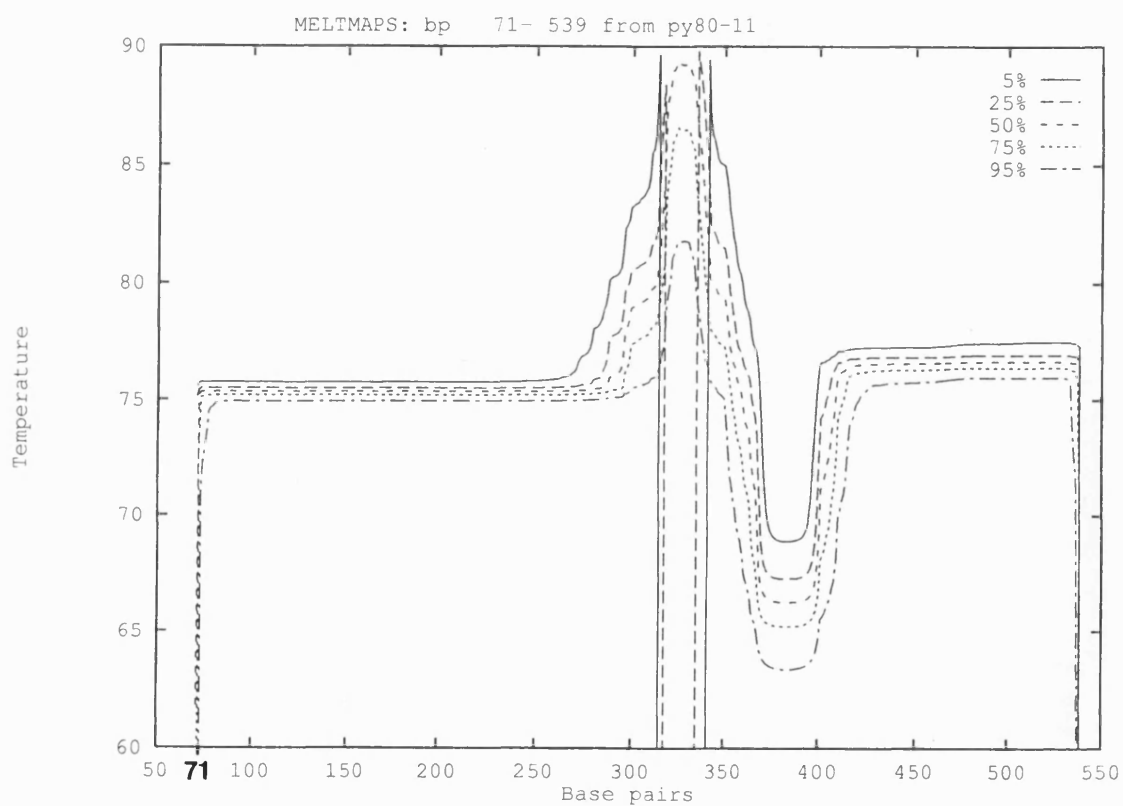
Primer sequences are underlined. The sequence was divided into two overlapping regions, Y1 and Y2. The primer pair corresponding to region Y1 are shown in bold and the primer pair corresponding to region Y2 is shown in bold italics.

are underlined. Additional primers containing a single base pair substitution were also synthesised for use in PCR to introduce a positive control for mutation detection (section 2.2.3.2).

The melting profile of the entire pY-80 sequence from base position 1-470, was considered using the MELT87 algorithm. (Fig 3.2.2). This program calculates the temperature required to bring each base-pair to a melting equilibrium in relation to its complete disassociation. The 50% melting probability is taken as standard and this profile is called the melting map of a sequence. Higher and lower levels of base-pair association (from 5%, 25%, 75% and 95%) as indicated by the key in Fig 3.2.2 can also be calculated by this program. 5% represents the temperature curve giving rise to 95% disassociation (single stranded) and 5% non-disassociation (double-stranded).

In this program the first base position in the sequence is given an arbitrary starting position, which in this instance corresponds to position 71 (see arrow). The melting profile of the pY-80 sequence divides it into three parts, the first and the last corresponding to base positions 71-260 (base positions 1-189 in the sequence) and 400-540 (base positions 329-469) and the middle part corresponding to base positions 260-400 (base positions 189-329 in the sequence). The first and last parts of the sequence show good co-operativity with a melting map of approximately 75°C which can be equated to a specific percentage of denaturant (Table 2.2). However the central portion exhibits considerable variation in its melting map and this lack of co-operativity precludes the screening of the entire fragment in a single DGGE experiment.

The introduction of an artificially high melting domain in the form of a GC clamp may alter the melting map of a sequence.

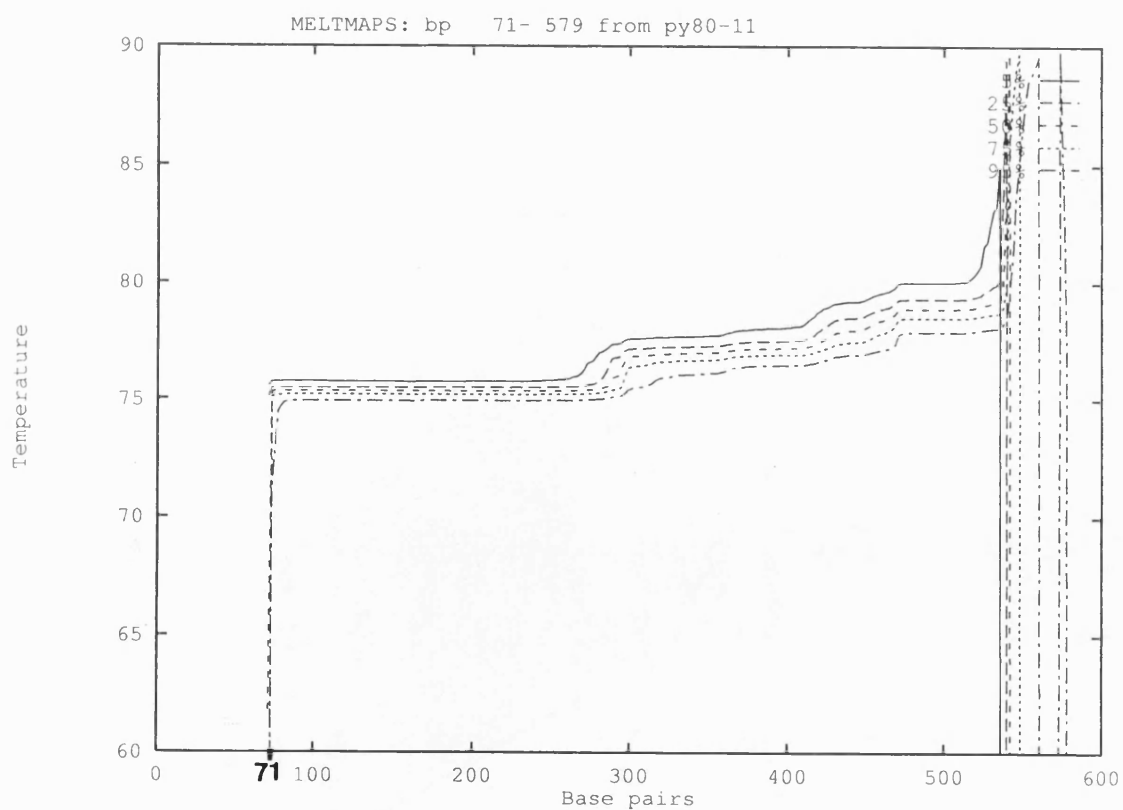
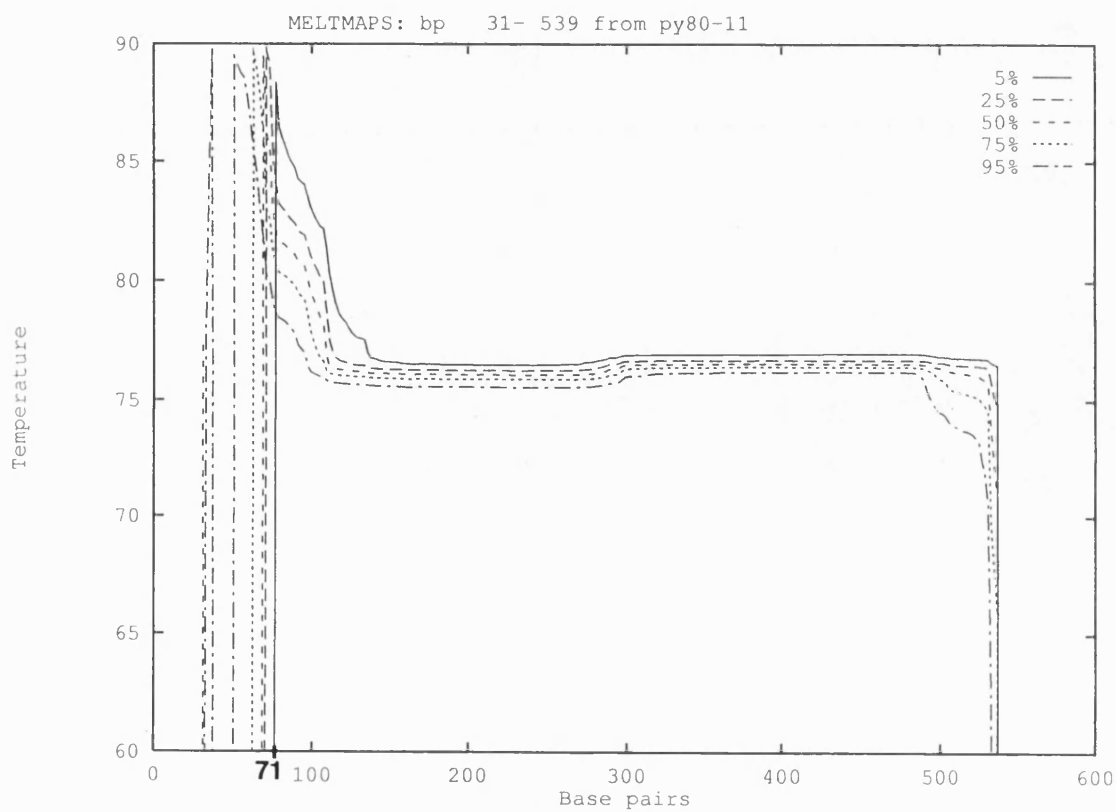


**Fig 3.2.2** Melting map of the pY-80 sequence as determined by the MELT87 program. Position 71 in the horizontal axis corresponds to nucleotide position 1 in the sequence.

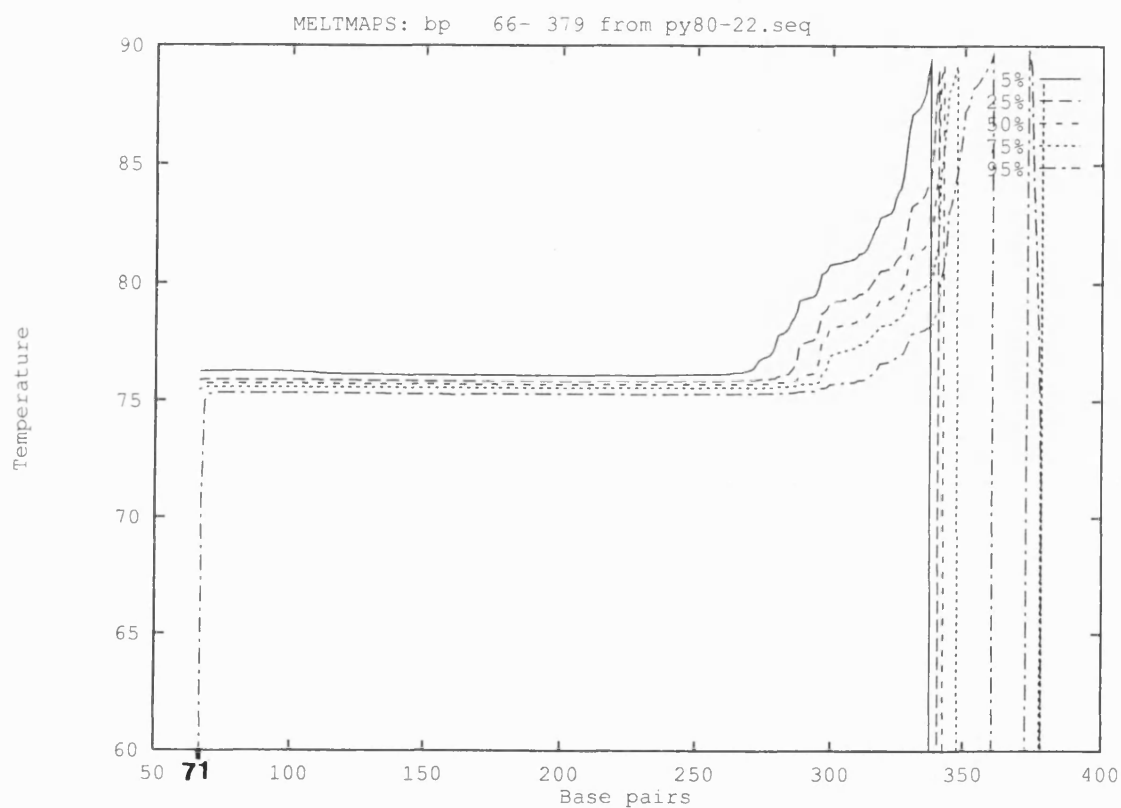
Fig 3.2.3a and 3.2.3b show the effect of introducing a GC clamp of 40 nucleotides at either end of the pY-80 sequence. The presence of a GC clamp at the 5' end transforms the melting profile and removes the irregularity of the central region of the sequence. The whole of this region with the exception of the most 5' end shows a consistent melting temperature of approximately 76°C. As shown in Fig 3.2.3a a 49 bp stretch closest to the clamp (position 71-120/corresponding to positions 1-49 of the pY-80 sequence), will not be resolved under these conditions as it has a much higher melting temperature. Similarly the addition of a GC clamp at the 3' end of the sequence reduces the variation in the melting map of the central region. As was the case however with a 5' clamp, the melting map close to the clamp, will not be resolved as it has a much greater melting temperature. Therefore, in order to completely screen the sequence for mutation different regions of the sequence had to be considered separately.

By introducing a GC clamp at successively greater distances (in general 10-20 bp) away from the 5' end of the sequence it was possible to determine the maximal size of the pY-80 fragment that could be resolved into a single melting domain. Fig 3.2.4 shows the melting profile for positions 71-341 (corresponding to positions 1-270 of the pY-80 sequence). The longest fragment that could be resolved into a single melting domain is from positions 71-300 (positions 1-229 of the pY-80 sequence). The presence of a GC clamp shifts the melting temperature of almost the entire fragment to the same melting temperature (approximately 75°C).

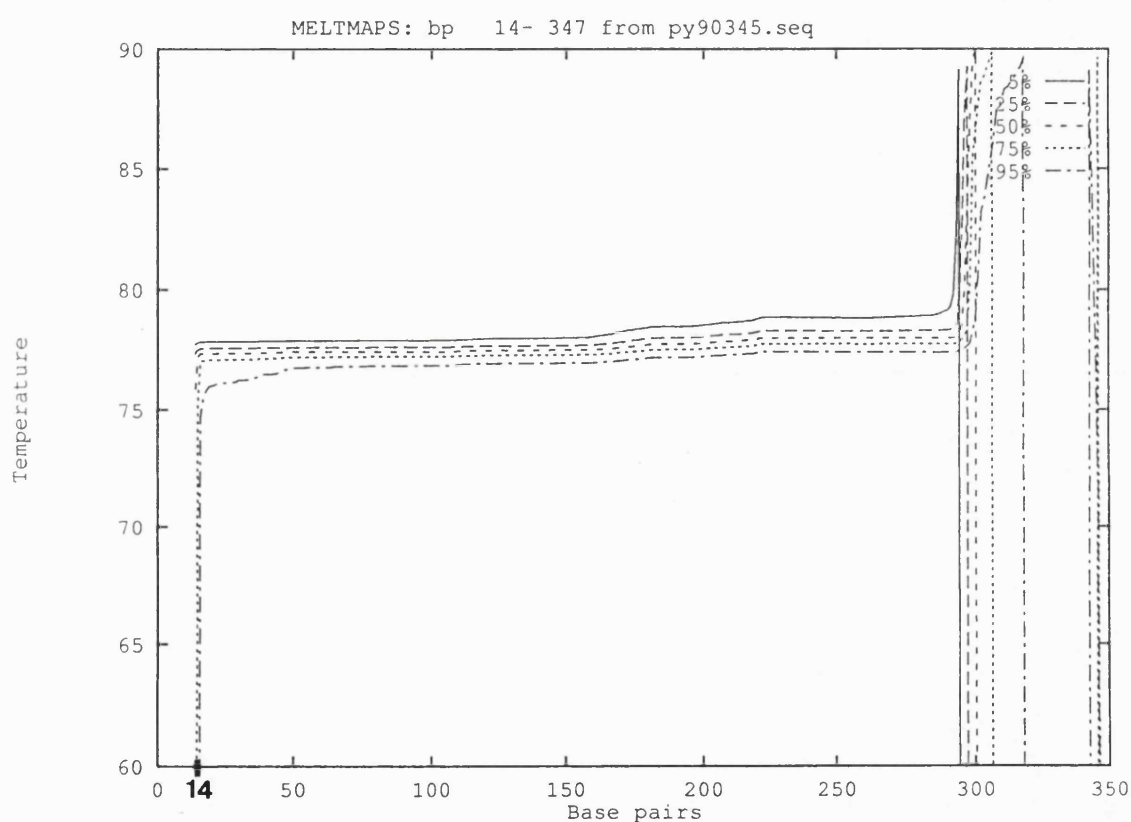
A similar calculation was performed on the 3' end terminal side from positions 201 to 470 of the pY-80 sequence (Fig 3.2.5). In this instance base pair position 201 and 470 correspond to positions 14 and 283 on the graph respectively.



**Fig 3.2.3 a and b** Melting map of the pY-80 sequence determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 5' and 3' end of the sequence respectively.



**Fig 3.2.4** Melting map of the pY-80 sequence, position 1-270 bp determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 3' end of the sequence. Position 71 in the horizontal axis corresponds to nucleotide position 1 in the sequence.



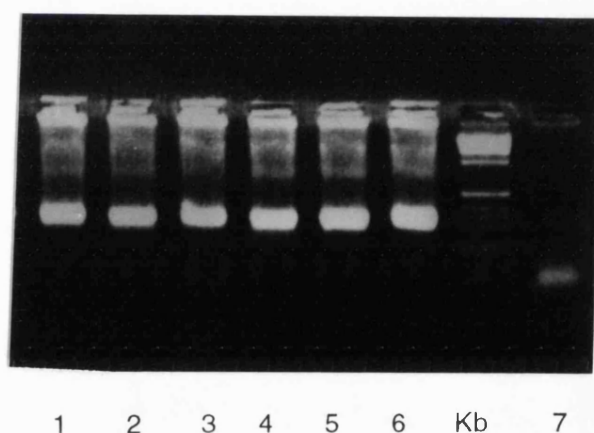
**Fig 3.2.5 Melting map of the pY-80 sequence,** position 201-470 bp determined by the MELT87 program following the introduction of a 50 bp GC clamp at the 3' end of the sequence. Position 14 in the horizontal axis corresponds to nucleotide position 201 in the sequence.

This generated a melting profile containing a single melting domain for Y2 (positions 201-470 of the pY-80 sequence) (Fig 3.2.5). In this case, the melting temperature was slightly higher (about 77°C) and it was possible to resolve the entire sequence.

The SQHTX algorithm calculates the difference in displacement for a specified melting temperature between the wild-type homoduplex fragment and a heteroduplex with a single base mismatch. The difference in displacement will reliably predict if all potential mismatches along the DNA fragment can be resolved. The temperature corresponding to the appropriate displacement value can be approximately equated to a particular denaturing gradient, and electrophoretic conditions for its analysis can be inferred to allow the detection of all single base changes at any position in the fragment. The melting temperatures of 75°C and 77°C for fragments Y1 and Y2 respectively were approximately equated to 65 and 70% of chemical denaturant as discussed in Table 2.2. Denaturing gradient ranges of 60-80% and 65-80% for the Y1 and Y2 regions respectively were chosen for the subsequent experimental analysis by DGGE.

Fig 3.2.6 shows the PCR analysis of samples using the primer pairs outlined in Fig 3.2.1 for regions Y1 and Y2. No amplification was observed in the female DNA lane consistent with the amplification of a Y specific chromosomal fragment. In order to obtain a positive control for mutation detection, a PCR reaction using an additional primer where a single base pair substitution had been introduced was used to generate a mutated Y1 and Y2 templates. The mutations introduced were a (G->A) and a (T->G) substitution for Y1 and Y2 respectively.

Fig 3.2.7 shows an ethidium bromide stained DGG of PCR



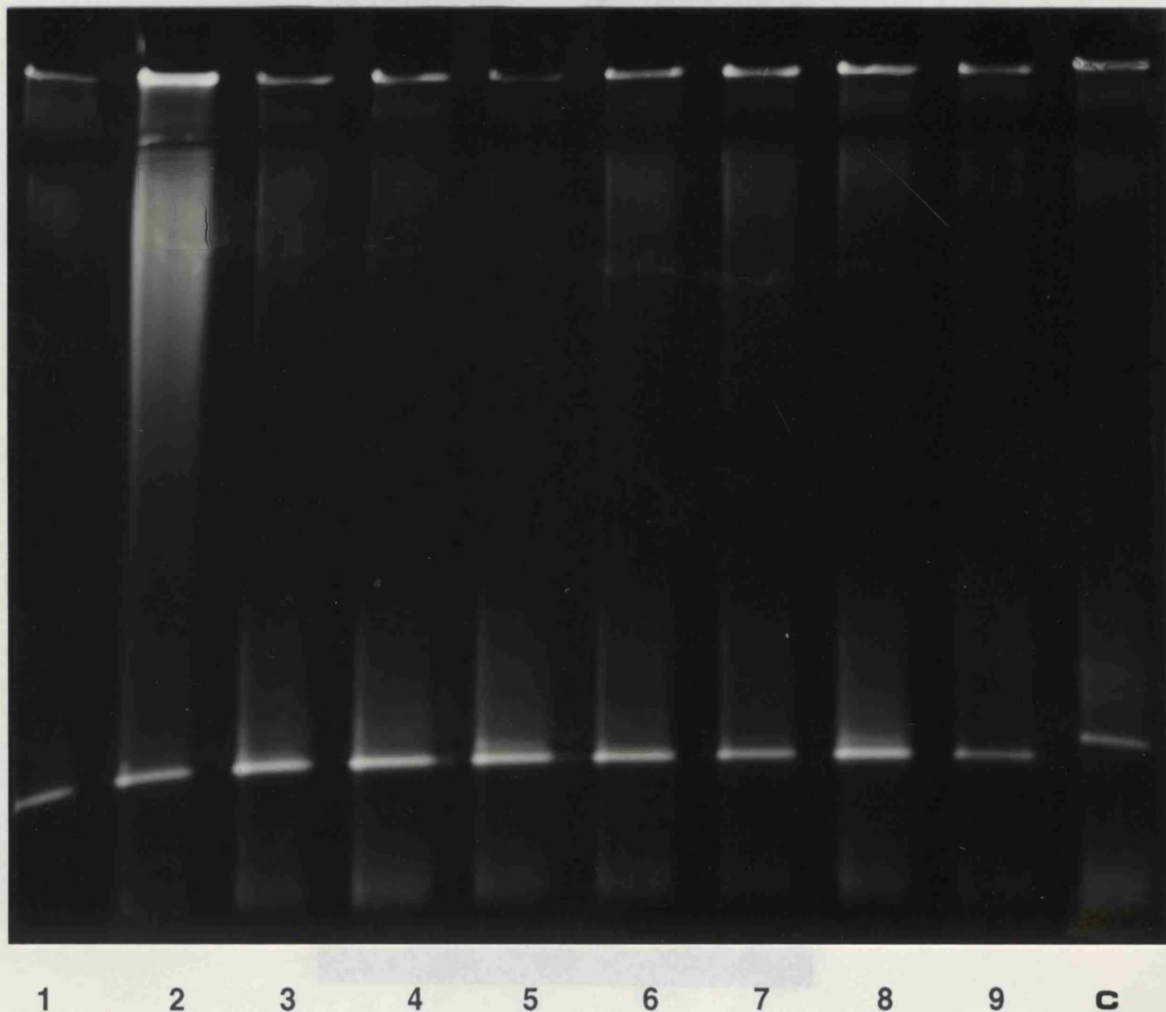
**Fig 3.2.6 PCR amplification of the Y1 and Y2 regions of pY-80 corresponding to regions positions 1-270 and 201-470 respectively**

PCR amplification products of Y1 (lanes 1-3) and Y2 (lanes 4-6) of 310 and 320 bp respectively were electrophoresed through a 2% agarose gel. No amplification was observed in the female control (lane 7). There was no evidence for non-specific amplification.

products of the Y1 region in 9 male samples (lanes 1-9) and a positive control (lane 10). There was no variation observed in these 9 male samples or in the total population of 20 males analysed. The positive control focusses at an earlier position in the gel as anticipated for a G to A substitution. Fig 3.2.8 shows an ethidium bromide stained DGG of the analysis of 2 male and 2 control samples for sequence variation within the Y2 region. As was observed in the study of Y1, no variation was found in these two individuals or in the entire male panel analysed. In this instance the positive control resolves later in the gel as would be expected for a T to G substitution

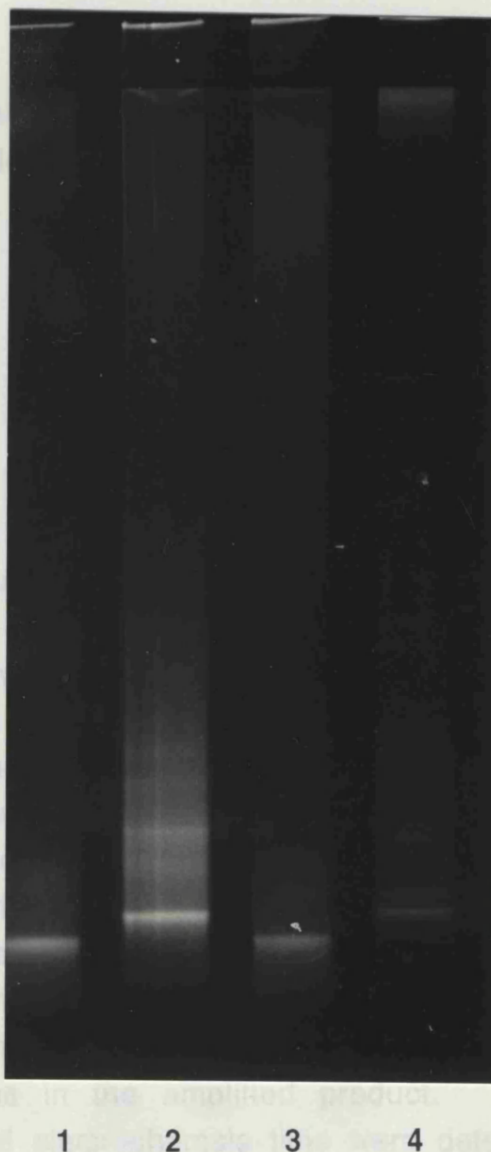
## **Conclusion**

These Y chromosome studies provided the first opportunity in this project to apply the SQHTX and MELT87 Lerman algorithms to optimise conditions for mutation detection using DGGE. The experimental conditions inferred from these algorithms allowed good separation between normal and mutated templates. Furthermore, the melting changes observed following the introduction of particular mutations were consistent with those expected from a knowledge of base pair associations.



**Fig 3.2.7** An ethidium bromide stained DGG of nine male samples and a positive control for the Y1 region, position 1-270bp, using a 60-80% denaturing gel

An ethidium bromide stained DGG showing 9 male samples (lanes 1-9) amplified using the primer pair Y1/Y1 and a single positive control (c) generated using the primer pair Y1/Y1\* (lane 10). A constant band was resolved in the nine sample lanes. A higher band was resolved in the control lane which contains an artificial mutation (G→A) incorporated into the primer sequence.



**Fig 3.2.8** An ethidium bromide stained DGG of 2 male samples and 2 positive controls for the Y2 region, position 201-470bp, using a 65-80% denaturing gel

An ethidium bromide stained DGG showing 2 male samples (lanes 2 and 4) amplified using the primer pair Y2/Y2 and two single positive control generated using the primer pair Y2/Y2\*. A constant band was resolved in the two sample lanes. A lower band was resolved in the control lanes (1 and 3) which contain an artificial mutation (T->G) incorporated into the primer sequence.

### **3.3 Mutation analysis of the PAH gene by Denaturing Gradient Gel Electrophoresis.**

#### **3.3.1 Summary**

In total 5 of the 13 exons of the PAH gene were evaluated for the presence of mutations by DGGE in a group of 50 independent affected patients, living in the South-East of England and attending the University College Hospital London PKU clinic. Each of the exons studied were considered separately. Prior to designing oligonucleotide primers for PCR amplification of genomic DNA, the melting characteristics of each exon with adjacent intron segments were evaluated theoretically using the computer algorithms MELT87 and SQHTX, kindly provided by Dr. Leonard Lerman. An artificially high melting domain, in the form of a GC clamp, was introduced at the 5' end of one of the amplification primers to allow for the identification of all potential mutations in the amplified product. The optimum gradient range and electrophoresis time were determined using the SQHTX algorithm for each fragment being investigated.

The exons chosen for analysis in this study were primarily those which had previously been shown in other North European population to exhibit a high frequency of mutation. The study of exons 3, 5, 10, 11 and 12 by DGGE from 50 independent patients identified ten distinct mutations corresponding to 42% of the mutated chromosomes in this population and a further two mutations, observed in 20% of patients, which seemingly have no obvious effect on the function of the PAH protein. As observed in previous studies certain mutations,

for example R408W and IVS12nt1, were observed at relatively high frequencies of 14 and 12% respectively. In addition three mutations, K341R, L347F and IVS10nt-3c->t represented mutations not previously described.

### **3.3.2 Sample preparation**

Blood samples from affected individuals and where possible parents and sibs were provided by Dr David Brenton from the PAH unit of University College Hospital. Patients were clinically sub-divided into mild and severe forms of PKU according to the level of blood phenylalanine. Genomic DNA was extracted from 10 ml of EDTA-anticoagulated blood (section 2.2.1.1) from 50 independent affected individuals and 34 family members. Samples were obtained over a period of time and numbered from 1-84 in accordance with their arrival date.

### **3.3.3 Generation of exon specific PCR products**

The polymerase chain reaction was used to selectively amplify exon 3, 5, 10, 11 and 12 of the PAH gene for all 50 unrelated affected individuals. Primers were designed from intron sequences flanking each exon in accordance with the computer algorithms MELT87 and SQHTX. An artificially high melting domain, in the form of a GC clamp, was introduced at the 5' end of one primer of each of the primer pairs. Amplification conditions for each exon were optimised to give a specific product of the appropriate size, corresponding to the region under investigation (section 2.2.3.3). Gel electrophoresis of the PCR amplification products for exons 3, 5, 10, 11 and 12 corresponding to product sizes of 339, 313, 272, 341 and 251

bp respectively are shown in Fig 3.3.1. There was no evidence of nonspecific amplification for any of the 5 exons studied.

### **3.3.4 Denaturing Gradient Gel Electrophoresis**

The computer algorithms MELT87 and SQHTX analyse the available sequence spanning a particular exon. Primers are chosen to optimise the resolving capacity of the sequence. These programs allowed the optimal denaturing and electrophoretic conditions to be inferred.

After DGGE electrophoresis (section 2.2.5) samples were visualised by either ethidium bromide or silver staining of the gel (section 2.2.5.3). The altered band pattern of the mutations identified in exon 3, 10, 11 and 12 are shown through Figs 3.3.5-3.3.30. No mutations were identified in any of the 50 affected individuals screened for exon 5. Control amplification products from normal individuals were included on each gel although normal banding patterns are not necessarily shown in each figure.

The following section is divided into the DGGE analysis for each exon in turn. It should be noted that in some cases constant denaturing gels (CDG) are shown in preference to DGG. In these instances DGGE analysis was initially used to determine exact denaturing conditions and then constant denaturing conditions were used in subsequent analyses. In general CDG gave better resolution of bands than DGGE. The molecular changes causing the altered mobility of PCR fragments were determined by cloning of the mutated allele and subsequent sequence analysis. In all instances the cloned fragment was re-evaluated by DGGE to ensure that no PCR mismatches had been incorporated during the cloning step.

### 3.3.5 Screening of Individual Exons for Mutation

#### 3.3.5.1 Exon 3

DGGE was carried out under the conditions described in Table 2.4 as defined from the analysis using the Lerman programs, MELT87 and SOHIX. Fig 3.3.2 shows the melt map for the exon

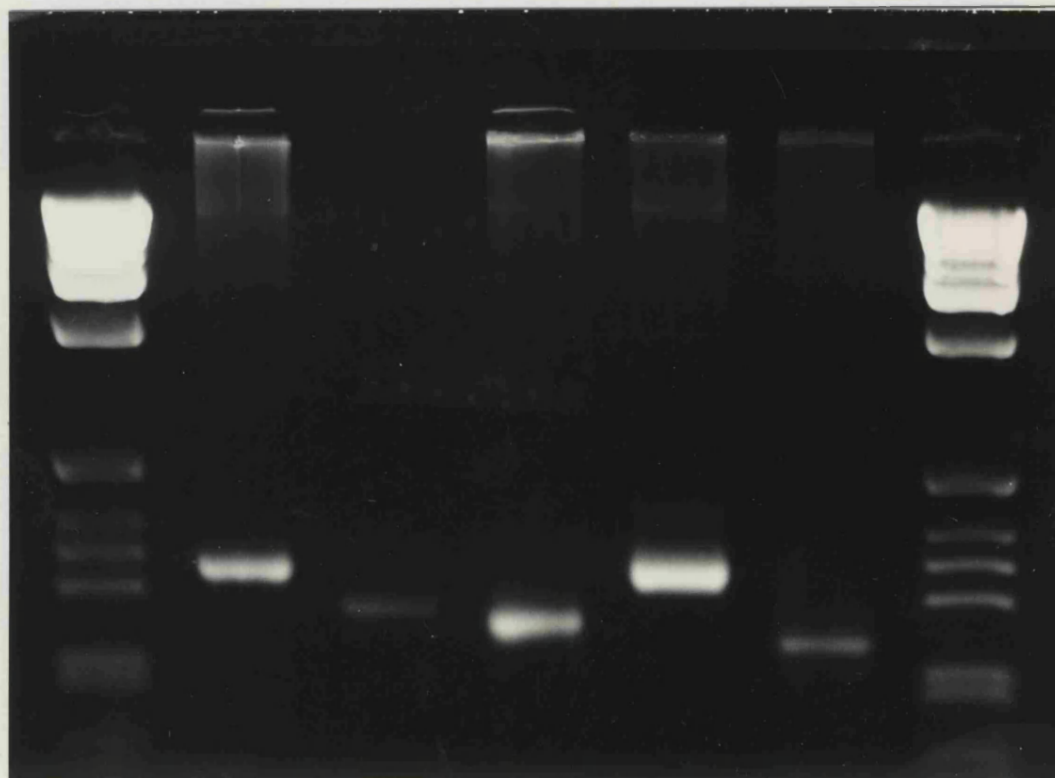


Fig 3.3.1 PCR amplification of exon/intron sequences of exon 3, 5, 10, 11 and 12 of the PAH gene.

PCR products of 339, 313, 272, 341 and 251 bp corresponding to exons 3, 5, 10, 11 and 12 were electrophoresed through a 2% agarose gel. There was no evidence for non-specific amplification.

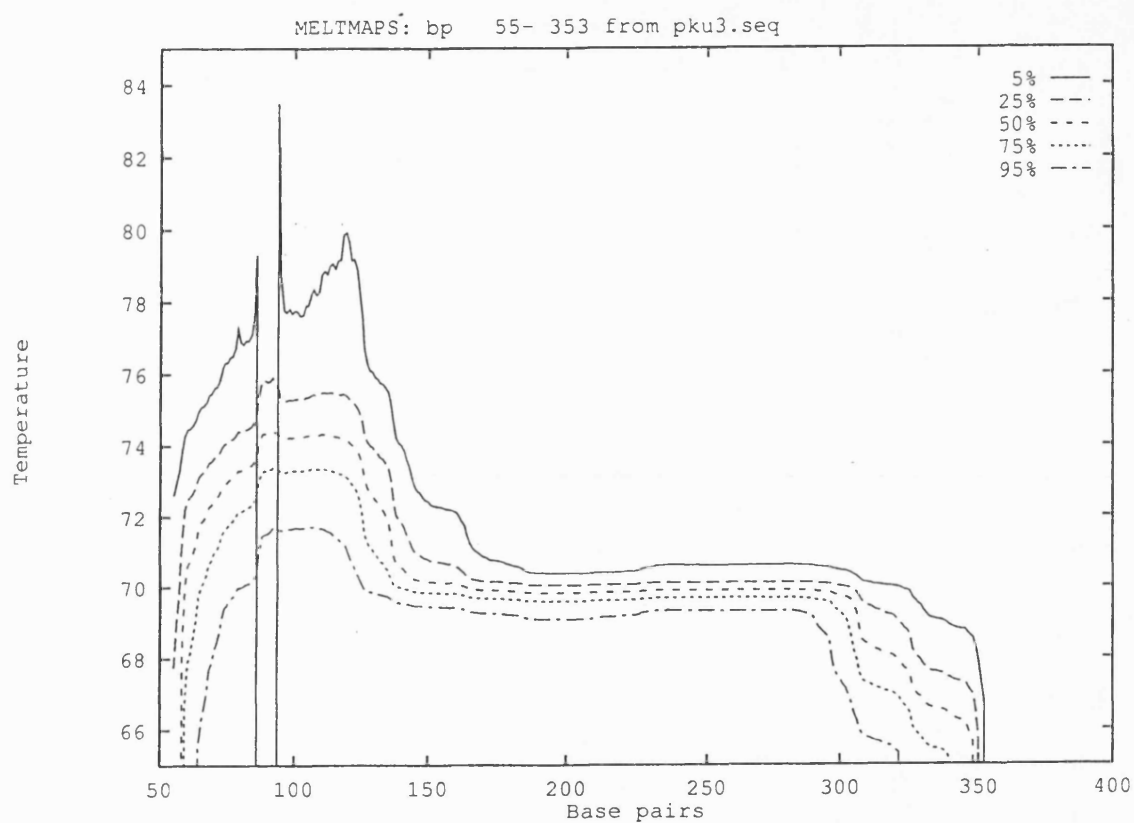
### 3.3.5 Screening of Individual Exons for mutation

#### 3.3.5.1 Exon 3

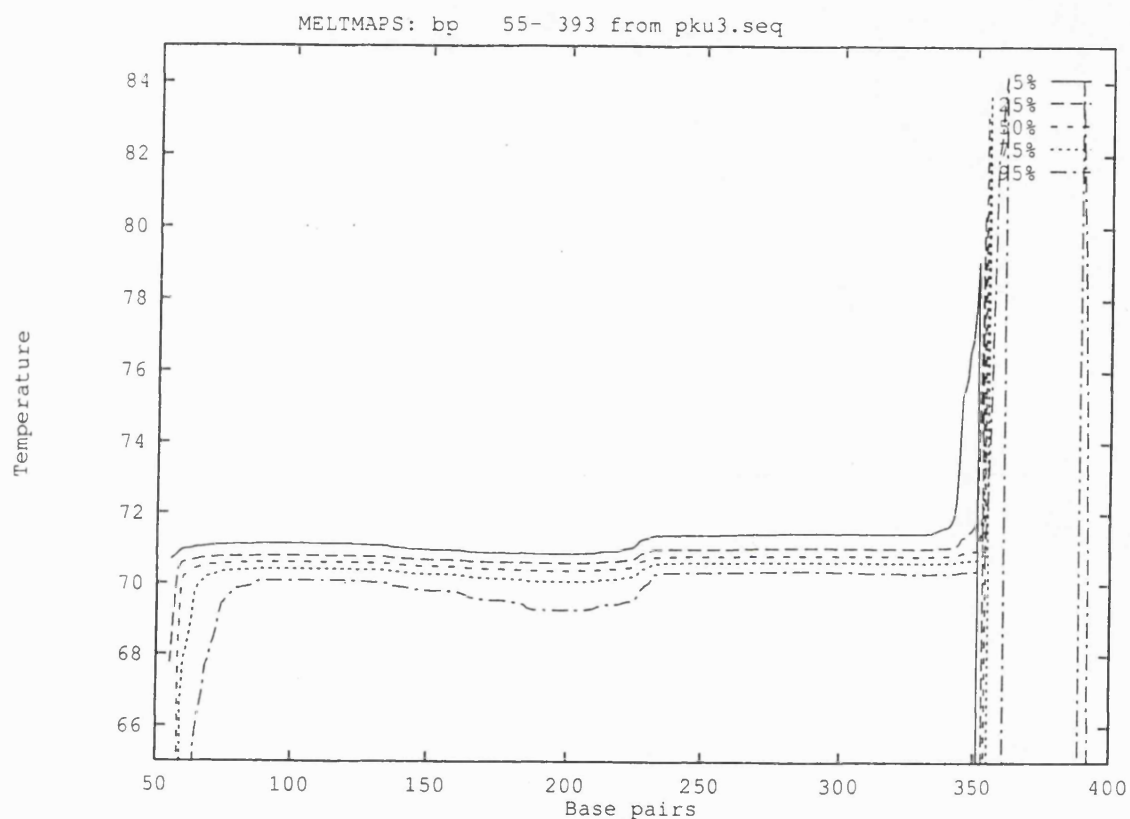
DGGE was carried out under the conditions described in Table 2.4 as defined from the analysis using the Lerman programs, MELT87 and SQHTX. Fig 3.3.2 shows the melt map for the exon 3 and flanking intronic sequences without the presence of any GC clamp. This map shows a high melting region towards the 5' end of the sequence. Mutations present in this region would not be resolved in subsequent DGGE analyses.

To facilitate the screening of the fragment for all possible mutations using a single set of conditions it is sensible to generate a single melting domain covering the entire fragment. The effects of placing a GC clamp at the 5' and 3' ends of the sequence were carefully considered. Fig 3.3.3 shows the melting map for the sequence following the introduction of a GC clamp at the 3' end of the exon 3 sequence. As the melting temperature at each base position is comparable, ie there is a melting co-operativity between each base in the sequence, any mutation present in the fragment can be resolved. This approach was used in the analysis of patients samples for exon 3. The alternative approach of placing a GC clamp at the 5' end of the sequence would have failed to resolve the sequence into a single domain. There are two separate melting domains with temperatures varying from 75°C at the 5' end, to 70°C towards the 3' end (Fig 3.3.4). The entire sequence could not therefore be screened under the same set of conditions and this was not considered as a practical approach.

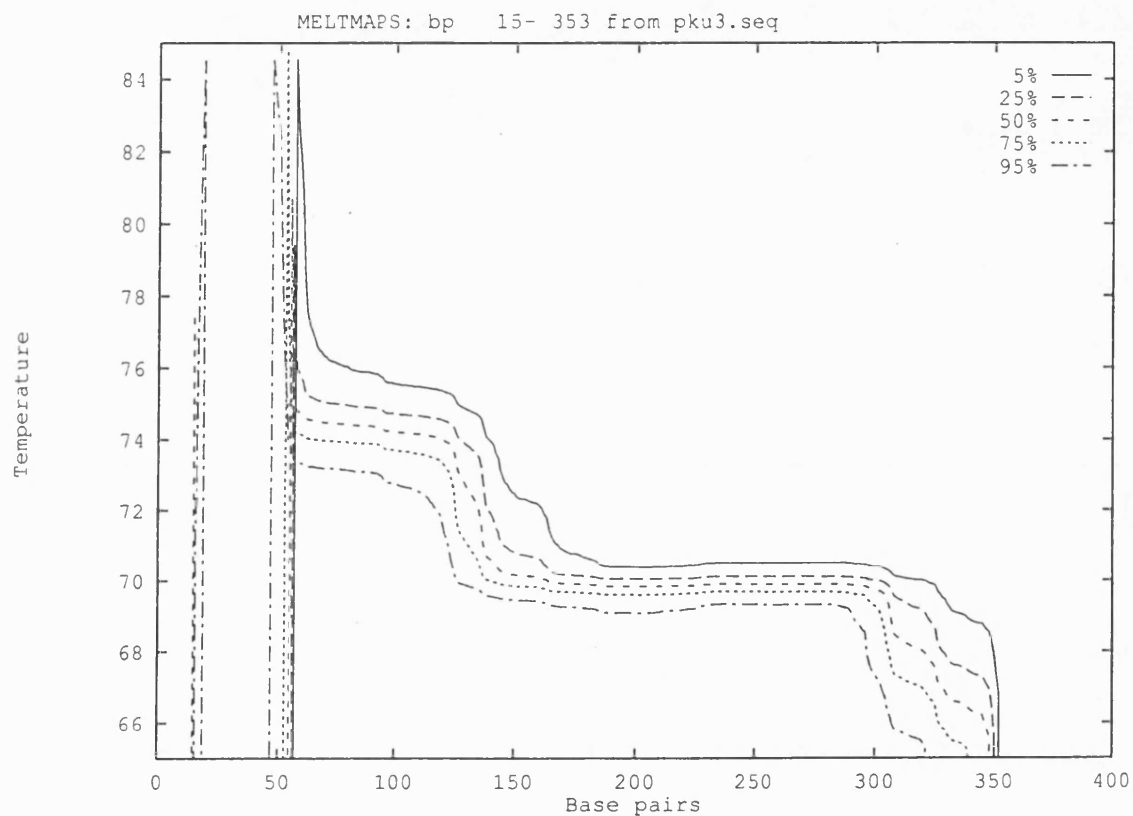
The complete screening of the 50 PKU patients identified 2 types of mutation at a frequency of 5% and 2% respectively. Fig 3.3.5 shows a silver stained gel of these two mutations:



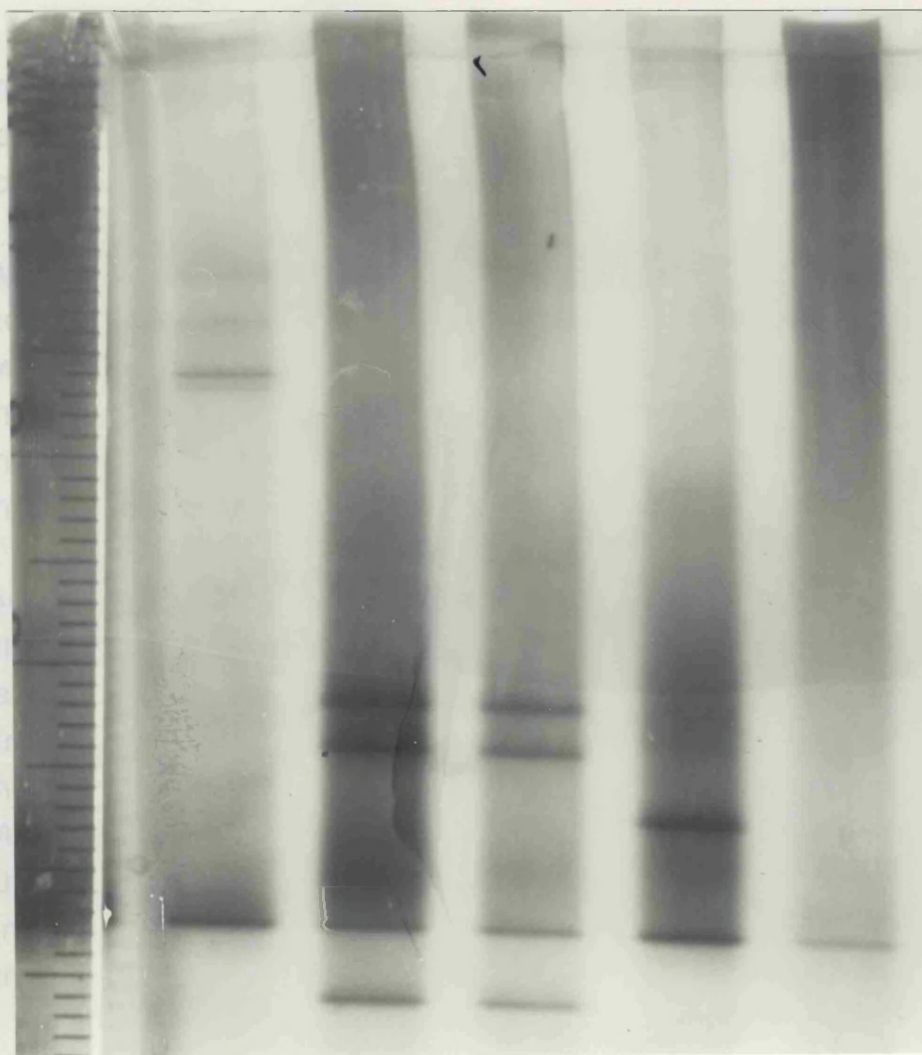
**Fig 3.3.2** Melting map of the sequence containing the exon 3 and splice site sequence of the PAH gene as determined by the MELT87 program.



**Fig 3.3.3** Melting map of the sequence containing the exon 3 and splice site sequence of the PAH gene determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 3' end of the sequence.



**Fig 3.3.4** Melting map of the sequence containing the exon 3 and splice site sequence of the PAH gene determined by the MELT87 program following the introduction of a 40 bp GC clamp at the 5' end of the sequence.



N 47 55 49 N

**Fig 3.3.5** A silver stained CDG of the two mutation types identified in exon 3 using a 52% constant denaturing gel.

A silver stained CDG showing the patient samples 47 and 55 (Mutation 1), and 49 (Mutation 2) which represent the two different banding patterns observed in the patients studied for mutation in exon 3. This pattern is described in the text. Sample N represents a normal control. The bands resolving at a lower percentage of denaturant in control lane 1 were not reproducible and presumably reflect contamination or non-specific amplification.

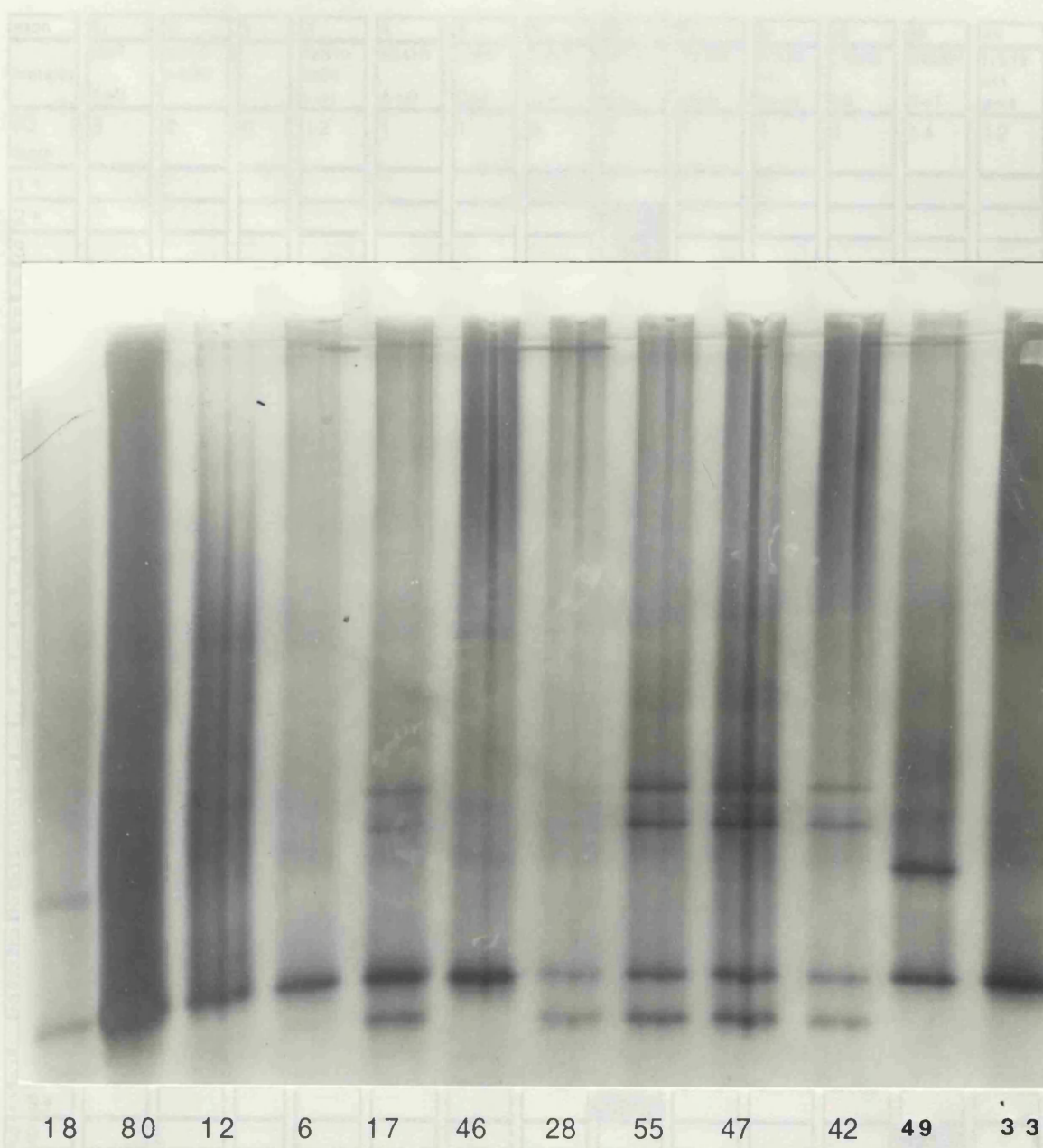
mutation 1, present in patients 47 and 55, and mutation 2, present in patient 49. The gel also includes two normal controls, N.

Mutation 1 gives four resolvable bands. One band is identical to that present in normal individuals and correspond to the wildtype allele. The two upper bands correspond to heteroduplexes formed through the association of mutated and normal alleles. The lowest band corresponds to the homoduplex of the mutated allele.

Mutation 2 appears as two resolvable bands. It appears that the two homoduplex forms from the normal and mutated alleles and the two heteroduplex forms cannot be distinguished from each other, the lower band corresponding to the homoduplex and the upper band to the heteroduplex forms. Alternatively each band may represent a homoduplex form and the heteroduplex forms failed to be resolved. The former interpretation was confirmed following re-PCR and DGGE of the lower band as the two resolvable bands were reformed.

No patient was detected as homozygous for mutations on exon 3. In total 5 patients, 17, 28, 42, 47 and 55 had mutation 1 and 2 patients, 18 and 49 had mutation 2 (Fig 3.3.6). The pattern of mutation detection for all 50 affected individuals is shown in Table 3.3.1. As samples were obtained over a period of time and numbered from 1 to 84 in accordance with their date of arrival, Table 3.3.2 provides a key to identify the corresponding family members of each patient. In general, samples were available from just one of the parents of an affected child, usually the mother.

### **Cloning and sequencing of the mutated alleles.**






**Fig 3.3.6** A silver stained CDG of the two mutation types identified in exon 3 using a 52% constant denaturing gel.

A silver stained CDG showing the patient samples, 17, 28, 55, 47 and 42 (mutation 1) and patient samples 18 and 49 (mutation 2) who exhibit mutation pattern 2. These banding patterns are described in the text. Patient samples 80, 12, 6, 46 and 33 represent normals for exon 3.

exon	III	III	V	X	X	X	XI	XI	XI	XI	XI	XII	XII
mutatio	I65T T->C	ΔATC 94/95		IVS10 nt66 c->t	K341R A->G	L347F C->T	L385L G->C	IVS10 nt-11 g->a	Y356X C->G	IVS10 nt- 3c->t	ΔK363 ΔG	R408W C->T	IVS12 nt1 g->a
FQ	5	2	0	12	1	1	8	4	1	1	1	14	12
Sam													
1 •													
2 •													
3													
4													
5 •													
6													
7 •													
8 •													
9 •													
10 •													
11													
12 •													
13 •													
14													
15 •													
16 •													
17 •													
18 •													
19 •													
20													
21													
22													
23													
24 •													
25 •													
26 •													
27													
28 •													
29 •													
30													

**Table 3.3.1** Distribution of the identified PAH mutations in 50 unrelated PKU patients (•) and close relatives.

key	
	polymorphism
	heterozygous mutation
	homozygous mutation

exon	III	III	V	X	X	X	XI	XI	XI	XI	XI	XII	XII
muta	I65T T->C	ΔATC 94/95		IVS10 nt66 c->t	K341R A->G	L347F C->T	L385L G->C	IVS10 nt-11 g->a	Y356X C->G	IVS10 nt3c- >t	ΔK363 ΔG	R408W C->T	IVS12 nt1 g->a
FQ	5	2	0	12	1	1	8	4	1	1	1	14	12
Sam													
31													
32•													
33													
34													
35•													
36•													
37•													
38													
39													
40•													
41•													
42•													
43•													
44•													
45•													
46•													
47•													
48•													
49•													
50													
51													
52•													
53													
54•													
55•													
56													
57•													
58													
59•													
60													

**Table 3.3.1** Distribution of the identified PAH mutations in 50 unrelated PKU patients (•) and close relatives. (continued)

key	
	polymorphism
	heterozygous mutation
	homozygous mutation

exon	III	III	V	X	X	X	XI	XI	XI	XI	XI	XII	XII
muta	I65T T->C	ΔATC 94/95		IVS10 nt66 c->t	K341R A->G	L347F C->T	L385L G->C	IVS10 nt-11 g->a	Y356X C->G	IVS10 nt3c- >t	ΔK363 ΔG	R408W C->T	IVS12 nt1 g->a
FQ	5	2	0	12	1	1	8	4	1	1	1	14	12
Sam													
61													
62•													
63•													
64•													
65•													
66•													
67													
68•													
69•													
70													
71•													
72													
73													
74•													
75•													
76													
77													
78													
79•													
80													
81•													
82													
83													
84													

**Table 3.3.1 Distribution of the identified PAH mutations in 50 unrelated PKU patients (•) and close relatives. (continued)**

In total 84 samples (sam) were collected from PKU individuals and family members and screened for mutation in five exons of the PAH gene. Each individual sample was given a number from 1 to 84 and the 50 affected independent patients which were considered are marked with a black dot, •, shown along the vertical axis. The relationship between affected individuals and samples corresponding to family members is shown in Table 3.3.2. The names of the twelve mutations identified and the exons where they were found (in Roman numerals) are shown along the horizontal axis. The frequency (FQ) of each mutation is shown directly below it's name. The various shadings indicate the presence or absence of a particular mutation in an individual (see key to table).

FAMILY	PATIENT	SIBLING(S)	PARENT(S)
1	1		
2	2	3	4
3	5		6
4	7	38	39
5	8		
6	9	30	31
7	10		11
8	12		
9	13	14	
10	15		
11	16		
12	17		
13	18		
14	19		20
15	24	21	22/23
16	25		
17	26		27
18	28		
19	29		
20	32		33/34
21	35		
22	36		
23	37		
24	40		
25	41		

**Table 3.3.2 Correlates the 50 independent affected individuals chosen for analysis with the other family members available for study.**

Affected individuals considered as one of the fifty independent affected patients in the study are marked (•) in Table 3.3.1 and were assigned a family (shown above) and patient number. Sibling and parent samples which correspond to the same family number are shown adjacent to each patient column.

FAMILY	PATIENT	SIBLING(S)	PARENT(S)
26	42		
27	43		
28	44		
29	45		
30	46		
31	47		
32	48	58	
33	49		50/51
34	52		53
35	54		
36	55		56
37	57		
38	59	60	61
39	62		
40	63		
41	64		
42	65		
43	66		67
44	68		
45	71	72	73
46	74		
47	75	76	77/78
48	79		
49	69	80	70
50	81	82	83/84

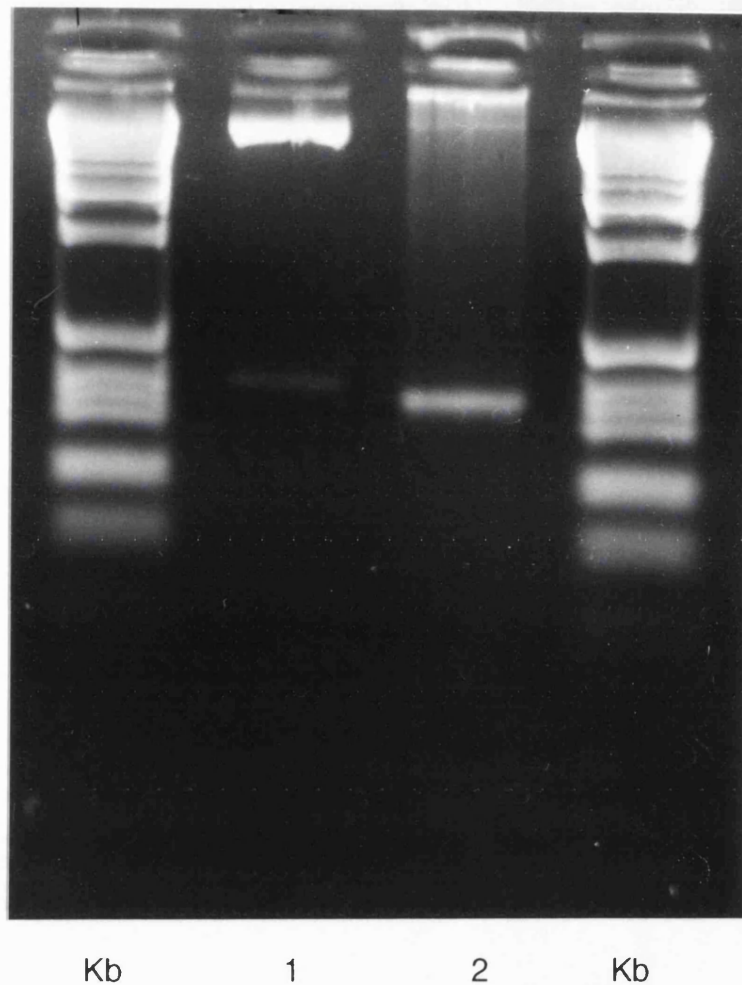
**Table 3.3.2 Correlates the 50 independent affected individuals chosen for analysis with the other family members available for study.**

Affected individuals considered as one of the fifty independent affected patients in the study are marked (•) in Table 3.3.1 and were assigned a family (shown above) and patient number. Sibling and parent samples which correspond to the same family number are shown adjacent to each patient column.

To determine the actual molecular changes underlying these 2 mutations the mutated allele was cloned and sequenced. Mutant alleles can be distinguished by their altered banding patterns after electrophoresis (Fig 3.3.5). The DNA bands corresponding to mutant alleles were excised from the DGG, eluted into a small volume of water, reamplified and cloned into the TA cloning vector, pCRT<sup>TM</sup> (section 2.2.8). Recombinant clones were identified after transformation by blue/white selection. DNA was prepared from these clones using the Magic mini preparation method. The presence of insert DNA was confirmed by digesting with the restriction enzyme EcoRI (Fig 3.3.7) which flanks the cloning site of the TA vector (Fig 2.1).

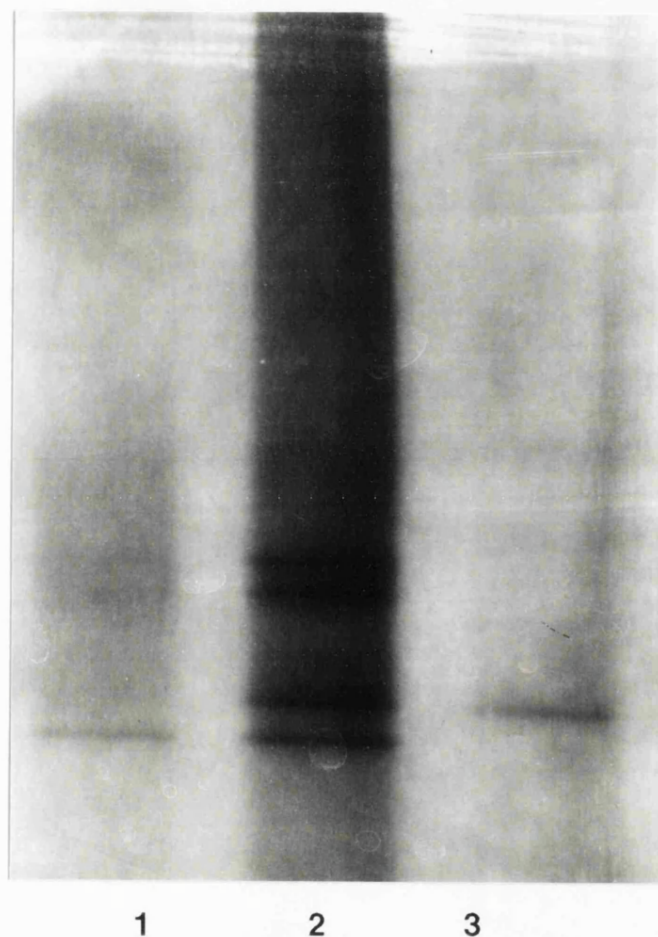
The fragment was sequenced (section 2.2.9) using the vector primers T7 and SP6 which flank the cloning site in the TA vector. In addition the cloned material was amplified and run on a denaturing gel under the same conditions as before to confirm that no additional mutations had been incorporated by PCR errors. Fig 3.3.8 gives as example a silver staining gel of a PCR amplified clone of mutation 1 present in exon 3. The band pattern of the cloned product corresponds to the mutant homoduplex band identified in the initial denaturing gel performed on this patient sample.

Fig 3.3.9 compares the sequence of mutated and normal alleles for this two exon 3 mutation which involves a T to C substitution at base position 194 and results in a Ile-Thr amino acid substitution at position 65. This mutation has previously been observed (John *et al.*,1992) and has been called I65T. A comparison of the amino-acids at position 65 in the PAH genes of the mouse and rat show that this position is conserved in mammals. In *Drosophila* the Ile at position 65 is substituted to Leu (Fig 3.3.10).



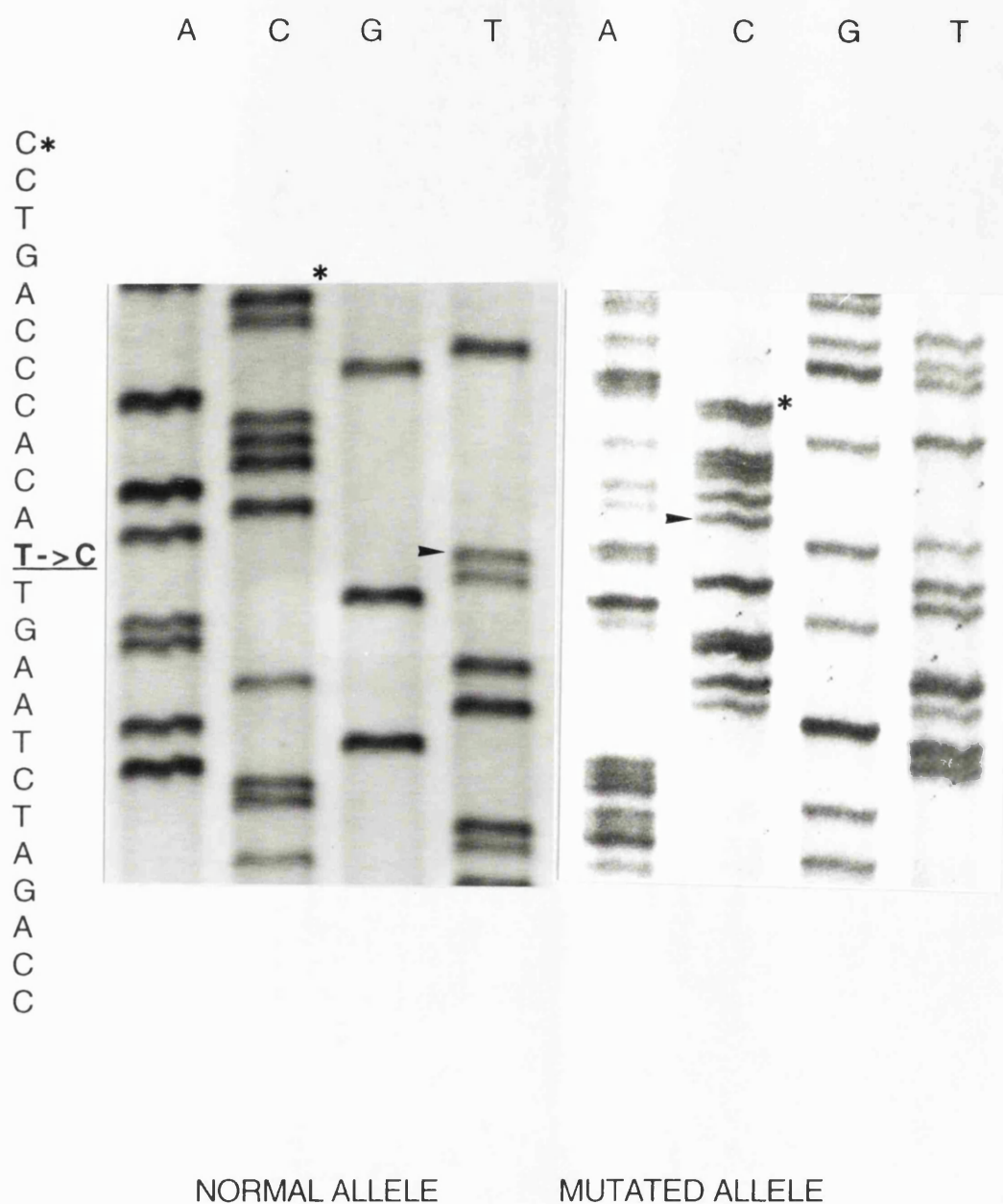
**Fig 3.3.7 EcoRI digestion of a cloned exon 3 PCR product of the PAH gene.**

One ug of plasmid DNA from a cloned exon 3 containing PCR fragment in the TA cloning vector was digested with EcoRI and electrophoresed on a 0.8% agarose gel (lane 1). An exon 3 PCR product, amplified from genomic DNA, was run as a control (lane 2). The larger size of the cloned product is due to the presence of vector sequence adjacent to the cloning site.



**Fig 3.3.8 Silver stained CDG of exon 3 mutation 1, identified using a 52% constant denaturing gel.**

CDG analysis of the PCR amplification products of the plasmid clone containing the mutated allele 1 (lane 1), a heterozygous patient for mutation 1 (lane 2) and a normal control (lane 3). The PCR of the plasmid clone shows identical pattern to that of the variant homoduplex obtained from genomic DNA.



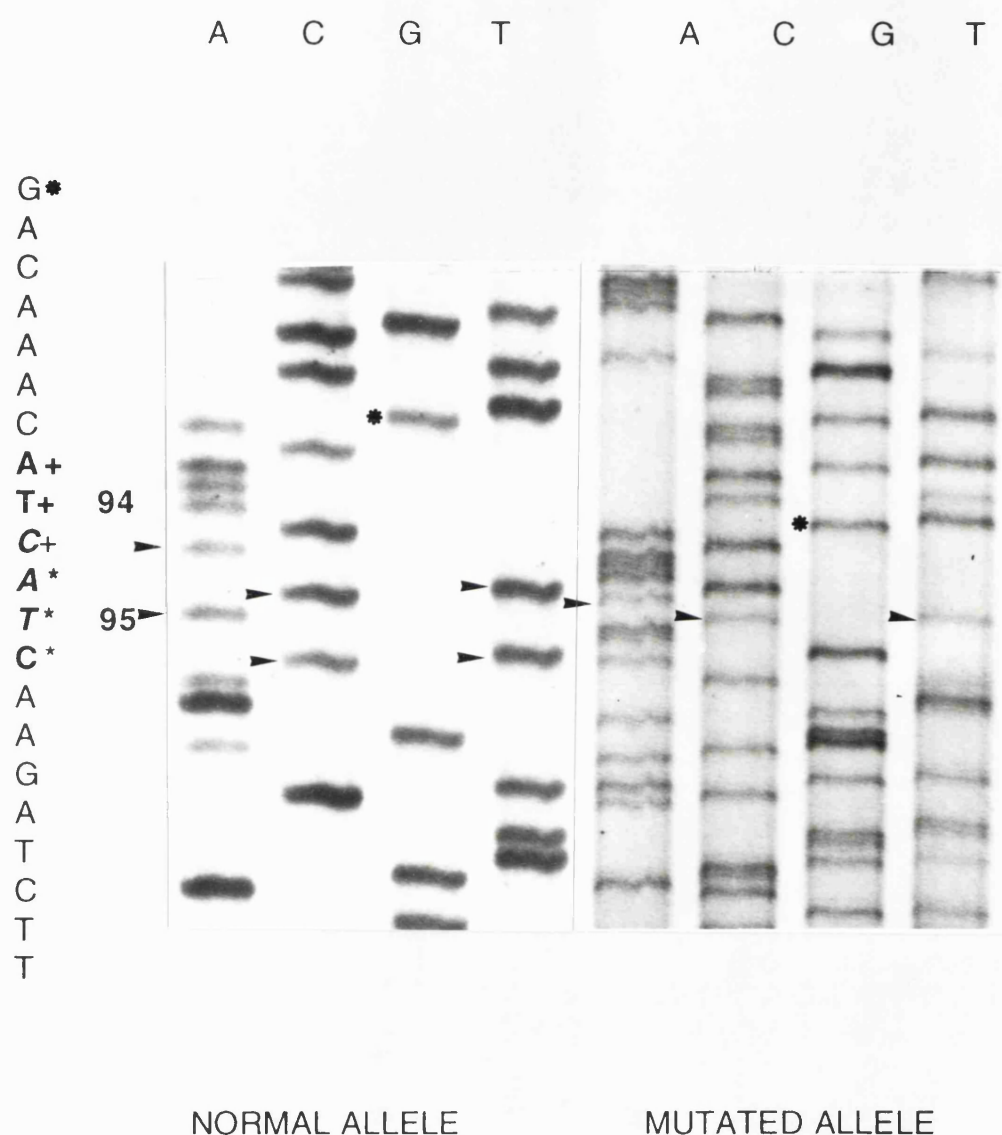
**Fig 3.3.9 Sequence comparison of mutated and normal alleles of the exon 3 region containing mutation 1.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 3 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a T to C transition which results in the substitution of the normal amino-acid Ile<sup>65</sup> with Thr<sup>65</sup>.

Species	61	62	63	64	65	66	67	68	69	70
Human	N	L	T	H	I	E	S	R	P	S
Rat	N	L	T	H	I	E	S	R	P	S
Mouse	N	L	T	H	I	E	S	R	P	S
Drosophila	N	L	C	I	L	S	R	I	L	A

**Fig 3.3.10 Comparison of the coding region containing the I65T mutation in the human, rat, mouse and drosophila PAH protein.**

The amino acid positions 61 to 70 of the PAH protein are compared between human, rat, mouse and drosophila. Amino-acids which are variant between species are shaded. The Ile residue at amino acid position 65, which is mutated in PKU, is conserved in all three mammalian species. The substitution of Ile for Leu in drosophila would be expected to be a conservative substitution.



**Fig 3.3.11 Sequence comparison of the mutated and normal alleles of the exon 3 region containing mutation 2.**

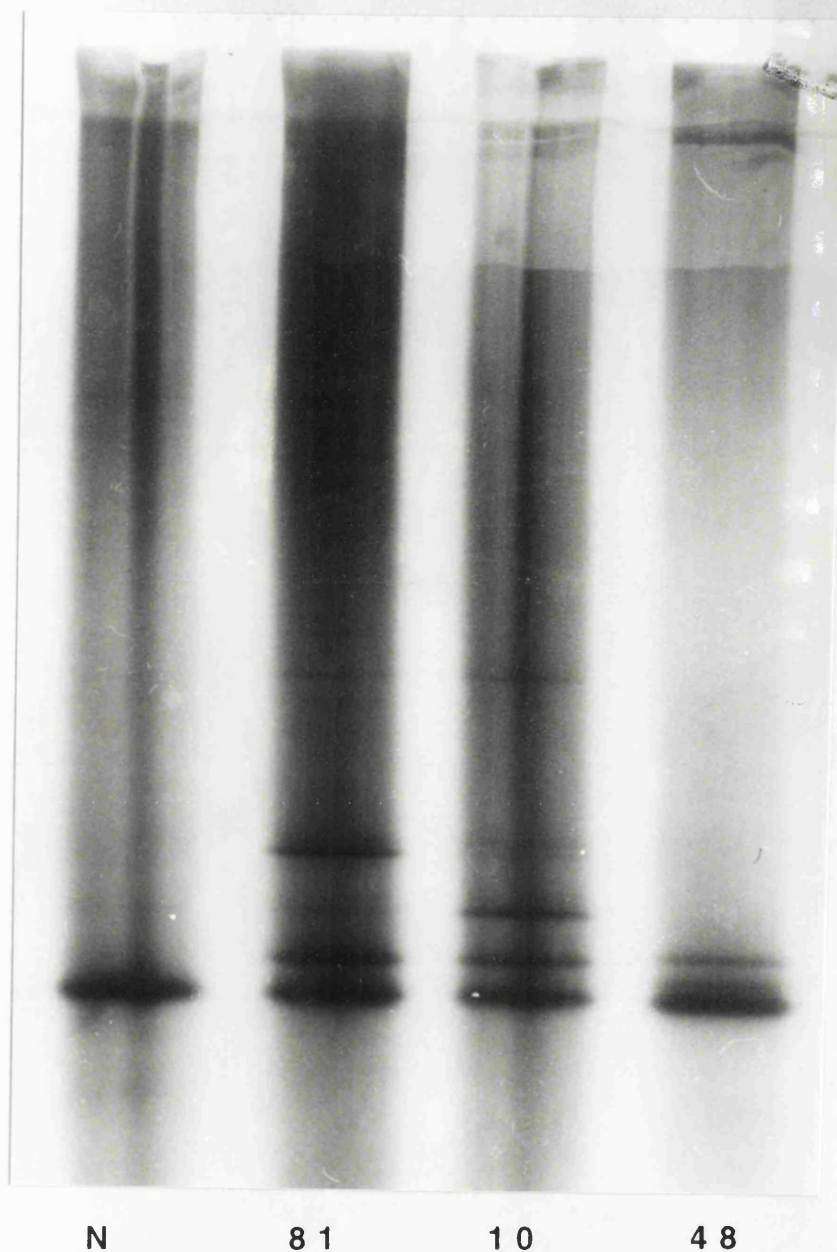
Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 3 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. A deletion of the triplet ATC indicated with + or\* results in the loss of Ile residue at position Ile<sup>94,95</sup>. Alternatively the same mutation could have arisen from deletion of the triplet CAT, shown in italics.

Fig 3.3.11 compares the sequence of mutated and normal alleles for a second exon 3 mutation. Here it can be seen that the mutation involves a deletion of three nucleotides ATC and results in the deletion of an Ile residue at amino acid position 94 or 95. This mutation has previously been observed (Caillaud *et al.*, 1991) and has been called delta ATC. A careful examination of the sequence suggest that this mutation may also be caused by a deletion of a CAT triplet (see figure 3.3.11). I will however, refer to this mutation as delta ATC in accordance with its naming in the literature.

### 3.3.5.2 Exon 10

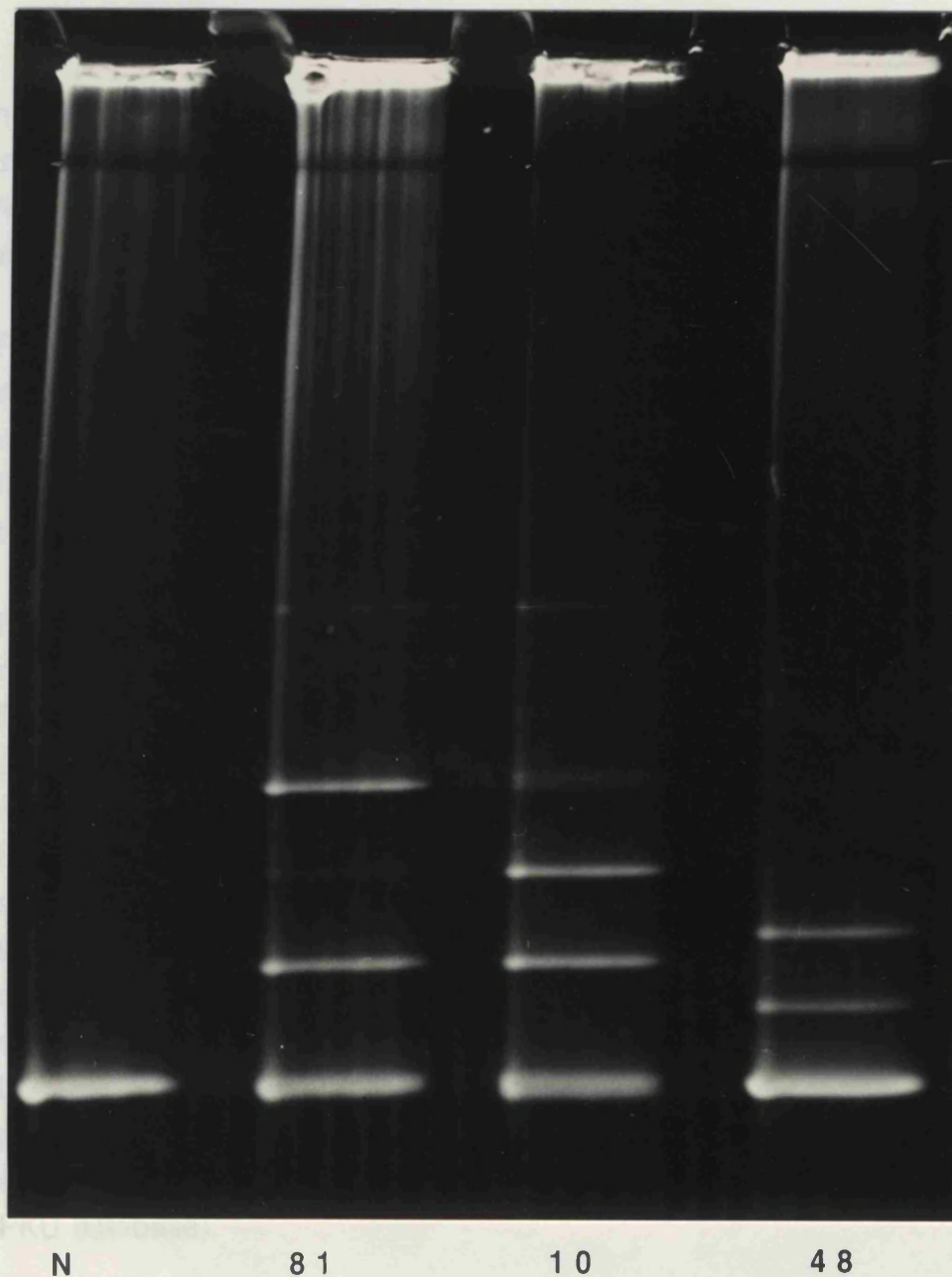
Analysis of exon 10 sequence using the Lerman algorithms suggested that the whole sequence could be resolved using a 48-52% denaturing gradient and a GC clamp on the upstream primer. DGGE was carried out under the conditions described in Table 2.4. The complete screening of the 50 PKU patients identified 3 different mutations in this exon at frequencies of 12%, 1% and 1% respectively. Fig 3.3.12 shows a silver stained gel containing these three mutation: mutation 1, present in patient 81, mutation 2 present in patient 10 and mutation 3 present in patient 48. The gel also includes a normal control, N.

Mutation 1 gives three resolvable bands. The lower band is identical to the band present in the normal control and is therefore the normal homoduplex form. It was anticipated that the middle band corresponds to the mutated homoduplex and the uppermost band correspond to the two heteroduplex forms. Naturally it was also possible that both homoduplexes co-migrate and therefore the two uppermost bands would correspond to the heteroduplex forms. It was established



**Fig 3.3.12 A silver stained DGG of the three mutation types identified in exon 10 using a 48-52 % denaturing gradient.**

A silver stained DGG showing the patient samples 81 (mutation 1), 10 (mutation 2) and 48 (mutation 3) which represent the three different banding patterns observed in the patients studied for mutation in exon 10. This pattern is described in the text. Sample N represents a normal control.



**Fig 3.3.13** An ethidium bromide stained CDG of the three mutation types identified in exon 10 using a 49% constant denaturing gel.

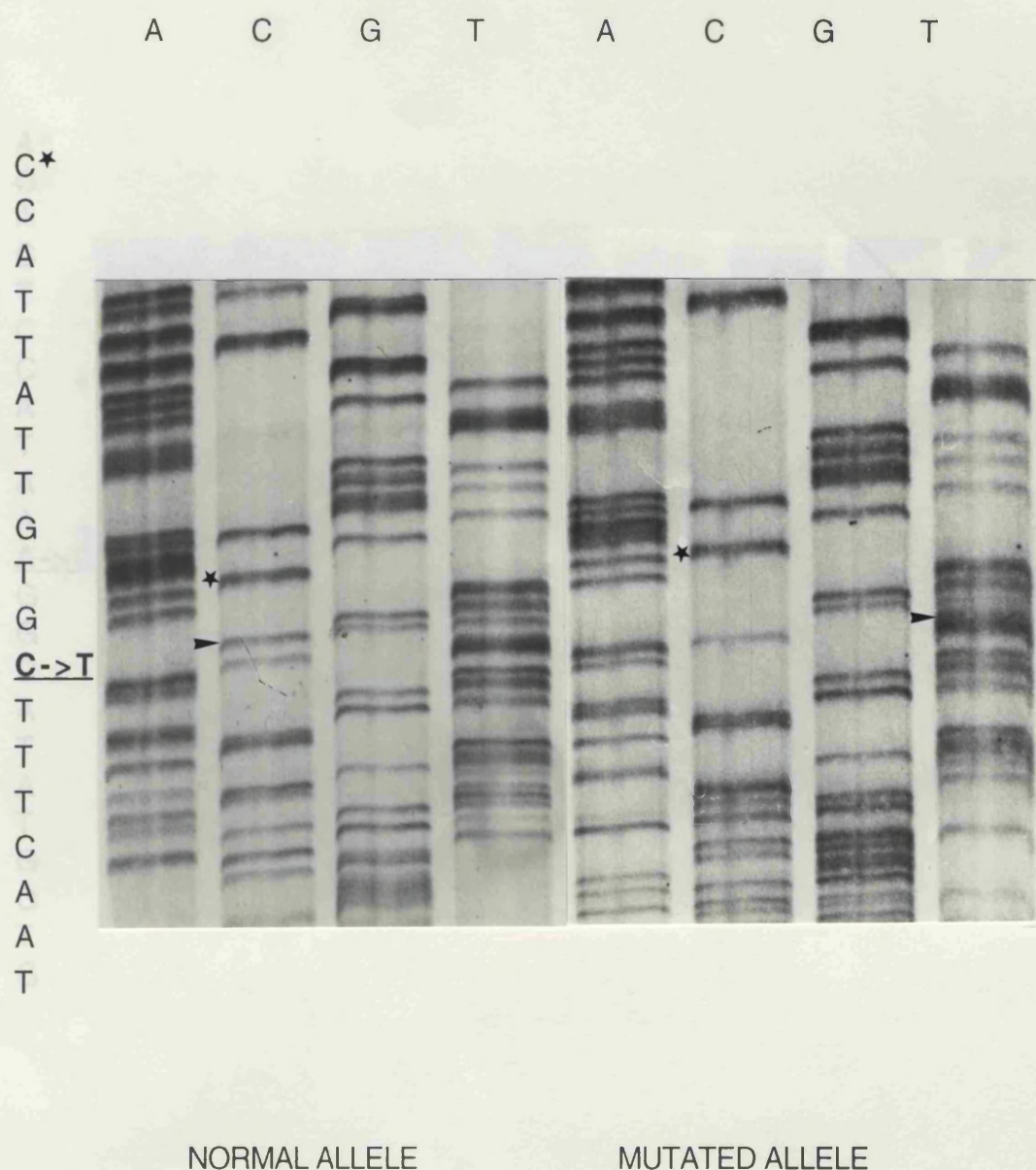
An ethidium bromide stained CDG showing the patient samples 81 (mutation 1), 10 (mutation 2) and 48 (mutation 3) which represent the three different banding patterns observed in the patients studied for mutation in exon 10. This pattern is described in the text. Sample N represents a normal control.

however following PCR and subsequent DGGE analysis of the lowermost band that this band consists solely of the normal homoduplex as only one resolvable band was observed. This was also found to be the case for mutation 2 and 3.

Fig 3.3.13 gives an example of an ethidium bromide stained gel of CDG analysis of each of the three mutation type as before with a normal control, N. In this example the homoduplex forms for patient 48 who is a carrier for mutation 3 are more easily resolved from each other.

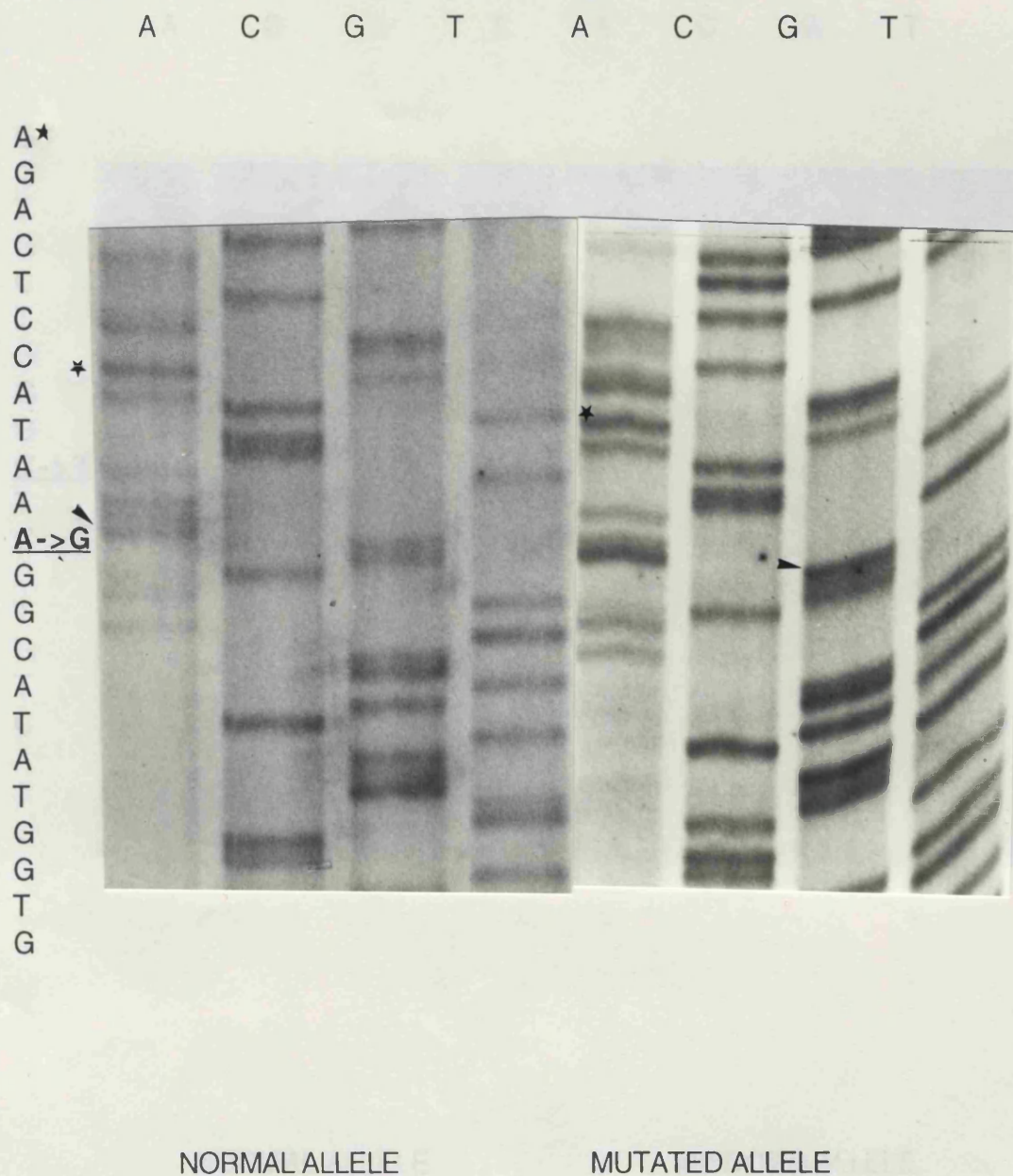
Fig 3.3.14-3.3.16 compares the sequence of mutated and normal alleles for these three exon/intron 10 mutations. DNA for each mutated allele was cloned and sequenced as previously described for exon 3. Mutation 1 is a previously unreported C to T substitution at position 66 in intron 10 and has been called IVS10nt66c->t. Mutation 2 is an A to G substitution at base position 1022 and results in a Lys to Arg amino acid substitution at position 341. This is also a novel mutation and has been called K341R. Mutation 3 is a substitution of a C to T at base position 1039 and results in a Leu to Phe substitution at position 347. This also has not previously been identified and has been called L347F (in agreement with Dr. C. Scriver responsible for the international PKU database).

A comparison of the amino acids at position 341 (mutation 2) and 347 (mutation 3) in the PAH genes for mouse, rat and drosophila shows that the amino acids at these positions are conserved across species (Fig 3.3.17). Mutation 1 presumably reflects a non pathological polymorphism since it is remote from the splice site region.



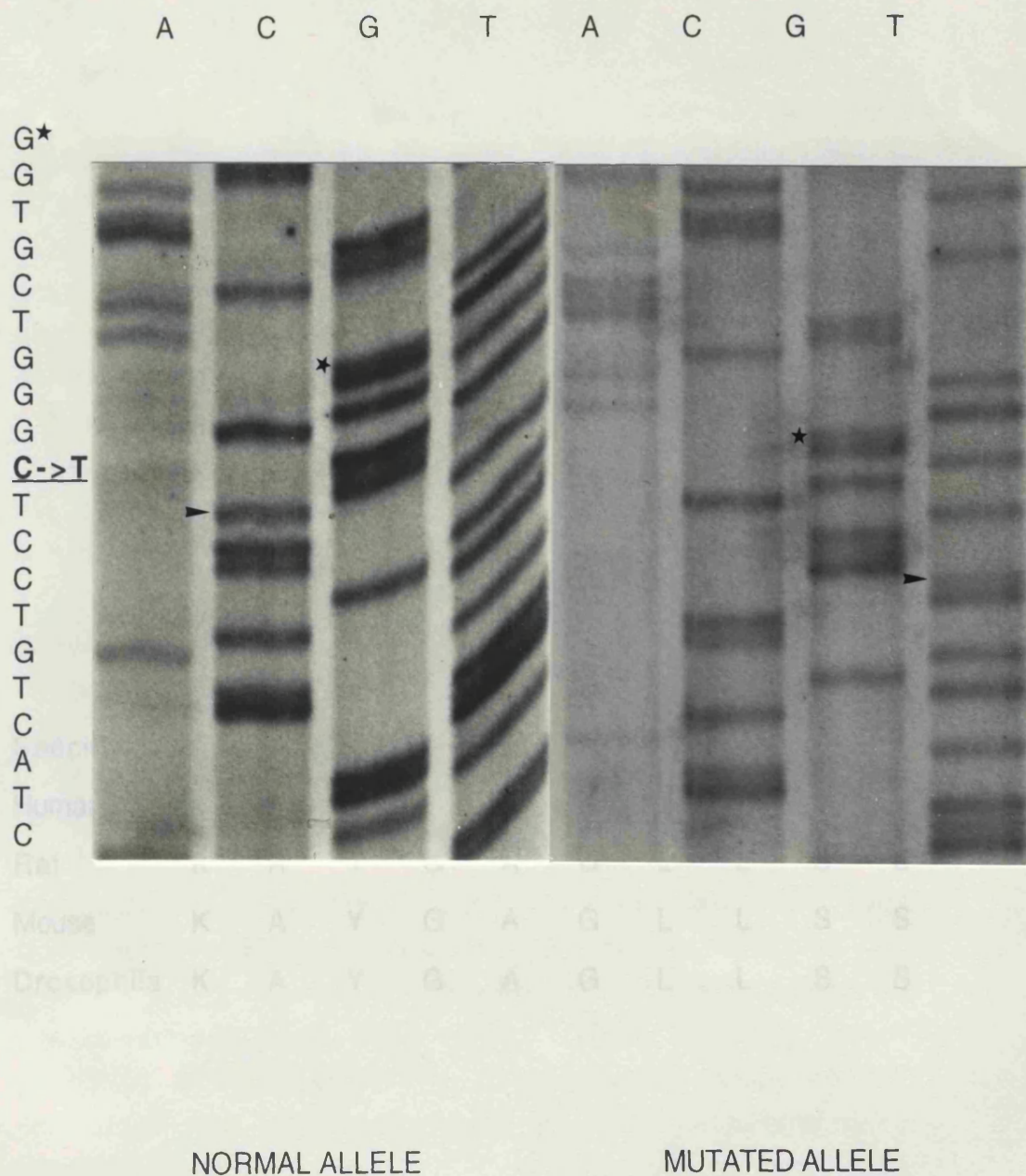
**Fig 3.3.14** Sequence comparison of the mutated and normal alleles of the exon 10 region containing mutation 1.

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of intron 10 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is an C to T transition at position 66 in the intron 10 sequence.



**Fig 3.3.15 Sequence comparison of mutated and normal alleles of the exon 10 region containing mutation 2.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 10 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is an A to G transition which results in the substitution of the normal amino-acid Lys<sup>341</sup> with Arg<sup>341</sup>.



**Fig 3.3.16 Sequence comparison of mutated and normal alleles of the exon 10 region containing mutation 3.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 10 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a C to T transition which results in the substitution of the normal amino-acid Leu<sup>347</sup> with Phe<sup>347</sup>.

Species	<u>341</u>	342	343	344	345	346	<u>347</u>	348	349	350
Human	K	A	Y	G	A	G	L	L	S	S
Rat	K	A	Y	G	A	G	L	L	S	S
Mouse	K	A	Y	G	A	G	L	L	S	S
Drosophila	K	A	Y	G	A	G	L	L	S	S

**Fig 3.3.17 Comparison of the coding region containing the K341R and L347F mutations in the human, rat, mouse and drosophila PAH protein.**

The amino acid positions 341 to 350 of the PAH protein are compared between human, rat mouse and drosophila. All the amino-acids in this region were conserved.

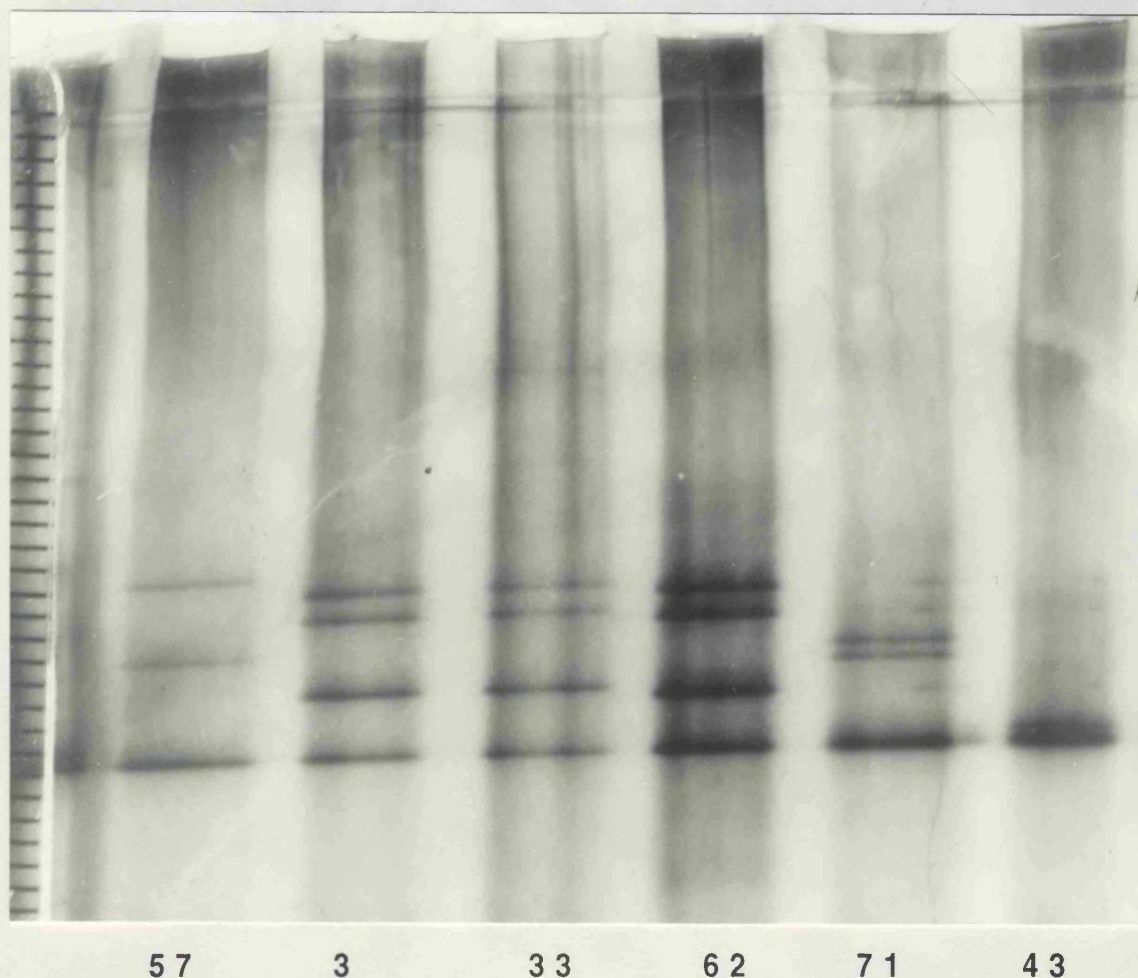
### 3.3.5.3 Exon 11

Analysis of exon 11 sequence using the Lerman algorithms suggested that the whole sequence could be resolved using a 50-60% denaturing gradient which was further refined to a CDG of 56%. DGGE was carried out under the conditions described in Table 2.4. The complete screening of the 50 PKU patients identified 5 types of mutations in this exon designated as mutations 1 to 5, at frequencies of 8%, 4%, 1%, 1% and 1% respectively.

Fig 3.3.18 a shows a silver stained gel containing these five mutation types: mutations 1-5 corresponding to patients 43, 3, 57, 62 and 71 respectively. Mutation 2 is represented by both patient 3 and 33. Fig 3.3.18 b shows an ethidium bromide stained gel of these five mutations.

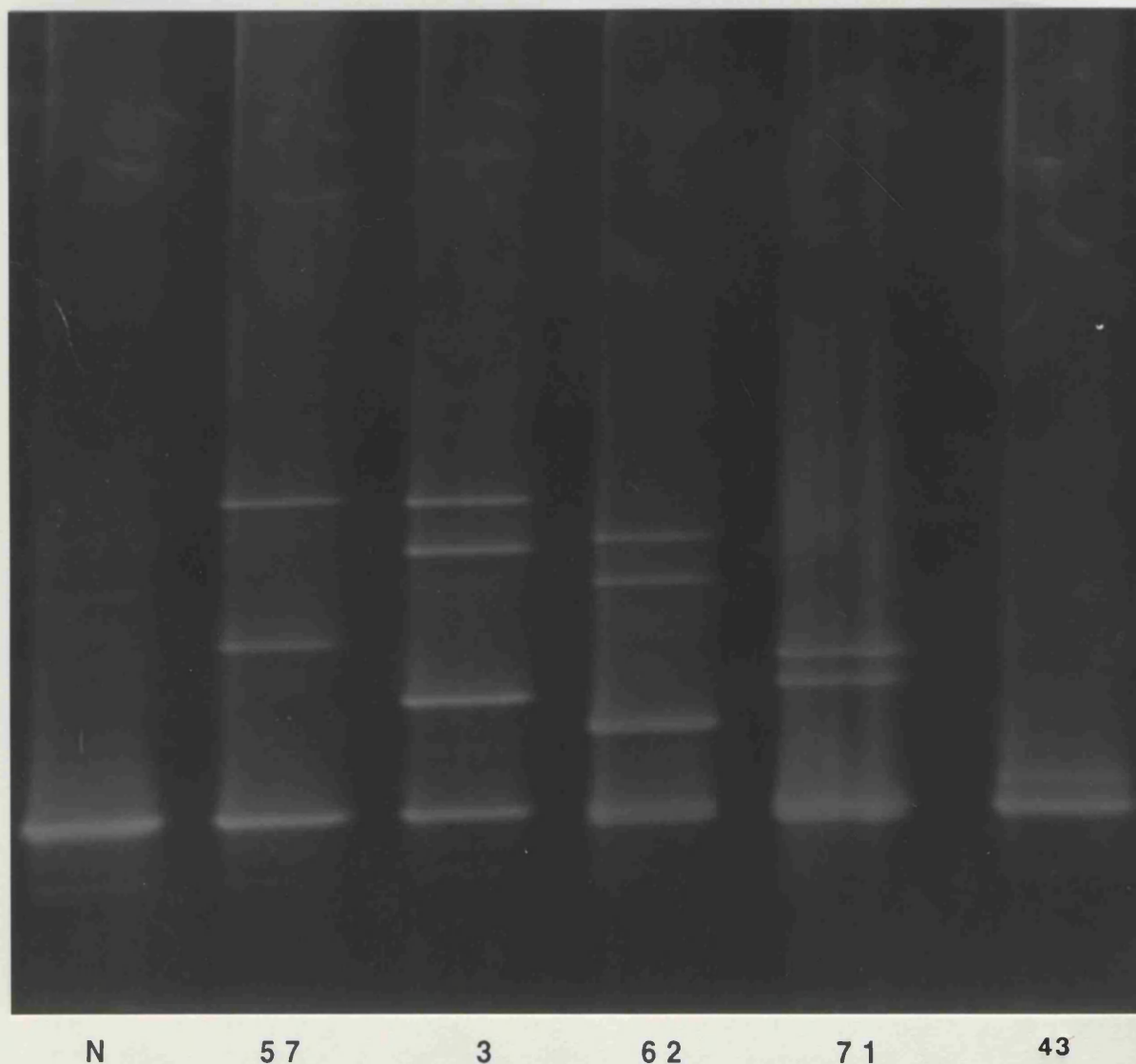
Mutations 2, 4 and 5 give four resolvable bands which correspond to the two homo- and heteroduplex forms. In contrast, mutation 3 gives only three resolvable bands. As previously suggested the absence of a fourth band may result from either the co-migration of the homoduplex or heteroduplex forms. This was considered by amplification of the lowermost band followed by DGGE analysis which generated an identical banding pattern to the normal control, consistent with the co-migration of the two heteroduplex forms. In contrast, mutation 1 shows only two resolvable bands. It appears that both homoduplexes and heteroduplex forms are not resolved from each other under these conditions. Alternatively the stability of the heteroduplex molecules may prohibit their resolution and so the two bands seen could correspond to the two homoduplex forms.

Fig 3.3.19 shows a CDG for mutation 1 (patient 43) using two



**Fig 3.3.18a** A silver stained CDG of the five mutation types identified in exon 11 using a 56% constant denaturing gel.

A silver stained CDG showing the patient samples 43 (Mutation 1), 3 and 33 (Mutation 2), 57 (Mutation 3), 62 (Mutation 4) and 71 (Mutation 5) which represent the five different banding patterns observed in the patients studied for mutation in exon 11. This pattern is described in the text.



**Fig 3.3.18b** An ethidium bromide stained CDG of the five mutation types identified in exon 11 using a 56% constant denaturing gel.

A silver stained CDG showing the patient samples 43 (Mutation 1), 3 (Mutation 2), 57 (Mutation 3), 62 (Mutation 4) and 71 (Mutation 5) which represent the five different banding patterns observed in the patients studied for mutation in exon 11. These patterns are described in the text. Sample N represents a normal control.

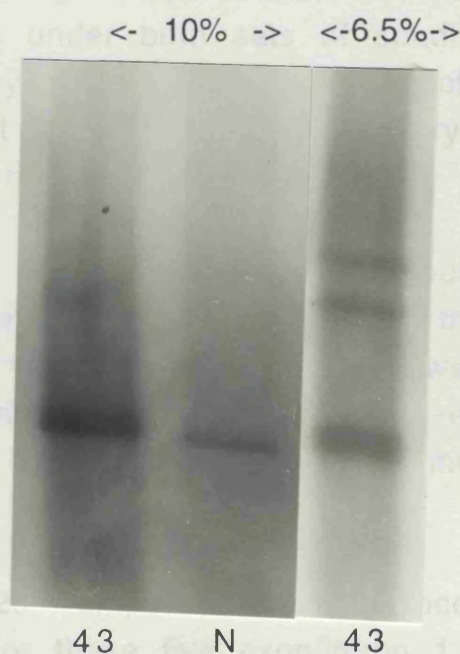
different percentages of acrylamide, 10% (as shown in Fig 3.3.19a and b) and 6.5%.

In the former conditions the resolution between the bands is poor but using a lower percentage of acrylamide, in this instance 6.5%, it was possible to resolve the two heteroduplex as well as the homoduplex forms originally in the 10% gel. In both cases an identical percentage of denaturing reagents was used. Thus, although it was possible to recognise the presence of the mutation under

was improved by the use of acrylamide. This suggests that at low concentrations of acrylamide the stability of heteroduplex

PCR products described for mutations in exon 11. In other mutations the results were established by preliminary sequencing. An example of a silver stained CDG of mutation 2 and 4 in exon 11.

Fig 3.3.21-3.3.22. Silver staining of mutation 2 and control allele for exon 11 mutation 11. DNA for each mutation was stained and sequenced as previously described for exon 3.



**Fig 3.3.19 Silver staining of a CDG of mutation 1 under identical denaturing conditions using different percentages of acrylamide**

Silver staining of patient sample 43 using 10 and 6.5% acrylamide under identical denaturing and electrophoretic conditions. The resolution of the heteroduplex and homoduplex forms (arrowed) is improved through a reduction in the acrylamide concentration used.

different percentages of acrylamide, 10% (as shown in Fig 3.3.18 a and b) and 6.5%.

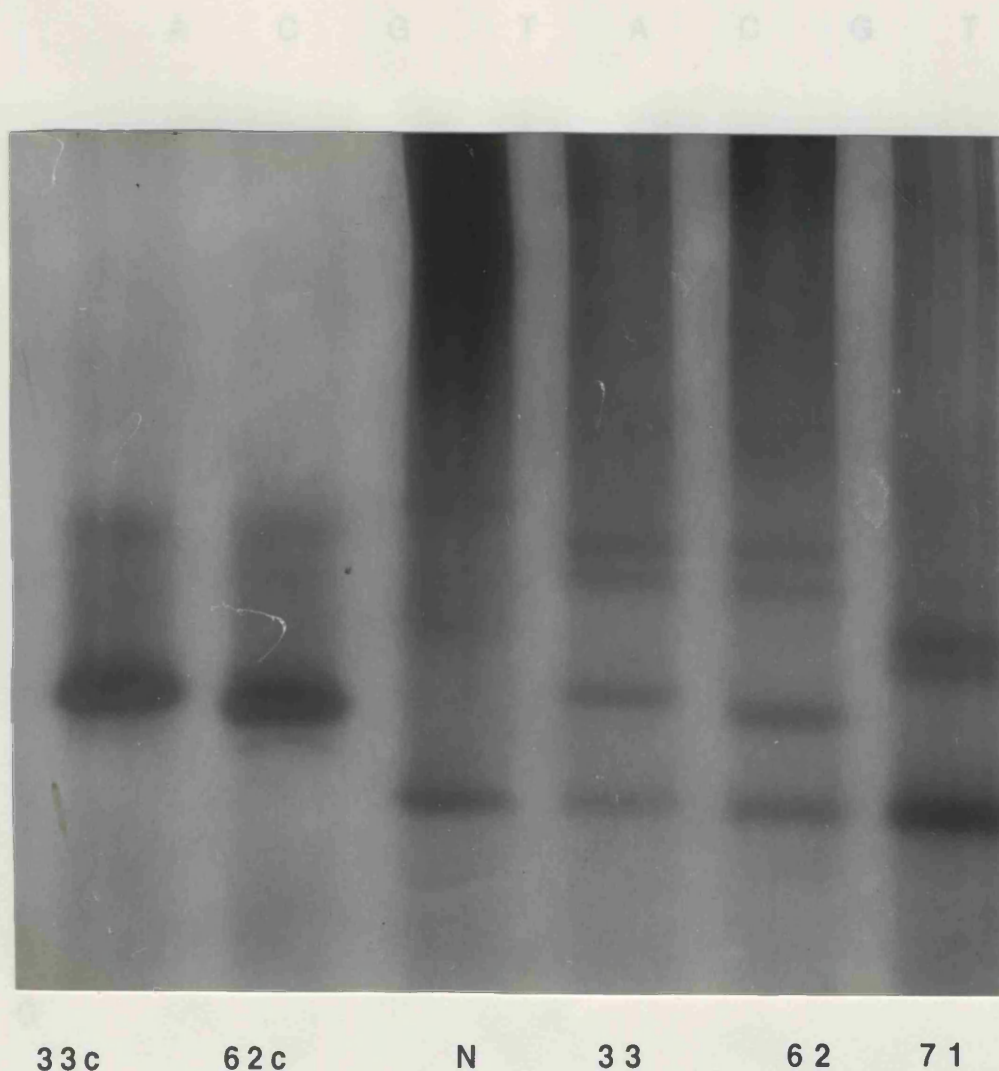
In the former conditions the resolution between the bands is poor but using a lower percentage of acrylamide, in this instance 6.5%, it was possible to resolve the two heteroduplex as well as the homoduplex forms originally in the 10% gel. In both cases an identical percentage of denaturing reagents were used. Thus, although it was possible to recognise the presence of the mutation under both sets of conditions its resolution was improved by altering the percentage of acrylamide. This suggests that at lower percentages of acrylamide the stability of heteroduplex molecules are increased.

PCR products were cloned as previously described for mutations in exon 3. As was the case in the analysis of other mutations the veracity of the clone was established by preliminary C DGE. Fig 3.3.20 gives an example of a silver stained CDG of PCR amplified clones for mutations 2 and 4 in exon 11.

Fig 3.3.21-3.3.25 compares the sequence of mutated and normal alleles for these five exon/intron 11 mutations. DNA for each mutated allele was cloned and sequenced as previously described for exon 3.

Mutation 1 is a G to C substitution in the coding sequence of exon 11 which results in a synonymous mutation at amino acid position 385. This mutation has previously been identified (Kalaydjieva *et al.*,1991) and has been called L385L.

Mutation 2 was found to be a G to A substitution in the intron 10-exon 11 splice site junction. This mutation has previously been identified (Dworniczak *et al.*,1991; Dasovich *et al.*,1991)



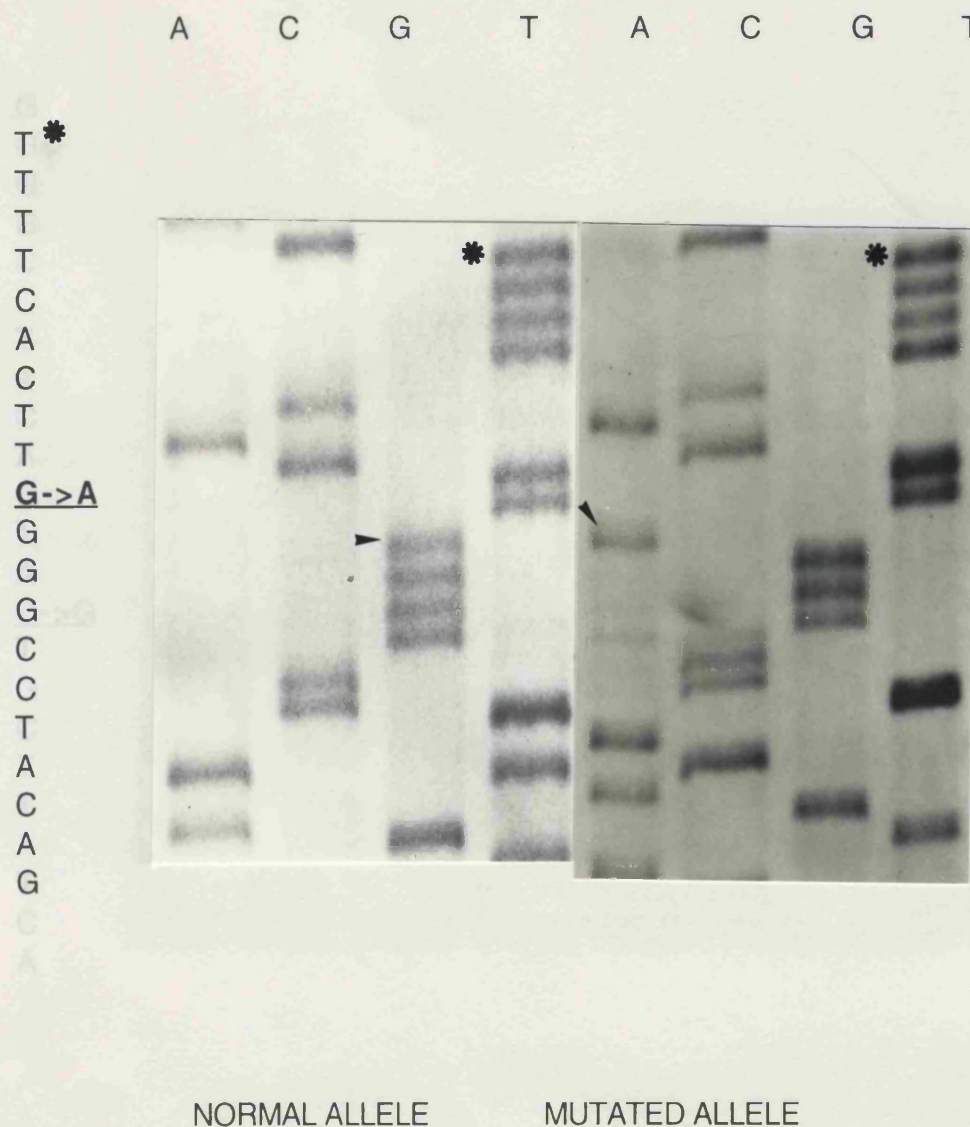
**Fig 3.3.20 Silver stained 56% CDG of exon 11 mutation 2, 4 and 5.**

CDG analysis of the PCR amplification products of the plasmid clones containing the mutated allele 2 and the mutated allele 4 (lane 1 and 2 respectively), a control (lane 3), a heterozygous patient for mutation 2 (lane 4), a heterozygous patient for mutation 4 (lane 5) and a heterozygous control corresponding to mutation 5 (lane 6). Each of the cloned DNA samples shows a band identical in mobility to that of the variant homoduplex obtained from genomic DNA.



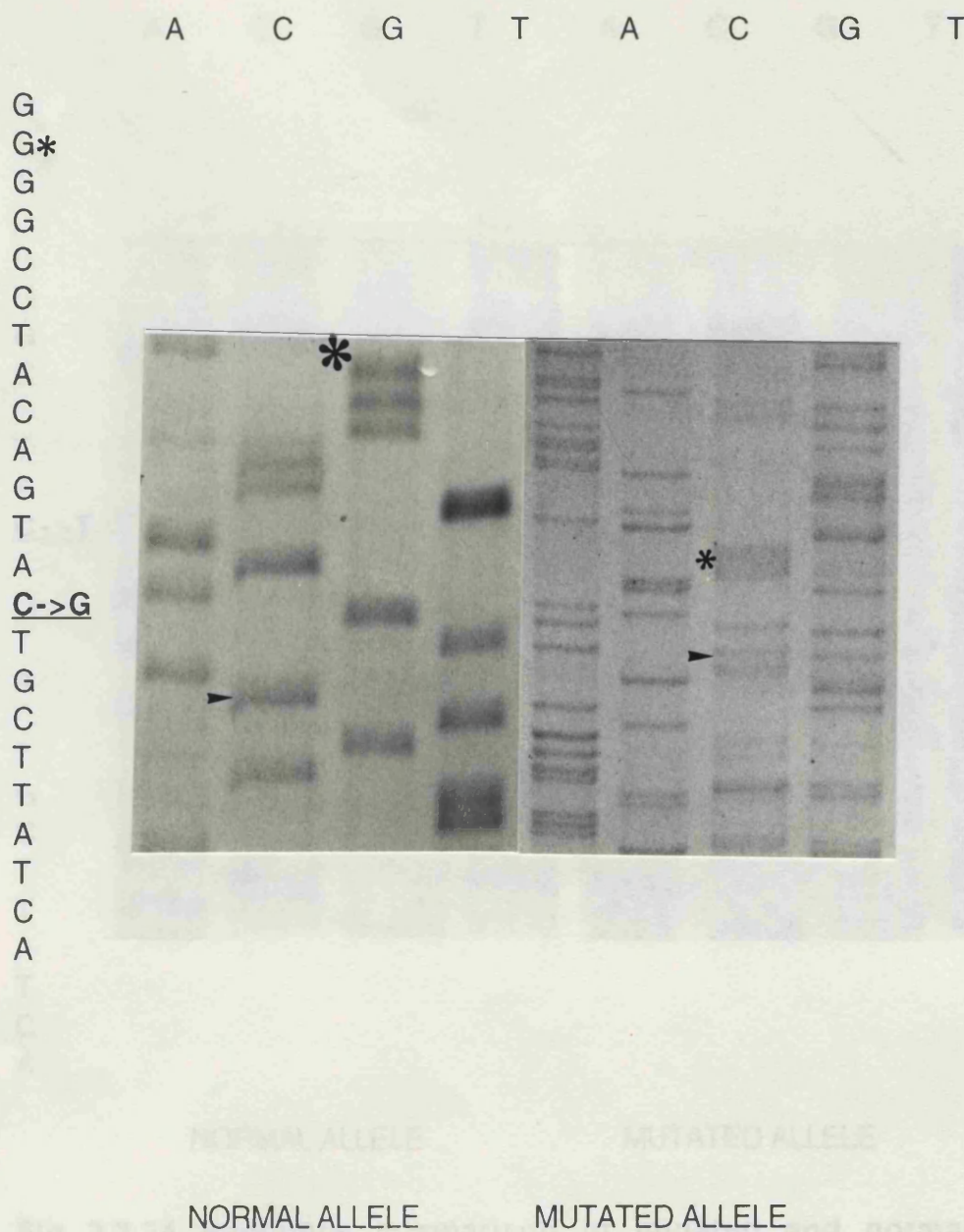
**Fig 3.3.21 Sequence comparison of mutated and normal alleles of the exon 11 region containing mutation 1.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of the exon 11 region of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a G to C transversion which results in the alteration of the 1155 nucleotide position in exon11.



**Fig 3.3.22 Sequence comparison of mutated and normal alleles of the intron 10 splice site region containing mutation 2.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of the intron 10 splice site region of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a G to A transition which results in the alteration of the 546 nucleotide position in the intron 10 splice acceptor site.



**Fig 3.3.23 Sequence comparison of mutated and normal alleles of the exon 11 region containing mutation 3.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 11 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a C to G transversion which results in the generation of a termination codon at amino acid position 356.



**Fig 3.3.24** Sequence comparison of mutated and normal alleles of the exon 11 region containing mutation 4.

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of intron 10 splice site region of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a C to T transition which results in the alteration of the 554 nucleotide position in the intron 10 splice acceptor site.

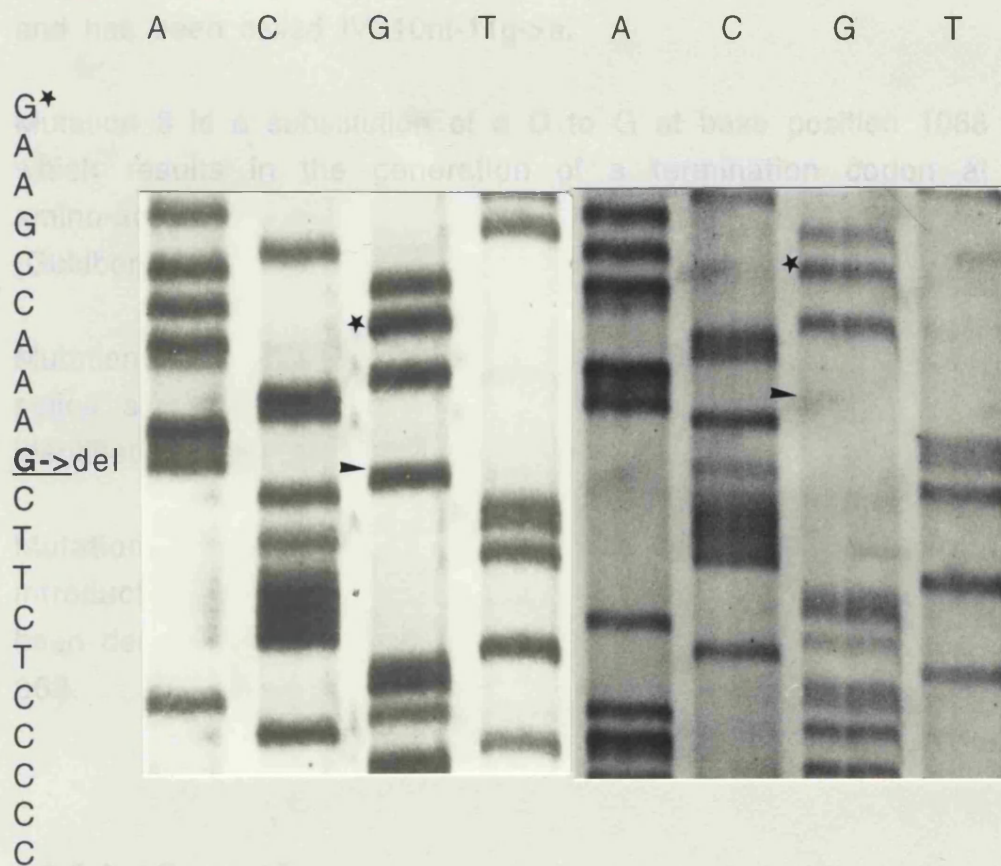


Fig 3.3.25 Sequence comparison of mutated and normal alleles of the exon 11 region containing mutation 5.

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 11 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. A deletion of a G at nucleotide position 1089 results in a frameshift mutation.

and has been called IVS10nt-11g->a.

Mutation 3 is a substitution of a C to G at base position 1068 which results in the generation of a termination codon at amino-acid position 356. This has previously been described (Guldborg *et al.*,1993) and has been called Y356X.

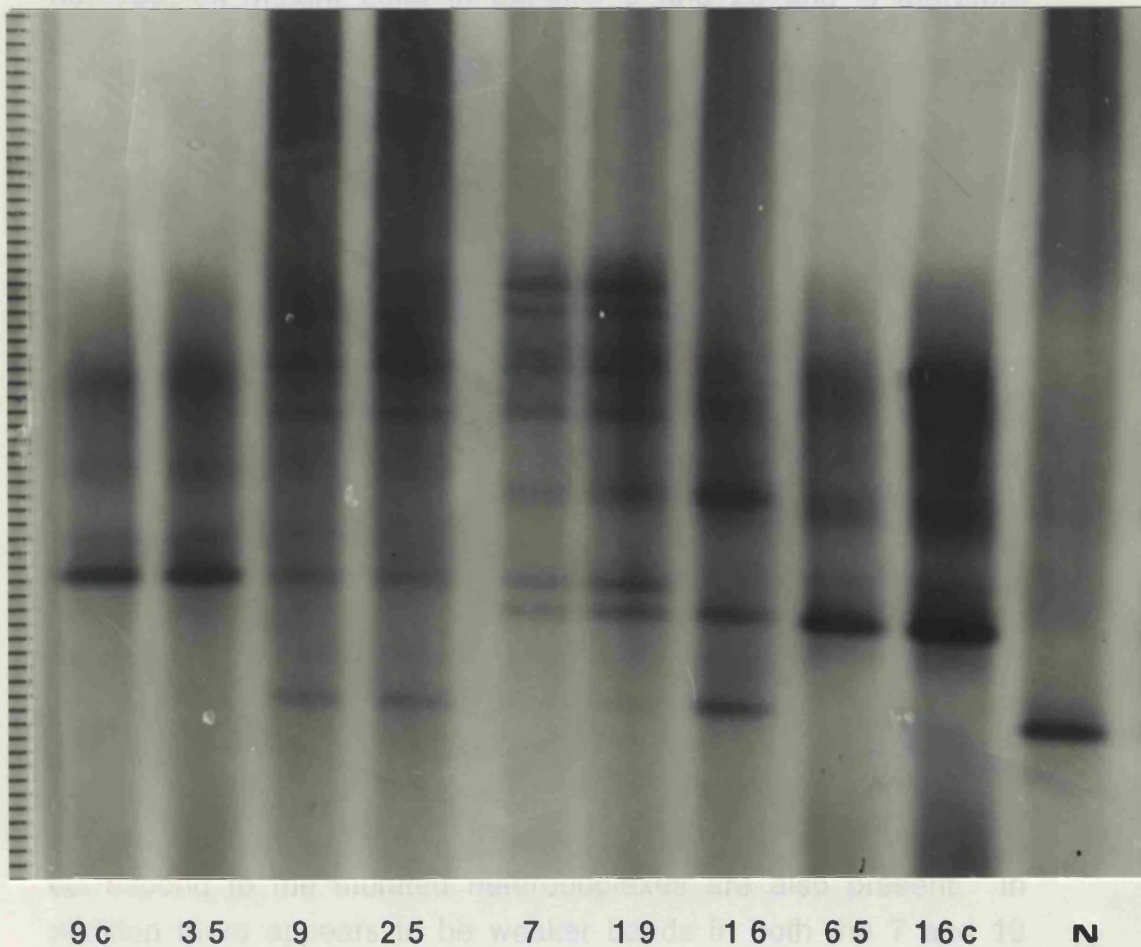
Mutation 4 is a C to T substitution in the intron 10-exon 11 splice site junction. This mutation is a third novel mutation identified in this study and has been called IVS10nt-3c->t.

Mutation 5 is a deletion of a G which results in the introduction in a frameshift mutation. This has previously been described (Guldborg *et al.*,1994) and has been called del K 363.

#### **3.3.5.4 Exon 12**

Analysis of exon 12 sequence using the Lerman algorithms suggested that the whole sequence could be resolved using a 54-57% denaturing gradient. DGGE was carried out under the conditions described in Table 2.4. The complete screening of the 50 PKU patients identified 2 mutation types at a frequency of 14% and 12% respectively. Fig 3.3.26 shows a silver stained gel containing these two mutation types:

Patients 9 and 25 are heterozygous for mutation 1. The DGGE pattern is composed of four separate bands corresponding to the normal and mutant homoduplex and heteroduplex forms. The lower band in each lane corresponds to the normal homoduplex form, observed in the normal control, N. Patient 35 shows a single DGGE band which corresponds to the



**Fig 3.3.26 A silver stained DGG of the two mutation types identified in exon 12 using a 54-57% denaturing gradient.**

A silver stained DGG of exon 12 PCR products from a number of patients having different combinations of exon 12 mutations. Patients 9 and 25 are heterozygous, patient 35 is homozygous for mutation 1. Patients 16 and 65 are heterozygous and homozygous for mutation 2 respectively and patients 7 and 19 compound heterozygotes for both mutations. The pattern of bands for each mutation are described in the text. Patient 3 provides a normal control (N). 9c and 16c represent DGGE analysis of cloned DNA of exon 12 containing the mutated allele for mutations 1 and 2 from patients 9 and 16 respectively.

homoduplex mutant band in patients 9 and 25 and is therefore homozygous for mutation 1.

Patient 65 shows a single mutant DGGE band which is also observed in patient 16 together with the "normal" band. Thus it seems that patient 65 is homozygous for mutation 2 and patient 16 is a heterozygote. In the DGGE of patient 16 only three bands were resolved, the two homoduplex forms corresponding to the two lower bands and the higher band corresponding to the two heteroduplex forms. The heteroduplexes presumably cannot be distinguished from each other.

Patients 7 and 19 seem to be compound heterozygotes for mutations 1 and 2 in exon 12. Neither patient has a band corresponding to the normal homoduplex form. In contrast, each patient has bands corresponding to both the mutated homoduplex forms observed in DGGE for mutations 1 and 2. Two further bands, not present in any of the other patients, which correspond to the mutated heteroduplexes are also present. In addition there appears to be weaker bands in both the 7 and 19 patient lanes whose presence are difficult to explain.

DGGE analysis for both mutations 1 and 2 are shown in lanes 9c and 16c respectively. The letter c refers to the use of cloned DNA to generate the amplified product and the number refers to the patient DNA used for the cloning experiment. These single bands are identical to those observed in homozygous 35 and 65 (for mutations 1 and 2) respectively.

Twelve patients had mutation 1 and nine patients had mutation 2. Of these, only two patient, 7 and 19, were compound

heterozygotes. The pattern of mutation detection for all 50 affected individuals is shown in Table 3.3.2.

These mutations follow a mendelian pattern of segregation in the small number of families available for analysis. In general, samples were only available from just one parent, usually the mother and the affected offspring. Fig 3.3.27 shows the results in an interesting family (38, Table 3.3.3), where both children (sample 59 and 60) are homozygous for mutation 1 in exon 12. The mother is heterozygous for this mutation and the father was not available for study. The observation that both offspring are homozygous suggested that the father was also a carrier for the same mutation. The parents were unrelated and therefore the chance occurrence of homozygotes must reflect the high frequency of this mutation (14%), in this population group.

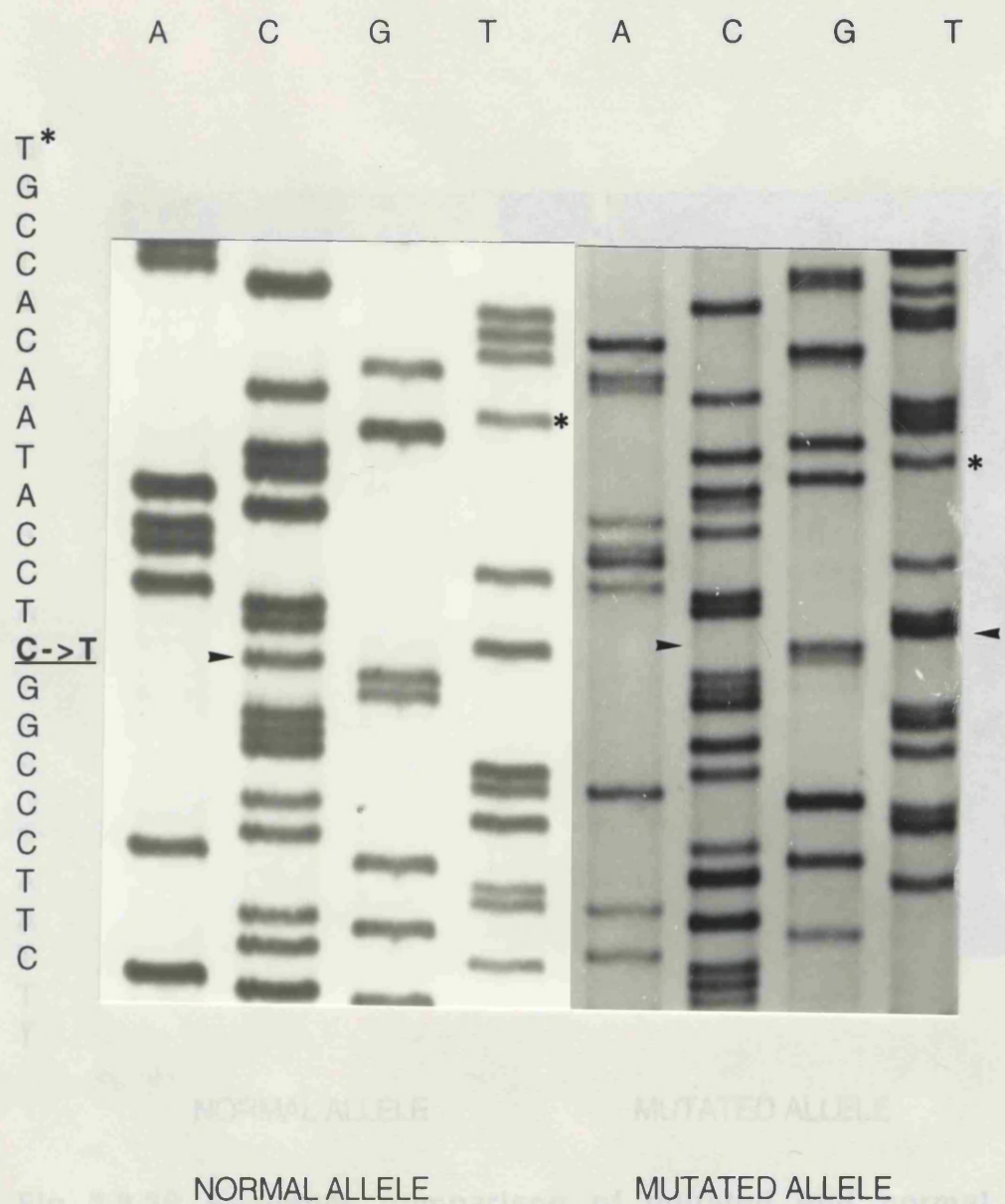
Fig 3.3.28-3.3.29 compares the sequence of mutated and normal alleles for these two exon/intron 12 mutations. DNA for each mutated allele was cloned and sequenced as previously described (section 3.3.5.1). Fig 3.3.26 shows the DGGE analysis for clones 9c and 16c which corresponds to the mutant alleles 1 and 2 respectively. Each clone gives a single resolvable band corresponding to the mutated homoduplex form.

Mutation 1 is a C to T substitution at position 1222 in the PAH sequence. This results in a Arg-Trp substitution at amino acid position 408. This change has been previously described (DiLella *et al.*, 1987) and is referred to as R408W. A comparison of the amino acid at position 408 of the PAH genes in the mouse, rat and *Drosophila* show that this position is conserved across species (Fig 3.3.30).



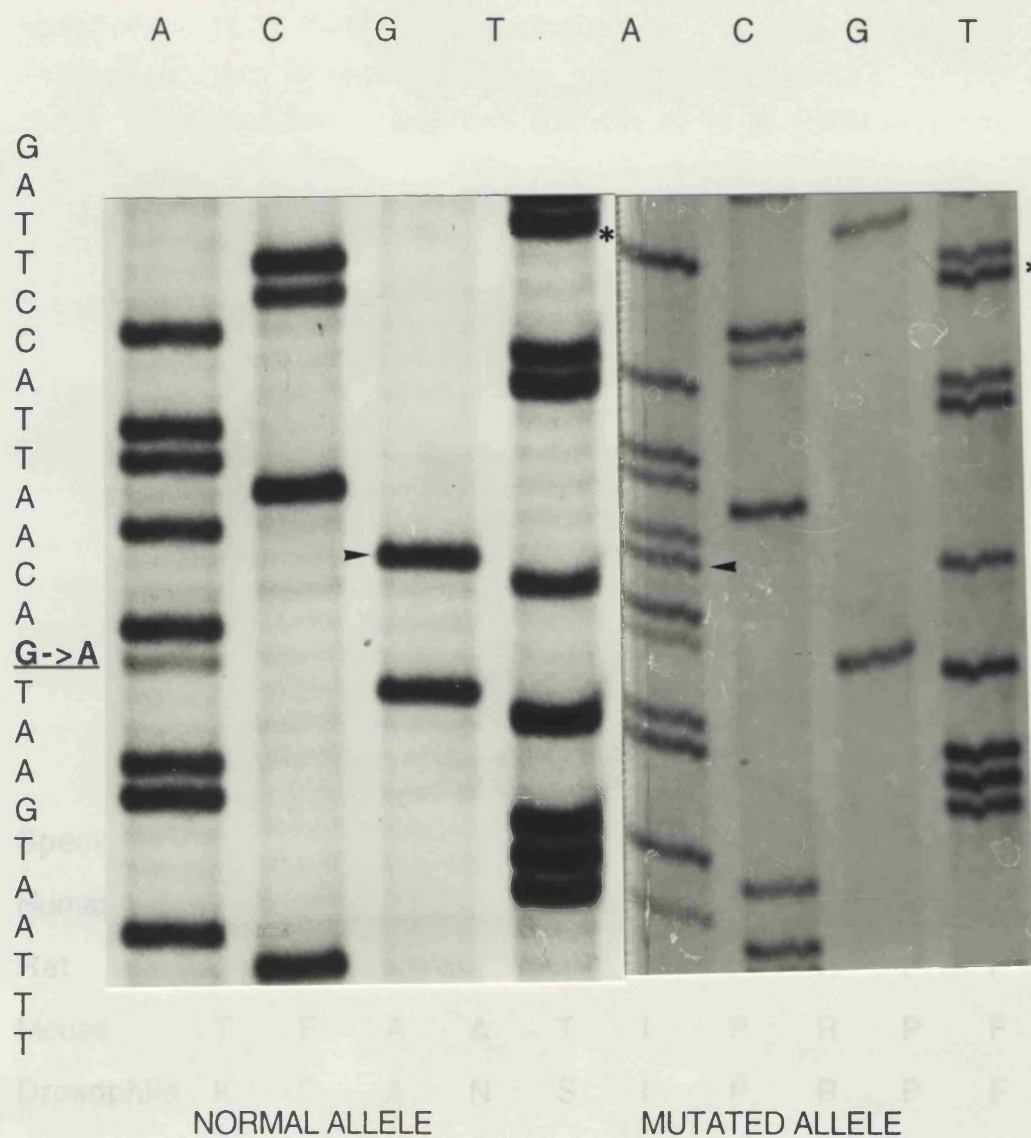
**Fig 3.3.27 DGGE analysis of the PAH exon 12 in a PKU family showing mendelian inheritance of mutation 1**

DGGE analysis for exon 12 in family 38 (Table 3.3.3) consisting of two parents and two affected offsprings. The mother(61) is a carrier for the mutation. The father was not available. The affected offsprings (59 and 60) have inherited the mutation 1 allele from their mother and it would appear the same mutated allele from their father. N refers to normal control.



**Fig 3.3.28 Sequence comparison of mutated and normal alleles of the exon 12 region containing mutation 1.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of exon 12 of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a C to T transition which results in the substitution of the normal amino-acid Arg<sup>408</sup> with Trp<sup>408</sup>.



**Fig 3.3.29 Sequence comparison of mutated and normal alleles of the intron 12 splice donor position containing mutation 2.**

Partial sequence ladders of cloned PCR products showing normal and mutated alleles of the intron 12 splice donor site of the PAH gene. The sequencing lanes are labelled A, C, G and T respectively. The mutation is a G to A transition which results in the alteration of the first nucleotide position in the intron 12 splice site sequence.

Mutation 2 is a G to A substitution at the first nucleotide position in the intron 12 donor splice site junction. This change has also been described (DiLella *et al.* in 1986) and has been called IVS-12nt1.

### 3.3.5.5 Exon 5

DGGE was carried out under the conditions described in Table 2.4. The complete screening of the 50 PKU patients identified no mutations. Fig 3.3.31 shows a silver stained gel of a DGGE analysis of 19 patients and normal controls (lanes 2, 19). Only a single band was resolvable in all instances. This band was identical in all the patient and control samples suggesting that there were no mutations present in exon 5 in the PKU population studied.

Species	401	402	403	404	405	406	407	408	409	410
Human	N	F	A	A	T	I	P	R	P	F
Rat	T	F	A	A	T	I	P	R	P	F
Mouse	T	F	A	A	T	I	P	R	P	F
Drosophila	K	F	A	N	S	I	P	R	P	F

**Fig 3.3.30 Comparison of the coding region containing the R408W mutation in the human, rat, mouse and drosophila PAH protein.**

The amino acid positions 401 to 410 of the PAH protein are compared between human, rat mouse and drosophila. Amino-acids which are variant between species are shaded. The Arg residue at amino acid position 408, which is mutated in PKU, is conserved in all four species.

Mutation 2 is a G to A substitution at the first nucleotide position in the intron 12 donor splice site junction. This change has also been described (DiLella *et al.* in 1986) and has been called IVS-12nt1.

#### **3.3.5.5 Exon 5**

DGGE was carried out under the conditions described in Table 2.4. The complete screening of the 50 PKU patients identified no mutations. Fig 3.3.31 shows a silver stained gel of a DGGE analyse of 19 patients and normal controls (lanes 2, 19). Only a single band was resolvable in all instances. This band was identical in all the patient and control samples suggesting that there were no mutations present in exon 5 in the PKU population studied.

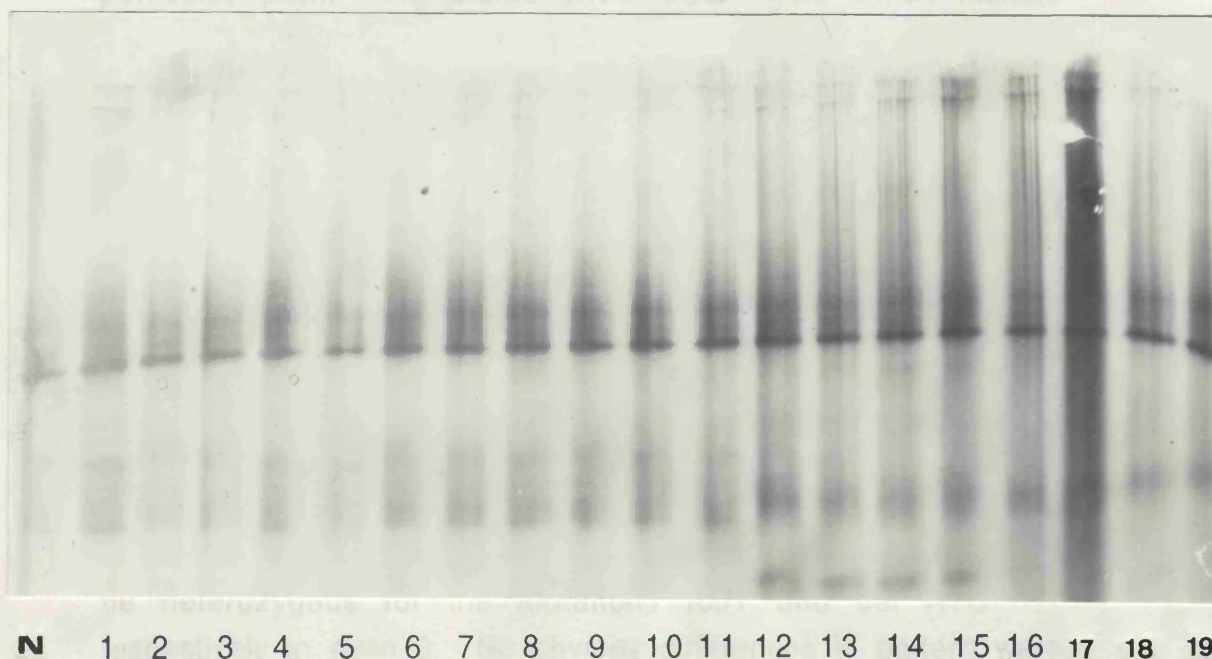
#### **3.3.6 Evaluation of SSCP as a method of mutation detection.**

##### **3.3.6.1 Evaluation of SSCP as a method of mutation detection for resolving the 12 different mutation types identified by DGGE.**

All 12 mutations identified by DGGE analysis were also examined by SSCP using the Pharmacia Phast system. Where possible at least two patient DNAs for each mutation were investigated. PCR of these DNAs were checked for specific amplification as shown in Fig 3.3.1. Samples were denatured in formamide at 95°C and run as described (section 2.2.7) under a variety of different electrophoretic conditions. Following electrophoresis, DNA was visualised by silver

staining (section 2.2.5.3) and the single strand conformation patterns examined.

In total SSCP analysis led to the detection of 4 of the 12 mutations, K341R, L347F, IVS12nt1 and IVS10156 that had previously been identified using DGGE. The approach was not very satisfactory as bands present in normal controls were not always evident in patient samples although these patients were known to be heterozygous for mutations for that particular exon. Fig 3.3.32 shows SSCP gels which include



**Fig 3.3.31 A silver stained CDG of exon 5 using a 55-65% denaturing concentration.**

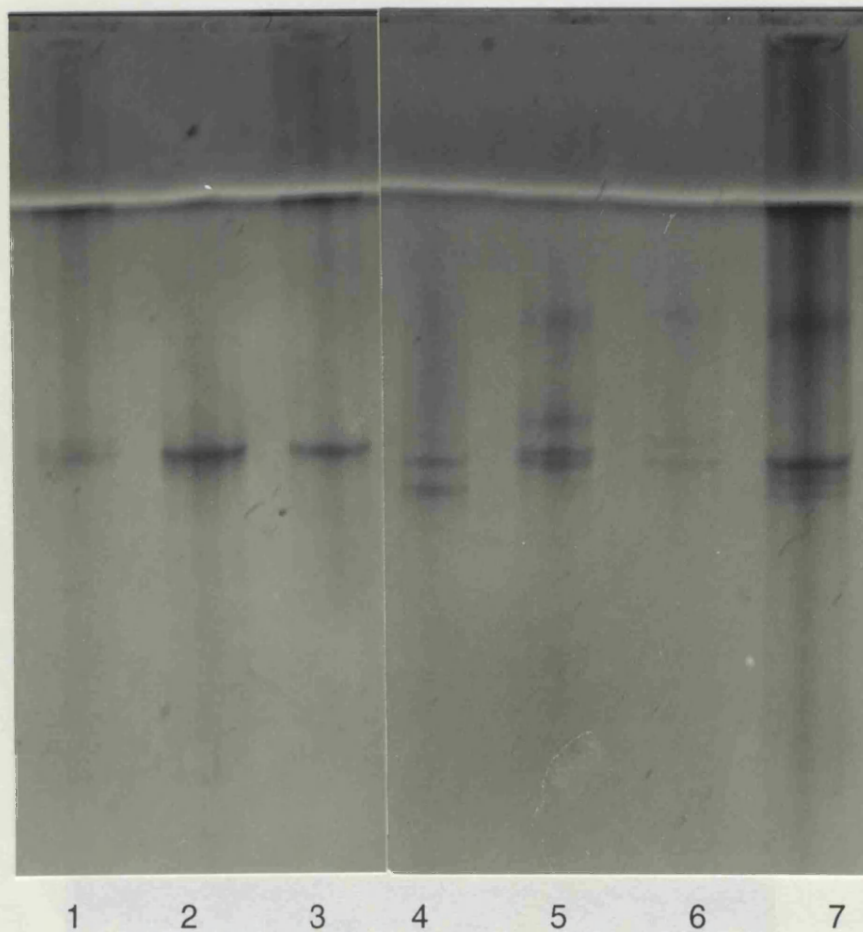
All products from the 19 patient samples and the normal control resolve at the same position in the gel. No mutations were identified.

staining (section 2.2.5.3) and the single strand conformation patterns examined.

In total SSCP analysis led to the detection of 4 of the 12 mutations, K341R, L347F, IVS12nt1 and IVS10nt66 that had previously been identified using DGGE. The approach was not very satisfactory as bands present in normal controls were not always evident in patient samples, although these patients were known to be heterozygous for mutations for that particular exon. Fig 3.3.32 show SSCP gels which include these four mutations along with suitable controls of normal DNA. The gels can be divided into four parts, lanes 1-3, 4-7, 8-11, and 12-13, corresponding to analysis of exon 3, 10, 11 and 12, of the PAH gene respectively.

#### Lanes 1-3 (Exon 3):

Lanes 1-3 show SSCP analysis of PCR products from exon 3 of the PAH gene, corresponding to PCR amplifications for normal and 2 patient samples. Patients 17 (lane 2) and 49 (lane 3) have previously been shown by DGGE and sequence analysis to be heterozygous for the mutations I65T and del ATC<sup>94,95</sup> respectively in exon 3. No obvious differences in pattern were observed between these two mutated alleles. It was surprising however that the banding pattern observed for the normal control was not always observed in patient lanes, given that these patients were heterozygous for the mutations studied. In this instance there is a weaker additional band in the normal control (not present in either of the two mutated samples). This is a phenomenon frequently associated with the Phast SSCP system and was observed throughout the course of this study in addition to separate experiments carried out by other individuals. SSCP analyses over a range of different

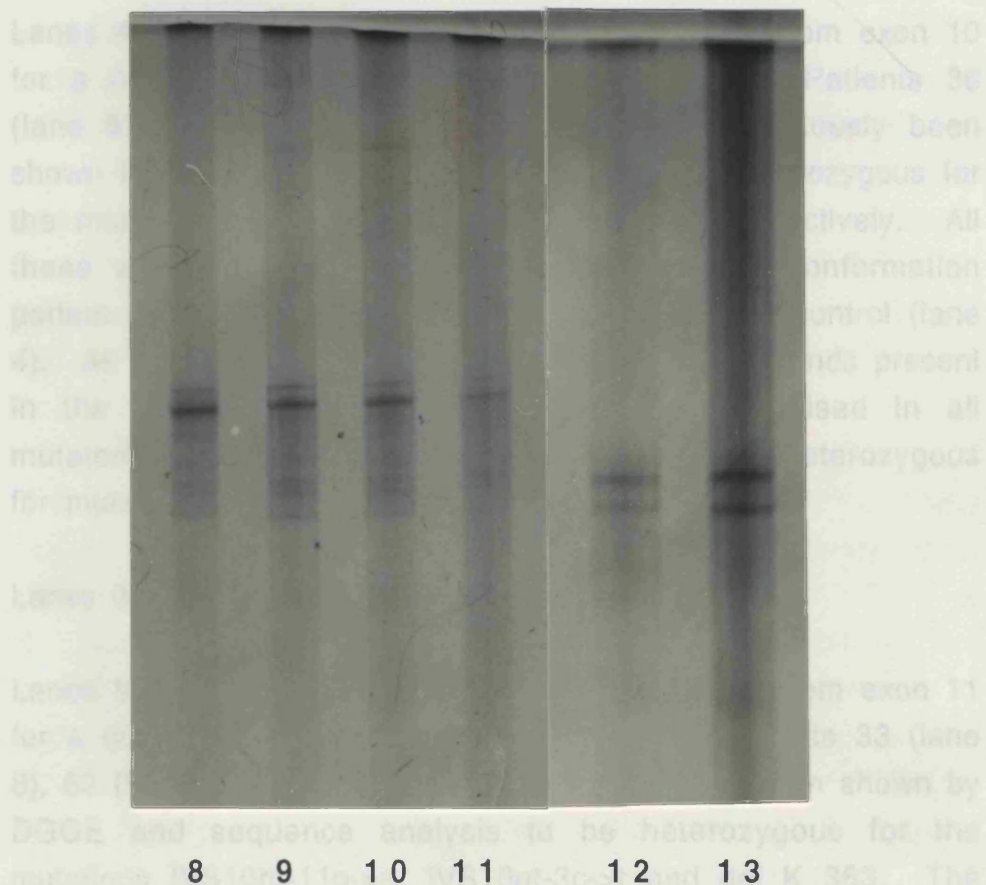


**Fig 3.3.32** A silver stained SSCP gel of PCR products containing known mutations identified by DGGE analysis for exons 3, 10, 11 and 12 of the PAH gene.

SSCP studies of exon 3 (lanes 1-3), exon 10 (lanes 4-7), exon 11 (lanes 8-11) and exon 12 (lanes 12-13) respectively were carried out at 10°C for 40 min at 400 V. SSCP analysis for nine mutations identified by DGGE are shown in addition to suitable controls. The banding patterns observed allowed four of these mutations to be resolved although the SSCP patterns were difficult to interpret. The gels and these observations are described in the text.

temperatures and electrophoretic times (Table 2.2.7) did not succeed in resolving either of these two mutations.

Lanes 4-7 (Exon 10) (lanes 8-11 (Exon 11) and lanes 12-13 (Exon 12))



**Fig 3.3.32** A silver stained SSCP gel of PCR products containing known mutations identified by DGGE analysis for exons 3, 10, 11 and 12 of the PAH gene.

SSCP studies of exon 3 (lanes 1-3), exon 10 (lanes 4-7), exon 11 (lanes 8-11) and exon 12 (lanes 12-13) respectively were carried out at 10°C for 40 min at 400 V. SSCP analysis for nine mutations identified by DGGE are shown in addition to suitable controls. The banding patterns observed allowed four of these mutations to be resolved although the SSCP patterns were difficult to interpret. The gels and these observations are described in the text.

temperatures and electrophoretic times (Table 2.2.7) did not succeed in resolving either of these two mutations.

#### Lanes 4-7 (Exon 10):

Lanes 4-7 show SSCP analysis of PCR products from exon 10 for a normal individual and 3 patient samples. Patients 36 (lane 5), 10 (lane 6) and 48 (lane 7) have previously been shown by DGGE and sequence analysis to be heterozygous for the mutations IVS10nt66, K341R and L347<sup>-</sup> respectively. All these alleles show a distinctive single strand conformation pattern compared with each other and the normal control (lane 4). As observed for SSCP studies of exon 3 the bands present in the normal control lane could not be visualised in all mutated samples, although these patients were heterozygous for mutations in exon 10.

#### Lanes 8-11 (Exon 11):

Lanes 8-11 show SSCP analysis of PCR products from exon 11 for a normal person and 3 patient samples. Patients 33 (lane 8), 62 (lane 9) and 71 (lane 10) have previously been shown by DGGE and sequence analysis to be heterozygous for the mutations IVS10nt-11g->a, IVS10nt-3c->t and del K 363. The single strand conformation patterns of these mutations were indistinguishable from each other and the normal control (lane 11). Patients having the Y356X and L385L mutations (not shown) also gave an identical banding pattern to the normal controls.

SSCP analysis over a range of different temperatures and electrophoretic times (section 2.2.7) did not succeed in resolving any of these mutations.

Lanes 12-13 (exon 12):

Lanes 12-13 show SSCP analysis of exon 12 PCR products from a normal control and a single patient sample patient 65 (lane 12) which has been shown by DGGE and sequencing analyses to be heterozygous for mutation IVS-12nt1. This sample shows a different single strand conformation band patterns to the normal control (lane 13). The normal and mutated lanes have one band in common and a second band where the band in the mutated sample has migrated fractionally slower. This very slight difference was found to be reproducible. Patients with the R408W mutation in exon 12 did not show any change in single strand mobility (not shown) in relation to the normal control.

The paucity of mutations identified, four out of twelve, using SSCP in these experiments may reflect the relatively large size of the PCR products used in the analysis. This is illustrated by the resolution of four out of a total of five mutations present in exons 10 and 12 (PCR products of approximately 250 bp each), in contrast to none out of a total of six mutations for exons 3 and 11 (PCR product sizes of approximately 350 bp), Table 3.3.3. The reduced efficiency of SSCP to identify mutations with increasing size of PCR products, has been discussed by Sheffield *et al.* 1993.

#### **3.3.6.2 SSCP as a technique to identify mutation other than those resolved by DGGE**

In the DGGE analysis of exon 12, two mutations were identified (section 3.3.5.4). As six other mutations in Caucasians have previously been reported in exon 12 (PKU consortium), SSCP gels were also run using the conditions outlined in section

Exon	PCR product length	Identification of mutation by SSCP
3	338 bp	0/2
1 0	272	3/3
1 1	341	0/5
1 2	251	1/2

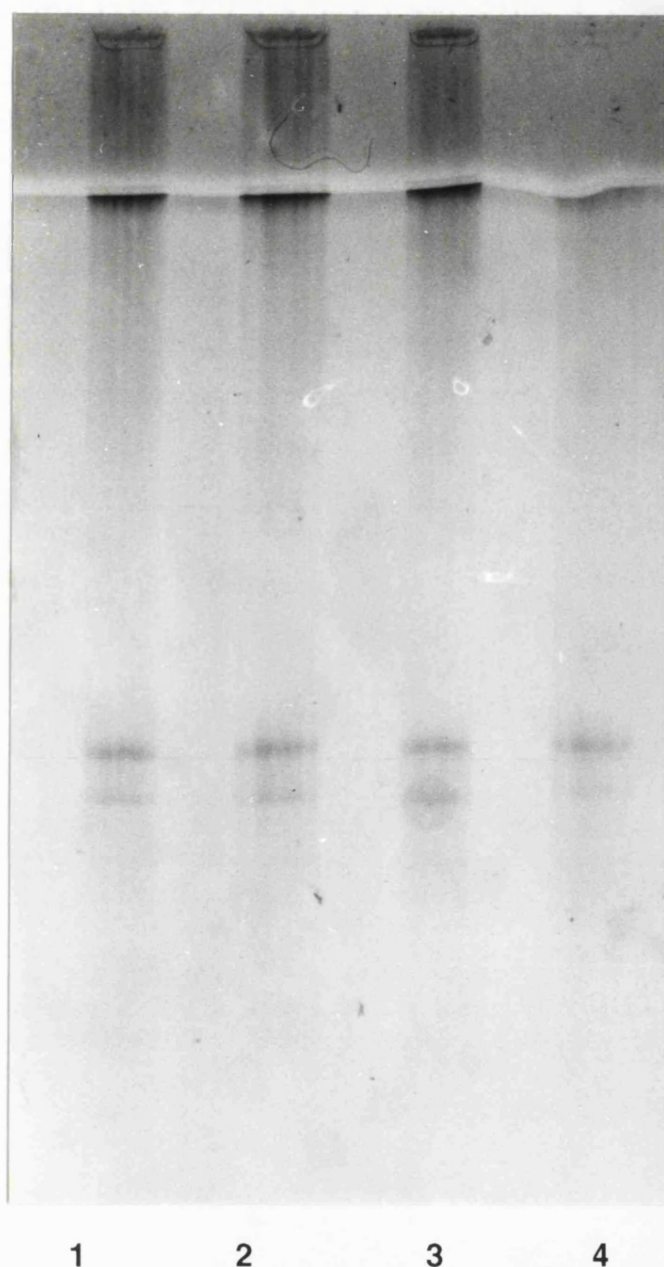
**Table 3.3.3 Summary of the results of the SSCP analysis for the 12 different mutations identified by DGGE**

Table 3.3.3 shows the PCR product sizes used in the SSCP studies of each exon and the success in the resolution of the mutations identified by DGGE. The relationship between PCR size and sensitivity of the SSCP technique is discussed in the text.

2.2.7 for 24 patient samples which had shown no mutation for exon 12 by DGGE (Table 3.3.1)

Although the SSCP approach had succeeded in resolving the IVS12nt1 no additional mutations were identified by SSCP in this study. For example, Fig 3.3.33 shows a SSCP gel of 4 individuals out of the 24 patients tested. There is no difference between the single strand conformation pattern of any of these samples and the normal control. This result is consistent with the observation obtained by DGGE analysis that this particular patient population contains only two exon 12 mutations. This is not very surprising as these other six mutations were reported to occur at rather low frequencies.

SSCP was used in an effort to resolve the 12 different mutations identified in exons 3, 10, 11 and 12 of the PAH gene by DGGE. In total 4 of these 12 mutations were resolved, all three mutations in exon 10 and one of the two mutations identified in exon 12. None of the mutations which occur in exons 3 and 11 were resolved.



**Fig 3.3.33** A silver stained SSCP gel of PCR products from four patient samples which showed no mutations in exon 12 by DGGE analysis.

SSCP studies for four samples were carried out at 5°C for 40 min at 400 Volts. The banding patterns were the same in all cases.

## Chapter 4. Discussion

### 4.1 Summary

Recent advances in molecular genetics have resulted in a growing requirement for the screening of known disease causing genes and of candidate genes for mutations. As a result greater emphasis has been placed on devising new methods for mutation detection to complement the conventional and laborious DNA sequencing strategies.

I chose to consider the suitability of DGGE as a mutation detection technique in favour of other screening strategies as the melting theory at the root of this approach allowed the optimisation of screening conditions for mutation detection. This is in contrast to more empirical techniques such as SSCP analysis and HA. It was also appealing, that initial studies using DGGE had reported a high sensitivity of mutation detection (Theophilus *et al.*, 1989).

In total three separate test systems were considered. The technique was initially set up using the human alpha-1-antitrypsin gene. The six major alleles of the gene were resolved. The second study attempted to identify variation in a short Y chromosomal fragment. This also provided an opportunity to integrate the use of the Lerman algorithms to optimise screening conditions. The concluding study utilised the DGGE technique in conjunction with these algorithms to screen for mutations in the PAH gene of PKU patients. All three studies, as well as other projects published recently (Table 4.1) confirmed the suitability of DGGE as a mutation detection system.

Gene	Disease	References
Androgen receptor	Androgen Insensitivity	De Bellis <i>et al.</i> , 1992
Androgen receptor	Prostate cancer	Newmark <i>et al.</i> , 1992
AAT	alpha one antitrypsin	Dubel <i>et al.</i> , 1991
APC	Familial adenomatous polyposis (FAP)	Fodde <i>et al.</i> , 1992; Olschwang <i>et al.</i> , 1992, 1993; Wallis <i>et al.</i> , 1992
Apolipoprotein E	Hyperlipidemia	Van de Maagdenberg <i>et al.</i> , 1993
CFTR	Cystic fibrosis (CF)	Vidaud <i>et al.</i> , 1990; Devoto <i>et al.</i> , 1991; Ferec <i>et al.</i> , 1992
Collagenase	Recessive dystrophic epidermolysis bullosa RDEB	Hovnanian <i>et al.</i> , 1991
COL3A1	Ehlers-Danlos syndrome type IV	Johnson <i>et al.</i> , 1992
Dopamine D2 receptor	Schizophrenia, delusional disorders	Catalano <i>et al.</i> , 1992
Electron transfer flavoprotein	Glutaric acidemia type II	Freneaux <i>et al.</i> , 1992
Factor VIII	Haemophilia A	Traystman <i>et al.</i> , 1990; Higuchi <i>et al.</i> , 1991; Aly <i>et al.</i> , 1992
Factor IX	Haemophilia B	Attree <i>et al.</i> , 1989; Vidaud, 1993
$\beta$ -globin	$\beta$ -Thalassemia	Cai and Kan, 1990; Losekoot <i>et al.</i> , 1990, 1991, 1992; Abrams <i>et al.</i> , 1990; Rosatelli <i>et al.</i> , 1992
gamma -globin	Hereditary persistence of fetal hemoglobin (HPFH)	Gottardi <i>et al.</i> , 1991

Gene	Disease	References
		(continued)
$\beta$ -Glucocerebrosidase		Laubscher <i>et al.</i> , 1994
$\beta$ -Glucosidase	Gaucher disease (GD)	Theophilus <i>et al.</i> , 1989
Growth hormon receptor gene	Laron syndrome	Berg <i>et al.</i> , 1992
HPRT	Gout	Cariello <i>et al.</i> , 1988
hTR- $\beta$	Generalised resistance to thyroid hormone (GRTH)	Takeda <i>et al.</i> , 1992
Insulin receptor	Leprechaunism	Barbetti <i>et al.</i> , 1992
Insulin receptor		Desbois, <i>et al.</i> , 1993
Insulin receptor	Non-insulin-dependent diabetes mellitus (NIDDM)	Cocozza <i>et al.</i> , 1992
Low density lipoprotein (LDL) receptor	Familial hypercholesterolemia	Top <i>et al.</i> , 1992; 1993
Neurofibromatosis type 1 (NF1) gene	NF1	Valero <i>et al.</i> , 1994
Ornithine aminotransferase (OAT)	Gyrate atrophy	Mashima <i>et al.</i> , 1994
Ornithine transcarbamylase (OTC)	Ornithine transcarbamylase deficiency	Finkelstein <i>et al.</i> , 1990
p53		Beck <i>et al.</i> , 1993
p53	Gastric cancer	Renault <i>et al.</i> , 1993
p53	Colorectal cancer	van den Broek <i>et al.</i> , 1993
p53	Acute Leukemias	Pignon <i>et al.</i> , 1994
Phenylalanine hydroxylase (PAH)	Phenylketonuria (PKU)	Guldborg <i>et al.</i> , 1993
PKD	Polycystic kidney disease	Saris <i>et al.</i> , 1990
Porphobilinogen deaminase (PBGD)	Acute intermittent porphiria (AIP)	Gu <i>et al.</i> , 1994
k-ras	Pancreatic cancer	Pellegata <i>et al.</i> , 1992

Gene	Disease	References
(continued)		
Rhodopsin	Autosomal dominant retinitis pigmentosa (ADRP)	Sheffield <i>et al.</i> , 1991
ROM1	Vitelliform macular dystrophy (Best disease)	Nichols <i>et al.</i> , 1994
vWF	von Willebrand's disease type IIB	Ribba <i>et al.</i> , 1991

**Table 4.1 DGGE applications in Human Molecular Genetics (Fodde and Losekoot, 1994 and added applications published since review submitted)**

In the following section, these experiments will be discussed in greater detail. In subsequent sections I will review and relate this approach to the other strategies commonly used to screen for mutation. In doing so, the importance and various roles for mutation studies will also be developed.

## **4.2 DGGE studies**

### **4.2.1 Alpha 1 antitrypsin**

Initially DGGE was applied to the identification of known mutations in the human alpha-1-antitrypsin gene. Patient samples had previously been examined by isoelectric focussing and their mutations classified. In total, all the six major AAT alleles considered: M1Ala, M1Val, M2, M3, S and Z were resolved. These are described in Table 3.1.1. The Lerman computer algorithms however were not available, initially, to optimise screening conditions. Thus, this part of my work provided an opportunity to become familiar with the apparatus and the methods required for the use of the DGGE technique as well as giving an initial indication as to the sensitivity of the approach.

The six variants occurred in three separate exons of the AAT gene. Perpendicular gels were run for each exon to evaluate if and where a GC clamp was necessary and to determine the appropriate range in denaturing gradient required for mutation detection (Fig 3.1.1). Furthermore it gave some indication of the melting map of the fragment by alluding to the number of melting domains present within the fragment. Unfortunately this meant that primer pairs were designed without prior knowledge of their impact on the melting map and their efficiency for subsequent screening studies.

This is borne out by the difficulties encountered in resolving the AAT mutations. For example it was not possible to resolve the M1Val/M1Ala and S alleles on exon 3 using the same primer pair (section 3.1.5, Fig 3.1.2 a and b). Indeed, two completely independent primer pairs had to be used to resolve each allele. This was also the case when resolving the M3 and Z alleles present in exon 5 (section 3.1.5, Fig 3.1.3 a and b). Fortunately both the MELT 87 and SQHTX algorithms were available for subsequent analyses on the Y chromosome and PAH gene. It was encouraging however that all six variants were resolved as a result of the trial and error approach adopted. Moreover, the distinctive banding patterns allowed any allele combination to be readily identified.

#### **4.2.2 Y chromosome**

The study of a 470 bp sequence, called pY-80, on the Y chromosome provided the first real opportunity to apply the MELT87 and SQHTX algorithms to optimise DGGE electrophoretic conditions. Although DGGE can be used to screen for mutation in fragments in excess of 500 bp, simulation of the melt map using the MELT87 program showed that there was no co-operativity in melting between different segments of the sequence (Fig 3.2.2). The introduction of GC clamps at either end of the sequence although increasing the general co-operativity across the sequence, still delimited the screening of some parts of the fragment in closest proximity to the GC clamps (Fig 3.2.3 a and b).

By simulating the affect of incorporating a GC clamp at different positions along the sequence and looking at its impact on the melt map the sequence could be subdivided into two overlapping fragments, each fragment consisting of a

single melting domain (Fig 3.2.4, Fig 3.2.5). The SQHTX algorithm complemented the MELT87 program analysis by predicting the level of displacement that would be achieved between the wild-type sequence and a mutated heteroduplex for every conceivable position contained within the sequence. This ensured that the most appropriate set of primers were synthesised. In addition, optimal electrophoretic conditions were provided so that the actual displacement values would be sufficient to allow any possible mutation to be resolved.

In the study of 20 individual male samples no sequence variation was discovered. This is consistent with the observations from other studies which have suggested that there is a very low level of polymorphism on the Y chromosome (Jakubiczka *et al.*, 1989; Malaspina *et al.*, 1990; Spurdle *et al.*, 1994). Indeed only seven reliably documented Y-specific RFLPs have been reported (reviewed in Jobling, 1994). As DGGE identifies all possible sequence variation within a sequence, unlike RFLP analysis which looks solely for loss or gain of a restriction site, it is a far more efficient screening approach. In an effort to identify sequence variation, a more extensive search using DGGE analysis would have to be carried out.

The introduction however, of a single base change in one primer gave us the opportunity to incorporate a positive control for mutation detection into the experiment (Fig 3.2.7; 3.2.8). The single base substitutions within the primer (see Table 2.3), of a (G->A) for Y1 and a (T->G) for Y2 gave the expected changes in banding pattern. The mutated PCR product focussed higher in the gradient gel for Y1 and lower for Y2 as a result of an increase and decrease respectively in the sequence stability caused by the introduction of a G->A or T->G mismatch.

The identification of these single base alterations suggested that the use of the Lerman algorithms had successfully optimised conditions for the resolution of these specific mutations and consequently other variations if they had been present in this particular sequence. This is in marked contrast to the difficulties encountered in optimising appropriate conditions as apparent in the alpha-1-antitrypsin studies when the computer programs were not available.

The main objective of this study in addition to setting up and utilising the Lerman algorithms for mutation detection was to identify additional variation which would allow the paternal lineage in human evolution to be analysed. The paucity of Y polymorphisms has hindered the analysis of its evolution. Fortunately recent investigation have uncovered variation in the repeat lengths of Y specific hypervariable repeating units (Jobling,1994; Mathias *et al.*,1994) which will allow this evolutionary question to be addressed.

### **4.2.3 Studies on Phenylketonuria**

#### **4.2.3.1 Family studies: DGGE analyses**

The first mutations described for PKU required the rather laborious task of cloning individual genes from patients and their subsequent sequence analysis to identify the molecular lesions involved (DiLella *et al.*,1986). Since then several different methods of screening that do not involve cloning have been developed in combination with PCR to identify additional mutations. The PAH gene region of interest for mutational analyses has been restricted to exonic and splice site junction sequences as most functionally significant mutations could be expected to fall into these regions. This has been borne out by

the PAH mutation studies described during the period of my work. For example, DGGE analysis of the Danish and Sicilian PKU population for mutations within the coding and splice site regions have identified 99% of mutations (Gulberg *et al.*, 1993a; 1993b). Indeed, an international database for PKU mutations set up by the PAH mutation analysis consortium report the existence of approximately 180 individual mutations, all of which fall into this category.

These studies have demonstrated a high degree of molecular heterogeneity in PKU consistent with the high phenotypic heterogeneity of the disease. This heterogeneity is also reflected in my results from a South-East of England population where 12 different mutations have been identified (section 3.3, Table 3.3.1), ten loss of function mutations, one synonymous change and one intron polymorphism. The former account for 42% of mutated chromosomes in the population sample of 50 unrelated PKU patients studied.

These loss of function mutations are expected to impair PAH activity by several mechanisms: missense mutations presumably affecting the activity of the PAH protein (I65T, K341R, L347F, R408W), mutations at different intron-exon boundaries that impair normal RNA splicing (IVS10nt-11, IVS10nt-3, IVS12nt1), nonsense (Y356X) and frameshift mutations (delta K 363) codons which alter the amount of the functional protein and deletions which result in the truncation of the protein (delta ATC<sup>94,95</sup>). The synonymous nucleotide change resulting in a substitution of a G to C change at amino acid position 385 is a common polymorphism which has previously been observed in the caucasian population (Kalaydjieva *et al.*, 1991). The second polymorphism, the substitution of a C to T nucleotide at position 66 in intron 10, has not previously been reported.

Three novel loss of function mutations, K341R, L347F and IVS10nt-3c->t, have been identified in the study described here. As different screening approaches, in particular DGGE have now been used extensively in the study of PKU, over 100 novel mutations have been identified in the last two years (Guldberg *et al.*, 1993a, 1993b). It seems likely therefore, that these mutations are specific to the South-East England PKU population.

The identification of novel mutations is not surprising as this population had not previously been screened. As the frequency of each of these three mutations is only 1%, one mutated allele in the hundred tested, it seems unlikely that these mutations represent common alleles in the population. A comparison of the amino acids across the three mammalian species and *Drosophila* at the two amino acid positions (Fig 3.3.17) confirm that they are conserved during evolution which is consistent with their involvement in the cause of the disease.

It could be argued that these mutations may represent PCR errors incorporated during amplification. However as with all the mutations identified in the course of this study, their authenticity was confirmed by a repeat PCR and DGGE analysis to confirm that the band pattern of the clone and the mutated allele from the patient were identical. There was no evidence for any of the 12 mutations identified here that such misincorporation errors had occurred.

In a study analogous to that carried out by Condie *et al.*, (1993) and Theophilus *et al.*, (1989) (see section 4.2.3) the detection sensitivity of SSCP was compared with DGGE. The ability of SSCP to resolve the 12 mutations identified in this study was considered. Using a number of different electrophoretic and temperature parameters (section 2.2.7), SSCP succeeded in

resolving only 4 of the 12 mutations identified using DGGE (3.3.32).

This suggests that Phast SSCP system in my hands could only identify 33% of the mutation identified using DGGE. This may in part be due to the relatively large size of the products screened (Fig 3.3.1) which ranged in size from 251-339 bp somewhat greater than the optimal size for SSCP screening (Sheffield *et al.*,1993; Grompe, 1993). It is interesting that SSCP succeeded in resolving mutations in exons 10 and 12 which were amplified using the smaller PCR products (Table 3.3.2). As the PCR templates for SSCP, utilised the same primers as those used for DGGE analysis, it could also be envisaged that the presence of GC clamp at either end of the product would affect the sensitivity of the approach. However, as new primers were not synthesised, this possibility was not evaluated. The possibility that SSCP analysis could resolve mutations not identified using DGGE was also considered. Twenty-five patients which showed no mutation for exon 12 by DGGE were evaluated but no new mutations were identified (Fig 3.3.33).

#### **4.2.3.2 Population studies**

PKU is the most common autosomal recessively inherited disorder of amino acid metabolism with a frequency of approximately 1 in 10,000 and a carrier frequency of 1 in 50. This high carrier frequency suggests that the mutant chromosomes have been maintained because of some selective advantage to heterozygote individuals. A number of examples of environmental effects on gene frequency have been documented, for instance the high incidence of sickle cell anaemia in Africa results from a heterozygote advantage to

carrier individuals against malaria.

The nature of the heterozygote advantage in PKU, if indeed it exists, is not known. There is some evidence to suggest that carriers for the disease may have increased resistance to ochratoxin A, a mycotoxin found in fungus contaminated grains and lentils. Resistance against this compound, which is a known abortifacient, would increase the viability of offspring and the fertility of female carriers (Woo, 1989). As some mutations appear to predominate, for example: the R408W mutation in my study accounts for 14% of mutations, in the Danish population the IVS12nt1g->a accounts for 37% of mutations and in the Sicilian population the IVS10nt-11g->a accounts for 15% of mutations, the higher levels of these mutations may correspond to this selective advantage. Mutations which occur at much lower frequencies may retain no or an intermediary selective pressure and hence their reduced frequencies. This may be a phenomenon associated with other disorders, for example in Cystic Fibrosis the DF508 mutation is responsible for 70% of CF chromosomes (Dork *et al.*, 1994). Extensive examination of these changes which appear to allow some selective advantage may further our understanding on the mechanisms of environmental-genetic interactions.

Some authors have also suggested that hypermutability at CpG dinucleotides could account for the high frequency of mutant PAH chromosomes (Abadie *et al.*, 1989). This is difficult however to reconcile with the observations that only a quarter of the reported PAH mutations involve CpG dinucleotide mutational hot spots and most of these are not especially common among Caucasians (Eisensmith and Woo, 1991). Furthermore, as many other genes associated with less prevalent genetic disorders contain similar numbers of CpG

dinucleotides this fact alone cannot explain the prevalence of PKU. In my studies three out of the twelve mutations identified arose from mutations at CpG dinucleotide positions in agreement with these general findings.

The association with particular mutations and different mutant haplotypes suggests that there has been multiple founding populations for the PAH mutations. These observations suggest that several founder effects of the same mutation and heterozygote advantage may be the mechanisms that best explain the present distribution of mutant PAH chromosomes. Nevertheless it has been possible to utilise haplotype analysis to look at the historical demography of particular ethnic groups. It has been possible for example to trace a deletion mutation in exon 3 of the PAH gene which is observed in a large proportion of Yemenite Jews to a single ancestor who originated from Yemen and who's offspring later moved to Israel around 1762 (Avigad *et al.*,1990). It has also been possible to identify several space clusters of a particular founding mutation in both the French-Canadian and French populations for the MIV mutation (Lyonnet *et al.*,1992). Mutation studies can therefore provide an insight into historical demographic changes that have occurred.

#### **4.2.3.3 Clinical aspects of PKU mutation analysis**

The introduction of neonatal screening for PKU in the 1960s reflected the prevalence and severity of this disease. The optimisation of the diet in an effort to reduce hyperphenylalaninemia levels in patients has continued to cause problems, particularly as patients are now encouraged to remain on diet throughout their life. This has given rise to studies attempting to correlate the phenotype (level of

hyperphenylalaninemia off diet) and genotype of a patient to his or her dietary response. It is anticipated that by comparing the response of individuals to their diet and relating this to their genotype it will be possible to determine the optimum dietary restrictions for other patients with similar genetic backgrounds.

In practice this would allow the dietician to prescribe the most suitable diet for any given patient. The diet could therefore be more personalised right from the beginning. Before and throughout pregnancy it is also essential to control the phenylalaninemia to within the normal range. It is hoped that a comparison of the genotypes of these patients with their particular diets while pregnant will allow appropriate diets for other patients to be more easily devised.

In order to evaluate the impairment of a particular mutation on enzyme function, *in vitro* expression studies have also been carried out on over thirty different mutant enzymes (Flatmark *et al.*, 1994). This allows the affect of a particular mutation on the activity of the PAH protein to be quantified. Regardless of whether *in vitro* analysis truly reflects the conditions *in vivo*, the results thus far clearly indicate that these values are at least a relative reflection of the *in vivo* activity. Expression analyses can therefore be a valuable aid in the establishment of biochemical and clinical phenotypes in PKU patients.

As in a lot of genetic diseases, there is a demand for prenatal diagnosis and carrier detection services. The screening of all 13 exons of the PAH gene using DGGE makes this service a reality. In most instances parents who have one affected child would like to be offered a prenatal diagnosis when starting a second pregnancy. Alternatively, close relatives may wish to

know their carrier status. This would also mean that genetic counselling would be provided not in terms of probability but with a definite prognosis.

### **4.3 Mutation detection**

#### **4.3.1 Choosing a detection strategy**

Table 4.2 summarises the advantages and disadvantages observed for the most commonly used screening methods (see also section 1.2.3.3). As reviewed in Grompe, 1993 the requirements for mutation detection analyses will depend very much on whether a particular gene is known to be defective in a disease or being considered as a candidate. In the former case it would seem appropriate to develop a specialised set up to consider a large number of individuals at a high degree of accuracy. In the latter case, where a number of genes are being evaluated as candidates for the disease, the degree of accuracy of a technique can be compensated by the availability of enough patient samples and by the speed and simplicity of the approach.

It is therefore worth considering not only the sensitivity of a given approach but also other criteria for example: speed, simplicity, suitability for evaluating both DNA and RNA templates and limitations of available sequence, size and structural characterisation of the gene before deciding on the most appropriate technique to use.

#### **4.3.2 Applications of the different screening strategies**

Method	sensitivity	maximum size	localizes	toxic chemicals	mRNA scanning	best reference
SSCP	80%	250bp	no	none	+	Michaud <i>et al</i> , 1992
DGGE	95%	600	no	none	++	Sheffield <i>et al</i> , 1991
CMC	>95%	1700	yes	yes	+++	Rossiter <i>et al</i> , 1991
PCR.DS	>99%	500	yes	none	++	Bevan <i>et al</i> , 1992
HA	80%	300	no	none	+	White <i>et al</i> , 1992
RFLP	<50%	tens of Kb	yes	none	-	Botstein <i>et al</i> , 1980

Sensitivity : the approximate % of mutations detectable with the method is given.

Localizes : indicates whether a method provides the exact location of the mutation within the examined fragment.

+++ , indicates a high, ++ moderate and + limited usefulness of the method for this application.

**Table 4.2 Advantages and disadvantages of screening methods for detection of single base mutations.**

(R.R.H. Cotton, 1989), (M.Grompe, 1993).

Sequence analysis will always be required to confirm and characterise any mutation. Although it allows the greatest sensitivity for mutation detection it may not be suitable for screening a large number of patients particularly if the structure of the gene precludes its analysis to a small number of PCR products, for example the type VII collagen gene which is mutated in dystrophic epidermolysis bullosa and the fibrillin genes mutated in Marfan's syndrome have 118 (Christiano *et al.*, 1994) and 86 (Periera *et al.*, 1993) exons respectively.

The introduction of automated sequencers (Applied Biosystems ABI and Pharmacia ALF) which allow sequencing of approximately 0.8 kb of DNA at high fidelity has made this strategy more applicable as a mutation detection system. As large scale sequencing is not a cost effective approach for mutation detection most investigators choose to target the region containing the mutation prior to any sequence analysis.

Sequence analysis is I expect, the more appropriate method for mutation detection when only a few samples are to be screened and in cases where the gene is small and has a simple structure. Consequently it will be possible to sequence the entire gene in a minimum of reactions. Naturally in larger scale screening operations it may be more applicable to use a mutation detection approach where a distinctive phenotype in an individual can be correlated with a particular mutation if that phenotype has been observed previously and the mutation characterised. Although the molecular changes causing most disorders are very heterogeneous certain mutation types may predominate, for example although over 375 different mutations have been observed in the CFTR gene, the  $\Delta F508$  mutation is responsible for almost 70% of mutation cases (Mercier *et al.*, 1993). This is a phenomenon which is also

commonly observed in different PKU populations which have so far been analysed in depth (section 4.2.3.2).

The DGGE technique which gives good reproducibility and allows clear resolution of different mutations is applicable to such analyses. A distinctive banding pattern is observed for each mutation described in the studies presented here (section 3.1 and 3.3) and is a phenomenon characteristic of the technique (Fodde and Losekoot, 1994). In theory DGGE analysis of a heterozygote individual will generate four distinct bands corresponding to the two homoduplex and heteroduplex forms. The homoduplexes by definition will have a greater stability than heteroduplex forms and are therefore expected to resolve at a higher level of denaturant (Table 4.3). The occurrence of three distinct bands is explained by two possibilities. It may be that the homoduplexes can be resolved from each other but the two heteroduplexes resolve at the same position and are therefore indistinguishable. Alternatively the opposite may be the case and the two homoduplexes resolve at the same position. It follows that the existence of only two bands would suggest that neither the two homoduplex forms nor the two heteroduplex forms are resolvable from each other. It could also be envisaged that the two bands each represent a single homoduplex and that the heteroduplex forms are not resolvable by the gradient.

CMC analysis has the advantage of directing sequencing to a particular region of the DNA segment since the size of the cleavage product indicates the position of the mutation. Each variant will therefore show a characteristic banding pattern. Other screening strategies like SSCP analysis and in particular HA where the resolution between individual variants is adequate to confirm the presence of a mutation but not always to discriminate the mutation present suggests that

Number of bands	4	3		2	
		hypothesis 1	hypothesis 2	hypothesis 1	hypothesis 2
Homo(____)	-----				
and hetero					
(-- )duplex	-----	-----	-----		
distribution					
within the	-----	-----	-----	-----	-----
banding					
pattern	-----	-----	-----	-----	-----

**Table 4.3 Diagrammatical representation of the different types of DGGE banding patterns observed in the DGGE studies presented.**

these approaches are not as appropriate for large scale population screening.

In screening studies where several candidate genes need to be screened in a large patient panel to identify genes involved in a disease, the simplicity of the SSCP and HA protocols might overcome the disadvantages of reduced sensitivity and reproducibility. These approaches only require a specific PCR reaction in order to screen any DNA sequence whereas DGGE for instance, requires prior knowledge of the sequence to optimise protocols for screening. A review of the literature would suggest that, where limited sequence information is available, for example in the screening of candidate cDNAs the application of SSCP/HA is most commonly used. In these instances the loss in sensitivity is overcome by screening a relatively large number of unrelated patient samples.

As positional cloning approaches are limited by the obvious difficulties in identifying and screening genes (obtained from within the region of interest), the screening of candidate genes has grown in popularity. When a strong candidate gene exists eg the ROM1 as a cause of Best's macular degeneration (Nichols *et al.*, 1994) direct sequencing or the use of DGGE/CMC is far more appropriate in these instances than SSCP or HA. The almost 100% sensitivity of the former strategies is crucial as the gene can be discounted as being as strong candidate if no mutations are identified. As SSCP and HA are wholly empirical and have sensitivities at best of up to 80%, it would be more difficult to confidently exclude a good candidate on the absence of mutation using these approaches.

All these approaches are also applicable to the use of mRNA as the starting material for mutation detection. The advantages

and disadvantages of using DNA and RNA templates are summarised in Table 4.4 (taken from Grompe, 1993). The use of mRNA as starting template has the advantage that it allows a gene to be examined without prior information of its genetic structure. This allows the screening to be condensed into far fewer PCR reactions which is particularly advantageous where a gene contains a large number of introns.

Naturally, the use of mRNA screening will only be applicable to genes whose mRNA is easily accessible. In certain cases, protocols allow for the amplification of mRNA from tissues where the gene is expressed at a low level (illegitimate transcription) (Chelly *et al.*, 1989; Abadie *et al.*, 1993). However this usually requires the use of nested PCR and there is also the possibility of introducing PCR errors. It is therefore important to confirm any mutations subsequently, using genomic DNA. This strategy will of course direct genomic screening to a much smaller region.

The obvious drawback of mRNA analysis is that mutations resulting in altered splicing will generally not be identified. This is particularly important in diseases such as PKU where the number and occurrence of splice site mutations is quite high (section 4.2.3.2). Although, the use of genomic DNA overcomes many of these difficulties, the exon-intron structure of many genes is quite complicated. In addition to introducing a requirement for defining these boundaries a knowledge of flanking intronic sequence is necessary to allow primers to be designed from outside the coding and splice site junction sequences. Although genomic DNA is required to screen the promoter regions of genes it would appear that the prevalence of promoter mutations is fairly infrequent.

	Advantages	Disadvantages
Genomic DNA	<p>Easily accessible (blood)</p> <p>In autosomal loci both alleles are equally represented</p> <p>Mutations in the promotor and intronic splice junctions can be detected</p> <p>First choice in autosomal dominant traits</p>	<p>Genomic sequence and gene structure information are needed</p> <p>Only small segments of coding region (exons) are analysed</p> <p>More PCR reactions</p>
mRNA	<p>Long segments of peptide coding region can be analysed</p> <p>Gene structure information not needed</p> <p>Fewer PCR reactions</p> <p>Aberrant mRNA sizes can be seen</p> <p>First choice in X-linked traits</p>	<p>Gene may not be expressed in accessible specimens</p> <p>In autosomal loci only one allele may be represented</p> <p>Mutations in the promotor and intronic splice junctions are not detected</p>

**Table 4.4 Genomic DNA versus mRNA as starting material for mutation analysis. (Grompe, 1993)**

### 4.3.3 Studies evaluating different screening approaches

It is surprising that experimental comparisons of the efficiency of different strategies have not been evaluated in more detail. Indeed only two reported studies have attempted to evaluate the ability of different screening strategies to identify the same set of mutations.

The first study addressed the efficiency of mutation detection in patients affected with Gaucher disease using the DGGE, CMC and RNase screening strategies. Five mutations in the  $\beta$ -glucosidase gene were considered (Theophilus *et al.*, 1989). DGGE was the only method which detected all the mutations in the samples analysed. RNase protection was considered the easier approach although it was not as sensitive. It is worth bearing in mind that the use of RNA as a starting template is not always available or appropriate (section 4.3.2). They encountered difficulties in the use of osmium tetroxide to identify T mismatches by CMC. The use of CMC for subsequent study of a large number of other diseases would suggest that these problems can be overcome. It could also be argued that as only five mutations were screened in total it would be difficult to make any firm judgements on any of the techniques.

A similar report, describing mutation screening of patient samples with known mutations in the p53 gene using SSCP, DGGE and CMC has also been described (Condie *et al.*, 1993). In total the ability to identify twenty separate mutations were considered. They found that SSCP detected 90% of mutations while DGGE and CMC successfully identified all the mutations studied. They concluded that a 90% detection rate for SSCP was not a great disincentive for its use in view of the speed

and simplicity of the approach. The PCR fragment sizes used in the SSCP analysis however were small, between 200-300 bp. Recent studies have shown that as the size of the product increases the efficiency of SSCP is dramatically reduced (Sheffield *et al*, 1993).

There were difficulties encountered in optimising conditions for their detection of all twenty mutations by DGGE. It appears that the conditions initially used in this study were not optimally defined. In part this can be attributed to not using the SQHTX algorithm for defining optimal electrophoretic conditions. As suggested in a recent review by Fodde and Losekoot (1994) it would be helpful if the Lerman algorithms, especially the SQHTX were made more user friendly. An exhaustive analysis of the p53 gene coding sequence, using the full potential of the Lerman programs, has since successfully been completed by a different group (Pignon *et al.*, 1994). The CMC approach was also used very successfully. It appears that the relatively labour intensive and the hazardous nature of some of the chemicals used has limited the appeal of CMC for people requiring mutation detection analysis (Fodde and Losekoot, 1994). These disadvantages must however be weighed against the obvious strengths of the approach: its sensitivity, ability to screen fragments in excess of 1 kb and to pin point the region in which the sequence variation has occurred.

As the identification of disease causing genes is likely to rise considerably in the near future it is imperative that more studies evaluating the sensitivities of different approaches be undertaken. Presently there seems to be some confusion as to the most appropriate approach although as summarised above a lot depends on the purpose of a given study.

#### **4.4 Conclusion.**

It seems obvious that in order to develop novel approaches to reduce the affects of a disease or to entertain any likelihood of cure, our knowledge of the aberrant gene and its protein must be better understood. The use of mutation analysis is an essential component to advance our understanding of genetic disease.

In the same way as the use of dinucleotide repeats has revolutionised linkage analysis, mutation detection strategies have allowed the molecular lesions underlying genetic disease to be uncovered in a far more efficient manner than could have previously been envisaged. Naturally the most appropriate strategy will depend on a number of parameters as were discussed earlier. The strong theoretical background of DGGE however, which allows optimisation of experimental conditions has increased the popularity of this approach. As a result following its extensive usage in the last few years its suitability for mutation detection is now readily accepted. This is borne out by the studies presented here.

## Bibliography

Abadie, V., Lyonnet, S., Maurin, N., Berthelon, M., Caillaud, C., Giraud, F., Mattei, J.-F., Rey, J., Rey, F., Munnich A. (1989). CpG dinucleotides are mutation hot spots in phenylketonuria. *Genomics* 5: 936-939.

Abadie, V., Jaruzelska, J., Lyonnet, S., Millasseau, P., Berthelon, M., Rey, F., Munnich, A., Rey, J. (1993). Illegitimate transcription of the phenylalanine hydroxylase gene in lymphocytes for the identification of mutations in phenylketonuria. *Human Molecular Genetics* 2(1): 31-34.

Abrams, E. S., Murdaugh, S.E., Lerman, L.S. (1990). Comprehensive detection of single base changes in human genomic DNA using DGGE and a GC clamp. *Genomics* 7: 463-475.

Alexander, B., and Goldstein, R. (1953). Dual hemostatic defect in pseudohemophilia. *J. Clin. Invest.* 32: 551.

Aly, A. M., Higuchi, M., Kasper, C.K., Kazazian, H.H., Antonarakis, S.E., Hoyer, L.W. (1992). Haemophilia A due to mutations that create new N-glycosylation sites. *Proc. Natl. Acad. Sci* 89: 4933-4937.

Arber, W. (1974). DNA modification and restriction. *Proc. Nucleic. Acid. Res. Mol.Biol.* 14: 1-37.

Armstrong, M. D., Binkley, E.L. (1956). Studies on phenylketonuria. V. Observations on a new born infant with phenylketonuria. *Proc. Soc. Expt. Biol. Med.* 93: 418-420.

Artlich, A., Horn, M., Lorenz, B., Bhattacharya, S., Gal, S. (1992). Recurrent 3-bp deletion at codon 255/256 of the rhodopsin gene in a German pedigree with autosomal dominant retinitis pigmentosa. *Am.J.hum. Genet.* 50: 876-878.

Attree, O., Vivaud, D., Vivaud, M., Amselem, S., Lavergne, J.-M., Goossens, M. (1989). Mutations in the catalytic domain of human coagulation factor IX: rapid characterization by direct genomic sequencing of DNA fragments displaying an altered melting behavior. *Genomics* 4: 266-272.

Avigad, S., Cohen, B.E., Bauer, R., Schwartz, G., Frydman, M., Woo, S.L.C. (1990). A single origin of PKU in Yemenite Jews. *Nature* 344: 168-170.

Barbetti, F., Gejman, P.V., Taylor, S.I., Raben, N., Cama, A., Bonora, E., Pizzo, P., Moghetti, P., Muggeo, M., Roth, J. (1992). Detection of

## Bibliography

mutations in Insulin receptor gene by DGGE. *Diabetes* **41**: 408-415.

Beck, J. S., Kwitck A.E., Cogen P.H., Metzger A.K., Duyk G.M., Sheffield V.C. (1993). A denaturing gradient gel electrophoresis assay for sensitive detection of p53 mutations. *Hum. Genet.* **91**: 25-30.

Berg, M. A., Guevara-Aguirre, J., Rosenbloom, A.L., Rosenfeld, R.G., Francke, U. (1992). Mutation creating a new splice site in the growth hormone receptor gene of 37 Ecuadorean patients with Laron Syndrome. *Hum. Mut.* **1**: 24-34.

Borrensens, A. L., Hovig, E., Brogger, A., (1988). Detection of base mutations in genomic DNA using DGGE followed by transfer and hybridisation with gene-specific probes. *Mut. Res.* **202**: 77-83.

Botstein, D., White, R.L., Skolnick, M., Davies, R.W. (1980). Construction of a linkage map using restriction fragment length polymorphism. *Am. J. Hum. Genet.* **32**: 314-331.

Boyd (1961). Phenylketonuria, City of Birmingham screening survey. *Brit. Med. J.* **1**: 771-773.

Burmeister, M., G. diSibio, D. Cox and R. Myers (1991). Identification of polymorphisms by genomic denaturing gradient gel electrophoresis: application to the proximal region of human chromosome 21. *Nucleic Acids Res* **19**(7): 1475-1481.

Cai, S. P., Kan Y.W. (1990). Identification of the multiple beta-thalassemia mutations by DGGE. *J. Clin. Invest.* **85**: 550-553.

Caillaud, C., Lyonnet, S., Rey, F., Melle, D., Frebourg, T., Munnich, J., et al. (1991). A 3-base pair in-frame deletion of the PAH gene results in a kinetic variant of PKU. *J. Biol. Chem.* **266**: 9351-9354.

Cariello, N. F., Scott, J.K., Kat, A.G., Thilly, W.G., Keohavong, P. (1988). Resolution of a missense mutant in human genomic DNA by denaturing gradient electrophoresis and direct sequencing using in vitro DNA amplification: HPRT<sub>Munich</sub>. *Am J Hum Genet* **42**: 726-734.

Carter, N. P., Ferguson-Smith, M.A., Perryman, M.T., Telenius, H., Pelmear, A.H., Leversha, M.A., Glancy, M.T., Wood, S.L., Cook, K., Dyson, H.M., et al (1992). *J. Med. Genet.* **29**: 299-307.

## Bibliography

Catalano, M., Nobile, M., Novelli, E., Smeraldi, E. (1992). Use of PCR and DGGE to identify polymorphisms in three exons of dopamine D2 receptor gene in schizophrenic and delusional patients. *Neuropsychobiology* **26**: 1-3.

Chelly, J., Concordet, J.P., Kaplan, J.C., and Kahn, A., (1989). Illegitimate transcription: transcription of any gene in any cell type. *Proc. Nat. Acad. Sci* **86**: 2617-2621.

Christiano, A. M., Hoffman ,G.G., Chung-Honet, L. C., Seungbok, L., Cheng ,W., Uitto, J., Greenspan ,D.S. (1994). Structural organization of the human type VII collagen gene (COL7A1), composed of more exons than any previously characterised gene. *Genomics* **21**: 169-179.

Cocozza, S., Porcellini, A., Riccardi, G., Monticelli, A., Condorelli, G., Ferrar, A., Pianese, L., Miele, C., Capaldo, B., Beguinot, F. (1992). NIDDM associated with mutation in tyrosine kinase domain of insulin receptor gene. *Diabetes* **41**: 521-526.

Collins, F. (1992). Positional cloning. Lets not call it reverse anymore. *Nature Genet.* **1**: 3-6.

Condie, A., Eeles, R., Borrensens, A., Coles, C., Cooper, C., Prosser, J. (1993). Detection of point mutations in the p53 gene:comparison of SSCP,CDGE, Hydroxylamine osmium tetroxide techniques. *Human Mutation* **2**: 58-66.

Cooper, D. and H. Youssoufian (1988). The CpG dinucleotide and human genetic disease. *Hum Genet* **78**: 151-155.

Cotton, R. G. H., Rodrigues, N.R.,and Campbell, R.D. (1988). Reactivity of cytosine and thymidine in single base pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc. Natl. Acad. Sci. U.S.A.* **85**: 4397-4401.

Cotton, R. G. H. (1989). Detection of single base changes in nucleic acids. *Biochem.J.* **263**: 1-10.

Cotton, R. G. H. (1990). Heterogeneity of phenylketonuria at the clinical, protein and DNA levels. *J. Inherit. Metab. Dis.* **13**: 739-750.

Cotton, R. G. H. (1993). Current methods of mutation detection. *Mutation Research* **285**: 125-144.

## Bibliography

Daniels, G. ;. M., J., Anstee, D., Bird, G., Brodheim, E., et al., (1993). ISBT Working Party on Terminology for Red Cell Surface Antigens. *Vox Sanguinis* **65**: 77-80.

Darlington, G. H., Astrin, K.H., et al (1982). Assignment of human alpha-one-antitrypsin to chromosome 14 by somatic cell hybrid analysis. *Proc. Nat. Acad. Sci.* **79**: 870-873.

Dasovich, M., Konecki, D., Lichter-Konecki, U., Eisensmith, R., Guttler, F., Woo, S.L.C., et al. (1991). Molecular characterisation of PKU allele prevalent in Southern Europe and Ireland. *Somatic Cell. Mol. Genet.* **17**: 303-309 :

Dausset, J., Colombani, J., (1972). Histocompatibility testing.

De Bellis, A., Quigley, C.A., Cariello, N.F., El-Awady, M.K., Sar, M., Lane, M.V., Wilson, E.M., French, E.S. (1992). Single base mutation in the human androgen receptor gene causing complete androgen insensitivity: Rapid detection by a modified denaturing gradient gel electrophoresis technique. *Mol Endocrinol* **6**: 1909-1920.

Desbois, C., Magre, J., Blanquet, V., Capeau, J., Goossens, M., Besmond, C. (1993). Detection of sequence variations in the human Insulin-receptor gene by parallel DGGE. *Human Mutation* **2**: 395-403.

Devoto, M., Ronchetto, P., Fanen, P., Oriols, J.J., Romeo, G., Goossens, M., Ferrari, M., Magnani, C., Seia, M., Cremonesi, L. (1991). Screening for non-delta F508 mutations in five exons of the cystic fibrosis transmembrane conductance regulator (CFTR) gene in Italy. *Am. J. Hum. Genet.* **48**: 1127-1132.

DiLella, A., Kwok, S.C.M., Ledley, F.D., Marvit, J., Woo, S.L.C. (1986). Molecular structure and polymorphic map of the human phenylalanine hydroxylase gene. *Biochemistry* **25**(4): 733-749.

DiLella, A. G., Marvit, J., Lidsky, A.S., Guttler, F., Woo, S.L.C. (1986). Tight linkage between a splicing mutation and a specific DNA haplotype in phenylketonuria. *Nature* **322**: 799-803.

Dillela, A., Marvit, J., Brayton, K., and Woo, S.L.C. (1987). An amino-acid substitution involved in PKU is in linkage disequilibrium with DNA haplotype 2. *Nature* **327**: 333-336.

## Bibliography

Dockhorn-Dworniczak, B., B. Dworniczak, L. Brömmelkamp, J. Bülles, J. Horst and W. W. Böcker (1991). Non-isotopic detection of single strand conformation polymorphism (PCR-SSCP): a rapid and sensitive technique in diagnosis of phenylketonuria. *Nucleic Acids Res* **19**(9): 2500.

Dork, T., Will, K., Grade, K., Krawczak, M., Tummeler, B. (1994). A 32-bp deletion (2991del32) in the Cystic Fibrosis gene associated with CFTR mRNA reduction. *Human Mutation* **4**: 65-70.

Dubel, J. R., Finwick, R., Hejtmancik, J.F. (1991). DGGE of the alpha-one-antitrypsin gene: Application to prenatal diagnosis. *Am J Med Genet* **41**: 39-43.

Dunn, J. M., Phillips, R.A., Zhu, X., Becker, A., Gallie, B.L. (1989). Mutations in the RB1 gene and their effects on transcription. *Mol Cell Biol* **9**: 4596-604.

Dworniczak, B., Aulehla-Scholz, C., Kalaydjieva, L., Bartholome, K., Grudde, K., and Horst, J. (1991). Aberrant splicing of PAH mRNA: The major cause for PKU in parts of Southern Europe. *Genomics* **11**: 242-246.

Eisensmith, R. C. and S. L. C. Woo (1991). Phenylketonuria and the phenylalanine hydroxylase gene. *Mol Biol Med* **8**: 3-18.

Fagerhol, M. K., and, Cox D.W. (1981). The Pi polymorphism: Genetic biochemical and clinical aspects of alpha-1-antitrypsin. New York and London, Plenum Press.

Fanen, P., Ghanem, N., Vidaud, M., Besmond, C., Martin, J., Costes, B., Plassa, F., Goossens, M. (1992). Molecular characterisation of cystic fibrosis: 16 novel mutations identified by analysis of the whole cystic fibrosis conductance transmembrane regulator (CFTR) coding regions and splice site junctions. *Genomics* **13**: 770-774.

Ferec, C., Audrezet, M.P., Mercier, B., Guillermit, H., Moullier, P., Quere, I., Verlingue, C. (1992). Detection of over 98% cystic fibrosis mutations in a celtic population. *Nature Genetics* **1**: 188-191.

Finkelstein, J. E., Francomano, C.A., Brusilow, S.W., Traystman, M.D. (1990). Use of denaturing gradient gel electrophoresis for detection of

## Bibliography

mutation and prospective diagnosis in late onset ornithine transcarbamylase deficiency. *Genomics* **7**: 167-172.

Fischer, S. G., Lerman, L.S. (1979). Two-dimensional electrophoresis separation of restriction enzymes fragments of DNA. *Methods in Enzymology* **68**: 183-191.

Fischer, S. G., Lerman, L.S. (1983). DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci USA* **80**: 1579-1583.

Flatmark, T., Knappskog, P., Martinez, A., Doskeland, A.P., Apold, J. (1994). Characterisation of the wild-type and mutant forms of the recombinant human phenylalanine hydroxylase.

Fodde, R. (1992). Eight novel inactivating germ line mutations at the APC gene identified by DGGE. *Genomics* **13**: 1162-1168.

Fodde, R., and Losekoot, M. (1994). Mutation detection by Denaturing Gradient Gel Electrophoresis. (DGGE). *Human Mutation* **3**: 83-94.

Folling A. (1934). Über Ausscheidung von Phenylbrenztraubensaure in den Harn als Stoffwechselanomalie in Verbindung mit Imbezilitat. *Z Physiol Chem* **227**: 169-176.

Freneaux, E., Sheffield, V.C., Molin, L., Shires, A., Rhead, W.J. (1992). Heterogeneity in beta-oxidation flux, polypeptide synthesis and complementarity DNA mutations in the subunit of electron transfer flavoprotein in eight patients. *J. Clin. Invest.* **90**: 1679-1686.

Garland, H. G. (1934). Hereditary scoliosis. *Brit. Med. J.* **1**: 328.

Garrod, A. E. (1902). The incidence of alkaptonuria: a study in clinical individuality. *Lancet* : 1616-1620.

Gibbs, N. K., Woolf, L.I. (1959). Tests for phenylketonuria, results of a one year programme for its detection in infancy and among mental defectives. *Brit. Med. J.* **2**: 532-535.

Gibbs, R. A., and, Caskey, C.T. (1987). Identification and localisation of mutations at the Lesch-Nyhan locus by ribonuclease A cleavage. *Science* **236**: 303-305.

## Bibliography

Gottardi, E., Losekoot, M., Fodde, R., Saglio, G., Camaschella, C., Bernini, L.F., (1991). Rapid identification by DGGE of mutations in the gamma-globin gene promoters in non-deletion type HPFH. *Br. J. Haematol* **80**: 533-538.

Gray, M. R. (1992). Detection of DNA sequence polymorphisms in human genomic DNA by using denaturing gradient gel blots. *Am.J.Hum.Genet.* **50**: 331-346.

Grompe, M. (1993). The rapid detection of unknown mutations in nucleic acids. *Nature Genetics* **5**: 111-117.

Gu, X., de Rooij, F., Voortman, G., Velde, K.T., Deybach, J.C., Nordmann, Y., Grandchamp, B. (1994). Detection of eleven mutations causing acute intermittent porphyria using denaturing gradient gel electrophoresis. *Human Genetics* **93**: 47-52.

Guldborg, P., Lou, H. C., Henriksen, K.F., Mikkelsen, I., Olsen, B., Holck, B., Guttler, F. (1993). A novel missense mutation in the phenylalanine hydroxylase gene of a homozygous Pakistani patient with non-PKU hyperphenylalaninemia. *Human Molecular Genetics* **2**(7): 1061-1062.

Guldborg, P., Henriksen, K. F., Guttler, F. (1993a). Molecular analysis of Phenylketonuria in Denmark : 99% of the mutations detected by Denaturing Gradient Gel Electrophoresis. *Genomics* **17**: 141-146.

Guldborg, P., Romano, V., Ceratto, N., Bosco, P., Cinna, M., Guttler, F., et al (1993b). Mutational spectrum of PAH deficiency in Sicily. *Human Molecular Genetics* **2**(10): 1703-1707.

Guthrie, R. (1961). Blood screening for phenylketonuria. *J. Amer. Med. Assoc.* **178**: 338-343.

Guthrie, R., Susie, A. (1963). A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics* **32**: 338-343.

Guttler, F. (1980). Hyperphenylalaninemia: diagnosis and classification of the various types of phenylalanine hydroxylase deficiency in childhood. *Acta Pediat. Scand* **280**: 1-80.

Harris, H., Hopkinson, D.A. (1972). Average heterozygosity per locus in

## Bibliography

man: an estimate based on the incidence of enzyme polymorphisms. *Ann. Hum. Genet., London* **36**: 9.

Harris, H. (1975). The principles of Human Biochemical Genetics. North-Holland/ American Elsevier.

Heidtmann, H., Travis, J., (1986). Human alpha one proteinase inhibitor. Proteinase inhibitors: research monographs in cell and tissue physiology. Amsterdam : Elsevier.

Higuchi, M., Antonarakis, S.E., Kasch, L., Oldenburg, J., Economou-Petersen, E., Olek, K., Arai, M., Inaba, H., Kazanian, H.H. (1991). Molecular characterisation of mild to moderate haemophilia A: Detection of the mutation in 25 of 29 patients by DGGE. *Proc. Natl. Acad. Sci.* **88**: 8307-8311.

Hovig, E. (1991). Constant denaturant gel electrophoresis, a modification of denaturing gradient gel electrophoresis, in mutation detection. *Mutation research* **262**: 63-71.

Hovnanian, A., Duquesnoy, P., Amselem, S., Blanchet-Bardon, C., Lathrop, M., Dubertret, L., Goossens, M. (1991). Exclusion of linkage between the collagenase gene and generalised recessive dystrophic epidermolysis bullosa phenotype. *J. Clin. Invest.* **88**: 1716-1721.

Jakubiczka, S., Arnemann, J., Cooke, H. J., Krawczak, M., and Schmidtke, J., (1989). A search for restriction fragment length polymorphism on the human Y chromosome. *Hum. Genet.* **84**: 86-88.

Janocha, S., Wolz, W., Srsen, S., Srsnova, K., Montagutelli, X., Guenet, J.-L., Grimm, T., Kress, W., Muller, C.R., (1994). The human gene for Alkaptonuria (AKU) maps to chromosome 3q. *Genomics* **19**: 5-8.

Jeffreys, A. J., Wilson, V., Thein, S.L. (1985). Hypervariable "minisatellite" regions in human DNA. *Nature* **314**: 67-73.

Jervis, G. A. (1947). Studies on phenylpyruvic oligophrenia. *J. Biol. Chem.* **169**: 651-656.

Jobling, M. A. (1994). A survey of long-range DNA polymorphisms on the human Y chromosome. *Human Molecular Genetics* **3**(1): 107-114.

John, S., Scriver, C., Laframboise, R., Rozen, R. (1992). In vitro and in

## Bibliography

vivo correlations for I65T and MIV mutations at the phenylalanine hydroxylase locus. *Hum. Mutat.* **1**: 147-153.

Johnson, P., H. Cadiou and D. Hopkinson (1991). Detection of the common alpha-1-antitrypsin variants by denaturing gradient gel electrophoresis. *Ann Hum Genet* **55**: 183-198.

Johnson, P. H., Richards, A.J., Pope, F.M., Hopkinson, D.A. (1992). Analysis of mutations in Ehlers-Danlos syndrome type IV (EDS IV) by denaturing gradient gel electrophoresis (DGGE). *J Inherit Metabol Dis* **15**: 426-430.

Kalaydjieva, L., B. Dworniczak, C. Aulehla-Scholz, I. Kremensky, J. Bronzova, A. Eigel and J. Horst (1990). Classical phenylketonuria in Bulgaria: RFLP haplotypes and frequency of the major mutations. *J Med Genet* **27**: 742-745.

Kalaydjieva, L., B. Dworniczak, V. Kucinskas, V. Yurgeliavicius, E. Kunert and J. Horst (1991). Geographical distribution gradients of the major PKU mutations and the linked haplotypes. *Hum Genet* **86**: 411-413.

Kan, Y. M., and Dozy, A.M. (1978). Polymorphism of the DNA adjacent to the beta-globin gene region: relationship to sickle mutation. *Proc. Natl. Acad. Sci. USA* **75**: 5631-5635.

Kaufman, S. (1959). Studies on the mechanism of the enzymatic conversion of phenylalanine to tyrosine. *J.Biol. Chem.* **234**: 2677-2682.

Kaufman, S. (1963). The structure of the phenylalanine-hydroxylation cofactor. *Proc. Nat. Acad. Sci.* **50**: 1085-1093.

Kaufman, S. (1976). Phenylketonuria: biochemical mechanisms. *Adv. Neurochem.* **2**: 1-132.

Keen, J., Lester, D., Inglehearn, C., Curtis, A., Bhattacharya, S. (1991). Rapid detection of single base mismatches as heteroduplexes on Hydrolink gels. *Trends Genet.* **7**: 5.

Kidd, V. J., Wallace, R.B., Itakura, K., Woo, S.L.C. (1983). alpha-one-antitrypsin deficiency detection by direct analysis of the mutation in the gene. *Nature* **304**: 130-234.

Knight, S., Flannery, A., Hirst M., Davies, K., et al, (1993). FRAXE mental

## Bibliography

retardation associated with trinucleotide repeat expansion and hypermethylation of a CpG island in XQ28. *Am J Hum Genet* **53**(3): 79.

Kogan, S. and J. Gitschier (1990). Mutations and a polymorphism in the factor VIII gene discovered by denaturing gradient gel electrophoresis. *Proc Natl Acad Sci USA* **87**: 2092-2096.

Konecki, D. S. and U. Lichter-Konecki (1991). The phenylketonuria locus: current knowledge about alleles and mutations of the phenylalanine hydroxylase gene in various populations. *Hum Genet* **87**: 377-388.

Konecki, D. S., M. Schlotter, F. Trefz and U. Lichter-Konecki (1991). The identification of two mis-sense mutations at the PAH gene locus in a Turkish patient with phenylketonuria. *Hum Genet* **87**: 389-393.

Krolewski, A. S., Krolewski, B., Gray, M., Stanton, V., Warram, J.H., Housman, D. (1992). High-frequency DNA sequence polymorphisms in the Insulin receptor gene detected by denaturing gradient gel blots. *Genomics* **12**: 705-709.

Kwok, S. C. M., Ledley, F.D., Dilella, A.G., Robson, K.J.H., Woo, S.L.C. (1985). Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalanine hydroxylase. *Biochemistry* **24**: 556-561.

La Spada, A. R., Wilson, E.A., Lubahn, D.B., Harding, A.E., Fischbeck, K.H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**: 677-79.

Lahiri, D. K., and Nurnberger, Jr. (1991). A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucl. Acid. Res.* **19**(19): 5444.

Landsteiner (1900). *Zbl. Bakt.* **27**: 357-362.

Landsteiner (1901). *Wien. Klin. Wschr.*, **14**,: 1132-1134.

Laubscher, K. H., Glew, R.H., Lee, R.E., Okinaka, R.T. (1994). Use of denaturing gradient gel electrophoresis to identify mutant sequences in the beta-glucocerebrosidase gene. *Human mutation* **3**: 411-415.

Lerman, L. S., Silverstein, K. (1987). Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis.

## Bibliography

*Methods in Enzymol* **155**: 482-501.

Levitt, R. C. (1991). Polymorphisms in the transcribed 3' untranslated region of eukaryotic genes. *Genomics* **11**: 484-489.

Lidsky, A. S., Ledley, F.D., Dilella, A.G., Kwok, S.C.M., Daiger, S.P., Robson, K.J.H., Woo, S.L.C. (1985). Extensive restriction site polymorphism at the human phenylalanine hydroxylase locus and application in prenatal diagnosis of PKU. *Am. J. Hum. Genet.* **37**: 619-634.

Litt, M., and Luty, J.A., (1989). A hypervariable microsatellite revealed by in vitro amplification of dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44**: 397-401.

Losekoot, M., Fodde, R., Harteveld, C.L., van Heeren, H., Giordano, P.C., Bernini, L.F. (1990). DGGE and direct sequencing of PCR amplified genomic DNA: A rapid and reliable diagnostic approach to beta-thalassemia. *Br. J. Haematol.* **76**: 269-274.

Losekoot, M., Fodde, R., Harteveld, C.L., van Heeren, H., Giordano, P.C., Bernini, L.F. (1991). Homozygous beta thalassaemia owing to a mutation in the cleavage-polyadenylation sequence of the human beta-globin gene. *J. Med. Genet.* **28**: 252-255.

Losekoot, M., van Heeren, H., Schipper, J.J., Giordano, P.C., Bernini, L.F., Fodde, R. (1992). Rapid detection of highly polymorphic fragment in the beta-globin gene by DGGE of PCR amplified material. *J. Med. Genet.* **29**: 574-577.

Lyonnet, S., C. Caillaud, F. Rey, M. Berthelon, J. Frézal, J. Rey and A. Munnich (1989). Molecular genetics of phenylketonuria in mediterranean countries: a mutation associated with partial phenylalanine hydroxylase deficiency. *Am J Hum Genet* **44**: 511-517.

Lyonnet, S., Melle, D., Debraekeleer, M., Rey, J., Munnich, A., Scriver, C.R., (1992). Time and space clusters of the french-canadien MIV PKU mutation in France. *Am J Hum Genet* **51**(1): 191-196.

Mahadevan, M. S., Amemiya, C., Jansen, G., Sabourin, L., Baird, S., Neville, C.E., Wormskamp, N., Segers, B., Batzer, M., Lamerdin, J., et al (1993). Structure and genomic sequence of the myotonic dystrophy (DM) kinase gene. *Hum. Mol. Genet.* **2**: 299-304.

## Bibliography

Malaspina, P., Persichetti, F., Novelletto, A., Iodice, C., Terrenato, L., Wolfe, J., Ferraro, M., Prantero, G. (1990). The human Y chromosome shows a low level of DNA polymorphism. *Ann. Hum. Genet.* **54**: 297-305.

Maniatis, T. (1991). Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press.

Mashima, Y., Shiono, T., Inana G., (1994). Rapid and efficient molecular analysis of Gyrate Atrophy using DGGE. *Investigative Ophthalmology and Visual Science* **35**(3): 1065-1070.

Mathias, N., Bayes, M., Tyler-Smith, C. (1994). Highly informative compound haplotypes for the human Y chromosome. *Hum. Mol. Genet.* **3**(1): 115-123.

Mercier, B., Lissens, W., Audrezet, M.P., Bonduelle, M., Liebaers, I., Ferec, C. (1993). Detection of more than 94% cystic fibrosis mutations in a sample of Belgian population and identification of four novel mutations. *Human Mutation* **2**: 16-20.

Mitoma, C. (1956). Studies on partially purified phenylalanine hydroxylase. *Arch. Biochem. Biophys.* **60**: 476-484.

Mullis, K. B., Faloona, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. *Meth. Enzymol.* **155**: 335-350.

Myers, R. M., Larin, Z., Maniatis, T., (1985). Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA: DNA duplexes. *Science* **230**: 1242.

Myers, R. M., Maniatis, T., Lerman, L.S. (1987). Detection and localisation of single base changes by denaturing gradient gel electrophoresis. *Methods in Enzymol* **155**: 501.

Nagafuchi, S., Yanagisawa, H., Sato, K., Shirayama, T., Ohsaki, E., Bundo, M., Yamada, M., et al., (1994). Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nature Genetics* **6**: 14-16.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culverm, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., White, R. (1987). Variable number of tandem repeats (VNTR) markers for human gene mapping. *Science* **235**: 1616-1622.

## Bibliography

Newmark, J. R., Hardy, D.O., Tonb, D.C., Carter, B.S., Epstein, J.I., Isaacs, W.B., Brown, T.R., Barrack, E.R. (1992). Androgen receptor gene mutations in human prostate cancer. *Proc Natl Acad Sci USA* **89**: 6319-6323.

Nichols, B. E., Bascom, R., Litt, M., McInnes, R., Sheffield, V.C., Stone, E. M. (1994). Refining the locus for Best vitelliform macular dystrophy and mutation analysis of the candidate gene ROM1. *Am. J. Hum. Genet.* **54**: 95-103.

Okano, Y., Hase, Y., Shintaku, H., Araki, K., Furuyama, J., Oura, T., Isshiki, G. (1994). Molecular characterization of phenylketonuric mutations in Japanese by analysis of phenylalanine hydroxylase mRNA from lymphoblasts. *Human Molecular Genetics* **3**(4): 659-660.

Olschwang, S., Fabre, R., Laurent-Puig, P., Vassal, A., Hamelin, B., Nakamura, Y., Thomas, G. (1992). Detection by DGGE of a new polymorphism closely linked to the adenomatous polyposis coli region. *Genet* **88**: 658-660.

Orita, M., Suzuki, Y., Sekiya, T., Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the PCR. *Genomics* **5**: 874-879.

Orkin, S. H. (1986). Reverse genetics and human disease. *Cell* **47**: 845-850.

Orr, H., Chung, M., Banfi, S., Zoghbi H., et al., (1993). Expansion of an unstable trinucleotide CAG repeat in Spinocerebellar ataxia type -1. *Nature genetics* **4**(3): 221-226.

Pauling, L., Itano, H. A., Singer, S.J. and Wells, I.C. (1949). Sick cell anaemia, a molecular disease. *Science* **110**: 543.

Pellegata, N. S., Losekoot, M., Fodde, R., Pugliese, V., Caccomanno, S., Renault, B., Bernini, L.F., Ranzani, G.N. (1992). Detection of k-ras mutations by DGGE: A study on pancreatic cancer. *Anticancer Res* **12**: 1731-1736.

Penrose, L. S. (1935). Inheritance of phenylpyruvic amentia (PKU). *Lancet* **II**: 192-194.

## Bibliography

Pereira, L., D'Alessio, M., Ramirez, F., Lynch, J. R., Sykes, B., Pangilinan, T., Bonadio, J. (1993). Genomic organization of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Human Molecular Genetics* 2(7): 961-968.

Pignon, J. M., Vinatier, I., Fanen, P., Jonveaux, P., Tournilhac, O., Imbert, M., Rochant, H., and Goossens, M. (1994). Exhaustive analysis of the p53 gene coding sequence by Denaturing Gradient Gel Electrophoresis : Application to the detection of point mutations in acute leukemias. *Human Mutation* 3: 126-132.

Renault, B., van den Broek, M., Fodde, R., Wijnen, J., Pellegata, N.S., Amadori, D., Meera Khan, P., Ranzani, G.N. (1993). Base transitions are the most frequent genetic changes at p53 in gastric cancer. *Cancer Res.* 53: 2614-2617.

Ribba, A. S., Lavergne, J.M., Bahnak, B.R., Derlon, A., Pietu, G., Meyer, D. (1991). Duplication of a methionine within the glycoprotein Ib binding domain of von Willebrand factor detected by DGGE in a patient with type IIB von Willebrand disease. *Blood* 78: 1738-1743.

Riesner, D., Steger, G., Zimmat, R., Owens, R.A., Wagenhofer, M., Hillen, W., Vollbach, S., and Henco, K. (1989). TGGE of nucleic acids: analysis of conformational transitions, sequence variations and protein-nucleic acid interactions. *Electrophoresis* 10: 377-389.

Riess, O., Michel, A., Speer, A., Cobet, G., Coutelle, C. (1987). Introduction of genomic diagnosis of classical PKU to the health care system of the German Democratic Republic. *Clinic. Genet.* 32: 209-215.

Rosatelli, M. C., Dozy, A., Faa, V., Meloni, A., Sardu, R., Saba, L., Kan, Y.W., Cao, A. (1992). Molecular characterisation of beta-thalassemia in the Sardinian population. *Am. J. Hum. Genet.* 50: 422-426.

Royle, N. J., Clarkson, R.E., Wong, Z., Jeffreys, A.J. (1990). Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. *Genomics* 3: 352-360.

Rubbrecht, O. (1910). L'origine du type familial de la maison de Hapsbourg. *Bruxelles Med.* :

Sanger, F., Nicklen, S., and Coulson, A.R., (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acd. Sci. U.S.A.* 74: 5463-5467.

## Bibliography

Saris, J. J., Breuning, M.H., Dauwerse, H.G., Snijdwint, F.G.M., Top, B., Fodde, R., van Ommen, G.J.B. (1990). Rapid detection of polymorphism near gene for adult polycystic kidney disease. *Lancet* **335**: 1102-1103.

Scriber, C. R., Kaufman, S., Woo, S.L.C. (1989). The hyperphenylalaninemias. The Metabolic Basis of Inherited Disease. New York, McGraw-Hill. 495-546.

Sheffield, V. C., Cox, D. R., Lerman, L. S., Myers, R.M. (1989). Attachment of a 40-base pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *PNAS* **86**: 232-236.

Sheffield, V. C., Fishman, G.A., Beck, J.S., Kimura, A.E., Stone, E.M. (1991). Identification of novel rhodopsin mutations associated with retinitis pigmentosa by GC-clamped denaturing gradient gel electrophoresis. *Am J Hum Genet* **49**: 699-706.

Sheffield, V. C., Beck, J.S., Kwitek, A.E., Sandstrom, D.W., Stone, E.M. (1993). The sensitivity of SSCP analysis for the detection of single base substitutions. *Genomics* **16**: 325-332.

Snell, G. D. (1948). *J.Genet* **49**: 87-108.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Molec. Biol.* **98**: 503-517.

Spurdle, A., Hammer, M., and Jenkins, T., (1994). The Y-Alu polymorphism in Southern African populations and its relationship to other Y-specific polymorphisms. *Am J Hum Genet* **54**(2): 319-330.

Stallings, R. L., Torney, D.C., Hildebrand, C.E., Longmire, J.L., Deaven, L.L., Jett, J.H., Doggett, N.A., Moyzis, R.K. (1991). Physical mapping of human chromosomes by repetitive sequences fingerprinting. *Proc. Natl. Acad. Sci. USA* **87**: 6218-6222.

Takeda, K., Weiss, R.E., Reteroff, S. (1992). Rapid localisation of mutations in the thyroid hormone receptor beta gene by DGGE in 18 families with thyroid hormon resistance. *J. Clin. Endocrinol. Metab.* **74**: 712-719.

## Bibliography

Tassabehji, M., Read, A.P., Newton, V.E., Harris, R., Ballings, R., Gruss, P., Stractan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the PAX-3 paired box gene. *Nature* **355**: 635-636.

Teraski, P. (1970). Histocompatibility testing.

The Huntington's Disease Collaborative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**: 971-983.

Theophilus, B. D. M., Latham, T., Grabowski, G.A., Smith, F.I. (1989). Comparison of RNase A, chemical cleavage, and GC-clamped denaturing gradient gel electrophoresis for the detection of mutations in exon 9 of the human acid B- glucosidase gene. *Nucl Acids Res* **17**: 7707-7722.

Top, B., Uitterlinden, A.G., Van der Zee, A., Kastelein, J.J.P., Gevers Leuven, J.A., Havekes, L.M., Frants, R.R. (1992). Absence of mutations in the promotor region of the low density lipoprotein receptor gene in a large number of familial hypercholesterolaemia patients as revealed by DGGE. *Hum. Genet.* **89**: 561-565.

Top, B. (1993). Identification of a splice site-mutation in the low density lipoprotein receptor gene by DGGE. *Human Genetics* **91**: 480-484.

Traystman, M. D., Higushi, M., Kasper, C.K., Antonarakis, S.E., Kazazian, H.H. (1990). Use of denaturing gradient gel electrophoresis to detect point mutations in the factor VIII gene. *Genomics* **6**: 293-301.

Tsukahara, M., Maisuura, S., Kishi, F., Yoshida, A., Kajh, F., (1990). Isolation of a Y chromosomal DNA sequence and its clinical application. *Jpn. J. Human Genet.* **35**(331-339):

Udenfriend, S., and Cooper, J.R. (1952). The enzymatic conversion of phenylalanine to tyrosine. *J. Biol. Chem.* **194**: 503-511.

Valero, M. C., Velasco, E., Moreno, F., Hernandez-Chico C. (1994). Characterization of four mutations in the neurofibromatosis type 1 gene by DGGE. *Human Molecular Genetics* **3**(4): 639-641.

van den Broek, M. H., Renault, B., Fodde, R., Verspaget, H., Griffioen, G., Meera Khan, P. (1993). Sites and types of p53 mutations in an

## Bibliography

unselected serie of colorectal cancers in The Netherlands. *Anticancer Res* **13**: 587-592.

van den Maagdenberg, A. M. J. M., Weng, W., de Bruijn, I.H., de Knijff, P., Funke, H., Smelt, A.H.M, GeversLeuven, J.A., van't Hooft, F.M., Assmann, G., Hofker, M.H., Havekes, L.M., Frants, R.R. (1993). Characterisation of five new mutants in the carboxylterminal domain of human Apolipoprotein E. *Am. J. Hum. Genet.* **52**: 937-946.

Vetrie, D., Boye, E., Flinter, F., Bobrow, M., Harris, A. (1992). DNA rearrangements in the alpha 5 (IV) Collagen Gene (Col4A%) of individuals with Alport Syndrome: Further refinement using pulsed-field gel electrophoresis. *Genomics* **14**: 624-633.

Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammerstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C.I.D., Bentley, D.R. (1993). The gene involved in X-linked agammaglobulinaemia is a member of the src family of protein-tyrosine kinases. *Nature* **361**: 226-233.

Vidaud, M., Fanen, P., Martin, J., Ghanem, N., Nicolas, S., Goossens, M. (1990). Three point mutations in the CFTR gene in French cystic fibrosis patients: identification by denaturing gradient gel electrophoresis. *Hum Genet* **85**: 446-449.

Vidaud D. (1993). Nucleotide substitutions at the -6position in the promotor region of the factor IX gene result in different severity of hemophilia B Leyden: consequences for genetic counseling. *Human Genetics* **91**: 241-244.

Vidaud M. (1990). Frequency of the cystic fibrosis delta F508 mutation in a large sample of the french population./DGGE. *Human Genetics* **85**: 391-445.

Wallis, Y., MacDonald, F., Rindl, P.M., Hulten, M., McKeown, C., Morton, D.G., Keighley, M.R.B., Fodde, R., Van der Luijt, R., Meera, Khan, P. (1992). Germline APC mutation responsible for familial adenomatous polyposis in Indian family. *Lancet* **340**: 1035.

Weatherall, D. (1985). The new Genetics and clinical practice.

Weber, J. L., and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymorphism chain

## Bibliography

reaction. *Am. J. Hum. Genet.* **44**: 388-396.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G., Lathrop, M., (1992). A second-generation linkage map of the human genome. *Nature* **359**: 794-801.

White, M. B., Carvalho, M., Derse, D., O'Brien, S.J., Dean, M. (1992). Detecting single base substitutions as a heteroduplex polymorphisms. *Genomics* **12**: 301-306.

Whitehouse, D. B., Lovegrove, J.U., Hopkinson, D.A. (1989). Variation in alpha-one-antitrypsin phenotypes associated with penicillamine therapy. *Clin. Chim. Acta* **179**: 109-116.

Winter, E., Yamamoto, F., Almoguera, C., Perucho, M. (1985). A method to detect and characterise point mutations in transcribed genes, amplification and overexpression of the mutant c-Ki-ras allele in human tumor cells. *Proc. Natl. Acad. Sci. USA* **82**: 7575-7579.

Woo, S. L. C. (1989). Molecular basis and population genetics of phenylketonuria. *Biochemistry* **28**: 1-7.

Woolf, L. I. (1986). The heterozygote advantage in PKU. *Am. J. Hum. Genet.* **38**: 773-775.

Wu, D. Y., Wallace, R.B., (1989). The ligation amplification reaction (LAR) amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics* **4**: 560-9.

Wyman, A. R., and White, R. (1980). A highly polymorphic locus in human DNA. *Proc. Natl. Acad. Sci. U.S.A.* **77**: 6754-6758.

Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., (1991). Fragile X genotype characterised by an unstable region of DNA. *Science* **252**: 1179-1181.