



<http://researchspace.auckland.ac.nz>

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.

<http://researchspace.auckland.ac.nz/feedback>

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form.

**MOLECULAR GENETICS OF RESTRICTION FRAGMENT LENGTH
POLYMORPHISMS LINKED TO THE HUNTINGTON DISEASE LOCUS**

Patricia M. Stapleton
Department of Cellular and Molecular Biology
University of Auckland

Thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
November 1988

UNIVERSITY OF AUCKLAND LIBRARY
BIOLOGY

THESIS

THESIS

89-13

cop. 2

TABLE OF CONTENTS

PAGE

ACKNOWLEDGEMENTS

ABSTRACT

ABBREVIATIONS

CHAPTER ONE INTRODUCTION

1.1 GENERAL INTRODUCTION	1
1.2 RESTRICTION FRAGMENT LENGTH POLYMORPHISMS	3
1.2.1 Detection of Restriction Fragment Length Polymorphisms	3
a) Probes of unique sequence	3
b) Probes of repeat sequence	4
1.3 THE USE OF RFLPS IN THE LOCATION OF DISEASE GENES	5
1.3.1 Chromosomal Location of Monogenic Disease	6
1.3.2 Chromosomal Location of Polygenic Disease	8
1.4 RFLPS IN THE DIAGNOSIS OF GENETIC DISEASE	9
1.4.1 Direct Detection of Monogenic Disease	9
1.4.2 Indirect Detection of Monogenic Disease	10
a) The use of phenotypically neutral RFLPs from within disease loci and control regions	10
b) The use of RFLPs linked to disease loci	11
1.4.3 Indirect Detection of Polygenic Disease using Linked RFLPs	11
1.4.4 Direct Identification of Somatic Mutations	12
1.5 THE CONSTRUCTION OF A GENETIC LINKAGE MAP OF THE HUMAN GENOME USING RFLPS	12
1.6 OTHER USES FOR RFLPS	14
1.6.1 Individual Identifications	14
1.6.2 Evolutionary Relationships and Population Genetics	15
1.7 OTHER METHODS FOR THE DETECTION OF GENOMIC VARIABILITY	15
1.7.1 Differential Hybridisation using Oligonucleotide Probes	16

1.7.2 Differential DNA Melting During Denaturing Gradient Gel Electrophoresis	16
1.7.3 Electrophoresis of Partial RNA-DNA Hybrids	16
1.7.4 The Polymerase Chain Reaction	17
1.8 RFLPS LINKED TO THE HD LOCUS ON CHROMOSOME 4	18
1.9 AIMS OF THIS THESIS	20

CHAPTER TWO MATERIALS AND METHODS

2.1 GROWTH MEDIA	22
2.1.1 Liquid Media	22
2.1.2 Solid Media	22
2.2 BACTERIAL STRAINS	23
2.3 PLASMIDS AND BACTERIOPHAGE	23
2.4 BUFFERS AND SOLUTIONS	23
2.5 CHEMICALS	24
2.6 ANTIBIOTICS	24
2.7 ENZYMES	24
2.8 RADIOISOTOPES	25
2.9 GROWTH OF BACTERIA	25
2.10 PREPARATION OF COMPETENT <i>E. COLI</i> CELLS, TRANSFORMATION AND TRANSFECTION	25
2.11 GROWTH AND PURIFICATION OF BACTERIOPHAGE LAMBDA	25
2.11.1 Small Scale Liquid Lysates	25
2.11.2 Large Scale Liquid Lysates	26
2.11.3 Purification of Bacteriophage from Large Scale Liquid Lysates	26
2.11.4 Plaque Purification	26
2.11.5 Preparation of Phage Stock	27

2.12 DNA ISOLATION	27
2.12.1 Isolation of High Molecular Weight DNA from Whole Blood	27
2.12.2 Rapid Isolation of Crude DNA from Blood, Hair Roots and Tissue	28
a) Blood	28
b) Hair roots	28
c) Formalin-fixed tissue	28
2.12.3 Plasmid DNA	28
a) Alkaline lysis	28
b) Lysis by boiling	29
c) Cleared lysates	29
d) Caesium chloride density gradient purification	29
2.12.4 Phage DNA Isolation	30
a) Rapid isolation of λ DNA from small scale liquid lysates	30
b) Isolation of λ DNA from large scale liquid lysates	30
2.12.5 Purification of DNA by Phenol-Chloroform Extraction and Alcohol Precipitation	30
2.13 ELECTROPHORESIS OF DNA	31
2.13.1 Agarose Gel Electrophoresis	31
2.13.2 Polyacrylamide Gel Electrophoresis	31
2.14 RECOVERY OF DNA FRAGMENTS FROM AGAROSE AND POLYACRYLAMIDE GELS	31
2.14.1 Electroelution of DNA from Agarose and Polyacrylamide Gels	32
2.14.2 Freeze-Thaw Method for the Recovery of DNA from Agarose Gels	32
2.14.3 Isolation of DNA by Binding to Powdered Glass	32
2.14.4 Use of Low-Melting-Point Agarose	32
2.14.5 Elution of DNA from Polyacrylamide	32
a) Double-stranded DNA	32
b) Oligonucleotides	32
2.15 ENZYME REACTIONS	33
2.15.1 Restriction Endonuclease Digestions	33
2.15.2 Ligation Reactions with T4 DNA Ligase	33

2.15.3 <i>E. coli</i> DNA Polymerase I reactions	34
a) <i>E. coli</i> DNA Polymerase I Large Fragment (Klenow fragment)	34
b) Nick translation	34
2.15.4 SP6 RNA and T7 RNA Polymerases	35
2.16 DETERMINATION OF INCORPORATION OF RADIOACTIVITY INTO ACID-PRECIPIITABLE MATERIAL	35
2.17 DNA TRANSFER AND HYBRIDISATION	36
2.17.1 Colony Transfer to Whatman 540 Paper and Hybridisation	36
2.17.2 Plaque Transfer to Nitrocellulose	36
2.17.3 DNA Dot Blots	36
2.17.4 Transfer of DNA Fragments from Gels to Membranes	37
a) Nitrocellulose	37
b) Diazobenzylxymethyl (DBM) paper	37
c) Hybond N	38
d) Zetaprobe	38
e) Gene Screen Plus	38
2.17.5 Hybridisation , Washing and Radioautography of Membranes	39
2.18 CONSTRUCTION OF PARTIAL GENOMIC LIBRARIES	40
2.18.1 DNA Preparation, Ligation and Packaging	40
a) λ EMBL3 as vector	40
b) λ gt10 as vector	40
2.18.2 Library Storage and Amplification	41
2.19 DNA SEQUENCING	41
2.19.1 Use of the M13 Vector System	41
a) Intoduction	41
b) Transfection of JM101 or JM103 cells with M13 DNA	41
c) Preparation of M13 DNA and phage storage	41
d) Sizing of single-stranded DNA	42
e) Annealing of complementary strands of M13 DNA	42
f) M13 DNA sequencing	42
2.19.2 Plasmid Sequencing	43
2.19.3 Electrophoresis and Radioautography of Sequencing Products	44

2.20	DNA POLYMERASE ASSAY	44
2.20.1	Introduction	44
2.20.2	M13 Single-Stranded DNA as Substrate	44
2.20.3	"Gapped" Calf Thymus DNA as Substrate	45
2.21	THE POLYMERASE CHAIN REACTION	45
2.21.1	Amplification using the Klenow Fragment	45
a)	Incorporation of labelled radionucleotides	45
b)	"Cold" reactions	45
2.21.2	Amplification using <u>Taq</u> Polymerase	46
2.22	COMPUTER ANALYSES	46
2.22.1	Sequence Information	46
2.22.2	LOD Scores	47

CHAPTER THREE LINKAGE OF G8 TO THE HUNTINGTON DISEASE LOCUS IN NEW ZEALAND FAMILIES

3.1	INTRODUCTION	48
3.1.1	Background to the Discovery of the G8 Marker for Huntington Disease (HD)	48
3.1.2	Problems Associated with the use of G8 in Predictive Testing for HD	49
a)	Problems associated with the G8 marker	49
b)	Problems associated with the pedigrees	49
3.2	ASSESSMENT OF FAMILIES SEGREGATING HUNTINGTON DISEASE FOR LINKAGE OF THE LOCUS TO G8	50
3.2.1	Introduction	50
3.2.2	Haplotype Analysis of the Huntington Disease Family Members	50
3.2.3	Frequency of *H1, *H2, *E and *P Sites and the Level of Heterozygosity Observed in the General Population of New Zealand	51
3.2.4	Linkage Analyses	51

3.3 THE USEFULNESS OF G8 AS A MARKER FOR HUNTINGTON DISEASE IN PREDICTIVE TESTING OF AT RISK PEOPLE AND FOR EXCLUSION TESTS IN PREGNANCY	53
3.3.1 Predictive Testing	53
3.3.2 Exclusion Testing	56
3.4 GENOTYPE INCONSISTENCIES UNRELATED TO HUNTINGTON DISEASE	57
3.5 CONCLUSIONS	59

CHAPTER FOUR SEQUENCE ANALYSIS OF THREE POLYMORPHIC GENOMIC REGIONS DETECTED BY G8

4.1 INTRODUCTION	62
4.2 ANALYSIS OF THE *H1 REGION	63
4.2.1 Introduction	63
4.2.2 Screening Genomic DNA for an Invariant Fragment which Contains the *H1 Region	63
4.2.3 Isolation of Fragments Encompassing *H1- and *H2+	64
a) Library construction and screening	64
b) Subcloning from the G8 homologous recombinants	65
4.2.4 Sequence Comparison of *H1- and *H1+ DNAs	66
4.2.5 Sequence Analysis of DNA in the *H1 Region from a Third Individual	66
4.3 ANALYSIS OF THE *E REGION	67
4.3.1 Isolation of Fragments Encompassing *E+ and *E-	68
a) Library screening	68
b) Subcloning from the L1*E+ and L2*E- DNAs	68
4.3.2 Sequence Comparisons of *E+ and *E- DNAs	69
4.4 ANALYSIS OF THE *H2 REGION	70
4.4.1 Isolation of Fragments Encompassing *H2+ and *H2-	70
a) Subcloning from G8 to obtain *H2+	70
b) Identification of a 1.5Kb <u>Eco</u> RI fragment in *H2- DNA which is homologous with the 1.5Kb <u>Eco</u> RI fragment of G8	71

c) Library construction and screening for recombinants containing the *H2- site	71
d) Subcloning from L3*H2- DNA	71
e) Sequence comparison of *H2+ and *H2- DNAs	71
4.5 SCREENING DNA FROM OUTSIDE OF G8 FOR REPEAT SEQUENCES	72
4.6 DISCUSSION	74

CHAPTER FIVE ATTEMPTS TO DEVELOP AN ALTERNATIVE METHOD OF SCREENING FOR RFLPS LINKED TO THE HD LOCUS

5.1 INTRODUCTION	76
5.2 THE POLYMERASE CHAIN REACTION (PCR)	76
5.3 PCR AMPLIFICATION OF PLASMID/PHAGE DNA	78
5.3.1 Preliminary Experiments with Test Systems	78
a) Oligonucleotide synthesis and purification	78
b) Plasmids/phage containing DNA homologous with the oligonucleotide primers	78
c) DNA amplification	79
5.4 PCR AMPLIFICATION OF GENOMIC DNA	81
5.4.1 Investigation of the *H2, *H1 and *E Genomic Regions using PCR and Subsequent Restriction Enzyme Digestion of the Amplified Products	81
5.4.2 Simultaneous Amplification of *H1, *H2 and *E using the PCR and Subsequent Restriction Enzyme Digestion of the Amplified Products	82
a) Test system	83
b) Genomic DNA	83
5.4.3 PCR Amplification of a Target Sequence using either Low Molecular Weight (degraded) DNA or Crude DNA as Template	84
5.4.4 A Consideration of Some Experimental Results in an Attempt to Resolve the Problems	85
5.4.5 Conclusion	87
5.5 SEQUENCE ANALYSIS OF PRODUCTS AMPLIFIED IN THE POLYMERASE CHAIN REACTION	87
5.5.1 *E DNA	87

5.5.2 *H1 DNA	89
5.6 DISCUSSION	90

CHAPTER SIX CONCLUDING DISCUSSION

6.1 FAMILIAL STUDIES FOR HD STATUS	92
6.2 PCR-AIDED RFLP ANALYSIS IN HD	93
6.2.1 Polymorphic Sites *H1, *H2 and *E	93
6.2.2 Other Polymorphic Sites	95
6.2.3 Advantages Over Conventional RFLP Analysis	95
6.3 PCR AND DNA "FINGERPRINTING"	96
6.4 HIGH EFFICIENCY CLONING AND SCREENING	97
6.5 SUMMARY	98

REFERENCES	99
------------	----

APPENDIX I

A LIST OF GENOTYPE RESULTS (TABLE I 1)
 FREQUENCIES FOR G8 RFLPS (TABLE I 2)
 HETEROZYGOSITY OF G8 RFLPS (TABLE I 3)

APPENDIX II

RESTRICTION ENZYME MAPS OF λ EMBL3 AND λ gt10 (FIGURE II 1)

APPENDIX III

THE USE IN THIS THESIS OF DNA "FINGERPRINTING"

APPENDIX IV

ATTEMPTS TO ISOLATE A THERMOPHILIC DNA POLYMERASE FOR USE IN THE PCR

APPENDIX V

PRENATAL DIAGNOSIS AND CARRIER DETECTION OF CYSTIC FIBROSIS

APPENDIX VI

PUBLISHED PAPER

ACKNOWLEDGEMENTS

I wish to thank sincerely the following people and acknowledge their contributions, both direct and indirect to this thesis:

The late Professor A.M.O. Veale for initiating the research project.

Professor Peter Bergquist for discussion throughout the work and guidance with the preparation of the manuscript.

Doctor David Lane for constructive criticism of the manuscript.

Doctor Cyril Chapman for calculating the LOD scores.

Doctor Joanne Dixon for discussion and providing an interface between the research and the patients.

Members of the Huntington disease families for providing blood samples.

Doctor David Saul, Moreland Gibbs and Gregory Albertson for help and advice with computing.

Liam Williams for help with proof reading and Andrew Spiers for reading Chapter 1 and providing useful comments.

Barry Benning for always prompt repair of any of my equipment when it failed.

All students and staff of the department both past and present, but in particular Julie Allan, Andrew Fidler, Martin Kennedy, Gaetan Limsowtin, Rovine Knight, Judy McAnulty, Lyn McMillan, Dave Reddy, Jan Robinson, Carolyn Wallace and Liam Williams for their friendship and support.

I owe very special thanks to the following people:

Julie Whitefield for her direct help with typing and many discussions and for her indirect help as a valued friend, advisor and for always being available to give unlimited support with full empathy.

Joanna Putterill and Don Love for their friendship and sound perceptive advice at times of frequent crisis. In addition Don for helping me to begin the writing.

John Croft for the very long term use of his computer.

My family who have provided a loving supportive environment and allowed me to give total commitment to my work.

Lastly I thank the Medical Research Council of New Zealand, the Auckland Medical Research Foundation and the University Grants Committee for funding.

ABSTRACT

New Zealand families segregating the Huntington disease (HD) phenotype were investigated for linkage of the HD locus to the anonymous DNA locus G8, from chromosome 4p16.3. Linkage of the two loci was indicated in the largest family assessed. The results from this family together with those from 10 other smaller families were consistent with the existence of a single locus (that is, genetic homogeneity for HD). One crossover between the HD locus and the G8 marker locus was detected. Overall, the results of the linkage analysis indicate a distance of 4cM separating HD from G8.

The usefulness of G8 as a marker in predictive testing for HD was examined in the New Zealand families. In agreement with overseas findings, G8 and DNA contiguous with G8 are useful for predictive testing in some families. The main limitation for predictive testing is the family structure which often results in the unavailability of key individuals for testing and thus prevents prediction in others.

Sequence analysis of three sites which produce restriction fragment length polymorphisms (RFLPs) detected by G8 revealed that single point mutations were responsible for the presence or absence of the polymorphic sites *H1, *H2 and *E. The sequences were highly conserved between individuals in the regions investigated. The conservation of sequence provided the potential for the use of the polymerase chain reaction (PCR) to amplify each polymorphic region for more rapid assessment of these RFLPs.

The three regions were amplified successfully from genomic DNA. In addition, amplification of the three regions was possible with template DNA which was degraded or crude or isolated from tissue which had been fixed in formalin. However, the results of the subsequent analyses of the amplified products by restriction enzyme digestion showed that there are problems that can render the PCR unreliable. The presence of non-target sequences hindered detection of the genotype in the *H1 and *H2 regions and masked the true genotype of a person in the *E region. Thus, the potential of the PCR for presymptomatic diagnosis of HD remains to be realised. However, when the problems are overcome a very rapid analysis of RFLPs linked with HD using the PCR will be possible, as will be retrospective analyses.

ABBREVIATIONS

ABM	aminobenzyloxymethyl
Ap	ampicillin
ASO	allele specific oligonucleotide
ATP	adenosine triphosphate
BCIG	5-bromo-4-chloro-3-indolyl- β -galactopyranoside
bp	base pair (s)
BMD	Beckers muscular dystrophy
BRL	Bethesda Research Laboratories
BSA	bovine serum albumin
$^{\circ}\text{C}$	degrees Celsius
CF	cystic fibrosis
CGD	chronic granulomatosis disease
Cm	chloramphenicol
cM	centimorgan (1cM is a map distance over which 1% recombination occurs and which equates approximately with a physical distance of 1×10^6 bp)
cpm	counts per minute
CsCl	caesium chloride
CTP	cytidine 5'-triphosphate
CTD	calf thymus DNA
CVS	chorionic villus sample
dATP	2'-deoxyadenosine 5'-triphosphate
ddCTP	dideoxycytidine 5'-triphosphate
ddGTP	dideoxyguanosine 5'-triphosphate
ddNTP	dideoxyribonucleoside triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
DMD	Duchenne muscular dystrophy
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
dpm	decays per minute
ds	double stranded
DTT	dithiothreitol
dTTP	deoxythymidine 5'-triphosphate
EDTA	ethylene diamine tetra-acetic acid (disodium salt)
EtBr	ethidium bromide
GF/C	glass fibre filter
GTP	guanosine 5'-triphosphate
HD	Huntington disease
<i>hfl</i>	high-frequency lysogeny
HGM	human gene mapping
HPRT	hypoxanthine guanine phosphoribosyl transferase
HVR	hypervariable region
IPTG	isopropyl- β -D-thio-galactopyranoside
Kb	1000 base pairs
KF	Klenow fragment
Km	kanamycin
LB	Luria broth
LOD	log of the odds
M	moles per litre
mA	milliampere
mcs	multiple cloning site
mg	milligram
ml	millilitre
mM	millimoles per litre
MMD	myotonic muscular dystrophy

NEB	New England Biolabs
NEN	New England Nuclear
ng	nanogram
nM	nanomoles per litre
OD	optical density (absorbance)
OR	oligonucleotide restriction
ori	orientation
OTC	ornithine transcarbamylase
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PFGE	pulsed field gel electrophoresis
pfu	plaque forming unit
PIC	polymorphism information content
PKU	phenylketonuria
PMSF	phenyl methane sulphonyl fluoride
PVP	polyvinyl pyrrolidine
RF	replicative form
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rp	reverse primer
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
Sm	streptomycin
ss	single stranded
SSC	sodium chloride, sodium citrate
Tc	tetracycline
TCA	trichloroacetic acid
Td	dissociation temperature
TE	10mM Tris-pH 8, 1mM EDTA
TLC	thin layer chromatography
Tm	melting temperature
Tris	Tris (hydroxymethyl) aminomethane
TTP	thymidine 5'-triphosphate
μCi	microcurie
μg	microgram
μl	microlitre
μM	micromoles per litre
UTP	uridine 5'-triphosphate
UV	ultra-violet
V	volt
VNTR	variable number of tandem repeats
YT	yeast tryptone broth
17mer	universal primer, 17 bases