LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

GENETIC ANALYSIS OF THE HUMAN PROHIBITIN GENE (*PHB*) IN BREAST AND OTHER CANCERS

A Thesis submitted for the degree of

Doctor of Philosophy (Ph. D.)

by

Richard Hagan, B. A. (Moderatorship) in Genetics

at

University of Dublin, Trinity College

Supervisor:

Professor Peter Daly,

Department of Haematology/ Oncology

St. James's Hospital,

Dublin 8.

October 1998

TRINITY COLLEGE
16 MAY 2000
LIBRARY DUBLIN
THESIS
5626

CONTENTS

Table of Contents		Page
Declaration		**i**
Acknowledgements		**ii**
Abbreviations		**iii**
Figures and Tables		**v**
Summary		**ix**

1 GENERAL INTRODUCTION

1.1	Introduction	**1**
1.2	Genes Involved in Inherited Breast Cancer	
1.2.1	*BRCA1*: Linkage to Chromosome 17q21	**2**
1.2.2	*BRCA2*: Mapping to Chromosome 13q12	**3**
1.2.3	Association of Familial Breast Cancer and Other Cancers	**4**
1.2.4	Germline Mutations in *BRCA1* and *BRCA2*	**5**
1.2.5	Somatic Mutations in *BRCA1* and *BRCA2*	**7**
1.2.6	Risk of Breast Cancer for *BRCA1* and *BRCA2* Mutation Carriers	**8**
1.2.7	Genetic Heterogeneity of Familial Breast Cancer	**10**
1.2.8	Other Genes Associated with an Inherited Predisposition to Breast Cancer	**11**
1.3	Somatic Genetic Aberrations in Breast Cancer	
1.3.1	Introduction	**15**
1.3.2	Flow Cytometry and Cytogenetic Analysis of Breast Tumours	**16**
1.3.3	Gene Amplification	**18**
1.3.4	Loss of Heterozygosity (LOH)	**20**
1.3.5	Genetic Evolution of Breast Cancer	**22**
1.4	Tumour Suppressor Genes	
1.4.1	Definition of a Tumour Suppressor Gene	**23**
1.4.2	*RB* as a Paradigm of a Tumour Suppressor Gene	**24**

1.5 The Cell Cycle **25**

1.5.1 Cyclins, Cyclin-Dependent Kinases and the G1/S Transition **26**

1.5.2 Regulation of CDKs **28**

1.5.3 Cell Cycle Checkpoints **29**

1.5.4 Genes that Control the Cell Cycle and Tumour Suppressor Genes **30**

1.5.5 Cancer and Genes Involved in the Cell Cycle **32**

1.6 The Prohibitin Gene (*PHB*)

1.6.1 Isolation and Identification of the Prohibitin Gene (*PHB*) **33**

1.6.2 Amino Acid Composition **34**

1.6.3 Subcellular Location **36**

1.6.4 Chromosomal Location **37**

1.6.5 Gene Structure **39**

1.6.6 Prohibitin and the Cell Cycle **42**

1.6.7 Functional Activity **42**

1.6.8 Mutational Analysis **44**

1.7 Replicative Senescence **46**

1.8 Techniques used to Screen for Unknown Mutations **49**

1.9 Summation **50**

2 MATERIALS AND METHODS

2.1 Individuals Included in these Studies

2.1.1 Patient Population **52**

2.1.2 Control Population **52**

2.2 Collection and Storage of Peripheral Blood **53**

2.3 Isolation of total RNA, cDNA Synthesis and Isolation of Genomic DNA

2.3.1 Isolation of total RNA from Peripheral Blood **53**

2.3.2 Reverse Transcription of mRNA **54**

2.3.3 Isolation of DNA from Peripheral Blood **54**

2.3.4 Isolation of DNA from Buccal Cells **54**

2.3.5 Isolation of DNA from Formalin-Fixed Paraffin-Embedded Tissue **55**

2.3.6 Microdissection of Tumour Tissue **55**

2.4 Denaturing Gradient Gel Electrophoresis (DGGE)
2.4.1 Introduction 56
2.4.2 Computational Tools 57
2.4.3 MELT: The Melting Map Program 58
2.4.4 SQHTX: Program for the Prediction of Mobility and Displacement of Molecules in a DGG 60
2.4.5 Perpendicular DGGE 61
2.4.6 Parallel DGGE 63
2.4.7 Generation of Higher Melting Domains with a GC-clamp 64
2.4.8 Preparation and Electrophoresis of Denaturing Gradient Gels 65
2.4.9 Excision of Bands from Denaturing Gradient Gel for use in PCR 66
2.5 The Polymerase Chain Reaction
2.5.1 Introduction 66
2.5.2 Nested PCR Analysis of Prohibitin from cDNA 67
2.5.3 PCR Analysis of Prohibitin from Genomic DNA 68
2.5.4 Microsatellite Analysis 69
2.6 Restriction Enzyme Reactions 70
2.7 DNA Sequence Analysis 70

3 DGGE ANALYSIS OF *PHB* FROM cDNA

3.1 Aim of the Study 74
3.2 Results
3.2.1 Fragment and Primer Design for DGGE Analysis of the *PHB* cDNA 76
3.2.2 DGGE Analysis of the cDNA of *PHB* 81
3.2.3 Detection of Nucleotide Changes in the Coding Sequence of *PHB* in Women with Breast Cancer 84
3.2.4 Segregation of the Arg43Leu Allele in Family 1 85
3.2.5 Loss of Heterozygosity (LOH) Analysis of Tumour Samples 87
3.3 Discussion 88

4 DGGE ANALYSIS OF *PHB* FROM cDNA AND GENOMIC DNA

4.1 Introduction **93**

4.2 Results

4.2.1 Fragment and Primer Design for DGGE Analysis of the *PHB* Genomic DNA sequence **94**

4.2.2 DGGE Analysis of Genomic DNA **95**

4.2.3 Detection of Nucleotide Changes in *PHB* in Women with Breast Cancer **95**

4.2.4 Screening of Women with Breast Cancer for the Arg43Leu and Leu145Leu Alleles **98**

4.2.5 Screening of Women with Breast Cancer for the *Eco*R1 Polymorphism **99**

4.2.6 Segregation of the Arg43Leu Allele in Families 2-6 **99**

4.2.7 Statistical Analysis **102**

4.3 Discussion **104**

5 STUDY OF THE ARG43LEU ALLELE IN COLON CANCER

5.1 Introduction **109**

5.2.1 Selection of Patients with Colon Cancer **109**

5.2.2 Individual Results **110**

5.3 Discussion **111**

6 HAPLOTYPE ANALYSIS OF THE ARG43LEU ALLELE

6.1.1 Introduction **114**

6.1.2 Microsatellite DNA **115**

6.1.3 Location of Prohibitin **115**

6.2 Results **116**

6.2.1 Markers **116**

6.2.2 Haplotype Sharing in Affected and Unaffected Family Members **117**

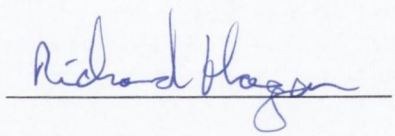
6.3 Discussion **119**

7 GENERAL DISCUSSION 122

REFERENCES 131

DECLARATION

I hereby declare that, except where otherwise acknowledged, this Thesis is entirely my own work. It has not been submitted for a higher degree on a previous occasion at this or any other university. The library may lend or make copies of this Thesis upon request.



Richard Hagan.

Acknowledgements

In order of appearance, I would like to thank Ross McManus for his guidance, supervision and critical reviews of this thesis, without which I would never have got this far. Prof. Peter Daly for his time, patience and words of comfort when I most needed them. Wilma and Kate for their perseverance with me and the collection of information and samples. The secretaries in Hospital 1 Ground Floor who have cheered me up every time I walk into their office.

I most sincerely thank the Health Research Board for their financial support and funding for this project.

From a technical point of view I wish to thank Claude Besmond and Emmanuelle Girdon for their advice in the use of software. David Noone for technical assistance with the sequencing of PCR products and Denis Sheils for statistical advice.

On a more personal note I have had the great pleasure of working with some of my best friends, Maria, Amanda, Orla, Suzanne, Sara, Jennifer, Linda and Susan. I would also like to acknowledge the help and friendship of those in the Departments of Haematology, Immunology, Histopathology, Biochemistry, Microbiology, Genetics and Nutrition that are unfortunately far too numerous to list.

Most importantly I could not have done this without the support of my family, Ian, Anne, John, Darryl, Leigh, and my long-suffering friend and partner Emma.

Abbreviations

AM	**Amplitude Modulation**
APS	**Ammonium Persulphate**
AT	**Ataxia-telangiectasia**
ATP	**Adenosine Triphosphate**
cDNA	**Complementary DNA**
cM	**Centimorgan**
DEPC	**Di-Ethyl Pyrocarbonate**
DGG	**Denaturing Gradient Gel**
DGGE	**Denaturing Gradient Gel Electrophoresis**
dNTP	**Deoxyribonucleoside Triphosphate**
DNA	**Deoxyribonucleic Acid**
FAP	**Familial Adenomatous Polyposis**
FFPE	**Formalin-Fixed Paraffin-Embedded**
GAP	**GTPase Activating Protein**
HDF	**Human Diploid Fibroblasts**
HNPCC	**Hereditary Non-Polyposis Colon Cancer**
kb	**Kilobases**
kDa	**Kilodaltons**
Kd	**Dissociation constant**
LFL	**Li-Fraumeni-like Syndrome**
LFS	**Li-Fraumeni Syndrome**
LOD	**Logarithm of the Odds**
LOH	**Loss of Heterozygosity**
mRNA	**messenger Ribonucleic Acid**
mtDNA	**Mitochondrial DNA**
PCNA	**Proliferating Cell Nuclear Antigen**
PCR	**Polymerase Chain Reation**
RFLP	**Restriction Fragment Length Polymorphism**
RNA	**Ribonucleic Acid**
rpm	**Revolutions Per Minute**

RT-PCR	**Reverse Transcription PCR**
SSCP	**Single-Stranded Conformational Polymorphism**
T	**Temperature**
Tm	**Melting Temperature**
UTR	**Untranslated Region**
UV	**Ultra-Violet**

Figures and Tables

Chapter 1:

Figure 1.1	Location of polymorphic markers and genes on chromosome 17q
Figure 1.2	Diagrammatic representation of the genomic organisation of *BRCA1* and *BRCA2*
Figure 1.3	Germline mutations that have been reported in the *BRCA1* gene
Figure 1.4	The Cell Cycle
Figure 1.5	Comparison of prohibitin amino acid sequences
Figure 1.6	Amino acid sequences of the human prohibitin and BAP37 proteins
Figure 1.7	The conserved 350 base pair region spanning the 5'-ends of the rat and human *prohibitin* genes

Chapter 2:

Figure 2.1	Input file M87
Figure 2.2	SQHTX output file
Figure 2.3	Perpendicular Denaturing Gradient Gel of an amplified fragment with a polymorphism
Figure 2.4	Perpendicular Denaturing Gradient Gel showing an amplified fragment with two melting domains
Figure 2.5	Parallel Denaturing Gradient Gel showing a time course
Figure 2.6	Schematic representation of *PHB* cDNA showing the location of the primers 1A-4B used to generate Fragments 1-4
Figure 2.7	Map of the plasmid vector pCR™II
Table 2.1	Primers used for amplification and DGGE of *PHB* cDNA
Table 2.2	Primers used for amplification and DGGE of *PHB* genomic DNA
Table 2.3	Microsatellite primer sequences
Table 2.4	Primers used for the amplification of fragments for DNA sequence analysis

Chapter 3:

Figure 3.1	Melt map of the prohibitin cDNA sequence
Figure 3.2	Melt map of Fragment 1 without a GC-clamp
Figure 3.3	Melt maps of Fragment 1 with and without a GC-clamp
Figure 3.4	Melt maps of Fragment 2 with and without a GC-clamp
Figure 3.5	Melt maps of Fragment 3 with and without a GC-clamp
Figure 3.6	Melt maps of Fragment 1 with a GC-clamp for 5%, 50% and 95% of the molecule in a helical state
Figure 3.7	Melt maps of Fragment 1 wild type and single base substitution variant sequences
Figure 3.8	Perpendicular Denaturing Gradient Gel of a fragment with a single melting domain
Figure 3.9	Parallel Denaturing Gradient Gel time course
Figure 3.10	Parallel Denaturing Gradient Gel comparing amplified fragments from cDNA and genomic DNA
Figure 3.11	Parallel Denaturing Gradient Gel of the Arg43Leu allele
Figure 3.12	Sequence analysis of the Arg43Leu allele
Figure 3.13	Creation of the *Sty1* restriction site by the Arg43Leu allele
Figure 3.14	Pedigree of Family 1
Figure 3.15	Analysis of the PCR products of exon 3 from tumour and normal tissue when digested with the restriction endonuclease *Sty1*
Table 3.1	Denaturing Gradient Gel (DGG) conditions for fragments amplified from *PHB* cDNA
Table 3.2	History of cancer in the 35 women with an affected relative

Chapter 4:

Figure 4.1	Creation of the *Sac1* restriction site by the Leu145Leu allele
Figure 4.2	Parallel Denaturing Gradient Gel demonstrating intron 5 and intron 6 variants
Figure 4.3	Electropherograms of intron 6 polymorphisms

Figure 4.4 Restriction endonuclease, *EcoR1*, digestions of PCR products containing intron 2

Figure 4.5 Pedigree of Family 1

Figure 4.6 Pedigree of Family 2

Figure 4.7 Pedigree of Family 3

Figure 4.8 Pedigree of Family 4

Figure 4.9 Pedigree of Family 5

Figure 4.10 Amino acid sequence homology of the central region of prohibitin and BAP37

Table 4.1 Denaturing Gradient Gel (DGG) conditions for fragments amplified from *PHB* genomic DNA

Table 4.2 History of cancer in the 72 women with an affected relative

Table 4.3 Incidence of intron polymorphisms

Table 4.4 Summary of individuals identified with Arg43Leu or Leu145Leu

Table 4.5 Summary of *PHB* genotypes

Chapter 5:

Table 5.1 Personal history of cancer for individuals with colon cancer, as given in their medical records

Chapter 6:

Figure 6.1 Examples of the microsatellite alleles for D17S855

Figure 6.2 Examples of the microsatellite alleles for D17S1820

Figure 6.3 Order of microsatellite markers on chromosome 17

Figure 6.4 Control sequence and microsatellite alleles for D17S943 and D17S951

Figure 6.5 Pedigree of Family 1 with haplotype of 17q markers

Figure 6.6 Microsatellite alleles for D17S951

Figure 6.7 Microsatellite alleles for D17S855

Figure 6.8 Pedigree of Family 2 with haplotype of 17q markers

Figure 6.9 Pedigree of Family 3 with haplotype of 17q markers

Figure 6.10 Pedigree of Family 4 with haplotype of 17q markers

Figure 6.11 Pedigree of Family 5 with haplotype of 17q markers

Figure 6.12 Haplotypes associated with Arg43Leu

Summary

Prohibitin, an evolutionarily conserved gene situated on chromosome 17q21, was originally identified as a gene with antiproliferative properties. Studies of a Japanese population have shown *prohibitin* to be somatically mutated in a proportion of breast tumours. The gene has not heretofore been shown to have an association with inherited forms of breast or other cancers. In this thesis the technique of denaturing gradient gel electrophoresis (DGGE) was developed to analyse the complete coding sequence of the *prohibitin* gene from fragments generated by the polymerase chain reaction (PCR). This was achieved by examining the melting profiles of different regions of the *prohibitin* sequence with the melt map program MELT87. Four overlapping fragments were designed and subsequently amplified by reverse transcription PCR thus enabling analysis of the prohibitin cDNA sequence by DGGE. A further five fragments were developed for analysis of *prohibitin* from genomic DNA. These five fragments were generated for exons 3-7 inclusive. By investigating the *prohibitin* gene in this way a total of 5 novel mutations were identified; one missense mutation in exon 3 (Arg43Leu), one silent mutation in exon 5 (Leu145Leu) and three intronic base substitutions, two in intron 6 and one in intron 3. Nucleotide alterations were determined by sequencing of PCR products.

Both of the coding sequence variants affected the recognition sites of restriction endonucleases. This afforded a straightforward technique by which a large population could be screened for the nucleotide differences. The missense mutation, Arg43Leu, was found in six unrelated individuals among 221 women with primary breast cancer and who were not selected for a family history of the disease. The Arg43Leu allele was found in

one individual from a control group of 200 women who did not have breast cancer and who had no first or second degree relatives with the disease. Four of the six patients with this novel *prohibitin* variant had at least one relative with breast cancer. For three of the six cases analysis of the extended family members verified that the mutation was not *de novo* and in these families there was a significant association between the Arg43Leu allele and cancer. Sites involved were breast and colon in females and colon, prostate and kidney in males. Further information on the family history of the remaining three patients was incomplete or unobtainable. In one individual with colon cancer where both normal and tumour tissue was available, the tumour sample showed loss of heterozygosity for the wildtype *prohibitin* allele.

A study of 119 formalin-fixed paraffin-embedded normal tissue samples from individuals with colon cancer revealed one further person with the Arg43Leu allele. Haplotype analysis using 8 dinucleotide repeat markers on chromosome 17q21 indicates that the correlation of Arg43Leu with cancer is independent of *BRCA1*. In conclusion, in some families the novel Arg43Leu variant of the *prohibitin* gene, or a very tightly-linked gene, appears to be associated with cancer development at multiple sites.

CHAPTER 1

GENERAL INTRODUCTION

GENERAL INTRODUCTION

1.1 Introduction

Breast cancer is an extremely common disease affecting women. In America and Western Europe between one in eight and one in twelve women develop breast cancer during their lifetime (American Cancer Society, Cancer Facts & Figures, 1995). A feature of the high incidence of female breast cancer is that approximately 10-15 % of breast cancers are associated with a positive family history of the disease (Weber *et al.*, 1993). A positive family history has long been known to be associated with an increased risk of breast cancer (Broca, 1866 cited in King *et al.*, 1993; Anderson 1974; Ottman *et al.*, 1983; Sattin *et al.*, 1985). Claus *et al.* (1990) showed that, in addition to a positive family history of breast cancer in first degree relatives, age at onset of disease was the strongest indicator of a possible genetic predisposition to breast cancer, with the risk of breast cancer to a mother or sister of a case increasing as age at onset of the case decreased. In another report, however, Mettlin *et al.* (1990) found that among patients aged 55 years, or older, greater risk was associated with diagnosis of breast cancer in a first degree relative at an older age. Evidence for the existence of a rare autosomal dominant allele segregating with increased susceptibility to breast cancer was provided by Claus *et al.* (1991). The allele appeared to confer particularly high risk at young ages. Most studies have shown that the relative risk to first degree relatives of individuals with breast cancer without selection for age at diagnosis or laterality of disease is between 2 and 3 (Slattery and Kerber, 1993).

There is much statistical and empirical evidence that genes are involved in predisposing to the development of breast cancer but identifying such genes has been

fraught with epidemiological difficulties. Families with multiple cases of breast cancer may not have a genetic predisposition to the disease and families with an inherited susceptibility may contain sporadic cases of the disease, not due to an inherited predisposition. In families with predisposing mutations the penetrance is not necessarily complete in carriers and age at onset of disease can vary widely amongst affected family members. The genes involved in hereditary breast cancer which have been identified to date have been found in families containing a large number of affected individuals diagnosed at a young age (Miki *et al.*, 1994; Wooster *et al.*, 1995).

1.2 Genes Involved in Inherited Breast Cancer

1.2.1 *BRCA1*: Linkage to Chromosome 17q21

The first evidence of the existence of a gene predisposing to early-onset breast cancer was reported by Hall *et al.* (1990). Linkage analysis was carried out on 23 families with 183 polymorphic markers in families with multiple cases of women with breast cancer. No other cancers were included. They estimated that 40 % of the families were linked to D17S74, a marker which is located on chromosome 17q21 (Figure 1.1). In this original study families with early-onset disease were defined as those with an average age at diagnosis of breast cancer of 45 years of age or less. Confirmation of linkage to 17q21 came from Narod *et al.* (1991), who reported the same chromosomal region for families with multiple cases of breast and ovarian cancer, though no association between positive linkage and age at onset of disease was found in this study. This breast and ovarian cancer locus was named *BRCA1* in a report by Solomon and Ledbetter (1991).

The region of chromosome 17q which was initially linked to *BRCA1* was over 50 cM in length. A number of groups narrowed this region by genotyping the families for an

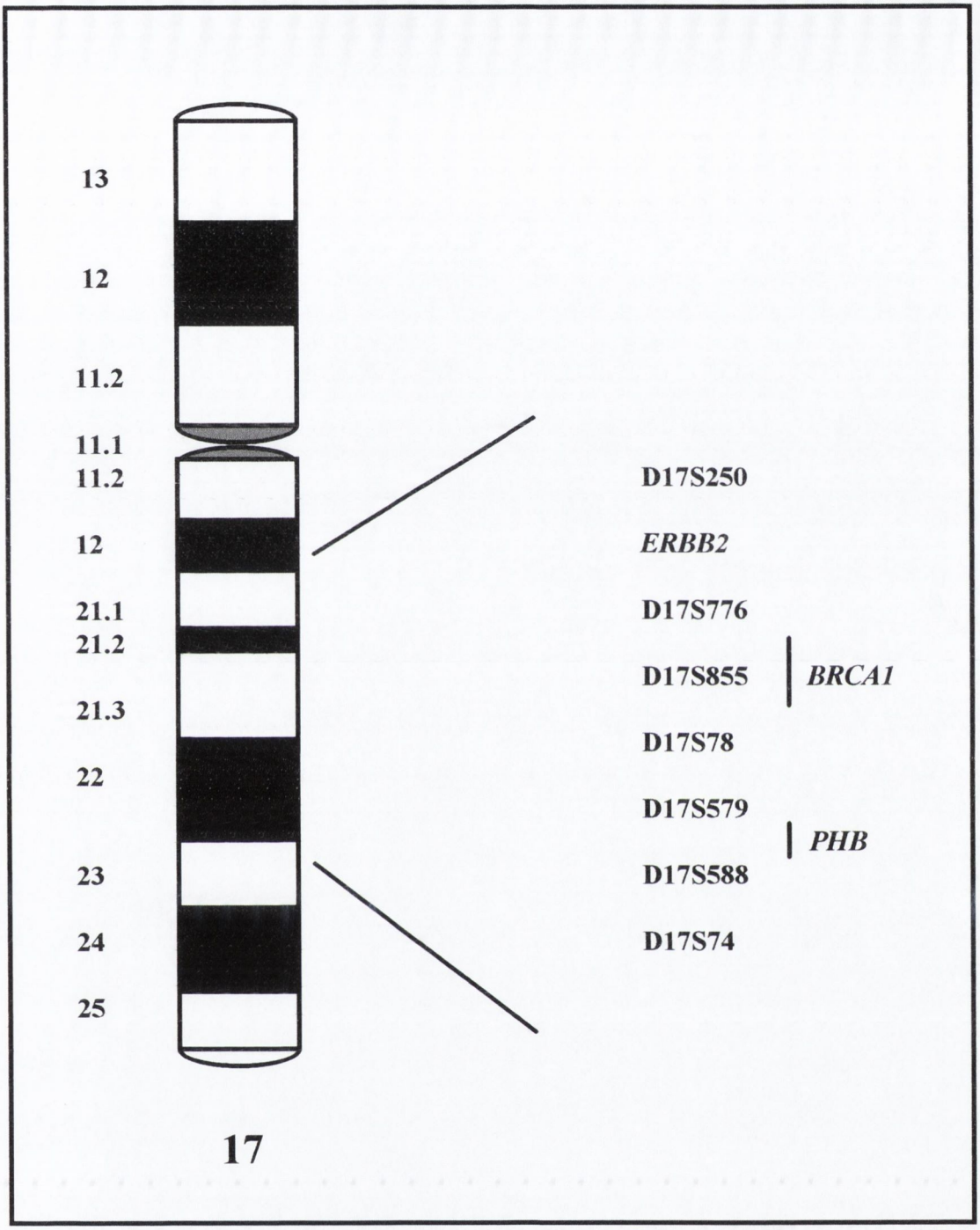


Figure 1.1. Location of polymorphic markers and genes on chromosome 17q

increasing number of markers in the region. Hall *et al.* (1992) defined an 8 cM interval, using 11 markers, which they described as being "very likely to include the disease gene *BRCA1*". In a collaborative linkage study of 214 breast or breast-ovarian families the strongest linkage evidence was obtained with the marker D17S588 (Easton *et al.*, 1993a). The predisposing gene was determined as being proximal to D17S588 and distal to D17S250, a length of 8.3 cM in males and 18.0 cM in females, thus confirming the results of Hall *et al.* (1992). In all of these studies there was evidence of genetic heterogeneity for 'breast-only' cancer families. Further studies reduced the locale of *BRCA1* to less than 4 cM (Bowcock *et al.*, 1993) and then to a minimal region of 1-2 cM proximal to D17S78 and distal to D17S776 (Simard *et al.*, 1993; Goldgar *et al.*, 1994; Albertsen *et al.*, 1994). The *BRCA1* gene was finally isolated by positional cloning methods and predisposing mutations were detected which confirmed that it was the gene associated with breast cancer susceptibility (Miki *et al.*, 1994).

1.2.2 *BRCA2*: Mapping to Chromosome 13q12

Not all breast-only cancer families were linked to the 17q21 locus and this indicated that further genes remained to be detected which predisposed to early-onset breast cancer. A second breast cancer susceptibility gene, *BRCA2*, was mapped to a 6 cM region of chromosome 13q12-13 by Wooster *et al.* (1994) between the markers D13S289 and D13S267, on the basis of meiotic recombinants in early-onset breast cancer cases from linked families. A combination of recombinants from other families, the use of additional markers in the region and the discovery of a homozygous somatic deletion in a pancreatic cancer (Schutte *et al.*, 1995) precipitated the identification of the *BRCA2* gene and germline mutations of the gene (Wooster *et al.*, 1995). Mutations in *BRCA2* confer roughly the same lifetime risk of female breast cancer as *BRCA1* of between 70-90 % (Ford *et al.*, 1998). This value for the lifetime risk has been based on mutations found in

large families with many affected individuals. Families that share these same mutations or other mutations in *BRCA1* or *BRCA2* but which result in a smaller number of affected family members are not included. Consequently, this value for the lifetime risk may be an overestimation. The genes *BRCA1* and *BRCA2* were predicted to account for approximately 70-80 % of inherited breast cancers (Szabo and King, 1995), whilst the remaining 20-30 % of high-risk families have no detected mutations in either *BRCA1* or *BRCA2*. Together *BRCA1* and *BRCA2* account for 6-10 % of breast and ovarian cancers unselected for family history (Szabo and King, 1997). The findings of more recent studies show that the estimates for high-risk families due to *BRCA1* and *BRCA2* are decreasing and possibly the majority of families with five or fewer cases of breast cancer and no ovarian or male breast cancer cases are not due to either *BRCA1* or *BRCA2* (Ford *et al.*, 1998).

1.2.3 Association of Familial Breast Cancer and Other Cancers

Families with mutations in either *BRCA1* or *BRCA2* are at an increased risk of other cancers apart from female breast cancer. *BRCA1* accounts for nearly 80 % of families with both female breast and ovarian cancer whereas the *BRCA2* gene confers a lower risk of ovarian cancer but a higher risk of male breast cancer than *BRCA1*. When the risk of other cancers was examined in *BRCA1* families a significant excess of colon and prostate cancers were found but no significant excess or deficiency of any other cancer (Ford *et al.*, 1994). An increase in the frequency of prostatic cancer, pancreatic cancer and possibly ocular melanoma has been reported for *BRCA2* carriers (Szabo and King, 1995; Phelan *et al.*, 1996).

1.2.4 Germline Mutations in *BRCA1* and *BRCA2*

The *BRCA1* and *BRCA2* genes share a number of features which make the search for mutations and their significance more difficult to determine. Both genes have a complex genomic organisation spanning approximately 70 kb of genomic DNA and long coding sequences with a large central exon, exon 11, (4, 932 bp for *BRCA2* and 3, 426 bp for *BRCA1*) (Figure 1.2). The *BRCA1* gene has 5, 592 nucleotides and 22 coding exons, and *BRCA2* is 10, 443 nucleotides long encoded by 26 exons (Miki *et al.*, 1994; Tavtigian *et al.*, 1996). Other similarities include AT-rich coding sequences which code for highly charged proteins and both have translational start sites in exon 2. Neither gene shows significant homology to other known genes although a number of motifs have been identified for both genes. *BRCA1* has a RING finger motif and both *BRCA1* and *BRCA2* have putative granin consensus sequences (Holt *et al.*, 1996; Jensen *et al.*, 1996).

A wide range of mutation-detection techniques have been used to screen the gene completely or partially for mutations in both *BRCA1* and *BRCA2*. These include direct sequencing, the protein truncation test, heteroduplex analyses, SSCP and DGGE (Shattuck-Eidens *et al.*, 1995; Gayther *et al.*, 1997). No single technique can easily detect all mutations and both genomic DNA and cDNA are required for a comprehensive screening of either gene. The majority of mutations identified to date are nonsense, regulatory, alteration of splice sites or frameshift, all of which result in a truncated protein or no product at all.

Couch and Weber (1996) reported that of the 254 mutations identified in *BRCA1* by July 1996, 132 were unique with 81 % of these mutations being small deletions, insertions, nonsense point mutations, splice variants or regulatory mutations, resulting in absence or premature termination of the BRCA1 protein (Figure 1.3). Four percent were described as disease-associated missense mutations and 15 % of the described alterations were of unknown biological significance. This data is collected and published by the

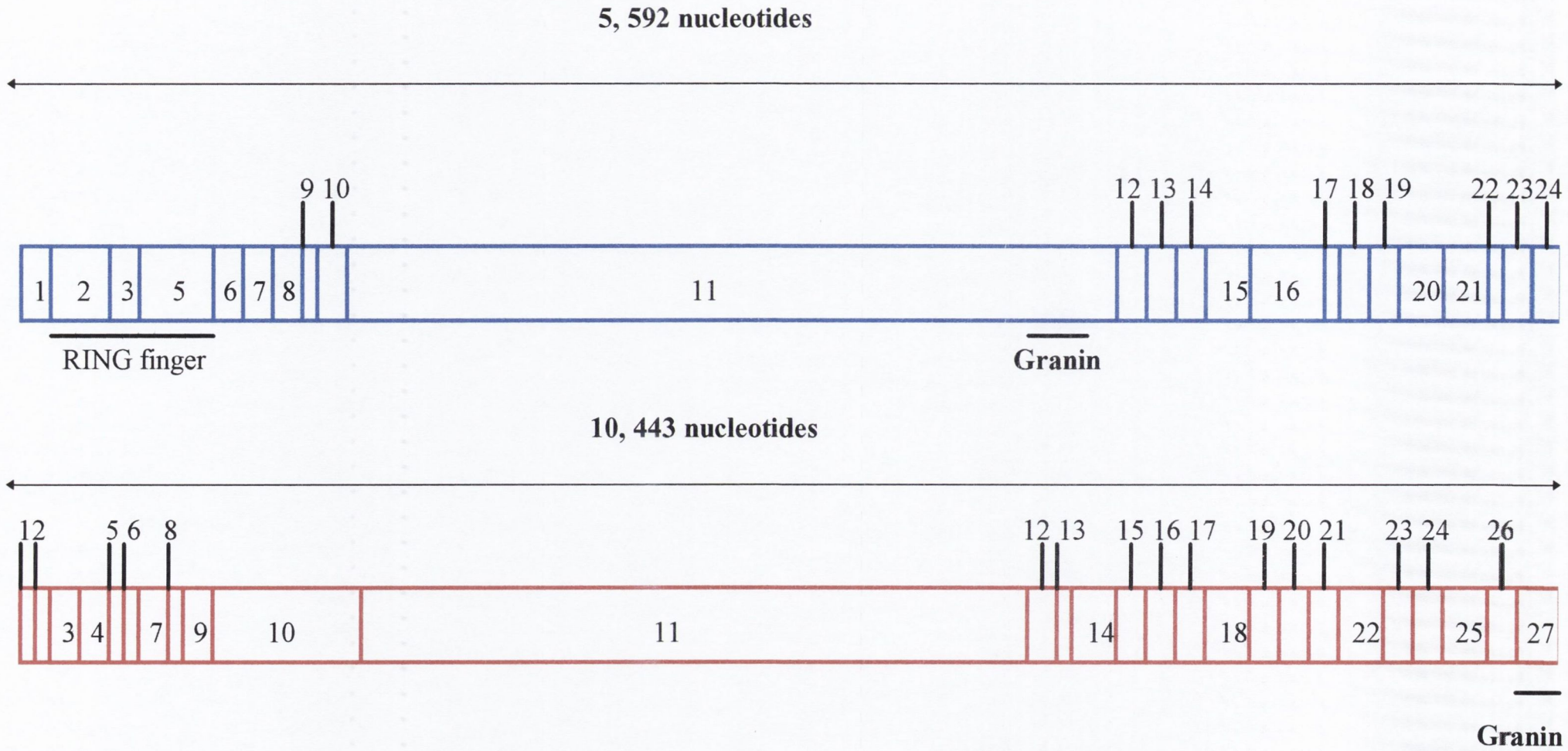


Figure 1.2. Diagrammatic representation of the genomic organisation of *BRCA1* and *BRCA2*. *BRCA1* is the upper figure in blue, *BRCA2* is the lower figure in red.

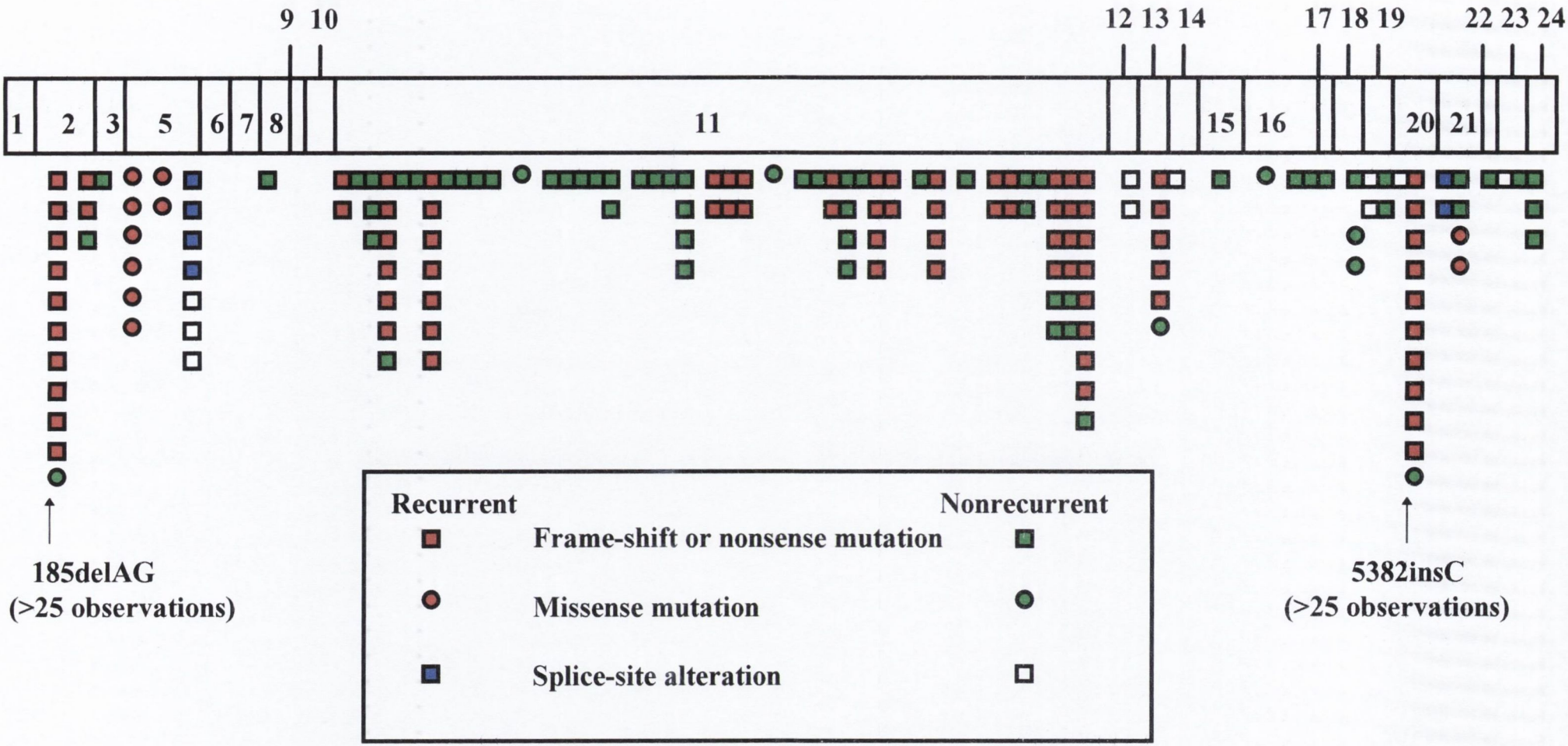


Figure 1.3. Germline mutations that have been reported in the *BRCA1* gene. The numbers refer to the exons of the gene; there is no exon 4 (Figure adapted from Collins, 1996; mutations described by Szabo and King, 1995; Couch *et al.*, 1996).

Breast Cancer Information Core (BIC) for mutations/polymorphisms in the *BRCA1* gene (http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic), a similar database does not exist for *BRCA2* mutations/polymorphisms and the percentage of each type of mutation varies between studies depending on the techniques used, ascertainment of individuals and regions screened.

The selection of families for screening of the *BRCA1* and *BRCA2* genes for mutations, however, has been biased towards families with a large number of affected individuals diagnosed at an early age. Therefore the mutational spectrum may be weighed towards mutations that cause maximum disruption of the functional protein and result in a high penetrance level. The level of penetrance may also be influenced by environmental factors or other genes that modify the effect of BRCA1. For example, the 185delAG mutation in *BRCA1* has been found at approximately equal frequencies in Iraqi/Iranian and Ashkenazi Jewish families with ovarian cancer (Modan *et al.*, 1996) yet both breast and ovarian cancer incidence are significantly lower for Iraqi/Iranian women than for Ashkenazi Jewish women (cited in Szabo and King, 1997). The comparison of women carrying 185delAG who develop cancer with those who are disease-free could identify factors, genetic or environmental, that contribute to the development of the disease. The determination of the full range of mutations awaits screening of the general population, families with a small number of affected individuals and cases of breast cancer where a family history is negative for other breast cancer cases. A functional test for BRCA1 and BRCA2 will be required before all mutations can be interpreted as being disease-associated or not.

1.2.5 Somatic Mutations in *BRCA1* and *BRCA2*

Following the identification of the breast cancer susceptibility genes, *BRCA1* and *BRCA2*, several studies investigated the frequency of somatic mutations in these genes. The hereditary breast/ovarian cancer gene *BRCA1* is located on 17q21, near the *prohibitin* gene (*PHB*). The chromosomal region 17q21 displays loss of heterozygosity (LOH) in sporadic breast and ovarian cancers and breast tumours unselected for family history with frequencies ranging from 30-70% (Eccles *et al.*, 1990; Sato *et al.*, 1991; Futreal *et al.*, 1992; Jacobs *et al.*, 1993; Cropp *et al.*, 1994; Beckmann *et al.*, 1996). Similarly somatic loss of markers for the *BRCA2* region, *RB* was simultaneously lost, has been found in 25-30 % of sporadic breast tumours (Cleton-Jansen *et al.*, 1995) and in other tumour types (Foster *et al.*, 1996).

Several studies investigating whether or not LOH specifically affects the *BRCA1* and *BRCA2* loci have been conducted. These have taken the approach of screening a number of breast and/or ovarian tumours which were unselected for family history of disease but which exhibited LOH at the locus. The results described to date for *BRCA1* and *BRCA2* are analogous. An examination of the *BRCA1* coding region for mutations revealed no somatic mutations in primary breast and ovarian tumours selected on the basis of allelic loss at the *BRCA1* locus (Futreal *et al.*, 1994) but in two other studies approximately 9 % of sporadic ovarian tumours were shown to have somatic *BRCA1* mutations (Merajver *et al.*, 1995; Hosking *et al.*, 1995). Four of the five somatic mutations discovered in these sporadic ovarian tumours result in a premature truncation of the *BRCA1* reading frame and the fifth alteration is a missense mutation in the zinc-finger motif (Merajver *et al.*, 1995; Hosking *et al.*, 1995). For *BRCA2* a single missense mutation was identified from 70 primary breast tumours and no mutations in 55 primary

epithelial ovarian carcinomas (Lancaster *et al.*, 1996). Therefore neither *BRCA1* or *BRCA2* are frequently inactivated in sporadic tumours by somatic mutational events.

Several reports have indicated that LOH of 17q does not reflect somatic inactivation of *BRCA1* but may indicate the presence of other tumour suppressor genes on chromosome 17 (Futreal *et al.*, 1994). Certainly a number of LOH studies of sporadic breast and ovarian tumours have detected deletion of regions distal and proximal to *BRCA1* (Jacobs *et al.*, 1993; Saito *et al.*, 1993; Cropp *et al.*, 1993, 1994). Cropp *et al.* (1993) identified three distinct regions on chromosome 17q affected by LOH, the most proximal of these includes the *prohibitin* region. Selection of tumours for LOH of 17q21 and mutational analysis of *BRCA1* has not identified a high percentage of positive cases (Futreal *et al.*, 1994; Merajver *et al.*, 1995; Hosking *et al.*, 1995), suggesting again, that other genes in the region are inactivated by the LOH identified.

1.2.6 Risk of Breast Cancer for *BRCA1* and *BRCA2* Mutation Carriers

The risk to female carriers of developing breast cancer for specific mutations in the breast cancer susceptibility genes *BRCA1* and *BRCA2* has been estimated through linkage studies. The penetrance can be calculated by maximisation of the LOD (logarithm of the odds) scores. Therefore, for genes that show a variable penetrance between families, the estimation does not apply to all possible mutations. Using this method the estimated breast cancer risk conferred by a mutation in *BRCA1* is 51 % by age 50 and 85 % by age 70, with the estimated risk of either breast or ovarian cancer being 95 % by age 70 (Easton *et al.*, 1995). For *BRCA2* the lifetime risk is similar to that found for *BRCA1* but has an older distribution of ages at onset (Schubert *et al.*, 1997). The estimated breast cancer risk conferred by a mutation in *BRCA2* is 28 % by age 50 and 84 % by age 70 (Ford *et al.*, 1998). One-hundred and thirty-two distinct mutations in *BRCA1* have been reported to date from over 300 families (Couch and Weber, 1996). In an examination of analyses of

inherited mutations in *BRCA1* and/or *BRCA2* in high-risk families (those with either at least three female relatives with breast cancer or at least two affected relatives if one case is either ovarian cancer or male breast cancer) from Finland, France, Holland and Belgium, Hungary, Iceland, Israel, Russia, Sweden and Denmark, and the United States, *BRCA1* mutations were found to be present at 1.5-2.0 fold excess to mutations in *BRCA2* in each region except Iceland (Szabo and King, 1997). Comparing these findings to other studies the authors felt that this represented a true difference in *BRCA1*- versus *BRCA2*-mutation frequency and was not due to ascertainment bias. The percentage of familial male breast cancer due to *BRCA2* has been estimated to be about 19 % from a combination of results from studies in the United States (Couch *et al.*, 1996b; Friedman *et al.*, 1997; Serova *et al.*, 1997).

The penetrance of identical mutations varies considerably as does the heterogeneity of ovarian and breast cancer risk between families. Two of the most common mutations in *BRCA1* are 185delAG and 5382insC, which account for approximately 11 % of all mutations reported thus far (Couch and Weber, 1996). For the 185delAG mutation expression of the phenotype has been shown to range from early-onset bilateral breast cancer and ovarian cancer to later onset of breast cancer, although the overall penetrance remains similar for all these families (Friedman *et al.*, 1995). In a study of 30 males diagnosed with breast cancer in Iceland, 12 were found to carry the same frameshift mutation in *BRCA2*, 999del5, but 3 of them had no family history of breast cancer (Thorlacius *et al.*, 1996). Also in this study, three other families with this frameshift mutation did not have any cases of male breast cancer. These findings indicate that the same mutation can result in different cancer phenotypes and that the variable expression indicates an effect from modifying genes or environmental factors.

It has been suggested that the location of *BRCA1* mutations is correlated with the incidence of breast and ovarian cancer (Gayther *et al.*, 1995). Mutations that occur in the

first two-thirds of the gene and produce a truncated protein, result in a higher risk of ovarian cancer relative to breast cancer than when the truncating mutation occurs in the last one-third of the gene. In an analysis of families with mutations in *BRCA2* a high proportion of ovarian cancers were found to cluster in a 3.3 kb region of exon 11 (Gayther *et al.*, 1997). Mutations outside this region are found in families with a greater ratio of breast to ovarian cancers. Further genotype-phenotype associations will require analysis of the distribution of mutations along the complete genes from a larger number of individuals with breast and/or ovarian cancer.

1.2.7 Genetic Heterogeneity of Familial Breast Cancer

Despite original estimates predicting that 80 % of families with at least one case of ovarian cancer and nearly 50 % of breast-only cancer families were likely to be linked to *BRCA1* (Easton *et al.*, 1993a; Narod *et al.*, 1995), more recent studies have found that only between 20-25 % of high risk families have identifiable *BRCA1* mutations (Szabo and King, 1997). High risk families have been classified as those with at least 3 females with breast cancer or at least 2 cases if one case is either ovarian cancer or male breast cancer (Szabo and King, 1997). The number of high risk families due to mutations in *BRCA2* is substantially less than that due to *BRCA1* mutations in all geographical regions examined except Iceland. This probably represents the relative contribution of each gene to the number of high risk families rather than a bias due to ascertainment. Numerous studies have screened high risk families for mutations in *BRCA1* and *BRCA2* but have failed to find mutations in either gene in roughly 30 % of families (Serova *et al.*, 1997; Ramus *et al.*, 1997; Hakansson *et al.*, 1997; Schubert *et al.*, 1997). Serova *et al.* (1997) did not find any mutations in *BRCA1* or *BRCA2* in 15 out of 23 families that were specific for female breast cancer even though, for those analysed, 7 had positive LOD scores for *BRCA1* or *BRCA2* and only 3 of the 15 had negative LOD scores. Many possible explanations for

this low detection rate of mutations in *BRCA1* and *BRCA2* have been proposed. These include not detecting mutations that involve large insertions or deletions, alterations that affect the regulatory aspects of the genes or that the case screened from the high risk family was a sporadic case. Even if this is the situation for some families the results to date suggest that there are other major breast cancer susceptibility genes.

1.2.8 Other Genes Associated with an Inherited Predisposition to Breast Cancer

Whilst the genes *BRCA1* and *BRCA2* are predicted to account for a large percentage of breast cancer families, germline mutations of *BRCA1* are predicted to account for just 1-2 % of all breast cancers and around 3 % of ovarian cancers. Apart from these highly-penetrant examples, a small number of other specified genes have also been associated with an inherited predisposition to breast cancer. These include the *p53* gene located on chromosome 17p, the ataxia-telangiectasia gene, *ATM*, the gene for Cowden syndrome, *PTEN/MMAC1*, the gene for Peutz-Jeghers syndrome, *LKB1*, the oestrogen receptor gene, the *HRAS1* locus and the androgen receptor gene.

Germline mutations of *p53* are associated with the Li-Fraumeni syndrome (LFS), a classification used to describe families with a proband aged under 45 years with a sarcoma having a first degree relative under 45 years of age with any cancer and an additional first or second degree relative under 45 years of age in the same lineage with any cancer or a sarcoma at any age (Li *et al.*, 1988). Germline mutations in *p53* are also associated with the Li-Fraumeni-like syndrome (LFL), a description of families with incomplete Li-Fraumeni syndrome (Pearson *et al.*, 1982; Birch *et al.*, 1994), although around 30 % of LFS and a higher percentage of LFL families are not due to germline mutations in *p53* (Varley *et al.*, 1997). To date 114 germline mutations of *p53* have been reported (Varley *et al.*, 1997) but the numbers are too small to determine whether penetrance and phenotypic expression are influenced by the type of germline mutation. Transitions and

transversions represent 66 % and 21 % of germline *p53* mutations respectively, with the remainder being deletions, insertions or splice variants. The spectrum of tumours within the Li-Fraumeni syndrome includes acute leukaemia, premenopausal breast cancer, brain and adrenocortical tumours, as well as bone and soft-tissue sarcomas (Li *et al.*, 1988). Premenopausal breast cancer is particularly common in Li-Fraumeni families. Other cancers that have been described at a higher frequency in Li-Fraumeni families are melanoma, germ cell tumours, Wilms' tumours, gastric carcinomas, pancreatic cancers and lung cancers (Hartley *et al.*, 1987; Strong *et al.*, 1987; Hartley *et al.*, 1989; Varley *et al.*, 1997).

In 3 studies women with breast cancer who had been either, in one case completely unselected or, in the other two studies, selected for early age at onset were examined for germline mutations of *p53* (Borresen *et al.*, 1992; Prosser *et al.*, 1992; Sidransky *et al.*, 1992). Two germline mutations were found in the unselected group of 167 women and one in each of the other two studies of a total of 262 women. In each of the four cases the individual had a strong family history of breast and other cancers. For two of these cases there was a family history suggestive of a Li-Fraumeni-like syndrome. On this basis, germline mutations of *p53* in women presenting with breast cancer would appear to be a rare occurrence, yet a small but significant number, 1 %, of breast cancers in women between the ages of 30 and 40 years occur in Li-Fraumeni families (Easton *et al.*, 1993b). In two of the four reports described above only exons 5-8 of the 11 exons that comprise the *p53* gene were analysed. The coding region spans exons 2-11 and it is estimated that over 20 % of all mutations remain undetected due to analysis by many groups of only the core domain of exons 5-8 (Greenblatt *et al.*, 1994).

Ataxia-telangiectasia (AT) is an autosomal recessive disorder distinguished by progressive cerebellar ataxia and oculo-cutaneous telangiectasia. The disorder is associated with an increased risk of lymphomas and leukaemias in individuals homozygous

for a mutation in the AT gene (*ATM*) (Swift *et al.*, 1991). Studies of AT family members resulted in the hypothesis that individuals who are heterozygous for an *ATM* mutation are at an increased risk of developing breast cancer. In one study an *ATM* mutation was shown to be associated with an increased risk of breast cancer in both homozygotes and heterozygotes (Stankovic *et al.*, 1998). Recent attempts to define the risk of breast cancer in female *ATM* heterozygotes have been made using two different approaches (Athma *et al.*, 1996; Fitzgerald *et al.*, 1997). In one, the estimated risk to heterozygotes was calculated to be nearly four times that for the non-carriers (Athma *et al.*, 1996). This estimate was determined by following identified *ATM* mutations in relatives of those with AT. In the other study *ATM* mutations were found in 2/401 women with early-onset breast cancer and in 2/202 controls. From this they concluded that there was no evidence for an increased risk of breast cancer to heterozygotes (Fitzgerald *et al.*, 1997). It has been pointed out that in terms of confidence intervals the data from these reports are consistent with no increased risk to heterozygotes but also consistent with an increased risk of up to seven fold (Bishop and Hopper, 1997). This paradox is due to the small number of mutation carriers found in one study and the small number of breast cancers examined in the other study. As a result the precision of the estimated risk is reduced and the variable of chance plays a greater role.

Cowden syndrome is an autosomal dominant disorder associated with the development of hamartomas and benign tumours in a variety of tissues, including the breast. Women with Cowden syndrome have been reported to be at an increased risk of breast cancer and this has been supported by the identification of germline mutations in the *PTEN/MMAC1* gene which is located in the Cowden syndrome critical interval on chromosome 10 (Tsou *et al.*, 1997). In a mutational analysis of 54 unselected primary breast cancers two mutations were identified (Rhei *et al.*, 1997). One was a somatic 2-bp deletion in a sporadic breast cancer and the second was a 4-bp deletion in a breast cancer

patient with a clinical history consistent with Cowden syndrome. Another syndrome, Peutz-Jeghers syndrome, is due in most cases to the presence of germline mutations in the *LKB1* gene on chromosome 19p. The syndrome is described by intestinal hamartomas with a greater risk of cancer types, including breast cancer. Mutation of the *LKB1* gene does not, however, appear to be associated with the development of sporadic breast cancer (Bignell *et al.*, 1998). One gene that has been associated with the development of sporadic breast cancer is *prohibitin* (Sato *et al.*, 1992). The authors identified 3 cases of somatic mutation in the coding sequence from an initial sample of 23 sporadic breast cancers which showed LOH on 17q or had developed in patients 35 years of age or younger. Two were missense mutations and one was a 2 base pair deletion which created a frameshift resulting in a premature stop codon 70 base pairs downstream. The tumour with the frameshift mutation did not possess a normal allele suggesting that the *PHB* product may be associated with development and/or progression of breast cancer.

The oestrogen receptor locus on chromosome 6q is frequently lost in tumours and has been proposed as a candidate gene for inherited breast cancer (Fuqua *et al.*, 1993). In two families where linkage has been shown against *BRCA1* and *BRCA2*, breast cancer is coinherited with the oestrogen receptor (Zuppan *et al.*, 1991; Szabo and King, 1995). Although the total lod score for these families is 3.7 (1.8 and 1.9 for the two families), no mutations of functional significance have as yet been identified. Unlike the oestrogen receptor, mutations in the androgen receptor have been identified that are responsible for inherited predisposition to breast cancer. In this report the mutations predisposed to male breast cancer in men with the rare, X-linked Reifenstein syndrome (Wooster *et al.*, 1992).

From epidemiological studies an association between certain rare variants of the *HRAS1* minisatellite locus and an increased risk of breast cancer was identified (Krontiris *et al.*, 1993). The locus is within the 3' untranslated region (UTR) of the *HRAS* oncogene on chromosome 11. The rare alleles are associated with a relative risk of 2.0 for breast

cancer which translates to the population as a whole as an attributable risk of about 9 %. The way in which these variant alleles increase the risk of cancer has not been determined but it may be due to interference with the regulation of gene expression.

Genes that predispose to breast cancer in a less penetrant manner than *BRCA1* and *BRCA2* are likely to result in relatively unremarkable family histories but may in total account for a larger number of breast cancers. The existence of low penetrance genes could explain a substantial percentage of all cancers and would therefore be of great importance. At the lower end of the penetrance scale such genes may not confer a detectable increased risk to relatives, thus making these genes very difficult to detect.

1.3 Somatic Genetic Aberrations in Breast Cancer

1.3.1 Introduction

Multiple genetic aberrations have been identified by molecular genetic analysis of tumours from both sporadic and inherited cases that may represent important steps in tumourigenesis. These aberrations range from single nucleotide point mutations and structural chromosomal changes to gross change in chromosome copy number, termed aneuploidy (Devilee and Cornelisse, 1994). The most common and extensively covered aberrations are gene amplification and loss of heterozygosity. The amplification of genes provides a mechanism whereby the tumour cells gain an increased level of expression of dominantly acting oncogenes whilst loss of heterozygosity indicates that loss of function of tumour suppressor genes is also important for the process of tumourigenesis. The determination of the genes affected by these aberrations is fundamental to understanding the genetic pathways involved in breast carcinogenesis.

The evolution of somatic genetic aberrations underlying the progression of breast cancer is not fully understood. This is, in part, due to the relatively small amount of information available from premalignant and preinvasive breast lesions. The histological stages of breast carcinogenesis are not clearly defined as is the case with colorectal carcinogenesis for which the critical genetic steps from premalignant lesions to primary tumours and metastases have been identified (Fearon and Vogelstein, 1990). Therefore, the genetic aberrations important for breast tumour initiation and progression have been difficult to distinguish from those of background genetic instability of the tumour DNA. Breast cancers can be distinguished on the basis of invasive carcinoma or non-invasive carcinoma. The most common invasive type is infiltrating ductal carcinoma whilst the two most important non-invasive carcinomas are ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS). Benign lesions are also found which have some of the signs of in situ carcinoma (atypical hyperplasia). It would appear that most invasive carcinomas are preceded by a carcinoma in situ although it has also been postulated that a direct transition from normal epithelium to malignancy can occur (Ponten *et al.*, 1990).

1.3.2 Flow Cytometry and Cytogenetic Analysis of Breast Tumours

Flow cytometry has been used extensively to examine the DNA content of breast carcinomas. Between 60 -80 % of breast tumours present clear evidence of an aneuploid DNA content by image and flow cytometry (Kallioniemi *et al.*, 1987; Beerman *et al.*, 1990). The most frequently observed aneuploid state is that with a hypotetraploid DNA content. A single aneuploid stemline is usually observed with flow cytometry but multiple aneuploid stemlines have been identified in 10-15% of cases (Devilee *et al.*, 1994). This clearly shows that intratumour genetic heterogeneity exists and that this is likely to be underestimated when only a small number of tumour samples are analysed. In two studies

40-60% of primary tumours showed multiple DNA aneuploid tumour cell populations (Beerman *et al.*, 1991; Bonsing *et al.*, 1993).

The identification of genetic aberrations in breast cancer from cytogenetic studies is limited given the many technical difficulties. Of the tumours that have been karyotyped the most common abnormalities are changes affecting whole chromosomes and nearly all chromosomes are affected at different frequencies. Cytogenetic studies reveal that the majority of tumours are composed of more than one stemline and the resultant karyotypes are a composition of these. This is in agreement with cytometric analysis showing intratumour heterogeneity.

Comparative Genomic Hybridization (CGH) is a recently developed molecular cytogenetic method used to screen entire genomes to determine their DNA-sequence copy-number variation (Kallioniemi *et al.*, 1992). CGH identifies regions of increased or reduced DNA-sequence copy-number by the intensity of flourescence of tumour DNA, or any test genome, and normal reference DNA hybridized to normal metaphase chromosomes; in this analysis no specific probes or prior knowledge of chromosomal aberrations is required. The genetic imbalance is quantified by digital image analysis of the difference in fluorescence intensity ratio. One significant advantage of CGH over other techniques for the detection of genetic aberrations in tumours, such as karyotyping, is that genomic DNA is the only source material required and DNA extracted from archived tumour specimens can be used (Reid *et al.*, 1995). The activation of oncogenes and the inactivation of tumour suppressor genes are frequently associated with DNA amplifications and deletions respectively. Consequently CGH is an important process by which chromosomal regions containing tumour-associated genes can be identified.

The technique of CGH has been used extensively to study breast carcinomas in recent years (Isola *et al.*, 1994; Kallioniemi *et al.*, 1994; Muleris *et al.*, 1994; Isola *et al.*, 1995; Reid *et al.*, 1995; Courjal and Theillet, 1997; Nishizaki *et al.*, 1997; Schwendel *et*

al., 1998). Kallioniemi *et al.* (1994) were first to use CGH to examine breast cancer, both cell lines and primary tumours. In total they identified 26 chromosomal regions, ranging from very small regions to whole-arm or whole-chromosomes, which showed amplification. Seventy-eight percent of the amplification events occurred at loci that were previously unknown to be amplified in primary breast tumours. The subsequent studies have revealed a complex pattern of gains and losses of many chromosomes in breast carcinomas. The most common regions of DNA copy-number increase have been on 1q, 3q, 6p, 8q, 11q, 12q, 17q and 20q. DNA losses have commonly been found for 3p, 6q, 8p and 16q. The regions of amplification on chromosomes 8q, 11q and 17q have been shown to include the oncogenes *MYC*, *cyclin D1* and *ERBB2* respectively (Kallioniemi *et al.*, 1994; Reid *et al.*, 1995; Courjal and Theillet, 1997; Schwendel *et al.*, 1998). The target genes located on the additional regions of increased DNA copy-number are unknown. The increased copy-number at 20q has become recognised as one of the most common genetic aberrations in breast cancer only since the application of CGH analysis (Kallioniemi *et al.*, 1994). Recently a candidate target gene on 20q, *AIB1* (amplified in breast cancer-1), has been identified which is amplified in approximately 10% and overexpressed in 64% of primary breast tumours (Anzick *et al.*, 1997).

1.3.3 Gene Amplification

Amplification of specific chromosomal regions is associated with the activation or over-expression of the target gene within the amplification unit. Molecular analysis of primary breast tumours has shown that amplification of three chromosomal regions, 8q24, 11q13 and 17q12, is centred on the known oncogenes *MYC*, *cyclin D1* (*CCND1*) and *ERBB2* respectively. The *MYC* gene has been found to be over-expressed, by gene amplification or regulatory means, in 6-23% of breast cell lines or breast tumours (Escot *et al.*, 1986; Mariani-Costantini *et al.*, 1988; van de Vijver, 1993). *MYC* codes for a protein

which, when dimerised to other proteins, transcriptionally activates genes that promote progression of the cell cycle (Amati *et al.*, 1992; Kretzner *et al.*, 1992). Patients with *MYC* amplified tumours have been shown to have a very poor prognosis (Berns *et al.*, 1992a, 1992b).

ERBB2 was originally identified as a result of transfection studies with DNA from chemically induced rat neuroglioblastomas (Schechter *et al.*, 1984). This gene was named *neu* and was shown to be related but distinct from the proto-oncogene *c-erbB*. Two other groups also isolated *erbB*-related genes and named them *HER-2* and *c-erbB-2* (Coussens *et al.*, 1985; Semba *et al.*, 1985). Comparisons of sequence analysis and chromosomal mapping studies identified all three (*neu*, *HER-2* and *c-erbB-2*) as the same gene, *ERBB2*. Importantly, King *et al.* (1985) determined that the same gene was amplified 5-10 fold in a breast cancer cell line. *ERBB2* is a member of the tyrosine kinase family and is closely related to the gene encoding the epidermal growth factor receptor gene, *ERBB1*. The prognostic significance of the amplification of *ERBB2* has been the focus of much interest since it was reported to be amplified in 30% of human breast cancers and an important predictor of disease recurrence and death for patients with axillary node-positive disease (Slamon *et al.*, 1987). Amplification rates of *ERBB2* from several studies range from 10-46% with an overall rate of 20% (Slamon *et al.*, 1989; Clark and McGuire 1991). The correlation between amplification and clinical outcome, however, has not been conclusively confirmed as some reports claimed that amplification of *ERBB2* is not related to clinical outcome (Ali *et al.*, 1988; Zhou *et al.*, 1989). In their follow-up to the Slamon *et al.* (1987) study, Clark and McGuire (1991) stated that amplification of the *ERBB2* gene did not correlate with either disease-free or overall survival in patients with primary breast cancer. Their results for 185 node-positive patients were not significant in

multivariate analysis and that while *ERBB2* amplification may be of biological importance it would only be of marginal use as a prognostic factor for predicting clinical outcome.

The chromosomal region 11q13 is amplified in 15-20% of breast tumours (Lidereau *et al.*, 1988; Theillet *et al.*, 1989, 1990). Amplification of this region in breast cancer is associated with poor prognosis, lymph node involvement, and oestrogen and progesterone receptor positivity (Borg *et al.*, 1991; Berns *et al.*, 1992c; Schuuring *et al.*, 1992a). This region contains several candidate genes as the target for amplification. Two of these genes, identified as genes activated by mouse mammary tumour virus, are members of the fibroblast growth factor family, *FGF3* (*INT-2*) and *FGF4* (*HST1*), which lie 40 kb apart on chromosome region 11q13 (Casey *et al.*, 1986; Yoshida *et al.*, 1988). Both are amplified in primary human breast cancers (Zhou *et al.*, 1988; Theillet *et al.*, 1989) but they are transcriptionally silent in normal breast tissue and breast tumours and it would appear that they are merely carried on the amplification unit with the actual oncogene(s) (Liscia *et al.*, 1989; Theillet *et al.*, 1989). The search for genes present on the 11q13 amplicon that are co-amplified and overexpressed in all tumours with the 11q13 amplification precipitated the identification of the genes *CCND1* and *EMS1* (Schuuring *et al.*, 1992b). Cyclin D1, which is encoded by *CCND1*, promotes progression through the G1 phase of the cell cycle by regulating the activity of two cyclin-dependent protein kinases, CDK4 and CDK6 (Scherr 1993). *EMS1* is involved in regulating the interactions between components in the adherens junctions (Schuuring *et al.*, 1993). The complex pattern of amplification of the 11q13 region suggests that there are several foci of amplification which confer a selective advantage to the tumour cells, both *CCND1* and *EMS1* fulfil the criteria as key oncogenes on 11q13 (Schuuring, 1995).

1.3.4 Loss of Heterozygosity (LOH)

Knudson's "two hit" hypothesis for the functional inactivation of a tumour suppressor gene describes how the first hit is recessive and may be carried in the germline or be acquired somatically and is often a point mutation (Knudson, 1971). The second mutation, always occurring at the somatic level, affects the remaining normal allele and usually involves a chromosomal mechanism resulting in the loss of genetic material. In a tumour, a reduction from heterozygosity to hemi- or homozygosity by this process is known as loss of heterozygosity (LOH) and can be determined by investigating polymorphic markers which allow the distinction of paternal and maternal chromosomes. The genetic mechanisms proposed by Knudson have been identified for numerous tumour suppressor genes and as a consequence, frequently observed LOH at specific chromosomal regions is indicative of the presence of a tumour suppressor gene. Therefore deletion of genomic regions may be important in the initiation and progression of breast cancer due to the inactivation of tumour suppressor genes.

The analysis of all chromosomes by LOH in a series of tumour specimens is termed allelotyping (Vogelstein *et al.*, 1989). In breast cancer LOH is very common and occurs in nearly all chromosomes (Devilee and Cornelisse, 1994). Those exhibiting the most frequent losses include chromosomes 1, 3p, 6q, 8p, 11q, 13q, 16q, 17, and 22q (Devilee *et al.*, 1989; Sato *et al.*, 1990; Devilee and Cornelisse, 1994). Only a small number of tumour suppressor genes, eg. *RB1* and *p53*, are currently known to be located on these chromosomal regions. Loss of chromosome 17p13, the chromosomal location of *p53*, is one of the most frequent in breast carcinomas (Coles *et al.*, 1990; Anderson *et al.*, 1992). Mutation of the *p53* gene has been observed in up to 50% of breast tumours (Osborne *et al.*, 1991; Coles *et al.*, 1992) and is associated with a poor prognosis (Thor *et al.*, 1992; Allred *et al.*, 1993). Chromosomal regions 11q22-23.1 and 11q25-ter show frequent allelic loss in sporadic breast cancer (Koreth *et al.*, 1997) and patients with primary breast

cancers with LOH of 11q23 have a more aggressive post-metastatic disease course than patients without these deletions (Winquist *et al.*, 1995). LOH on chromosome 17 has been particularly well studied and multiple regions of LOH have been identified (Cropp *et al.*, 1993). *BRCA1* and *p53* are likely to be two of the target genes for LOH on chromosome 17 but other genes such as *NME*, *MDC* and prohibitin (*PHB*) may also be targets for LOH. The *NME1* gene at chromosome 17q23 encodes components of a protein whose loss has been associated with metastasis (Steeg *et al.*, 1988) and the *MDC* gene has been shown to be somatically rearranged in 2 primary breast tumours (Emi *et al.*, 1993). Thus determining the genes involved the LOH events is as complex as identifying those targetted in gene amplification.

1.3.5 Genetic Evolution of Breast Cancer

The molecular genetic events underlying breast carcinogenesis are not completely understood and describing any pathway of evolution suffers from the limited information available from precursor lesions as well as the genetic heterogeneity of the disease. Data from a cytogenetic and flow cytometric study indicates that the initial events in the process of tumourigenesis include structural chromosomal aberrations (LOH and gene amplification) and losses of whole chromosomes (Dutrillaux *et al.*, 1991). This net loss of material leads to a state termed hypodiploid which refers to a cellular DNA content less than that found in normal cells. It has been proposed that this state then gives rise to hyperdiploidy, the most common chromosomal abnormality in breast tumours, by a process of endoreduplication (Devilee and Cornelisse, 1994). Alternatively, endoreduplication may be an early event in tumour development and only after this step do chromosomal gains and losses influence tumourigenesis (Ewers *et al.*, 1984). Hypotetraploid DNA content is the most common form of hyperdiploidy with near-diploid, triploid, tetraploid and hypertetraploid tumours constituting 5-15% of aneuploid

breast cancers each. Only 1-2 % of breast cancers persist as hypodiploid (Devilee *et al.*, 1994). These hypodiploid tumours often show the presence of endoreduplicated subclones and suggests that progressive chromosomal loss in hypodiploid tumours increases the likelihood of endoreduplication (Dutrillaux *et al.*, 1991). The resultant hyperdiploid clone would then have a growth advantage due to a gene dosage effect rather than the gain of important structural rearrangements. Recent investigations of ductal carcinoma *in situ* (DCIS) of the breast by CGH have identified a variety of gains and losses of chromosomal regions which may help to pinpoint specific events involved with disease initiation or progression (James *et al.*, 1997). Although the number of samples studied was small, it was clear that fewer gains/losses were observed than in invasive carcinomas and that the aberrations were similar to those found in invasive breast cancer. Similarly, the study of hypodiploid breast cancer has shown gains and losses of chromosomal regions common to those found in invasive carcinomas indicating that they share a common genetic evolution pathway where structural chromosomal aberrations precede gross DNA ploidy changes (Tanner *et al.*, 1998). Further evaluation of preinvasive and premalignant lesions may help to elucidate the genetic aberrations essential to the initiation and progression of breast cancer.

1.4 Tumour Suppressor Genes

1.4.1 Definition of a Tumour Suppressor Gene

Many of the genes associated with inherited cancer syndromes, including *BRCA1* and *BRCA2*, have been classed as tumour suppressor genes. What defines a tumour suppressor gene? No single characteristic is sufficient to define a tumour suppressor gene but three features are central to our current understanding. Firstly, the presence in the

tumour of loss-of-function mutations of both alleles of the gene. Secondly, there should be gene inactivation in inherited as well as sporadic tumours and thirdly, the wild type allele should have the ability to rescue the tumour phenotype *in vitro*. Many tumour suppressor genes such as *RB*, *p53*, *WT1*, *APC* and *p16* fulfil these criteria. Yet *BRCA1* and *BRCA2* are rarely mutated in sporadic breast tumours. Also, the tumour suppressor genes that predispose to Hereditary Non-Polyposis Colorectal Cancer (HNPCC), *hMLH1* and *hMLH2*, are involved in mismatch repair and create a mutator phenotype and are unlikely to restore finite proliferation by their reintroduction. To date tumour suppressor genes have been found to act as a crux in pathways that regulate proliferation, differentiation, apoptosis and the response to DNA damage.

1.4.2 *RB* as a Paradigm of a Tumour Suppressor Gene

The retinoblastoma gene (*RB*) predisposes to retinoblastoma, a cancer of the eye in children arising from cells of the embryonal neural retina, and was the first tumour suppressor gene to be isolated. The retinoblastoma gene has been used as a paradigm of the tumour suppressor gene. Approximately 10 % of cases have a family history of the disease and are termed familial, with the remaining 90 % occurring at an older age and being sporadic, a term used when there is no family history. Given these observations, Knudson proposed that all types of retinoblastoma require two rate-limiting genetic hits (mutations) that are present in every retinoblastoma cell (Knudson, 1971). By extension the hypothesis proposes that both alleles of a tumour suppressor gene need to be inactivated in order to contribute to oncogenesis. In the case of an individual with a germline defect in one allele, regarded as the first hit, only a second somatic hit in the normal allele is required to lead to the development of cancer. This second mutation often proceeds via a chromosomal mechanism resulting in the loss of the wild-type allele and is detectable as loss of heterozygosity (LOH) of polymorphic DNA markers in the vicinity of

the gene. Such LOH is therefore interpreted as evidence for the presence of a tumour suppressor gene close to the deleted markers. Confirmation of the first hit came from studies of alleles of the esterase D gene, mapped to chromosome region 13q14, which were linked to familial predisposition to retinoblastoma (Sparkes, 1983). The second hit was confirmed to be in the other allele at this locus by loss of heterozygosity analysis which involved loss of sequence from the chromosome not carrying the affected retinoblastoma locus (Benedict *et al.*, 1983; Cavanee *et al.*, 1983; Cavanee *et al.*, 1985). With the subsequent cloning of the *RB* gene, Knudson's proposals were corroborated by detecting a single germline mutation in children with an inherited predisposition and two somatic mutations in sporadic cases (Friend *et al.*, 1986; Yandell *et al.*, 1989; Hogg *et al.*, 1993). This demonstrated that inactivation of the gene alone by loss-of-function mutations caused the development of both familial and sporadic cancers.

Another fundamental characteristic of tumour suppressor genes is that reintroduction of a wild type gene into cells lacking the functional tumour suppressor product would reverse the tumour phenotype. This proposal comes from studies where cancer cells were fused with non-transformed cells and the malignant phenotype was suppressed. This has been shown for *RB*, whereby introduction of an intact chromosome 13 into cells lacking a functional *RB* has prevented cells from proliferating (Stanbridge, 1992). From the studies outlined above *RB* has become the model for a tumour suppressor gene.

1.5 The Cell Cycle

The cell cycle can be described in terms of four different stages, G1, S (DNA synthesis) phase, G2 and M (mitosis) phase (Figure 1.4). The cell cycle ends with completion of M phase and is followed by G1 progression, S phase, G2 progression and

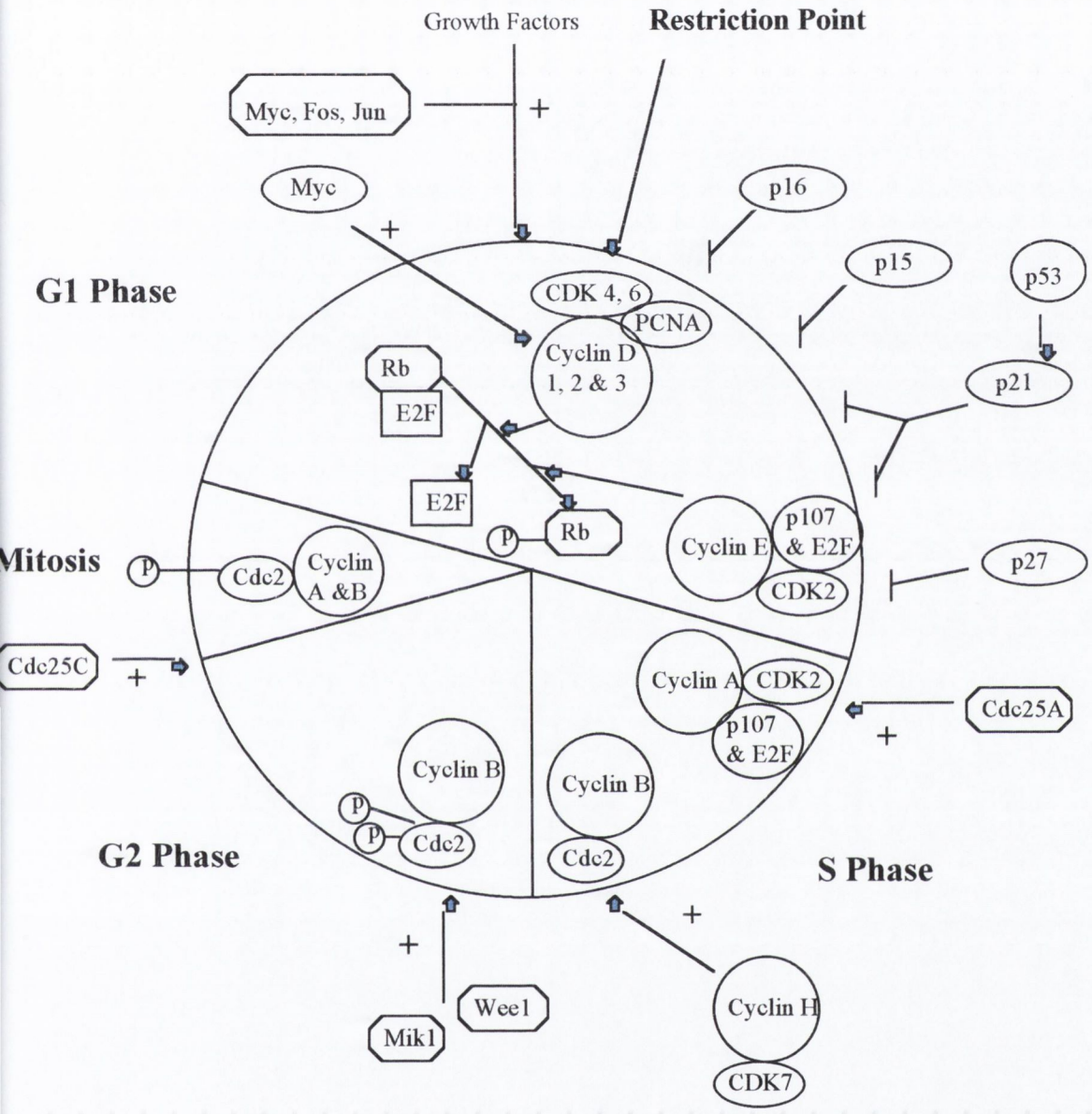


Figure 1.4. The Cell Cycle. A highly schematic view of points of action of mammalian cyclin-CDK complexes in the cell cycle (adapted from Hunter and Pines, 1994).

then M phase, if the conditions are propitious for a further cell cycle. When the cell is not involved in a cell cycle it enters a quiescent stage known as G0. The fundamental task of the cell cycle is to ensure that during the S phase chromosomes are replicated once and during the M phase that identical chromosomal copies are segregated equally to two daughter cells. These events form an ordered developmental sequence that ensures that each newly divided cell receives a full complement of the hereditary material. Therefore it is critical for the cell to ensure that these episodes occur in the correct order and at the correct time. This is maintained by a number of complex controls acting at different stages.

Under normal circumstances the S phase always occurs after the M phase of the previous cell cycle. This ensures that G2 cells do not enter another S phase which would result in a second round of DNA replication and therefore a change in cell ploidy. For cells to enter S phase they must complete progression through G1. Cells entering G1 from mitosis require serum mitogens continuously until several hours before onset of the S phase; after this point they become relatively serum independent. This point is termed the restriction point (Pardee, 1989). Having passed the restriction point the cell is commited to a complete cycle through to the end of the M phase although this can be affected by DNA damage or metabolic disturbances.

1.5.1 Cyclins, Cyclin-Dependent Kinases and the G1/S Transition

Positive regulators of the cell cycle include a group of related proteins called the cyclins (D, E, A, B). The cyclins are the positive regulatory subunits of a class of related protein kinases named cyclin-dependent kinases, CDKs. At least 10 different cyclins and 7 CDKs are known and can form at least 15 different cyclin-CDK complexes (Lew and Reed, 1992; Sherr, 1993). Having been stimulated to enter the cell cycle from quiescence (G0 phase), the first group of cyclins to be expressed are the D-type cyclins. There are three D-type cyclins (D1, D2 and D3) and they are, in part, expressed in a cell-type

specific manner with most cells expressing D3 and either D1 or D2. Acting as growth factor sensors, the expression of the D-type cyclins depends on mitogenic stimulation rather than on the cell's position in the cycle (Sherr, 1993). The induction of these cyclins is part of the delayed early response to growth factor stimulation whereas the immediate early response genes include *c-fos*, *c-jun* and *c-myc*. Once expressed, the D-type cyclins form active kinase complexes in mid- and late G1 with their catalytic partners, CDK4 and CDK6 (Sherr, 1993).These complexes then phosphorylate key substrates that are required for the progression of the cell cycle and their catalytic activity continues as long as there is mitogenic stimulation. Removal of such extracellular stimuli results in cessation of cyclin D synthesis and the cell exits the cycle. If this occurs before the restriction point the cell is incapable of entering the S phase (Matsushime *et al.*, 1991) but loss of cyclin D1 activity later in the cell cycle is without effect (Quelle *et al.*, 1993; Baldin *et al.*, 1993).

The expression of cyclin E comes after the synthesis of cyclin D1 in G1, is normally periodic and reaches a maximum at the G1 - S phase transition (Dulic *et al.*, 1992; Koff *et al.*, 1992). The catalytic partner of cyclin E is CDK2. Over-expression of both cyclin D1 and cyclin E causes a reduction in the G1 interval greater than that caused by over-expression of either one alone, indicating that they regulate different processes in the G1 - S phase transition (Resnitzky and Reed, 1995). Cyclin A is induced shortly after cyclin E and forms a complex with CDK2 in S phase and CDC2 in late S and G2 phase. Cyclin A is thought to regulate entry into S phase, inactivation of cyclin A by antibody injection blocks entry into S phase (Girard *et al.*, 1991; Pagano *et al.*, 1992). The mitotic cyclin, cyclin B, regulates mitosis in association with CDC2 (Dunphy, 1994). Cyclin B is induced in S phase, cyclin B - CDC2 complexes accumulate towards M phase and then cyclin B rapidly degrades during mitosis.

1.5.2 Regulation of CDKs

The cyclins are expressed in a cell cycle-dependent manner and in this way they control the activation of their associated CDKs. CDK activity also depends upon both negative and positive phosphorylations, and can be affected by at least two families of CDK inhibitory proteins (CKIs).

Phosphorylation of specific residues of CDKs is required for the activation of its kinase activity by making their catalytic pocket accessible for the protein substrate (Jeffrey *et al.*, 1995). This phosphorylation is controlled by a CDK Activating Kinase (CAK) which requires cyclin H for its activity (Makela *et al.*, 1994; Fisher and Morgan, 1994). CAK phosphorylates cyclin-CDK2, cyclin-CDC2 and cyclin D2-CDK4 complexes. CDKs are also regulated by another phosphorylation event through the *wee-1/mik-1* related protein kinases (Dunphy, 1994).

The CDK inhibitors (CKIs) can be divided into two groups, those that associate with a complex containing a cyclin, a CDK and the proliferating cell nuclear antigen (PCNA) (Xiong *et al.*, 1992; Zhang *et al.*, 1993) and those that act as competitive inhibitors of D-type cyclins by binding to CDK4 and CDK6 (Serrano *et al.*, 1993; Hirai *et al.*, 1995; Hannon and Beach, 1994). The first group includes the proteins $p21^{CIP1}$, $p27^{KIP1}$ and $p57^{KIP2}$. The first of these to be identified, $p21^{CIP1}$, binds to and inhibits a wide variety of cyclin-CDK complexes including cyclin D-CDK4, cyclin E-CDK2 and cyclin A-CDK2 (Harper *et al.*, 1993; Gu *et al.*, 1993; Xiong *et al.*, 1993; El-Deiry *et al.*, 1994; Noda *et al.*, 1994). $p21^{CIP1}$ expression is controlled, to a certain extent, by p53 (El-Deiry *et al.*, 1994). DNA damage results in the increase of p53 levels, which in turn induce synthesis of $p21^{CIP1}$. The increase of $p21^{CIP1}$ levels inhibits cyclin-CDK kinase activity and allows time for DNA repair to take place before the cell enters the S phase. As is the case for $p21^{CIP1}$, $p27^{KIP1}$, inhibits a wide variety of cyclin-CDK complexes (Polyak *et al.*, 1994; Toyoshima and Hunter, 1994). Levels of $p27^{KIP1}$ are high in quiescent cells and low during the cell

cycle. Residual $p27^{KIP1}$ is sequestered by cyclin D-CDK complexes thus removing repression of cyclin E-CDK2 and cyclin A-CDK2 activity in cycling cells (Toyoshima and Hunter, 1994).

The second group of CKIs are specific polypeptide inhibitors of CDK4 and CDK6, but not other CDKs, and are known as the INK4 proteins. The four known INK4 proteins are $p16^{INK4a}$, $p15^{INK4b}$, $p18^{INK4c}$ and $p19^{INK4d}$. By binding to and inhibiting CDK4 and CDK6 they can cause a G1 phase arrest thus indicating that they are potent inhibitors of cell cycle progression (Serrano *et al.*, 1993; Hirai *et al.*, 1995; Hannon and Beach, 1994).

1.5.3 Cell Cycle Checkpoints

Of fundamental importance to the understanding of the initiation of cancer and its progression is how specific checkpoints operate within the cell cycle and what are the consequences of their failure. Checkpoints interrupt cell cycle progression when damage to the genome or spindle is detected, or when cells have failed to complete an event. Three checkpoints have been described. These are the DNA damage checkpoint, a spindle checkpoint and a spindle pole body duplication checkpoint. The DNA damage checkpoint involves either excision repair pathways that mend modified bases, or breaks in the DNA that are repaired by non-homologous end joining or by the creation of new telomeres. The DNA damage checkpoint occurs at three different stages of the cell cycle. One is at the G1/S transition, a second ensures successful progression through the S phase and a third at the G2/M transition (Paulovich *et al.*, 1997). Many of the genes involved at one arrest point are also required for arrest at the other stages. Failure of one or more checkpoints to arrest the cycle at any of these stages of the cell cycle is usually found in cancer cells. The consequences of a failed DNA damage checkpoint depend on the type of damage and on the stage of the cycle at which it presents.

To avoid replicating damaged DNA the DNA damage checkpoint at the G1/S stage arrests the cells by responding to certain types of damage and whilst they still have a G1 DNA content. During this period time is allowed for DNA to be repaired before proceeding into the S phase of the cell cycle. Again, in response to DNA damage, the rate of DNA replication during the S phase can be slowed so as to allow repair to take place. Cells at the S phase risk not completing replication and misincorporating bases. The consequences of not arresting the cell cycle in response to DNA damage at either of these stages is determined, in part, by the type of DNA damage and the effect that replicating the damaged template incurs. For the case of double-stranded breaks there is a high risk of loss of heterozygosity, DNA amplification, chromosome loss/gain and chromosomal rearrangement or truncation. Also base mismatches that are not repaired by the mismatch repair system or other lesions that are not corrected result in the fixation of the mutation after replication or subsequent repair synthesis. Initial lesions may also result in the formation of secondary lesions.

For G2/M transition the DNA damage checkpoint is activated only by double-stranded breaks and progression from G2 to M is dependent on an intact genome. At this stage failure to arrest and create a delay in which to repair the damage can result in chromosomal loss, creation of a truncated chromosome or lead to chromosomal rearrangement and gene amplification. The DNA damage checkpoint appears to react to DNA damage that would cause the most serious consequences if allowed to continue on to the next stage of the cell cycle.

1.5.4 Genes that Control the Cell Cycle and Tumour Suppressor Genes

Two of the most extensively investigated tumour suppressor genes, *p53* and the retinoblastoma gene, *RB*, play a critical role in the maintenance and regulation of the cell cycle. The p53 protein is a transcription factor that increases the transcription of several

genes that have a variety of p53-dependent functions in the cell (Ko and Prives, 1996). Among the products of these genes is p21, which can inhibit cyclin-cyclin-dependent kinases (cyclin-CDKs), bind CDKs and cyclins and arrest the cell cycle. The human MDM2 protein, the product of an oncogene, is also transcriptionally activated by p53 but inactivates p53 mediated transcription and creates an autoregulatory loop with p53 activity (Levine, 1997). The *RB* gene product acts as a signal transducer that connects the cell cycle "clock" with the transcriptional machinery. In this way Rb controls the expression of many genes that are involved in governing the progression of the cell cycle.

Central to the regulation of the G1/S transition in the cell cycle is the p16-cyclinD1-CDK4-Rb pathway. Once p53 is activated, by DNA damage, it initiates the transcription of *p21*. p21 then interacts with the p16-cyclin D1-CDK4-Rb pathway by binding to the cyclin D1-CDK4 complex. A single p21 bound per complex allows CDK activity but if two molecules of p21 bind per complex then CDK activity is inhibited and cell cycle progression is halted. Both p16 and p21 are inhibitors of cyclin-dependent kinases. p21 is also capable of binding to PCNA and by doing so stops some of its functions and prevents DNA replication. Hence *p53* is central to cell cycle regulation. Rb also connects with the pathway but at a stage further downstream in the process. The cyclin D1-CDK4 complex's main function is the regulation of the retinoblastoma protein by phosphorylation. The cyclin-dependent kinase, CDK4, is directed by cyclin D1 to the retinoblastoma protein, which it then phosphorylates. The majority of the retinoblastoma protein is hyperphosphorylated during the last hours of G1 as compared to the under phosphorylated form in the preceding period of G1. Only under favourable conditions is Rb phosphorylated, thus allowing completion of the cell cycle. By existing in either the hyperphosphorylated or hypophosphorylated state Rb can then regulate E2F-DP transcription factor complexes which go on to regulate genes that are required for the S phase of the cell cycle. Once Rb is phosphorylated, E2F-DP proteins are released from the

Rb complex therefore relieving their repression or activating their transcription. Rb remains in the hyperphosphorylated form until it emerges from the M phase of the cell cycle.

p53 has also been associated with the G2/M phase checkpoint and the G0/G1/S phase transition. Mutation of p53 could result in genomic instability at the G2/M phase checkpoint and experiments with mutant p53 by Del Sal *et al.* (1995) suggest that p53 acts at the G0/G1/S phase transition by direct protein-protein signalling and not as a transcription factor.

1.5.5 Cancer and Genes Involved in the Cell Cycle

The p16-cyclin D1-CDK4-Rb pathway is central to the regulation of the cell cycle and alteration or mutation of its constituent genes and gene products has provided a great insight to the understanding of human malignancies. In retinoblastomas, small cell lung carcinomas, sarcomas and bladder carcinomas loss of the Rb function is ascribable to mutations in the *RB* gene (Horowitz *et al.*, 1990). In small cell lung carcinomas where normal Rb and p16 expression were investigated, the carcinomas lacking normal Rb expression had the wild type p16 gene product but those lacking the p16 protein had normal levels of Rb (Otterson *et al.*, 1994). p16 is deleted in many cancers such as glioblastomas and pancreatic carcinomas (Jen *et al.*, 1994; Schmidt *et al.*, 1994; Caldas *et al.*, 1994) and is turned off by heavy methylation in some cancer cells (Levine, 1997). The cyclin D1 gene is amplified and overexpressed in oesophageal cancer, squamous cell carcinomas and about 16 % of breast cancers (Lammie *et al.*, 1991; Jiang *et al.*, 1992). The *CDK4* gene has also been found to be mutated or amplified, as in glioblastomas (Schmidt *et al.*, 1994).

Mutation of the *p53* gene is found in over 50 % of all human cancers (Hollstein *et al.*, 1994) although some tumour types never appear to acquire *p53* mutations (Lutzker

and Levine, 1996). The selection of *p53* mutations in tumours is predominantly for mutations located in regions that are involved in sequence specific DNA binding. The combination of selection and environmental mutagens produces the mutational spectrum seen for *p53* in human cancers.

1.6 The Prohibitin Gene (*PHB*)

1.6.1 Isolation and Identification of the Prohibitin Gene (*PHB*)

The genes that negatively regulate cell proliferation are of vast importance to the processes of development and cancer. Such antiproliferative activity has been demonstrated in a variety of cell or tissue types (Stanbridge, 1992). The existence of an antiproliferative activity was demonstrated in human diploid fibroblasts (HDF) by fusion of dividing and non-dividing, senescent HDF. This resulted in inhibition of cell division in the dividing cells (Norwood *et al.*, 1985). Microinjection of poly (A) enriched RNAs from a number of sources, including non-dividing T cells, normal rat liver and senescent HDF, have all been shown to inhibit DNA synthesis in dividing cells (Lumpkin *et al.*, 1985; Lumpkin *et al.*, 1986; Pepperkok *et al.*, 1988). By fractionating rat and human liver mRNA on a sucrose gradient and using the fractions in microinjection studies a peak of inhibitory activity was associated with mRNA transcripts of 2 kb in length (Lumpkin *et al.*, 1985; Pepperkok *et al.*, 1988). Nuell *et al.* (1989) subsequently isolated cDNA clones associated with this inhibitory activity by differential hybridization to RNA from normal liver as opposed to RNA from regenerating rat liver. The cDNAs from these clones were used to identify their cognate RNA which were then tested for antiproliferative activity by microinjection assays. From these experiments a single clone of 122 bp was determined to code for part of a growth inhibitor and was designated as "prohibitin". It was further

shown that the prohibitin mRNA could inhibit DNA synthesis in normal fibroblasts in culture (McClung *et al.*, 1989).

A prohibitin cDNA sequence with a complete open reading frame was subsequently isolated from a rat intestine cDNA library (Nuell *et al.*, 1991). The clone was predicted to code for a protein of 272 amino acids and a molecular weight of approximately 30 kDa. Prohibitin was shown to be an intracellular protein which is expressed in all rat tissues tested and its mRNA has the ability to block DNA synthesis in cancer cells. Based on the observation that regions of the rat *prohibitin* gene are highly conserved with the *Drosophila* gene *Cc*, a PCR product from a region of the human *prohibitin* gene was generated. This product was then used to screen both human cDNA and genomic cosmid and phage libraries to isolate the whole coding sequence of the human *prohibitin* gene (Sato *et al.*, 1992). As with the rodent *prohibitin* gene, the human protein was predicted to have 272 residues and was expressed in all tissues examined.

1.6.2 Amino Acid Composition

The human *prohibitin* gene shares a 99.6 % identity with the rat *prohibitin* gene at the amino acid level and is 90.3 % identical at the nucleotide level (Sato *et al.*, 1992). Most of the differences at the nucleotide level are third position changes and the only amino acid substitution is a conservative change of the phenylalanine residue at codon 107 for a tyrosine residue. A high level of conservation at the amino acid level is seen for the *prohibitin* gene in all species that have been examined. The human sequence is 75 % identical to *Drosophila* and 52 % identical to yeast at the protein sequence level. The full or partial gene sequence of homologues of the *prohibitin* gene have been detected in a wide range of other species. These include yeast (McClung *et al.*, 1995), rat (Nuell *et al.*,1991), mouse (Terashima *et al.*, 1994), *Drosophila* (Eveleth and Marsh, 1986), *Arabidopsis thaliana* and *Nicotiana tabacum* (tobacco) (Snedden and Fromm, 1997) and

a selection of mammals (Matsuyama *et al.*, 1997). Homology to prohibitin has even been found in sequences from the ancient cyanobacterium *Synechocystis* (Coates *et al.*, 1997). The mouse, rat and human prohibitin proteins contain 272 amino acids, the *Drosophila*, *Arabidopsis*, tobacco and yeast proteins have 276, 277, 280 and 287 amino acids respectively (Figure 1.5). The differences are explained by a small number of amino acids at the amino and carboxy-terminals of the protein, in the case of yeast this constitutes a region of 13 amino acid residues close to the C-terminal of the protein and 2 additional amino acids at the amino-terminus. Such a high level of conservation in many different organisms suggests that throughout evolution *prohibitin* has remained an important gene.

Computer-generated models of the secondary structure of the prohibitin protein by McClung *et al.* (1995) indicate a number of features of note. A hydrophobic α-helix is predicted to be present at the amino-terminus. This is followed by an internal β-sheet region and then another α-helical region is predicted towards the carboxy-terminus. Ikonen *et al.* (1995) noted that if the 18 amino-terminal residues are plotted on a helical wheel they resemble a mitochondrial targeting signal which forms an amphiphilic α-helix after insertion in lipid membranes. This agreed favourably with their report that prohibitin is localised to the mitochondria. Prohibitin also contains a number of possible sites for *N*-glycosylation, *N*-myristolation, tyrosine sulfation, and phosphorylation by protein kinase C and casein kinase II. The significant alteration in mobility of the protein that would be induced by *N*-glycosylation and *N*-myristolation is not seen when the prohibitin protein is analysed by two-dimensional Western blots (McClung *et al.*, 1995), indicating that these sites may not be active. Three isoforms were, however, detected by this method and these could be explained singularly or by a combination of; phosphorylation, sulfation and/or mono(ADP-ribosyl)ation. Prohibitin has been isolated from cultured rat ovarian granulosa cells in a phosphorylated and unphosphorylated form of 28 kDa (Thompson *et al.*, 1997). A number of other observations of a phosphorylated form of prohibitin have been made

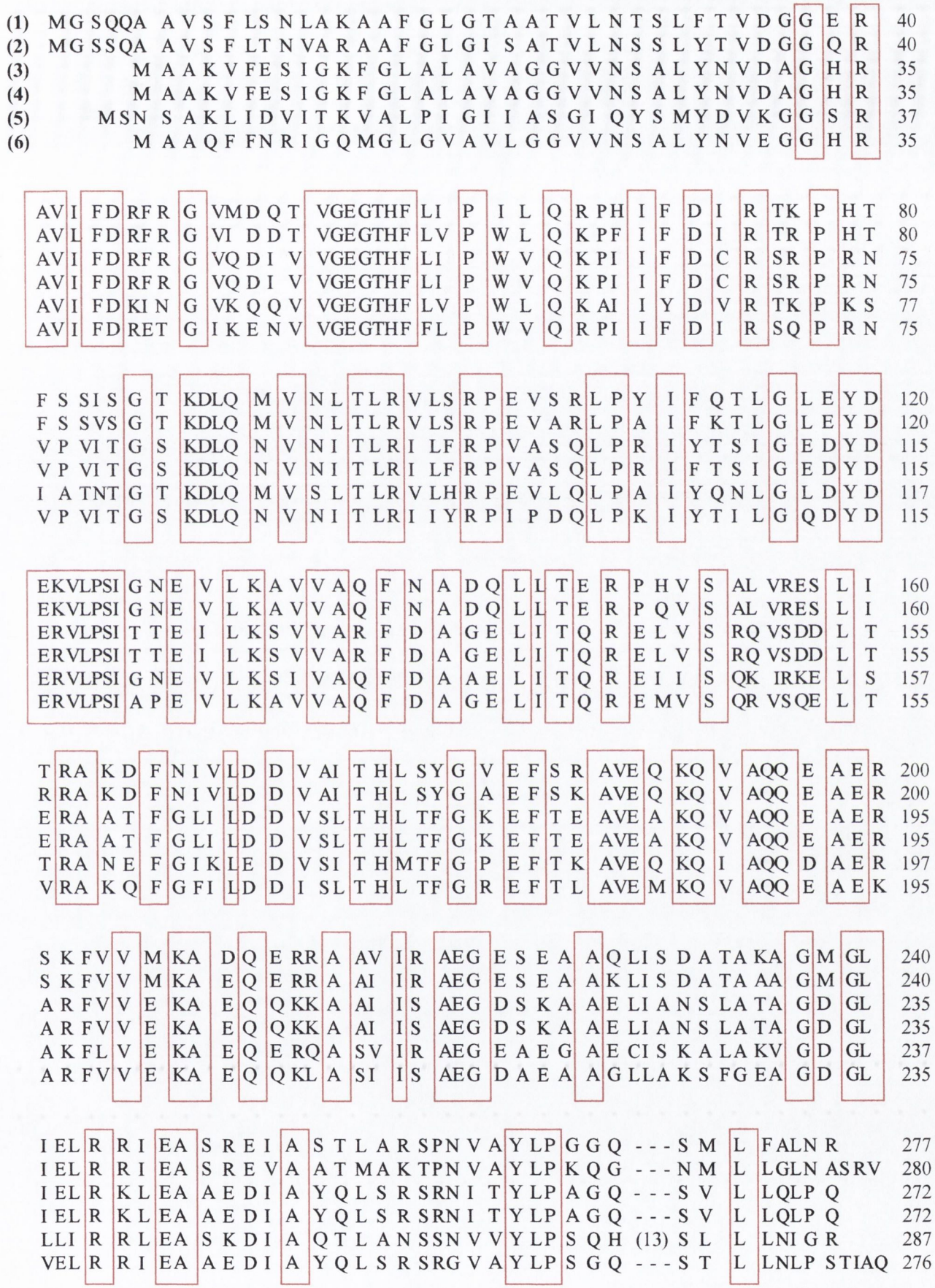


Figure 1.5. Comparison of prohibitin amino acid sequences: (1)=Arabidopsis, (2)=Tobacco, (3)=Mouse and Rat, (4)=Human, (5)=Yeast, (6)=Drosophila. Identical amino acids in all 6 sequences are boxed in red. Sequence analysis from Snedden and Fromm, 1997, and Lamers and Bacher, 1997.

(McClung *et al.*, 1995). Whether prohibitin is modified in other ways is currently unknown.

1.6.3 Subcellular Location

Early reports suggested that prohibitin existed in 2 different forms, one in the nucleus and one in the cytoplasm, and that it is associated with a large protein complex of about 500 kDa (Wood *et al.*, 1991). The bulk of the protein was reported to be in the cytoplasm (Nuell *et al.*, 1991). Several reports have since described diverse locations for prohibitin within the cell. Whilst investigating the identity of proteins that are intracellularly associated with mIgM and mIgD in mouse B lymphocytes, Terashima *et al.* (1994) found 3 proteins of 32, 37 and 41 kDa, that are associated with the membrane bound IgM (mIgM) antigen receptor complex but not with the mIgD complex. The 32 kDa protein was identified as prohibitin on the basis of its amino acid sequence and the 37 kDa protein was named as a B-cell receptor associated protein, BAP37. Terashima *et al.* (1994) concluded that prohibitin and BAP37 are intracellular proteins that associate with the transmembrane domain of mIgM in B cells. They also found prohibitin and BAP37 in protein complexes of ~120-200 kDa but did not comment on any possible interaction between prohibitin and BAP37 in these complexes. Coates *et al.* (1997) reported that prohibitin and BAP37 are physically associated although the mode or mechanism of their interaction with each other and/or separate proteins was not shown.

A second location for the prohibitin protein was found when it was localised to mitochondria. This was demonstrated by expressing rat prohibitin with a carboxy-terminal c-Myc epitope in baby hamster kidney cells and visualising the proteins by using a Myc-antibody (Ikonen *et al.*, 1995). The mitochondrial location for prohibitin has been confirmed in rat liver, *Drosophila* (McClung *et al.*, 1995; Dell'Orco *et al.*, 1996) and human, rodent and *Xenopus laevis* cell lines (Coates *et al.*, 1997). The mitochondrial

location was also confirmed in plants by Snedden and Fromm (1996). In these experiments the majority, but not all, of the prohibitin protein localised to the mitochondria. The region of localisation within the mitochondria was narrowed to the mitochondrial inner membrane by Western blot analysis after the mitochondria had been fractionated into constituent matrix and inner and outer membranes (McClung *et al.*, 1995). Such varying cellular locations as the mitochondria and the plasma membranes of B cells suggest that prohibitin is a protein that can exert its function(s) at distinct levels. Lamers and Bacher (1997) concluded that the cellular location of prohibitin is determined by the differentiation and activation state of the cell, although the possibility that either of the observed locations was incorrect was not ruled out. They also stated that the alternate locations may, in part, be explained by the finding that prohibitin exists almost exclusively in large complexes in the detergent lysates of fibroblast and B cell lines and is rarely found uncomplexed.

1.6.4 Chromosomal Location

The human *prohibitin* gene was localised to chromosome 17q21-q22 by probing *Eco*R1 digested genomic DNA from 33 mouse-human hybrid cell lines with a fragment of the human prohibitin genomic DNA (White *et al.*, 1991). No discordance was seen between *prohibitin* and a location on human chromosome 17. Sato *et al.* (1992) determined the gene location as 17q12-21 by fluorescent *in situ* hybridization and, by comparison to the previous location of 17q21-q22, assigned the gene to the chromosomal region 17q21.

Southern blot analysis of genomic DNA indicated that the *prohibitin* gene could be a member of a gene family. Screening of a human genomic cosmid library and a human fetal-brain cDNA library with a human prohibitin cDNA insert as a probe revealed five cosmid clones carrying sequences homologous to the *prohibitin* gene (Sato *et al.*, 1993). Four of these were intronless pseudogenes whilst the fifth was identified as the functional

prohibitin gene. The four pseudogenes (*PHBP1-4*) are located on chromosomes 6q25, 11p11.2, 1p31.3 and 2q21 respectively and were found to deviate in sequence from the functional *prohibitin* gene in the regions examined. For example, the DNA sequence of *PHBP1* contains 84 nucleotide substitutions, one single base pair deletion and one three base pair deletion in the region of *PHBP1* corresponding to the coding sequence of the functional *prohibitin* gene. The large number of nucleotide substitutions and base pair deletions identified in *PHBP1* and the other pseudogenes indicate that they are unlikely to be functional. Furthermore, sequence analysis of the 5' region of *PHBP1* shows that it does not contain the transcriptional regulatory elements seen in the functional *PHB* gene. Finally, *PHBP1* is flanked by AT-rich sequences, a feature which has been reported for other processed pseudogenes (Vanin, 1985). From this analysis Sato *et al.* (1993) concluded that the *prohibitin* gene family consisted of only one functional *PHB* gene on 17q21 and four untranscribed pseudogenes.

A protein which shows a 50.4 % similarity to prohibitin at the amino acid level was, however, discovered as one of the three proteins isolated by Terashima *et al.* (1994). This 37 kDa protein (BAP37), later named prohibitone, was revealed to be very closely related to prohibitin (Figure 1.6) (Lamers and Bacher, 1997). Homologues of *BAP*37 have now been identified and sequenced in organisms ranging from yeast, *C.elegans*, *Drosophila*, *Arabidopsis* and human. The chromosomal location of *BAP*37 has yet to be determined. These homologues, as is the situation for *PHB*, are highly conserved through evolution. The identification of these two very similar and highly conserved genes leaves open the possibility that they constitute members of a multi-gene family of evolutionarily-conserved genes.

Figure 1.6. Amino acid sequences of the human prohibitin and BAP37 proteins. Human prohibitin is the lower strand. Identical residues are in bold and underlined.

1.6.5 Gene Structure

The human prohibitin cDNA sequence isolated by Sato *et al.* (1992) is 1024 nucleotides in length and has 7 exons, 6 of which code for the prohibitin protein which is composed of 272 amino acids. Further sequence analysis extended information concerning the 3' untranslated region (UTR) (Jupe *et al.*, 1995) whilst the 5' regulatory region of the human *prohibitin* gene has been cloned and sequenced independently by two groups (Sato *et al.*, 1993; Altus *et al.*, 1995). Comparison of these two independent sequences shows 8 base differences and 3 single base insertions. A comparative analysis of the rat and human promoter sequences revealed a conserved region of 350 base pairs, as shown in Figure 1.7 (McClung *et al.*, 1995). Seven regions upstream of the first exon are particularly highly-conserved. These include regions resembling the consensus sequence for a CCAAT box and an Sp1 site. The CCAAT box is located approximately 64 bases 5' of the first transcription start site identified in rat, a similar position to that found in most genes. The Sp1 site is represented as a "GC box" matching that of the consensus sequence and a second Sp1 site is found in intron 1. A site resembling a region of the SV40 enhancer that binds regulatory proteins found in nuclear extracts from HeLa and lymphoid cells is also present. Two inverted repeats and two more regions which showed high sequence identity between the rat and human sequences are also present but have no known transcriptional function. Sato *et al.* (1993) found no typical TATA box sequence but did identify a TATG sequence which may be used to initiate transcription; however the functional significance of this is unclear as multiple initiation sites for transcription in the 5' regulatory sequence have been identified and the proposed TATA box (TATG sequence) is present in the first exon of many of these transcripts (McClung *et al.*, 1995).

The intron-exon boundaries of the human *prohibitin* gene were sequenced (Sato *et al.*,1993) and are identical to those of the rat *prohibitin* gene (Altus *et al.*, 1995). Variation in the length of all the introns is found between the two species which contribute

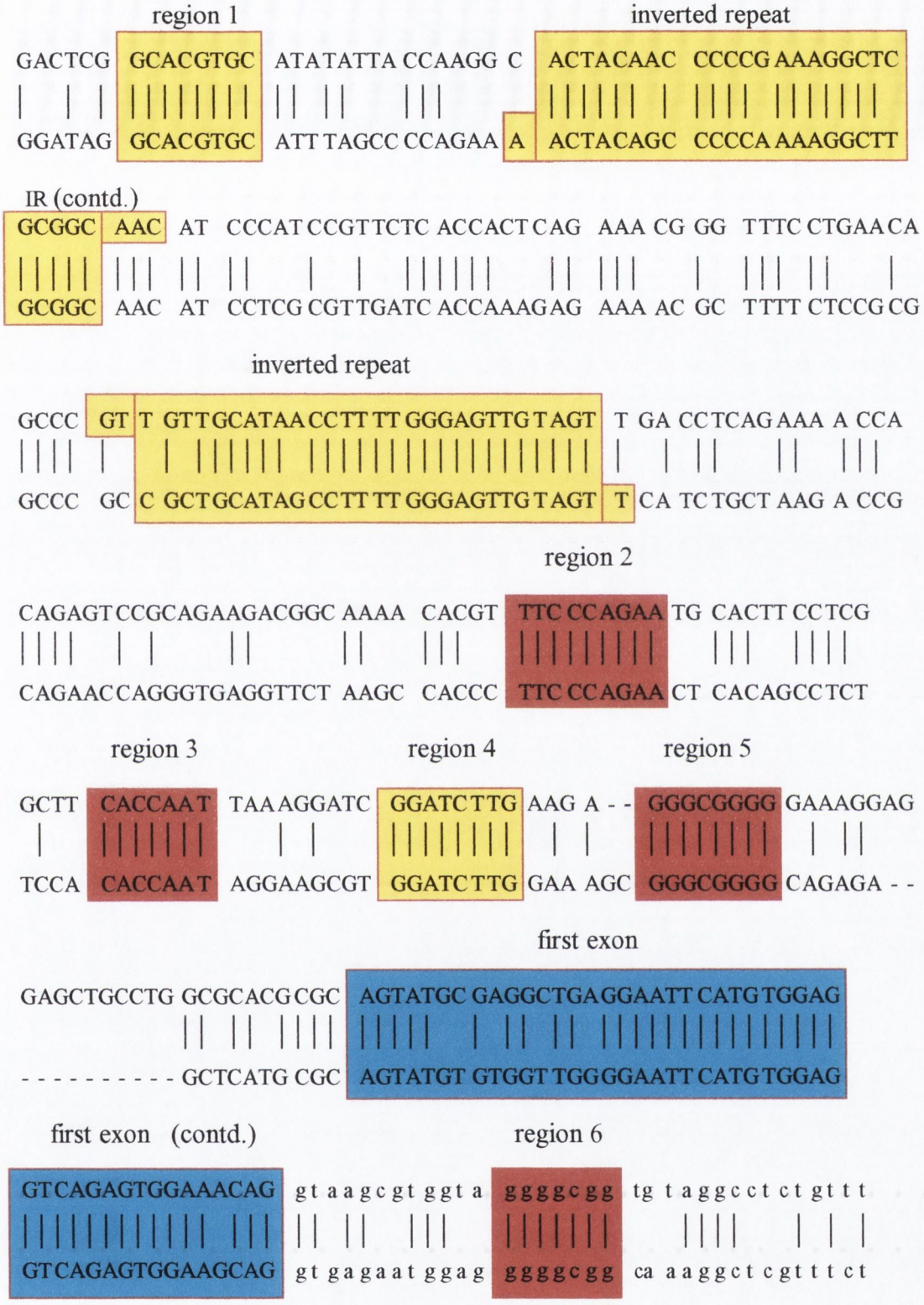


Figure 1.7. The conserved 350 base pair region spanning the 5' -ends of the rat and human *prohibitin* genes. The human sequence is the lower strand. Regions 1 and 4, and the inverted repeats, do not match any known consensus sequences. Regions 2 and 3 resemble a transcriptional control protein binding site in the SV40 enhancer and the CCAAT box consensus sequence respectively. Regions 5 and 6 resemble the GC box bound by the Sp1 transcription factor. Intron 1 sequence is presented in lower case letters (McClung *et al.*, 1995).

to the overall increase in size of the genomic DNA sequence of the rat gene as compared to that of the human *prohibitin* gene. The length of the exons in the rat and human genes are conserved with the exception of exon 1 where alternative transcriptional initiation sites are found. Translation of the prohibitin protein begins in exon 2 in rat, human and *Drosophila*. The seventh exon is the largest of the 7 exons and codes for 70 of the 272 amino acids in the rat and human proteins. Exon 7 contains a long 3' UTR that possesses two polyadenylation signals (AATAAA) positioned 737 base pairs apart. These sites result in two mRNA transcripts that differ in size by approximately 700 bp and which are 1.2 kb and 1.9 kb in length (Nuell *et al.*, 1991). Both are observed in rat and humans, are expressed in all tissues and cell lines examined and code for the same protein of ~30 kDa (McClung *et al.*, 1989; Nuell *et al.*, 1991; Liu *et al.*, 1994; Dell'Orco *et al.*, 1994).

The original identification of prohibitin as one of a number of antiproliferative mRNAs from rat liver detected a single prohibitin mRNA of 1.9 kb, since the probe used in those experiments was comprised of a part of the sequence that is only present in the 1.9 kb mRNA (McClung *et al.*, 1989). Nuell *et al.* (1991) observed two mRNA transcripts in rat cells and human fibroblasts, one of 1.2 kb and one of 1.9 kb, both with a poly(A) tail of about 200 bases. Different-sized mRNA transcripts, however, have been detected in rat germ cells (Choongkittaworn *et al.*, 1993) but not in any other cells analysed from rat or human. The alternate-sized transcripts of 1.5 kb and 0.7 kb were detected in rat germ cells along with the 1.2 kb transcript but not the longer 1.9 kb transcript. Additionally a transcript of 1.6 kb was reported at some stages of spermatogenesis and was also present in the uterus. The authors did not determine if the variant length transcripts were due to alternative splicing, polyadenylation or degradation of larger transcripts. However, analysis by Western blots identified only the ~30 kDa prohibitin protein in all the cells examined.

The level of expression of the 1.2 kb and 1.9 kb transcripts varies between tissue types and the ratio of 1.2 kb to 1.9 kb is different from one tissue type to the next. During the cell cycle endogenous prohibitin mRNA and protein levels fall as the cells enter S phase from G1 and rise as they enter G2 then fall again during M phase, although throughout the cell cycle the levels only vary by 3-4 fold (Roskams *et al.*, 1993). The protein levels in each tissue type generally parallel the amount of mRNA expressed in each cell. Contrastingly, for all immortalised cells analysed the level of mRNA expression is elevated as is the protein level when compared to that of normal cells (Jupe *et al.*, 1995; Jupe *et al.*, 1996b). The elevated mRNA expression is almost exclusively due to the 1.9 kb transcript. Levels of the 1.2 kb transcript remain virtually the same in normal and immortalised cells. The higher expression of the 1.9 kb mRNA and protein is a feature that has been associated with human tissues that have increased cellular proliferation or have proliferative potential (Liu *et al.*, 1994; McClung *et al.*, 1995; Jupe *et al.*, 1996b). This creates the paradoxical situation whereby the levels of a messenger RNA with antiproliferative activity are increased in cells that are proliferating. It also conflicts with the initial isolation of a partial prohibitin mRNA which was found to be more highly expressed in normal liver as compared to regenerating liver (McClung *et al.*, 1989). In general the expression levels of prohibitin mRNA and protein in either normal or immortalised cells do not appear to change significantly during the cell cycle except in the case of rat testis where the *prohibitin* gene is expressed differentially (Roskams *et al.*, 1993; Choongkittaworn *et al.*, 1993). Only one study has investigated the levels of prohibitin mRNA expression directly from tumour tissue (Asamoto *et al.*, 1994). In this case the steady state level of prohibitin mRNA was found to be increased in rat bladder carcinomas as compared with that of normal rat bladder or liver. No comment was made on the relative levels of expression of the 1.2 kb and 1.9 kb transcripts.

1.6.6 Prohibitin and the Cell Cycle

Another gene postulated to be involved in the regulation of the cell cycle is *prohibitin*. The first prohibitin cDNA clone was originally isolated from mRNA found in normally-nonproliferating tissues from intact animals. Prohibitin was demonstrated to have antiproliferative activity when microinjected into human diploid fibroblasts (HDF) which had been stimulated to enter the cell cycle (McClung *et al.*, 1989). Further microinjection studies of the prohibitin message at different times during the cell cycle showed that the antiproliferative activity was specific to a point in the transition from G0 to S phase in the cell cycle (Roskams *et al.*, 1993). Prohibitin's involvement in cell cycle regulation is unlikely to be orchestrated by altering the expression of *PHB* mRNA or protein as the levels of both alter only slightly throughout the cell cycle (Lumpkin *et al.*, 1985; Roskams *et al.*, 1993; Liu *et al.*, 1994). It may be mediated by a post-synthetic modification such as phosphorylation or a small molecular weight post-translational modification. These proposed modifications could explain the prohibitin isoforms that have been detected by two-dimensional Western blot analysis in rat liver and HDF (McClung *et al.*, 1995).

1.6.7 Functional Activity

The functions of prohibitin are currently unknown but several studies of its cellular location and properties indicate that it may be involved in processes as diverse as senescence, development, tumour suppression and regulation of the cell cycle. The majority of the prohibitin protein appears to locate to the mitochondria (Ikonen *et al.*, 1995). Subcellular fractionation techniques further localised the protein to the inner membrane of the mitochondria. Prohibitin was not detected in the nuclear fraction, or in the cytosol, showing that it is not soluble (McClung *et al.*, 1995). A previously identified inner mitochondrial protein of 32 kDa, which controls Ca^{2+} efflux from the mitochondrial matrix, has its activity regulated by mono(ADP-ribosyl)ation (Richter and Kass, 1991).

This protein is similar to prohibitin in so far as prohibitin is a protein of approximately 30 kDa, it is associated with the mitochondrial inner membrane and a mono(ADP-ribosyl)ated protein has been shown to coincide with prohibitin by Western analysis. In mitochondrial energy metabolism, ATP generation can be activated by high concentrations of Ca^{2+}. Regulation of this process is determined by the ability to control the release of Ca^{2+} from the mitochondrial matrix. Microinjection of prohibitin mRNA into cells thereby causing inappropriate/overexpressed levels of the protein could lead to the alteration of ATP levels at critical times during the cell cycle and this may explain the observed antiproliferative activity.

In contrast to the mitochondrial location of prohibitin, Terashima *et al.* (1994) showed that prohibitin and another very similar protein, BAP37, were non-covalently associated with the mIgM antigen receptor of B lymphocytes in mice. Hence they proposed the most likely location for prohibitin to be in or adjacent to the plasma membrane. This location for prohibitin appears to be unique to B cells or possibly haemopoietic cells. In these experiments proteins that coprecipitated with mIgM were identified but other cellular locations for prohibitin were not investigated. It was hypothesized that the different biological effects observed after antigen binding to either mIgM or mIgD could be explained by signals transduced through the association of prohibitin and BAP37 to mIgM but not mIgD. They also speculated that prohibitin could have a role in the internalisation of the IgM B-cell receptor and that a signal, possibly mediated through prohibitin, could initiate programmed cell death or apoptosis. Interestingly, the mitochondria have also been suggested to have a role in the control of apoptosis and members of the Bcl-2 family of proteins, which are associated with the mitochondria, possibly promote this by regulation of mitochondrial Ca^{2+} (Baffy *et al.*, 1993; Zamzami *et al.*, 1996).

One process that may be shared by prohibitin at these two distinct localities is that of Ca^{2+} fluxes. Dolmetsch *et al.* (1997) have demonstrated that differential gene transcription in B lymphocytes is achieved by amplitude modulation (AM) of the Ca^{2+} signalling system. When B lymphocytes respond to the arrival of an antigen at the B-cell receptor (BCR), this external information is translated into internal messengers such as Ca^{2+}. The BCR uses Ca^{2+} to control separate cellular processes by using an AM form of signalling. The association of prohibitin and the BCR could provide a means of moderating the level of Ca^{2+} signals.

Other observations that *PHB* is essential for cellular function have been described for the homologues of *prohibitin* found in *Drosophila* and yeast (Eveleth and Marsh, 1986; Franklin and Jazwinski, 1993). In *Drosophila* the *prohibitin* homologue *Cc* is expressed throughout embryogenesis, at pupariation and in adults. Homozygous mutations in *Cc* have been shown to be lethal in late larvae or during the progression to the pupae (Eveleth and Marsh, 1986). Therefore prohibitin appears to be of critical importance and essential during development in *Drosophila*. The connection between prohibitin and the developmental processes is supported by its detection as a differentially expressed mRNA in rat seminiferous epithelium (Choongkittoworn *et al.*, 1993). In rat testis, prohibitin levels are high before and after S phase in spermatocytes but low during S phase and absent in M phase. This expression pattern correlates with prohibitin as a negative regulator of proliferation and differentiation of germ cells.

1.6.8 Mutational Analysis

The initial investigation for mutations in *PHB* was prompted by reports of its antiproliferative activity, its location on chromosome 17q21 and the resulting postulation that it might be a tumour suppressor gene. At the time *PHB* was a candidate gene for inherited susceptibility to breast cancer in families with early-onset of disease that had

been linked to chromosome 17q21. Several studies have looked for somatic and/or germline mutations of the *prohibitin* gene in a variety of cancers, including breast, ovary, lung and liver (Sato *et al.*, 1992; Sato *et al.*, 1993; Tokino *et al.*, 1993; Cliby *et al.*, 1993). To date, mutations in the coding sequence have only been found in tumours of the breast (Sato *et al.*, 1992; Sato *et al.*, 1993). In total 4 somatic mutations have been identified which affect the coding sequence of *PHB* (Sato *et al.*, 1993). The RNase protection assay was used to screen 120 primary breast tumours which had shown loss of heterozygosity (LOH) for the long arm of chromosome 17 and/or had been derived from patients younger than 35 years of age. One tumour from a 35 year old woman contained no normal prohibitin protein since one *prohibitin* allele had been lost by deletion whilst the other allele carried a frameshift mutation resulting in a premature stop codon 70 base pairs downstream of the mutation.

Three of the four known somatic mutations of the *prohibitin* gene are missense mutations and were identified in tumours which had exhibited LOH for the long arm of chromosome 17. The age at diagnosis for two of these women was over 35 years whilst the age of the third woman was not given (Sato *et al.*, 1993). In the absence of a functional test for the protein certain criteria may aid in deciding if a missense mutation is likely to have a deleterious effect or not. The criteria that have been used to distinguish cancer-predisposing alleles from common polymorphisms, in the germline, in a known or putative tumour suppressor gene include (1) truncation or non-conservative amino acid substitution in the protein caused by a DNA sequence variation; (2) cosegregation of the allelic variant with the disease; (3) low frequency of the variant in control chromosomes and (4) conservation or similarity of the altered amino acid across species (Friedman *et al.*, 1994a; Stoppa-Lyonnet *et al.*, 1997). Two of the three missense mutations occur at residues that are conserved in all species in prohibitin and BAP37. These are a valine at codon 88 which is replaced by an arginine residue and an arginine at codon 214 which is

replaced by a valine residue. Both constitute non-conservative substitutions. Valine is an aliphatic hydrophobic amino acid whilst arginine is a positively charged basic residue. The third substitution is a conservative change which results in an arginine at codon 105 being replaced by histidine. This arginine residue is conserved amongst mammals but not in any other species examined. From the criteria outlined above, two out of the three amino acid substitutions identified in sporadic breast tumours fit the criteria for deleterious mutations.

1.7 Replicative Senescence

Normal cells are described as having a finite replicative life span as they cannot divide indefinitely. The process that limits the proliferative potential of cells is called replicative senescence and was first identified in human fibroblasts in culture (Hayflick, 1965). Although most proliferatively-competent cells in mammals do undergo senescence, there are three exceptions; (1) the germline, (2) certain stem cells and (3) many malignant tumours which appear to have cells that have an immortal replicative life span. This third example suggests that replicative senescence is a tumour suppressive mechanism that is lost in tumour cells.

Once a cell has completed a finite number of divisions, cell growth is arrested with a G1 DNA content. This growth arrest is permanent, the cell cannot enter S phase under any stimuli, and indicates that cells in replicative senescence express inhibitors which act in a dominant way. The cell in replicative senescence represses positive-acting, growth-regulatory genes that are important for G1 progression. These include E2F and Rb, which is constitutively hypophosphorylated (Stein *et al.*, 1990). In contrast to the situation for positive growth regulators, negative growth regulators such as p16 and p21 are overexpressed (Noda *et al.*, 1994; Hara *et al.*, 1996) and possibly other growth inhibitors not yet identified (Campisi, 1997). Their overexpression inhibits the CDKs which then

maintain Rb in an underphosphorylated form in senescent cells. Along with the arrest of cell proliferation, senescent cells have two other remarkable phenotypes. The first is resistance to apoptosis and the second is differentiated gene expression, the effects of which vary between cell types. This second phenotypic change is not well understood and may be a result of the senescence-associated growth arrest.

Characteristics that suggest replicatively senescent differentiated cells are not conducive to tumour formation include the following: cells that have a finite replicative life span are considerably less likely to form tumours than immortal cells that are replicating. Although being immortal does not necessarily go hand in hand with tumour formation, it does allow multiple cell divisions during which deleterious mutations can be acquired which may lead to tumourigenesis. Two genes that are responsible for establishing and maintaining replicative senescence are the tumour suppressor genes *RB* and *p53*. If either of these genes is lost in human fibroblasts the life span of the cells is increased but they eventually undergo senescence. When both genes are absent human fibroblasts become genomically unstable and either die or become immortal. Thus, *p53* and *RB* are essential for the normal replicative senescence of cells. These findings suggest that replicative senescence may have evolved to diminish tumourigenesis.

Experiments by Franklin and Jazwinski (1993) in yeast identified *prohibitin* as being a longevity-assurance gene that affects replicative lifespan. This was demonstrated by deleting or overexpressing the *prohibitin* gene in yeast haploid cells. In yeast prohibitin is differentially expressed during the life span and interacts with *RAS2*, a gene that extends cell longevity and that is also differentially expressed during the yeast life span (Jazwinski, 1996). In human diploid fibroblasts (HDF) this direct relationship between prohibitin and cell life span has not been reported but Liu *et al.* (1994) have suggested a possible involvement for prohibitin in senescence of HDF. They showed that approximately 10 % of the total prohibitin protein present in younger cells but not in older cells was post-

synthetically modified, possibly by phosphorylation. This could allow a mechanism whereby prohibitin may function as an antiproliferative protein in senescent populations.

Senescence, or finite proliferation, is considered to be a dominant phenotype. Conversely, cellular immortalisation is recessive (Pereira-Smith and Smith, 1983). This was concluded from the observation that when proliferating normal cells are fused with immortal tumour derived cells the majority of resultant hybrid cells eventually undergo senescence. Also, fusion studies between different immortal human tumour cells frequently produced hybrid cells that underwent senescence. Subsequent experiments identified 4 complementation groups, A, B, C and D, involved in immortalisation (Pereira-Smith and Smith, 1988) which suggests that senescence can occur by a number of pathways. Normal human chromosomes 4, 1 and 7 have been shown to reverse immortality for cell lines from the complementation groups B, C and D, respectively (Ning *et al.*, 1991; Ogata *et al.*, 1993; Hensler *et al.*, 1994). A gene on chromosome 4, Mortality Factor 4 (MORF 4), has recently been identified which induces senescence in cell lines from complementation group B whilst having no effect on cell lines from the other complementation groups (Pereira-Smith, 1998).

It has been proposed that immortalisation results from the alteration of processes (or genes) related to a normal cell's ability to create or react to an inhibitor of DNA synthesis. Hence the involvement of prohibitin, an inhibitor of the initiation of DNA synthesis, in cellular immortalisation was examined. Prohibitin mRNA was demonstrated to be capable of blocking cell cycle progression in cell lines from the complementation group B when microinjected into cells. What part *PHB* plays in this process is still unclear as the senescence gene(s) associated with complementation group B have previously been shown to be located on chromosome 4 (Ning *et al.*, 1991) and *prohibitin* is located on chromosome 17. A possible explanation is that prohibitin's activity is downstream in the

senescence pathway associated with the gene(s) on chromosome 4. These findings suggest that prohibitin may have a senescence-related tumour suppressor activity.

1.8 Techniques used to Screen for Unknown Mutations

The detection of mutations in identified genes is a critical step in linking the gene to a particular disease phenotype. With the development of the polymerase chain reaction (PCR) (Mullis *et al.*, 1986) the number and range of mutation-detection techniques has grown. The choice of technique for mutational analysis requires weighing up its disadvantages against its advantages and how they might apply to the gene to be screened. In general, screening techniques are based on either enzymatic or chemical cleavage of mismatched bases in a duplex molecule, or on comparison of the electrophoretic behaviour of wild type and mutant DNA/RNA sequences.

Enzyme cleavage involves heteroduplex formation between any transcribed RNA and RNA or DNA (Myers *et al.*, 1985a). The technique is not PCR-based and requires cloning. The efficiency of detection is only 60-70 % but precisely locates the mutation and does not involve use of hazardous chemicals. The techniques of chemical cleavage or modification of mismatched bases are capable of detecting 100 % of mutations and their location (Cotton *et al.*, 1988). These techniques have found most use in the screening of large genes. They are, however, arduous with many steps involving hazardous chemicals.

The techniques based on electrophoretic behaviour include single-strand conformation polymorphism (SSCP) analysis (Orita *et al.*, 1989; Hayashi *et al.*, 1992), heteroduplex analysis (White *et al.*, 1992) and denaturing gradient gel electrophoresis (DGGE) (Fischer and Lerman, 1983; Myers *et al.*, 1985b; Sheffield *et al.*, 1989). SSCP and heteroduplex analysis are both frequently used because they are easily established techniques which are simple to use and can detect mutations in a relatively short period of

time. The efficiency of detection is about 80-90 % for fragments around 200 bp in length but the efficiency drops for longer fragments and requires a variety of electrophoresis conditions. DGGE is a non-radioactive technique that theoretically can detect 100 % of mutations. The technique is most suited to fragments of 200-600 bp and is the best technique for routine screening of such fragments. The initial setting up and design of gel-running conditions is time-consuming and requires effort but the technique is then relatively uncomplicated.

1.9 Summation

Prohibitin is located on chromosomal region 17q21, a region that frequently undergoes LOH in sporadic breast tumours and in tumours from individuals with a family history of the disease (Borg *et al.*, 1994). This loss of heterozygosity is associated with regions of 17q other than that occupied by the breast cancer susceptibility gene, *BRCA1*, suggesting that other tumour suppressor genes are located in the region (Cropp *et al.*, 1993; Cornelis *et al.*, 1993). Futhermore, *BRCA1* is rarely found to be somatically mutated in sporadic tumours, supporting the premise that additional genes on 17q are involved in the development of breast cancer (Futreal *et al.*, 1994; Merajver *et al.*, 1995; Hosking *et al.*, 1995). Evidence indicating that *prohibitin* is a tumour suppressor gene that is involved with the development of breast cancer has come from a study of primary breast tumours in Japan (Sato *et al.*, 1992). In a number of tumours both *prohibitin* genes were somatically inactivated.

Although the functions of prohibitin are not known, *prohibitin* transcripts have been shown to possess antiproliferative activity by inhibiting the progression of the cell cycle in HDF and HeLa cells. This antiproliferative activity has been postulated to function as a tumour suppressor activity in cellular senescence and may act by post-synthetic

modification. *Prohibitin* is also a highly conserved gene and is expressed in all human tissues examined (McClung *et al.*, 1995). In view of the above observations we undertook the analysis of the *prohibitin* gene in a cohort of breast cancer patients.

CHAPTER 2

MATERIALS AND METHODS

2.1 Individuals Included in these Studies

2.1.1 Patient Population

Patients with breast cancer diagnosed at any age, who presented to the Department of Oncology, St. James' Hospital, Dublin for new evaluation or ongoing management were eligible to participate. All participating patients provided consent, 15 ml of peripheral blood for total RNA and genomic DNA isolation and completed a questionnaire. The questionnaire asked the consenting patients for information on all their 1st, 2nd and 3rd degree relatives pertaining to their date of birth, date of death, cancer, site of cancer, age at diagnosis and any other serious illnesses.

Genomic DNA samples from the patients were assigned a three digit case number prefixed by the letter D, or L if the patient was diagnosed with breast cancer at 50 years of age or older eg. D123, L010. Total RNA samples were also assigned a three digit case number and this was prefixed by the letter R eg. R056. This case number was also used to describe the corresponding cDNA sample. For the 6 families identified in these studies the family members are identified by their unique identification code given in their family pedigree.

2.1.2 Control Population

Women without a previous history of breast cancer and without a known 1st or 2nd degree relative with breast cancer were asked to participate in the study in order to provide a comparative group to the breast cancer population. Consenting controls provided either 5 ml of peripheral blood or a mouth swab of buccal cells for DNA isolation. Controls were frequency matched with the breast cancer population for being under 50 years of age or 50 years of age or older. Genomic DNA samples from the control population were given a three digit number prefixed by the letter C or K eg. C128, K057.

2.2 Collection and Storage of Peripheral Blood

Peripheral blood was collected in 5 ml tubes (VACUETTE® greiner labortechnik) containing lithium heparin to prevent coagulation. All the peripheral blood samples used in these studies were collected by the staff of the phlebotomy unit or nurses of St. James' Hospital, Dublin. In exceptional circumstances blood samples were acquired from individuals living in England and Australia by Diane Averill of the Cancer Research Campaign Section of Molecular Carcinogenesis, Institute of Cancer Research, U.K. and the individual's General Practitioner, respectively. All of these individuals were related to women in the breast cancer population used in these studies. The tubes were then stored at -20 °C or -70 °C before isolation of genomic DNA or total RNA.

2.3 Isolation of total RNA, cDNA Synthesis and Isolation of Genomic DNA

2.3.1 Isolation of total RNA from Peripheral Blood

Peripheral blood, 5 ml, stored at -70 °C was thawed, transferred to a 20 ml universal container and centrifuged at 3, 500 r.p.m. for 5 minutes. The pellet was retained and total RNA extracted by a guanidinium-thiocyanate phenol-chloroform method as described by Ausubel et al. (1988). The resulting nucleic acid solution was then treated with RQ1 RNase-Free DNase (Promega) according to the manufacturer's instructions. Great care was taken to ensure that all solutions used were free from ribonucleases; all solutions used in the preparation of the total RNA were treated with Di-Ethyl Pyrocarbonate (DEPC). Where applicable the solutions and eppendorf tubes used in the extraction or storage of RNA were autoclaved following the treatment with DEPC.

2.3.2 Reverse Transcription of mRNA

Reverse transciption of mRNA was performed under the following conditions: A 30 µl reaction volume contained 100 ng total RNA, 300 U Moloney-Murine Leukaemia Virus Reverse Transcriptase, 10 mM DTT, 1 X First Strand Buffer (Gibco BRL), 0.5 µg Oligo (dT) 15 Primer, 20 U rRNasin and 1 mM of each dNTP (Promega). cDNA synthesis was carried out at 37 °C for 1 hour. Aliquots of the resultant cDNA were then used in the polymerase chain reaction (PCR). The complete process is referred to as reverse transcription- PCR (RT-PCR).

2.3.3 Isolation of DNA from Peripheral Blood

Peripheral blood stored at -20 °C or -70 °C was thawed and centrifuged at 3, 500 r.p.m. for 5 minutes. The supernatant was poured off and the pellet retained. Genomic DNA was extracted from the pellet as described by Rossiter *et al.* (1991) in an extraction buffer containing 10 mM Tris-HCl (pH 7.5), 400 mM NaCl, 2mM Na_2 EDTA, SDS at a concentration of 0.8 % (w/v) and 100 µg/ml proteinase K. The extracted genomic DNA was redissolved in dH_2O at a concentration of ~0.2 µg/µl and stored at -20 °C.

2.3.4 Isolation of DNA from Buccal Cells

Buccal cells were collected on a cytology brush (Cytosoft™ Cytology Brush, Medical Packaging Corporation) and replaced in the original packet. The following protocol was then used to isolate genomic DNA from the cells. The brush was removed from the packet and the brush head cut off and placed in a 1.5 ml eppendorf tube. Five hundred microlitres of 50 mM NaOH was added to the tube and vortexed for 1 minute. The tube was heated at 95 °C in a water bath for 10 minutes, followed by the addition of 50 µl of 1M Tris pH 8 and then vortexed for 30 seconds. Then 200 µl of Instagene™ Matrix (Bio-Rad) were added to the sample and the tube was incubated at 56 °C for 30

minutes in a water bath. The Instagene™ Matrix (Bio-Rad) is used to absorb cell lysis products that could interfere with the subsequent DNA manipulations ie. PCR. The tube was removed from the water bath and vortexed for 10 seconds, followed by heating at 100 °C for 8 minutes and vortexed again for 10 seconds. Finally the brush head was removed from the eppendorf tube using sterile forceps. The tube was then spun for 3 minutes at 12,000 rpm and the supernatant transferred to a new tube for use in the PCR.

2.3.5 Isolation of DNA from Formalin-Fixed Paraffin-Embedded Tissue

Three 5 μm thick sections were cut from the formalin-fixed paraffin-embedded tissue using a Leitz 1512 microdissector and placed in a sterile 1.5 ml eppendorf tube. Sections were then dewaxed and incubated overnight at 55 °C in 150 μl of 50 mM Tris pH 8.5, 1 mM EDTA pH 8.0, 0.5 % Tween 20 and 200 μg/ml of proteinase K (Innis *et al.*, 1990). This was followed by a phenol, chloroform-isoamyl alcohol (25: 24: 1) extraction and a chloroform-isoamyl alcohol (24: 1) extraction of the solution (Sambrook *et al.*, 1989). The DNA was resuspended in 50 μl of sterile dH_2O and stored at -20 °C.

2.3.6 Microdissection of Tumour Tissue

The microdissection of tumour tissue from formalin-fixed paraffin-embedded material was performed by Orla Sheils (Department of Histopathology and Morbid Anatomy, Trinity College Dublin) according to the protocol described by Pan *et al.* (1994).

2.4 Denaturing Gradient Gel Electrophoresis (DGGE)

2.4.1 Introduction

The technique of Denaturing Gradient Gel Electrophoresis (DGGE) exploits the effect of helicity, partial melting and complete melting with dissociation of the strands, on the mobility of DNA when electrophoresed through a polyacrylamide gel. The term "melting" when applied to double stranded DNA indicates the transition from a double stranded, helical structure to a single stranded, nonhelical or random coil conformation. In a partially melted state the molecule moves much more slowly through the pores in the denaturing gradient gel than when it is present in a fully helical structure or as completely-dissociated single strands.

It is when the molecule is present as a partially melted structure that its electrophoretic mobility is affected. DGGE utilises this characteristic to separate DNA sequences which differ by as little as a single base substitution. It is proposed that the partially melted molecule can be viewed as having a large number of ends as compared to the double-helix or single-stranded DNA molecule, which is viewed as having only two ends (Lerman, 1982). The numerous ends of the partially melted molecule enter several pores within the polyacrylamide gel. Unless all the ends of the molecule enter the one pore, which is unlikely, the molecule will become retarded. The double-helix can easily migrate through the one pore and avoids being retarded. In the case of DGGE (which is run at a temperature of 60 °C) increasing concentrations of urea and formamide in the polyacrylamide gel are used to produce denaturing conditions equivalent to a temperature gradient of 60-100 °C. A range which is sufficient to cause the denaturation of any DNA molecule.

Melting occurs in a cooperative way within discrete regions of contiguous base pairs as the DNA molecule passes through a relatively small increase in temperature. Such

regions are termed domains and any alteration to the base stacking changes the thermodynamic melting properties of the domain. For an example, this means that replacing a dA-dT base pair with a dT-dA base pair affects the stacking interactions and consequently alters the melting behaviour of the domain. As the temperature continues to rise, melting proceeds as several domains of discrete transition from the double helix form to the random, non-helical conformation. Alterations can be detected in a domain as long as melting of the domain with the highest melting temperature has not occurred. Melting of the domain with the highest melting temperature results in complete strand dissociation. This consequently produces the two native single strands and these are no longer retarded in the gel in the same way as partially melted molecules.

The separation of molecules by DGGE is dependent on the sequence of the molecule and not its length. The degree of mobility retardation, however, is roughly-determined by the length of the melted and unmelted regions. DGGE is a highly sensitive technique which provides for ease of detection of sequence alterations, reproducibility of bands within the gel and reliability of results (Abrams *et al.*, 1990; Blanquet *et al.*, 1993; Stoppa-Lyonnet *et al.*, 1997).

2.4.2 Computational Tools

Two computer programs have been designed to aid in the planning of DGGE experiments for double stranded DNA fragments where the sequence of the DNA to be analysed is known (Lerman *et al.*, 1984; Myers *et al.*, 1985c). These programs are known as MELT and SQHTX and both were originally written in Fortran 77 by Leonard Lerman (Lerman and Silverstein, 1987). The Macintosh version of the MELT program was translated by E. Hovig (University of Oslo) and provided for this work by L. Lerman and E. Hovig.

2.4.3 MELT: The Melting Map Program

The MELT program calculates the temperature, T, at which each base pair, in a given sequence, has a probability, π, of being in a helical state. At a π value of 0.5 or a 50 % helicity, the corresponding temperature is known as the melting temperature (T_m) of a domain. For any DNA sequence a plot of T versus base position in the sequence for a set value of π is known as a melt map. MELT can be used to calculate T_m values corresponding to the probability of a molecule being 5 %, 25 %, 75 % and 95 % melted. The MELT calculations are based on the theory that only two possible states exist for the base pairs. These states are either stacked helical or unpaired random conformation. Other types of interaction between the two strands when in the original helix or separated are not considered.

The program requires two input files before initiating a run. The first contains the DNA sequence to be analysed and the second is the setup file called M87 (Figure 2.1). The relevant details of the sequence to be analysed are inserted into the setup file. These include; the length of the sequence, the position of the first and last base, the starting temperature of the calculation (taken as 60 °C since the top of the gel is immersed in recirculating buffer at 60 °C), the temperature increments used to reach the highest melting domain in the sequence (set at 0.2 °C for all calculations) and the format and name of the sequence file.

```
1, 480          1 'MMIN (position of first base), MEND (position of last base)'
60.0            2 'TINIT (initial temperature)'
0.2             3 'TSTEP (increments in temperature)'
200             4 'NMN (total number steps of rise in temperature)'
'pro2gc'        5 'sequence file which consists of letters A, T, G, C.'
'(5 (10a, 1x))' 6 'format of sequence file'
'pro2gc'        7 'output file root (prefix), precisely six characters'
```

Figure 2.1. Example of the setup file M87.

MELT also provides a variety of other useful information on the characteristics of the sequence being examined, which is presented in the files *pbr*, *dtht*, *coop* and *dsklg*. The *pbr* file gives the number of nonhelical, melted, base pairings at each temperature. Therefore if the complete fragment is melted then this number will equal the number of bases in the fragment. The output of *dtht* shows the change in the fraction of nonhelical bases at each temperature increment. The file *coop*, coherence, represents the amount of close coupling in the melting transition. The inverse of this figure represents the breadth of temperature interval over which a base pair changes from fully helical to fully melted. The log of the dissociation constant (K_d) for complete strand separation at each temperature is given in *dsklg*. This final output file is particularly important in the design of experiments as the T_m of the domain of interest should be at least 0.5 °C lower than the temperature at which the log K_d is equal to -5. This ensures that the domain melts before there is complete strand dissociation.

The MELT program predicts the melting properties of a known DNA sequence and these computational analyses have been substantiated by a considerable number of empirical results (Fischer *et al.*, 1983; Myers *et al.*, 1985c; 1985d). This information can

then be used to determine the choice of primers for regions of the sequence which can be used for mutational screening by DGGE.

2.4.4 SQHTX: Program for the Prediction of Mobility and Displacement of Molecules in a DGG

The program SQHTX can be used in collaboration with the MELT program to establish the optimum electrophoresis conditions to obtain the maximum resolution in a gel of DNA fragments differing by a single base change. The SQHTX program is used to calculate the effect on the mobility of a molecule for a base change at every base pair in the molecule for a series of stipulated electrophoresis times. The change in stability for a base change has a set value although this value may be considerably different in real terms. The displacement between the wild type sequence and the variant sequence is calculated as a temperature difference which can be converted to a difference in denaturant concentration.

The set up file of SQHTX is called SETSQHT and requires an input of 16 lines of information regarding the sequence to be analysed and electrophoresis and gel parameters. These include identification of the first and last bases of the sequence, initial and end temperature of the gel, increments in temperature (usually 0.2 °C), the time interval between calculations for gel running times and an output file name for the displacement results.

An example of the output of SQHTX, the displacement file, is shown in Figure 2.2. The top row is a list of electrophoresis times, in hours (1.00...2.00...), which start and increase at intervals specified in SETSQHT (usually 1 hour). The column on the left of the figure, headed SITE, indicates the base position in the sequence (2...22...42...) and this number increases at an interval also previously specified in SETSQHT (in this example the interval is 20 base pairs). The calculated differences in the temperature reached by the wild

SITE	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00
2	-0.03	-0.04	-0.04	-0.04	-0.04	-0.05	-0.06	-0.07	-0.07	-0.07
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
42	-0.00	-0.00	-0.00	-0.00	-0.01	-0.02	-0.04	-0.09	-0.17	-0.27
62	-0.00	-0.00	-0.01	-0.02	-0.05	-0.09	-0.16	-0.25	-0.38	-0.50
82	-0.00	-0.00	-0.01	-0.01	-0.03	-0.06	-0.11	-0.19	-0.31	-0.45
102	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01	-0.02	-0.08	-0.19	-0.34
122	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01	-0.02	-0.06	-0.17	-0.32
142	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.02	-0.06
162	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.03
182	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.02	-0.06
202	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01	-0.02	-0.04	-0.09
222	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01	-0.02	-0.04	-0.10
242	0.00	-0.01	-0.11	-0.21	-0.27	-0.32	-0.36	-0.39	-0.42	-0.43
262	-0.00	-0.22	-0.59	-0.75	-0.81	-0.83	-0.83	-0.83	-0.81	-0.77
282	-0.01	-0.25	-0.62	-0.76	-0.80	-0.81	-0.81	-0.80	-0.78	-0.74
302	-0.11	-0.52	-0.72	-0.70	-0.67	-0.64	-0.62	-0.61	-0.59	-0.55
322	-0.32	-0.69	-0.73	-0.62	-0.54	-0.51	-0.49	-0.48	-0.46	-0.43
342	-0.58	-0.76	-0.61	-0.45	-0.39	-0.36	-0.35	-0.34	-0.33	-0.30
362	-0.56	-0.55	-0.36	-0.25	-0.22	-0.20	-0.20	-0.19	-0.18	-0.17
	62.65	64.33	65.43	66.15	66.74	67.28	67.80	68.30	68.78	69.23

Figure 2.2. SQHTX output file.

type homoduplex and the heteroduplex for a single base mismatch are given in the columns beneath each time interval (+/- 0.00 indicates no difference). The number at the foot of each of these columns corresponds to the temperature reached for the wild type sequence after the length of gel running time indicated at the head of the column has elapsed eg. 65.43 °C after 3 hours. The program assumes that there is a linear gradient of denaturant. A displacement value of 0.2 is taken as significant in terms of visualising a difference on the denaturing gradient gel. For the example in Figure 2.2 this equates to the ability to detect a base difference in the region of nucleotides 322-362 after 1 hour, nucleotides 262-362 after 2 hours and nucleotides 242-362 after 4 or more hours. Similarly after 10 hours of electrophoresis a single base difference can be detected at any position between nucleotides 42-122. Finally, as with MELT, the log of the dissociation constant as a function of temperature is given for values of -4 to -8. The analysis of any fragment is possible up to the temperature where complete strand dissociation takes place; this temperature is taken as the log of the dissociation constant for a value of -5.

2.4.5 Perpendicular DGGE

Denaturant gradient gels can be used in two different orientations. Both types of gel are constructed and electrophoresed using the same equipment. In one of these configurations the gradient of denaturant in the gel is perpendicular to the direction of electrophoresis and these are termed perpendicular gels. In a perpendicular gel the denaturant concentration increases linearly across the gel. The sample is loaded as a single aliquot in one well spanning the full denaturant range at the top of the gel. This ensures that during electrophoresis each DNA molecule migrates along a line of uniform denaturant concentration and that some of the sample is present at all concentrations of denaturant. At the 0 % or low denaturant concentration edge of the gel, the DNA sample migrates as a double stranded molecule at a velocity, determined by its length, equivalent

to that in a standard non-denaturing polyacrylamide gel of equivalent concentration. The loading dye used with these molecules allows an estimation of migration distance in the gel and therefore the length of time required for electrophoresis.

Figure 2.3 is an example of a perpendicular gel for a section of the *prohibitin* gene. Migration is from top to bottom and the denaturant concentration increases from left to right. The fragment generates a reproducible sigmoidal curve with an abrupt inflection, the position of which, in terms of denaturant concentration, is characteristic to the fragment under scrutiny. This effect on mobility is the result of the retardation of the sample in the gel caused by the molecule being in a melted or partially melted state. The amplified fragment in Figure 2.3 contains a polymorphism in its lowest melting domain. The melting occurs between a concentration of 40-60 % denaturant and multiple bands are seen in this region of the curve due to the different alleles which affect the melting temperature of the domain. Strand dissociation occurs at ~65 % denaturant; where the DNA becomes single stranded a discontinuity in the curve is evident in the gel.

The rise in the contour seen in Figure 2.3 is not always so steep and may also be composed of more than one mobility transition. Figure 2.4 shows the "profile" of a fragment with 2 melting domains; melting of domain one occurs at between 35-45 % denaturant and the second domain melts at between 50-65 % denaturant. Single base differences between otherwise identical fragments cause a shift in the inflection to the left or right thereby indicating that the melting temperature of the fragment has increased or decreased (Figure 2.3). Therefore by this method of perpendicular DGGE the number of melting domains and the breadth of their melting temperatures are revealed.

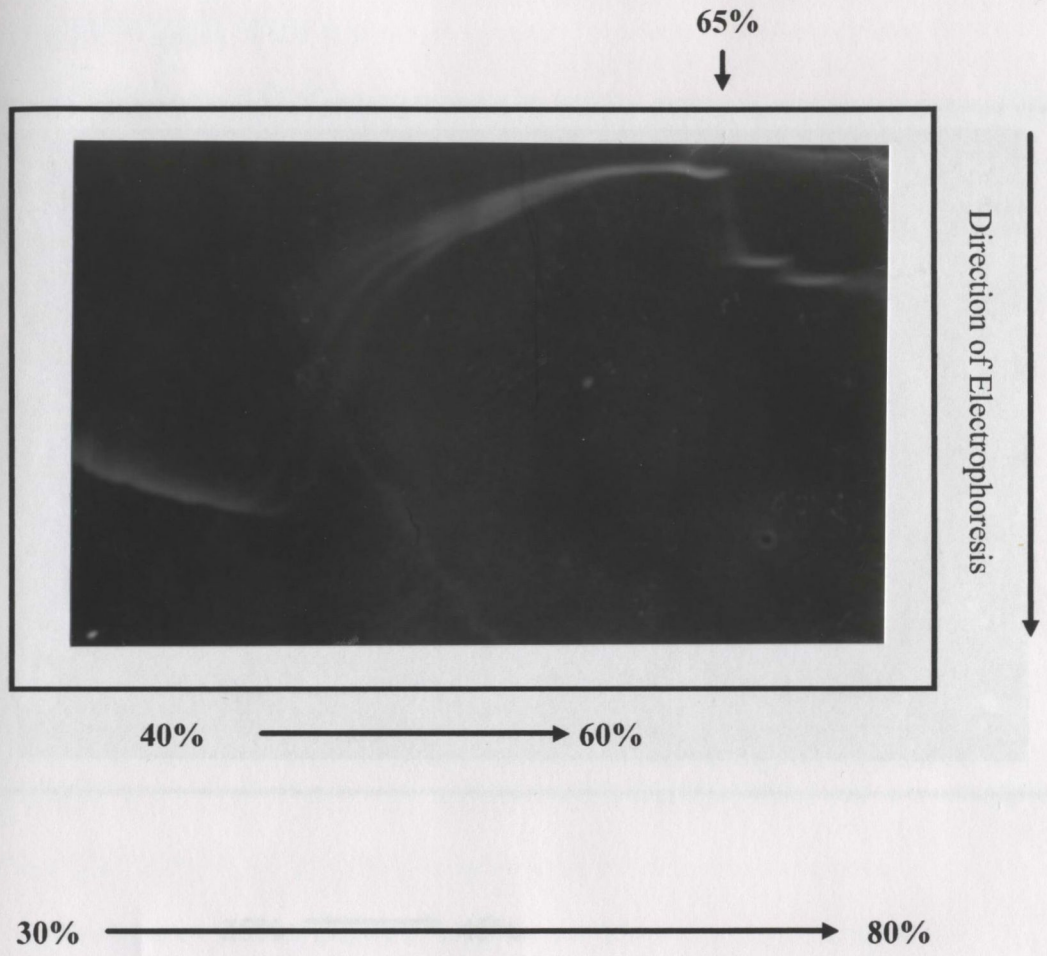


Figure 2.3. Perpendicular Denaturing Gradient Gel of an amplified fragment with a polymorphism. The denaturant concentration ranges from 30 - 80% with melting of the lowest melting domain occurring between 40 - 60% denaturant. Strand dissociation is indicated at ~65% denaturant concentration. The direction of electrophoresis is shown as being perpendicular to the denaturant concentration.

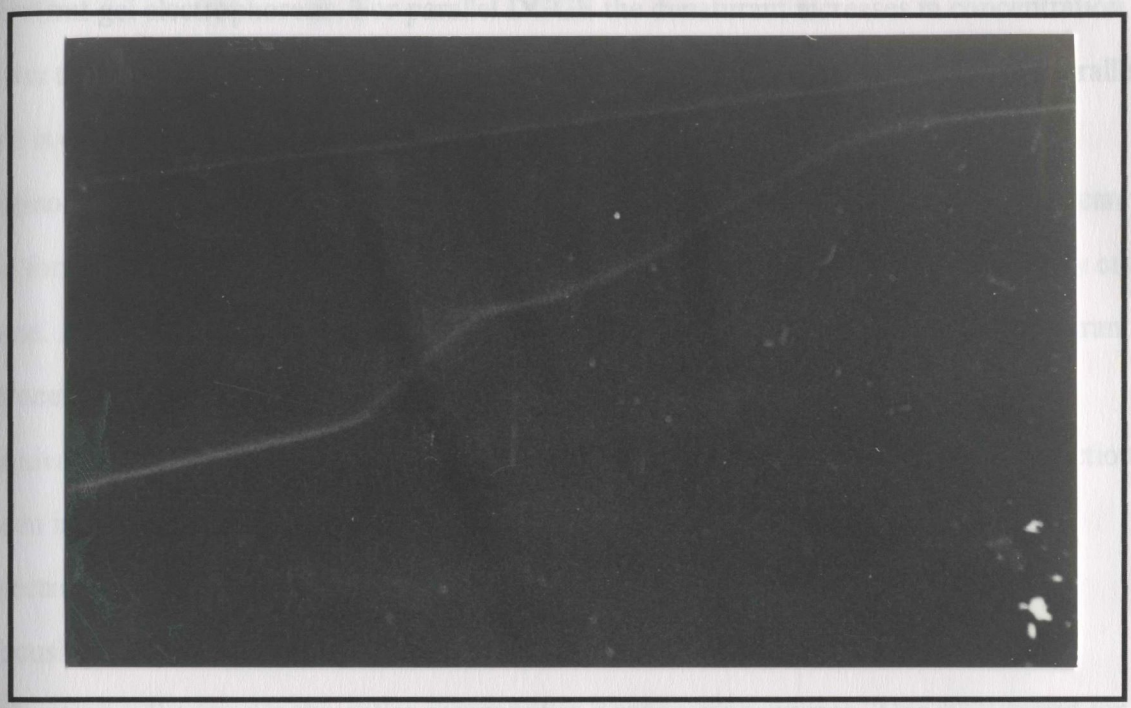


Figure 2.4. Perpendicular Denaturing Gradient Gel showing an amplified fragment with two melting domains. The denaturant concentration of the gel ranges from 20 - 70% with the first domain melting at between 35 - 45%. The second domain melts at between 50 - 65% denaturant.

2.4.6 Parallel DGGE

The second type of denaturant gradient gel involves application of an electric field parallel to the direction of the denaturant gradient. This is termed parallel denaturant gradient gel electrophoresis. For parallel DGGE the denaturant increases in concentration from the top to the bottom of the gel. The range of denaturant concentration in the parallel gel is determined from the denaturant concentration at which the inflection of the sigmoidal curve for the fragment is found in the perpendicular gel. A number of wells can be formed at the top of the parallel gel allowing multiple samples to be screened at any one time. As the sample migrates through the gel it encounters a linearly increasing denaturant concentration. When the sample fragment reaches the denaturant concentration that is equivalent to its melting temperature (which corresponds to the beginning of the inflection seen in the perpendicular gel) its mobility becomes reduced (Figure 2.5). Continued electrophoresis causes further reduction in the mobility of the sample and increased focusing of the band in the gel. Sequence variants are identified as having altered migrational distances in the gel as compared to a wild type sequence. This is observed for even very small differences in melting temperatures between the sequences.

In the case of a DNA fragment which is amplified from an individual who is homozygous for a particular sequence a single band represented by two identical homoduplexes is seen in the parallel gel. For an individual who is heterozygous for the amplified segment, for example a single base substitution in one allele, two different homoduplexes are produced. The homoduplexes migrate as two separate bands in the parallel gel. For certain base alterations the melting temperature of the domain may not be altered significantly enough for the difference to be detected by the parallel gel. However, during the polymerase chain reaction two heteroduplexes are also formed by the annealing of the complementary strands of the two alleles. A further denaturing and annealing step at the end of the PCR is included to maximise the proportion of heteroduplex molecules.

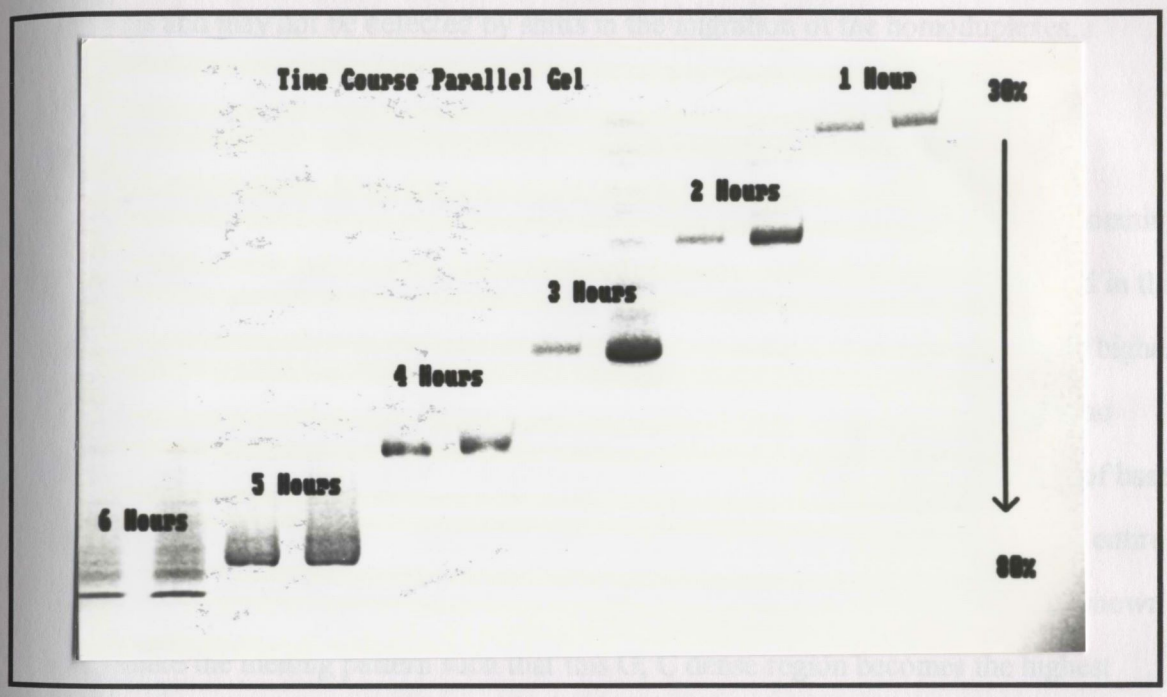


Figure 2.5. Parallel Denaturing Gradient Gel showing a time course.

These heteroduplexes contain base mismatches which cause them to be less stable than the analogous homoduplexes. As a result of this decreased stability the heteroduplexes melt at much lower temperatures (corresponding to much lower denaturant concentrations) in the parallel gel and 2 extra bands can be seen positioned above the homoduplexes in the gel. These additional bands increase the sensitivity of the technique for the detection of sequence differences which cause minimal change in the temperatures of individual domains and may not be detected by shifts in the migration of the homoduplexes.

2.4.7 Generation of Higher Melting Domains with a GC-clamp

As described above, when a molecule migrates through a parallel gel the domain with the lowest melting temperature melts first and the molecule becomes retarded in the gel because its electrophoretic mobility is severely affected. Further melting of the highest melting domain causes complete strand dissociation into single strands which are no longer retarded in the gel due to sequence specific mobility. Thus, only detection of base changes within the lower melting domains is possible. Melting is a property of the entire fragment and the inclusion of a region of guanines, G, and cytosines, C, has been shown to manipulate the melting pattern such that this G, C dense region becomes the highest melting domain (Sheffield *et al.*, 1989). This "GC-clamp" overcomes the situation where the sequence of interest could not be screened for base alterations because it was originally in the highest melting domain of the fragment.

The polymerase chain reaction can be utilised to attach a GC-clamp, usually composed of 25-65 Gs and Cs, to the 3' or 5' end of the PCR product. This is done by incorporating the GC-clamp sequence to the 5' end of one of the primers at the time of its synthesis. The effect of such an addition can be evaluated through the use of MELT software. Therefore the manipulation of primer sequences in conjunction with the use of

MELT software allows the design of PCR products which contain a single low melting domain.

2.4.8 Preparation and Electrophoresis of Denaturing Gradient Gels

Denaturant Gradient Gels (DGG) are prepared by mixing two solutions, which differ in denaturant concentration, with a gravitational gradient maker. This ensures that a linear gradient is formed in the gel. The gradient always ranges, either horizontally for perpendicular gels or vertically for parallel gels, from the low concentration of denaturant to the higher concentration. Solutions of 0 % denaturant and 100 % denaturant have a concentration of 6.5 % acrylamide (37.5:1 acrylamide:bisacrylamide) (BDH) in 1 X TAE buffer (40 mM Tris acetate, 20 mM sodium acetate, 1 mM EDTA pH 7.4). The 100 % denaturant contains 40 % formamide and 7 M urea. These two stock denaturant solutions are combined to obtain the upper and lower denaturant concentration limits of the gel.

A 12 ml volume of each of the resultant denaturant solutions is then mixed with 80 µl of 10 % APS and 7.5 µl of TEMED and the gel poured using the gradient maker. Gels were allowed to set for a minimum of 1 hour before electrophoresis. The PCR products were mixed 1:1 with a loading solution (20 % Ficoll, 10 mM Tris-HCl, pH 7.8, 1 mM EDTA, pH 8, 0.1 % Bromophenol Blue) and 15 µl of the mixture was loaded per lane for a parallel DGG. The samples were then electrophoresed through the gel at 150 V at a constant temperature of 60 °C in 1 X TAE buffer ("The Denaturing Gradient Gel Electrophoresis System" was supplied by C.B.S. Scientific Company, Inc., Del Mar, California, U.S.A.). The gels are run vertically and are completely immersed in the recirculating buffer. The buffer is kept at a constant 60 °C and only in sequences of exceptionally high As and Ts does the melting temperature drop to temperatures below 60 °C. Electrophoresis running time was determined for each fragment. Following electrophoresis the gels were stained with Ethidium Bromide (0.5 µg/ml) for 10 minutes

and then washed with double distilled H_2O. The gels were visualised on a UV transilluminator and documented by polaroid photography or a video gel-documentation system (UVP).

2.4.9 Excision of Bands from Denaturing Gradient Gel for use in PCR

Using a sterile blade the fragment band of interest was excised from the Denaturing Gradient Gel which had been stained with Etidium Bromide and visualised using a UV transilluminator. The thin gel slice was placed in a sterile 1.5 ml eppendorf tube and 40 µl of sterile deionised H_2O added. The tube was then placed in the -70 °C freezer for 1 hour, removed and thawed at room temperature, replaced in the -70 °C for 1 hour, removed and thawed at 4 °C overnight and from this solution 1 µl was removed and used in the PCR (Andersson *et al.*, 1993).

2.5 The Polymerase Chain Reaction

2.5.1 Introduction

The polymerase chain reaction (PCR) amplifies DNA sequences in a selective way and generates large amounts of a discrete DNA product which corresponds to the original target site (Mullis *et al.*, 1986). PCR is a technique for the *in vitro* enzymatic amplification of a specific segment of DNA, which since its introduction has been employed in an enormous number of applications and experiments.

As is well known, the required constituents for amplification of a specific segment of DNA by the polymerase chain reaction are: deoxyribonucleoside triphosphates (dNTPs), a DNA polymerase, 2 single-stranded oligonucleotide primers, a DNA template (genomic DNA or cDNA) and a buffer containing magnesium. The primers are designed

to hybridise to opposite strands of the DNA and are oriented so that synthesis of the new strands extends across the DNA towards the other primer. Both the primers and the dNTPs are supplied in vast excess so as to allow repeated synthesis. For most reactions the polymerase chain reaction can be described in terms of a cycle that is composed of 3 steps (denaturing, annealing and synthesis) and which is performed 25-40 times. In the first step the DNA template is denatured by heating the sample to between 90-95 °C. During the second step the sample is cooled to an optimal temperature for the pair of primers to anneal to their complementary DNA sequences and the third step allows synthesis of the new strands by the DNA polymerase in a 5' to 3' direction. In the first cycle synthesis continues past the sequence complementary to the other primer. However, following the second cycle products are created that are exactly the length from one primer end to the other primer end. In the subsequent cycles the quantity of this product increases exponentially, producing a several million fold amplification of the desired DNA segment for a standard PCR of 25-40 cycles.

2.5.2 Nested PCR Analysis of Prohibitin from cDNA

cDNA was amplified in a 50 µl reaction volume using the standard conditions of 50 mM KCl, 10 mM Tris-HCl, pH 9.0, 1-3 mM $MgCl_2$ (Promega), 0.1 % Triton® X-100 (Union Carbide Chemicals and Plastics Co., Inc.), 0.2 mM each dNTP (Promega), 1 U *Taq* polymerase (Promega) and 15 pmol of each primer (1A and 3B) (Figure 2.6) under the following parameters: initial denaturation of 5 minutes at 94 °C followed by 25 cycles at 94 °C for 1 minute, 56 °C for 1 minute, 72 °C for 1 minute and ended with an extension at 72 °C for 5 minutes. An aliquot of 0.5 µl was then used in a second round of the PCR for 35 cycles using the reaction conditions as described above but with the primer pairs corresponding to the specific fragments (Table 2.1). All products from the PCR were analysed by the electrophoresis of a 5 µl sample of the PCR product through a 2 %

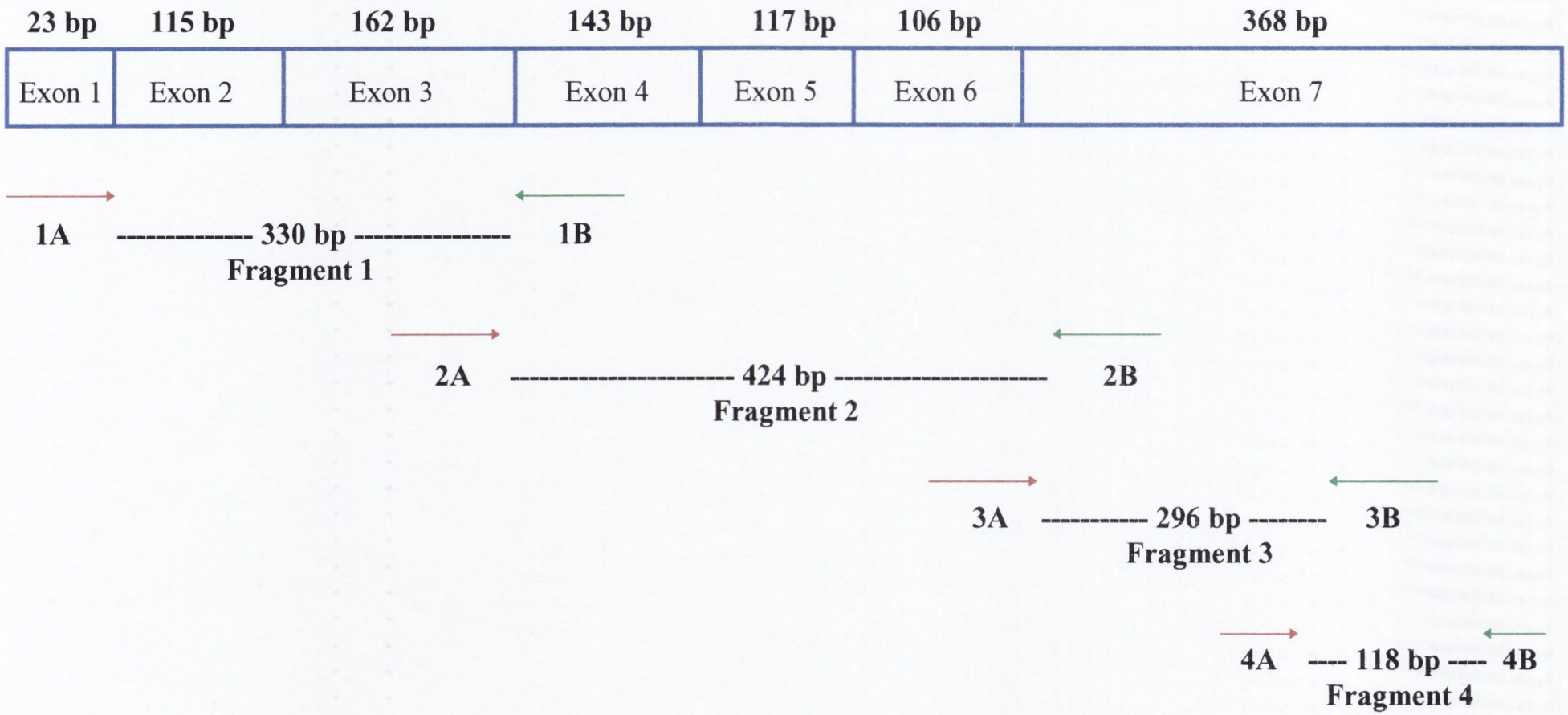


Figure 2.6. Schematic representation of *PHB* cDNA showing the location of the primers 1A-4B used to generate Fragments 1-4. Exons are represented by blue boxes, forward primers by red arrows and reverse primers by green arrows.

Primer Name	Primer sequence (5' - 3')	Amplified Product Name
1A (forward)	TGTGGAGGTCAGAGTGGAAGCAG	Fragment 1
1B (reverse)	40 (GC) - TGCGCAGTGTGATGTTGACATTC	
2A (forward)	50 (GC) - GTGCCAGTCATCACTGGTAGC	Fragment 2
2B (reverse)	CTCAGCAGAGATGATGGCCG	
3A (forward)	60 (GC) - GGAAAAGGCTGAGCAACAGA	Fragment 3
3B (reverse)	20 (GC) - GCAGAAGGAAGGCTGTGT	
4A (forward)	CCACAGCCCCGATGATTCTT	Fragment 4
4B (reverse)	30 (GC) - AGAGATCTGAAGTGATTTTACC	

Table 2.1. Primers used for amplification and DGGE of *PHB* cDNA.

agarose gel which contained Ethidium Bromide (0.5 μg/ml). The products were then visualised on a UV transilluminator and documented.

For all PCRs, from either cDNA or genomic DNA, a negative control reaction which did not contain any template was set up. Specific PCR controls for the detection of genomic DNA contamination of the cDNA were also set up. For each cDNA sample a segment of the β-actin gene was generated by PCR. The primers were chosen so that the product generated from the β-actin cDNA template differed in size from that produced from the β-actin genomic DNA template. This distinction is achieved by designing primers from exon sequence which amplify a product that also contains intron sequence, therefore the genomic DNA product is larger than the cDNA product. The primers used were: forward primer 5' TAC CTC ATG AAG ATC CTC A 3' and reverse primer 5' TTC GTG GAT GCC ACA GGA C 3', these amplify a fragment of 267 bp from cDNA and a 362 bp fragment from genomic DNA (Nakajima-Iijima *et al.*, 1985; Stark *et al.*, 1993; Taylor and Heasman, 1994). The human genome has a number of processed β-actin pseudogenes which do not possess intron sequence (Leavitt *et al.*, 1984). Consequently, contaminating traces of genomic DNA could result in the co-amplification of a product equivalent in size to that generated from the cDNA. Therefore, good laboratory practice as well as treatment of each total RNA sample with RQ1 RNase-Free DNase (Promega) is essential to prevent contamination of the cDNA sample with genomic DNA.

2.5.3 PCR Analysis of Prohibitin from Genomic DNA

Genomic DNA (0.1 μg) was amplified in a 50 μl reaction volume under the standard conditions set out above in section 2.5.2 using the required set of primer pairs (Table 2.2). The reaction consisted of an initial denaturation of 5 minutes at 94 °C followed by 35-40 cycles at 94 °C for 1 minute, 55-58 °C for 1 minute, 72 °C for 1 minute and ended with an extension at 72 °C for 5 minutes. For the amplification of products used

Primer Name	Primer Sequence (5' - 3')	Amplified Product Name
5A (forward) 5B (reverse)	TTATGACATGTCATTGTCCT 40 (GC) - CTAAGATTCCATATCAAGCTG	Fragment 5
6A (forward) 6B (reverse)	ACACTTGTTTTCCTACAGTGG 45 (GC) - GGCAGGTCAGGTTAATGAGG	Fragment 6
7A (forward) 7B (reverse)	45 (GC) - CCTCATTAACCTGACCTGCC AGGCACCTAAACGAGAACTGC	Fragment 7
8A (forward) 8B (reverse)	GTGCTGTCGAAGCAATCACAC 40 (GC) - CTCTGGAACCAATTGCAGAAG	Fragment 8
9A (forward) 9B (reverse)	40 (GC) - TAGCCATCAGGAGGAAGTAAAC AGTTGGCAATCAGCTCAGCTG	Fragment 9

Table 2.2. Primers used for amplification and DGGE of *PHB* genomic DNA.

for DGGE the reaction was followed by denaturation at 94 °C for 20 minutes and a 1 hour incubation at the relevant annealing temperature to ensure maximum heteroduplex formation.

2.5.4 Microsatellite Analysis

The polymorphic markers D17S855, D17S951, D17S579, D17S791, D17S1827, D17S943, D17S588 and D17S1820 were chosen from the Genome Database (GDB) (Anderson *et al.*, 1993; Gyapay *et al.*, 1994; Dib *et al.*, 1996; personal communications cited in the GDB) and were supplied by Research Genetics, Inc., AL, U.S.A.. PCR was performed, using 10 pmoles of each primer (Table 2.3) under the following conditions containing approximately 100 ng genomic DNA in 50 mM KCl, 10 mM Tris-HCl, pH 9.0, 1.5 mM $MgCl_2$ (Promega), 0.1 % Triton® X-100 (Union Carbide Chemicals and Plastics Co., Inc.), 0.2 mM dTTP, dATP and dGTP, 0.05 mM dCTP (Promega), and (α-P^{32}) dCTP (Amersham) at 1 µCi/reaction and 1 U of *Taq* DNA Polymerase (Promega) in a final volume of 25 µl. Amplification conditions, for all the polymorphic markers, consisted of denaturation at 94 °C for 5 minutes, followed by 35 cycles of 1 minute at 94 °C, 1 minute at 56 °C, 1 minute at 72 °C, followed by a 5 minute extension at 72 °C in a PTC-100™ Programmable Thermal Controller (MJ Research, Inc.). The resultant products were diluted 1:1 with loading buffer (98 % Formamide, 10 mM EDTA, 0.025 % Xylene Cyanol, 0.025 % Bromophenol Blue), heated at 94 °C for 5 minutes and rapidly cooled on ice prior to loading on the gel.

A 5 µl sample of the PCR product was electrophoresed through an 8 % polyacrylamide gel, containing 1X TBE and 7 M Urea, at 35 - 40 mAmps for 3 - 4 hours. Gels had been polymerised and cross-linked with 250 µl of 10 % APS, 25 µl of TEMED and left to set overnight. Gels were vacuum dried at 80 °C (Slab Gel Dryer, Savant) and

Locus Symbol	Primer Sequence (5' - 3')	Fragment Size (bp)
Forward	GGATGGCCTTTTAGAAAGTGG	
D17S855		143-155
Reverse	ACACAGACTTGTCCTACTGCC	
Forward	GGCCTCCCAAACTGCTT	
D17S951		170-188
Reverse	TCTACCCCGATGAGCCA	
Forward	AGTCCTGTAGACAAAACCTG	
D17S579		111-133
Reverse	CAGTTTCATACCAAGTTCCT	
Forward	GTTGGCTCCAGTTATTCCCC	
D17S791		165-199
Reverse	GCTCGTCCTTTGGAAGAGTT	
Forward	GTGAGTGGAGATTGCACC	
D17S1827		130-148
Reverse	CACTGGGCTAAGGCTGT	
Forward	TGACTGTAGCCCTCACCC	
D17S943		181-199
Reverse	TTCCAACCATCGTTTGTGTA	
Forward	CCTGGTCTAGGAAGAGTGTCA	
D17S588		154-174
Reverse	GTGTAAGCATCTGTGTATACTAC	
Forward	CATGAGGTCTTCCAGAAGG	
D17S1820		260-278
Reverse	AACACACTTGCTGATGTGC	

Table 2.3. Microsatellite primer sequences.

exposed to film (AGFA, Medical X-Ray Film, 35x43 cm) for 1 - 48 hours at room temperature or at -70 °C with an intensifying screen for optimum exposure.

2.6 Restriction Enzyme Reactions

PCR products were digested under the conditions described by the suppliers (Gibco BRL, New England Biolabs and Boehringer Mannheim). Briefly, 5 µl of PCR product were digested in a total reaction volume of 10-20 µl with 1 unit of enzyme/µg of product, 1 X digestion buffer and incubated at the relevant temperature for 2 - 16 hours.

2.7 DNA Sequence Analysis

PCR products were generated for sequence analysis for Fragments 5, 7 and 8 from genomic DNA and for Fragment 1 from a cDNA template. The primers used for these PCR reactions were the same as those used to generate the PCR products for DGGE analysis except that the GC-clamp regions were not included (Table 2.4). For Fragment 1 PCR products with an altered electrophoretic mobility were excised from a DGG and used to generate more of the variant PCR product (section 2.4.9). After amplification of the specific PCR product the fragments were purified using the Wizard PCR Prep DNA Purification System (Promega), by this method dNTPS and primers are removed. The DNA was then quantitated by measuring the absorbance at 260 nm.

The sequencing reactions for the fragments generated from genomic DNA, were performed using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit with Amplitaq®DNA Polymerase, FS. For each reaction the following reagents were mixed in a labelled 0.5 ml eppendorf tube: 8.0 µl Terminator Ready Reaction Mix, 3-6 µl PCR product at a concentration of 10-30 ng/µl, 3.2 pmole Primer (Forward or Reverse)

Amplified Product Name	Forward Primer (5' - 3')	Reverse Primer (5' - 3')
Fragment 1 (-GC)	TGTGGAGGTCAGAGTGGAAGCAG	TGCGCAGTGTGATGTTGACATTC
Fragment 5 (-GC)	TTATGACATGTCATTGTCCT	CTAAGATTCCATATCAAGCTG
Fragment 7 (-GC)	CCTCATTAACCTGACCTGCC	AGGCACCTAAACGAGAACTGC
Fragment 8 (-GC)	GTGCTGTCGAAGCAATCACAC	CTCTGGAACCAATTGCAGAAG

Table 2.4. Primers used for the amplification of fragments for DNA sequence analysis.

and deionised H_2O was added for a final reaction volume of 20 µl. The tubes were then transferred to the thermal cycler and the following cycling parameters were followed: 25 cycles of 10 seconds at 96 °C, 5 seconds at 50 °C, 4 minutes at 60 °C. After the 25 cycles the tubes were held at 4 °C. The products then required an ethanol precipitation protocol to remove the dye terminators from the solution. For each reaction a 1.5 ml eppendorf tube was labelled and the following reagents added: 2.0 µl 3M sodium acetate, pH 5.2, and 50 µl 95% ethanol. The entire 20 µl of the sequencing reactions was transferred to the 1.5 ml eppendorf tubes, vortexed and placed on ice for 10 minutes. The tubes were then spun for 30 minutes at maximum speed. The supernatant was removed and the pellet was washed with 250 µl of 75% ethanol. Pellets were then dried using a vacuum centrifuge. These samples were then resuspended in a loading buffer containing deionised formamide and 50 mM EDTA at a ratio of 5:1 formamide to EDTA and analysed on the ABI 373 (Applied Biosystems Division, Perkin Elmer) by David Noone of the Department of Genetics, T. C. D.

The Original TA Cloning® Kit (Invitrogen) was used to clone the PCR products that had been generated from cDNA templates. By this method the PCR product is ligated into pCR™II, a linearised plasmid vector, and transformed into One Shot™ (INV αF') competent cells. The plasmid is then isolated and the cloned insert can be sequenced.

Taq polymerase has a non-template-dependent activity which adds a single deoxyadenosine (A) to the 3' ends of PCR products. The pCR™II vector has single 3' deoxythymidine (T) residues which allows PCR products to be effectively ligated with the vector. The ligation reaction was set up using the following procedure: 1 µl of fresh PCR product (preferably less than 1 day old because the single 3' A-overhangs on the PCR products are removed over time, thus reducing the efficiency of the ligation reaction) was added to 1 µl 10X Ligation Buffer (composed of 60 mM Tris-HCl, pH 7.5, 60 mM $MgCl_2$, 50 mM NaCl, 1 mg/ml bovine serum albumin, 70 mM β-mercaptoethanol, 1 mM

ATP, 20 mM dithiothreitol, 10 mM spermidine), 2 μl pCR™II vector (25ng/μl), 5 μl sterile dH_2O and 1 μl T4 DNA ligase (4.0 Weiss units). This reaction was then incubated at 14 °C overnight. The pCR™II vector, sterile dH_2O and 10X Ligation Buffer are supplied in the Original TA Cloning® Kit (Invitrogen).

Once the ligation reaction has been completed the pCR™II vector containing the insert can then be transformed into One Shot™ (INVαF') cells. The Original TA Cloning® Kit (Invitrogen): INVαF' One Shot™ Kit contains: SOC Medium (composed of 2% Tryptone, 0.5% Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose), 0.5 M β-mercaptoethanol and INVαF' cells. Prior to transformation Luria-Bertani (LB) agar plates have to be prepared. LB medium, pH 7.0, is composed of 1.0% Tryptone, 0.5% Yeast Extract and 1.0% NaCl dissolved in dH_2O. For the LB agar, 15g/litre of agar was added to the LB medium and the solution was autoclaved for 20 minutes at 15 lb/sq. in. After autoclaving, the solution was allowed to cool to ~55 °C and ampicillin was added at a concentration of 50 μg/ml (the pCR™II vector has an ampicillin resistance gene which allows for its selection and maintenance in *E. coli*). The plates were then poured, left to harden and once set inverted and stored at 4 °C.

The pCR™II vector contains a *lac* promoter and a *lacZα* fragment, the *lac* promoter is required for bacterial expression of the *lacZα* fragment (the first 146 amino acids of β-galactosidase). Complementation in *trans* with the Ω fragment gives active β-galactosidase and results in the growth of blue colonies in the presence of X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside). The PCR product inserted during the ligation reaction is positioned so as to interrupt the reading frame of the *lacZ* gene (Figure 2.7). This results in the formation of white colonies and allows for the selection of colonies containing inserts. X-Gal is prepared at a concentration of 40 mg/ml in

M13 Reverse Primer

CAG GAA ACA GCT ATG AC C ATG ATT ACG CCA AGC TAT TTA GGT

GAC ACT ATA GAA TAC TCA AGC TAT GCA TCA AGC TTG GTA CCG

AGC TCG GAT CCA CTA GTA ACG GCC GCC AGT GTG CTG GAA TTC

GGC TT **PCR PRODUCT** AAG CCG AAT TCT GCA GAT AT CCA TCA

CAC TGG CGG CCG CTC GAG CAT GCA TCT AGA GGG CCC AAT TCG

CCC TAT AGT GAG TCG TAT TAC AAT TCA CTG GCC GTC GTT TTA C

M13 Forward Primer

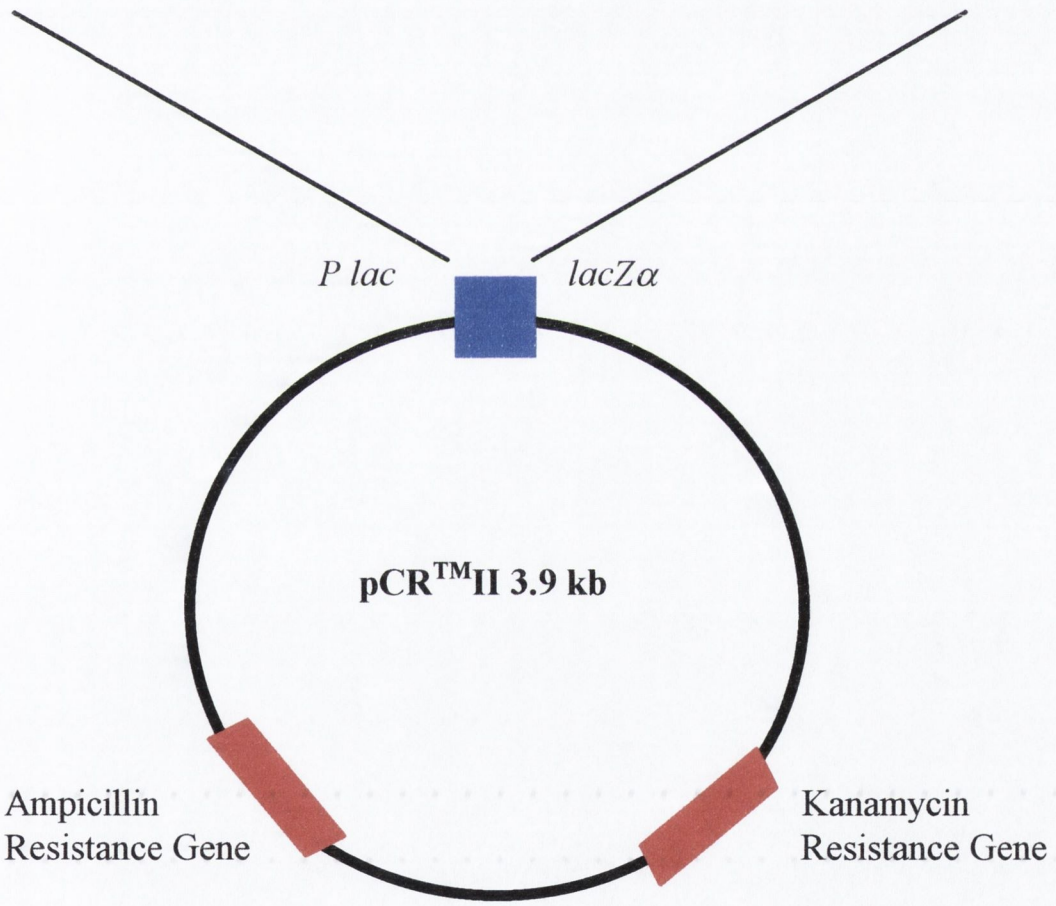


Figure 2.7. Map of the plasmid vector pCR™II. The sequence of the multiple cloning site is shown with a PCR Product inserted. The M13 reverse and forward primer sites are indicated in the green boxes. The *lac* promoter (*P lac*) and *lacZα* fragment are located on either side of the PCR product insertion site. Ampicillin and Kanamycin resistance genes are shown as red boxes.

dimethylformamide. To the LB agar plates with ampicillin, 40 µl of X-Gal was added, spread evenly over the surface of the plates and allowed to dry.

For each transformation reaction 2 µl of 0.5 M β-mercaptoethanol was added to 50 µl of One Shot™ INVαF' competent cells and mixed gently (lysis is caused by pipetting up and down). From the ligation reaction, 2 µl were taken and mixed gently with the competent cells. This was then kept on ice for 30 minutes before being heat shocked at 42 °C for 30 seconds. The vials were then removed and placed on ice for 2 minutes before 450 µl of the SOC medium (at room temperature) was added. The vials were then shaken horizontally at 37 °C for 1 hour at 225 rpm and placed on ice. The LB agar plates containing ampicillin and X-Gal were then labelled and both 50 µl and 200 µl from each transformation vial were spread on separate plates. Once all the liquid was absorbed the plates were inverted and incubated at 37 °C for at least 18 hours. Following this incubation the plates were transferred to a 4 °C fridge to allow for full colour development of the colonies.

To determine the presence and orientation of the insert in the pCR™II vector, 10 "white" colonies were picked using sterile wooden tooth picks (the "white" colonies may have a light blue colour or appear white with blue centres. This is caused by a leaky expression of the *lacZ* fragment or a partial disruption of *lacZ* by the insert). The colonies were then grown overnight in 2-5 ml LB broth containing 50 µg/ml of ampicillin. The plasmid was then isolated (Ausubel *et al.*, 1988) and analysed by restriction endonuclease digests (section 2.6) to determine the size and orientation of the insert. Plasmids containing inserts of the correct size were sequenced by David Noone of the Department of Genetics, T. C. D. with the ABI 373 (Applied Biosystems Division, Perkin Elmer). The sequencing reactions were performed with M13 forward and reverse primers. The pCR™II vector contains these sites on either side of the PCR product insertion site.

CHAPTER 3

DGGE ANALYSIS OF *PHB* FROM cDNA

3.1 Aim of the Study

As described earlier the importance of prohibitin to the cell can be implied from a number of observations and results from varying organisms and experimental systems. *Prohibitin* is an evolutionarily conserved gene with homologues found in all organisms examined. It is expressed in all tissues and cell lines tested. Evidence has been collated indicating a role for prohibitin in regulation of normal cell cycle progression, senescence, development and tumour suppression. Support for the concept that prohibitin may be a tumour suppressor comes from several sources. *Prohibitin* was originally detected as part of a strategy to isolate genes with antiproliferative activity (McClung *et al.*, 1989). Its antiproliferative activity was further demonstrated by the inhibition of DNA synthesis in normal human diploid fibroblasts and cancer cells (Nuell *et al.*, 1991). Tumour suppressor genes represent a subset of genes that negatively regulate cell proliferation and are defined by the connection between the loss of their activity and the development of neoplasia. Prohibitin may have such an association since a probable loss-of-function mutation and three missense mutations have been found in the *prohibitin* gene in sporadic breast tumours (Sato *et al.*, 1992; Sato *et al.*, 1993). *PHB* was localised to chromosome 17q21 and generated intense interest since *BRCA1* had been mapped to this region. This chromosomal region is postulated to contain one or more tumour suppressor genes in addition to *BRCA1* (Futreal *et al.*, 1994). As discussed, the functions of prohibitin and its relationship with the development of breast cancer are currently unknown. Therefore the detection of germline mutations in *prohibitin* in women with breast cancer provides an opportunity to determine prohibitin's role and its place, if one exists, in the pathway to carcinogenesis.

In this study the *prohibitin* gene was analysed for germline mutations in women diagnosed with breast cancer. The women selected from the patient population (*Materials and Methods:* section 2.1.1) were all diagnosed premenopausally with breast cancer. This

group was chosen because women who develop breast cancer at an early age are considered to have a greater likelihood of carrying a germline mutation in a gene that predisposes to breast cancer than women diagnosed with breast cancer at an older age. The patients were divided into two groups, those with a family history of at least one first or second degree relative with breast cancer and those with no family history of breast cancer, termed sporadic cases. The aim of the study was to determine if germline mutations of the *prohibitin* coding sequence could be a causative factor in the onset of premenopausal breast cancer.

To screen the *PHB* gene for mutations, the complete coding sequence was amplified by RT-PCR as four overlapping fragments. The four fragments of the gene were then screened for mutations using Denaturing Gradient Gel Electrophoresis (DGGE) for each individual. Several factors influenced the choice of mutation detection technique; *prohibitin* has a relatively small coding sequence, the full coding sequence was known and the somatic mutations identified to date only involved alterations of one or two bases. Given the type and rarity of mutations it was regarded that a proven highly-sensitive technique for detecting mutations should be chosen. Therefore DGGE was selected for its fidelity in mutation detection, capable of detecting 100 % of single base substitutions and insertions or deletions of a small number of bases (Abrams *et al.*, 1990). Once set up DGGE can be used to screen a large number of samples in a relatively short time. It was also selected as it does not involve the use of radioactive substances or the hazardous chemicals required for other non-radioactive mutation detection methods.

3.2 Results

3.2.1 Fragment and Primer Design for DGGE Analysis of the *PHB* cDNA

The starting point for analysing the *PHB* sequence by DGGE was to generate a melting profile for the prohibitin cDNA sequence. The melting behavior of the whole coding sequence, as well as some of the 5' UTR and 3' UTR, of the *prohibitin* gene (Sato *et al.*, 1993) was analysed by the computer program MELT (Figure 3.1) (Lerman and Silverstein, 1987). MELT predicts the melting behaviour of any given nucleotide sequence and its predictions have been corroborated by a large body of evidence from electrophoretic experiments (Fischer *et al.*, 1983; Myers *et al.*, 1985c; 1985d). Figure 3.1 represents in a graphical format the information contained in the output file of MELT for the complete *prohibitin* gene. This is a plot of the numerical position of the nucleotide within the sequence analysed against the temperature at which the sequence is calculated to have a 50 % probability of being in a helical state. Such a graph is termed a melt map or the melting profile of the sequence.

The melting of neighbouring base pairs, as the temperature is gradually increased, is closely coupled over a substantial number of bases despite the individual base pairs having differing stabilities. The cooperative melting of between roughly 50-300 contiguous bases represents what is termed a "melting domain" or "domain". The base pairs comprising a domain all melt within a narrow temperature interval and each domain is generally separated from other domains on either side by a noticeable difference in melting temperature. For the nucleotide sequence of prohibitin the melt map can be seen to have a number of significant characteristics (Figure 3.1). The first ~320 bases melt as a single domain at 73.5 °C. This is followed by a sharp rise in melting temperature of nearly 3 °C in less than 10 bases marking the end of the first domain. At this point another region of

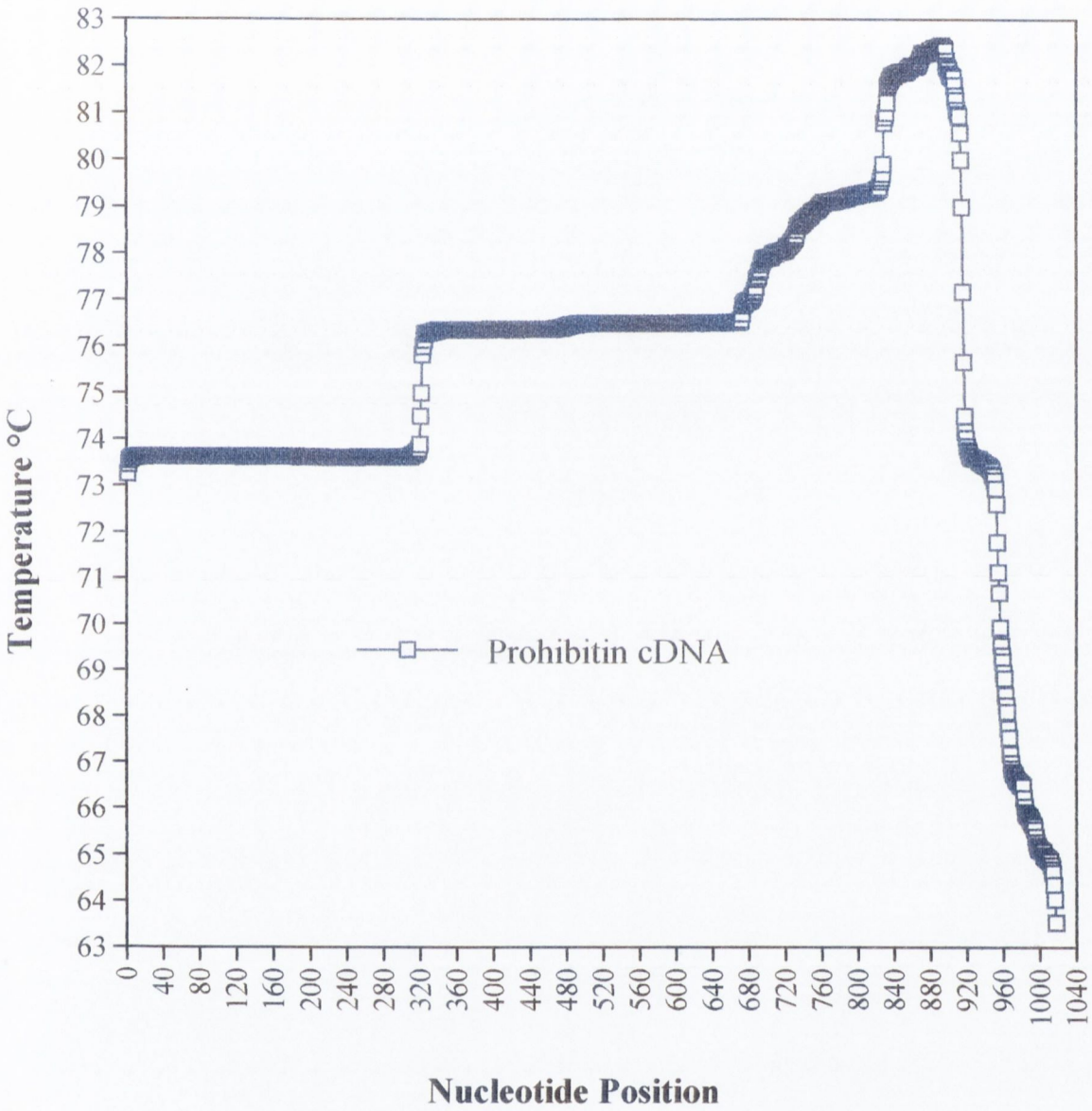


Figure 3.1. Melt map of the prohibitin cDNA sequence.

~350 bases melts as a second domain at 76.5 °C. These first two domains correspond almost exactly to the first 6 exons of the *prohibitin* gene.

The remainder of the melt map represents exon 7 and part of the 3' UTR of the gene. Starting at around nucleotide position 675, the end of the second domain, the sequence melts progressively over an increasing temperature without any domains of significant length for about 150 bases. This region has a high GC content of 61.2 %. The melting temperature then rises sharply over a short sequence of bases before a small incremental rise to the highest melting temperature of the sequence at 82.5 °C due to a GC content of 75 % for this region. This point roughly coincides with the end of the coding sequence of the gene. Finally for the last 125 bases, the 3' UTR, there is a dramatic fall in melting temperature due to the high AT composition (56.8 %) of the 3' UTR sequence as compared to the lower AT concentration in the previous part of the sequence (43.6 %).

This first melt map gives a rough guide as to which regions of the sequence could be amplified together for analysis by DGGE. The melt map outlined indicates that the first two domains defined above should be analysed as individual fragments (Fragments 1 and 2) and the rest of the sequence separated into two further domains (Fragments 3 and 4), one with a high melting temperature profile and the other a low melting temperature fragment. The next step in the process is to analyse shorter regions of the sequence separately. When only the first 330 bases are processed by MELT several features of the way in which double stranded DNA melts can be highlighted. Figure 3.2 shows the melt map for these 330 bases (Fragment 1). Here the configuration of the melt map is quite different to that when the 330 bases were analysed as part of the complete prohibitin sequence, again emphasising the cooperative nature of the process. The initial 110 bases now melt at a higher temperature and are succeeded by a fall in temperature of nearly two degrees. The ensuing 200 nucleotides then melt over a progressively lower temperature. It can be seen that the 700 bases removed from the prohibitin sequence were having a

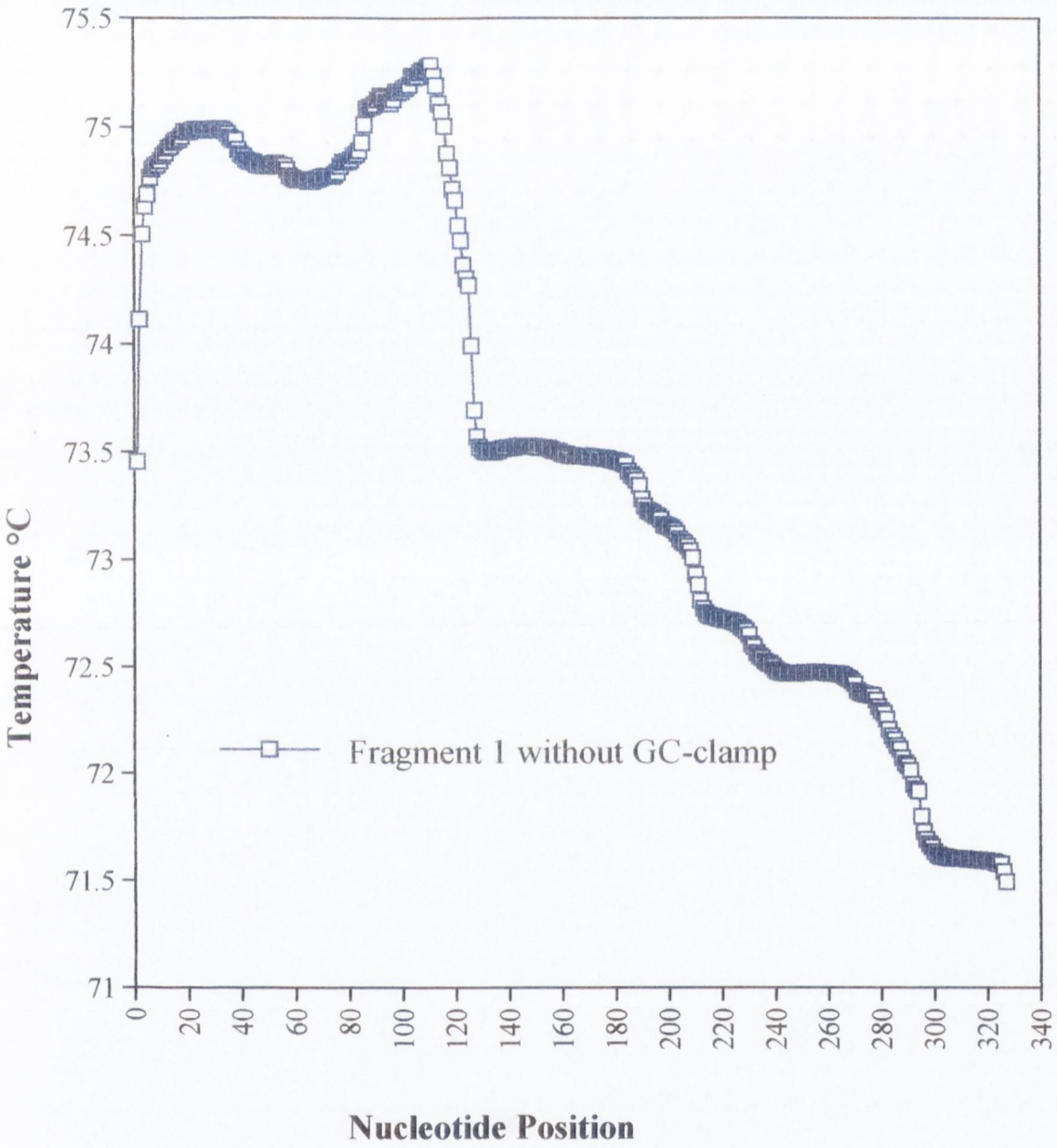


Figure 3.2. Melt map of Fragment 1 without a GC-clamp.

profound effect on the melting behaviour of the first 330 bases and that, without them, the thermodynamic characteristics of the domain are severely altered.

As has been shown the melting behaviour of a given sequence is determined not only by the neighbouring bases of any particular nucleotide but is a property of the whole sequence. To rectify the melting behaviour of the first 330 bases a strategy has been developed whereby a sequence can have additional bases added onto either end which then influence the melting behaviour in a predictable way. This strategy utilises the technique of the polymerase chain reaction, PCR. In order to perform a PCR from any DNA sequence specific primers are made that recognise the desired target sequence. When the primers are synthesised additional bases can be included at the 5' end of the primer. For the scenario described here it is advantageous that the sequence of interest is modified such that it forms a single domain of almost uniform melting temperature. This is accomplished by adding a number of G and C bases to the 5' end of the primer (Sheffield *et al.*, 1989). The PCR then generates products with this "GC-clamp" included at one end of the fragment. As the energy required to disrupt hydrogen bonds between G:C residues is greater than between A:T residues, the addition of a G:C stretch has the effect of raising the temperature required to dissociate the two strands. Adding a GC-clamp to the sequence alters the melting pattern of the entire fragment. Figure 3.3 shows the 330 bases with and without the addition of a GC-clamp. The 330 bases now melt as a single domain of uniform temperature whilst the region incorporating the GC-clamp melts at a significantly higher temperature. This ensures that the lower melting domain can dissociate and thus be analysed by DGGE whilst the GC-clamp prevents the two strands from dissociating completely.

The required length of the GC-clamp is determined by the melting temperature, Tm, of the domain and the log of the dissociation constant, Kd, for the complete strand. In general, as long as the temperature for a Kd of less than -5 is above the Tm then it is

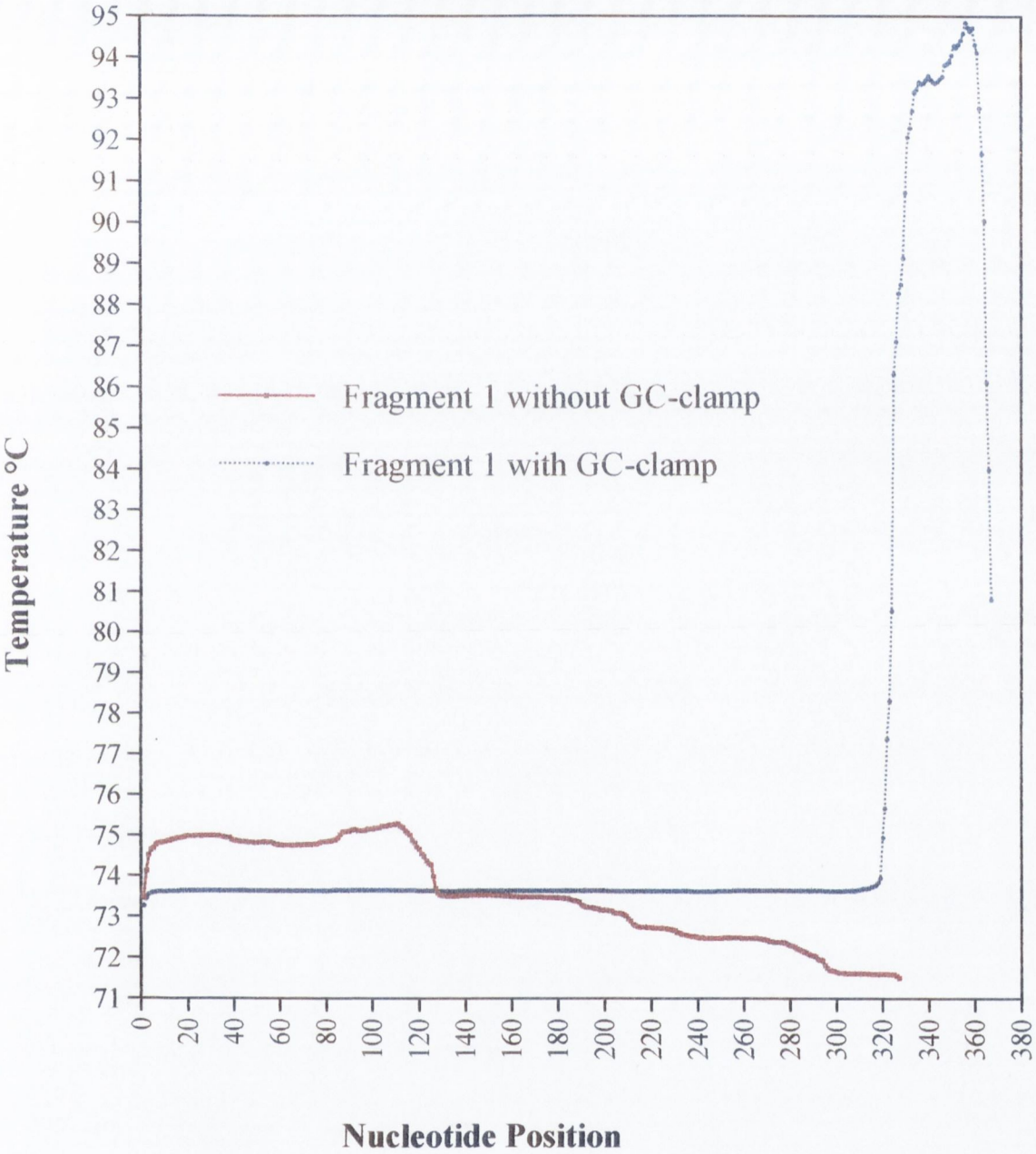


Figure 3.3. Melt maps of Fragment 1 with and without a GC-clamp.

regarded that the partially melted molecule will remain stable in the gel. This has been substantiated in a number of experiments (Lerman and Silverstein, 1992). By increasing the length of the GC-clamp the stability of the molecule increases when in a partially-melted state in the gel.

Having completed analysis of the design of a pair of primers for the first 330 bases of the sequence a second overlapping domain was designed. A similar approach was used for the second domain, this time a GC-clamp was attached to the 5' primer and a small GC-clamp of 10 bases was attached to the 3' primer. This ensured that the complete sequence of interest was confined to a single melting domain. Figure 3.4 shows the second domain with and without the GC-clamp (Fragment 2). The rest of the sequence required a greater amount of analysis due to the range of temperature over which the sequence was predicted to melt. The relatively high melting temperature of bases 700 to 900 compared to the lower melting temperatures of bases 900 to 1020 created a situation where the inclusion of a small number of bases from either region affected the melt map significantly. This situation was resolved by introducing a GC-clamp at both ends of a PCR fragment (Figure 3.5), for a third domain (Fragment 3), which corrected for the inclusion of a second lower melting domain at the 3' end. The fourth domain was then designed with a smaller GC-clamp as the melting domain constituted the last 148 bases of the sequence (Fragment 4).

The melt maps described and illustrated in figures 3.1-3.5 were constructed using calculations predicting that 50 % of the DNA molecule is in a helical state. MELT also produces calculations for conditions where 5 %, 25 % and 95 % of the molecule are in a helical state. These are important as they establish the temperature interval over which a domain melts. In Figure 3.6 three melt maps for Fragment 1 at 5 %, 50 % and 95 % are illustrated on the one graph. The region shown corresponds to the single melting domain and does not include the GC-clamp region of the fragment so as to emphasise the small

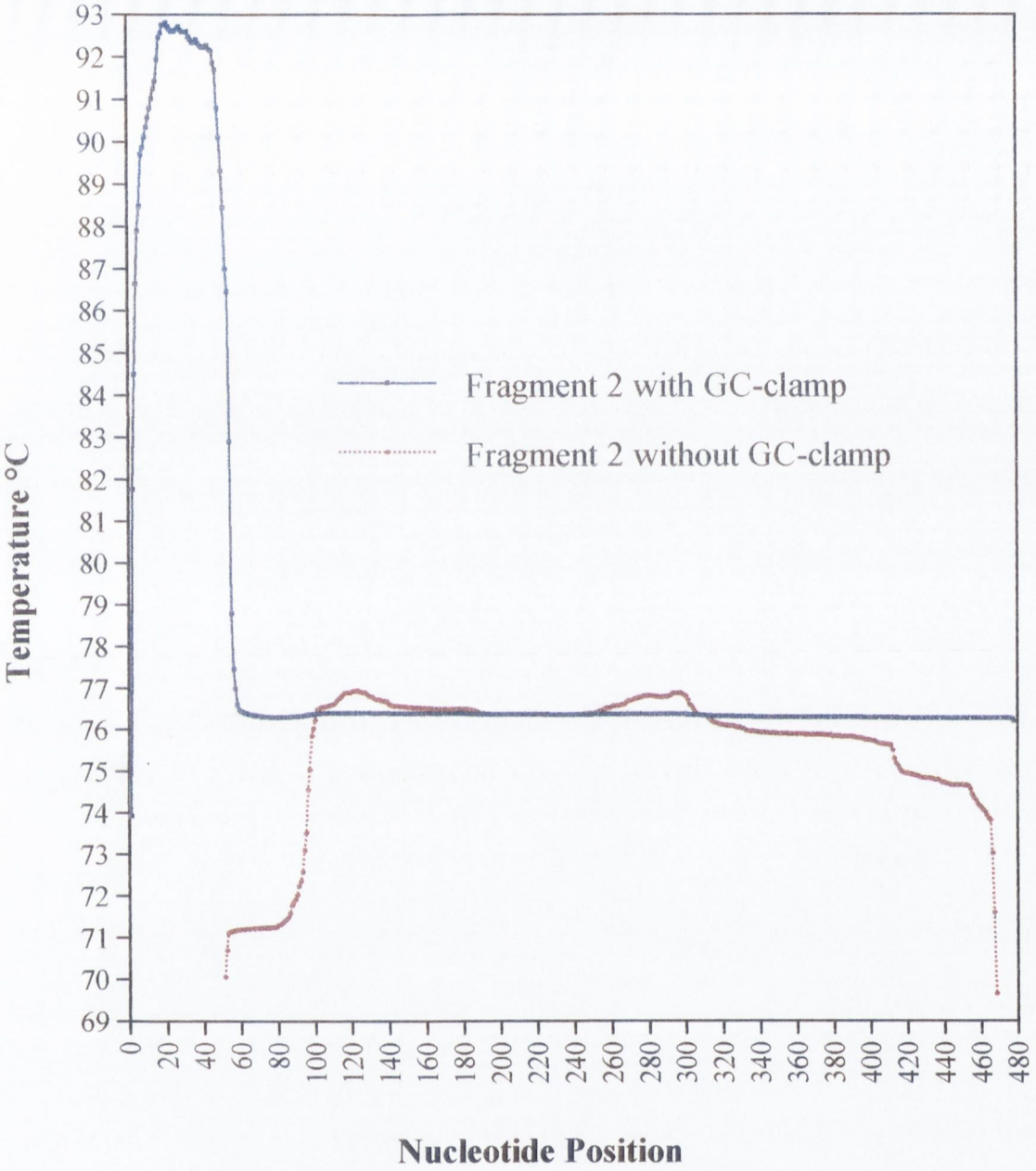


Figure 3.4. Melt maps of Fragment 2 with and without a GC-clamp.

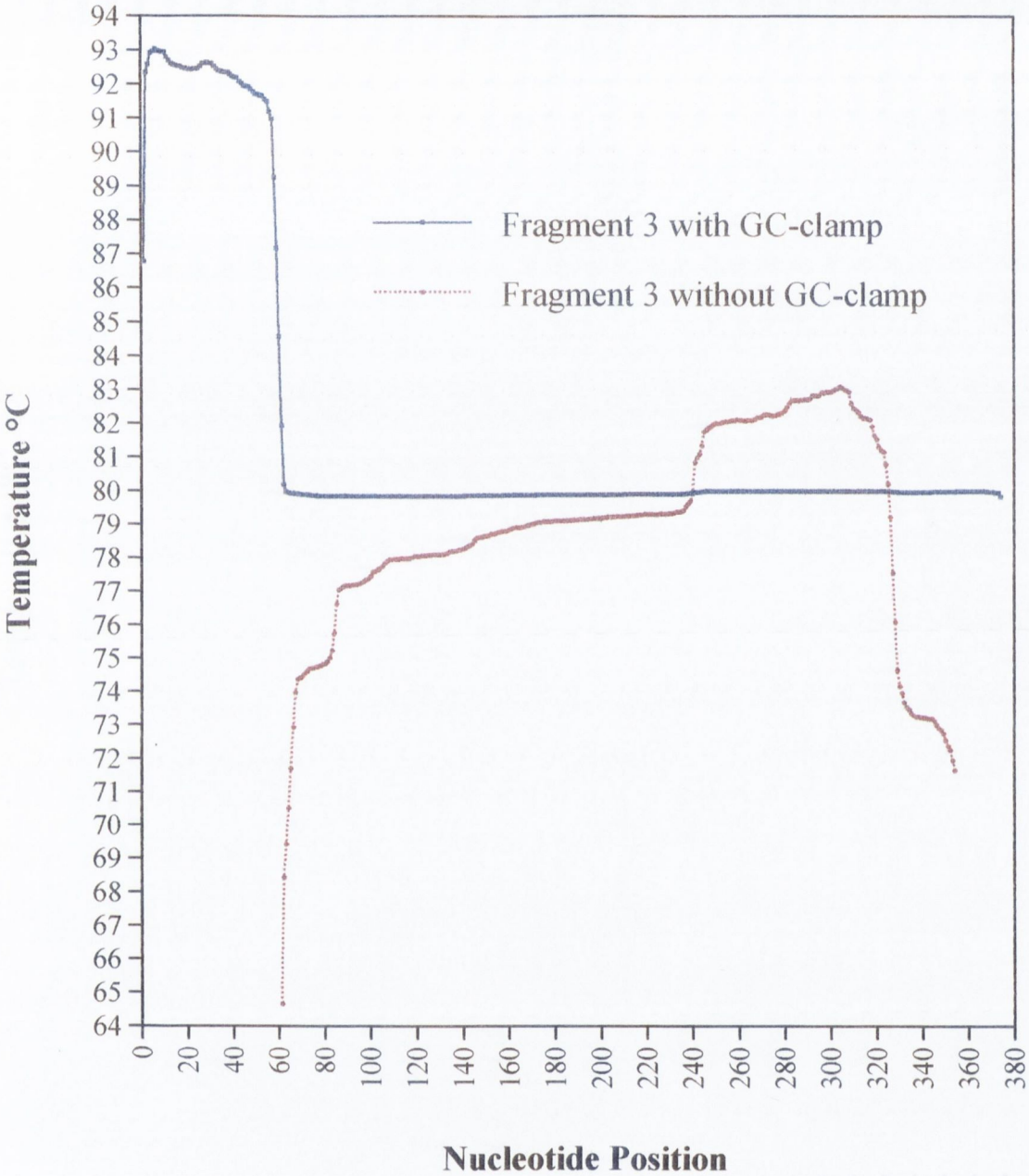


Figure 3.5. Melt maps of Fragment 3 with and without a GC-clamp.

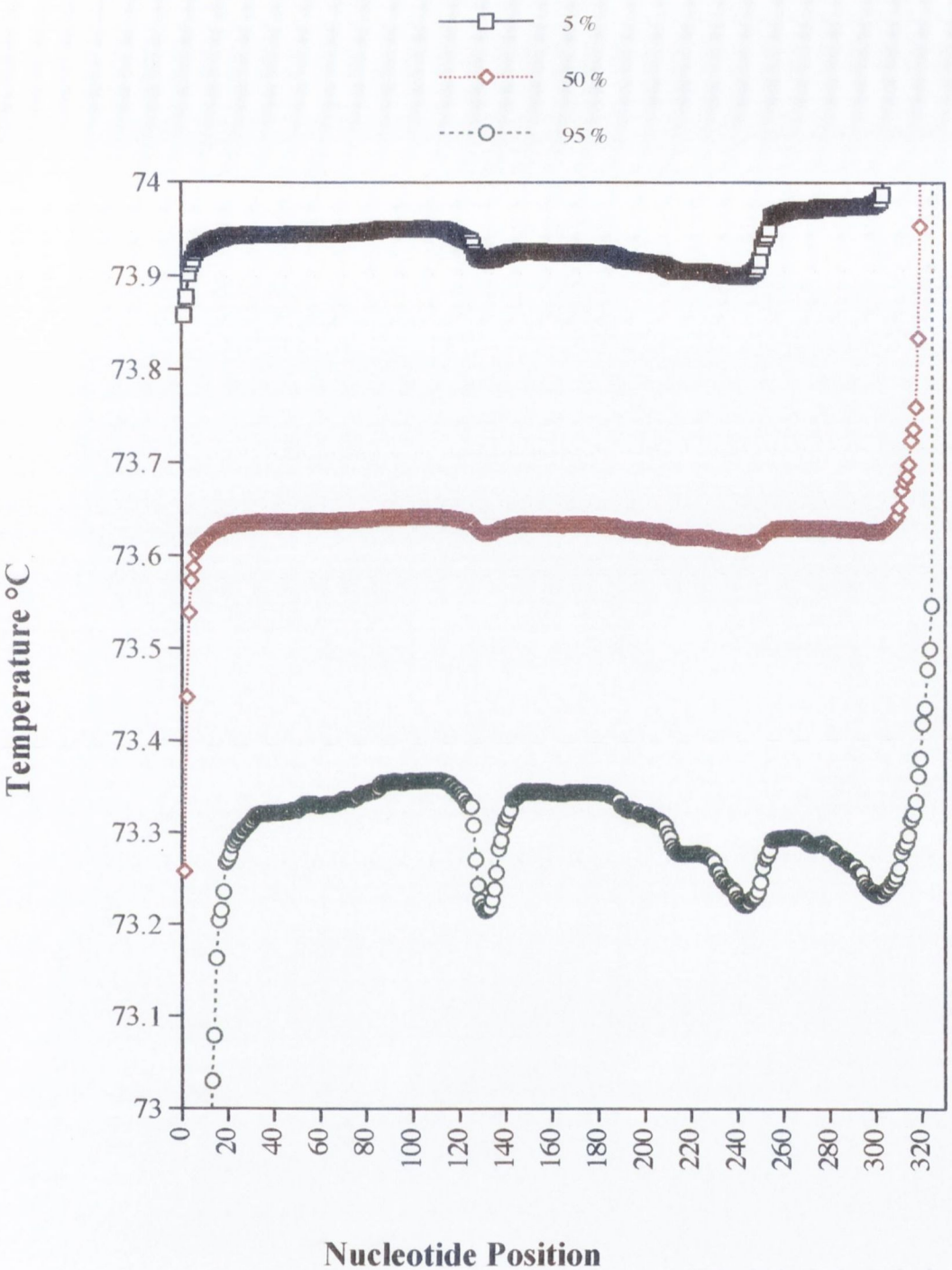


Figure 3.6. Melt maps of Fragment 1 with a GC-clamp for 5%, 50% and 95% of the molecule in a helical state.

temperature difference between the melt maps. The melt maps show that at a Tm of 73.2 °C 95 % of the molecule is still in a helical state and as the temperature increases the proportion of the molecule in a helical state decreases until at nearly 74 °C only 5 % of the molecule is in a helical state. Only for long domains with narrow melting bandwidths is the 50 % contour an accurate guide to melting. Hence considerable time and effort were applied to the construction of fragments with an almost constant melting temperature.

The alteration of Tm by mutations within a sequence can be simulated by modifying the sequence being analysed by MELT. Figure 3.7 depicts the melting profiles of Fragment 1 for the wild type sequence and a single A to G base substitution at position 133. In this case the Tm of the altered sequence has increased relative to that of the wild type sequence due to a single base change in a sequence of 330 bases. Each fragment was analysed in this way for various single base substitutions near the 5' and 3' ends of the particular sequence as well for at least two other sites within the sequence. Sequences with a large number of consecutive Gs and Cs or As and Ts were particularly well scrutinised. For example the A at position 133 in Fragment 1 is in the middle of a 9 bp A/T sequence. A change in Tm was seen for base substitutions throughout the sequences and these simulations are an important step in determining primer selection as it establishes, in theory, that a base change anywhere in the lowest melting domain can be distinguished from the wild type sequence.

The melting profiles of 4 overlapping fragments were determined and primers for PCR amplification were designed. The overlapping nature of the fragments means that the primer sequences for one fragment did not form the primer sequence for the neighbouring fragment. In this way the complete coding sequence was contained in the four fragments and was subject to analysis by DGGE. A review of the MELT analysis shows that Fragment 1 contains exons 1, 2, and 3 (284 nucleotides screened), Fragment 2 contains exons 4, 5, 6 and some of 7 (379 nucleotides screened), Fragment 3 contains exon 7 and

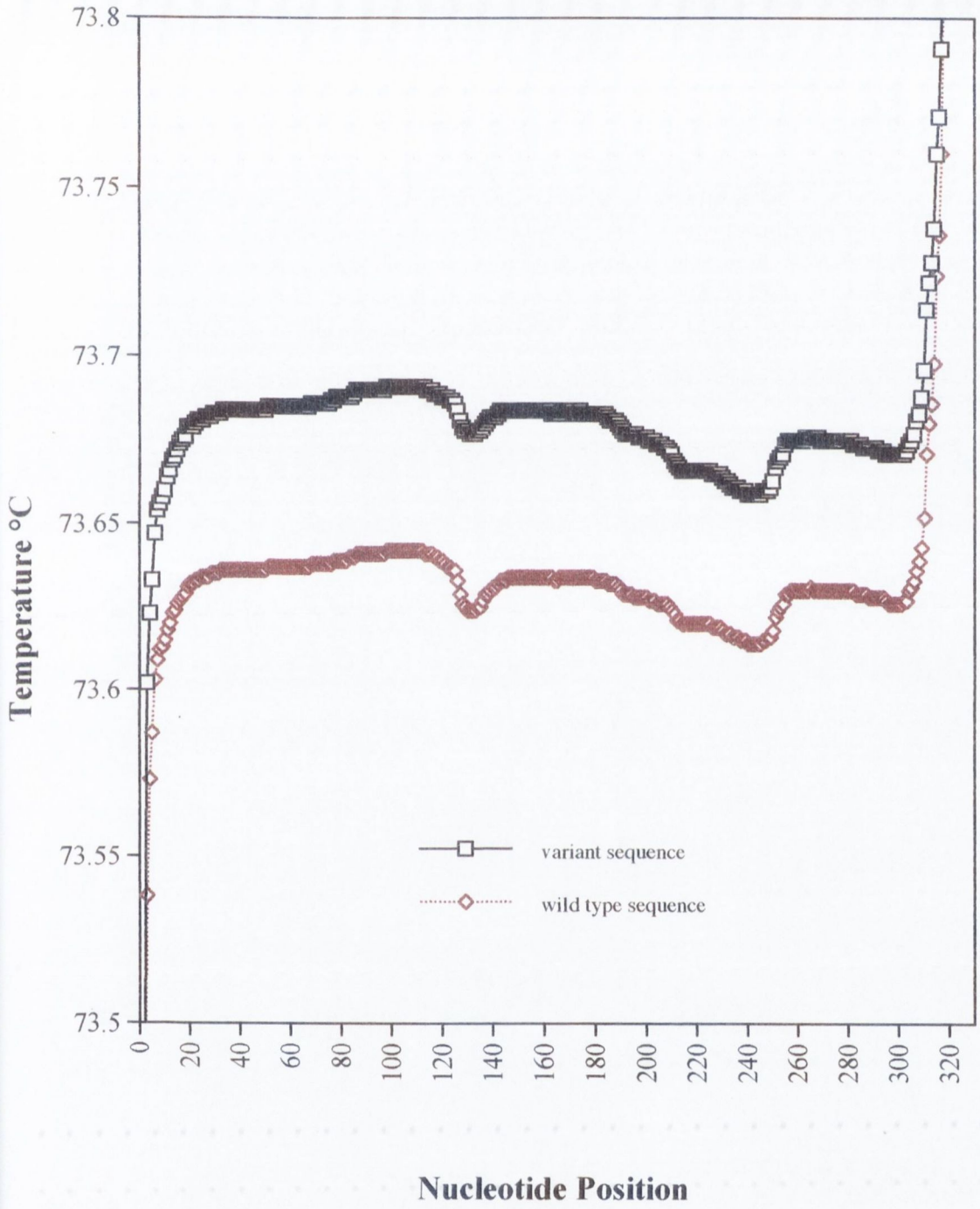


Figure 3.7. Melt maps of Fragment 1 wild type and single base substitution variant sequence.

the 3' UTR (258 nucleotides screened), and Fragment 4 contains the remainder of the 3' UTR (76 nucleotides screened). Therefore by using MELT, primers were designed which enabled all of the coding sequence, 27 bases of the 5' UTR and 158 bases of the 3' UTR (a total of 980 base pairs) of the *prohibitin* gene to be amenable to screening for nucleotide changes by DGGE.

3.2.2 DGGE Analysis of the cDNA of *PHB*

Fragments 1, 2 and 3 were amplified together in a first round PCR and then each fragment was individually amplified in a second round nested PCR. This nested PCR strategy was necessitated by the inclusion of GC-clamps on both primers used for Fragments 2 and 3. When the two primers with GC-clamps were used in the same PCR for either Fragment 2 or 3 it was found empirically that the amount of product yielded was considerably less than when using only one primer with a GC-clamp, as for Fragments 1 and 4. This reduction in quantity of final product is probably due to the formation of "primer dimers" which result from the regions of complementarity in the GC-clamps annealing to each other instead of the target sequence.

Each fragment was subjected to electrophoresis on a perpendicular denaturing gradient gel (DGG), (*Materials and Methods:* section 2.4.5), of 30 %-80 % for 3 hours. An estimate of the percentage denaturant required can be calculated from the Tm predicted by MELT and the equation:

$$\frac{\text{Melt Map Temperature, Tm (°C) - 60 °C}}{0.3\text{°C/\% gradient}} = \text{x \% gradient}$$

This confirmed that each fragment had only one lower melting domain (Figure 3.8) and revealed the range of denaturant over which the fragment melted. Figure 3.8 shows the three sections corresponding to the different stages that a fragment with a single low melting temperature domain being electrophoresed through a gradually increasing

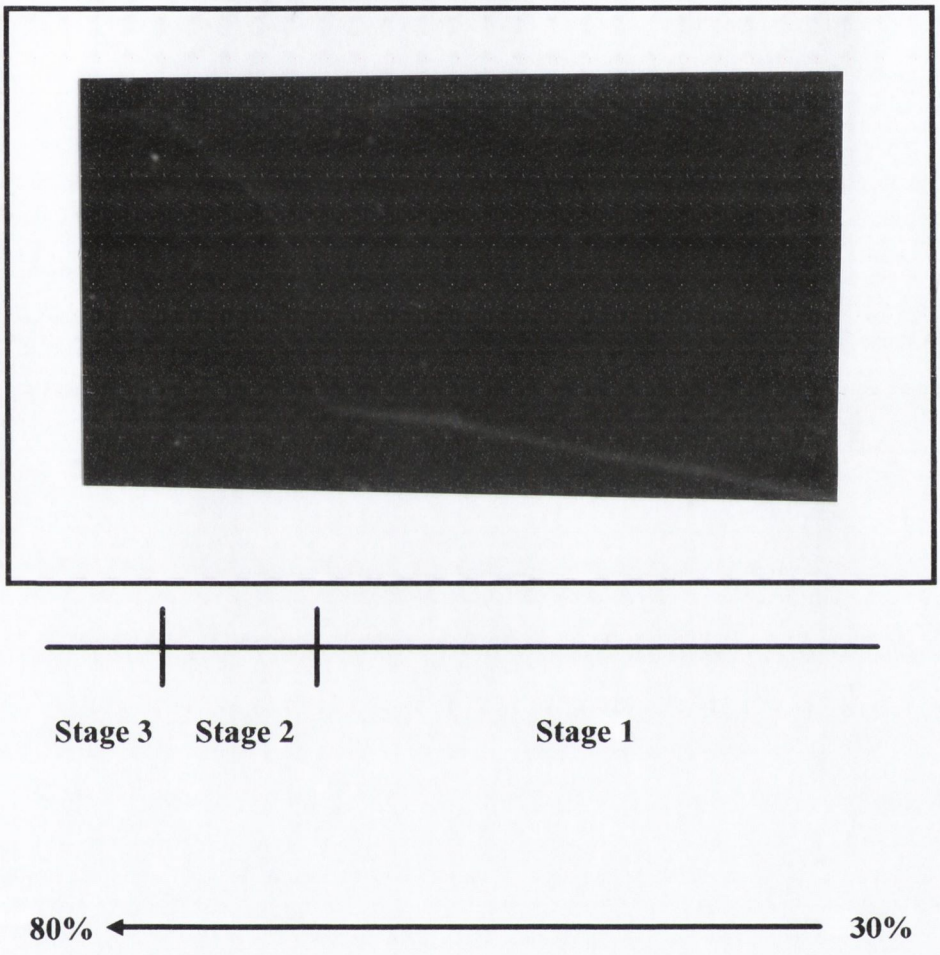


Figure 3.8. Perpendicular Denaturing Gradient Gel of a fragment with a single melting domain. Stage 1: PCR product (case # R039) is double stranded and migrates through the gel without being retarded; Stage2: The lowest melting domain of the PCR product melts and the mobility of the fragment is reduced; Stage 3: The lowest melting domain of the PCR product is completely melted but the fragment remains intact even at the highest denaturant concentraions due to the GC-clamp.

concentration of denaturant finds itself. In the first stage the PCR product is a double stranded fragment and travels through the low concentration of denaturant without being retarded. Subsequently, as the denaturant concentration increases the lowest melting domain begins to melt and the electrophoretic mobility of the fragment is reduced causing the fragment to be retarded in the gel. The lowest melting domain melts completely over a small increase in denaturant concentration and causes the fragment to become severely retarded in the gel. This second stage represents the range of melting temperatures predicted by MELT for the interval between the melt maps for a 5 % and 95 % helical state. The third section of Figure 3.8 represents the molecule with a completely melted lower domain but with the higher domain, in this situation represented by the GC-clamp, remaining double stranded even at very high denaturant concentrations.

The optimal electrophoretic run times for each fragment were determined by performing time courses using parallel DGGs. These were performed by loading a sample of each fragment at time intervals of 1 hour. The optimal time was determined by the time at which fragment retardation occurred in the gel. An example of a time course DGG is shown in Figure 3.9. For the first hour the fragment migrates through the gel without being significantly retarded and does not form a focused band. After 2 hours the rate of migration has decreased considerably. Over the next 3 hours the fragment is only capable, if at all, of migrating through a very small distance in the gel. Details of DGGE conditions for each fragment are given in Table 3.1.

The design of these experiments took into consideration the existence of four intronless pseudogenes *PHB1-4*, located on the chromosomal regions 6q25, 11p11.2, 1p31.3 and 2q21 respectively, as well as the functional *PHB* gene on 17q21 (Sato *et al.*, 1993). From sequence analysis it is presumed that none of the processed pseudogenes are functional and therefore they do not produce an mRNA transcript which could be reverse transcribed into cDNA. *PHB1* is 91.3 % identical to *PHB* at the nucleotide level and any

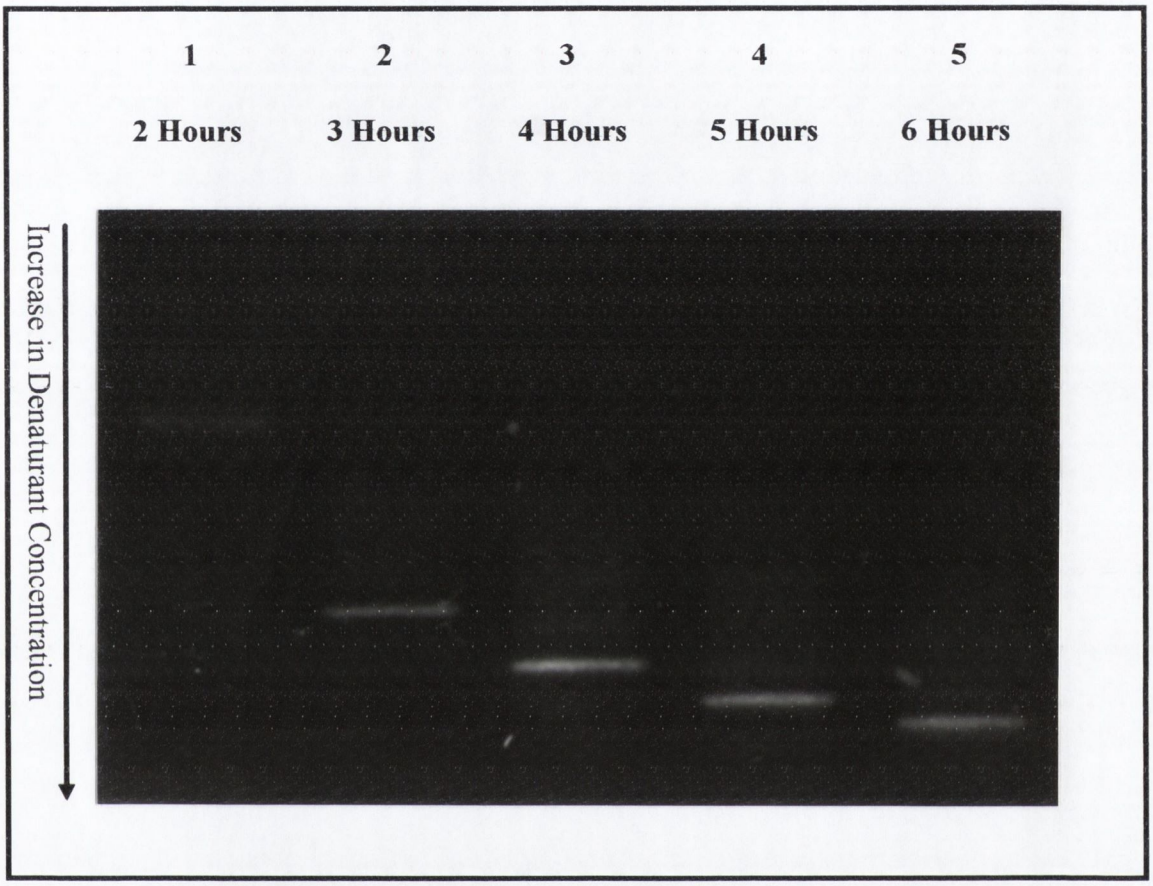


Figure 3.9. Parallel Denaturing Gradient Gel time course. Lane 1: Fragment migrates through gel without being retarded; Lane 2: Fragment melts and migration becomes slower. The band becomes more focused at this stage; Lanes 3-5: Migration of the fragment is severely retarded with minimal migration through the gel. The direction of electrophoresis is shown as being parallel to the denaturant concentration.

Amplified Product Name	Product Size (bp)	Range of DGG Concentration (%)	Total DGG Electrophoresis Time (Hours)
Fragment 1	330	30 - 80	5
Fragment 2	424	30 - 80	6
Fragment 3	296	50 - 100	4
Fragment 4	118	20 - 70	3

Table 3.1. Denaturing Gradient Gel (DGG) conditions for fragments amplified from *PHB* cDNA.

contamination of the cDNA with genomic DNA could cause false positives with regard to identifying sequence alterations in *PHB*. Figure 3.10 shows a DGG of PCR products generated for Fragment 1 from both the cDNA and genomic DNA of two unrelated individuals with breast cancer. This photograph shows that the PCR products generated from cDNA (case # R024 and case # R034) produce a single band retarded in the DGG at the same position. The presence of a single band and absence of heteroduplex bands indicates that the single band represents two homoduplexes with identical sequences. The sequences amplified from the corresponding genomic DNA of these two unrelated individuals with breast cancer (case # D132 and case # D029) appear quite different on the DGG from the PCR products generated from the cDNA. For both samples amplified from genomic DNA more than one band is present, a single focused band (homoduplexes) and above each of these, two slightly diffuse bands (heteroduplexes). All are retarded in the gel at a much lower denaturant concentration than the PCR products generated from cDNA. The heteroduplexes indicate that, for each sample, the single focused band represents two homoduplexes with differing sequences which are retarded at the same position in the DGG. From the position of the homoduplexes in the DGG it is clear that the two samples do not share a common pseudogene sequence. The inability to distinguish two homoduplexes for each sample on the DGG may be due to a very small difference in the melting temperature between the pseudogene sequences or that the denaturant concentration required to separate the two alleles has not yet been reached.

Figure 3.10 demonstrates that the *PHB* pseudogene sequences amplified from genomic DNA differ from the sequence of the functional *PHB* gene and that these differences are detectable by DGGE. For these two individuals the sequence differences between the functional *PHB* and the pseudogenes alter the melting temperature of the fragment significantly whereas for other cases the melting temperature may differ to a

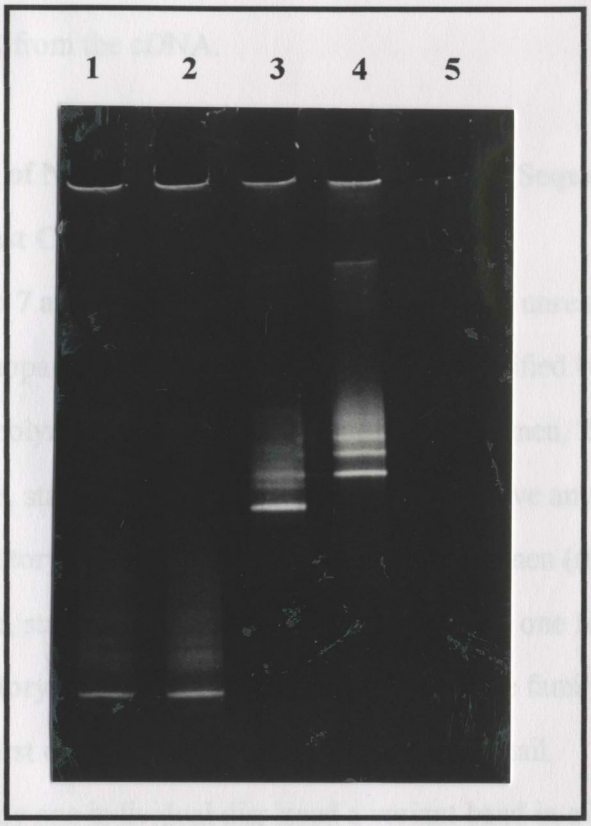


Figure 3.10. Parallel Denaturing Gradient Gel comparing amplified fragments from cDNA and genomic DNA. Lanes 1 - 2: Fragment 1 amplified from the cDNA of two unrelated individuals (case # R024 and case # R034); Lanes 3 - 4: Fragment 1 amplified from the genomic DNA of the two unrelated individuals in Lanes 1 -2 respectively (case # D132 and case # D029). Lane 5: No PCR product loaded.

lesser degree. Therefore it is important that genomic DNA is not co-amplified when performing a PCR from the cDNA.

3.2.3 Detection of Nucleotide Changes in the Coding Sequence of *PHB* in Women with Breast Cancer

Exons 1 to 7 and 158 bases of the 3' UTR from 60 unrelated women, who were diagnosed premenopausally with breast cancer, were amplified by the PCR from cDNA and screened for polymorphisms by DGGE. Of the 60 women, 25 (mean age at diagnosis: 40.12 years of age, standard deviation = 10.17) did not have any first or second degree relatives with a history of breast cancer. The other 35 women (mean age at diagnosis: 41.54 years of age, standard deviation = 7.03) had at least one first or second degree relative with a history of breast cancer. Table 3.2 gives the family history of the 35 women with an affected first or second degree relative in more detail.

In this study one individual displayed a variant band in a heterozygous state in Fragment 1 (Figure 3.11). The aberrant band was excised from the DGG and reamplified under the original PCR conditions with the same primer sequences that had been used to generate Fragment 1 except that the 3' primer no longer contained the GC-clamp (*Materials and Methods:*sections 2.4.9 and 2.7) . The product was then cloned using the TA cloning kit (*Materials and Methods:* section 2.7) and sequenced with M13 primers as the plasmid vector contains M13 priming sites either side of the PCR product insertion site. The fragment was also sequenced from a healthy control in a similar process. The sequence analysis (Figure 3.12) of the aberrant band revealed a G to T transversion, which produces a missense mutation at codon 43, in exon 3, with an arginine residue being replaced by a leucine amino acid residue (Arg43Leu). This sequence variant results in the creation of a site for the restriction enzyme *Sty*1 (Figure 3.13), thus affording a convenient mechanism of screening for the alteration. The individual carrying this germline sequence

Cases of Breast (Br), Ovarian (Ov), Breast-Ovarian (BrO) and Bilateral Breast (BiBr) Cancer				Number of Families: Total = 35 (first, second and third degree relatives)	Number of Families: Total = 35 (first and second degree relatives)
Br	BiBr	Ov	BrO		
> 4				2	1
2	1	2		1	0
4				2	1
3	1			1	0
3		1		1	0
3				6	8
2	1			1	2
1	2			1	1
2			1	2	1
2		1		1	1
2				15	16
1	1			1	2
1			1	1	2

Table 3.2. History of cancer in the 35 women with an affected relative.

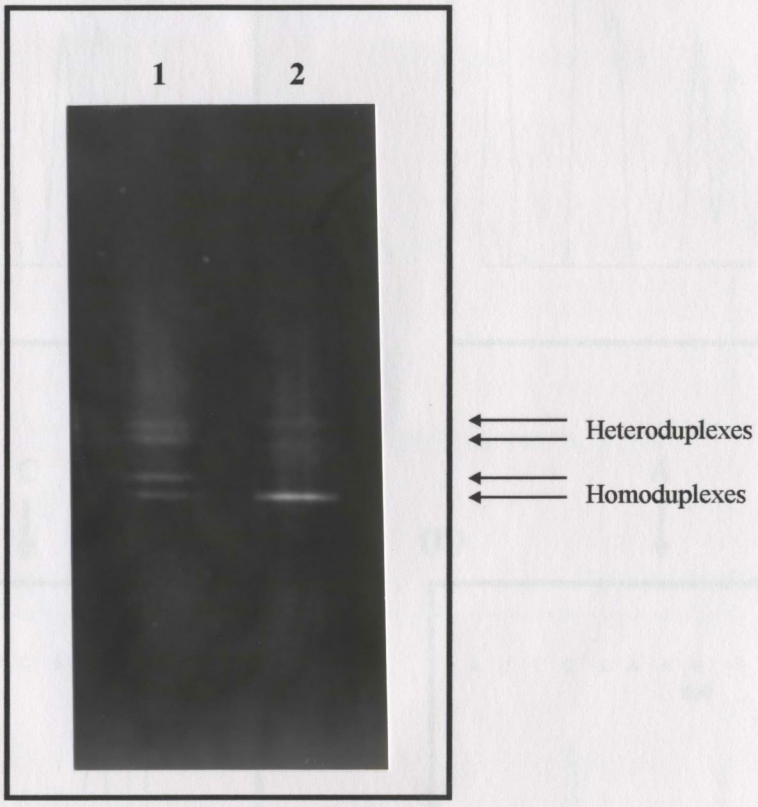


Figure 3.11. Parallel Denaturing Gradient Gel of Arg43Leu allele.

Lane 1 Heterozygous individual for Arg43Leu allele (case # R086).
Lane 2 Homozygous indivdual for wild type allele (case # R116). The shadow above the wild type homoduplex band is a characteristic of electrophoresis of PCR products through a Denaturing Gradient Gel.

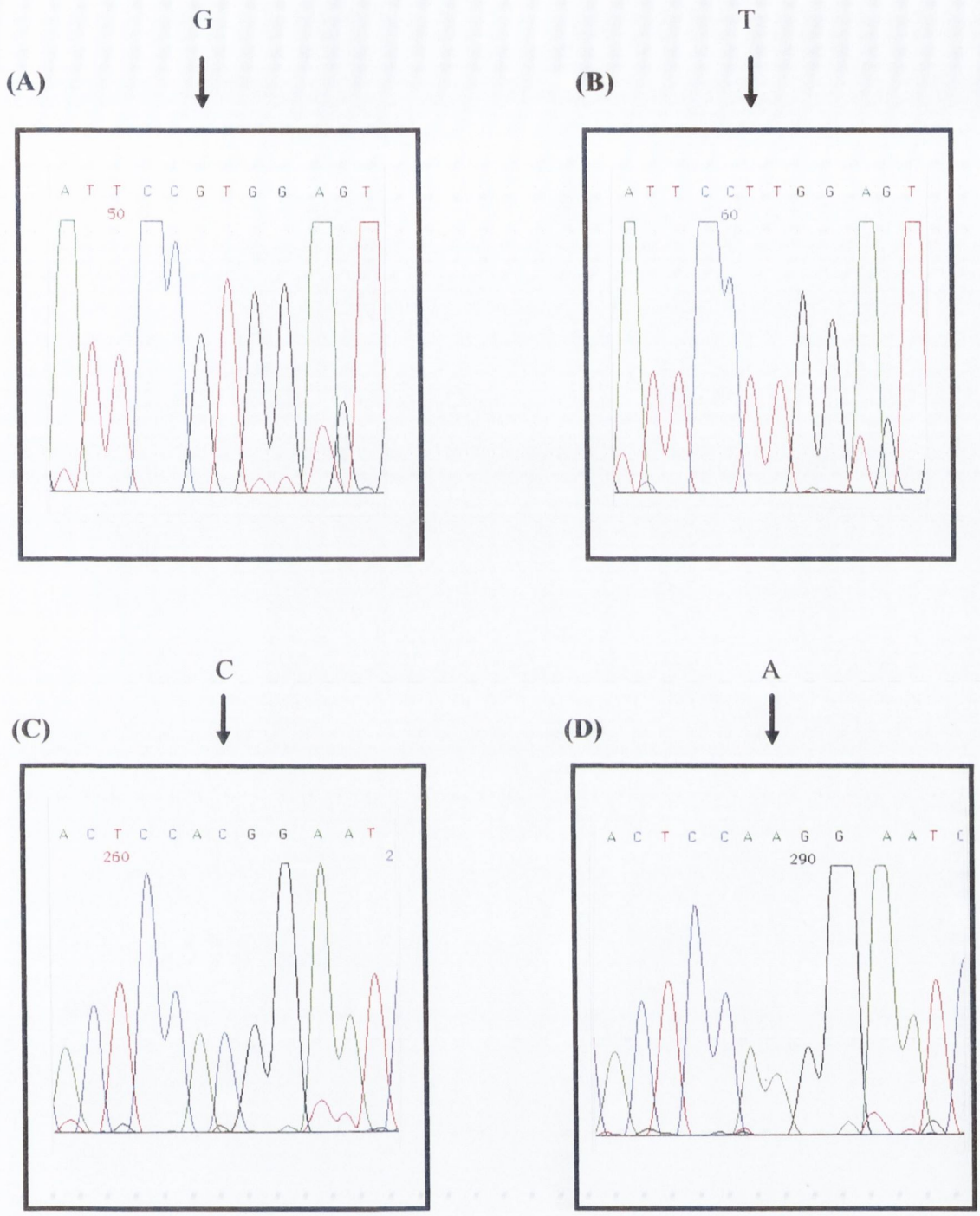


Figure 3.12. Sequence analysis of the Arg43Leu allele. Sequence analysis was performed with the Original TA Cloning® Kit and on the ABI 373. (A) Wild type sense strand sequence (case # R116) (B) Arg43Leu sense strand sequence (case # R086) (C) Wild type antisense strand sequence (case # R116) (D) Arg43Leu antisense strand sequence (case # R086). The arrows indicate the position of the base difference on the sequence trace diagram.

variant was diagnosed with breast cancer at 51 years of age and had a family history of breast cancer with her mother being diagnosed at 59 years of age.

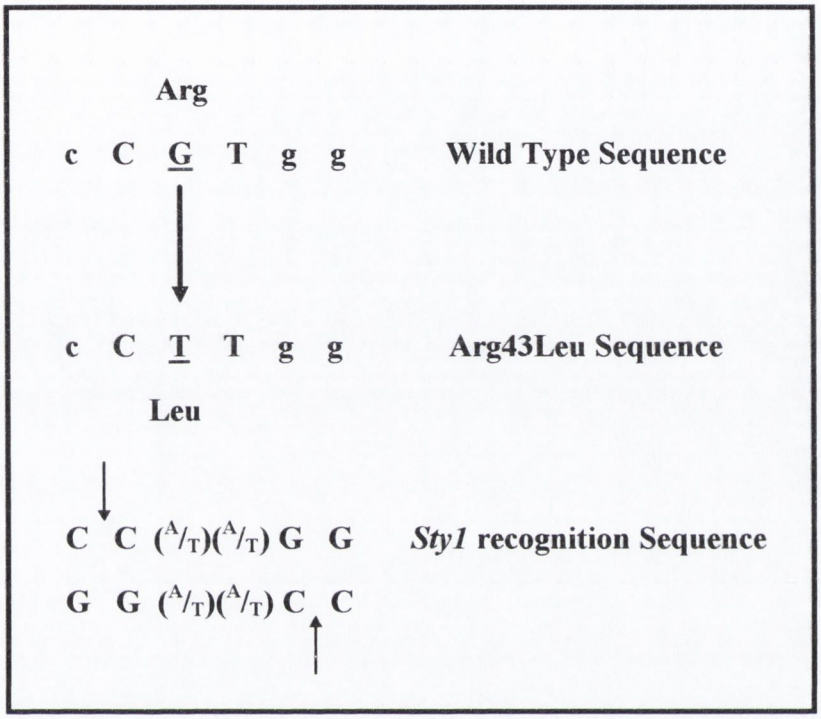


Figure 3.13. Creation of the *Sty1* restricton site by the Arg43Leu allele.

3.2.4 Segregation of the Arg43Leu Allele in Family 1

To determine if the Arg43Leu allele cosegregated with cancer in Family 1 (Figure 3.14) blood and tissue samples were collected from as many family members as possible. The presence or absence of the Arg43Leu variant was confirmed by DGGE, restriction enzyme digest and/or sequencing of PCR products. The PCR product which was used in the restriction enzyme digests was amplified from genomic DNA with primers designed for sequences in intron 2 and intron 3. In this way no pseudogene sequences would be co-

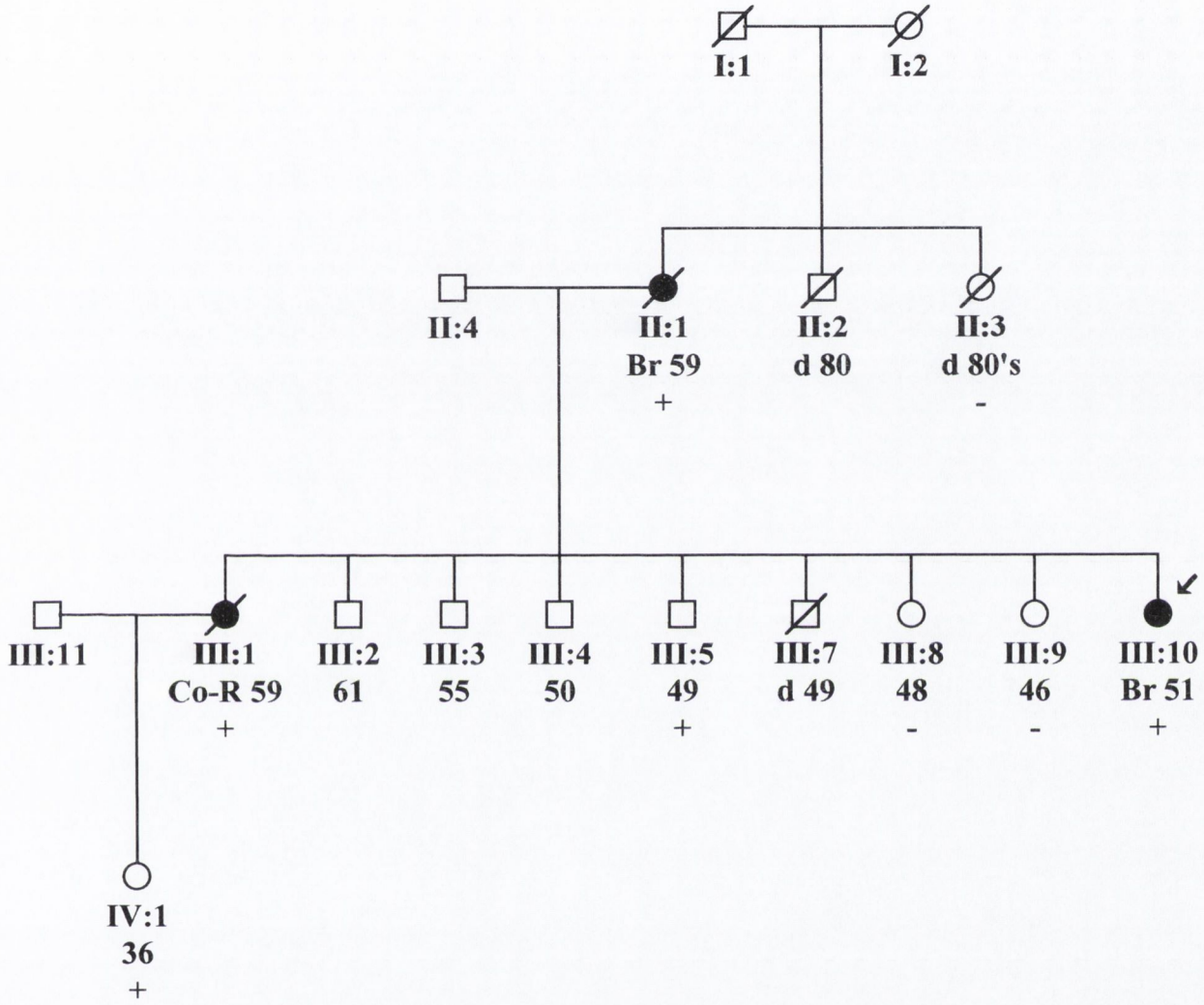


Figure 3.14. Pedigree of Family 1. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer.

amplified as they do not possess any intron sequences. The sequences of the primers used were, forward primer Exon3F: 5' AGT GAC TGC TTA GTT TGC AGC TAG 3' and reverse primer, Exon3R: 5' GCA GTA GTC TAG TCA TTC CTA TG 3'. The size of the PCR product generated using these primers is 477 bp.

It is important when performing restriction endonuclease digests to confirm that the enzyme is active under the given reaction conditions and that digestion has taken place where recognition sequences are present. One way of doing this is by including a recognition sequence for the restriction endonuclease in the PCR product generated. This can be achieved by designing a recognition sequence for the enzyme within one of the primers, or if convenient, a nearby recognition sequence can be included within the sequence amplified. Another approach would be to include a second PCR product within the same digestion reaction that has a known recognition sequence for the restriction endonuclease. Comparison of the PCR products before and after the digestion reaction would then be used to assess the performance of the enzyme and determine if the enzyme had cleaved the PCR product of interest. In this study, and all other restriction endonuclease digests, the PCR product for each individual was generated on three separate occasions and a digestion reaction performed on each of these in order to ascertain the presence or absence of the restriction endonuclease recognition sequence. The nearest *Sty*1 recognition sequence, to the one created by the Arg43Leu variant, in the *prohibitin* genomic DNA sequence is several hundred base pairs downstream of exon 3. For some family members genomic DNA was extracted from formalin-fixed paraffin-embedded (FFPE) tissue. Considering the size of the PCR product required to incorporate the downstream recognition sequence this option was not pursued.

Only tissue in the form of FFPE sections was available from family members II:1 (mother), II:3 (maternal aunt) and III:1 (sister) for analysis from genomic DNA. Blood

samples were obtained from family members III:5, 8, 9 (siblings), and IV:1 (niece) for analysis by cDNA and genomic DNA.

Family members III:1, diagnosed with an adenocarcinoma of the large bowel at 59 years of age, II:1, diagnosed with an infiltrating ductal carcinoma at the age of 59, III:5, unaffected at 49, and IV:1, daughter of family member III:1 and unaffected at age 36 are all carriers of the Arg43Leu variant. The other 4 relatives tested did not have the variant allele and were all unaffected by cancer at ages 36, 46, 48 and 80. The other 5 brothers of the index case were unavailable for testing and were all reported, by the other family members, to have no history of cancer.

3.2.5 Loss of Heterozygosity (LOH) Analysis of Tumour Samples

Of the three cancers presented in Family 1 normal and tumour tissue was only available from the index case's sister (Individual III:1) who had bowel cancer. Microdissection of the normal and tumour tissue was performed as described (*Materials and Methods:* section 2.3.6). Following microdissection of the normal and tumour tissue and amplification of exon 3 by PCR, with the primers described in section 3.4.4, the PCR products were digested with the restriction endonuclease *Sty*1 (Figure 3.15). Three bands are present when the PCR product from the normal tissue is digested, a single full-length product (477 bp) indicating the wild type sequence and two smaller products (262 bp and 215 bp) resulting from the sequence variant product being cleaved in two by *Sty*1. When the PCR product amplified from the tumour tissue is digested no full-length product is seen, only the two smaller products resulting from the digestion of the sequence variant product, thus demonstrating loss of heterozygosity for the *prohibitin* allele Arg43Leu. The negative image from the photograph of the agarose gel exhibited in Figure 3.15 strikingly shows the loss of the wild type allele in the tumour tissue when digested with the restriction endonuclease *Sty*1. This amplification and digestion of the tumour tissue was

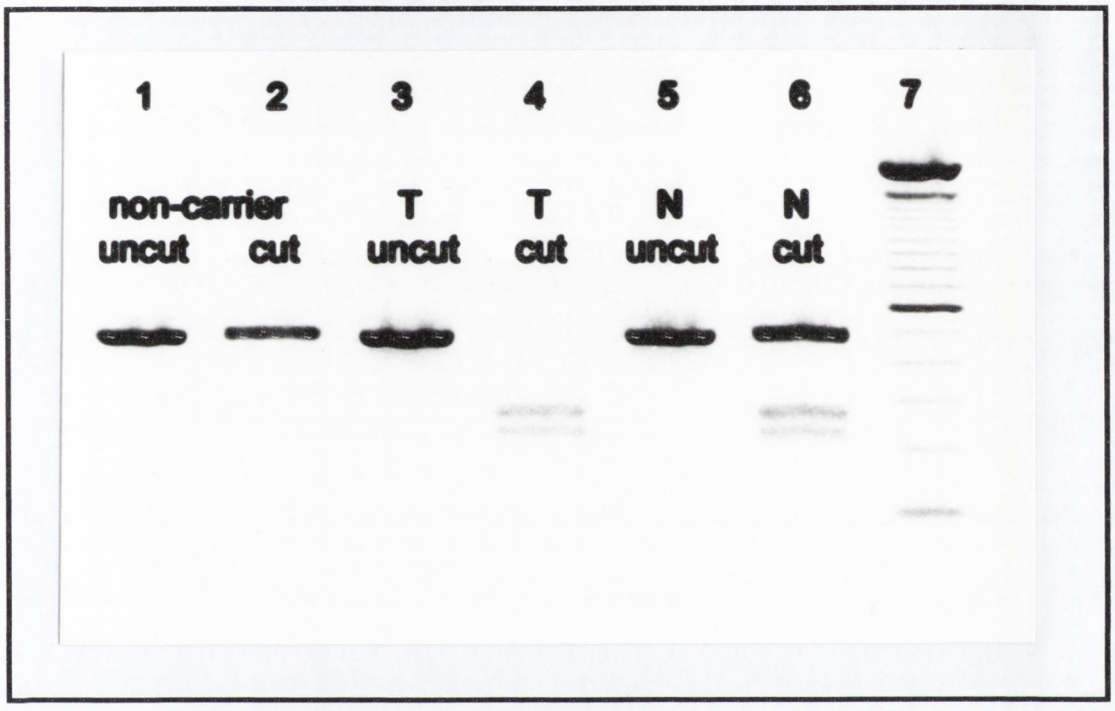


Figure 3.15. Analysis of the PCR Products of exon 3 from Tumour and Normal tissue when digested with the restriction endonuclease *Sty1*. Lanes 1-2: Amplified and digested PCR products of an individual without the Arg43Leu allele (case # D017); Lanes 3-4: Amplified and digested PCR products of Tumour (T) tissue from Family 1: III:1; Lanes 5-6: Amplified and digested PCR products of Normal (N) tissue from Family 1: III:1; Lane 7: 100 bp size marker.

repeated three times and on each occasion no full-length product or only a small percentage of the full-length product remained undigested, indicating that the microdissected tumour tissue did not contain a significant proportion of non-neoplastic cells.

3.3 Discussion

This work represents a comprehensive screening of the complete human *prohibitin* gene (*PHB*) coding sequence for germline mutations in 60 women with breast cancer by Denaturing Gradient Gel Electrophoresis (DGGE) (thirty-five women with breast cancer and a family history of breast cancer, and twenty-five women with breast cancer but no reported family history of breast cancer). In this study the index cases were selected on the basis of having been diagnosed premenopausally with breast cancer. Age or menopausal status were otherwise not used as criteria for classification of a family history of disease. Previous studies attempting to detect germline or somatic mutations of *prohibitin* in breast cancer have either selected women with breast cancer diagnosed at <40 years of age (Tokino *et al.*, 1993) or breast tumours with evidence of LOH on chromosomal region 17q and/or breast cancer in patients 35 years old or younger (Sato *et* al., 1992; Sato *et al.*, 1993). No germline alterations affecting the exon sequences of *PHB* were detected in these studies. Sato *et al.* (1993) used the RNase protection assay to screen 120 primary breast tumours and discovered 4 somatic mutations affecting the *PHB* coding sequence. Since mutations appeared to be very rare, a technique that showed maximum sensitivity was required to screen *PHB* for germline alterations. DGGE analysis was chosen because it is the most sensitive screening method for detecting point mutations and deletions or insertions of a small number of nucleotides. It has also been widely used to screen genes involved in genetic diseases (Blanquet *et al.*, 1993; Stoppa-Lyonnet *et al.*, 1997).

In this investigation sequence alterations were identified in 1/60 women with breast cancer or 1/35 women with a family history of breast cancer and 0/25 women with sporadic breast cancer. Not surprisingly, due to the low frequency of the Arg43Leu allele and the small numbers screened the difference in the incidence of mutation between the two groups does not reach statistical significance. The woman carrying the Arg43Leu variant was diagnosed with an infiltrating ductal carcinoma at 51 years of age and represented the oldest affected individual in the population screened. Her mother, the only other family member affected by breast cancer, was diagnosed at 59 years of age with a poorly differentiated invasive adenocarcinoma.

The DNA sequence alteration was a single base substitution and resulted in an arginine to leucine amino acid change at codon 43 (Arg43Leu) in exon 3 of *PHB*. This is the first documented germline alteration of the *PHB* gene. *PHB* is a highly-conserved gene through evolution. Rat and mouse *prohibitin* genes are identical and only differ from the human *PHB* by a single conservative amino acid substitution at codon 107 in exon 4. As described, homologues of *prohibitin* have been found and sequenced in a wide range of species (Eveleth and Marsh, 1986; Nuell *et al.*, 1991; Sato *et al.*, 1992; Terashima *et al.*, 1994; McClung *et al.*, 1995; Snedden and Fromm, 1997). In *Arabidopsis thaliana*, *Nicotiana tabacum*, rat, mouse and human the arginine residue corresponding to the human codon 43 is conserved in all 5 species. In *Drosophila* and yeast the arginine residue is replaced by the neutral, polar amino acids threonine and asparagine respectively. The germline alteration identified here, Arg43Leu, replaces the positively charged, polar arginine residue with the aliphatic hydrophobic amino acid leucine, a non-conservative change. The effect of substituting amino acids with such different characteristics on the structure and/or function of the prohibitin protein is unknown. From the predicted secondary structure codon 43 resides in the middle of a β-sheet but does not coincide with any putative sites of post-translational modification (McClung *et al.*,1995) or residues

involved in a possible signal sequence. However, the amino acid, leucine, is more commonly found in α-helices than β-sheets.

As stated before, *PHB* is a highly conserved gene and is 99.6 % identical to the mouse and rat sequences at the protein level. A study of protein sequence conservation between orthologous mouse and human sequences shows that these vary between 36-100 % identity, with an average value of 85 % (Makalowski *et al.*, 1996). Of the 1196 protein sequences included in the study only 27 were 100 % identical at the protein level, including the very-slowly evolving genes such as the actins and histones. Therefore, on this scale of measurement *prohibitin* is one of the most highly conserved genes known. Interestingly BRCA1 was amongst the most divergent proteins between mouse and human with only 57.2 % identity for its protein sequence. The low divergence rate between the mouse and human *prohibitin* genes may indicate that germline alterations are largely unfavourable and would therefore be uncommon in humans.

The spectrum of somatic mutations in *PHB*, as currently known, was identified from a Japanese population (Sato *et al.*, 1992; Sato *et al.*, 1993). In total, 3 missense mutations and one 2 base pair deletion which consequently produces a premature stop codon have been documented. In both studies tumours that exhibited LOH for 17q and/or were diagnosed in women ≤35 years of age were selected. The results of our study, the identification of a single missense mutation, Arg43Leu, in exon 3 in 120 chromosomes, indicate that germline mutations in *prohibitin* are not common. The findings of Tokino *et al.* (1993), who did not detect any germline mutations in *PHB* in women diagnosed with breast cancer at ≤40 years of age, are in agreement with these results. Tokino *et al.* (1993) did not find any germline mutations in exons 4, 5, 6 and 7 of 76 women. Additionally exons 2 and 3 were sequenced in 9 women with an affected first degree relative with breast or ovarian cancer. Therefore, Tokino *et al.* (1993) concluded that *prohibitin* is not an early-onset breast cancer susceptibility gene. Given our results, the Arg43Leu allele

would be unlikely to be found in such a small population. Moreover, if the Arg43Leu variant predisposes an individual to breast cancer at a later age of onset of disease then, due to the ascertainment criteria used, the other studies may have selected against the presence of the mutation in their populations regardless of the number of individuals or tumours studied.

The missense mutation, Arg43Leu, in exon 3 cosegregates with the 3 individuals affected by cancer in Family 1. With only a small number of individuals in the family affected by cancer there are several possible explanations as to why such a segregation pattern is observed. Firstly, the Arg43Leu allele may represent a cancer-predisposing mutation that is associated with a late age at onset of disease. Most known cancer susceptibility genes have been identified by linkage studies of families with large numbers of individuals affected at a young age and for genes which possess highly penetrant mutations. Genes that have less penetrant alleles are less likely, by definition, to cause large numbers of cancers within a family. The difficulty of detecting genes with less-penetrant mutations by linkage analysis is increased by the occurrence of sporadic cases of cancer. Also, the existence or exclusion of genetic heterogeneity is difficult to establish. Therefore the *prohibitin* allele, Arg43Leu, may be an example of a mutation in a cancer susceptibility gene which has a lower penetrance with a later average age at onset of disease. Evidence in support of this idea comes from the discovery of loss of heterozygosity (LOH) in one of the tumours in Family 1. LOH is the molecular genetic hallmark of tumour suppressor inactivation and has been observed for the bowel cancer, the one tumour where both normal and tumour tissue samples were available.

Secondly, for many common cancers, first degree relatives of patients are at an increased risk for cancer at the same site of around two- to five-fold that in the general population (Houlston *et al.*, 1991; Goldgar *et al.*, 1994). Such population-based calculations of risk include individuals with an inherited predisposition and those that

develop cancer due to nongenetic influences. The mother and daughter in Family 1 both have breast cancer and this may be due to the increased risk associated with diet, lifestyle or other environmental factors, or to a predisposing gene. Breast and bowel cancer are the second (~17 %) and fourth (~6 %) most common cancers respectively to affect Irish women (National Cancer Registry figures for 1994). The figures include non-melanoma skin cancer. The age at onset of disease for all 3 women is, not surprisingly, older than the average age at onset of disease found in families that have been selected for multiple cases of either cancer type as these tend to contain cases at a younger age at diagnosis. Therefore distinguishing between a rare mutation that causes a later age at onset of disease and a chance clustering of common cancers associated with the allele is particularly difficult.

Thirdly, the *PHB* allele may be tightly linked to another gene which is actually responsible for cancer susceptibility. An example illustrating such a case was that of a missense mutation in the adenovirus E1A enhancer binding protein on 17q21 which was found to cosegregate with cancer in a breast cancer family. A later report, however, showed that the family also carried a nonsense mutation in exon 11 of another gene located at 17q21, *BRCA1*, which is assumed to be responsible for the development of disease (Friedman *et al.*, 1994b). The missense mutation in the adenovirus E1A enhancer binding protein was therefore in linkage disequilibrium with *BRCA*1.

In conclusion, the results from these preliminary investigations suggest that if mutations in *PHB* are involved in inherited forms of breast cancer they are rare and may be associated with a later age at onset of disease.

CHAPTER 4

DGGE ANALYSIS OF *PHB* FROM cDNA AND GENOMIC DNA

4.1 Introduction

The results from chapter 3 inferred that germline mutations in *PHB* may result in hereditary breast cancer which manifests at an older age and/or at a lower level of penetrance than that identified for mutations in tumour suppressor genes which are currently known to predispose to breast cancer. The detection of highly-penetrant, early-onset, breast cancer susceptibility genes has resulted from the selection criteria used in such studies; usually 3 or more affected first or second degree relatives with at least one being diagnosed before 50 years of age. In practice, however, families with more individuals affected at a young age have been found to be at greater risk of having a predisposing mutation.

To examine the hypothesis that germline mutations of *PHB* predispose to a later age at onset of disease and/or a lower degree of penetrance the coding sequence and the splice sites of *PHB* were screened by DGGE in a larger population of women with breast cancer who were selected using broader criteria. The mutational analysis of *PHB* from genomic DNA allowed for detection of mutations that would not be detected by screening the cDNA. Women with breast cancer were selected on the basis of having at least one first, second or third degree relative affected by breast cancer at any age. Individuals were ascertained on this basis to increase the likelihood of detecting mutations which result in a late age at onset of disease and a smaller number of affected individuals in a family.

4.2 Results

4.2.1 Fragment and Primer Design for DGGE Analysis of the *PHB* Genomic DNA sequence

The genomic clone sequence of the human *prohibitin* gene, as published by Sato *et al.* (1993), encompasses 3.5 kb of genomic DNA sequence and contains 7 exons coding for 272 amino acids. In comparison, the genomic organisation of the human *prohibitin* gene is illustrated as being over 8 kb in length (McClung *et al.*, 1995). Although this sequence has not been published it contains the same 7 exons and codes for the same 272 amino acids as reported by Sato *et al.* (1993). The difference in size appears to be explained by an increase in the length of the intron sequences. The two genomic DNA sequences are in full agreement for information concerning the exon-intron junctions (Altus *et al.*, 1995).

Following the same procedural steps as were performed for the cDNA sequence, the genomic DNA sequence was analysed for melting profiles and fragment design by MELT (Lerman and Silverstein, 1987). Four pairs of primers were designed to amplify products which contained exons 3, 4, 5 and 6 and their corresponding splice acceptor/donor sites. Exons 1 and 2 were not amplified from genomic DNA due to the lack of sequence information surrounding each exon as supplied from the human genomic clone sequence. A fragment was also designed to incorporate a region of exon 7 and its splice acceptor site. Due to the complex melting profile of exon 7 (as described in *Chapter 3*: section 3.2.1) two fragments were required in order to screen the full exon for sequence differences by DGGE. However, both primers for the second fragment are located in the exon sequence and due to the existence of processed pseudogenes of *PHB* not all of exon 7 was analysed from genomic DNA; this may have resulted in the amplification of pseudogene sequences as well as the functional *prohibitin* gene sequence which would

compromise detection of alterations of the *PHB* sequence. For each fragment only one primer required the attachment of a GC-clamp to its 5' end to ensure that the sequence of interest was contained solely in the lowest melting domain (Sheffield *et al.*, 1989).

4.2.2 DGGE analysis of Genomic DNA

Following generation of the products by PCR each fragment was subjected to electrophoresis on a perpendicular denaturing gradient gel (DGG) of 30 %-80 % for 3 hours. From this the optimal denaturant conditions for the 5 fragments were defined. The electrophoresis times for each fragment were then determined by performing time courses using parallel DGGE. Reduced electrophoretic mobility of the fragment in the DGG indicated that melting of the lowest melting domain had occurred. As at least one primer from each primer pair was designed from intron sequence, possible contamination by coamplification of a prohibitin pseudogene sequence was obviated; the processed pseudogenes do not possess introns. Details of DGGE conditions for each fragment amplified from genomic DNA are given in Table 4.1.

4.2.3 Detection of Nucleotide Changes in *PHB* in Women with Breast Cancer

In this study 72 women with breast cancer and at least one first, second or third degree relative with breast cancer were screened for sequence alterations in *PHB* (Table 4.2). This included the 35 women with a family history who were described in chapter 3. Using genomic DNA as a template, the sequence of exons 3, 4, 5, 6 and the first 53 bases of exon 7 along with their corresponding intron/exon borders, were investigated by DGGE in all 72 women. The cDNA sequence was also analysed by DGGE from the additional 37 women who were not part of the group screened as described in chapter 3. Two different single base substitutions were found in the *PHB* coding sequence in a total of 5 individuals. The first was the previously reported Arg43Leu variant in exon 3 which

Amplified Product Name	Product Size (bp)	Range of DGG Concentration (%)	Total DGG Electrophoresis Time (Hours)
Fragment 5	357	30 - 80	4
Fragment 6	349	30 - 80	7
Fragment 7	221	30 - 80	5
Fragment 8	310	30 - 80	5
Fragment 9	210	30 - 80	5

Table 4.1. Denaturing Gradient Gel (DGG) conditions for fragments amplified from *PHB* genomic DNA.

Cases of Breast (Br), Ovarian (Ov), Breast-Ovarian (BrO) and Bilateral Breast (BiBr) Cancer				Number of Families: Total = 72 (Family = first, second and third degree relatives)
Br	BiBr	Ov	BrO	
>4				2
2	1	2		1
4				5
3	1			1
3		1		1
3				14
2	1			5
1	2			1
2			1	2
2		1		1
2				36
1	1			2
1			1	1

Table 4.2. History of cancer in the 72 women with an affected relative.

was found in 4 individuals. The other was found in exon 5 in a single individual. Three previously unreported intronic polymorphisms, in a total of 11 individuals, were also discovered. Of these new polymorphisms 2 are in intron 6 and 1 is in intron 3. A single polymorphism in intron 5, previously described by Tokino *et al.* (1993) was detected by DGGE in either a heterozygote or homozygote state in 37 women.

The Arg43Leu allele was confirmed to be carried by 4 women by DGGE analysis, restriction enzyme digestion and sequence analysis. The average age at diagnosis of the 4 women is 49.3 years as compared to an average age of 44.3 years for those without the variant allele. The single alteration detected in Fragment 2 which corresponds to the PCR product for exon 5 was sequenced and found to be a G to C base substitution at the third position of codon 145 (Figure 4.1). No amino acid change resulted from the base change and therefore *PHB* Leu145Leu is a silent polymorphism. The Leu145Leu variant does, however, create a cleavage site for the restriction enzyme *Sac*1 and allows for a simple screening procedure of a large number of individuals. The individual carrying this germline mutation was diagnosed with breast cancer at 44 years of age. Her family history showed a maternal aunt who was diagnosed with breast cancer at 44 years of age.

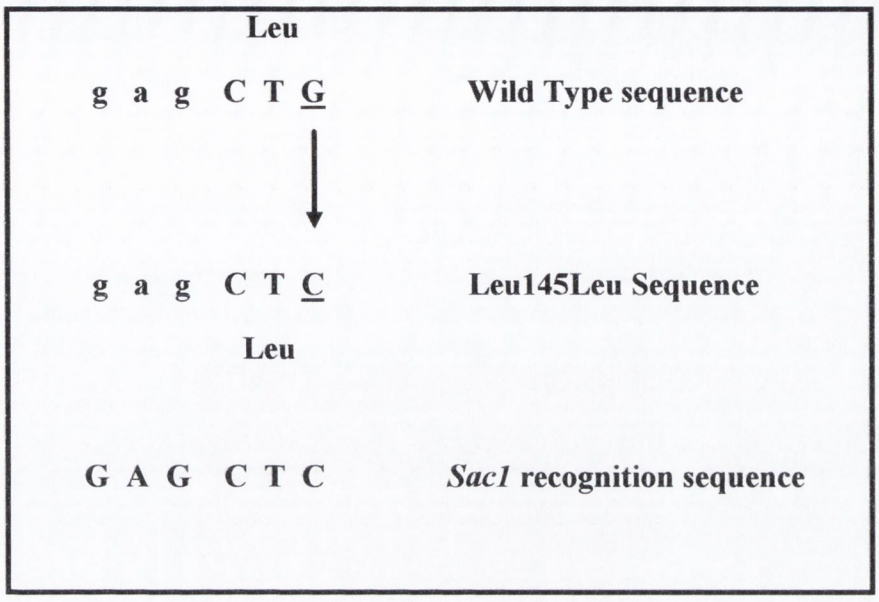


Figure 4.1. Creation of the *Sac1* restriction site by the Leu145Leu allele.

DGGE analysis of Fragments 5 and 8 detected 4 variant bands that were not present in the fragments amplified from their corresponding cDNA. Sequencing of PCR products revealed that the 4 polymorphisms were found in the intervening sequences, one in intron 3, one in intron 5 and two in intron 6. The polymorphism in intron 3, observed in 6 women, is located 127 bases downstream of the 3' end of exon 3 and substitutes a T for a C. The polymorphism in intron 5, which was detected in 37 women, is 12 bases upstream of the 5' end of exon 6 and is one of the polymorphisms originally described by Tokino *et al.* (1993). The polymorphism replaces a C with a G and produces a cut site for the restriction enzyme *Bsm*A1. Of these 37 women, 27 were heterozygous for the site whilst the other 10 were homozygotes. The intron 6 polymorphisms are located (1) 71 bases and (2) 118 bases downstream of the 3' end of exon 6 (Figure 4.2) and both are G to A transitions (Figure 4.3). Three women were heterozygous for both the polymorphism in intron 3 and the second polymorphism in intron 6. Table 4.3 summarises the findings for

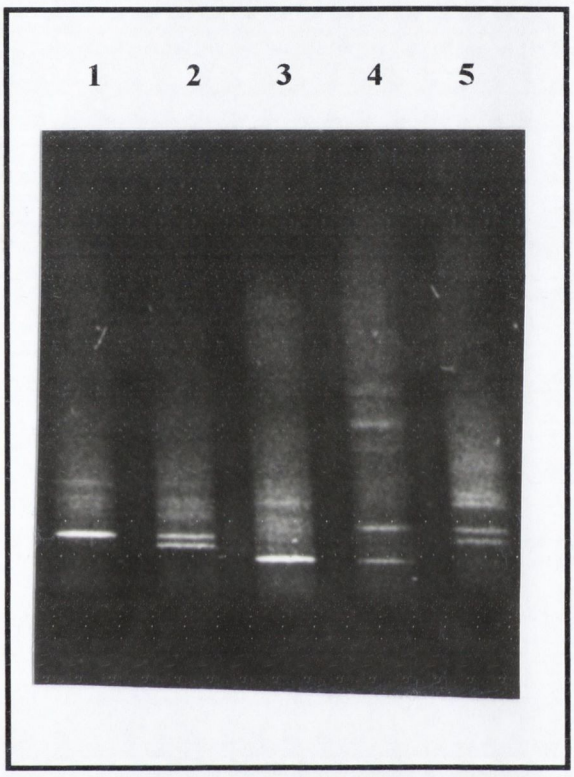


Figure 4.2. Parallel Denaturing Gradient Gel demonstrating intron 5 and intron 6 variants.

Lane 1: Homozygous wild type (case # D116)
Lane 2: Heterozygous wild type and intron 6 polymorphism (2) (case # D118)
Lane 3: Homozygous intron 5 polymorphism (*Bsm*A1$^{+/+}$)(case # D101)
Lane 4: Heterozygous intron 5 polymorphism (*Bsm*A1$^{-/+}$) and intron 6 polymorphism (1) (case # D008)
Lane 5: Heterozygous wild type and intron 6 polymorphism (1) (case # D162)

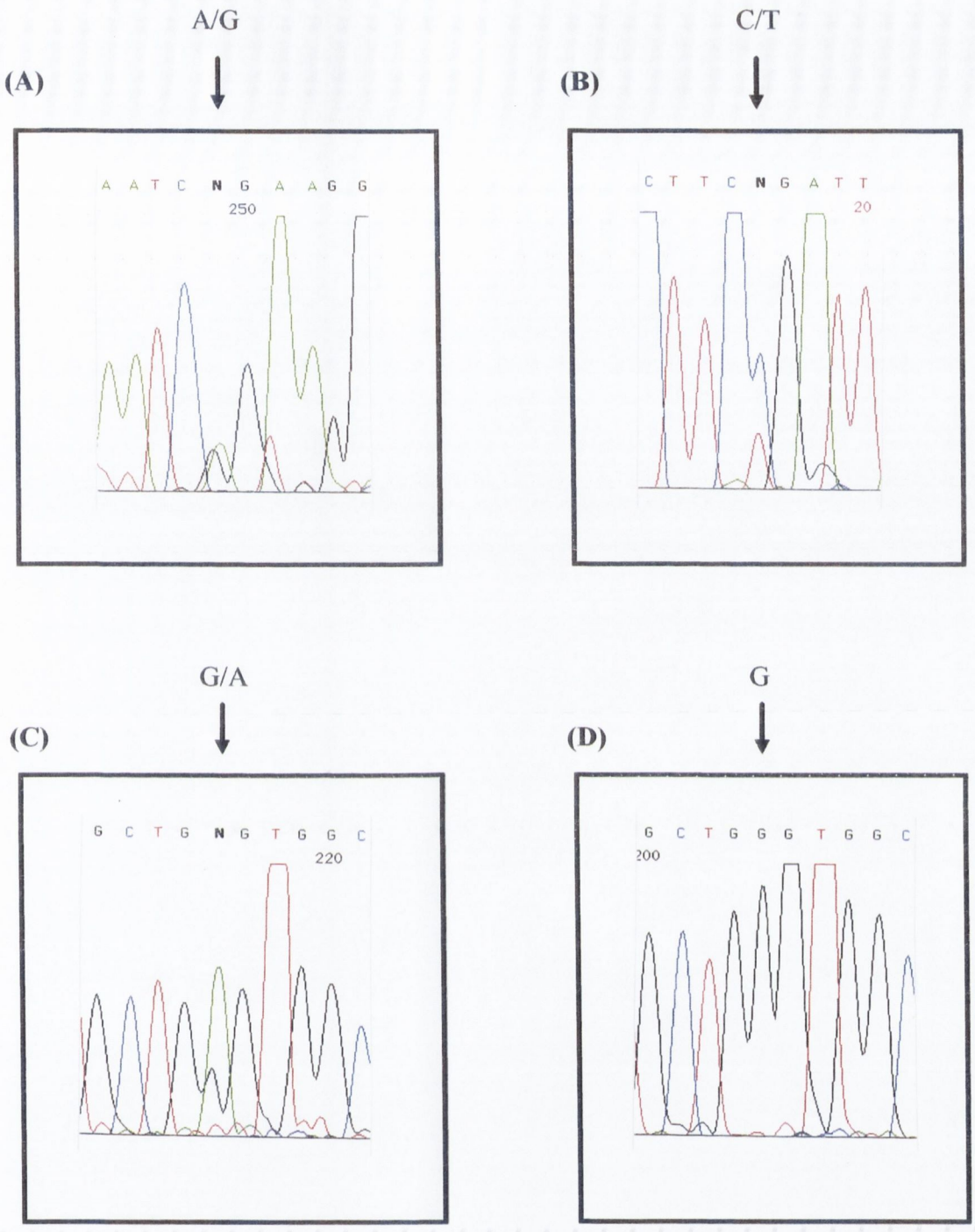


Figure 4.3. Electropherograms of intron 6 polymorphisms. Sequence analysis was performed with the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit with Amplitaq®DNA Polymerase, FS, and on the ABI 373. (A) Forward sequence analysis of heterozygous intron 6 polymorphism (2) (case # D118). (B) Reverse sequence analysis of heterozygous intron 6 polymorphism (2) (case # D118). (C) Forward sequence analysis of heterozygous intron 6 polymorphism (1) (case # D162). (D) Forward sequence analysis of homozygous wild type intron 6 (case # D116). The arrows indicate the position of the base difference on the sequence trace diagram and N represents A, T, G or C.

the intronic polymorphisms described above. None of these polymorphisms were found in association with the Arg43Leu allele.

4.2.4 Screening of Women with Breast Cancer for the Arg43Leu and Leu145Leu Alleles

To assess the frequency of the Arg43Leu and Leu145Leu alleles in the breast cancer population a total of 221 women diagnosed with breast cancer at any age were screened for the variant alleles. This was accomplished by restriction enzyme digestion of PCR generated sections of exon 3 with *Sty*1 for the Arg43Leu allele and of exon 5 with *Sac*1 for the Leu145Leu allele. The breast cancer population consisted of 141 women with no reported family history of breast cancer, 8 women with a family history of breast cancer but who had not given a blood sample at the time of selection for screening of *PHB* by DGGE, and the 72 women with a family history who had already been screened, as described above, for nucleotide changes in *PHB* by DGGE.

Of these 221 women a total of 6, including the 4 previously identified, carried the Arg43Leu allele. The two newly found cases were both women without a reported family history of breast cancer, 2/141 (1.4 %), and were diagnosed at 44 years of age and 47 years of age. This compares to 4/80 (5 %) of the women with a reported family history. No other carriers of the Leu145Leu polymorphism were detected in this population. A control group of 200 women with no reported family history of breast cancer were also screened for the two alleles. From this control population one 27 year old women was identified as having the Arg43Leu substitution and none had the Leu145Leu variant (Table 4.4).

Polymorphisms	Number of Cases
intron 3 polymorphism $^{+/-}$ and intron 6 polymorphism (2) $^{+/-}$	3
intron 3 polymorphism $^{+/-}$	3
intron 6 polymorphism (1) $^{+/-}$	2
intron 6 polymorphism (2) $^{+/-}$	3
*Bsm*A1$^{-/-}$	35
*Bsm*A1$^{+/-}$	27
*Bsm*A1$^{+/+}$	10

Table 4.3. Incidence of intron polymorphisms.

Allele Name	Age at Diagnosis of Breast Cancer	Family History of Breast Cancer
Arg43Leu	51	Yes
Arg43Leu	41	Yes
Arg43Leu	53	Yes
Arg43Leu	51	Yes
Arg43Leu	44	No
Arg43Leu	47	No
Leu145Leu	44	Yes
Arg43Leu	Unaffected at 27	No

Table 4.4. Summary of individuals identified with Arg43Leu or Leu145Leu.

4.2.5 Screening of Women with Breast Cancer for the *Eco*R1 Polymorphism

An *Eco*R1 polymorphism in *PHB* was originally discovered by White *et al.* (1991) by a Southern hybridization analysis of *Eco*R1 digested genomic DNA. It was not until Tokino *et al.* (1993) attempted to clone an amplified segment containing exons 2 and 3 that the *Eco*R1 polymorphism was found to be within intron 2. A PCR product of approximately 1.5 kb which contained intron 2 was generated for the 72 women with a family history of breast cancer who had been screened for mutations in *PHB* using DGGE (the primers used were EcoRF: 5' ATG GAG GGA CAG TGG GTC A 3' and EcoRR: 5' TCC CTA CCA CAA TGT CCT GC 3'). The product was then cut with the restriction enzyme *Eco*R1 and the presence or absence of the *Eco*R1 recognition site was determined (Figure 4.4). The *Eco*R1 site was found to be genetically distinct to the polymorphism in intron 5 which affects the target site for the restriction enzyme *Bsm*A1. Each chromosome examined possessed either the *Eco*R1 site or the *Bsm*A1 site. Therefore, the 10 women who were homozygous for the *Bsm*A1 site ($Bsm A1^{+/+}$) were homozygous for not having the *Eco*R1 site ($Eco R1^{-/-}$) and the 27 women who were heterozygous for the *Bsm*A1site ($Bsm A1^{-/+}$) were also heterozygous for the *Eco*R1 site ($Eco R1^{+/-}$). Table 4.5 details the various genotypes identified.

4.2.6 Segregation of the Arg43Leu Allele in Families 2-6

In Family 1 the Arg43Leu allele cosegregated with all the cases of cancer in the family (Figure 4.5). The pattern of allelic segregation was now investigated as far as possible in the other 5 families. Families 2, 3 and 4 contained at least 2 confirmed cases of female breast cancer whereas in Families 5 and 6 the only reported cases of breast cancer were the index cases themselves.

Family 2 constituted the largest family in this study with 11 members providing samples for DNA analysis and 11 members with a confirmed cancer (Figure 4.6). The

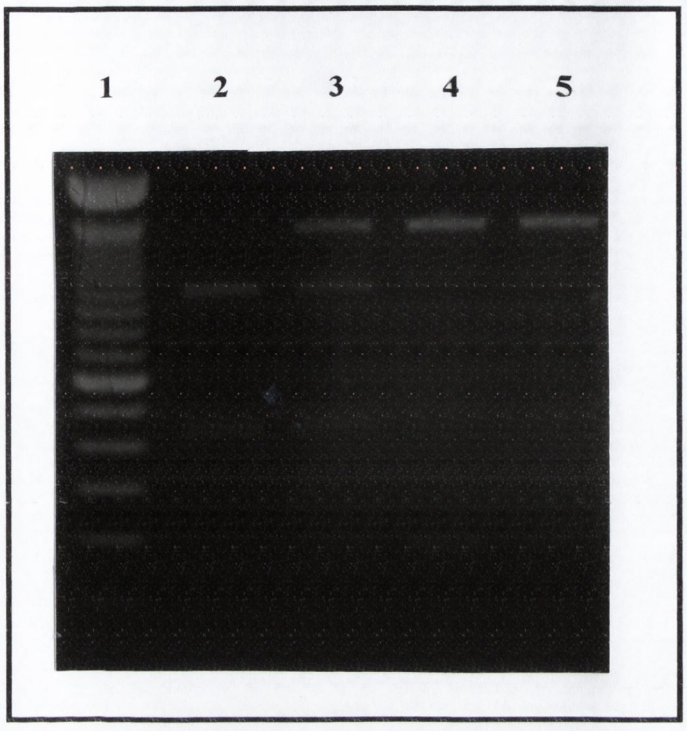


Figure 4.4. Restriction endonuclease, *Eco*R1, digestions of PCR products containing intron 2. Lane 1: 100 bp size marker; Lane 2: Individual homozygous for the *Eco*R1 recognition sequence (case # D009); Lane 3: Individual heterozygous for the *Eco*R1 recognition sequence (case # D017); Lanes 3-4: Individuals homozygous for the wild type sequence (case # D027 and case # D129).

Prohibitin Genotype	Number of Patients
*Eco*R1$^{+/-}$, *Bsm*A1$^{-/+}$	24
*Eco*R1$^{+/+}$, *Bsm*A1$^{-/-}$	26
*Eco*R1$^{-/-}$, *Bsm*A1$^{+/+}$	6
*Eco*R1$^{+/-}$, *Bsm*A1$^{-/+}$, *Sac*1$^{+/-}$	1
*Eco*R1$^{-/-}$, *Bsm*A1$^{+/+}$, *Sty*1$^{+/-}$	4
*Eco*R1$^{+/-}$, *Bsm*A1$^{-/+}$, intron 6 polymorphism (1)$^{+/-}$	1
*Eco*R1$^{+/+}$, *Bsm*A1$^{-/-}$, intron 6 polymorphism (1)$^{+/-}$	1
*Eco*R1$^{+/+}$, *Bsm*A1$^{-/-}$, intron 6 polymorphism (2)$^{+/-}$	3
*Eco*R1$^{+/+}$, *Bsm*A1$^{-/-}$, intron 6 polymorphism (2)$^{+/-}$, intron 3 $^{+/-}$	3
*Eco*R1$^{+/+}$, *Bsm*A1$^{-/-}$, intron 3 $^{+/-}$	2
*Eco*R1$^{+/-}$, *Bsm*A1$^{-/+}$, intron 3 $^{+/-}$	1
Total	72

Table 4.5. Summary of *PHB* genotypes.

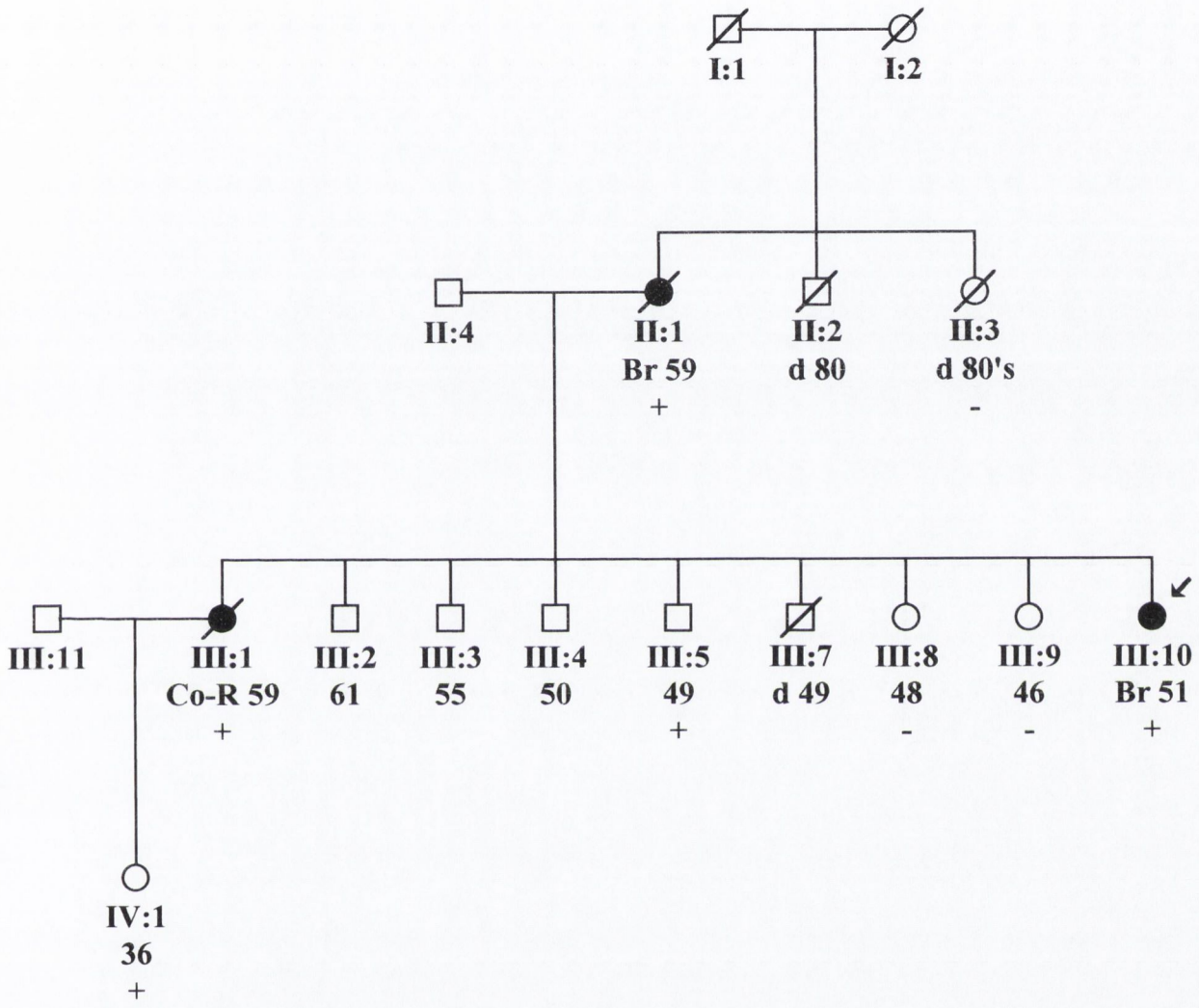


Figure 4.5. Pedigree of Family 1. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer.

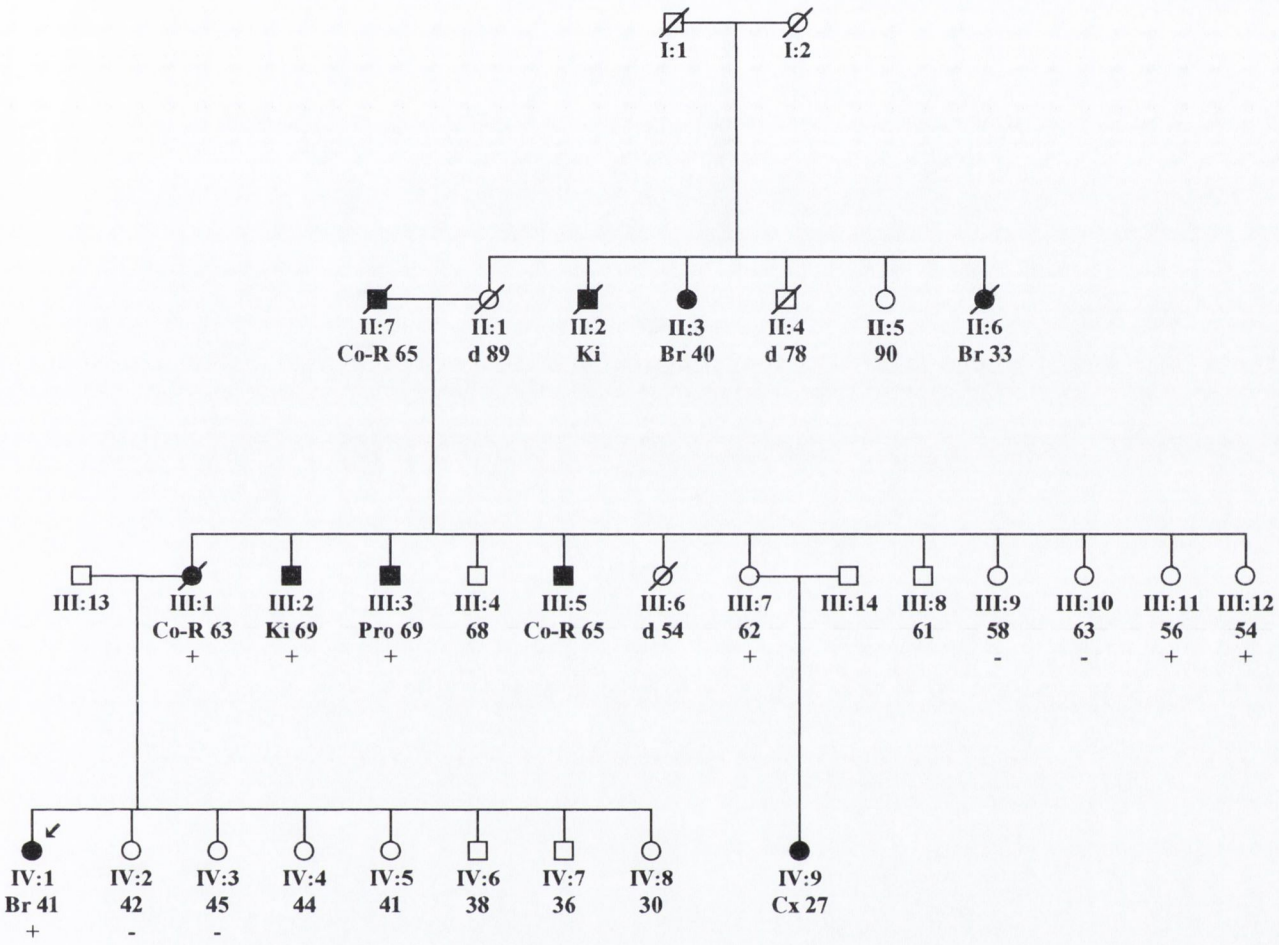


Figure 4.6. Pedigree of Family 2. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer, Ki = renal cancer, Pro = prostate cancer, Cx = cervical cancer.

index case in this family was the youngest individual with the Arg43Leu allele and cancer, of any type, in all 6 families. Two sisters of the index case were unaffected at 42 and 45 years of age and did not carry the Arg43Leu variant. Individuals that had the Arg43Leu allele include the proband's mother, who had a liver biopsy suggesting a primary stomach or colon cancer at 63 years of age although cholangiocarcinoma could not be entirely ruled out. The maternal uncles III:2 and III:3 with renal cell carcinoma at 69 years of age and prostate cancer at age 69 respectively were also carriers of the Arg43Leu variant. A blood or tissue sample was not available for DNA analysis from the maternal uncle, III:5, who was diagnosed at 65 years of age with bowel cancer. Five unaffected maternal aunts were screened for Arg43Leu; 3 were carriers (III:7, 11, 12) and 2 were not (III:9, 10). A daughter, IV:9, of the maternal aunt, III:7, had cervical cancer at age 27 but no blood or tissue could be procured for examination. Interestingly, from analysis of their genomic DNA, the maternal uncle, III:2, and the maternal aunt, III:12, were found to carry only the Arg43Leu variant and were therefore homozygous or hemizygous for the allele. Other confirmed cancer cases included the maternal grandfather, bowel at age 65, and the maternal grandmother's sisters who had breast cancer at 33 and 40 years of age respectively. Both the maternal grandmother and grandfather could be carriers of the Arg43Leu allele. The maternal grandmother's brother was reported to have had a renal cell carcinoma but the age at diagnosis was not known. Neither of the grandparents or their siblings could be assessed for the Arg43Leu variant as they all lived in the U.S.A. and personal contact could not be made with them.

For Family 3 only 3 siblings of the proband were available for testing for the presence of Arg43Leu (Figure 4.7). Formalin-Fixed Paraffin-Embedded (FFPE) tissue from the sister with breast cancer, diagnosed at 33 years of age, was obtained (U.C.H., London, U. K.) for extraction of genomic DNA and restriction enzyme analysis. The family history was otherwise unremarkable. Blood samples were given by the other sister

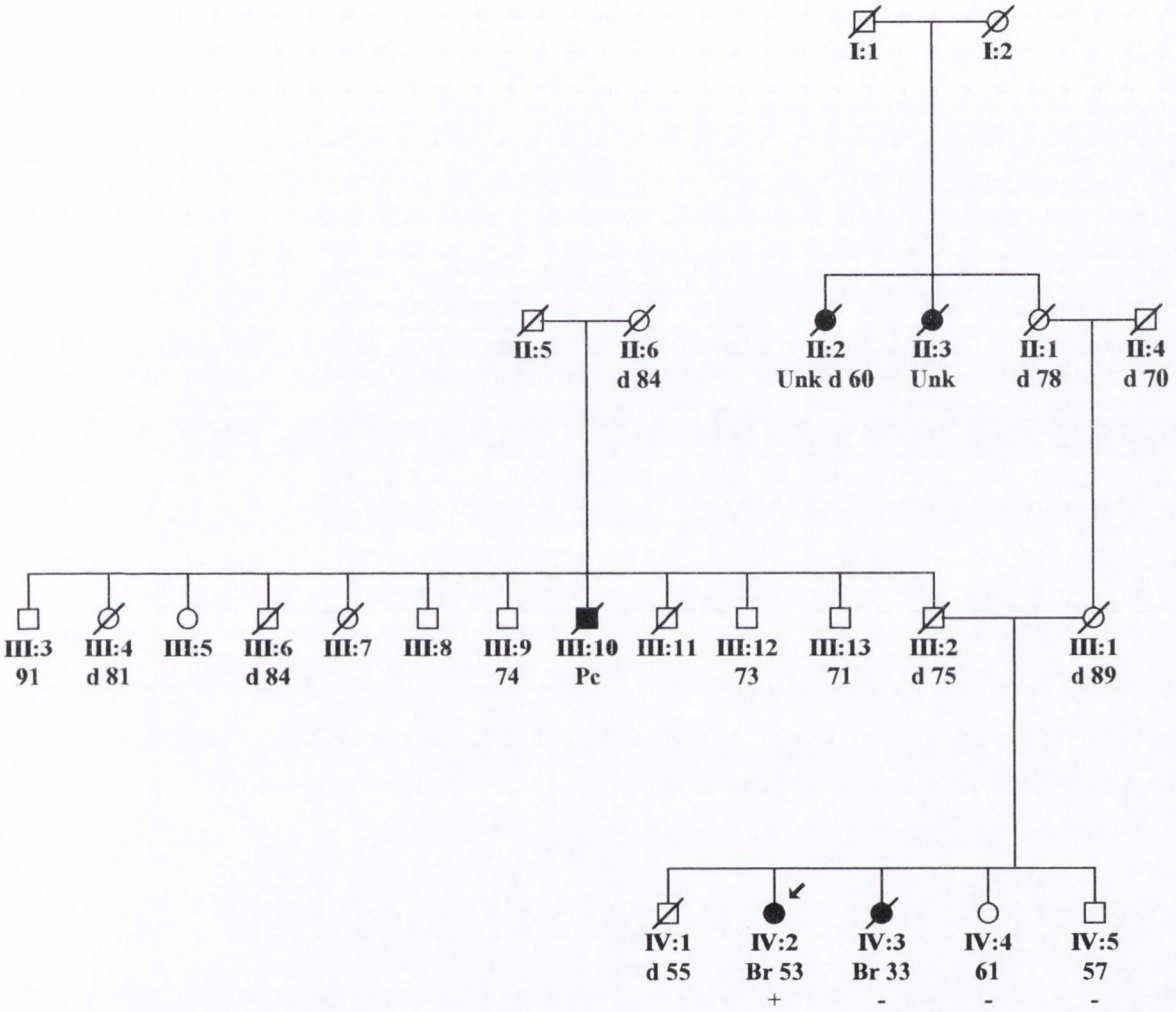


Figure 4.7. Pedigree of Family 3. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Pc = pancreatic cancer, Unk = cancer of unknown origin.

and brother who were both unaffected. None of the siblings carried the sequence variant, therefore it could not be determined if the sequence variant was inherited or *de novo*.

In Family 4 nine individuals, including the proband, were screened for the missense mutation Arg43Leu (Figure 4.8). Of these nine, 2 of the 3 with cancer were positive for the Arg43Leu variant, 4 of the 5 unaffected individuals were positive for the Arg43Leu variant and 3 did not carry the sequence variant. One paternal uncle, III:5, with the Arg43Leu allele had an adenocarcinoma of the large bowel at the age of 55. The father of the index case is an obligate carrier but was unaffected by cancer at the time of his death from a cerebral-vascular accident at 58 years of age.

Four other unaffected carriers were identified, 2 brothers of the proband, aged 53 and 54, 1 sister, aged 45, and a paternal aunt aged 61. The two other cases of breast cancer in Family 4 are a paternal aunt, III:4, diagnosed at 45 and her daughter, IV:3, diagnosed at 52. The daughter, IV:3, did not have the Arg43Leu variant, nor did her brother but determination of the paternal aunt's status could not be ascertained as no tissue sample could be traced due to the date of surgery (early 1970's) and closure of the hospital in which the operation was performed. Both paternal grandparents had cancer; the grandmother had cancer of the rectum at 63 years of age and the grandfather had bowel cancer at 70 years of age but neither could be tested for the Arg43Leu variant. Several cases of other cancers, including cancer of the cervix and bowel, were reported in the siblings of the paternal grandmother but ascertainment by hospital records or confirmation by personal contact could not be achieved.

For Family 5 the index case had 2 primary cancers, breast cancer diagnosed at the age of 44 and AML-M4Eo (Acute Myelomonocytic Leukaemia with eosinophilia) at 45 years of age (Figure 4.9). Only one unaffected sister could be tested for the allele and she proved to be negative. The one other cancer reported for this family was a brain tumour in

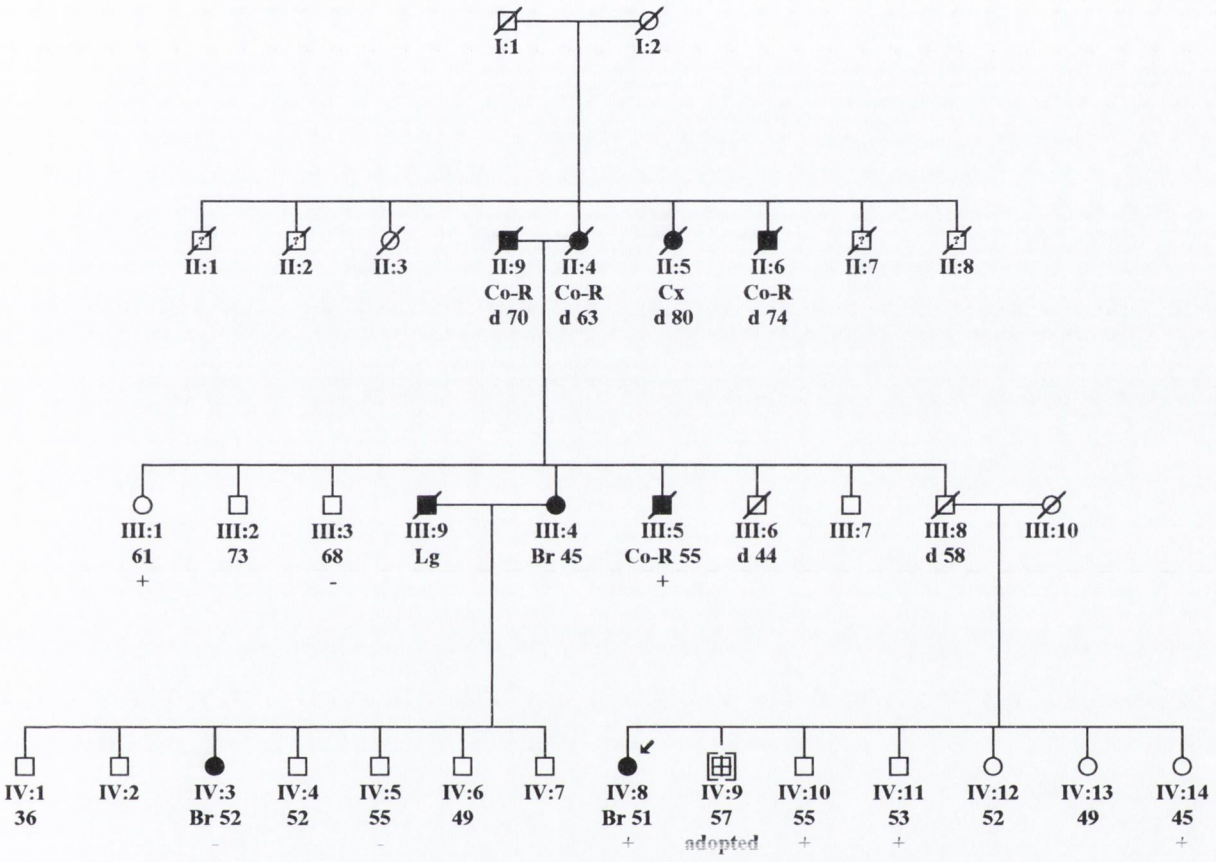


Figure 4.8. Pedigree of Family 4. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer, Lg = lung cancer, Cx = cervical cancer, ? = possibly affected.

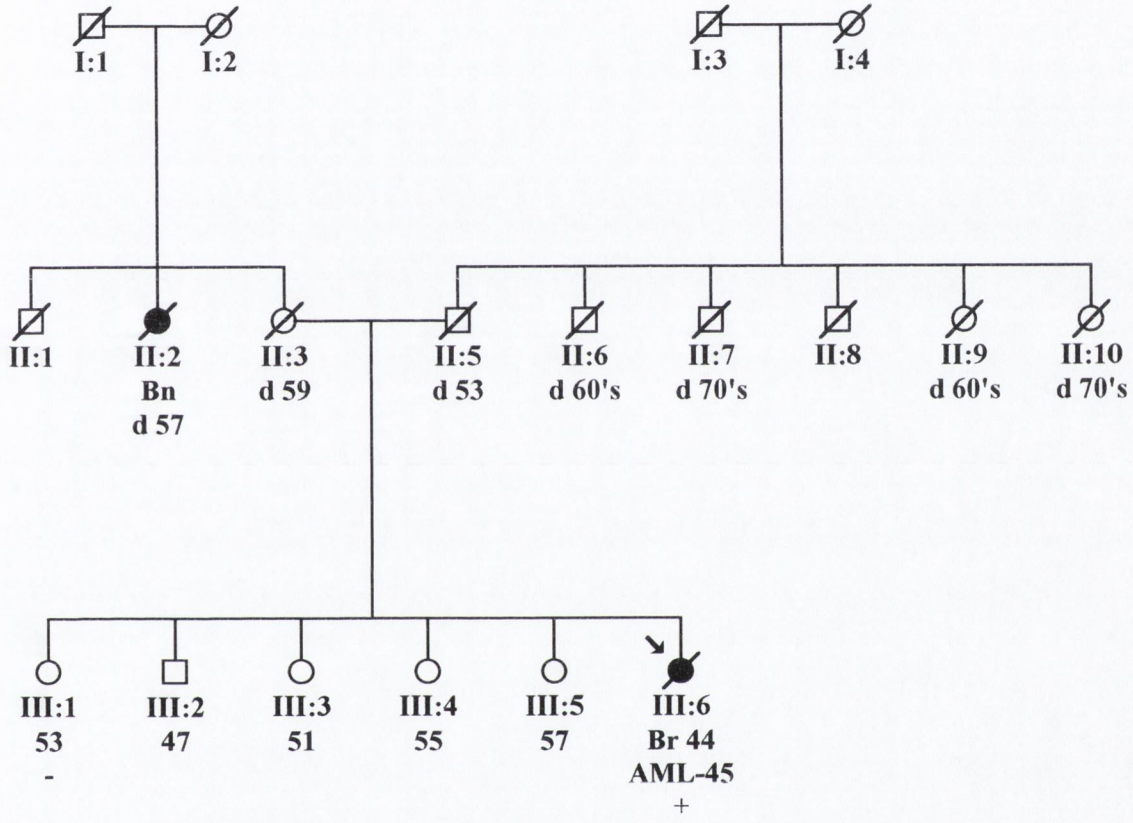


Figure 4.9. Pedigree of Family 5. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, AML = acute myelomonocytic leukaemia, Bn = brain tumour.

the paternal aunt at 57 years of age but no death certificate could be ascertained to verify this report.

In Family 6 the proband did not have any siblings and was of Eastern European descent and now living in Ireland. As a consequence the only accurate information regarding family history is limited to that of her parents. Her father had stomach cancer at 69 years of age and her mother was unaffected. Neither could be tested for the Arg43Leu allele.

4.2.7 Statistical Analyses

The data analysis looked at the difference in mutation level for the Arg43Leu allele in cases (6/221) and controls (1/200). The small numbers of individuals in each group meant that exact permutation tests were applied. Using Barnard's unconditional test for the difference of two binomial proportions the exact test is significant at the 7 % level. There is also a first order of magnitude difference between the proportion of the Arg43Leu variants in cases (0.027) and that in the controls (0.005).

In the six families identified, apart from the 6 index cases, 3 women with breast cancer were screened for the Arg43Leu allele but only 1 of these 3 women was a carrier of the variant allele. Consequently, within these families, a significant association between the Arg43Leu variant and breast cancer was not found. Given the low incidence of breast cancer but considerable number of all cancer types in the families, the likelihood of family members inheriting the Arg43Leu allele and having cancer of any type was analysed. The probands were excluded from this analysis on the basis that they were not independent observations because they had been selected for cancer and the Arg43Leu allele. In Families 1, 2 and 4 the inheritance of the Arg43Leu allele had been established by detection of the allele in a parent or aunt or uncle of the proband. In Families 3, 5 and 6 the variant allele was only found to be carried by the proband and the possibility that the

variant had arisen *de novo* could not be excluded. Therefore, in the families where inheritance of the variant allele had been established, the significance of the association was examined for seven cases, six of which were carriers.

These seven cases are: Family 1 members II:1 and III:1; Family 2 members III:1, III:2 and III:3 and Family 4 members III:4 and III:5. In Family 4, the aunt (III:4) of the proband has been shown by haplotype analysis of her daughter (IV:3) (*Chapter 6*: section 6.2) not to have inherited the Arg43Leu allele. Therefore her daughter (IV:3), who also had breast cancer, was not included in the analysis as she could not have inherited the Arg43Leu variant from her mother.

Based on the *a priori* hypothesis that the Arg43Leu allele is associated with cancer a one-sided exact test can be performed. The probability of each one of the seven cases possessing an Arg43Leu allele is calculated based on the genotype, if known, of the closest relative through whom the proband relationship is mediated. For the individuals in Families 1 and 2 the probability of having the Arg43Leu allele, (+), is 0.5 and for Family 4 this probability is 0.25. The probability of not having the Arg43Leu allele, (-), is then 1 minus 0.5 or 1 minus 0.25. For each of the seven individuals there are 2 possible outcomes, either (+) or (-), which means that there are a total of 128 (2^7) possible outcomes. Six of the seven individuals are (+) and one is (-), therefore only 8 outcomes are as significant or more significant than the observed. These outcomes can be represented as +++++++, -++++++, +-+++++ etc for the seven individuals. The probability for each of these outcomes is calculated by multiplying the individual probabilities together ie. for +-+++++ the probability is 0.5 x 0.5 x 0.5 x 0.5 x 0.5 x 0.25 x 0.25 = 0.001953. The sum of the probabilities for all 8 outcomes is 0.0234. This is the probability, p = 0.0234, that an association as large as that found between cancer and the Arg43Leu variant has occurred by chance.

4.3 Discussion

In chapter 3 an exhaustive analysis of the *PHB* coding sequence was undertaken. However, other mutations can occur outside of the coding sequence that affect the protein product. Approximately 6 % of mutations identified in *BRCA1* and *APC* are splice site mutations (Harris, 1996). Therefore the initial study was expanded to examine both the intron/exon borders as well as the full cDNA sequence of *PHB* for base changes. Women with breast cancer diagnosed at any age and whose family history included a first, second or third degree relative with breast cancer were included in addition to those chosen as having a family history of breast cancer in the initial study. Using a wider set of selection criteria increased the probability of including women who may be predisposed to breast cancer through the inheritance of a disease-causing allele of lower penetrance.

This comprehensive mutational study resulted in the discovery of three more women with the Arg43Leu variant in exon 3 and one woman with the silent polymorphism Leu145Leu in exon 5. Three new polymorphisms, one in intron 3 and two in intron 6, were also discovered. Further analysis of sporadic breast cancer cases revealed two additional women with the Arg43Leu variant. The analysis of *PHB* from genomic DNA did not include exons 1 and 2 and their exon-intron junctions, although exon 2 was analysed from cDNA. Exon 1 is only 23 nucleotides in length and is transcribed but is not translated. To exclude the possibility that functionally significant sequence alterations are present in these regions the sequence surrounding exons 1 and 2 would need to be determined to allow for analysis by DGGE.

Examination of the families of the 6 index cases found to carry the Arg43Leu variant identified a total of 36 people who could be tested for the Arg43Leu allele. In one "family", Family 6, only the proband was screened. DNA was available from 14 individuals with cancer in the 6 families and 12 of these were found to carry the Arg43Leu variant. If the index cases are removed, since they were selected on the basis of having

Arg43Leu and cancer and are therefore not independent observations, then Arg43Leu segregates with 6 out of 8 cancers. The average age at diagnosis of breast cancer for the 7 women in the families is 49.6 years and for the other cancers is 63 years. In Families 1 and 2 the Arg43Leu variant cosegregates with every affected person analysed. In Families 3, 5 and 6 only the index case has the Arg43Leu allele and as a result it cannot be determined whether the variant occurred *de novo* or was inherited in these families.

Examination of the difference between the mutation level of the Arg43Leu variant in women with breast cancer (6/221) and the control group (1/200) reveals a p-value of $p = 0.07$. Although this does not reach the recognised statistically significant level of $p = 0.05$, it does indicate a possible association between the Arg43Leu variant and breast cancer. The control individual that was identified as carrying the Arg43Leu allele is 27 years of age and, given our results, is unlikely to have developed a cancer at this age even if the Arg43Leu variant does confer susceptibility to cancer. Selection of the controls required that there was no reported family history of breast cancer in first or second degree relatives.

Statistical analysis of the three families, in which the Arg43Leu variant has been shown to be inherited, was carried out for the affected relatives of the probands. This revealed a statistically significant association for the Arg43Leu allele with cancer ($p = 0.0234$). This level of significance shows that the probability of the observed association of Arg43Leu with cancer occurring by chance is remote. Interpretation of this association must be viewed in the context that analyses could not be performed with the inclusion of all the families and that the significance of the findings is based on a small number of individuals with any type of cancer. Of the seven individuals included in this analysis only one had breast cancer and the Arg43Leu variant.

Five of the family members from the six families with Arg43Leu had cancers at sites apart from the breast and 2 more were possible carriers of the Arg43Leu variant. No

mutations (somatic or germline), to date, have been identified in *PHB* from a variety of other tumours. No mutations were identified in tumours of the ovary, liver and lung, with over 30 samples of each primary malignancy being examined (Sato *et al.*, 1993). Mutations were not detected in exons 4 and 5 of 20 epithelial ovarian tumours selected for LOH of the *PHB* chromosomal region (Cliby *et al.*, 1993) and no mutations in 7 rat bladder tumours and 10 rat bladder cell lines were found (Asamoto *et al.*, 1994). The other cancers in the family members with Arg43Leu or who are obligate carriers of the sequence variant include 5 cases of colorectal cancer, 1 renal cell carcinoma, 1 prostate cancer, and 1 AML-M4Eo (Acute Myelomonocytic Leukaemia with Eosinophilia). If the range of cancers in these families is caused by the Arg43Leu variant in *PHB*, it would appear to indicate that *PHB* represents a tumour suppressor gene associated with a broad spectrum of cancers.

The *PHB* gene in cultured human cell lines has been scrutinised for mutations in all exons using single stranded conformational polymorphism analyses (Jupe *et al.*, 1996b). Using this approach only the *Bsm*A1 restriction enzyme site polymorphism in intron 5 was detected. The authors also reported that the human immortalised cells classified as belonging to complementation group B but not those of the other complementation groups, A, C and D, are sensitive to the antiproliferative activity of wild type prohibitin mRNA transcripts (Pereira-Smith and Smith, 1988; McClung *et al.*, 1995; Jupe *et al.*, 1995). The group B cells examined were homozygous for the *Eco*R1 site (*Eco*R1$^{+/+}$), discussed in section 4.2.5 (named the B allele by Jupe *et al.*, 1996a), but all other cell lines examined in complementation groups A, C and D were *Eco*R1 heterozygotes or were homozygous for the absence of the *Eco*R1 site (named the non-B allele by Jupe *et al.*, 1996a). More extensive analysis of 22 human breast cancer cell lines from the B complementation group found that 17/22 (77.3 %) were homozygous for the B allele and 5/22 (22.7 %) were homozygous for the non-B allele (Jupe *et al.*, 1996b). Most notably,

however, none were heterozygous. This is in contrast to the unpublished results for normal cell lines where about 50 % were reported to be heterozygous (Jupe *et al.*, 1996b). This led the authors to suggest that loss of heterozygosity may explain to a degree why some of the cell lines were not heterozygous for the *Eco*R1 polymorphism.

Seventeen out of the 22 human breast cancer cell lines were homozygous for the B allele and most of the cell lines were created from tumours in women diagnosed at ages older than 50 years of age. Therefore the B allele may have an association with an allele that causes late age at onset of disease. From our results 35 (48.6 %) of the 72 women screened by DGGE are homozygous for the B allele, twenty seven (37.5 %) are heterozygous for the B allele, and 10 (13.9 %) are homozygous for the non-B allele. In the 72 women screened in this study no association was found between age at diagnosis and the B allele. For the Arg43Leu allele, however, 5 out of 6 of the index cases are homozygous for the non-B allele, whilst the sixth is a heterozygote.

Terashima *et al.* (1994) described 3 proteins of 32, 37 and 41 kDa, that are intracellularly associated with the mIgM antigen receptor complex of B cells. The 32 kDa protein was identified as prohibitin and the 37 kDa protein as being very closely related to prohibitin on the basis of its amino acid sequence. The authors stated that due to the high sequence similarity and the high degree of conservation of the central region of the protein that both have a similar function and belong to a conserved gene family. Both prohibitin and BAP37 were estimated to be present at an equimolar ratio with mIgM which led the authors to suggest that the two proteins might form a dimer. Coates *et al.* (1997) identified a series of repeated leucine residues and a predicted coiled coil region in prohibitin and BAP37 (Figure 4.10) corresponding to residues 55 to 218 of the human prohibitin protein sequence. Of the 14 leucine residues present in this portion of the human prohibitin sequence 11 are conserved in the human BAP37 protein. Nine are conserved in yeast, *Drosophila*, mammals and *Arabidopsis* in prohibitin and BAP37. On two occasions

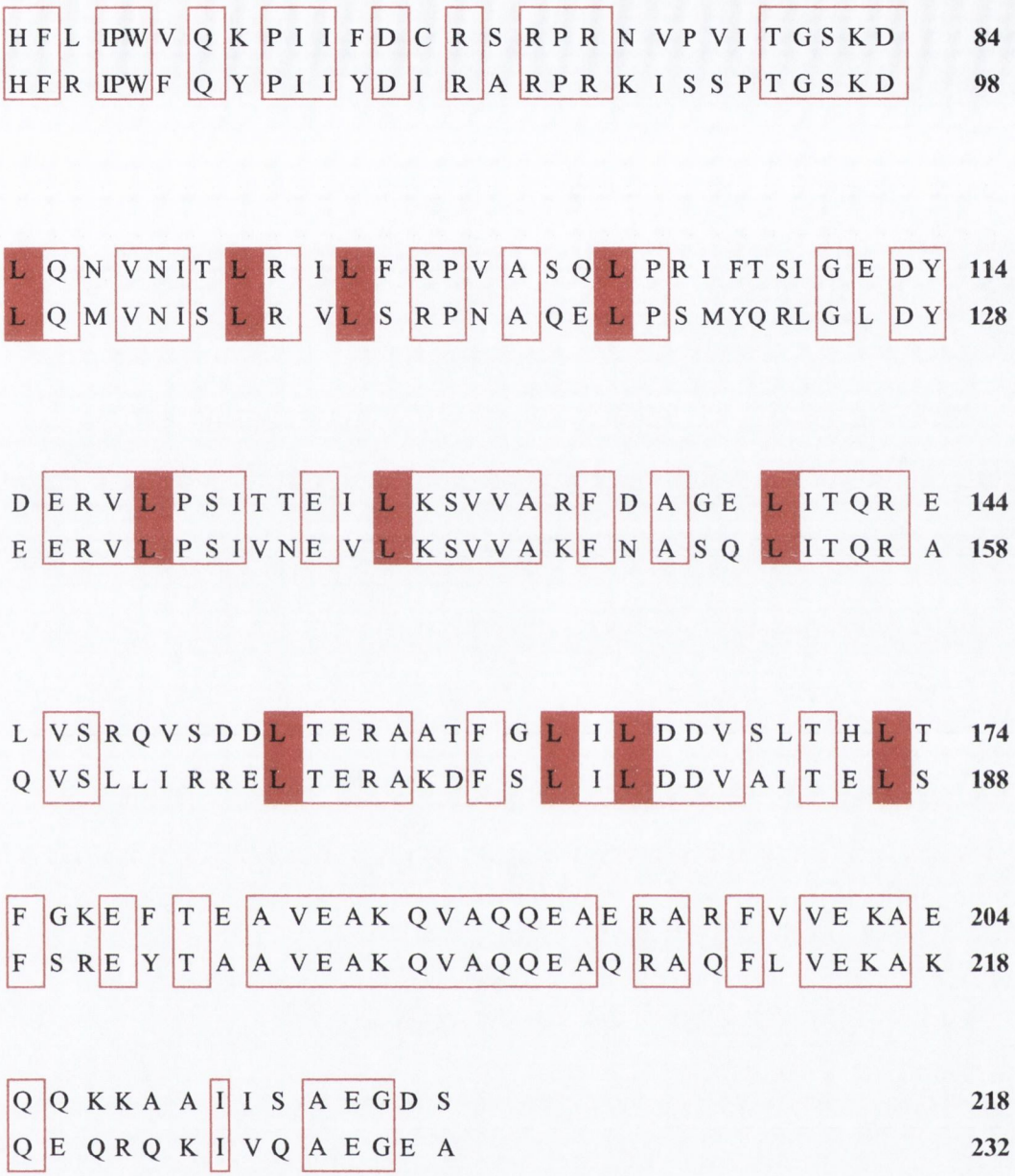


Figure 4.10. Amino acid sequence homology of the central region of prohibitin and BAP37. Upper strand is the prohibitin sequence. The human prohibitin sequence is from residues 55 to 218 and the human BAP37 sequence is from residues 69 to 232. Identical amino acids are boxed and the conserved leucine residues are high-lighted in red (adapted from Coates *et al.*, 1997).

the leucine residues are replaced conservatively by isoleucine or methionine in the corresponding *PHB* and *BAP*37 genes. At the other 3 positions non-conservative changes are only found in the BAP37 sequences. These 3 positions may represent amino acids of differing functional or structural significance between the two proteins whilst the other leucine residues correspond to regions of functional identity. Therefore it is quite striking that the variation Arg43Leu involves the creation of a leucine residue. The Arg43Leu variant lies outside the aforementioned region of prohibitin but may cause a premature structural interaction with the remaining leucine residues. The silent Leu145Leu polymorphism would seem to be an unlikely candidate to cause a conformational change. One possible mode of interaction would be the incorporation of a less abundant tRNA, which may result in a slower translation of product. The effect of this would not become apparent until loss of the wild type allele occurred.

The functional impact of a missense mutation is often due to alteration of a site or structure that is involved with contact or interaction with another protein or DNA sequence. The p53 protein has four main functional domains and the sequence specific DNA binding domain is most frequently mutated in cancer cells. Nearly 80 % of p53 mutations are missense with 40 % of those being found at 6 residues in this DNA-binding domain. The identification of the type of association that prohibitin has with BAP37 could reveal the critical structural areas of both proteins and aid in determining how they function. The identification of two individuals in Family 2 with only the Arg43Leu allele, one diagnosed with cancer of the prostate at 69 years of age and the other an unaffected female at 54 years of age, suggests that the altered prohibitin protein can function adequately under normal cirumstances. In yeast prohibitin is known to interact with other proteins in different systems, including Ras and mitochondrial proteins, and may have similar properties in the human. Thus prohibitin may require the alteration of several other genes or proteins before an effect of the Arg43Leu variant is realised.

CHAPTER 5

STUDY OF THE ARG43LEU ALLELE IN COLON CANCER

5.1 Introduction

In chapter 4 a number of individuals with bowel cancer were identified as carrying the Arg43Leu allele and in the tumour tissue from one case LOH analysis had shown that the wild type allele had been lost. This prompted an investigation into the frequency of the variant in a population with cancer of the colon. The presence of mutations in *PHB* in colon cancer had not been examined before. The chromosomal region of 17q21 has been analysed in colorectal tumours by Leggett *et al.* (1995). They demonstrated by allelic loss on chromosome 17q that the smallest region of overlap of deletions, not including *p53* on 17p, contained *prohibitin*.

The consent of all the individuals involved in the breast cancer studies had been obtained by interview and required a mouth swab or blood sample for DNA analysis. In order to gather a large number of people with colon cancer and collect a sample for extraction of DNA, a different approach was used. In this study, sections from formalin-fixed paraffin-embedded tissue blocks which corresponded to non-malignant tissue were used to extract DNA. In order to assure anonymity, random codes were assigned to the cut sections so that any section could not be traced directly to the person it had originated from. This study was sanctioned by St James' Hospital Research Ethics Committee.

5.2.1 Selection of Patients with Colon Cancer

Patients attending St. James' Hospital and who were newly diagnosed with colon cancer in 1994 or 1995 were selected for inclusion in this study. The selection of patients was not biased for gender, age at diagnosis, type of tumour or any other criteria. In order to maintain anonymity between the laboratory and the patient, Dr. Maeve Doyle compiled information, where available from the medical records, pertaining to 170 patients. The final number of individuals included in the study was reduced to 119 due to practical and technical concerns.

A list of the pathology numbers for the patients was drawn up and the specimens relating to the normal tissue, as determined by histopathology reports, were selected. Tissue samples were dissected and each was assigned a random three digit code. A copy of the random codes and the histopathology numbers was kept separate from the patients' personal details. By this method the anonymity of the individual was strictly maintained.

5.2.2 Individual Results

Genomic DNA was extracted from the dissected tissue samples and exon 3 of the *prohibitin* gene was amplified by PCR (the primers used were 3A: 5' TTA TGA CAT GTC ATT GTC CT 3' and REV3: 5' AGT GAT GAC TGG CAC ATT AC 3', generating a product of 182 bp). The amplified products of these PCRs were analysed for the presence of the Arg43Leu allele by incubation with the restriction enzyme *Sty*1. Details from 170 patients were originally collected but due to the number of cases where either only malignant tissue from the biopsy was available or tissue blocks were not available or the extracted DNA was of poor quality a total of 119 patients were screened for the Arg43Leu variant. From this screening one individual was found to carry the Arg43Leu allele. This patient was identified as a man diagnosed at 78 years of age with an adenocarcinoma of the ascending colon. The medical chart recorded that he had no personal medical history of note and that there was inadequate information to determine the history of cancer in family members.

The personal history of cancer for the 119 individuals was documented and is listed in Table 5.1. Eleven reported a previous malignancy and 2 reported 2 previous malignancies. Three of those with a previous cancer also had a family history of cancer.

Site of Cancer	Number of Incidences
Basal Cell Carcinoma	2
Basal Cell Carcinoma and Breast	1
Bladder	1
Breast	1
Breast and Lung	1
Chronic Lymphocytic Leukaemia	1
Oesophageal	1
Prostate	1
Squamous Cell Carcinoma	1
Testicular	1

Table 5.1. Personal history of cancer for individuals with colon cancer, as given in their medical records.

5.3 Discussion

This study shows that the Arg43Leu variant occurred in only one person from the group of 119 individuals examined. The result (1/119) represents less than 1 % of the colon cancers screened and is over 3-fold less common than the Arg43Leu allele in the breast cancer population screened (6/221). However, in those families analysed for segregation of Arg43Leu, breast cancer occurred over 3 times more often than colon cancer in carriers of the variant allele. In this regard, the incidence of the Arg43Leu allele in the population with colon cancer is consistent with that derived from the families shown to have inherited the sequence variant. To establish the significance of these findings much larger numbers, outside the scope of this work, would be required. With the small number of cases these results are difficult to interpret, however, they may suggest that the Arg43Leu variant of *PHB* is not a major contributor to the number of colon cancer cases but they do not exclude a causative association.

The selection of individuals with colon cancer for screening of the Arg43Leu allele included affected men and women who had not been selected for family history of disease or age at diagnosis. The National Cancer Registry of Ireland has published its first figures for the incidence and type of cancer in the Irish population for 1994. In 1994, 559 females and 592 males in Ireland were diagnosed with colon cancer. This represented 6.15 % of all cancers affecting females and 6.27 % of all cancers affecting males. For both sexes colon cancer was the fourth most common type of cancer after skin, breast and cancer of the cervix in females and after skin, prostate and lung cancer in males. In developed countries colorectal cancer is the second and third highest cause of death from malignancy in males and females respectively. Therefore, the discovery of further genes which predispose to the development of colon cancer could be of benefit mainly towards a greater understanding of the disease.

For common cancers, such as colorectal cancer, first degree relatives of patients are at an increased risk for cancer at the same site (St. John *et al.*, 1993). Several genes have already been identified which contribute to the increased risk of colorectal cancer in these families. These genes are predominantly associated with one of two syndromes, Familial Adenomatous Polyposis (FAP) and Hereditary Non-Polyposis Colorectal Cancer (HNPCC). The gene for FAP was linked to chromosomal region 5q, then cloned and termed the *APC*, adenomatous polyposis coli, gene (Joslyn *et al.*, 1991). HNPCC has recently been found to be the result of mutation in any one of a number of mismatch repair (MMR) genes on chromosomes 2, 3 or 7 (Fishel *et al.*, 1993; Leach *et al.*, 1993; Bronner *et al.*, 1994; Papadopoulos *et al.*, 1994; Palombo *et al.*, 1994). The majority of germline alterations are found in the *hMSH2* and *hMLH1* genes and only 3 germline mutations have been found in *hPMS1* and *hPMS2* (Liu *et al.*, 1996). Both syndromes are represented by a near-complete penetrance of the phenotype if the allele is inherited but account for only 5-10 % of all colorectal cancers. The existence of lower penetrance genes may be responsible for a larger number of cases and small familial clusterings of colorectal cancer.

Over 90 % of colorectal cancers are diagnosed over the age of 50 years (Peltomaki *et al.*, 1993; Kim *et al.*, 1994). The genes currently known to predispose to colon cancer contribute to a large percentage of early-onset colon cancers but a smaller number of tumours found in individuals over 50 years of age. The "Amsterdam" criteria for HNPCC, as defined by the International Collaborative Group on HNPCC, include having at least one family member affected before 50 years of age (Vasen *et al.*, 1991). Therefore, the identification of a gene which predisposed to even a small percentage of colon cancers of later age at onset of the disease could contribute to as many cases as one of the genes responsible for early age at onset of disease.

An estimated 20-25 % of all colorectal cancers are associated with a family history of the disease (Bonelli *et al.*, 1988). Between 5-10 % of all cases of colorectal cancer are

due to HNPCC and another 0.2% are due to FAP (Lynch *et al.*,1993). If these frequencies are found in the population in this study (n=119), between 24-30 individuals would be expected to have a family history of disease and 6-12 would either have HNPCC or FAP. The results of our study indicate that the incidence of the Arg43Leu allele in the colon cancer population is small and appears, from the breast cancer control group, not to be significantly different from the frequency in the normal population. By comparison to the expected number of affected individuals due to an inherited predisposition in one of the known colon cancer susceptibility genes the Arg43Leu allele is present at an equivalent frequency to individuals with FAP.

The significance of this result is difficult to evaluate since information pertaining to the patient's family history is inadequate. Pedigree construction is central to exploring the possibility that there is a genetic susceptibility and assessing risk to unaffected members of the family. The age of the individual identified, 78 years old, suggests that the cancer is sporadic but without a described family medical history an inherited predisposition cannot be ruled out. The average age at diagnosis of colon cancer in Ireland in 1994 was 70 years in males and 72 years in females. In the families analysed for segregation of the Arg43Leu allele the youngest individual affected with colon cancer was 55 years of age and several elderly relatives were reported to have malignancies affecting the bowel. From these results an association between the Arg43Leu allele and the development of colon cancer is unlikely although it does not exclude the possibility that carriers of the Arg43Leu variant are predisposed to a late age at onset of colon cancer.

CHAPTER 6

HAPLOTYPE ANALYSIS OF THE ARG43LEU ALLELE

6.1.1 Introduction

The discovery of the Arg43Leu allele and the finding that it cosegregates with cancer in some families raises a number of intriguing possibilities. If the allele is associated with an increased risk of disease, then either the allele is a disease-causing mutation or it is in linkage disequilibrium with another mutation that causes the disease. The Arg43Leu allele may have arisen in one of two ways. Firstly, Arg43Leu may have arisen only once and the allele present in all those identified in these studies is identical by descent. Alternatively, it may have arisen independently on a number of occasions. In a number of instances the same mutation in *BRCA1* has been discovered in multiple families and this has usually been considered to represent a single mutational event. However, for the 185delAG mutation, haplotype analysis has indicated that it has arisen independently in at least two families (Xu *et al.*, 1997). For a recent mutation the disease will be associated with marker alleles spanning a larger chromosomal area than that for a more ancient mutation. In the case of linkage disequilibrium the effect disappears within a few generations of the founding mutation unless the marker alleles are very tightly linked. Therefore, the aim of this study was to determine the haplotype of carriers of the Arg43Leu allele using polymorphic markers on chromosome region 17q21. From this information it was hoped to ascertain if Arg43Leu arose once or multiple times, to establish the smallest common region that is inherited with Arg43Leu in the individuals with cancer and to rule out other genes located on 17q21 which are known to be involved in cancer predisposition.

6.1.2 Microsatellite DNA

The ability to map genes responsible for inherited diseases has been accelerated with the identification of DNA polymorphisms located throughout the genome. The first type of these were single base pair changes that altered, thereby creating or removing, the recognition site for restriction enzymes. The pattern of restriction fragment length polymorphisms (RFLPs) was used successfully to map genes for long established cancer syndromes such as familial adenomatous polyposis (Bodmer *et al.*, 1987). A disadvantage with RFLPs is their lack of informativeness, as in most cases the RFLP has only two alleles. Therefore an individual can only possess one of three possible variations; homozygous wild type, homozygous variant or heterozygous. Jeffreys *et al.* (1985) then made a significant breakthrough by identifying hypervariable "minisatellites" that constituted short, tandemly repeated DNA sequences that were highly variable within and between individuals. Following the identification of minisatellites came the discovery of repetitive dinucleotide sequences (Weber and May, 1989). These "microsatellites" are most often composed of repeated CA dinucleotides and are found throughout the genome. The highly polymorphic nature of microsatellites, their frequency in the genome and the advent of the polymerase chain reaction as a means of amplifying the repeat has made them the choice of marker in linkage studies, especially the identification of genes involved in cancer syndromes.

6.1.3 Location of Prohibitin

The *prohibitin* gene (*PHB*) was located to the chromosomal region 17q21 by Sato *et al.* (1992). At the time, *prohibitin* was one of a number of candidate genes for the early-onset breast/ovarian gene *BRCA1* which had been linked to this region of chromosome 17q by the marker D17S74 (Hall *et al.* 1990). Bowcock *et al.* (1993) then mapped *BRCA1* to an interval of <4 cM which excluded *PHB* and a number of other candidate

genes. A high-density map of chromosome 17q places *PHB* at between 76 cM and 77 cM, over 6 cM distal to the location of *BRCA1* at 70 cM (Anderson *et al.*, 1994). A considerable number of markers have now been identified within and around the genomic DNA sequence of *BRCA1* in comparison to other regions in the human genome which have not been as well-mapped for polymorphic markers.

6.2 Results

6.2.1 Markers

A total of 34 DNA samples were available for haplotype analysis from the probands identified as carriers of the Arg43Leu allele and their relatives. The haplotype analysis was performed using 8 polymorphic markers that covered a region of chromosome 17q21 from the *BRCA1* internal repeat, D17S855 (Figure 6.1), to the microsatellite repeat D17S1820 (Figure 6.2) which is located 9 cM distal to D17S855. Figure 6.3 shows the order of these markers on chromosome 17 as published in the Genome Data Base by Généthon. Distances in cM are sex-averaged distances for the chromosomal location of the polymorphic markers. No markers were available for the chromosomal region at 74 cM and 75 cM. All the markers are CA-repeat polymorphisms and generate PCR fragments that range from 121 bp to 280 bp in size. The polymorphic markers were chosen for their high percentage of heterozygosity and location relative to *BRCA1* and *PHB*. The size of the alleles was determined by use of a control sequence from bacteriophage M13mp18, a single-stranded circular DNA of 7.3 kb (USB) (Figure 6.4).

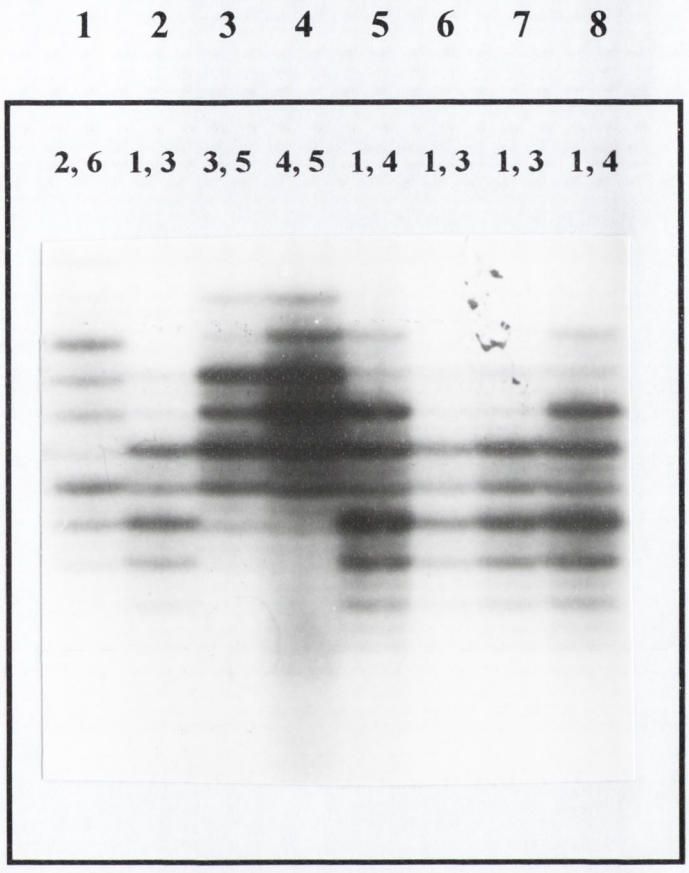


Figure 6.1. Examples of the microsatellite alleles for D17S855

Lane 1: Family 3: IV:2	alleles 2, 6
Lane 2: Family 2: III:11	alleles 1, 3
Lane 3: Family 2: III:7	alleles 3, 5
Lane 4: Family 2: III:12	alleles 4, 5
Lane 5: Family 2: III:2	alleles 1, 4
Lane 6: Family 2: III:9	alleles 1, 3
Lane 7: Family 2: III:10	alleles 1, 3
Lane 8: Family 2: IV:3	alleles 1, 4

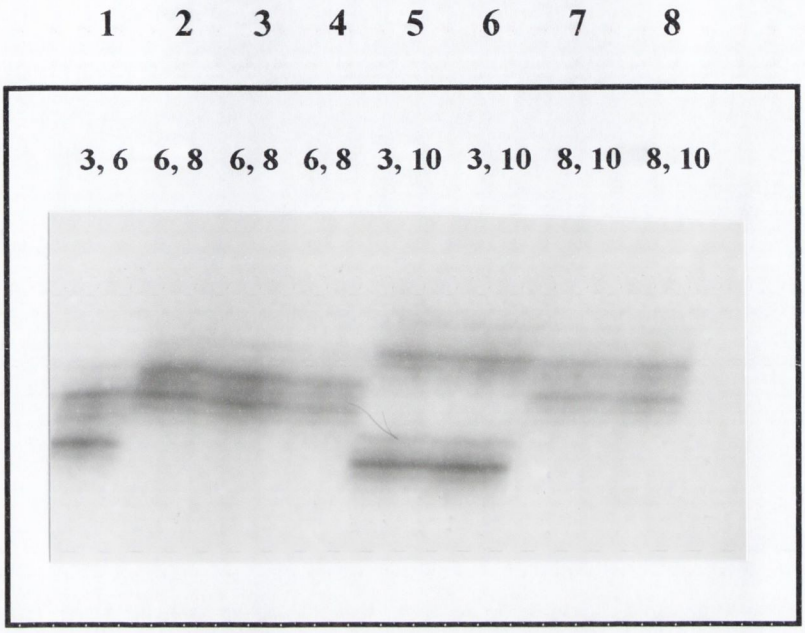


Figure 6.2. Examples of the microsatellite alleles for D17S1820

Lane 1: Family 2: IV:3	alleles 3, 6
Lane 2: Family 2: III:3	alleles 6, 8
Lane 3: Family 2: III:10	alleles 6, 8
Lane 4: Family 2: III:9	alleles 6, 8
Lane 5: Family 2: III:2	alleles 3, 10
Lane 6: Family 2: III:12	alleles 3, 10
Lane 7: Family 2: III:7	alleles 8, 10
Lane 8: Family 2: III:11	alleles 8, 10

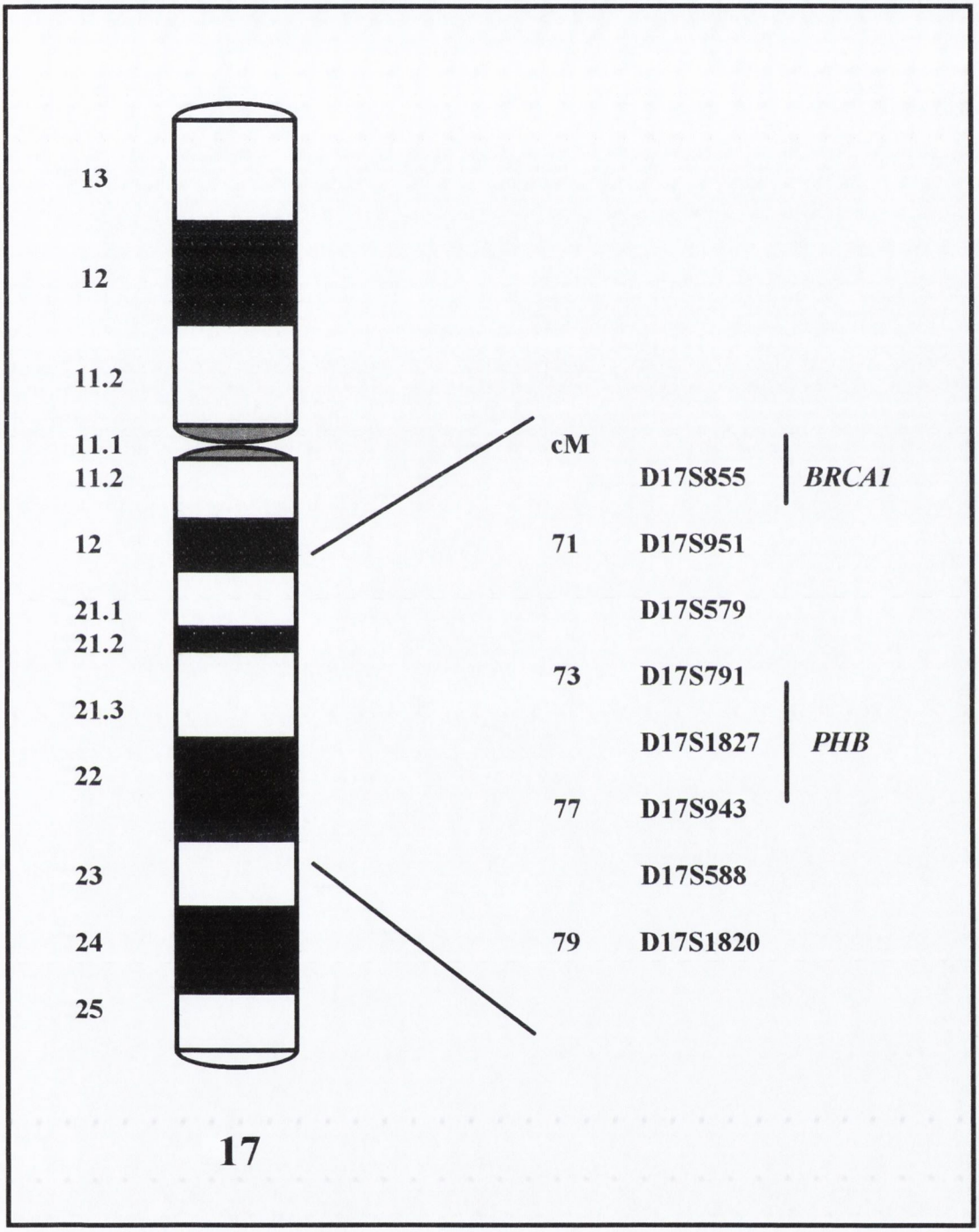


Figure 6.3. Order of microsatellite markers on chromosome 17

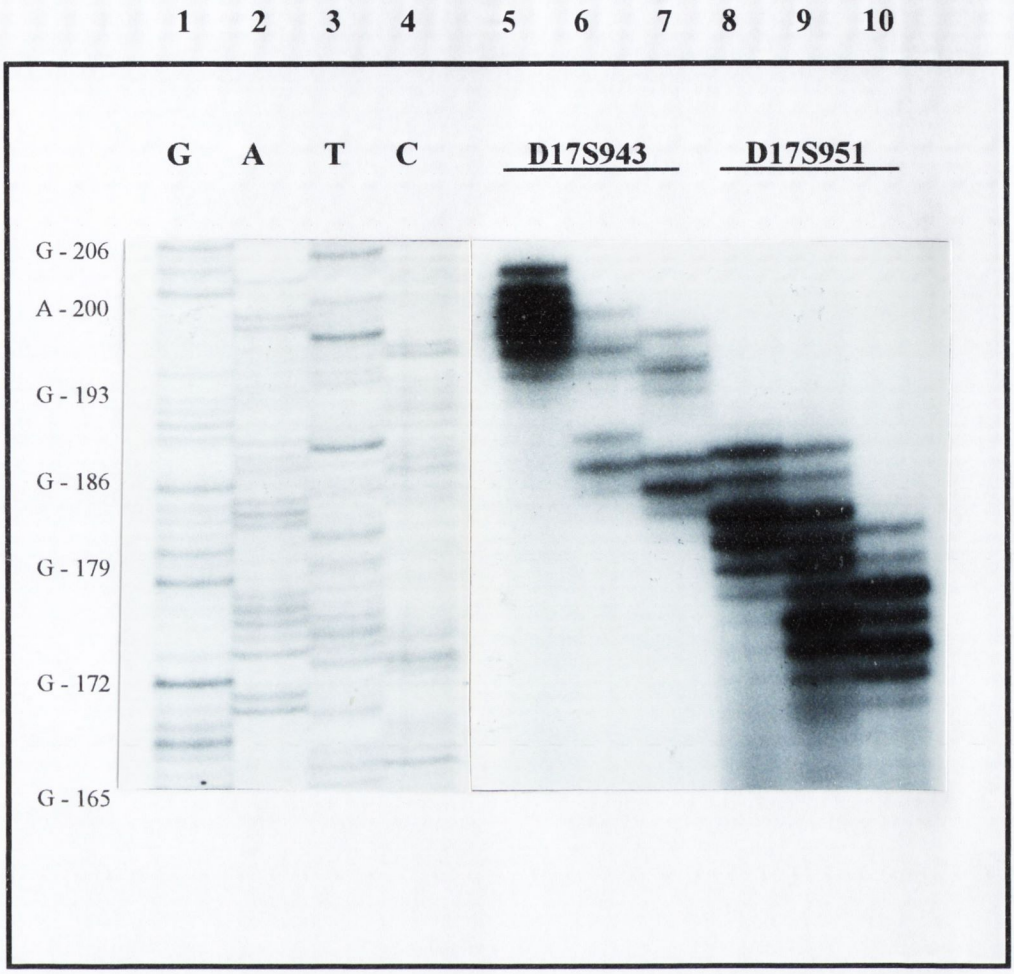


Figure 6.4. Control sequence and microsatellite alleles for D17S943 and D17S951

Lanes 1 - 4: G/A/T/C - Control sequence from bacteriophage M13mp18
Lane 5: Family 3: IV:4 alleles 5, 6
Lane 6: Family 2: III:9 alleles 2, 4
Lane 7: Family 4: IV:8 alleles 1, 3
Lane 8: Family 3: IV:3 alleles 6, 6
Lane 9: Family 5: III:6 alleles 2, 6
Lane 10: User Control alleles 1, 3

6.2.2 Haplotype Sharing in Affected and Unaffected Family Members

Five members of Family 1 were typed for the 8 microsatellite repeats, these included the 3 affected individuals in the family, the maternal aunt (II:3) and one of the proband's sisters (III:9) (Figure 6.5). The haplotype shared by both the individuals with breast cancer is not common to either of the unaffected relatives. The haplotype for the sister with colon cancer (III:1) is derived from a tissue block composed of predominantly tumour tissue. The amount of normal tissue present in the FFPE block was very small and the quantity of DNA extracted was insufficient for haplotype analysis. For the 4 markers D17S855, D17S951 (Figure 6.6), D17S579 and D17S1827, only one allele is observed. This case had previously been shown to exhibit loss of heterozygosity for *PHB* (*Chapter 3*: section 3.4.5). None of the alleles for these four markers were the same size as either of the maternal (II:1) alleles and the LOH described had shown that the maternal allele was retained and the paternal allele was lost. Two of the markers, D17S855 (Figure 6.7) and D17S1827, exhibit alleles that are shorter than the alleles described by Généthon and are not found in any of the other individuals investigated. This may indicate genomic instability as an explanation for the observed alleles at the 4 markers. Genomic instability is characterised by the expansion or contraction of short repeated DNA sequences, "microsatellite instability" (Ionov *et al.*, 1993), a feature of the inactivation of MMR genes. Genetic instability has been shown to be due to the products of the mismatch repair genes and is well described for colorectal cancers in individuals over 50 years of age where no other family member is affected (Liu *et al.*, 1995). An alternative explanation for the observed findings would be that the incorrect sample was used for the microsatellite analysis of these markers. Resolution of these issues would require DNA extracted from a normal tissue sample for this case (III:1) or a DNA sample from the individual's father (II:4). Unfortunately neither of these were available for microsatellite analysis.

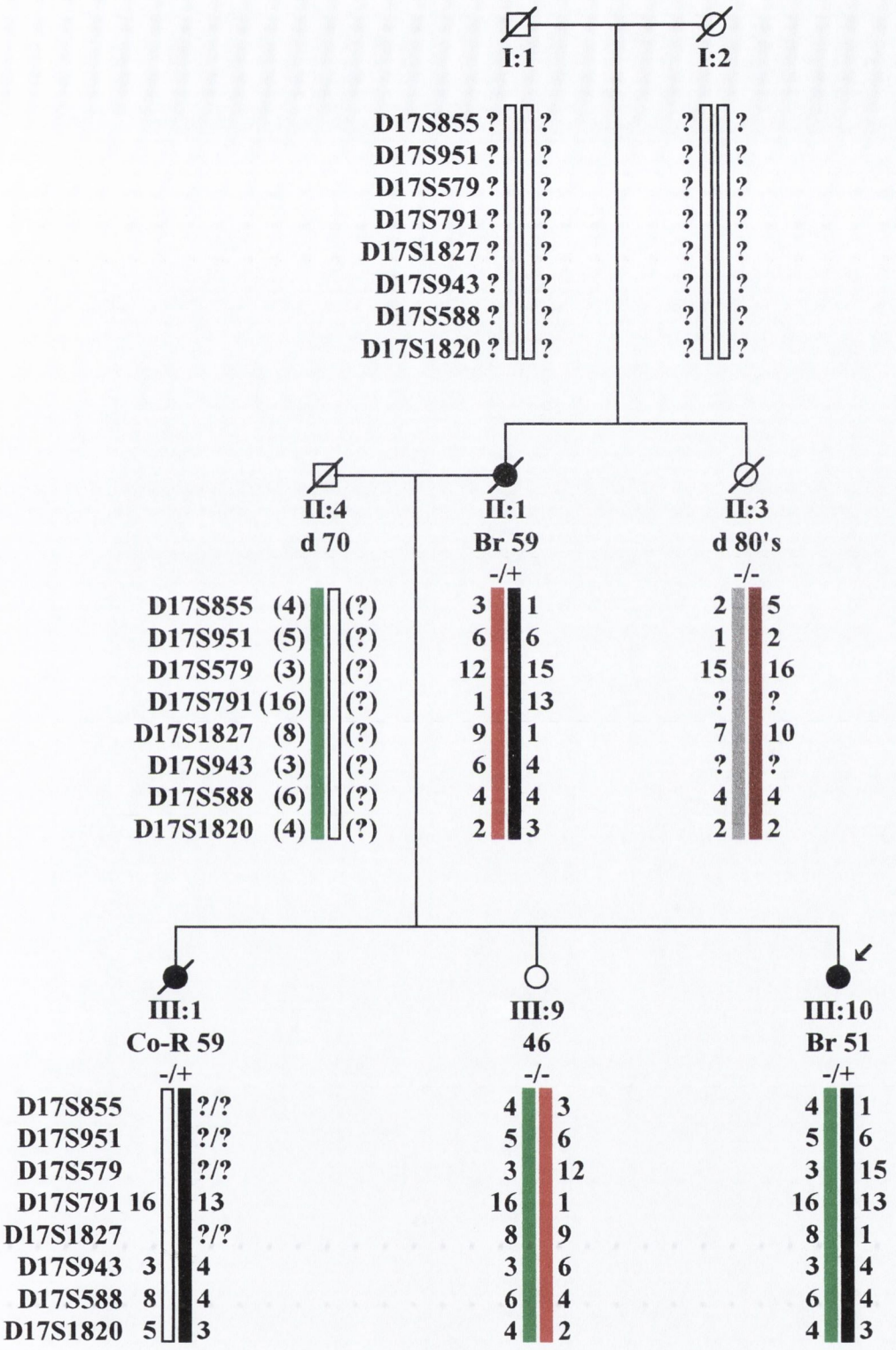


Figure 6.5. Pedigree of Family 1 with haplotype of 17q markers. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer, ? = allele not determined, (#) = allele inferred from offspring, ?/? = allele present but origin cannot be determined.

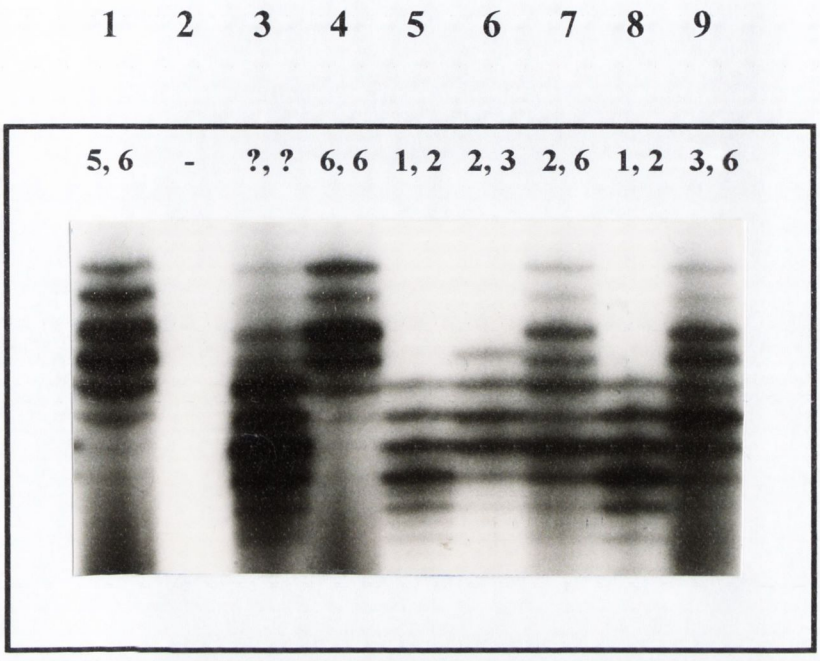


Figure 6.6. Microsatellite alleles for D17S951.

Lane 1: Family 1: III:9	alleles 5, 6
Lane 2: No product	
Lane 3: Family 1: III:1	alleles ?, ?
Lane 4: Family 1: II:1	alleles 6, 6
Lane 5: Family 1: II:3	alleles 1, 2
Lane 6: Family 5: III:1	alleles 2, 3
Lane 7: Family 5: III:6	alleles 2, 6
Lane 8: Colon study case	alleles 1, 2
Lane 9: Control case	alleles 3, 6

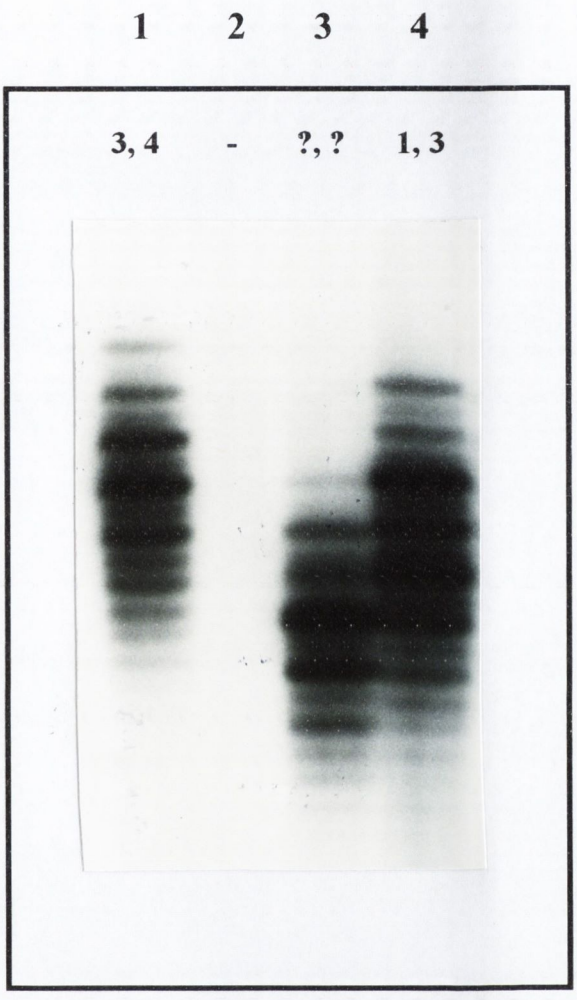


Figure 6.7. Microsatellite alleles for D17S855.

Lane 1: Family 1: III:9	alleles 4, 3
Lane 2: No product	
Lane 3: Family 1: III:1	alleles ?, ?
Lane 4: Family 1: II:1	alleles 1, 3

Family 2 is the largest family for which DNA was available for haplotype studies (Figure 6.8), with a total of 11 individuals being typed for the polymorphic repeats. Somewhat surprisingly, given the rarity of the allele, Arg43Leu is found on two distinct haplotypes in this family. Two individuals, the maternal aunt, III:12, and the maternal uncle, III:2, are Arg43Leu homozygotes. Haplotype 1 is shared by 5 individuals, the index case, IV:1, her mother, III:1, both affected maternal uncles, III:2 and III:3, and one unaffected maternal aunt, III:12. In the case of the maternal uncle, III:3, with renal cell carcinoma a recombination event below D17S943 does not alter the status of the *prohibitin* allele. Haplotype 2 is shared by 4 family members, one affected maternal uncle, III:2, and 3 unaffected maternal aunts, III:7, -11, -12. In two of these cases, III:2, -11, the haplotype has resulted from a recombination event that occurred distal to D17S791. The recombination event may have occurred distal to D17S1827 but the parent was homozygous for this allele and as such no information can be gained.

For Family 3 all 4 members that were screened for Arg43Leu were typed for the 8 polymorphic markers (Figure 6.9). Although both affected individuals, IV:2 and IV:3, share one chromosome this does not carry Arg43Leu as only the proband has the Arg43Leu allele. Therefore the proband's other chromosome represents the Arg43Leu associated haplotype and it only segregates with one cancer in this family. The 2 other relatives, IV:4 and IV:5, share identical haplotypes for both chromosomes, neither of which is shared by the index case.

In Family 4 nine of the family members were analysed for their haplotypes (Figure 6.10). The index case, IV:8, and her 3 siblings, IV:10, -11, -14, carry the Arg43Leu variant and all 4 share the same haplotype at the 8 marker loci for this chromosome. The index case's other chromosome is common to herself and 2 of her siblings whilst the other brother has inherited the other maternal chromosome. The other 2 family members that carry Arg43Leu, III:1 and III:5, share the same markers as the other carriers in the family.

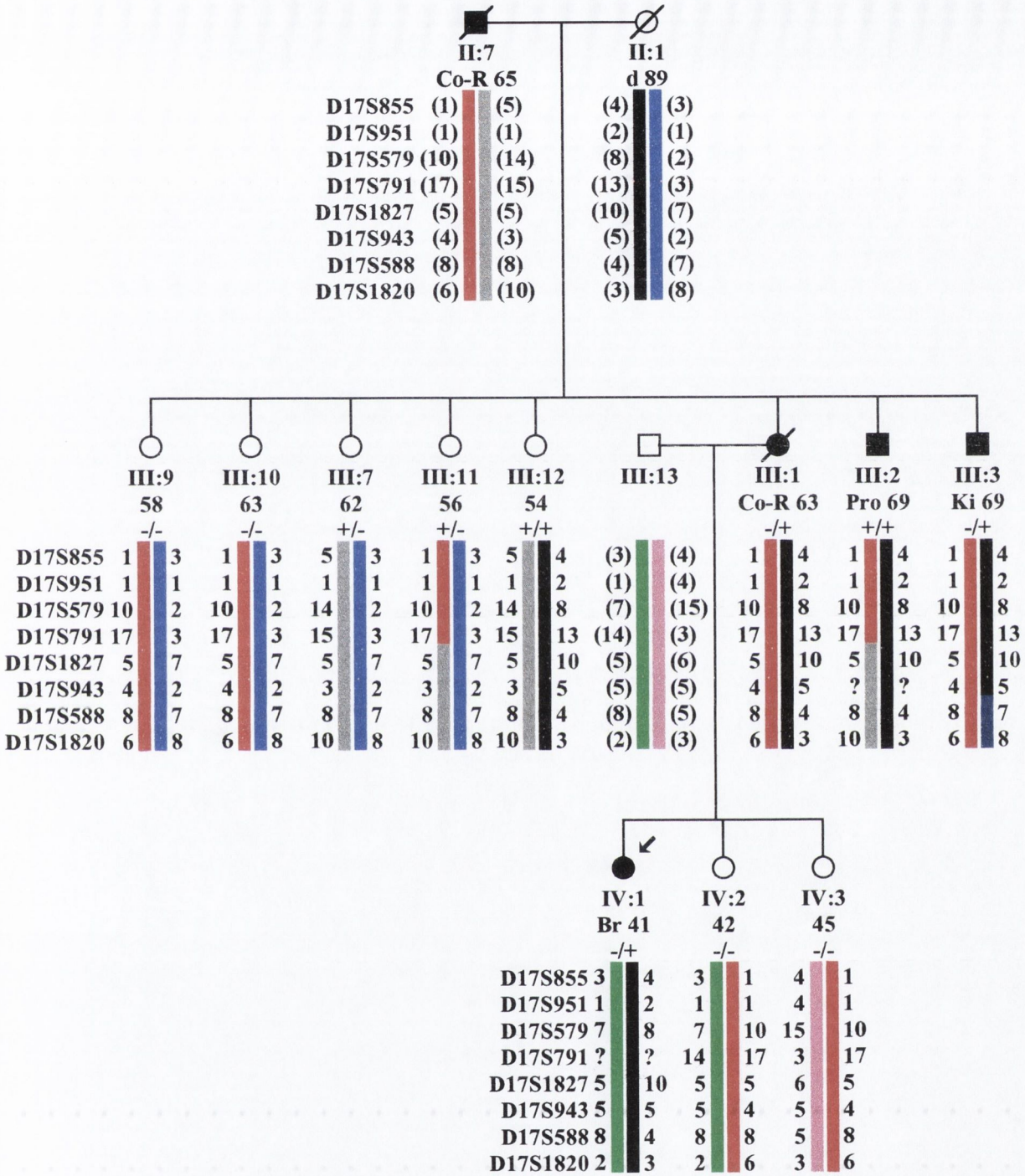


Figure 6.8. Pedigree of Family 2 with haplotype of 17q markers. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer, Ki = renal cancer, Pro = prostate cancer, ? = allele not determined, (#) = allele inferred from offspring.

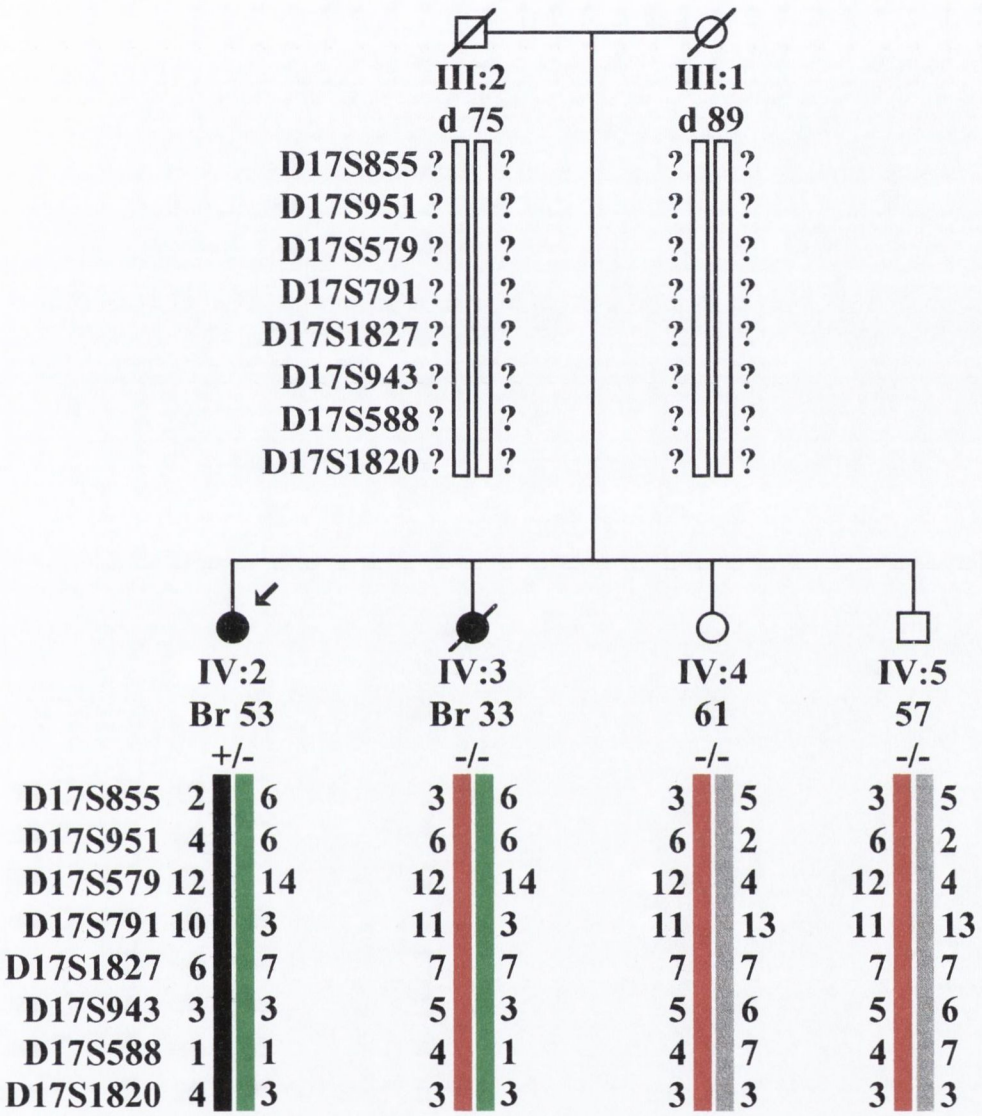


Figure 6.9. Pedigree of Family 3 with haplotype of 17q markers. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, (#) = allele inferred from offspring.

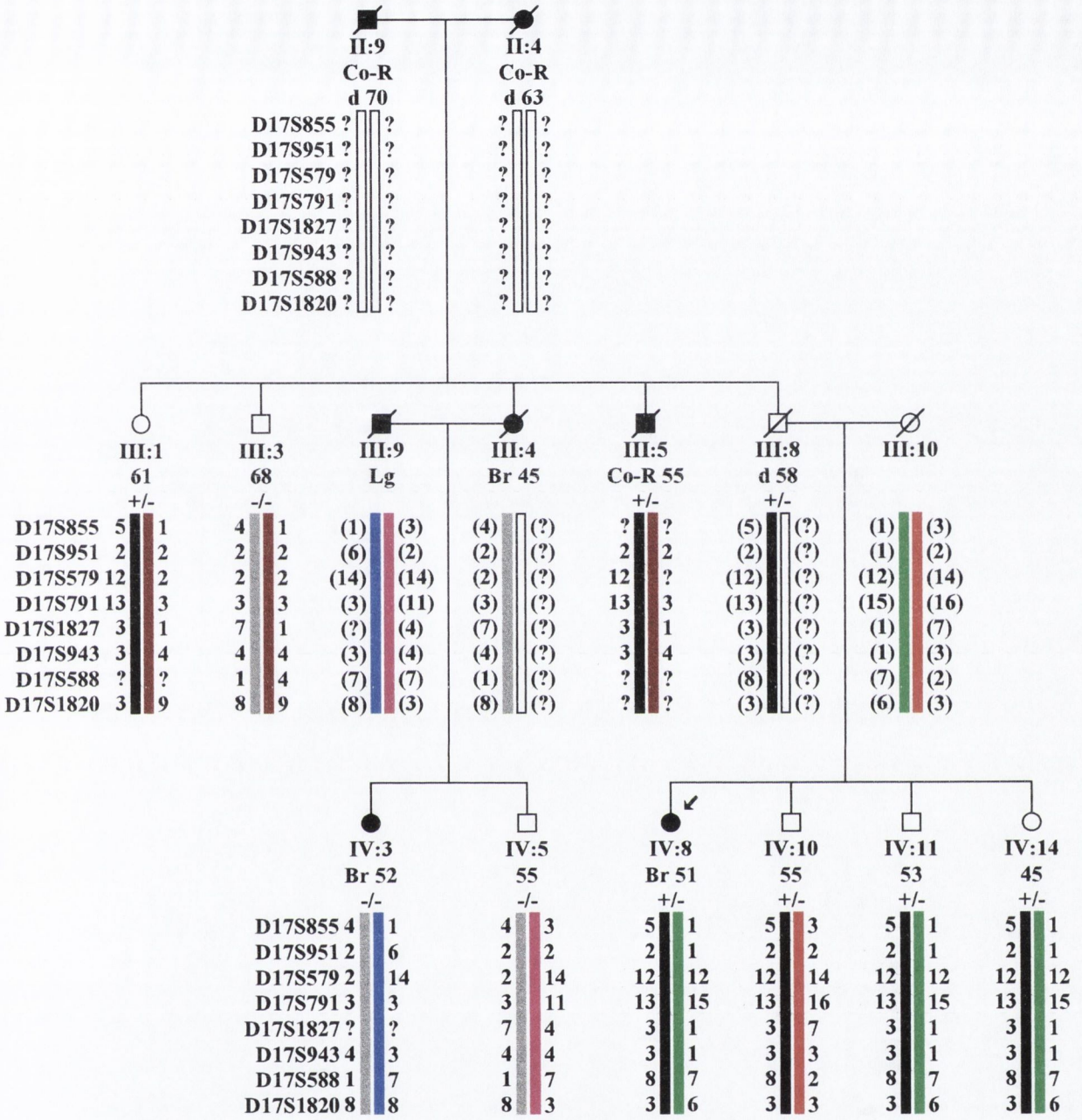


Figure 6.10. Pedigree of Family 4 with haplotype of 17q markers. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, Co-R = colo-rectal cancer, ? = allele not determined, (#) = allele inferred from offspring.

For the paternal uncle, III:5, the alleles for D17S855, D17S588 and D17S1820 were not determined. The cousins of the index case, IV:3 and IV:5, do not have the Arg43Leu associated haplotype. From these haplotypes it can be deduced that the maternal aunt with breast cancer, III:4, does not have the Arg43Leu associated haplotype. By this analysis it has been shown that the Arg43Leu allele is only associated with breast cancer in the index case in Family 4.

Two individuals from Family 5 were typed and shown to share one chromosome. The Arg43Leu allele was therefore associated with the other chromosome (Figure 6.11). The other proband identified from the breast cancer population as carrying the Arg43Leu allele was typed but the chromosomal order could not be determined as no relatives were screened. The individual identified in the colon cancer study was also typed for all markers as was the single individual identified from the control group. Phase discrimination of the *PHB* alleles with respect to the haplotypes was not possible for these individuals.

6.3 Discussion

The purpose of this study was to establish the haplotype of the individuals who had been identified as carrying the Arg43Leu allele, and their relatives, and to determine the smallest common region shared by the carriers. By examining the haplotypes of the carriers a number of inferences can be made about the Arg43Leu allele and other genes located in the chromosomal region typed. Firstly, the haplotype analysis of those with the Arg43Leu allele showed no common haplotype when viewed over all 8 polymorphic markers. Only affected individuals are informative for recombinant events, the recombinant events in Family 2 localised the possible disease-causing locus associated with the Arg43Leu allele to a region proximal to the marker D17S588.

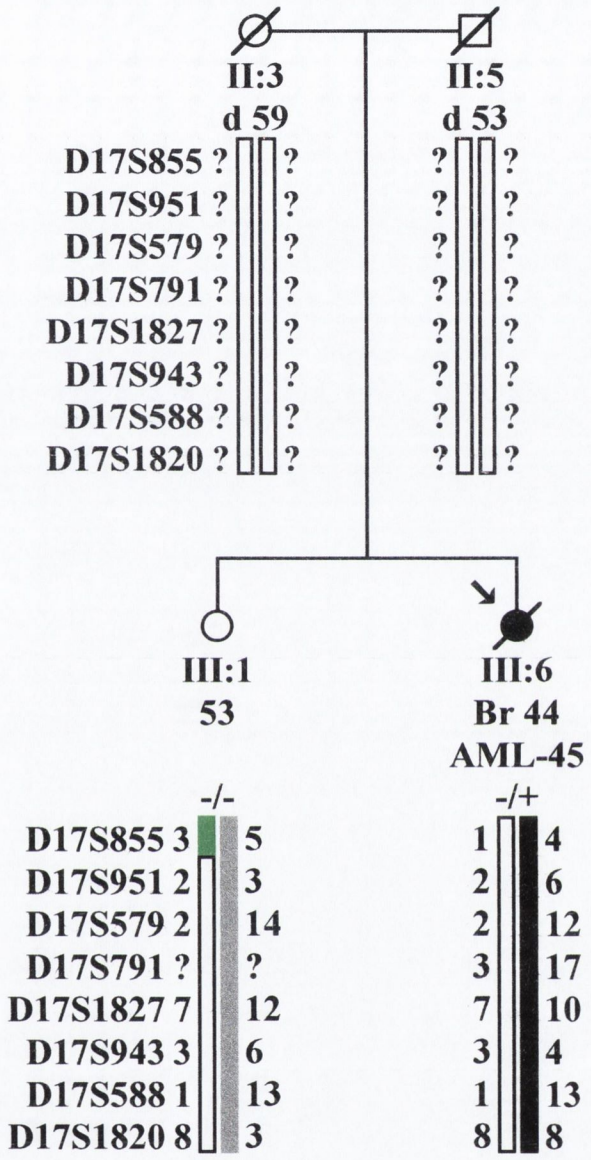


Figure 6.11. Pedigree of Family 5 with haplotype of 17q markers. Age at diagnosis, death or present age is shown when known. The Arg43Leu allele is indicated as being present (+) or absent (-) for those that have been tested. Affected individuals are illustrated by blackened symbols. Br = breast cancer, AML = acute myelomonocytic leukaemia, ? = allele not determined.

The haplotype for the two markers nearest to *PHB*, D17S1827 and D17S943, which are located in the chromosomal region between the markers D17S791 and D17S588, was not shared by any of the carriers of Arg43Leu when compared between families. When examined individually, common alleles are found for D17S1827 between Family 5 and haplotype 1 in Family 2 (Figure 6.12). Common alleles are also found for D17S943 between Family 3, Family 4 and haplotype 2 in Family 2, and between Family 1 and Family 5. In these cases the alleles described for D17S943 represent two of the most common alleles in the population. For those with the Arg43Leu allele but where no DNA from relatives was available it is possible that they share the same two alleles at D17S1827 and D17S943 as well as Arg43Leu. These three individuals are (1) the individual identified from the colon cancer study, (2) the woman from Family 6 of Russian descent with breast cancer and (3) the individual identified from the control group. However, without haplotype analysis from relatives of these individuals the chromosomal order cannot be verified.

Secondly, it is not apparent whether the Arg43Leu variant is the result of a number of independent mutations or the consequence of a founder effect. If the sequence variant has arisen multiple times then it would appear more likely than not to be the causative factor in the cancers described. The association of the Arg43Leu allele with cancer may also be due to linkage disequilibrium with another closely linked allele. In the case of a founder mutation, the Arg43Leu variant could be causative of the disease or very closely linked to either another gene or a mutation in the *prohibitin* gene outside of the coding sequence which is involved in altering the regulation of the gene. The identification of a carrier of Arg43Leu of Russian descent can be viewed as contributing to either possibility. A founder mutation would imply that Arg43Leu is an ancient mutation that has migrated a great geographical distance and that the shared haplotype would be confined to a very small chromosomal region, thus indicating that Arg43Leu is a strong candidate as the most

● IV:1 Proband Family 2 Haplotype 1			● III:6 Proband Family 5		● III:10 Proband Family 1	
D17S855	4	3	4	1	1	4
D17S951	2	1	6	2	6	5
D17S579	8	7	12	2	15	3
D17S791	13	14	17	3	13	16
D17S1827	10	5	10	7	1	8
D17S943	5	5	4	3	4	3
D17S588	4	8	13	1	4	6
D17S1820	3	2	8	8	3	4

● IV:2 Proband Family 3			● IV:8 Proband Family 4		○ III:7 Maternal Aunt Family 2 Haplotype 2	
D17S855	2	6	5	1	5	3
D17S951	4	6	2	1	1	1
D17S579	12	14	12	12	14	2
D17S791	10	3	13	15	15	3
D17S1827	6	7	3	1	5	7
D17S943	3	3	3	1	3	2
D17S588	1	1	8	7	8	7
D17S1820	4	3	3	6	10	8

■ N138 Colon Study			● II:1 Proband family 6		○ C58 Unaffected Control	
D17S855	2	4	1	3	4	3
D17S951	1	2	1	2	6	3
D17S579	12	15	?	?	?	?
D17S791	3	14	14	17	17	15
D17S1827	7	9	7	6	7	7
D17S943	3	6	3	4	3	2
D17S588	7	8	1	15	7	4
D17S1820	3	8	3	4	9	1

Figure 6.12. Haplotypes associated with Arg43Leu. For Families 1-5, the Arg43Leu allele is associated with the left hand column. For the proband from Family 6, the unaffected control and the individual from the colon study, phase discrimination was not possible.

probable cause of the cancers. The most common *BRCA1* allele in Russia is 5382insC and it is also the most common allele in Europeans as a whole, including a number of British families (Szabo and King, 1997). The 5382insC allele is thought to have originated in the Baltic region approximately 38 generations ago (Neuhausen *et al.*, 1996). If the Arg43Leu carried by the woman of Russian descent is independent of the other cases then it would suggest that it is causative but without an examination of the allele frequency in the Russian population it may represent a chance association with cancer. This raises the possibility of investigating other populations for Arg43Leu and determining if the allele is associated with cancer in these groups.

Thirdly, the haplotypes show that the affected families do not share a common *BRCA1* allele. No recombinant events between the intragenic *BRCA1* marker, D17S855, occur in affected individuals in any of the families, and only in one unaffected individual. Therefore it cannot be excluded that in some of the families different *BRCA1* mutations are responsible for the cancers, although this would seem unlikely, considering the distance between *PHB* and *BRCA1* and the high penetrance of known *BRCA1* mutations, which is not seen in the families with Arg43Leu.

In order to determine if the Arg43Leu alleles are identical by descent or have arisen on separate occasions a smaller region around the *PHB* locus would need to be typed with a higher density of markers. In chapter 4 several RFLPs were described that allowed for the distinction of at least 7 *PHB* genotypes. In 5 out of 6 of the index cases with the Arg43Leu allele an association was identified with the less common *Bsm*A1^{+} polymorphism and in the sixth case it could not be resolved. This suggests that the Arg43Leu allele was more likely to have arisen once or that it has coincidently occurred on the same *PHB* genotype a number of times. In conclusion, if the Arg43Leu allele is associated with cancer then the cancer predisposition is most likely due to the Arg43Leu variant itself or a very tightly-linked alteration in another gene.

CHAPTER 7

GENERAL DISCUSSION

GENERAL DISCUSSION

Most breast cancers are sporadic and by definition are the result of somatic mutation but a proportion of incidents, estimated to be between 5-10 %, are caused by an inherited mutation in the germline which predisposes the carrier to the disease. Such germline mutations result in family histories of disease that indicate an autosomal dominant pattern of inheritance. The disease itself, however, is recessive and requires, with the exception of dominant-negative mutations, a loss-of-function in both alleles to establish the cancer phenotype. The recent identification of two breast cancer susceptibility genes, *BRCA1* and *BRCA2*, has been shown to explain the majority of families with a strong family history of breast and/or ovarian cancer. *BRCA1* and *BRCA2* were isolated by an approach based on the genetic analyses of families with multiple cases of breast cancer with a high incidence of early age at onset. Therefore the estimated penetrance of all disease-associated alleles found in such pedigree-based studies was very high, between 70-90 % for breast cancer by age 70 (Ford *et al.*, 1998). Recent findings suggest that, in Europe, up to 4 % of women diagnosed with breast cancer before the age of 50 will carry *BRCA1* or *BRCA2* mutations (Foulkes, 1998). The findings that not all families with multiple cases of breast cancer are linked to these genes, along with other evidence, suggests that there is at least one other gene that predisposes to familial breast cancer. Other families with breast cancer and less-penetrant mutations may exist and as a consequence may contribute to a substantial proportion of breast cancer cases.

The human *prohibitin* (*PHB*) gene located on chromosome 17q21 was the focus for this thesis. *PHB* was proposed as a candidate gene for *BRCA1* but was ruled out at an early stage by Bowcock *et al.* (1993). Further LOH studies, however, determined that other tumour suppressor genes are most likely present in the region (Futreal *et al.*, 1994). Experimental observations of prohibitin in a number of different fields of investigation indicated that *prohibitin* could be a tumour suppressor gene which may be involved in the development of breast cancer (Sato *et al.*, 1992; McClung *et al.*, 1995). In this thesis a

germline mutational analysis of the *PHB* gene by denaturing gradient gel electrophoresis (DGGE) was undertaken in a population of 72 women diagnosed with breast cancer. These women also had a family history of breast cancer in a first, second or third degree relative. Among the families of the 72 women, 33 had 3 or more cases of breast or ovarian cancer and 39 were composed of only the index case and one other relative with breast cancer. Women with a family history of breast cancer have a greater likelihood of being predisposed to the disease than individuals with no affected relatives.

In this group of 72 women, four were identified as carriers of the Arg43Leu variant in exon 3 and one woman carried the silent polymorphism, Leu145Leu in exon 5. Both sequence variations create cleavage sites for restriction endonucleases. This provided a convenient and relatively quick method for determining their presence in a large number of individuals. A total of 221 women with breast cancer were screened for both of these alleles and 2 further women with the Arg43Leu variant were identified. Neither of these women had a family history of breast cancer although one did have a second primary malignancy, AML-M4Eo. The information pertaining to the family history of the other woman was limited to her first degree relatives. As well as the two novel coding sequence variants, three new intronic sequence variants were identified which are not predicted to affect normal splicing.

The G to T transversion which produces the Arg43Leu missense mutation was present in 6/221 (2.7 %) breast cancer patients unselected for family history or age at onset. A similar frequency of 3 % has been reported for the common founder mutations 185delAG in *BRCA1* and 6174delT in *BRCA2* in a series of 268 Ashkenazi Jewish women with breast cancer who were also unselected for family history or age at onset (Fodor *et al.*, 1998). The lifetime risk of developing breast cancer in Ashkenazi Jewish carriers of either of these mutations was calculated to be 36 %, a figure much lower than the estimates of 70-90 % risk of breast cancer based on studies of families at high risk of breast cancer (Ford *et al.*, 1998). The risk, however, is still approximately three times the overall risk for the general public. A study in Iceland of the common founder mutation

999del5 in *BRCA2*, which occurs in 0.6 % of the general population and 7.7 % of unselected female breast cancer patients, estimated that the risk of breast cancer for female carriers was 37.2 % at age 70 (Thorlacius *et al.*, 1998). All of these studies support the finding that the lifetime risk for carriers of *BRCA1* or *BRCA2* mutations of developing cancer is considerably lower than the estimates based on high-risk breast cancer families.

The Arg43Leu allele frequency in the breast cancer population is over 5-fold greater than that of the control group. Statistical analysis of the difference between the mutation level of the Arg43Leu variant in women with breast cancer (6/221) and the control group (1/200) revealed a p-value of $p = 0.07$. The significance of this association may be greater given the young age of the control carrier relative to the age at onset of disease for those with the Arg43Leu allele. Analysis of the association of the Arg43Leu variant with cancer in 3 families revealed a significance of $p = 0.0234$. Both of these results support the hypothesis that the Arg43Leu allele is associated with an increased risk of cancer. The value of $p = 0.07$, however, is not statistically significant and the value of $p = 0.0234$ is based on 3 out of 6 families in which a *de novo* Arg43Leu allele could be excluded. Limiting information and samples were available for the other 3 families. Only a small number of cancers were reported in these 3 families and the one case of breast cancer, not inclusive of the index cases, did not have the Arg43Leu variant. In the six families studied the predominant cancer, excluding breast cancer, was colorectal cancer. The study of normal tissue from individuals with colon cancer revealed a frequency for the Arg43Leu allele similar to that found in the control population. This would suggest that the cases of colorectal cancer in the families studied are not due to the Arg43Leu haplotype. It is important, therefore, to view an association between the Arg43Leu variant and breast or other cancers in the above context.

Haplotype analysis performed using microsatellites shows that Arg43Leu is present on different extended haplotypes for each of the index cases. This does not clarify if the variant has arisen on a number of occasions or that a founder mutation occurred a significant number of generations previously. A founder mutation would result in a

haplotype that is identical by descent but which does not extend very far from the Arg43Leu allele. A denser microsatellite map covering a narrow region, less than 1 cM, would be required to discriminate between a founder mutation or multiple events. The polymorphisms discovered in this work indicate that the Arg43Leu appears to be present on a common *PHB* genotype. Consequently an association between Arg43Leu and the development of cancer would most likely be due to the Arg43Leu variant, another prohibitin alteration outside of the regions analysed or a very tightly-linked gene.

The Arg43Leu amino acid substitution is a non-conservative alteration of a hydrophilic residue with a hydrophobic residue. The structural impact of this change on the protein and any subsequent functional alterations are unknown for the Arg43Leu variant. The known range of mechanisms by which germline mutations predispose to cancer is expanding. A functionally silent missense mutation in the coding region of the *APC* gene results in a moderately increased risk of colorectal cancer (Laken *et al.*, 1997). The mutation, I1307K, was found in individuals with a less well-defined family history of colorectal cancer and causes hypermutability in the surrounding region of the gene. The mutation was present at 6 % in the general Ashkenazi Jewish population and as such represents the most common cancer-associated polymorphism, known so far, in a specific population.

Pedigree analysis of the families with the Arg43Leu allele in this study showed a variation in age at diagnosis of cancer, cancer types and the number of affected individuals in each family. In the Icelandic population the only discernable *BRCA2* mutation, 999del5, is associated with both high-risk and low-risk families. Previous study of the 999del5 mutation had shown that 25 out of 61 carriers of this mutation had no first or second degree relatives with breast cancer (Thorlacius *et al.*, 1997). Such variation between families and risk to individuals with the mutation may be explained by the existence of modifying genes or environmental factors. For individuals with a *BRCA1* mutation the risk for ovarian cancer was 2.11 times greater when also carrying one or two rare *HRAS1* alleles, as compared to *BRCA1* carriers with only the common *HRAS1* alleles (Phelan *et*

al., 1996). The I1307K variant of *APC* may also be a genetic modifier of risk in *BRCA* heterozygotes (Redston *et al.*, 1998). Based on the assumption that Arg43Leu is having an effect, similar elements may be involved that affect the risk of cancer for the family or individual.

The 3 novel intronic sequence variants identified in this study are all single base substitutions that do not affect the splice site donor/acceptor sequences and are found in association with the *Eco*R1$^+$ RFLP in intron 2. Reports have proposed that introns have gene regulatory roles which result in susceptibility to cancer (Malkinson and You, 1994). Intronic mutations that do not directly affect the splice site donor/acceptor sequences have been found to be associated with cancer. An intronic deletion of 36bp in intron 5 of *p53* causes the skipping of exon 6 which results in the translated mRNA producing an inactive truncated protein (Voglino *et al.*, 1997) and a single germline base substitution in intron 6 of *p53* has been shown to be associated with paediatric malignancies characteristic of those found in the Li-Fraumeni syndrome (Avigad *et al.*, 1997).

As discussed earlier the functions of prohibitin are unknown and its cellular location is that of a mitochondrial protein as well as having an association with the mIgM of B cells. The 3' UTR of eukaryotic mRNAs has important biological functions and can control gene expression by affecting the localisation, stability and translation of mRNAs (Jackson, 1993; Decker and Parker, 1995). This regulatory potential may explain the localisation of prohibitin as a mitochondrial inner membrane protein and a mIgM-associated protein. The antiproliferative activity of the wild type prohibitin 1.9 kb transcript has been shown to be effective in normal cells and immortalised human cells belonging to complementation group B but not groups A, C and D (Jupe *et al.*, 1996a). The wild type prohibitin 1.9 kb mRNA has a longer 3' UTR than the 1.2 kb message and functional assays by Jupe *et al.* (1996a) showed that the antiproliferative activity is due in part to the 3' UTR exclusive to the 1.9 kb transcript. Jupe *et al.* (1996b) have identified sequence variants in this region of the 3' UTR of *PHB* in 3 cell lines derived from breast tumours that belong to complementation group B. In one cell line, SK-BR-3, thirty

mutations were detected in this region whilst in the other two cell lines (MCF-7 and BT-20) only one or two mutations were found, with one mutation from each of these cell lines also being found in SK-BR-3. One breast cancer cell line which was homozygous for the non-B allele (BT-549) and not in complementation group B was also screened for mutations in the 3' UTR but its sequence did not differ from the wild type sequence. Jupe *et al.* (1996b) concluded that the 3' UTR could act as a *trans*-acting RNA regulator and that in breast cancer cell lines homozygous for the B allele sequence alteration of the 3' UTR causes functional impairment. The small number of breast cancer cell lines investigated for the B/non-B allele and mutations in the 3' UTR requires further study to confirm an association between the two.

In this study the Arg43Leu variant was found to be associated with the non-B allele. Therefore there is no *a priori* reason to suggest that the association between cancer and the Arg43Leu allele is due to mutation in the 3' UTR. The 3' UTR of *PHB* is less highly-conserved than the coding sequence (Snedden and Fromm, 1997) and the base substitutions present in the breast cancer cell lines MCF-7 and BT-20 may represent common polymorphisms in linkage with the B allele. Other undetected coding or non-coding sequence alterations may be responsible for the loss of antiproliferative activity. The finding that the Arg43Leu allele is associated with the non-B allele supports the hypothesis that prohibitin may function as a tumour suppressor in more than one way; at both the RNA transcript and the protein level, assuming that Arg43Leu is having an effect.

An alternative pathway by which prohibitin could apply its antiproliferative activity is by regulating Ras. In yeast the *RAS2* gene is involved in longevity and extends the life of the cell by overexpression (Jazwinski, 1996). Further examination of mutants of *RAS2* and *PHB1* and their effect on longevity showed that they interact. The pathway so determined places *PHB1* upstream of *RAS2* and Phb1p may act as a GTPase activating protein (GAP) to down-regulate Ras2p activity. Ras proteins are functionally active when bound to GTP and regulate cell growth as well as differentiation and are often mutated in tumours (McCormick, 1995). Franklin and Jazwinski (1993) presented data showing that the

lifespan of the cell is extended by 30 % when *PHB1* is deleted in haploid cells but shortened by 20 % when *PHB1* was overexpressed. Hence this provides an engaging pathway through which prohibitin would perform its antiproliferative role.

The work of Coates *et al.* (1997) would appear to contradict these lifespan findings, they found that in the situation for haploid cells deleted for *PHB1* the mean and maximum lifespan are reduced. Lack of prohibitin protein in the cell may have severe implications for the lifespan of the cell. Cells lacking either *PHB1* or *PHB2*, or both, were constructed and their mean and maximum lifespan compared to that of the wild type. The double mutants showed a significant decrease in mean and maximum lifespan whereas the cells deleted for either *PHB1* or *PHB2* showed an intermediate value for the mean and maximum lifespan. For the double mutants characteristic signs of premature ageing were noted as opposed to cells dying early. Although the results of Jazwinski (1996) and Coates *et al.* (1997) appear irreconcilable both show that the consequences of functional alterations to the *prohibitin* gene, its product or other genes interacting with prohibitin are important to longevity and further substantiate that prohibitin could function as a tumour suppressor.

In the yeast, *Saccharomyces cerevisiae*, loss of function of either *prohibitin*, *PHB1*, or the *prohibitin*-related gene, *PHB2*, in cells with mitochondrial DNA deleted led to alteration of the mitochondrial morphology (Berger and Yaffe, 1998). Coates *et al.* (1997) described how deletion of *PHB1* and *PHB2* in the same cell reduced the mitochondrial membrane potential by ~80 % compared to that of cells with the wild type genes. From these observations it can be concluded that, in yeast, prohibitin is involved with mitochondrial functions including mitochondrial inheritance and the regulation of mitochondrial morphology. The interaction of prohibitin with mtDNA encoded and nuclear encoded mitochondrial proteins provides one mechanism by which prohibitin could function as a growth regulator. This may involve regulation of the levels of ATP generated within the cell thus determining whether a cell has a sufficient energy supply for cell division.

The *prohibitin* sequences of yeast and human are highly conserved and subsequently the interactions identified in yeast are possibly similar to those in the human cell. The human mitochondrial DNA (mtDNA) genome encodes 13 polypeptides all of which are present in the mitochondrial energy-generating pathway, oxidative phosphorylation. The enzyme complexes involved in oxidative phosphorylation are located within the mitochondrial inner membrane. The nuclear genes contribute a further 300-400 mitochondrial proteins required for mitochondrial structure and function. Every cell possesses hundreds of mitochondria and thousands of mtDNAs, therefore mtDNA mutations can occur in all tissue types and be somatic or germline in origin. The human mtDNA has a very high mutation rate, reported rates vary between 10 times that of nuclear genes to more than two orders of magnitude higher than that observed in nuclear sequences, both *in vitro* and *in vivo* (Wallace, 1992; Khrapko *et al.*, 1997). This rate of sequence evolution is due to both to a high mutation rate and a high mutation fixation rate.

The association between the alteration of mtDNA and cancer has been documented, albeit, on a relatively small number of occasions (Baggetto, 1993). The age-related accumulation of somatic mtDNA mutations, deletions and point mutations, is paralleled by the decrease in the capacity for oxidative phosphorylation. To date no advantageous mtDNA mutations have been detected. Germline mtDNA mutations are associated with an array of degenerative diseases such as Alzheimer disease, Leber hereditary optic neuropathy (LHON) and myoclonic epilepsy and ragged-red fibre disease (Wallace, 1994).

An increased mtDNA mutation rate has been observed in colorectal and gastric tumours as compared to their paired normal tissue (Alonso *et al.*, 1997) and the mitochondria of colonic tumours have been found to be altered in structure and function (Modica-Napolitano *et al.*, 1989). In leukaemia, Boultwood *et al.* (1996) described the observation that mtDNA is always amplified in AML. This may be an attempt to increase ATP production due to the demands for energy precipitated by the proliferative state (Gattermann and Aul, 1997). To identify the genes involved in the progression of human

breast cancer, differential gene expression in two mammary carcinoma cell lines was investigated (Schiemann *et al.*, 1998). The results showed that tumour progression is mediated by transcriptional regulation of mitochondria-related genes as well as cell cycle-related proteins. Such empirical evidence is of interest considering (1) the possible association of the Arg43Leu allele with breast and colorectal cancer in the families in this study; (2) the index case in Family 5 with AML-M4Eo and breast cancer and (3) prohibitin's interaction with mitochondrial proteins that are required for mitochondrial inheritance and morphology.

In this thesis 5 novel germline mutations of the human *PHB* gene have been identified. Two of these are coding sequence variants and the other three are intronic base substitutions. One of the coding sequence alterations, Arg43Leu, is found at an increased frequency in the breast cancer population as compared to controls. This association is not statistically significant ($p=0.07$) but in the three families where the Arg43Leu allele was demonstrated not to be the result of a *de novo* mutation it was significantly associated with the development of breast and other cancers. No association, however, was found between the Arg43Leu variant and a population with colorectal cancer. If the colorectal cancers in the three families are not due to the Arg43Leu variant then there is no evidence for an association of the Arg43Leu variant with cancer. Haplotype analysis has determined that if an association between the Arg43Leu variant and cancer exists, then it is unlikely to be due to *BRCA1* and that *prohibitin* or a very tightly-linked gene would be the most probable causative factor. If the hypothesis that the G to T transversion resulting in the Arg43Leu missense mutation predisposes its carriers to cancer is correct then this must be reconciled with the findings that prohibitin is a mitochondrial and mIgM-associated protein which functions upstream of Ras and has antiproliferative activity. Future studies of the expression and the cellular localisation of the prohibitin variants in normal and tumour tissue may help to resolve the pathways involved.

REFERENCES

Abrams, E. S., Murdaugh, S. E., and Lerman, L. S. Comprehensive detection of single base changes in human genomic DNA using denaturing gradient gel electrophoresis and a GC-clamp. *Genomics* 7, 463-475, 1990.

Albertsen, H. M., Smith, S. A., Mazoyer, S., Fujimoto, E., Stevens, J., Williams, B., Rodriguez, P., Cropp, C. S., Slijepcevic, P., Carlson, M., Robertson, M., Bradley, P., Lawrence, E., Harrington, T., Mei Sheng, Z., Hoopes, R., Sternberg, N., Brothman, A., Callahan, R., Ponder, B. A. J. and White, R. A physical map and candidate genes in the *BRCA1* region on chromosome 17q12-21. *Nat. Genet.* 7, 472-479, 1994.

Ali, I. U., Campbell, G., Lidereau, R., and Callaghan, R. Amplification of *c-erbB-2* and aggressive human breast tumours? *Science* 248, 1795-1796, 1988.

Allred, D. C., Clark, G. M., Elledge, R., Fuqua, S. A., Brown, R. W., and Chambness, G. C. Association of p53 protein expression with tumour cell proliferation rate and clinical outcome in node-negative breast cancer. *J. Natl. Cancer Inst.* 85, 200-206, 1993.

Alonso, A., Martin, P., Albarran, C., Aquilera, B., Garcia, O., Guzman, A., Oliva, H., and Sancho, M. Detection of somatic mutations in the mitochondrial DNA control region of colorectal and gastric tumors by heteroduplex and single-strand conformation analysis. *Electrophoresis* 18, 682-685, 1997.

Altus, M. S., Wood, C. M., Stewart, D. A., Roskams, A. J. I., Friedman, V., Henderson, T., Owens, G. A., Danner, D. B., Jupe, E. R., Dell'Orco, R. T., and McClung, J. K. Regions of evolutionary conservation between rat and human *prohibitin*-encoding genes. *Gene* 158, 291-294, 1995.

Amati, B., Dalton, S., Brooks, M. W., Littlewood, T. D., Evan, G. I., and Land, H. Transcriptional activation by the human c-Myc oncoprotein in yeast requires interaction with Max. *Nature* 359, 423-426, 1992.

Anderson, D. E. Genetic study of breast cancer: identification of a high risk group. *Cancer* 34, 1090-1097, 1974.

Anderson, L. A., Friedman, L., Osborne-Lawrence, S., Lynch, E., Weissenbach, J., Bowcock, A., and King, M. -C. High-density map of the *BRCA1* region of chromosome 17q12-q21. *Genomics* 17, 618-623, 1993.

Anderson, T. I., Gaustad, A., Ottestad, L., Farrants, G. W., Nesland, J. M., Tveit, K. M., and Borresen, A. L. Genetic alterations of the tumour suppressor gene regions 3p, 11p, 13q, 17p and 17q in human breast carcinomas. *Genes Chromosom Cancer* 4, 113-121, 1992.

Andersson, B., Ying, J. -H., Lewis, D. E., and Gibbs, R. A. Rapid characterisation of HIV-1 sequence diversity using denaturing gradient gel electrophoresis and direct automated DNA sequencing of PCR products. *PCR Methods Applic.* 2, 293-300, 1993.

Anzick, S. L., Kononen, J., Walker, R. L., Azorsa, D. O., Tanner, M. M., Guan, X. -Y., Sauter, G., Kallioniemi, O. -P., Trent, J. M., and Meltzer, P. S. AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* 277, 965-968, 1997.

Asamoto, M., and Cohen, S. *Prohibitin* gene is overexpressed but not mutated in rat bladder carcinomas and cell lines. *Cancer Lett.* 83, 201-207, 1994.

Athma, P., Rappaport, R., and Swift, M. Molecular genotyping shows that ataxia-telangiectasia heterozygotes are predisposed to breast cancer. *Cancer Genet. Cytogenet.* 92, 130-134, 1996.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K (Eds) Current protocols in molecular biology. John Wiley & Sons,1988.

Avigad, S., Barel, D., Blau, O., Malka, A., Zoldan, M., Mor, C., Fogel, M., Cohen, I. J., Stark, B., Goshen, Y., Stein, J., and Zaizov, R. A novel germline p53 mutation in intron 6 in diverse childood malignancies. *Oncogene* 14, 1541-1545, 1997.

Baffy, G., Miyashita, T., Williamson, J., and Reed, J. Apoptosis is induced by withdrawal of interleukin-3 (IL-3) from an IL-3-dependent hematopoietic cell line associated with repartioning of intracellular calcium and is blocked by enforced Bcl-2 oncoprotein production. *J. Biol. Chem.* 268, 6511-6519, 1993.

Baggetto, L. G. Role of mitochondria in carcinogenesis. *Eur. J. Cancer* 29, 156-159, 1993.

Baldin, V., Lukas, J., Marcote, M. J., Pagano, M., and Draetta, G. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev.* 7, 812-821, 1993.

Beckmann, M. W., Picard, F., An, H. X., van Roeyen, C. R. C., Dominik, S. I., Mosny, D. S., Schnurch, H. G., Bender, H. G., and Niederacher, D. Clinical impact of loss of heterozygosity of BRCA1 and BRCA2 markers in sporadic breast cancer. *Br. J. Cancer* 73, 1220-1226, 1996.

Beerman, H., Smit, V. T. H. B. M., Kluin, P. M., Bonsing, B. A., Hermans, J., and Cornelisse, C. J. Flow cytometric analysis of DNA stemline heterogeneity in primary and metastatic breast cancer. *Cytometry* 12,147-154, 1991.

Beerman, H., Kluin, P. M., Hermans, J., van de Velde, C. J., Cornelisse, C. J. Prognostic significance of DNA-ploidy in a series of 690 primary breast cancer patients. *Int. J. Cancer* 15, 34-39, 1990.

Benedict, W. F., Murphree, A. I., Banerjee, A., Spina, C. A., Sparkes, M. C., and Sparkes, R. S. Patient with 13 chromosome deletion: evidence that the retinoblastoma gene is a recessive cancer gene. *Science* 219, 973-975, 1983.

Berger, K. H., and Yaffe, M. P. Prohibitin family members interact genetically with mitochondrial inheritance components in *Saccharomyces cerevisiae. Mol. Cell. Biol.* 18, 4043-4052, 1998.

Berns, P. M. J. J., Foekens, J. A., Van Putten, W. L. J., Van Staveren, I. L., Alexieva-Figusch, J., and Klijn, J. G. M. Prognostic factors in human primary breast cancer: comparison of c-myc and HER2/neu amplification. *J. Steroid Biochem. Mol. Biol.* 43, 13-19, 1992a.

Berns, P. M. J. J., Foekens, J. A., Van Putten, W. L. J., Van Staveren, I. L., Portengen, H., and Foekens J. A. c-Myc amplification is a better prognostic factor than HER2/neu amplification in primary breast cancer. *Cancer Res.* 52, 1107-1113, 1992b.

Berns, P. M. J. J., Klijn, J. G. M., Van Staveren, I. L., Portengen, H., Noordegraaf, E., and Foekens, J. A. Prevalence of amplification of the oncogenes *c-myc*, *her/neu*, and *int-2* in thousand human breast tumours: correlation with steroid recptors. *Eur. J. Cancer* 28, 697-700, 1992c.

Bhagirath, T. H., Condie, A., Dunlop, M. G., Wyllie, A. H., and Prosser, J. Exclusion of constitutional p53 mutations as a cause of genetic susceptibility to colorectal cancer. *Br. J. Cancer* 68, 712-714, 1993.

Bignell, G. R., Barfoot, R., Seal, S., Collins, N., Warren, W., Stratton, M. R. Low frequency of somatic mutations in the *LKB1*/Peutz-Jeghers syndrome gene in sporadic breast cancer. *Cancer Res.* 58, 1384-1386, 1998.

Birch, J. M., Hartley, A. L., Tricker, K. J., Prosser, J., Condie, A., Kelsey, A. M., Harris, M., Morris Jones, P. H., Binchy, A., Crowther, D., Craft, A. W., Eden, O. B., Evans, D. G. R., Thompson, E., Mann, J. R., Martin, J., Mitchell, E. L. D., and Santibanez-Koref, M. F. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res.* 54, 1298-1304, 1994.

Bishop, D. T., and Hopper, J. At-tributable risks? *Nat. Genet.* 15, 226, 1997.

Blanquet, V., Turlleau, C., Gross, M. -S., Goossens, M., and Besmond, C. Identification of germline mutations in the *RB1* gene by denturant gradient gel electrophoresis and polymerase chain reaction direct sequencing. *Hum. Mol. Gen.* 2, 975-979, 1993.

Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J. R., Ellis, A., Gorman, P., Lucibello, F. C., Murday, V. A., Rider, S. H., Scambler, P., Sheer, D., Solomon, E., and Spurr, N. K. Localisation of the gene for familial adenomatous polyposis on chromosome 5. *Nature* 328, 614-616, 1987.

Bonelli, L., Martines, H., Conio, M., Bruzzi, P., and Aste, H. Family history of colorectal cancer as a risk for benign and malignant tumours of the large bowel. A case control study. *Int. J. Cancer* 41, 513-517, 1988.

Bonsing, B. A., Beerman, H., Kuipers-Dijkshoorn, N., Fleuren, G. J., and Cornelisse, C. J. High levels of DNA index heterogeneity in advanced breast cancers. Evidence for DNA ploidy differences between lymphatic and hematogenous metastses. *Cancer* 71, 382-391, 1993.

Borg, A., Sigurdsson, H., Clark, G. M., Ferno, M., Fuqua, S. A. W., Olsson, H., Killander, D., and McGuire, W. L. Association of *int-2/hst-1* coamplification in primary breast cancer with hormone-dependent phenotype and poor prognosis. *Br. J. Cancer* 63, 136-142, 1991.

Borg, A., Zhang, Q. X., Johannsson, O., and Olsson, H. High frequency of allelic imbalance at the *BRCA1* region on chromosome 17q in both familial and sporadic ductal breast carcinomas. *J. Natl. Cancer Inst.* 86, 792-794, 1994.

Borresen, A. -L., Andersen, T. I., Garber, J., Barbier-Piraux, N., Thorlacius, S., Evfjord, J., Ottestad, L., Smith-Sorensen, B., Hovig, E., Malkin, D., and Friend, S. H. Screening for germline *TP53* mutations in breast cancer patients. *Cancer Res.* 52, 3234-3236, 1992.

Boultwood, J., Fidler, C., Mills, K. I., Frodsham, P. M., Kusec, R., Gaiger, A., Gale, R. E., Linch, D. C., Littlewood, T. J., Moss, P. A. H., and Wainscoat, J. S. Amplification of mitochondrial DNA in acute myeloid leukaemia. *British Journal of Haematology* 95, 426-431, 1996.

Bowcock, A. M., Anderson, L. A., Friedman, L. S., Black, D. M., Osborne-Lawrence, S., Rowell, S. E., Hall, J. M., Solomon, E., and King, M. -C. THRA1 and D17S183 flank an interval of <4 cM for the breast-ovarian cancer gene (*BRCA1*) on chromosome 17q21. *Am. J. Hum. Genet.* 52, 718-722, 1993.

Broca, P. O. *Traite des Tumeurs*, Vol 1. Paris: P. Asselin, 1866.

Bronner, C. E., Baker, S. M., Morrison, P. T., Warren, G., Smith, L. G., Lescoe, M. K., Earabino, C., Lipford, J., Lindblom, A., Tannergard, P., Bollag, R. J., Godwin, A. R., Ward, D. C., Nordenskjold, M., Fishel, R., Kolodner, R., and Liskay, R. M. Mutation in the DNA mismatch repair gene homolog *hMLH1* is associated with hereditary nonpolyposis colon cancer. *Nature* 368, 258-261, 1994.

Caldas, C., Hahn, S. A., da Costa, L. T., Redston, M. S., Schutte, M., Seymour, A. B., Weinstein, C. L., Hruban, R. H., Yeo, C. J., and Kern, S. E. Frequent somatic mutations and homozygous deletions of the *p16* (*MTS1*) gene in pancreatic adenocarcinoma. *Nat. Genet.* 8, 27-32, 1994.

Campisi, J. The biology of replicative senescence. *Eur. J. Cancer* 33, 703-709, 1997.

Casey, G., Smith, R., McGillivray, D., Peters, G., and Dickson, C. Characterisation and chromosome assignment of the human homologue of *int-2*, a potential proto-oncogene. *Mol. Cell. Biol.* 6, 502-510, 1986.

Cavanee, W. K., Drya, T. P., Phillips, R. A., Benedict, W. F., Godbout, R., Gallie, B. L., Murphee, A. L., Strong, L. C., and White, R. L. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* 305, 779-789, 1983.

Cavanee, W. K., Hansen, M. F., Nordenskjold, M., Kock, E., Maumanee, I., Squire, J. A., Phillips, R. A., and Gallie, B. L. Genetic origin of mutations predisposing to retinoblastoma. *Science* 228, 501-503, 1985.

Choongkittaworn, N. M., Kim, K. H., Danner, D. B., and Griswold, M. D. Expression of prohibitin in rat seminiferous epithelium. *Biol. Reproduct.* 49, 300-310, 1993.

Clark, G. M., and McGuire, W. L. Follow-up study of HER-2/*neu* amplification in primary breast cancer. 51, 944-948, 1991.

Claus, E. B., Risch, N., and Thompson, W. D. Age of onset as an indicator of familial risk of breast cancer. *Am. J. Epidemiol.* 131, 961-972, 1990.

Claus, E. B., Risch, N., and Thompson, W. D. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am. J. Hum. Genet.* 48, 232-242, 1991.

Cleton-Jansen, A. M., Collons, N., Lakhani, S. R., Weissenbach, J., Devilee, P., Cornelisse, C. J., and Stratton, M. R. Loss of heterozygosity in sporadic breast tumours at the *BRCA2* locus on chromosome 13q12-13. *Br. J. Cancer* 72, 1241-1244, 1995.

Cliby, W., Sarkar, G., Ritland, S. R., Hartmann, L., Podratz, K. C., and Jenkins, R. B. Absence of *prohibitin* gene mutations in human epithelial ovarian tumours. *Gynecol. Oncol.* 50, 34-37, 1993.

Coates, P. J., Jamieson, D. J., Smart, K., Prescott, A. R., and Hall, P. A. The prohibitin family of mitochondrial proteins regulate replicative lifespan. *Curr. Biol.* 7, 607-610, 1997.

Coles, C., Thompson, A. M., Elder, P. A., Cohen, B. B., Mackenzie, I. M., Cranston, G., Chetty, U., Mackay, J., Macdonald, M., Nakamura, Y., et al., Evidence implicating at least two genes on chromosome 17p in breast carcinogenesis. *Lancet* 336, 761-763, 1990.

Coles, C., Condie, A., Chetty, U., Steel, C. M., Evans, H. J., and Prosser, J. p53 mutations in breast cancer. *Cancer Res.* 52, 5291-5298, 1992.

Collins, N., McManus, R., Wooster, R., Mangion, J., Seal, S., Lakhani, S. R., Ormiston, W., Daly, P. A., Ford, D., Easton, D. F., and Stratton, M. R. Consistent loss of the wild type allele in breast cancer from a family linked to the *BRCA2* gene on chromosome 13q12-13. *Oncogene* 10, 1673-1675, 1995.

Collins, F. S. *BRCA1* - Lots of mutations, lots of dilemmas. *New Eng. J. Med.* 334, 186-188, 1996.

Cornelis, R. S., Devilee, P., van Vliet, M., Kuipers-Dijkshoorn, N., Kersenmaeker, A., Bardoel, A., Khan, P. M., and Cornelisse, C. J. Allele loss patterns on chromosome 17q in 109 breast carcinomas indicate at least two distinct target regions. *Oncogene* 8, 781-785, 1993.

Cotton, R. G. H., Rodrigues, N. R., and Campbell, R. D. Reactivity of cytosine and thymine in single-base-pair mismatches and hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc. Natl. Acad. Sci. USA* 85, 4397-4401, 1988.

Couch, F. J., and Weber, B. L. Mutations and polymorphisms in the familiar early-onset breast cancer (*BRCA1*) gene: Breast Cancer Information Core. *Hum. Mutat.* 8, 8-18, 1996a.

Couch, F. J., Farid, L. M., Deshano, M. L., Tavtigian, S. V., Calzone, K., Campeau, L., Peng, Y., Bogden, B., Chen, Q., Neuhausen, S., Shattuck-Eidens, D., Godwin, A. K., Daly, M., Radford, D. M., Sedlacek, S., Rommens, J., Simard, J., Garber, J., Merajver, S., and Weber, B. L. *BRCA2* germline mutations in male breast cancer cases and breast cancer families. *Nat. Genet.* 13, 123-125, 1996b.

Courjal, F., and Theillet, C. Comparative genomic hybridization analysis of breast tumours with predetermined profiles of DNA amplification. *Cancer Res.* 57, 4368-4377, 1997.

Cropp, C. S., Champene, M. -H., Lidereau, R., and Callahan, R. Identification of three regions on chromosome 17q in primary human breast carcinomas which are frequently deleted. *Cancer Res.* 53, 5617-5619, 1993.

Cropp, C. S., Nevanlinna, H. A., Pyrhonen, S., Stenman, U. -H., Salmikangas, P., Albertsen, H., White, R., and Caliahan, R. Evidence for involvement of *BRCA1* in sporadic breast carcinomas. *Cancer Res.* 54, 2548-2551, 1994.

Decker, C., and Parker, R. Diversity of cytoplasmic functions for the 3' untranslated region of eukaryotic transcripts. *Curr. Op. Cell Biol.* 7, 386-392, 1995.

Dell'Orco, R. T., Liu, X. T., Jupe, E. R., Kiehlbauch, J. L., and McClung, J. K. Prohibitin expression in immortalised cell populations. *In Vitro: Cell. Develop. Biol.* 30A, 110, 1994.

Dell'Orco, R. T., McClung, J. K., Jupe, E. R., and Liu, X. T. Prohibitin and the senescent phenotype. *Exp. Gerontol.* 31, 245-252, 1996.

Del Sal, G. D., Ruaro, E. M., Utrera, R., Cole, C. N., Levine, A. J., and Schneider, C. Gas1-induced growth suppression requires a transactivation -independent p53 function. *Mol. Cell. Biol.* 15, 7152-7160, 1995.

Devilee, P., Van den Broek, M., Kuipers-Dijkshoorn, N., Kolluri, R., Meera, Khan, P. M., Pearson, P. L., and Cornelisse, C. J. At least four different chromosomal regions are involved in loss of heterozygosity in human breast cancer. *Genomics* 5, 554-560, 1989.

Devilee, P., and Cornelisse, C. J. Somatic genetic changes in human breast cancer. *Biochim Biophys Acta* 1198, 113-130, 1994.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., and Weissenbach, J. A comprehensive genetic map of the human genome based on 5, 264 microsatellites. *Nature* 380, 152-154, 1996.

Dolmetsch, R. E., Lewis, R. S., Goodnow, C. C., and Healy, J. I. Differtial activation of transcription factors induced by Ca^{2+} response amplitude and duration. *Nature* 386, 855-858, 1997.

Dulic, V., Lees, E., and Reed, S. I. Association of human cyclin E with a periodic G1-S phase protein kinase. *Science* 257, 1958-1961, 1992.

Dunphy, W. G. The decision to enter mitosis. *Trends Cell Biol.* 4, 202-207, 1994.

Dutrillaux, B., Gerbault-Seureau, M., Remvikos, Y., Zafrani, B., Prieur, M. Breast cancer genetic evolution. 1. Data from cytogenetics and DNA content. *Breast Cancer Res. Treat.* 19, 245-255, 1991.

Easton, D. F., Bishop, D. T., Ford, D., Crockford, G. P., and the Breast Cancer Linkage Consortium. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. *Am. J. Hum. Genet.* 52, 678-701, 1993a.

Easton, D., Ford, D., and Peto, J. Inherited breast cancer susceptibility. *Cancer Res.* 18, 95-113, 1993b.

Easton, D. F., Ford, D., Bishop, D. T., and the Breast Cancer Linkage Consortium. Breast and ovarian cancer incidence in *BRCA1* mutation carriers. *Am. J. Hum. Genet.* 56, 265-271, 1995.

Eccles, D. M., Cranston, G., Steel, C. M., Nakamura, Y., and Leonard, R. C. Allele losses on chromosome 17 in human epithelial ovarian carcinoma. *Oncogene* 5, 1599-1601, 1990.

El-Deiry, W. S., Harper, J. W., O'Connor, P. M., Velculescu, V. E., Canman, C. E., Jackman, J., Pietenpol, J. A., Burrell, M., Hill, D. E., Wang, Y., Wiman, K. G., Mercer, W. E., Kastan, M. B., Kohn, K. W., Elledge, S. J., Kinzler, K. W., and Vogelstein, B. WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer Res.* 54, 1169-1174, 1994.

Emi, M., Katagiri, T., Harada, Y., Saito, H., Inazawa, J., Ito, I., Kasumi, F., and Nakamura, Y. A novel metalloprotease disintegrin-like gene at 17q21 is somatically rearranged in 2 primary breast cancers. *Nat. Gen.* 5, 151-157, 1993.

Escot, C., Theillet, C., Lidereau, R., Spyratos, F., Champeme, M. H., Gest, J., and Callahan, R. Genetic alterations of the *c-myc* proto-oncogene in human primary breast carcinomas. *Proc. Natl. Acad. Sci.* USA 83, 4834-4838, 1986.

Eveleth, D. D., Jr. and Marsh, J. L. Sequence and expression of the *Cc* gene, a member of the dopa decarboxylase gene cluster of *Drosophila*: Possible translational regulation. *Nucleic Acids Res.* 14, 6169-6183, 1986.

Ewers, S. B., Langstrom, E., Baldetorp, B., Killander, D. Flow-cytometric DNA analysis in primary breast carcinomas and clinicopathological correlations. *Cytometry* 5, 408-419, 1984.

Fearon, E. R., and Vogelstein, B. A genetic model for colorectal tumourigenesis. *Cell* 61, 759-767, 1990.

Fishel, R., Lescoe, M. K., Rao, M. R. S., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M., and Kolodner, R. The human mutator gene homolog *MSH2* and its association with hereditary nonpolyposis colon cancer. *Cell* 75, 1027-1038, 1993.

Fischer, S. G., and Lerman, L. S. 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, 1983.

Fitzgerald, M. G., Bean, J. M., Hegde, S. R., Unsal, H., MacDonald, D. J., Harkin, D. P., Finkelstein, D. M., Isselbacher, K. J., and Haber, D. A. Heterozygous *ATM* mutations do not contribute to early-onset of breast cancer. *Nat. Genet.* 15, 307-310, 1997.

Fodor, F. H., Weston, A., Bleiweiss, I. J., McCurdy, L. D., Walsh, M. M., Tartter, P. I., Brower, S. T., and Eng, C. M. Frequency and carrier risk associated with common *BRCA1* and *BRCA2* mutations in Ashkenazi Jewish breast cancer patients. *Am. J. Hum. Genet.* 63, 45-51, 1998.

Ford, D., Easton, D. F., Bishop, D. T., Narod, S. A., Goldgar, D. E., and the Breast Cancer Linkage Consortium. Risks of cancer in *BRCA1* mutation carriers. *Lancet* 343, 692-695, 1994.

Ford, D., Easton, D. F., Statton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D. T., Weber, B., Lenoir, G., Chang-Claude, J., Sobol, H., Teare, M. D., Struewing, J., Arason, A., Scherneck, S., Peto, J., Rebbeck, T. R., Tonin, P., Neuhausen, S., Barkardottir, R., Eyfjord, J., Lynch, H., Ponder, B. A. J., Gayther, S. A., Birch, J. M., Lindbolm, A., Stoppa-Lyonnet, D., Bignon, Y., Borg, A., Hamann, U., Haites, N., Scott, R. J., Maugard, C. M., Vasen, H., Seitz, S., Cannon-Albright, L. A., Schofield, A., Zelada-Hedman, M., and the Breast Cancer Linkage Consortium. Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. *Am. J. Hum. Genet. 62:676-689*, 1998.

Foster, K. A., Harrington, P., Kerr, J., Russell, P., DiCioccio, R. A., Scott, I. V., Jacobs, I., Chevenix-Trench, G., Ponder, B. A., and Gayther, S.A. Somatic and germline mutations of the *BRCA2* gene in sporadic ovarian cancer. *Cancer Res.* 56, 3622-3625, 1996.

Foulkes, W. D. *BRCA1* and *BRCA2*: penetrating the clinical arena. *Lancet* 352, 1325-1326, 1998.

Franklin, D. S. and Jazwinski, S. M. A yeast homolog of the rat *prohibitin* antiproliferative gene is differentially expressed and determines longevity in *Saccharomyces cerevisiae*. *J. Cell. Biochem. Suppl.* 17D, 159, 1993.

Friedman, L. S., Ostermeyer, E. A., Szabo, C. I., Dowd, P., Lynch, E. D., Rowell, S. E., and King, M. -C. Confirmation of *BRCA1* by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nat. Genet.* 8, 399-404, 1994a.

Friedman, L. S., Ostermeyer, E. A., Lynch, E. D., Szabo, C. I., Anderson, L. A., Dowd, P., Lee, M. K., Rowell, S. E., Boyd, J., and King, M. -C. The search for *BRCA1*. *Cancer Res.* 54, 6374-6382, 1994b.

Friedman, L. S., Szabo, C. I., Ostermeyer, E. A., Dowd, P., Butler, L., Park, T., Lee, M. K., Goode, E. L., Rowell, S. E., and King, M. -C. Novel inherited mutations and variable expressivity of *BRCA1* alleles, including the founder mutation 185delAG in Ashkenazi Jewish families. *Am. J. Hum. Genet.* 57, 1284-1297, 1995.

Friedman, L. S., Gayther, S. A., Kurosaki, T., Gordon, D., Noble, B., Casey, G., Ponder, B. A. J., and Anton-Culver, H. Mutation analysis of BRCA1 and BRCA2 in a male breast cancer population. *Am. J. Hum. Genet.* 60, 313-319, 1997.

Friend, S. H., Bernards, R., Rogeli, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M., and Dryja, T. P. A human DNA segment with properties of the gene that predispose to retinoblastoma and osteosarcoma. *Nature* 323, 643-646, 1986.

Fuqua, S. A., Chamness, G. C., and McGuire, W. L. Oestrogen receptor mutations in breast cancer. *J. Cell Biochem.* 51, 135-139, 1993.

Futreal, P. A., Soderkvist, P., Marks, J. R., Iglehart, J. D., Cochran, C., Barrett, J. C., and Wiseman, R. W. Detection of frequent allelic loss on proximal chromosome 17q in sporadic breast carcinomas using microsatellite length polymorphisms. *Cancer Res.* 52, 2624-2627, 1992.

Futreal, P. A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L. M., Haugen-Strano, A., Swensen, J., Miki, Y., Eddington, K., McClure, M., Frye, C., Weaver-Feldhaus, J., Ding, W., Gholami, Z., Soderkvist, P., Terry, L., Jhanwar, S., Berchuck, A., Iglehart, J. D., Marks, J., Ballinger, D. G., Barrett, J. C., Skolnick, M. H., Kamb, A., and Wiseman, R. *BRCA1* mutations in primary breast and ovarian carcinomas *Science* 266, 120-122, 1994.

Gattermann, N., and Aul, C. Mitochondrial DNA amplification in AML: A hypothesis. *British Journal of Haematology* 97, 242, 1997.

Gayther, S. A., Warren, W., Mazoyer, S., Russell, P. A., Harrington, P. A., Chiano, M., Seal, S., Hamoudi, R., van Rensburg, E. J., Dunning, A. M., Love, R., Evans, G., Easton, D., Clayton, D., Stratton, M. R., and Ponder, B. A. J. Germline mutations of the *BRCA1* gene in breast/ovarian cancer families: evidence for a genotype/phenotype correlation. *Nat. Genet.* 11, 428-433, 1995.

Gayther, S., Mangion, J., Russell, P., Seal, S., Barfoot, R., Ponder, B., Stratton, M., and Easton, D. Variation of risks of breast and ovarian cancer associated with different germline mutations of the *BRCA2* gene. *Nat. Genet.* 15, 103-105, 1997.

Girard, F., Starausfeld, U., Fernandez, A., and Lamb, N. Cyclin A is required for the onset of DNA replication in mammalian fibroblasts. *Cell* 67, 1169-1179, 1991.

Goldgar, D. E., Easton, D. F., Cannon-Albright, L. A., and Skolnick, M. H. A systematic population based assessment of cancer risk in first-degree relatives of cancer probands. *J. Natl. Cancer Inst.* 86, 1600-1608. 1994.

Greenblatt, M. S., Bennett, W. P., Hollstein, M., and Harris, C. Mutations in the *p53* tumour suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res.* 54, 4855-4878, 1994.

Gu, Y., Turck, C. W., and Morgan, D. O. Inhibition of CDK2 activity in vivo by an associated 20K regulatory subunit. *Nature* 366, 707-710, 1993.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., and Weissenbach, J. The 1993-94 Généthon human genetic linkage map. *Nat. Genet.* 7, 246-339, 1994.

Hakansson, S., Johannsson, O., Johansson, U., Sellberg, G., Loman, N., Gerdes, A. -M., Holmberg, E., Dahl, N., Pandis, N., Kristoffersson, U., Olsson, H., and Borg, A. Moderate frequency of *BRCA1* and *BRCA2* germline mutations in Scandinavian familial breast cancer. *Am. J. Hum. Genet.* 60, 1068-1078, 1997.

Hall, J. M., Lee, M. K., Newman, B., Horrow, J. E., Anderson, L. A., Huey, B., and King, M. C. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 250, 1684-1689, 1990.

Hall, J. M., Friedman, L., Guenther, C., Lee, M. K., Weber, J. L., Black, D. M., and King, M. -C. Closing in on a breast cancer gene on chromosome 17q. *Am. J. Hum. Genet.* 50, 1235-1242, 1992.

Hannon, G. J., and Beach, D. $p15^{INK4b}$ is a potential effector of TGF-β-induced cell cycle arrest. *Nature* 371, 257-261, 1994.

Hara, E., Smith, R., Parry, D., Tahara, H., and Peters, G.Regulation of p16 (CdkN2) expression and its implications for cell immortalisation and senescence. *Mol. Cell Biol.* 16, 859-867, 1996.

Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K., and Elledge, S. J. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75, 805-816, 1993.

Harris, C.C. The 1995 Walter Hubert Lecture-molecular epidemiology of human cancer: insights from the mutational analysis of the p53 tumour-suppressor gene. *Br. J. Cancer* 73, 261-269, 1996.

Hartley, A. L., Birch, J. M., Marsden, H. B., and Harris, M. Malignant melanoma in families of children with osteosarcoma, chondrosarcoma and adrenal cortical carcinoma. *J. Med. Genet.* 24, 664-668, 1987.

Hartley, A. L., Birch, J. M., Kelsey, A. M., Marsden, H. B., Harris, M., and Teare, M. D. Are germ tumours part of the Li-Fraumeni cancer family syndrome. *Cancer Genet. Cytogenet.* 42, 221-226, 1989.

Hayashi, K. PCR-SSCP: A method for detection of mutations. *Gen. Anal. Tech. Appl.* 9, 73-79, 1992.

Hayflick, L. The limited *in vitro* lifetime of human diploid cell strains. *Exp. Cell Res.* 37, 614-636, 1965.

Hensler, P. J., Annab, L. A., Barrett, J. C., and Pereira-Smith, O. M. A gene involved in control of human cellular senescence on human chromosome 1q. *Mol. Cell Biol.* 14, 2291-2297, 1994.

Hirai, H., Roussel, M. F., Kato, J. Y., Ashmun, R. A., and Sherr, C. J. Novel INK4, proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Mol. Cell Biol.* 15, 2672-2681, 1995.

Hogg, A., Bia, B., Onadim, Z., and Cowell, J. K. Molecular mechanisms of oncogenic mutations in tumours from patients with bilateral and unilateral retinoblastoma. *Proc. Natl. Acad. Sci. USA* 90, 7351-7355, 1993.

Hollstein, M., Rice, K., Greenblatt, M. S., Soussi, T., Fuchs, R., Sorlie, T., Hovig, E., Smith-Sorensen, B., Montesano, R., and Harris, C. C. Database of *p53* gene somatic mutations in human tumours and cell lines. *Nucl. Acids Res.* 22, 3547-3551, 1994.

Holt, J. T., Thompson, M. E., Szabo, C., Robinson-Benion, C., Arteaga, C. L., King, M. -C., and Jensen, R. A. Growth retardation and tumour inhibition by *BRCA1*. *Nat. Genet.* 298-302, 1996.

Horowitz, J. M., Park, S. -H., Bogenmann, E., Cheng, J. C., Yandell, D. W., Kaye, F. J., Minna, J. D., Dryja, T. P., and Weinberg, R. A. Frequent inactivation of the retinoblastoma anti-oncogene is restricted to a subset of human tumour cells. *Proc. Natl. Acad. Sci. USA* 87, 2775-2779, 1990.

Hosking, L., Trowsdale, J., Nicolai, H., Solomon, E., Foulkes, W., Stamp, G., Signer, E., and Jeffreys, A. A somatic *BRCA1* mutation in an ovarian tumour. *Nat. Genet.* 9, 343-344, 1995.

Houlston, R. S., McCarter, E., Parhboo, S., Scurr, J., and Slack, J. Family history and risk of breast cancer. *J. Med. Genet.* 29, 154-157, 1991.

Ikonen, E., Fiedler, K., Parton, R. G., and Simons, K. Prohibitin, an antiproliferative protein, is localised to mitochondria. *FEBS Lett.* 358, 273-277, 1995.

Ionov, Y., Peinado, M. A., Malkbosyan, S., Shibata, D., Perucho, M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colon carcinogenesis. *Nature* 363, 558-561, 1993.

Isola, J., De Vries, S., Chu, I., Ghazvini, S., and Waldman, F. Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumour samples. *Am. J. Pathol.* 145, 1301-1308, 1994.

Isola, J. J., Kallioniemi, O. -P., Chu, L. W., Fuqua, S. A., Hilsenbeck, S. G., Osborne, C. K., Waldman, F. M. Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer. *Am. J. Pathol.* 147, 905-911, 1995.

Jackson, R. Cytoplasmic regulation of mRNA function: the importance of the 3' untranslated region. *Cell* 74, 9-14, 1993.

Jacobs, I. J., Smith, S. A., Wiseman, R. W., Futreal, P. A., Harrington, T., Osborne, R. J., Leech, V., Molyneux, A., Berchuck, A., Ponder, B. A. J., and Bast, R. C., Jr. A deletion unit on chromosome 17q in epithelial ovarian tumours distal to the familial breast/ovarian cancer locus. *Cancer* Res.53, 1218-1221, 1993.

James, L. A., Mitchell, E. L. D., Menasce, L., and Varley, J. M. Comparative genomic hybridisation of ductal carcinoma *in situ* of the breast: identification of regions of DNA amplification and deletion in common with invasive breast carcinoma. *Oncogene* 14, 1059-1065, 1997.

Jazwinski, S. M. Longevity, genes and aging. *Science* 273, 54-59, 1996.

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

Jen, J., Harper, J. W., Bigner, S. H., Bigner, D. D., Papadopoulos, N., Markowitz, S., Willson, J. K. V., Kinzler, K. W., and Vogelstein, B. Deletion of *p15* and *p16* genes in brain tumours. *Cancer Res.* 54, 6353-6358, 1994.

Jensen, R. A., Thompson, M. E., Jetton, T. L., Szabo, C. I., van der Meer, R., Helou, B., Tronick, S. R., Page, D. L., King, M. -C., and Holt, J. T. BRCA1 is secreted and exhibits properties of a granin. *Nat. Genet.* 12, 303-308, 1996.

Jiang, W., Kahan, S., Tomita, N., Zhang, Y., Lu, S., and Weinstein, B. Amplification and expression of the human *cyclin D* gene in oesophageal cancer. *Cancer Res.* 52, 2980-2983, 1992.

Joslyn, G., Carlson, M., Thliveris, A., Albertsen, H., Gelbert, L., Samonitz, W., Groden, J., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J. P., Warrington, J., McPherson, J., wasmuth, J., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M., and White, R. Identification of deletion mutations and three new genes at the familial polyposis locus. *Cell* 66, 601-613, 1991.

Jupe, E., Liu, X. -T., Kiehlbauch, J. L., McClung, J. K., and Dell'Orco, R. T. Prohibitin antiproliferative activity and lack of heterozygosity in immortalised cell lines. *Exp. Cell Res.* 218, 577-580, 1995.

Jupe, E., Liu, X. -T., Kiehlbauch, J. L., McClung, J. K., and Dell'Orco, R. T. The 3' untranslated region of *prohibitin* and cellular immortalisation. *Exp. Cell Res.* 224, 128-135, 1996a.

Jupe, E. R., Liu, X. -T., Kiehlbauch, J. L., McClung, J. K., and Dell'Orco, R. T. Prohibitin in breast cancer cell lines: Loss of antiproliferative activity is linked to 3' untranslated region mutations. *Cell Growth and Differentiation* 7, 871-878, 1996b.

Kallioniemi, O. -P., Blanco, G., Alavaikko, M., Hietanen, T., Mattila, J., Lauslahti, K., Koivula, T. Tumour DNA ploidy as an independent prognostic factor in breast cancer. *Br. J. Cancer* 56, 637-642, 1987.

Kallioniemi, A., Kallioniemi., O. -P., Sudar., D., Rutovitz, D., Gray, J. W., Waldman, F., and Pinkel, D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumours. *Science* 258, 818-821, 1992.

Kallioniemi, A., Kallioniemi, O. -P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W., and Waldman, F. M. Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc. Natl. Acad. Sci.* 91, 2156-2160, 1994.

King, M. -C., Rowell, S. E., and Love, S. M. Inherited breast and ovarian cancer: What are the risks? What are the choices? *J. Am. Med. Assoc.* 269, 1975-1980, 1993.

Knudson, A. G. Mutation and cancer: a statistical study of retinoblastoma. *Proc. Natl. acad. Sci. USA* 68, 820-823, 1971.

Ko, L. J., and Prives, C. p53: puzzle and paradigm. *Genes Dev.* 10, 1054-1072, 1996.

Koreth, J., Bakkenist, C. J., and McGee, J. Allelic deletions at chromosome 11q22-q23.1 and 11q25-qter are frequent in sporadic breast but not colorectal cancer. *Oncogene* 14, 431-437, 1997.

King, C. R., Kraus, M. H., and Aaronsen, S. A. *Science* 229, 974-976, 1985.

Koff, A., Giordano, A., Desai, D., Yamashita, K., Harper, J. W., Elledge, S., Nishimoto, T., Morgan, D. O., Franza, B. R., and Roberts, J. M. Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle. *Science* 257, 1689-1694, 1992.

Khrapko, K., Coller, H. A., Andre, P. C., Li, X.-C., Hanekamp, J. S., and Thilly, W. G. Mitochondrial mutational spectra in human cells and tissues. *Proc. Natl. Acad. Sci.* 94, 13798-13803, 1997.

Kretzner, L., Blackwood, E. M., and Eisenman, R. N. Myc and max proteins possess distinct transcriptional activities. *Nature* 359, 426-429, 1992.

Krontiris, T. G., Devlin, B., Karp, D. D., Robert, N. J., and Risch, N. An association between the risk of cancer and mutations in the *HRAS1* minisatellite locus. *New Engl. J. Med.* 329, 517-523, 1993.

Laken, S. J., Petersen, G. M., Gruber, S. B., Oddoux, C., Ostrer, H., Giardello, F. M., Hamilton, S. R., Hampel, H., Markowitz, A., Klimstra, D., Jhanwar, S., Winawer, S., Offit, K., Luce, M. C., Kinzler, K. W., and Vogelstein, B. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in *APC*. *Nat. Genet.* 17, 79-83, 1997.

Lamers, M. C., and Bacher, S. Prohibitin and prohibitone, ubiquitous and abundant proteins that are reluctant to reveal their real identity. *Int. Arch. Allergy. Immunol.* 113, 1446-149, 1997.

Lammie, G., Fantl, V., Smith, R., Schuuring, E., Brookes, S., Michalides, R., Dickson, C., Arnold, A., and Peter, G. D11S287, a putative oncogene on chromosome 11q13, is amplified and expressed in squamous cell and mammary carcinomas and linked to *BCL-1*. *Oncogene* 6, 439-444, 1991.

Lancaster, J. M., Wooster, R., Mangion, J., Phelan, C. M., Cochran, C., Gumbs, C., Seal, S., Barfoot, R., Collins, N., Bignell, G., Patel, S., Hamoudi, R., Larsson, C., Wiseman, R. W., Berchuck, A., Iglehart, J. D., Marks, J. R., Ashworth, A., Stratton, M. R., and Futreal, P. A. *BRCA2* mutations in primary breast and ovarian cancers. *Nat. Genet.* 13, 238-240, 1996.

Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L. A., Nystrom-Lahti, M., Guan, X. -Y., Zhang, J., Meltzer, P. S., Yu, J. -W., Kao, F. -T., Chen, D. J., Ceroselatti, K. M., Fournier, R. E. K., Todd, S., Lewis, T., Leach, R. J., Naylor, S. L., Weissenbach, J., Mecklin, J. -P., Jarvinen, H., Petersen, G. M., Hamilton, S. R., Green, J., Jass, J., Watson, P., Lynch, H. T., Trent, J. M., de la Chapelle, A., Kinzler, K. W., and Vogelstein, B. Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75, 1215-1225, 1993.

Leavitt, J., Gunning, P., Porreca, P., Ng, S. -Y., Lin, C. -S., and Kedes, L. Molecular cloning and characterisation of mutant and wild type human β-actin genes. *Mol. Cell. Biol.* 4, 1961-1969, 1984.

Leggett, B., Young, J., Buttenshaw, R., Thomas, L., Young, B., Chenevix-Trench, G., Searle, J., and Ward, M. Colorectal carcinomas show frequent allelic loss on the long arm of chromosome 17 with evidence for a specific target region. *Br. J. Cancer* 71, 1070-1073, 1995.

Lerman, L. S., and Silverstein, K. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. In: Wu, R., (ed) *Methods Enzymol.* 155, 482-501, 1987.

Lerman, L. S., and Frisch, H. L. Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule? *Biopolymers* 21, 995-997, 1982.

Lerman, L. S., Fischer, S. G., Hurley, I., Silverstein, K., and Lumelsky, N. Sequence-determined DNA separations. *Annu. Rev. Biophys. Bioeng.* 13, 399-423, 1984.

Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* 88, 323-331, 1997.

Lew, D. J., and Reed, S. I. *Trends Cell Biol.* 2, 77-81, 1992

Li, F. P., Fraumeni, J. F., Mulvihill, J. J., Blattner, W. A., Dreyfus, M. G., Tucker, M. A., and Miller, R. W. A cancer family syndrome in twenty-four kindreds. *Cancer Res.* 48, 5358-5362, 1988.

Lidereau, R., Callahan, R., Dickson, C., and Peters, G. Amplification of the *int-2* gene in primary human breast tumours. *Oncogene Research* 2, 286-291, 1988.

Liscia, D. S., Merlo, G. R., Garrett, C., French, D., Mariani-Costantini, R., and Callahan, R. Expression of *int-2* mRNA in human tumours amplified at the *int-2* locus. *Oncogene* 4, 1219-1224, 1989.

Liu, B., Parsons, R., Papadopoulos, N., Nicolaides, N. C., Lynch, H. T., Watson, P., Jass, J. R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. Analysis of mismatch repair genes in hereditary nonpolyposis colorectal cancer patients. *Nature Med.* 2, 169-174, 1996.

Liu, B., Nicolaides, N. C., Markowitz, S., Willson, J. K. V., Parsons, R. E., Jen, J., Papadopolous, N., Peltomaki, P., de la Chapelle, A., Hamilton, S. R., Kinzler, K. W., and Vogelstein, B. Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat. Genet.* 9, 48-55, 1995.

Liu, X. -T., Stewart, C. A., King, R. L., Danner, D. A., Dell'Orco, R. T., and McClung, J. K. Prohibitin expression during cellular senescence of human diploid fibroblasts. *Biochem. Biophys. Res. Comm.* 201, 409-414, 1994.

Lumpkin, C. K., Jr., McClung, J. K., and Smith, J. R. Entry into S phase is inhibited in human fibroblasts by rat liver poly(A) +RNA. *Exp. Cell Res.* 160, 544-549, 1985.

Lumpkin, C. K., Jr., McClung, J. K., Pereira-Smith, O. M., and Smith, J. R. Existence of high abundance antiproliferative mRNA's in senescent human diploid fibroblasts. *Science* 232, 393-395, 1986.

Lutzker, S., and Levine, A. J. A functionally inactive p53 protein in embryonal carcinoma cells is activated by DNA damage or cellular differentiation. *Nature Med.* 2, 804-810, 1996.

Lynch, H. T., Smyrk, T. C., Watson, P., Lanspa, S. J., Lynch, J. F., Lynch, P. M., Cavalieri, R. J., and Boland, C. R. Genetics, natural history, tumour spectrum, and pathology of hereditary nonpolyposis colorectal cancer: An updated review. *Gastroenterology* 104, 1535-1549, 1993.

McClung, J. K., Danner, D. B., Stewart, D. A., Smith, J. R., Schneider, E. L., Lumpkin, C. K., Dell'Orco, R. T., and Nuell, M. J. Isolation of a cDNA that hybrid selects antiproliferative mRNA from rat liver. *Biochem. Biophys. Res. Comm.* 164, 1316-1322, 1989.

McClung, J. K., Jupe, E. R., Liu, X. -T., and Dell'Orco, R.T. Prohibitin: Potential role in senescence, development, and tumour suppression. *Exp. Gerontol.* 30, 99-124, 1995.

McCormick, F. Ras signaling and NF1. *Curr. Opin. Genet. Dev.* 5, 51-55, 1995.

Makalowski, W., Zhang, J., and Boguski, M. S. Comparative analysis of 1,196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Res.* 6, 846-857, 1996.

Malkinson, A. M., and You, M. The intronic structure of cancer-related genes regulates susceptibility to cancer. *Mol. Carcin.*, 10, 61-65, 1994.

Mariani-Costantini, R., Escot, C., Theillet, C., Gentile, A., Merlo, G., Lidereau, R., and Callahan, R. In situ *myc* expression and genomic status of the *c-myc* locus in infiltrating ductal carcinomas of the breast. *Cancer Res.* 48, 199-205, 1988.

Matsushime, H., Roussel, M. F., Ashmun, R. A., and Sherr, C. J. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. *Cell* 65, 701-713, 1991.

Matsuyama, S., Kubo, K., Ohashi, F., and Takamori, Y. Partial cloning of *prohibitin* cDNA from canine, feline, bovine, equine, and rabbit liver mRNA by RT-PCR. *J. Vet. Med. Sci.* 59, 201-203, 1997.

Merajver, S. D., Pham, T. M., Caduff, R. F., Chen, M., Poy, E. L., Cooney, K. A., Weber, B. L., Collins, F. S., Johnston, C., and Frank, T. S. Somatic mutations in the *BRCA1* gene in sporadic ovarian tumours. *Nat. Genet.* 9, 439-443, 1995.

Mettlin, C., Croghan, I., Natarajan, N., and Lane, W. The association of age and familial risk in a case-control study of breast cancer. *Am. J. Epidemiol.* 131, 973-983, 1990.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayanath, P., Ward, J., Tonin, P., Narod, S., Bristow, P. K., Norris, F. H., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J. C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M. H. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266, 66-71, 1994.

Modan, B., Gak, E., Bar Sade-Bruchim, R., Hirsch-Yechezkel, G., Theodor, L., Lubin, F., Ben-Baruch, G., Beller, U., Fishman, A., Dgani, R., Menczer, J., Papa, M., and Friedman, E. High frequency of BRCA1 185delAG mutation in ovarian cancer in Israel. *JAMA* 276, 1823-1825, 1996.

Modica-Napolitano, J. S., Steele, G. D., and Chen, L. B. Aberrant mitochondria in two human colon carcinoma cell lines. *Cancer Res.*49, 3369-3373, 1989.

Muleris, M., Almeida, A., Gerbault-Seureau, M., Malfoy, B., and Dutrillaux, B. Detection of DNA amplification in 17 primary breast carcinomas with homogeneously staining regions by a modified comparative genomic hybridisation technique. *Genes Chrom. Cancer* 10, 160-170, 1994.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Quant. Biol.* 51, 263-273, 1986.

Myers, R. M., Fischer, S. G., Lerman, L. S., and Maniatis, T. Modification of the melting properties of duplex DNA by attachment of a GC-rich DNA sequence as determined by denaturing gradient gel electrophoresis. *Nucleic Acids Res.* 13, 3111-3129, 1985 a.

Myers, R. M., Fischer, S. G., Maniatis, T., and Lerman, L. S. Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Res.* 13, 3111, 1985 b.

Myers, R. M., Lumelsky, N., Lerman, L. S., and Maniatis, T. Detection of single base substitutions in total genomic DNA. *Nature* 313, 495-498, 1985 c.

Myers, R. M., Larin, Z., and Maniatis, T. Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science* 230, 1242-1246, 1985 d.

Myers, R. M., Maniatis, T., and Lerman, L. S. Detection and localisation of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol.* 155, 501-527, 1987.

Nakajima-Iijima, S., Hamada, H., Reddy, P., and Nakunaga, T., Molecular structure of the human cytoplasmic beta-actin gene: interspecies homology of sequences in the introns. *Proc. Natl. Acad. Sci.USA* 82, 6133-6137, 1985.

Narod, S. A., Feunten, J., Lynch, H. T., Watson, P., Conway, T., Lynch, J., and Lenoir, G. M. Familial breast-ovarian cancer locus on chromosome 17q12-23. *Lancet* 338, 82-83, 1991.

Narod, S. A., Ford, D., Devilee, P., Barkadottir, R. B., Lynch, H. T., Smith, S. A., Ponder, B. A. J., Weber, B. L., Garber, J. E., Birch, J. M., Cornelis, R. S., Kelsell, D. P., Spurr, N. K., Smyth, E., Haites, N., Sobol, H., Bignon, Y.-J., Chang-Claude, J., Hamann, U., Lindblom, A., Borg, A., Piver, M. S., Gallion, H. H., Struewing, J. P., Whittemore, A., Tonin, P., Goldgar, D. E., Easton, D. F., and the Breast Cancer Linkage Consortium. An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. *Am. J. Hum. Genet.* 56, 254-264, 1995.

Neuhausen, S. L., Mazoyer, S., Friedman, L., Stratton, M., Offit, K., Caligo, A., Tomlinson, G., Cannon-Albright, L., Bishop, T., Kelsell, D., Solomon, E., Weber, B., Couch, F., Struewing, J., Tonin, P., Durocher, F., Narod, S., Skolnick, M. H., Lenoir, G., Serova, O., Ponder, B., Stoppa-Lyonnet, D., Easton, D., King, M. -C., and Goldgar, D. E. Haplotype and phenotype analysis of six recurrent *BRCA1* mutations in 61 families: Results of an international study. *Am. J. Hum. Genet.* 58, 271-280, 1996.

Ning, Y., Weber, J.L., Killary, A. M., Ledbetter, D. H., Smith, J. R., and Pereira-Smith, O. M. Genetic analysis of indefinite division in human cells: evidence for a cell senescence-related gene(s) on human chromosome 4. *Proc. Natl. Acad. Sci. USA* 88, 55635-55639, 1991.

Nishizaki, T., De Vries, S., Chew, K., Goodson, W. H. 3rd, Ljung, B. M., Thor, A., and Waldman, F. M. Genetic alterations in primary breast cancers and their metastases: direct comparison using modified comparative genomic hybridization. *Genes Chromosom Cancer* 19, 267-272, 1997.

Noda, A., Ning, Y., Venable, S. F., Pereira-Smith, O. M., and Smith, J. R. Cloning of senescent cell derived inhibitors of DNA synthesis using an expression screen. *Exp. Cell Res.* 211, 90-98, 1994.

Norwood, T. H., and Smith, J. R. Handbook of the biology of aging, second edition, Van Nostrand Reinhold Co., New York, NY (Finck, C. E., and Schneider, E. L., Eds.) 303-306, 1985.

Nuell, M. J., McClung, J. K., Smith, J. R., and Danner, D. B. Approach to the isolation of antiproliferative genes. *Exp. Gerontol.* 24, 469-476, 1989.

Nuell, M. J., Stewart, D. A., Walker, L., Friedman, V., Wood, C. M., Owens, G. A., Smith, J. R., Schneider, E. L., Dell'Orco, R. T., Lumpkin, C. K., Danner, D. B., and McClung, J. K. Prohibitin, an evolutionarily conserved intracellular protein that blocks DNA synthesis in nomal fibroblasts and HeLa cells. *Mol. Cell. Biol.* 11, 1372-1381, 1991.

Orita, M., Suzuki, Y., Sekija, T., and Hayashi, K. A rapid and sensitive detection of point mutations and genetic polymorphisms using polymerase chain reaction. *Genomics* 5, 874-879, 1989.

Osborne, R. J., Merlo, G. R., Mitsudomi, T., Venesio, T., Liscia, D. S., Cappa, A. P., Chiba, I., Takahashi, T., Nau, M. M., Callahan, R., et al., Mutations in the *p53* gene in primary human breast cancers. *Cancer Res.* 51, 6194-6198, 1991.

Otterson, G. A., Kratzke, R. A., Coxon, A., Kim, Y. W., and Kaye, F. J. Absence of p16INK4 protein is restricted to the subset of lung cancer lines that retains wild type Rb. *Oncogene* 9, 3375-3378, 1994.

Ottman, R., Pike, M. C., King, M. -C., and Henderson, B. E. Practical guide for estimating risk for familial breast cancer. *Lancet* 3, 553-558, 1983.

Pagano, M., Pepperkok, R., Verde, F., Ansorge, W., and Draetta, G. Cyclin A is required at two points in the human cell cycle. *EMBO J.* 11, 961-971, 1992.

Palombo, F., Hughes, M., Jiricny, J., Truong, O., and Hsuan, J. Mismatch repair and cancer. *Nature* 367, 417, 1994.

Pan, L. X., Diss, T. C., Peng, H. Z., and Isaacson, P. G. Clonality analysis of defined B-cell populations in archival tissue sections using microdissection and the polymerase chain reaction. *Histopathology* 24, 323-327, 1994.

Papadopoulos, N., Nicolaides, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., Adams, M. D., Venter, J. C., Hamilton, S. R., Petersen, G. M., Watson, P., Lynch, H. T., Peltomaki, P., Mecklin, J.-P., de la Chapelle, A., Kinzler, K. W., and Vogelstein, B. Mutation of a *mutL* homolog in hereditary colon cancer. *Science* 263, 1625-1629, 1994.

Pardee, A. B. G1 events and regulation of cell proliferation. *Science* 246, 603-608, 1989.

Paulovich, A. G., Toczyski, D. P., and Hartwell, L. H. When checkpoints fail. *Cell* 88, 315-321, 1997.

Pearson, A. D. J., Craft, A. W., Ratcliffe, J. M., Birch, J. M., Morris Jones, P. H., and Roberts, D. F. Two families with the Li-Fraumeni cancer family syndrome. *J. Med.Genet.* 19, 362-365, 1982.

Peltomaki, L., Aaltonen, L. A., Sistonen, P., Pylkkanen, L., Mecklin, J.-P., Jarvinen, H., Green, J. S., Jass, J. R., Weber, J. L., Leach, F. S., Petersen, G. M., Hamilton, S. R., de la Chapelle, A., and Vogelstein, B. Genetic mapping of a locus predisposing to human colorectal cancer. *Science* 260, 810-812, 1993.

Pepperkok, R., Zanetti, M., King, R., Delia, D., Ansorge, W., Philipson, L., and Schneider, C. Automatic microinjection system facilitates detection of growth inhibitory mRNA. *Proc. Natl. Acad. Sci. USA*. 85, 6748-6752, 1988.

Pereira-Smith, O. Identification of a gene on human chromosome 4 that induces senescence in a subset of immortal cells. *Br. J. Cancer* 78 (Suppl 1), 4, 1998.

Pereira-Smith, O. M., and Smith, J. R. Evidence for the recessive nature of cellular immortality. *Science* 221, 964-966, 1983.

Pereira-Smith, O. M., and Smith, J. R. Genetic analysis of indefinite division in human cells: identification of four complementation groups. *Proc. Natl. Acad. Sci. USA* 85, 6042-6046, 1988.

Phelan, C. M., Rebbeck, T. R., Weber, B. L., Devilee, P., Ruttledge, M. H., Lynch, H. T., Lenoir, G. M., Stratton, M. R., Easton, D. F., Ponder, B. A., Cannon-Albright, L., Larsson, C., Goldgar, D. E., and Narod, S. A. Ovarian cancer risk in *BRCA1* carriers is modified by the *HRAS1* variable number of tandem repeat (VNTR) locus. *Nat. Genet.* 12, 309-311, 1996.

Phelan, C. M., Lancaster, J. M., Tonin, P., Gumbs, C., Cochran, C., Carter, R., Ghadirian, P., Perret, C., Moslehi, R., Dion, F., Faucher, M. -C., Dole, K., Karimi, S., Foulkes, W., Lounis, H., Warner, E., Goss, P., Anderson, D., Larsson, C., Narod, S. A., and Futreal, P. A.. Mutation analysis of the *BRCA2* gene in 49 site-specific breast cancer families. *Nat. Genet.* 13, 120-122, 1996.

Polyak, K., Lee, M. -H., Erdjument-Bromage, H., Tempst, P., and Massague, J. Cloning of $p27^{Kip1}$, a cyclin-dependent kinase inhibitor and a potential mediator of extracellular antimitogenic signals. *Cell* 78, 59-66, 1994.

Ponten, J., Holmberg, L., Trichopoulos, D., Kallioniemi, O. -P., Kvale, G., Wallgren, A., and Taylor-Papadimitriou, J. Biology and natural history of breast cancer. *Int. J. Cancer* 5, 5-21, 1990.

Prosser, J., Elder, P. A., Condie, A., MacFadyen, I., Steel, C. M., and Evans, H. J. Mutations in *p53* do not account for heritable breast cancer: a study in five affected families. *Br. J. Cancer* 63, 181-184, 1991.

Prosser, J., Porter, D., Coles, C., Condie, A., Thompson, A. M., Chetty, U., Steel, C. M., and evans, H. J. Constitutional *p53* mutation in a non-Li-Fraumeni cancer family. *Br. J. Cancer* 65, 527-528, 1992.

Quelle, D. E., Ashmun, R. A., Shurtleff, S. A., Kato, J., Bar-Sagi, D., Roussel, M. F., and Sherr, C. J. Overexpression of mouse D-type cyc,ins accelerates G1 phase in rodent fibroblasts. *Genes Dev.* 7, 1559-1571, 1993.

Ramus, S. J., kote-Jarai, Z., Friedman, L. S., van der Looij, M., Gayther, S. A., Csokay, B., Ponder, B. A. J., and Olah, E. Analysis of *BRCA1* and *BRCA2* mutations in Hungarian families with breast or breast-ovarian cancer. *Am. J. Hum. Genet.* 60, 1242-1246, 1997.

Redston, M., Nathanson, K. L., Yuan, Z. Q., Neuhausen, S. L., Satagopan, J., Wong, N., Yang, D., Nafa, D., Abrahamson, J., Ozcelik, H., Antin-Ozerkis, D., Andrulis, I., Daly, M., Pinsky, L., Schrag, D., Gallinger, S., Kaback, M., King, M. -C., Woodage, T., Brody, L. C., Godwin, A., Warner, E., Weber, B., Foulkes, W., and Offit, K. The APCI1307K allele and breast cancer risk. *Nat. Genet.* 20, 13-14, 1998.

Reid, T., Just, K., Holtgreve-Grez, H., du Manoir, S., Speicher, M., Schrock, E., Latnam, C., Biegen, H., Zetterberg, A., Cremer, T., Auer, G. Comparative genomic hybridization of formalin-fixed paraffin-embedded breast tumours reveals different patterns of chromosomal gains and losse in fibroadenomas and diploid and aneuploid breast carcinomas. *Cancer Res.* 55, 5415-5423, 1995.

Resnitzky, D., and Reed, S. I. Different roles of cyclins D1 and E in regulation of the G1-to-S transition. *Mol. Cell Biol.* 14, 3436-3469, 1995.

Rhei, E., Kang, L., Bogomolniy, F., Federici, M. G., Borgen, P. I., and Boyd, J. Mutation analysis of the putative tumour suppressor gene PTEN/MMAC1 in primary breast carcinomas. *Cancer Res.* 57, 3657-3659, 1997.

Richter, C. and Kass, G. E. N. Oxidative stress in mitochondria: Its relationship to cellular Ca^{++} homeostasis, cell death, proliferation and differentiation. *Chem. Biol. Interact.* 77, 1-23, 1991.

Roskams, A. J. I., Friedman, V., Wood, C. M., Walker, L., Owens, G. A., Stewart, D. A., Altus, M. S., Danner, D. B., Liu, X. -T., and McClung, J. K. Cell cycle activity and expression of prohibitin mRNA. *J. Cell. Physiol.* 157, 289-295, 1993.

Rossiter, B. J. F., Grompe, M., and Caskey, C. T. Detection of deletions and point mutations. Chapter 5, PCR: A pratical approach. Oxford University Press (McPherson, M. J., Quirke, P., and Taylor, G.R. Eds).

Saito, H., Inazawa, J., Saito, S., Kasumi, F., Koi, S., Sagae, S., Kudo, R., Saito, J., Noda, K., and Nakamura., Y. Detailed deletion mapping of chromosome 17q in ovarian and breast cancers: 2-cM region on 17q21.3 often and commomnly deleted in tumours. *Cancer Res.* 53, 3382-3385, 1993.

Sato, T., Daito, H., Swensen, J., Olifant, A., Wood, C., Danner, D., Sakamoto, T., Takita, K., Kasumi, F., Miki, Y., Skolnick, M., and Nakamura, Y. The human *prohibitin* gene located on chromosome 17q21 is mutated in sporadic breast cancer. *Cancer Res.* 52, 1643-1646, 1992.

Sato, T., Sakamoto, T., Takita, K. -I., Saito, H., Okui, K., and Nakamura, Y. The human *prohibitin* (*PHB*) gene family and its somatic mutations in human tumours. *Genomics* 17, 762-764, 1993.

Sato, T., Tanigami, A., Yamakawa, K., Akiyama, F., Kasumi, F., Sakamoto, G., and Nakamura, Y. Allelotype of breast cancer: Cumulative allele losses promote tumour progression in primery breast cancer. *Cancer Res.* 50, 7184-7189, 1990.

Sattin, R. W., Rubin, G. L., Webster, L. A., Huezo, C. M., Wingo, P. A., Ory, H. W., and Layde, P. M. Family history and the risk of breast cancer. *JAMA* 253, 1908-1913, 1985.

Schechter, A. L., Stern, D. F., Vaidyanathan, L., Decker, S. J., Drebin, J. A., Greene, M. I., and Weinberg, R. A. The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature* 312, 513-516, 1984.

Scherr, C. Mammalian G1 cyclins. *Cell* 73, 1059-1065, 1993.

Schiemann, S., Schwirke, M., Brunner, N., and Weidle, U. H. Molecular analysis of two mammary carcinoma cell lines at the transcriptional level as a model for progression of breast cancer. *Clin. Exp. Metastasis* 16, 129-139, 1998.

Schmidt, E. E., Ichimura, K., Reifenberger, G., and Collins, V. P. *CDKN2* (*p16/MTS1*) gene deletion or *CDK4* amplification occurs in the majority of glioblastomas. *Cancer Res.* 54, 6321-6324, 1994.

Schubert, E. L., Lee, M. K., Mefford, H. C., Argonaza, R. H., Morrow, J. E., Hull, J., Dann, J. L., and King, M. -C. *BRCA2* in American families with four or more cases of breast or ovarian cancer: recurrent and novel mutations, variable expression, penetrance, and the possibility of families not attributable to *BRCA1* or *BRCA2*, *Am. J. Hum. Genet.* 60, 1031-1040, 1997.

Schutte, M., daCosta, L. T., Hahn, S. A., Moskaluk, C., Hoque, A. T., Rozenblum, E., Weinstein, C. L., Bittner, M., Meltzer, P. S., Trent, J. M., Yeo, C. J., Hruban, R. H., and Kern, S. E. Identification by representational difference analysis of a homozygous deletion in pancreatic carcinoma that lies within the *BRCA2* region. *Proc. Natl. Acad. Sci. USA* 92, 5950-5954, 1995.

Schuuring, E., Verhoeven, E., Van Tinteren, H., Peterse, J. L., Nunnink, B., Thunnissen, F. B. J. M., Devilee, P., Cornelisse, C. J., Van de Vijver, M. J., Mooi, W. J., and Michalides, R. J. A. M. Amplification of genes within the chromosome 11q13 region is indicative of poor prognosis in patients with operable breast cancer. *Cancer Res.* 52, 5229-5234, 1992a.

Schuuring, E., Verhoeven, E., Mooi, W. J., and Michalides, R. J. A. M. Identification and cloning of two overexpressed genes, *U21B31/PRAD1* and *EMS1*, with the amplified chromosome 11q13 region in human carcinomas. *Oncogene* 7, 355-361, 1992b.

Schuuring, E., Verhoeven, E., Litvinov, S., and Michalides, R. J. A. M. The product of the *EMS1* gene, amplified and overexpressed in human carcinomas, is homologous to a *v-src* substrate and is located in cell-substratum contact sites. *Mol. Cell Biol.* 132, 2891-2898, 1993.

Schuuring, E. The involvement of the chromosome 11q13 region in human malignancies: *cyclin D1* and *EMS1* are two new candidate oncogenes - a review. *Gene* 159, 83-96, 1995.

Schwendel, A., Richard, F., Langreck, H., Kaufmann, O., Lage, H., Winzer, K. J., Petersen, I., and Dietel, M. Chromosome alterations in breast carcinomas: frequent involvement of DNA losses including chromosomes 4q and 21q. *Br. J. Cancer* 76, 806-811, 1998.

Serova, O. M., Mazoyer, S., Puget, N., Dubois, V., Tonin, P., Shugart, Y. Y., Goldgar, D., Narod, S. A., Lynch, H. T., and Lenoir, G. M. Mutations in *BRCA1* and *BRCA2* in breast cancer families: Are there more breast cancer-susceptibility genes? *Am. J. Hum. Genet.* 60, 486-495, 1997.

Serova-Sinilnikova, O. M., Boutrad, L., Stoppa-Lyonnet, D., Bres-sac-de-Paillerets, B., Dubois, V., Lasset, C., Janin, N., Bignon, Y. J., Longy, M., Maugard, C., Lidereau, R., Leroux, D., Frebourg, T., Mazoyer, S., and Lenoir, G. M. *BRCA2* mutations in hereditary breast and ovarian cancer in France. *Am. J. Hum. Genet.* 60, 1236-1239, 1997.

Serrano, M., Hannon, G. J., and Beach, D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* 366, 704-707, 1993.

Sheffield, V. C., Cox, D. R., Lerman, L. S., Myers, R. M. 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. *Proc. Natl. Acad. Sci. USA* 86, 232-236, 1989.

Sherr, C. J. Mammalian G1 cyclins. *Cell* 73, 1059-1065, 1993.

Sidransky, D., Tokino, T., Helzlsouer, K., Zehnbauer, B., Rausch, G., Shelton, B., Prestigiacomo, L., Vogelstein, B., and Davidson, N. Inherited *p53* gene mutations in breast cancer. *Cancer Res.* 52, 2984-2986, 1992.

Simard, J., Feunten, J., Lenoir, G., Tonin, P., Normaud, T., Luu The, V., Vivier, A., Lasko, D., Morgan, K., Rouleau, G. A., Lynch, H., Labrie, F., and Narod, S. Genetic mapping of the breast-ovarian cancer syndrome to a small interval on chromosome 17q12-21: exclusion of candidate genes *EDH17B2* and *RARA*. *Hum. Molec. Genet.* 2, 1193-1199, 1993.

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., McGuire, W. L. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/*neu* oncogene. *Science* 177-182, 1987.

Slamon D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., Ullrich, A., and Press, M. F. Sudies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. *Science* 244, 707-712, 1989.

Slattery, M. L., and Kerber, R. A. A comprehensive evaluation of family history and breast cancer risk. *JAMA* 270, 1563-1568, 1993.

Snedden, W. A., and Fromm, H. Characterisation of the plant homologue of *prohibitin*, a gene associated with antiproliferative activity in mammalian cells. *Plant. Mol. Biol.* 33, 753-756, 1997.

Solomon, E., and Ledbetter, D. H. Report of the Committee on the Genetic Constitution of Chromosome 17. *Cytogenet. Cell Genet.* 58, 686-738, 1991.

Sparkes, R. S., Murphee, A. L., Lingua, R. W., Sparkes, M. C., Field, L. L., Funderburk, S. J., and Benedict, W. F. Gene for hereditary retinoblastoma assigned to human chromosome 13 by linkage to *esterase-D*. *Science* 219, 971-973, 1983.

Stanbridge, E. J. Functional evidence for human tumour suppressor genes: chromosome and molecular genetic studies. *Cancer Surv*. 12, 5-24, 1992.

Stankovic, T., Kidd, A. M. J., Sutcliffe, A., McGuire, G. M., Robinson, P., Weber, P., Bedenham, T., Bradwell, A. R., Easton, D. F., Lennox, G. G., Haites, N., Byrd, P. J., and Taylor, A. M. R. *ATM* mutations and phenotypes in ataxia-telangiectasia families in the British Isles: Expression of mutant *ATM* and the risk of leukemia, lymphoma, and breast cancer. *Am. J. Hum. Genet.*62, 334-345, 1998.

Stark, R., Andre, C., Thierry, D., Cherel, M., Galibert, F., and Gluckman, E. The expression of cytokine and cytokine receptor genes in long-term bone marrow culture in congenital and acquired bone marrow hypoplasias. 83, 560-566, 1993.

Steeg, P. S., Bevilacqua, G., Kopper, L., Thorgeirsson, U. P., Talmadge, J. E., Liotta, L. A., and Sobei, M. E. Evidence for a novel gene associated with low tumour metastatic potential. *J. Natl. Cancer Inst.* 80, 200-204, 1988.

Stein, G. H., Beeson, M., and Gordon, L. Failure to phosphorylate the retinoblastoma gene product in senescent human fibroblasts. *Science* 249, 666-669, 1990.

St. John, D. J., McDermott, F. T., Hopper, J. L., Debney, E. A., Johnson, W. R., and Hughes, E. S. R. Cancer risk in relatives of patients with common colorectal cancer. *Ann. Intern. Med.* 118, 785-790, 1993.

Stoppa-Lyonnet, D., Laurant-Puig, P., Essioux, L., Pages, S., Ithier, G., Ligot, L., Fourqet, A., Salmon, R. J., Clough, K. B., Pouillart, P., the Institut Curie Breast Cancer Group, Bonaiti-Pellie, C., and Thomas, G. *BRCA1* sequence variations in 160 individuals referred to a breast/ovarian family cancer clinic. *Am. J. Hum. Genet.* 60, 1021-1030, 1997.

Strong, L. C., Stine, M., and Norsted, T. L. Cancer in survivors of childhood soft tissue sarcomas and their relatives. *J. Natl. Cancer. Inst.* 79, 1213-1220, 1987.

Swift, M., Morrell, D., Massey, R., and Chase, C. Incidence of cancer in 161 families affected by ataxia-telangiectasia. *New Engl. J. Med.* 325, 1831-1836, 1991.

Szabo, C. I., and King, M. -C. Inherited breast and ovarian cancer. *Hum. Mol. Genet.* 4, 1811-1817, 1995.

Szabo, C. I., and King, M. -C. Population genetics of *BRCA1* and *BRCA2*. *Am. J. Hum. Genet.* 60, 1013-1020, 1997.

Takahashi, H., Behbakht, K., McGovern, P. E., Chiu, H. -C., Couch, F. J., Weber, B. L., Friedman, L. S., King, M. -C., Furusato, M., LiVolsi, V. A., Menzin, A. W., Liu, P. C., Benjamin, I., Morgan, M. A., King, S. A., Rebane, B. A., Cardonick, A., Mikuta, J. J., Rubin, S. C., and Boyd, J. Mutation analysis of the *BRCA1* gene in ovarian cancers. *Cancer Res.* 55, 2998-3002, 1995.

Tanaka, K., Oshimura, M., Kikuchi, R., Seki, M., Hayashi, T., and Miyaki, M. Suppression of tumourigenicity in human colon carcinoma cells by introduction of normal chromosome 5 or 18. *Nature* 349, 340-342, 1991.

Tanner, M. M., Karhu, R. A., Nupponen, N. N., Borg, A., Baldetrop, B., Pejovic, T., Ferno, M., Killander, D., and Isola, J. J. Genetic aberrations in hypodiploid breast cancer: frequent loss of chromosome 4 and amplification of *Cyclin D1* oncogene. *Am. J. Path.* 153, 191-199, 1998.

Tavtigian, S. V., Simard, J., Rommens, J., Couch, F., Shattuck-Eidens, D., Neuhausen, S., Merajver, S., Thorlacius, S., Offit, K., Stoppa-Lyonnet, D., Belanger, C., Bell, R., Berry, S., Bogden, R., Chen, Q., Davis, T., Dumont, M., Frye, C., Hattier, T., Jammulapati, S., Janecki, T., Jiang, P., Kehrer, R., Leblanc, J. -F., Mitchell, J. T., McArthur-Morrison, J., Nguyen, K., Peng, Y., Samson, C., Schroeder, M., Snyder, S. C., Steele, L., Stringfellow, M., Stroup, C., Swedlund, B, Swensen, J., Teng, D., Thomas, A., Tran, T., Tran, T., Tranchant, M., Weaver-Feldhaus, J., Wong, A. K. C., Shizuya, H., Evfjord, J. E., Cannon-Albright, L., Labrie, F., Skolnick, M. H., Weber, B., Kamb, A., and Goldgar, D. E. The complete *BRCA2* gene and mutations in chromosome 13q-linked kindreds. *Nat. Genet.* 12, 333-337, 1996.

Taylor, J. J., and Heasman, P. A. Control genes for reverse transcriptase/polymerase chain reaction (RT-PCR). *Br. J. Haem.* 444-445, 1994.

Terashima, M., Kim, K. -M., Adachi, T., Nielsen, P. J., Reth, M., Kohler, G., and Lamers, M. C. The IgM antigen receptor of B lymphocytes is associated with prohibitin and a prohibitin-related protein. *EMBO J.* 13, 3782-3792, 1994.

Theillet, C., Le Roy, X., De Lapeyriere, O., Grosgeorges, J., Adnane, J., Raynaud, S. D., Simony-Lafontaine, J., Goldfarb, M., Escot, C., and Birnbaum, D. Amplification of *FGF*-related genes in human tumours: possible involvement of *HST* in breast carcinomas. *Oncogene* 4, 915-922, 1989.

Theillet, C., Adnane, J., Szepetowski, P., Simon, M. P., Jeanteur, P., Birnbaum, D., and Gaudray, P. BCL-1 participates in the 11q13 amplification found in breast cancer. *Oncogene* 5, 147-149, 1990.

Thompson, W. E., Sanbuissho, A., Lee, G. Y., and Anderson, E. Steroidogenic acute regulatory (StAR) protein (p25) and prohibitin from cultured rat ovarian granulosa cells. *J. Reprod. Fertil.* 109, 337-348, 1997.

Thor, A. D., Moore, D. H. I., Edgerton, S. M., Kawasaki, E. S., Reihaus, E., Lynch, H. T., Marcus, J. N., Schwartz, L., Chen, L. C., Mayall, B. H., et al., Accumulation of p53 tumour suppressor gene protein: an independent marker of prognosis in breast cancer. *J. Natl. Cancer Inst.* 84, 845-854, 1992.

Thorlacius, S., Olafsdottir, G., Tryggvadottir, L., Neuhausen, S., Jonasson, J. G., Tavtigian, S. V., Tulinius, H., Ogmundsdottir, H. M., and Eyfjord, J. E. A single *BRCA2* mutation in male and female breast cancer families from Iceland with varied cancer phenotypes. *Nat. Genet.* 13, 117-119, 1996.

Thorlacius, S., Struewing, J. P., Hartge, P., Olafsdottir, G. H., Sigvaldason, H., Tryggvadottir, L., Wacholder, S., Tulinius, H., and Evfjord, J. E. Population-based study of risk of breast cancer in carriers of *BRCA2* mutation. *Lancet* 352, 1337-1339, 1998.

Tokino, T., Davidson, N., Helzlsouer, K., Zehnbauer, B., Nakamura, Y., Vogelstein, B., and Sidransky, D. Absence of germline *prohibitin* mutations in early-onset breast cancer. *Internatl. J. Oncol.* 3, 769-772, 1993.

Toyoshima, H., and Hunter, T. p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related to p21. *Cell* 78, 67-74, 1994.

Tsou. H. C., Teng, D. H.-F., Ping, X. L., Brancolini, V., Davis, T., Hu, R., Xie, X. X., Gruener, A. C., Schrager, C. A., Christiano, A. M., Eng, C., Steck, P., Ott, J., Tavtigian, S. V., and Peacocke, M. The role of *MMAC1* mutations in early-onset breast cancer: Causative in association with Cowden syndrome and excluded in *BRCA1*-negative cases. *Am. J. Hum. Genet.* 61, 1036-1043, 1997.

Van de Vijver, M. J. Molecular genetic changes in human breast cancer. *Adv. Cancer Res.* 61, 25-56, 1993.

Vanin, E. F. Processed pseudogenes: Characteristics and evolution. *Annu. Rev. Genet.* 19, 253-272, 1985.

Varley, J. M., Evans, D. G. R., and Birch, J. M. Li-Fraumeni syndrome - a molecular and clinical review. *Br. J. Cancer* 76, 1-14, 1997.

Varley, J. M., McGown, G., Thorncroft, M., Santibanez-Koref, M. F., Kelsey, A. M., Tricker, K. M., Evans, D. G. R., and Birch, J. M. Germline mutations of *TP53* in Li-Fraumeni families: an extended study of 39 families. *Cancer Res.* 57, 3245-3252, 1997.

Vasen, H. F., Mecklin, J. P., Meera Khan, P., and Lynch, H. T. The international collaborative group on hereditary nonpolyposis colorectal cancer (ICG-HNPCC). *Dis. Colon Rectum* 34, 424-425, 1991.

Vogelstein, B., Fearon, E. R., Kern, S. E., Hamilton, S. R., Preisinger, A. C., Nakamura, Y., and White, R. Allelotype of colorectal carcinomas. *Science* 244, 207-211, 1989.
Wallace, D. C. Mitochondrial genetics: a paradigm for aging and degenerative diseases? *Science* 256, 628-632, 1992.

Wallace, D. C. Mitochondrial DNA sequence variation in human evolution and disease. *Proc. Natl. Acad. Sci.*91, 8739-8746, 1994.

Weber, B. L., and Garber, J. E. Family history and breast cancer, probabilities and possibilities. *JAMA* 270, 1602-1603, 1993.

Weber, J. L., and May, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* 44, 388-396, 1989.

White, M. B., Calvalho, M., Derse, D., O'Brien, S. J., and Dean, M. Detecting single base substitutions as heteroduplex polymorphisms. *Genomics* 12, 301-306, 1992.

White, J. J., Ledbetter, D. H., Eddy, R. L., Jr., Shows, T. B., Stewart, D. A., Nuell, M. J., Friedman, V., Wood, C. M., Owens, G. A., McClung, J. K., Danner, D. B., and Morton, C. C. Assignment of the human *prohibitin* gene (*PHB*) to chromosome 17 and identification of a DNA polymorphism. *Genomics* 11, 228-230, 1991.

Winquist, R., Hampton, G. M., Mannermaa, A., Blanco, G., Alavaikko, H., Kiviniemi, H., Taskinen, P. J., Evans, F. A., Newsham, I., and Cavanee, W. K. Loss of heterozygosity for chromosome 11 in primary human breast tumours is associated with poor survival after metastasis. *Cancer Res.* 55, 2660-2664, 1995.

Wood, C. M., White, J. J., Stewart, D. A., Nuell, M. J., Ledbetter, D. H., Eddy, R. L. Jr., Shows, T. B., Morton, C. C., and Danner, D. B. A novel intracellular antiproliferative protein, prohibitin, is the mammalian equivalent of the *D. melanogaster* gene product, *Cc*, essential for development. *Am. J. Hum. Gen.* (supplement). 49, 1991.

Wooster, R., Mangion, J., Eeles, R., Smith, S., Dowsett, M., Averill, D., Barrett-Lee, P., Easton, D. F., Ponder, B. A., and Stratton, M. R. A germline mutation in the androgen receptor in two brothers with breast cancer and Reifenstein syndrome. *Nat. Genet.* 2, 132-134, 1992.

Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., Fields, P., Marshall, G., Narod, S., Lenoir, G., Lynch, H., Feunten, J., Devilee, P., Cornelisse, C. J., Menko, F. H., Daly, P. A., Ormiston, W., McManus, R., Pye, C., Lewis, C. M., Cannon-Albright, L. A., Peto, J., Ponder, B. A. J., Skolnick, M. H., Easton, D. F., Goldgar, D. E., and Stratton, M. R. Localisation of a breast cancer susceptibility gene *BRCA2*, to chromosome 13q12-13. *Science* 265, 2088-2090, 1994.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Patel, S., Rico, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., Gudmundsson, J., Ficenec, D., Kelsell, D., Ford, D., Tonin, P., Bishop, D. T., Spurr, N. K., Ponder, B. A. J., Eeles, R., Peto, J., Devilee, P., Cornelisse, C., Lynch, H., Narod, S., Lenoir, G., Egilsson, V., Barkadottir, R. B., Easton, D. F., Bentley, D. R., Futreal, P. A., Ashworth, A., and Stratton, M. R. Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* 378, 789-792, 1995.

Xiong, Y., Zhang, H., and Beach, D. D type cyclins associate with multiple protein kinases and the DNA replication and repair factor PCNA. *Cell* 71, 505-514, 1992.

Xu, C. -F., Chambers, J., Nicolai, H., Brown, M., Hujeirat, Y., Mohammed, S., Hodgson, S., Kelsell, D. P., Spurr, N. K., Bishop, D. T., and Solomon, E. Mutations and alternative splicing of the *BRCA1* gene in UK breast/ovarian cancer families. *Genes Chromosom. Cancer* 18, 102-110, 1997.

Yandell D. W., Campbell, T. A., Dayton, S. H., Petersen, R., Walton, D., Little, J. B., McConkie-Rosell, A., Buckley, E. G., and Dryja, T. P. Oncogenic point mutations in the human retinoblastoma gene: their applications to genetic counselling. *N. Engl. J. Med.* 321, 1689-1695, 1989.

Yoshida, M. C., Wada, M., Satoh, H., Yoshida, T., Sakamoto, H., Miyagawa, K., Yokota, J., Koda, T., Kakinuma, M., Sugimara, T., and Terada, M. Human *HST1* (*HSTF1*) gene maps to chromosome band, 11q13 and coamplifies with the *INT2* gene in human cancer. *Proc. Natl. Acad. Sci.* USA 85, 4861-4864, 1988.

Zamzami, N., Susin, A., Marchetti, P., Hirsch, T., Gomez-Monterrey, I., Castedo, M., and Kroemer, G. Mitochondrial control of nuclear apoptosis. *J. Exp. Med.* 183, 1533-1544, 1996.

Zhang, H., Xiong, Y., and Beach, D. Proliferating cell nuclear antigen and p21 are components of multiple cell cycle complexes. *Mol. Cell Biol.* 4, 897-906, 1993.

Zhou, D. -J., Ahuja, H., and Cline, M. J. Proto-oncogene abnormalities in human breast cancer: *c-ERBB-2* amplification does not correlate with recurrence of disease. *Oncogene* 4, 105-108,1989.

Zuppan, P. J., Hall, J. M., Lee, M. K., Ponglikitmongkol, M., and King, M. -C. Possible linkage of the oestrogen receptor gene to breast cancer in a family with late-onset disease. *Am. J. Hum. Genet.* 48, 1065-1068, 1991.