MOLECULAR GENETIC STUDY OF HEREDITARY RETINOBLASTOMA

ZERRIN ONADIM

A thesis submitted for the degree of Doctor of Philosophy in the University of London

March 1993

The Imperial Cancer Research Fund Laboratory of Molecular Genetics
Department of Haematology and Oncology, Institute of Child Health
The University of London
ABSABTRACT

This thesis describes a molecular genetic analysis of retinoblastoma (Rb), the commonest ophthalmic malignancy in childhood, that has both familial and non-familial forms. In familial cases the disease is transmitted as an autosomal dominant trait with incomplete penetrance. At the start of the study the gene responsible, RB1, which was also the first tumour suppressor gene identified, was available together with several intragenic probes.

The aim of the study was to analyse mutations in patients predisposed to Rb to gain an insight into the mechanism of mutagenesis and to offer screening for carrier status. The first part of the study involved family linkage analysis with Restriction Fragment Length Polymorphisms (RFLP) and other intragenic polymorphisms of the RB1 gene using Southern blotting techniques and the Polymerase Chain Reaction (PCR). The results indicated that, using the combination of available polymorphisms, it is possible to offer screening to 95 % of Rb families. Using PCR, it was also possible to use formalin-fixed, paraffin-embedded tissue samples from archival material. During the course of this analysis, five pre-natal and six post-natal screenings were carried out the results of which indicated inheritance of the predisposing allele in three cases.

In search of a suitable technique to identify causative mutations in the RB1 gene various techniques were assessed for their usefulness. These techniques were Ribonuclease (RNase) protection, PCR sequencing of individual exons, Carbodiimide (CDI) modified heteroduplex analysis and Single Strand Conformation Polymorphism (SSCP) analysis. SSCP analysis proved to be the most successful in this study and six germ-line mutations were identified in the RB1 gene. Four of the mutations were from typical Rb families with bilaterally affected individuals and the other two from families with "mild" phenotypes. The analyses of these mutations provided insights into the mechanisms of mutagenesis in RB1 and the phenotype-genotype relationship in Rb.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Dr John K Cowell for his guidance, help and support without which this thesis would not have been completed. This project was financed by the David Allen Retinoblastoma Appeal and I would like to record my thanks and appreciation to them.

My thanks are due to Dr J Pritchard, Dr CD Mitchell, Dr F Katz, Dr S Al-Mahdawi, Dr PN Baird, Dr A Michalski, Miss A Hogg and Professor M Pembury for helpful discussions and advice. My special thanks go to Ms Sheila Giles for her patience and invaluable help in typing and preparation of the bibliography.

I would also like to thank Professor OB Eden who supported me in the final stages and through the writing up period and provided encouragement and opportunity for its completion.

I must also thank Mrs S Theodoridou whose friendship and encouragement kept me going through some very difficult periods.

Last but not least my thanks and appreciation go to my husband for the preparation of the graphics and for his patience and support throughout this project.

Zerrin ONADIM
London, 1993
To my Parents

Anne ve Baba'ma
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>1</td>
</tr>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>3</td>
</tr>
<tr>
<td>Dedication</td>
<td>4</td>
</tr>
<tr>
<td>Contents</td>
<td>5</td>
</tr>
<tr>
<td>Abbreviations and symbols</td>
<td>12</td>
</tr>
<tr>
<td><strong>1.0 INTRODUCTION</strong></td>
<td>16</td>
</tr>
<tr>
<td>1.1 Oncogenes</td>
<td>17</td>
</tr>
<tr>
<td>1.2 Knudson's two-hit hypothesis and childhood cancer incidence</td>
<td>24</td>
</tr>
<tr>
<td>1.3 Retinoblastoma</td>
<td>26</td>
</tr>
<tr>
<td>1.3.1 The clinical recognition and description of retinoblastoma</td>
<td>26</td>
</tr>
<tr>
<td>1.3.2 The role of heredity in retinoblastoma</td>
<td>28</td>
</tr>
<tr>
<td>1.3.3 Definition, histology and cell of origin of retinoblastoma</td>
<td>29</td>
</tr>
<tr>
<td>1.3.4 Spontaneous regression of retinoblastoma</td>
<td>32</td>
</tr>
<tr>
<td>1.3.5 The incidence and age distribution of retinoblastoma</td>
<td>33</td>
</tr>
<tr>
<td>1.3.6 Treatment of retinoblastoma</td>
<td>35</td>
</tr>
<tr>
<td>1.3.7 Survival rates in retinoblastoma</td>
<td>38</td>
</tr>
<tr>
<td>1.3.8 Second primary cancers in retinoblastoma</td>
<td>39</td>
</tr>
<tr>
<td>1.3.9 Trilateral retinoblastoma</td>
<td>43</td>
</tr>
<tr>
<td>1.4 Retinoblastoma - the gene</td>
<td>46</td>
</tr>
<tr>
<td>1.4.1 The location of the retinoblastoma gene</td>
<td>46</td>
</tr>
<tr>
<td>1.4.1.1 The cytogenetic evidence</td>
<td>46</td>
</tr>
<tr>
<td>1.4.1.2 Esterase D studies</td>
<td>48</td>
</tr>
<tr>
<td>1.4.1.3 The retinoblastoma gene is recessive</td>
<td>50</td>
</tr>
<tr>
<td>1.4.2 Cloning of the retinoblastoma gene</td>
<td>54</td>
</tr>
<tr>
<td>1.4.3 RB1 involvement in other tumours</td>
<td>57</td>
</tr>
<tr>
<td>1.4.4 Structure and genomic organisation of RB1</td>
<td>57</td>
</tr>
<tr>
<td>1.4.5 Function of RB1</td>
<td>60</td>
</tr>
<tr>
<td>1.4.6 RB1 - linkage and clinical applications</td>
<td>62</td>
</tr>
<tr>
<td>1.4.6.1 Linkage analysis - theory and applications</td>
<td>62</td>
</tr>
<tr>
<td>1.4.6.2 RB1, the clinical applications</td>
<td>66</td>
</tr>
<tr>
<td>1.4.7 The pattern of inheritance and variable phenotypic expression in retinoblastoma</td>
<td>67</td>
</tr>
<tr>
<td>1.5 Tumour suppressor genes</td>
<td>72</td>
</tr>
</tbody>
</table>
2.0 METHODS ............................................................................................ 82
  2.1 Preparation of solutions and buffers ...................................................... 83
  2.2 Enzymes ................................................................................................. 83
  2.3 Plasmids ................................................................................................. 84
  2.4 Preparation and handling of bacteria .................................................... 84
    2.4.1 Transformation of bacteria ........................................................ 84
  2.5 Isolation of plasmid DNA .................................................................. 85
    2.5.1 Small scale isolation ................................................................ 85
    2.5.2 Large scale isolation ................................................................ 88
  2.6 Isolation of DNA from human cells .................................................. 90
    2.6.1 Isolation of DNA from whole blood ....................................... 90
    2.6.2 Isolation of DNA from white blood cells and lymphoblastoid cell lines ............................................................. 90
    2.6.3 Isolation of DNA from chorionic villi ..................................... 91
    2.6.4 Extraction of DNA from paraffin-embedded tissue sections... 91
  2.7 Isolation of total RNA from lymphoblastoid cell lines ..................... 92
  2.8 Restriction enzyme digestion .............................................................. 93
  2.9 Agarose gel electrophoresis ................................................................ 93
    2.9.1 Photography ............................................................................. 94
  2.10 Polyacrylamide gel electrophoresis ................................................... 94
  2.11 Molecular weight markers .................................................................. 96
  2.12 Size estimation of DNA ...................................................................... 97
  2.13 Extraction of DNA from agarose gels ............................................... 97
  2.14 Oligo-primer extension labelling of DNA ......................................... 98
  2.15 Southern blotting ..................................................................................100
    2.15.1 DNA hybridisation ................................................................. 102
  2.16 RNase protection ................................................................................. 103
    2.16.1 Generation of RNA probes ...................................................... 103
    2.16.2 Purification of riboprobes ....................................................... 105
    2.16.3 Hybridisation and RNase digestion ......................................... 106
  2.17 Polymerase chain reaction (PCR) ........................................................107
  2.18 PCR based detection of polymorphic sites .......................................... 109
  2.19 Carbodiimide modification of heteroduplexes of PCR product and primer extension .............................................................. 109
    2.19.1 Generation of heteroduplexes .................................................. 110
    2.19.2 CDI modification of heteroduplexes ....................................... 110
    2.19.3 Primer extension .................................................................... 111
  2.20 Single strand conformation polymorphism (SSCP) ......................... 112
2.21 Direct sequencing of PCR products ......................................................112
  2.21.1 Separation of DNA strands .....................................................112
  2.21.2 Dideoxy sequencing .................................................................113

3.0 RESULTS ..............................................................................................115
3.1 Application of intragenic polymorphisms in screening for mutant
gene carrier status in retinoblastoma families .....................................116
  3.1.1 Studies using restriction fragment length polymorphisms .........117
  3.1.2 Analyses of families with the RB1.20 and RB1.3
      polymorphisms ...........................................................................125
  3.1.3 Lod score analysis of family linkage data ..................................128
  3.1.4 Application of PCR analysis to paraffin-embedded tissue
      sections .....................................................................................132
  3.1.5 Pre- and peri-natal screening using intragenic RB1
      polymorphisms ........................................................................137
  3.1.6 Summary of phenotypic and genotypic data of Rb families
      in this series .............................................................................150
  3.1.7 Expression of the Rb phenotype ...............................................151
3.2 Assessment of techniques to detect mutations .......................................175
  3.2.1 RNase protection experiments ..................................................175
  3.2.2 PCR amplification of individual exons sequencing ..................181
  3.2.3 Detection of single base mismatches by CDI modification of
       DNA heteroduplexes followed by primer extension ...............190
  3.2.4 PCR - SSCP analysis ...............................................................195
3.3 Identification of oncogenic mutations in SSCP analysed DNA by
   PCR-sequencing ...........................................................................204
  3.3.1 Mutations identified in bilateral Rb families .........................204
  3.3.2 Mutations identified in families showing incomplete
       penetrance and mild expression of the Rb phenotype .........218

4.0 DISCUSSION ......................................................................................229
4.1 Genetic screening in retinoblastoma families .......................................230
  4.1.1 Manifestation of the Rb phenotype ...........................................242
4.2 Mutation detection techniques for the identification of the RB1
   gene mutations ...........................................................................246
4.3 Mutations of the RB1 gene ................................................................257
  4.3.1 The physical changes in RB1 ..................................................257
  4.3.2 Possible mechanisms of mutagenesis in RB1 .........................266
4.3.3 The role of pRB in tumorigenesis and possible functional consequences of RB1 mutations ........................................... 273

References ...................................................................................................... 285

Appendix I Restriction endonucleases ..................................................... 348
Appendix II Molecular weight markers .................................................. 349
Appendix III Pedigrees of Rb families .................................................... 350
Appendix IV Results of Rb screening in 57 Rb families ......................... 353
Appendix V DNA probes used for identification of RB1 RFLPs .......... 355
Appendix VI The Genetic Code ............................................................... 356
Publications ............................................................................................... 357
LIST OF FIGURES

1.1 Chromosomal mechanisms of a second event that will allow the phenotypic expression of a recessive mutant allele at the Rb locus... 52
1.2 A schematic representation of the RB1 gene .......................... 58
3.1 A schematic representation of the RB1 gene showing the approximate, relative positions of polymorphic sites ............... 117
3.2 Examples of the polymorphic banding profiles for RFLPs .......... 120
3.3 PCR based detection of RFLPs ............................................. 122
3.4 DNA sequence of the 50 to 53 bp repeat of the VNTR (Intron 17) detected by the RS2.0 probe ........................................... 124
3.5 Representative gels showing the RB1.20 VNTR and the RB1.3 single base pair polymorphism ........................................ 127
3.6 Segregation of RS2.0 alleles in families RBF12 and RBF04 ....... 131
3.7 Linkage analysis of family RBF58 using PCR and DNA extracted from formalin fixed paraffin-embedded tissue sections .......... 134
3.8 Presymptomatic screening in families RBF06 and RBF13 ...... 138
3.9 Presymptomatic screening in families RBF29 and RBF32 ......... 141
3.10 Presymptomatic screening in families RBF33 and RBF41 ....... 143
3.11 Linkage analysis in family RBF34 ........................................... 146
3.12 Pre-natal screening in family RBF58 ..................................... 147
3.13 Segregation of RS2.0 alleles in RBF29 and R0.6 in RBF18 ....... 152
3.14 Spontaneously regressed Rbs (retinomas) ............................. 154
3.15 Pedigrees of and autoradiographs from families RBF02, RBF38 and the pedigree of RBF41 .......................................... 157
3.16 Two families, RBF32 and RBF56 with unilaterally affected or unaffected gene carrier members ................................. 161
3.17 Family pedigrees and the segregation of alleles with informative polymorphisms in RBF26 and RBF55 ......................... 164
3.18 Analysis of family RBF01 using a variety of chromosome 13 probes ............................................................... 167
3.19 Schematic representation of haplotype analysis on chromosome 13 in RBF01 ............................................................... 170
3.20 Pedigrees of and informative polymorphism autoradiographs and/or agarose gels from RBF24 and RBF31 ......................... 173
3.21 Typical agarose gel analysis of total cellular RNA ................. 176
3.22 Autoradiographs of two RNase protection gels ........................ 179
3.23 Agarose gels showing the PCR amplification products for a variety of RB1 exons and their flanking intron sequences ..................... 185
3.24 DNA sequence ladders from those regions in introns 14, 15 and 19 of the RB1 gene which differed from the published sequence ...... 187
3.25 Schematic diagram summarising the CDI modification procedure ... 191
3.26 Detection of T-G mismatches using the RB1.3 polymorphism ...... 193
3.27 SSCP analysis of exons 13 and 17 of the RB1 gene .................... 197
3.28 SSCP analysis of exon 20 of the RB1 gene .............................. 200
3.29 SSCP analysis of the RB1.3 region of the RB1 gene ................... 203
3.30 Sequence analysis of exon 3 from family RBF64 ...................... 206
3.31 Sequence analysis of exon 13 in family RBF58 ....................... 209
3.32 Sequence analysis of exon 17 in family RBF59 ....................... 212
3.33 Sequence analysis of exon 20 in family RBF62 ....................... 215
3.34 Restriction enzyme analysis of the mutation in exon 20 from family RBF62 ................................................................. 217
3.35 Sequence analysis of exon 20 in family RBF29 ....................... 220
3.36 Restriction enzyme analysis of the mutation in exon 20 from family RBF29 ................................................................. 222
3.37 Sequence analysis of exon 20 in family RBF18 ....................... 224
3.38 Nucleotide sequence of the coding strand of exon 20 and its flanking intron sequences ....................................................... 226

LIST OF TABLES

1.1 Comparison of treatment in the UK for Rb between periods
1.2 Survival rates for Rb in the Western countries .......................... 39
1.3 The incidence of second cancers in Rb patients ......................... 41
1.4 The cumulative incidence rates for second cancers in Rb patients calculated using life-tables .................................................. 42
1.5 Some of the phenotypic features observed in the "D-deletion"
syndrome patients ...................................................................... 46
2.1 Solutions used in the isolation of DNA and RNA ...................... 87
2.2 General buffers ........................................................................ 89
2.3 Solutions used in DNA labelling .............................................. 99
2.4 Solutions used in Southern blotting and DNA hybridisation ............ 101
2.5 Solutions used in the generation and purification of riboprobes ...... 104
2.6 Solutions used in RNase protection assays ...................................... 106
2.7 Reagents and conditions of PCR ...................................................... 108
2.8 Solutions used in heteroduplex formation, CDI modification and
    primer extension ................................................................................... 111
2.9 Solutions used in separation of DNA strands and dideoxy
    sequencing ............................................................................................. 114
3.1 Characteristics of the RFLP DNA probes with their allele
    frequencies .............................................................................................. 118
3.2 Description of the primers used for PCR amplification of the
    polymorphic sites in the RB1 gene ...................................................... 123
3.3 Summary of informativity of the polymorphic sites in Rb family
    studies ................................................................................................... 125
3.4 Lod scores calculated for linkage between the Rb phenotype and
    RB1 intragenic probes for 21 Rb families ............................................ 129
3.5 Lod scores at different values of θ at 90% penetrance for the two
    families that showed incomplete penetrance ........................................ 132
3.6 The results of five pre-natal and six post-natal screenings carried
    out over a period of 3.5 years (Jan 1989 - Jun 1992) involving
    eight different families .......................................................................... 149
3.7 The Rb phenotype of mutant RB1 gene carriers in 57 Rb families ... 150
3.8 Details of the RB1 subclones used to generate riboprobes and
    transcription vectors they were cloned into ........................................ 177
3.9 Details of the oligonucleotide primers designed for amplification of
    individual exon and flanking intron sequences of the RB1 gene .... 183
3.10 Details of internal primers used in PCR-sequencing of exons 17
    and 20 .................................................................................................. 204
3.11 Summary of mutations identified in the RB1 gene ....................... 228
4.1 Summary of some of the reported RB1 gene mutations and their
    consequences ......................................................................................... 262
ABBREVIATIONS AND SYMBOLS

Nucleic acids and their constituents

G, A, T, C, U  The bases guanine, adenine, thymine, cytosine, uracil
NTP (GTP, ATP etc) The nucleoside triphosphates; guanine triphosphate etc
dNTP (dGTP etc)  The deoxynucleoside triphosphates
ddNTP (ddGTP etc) The dideoxynucleoside triphosphates
bp Base pairs
kb Kilobase pairs
DNA Deoxyribonucleic acid
cDNA Complementary deoxyribonucleic acid
RNA Ribonucleic acid
mRNA Messenger RNA
tRNA Transfer RNA

Chromosomes

X, Y The sex chromosomes
1 - 22 The autosomes
p The short arm
q The long arm

Reagents/chemicals/enzymes/media

BSA Bovine serum albumen
CDI 1-cyclohexyl-3-(2-morpholino-ethyl) carbodiimide
metho p-toluene sulfonate
DEPC Diethyl pyrocarbonate
DMSO Dimethyl sulphoxide
DNase Deoxyribonuclease
DTT Dithiothreitol
EDTA Ethylene diamine tetra-acetic acid
HEPES N-2-hydroxyethylene piperazine
H₂O Water
HCl Hydrochloric acid
KCl Potassium chloride
L-agar Luria agar
LB  Luria broth
MgCl₂  Magnesium chloride
MgSO₄  Magnesium sulphate
MOPS  3-(N-morpholino)-propane sulphonic acid
NaCl  Sodium chloride
NaOH  Sodium hydroxide
Na₃C₆H₅O₇  Trisodium citrate
NH₄OAc  Ammonium acetate
-OH  Hydroxyl group
³²P  Phosphorus-32
PBSA  Phosphate buffered saline
PIPES  Piperazine-N,N-bis [2-ethane sulphonic acid]
PVP  Polyvinylpyrrollidone
RNase  Ribonuclease
RPMI  Roswell Park Memorial Institute
³⁵S  Sulphur-35
SDS  Sodium dodecyl sulphate (lauryl sulphate)
SSC  Sodium salt citrate
TAE  Tris acetate
TBE  Tris borate
TE  Tris-EDTA-buffer
TEMED  N,N,N',N'-tetramethyl ethylene diamine
Tris  Tris-hydroxymethyl-aminomethane

Amino acids

Ala  Alanine
Arg  Arginine
Asn  Asparagine
Asp  Aspartic acid
Cys  Cysteine
Gln  Glutamine
Glu  Glutamic acid
Gly  Glycine
His  Histidine
Ile  Isoleucine
Leu  Leucine
Lys  Lysine
Met  Methionine
Phe  Phenylalanine
Pro  Proline
Ser  Serine
Thr  Threonine
Trp  Tryptophan
Tyr  Tyrosine
Val  Valine

Units

kD  Kilodalton
cpm  Counts per minute
rpm  Revolutions per minute
°C  Degrees centigrade
l/ml/μl  Litre/millilitre/microlitre
g/mg/μg  Gram/milligram/microgram
X⁻¹ (eg ml⁻¹)  Per X (eg per millilitre)
A/mA  Ampere/milliampere
M/mM  Molar/millimolar
Ci/mCi/μCi  Curie/millicurie/microcurie
(v/v)  Volume to volume ratio
(w/v)  Weight to volume ratio

Other

Rb  Retinoblastoma
pRB  Retinoblastoma protein
PCR  Polymerase Chain Reaction
RFLP  Restriction Fragment Length Polymorphism
SSCP  Single Strand Conformation Polymorphism
α  Alpha
β  Beta
γ  Gamma
θ, θ  Theta
λ  Lambda
Pedigree symbols

Male  Female

Unaffected

Bilaterally affected

Unilateral tumour, right eye

Unilateral tumour, left eye

Bilateral regressed tumours

Regressed tumour, right eye

Regressed tumour, left eye

Deceased

Fetus / sex unknown
1.0 INTRODUCTION
1.1 ONCOGENES

Most, if not all, cancer cells contain genetic damage that appears to lie at the heart of tumorigenesis (Bishop 1987, 1989). This now well known fact was first recognised by Boveri in 1914 who proposed that cancer resulted from a disturbance of the genetic material. The genetic disturbance does not happen in a single step, however, and the process resulting in tumorigenesis is complex and occurs in a gradual and multistep fashion (Foulds 1958, Nowell 1976). In experimental models, tumorigenesis has been broken down into at least three steps: initiation, promotion and progression (Farber and Cameron 1980). Each of these steps represents a physiological barrier that must be breached in order for a cell to progress further toward the end point of malignancy. In an attempt to understand these events tumour cells have been studied extensively for the type of genetic changes they carry. Two particular kinds of genetic changes result in the cancer phenotype: dominant, where usually the targets are genes which promote cancer, and recessive, where the targets are genes whose normal function prevents cancer. Dominant mutations typically result in a gain of gene function, whereas the recessive mutations cause loss of function.

The initial insight into cellular oncogenes and the oncogene paradigm developed over the past decade with the discovery of a group of viruses that are capable of inducing cancers. The first was the Rous Sarcoma virus (RSV) found in 1911 by Peyton Rous (reviewed in Varmus 1989). In 1969, Huebner and Todaro developed the viogene-oncogene hypothesis and established that oncogenes carried by RNA tumour viruses could cause cancer. RSV is one of many acutely transforming retroviruses which have been studied extensively because of their unusual molecular biology involving reverse transcription and the high efficiency integration of their genomes into mammalian chromosomes. One other unusual trait of these viruses, namely, the ability to pick up and transduce cellular genes, opened up the study of cellular oncogenes (reviewed in Varmus 1989). The RSV genome has two distinct parts: the first part is composed of genes responsible for viral replication involving the processes of reverse transcription, integration and progeny virus formation and the second part contains the src gene, which enables the virus to induce sarcomas in vivo and to transform chicken fibroblasts in monolayer culture. In 1977 the product of the src gene was immunoprecipitated and termed pp60\textsuperscript{src} (Brugge and Erikson 1977). This
protein was identified as a protein kinase (Collett and Erikson 1978). Stehelin et al (1976) showed that the src oncogene arose from a closely related endogenous gene residing in the chicken genome. This antecedent gene, termed a proto-oncogene is a normal cellular gene and an integral part of the chicken genome. This work proved the existence in the cellular genome of genes, "proto-oncogenes", that can exhibit strong transforming properties when properly activated. Oncogenic retroviruses, by virtue of their tendency to transduce cellular sequences, have unveiled more than 30 potential cellular oncogenes. These viruses were identified in isolates from tumours derived from chicken, turkeys, mice, rats, cats and monkeys. Among the many proto-oncogenes discovered in this way, is the myc proto-oncogene, for example, the product of which was later localised in the nucleus (Abrams et al 1982).

A second group of cell-associated oncogenes were identified using gene-transfer experiments (Cooper 1982, Weinberg 1985). The initial experiments were designed to demonstrate the presence of molecular determinants that were responsible for transformation of cells exposed to chemical carcinogens. This involved introducing DNA extracted from chemically transformed cells into appropriate, untransformed recipients. The host cells traditionally used were mouse 3T3 cells which were immortal in culture but showed cell contact inhibition of growth. Transformed 3T3 cells pile up in culture and form distinct foci. When DNA from different human tumour cell lines and biopsies were introduced into 3T3 cells, transformed foci were observed (Cooper 1982). DNA from normal cells, however, lacked such detectable transforming activity. In these experiments, typically, 10-25% of human tumour DNAs assayed were able to induce transformed foci. When the gene responsible was isolated the majority were shown to be activated cellular homologs of the viral ras genes (Balmain 1985). Occasionally, however, other activated genes were detected in these assays which were also identified as cellular homologs of retroviral oncogenes, eg raf (Fukui et al 1985). Novel cellular oncogenes (eg B-lym, mel) were also detected in this way (Diamond et al 1983, Padua et al 1984), for which there does not appear to be a viral counterpart. The similarities between the retroviral oncogenes and those discovered in transfection assays suggested that certain cellular proto-oncogenes can become activated in two ways. They may become associated with retroviruses or they may become altered via mutational events that depend on non-viral mechanisms (Land et al 1983b). One of the best
examples is the ras oncogene. The presence of activated ras genes (H-ras, K-ras) in human tumours was first identified in a human bladder carcinoma and a human lung carcinoma cell lines (Der et al 1982). Subsequently, a third activated ras gene, N-ras, was detected in a human neuroblastoma cell line (Shimizu et al 1983).

The proto-oncogenes and the proteins that they specify form a structurally and functionally heterogeneous group and, therefore, it was only expected that various molecular mechanisms are involved in their activation. Among these mechanisms are over-expression of a proto-oncogene following acquisition of a novel transcriptional promoter (Hayward et al 1981); over-expression due to amplification of a proto-oncogene (Alitalo 1984); increase in transcription under the influence of enhancer sequences (Payne et al 1982); deregulation of a proto-oncogene due to chromosomal rearrangements (Rowley 1983) and alterations in the structure of a proto-oncogene (Tabin et al 1982). There are many examples for each of these mechanisms and any one proto-oncogene can be found to be activated by different mechanisms in different tumours. For example, various types of chromosomal rearrangements have shown to activate the myc gene (reviewed by Fahrlander and Marcu 1986) and amplification of the N-myc and L-myc genes has been found in neuroblastomas (Schwab et al 1983) and in small cell lung tumours (Nau et al 1985) respectively. In chronic myelocytic leukaemias, the t(9;22) translocation results in the shifting of the c-abl gene from chromosome 9 to the bcr (breakpoint cluster region) of chromosome 22 forming a fused bcr/abl gene (Heisterkamp et al 1983, Groffen et al 1984). The resulting fusion protein was found to possess a tyrosine kinase activity not present in the normal c-abl encoded protein but similar to that of the protein product of the transforming viral abl gene (Konopka et al 1984). Transforming ability as a result of structural alterations of an oncogene is best documented for the p21 oncogene proteins encoded by the ras genes. In the human bladder carcinoma cell line, T24/EJ, a point mutation converted the H-ras proto-oncogene into a potent oncogene. This mutation was a G to T transversion converting a glycine, normally present as the 12th residue of the 21000 dalton p21 protein, to a valine residue (Tabin et al 1982, Reddy et al 1982). In a human lung carcinoma, on the other hand, H-ras was found to carry a mutation affecting amino acid 61 of the p21 protein (Yuasa et al 1983). Indeed the vast majority of ras mutations are shown to occur at positions 12,
13 and 61 which turned out to be crucial in the function of the protein (Barbacid 1987).

Although the study of cellular oncogenes and the molecular mechanisms involved in their activation led to a better understanding of these agents of cellular transformation, it revealed little about the complex processes of tumorigenesis in vivo. The creation of a tumour cell within a tissue would seem to require far more that the activation of one of these oncogenes within the cell. Spontaneous or chemically induced tumorigenesis was known to be a multistep process (Foulds 1958), while the activation of an oncogene such as H-ras seemed to occur as a single, discrete event. This discrepancy led to a suspicion that activation of an oncogene may represent only one component of a multistep process.

An initial connection between oncogenes and the multistep nature of tumorigenesis was made through studies of two viral oncogenes, the middle T (MT) and large T (LT) genes of polyomavirus (Rassoulzadegan et al 1982). Neither was found to be able to transform rat embryo fibroblasts on its own. The two, working in collaboration, however, gave rise to a fully tumorigenic phenotype. This work suggested that each oncogene was specialised to induce part of the phenotype required for full transformation. With the extension of this model to a number of cellular oncogenes, it was soon demonstrated that neither a ras nor a myc oncogene was able to induce full transformation, while the two, co-introduced into rat embryo fibroblasts, achieved full transformation (Land et al 1983a). At the same time, a ras oncogene was found to collaborate with the adenovirus E1A oncogene in the full transformation of baby rat kidney cells (Ruley 1983). Further investigation of the collaborative transformation process showed that each oncogene acts in a distinct, complementary way on the cellular phenotype. In rodent cells, ras oncoproteins induced refractility, anchorage independence and growth factor secretion - expression that did not favour immortalisation in culture - whereas, the myc oncoproteins appeared more adept at immortalisation and less able to induce anchorage independence and growth factor secretion (Weinberg 1985, Land et al 1986). With more characterisation specific combinations of oncogenes were identified, which could cooperate in transforming normal cells, which placed them into two functional categories based on the abilities to complement either a ras or a myc oncogene in transformation assays of rat embryo cells. It was later...
found that those oncogenes that function like ras encode cytoplasmic proteins while those that function like myc specify nuclear proteins (Weinberg 1985). The "nuclear" category includes myc, N-myc, L-myc, myb, mutant p53 (see Section 1.5), fos, jun, adenovirus E1A, polyomavirus and SV40 LT, papillomavirus E7 and tax of human T cell leukaemia virus type 1. The "cytoplasmic" category includes H-ras, K-ras, N-ras, src, erbB, fps, ros, yes, sea and polyomavirus MT. Not all oncogenes, however, fit neatly in this scheme pointing to the possible existence of other pathways or targets that are not addressed by the ras/myc paradigm.

The collaboration observed between oncogenes gave rise to the rule that at least two activated oncogenes are required to transform normal cells into fully malignant ones, although some deviations from this scheme have also been observed, most frequently in transfection of ras-like oncogenes into established cell types like Rat-1 cells or NIH 3T3 cells (Lowy et al 1978). The reason for this apparent violation, however, lay with the fact that such cells deviate from primary embryo cells in their established, immortilised phenotype representing a pre-malignant state. Other, apparently contradictory, results were observed when a ras oncogene was co-introduced with a neomycin resistance marker into embryo fibroblast monolayers and also when embryo fibroblast monolayers were infected with a retrovirus which transduces a ras oncogene (Land et al 1986, Land 1986). In the first experiment, subsequent application of neomycin resulted in killing the great majority of cells in the culture and the descendants of the oncogene bearing neomycin-resistant transfected cells developed into tumorigenic cells. In the latter case, when the virus spread through the monolayer infecting the great majority of cells, full transformation was observed. Results like these suggested that neighbouring normal cells exert a normalising or inhibitory influence on the growth of ras transformants and drew attention to the importance of the environment of the oncogene-bearing cell (Weinberg 1989). This made the simple and important point that the growth properties of a cell depend not only on its own genotype (eg its complement of oncogenes) but on its environment as well. This point, in turn, emphasised the importance of a critical early step in tumorigenesis, a step involving early pre-neoplastic cell clone expansion in spite of the inhibitory influences of normal neighbours. Activated ras oncogenes have been found in a number of pre-neoplastic murine and human tumour models suggesting that ras activation might be a relatively early event in tumour formation (Balmain et al 1984, Liu et al
1987, Vogelstein et al 1988). It has been suggested that initiating carcinogens create a critical genetic change, for example, involving a \textit{ras} gene. Any resulting initiated cells may then expand clonally under the influence of a tumour promoter, for example TPA (12-0-tetradecanoylphorbol-13-acetate) which acts to confer special growth advantages on \textit{ras}-bearing cells until they form a large enough clone of descendants to permit the occurrence of low probability, secondary genetic changes (Weinberg 1989). These secondary genetic changes create alleles that collaborate with the initially induced \textit{ras} oncogene to produce a fully tumorigenic cell that is no longer dependent on the promoting agent for its continued growth (Moolgavkar and Knudson 1981, Quintanilla et al 1986). Obvious candidates for the collaborator oncogenes were the \textit{myc}-like oncogenes or any other gene(s) that can, when mutated, create phenocopies of the \textit{myc}-induced state, i.e. the required changes in cell physiology in the nucleus (Weinberg 1989).

Although the physiological effects of \textit{myc} and \textit{myc}-like oncogenes seemed to be the immortalisation of the cell, this trait might not be central to their role in tumorigenesis. If \textit{ras}-transformed cells are unable to induce foci in the presence of normal neighbours while \textit{ras+myc} cells grow strongly, then one mechanism of action of \textit{myc} is that \textit{myc} enables \textit{ras} transformants to ignore or override the inhibitory influences of normal neighbouring cells. This draws attention to the importance of negative regulatory mechanisms in the cell. Negative environmental influences may encompass signals involved in maintaining contact inhibition, signals through which growth-inhibitory hormones like interferon and transforming growth factor \(\beta\) (TGF-\(\beta\)) shut down cell growth, indeed even the signals used by normal cells to shut down the growth of \textit{ras} transformants. It was suggested that the importance of \textit{myc}-like genes may be that these genes can remain active in spite of the presence of antitumorigenic influences that normally operate to shut down their expression (Weinberg 1989). What allows \textit{myc}-like oncogenes behave in this way? It was observed that the tumours induced by \textit{ras+myc} grew to a substantial size and then stopped growing. In contrast the \textit{ras+LT} tumours grew until they killed the host animal (Land et al 1983b). This suggested that LT contributes multiple functions for transformation only one of which corresponds to a function provided by \textit{myc}. A tumour cell may thus require additional functions - a third type of oncogene function - beyond those provided by the \textit{ras} and \textit{myc}-like genes. This theory was reinforced by the
fact that "dominantly" acting oncogenes could only be found in 10 to 30% of human cancers. Another reinforcing factor was the fact that the expanding list of oncogene functions included; growth factors, eg v-sis / platelet-derived growth factor (Doolittle et al 1983), growth factor receptors, eg v-erbB / epidermal growth factor receptor (Downward et al 1984), signal transducers, eg ras / GTP-binding proteins (Hurley et al 1984), protein kinases, eg src / tyrosine kinase (Collett and Erikson 1978) and transcription activators, eg jun / GCN4, AP-1 (Vogt et al 1987, Bohmann et al 1987). It seemed plausible, therefore, that a third kind of oncogenic function is the inactivation of negative regulatory pathways. This could be achieved either by deletion or inactivation of cis-acting negative regulatory sequences or by other types of damage that may serve to delete from the cell the diffusible trans-acting factors that normally act to mediate shut down of genes like myc (Weinberg 1989). Tumour cell genomes may, therefore, contain activated oncogenes coexisting with inactivated versions of negative regulatory genes, the two sets of changes collaborating to confer tumorigenicity.

The existence of negative regulatory genes was first suggested in 1969 as a result of cell-fusion experiments by Harris et al (1969). Using somatic cell hybrids they showed that, when malignant cells were fused with normal cells, the resulting hybrid often lost its malignant phenotype. Tumorigenic revertants could emerge, however, which had lost specific chromosomes suggesting that they carried genes which could suppress the malignant phenotype - tumour suppressor genes (see Section 1.5). Thus, at the cellular level the genes responsible for the malignant phenotype were recessive (Stanbridge et al 1981). It was, in fact, the study of rare childhood tumours which established the true nature of recessive oncogenes, the paradigm being for the children's eye cancer - retinoblastoma. In 1971, Knudson had already predicted that the loss of both copies of such genes might be causal in certain types of cancer, particularly Rb, postulating his now classical "two-hit" theory.
1.2 KNUDSON'S TWO-HIT HYPOTHESIS AND CHILDHOOD CANCER INCIDENCE

The so-called "hit theory" was developed mainly in the 1930s and 1940s in radiation biology (Timofeeff-Ressovsky and Zimmer 1947, cited in Vogel 1979, 1992). This concept was adapted to cancer research (Armitage and Doll 1954, 1957) suggesting the idea that cancers are formed not by single mutations but as a result of multiple, sequentially occurring events. What was lacking, however, was direct evidence that cancer could arise in as few as two steps and that each could occur at a rate that was compatible with accepted values for mutation rates. Knudson, in 1971, proposed that Rb is caused by two mutational events (hits). These mutations can be any disruption of the genetic material including chromosome aberrations such as translocations and deletions. His theory states that Rb, and some other childhood cancers, originate after two successive mutations, of which the second mutation is always somatic and the first mutation may be germinal (hereditary cases) or somatic (non-hereditary cases). The fact that at least one other mutation was necessary for tumour formation stemmed from the fact that not all retinal cells in children carrying a germline mutation develop into tumours and that, on occasion, tumours fail to develop in some patients at all. There is also a significant difference between the mean age at diagnosis of tumours between the hereditary and sporadic cases, the former presenting earlier than the latter (see section 1.3.5). Bilateral cases not yet diagnosed by a given age were plotted against age at diagnosis and exhibited a linear decline on a semi-logarithmic plot, the slope of which was consistent with a single event (Knudson 1971). In contrast, the unilateral (mostly non-hereditary) cases followed a curvilinear distribution comparable with two events. If the second mutation in hereditary cases was a random event the number of tumours that develop in each eye should fit a Poisson distribution. Knudson (1971) was able to show that this in fact occurs. This theory explains the occasional gene carrier who gets no tumour, those who develop only unilateral tumours, those who develop bilateral tumours and instances of multiple tumours in one eye. In 1973, Comings extended Knudson's hypothesis by proposing that two hits served to inactivate both alleles of a single gene that essentially functioned to suppress Rb formation.

In 1975, Knudson, Hethcote and Brown, using Poisson distribution analysis, developed a probabilistic model for hereditary cases of Rb that described its
age-dependent incidence. Subsequently, a quantitative general model was developed by Hethcote and Knudson (1978) using ages at diagnosis for both hereditary and non-hereditary cases of all embryonal cancers. This model related age-specific incidence data explicitly to cellular processes.

Knudson also speculated about the nature of the second hit and proposed that, in both hereditary and non-hereditary cases, the second hit results in the loss or mutation of both copies of a particular gene by one of four mechanisms; local mutation, chromosomal deletion, non-disjunction and resultant monosomy 13, or somatic recombination (Knudson 1978). According to this hypothesis, and in agreement with Comings' proposal (1973), the Rb gene was recessive, at the cellular level, although dominant (because of the large probability of the second event) with respect to imparting susceptibility to the tumour. The formal proof of this hypothesis came in 1984 with the evidence suggesting that the Rb gene is indeed recessive at the cellular level (see Section 1.4.1.3). After the cloning of the Rb gene (see Section 1.4.2) mutations were also identified in both of its copies, providing the final proof of the 'two-hit' theory.

Although the two-hit theory of tumorigenesis is now widely accepted as the best explanation of observations in Rb, certain clinical irregularities still exist. For example, the existence of families with affected sibs born to unaffected parents and the incidence of families with Rb in distant relatives led others to postulate different models for retinoblastoma (see Section 1.4.7).

In their general model of embryonal cancers, Hethcote and Knudson (1978) proposed that the total incidence of tumours is a function of the initial number of differentiating cells, the number of cell divisions associated with neoplastic transformation and the rates of germinal and somatic mutations involved. This implied that there should be a minimal incidence of cancer, which, in the absence of demographic variation in mutation rates, should be constant from one part of the world to another. In fact this is the case for all childhood cancers except for leukaemia and lymphoma (Knudson 1976). The background mutation rates can be increased by environmental mutagens (Knudson 1978). This fact seems to be especially important in childhood cancers when X-irradiation occurs pre-natally. The recent report by Morris et al (1990) demonstrating an increased incidence of Rb in the areas surrounding nuclear power plants also points to environmental causes.
However, the even distribution of childhood cancers over the world suggests that either environmental mutagens play a very minor role in their origin or they play an equally prominent role elsewhere.

If most solid childhood tumours are indeed correctly attributable to mutations in germ and/or somatic cells, then, the prospect of prevention of childhood cancer becomes dim. In fact, the incidence of hereditary forms may increase as treatment improves. The burden of cancer attributable to the hereditary groups may be lessened by identifying individuals having such cancer genes and employing pre-natal intervention. However, there would still remain the larger non-hereditary group. If this hypothesis is correct, then this kind of childhood cancer cannot be prevented and the main effort against childhood cancer must be that of early diagnosis and treatment. The isolation of the Rb gene went a long way realising this objective.

1.3 RETINOBLASTOMA

1.3.1 The clinical recognition and description of retinoblastoma

The study of retinoblastoma (Rb) is a story that began nearly 400 years ago in the dissecting chamber of Petrus Pawius of Amsterdam who is credited with giving the first description in the literature of a tumour resembling Rb (Dunphy 1964, Albert 1987). His report (1597), republished by Bartolini in 1657 described a 3-year-old child with a large fungating tumour of the left eye. For the next 150 years, reports of Rb virtually disappeared from the medical literature. Then, in 1767, an article was published in Medical Observations and Inquiries entitled "The case of a diseased eye communicated to Mr William Hunter by Mr Hayes, surgeon". Hayes (1767, cited by Dunphy 1964, Albert 1987) described a bilateral tumour in a 3-year-old girl stating "The eye seemed to have lost its deep black appearance... and acquired a more clear bright look, something resembling a cat's eye in the dark". This could be the first description of the so-called "amaurotic cat's eye", although some attribute the term to Georg Joseph Beer of Vienna who emphasised it as a diagnostic sign of the tumour (Albert 1987).
In 1805 William Hey, senior surgeon at the General Infirmary at Leeds, coined the term *fungus haematodes* (Dunphy 1964, Albert 1987). This is also the term used by the famous James Wardrop of London in his 1809 monograph on the tumour *On Fungus Haematodes or soft cancer* (Wardrop 1809, cited by Dunphy 1964, Sang and Albert 1977, Albert 1987). His description is the first exact clinical and pathological description of Rb as a specific disease. Wardrop brought together the random, isolated facts and observations of earlier authors and, in light of his own experience, concluded that Rb is distinct from the more general classification of "soft cancers" and that the ocular form seen in children arose in most instances from the retina. He also demonstrated the extension of the tumour to the optic nerves and brain and described metastasis to different parts of the body.

Observations of Rb in these early days was restricted to those showing extraocular extension. With the development of the ophthalmoscope in 1851 (Helmholtz 1851, from Albert 1987) it became possible to examine the retina accurately thus permitting the study of smaller, developing tumours. As the diagnosis of Rb became easier many other features of the disease became apparent including the manner of intraocular and extraocular extension, patterns of metastasis and recurrence, ocular complications, and associated malignancies.

At presentation the most common symptom for Rb is leukokoria (a white reflex rather than a red reflex in the pupil from the white mass within the eye), a persistent squint (strabismus) and signs of inflammation (10% of cases). Among the diagnostic signs of Rb are seeding into the vitreous and calcification. The current diagnostic approaches are clinical history, binocular indirect ophthalmoscopy, ultrasonography, CT scan, cytology and aqueous tumour enzyme level assays. In Rb, the aqueous humor lactic dehydrogenase and phosphoglucose isomerase levels are usually increased. It is important to be aware of the other conditions which might be confused with Rb especially when taking a family history and particularly when some of this interview is anecdotal. Many of the irregularities in pedigree analysis at the turn of the century are due to misdiagnosis. The most difficult differential diagnosis is probably between Coat's disease and Rb. Also larval granulomatosis due to *Toxocara canis* and granulomatous uveitis usually due to toxoplasmosis in children can be confused with Rb. The diagnosis becomes particularly delicate when the posterior segment of the eye is
concealed by vitreous haemorrhage, retinal detachment or inflammatory or fibrous reaction. Many ocular diseases may simulate Rb in every way under these conditions.

1.3.2 The role of heredity in retinoblastoma

The heredity aspects of retinoblastoma were not appreciated until the 19th century because, prior to that, it had been ill-described and, without effective treatment, there were virtually no survivors. In 1821, 12 years after Wardrop's pioneering description of Rb, Lerche (cited by Dunphy 1964, Vogel 1979, Albert 1987) described a sibship in which four of seven children were affected, one bilaterally. The parents were apparently unaffected. Vertical transmission of Rb was supposedly first described in 1896 by De Gouvea (De Gouvea 1910, cited by Dunphy 1964, Vogel 1979, Albert 1987) in Brazil reporting the case of an Rb survivor having an affected child. This case was not appreciated, however, for many years.

As late as 1905, Parsons stated, "there is no case on record of a child from whom a gliomatous eye has been removed, growing up and having children with glioma" (Dunphy 1964, Albert 1987). 'Glioma' was one of the earlier names of Rb which was introduced by Virchow (1864, cited in Dunphy 1964, Tso 1980, Albert 1987) who believed that Rb was of glial origin (see Section 1.3.3). One report that refuted Parsons' contention was by CN Ridley in the 1904 *Royal London Ophthalmological Hospital Reports* of a surviving Rb patient whose son also developed Rb (Dunphy 1964). The unaffected sister of this patient also had two children who died as a result of Rb. Thus, in this family there was both direct and collateral inheritance. Similar reports appeared over the next 50 years and, with an accumulating number of Rb survivors and their offspring available, more scientific studies were undertaken concerning the pattern of heredity. The emergence of such Rb families was due to the success of two forms of treatment; firstly enucleation and, later on in the 1940 and 50s, the use of radiation (see Section 1.3.6). The reports in the first half of the 20th century have been reviewed by Dollfus and Auvert (1953, cited in Vogel 1979, Albert 1987) and Kaelin (1955, reviewed in Vogel 1979). It was concluded that sporadic cases occur either as a result of a somatic mutation (occurring post-zygotically in a retinal cell) or as a new germinal mutation (occurring in a germ cell). Somatic
mutation was recognised as the more common event. It also became evident that somatic mutations were usually associated with unilateral disease which often presented at a later age (Griffith and Sorsby 1944). In contrast, germinal mutations usually resulted in multiple tumours in both eyes which presented at an earlier age. This generalisation, however, did not explain incomplete penetrance and it was difficult to determine whether the abnormal gene was present or not in phenotypically normal offspring, sibling or parent of an affected individual. There were also discrepancies and disagreements regarding the unusually high mutation rate occurring in spontaneously affected individuals.

In 1971, Knudson provided strong evidence to support a "two-hit" model for the causation of Rb, hypothesising that all retinoblastomas arise as a result of only two mutational events (Section 1.2). Two observations suggested this theory. Firstly the age of onset of Rb and secondly the mean number of tumours observed in individual hereditary cases. Although virtually all cases of Rb present before the age of 5 years, there is a significant difference between the mean age of presentation of the hereditary group (mean=14 months) compared to the sporadic group (mean=32 months) (Section 1.3.5). Knudson (1971), reviewing all hereditary cases of Rb diagnosed in a single institution, found a mean of 3-4 tumours per patient. Using age of onset and incidence data, he demonstrated mathematically that the rate of onset of Rb was compatible with only two-rate limiting steps and proposed (1978), as first suggested by Comings (1973), that Rb might result from the inactivation of both alleles of a single critical locus. In familial cases, the first mutation was inherited from the affected parent (and hence was present in all the cells of the individual) and the second mutation occurred as a random, sporadic event in precursor retinal cells. In sporadic cases, both mutations were somatic and occurred in the same retinal cell. Formal proof of Knudson's hypothesis was to come in 1984 (see section 1.4.1.3).

1.3.3 Definition, histology and cell of origin of retinoblastoma

Retinoblastoma is a childhood tumour which arises from the retina as a consequence of abnormal embryological development. Although this was first suggested by James Wardrop in 1809 it was not generally accepted until Robin and Nysten (1815) and Langenbeck (1836) (cited by Albert 1987)
demonstrated this unequivocally using microscopic studies. Robin and Nysten believed the tumour arose from the ganglion cells. Iwanoff (1869) and Manfredi (1868) (cited by Dunphy 1964, Albert 1987) believed the tumour arose from the nerve fibre layer.Virchow (1864, cited in Tso 1980, Albert 1987) was the first to claim that Rb was of glial origin and he introduced the name 'glioma' in 1864.

Flexner in 1891 and Wintersteiner in 1897 (cited in Dunphy 1964, Tso 1980, Albert 1987) described the so called 'rosettes' within the tumour which now bear their name. Flexner noted structures within the tumour which resembled rods and cones and traced one tumour to the photoreceptor cell layer proposing the name "neuroepithelioma". Wintersteiner assumed that the neoplasm arose from misplaced rods and cones. A less specific structure described in Rb was the Homer-Wright rosette in which radial arrangements of cells surround a tangle of fibrils. All these features describe the differentiated areas in retinoblastoma tumours. The majority of cells in tumours, however, resemble undifferentiated embryonic retinal cells. This feature prompted Verhoeff to name the tumour, retinoblastoma, which was adopted as a general term by the American Ophthalmological Society in 1926 (Verhoeff and Jackson 1926, cited by Tso 1980, Albert 1987).

Retinoblastoma has been given many names since its description as a specific entity in 1809. Dollfus and Auvert (1953, cited in Dunphy 1964, Albert 1987) in their report on glioma of the retina cite a total of 40 different names given to the tumour. This reflects the fact that the cells of origin in Rb have not yet been identified. In 1941 Parkill and Benedict (cited by Dunphy 1964, Tso 1980) argued that the rosettes in Rb resemble the cells of the primitive neural tube and indicate that Rb is basically a glial neoplasm like an ependymoma. Electron microscopical analysis of Rb confirmed the association with photoreceptor elements, endocrine elements and glial cells. In 1969, Tso et al demonstrated primitive inner segments of photoreceptor cells, terminal bars and cilia with a 9 + 0 pattern. Sang and Albert (1977) demonstrated neurosecretory granules in Rb tumours which resembled those found in amacrine cells (one kind of neuronal cell in retina) which take up and metabolise neurotransmitters normally found in healthy retina tissue. In 1970, using light and electron microscopy, Tso et al (1970a, 1970b) documented a high degree of maturation with evidence of photoreceptor differentiation by individual tumour cells and small flower-like clusters of
benign-appearing tumour cells (which they termed 'fleurettes'). This suggested the origin of Rb cells to be immature neuroepithelial cells which are destined to become photoreceptor cells.

Electron microscopic evidence of glial differentiation in Rb was first presented by Popoff and Ellsworth (1971, cited by Tso 1980, Albert 1987). Tso, in 1980, described an area of glial differentiation in an Rb where the glial-like tumour cells were separate from the undifferentiated Rb cells and were located away from the detached retina. Later on Lane and Klintworth (1983), using an immunoperoxide technique, demonstrated glial elements within other cases of Rb but failed to observe glial differentiation in these tumour cells. Craft et al (1985) reported a case in which cells resembling glial cells were intermixed with more typical cuboidal Rb cells.

There have been numerous studies on the biochemical properties of Rb with the aim of both finding clues to the origin of Rb cells and establishing clinical diagnostic criteria. The study of catecholamines in Rb is one such study. Rb bears certain clinical and morphological similarities to neuroblastoma particularly in view of its neurogenic origins and histologic appearance. The majority of patients with neuroblastoma show increased levels of catecholamine metabolites in the urine. Although several studies in children with Rb have indicated that levels of such compounds were elevated in at least some patients (Brown 1966, Francois and Koliopoulos 1979), other reports have suggested that Rb is not a catecholamine-secreting tumour to any significant degree (Dias and Amarasiri 1978). Sang and Albert (1982) studied catecholamine in cultured Rb cells and showed insignificant levels to be present. They believe, however, that there is efficient active catecholamine uptake and metabolism within Rb cells in spite of the fact that there is no storage of the bioamine per se. This study provided clues that Rb might have a metabolism similar to that of the neuronal cells of the normal retina.

Much attention has been focussed on the measurement of lactic acid dehydrogenase (LDH), which is an enzyme found in tissues with a glycolytic metabolism, in both aqueous humor and serum. Increased aqueous humor LDH levels have been found in some Rb patients. Total aqueous humor LDH levels and isoenzyme patterns were suggested to be of clinical value and to have prognostic significance (Dias 1979a). Some authors have suggested
that LDH measurements may be related to the degree of tumour necrosis (Swartz et al 1974) or to the duration of tumour involvement (Dias 1979b). Relatively low values have been suggested to be related to such factors as a very advanced degree of differentiation or the presence of old, successfully treated non-viable disease (Dias 1979b).

The most dramatic demonstration of the multipotential nature of the Rb cell came from the study of Kyritsis et al (1984) on the cells of the human Rb cell line, Y79. They found that all undifferentiated cells contain both a neuronal marker, neurone-specific enolase (NSE), and a glial marker, glial fibrillary acidic protein (GFAP), whereas the differentiating neuronal and glial-like cells gradually lose one or other marker and selectively express the marker that correlates with their morphology. This result supported the notion that Rb originates from a primitive bipotential (or multipotential) neuro-ectodermal cell.

In 1988, Bogenmann et al reported detection of mRNA transcripts believed to be specific for cones in Rb cells cultured in vitro. It has been suggested that Rb is derived from cells destined to become cones. However, the clinical observation that there is no predilection of Rb for the fovea, that region of the retina with the highest density of cones, conflicts with this hypothesis. The question of the cell of origin of Rb, therefore, is still not settled although many clues point to a primitive neuroectodermal cell.

### 1.3.4 Spontaneous regression of retinoblastoma

The general assumption is that gene carriers will develop Rb with the frequency of incomplete penetrance being 10%. However, in some predisposed individuals the tumours have been shown to spontaneously regress, an interesting feature of Rb, first described simultaneously by De Kleijn (1911, cited in Albert 1987) and Knieper (1911, cited in Gallie et al 1982b, Albert 1987). The overall incidence of spontaneous regression is approximately 1:80000 and according to Cole (1974) has been more commonly reported in other tumours such as hypernephroma, neuroblastoma, malignant melanoma and choriocarcinoma. In Rb, however, spontaneous regression may be as high as 1% (Gallie et al 1977, 1982a), i.e. about 1000 times more frequent than in any other neoplasm, and may be due to missed
diagnosis or misdiagnosis. In general the net result is a scar on the retina for
which there may be many causes.

Spontaneously regressed tumours may have the appearance of one of two
distinct clinical phenotypes. The first is phthisis bulbi or a shrunken, non-
functional eye with intraocular calcification. The second is characterised by
translucent, irregular retinal mass, "cottage cheese" calcification and
disturbance of the pigment epithelium. The only other retinal lesion to have
all three of these characteristics are Rb tumours successfully treated with
radiation. Gallie et al (1982b) proposed the term "retinoma" to describe
these distinctive retinal lesions, often associated with hereditary Rb but
lacking malignant characteristics. The diagnosis of retinoma, when there is a
family history of Rb, strongly suggests the presence of a mutant Rb gene.
Phthisis bulbi, on the other hand, can be caused by other means, much more
common than Rb, such as trauma or infection. Gallie et al (1982b) proposed
that the mechanisms leading to phthisis bulbi and retinoma in Rb mutant gene
carriers are different, phthisis bulbi being the end result of tumour necrosis
and ocular ischaemia, probably due to occlusion of the central retinal vessels.
Retinomas, however, have a clinically observable blood supply making
ischaemia unlikely. It was proposed that retinoma is an end result, when the
second mutation occurs in a relatively mature retinoblast and is the benign
manifestation of the late occurrence of the second mutation rather than
spontaneous regression of full-blown malignant Rb tumour. There is
anecdotal evidence, however, of tumours which arose particularly during the
First World War, when treatment was not available, but which had
disappeared by the end of the war. The fact that neuroblastoma can clearly
regress spontaneously supports the suggestion that differentiation can override
transformation.

1.3.5 The incidence and age distribution of retinoblastoma

Rb, although rare, is the commonest ophthalmic malignancy of childhood and
comprises between 2-3% of all childhood cancers. The average world-wide
incidence is approximately 1:20000 (Knudson 1976). Incidence figures in the
literature, however, range between 1:34000 and 1:10000 (reviewed in Vogel
1979). There seems to be some geographical variation in incidence with
particularly high incidence reported from Malawi (Molyneaux 1979) and
from South America, an area of the world where Rb has been known to exist since 2000 BC (Gaitan-Yanguas 1978). Vogel (1979) suggested that figures of 1:28000 - 1:15000 give reliable limits for its incidence. In the US the annual incidence in children under 5 is 11 per million (Pendergrass and Davis 1980, Tamboli et al 1990) which represents a cumulative rate by age 5 of 55 per million or 1:18000 (Sanders et al 1988). For the UK Sanders et al (1988) estimated an incidence of 1:23000. This figure may be an underestimate by 10% since tumour registration in Britain is considered incomplete.

Variations in incidence may reflect problems of ascertainment, which usually results in under-reporting. Those countries with an apparently very high incidence also have poorly developed medical services and it seems likely that the high incidence of Rb, virtually all due to sporadic disease in such countries, reflects some generalised genetic predisposition or some environmental factor over and above the background mutation rates. Evidence from Finland (Tarkkanen and Touvinen 1971, reviewed in Vogel 1979) suggests that the incidence of Rb is rising, though this observations might simply have reflected better survival of patients with the hereditary form of the disease. A rising incidence of Rb has not been found in general (Vogel 1979).

Equally imprecise are figures in the literature for the hereditary versus non-hereditary proportion of Rb and for the distribution of cases by laterality. In the UK (Sanders et al 1988) 40% were reportedly hereditary (all patients with bilateral Rb and/or a positive family history). Tamboli et al (1990), from a study of 220 cases in the US suggested that 25% of cases were bilaterally affected. Among the 550 cases studied in Holland, 31% were found to be bilateral (Schappert-Kimmijser et al 1966, cited by Vogel 1979) and in a study of 899 cases of Rb in France, 34% were bilateral (Bonaiti-Pellie et al 1976). DerKinderen (1988) found 36% of the 403 cases in Holland Rb registry (1945-1970) to be hereditary. The Childhood Cancer Research Group in Oxford established a registry of Rb diagnosed in Britain, and, among 918 cases of Rb diagnosed between 1962 and 1985, 40% were bilateral (Draper et al 1992). A preliminary survey (Jay et al 1988) suggested 47% of cases in the UK were hereditary supporting earlier data by Cowell et al (1986) although only 60% ascertainment had been achieved. A more accurate study by the Oxford group indicated that 44% of cases are
hereditary. Only 28% of bilateral cases and 7% of unilateral cases had a previous family history, making the overall frequency of familial Rb 15%.

Bilateral cases of Rb tend to present at a much younger age, approximately 15 months, compared with unilateral cases at 24-32 months (Jensen 1965, Leelawongs and Regan 1968, Knudson 1971). These figures, however, depend on whether there is a family history (FH) or germline mutation. Bonaiti-Pellie et al (1976) and later the Childhood Cancer Group (Oxford) found the mean age at diagnosis for bilateral cases with a positive FH to be 12.3 and 7.2 months respectively compared with 18.8 (Bonaiti-Pellie et al 1976) and 20.3 (Draper et al 1992) months for unilateral cases with a FH. For patients with new germline mutations (all bilaterals) the corresponding value was 13.9 (Bonaiti-Pellie et al 1976) and 14 (Draper et al 1992) months. The mean age at diagnosis for sporadic unilateral cases was found to be 29.8 (Bonaiti-Pellie et al 1976) and 29.5 (Draper et al 1992) months.

1.3.6 Treatment of retinoblastoma

In the 1930s and 1940s the thought of patients with Rb tumours retaining an eye with useful vision was inconceivable. Today the majority of patients retain useful vision in one or both eyes. The method of treatment chosen depends on several factors: the size and location of the tumour, the number of tumours in each eye, the patient's age, whether the disease is unilateral or bilateral, and whether there is vitreous involvement. In many cases clinical management involves a combination of therapies. Other than enucleation, among the method of treatments available these days are cryotherapy, lens sparing laser treatment, radiation treatment (external beam and radioactive 'plaques') and chemotherapy for metastasis.

The traditional method of treatment for Rb has been removal of the eye - enucleation. It was first suggested in 1809 by James Wardrop (Sang and Albert 1977, Albert 1987, McCartney et al 1988) but only became the accepted form of treatment after the introduction of chloroform as a general anaesthetic and with the availability of the ophthalmoscope for earlier diagnosis (in 1851). For many decades after that, enucleation remained the only form of treatment for Rb. These days enucleation is the only treatment in unilateral cases when the tumour fills most of the globe and there is little
hope of retaining useful vision and a strong possibility the tumour will escape the orbit. In bilateral cases, the surgeon will try to save at least one eye and so the one with the more advanced tumour is usually enucleated; the less involved eye is usually treated. Sanders et al (1988) compared the treatment for Rb, in the UK, between periods 1962-1968 and 1969-1980 and found a trend which was towards more conservative treatment (table 1.1). 5% of children with unilateral tumours in the later period (1969-1980) did not have enucleation as compared with 1% in the earlier period. For the bilateral cases at least one eye was retained for 83% of children in the later period, as compared with 73% in the earlier.

**Table 1.1** Comparison of treatment in the UK for Rb between periods 1962-1968 and 1969-1980 (Sanders et al 1988).

<table>
<thead>
<tr>
<th>Years</th>
<th>Total cases</th>
<th>Unilateral</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Enucleation</td>
<td>Conservative treat.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No (%)</td>
<td>No (%)</td>
<td></td>
</tr>
<tr>
<td>1962-1968</td>
<td>157</td>
<td>155 (99)</td>
<td>2 (1)</td>
<td></td>
</tr>
<tr>
<td>1969-1980</td>
<td>282</td>
<td>269 (95)</td>
<td>13 (5)</td>
<td></td>
</tr>
</tbody>
</table>

| Bilateral |  |  |  |
|-----------| |  |  |
|           | Enucleation | Conservative treatment |  |  |
|           | One eye | Both eyes | No (%) | No (%) |  |  |
| 1962-1968 | 111 | 66 (59) | 30 (27) | 15 (14) |  |  |
| 1969-1980 | 149 | 101 (68) | 25 (17) | 23 (15) |  |  |

The move away from enucleation has been brought about mainly by the introduction of radiation as a form of therapy. The first attempt to treat Rb with X-rays occurred in 1903 (Hilgartner 1903, cited by Dunphy 1964, Sang and Albert 1977) although it was not until 1921 that the first well-documented cure of bilaterally affected patients was reported (Verhoeff 1921). Since then several techniques of external beam irradiation have been used successfully.
There are, however, complications with the use of external beam irradiation, including damage to the lachrymal gland (dry eyes), muscle wasting in the irradiation field and bone growth deformities. Nowadays radiation treatment is usually limited to bilateral tumours and unilateral tumours having good central vision potential. Consequently, efforts to optimise treatment have become important. In 1930 Moore and co-workers treated Rb by placing radon seeds through a puncture in the sclera (Moore et al 1931). Stallard (1958) refined this work by designing special $^{60}$Co "plaques" that followed the curvature of the sclera. The radiation was directed towards the tumour and was blocked from irradiating surrounding tissue. The radioactive plaque technique was found to be very advantageous when treating localised tumours as a high dose of radiation is delivered to the areas of interest while minimising radiation effects to the extraocular structures. Rb is probably the most radiation sensitive tumour known, usually a single treatment is sufficient to kill the tumour. A small scar is left on the retina following successful treatment. However, when tumours are very small, photocoagulation, laser therapy and cryotherapy are the treatments of choice for primary and supplementary treatment, although some small tumours might not be suitable for these treatments due to their location. Photocoagulation destroys small retinal tumours by obliterating the blood vessels supplying the tumour with the use of a powerful light beam (Shields and Augsburger 1981). Cryotherapy for Rb was first reported in 1968 (Lincoff 1968) and involves freezing of tumours using a cryoprobe pushed up against the tumour. As a result the tumours cells are killed by ice crystal formation, dehydration and obliteration of the capillary circulation. Cryotherapy is occasionally used as primary treatment but more frequently employed to supplement other treatment modalities.

Chemotherapy is used in advanced (usually metastasing) cases of Rb where the prognosis is poor. It was first described by Kupfer in the treatment of Rb in 1953. In the 1950s and 60s nitrogen mustard and its analogue triethylene melamine were used. Today other chemotherapeutic agents inducing cyclophosphamide, vincristine doxorubicin, 5-fluoracil are used singly or in combination in cases involving metastasis. Metastases of Rb are uncommon in countries where the tumour is diagnosed and treated early. Dissemination of the tumour may be haematogenous or lymphatic. The central nervous system (CNS) can become involved via continuous spread along the optic nerve (JE Kingston, personal communication). The majority of the patients
with metastases show involvement of the CNS (MacKay et al 1984). The main sites of metastases outside the CNS are the orbital and periorbital tissues including the cranial bones. Occasionally, distant areas of involvement include bones and bone marrow, lymph nodes, liver, and kidneys. Most of metastases are seen within three years after diagnosis of Rb (MacKay et al 1984). The prognosis of children with metastatic disease remains poor and they rarely survive more than a year although longer survivals have been reported (Judisch et al 1980).

In many cases, a combination of above mentioned treatment modalities are used in the management of Rb and a trend towards more conservative treatment has been reported (Sanders et al 1988). As a result of many years of painstaking research on the treatment of Rb the survival rate is the highest of all children's cancers. This fact is primarily responsible for the increased frequency of hereditary cases and their documentation.

1.3.7 Survival rates in retinoblastoma

Rb is a highly malignant tumour which, if untreated, follows a lethal course either through local spread or distant metastases. A century ago the mortality rate with this malignancy was almost 100% and continues to be so in the Third World countries where treatment is not available. In the Western countries the survival rates for Rb have improved dramatically over the years (table 1.2). The improved survival rate is due to detection of the disease at an early stage and improved techniques for tumour control not a change in the natural history of the disease. The high mortality rate in less developed countries emphasises the fact that when the disease presents beyond the eye the survival rate is less than 10%.

The great majority of deaths from Rb occur within 3 years of diagnosis. In their 1988 paper Sanders et al reported a 3 year survival rate of 88% in the UK. Tamboli et al (1990) reported a 5 year survival rate of 91% in the US. Patients with bilateral tumours had a better survival rate (90.5%) than those with unilateral tumours (86.5%) for the first few years but their long-term survival rate was lower because of later deaths from ectopic intracranial Rb or second primary neoplasms (Sanders et al 1988). For survivors of hereditary Rb, therefore, the risk of developing second primary neoplasms
assumes an increasing importance in childhood and adult life. The mortality rate with the second neoplasms is around 80-85%. It is now believed that most of survivors of hereditary Rb die as a result of developing second neoplasms in adult life (see below).

Table 1.2 Survival rates for Rb in the Western countries.

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Survival Rate (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1869</td>
<td>5</td>
<td>Hirschberg 1869*</td>
</tr>
<tr>
<td>1897</td>
<td>17</td>
<td>Wintersteiner 1897*</td>
</tr>
<tr>
<td>1911</td>
<td>57</td>
<td>Leber 1911*</td>
</tr>
<tr>
<td>1960-1967</td>
<td>81</td>
<td>Jenson and Miller 1971</td>
</tr>
<tr>
<td>1974-1985</td>
<td>91</td>
<td>Tamboli et al 1990</td>
</tr>
<tr>
<td>1990s</td>
<td>93</td>
<td>UKCCSG**</td>
</tr>
</tbody>
</table>

* (cited by Dunphy 1964, Albert 1987)
** (United Kingdom Children’s Cancer Study Group, JK Cowell, personal communication)

1.3.8 Second primary cancers in retinoblastoma

Numerous studies have indicated that survivors of genetic Rb have a substantial risk of developing second primary tumours. One of the first observations on this subject came from Reese et al (1949, cited by Draper et al 1986) who reported two cases of second tumours among 55 bilateral Rb patients treated with combined radiation and surgery. The high risk of second neoplasms in Rb survivors has subsequently been well documented (Jensen and Miller 1971, Kitchen and Ellsworth 1974, Abramson et al 1984, Draper et al 1986, Roarty et al 1988). The increased incidence of second tumours was initially attributed to radiation since radiotherapy is used either as the primary mode of therapy or as an adjunct to enucleation in many bilateral cases. Unilateral patients usually receive no radiation (Section 1.3.6). In fact, Sagerman et al (1969) demonstrated that the incidence of
second tumours increased with increasing radiation dose. The occurrence of second cancers arising outside the field of radiation was first pointed out by Jensen and Miller (1971) who reported cases of osteosarcoma in long bones, far removed from the radiation field. Many more cases of second tumours outside the field of irradiation (Kitchen and Ellsworth 1974, Abramson et al 1984), and in some cases where no radiation therapy had been given at all (Abramson et al 1979, Abramson et al 1984, Roarty et al 1988), were subsequently reported. Although the increased risk of second primary tumours among survivors of the genetic form of Rb is well recognised, the incidence and risk figures in the literature differ widely from each other (tables 1.3 and 1.4).

In the largest study, shown in table 1.3 (Kitchen and Ellsworth 1974), out of 45 second tumours 40 were sarcomas and all tumours, with one exception were confined to children with bilateral disease. Of these, 16 were osteosarcomas, 7 of which were detected outside the field of radiation. In the study by Abramson et al (1984) (table 1.3), 693 of 711 patients had bilateral Rb and of 688 patients who were treated with radiation, 9% developed tumours in the field of radiation and 4% outside the field. Of 23 patients who received no radiation, 5 (22%) developed second tumours. Francois et al (1980) (table 1.3) followed up 85 Rb patients in their series for periods between 4 and 30 years and out of 8 second tumours observed (all bilateral cases), 3 were in the field of radiation. Tucker et al (1984) presented data on Rb patients who had survived at least 2 years from diagnosis (table 1.3). Grouping together unilateral and bilateral cases and including second tumours of all types and sites, they found that the incidence of second tumours during an average further follow-up of 7 years was 60 times higher than expected from population rates for malignant neoplasms, the rate for bone tumours being 1000 times the population rate.

In some studies the incidence of second tumours in Rb patients has also been calculated using life-table methodology (table 1.4) to correct for variable length of follow-up. The highest figures reported were those of Abramson et al (1984) who, in contrast to the findings of Sagerman et al (1969), found no relationship between incidence of tumours and dose of therapeutic radiation when analysed with life tables. Draper et al (1986), however, reported a significantly lower incidence of second tumours - 8.4% at 18 years (table 1.4). The corresponding figure for osteosarcoma alone was 6%. Leuder et
al (1986) and DerKinderen et al (1988) also found a low rate of development of second tumours: 14\% at 30 years of age (Leuder et al 1986); 11\% (tumours in irradiated field) and 10\% (outside irradiated field) at 35 years of age (DerKinderen et al 1988). In these studies, however, the cumulative incidence were expressed in years of age rather than years after diagnosis and may not be comparable with the ones given in table 1.4. In the data from Roarty et al (1988) the rate outside the irradiated field (8.1\%) was similar to that observed in non-irradiated patients (5.8\%) (table 1.4). Both Draper et al (1986) and Roarty et al (1988) concluded that carriers of the mutant Rb gene have an increased incidence of second tumours and that the incidence rate is further increased in patients who receive radiation therapy. Both authors suggested that the discrepancy between their incidence rates and that of Abramson et al (1984) was probably due to Abramson and his colleagues being more successful in following up patients who developed second tumours than those who remained unaffected. In his study Draper et al (1986) also investigated the affect of the use of chemotherapy on Rb patients and suggested that the use of cyclophosphamide might increase the risk of second cancers in genetically predisposed Rb patients. The same suggestion was also made later by Tucker et al (1987) and Winther et al (1988).

Table 1.3 The incidence of second cancers in Rb patients.

<table>
<thead>
<tr>
<th>No of Rb patients</th>
<th>Second tumours No (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1130</td>
<td>45 (4)</td>
<td>Kitchen and Ellsworth 1974</td>
</tr>
<tr>
<td>85</td>
<td>8 (9.4)</td>
<td>Francois et al 1980</td>
</tr>
<tr>
<td>711</td>
<td>93 (13)</td>
<td>Abramson et al 1984</td>
</tr>
<tr>
<td>319</td>
<td>20 (6)</td>
<td>Tucker et al 1984</td>
</tr>
<tr>
<td>882</td>
<td>30 (3.4)</td>
<td>Draper et al 1986</td>
</tr>
<tr>
<td>215</td>
<td>24 (11)</td>
<td>Roarty et al 1988</td>
</tr>
</tbody>
</table>
Table 1.4 The cumulative incidence rates for second cancers in Rb patients calculated using life-tables.

<table>
<thead>
<tr>
<th>Rb patients</th>
<th>Incidence (%) at years</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>20 - 50 90</td>
<td>Abramson et al 1984</td>
</tr>
<tr>
<td>Non-irradiated + tumours outside irradiation field</td>
<td>10 - 30 68</td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>- 8.4 -</td>
<td>Draper et al 1986</td>
</tr>
<tr>
<td>All</td>
<td>4.4 - 18.3 26.1</td>
<td>Roarty et al 1986</td>
</tr>
<tr>
<td>Non-irradiated</td>
<td>- - -</td>
<td>5.8</td>
</tr>
<tr>
<td>All irradiated</td>
<td>- - -</td>
<td>35.1</td>
</tr>
<tr>
<td>tumours in irradiated field</td>
<td>- - -</td>
<td>29.3</td>
</tr>
<tr>
<td>tumours outside irradiated field</td>
<td>- - -</td>
<td>8.1</td>
</tr>
</tbody>
</table>

The types of second tumours found were similar in most of the studies, osteosarcoma being the most common among all second cancers which included soft tissue sarcomas (eg fibrosarcoma, rhabdomyosarcoma), brain tumours, leukaemia and some epithelial tumours. Draper et al (1986) found that 54% of all second tumours were osteogenic sarcomas in their study. In the study by Roarty et al (1988), however, the corresponding figure was 19%. The higher percentage of neoplasms other than osteosarcoma in this study was suggested to be the result of longer follow-up of the patients. They found that osteosarcoma had its peak incidence in the second decade of the
patient's life whereas other neoplasms were distributed more evenly in the second and third decades. Among the brain tumours which developed there were tumours in pineal and suprasellar regions. These tumours were excluded from the study by Draper et al (1986) as they were regarded as histologically similar to Rb ('trilateral Rb', see later) but included in the studies by other groups. Draper et al (1986) found a strong association between Rb and melanoma and Roarty et al (1988) reported that 5.4% of all second tumours in their study were melanomas and that melanoma was one of the most common neoplasms in the third decade of life. It was observed that osteogenic sarcoma occurred in Rb mutant gene carriers at about the same age as they do in the normal population. This raised the possibility that such individuals would have dramatically higher incidences of gastrointestinal, lung and breast carcinomas when they approach the usual age of onset of these neoplasms. Although there is not abundant reliable data on Rb patients who are older than 35 years of age, studies on Rb families pointed to a generally increased risk of cancer in close relatives of Rb patients (Bonaiti-Pellie and Briard-Guillemot 1980, DerKinderen et al 1988, Sanders et al 1989). In the largest of these studies, Sanders et al (1989) followed up a series of 1438 parents and 2663 other relatives of Rb patients to ascertain the incidence of non-ocular cancer among them. Among 117 of these relatives, who were known carriers of Rb mutant gene, 23 developed non-ocular cancers representing a relative risk of 9.9. These mutant gene carriers were found 15 times more likely to die from lung cancer than the general population. Other frequently observed cancers included melanoma and bladder cancer. The incidence of non-ocular cancers among relatives who are not definitely known to be carriers showed an excess risk of 1.6 and it was concluded that a proportion of these relatives were in fact mutant gene carriers. They found no excess risk of developing cancer for relatives who are not mutant gene carriers.

1.3.9 Trilateral retinoblastoma

Among the tumours developed by Rb patients are intracranial neoplasms which have a histopathological resemblance to Rb. The existence of these 'ectopic retinoblastomas' was first recognised by Jacobiec et al (1977) who reported 2 intracranial malignancies in patients with Rb. Three years later Bader et al (1980) suggested the term 'trilateral Rb' to describe the clinical
syndrome of bilateral Rb with an ectopic midline intracranial tumour. Bader et al (1982), subsequently reported 14 children with characteristics of trilateral Rb and suggested that the development of an ectopic midline neuroblastic tumour in a patient with bilateral Rb represents an additional focus of multicentric Rb rather than a second primary tumour and drew attention to the ontogenesis of the pineal gland (see also Zimmerman et al 1982). The mammalian pineal has a phylogenetic and ontogenetic relationship to photoreception. The pineal eyes of certain birds, fish and reptiles exhibit histologic similarities to those of a visual organ and have photoreceptor activity. Hence, the pineal has often been called the third or median eye. In most higher species, the pineal has lost most of its structural and functional similarities to a photoreceptor organ but retained its sensitivity to photic stimuli (Wurtman and Moskowitz 1977). It has been proposed that ectopic Rb might arise mostly in the infant's pineal but also more rarely elsewhere in the diencephalon from germinal matrix cells near those that normally give rise to the optic cup (Bullitt and Crain 1981). The comparatively small number of such cells in the pineal and their even smaller numbers in the parasellar tissues would account for the relative rarity of ectopic Rbs even in genetically susceptible individuals (Zimmerman 1985).

In a British study conducted by Kingston, Plowman and Hungerford (1985), 11 ectopic intracranial Rb were identified among 630 children who had been treated for Rb over a 30 year period. 10 of these patients had bilateral Rb and one unilateral Rb. They also recorded the case of a 5-month old girl who had a suprasellar tumour but no evidence of retinal disease. This girl had a sister who had Rb and a suprasellar tumour. Also some rare pineoblastomas with well differentiated, Flexner-Wintersteiner rosettes and fleurettes and individual cells exhibiting photoreceptor differentiation have been observed in children from families not known to have had members with Rb (Stefanko and Manschot 1979, Bader et al 1982, reviewed in Zimmerman 1985). In these cases it is possible that an association with Rb may have been overlooked or these cases may represent spontaneous non-genetic forms. Mean age of patients with pineoblastomas showing retinoblastomatous differentiation (but no known association with Rb) was found to be significantly less than (<3 years) that of other patients with pineoblastoma showing no such differentiation (Zimmerman 1985). Zimmerman (1985) reviewed pineoblastomas and suprasellar and parasellar tumours from various series. He found that, in cases of trilateral Rb in which the intracranial
tumour was found in the pineal, the patients were significantly older at the
time of recognition of the pineal tumour than in the cases in which the brain
tumour was suprasellar or parasellar. Patients were usually discovered to
have pineal tumours long after they had been successfully treated for their
Rbs. In striking contrast, patients with suprasellar or parasellar tumours
often presented initially with clinical manifestations attributable to their brain
tumours and were subsequently found to have small, clinically unsuspected
Rbs or their intracranial tumours were discovered soon after diagnosis of Rb.
In the past, some of these intracranial tumours were misinterpreted as
metastatic Rb to the brain. However, as explained by Bader et al (1982), the
absence of pathological features usually associated with the development of
metastases argue against this interpretation. The histology of the ectopic
tumours are indistinguishable from primary Rb arising in the eye. Also
ectopic tumours occur predominantly in children with the genetic form of Rb.

It has been difficult to determine the incidence rate of ectopic Rbs for several
reasons. Firstly, because it has recently been recognised as a clinopathologic
entity and as mentioned above, in the past some cases were misinterpreted as
metastatic Rb. Secondly, the availability of computerised tomography (CT)
has recently led to the earlier discovery of these tumours at a stage when
there is no other evidence of dissemination. Thirdly, most of the cases
represented isolated reports and the total number of Rb patients in the
population from which they were drawn is unknown. Only two large series
of Rb cases with long-term follow-up data recorded the prevalence of pineal
or other independent intracranial tumours. These are the series reported by
al 1984) reported 3 pineoblastomas among 688 genetic Rb patients,
representing 0.4%. The latter (Kingston et al 1985) reported 10 cases among
432 bilateral Rb children which represents 2.3%. These estimates are
thought to be the minimum rates as an increasing number of Rb patients are
being diagnosed with ectopic tumours since CT scanning became available.
A more realistic value of incidence rate is thought to be 4% (JE Kingston
personal communication). Although, a very high percentage of intraocular
Rbs are successfully treated these days, the prognosis for ectopic Rb tumours
remains poor (Holladay et al 1991).
1.4 RETINOBLASTOMA - THE GENE

1.4.1 The location of the retinoblastoma gene

Clues for the localisation of the gene responsible for Rb came from three distinct lines of study: cytogenetic analysis, esterase D studies and molecular genetic investigation of Rb tumours.

1.4.1.1 The cytogenetic evidence

The first clues as to the location of the gene responsible for Rb came from chromosome analysis of a group of rare Rb patients with other congenital abnormalities. These patients were observed to have constitutional deletions in the long arm of a D-group chromosome (chromosomes 13, 14 and 15). The first reports were that of Stallard in 1962 and Lele et al in 1963 both of which reported the case of a female infant with Rb who had a number of other congenital abnormalities (table 1.5). At that time, it was difficult to distinguish between the middle sized acrocentric chromosomes, originally called the 'D' group and the deletion chromosome was thought to be number 15. Additional patients with the "D-deletion syndrome" were subsequently reported (Thompson and Lyons 1965, Van Kempen 1966, Wilson et al 1969). The type and severity of phenotypic features observed in these patients (table 1.5) depended on the extent of the deletion.

Table 1.5 Some of the phenotypic features observed in the "D-deletion" syndrome patients.

<table>
<thead>
<tr>
<th>Phenotypic feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mental retardation</td>
<td></td>
</tr>
<tr>
<td>abnormal thumbs</td>
<td></td>
</tr>
<tr>
<td>typical facial appearance: bulbous nasal tip</td>
<td></td>
</tr>
<tr>
<td>flat nasal bridge</td>
<td></td>
</tr>
<tr>
<td>long philtrum</td>
<td></td>
</tr>
<tr>
<td>bushy eyebrows</td>
<td></td>
</tr>
</tbody>
</table>

46
The development of chromosome banding techniques identified chromosome 13 as the "culprit" chromosome (Orye et al 1971, 1974, Wilson et al 1973). As the data accumulated it soon became clear that, although the extent of the deletions varied, they all involved part of chromosome band 13q14 (Knudson et al 1976, Francke and Kung 1976, Yunis and Ramsay 1978). This observation strongly suggested band 13q14 as a possible locus for the Rb gene. Large constitutional deletions are seen in only 3-5% of cases, however (Vogel 1979, Dryja et al 1983, Cowell et al 1986). Despite this rarity the existence of such deletions strengthened the hypothesis that genetic alteration is a primary event in the predisposition to Rb.

Although deletions are by far the most frequent constitutional chromosomal abnormality in Rb a number of translocations have also been described (Strong et al 1981, Motegi et al 1982, Higgins et al 1989). Translocations arise by accidental recombination between non-homologous chromosomes during meiosis. In balanced, or reciprocal translocations, there is no visible loss of chromosome material whereas unbalanced rearrangements result in a loss of chromosomal material. It was observed that, in some cases, the inheritance of an unbalanced form of the rearrangement resulted in deletion of the 13q14 region (Strong et al 1981, Turleau et al 1983). Strong et al (1981), for example, reported a family showing a balanced chromosomal translocation in unaffected carrier parents and some unaffected sibs and an unbalanced chromosome 13 resulting in a deletion in affected individuals. The deleted region of chromosome 13 in the predisposed children was q13.1-q14.3. In reciprocal translocations involving the autosomes, 13 was always one of the chromosomes involved and the breakpoint was always in band q14 (Motegi et al 1982, Higgins et al 1989, Mitchell and Cowell 1989). In rearrangements involving the X chromosome, the breakpoint on chromosome 13 was not always in q14. However, because in females one of the X chromosomes is genetically silenced by chromosome inactivation (Lyon 1961), the spreading of 'inactivation' from the translocated X chromosome allows genes at some distance from the breakpoint to become inactivated (Ejima et al 1982).

Rearrangements involving 13q14 have also been reported in tumour cells of individuals with normal somatic karyotypes (Balaban et al 1982, Benedict et al 1983a). This finding implicated the 13q14 region in the development of sporadic Rb. In Rb tumours, a variety of other chromosomal abnormalities,
especially involving chromosome 1 and 6, have also been reported (Benedict et al 1983a, Squire et al 1985). One of the abnormalities frequently involved is chromosome 6p, with two equal short arms and no long arms (iso-6p). Two normal chromosome 6 are also present. Another abnormality involves extra copies of 1q (1q+). Iso-6p was reported to be present in 45-56% and 1q+ in 44-78% of Rb tumours (Squire et al 1985, Potluri et al 1986). Similar observations for chromosome 1q have been made in many other types of malignancy. Iso-6p, on the other hand, is less frequently observed in other tumours, being restricted largely to Rb and malignant melanomas (Becher et al 1983). These abnormalities, however, are thought to be related to tumour progression rather than causation, since none have been reported in somatic karyotypes. Evidence for the presence of a tumour suppressor gene on chromosome 6 has been presented by Trent et al (1990) but its exact significance in tumorigenesis is still unclear. Other, less frequently observed chromosome abnormalities (Potluri et al 1986) in Rb tumours include monosomy 16, 1p+ and the presence of double minutes and homogeneously staining regions. The latter two represent cytological manifestation of gene amplification (Cowell 1982). In Rb, amplification of the N-myc and int-1 oncogenes have been reported in a few tumours and/or cell lines (Lee et al 1984, Sakai et al 1985, Arheden et al 1988). Both of these oncogenes appear to be expressed in cells of neurogenic origin although, again, their role in tumorigenesis is unclear.

1.4.1.2 Esterase D studies

Constitutional chromosomal deletions in rare patients pointed to the 13q14 region as the possible locus for the Rb gene. Linkage studies in familial non-deletion cases also supported this location. Esterase D (ESD) is a ubiquitous tissue esterase possibly involved in the metabolism of sialic acid (Varki et al 1986). ESD and its associated polymorphic variants was first described by Hopkinson et al (1973). There are several different phenotypes which arise by the co-dominance of two autosomal alleles and which are distinguishable by their electrophoretic mobility in starch gels. The most common alleles are ESD1 and ESD2. ESD was initially assigned to chromosome 13 by Van Heyningen et al (1975). Sparkes et al (1979) developed an in vitro quantitative assay for ESD and showed patients with various deletions of chromosome 13 had reduced enzyme levels (Sparkes et al 1980). In patients
with 13q14 deletions ESD activity in somatic cells showed only 50% of that found in tissues from normal individuals, thus assigning the ESD gene to 13q14.

The ESD assay has now been applied to population screening to detect deletion carriers, some of which were not detected following routine chromosome analysis. Dryja et al (1983) assayed ESD activity in 51 Rb patients with normal karyotypes and found no evidence of deletion of the ESD locus. Cowell et al (1986), however, screened 200 patients with Rb, nine of whom were identified as having low ESD levels. Two of these patients had very small deletions which might have been missed using conventional chromosome analysis alone. This screening programme was then extended to 500 patients 15 of whom had reduced ESD activity and had deletions involving 13q14 (Cowell et al 1989). This survey suggested a figure of 3% for large chromosome deletions in Rb patients. However, these figures are always subject to referral bias as Dryja et al (1983) found no deletion among 51 Rb patients and Turleau et al (1985) found 7 deletions in 66 Rb patients.

Further sublocalisation of the ESD and Rb loci was achieved by studying patients with small deletions affecting only part of 13q14. Yunis and Ramsay (1978) described an Rb patient where the deletion appeared to be of 13q14.2 and possibly including material from 13q14.1 and 13q14.3. Sparkes et al (1984) suggested that the Rb gene was in the 13q14.1 region through analysis of the chromosomes from an Rb patient who reportedly had normal ESD levels and a deletion of 13q14.1-q22. When ESD levels in the lymphoblastoid cell line of the same patient were analysed by Cowell et al (1987a), however, it was found that the cell line had only 50% ESD levels. Cowell et al (1987a) described a patient with a congenital phenotype typical of a deletion patient (table 1.5) and a 13q14-q31 deletion, the proximal breakpoint being in 13q14.3. This patient had normal ESD levels suggesting that the Rb gene lay distal to ESD. This was later confirmed by Mitchell and Cowell (1988) who isolated the deletion chromosome in somatic cell hybrids and demonstrated that the ESD gene was present on the deletion chromosome but the Rb gene was not. The close physical proximity of the ESD and Rb loci meant that the electrophoretic variants of the ESD protein described by Hopkinson et al (1973) could be used as markers in linkage studies for the inheritance of non-deletion form of Rb. Sparkes et al (1983) were able to
show close linkage between the two loci in an analysis of three families thus confirming that the hereditary, non-deletion form of the tumour was in 13q14 also. Similar analyses were carried out by other groups (Halloran et al 1985, Cowell et al 1987b). To date, no recombination events have been observed between the two loci and the cumulative lod score is 13.69 (Cowell et al 1987b) at a recombination fraction of zero, assuming a penetrance of 0.9. ESD allele frequencies, however, showed wide variations in different populations (Paphia and Nahar 1977). In Caucasian families, the low frequency of the ESD2 allele - 0.116 (Cowell et al 1987b)- indicated that this polymorphism is of little value in pre-natal and pre-morbid screening of Rb. In South and East Asia populations ESD2 has a higher frequency - 0.30-0.40 (Paphia and Nahar 1977)- so the polymorphism is of greater use in these populations.

The ESD gene was subsequently cloned by Squire et al (1986) and restriction enzyme analysis identified an Apa I restriction fragment length polymorphism (RFLP) (Section 1.4.6.1). In this case the frequency of the rarer allele was found to be 0.2 (heterozygous frequency 32%) and, hence, slightly more useful than the ESD protein polymorphism for mutant gene carrier detection.

In summary, therefore, cytogenetic analysis of chromosome deletion cases pointed to 13q14 as the location of the gene predisposing to Rb. These deletion cases included loss of the adjacent ESD gene (Sparkes et al 1980) which was then used as a marker in linkage studies for the inheritance of the non-deletion form. ESD linkage studies located the hereditary, non-deletion form to the same region of chromosome 13.

1.4.1.3 The retinoblastoma gene is recessive

Linkage analysis confirmed the hereditary predisposition to Rb resulted from the mutation of a gene in 13q14. For sporadic Rb only circumstantial evidence implicating 13q14 was available. A deletion including region 13q14 was observed in tumours cells from patients with both sporadic and hereditary Rb (Balaban et al 1982, Benedict et al 1983a). This suggested that the same Rb locus was involved in the tumorigenic events in both forms of the tumour. With the identification of deletion forms of Rb, the basis for the 'first event' of Knudson's hypothesis was established. It was predicted that the second
event entailed the loss or inactivation of the same gene in the homologous chromosome (Comings 1973, Knudson 1978). The ESD gene was used by two groups to study this suggestion (Godbout et al 1983, Benedict et al 1983b). In one study a patient had a constitutional ESD activity only 50% of normal, suggesting a chromosomal deletion (Benedict et al 1983b). The tumour from this patient contained only one chromosome 13 and no ESD activity, indicating loss of the normal chromosome 13 within the tumour. Therefore, there had been a complete loss of genetic information encoded at the 13q14 locus providing direct confirmation of Knudson's prediction of Rb gene as a recessive cancer gene. In another study it was found that, in some patients who were constitutionally heterozygous at the ESD locus, the tumour was hemizygous or homozygous (Godbout et al 1983) again pointing to the recessive nature of the Rb gene. The same "loss of heterozygosity" (LOH) was found using DNA probes and RFLP analysis (Cavenee et al 1983, Dryja et al 1984a). This finding prompted the question of a genetic mechanism that would permit the phenotypic expression of recessive mutant alleles, i.e. the mechanism involved in the second event. It was suggested that a second event in a retinal cell of a predisposed individual resulting in homozygosity for the mutant allele, i.e. mutant at the Rb locus on both chromosome 13 homologues, would lead to the development of a tumour clone (Cavenee et al 1983). Figure 1 outlines the six possible chromosomal mechanisms which were suggested by Cavenee et al (1983) to account for the functional inactivation of the Rb gene.

Of these mechanisms mitotic non-disjunction with loss of the "normal" chromosome (figure 1b) will result in hemizygosity at all loci on chromosome 13, whereas mitotic non-disjunction with reduplication of the "mutant" chromosome (figure 1c) will result in homozygosity at all loci on chromosome 13. Mitotic recombination (figure 1d) will give rise to homozygosity around the Rb locus and heterozygosity for other proximal or distal loci depending on the exact site of the crossover. Gene conversion (a non-reciprocal recombination event) (figure 1e) might also give rise to LOH, although usually in a smaller region. Two other mechanisms are deletion (figure 1f) and point mutation (figure 1g) of the wild-type Rb allele. Both non-hereditary and hereditary Rb could arise through these mechanisms, the difference being two somatic events occur in the former and one germinal and one somatic event occur in the latter case.
**Figure 1.1** Chromosomal mechanisms of a second event that will allow the phenotypic expression of a recessive mutant allele at the Rb locus. (a) the chromosome 13 homologues of a predisposed individual carrying the mutant Rb allele (rb) in all his somatic cells. (b)-(g) illustrate the possible mechanisms leading to the loss or mutation of the remaining wild type allele (+) in a retinal cell giving rise to a tumour cell (Cavenee et al 1983).
To test these hypotheses, cloned DNA segments from various loci on chromosome 13 were analysed. Each probe identified a polymorphic restriction enzyme site (Cavenee et al 1983, Dryja et al 1984a). Using such RFLPs, which lay proximal as well as distal to the 13q14 region, it was possible to study certain chromosomal events occurring in tumour initiation in individuals who were heterozygous in constitutional cells. LOH of a particular marker indicated that one or both copies of chromosome 13 was lost in the tumour cells.

One patient, originally described by Benedict et al (1983b) and whose constitutional cells exhibited half the ESD activity in his tumour cells and karyotyping indicated that only one chromosome 13 was present. The interpretation was that the patient carried a constitutional deletion of the Rb and ESD loci which was invisible in the light microscope and the tumour contained only the deleted chromosome 13. Development of hemizygosity at two loci flanking the ESD and Rb loci was shown confirming the complete loss within the tumour of the normal chromosome with retention of the deleted chromosome 13. In another patient the tumour was shown to develop homozygosity for all loci tested. Although both alleles were present constitutionally, only one allele at each locus was present in the tumour and in each case this was derived from the paternally inherited chromosome 13. This data was consistent with the complete loss of the entire maternally derived chromosome, with reduplication of the paternally derived chromosome. Since the father was unaffected but the patient had bilateral tumours, the retained paternal chromosome was presumed to contain a \textit{de novo} germinal mutation. In a third case, there was evidence of mitotic recombination. This patient was heterozygous both at the ESD locus and for the quinacrine staining satellite heteromorphism on the short arm of chromosome 13. The tumour cells were heterozygous for the satellite region and homozygous for ESD, suggesting that both copies of chromosome 13 were present, but that a recombination event had occurred proximal to the ESD and Rb loci resulting in homozygosity at these loci. This suggestion was confirmed using many probes identifying RFLPs (Cavenee et al 1983). This was the first demonstration of the phenomenon of mitotic recombination in human cells. In 2 other cases tumours were shown to be homozygous for markers in the q14 region indicating either chromosome reduplication or mitotic recombination. In 3 further cases, however, tumours were shown to maintain
heterozygosity for q14 markers indicating that gross chromosomal mechanisms are not universally applicable and point mutations, small deletions or gene conversion could also be involved in inactivating the Rb gene. Other groups also performed similar studies. Dryja \textit{et al} (1984a) demonstrated the development of homozygosity in 4 out of 8 tumours but found no correlation between the development of homozygosity and the degree of histopathologic differentiation of the tumour. In all groups of patients described by Benedict \textit{et al} (1983b), Cavenee \textit{et al} (1983) and Dryja \textit{et al} (1984a) heterozygosity at loci other than those on chromosome 13 was retained. Finally Cavenee \textit{et al} (1985) demonstrated in two patients that in tumours showing LOH the chromosome retained in the tumour was derived from the affected parent. This established unequivocally that the phenomenon of LOH results in the duplication of the predisposing mutation. By inference, LOH in sporadic tumours points to the duplication of the first hit.

Later on Benedict \textit{et al} (1987) reported at least partial homozygosity for a single chromosome 13 in 75\% of the 16 tumours they studied and more recently, Zhu \textit{et al} (1992) demonstrated LOH in 63\% of the 30 Rb tumours studied. When mechanisms of LOH were investigated in 13 of these tumours, it was found that mitotic recombination was the most common (46\%) mechanism. LOH developed by non-disjunction with reduplication in 5 tumours (39\%), a small deletion in one tumour (8\%) and undefined rearrangement in one other (Zhu \textit{et al} 1992).

1.4.2 Cloning of the retinoblastoma gene

All of the genetic analysis pointed to the presence of a single gene in chromosome region 13q14. This fact drew several groups to the idea of cloning the Rb gene using the methods of "reverse genetics" (Botstein \textit{et al} 1980) which involves finding a gene whose location is known but whose product is not.

In an attempt to clone a candidate Rb gene, several groups isolated DNA sequences from chromosome 13 either by flow-sorting (Lalande \textit{et al} 1984) or by constructing libraries from somatic cell hybrids containing human chromosome 13 (Cavenee \textit{et al} 1984, Dryja \textit{et al} 1984b, Scheffer \textit{et al} 1986). Unique sequence probes isolated in these ways were then localised on
chromosome 13 either; by in situ hybridisation to metaphase chromosomes (Drijya and Morton 1985), identifying segregation patterns in panels of somatic cell hybrids with chromosome 13 - deletions (Scheffer et al 1986), determination of locus dosage in constitutional cells from Rb patients carrying cytogenetically visible deletions (Cavenee et al 1984) and linkage studies to create a genetic map of the chromosome (Leppert et al 1986). One DNA sequence, H3-8, isolated from a flow sorted library (Lalande et al 1984) was localized to 13q14 using somatic cell hybrid mapping panels (Lalande et al 1986). This probe detected a homozygous deletion in 3 of 37 different tumours studied (Dryja et al 1986a). H3-8 was then used to isolate larger, overlapping segments of DNA by chromosome walking. One of these segments detected a conserved DNA sequence - p7H30.7R - which was used to screen mRNA from adenovirus 12 transformed fetal retinal cells. A single transcript, 4.7 kb long, was identified and a cDNA, 4.7R, was isolated from this library (Friend et al 1986). Following essentially the same strategy others also independently isolated Rb cDNA clones (Lee et al 1987a, Fung et al 1987). Curiously, this gene, named RB1, was found to be expressed in all tissues examined which was somewhat surprising at the time since the expectation was that it was important in the control of retinal precursor cell differentiation.

Two distinct pieces of evidence were presented to confirm the identity of the isolated cDNA clones. Firstly, the 4.7 kb mRNA transcript was found in normal tissues but was often (20-30%) missing in Rb tumours supporting its involvement in Rb tumour initiation (Friend et al 1986, Lee et al 1987a, Fung et al 1987). Others (Goddard et al 1988) failed to find such a high frequency of abnormal transcripts. Secondly, gross structural rearrangements (mostly deletions) involving RB1 were found in approximately 12-30% of Rb tumours using Southern and/or Northern blot analysis (Friend et al 1986, Friend et al 1987, Lee et al 1987a, Fung et al 1987, Goddard et al 1988). Some of these deletions were contained within the gene, excluding the possibility that some other neighbouring gene(s) might be involved. One of strongest evidence for the authenticity of 4.7R was provided by Mitchell and Cowell (1989) who studied a constitutional 1;13 translocation from an Rb patient isolated in somatic cell hybrids and showed that the 13 breakpoint lay within the cloned gene. The translocation was associated with a small intragenic deletion which included sequences from 4.7R. This was the first demonstration that a predisposing constitutional mutation involved the RB1
gene. Similar rearrangements were also reported by Higgins et al (1989) using pulse field gel electrophoresis. Eventually subtle mutations including point mutations were identified using RNase protection assays (Dunn et al 1988, Dunn et al 1989) or by direct sequencing of parts of the RB1 gene (Yandell et al 1989). Yandell et al (1989) identified point mutations in seven Rb tumours, four of which representing non-hereditary cases because the mutation was identified only in the tumour DNA. In the remaining three cases the mutation was identified both in the tumour and in constitutional DNA from the patients. In these three cases, the mutation was not present in either parent and hence represent new germline mutations. Finally the two-hit hypothesis was confirmed by analysis of causative mutations in RB1.

The absolute proof of identity of the RB1 gene demanded functional analysis which demonstrated that the cloned gene suppressed tumorigenicity in neoplastic cell lines. Consequently it was shown that a functional RB1 gene under the control of a retroviral promoter could be introduced into cell lines lacking the RB1 gene. As a result tumorigenicity was suppressed in immunosuppressed mice (Huang et al 1988, Bookstein et al 1990a, Sumegi et al 1990, Madreperla et al 1991). Others, however, have failed to confirm this observation (Xu et al 1991, Muncaster et al 1992) casting doubt on the reliability of nude mouse assay systems. As pointed out by Xu et al (1991), however, none of these studies utilised the RB1 promoter or other regulatory elements that might be adjacent to the RB1 coding region in vivo. At the time of writing their 1991 paper, Xu and co-workers were transferring a normal chromosome 13 by microcell fusion into each of the cell types used in the previous studies and the preliminary results indicated suppression of tumorigenicity is complete in some of these cell types. Although suppression of tumorigenicity was not complete in some studies, introduction of wild type RB1 into cells lacking RB1 expression, nevertheless re-established some aspects of growth control.

The RB1 gene was the first cancer gene to be cloned using reverse genetics. Other human disease genes that have been cloned by variations of this approach include those for X-linked chronic granulomatous disease (Royer-Pokora et al 1986), Duchenne muscular dystrophy (DMD) (Monaco et al 1986), Cystic fibrosis (CF) (Riordan et al 1989, Kerem et al 1989), DCC (for deleted in colon carcinoma) (Fearon et al 1990), Wilms' tumour (Call et al 1990, Gessler et al 1990), von Recklinghausen's neurofibromatosis (Wallace

1.4.3 RB1 involvement in other tumours

The fact that LOH for 13q was observed in tumours other than Rb (Dryja et al 1986b) led many researchers to look for structural abnormalities of RB1 in other cancers. Indeed structural alterations of the RB1 gene have been detected in osteosarcoma (Friend et al 1986, Friend et al 1987, Lee et al 1987a, Fung et al 1987, Toguchida et al 1988, Weichselbaum et al 1988), breast carcinoma (Lee et al 1988, Varley et al 1989, Bookstein et al 1989), small cell carcinoma of the lung (Harbour et al 1988, Yokota et al 1988, Mori et al 1990, Hensel et al 1990), soft tissue sarcomas (Friend et al 1987, Weichselbaum et al 1988, Stratton et al 1989, Reissmann et al 1989), bladder carcinoma (Horowitz et al 1989, 1990, Cairns et al 1991, Ishikawa et al 1991), prostate carcinoma (Bookstein et al 1990a, 1990b) and leukaemia (Hansen et al 1990, Liu et al 1992). It appears, therefore, that the RB1 gene is important in tumorigenesis in many different tissues, although, since only a small percentage of these other tumours (except for osteosarcoma, Toguchida et al 1988, 1989) show RB1 abnormalities they more likely represent progression events rather than being related to causality. In fact, with the understanding of the role of the RB1 gene (Section 1.4.5) it became clear how RB1 might contribute to the malignant phenotype in a variety of cell types.

1.4.4 Structure and genomic organisation of RB1

RB1 is a fairly large and complex gene extending over approximately 200 kb of DNA (figure 1.2) for which a restriction map was constructed (Wiggs et al 1988). McGee et al (1989) reported 9.2% of the RB1 sequence including approximately 200 bp of intron sequence immediately flanking each exon. There are 27 exons (figure 1.2) ranging in size from 31 bp to 1873 bp (McGee et al 1989, Hong et al 1989) and 26 introns ranging from 80 bp to 70500 bp (McGee et al 1989). The exons are clustered into three groups separated by the two largest introns (McGee et al 1989).
Figure 1.2 A schematic representation of the RBI gene. The vertical bars indicate the approximate positions of the exons (numbered 1 to 27) which are clustered into three groups separated by the two largest (35 kb and 70 kb) introns (V).

The 5' region of the RBI gene is extremely G+C rich (McGee et al 1989); a G+C content of 70% was observed between the nucleotide position -568 (568 bp upstream from the initiation methionine codon) to the end of exon 1 (+137) (T'Ang et al 1989). The region contains many possible Hpa II sites. These criteria are consistent with the presence of an HTF (Hpa II tiny fragments) island (McGee et al 1989, Ford et al 1990), although transcriptional motifs such as CCAAT or TATA boxes, typically found in eukaryotic promoters and involved in binding of the CAT/Enhancer Binding Protein (C/EBP) and positioning of RNA polymerase II respectively, were not found at their characteristic positions (McGee et al 1989, Hong et al 1989, T'Ang et al 1989). However, sequences showing homology to the binding site for the Sp1 transcription factor have been identified in this region (Hong et al 1989, T'Ang et al 1989) which implied that this region contains a promoter for RBI transcription. The regions that are required for promoter activity have been defined by expression assays of deletion mutants (Hong et al 1989, Sakai et al 1991a) and the sequence extending from 196 to 249 (Hong et al 1989) or 185 to 206 (Sakai et al 1991a) bases upstream of the initiation methionine codon were found to be essential for transcription. One of the sequences containing homology to the Sp1 binding site showed weak binding to Sp1 itself but was found to associate with another nuclear factor, named RBF-1 (retinoblastoma binding factor 1), whose function is not known (Sakai et al 1991a). A sequence showing homology to the binding sites for the ATF/CREB (activating transcription factor/cAMP response element binding protein) family of transcription factors is also identified within the putative RBI promoter region and was found to be located adjacent to the RBF-1 and Sp1 binding site (Sakai et al 1991a). The 3' end of the RBI gene
contains the usual polyadenylation signal sequence AATAAA. 31 bp
downstream of this hexamer, another conserved consensus sequence generally
found 5-30 nucleotide downstream of the polyadenylation signal sequence is
also present (McGee et al 1989). These sequences are involved in
transcription termination.

Exons 17, 18, 20 and 21 of the RB1 gene encode potential 'zinc finger'
motifs (His/Cys pairs in proteins) characteristic of proteins with DNA binding
ability (Hong et al 1989) and exon 20 contains a 'leucine zipper' motif
associated with protein-protein interactions (McGee et al 1989, Hong et al
1989). Functional importance of these motifs in the RB1 gene, however, is
not known.

The RB1 gene contains many direct repeats, inverted repeats and dyad
symmetry elements (T'Ang et al 1989). It is possible that the existence of
these structures, which can be potential recombination hot spots (direct
repeats) and which can potentially form stem and loop structures (inverted
repeats and dyad symmetry elements), promote structural rearrangements.
Alu-like repetitive sequences were identified in introns 2, 8, 11, 13, 14 and
17 (McGee et al 1989). The Alu sequence located in intron 2 contains two
sequences that are highly conserved in Alu repetitive sequences. The first is
a sequence (5'-GAGGCNGAGGC) corresponding to the T-antigen binding
sequence of the SV40 replication origin and the second one is a symmetrical
sequence (5'-CCAGCCTGG) of no known function (McGee et al 1989). The
short symmetrical sequence is also present in both of the Alu sequences in
introns 13 and 14. The short intron (intron 14) which separates exon 14 and
15 is almost entirely composed of Alu sequence suggesting that, at one time,
exons 14 and 15 were possibly a single exon and were divided by the
insertion of an Alu element during evolution (McGee et al 1989).

In October 1991, the sequencing of the entire RB1 gene, which was started in
1989 by the Howe Laboratory of Ophthalmology Group (Harvard Medical
School), has been completed (DW Yandell, personal communication). The
180389 bp sequenced began with G in the EcoRI site 2059 bp 5' to the first
base of the ATG codon and continued to 3008 bp to the last base of TGA stop
codon in exon 27. Although the sequence obtained is still being analysed,
some additional information has already been obtained (J Toguchida, DW
Yandell, personal communication). The examination of the nucleotide
frequencies of the 180389 bp sequence gave a value of 1.7, for the 
A+T/G+C ratio, which is higher than the average value reported in the 
human genome (1.52) indicating that this locus is A+T rich (Ciccarelli et al
1991). This result may support the theory that RB1 lies within the G-band of
13q14.2, rather than in the R-band of 13q14.1 since R bands are presumed to
be GC rich. The highest density of CpG dinucleotide pairs was observed
surrounding exon 1 where the transcription control elements were identified
(Sakai et al 1991a). A relatively large region (approximately 3kb) rich in
CpG sequences was also observed in intron 2, although the data did not
directly suggest the presence of regulatory elements or additional
transcriptional units in this area. Additional Alu-like sequences (>75%
homology to Alu consensus sequence) were identified (bringing the total to
48) and most were found to cluster between exons 7-17 (a 20 kb region) which
might explain the frequent occurrence of deletions in this region (see
discussion Section 4.3.2). 14 KpnI-like sequences, 21 mononucleotide
[poly(A)] repeats (>15 bases) and 3 dinucleotide repeats [(CA)n] were also
identified. The (CA)n repeats (>10 repeats) were found in introns 2, 14 and
17. Intron 17 also contains a 50-53 bp VNTR sequence which was identified
by Wiggs et al (1988). Another VNTR sequence, which consists of
consecutive repeats of the sequence CTTT(T), is present in intron 21
(Yandell and Dryja 1989). Both of these VNTRs constitute useful DNA
polymorphisms as the number of repeat units vary between different
individuals (see Sections 3.1.1 and 3.1.2).

1.4.5 Function of RB1

The identification and cloning of the RB1 gene meant that its gene product
(pRB) could be isolated and monoclonal antibodies prepared. In Western
blots several species of protein ranging in size from 105-115 kD were
identified. pRB was present only when both genomic structure and mRNA
transcript appeared normal indicating that the absence of normal functional
pRB is associated with tumour formation (Lee et al 1987a, Horowitz et al
1990). The higher molecular weight species of pRB were shown to represent
highly phosphorylated, mainly on serine and threonine residues, proteins and
those with lower molecular weights represented the less well phosphorylated
or underphosphorylated proteins (Lee et al 1987b, Whyte et al 1988, Ludlow
et al 1989). It was subsequently shown that phosphorylation of pRB was cell
cycle dependent (Chen et al 1989, DeCaprio et al 1989, Buchkovich et al 1989, Mihara et al 1989). The underphosphorylated forms of pRB were found predominantly in quiescent cells whereas the phosphorylated forms appeared towards the onset of DNA synthesis (Chen et al 1989, DeCaprio et al 1989, Buchkovich et al 1989). The appearance of highly phosphorylated forms coincided with the transition of the cell from G1 to S-phase of the cell cycle (DeCaprio et al 1989, Buchkovich et al 1989). It was demonstrated that the injection of excess underphosphorylated pRB into cells in early G1 blocks their entrance into S phase (Goodrich et al 1991). It has also been shown that a complex containing p34cdc2 kinase, a serine/threonine kinase which is required for cell cycle progression in all eukaryotic cells (Booher and Beach 1986, Draetta and Beach 1988), phosphorylates sites on pRB in vitro similar to those phosphorylated in vivo (Lin et al 1991, Lees et al 1991). More recently, an actual physical association between pRB and p34cdc2 or a closely related kinase has been reported (Hu et al 1992) confirming the suggestion that this kinase is one of the major regulators of pRB. These observations strongly suggested that pRB was involved in the control of the cell cycle (DeCaprio et al 1989, Goodrich et al 1991) the underphosphorylated form promoting quiescence. Although pRB was shown to have DNA binding ability (Lee et al 1987b) and has a 'zinc finger' motif associated with DNA binding ability (Hong et al 1989), it was not shown so far to bind directly to any specific DNA sequences. pRB was therefore thought to exert any regulatory control through interactions with other cellular proteins. The observation that underphosphorylated pRB in G1 cells was bound tightly to some unidentified nuclear structure suggested that phosphorylation apparently regulates the association of pRB with some nuclear protein(s) and that this binding serves to attach pRB tightly to the nuclear structure (Mittnacht and Weinberg 1991, Templeton et al 1991). In fact pRB has been shown to associate with many cellular and viral proteins (Whyte et al 1988, DeCaprio et al 1988, Dyson et al 1989, Ludlow et al 1989, Bagchi et al 1991, Kaelin et al 1991). Particularly interesting were the viral transforming proteins, E1A of adenovirus, large T-antigen (LT) of SV40 and E7 protein of human papilloma virus, all of which bind pRB (Whyte et al 1988, DeCaprio et al 1988, Dyson et al 1989). These transforming proteins share conserved regions that are necessary for their transforming function and mutations in these regions were shown not only to prevent cellular transformation but also binding to pRB (Whyte et al 1988). As such, these observations also established a link between "dominant" and
"recessive" oncogenes. LT bound specifically to the underphosphorylated form of pRB present only during the G1-phase of the cell cycle (Ludlow et al 1989). All of these viruses need to induce cell division in order to replicate themselves. Thus, the suggestion was that by sequestering the underphosphorylated form of pRB suppression of cell growth was lifted allowing the cell to enter S-phase. Two non-contiguous regions of pRB were shown, by deletion analysis, to be necessary for binding E1A and LT (Hu et al 1990, Huang et al 1990) and were located in that part of the gene coding for amino acids 393-572 and 646-772 representing exons 13-17 and 18-22 respectively. It has been suggested that, with appropriate folding of the protein, a "pocket" is created which facilitates binding to the viral transforming proteins as well as a specific set of endogenous cellular proteins (Kaelin et al 1990). Many such proteins were later identified (Kaelin et al 1991, Huang et al 1991, Bandara and La Thangue 1991, Bagchi et al 1991, Chittenden et al 1991, Defoe-Jones et al 1991). It was shown that RBI mutations in the pocket produced defective proteins which failed to bind viral proteins or associate with cellular proteins (Kaelin et al 1990, 1991, Templeton et al 1991, Mittnacht and Weinberg 1991, Chittenden et al 1991). With the characterisation of some of the RBI associated proteins as transcription factors (Bagchi et al 1991, Chittenden et al 1991, Bandara et al 1991, Partridge and La Thangue 1991) RBI has been implicated in transcriptional control of certain genes. Recent reports by Lee et al (1992), Jacks et al (1992) and Clarke et al (1992) who investigated the effects of RBI mutations in the mouse shed new light on the role of RBI in development and in tumorigenesis and strengthened the belief that RBI is involved in a signalling pathway controlling cell proliferation and differentiation. Thus, mutation analysis of the RBI gene contributed substantially to our understanding of the function of the RBI gene and this is discussed at more length in Section 4.3.3 of the discussion.

1.4.6 RBI - linkage and clinical applications

1.4.6.1 Linkage analysis - theory and applications

The phenomenon of linkage is based on the fact that genes that are carried on the same chromosome close to each other tend to be inherited together. There is, however, exchange of material between homologous pairs of
chromosomes at meiosis which is termed "crossing-over" or "recombination" resulting in independent segregation of linked loci. The proportion of such recombination events out of all opportunities for recombination is called the recombination fraction (\( \Theta \)). For two gene loci (with alleles A, a and B, b) that are unlinked, the four possible gametes (AB, Ab, aB, ab) are transmitted in an approximate 1:1:1:1 ratio. According to this ratio, for such unlinked loci, the recombination fraction, \( \Theta \), is equal to 50\%, while for linked loci it is less than 50\% and the exact value depends on the distance between the loci. This means that linked loci are usually separated by a small physical distance when \( \Theta \) is close to zero and a greater physical distance when \( \Theta \) is larger. The recombination fraction, therefore, is taken as a measure of genetic distance. The genetic and physical distance are not linearly related as the genetic distance represented by 50\% \( \Theta \) corresponds to an infinite physical distance between two loci. By use of mapping functions which relate genetic to physical distance, however, \( \Theta \) (in percentage) is translated into map distance, measured in centimorgans (cM). 1\% recombination represents 1 cM and 1 cM is generally accepted to represent approximately \( 10^6 \) bp in man. The recombination frequency in females is higher than in males and it might be necessary to treat the data from these two groups separately especially in cases of loose linkage. There are also hot spots for recombination on some chromosomes, for example, on the tip of the X chromosome, where closely linked loci give high \( \Theta \) values.

The primary objective of a linkage analysis between two loci is to determine whether the value of \( \Theta \) is smaller than 50\%. Linkage analysis is made simpler if large three generation families can be identified. At least one affected parent should be heterozygous for the two loci under investigation. In this case the allele associated with the phenotype can be determined i.e. phase-known. For example, if an individual is doubly heterozygous for two loci A and B with alleles A1, A2 and B1, B2, two linkage phases are possible. One chromosome may carry A1B1 and the other A2B2 or one chromosome may carry A1B2 and the other A2B1. This can only be established by analysis of parents of such an individual. The number of three generation families available for analysis is limited, however, and data from two generation families need to be used in which the phase of the doubly heterozygous individual is unknown. The use of data from such families has been made possible by the lod score method of linkage analysis (Morton 1955). This method calculates the likelihood (L) of observing the phenotypes
in a given family if the test loci are linked with a recombination fraction \( \Theta \) (\( \Theta = 0 \) to 0.45) and compares it with that of observing the phenotypes if the loci are not linked (\( \Theta = 0.5 \)). The ratio of the two probabilities (\( L(\Theta)/L(0.5) \)) gives the odds for, or against, linkage and, taking the logarithm of the odds (lod score, \( Z = \log_{10} L(\Theta)/L(0.5) \)), allows data from different pedigrees to be combined by simple addition. The lods are calculated for \( \Theta \) from 0 to 0.45 and the value of \( \Theta \) with the highest lod score is taken as the most likely estimate of the recombination frequency between the two loci. The results of linkage analysis from different sources may be combined by simply adding the lod scores for a given \( \Theta \) value. When the total lod score, \( Z(\Theta) \), reaches or exceeds +3, corresponding to odds for linkage of at least 1000:1, linkage is considered significant. When \( Z(\Theta) < -2 \), linkage is said to be excluded at that \( \Theta \) value.

In practice linkage analysis requires polymorphic loci where at least two different alleles occur with an appreciable frequency. When frequencies are close to zero or one, a locus is not very useful since then there is a good chance that the parents in any one family will be homozygous for the same allele. For a parental mating to be informative for linkage, however, at least one parent must be doubly heterozygous. The more polymorphic a locus, the more suitable it is for linkage analysis. One of the ways to measure degree of usefulness of the polymorphism is the polymorphism information content (PIC) (Botstein et al 1980). The PIC is the probability that a given offspring of a random mating between a carrier of a rare dominant gene and a non-carrier is informative for linkage between the locus of the dominant gene and a marker locus. The PIC is calculated as

\[
1 - \left( \sum \frac{n-1}{n} \right) - \sum \frac{2}{n} \left( \frac{P_i^2}{P_j^2} \right)
\]

where \( P_i \) is the frequency of the \( i \)th allele and \( n \) is the number of alleles at the marker locus.

Early linkage studies depended on easily detectable markers such as blood groups (Mohr 1951). Many enzymes also have naturally occurring alleles which show variation of mobility after electrophoresis through starch gels. Important in the early analysis of Rb families was the polymorphic esterase D enzyme (Sparkes et al 1983). Cytogenetic variations such as regions of
variable heterochromatin is another form of polymorphism (Donahue et al 1968). The most useful kind of polymorphisms, however, are the natural genetic variations in DNA sequences between individuals. Some of these natural sequence variations resulting from point mutations, small deletions or insertions lead to the creation or loss of cleavage sites for restriction endonucleases giving rise to restriction fragment length polymorphisms (RFLPs) (Botstein et al 1980). Another source of variation is the presence of variable number of tandem repeat (VNTR) sequences in the DNA (Jeffreys et al 1985, Nakamura et al 1987). VNTR polymorphisms has been proved to be the most useful as they give rise to many different alleles at a particular locus as opposed to the existence of only two different alleles with other polymorphisms.

The discovery of these most useful polymorphisms followed the advent of recombinant DNA technology which made the isolation of DNA fragments from various chromosomal regions possible. These DNA fragments are used as "probes" which become genetic markers when they detect genetic variations, mutations and polymorphisms, at the level of the DNA sequence. The increasing availability of polymorphic probes has increased the potential of linkage analysis making it a useful tool in mapping of the human genome and in providing improved risk estimates in genetic counselling through "gene-tracking".

Gene-tracking is one particular, and now widespread, application of linkage analysis which was made possible by the isolation of DNA fragments from regions of chromosomes thought to be important in the development of particular diseases. When such isolated sequences identify RFLPs, they can be used as genetic markers to track disease related genes through successive generations in affected families, thus making it possible to offer pre-natal and pre-morbid diagnosis or prediction to those families and they may also lead to the isolation of the relevant gene itself (Section 1.4.2). Predictive genetic tests based on RFLP analysis could be performed for a number of genetic diseases including Duchenne muscular dystrophy (Davies 1985, Hejtmancik et al 1986), cystic fibrosis (Wainwright et al 1985) and Rb (Cavenee et al 1986) even before the relevant genes were cloned.
1.4.6.2 RBI, the clinical applications

One of the aims of localising and cloning any disease related gene is to establish diagnostic (predictive) DNA tests. These tests could be direct identifying the actual causative defects, or indirect as in the case of standard linkage analysis (Section 1.4.6.1) used for many hereditary conditions. Since the localisation of the Rb gene to 13q14, every opportunity has been taken to establish a test to identify mutant gene carriers. Before the cloning of the RBI gene, several means of identifying gene carriers were used including: karyotype analysis (Vogel 1979, Dryja et al 1983, Cowell et al 1986) looking for predisposing structural rearrangements; measurement of ESD levels (Sparkes et al 1980, Cowell et al 1989) to detect deletion carriers; linkage analysis using the ESD protein polymorphism (Sparkes et al 1983, Halloran et al 1985, Cowell et al 1987b); linkage analysis using the Apa I RFLP within the ESD gene (Squire et al 1986) and RFLPs using other chromosome 13 markers (Cavenee et al 1986). None of these tests, however, proved useful for extensive genetic screening for Rb carrier status. As explained in Sections 1.4.1.1 and 1.4.1.2, few Rb patients carry large constitutional chromosome deletions and the low frequency of the rare allele in ESD protein and ESD Apa I polymorphism limited their use to only a few informative families. The problem with using flanking markers which lie some distance from the RBI gene is the possibility of recombination occurring. A recombination event between ESD and Rb locus has never been observed but recombination has been a problem in the use of more distant markers. The very first study of pre-natal prediction in Rb (Cavenee et al 1986), using flanking markers for example, met with limited success due to recombination events between marker and Rb locus.

With the cloning of the RBI gene in 1986 (Friend et al 1986, Section 1.4.2) genetic screening for carrier status became a real possibility (Cowell and Onadim 1990). Horstemke et al (1987a), using a genomic clone, H3-8 (Lalande et al 1984), which lay within the RBI gene, detected submicroscopic deletions in Rb families. They also reported a Hind III polymorphism detected by H3-8 (Horstemke et al 1987b) but only 2/6 families were informative (Greger et al 1988). Clearly, for reliable genetic screening, polymorphic probes with a higher PIC values were required. Although the RBI cDNA could be used to detect deletions and rearrangements, it did not identify any RFLPs. Eventually a series of
intragenic unique sequence DNA probes were isolated which recognise RFLPs (Wiggs et al 1988, Bookstein et al 1988). These probes have formed the basis of linkage studies in the analysis of Rb families worldwide and have been successfully applied in 80-90% of families (Wiggs et al 1988, Scheffer et al 1989, Goddard et al 1990, Onadim et al 1990). The first report of a pre-natal screening carried out using these intragenic probes came from Mitchell et al (1988). Several other polymorphisms have since been identified in the RB1 gene (Yandell and Dryja 1989) and added to the armoury of tests available for family studies.

1.4.7 The pattern of inheritance and variable phenotypic expression in retinoblastoma

The patterns of inheritance of Rb and the phenotypic expression of the mutant gene shows considerable heterogeneity which can complicate the analysis of hereditary predisposition of Rb. Apart from vertical transmission, many cases of affected sibs with unaffected parents have been described. Indeed, the very first report of familial Rb was of this kind (Lerche 1821). Families with unaffected transmitting members were first reported in the second half of the 19th century (Von Graefe 1868, Thompson and Knapp 1874 reviewed by Weller 1941). These reports were followed by other examples where transmission of the predisposing gene has apparently skipped one generation (Vogel 1957 reviewed in Vogel 1979, Migdal 1976). In other families distant relatives are affected by Rb with no apparent family history (Macklin 1960). Such families represent more than one fourth of all families in the series of Briard-Guillemot et al (1974 cited by Bonaiti-Pellie et al 1976, 1990). Part of the phenotypic heterogeneity is that Rb tumours can spontaneously regress (Section 1.3.4). In familial Rb these individuals are mutant gene carriers (Connolly 1983). We (Onadim et al 1991, Onadim et al 1992a) reported one family where expression of the phenotype ranges from asymptomatic gene carriers, regressed tumours, through unifocal to bilateral, multifocal lesions (Section 3.1.5, 3.1.7).

Some of the above mentioned patterns of inheritance can be explained by incomplete penetrance. Indeed, many pedigrees directly support incomplete penetrance, since they are characterised by apparent skipping of one generation. However, if one attempts to attribute the existence of families
with distant affected relatives and more than one unaffected carrier, and families with affected sibs from unaffected parents, to decreased penetrance, an inconsistency develops. Prior to the appearance of affected cases in a branch of a family penetrance is very low and after its occurrence penetrance is very high. Knudson's two-hit hypothesis (1971) states that the manifestation of Rb in gene carriers is due to a second somatic mutation occurring at random and at a constant rate (Section 1.2). The existence of such families is poorly compatible with this hypothesis as, according to it, a gene carrier would have a very small probability of somatic mutation never occurring in any of his retinal cells.

Another possible explanation, however, is that many of these pedigrees appear in the older literature (Vogel 1979). The carriers of new mutations in these families had mild form of the disease or were cases of incomplete penetrance. This follows since, due to the lack of effective treatment anyone with severe disease died. Thus, unaffected carriers are probably overrepresented in the old literature although every series of families has examples of these cases. Matsunaga (1978), however, has shown the pedigrees with Rb in remote relatives to be slightly more common than expected on the hypothesis of incomplete penetrance. On the other hand, Ellsworth (1969) and Matsunaga (1976) suggested that the penetrance could vary according to the manifestation in the carrier parent. Indeed, a correlation was found between expressivity (unilateral vs bilateral involvement) and penetrance (segregation ration among offspring). Both were found to increase with increasing degree of expressivity in the carrier parent (Kaelin 1955 reviewed by Vogel 1979, Matsunaga 1976, Bonaiti-Pellie and Briard-Guillemot 1981).

In an effort to explain the different patterns of inheritance and variable phenotypic expression in Rb, various hypotheses have been put forward over the years. One of these hypotheses suggests the existence of delayed mutation, originally proposed by Auerbach (1956, cited by Hermann 1977) for a type of split-hand. According to this hypothesis a premutated allele could be transmitted through one or several generations and then changed into a mutated allele ('telomutation', Hermann 1977) either in the retinal cells of a subject who would then have Rb or in his germ cells, producing affected offspring. Hermann (1977) postulated that this phenomenon could entirely explain the differences in expressivity and penetrance among gene carriers.
According to Matsunaga (1978), however, such differences would be due to different degrees of "host resistance" among gene carriers, a variable which would be partly genetically determined with polygenic inheritance and with two thresholds subdividing the population of gene carriers into three possible phenotypes; unaffected, unilaterally and bilaterally affected. Matsunaga (1978) argued against a Poisson distribution pattern of tumours in gene carriers and suggested the possibility that variation in the number of tumours might be determined by host resistance, while second somatic mutation might be a random event of the Poisson type. Bonaiti-Pellie and Briard-Guillemot (1981) considered both delayed mutation and host resistance models in explanation of their results. The heterogeneity ("low" and "high" transmitters) they found among unilaterally affected and unaffected cases favoured the delayed mutation hypothesis; low transmitters presumably bearing a premutation and high transmitters bearing the complete mutation. However, they also found that, among high transmitters, the risk for the offspring was still different between bilaterally affected individuals on the one hand and unilaterally and unaffected individuals on the other hand which was counter to the hypothesis of delayed mutation. This difference in penetrance among high transmitters according to the phenotype supported the hypothesis of host resistance. Bonaiti-Pellie and Briard-Guillemot (1981) also pointed out that the correlation found in expressivity and penetrance between parent and offspring argue against the hypothesis of Knudson (1971). However, they suggested that this hypothesis still holds under these circumstances if the somatic mutation rate is not constant but rather varies from individual to individual and that if this rate is itself partly genetically determined.

Mutational mosaicism is yet another model that was put forward to explain variable transmissibility and manifestation of Rb (Carlson and Desnick 1979). It involves the origin of mutations occurring during mitosis, whether in the parent at some stage prior to reproductive maturity or in the offspring at some time following fertilisation. The phenotypic expression and transmission of these new mutations are then dependent on the proportion of cells bearing the mutant gene as well as the location of these cells in somatic and/or germinal tissues. Motegi (1981, 1982) suggested that some sporadic bilateral Rb cases occur post-zygotically and that the population with sporadic bilateral Rb is etiologically heterogeneous, although statistical studies generally assumed that approximately 100% of sporadic bilateral Rb cases are caused by de novo germ cell mutation. 13q14 deletion mosaicism has been reported in the
literature for some Rb patients and sometimes also involving one parent (Motegi 1981, 1982, Michalova et al 1982, Ribeiro et al 1988). Transmission in some pedigrees, where there are many affected children and unaffected parents, can be explained by mutational mosaicism in one of the parents. Michalova et al (1982) in one particular case, found a small clone in the mother's lymphocytes with an interstitial 13q deletion which was identical to that found in the affected child. They suggested that this deleted chromosome could have been present in the germ cell and thus cause the non-mosaic chromosomal constitution in the child. Although, mutational mosaicism in germ line of one parent could explain transmission in some pedigrees it cannot adequately provide an explanation to the whole question of expressivity in Rb. It cannot, as pointed out by Matsunaga (1981), explain those cases where bilaterally affected parents have an unaffected carrier child who, in turn, produces bilaterally affected offspring.

Another possible explanation for familial transmission of Rb through healthy parents was provided by cytogenetic studies. Some authors (Strong et al 1981, Turleau et al 1983) reported cases of balanced chromosomal insertions. If the offspring inherit the unbalanced form they are either trisomic for 13q14 or carry a 13q14 deletion. Some will also inherit the balanced insertion. These observations account in some families for the apparent "lack of penetrance" not explained by simple Mendelian inheritance. Bonaiti-Pellie et al (1990) tested the hypothesis that such a mechanism may explain all observations on Rb segregation among families using likelihood computations on nine series of Rb reported in the literature. It was found that this mechanism could not entirely explain the familial data even if a reduced viability of unbalanced gametes (i.e. gametes with a 13q14 deletion or insertion) are allowed for. They concluded that this is probably a rare event and other mechanisms must be involved. Bonaiti-Pellie et al (1990) considered two hypotheses which have also been suggested by other authors in one form or the other over the years. The first one states that, the probability of a second event, and consequently of tumour formation for a gene carrier, could be genetically determined depending on loci other than the Rb locus. In this hypothesis, which is similar to the host resistance theory of Matsunaga (1978), the probability of tumour formation in a gene carrier would depend on the genotype of both parents and would vary from one generation to another. They investigated this hypothesis by studying the proportion of cancer deaths in grandparents of Rb patients (Bonaiti-Pellie and
Briard-Guillemot 1980) and found an excess of cancer deaths. Similar findings were reported by other authors (DerKinderen et al 1988, Sanders et al 1989). Bonaiti-Pellie and Briard-Guillemot (1980) attributed this excess to the existence of a non-specific factor of susceptibility to cancer. Bonaiti-Pellie et al (1990) argued that this factor acted on the first mutation and had no effect on the rate of the second mutations as the excess of cancer deaths was found on both sides of the family in unilateral isolated cases and only on one side of the family in bilateral isolated cases. Also, the excess of cancer deaths was not higher in families with a high degree of penetrance than in families with several unaffected carriers (Bonaiti-Pellie and Briard-Guillemot 1980, Matsunaga 1980a).

The second hypothesis states that the probability of tumour formation could depend on the nature of the germinal mutation (Bonaiti-Pellie et al 1990). If, for example, a particular mutation induces a greater probability of somatic recombination, or of abnormal chromosome segregation in mitosis, then the probability of tumour formation will be greater for that particular mutation carrier. Similar suggestions involving the nature of the germinal mutations have been made by other authors. Knudson (1983a) postulated that if an allele, lethal at the cellular level, was present at some other locus on the chromosome 13 homologue containing the Rb mutation, the occurrence of a deletion including this locus at the other chromosome 13 copy would lead to cell death and no tumour would result. In families in which such a lethal allele is associated with Rb mutation, the number of affected individuals would be smaller than expected. Dryja et al (1984a) suggested a similar explanation for the observation made by Matsunaga (1980b) that the incidence of bilateral Rb was lower in patients with a constitutional, microscopically detectable, 13q14 deletion than in non-deletion patients. There are also reports of deletion patients who have never developed tumours (Wilson et al 1987, Cowell et al 1988). The age of onset of tumours from patients with microscopically detectable chromosome 13 abnormalities also appears to be later than those with other germinal mutations (Ejima et al 1988). A retinal cell that becomes homozygous for a substantial deletion may be non-viable because it lacks essential genes. Also Bonaiti-Pellie et al (1990) in their analysis of a balanced insertion, found that deletion carrying gametes of bilaterally affected individuals have a normal viability, whereas deletion gametes of unilaterally or unaffected individuals have a reduced viability. In this hypothesis, the probability of tumour formation would remain unchanged
from one generation to the next, in contrast to the previous one where it would vary from one generation to the next. There exist, in the literature, family pedigrees that could suggest either. However, a precise test of this probability changing or remaining constant in successive generations in families has not been carried out.

The answer to the question of variable expression and pattern of inheritance in Rb will probably come through molecular investigations of specific types of mutations. Many mutations in the Rb gene have been identified and there is already some evidence in support of the hypothesis that specific mutations may give rise to particular phenotypes and that epigenetic changes may also be involved (Section 4.3.3).

1.5 TUMOUR SUPPRESSOR GENES

The name "tumour suppressor gene" stems from the fact that, the presence of a single copy of these genes appears to be sufficient to ensure normal tissue development, i.e. prevents tumorigenesis. This being the case, the malignant phenotype can be thought of as recessive at the cellular level and the inactive genes responsible can also be referred to as "recessive oncogenes". A third term "antioncogene" was coined by Knudson (1983b) to emphasize that these genes are the antithesis of the dominantly acting oncogenes. Presumably the normal function of antioncogenes is to ensure that cells enter an appropriate developmental pathway. Loss of this function yields a cell defective in growth control and able to undergo transformation. The RBI gene, being the first of such genes identified and cloned, established many of the precedents for the identification and the study of other tumour suppressor genes. LOH was one particularly important observation which has been used as justification to locate the site(s) of other potential recessive cancer genes.

As far back as 1972, Knudson and Strong (1972) had suggested that Wilms' tumour (WT), an embryonic kidney tumour, like Rb, also results from two mutations. Unlike Rb, familial occurrence of this tumour is extremely rare and evidence for the chromosomal location came from the investigations of patients with the AGR triad (aniridia, genitourinary abnormalities and mental retardation) who have about a 50% chance of developing WT. Such patients
were observed to have a cytogenetically detectable deletion of chromosome region 11p13 (Andersen et al 1978, Riccardi et al 1978, Franke et al 1979). These data prompted a molecular analysis of this region and LOH was soon demonstrated for the short arm of chromosome 11 (Fearon et al 1984, Koufus et al 1984) in WT tissue compared to normal tissue from the same individual. In 15-20% of WTs and those associated with the Beckwith-Wiedemann syndrome (characterised by multiple congenital abnormalities and an increased susceptibility to the development of embryonal tumours), however, LOH was detected at 11p15 without involvement of 11p13, suggesting the existence of a second WT gene (Mannens et al 1988, Reeve et al 1989, Wadey et al 1990). The presence of yet another, a third WT gene, was suggested when linkage analysis of rare WT families failed to show any linkage to chromosome 11 (Grundy et al 1988, Huff et al 1988). An added complication was the fact that the allele retained in sporadic tumours was invariably of paternal origin (Schroeder et al 1987, Williams et al 1989) suggesting that genomic imprinting modified perhaps by DNA methylation may be of importance (Sapienza et al 1987, Wilkins 1988). Support for this view is also provided by the observation that a gene located in chromosome 11p15 and expressed in many WT (Scott et al 1985, Reeve et al 1985), namely insulin-like growth factor II (IGF-II), is imprinted in the mouse (DeChiara et al 1991). All these results meant that the genetics of WT is more complicated than the simple two-hit model in Rb has suggested. More recently, chromosome 16q, which showed LOH in 9 out of 45 informative WTs analysed in one study (Maw et al 1992) has been implicated as third WT locus, although it is not yet known whether the 16q locus represents the familial WT predisposing locus in families which do not show linkage to chromosome 11p. Of the three loci involved in the genesis of WT, a candidate gene, WT1, for the 11p13 locus was cloned (Call et al 1990, Gessler et al 1990) and was shown to have structural rearrangements in some WTs (Cowell et al 1991, Huff et al 1991, Ton et al 1991). The cloned gene was expressed in the developing kidney and gonads particularly when cells made the transition from mesenchyme to epithelium but not in adult human kidney (Pritchard-Jones et al 1990, Haber et al 1990), although it was present in adult kidney of baboon and mouse (Call et al 1990). This gene, therefore, had the characteristics of a tissue-specific transcription factor that is expressed only at certain developmental stages in contrast to the ubiquitously expressed RB1 gene. Such restriction of expression readily explains why mutation at the WT1 locus results in one type of tumour. Four zinc-finger
motifs were identified in the protein encoded by WT1 indicating a sequence specific DNA binding protein (Call et al 1990, Gessler et al 1990). The zinc-finger domains of WT1 share sequence similarity with a mammalian immediate-early protein EGR1 which appear during G0 and G1 phases of the cell cycle and both WT1 and EGR1 was shown to recognise similar DNA binding sites (Rauscher et al 1990). EGR1 is ubiquitously expressed and acts as a strong transcriptional activator (like myc and E2F) of promoters containing a specific sequence that it binds. Binding of WT1 to the same sequence element might act to suppress transcriptional activity by EGR1 (Weinberg 1991). In transient transfection assays, the WT1 protein was shown to function as a repressor of transcription when bound to EGR1 binding sites (Madden et al 1991). The fetal mitogen IGF-II has also been implicated in the genesis of Wilms' tumours. IGF-II is overexpressed in all WTs examined thus far (Scott et al 1985, Reeve et al 1985) and, more recently, it was reported that WT1 bound to multiple sites in the IGF-II promoter region and functions as a potent repressor of IGF-II transcription in vivo (Drummond et al 1992). It was suggested that WT1 negatively regulates blastemal cell proliferation by limiting the production of a fetal growth factor in the developing vertebrate kidney (Drummond et al 1992). In this explanation the apparent parallels to pRB function is striking in that in each case the suppressor protein may directly interact with and antagonise mitogen-induced transcription factors (see discussion Section 4.3.3). The fact that patients with 11p13 deletions show abnormal development of the genital system (Andersen et al 1978) and the patients with Denys-Drash syndrome (abnormal genitals) are also predisposed to WT (Jadresic et al 1991, Baird et al 1992) suggests that the genital system might be more sensitive to WT1 mutations than the kidneys which possibly reflects the fact that more than one gene is involved in Wilms' tumorigenesis (Cowell 1992).

NF1 (neurofibromatosis type 1) is another tumour suppressor gene which has been cloned recently (Cawthon et al 1990, Visckochil et al 1990, Wallace et al 1990). NF1 (also known as von Recklinghausen neurofibromatosis) is an autosomal dominant inherited disorder with a variable clinical phenotype ranging from hyperpigmented patches of skin to multiple tumours of the peripheral nerves and brain indicating the gene's importance in the differentiation pathway of nerve cells. Unlike Rb and WT, for which the first indications of the chromosomal site of the locus came from the cytogenetic identification of chromosomal deletions associated with disease, linkage
studies in affected families provided the chromosomal assignment of NF1 to 17q11.2 (Barker et al 1987, Seizinger et al 1987). A candidate gene was cloned and the presence of mutations in the cloned gene in predisposed individuals provided strong evidence for its authenticity (Cawthon et al 1990, Visckochil et al 1990, Wallace et al 1990). The NF1 protein was found to show significant homology to the ras GTPase-activating proteins (GAP) and the IRA genes (encoding GAP homologues) of yeast (Ballester et al 1990, Xu et al 1990a). The protein product of ras, p21, is a GTP-binding protein and appears to be involved in signal transduction pathways. p21ras can exist in either a physiologically quiescent GDP-binding state or a GTP-binding signal-emitting state. Oncogenic p21ras proteins are in a continuous signal-emitting state as GTP to GDP hydrolysis is dysfunctional (Trahey and McCormick 1987, Weinberg 1991). The hydrolysis of GTP negatively regulates p21ras by converting it into the inactive GDP-bound form. GAP proteins can bind to normal p21ras and stimulate this process. They can not, however, bind to mutant forms of p21ras (Trahey and McCormick 1987). This sequence of events are compatible with two alternative schemes. GAP may be a down-regulator of activated p21ras (Downward et al 1990) or a downstream effector (target) that becomes activated when it encounters activated p21ras (Yatani et al 1990), there is evidence to suggest both possibilities. The homology of NF1 protein to GAP proteins suggest that NF1 also participates in the signalling pathway triggered by ras p21. Although it is not yet clear how NF1 acts in this pathway, it was shown that GAP-related region of the NF1 protein can stimulate GTP hydrolysis on normal p21ras but not on mutated p21ras (Martin et al 1990, Xu et al 1990b). It is possible, therefore, that GAP and NF1 proteins are two proteins in this signal transduction pathway, one of which is a negative regulator of p21ras whose activity is inhibited in response to growth factor signals, while the other is the downstream target (Marshall 1991). Like RB1, NF1 is ubiquitously expressed (Wallace et al 1990). The genetic basis for the differences between the variable clinical phenotypes associated with NF1 is still not clear. It has been suggested that inactivation of one NF1 allele may be sufficient to generate benign neurofibromas as has been proposed for inactivation of one FAP allele in benign colorectal adenomas (see later) and that inactivation of the second allele contributes to malignant neurofibrosarcoma formation (Marshall 1991). Involvement of another gene in neurofibrosarcoma has also been suggested, as consistent loss of genetic material on 17p (p53, see below) has been observed in neurofibrosarcomas but not in neurofibromas (Menon et al 1990).
The p53 protein, the gene for which resides on chromosome 17p13, was first identified as a cellular protein that complexed with SV40 LT antigen and coimmunoprecipitated with LT antigen when extracts of SV40-transformed cells were mixed with anti-SV40 LT antibodies (Lane and Crawford 1979, Linzer and Levine 1979). Because LT is needed to maintain the transformed phenotype, it was suggested that the interaction of p53 and LT is important for transformation. Large quantities of p53 were detected in transformed cells in culture and in human tumours (Crawford et al 1981, 1984) but very low quantities in normal cells. In transformed cells p53 was found to have a much longer half-life than in nontransformed cells (Oren et al 1981, Reich et al 1983). The expression of p53 gene was shown to immortalise cells (Jenkins et al 1984) and in oncogene cooperation assays with ras, murine p53 cDNA clones cooperated with ras in transformation of rat embryo fibroblasts (Eliyahu et al 1984, Parada et al 1984). Thus, p53 came to be classified as an oncogene which fell into the same cooperating class as c-myc. This interpretation was proved to be wrong, however, as it turned out that all of these transforming p53 cDNA clones were mutant forms of p53 (Hinds et al 1989). It was then shown that p53 actually acts as a tumour suppressor in experimental systems. Expression of cDNA or genomic clones of wild-type p53 was shown to suppress the transformation of cells in culture by other oncogenes (Finlay et al 1989, Eliyahu et al 1989) and the growth of transformed cells in culture (Baker et al 1990). Also, mutations of wild-type p53 alleles were identified in several human tumours and tumour cell lines (Baker et al 1989, Nigro et al 1989, Takahashi et al 1989, Mulligan et al 1990, Hollstein et al 1991, Iavarone et al 1992, Mazars et al 1992). A large body of data from several groups now indicates that the p53 gene is the most commonly mutated gene known in human cancer. This may suggest a non-specific role for p53. Indeed in the development of colon cancer (see below), for example, it is probably only important in tumour progression. Germ-line p53 mutations, however, are found in patients with rare autosomal dominant Li-Fraumeni syndrome which is characterised by diverse tumours at multiple sites in the body (Malkin et al 1990, Law et al 1991, Borresen et al 1992). Although p53 has many similarities with RB1, both encode for growth suppressor proteins found in the nucleus, both are targeted by the viral oncoproteins and both can be passed through the germline where they serve as congenital determinants of cancer predisposition, mutant p53, unlike pRB and other tumour suppressor proteins can promote cell growth even in the presence of the wild-type protein. Several classes of p53 mutations exist and
all mutants so far examined have lost the ability to suppress transformation (Finlay et al 1989, Baker et al 1990). Some of the mutations, however, may be dominant negative mutations that inhibit the function of normal p53 by complex formation (Eliyahu et al 1988). Unlike the other suppressor proteins, normal cellular p53 appears to assemble into higher order homooligomeric structures (Milner et al 1991). It was proposed that defective units of mutant molecules may participate in forming a multi-subunit complex together with wild-type monomers and destroy the function of the complex as a whole (Herskowitz 1987). It was shown that a number of p53 mutations increase the stability of the protein and alter its conformation (Gannon et al 1990) and also result in complex formation with a 70 kD heat shock protein (Finlay et al 1988). These heat shock proteins may act as chaperons that bring p53 subunits together and help them oligomerise (Weinberg 1991). Another mechanism has been suggested by Milner and Medcalf (1991) who found that activated mutant p53 can influence the conformation of co-translated wild-type p53 and drives the latter into the mutant phenotypic form. p53 protein, like pRB, is phosphorylated in a cell cycle-dependent manner (Sturzbecher et al 1990). The level of p53 increases in G1 in the cytoplasm and the protein moves to the nucleus at the beginning of S phase (Shaulsky et al 1990) during which it becomes phosphorylated (Bischoff et al 1990). p53 has been shown to possess both cis and trans-activational properties. p53 binds DNA in a sequence specific manner (Kern et al 1991) and more recently it has been demonstrated that normal p53 activates transcription in vitro which is blocked when mutant p53 complexed to normal p53 (Farmer et al 1992). p53 also trans-activates the creatine phosphokinase promoter (Weintraub et al 1991). A cellular protein which forms a complex with both mutant and normal p53 protein has been identified as the mdm-2 oncogene product and this association was shown to inhibit p53-mediated trans-activation (Momand et al 1992). The mdm-2 gene was originally identified as a dominant transforming oncogene present on a 'mouse double minute' chromosome (Fakharzadeh et al 1991). Oliner et al (1992) found mdm-2 amplification in 17 out of 47 human sarcomas they studied. Interestingly none of the five mdm-2 amplified sarcomas searched for p53 mutations exhibited any such mutations suggesting that high levels of mdm-2 may, like the virus oncoproteins, inactivate the tumour suppressor activity of p53 by complexing to it (Lane 1992). From these observations p53 emerged as a transcription factor. A different set of experiments, on the other hand, shed light into the biological function of p53. It has been shown that DNA
damage, e.g. by irradiation, induces the accumulation of normal p53 and that this accumulation mediates arrest of the cell cycle at G1, which is in accord with the growth inhibitory properties of high levels of normal p53 (Kastan et al 1991). This G1 arrest following irradiation is lost, however, if a mutant p53 gene is transfected into cells with wild-type endogenous p53 genes (Kuerbitz et al 1992). These observations suggested a role for p53 as a cell cycle checkpoint determinant. A model was put forward by Lane (1992) whereby p53 acts as a 'molecular policeman' monitoring the integrity of the genome. In this model, p53 is not required for normal cell division but when DNA is damaged the genome-guarding function of p53 is induced and p53 either switches off replication to allow DNA repair before division or triggers apoptosis. Cells in which the p53 pathway is inactivated by mutation of p53, or by host (e.g. mdm-2) or viral oncoproteins, however, replicate damaged DNA as they can not carry out cell cycle arrest. These cells are, therefore genetically less stable and might accumulate mutations at an increased rate leading to rapid selection of malignant clones. This model is consistent with the results of a recent study on mice that are p53 deficient. Mice homozygous for a p53 null allele are developmentally normal, suggesting that p53 is not essential for normal cell division, but are prone to the formation of spontaneous tumours indicating that the absence of normal p53 predisposes to neoplastic disease (Donehower et al 1992).

Studies on colorectal tumours provide an excellent example of multistep carcinogenesis and the way in which genetic alterations accumulate and presumably cause tumour initiation, formation and progression. As a result of these studies, the involvement of multiple tumour suppressor genes in tumorigenesis has also become clear. One of the tumour suppressor genes involved has been identified through the studies of familial adenomatous polyposis (FAP) which is a dominantly inherited disorder that predisposes to colon cancer. The tumour phenotype was assigned to chromosome region 5q15-22 through linkage studies (Bodmer et al 1987). That the same locus may be involved in the non-familial forms of colon cancer has been suggested by frequent LOH in this region in non-familial tumours (Solomon et al 1987, Vogelstein et al 1988). Colon carcinoma progresses through a series of 4 clinically distinct stages: hyperplastic changes in epithelium, adenomatous polyps, non-invasive carcinoma and finally invasive carcinoma. Progression through these stages was shown to involve several other events besides LOH at 5q, one of the predominant ones being the activation of K-ras oncogene.
(Vogelstein et al 1988). In progression to carcinoma, LOH was also observed for sequences on chromosome 17 and 18. The previously cloned p53 gene was found to map to the region of chromosome 17 that is most commonly lost in colon cancers (Baker et al 1989). From the locus on chromosome 18 (q21), on the other hand, a cDNA was subsequently cloned (Fearon et al 1990) and termed DCC (deleted in colon cancer). DCC was shown to have characteristics of a cell surface adhesion molecule and suggested to be part of a signal transduction pathway loss of which confers a growth advantage on evolving tumour cells (Fearon et al 1990, Weinberg 1991). From the FAP locus on chromosome 5 (q21), several genes were identified (Kinzler et al 1991b). One of these genes, named MCC (mutated in colon carcinoma), exhibited mutations in some tumour cells (Kinzler et al 1991a) but it was later shown that MCC lay outside the critical region of FAP (Joslyn et al 1991). A second gene, APC (adenomatous polyposis coli), was shown to be mutated in the germ line of patients predisposed to FAP (Groden et al 1991, Nishisho et al 1991) strongly supporting its role in tumorigenesis. The function of either of these genes is not yet clear. MCC has a very short region of homology to the G protein-coupled m³ muscarinic acetylcholine receptor (Kinzler et al 1991a). APC has some structural morphology with cellular G proteins which are connected with cellular signal transduction through their association with ras and NF1 (Kinzler et al 1991b).

The number of tumour suppressor genes identified is already substantial and continues to grow. In addition to the ones discussed above, the presence of tumour suppressor genes are indicated in neuroblastoma on chromosome 1 (Michalski et al 1992), in multiple endocrine neoplasia type 1 and 2 (MEN1 and 2) on chromosomes 11 (Byström et al 1990) and 10 (Mathew et al 1987, Simpson et al 1987) respectively and in acoustic neurofibroma (NF2) on chromosome 22 (Menon et al 1991). In identification of many of the tumour suppressor genes to date family studies played an important role. Even for tumour suppressor genes that have been discovered through routes other than studying familial cancers, there might exist germ line mutations that predispose to cancer. Many common adult human cancers are now believed to have familial types. The interpretation of family studies in these cancers, however, are complicated by the effects of environmental mutagens on the genotype. In lung cancer, for example, smoking is strong causative factor and complicates the interpretation of genetic changes identified. Deletions and LOH involving chromosome band 3p21 is consistently observed in many
lung tumours. Recently a gene from this region has been identified (La Forgia et al. 1991) and found to be a protein tyrosine phosphatase, a class of molecules that have been implicated in the control of the cell cycle. In hereditary breast cancer, although a candidate gene is not yet available, a locus in chromosome region 17q21 has been identified (Hall et al. 1990). New interactions between the products of oncogenes and tumour suppressor genes are being discovered with an ever increasing speed allowing insights into the control of signalling pathways and the cell cycle. Indeed, the protein interactions identified and the study of signalling pathways, in turn, sometimes lead to the identification of new tumour suppressor genes. GAP, for example, identified through its interaction with ras might be a tumour suppressor gene as is TGF-β which can negatively regulate cell growth and division, although these genes have not turned up as tumour suppressor genes during genetic analysis. Recently, however, α-inhibin (the protein product of which forms part of a heterodimeric growth factor structurally related to TGF-β) has been shown to be a tumour suppressor gene with gonadal specificity in mice (Matzuk et al. 1992) and is the first extracellular ligand (secreted protein) identified to have tumour suppressor activity.

In summary, therefore, although at the start, oncogenes have held centre stage in attempts to understand and explain the origins of cancer, they have provided only part of the picture. The rest of the picture belongs to the tumour suppressor genes which provide a better understanding of inherited predisposition in cancer, cell or tissue specific cancers and regulation of the cell cycle. It has now become clear that multistep carcinogenesis involves cooperation between activated oncogenes and inactivated tumour suppressor genes. Through the analysis of various cancers, especially of colon cancer, it has also become clear that it is the accumulation of genetic changes rather than their order of occurrence that is important. The model, thus created, involving activation of oncogenes and inactivation of tumour suppressor genes, highlights an important difference between adult cancers and hereditary embryonic cancers. In embryonic cancers like Rb, as few as two steps seem to be sufficient for malignancy and an involvement of an oncogene has not been implicated, whereas in adult cancers like colon cancer many more steps involving both oncogenes and tumour suppressor genes are required. This difference probably reflects the fact that embryonic cells are not fully committed to a particular differentiation pathway and still express a large number of genes necessary for rapid cell growth. Fully differentiated
cells, on the other hand, have already turned off many of their growth stimulatory genes and both a "reversal" process to activate those "silent" genes and a "suppressor" process to inhibit or inactivate the growth suppressor genes are required. A better understanding of the "spontaneous regression" phenomenon (see Section 1.3.4) seen in tumours like Rb and neuroblastoma might shed new light on tumour progression processes. One suggestion for the phenomenon of spontaneous regression is that critical mutations have occurred in cells which are far too committed and although a brief period of uncontrolled growth is allowed, they are brought back under control by overriding differentiation signals. In uncommitted cells fewer genes are required to be silenced to place them on a pathway to malignancy. This does not mean, however, that other genetic alterations are not needed to achieve full malignancy. Even in cancers which seem to require only two mutations for tumour formation, other genetic alterations that provide developing tumour cells with growth advantages must be occurring. In neurofibrosarcomas, for example, alterations in p53 as well as NF1 are involved (Menon et al 1990). This example again stresses the importance of overcoming suppressor processes and differentiation signals. Few tumours have been observed with two dominantly activated oncogenes but there are numerous examples of two or more tumour suppressor genes within one tumour. The examples include inactivations of both RB1 and p53 in lung carcinomas, osteosarcomas and p53 and DCC in colon carcinomas. In tumour cell line Saos-2 both RB1 and p53 genes are inactivated. It has been shown that introduction of either normal p53 (Diller et al 1990) or normal RB1 (Huang et al 1988) into these cell lines suppresses cell proliferation suggesting that neoplastic phenotypes of the cells require the inactivation of both genes.

The study of hereditary cancers and especially of Rb has been one of the most fruitful lines of research in cancer. These studies have not only led to the identification of tumour suppressor genes and a better understanding of origins of cancer but, in the case of Rb, also provided immediate clinical benefits in the form of genetic screening (Section 1.4.6.2).

In this thesis, the application of genetic analysis to hereditary Rb has been investigated and, as a result of identification of causative mutations in some of these families insights into the structure/function relationship of RB1 have been analysed.
2.0 METHODS
2.1 PREPARATION OF SOLUTIONS AND BUFFERS

All chemicals used in this project were analytical grade and were purchased from a variety of companies including Sigma, BDH Lab Supplies, Bethesda Research Laboratories (BRL), Fluka, Pharmacia and Boehringer Mannheim. 'Cold' dNTPs and pD(N)_6 (random hexa-deoxynucleosides) were purchased from Pharmacia. Radio-labelled NTPs and dNTPs were from Amersham International. The solutions and buffers were made in double distilled or distilled water obtained from a Milli Q 4 BOWL and Milli RO4 system (Millipore Ltd) respectively. Where required pH adjustments (pH meter 240, Ciba Corning Diag Ltd.) were made by the addition of an appropriate acid or alkali. Formamide was deionised before use using an ion exchange resin (Bio-Rad AG 501-X8) and stored at -20°C. All solutions were either autoclaved or filter sterilised and stored according to manufacturer's instructions. The pipette tips, eppendorf and other tubes used were also autoclaved prior to use. In experiments involving RNA, all glassware was also autoclaved and gel electrophoresis apparatus was washed with water which was previously treated with a 0.2% solution of DEPC (diethyl pyrocarbonate), left to stand overnight and autoclaved (Maniatis et al 1982).

2.2 ENZYMES

Restriction endonucleases (Appendix I) were obtained from a variety of manufacturers including Northampton Biologicals Ltd (NBL), Anglian Biotec, Promega, BioExcellence Ltd, Pharmacia and BRL, together with the appropriate 10x concentrated buffers. Enzymes and buffers were stored at -20°C. The following enzymes were obtained from sources indicated in brackets: Klenow fragment of DNA polymerase I (Anglian Biotec); Taq DNA polymerase (Promega or NBL); Sequenase (United States Biochemical Corporation(USB)); RNase A (Boehringer Mannheim or BRL); RNase T1 (BRL); T7 and T3 RNA polymerases, RNasin ribonuclease inhibitor (Promega); RNase free DNase I (Boehringer Mannheim); proteinase K (BDH). RNase A (when obtained as powder) was prepared as a 10 mg ml\(^{-1}\) stock solution and was heated at 90°C for 15 minutes prior to use to destroy any residual DNase activity.
2.3 PLASMIDS

DNA probes and cDNA fragments used to generate RNA probes were supplied in plasmid vectors Bluescribe +, Bluescript + and pTZ18R, all of which conferred ampicillin resistance on their bacterial hosts. Stocks of plasmid were kept as DNA stored at -20°C, and as frozen cell stocks of transformed bacteria at -70°C.

2.4 PREPARATION AND HANDLING OF BACTERIA

*Escherichia coli* strain JM83 (λ-, ara, Δ(lac-proAB), rpsL, thi-1, φ80lacZΔM15) (Pharmacia) was used in all manipulations involving bacteria. Bacteria were grown at 37°C in the appropriate culture media either in a rotary shaker (controlled environmental incubator shaker, New Brunswick Scientific, Edison, NJ, USA) or dry incubater (LEEC CV2, Nottingham, England). Media and stock solutions were made with double distilled water and autoclaved at 121°C for 15 minutes to ensure sterility. The following media were used:

1. Luria Broth (LB) (gl⁻¹): Difco Bactotryptone, 15; Difco Bacto Yeast Extract, 5; NaCl, 5.
2. Luria Agar (L-agar) (gl⁻¹): Difco Bactotryptone, 15; Difco Bacto Yeast Extract, 5; NaCl, 5; Difco Bacto Agar, 14.

L-agar was melted in a microwave oven as required. After the LB was cooled to 65°C antibiotic ampicillin was added either from a freshly prepared solution or from stocks frozen at -20°C (for up to one month) at a concentration of 50 µg ml⁻¹. For short term use (up to 4 weeks) bacteria were maintained on L-agar plates (4°C) containing ampicillin; for long term storage, cell stocks were frozen at -70°C in LB containing 15% (v/v) glycerol.

2.4.1 Transformation of bacteria

In order to transform bacteria with supercoiled plasmid, they must be made "competent". This process involves soaking bacterial cells in an ice-cold salt solution; calcium or rubidium chloride. Exactly why this treatment works is
not understood. It is possible that salt solution causes the DNA to precipitate onto the outside of the cells, or perhaps the salt is responsible for some kind of change in the cell wall that improves DNA binding. After DNA is added to treated cells, the movement of DNA into competent cells is stimulated by briefly raising the temperature to 42°C. Once again, the exact reason why this heat-shock is effective is not understood. Competent *E. coli* JM83 cells were made by the method of Kushner (1978). 10 ml of LB was inoculated with 0.1 ml of a 16 hour LB culture containing *E. coli* JM83 and was incubated, shaking at 37°C, until the OD$_{600}$ was 0.15 (approximately 2 hours). 1.5 ml aliquots were spun at 13000 rpm for 2 minutes in eppendorf tubes and the cells recovered were resuspended in an ice-cold buffer containing 10mM MOPS, 10mM rubidium chloride. After 5 minutes on ice, the cells were reharvested and resuspended in 100mM MOPS, 50mM magnesium chloride, 100mM rubidium chloride. After a further 60 minutes on ice, the cells were harvested and resuspended in 0.2 ml of the MgCl$_2$ containing buffer. The competent cells were stored in 15% (v/v) glycerol at -70°C, although the transformation efficiency declines rapidly after 6-8 months. For transformation, 50 ng of plasmid DNA was added to ice-cold competent cells, mixed gently and left on ice for 45 minutes. Bacteria were then heat shocked at 42°C for 2 minutes whereafter 1 ml of pre-warmed LB was added and the cells left at 37°C in a water bath for 30 minutes to allow expression of antibiotic resistence. Aliquots of the transformed cells (10 µl and 100 µl) were then spread on L-agar containing 50 µg ml$^{-1}$ ampicillin and left at 37°C overnight.

2.5 ISOLATION OF PLASMID DNA

Plasmid DNA was isolated from either 1.5 ml cultures (small scale) or 400 ml cultures (large scale) of transformed bacterial cells.

2.5.1 Small scale isolation

Single colonies from an L-agar plate were innoculated into 5 ml of L-Broth containing 50 µg ml$^{-1}$ ampicillin and incubated in a rotary shaker incubator at 37°C overnight. 1.5 ml of these cultures were transferred to an eppendorf
tube and spun at 13000 rpm in a bench-top microfuge (MSE microcentaur) for 4 minutes. The cell pellet was resuspended in 200 μl Birnboim buffer (table 2.1) and a few grains of lysozyme added. After 10 minutes at room temperature, 400 μl of 0.2M NaOH, 0.1% SDS were added to lyse the cells. After 5 minutes on ice, 200 μl of 3M potassium acetate (pH 5) were added, yielding a thick white precipitate consisting of protein and high molecular weight non-supercoiled bacterial DNA. The solution was gently mixed by inverting the tube and left on ice for 15 minutes. The sample was spun for 15 minutes at 13000 rpm in a microfuge and the supernatant containing the supercoiled DNA was transferred to another eppendorf tube. Supercoiled plasmid DNA was precipitated by addition of 0.6 volumes (480 μl) of propan-2-ol and left at room temperature for 20 minutes. The plasmid DNA was recovered by centrifugation at 13000 rpm for 15 minutes. The pellet was washed in 70% ethanol, dried and resuspended in 500 μl TE (table 2.2). The DNA was treated with 5 μl (10 units) of DNase free RNase A at 37°C for 30 minutes followed by 50 μg of proteinase K at 37°C for a further 30 minutes. The solution was then extracted one time each with equal volumes of phenol, phenol/chloroform (50:50), and chloroform to remove proteins. At each stage, the sample was mixed thoroughly and spun for 2 minutes to separate the solvent and aqueous layers. The upper, aqueous layer containing DNA was transferred to another eppendorf tube leaving behind the coagulated proteins on the interface and in the bottom phenol or phenol/chloroform layer. The last extraction with chloroform removes last traces of phenol. The plasmid DNA was reprecipitated by addition of 0.1 volume 3M sodium acetate, pH 5.2, and 2 volumes 100% ethanol and left at -70°C for 2 hours. The precipitate was recovered by spinning at 13000 rpm for 15 minutes, then washed in 70% ethanol, and dried and resuspended in 50 μl TE. An aliquot was analysed on a 1% agarose gel (Section 2.9). About 3-5 μg DNA was usually recovered using this method.
**Table 2.1** Solutions used in the isolation of DNA and RNA.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Birnboim Buffer</strong></td>
<td>50mM Glucose</td>
</tr>
<tr>
<td></td>
<td>25mM Tris-HCl, pH 8</td>
</tr>
<tr>
<td></td>
<td>10mM EDTA, pH 8</td>
</tr>
<tr>
<td><strong>Solution A</strong></td>
<td>75mM NaCl</td>
</tr>
<tr>
<td></td>
<td>24mM EDTA, pH 8</td>
</tr>
<tr>
<td><strong>Cell Lysis Buffer</strong></td>
<td>50mM NaCl</td>
</tr>
<tr>
<td></td>
<td>50mM Tris-HCl, pH 8</td>
</tr>
<tr>
<td></td>
<td>10mM EDTA, pH 8</td>
</tr>
<tr>
<td><strong>DNA Extraction Solution</strong></td>
<td>100mM Tris-HCl, pH 8</td>
</tr>
<tr>
<td></td>
<td>4mM EDTA, pH 8</td>
</tr>
<tr>
<td><strong>Solution D</strong></td>
<td>4M Guanidinium thiocyanate</td>
</tr>
<tr>
<td></td>
<td>25mM Sodium citrate</td>
</tr>
<tr>
<td></td>
<td>0.5% Sarcosyl</td>
</tr>
<tr>
<td></td>
<td>0.1M 2-mercaptoethanol</td>
</tr>
</tbody>
</table>
2.5.2 Large scale isolation

400 ml of L-Broth containing 50 \( \mu \text{g ml}^{-1} \) ampicillin was inoculated with 500 \( \mu \text{l} \) of a 16 hour L-Broth culture and grown overnight with vigorous shaking. It was then divided into two 200 ml aliquots in two autoclaved plastic bottles and spun at 4°C, 4000 rpm (Beckman J2-21, rotor SER1257) for 10 minutes. The pellets were resuspended in 10 ml of Birnboim buffer (table 2.1) containing 2 mg ml\(^{-1}\) lysozyme and left at room temperature for 30 minutes. To each bottle two volumes (20 ml) of 0.2M NaOH containing 1% SDS (w/v) was added, mixed well and the bottles were kept on ice for five minutes. 3M potassium acetate, pH 5 (10 ml), was added and mixed resulting in a heavy white precipitate. The bottles were kept on ice for 15 minutes and then spun for 10 minutes at 4000 rpm at 4°C to sediment the white precipitate. The supernatant from each bottle was filtered through a muslin cloth into four Sarstedt centrifuge tubes (Sarstedt Ltd.). 0.6 volume isopropanol was added and kept at room temperature for 15 minutes to precipitate DNA. The DNA pellet was recovered by centrifugation (Beckman J2-21, rotor JA-20 SER E 14991) at 4°C, 4000 rpm for 20 minutes and washed in 70% ethanol. The pellets from four tubes were resuspended in TE (table 2.2) to give a total volume of 10 ml and transferred to a Falcon tube (Becton Dickinson UK Ltd). The DNA was treated with 100 \( \mu \text{g ml}^{-1} \) RNase A for 2 hours at 37°C. It was then treated with 100 \( \mu \text{g ml}^{-1} \) proteinase K either at 55°C for 2 hours or 37°C overnight. The DNA was extracted using the standard phenol/chloroform technique (Section 2.5.1). 0.1 volume of 3M sodium acetate, pH 5.2, and 2 volumes 100% ethanol was then added, the solutions mixed, and placed at -70°C for at least 1 hour or overnight. The DNA pellet was recovered by centrifugation in a desk centrifuge (Centra-4R, International Equipment Company (IEC), Dunstable, UK) at 4°C, 4000 rpm for 15 minutes, after which it was washed in 70% ethanol, dried and resuspended in 500 \( \mu \text{l} \) of TE. The concentration was estimated by measurement of the OD\(_{260}\) (Optical Density at 260 nm; absorbance of 1 = 50 \( \mu \text{g ml}^{-1} \) DNA). The DNA yield was usually between 1-2 mg. OD\(_{280}\) readings were also taken. Measurement of the OD\(_{260} : \) OD\(_{280}\) ratio provides an indication as to whether the DNA is contaminated by protein or phenol; a value of 1.7-1.9 indicates an acceptable degree of purity (Maniatis et al 1982).
Table 2.2  General buffers.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TE</strong></td>
<td>10mM Tris-HCl, pH 8</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA, pH 8</td>
</tr>
<tr>
<td><strong>TAE</strong> (Tris-acetate)</td>
<td>0.04M Tris-acetate</td>
</tr>
<tr>
<td></td>
<td>0.02M EDTA, pH 8</td>
</tr>
<tr>
<td><strong>TBE</strong> (Tris-borate)</td>
<td>0.089M Tris base</td>
</tr>
<tr>
<td></td>
<td>0.089M Boric acid</td>
</tr>
<tr>
<td></td>
<td>0.002M EDTA</td>
</tr>
<tr>
<td><strong>TBE</strong> (long run)</td>
<td>0.134M Tris base</td>
</tr>
<tr>
<td></td>
<td>0.045M Boric acid</td>
</tr>
<tr>
<td></td>
<td>0.002M EDTA</td>
</tr>
<tr>
<td><strong>Loading buffer</strong></td>
<td>0.24% Bromophenol blue</td>
</tr>
<tr>
<td>(for agarose gels)</td>
<td>15% Ficoll</td>
</tr>
<tr>
<td></td>
<td>100mM EDTA, pH 8</td>
</tr>
<tr>
<td><strong>Stop solution</strong></td>
<td>95% formamide</td>
</tr>
<tr>
<td>(for polyacrylamide gels)</td>
<td>20mM EDTA</td>
</tr>
<tr>
<td></td>
<td>0.05% Bromophenol blue</td>
</tr>
<tr>
<td></td>
<td>0.05% Xylene cyanol</td>
</tr>
</tbody>
</table>
2.6 ISOLATION OF DNA FROM HUMAN CELLS

DNA was prepared from a variety of sources including whole blood, white blood cells (lymphocytes), lymphoblastoid cell lines, chorionic villi and formalin-fixed, paraffin-embedded tissue samples.

2.6.1 Isolation of DNA from whole blood

Blood samples from family members were collected in either EDTA or lithium heparin anticoagulant tubes. DNA was extracted either from fresh blood or from blood stored at -70°C. The extraction was carried out in 50 ml Falcon tubes. The blood was made up to 50 ml with ice cold water to lyse residual red cells and spun at 3600 rpm, 4°C, for 20 minutes in a desk centrifuge (IEC CENTRA 4R). To the pellet containing white cells, 40 ml ice-cold 0.1% Nonidet P40 (detergent) was added and the pellet resuspended. After another centrifugation step, the resulting pellet was resuspended in 2.5 ml of Solution A (table 2.1), 100 μl (10 mg ml⁻¹) proteinase K and 0.125 ml of 10% SDS were added. The sample was incubated for 2 hours at 55°C or overnight at 37°C. After addition of a further 2.5 ml of Solution A (table 2.1), the sample was extracted using standard phenol/chloroform techniques (Section 2.5.1). The sample was then treated with RNase A (50-100 μg ml⁻¹) at 37°C for 2 hours and re-extracted. The DNA, which was precipitated by addition of 0.1 volume 3M Sodium acetate, pH 5.2, and 2 volumes 100% ethanol, was then spooled out of the solution using a sealed pasteur pipette. After a 70% ethanol wash, the DNA was dissolved in 0.5 ml of TE. An aliquot was checked on an agarose gel (Section 2.9) to estimate the concentration.

2.6.2 Isolation of DNA from white blood cells and lymphoblastoid cell lines

White cells or lymphoblastoid cells were washed in PBSA (Phosphate Buffered Saline) and resuspended in 9.5 ml of cell lysis buffer (table 2.1). 0.5 ml of 10% SDS was added and mixed by inverting the tube. It is important to add the SDS after resuspending the cells to avoid clumping. Proteinase K was added to a final concentration of 100 μg ml⁻¹ and the
sample was incubated for 2 hours at 55°C or overnight at 37°C whereafter the DNA was extracted using standard phenol/chloroform extraction techniques (Section 2.5.1). The top aqueous layer, containing the DNA, was removed and treated with RNase A (50-100 µg ml⁻¹) at 37°C for 2 hours and re-extracted. The DNA was precipitated with 0.1 volume of 3M sodium acetate, pH 5.2, and 2 volumes of 100% ethanol and spooled out using a sealed pasteur pipette. After a 70% ethanol wash DNA was resuspended in an appropriate volume of TE and an aliquot was checked on an agarose gel (Section 2.9).

2.6.3 Isolation of DNA from chorionic villi

Chorionic villus samples were provided in RPMI media which were removed by centrifugation. The tissue was washed in PBSA and divided into 2 eppendorf tubes. Each sample was then resuspended in 475 µl of cell lysis buffer (table 2.1). After addition of 25 µl of 10% SDS and 10 µl of 10 mg ml⁻¹ proteinase K, the samples were left at 55°C for 2 hours or at 37°C overnight. Phenol/chloroform extractions and RNase A (50-100 µg ml⁻¹) treatment at 37°C followed. Finally, DNA was precipitated by adding 0.1 volume of 3M sodium acetate, pH 5.2, and 2 volumes of 100% ethanol and washed in 70% ethanol. DNA was resuspended in an appropriate volume of TE, and an aliquot was checked on an agarose gel (Section 2.9).

2.6.4 Extraction of DNA from paraffin-embedded tissue sections

DNA from 10 µm thick formalin-fixed, paraffin-embedded tumour sections was prepared by a method modified from Shibata et al (1988). Individual paraffin sections were dissolved in 0.5 ml xylene in an eppendorf tube. The tissue was recovered by centrifugation in a microfuge for 4 minutes and then washed twice in ethanol and desiccated for 0.5 to 2 hours. 50 µl of DNA extraction solution (table 2.1) was added to each eppendorf. After addition of 25µg proteinase K, the solution was mixed well and left at 37°C overnight. The samples were vortexed for 30 seconds after overnight incubation and boiled for 7 minutes. 5 µl aliquots were checked on a 0.6% agarose gel with ethidium bromide staining (Section 2.9). 1 to 10 µl were used for PCR (Section 2.17).
2.7 ISOLATION OF TOTAL RNA FROM LYMPHOBLASTOID CELL LINES

A single step method involving acid guanidinium thiocyanate and phenol-chloroform extraction described by Chomczynski and Sacchi (1987) was used. This method was preferred to that involving ultracentrifugation of a guanidinium thiocyanate lysate through a CsCl cushion (Chirgwin et al 1979), because it resulted in an equally pure preparation of undegraded RNA in reasonably high yield but in a much shorter time. The method was simple and allowed simultaneous processing of a large number of samples. Lymphoblastoid cells grown in suspension were counted and 5-50 x 10^6 cells were used to isolate total RNA. The cell suspension was spun at 1000 rpm for 3 minutes and the pellet was resuspended in an appropriate volume (100 μl per 10^6 cells) of denaturing Solution D (table 2.1) containing guanidinium thiocyanate which is a strong protein denaturant and ribonuclease inhibitor. The suspension was transferred to an autoclaved Sarstedt tube. Sequentially, 0.1 volume 2M sodium acetate, pH 4; 1 volume water saturated phenol and 0.2 volume chloroform/isoamyl alcohol mixture (49:1) were added to the denatured cells which were mixed thoroughly by inverting the tube after each addition. The final suspension was shaken vigorously for 10 seconds and placed on ice for 15 minutes. Centrifugation (JA-20 Beckmann) at 10000 rpm for 20 minutes at 4°C achieved separation of two layers. The RNA remains in the top aqueous layer, the DNA and proteins being left, either in the lower phenol layer, or on the interface. The top layer, containing RNA, was carefully transferred to a fresh tube taking care not to disturb the interface and was mixed with 1 volume isopropanol and placed at -20°C for at least 1 hour to precipitate the RNA. The precipitate was recovered, at 10000 rpm for 20 minutes at 4°C. The resulting pellet was dissolved in 0.3 ml of Solution D, transferred to an eppendorf tube and reprecipitated in 1 volume isopropanol at -20°C for 1 hour. After centrifugation the pellet was washed in 70% ethanol and recentrifuged. The RNA pellet was vacuum dried (Genevac CE100, BRL) for 5 minutes and then dissolved in 50 μl DEPC-treated water (Section 2.1) at 65°C for 10 minutes. An aliquot (1-10 μl) was used for spectrophotometric quantitation. An absorbance of 1 OD unit at 260 nm represents 40 μg ml⁻¹ RNA. OD 260/280 ratio readings of between 1.8 and 2.0 indicated acceptable purity. Typical yields were between 3-5 μg RNA per 10^6 cells. An aliquot was run, after denaturing at 65°C for 15 minutes to prevent secondary structure formation, on a 1%
agarose gel made up in TAE (table 2.2) containing 0.2% SDS (Section 2.9) until the dye front approached the bottom of the gel. The gel was stained in water containing 5 µg ml\(^{-1}\) ethidium bromide for 15 minutes. After the excess ethidium bromide was removed by washing in water for 2 x 30 minutes, the gel was checked for the presence of 28S and 18S ribosomal RNA bands which are indicative of the integrity of RNA, using a UV light source (see Results section 3.2.1). After addition of 0.1 volume 3M sodium acetate, pH 5.2, and 2.5 volumes 100% ethanol, the RNA obtained were divided into aliquots and stored at -70°C until required.

2.8 RESTRICTION ENZYME DIGESTION

Restriction enzymes used in this project are listed in Appendix I. Reactions were carried out in 0.5 ml or 1.5 ml eppendorf tubes. Reaction mixtures contained the appropriate amount of DNA, 10x concentrated restriction enzyme buffer and an excess of restriction enzyme. For genomic DNA digestion 5-10 units of enzyme per microgram of DNA was used. The corresponding figure for plasmid DNA and PCR amplified DNA was 3-4 units per microgram. Where there was difficulty with incomplete digestion of genomic DNA, 2-4mM spermidine was also included in the reaction mixture. When release of cloned DNA fragments from plasmids required double digests, the two enzymes were used simultaneously if the buffer requirements were the same. If not, the low salt digestion was performed first and then the buffer adjusted appropriately for the second enzyme. Digestion was usually for a minimum of 4 hours but in some cases was left overnight. A small aliquot was run on a 1% agarose gel (Section 2.9) to check for complete digestion.

2.9 AGAROSE GEL ELECTROPHORESIS

Agarose (ultrapure electrophoresis grade, Gibco-BRL) gel powder was melted in TAE buffer (table 2.2) at the required concentration in a microwave oven. For DNA electrophoresis, 1 µg ml\(^{-1}\) ethidium bromide was added after cooling to 55°C and 4-6 mm gels were cast on a 20 x 20 cm perspex former.
The gel was placed in a horizontal gel tank and submerged in 2 litres of TAE buffer containing 1 μg ml⁻¹ ethidium bromide. Loading buffer (table 2.2) which contains 15% Ficoll was added to each sample to increase the density thus allowing it to sink into the well. Electrophoresis was carried out under varying conditions ranging from 0.5 V cm⁻¹ to 5 V cm⁻¹ (Kingshill 15AOIC stabilised power supply). The length of the electrophoresis run depended on the required resolution of the DNA fragments. When fragments differed from each other by only 50-100 bp, the electrophoresis run was overnight applying low voltage. For RNA electrophoresis, the running buffer contained 0.2% SDS but no ethidium bromide. The RNA samples were denatured at 65°C for 15 minutes before loading. The gels were stained after the run, in water containing 5 μg ml⁻¹ ethidium bromide for 15 minutes. They were then washed (2 x 30 minutes) in water to remove excess ethidium bromide before photography.

2.9.1 Photography

Ethidium bromide is an intercalating dye which fluoresces in ultraviolet (UV) light. The ethidium bromide stained gels were illuminated from below by a Chromotome TM-20 transilluminator (Ultraviolet Products Inc., California USA). Photographs of the gels were taken by either a Polaroid Cu5 hand camera (5in lens / CU5 88-48 hood providing 0.42x magnification onto Polaroid 667 film ISO 3000) or by a camera with 16mm Cosmicar television lens connected to a video copy processor (Model K61S, Mitsubishi) producing photographs on thermal paper. A ruler was photographed alongside with the gel to aid in the calculation of fragment sizes.

2.10 POLYACRYLAMIDE GEL ELECTROPHORESIS

Polyacrylamide gels were cast between two vertically mounted glass plates, one of which was shorter (33.3 x 39.4 cm) than the other (33.3 x 42 cm). Before assembly the plates were washed with detergent and warm water and rinsed well. One side of the shorter plate was siliconised by wiping the surface with dichlorodimethylsilane (BDH) in a fume cupboard. When dry the siliconised plate was rinsed in water. The larger plate was wiped with
ethanol and 0.3 mm thick side spacers were placed on either side. When assembling for field-gradient ('wedge') gels, two additional short spacers (2 cm) were placed on top of the long spacers at the bottom of the assembly. The shorter plate was then placed on top with siliconised side forming the inner surface. The edges of the glass plates were sealed with tape.

Polyacrylamide was polymerised from acrylamide during gel casting in a radical reaction started with ammonium persulphate and TEMED (N,N,N',N'-tetramethylethylenediamine) (BRL). 4-6% gels were made in TBE buffer (table 2.2) from a 40% (w/v) acrylamide stock (acrylamide:bisacrylamide 19:1) (NBL). Denaturing gels were made by addition of 7M urea (Sigma) to the gel mix. For polymerisation freshly made 10% ammonium persulphate (final concentration 0.08%) and 100 µl TEMED (for 100 ml gel volume) were added and mixed well. The gel mixture was then poured in between the assembled glass plates which were held at a 45° angle. The air bubbles were excluded by holding the plates vertically and gently tapping on the glass plates. After pouring, a comb was inserted between the plates to a depth of 2-3 mm below the short plate. The plates were then clamped at the sides and on the top using "bulldog" clamps and left horizontally on the bench for polymerisation. When the gel was set (usually after 30 minutes), the clamps and the tape were removed and the outer surface of the plates were wiped clean. The gel sandwich was then secured to a vertical electrophoresis apparatus (BRL) with the shorter plate facing inwards. The upper and lower reservoirs were filled with 1x TBE (500 ml to each). The comb was removed from the gel, and the wells were rinsed with buffer using a syringe and needle to remove any unpolymerised acrylamide. The apparatus was connected to a suitable power supply, either Feathervolt Power Pack 2000 (Stratagene) or LKB Bromma 3000 Xi (Bio-Rad). Denaturing gels were pre-run for 30 minutes to heat the gel prior to loading of the samples. Samples were loaded after addition of stop solution (USB) (table 2.2) using a Gilson pipettor and Mµliti-Flex flat tips (Anachem-England). The length of the actual run and the power used depended on the particular experiment. When electrophoresis was complete, denaturing gels were fixed and urea was removed in a solution containing 10% methanol and 10% acetic acid for 45 minutes in a fume cupboard. The gel was then transferred onto 3MM Whatman paper and dried for 2 hours in a gel drier (Model 583, Bio-Rad) connected to a vacuum pump (Genevac Pump CVP50). When the DNA or RNA samples were labelled with 32P, gels were exposed
to XAR-5 autoradiography film (Kodak) inside light proof cassettes (Genetic Research Instrumentation Ltd) for 12-72 hours usually with intensifying screens (Cronex Quanta 3, Dupont) at -70°C. Autoradiography was carried out at room temperature, without an intensifying screen, when the samples were labelled with $^{35}$S. The films were developed using a Fuji RG II X-Ray Film Processor.

2.11 MOLECULAR WEIGHT MARKERS

For agarose gel electrophoresis, a 1 Kb DNA ladder (size range 12 kb to 75 bp, Gibco-BRL) and λHHR (size range 23 kb to 140 bp, Anglian Biotec) were used as molecular weight markers (Appendix II). λHHR is a mixture of a bacteriophage Lambda DNA Hind III digest and Lambda DNA Hind III/EcoRI double digest (Appendix II). Usually 1 to 2 μg per lane were run with each gel and the bands were visualised by ethidium bromide staining.

For polyacrylamide gel electrophoresis, the 1 Kb DNA ladder mentioned above and/or Hinf I cut pAT 153 DNA (Anglian Biotec) or Hinf I cut pBR 322 DNA (Pharmacia) (size range: 1630 to 75 bp) (Appendix II) were used as molecular weight markers. These markers were visualised on polyacrylamide gels by radio-labelling with $^{35}$S dATP. The molecular weight marker bands have sticky ends and "end-labelling" resulted from filling them in with Klenow (large fragment of DNA polymerase I) and $^{35}$S dATP as follows; 1 μg of each DNA marker was incubated with 2 μl of 10x Nick translation buffer (table 2.3), 25mM each of dGTP, dTTP, dCTP, 20 μCi (50 picomoles) of $^{35}$S dATP and 1 μl (5 units) Klenow at room temperature for 30 minutes. The reaction was stopped by addition of 1 μl of 0.5M EDTA, pH 8 and the DNA was precipitated by addition of 5M Ammonium acetate to a final concentration of 2.5M and 3 volumes of 100% ethanol at -70°C for 20 minutes. Unincorporated nucleotides remained in solution. The precipitate was washed in 70% ethanol, dried and resuspended in 10 μl TE. 1 μl was used to measure cpm in a liquid scintillation counter (LKB Wallac 1218 Rackbeta). The specific activity was typically 2-4 x 10$^5$ cpm per μg DNA. Generally 2-4 x 10$^4$ cpm per lane were run on gels.
2.12 SIZE ESTIMATION OF DNA

The migration rates of the DNA molecules through the gel matrix are inversely proportional to the logarithms of the molecular weights (Aaij and Borst 1972). For size estimation, the logarithm of the size in kilobase pairs (kb) may be plotted against mobility (distance migrated) or the size may be plotted against the reciprocal of the mobility (Southern 1979). Although relating the size to the reciprocal of the mobility has been shown to give a straight line over a wider range than the semi-logarithmic plot (Southern 1979), both methods give a good fit to a line in the size range 0-9 kb for 0.8-1.2% agarose gels. The reciprocal mobilities of fragments of unknown sizes were used to calculate their size, by reference to the standard regression line for the sizes of marker fragments.

2.13 EXTRACTION OF DNA FROM AGAROSE GELS

DNA fragments to be used as probes in hybridisation experiments were size fractionated in 1% agarose gels after release from plasmids following digestion with the appropriate restriction enzymes. The gel piece containing the required band was cut out, sliced into small pieces and placed in an eppendorf tube. DNA was extracted and purified using the GeneClean kit (BIO 101, California, USA) according to the manufacturer's recommendations. A volume of sodium iodide equivalent to 3x the volume of agarose, was added and the suspension was incubated at 55°C for 2-3 minutes to dissolve agarose. "Glassmilk", a suspension of glass beads in TE, was added to the mixture which was then kept on ice for 5 minutes. The amount of "glassmilk" added depended on the amount of DNA present; 5 μl was added for 5 μg DNA or less. Glassmilk binds to the DNA which can be recovered by spinning for 10 seconds in a microfuge. The pellet was washed three times in "New Wash" (Tris/ethanol mixture), to remove traces of agarose and salt. After the last wash, the pellet was suspended in 10 μl TE and incubated at 55°C for 3 minutes to elute DNA from the glass beads. The beads were pelleted by spinning for 30 seconds and the supernatant was removed to a new eppendorf. This procedure was repeated once again with a further 10 μl TE which was then added to that from the first elution. The concentration of the DNA was estimated by running a small aliquot on a gel. Yields were generally in the range of 60-80%.
2.14 OLIGO-PRIMER EXTENSION LABELLING OF DNA

The oligo-primer extension method described by Feinberg and Vogelstein (1983) was used in labelling of all DNA probes. The method involves using random hexanucleotides as primers for the initiation of synthesis of a second DNA strand by Klenow in the presence of a radiolabelled nucleotide which is incorporated in the newly-synthesised DNA. 50-100 ng of probe DNA was denatured by heating in a boiling water bath for 3 minutes, quenched and kept on ice to prevent renaturation. 3 μl of 10x concentrated Oligolabelling buffer (table 2.3) were added to the DNA together with 1.2 μl of 10 mg ml⁻¹ BSA (Bovine Serum Albumin), 10 picomoles (equivalent to 30 μCi) of α³²P dCTP, 1 μl (5 units) of Klenow and then TE to a final volume of 30 μl. The mixture was incubated at room temperature for 4 hours. Unincorporated nucleotides were separated by elution through a Sephadex G-50 (medium) column prepared in a short pasteur pipette plugged with glass wool. The reaction mixture was added to the column, followed by TE in 150 μl aliquots. Each aliquot was collected in a separate eppendorf tube and monitored for radioactivity using a hand-held monitor. Labelled DNA was eluted in the first peak of radioactivity (usually in fractions 5-9), ahead of a second peak which contained the unincorporated nucleotides. The DNA containing fractions were pooled and the specific activity was calculated from a 5 μl sample using a Bioscan QC 2000 (V. A. Howe Co. Ltd., London, UK) counter. Specific activity of the probes was typically 2 x 10⁸ cpm per μg DNA but depended on the age of the radionucleotides used.
Table 2.3 Solutions used in DNA labelling.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition and Details</th>
</tr>
</thead>
</table>
| **10x Nick translation buffer** | 0.5M Tris-HCl, pH 7.5  
0.1M MgSO₄  
1mM DTT |
| **10x Oligolabelling Buffer** | A mixture of Solutions I, II, III in the ratio of 2:5:3 |
| **Solution I** | 1.25M Tris-HCl, pH 8  
125mM MgCl₂  
255mM 2-mercaptoethanol  
0.5mM each of dGTP, dATP, dTTP |
| **Solution II** | 2M HEPES pH 6 (with NaOH) |
| **Solution III** | 90 OD₂₅₀ units ml⁻¹ deoxynucleosides  
in 3mM Tris-HCl, pH 8; 0.2mM EDTA |
2.15 SOUTHERN BLOTTING

Genomic DNA fragments were transferred to nylon membranes by Southern blotting essentially as described by Southern (1975). Between 2-10 μg genomic DNA was digested with the appropriate restriction endonuclease and size fractionated by overnight electrophoresis through either 1.2% (for RS2.0 polymorphism) or 1% (for all the other polymorphisms) agarose gels. After the run, the gel was photographed and was left on the UV source for 0.5-3 minutes, to break up the DNA fragments in situ. The gel was treated with 2 volumes of denaturing solution (table 2.4) for 1 hour to make the DNA single stranded and then washed in water briefly to remove excess denaturing solution. The gel was then washed in 2 volumes of neutralising solution (table 2.4) twice for 30 minutes each with gentle agitation. For blotting, a platform (2cm high) was placed in a tray and covered with 2 layers of 3MM Whatman paper. The ends of the filter paper was left touching the bottom of the tray so as to form a wick. The tray was filled with 20x SSC (table 2.4) to just below the level of the platform and the gel placed on the platform with its bottom surface uppermost. Saran wrap (Dow-Corning Corp, USA) was used to cover the areas of exposed filter paper surrounding the gel to ensure that the capillary action is through the gel only and also to reduce evaporation of the 20x SSC. A Hybond N nylon membrane (Amersham) was cut to size, pre-soaked in 2x SSC and laid in position over the gel. Air bubbles between the gel and membrane were excluded by rolling the surface with a pipette. Three pieces of 3MM Whatman paper of the same size as the gel were pre-soaked in 2x SSC and laid in position over the gel membrane, again excluding air bubbles. A stack of paper towels (approximately 10 cm thick) was then placed over the filter papers. The blotting assembly was left for 16-24 hours for the SSC solution to be drawn up into the paper towels by capillary action, passing through the gel and filter paper. During this process, DNA is eluted from the gel and, in the presence of high salt concentration, binds to the membrane. After eluting was complete, the membrane was marked with the position of the wells, washed in 2x SSC to remove any pieces of gel and then baked at 80°C for 2 hours. It was possible to check for complete transference of DNA by re-staining the gel in ethidium bromide and viewing it over UV light.
Table 2.4  Solutions used in Southern blotting and DNA hybridisation.

<table>
<thead>
<tr>
<th>Solution Type</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>20x SSC (Salt Sodium Citrate)</td>
<td>3M NaCl</td>
</tr>
<tr>
<td></td>
<td>0.3M Na$_3$C$_6$H$_5$O$_7$</td>
</tr>
<tr>
<td>Denaturing Solution</td>
<td>1.5M NaCl; 0.5M NaOH</td>
</tr>
<tr>
<td>Neutralising Solution</td>
<td>0.5M Tris base</td>
</tr>
<tr>
<td></td>
<td>0.3M sodium citrate</td>
</tr>
<tr>
<td></td>
<td>3M NaCl</td>
</tr>
<tr>
<td></td>
<td>pH 5.5 with conc. HCl</td>
</tr>
<tr>
<td>DNA Hybridisation Buffer (65°C)</td>
<td>4x SSC</td>
</tr>
<tr>
<td></td>
<td>10x Denhardt's</td>
</tr>
<tr>
<td></td>
<td>0.1% SDS</td>
</tr>
<tr>
<td></td>
<td>0.01 mg ml$^{-1}$ poly A</td>
</tr>
<tr>
<td></td>
<td>0.05 mg ml$^{-1}$ boiled/sonicated salmon sperm</td>
</tr>
<tr>
<td>10x Denhardt's solution</td>
<td>0.2% (w/v) BSA</td>
</tr>
<tr>
<td></td>
<td>0.2% (w/v) Ficoll</td>
</tr>
<tr>
<td></td>
<td>0.2% (w/v) PVP</td>
</tr>
</tbody>
</table>

101
2.15.1 DNA hybridisation

Pre-hybridisation and hybridisation of membranes with radio-labelled DNA probes was carried out in plastic bags sealed using an electric bag sealer (Impulse Sealer, TEW Heating Equipment Co.). Membranes were pre-hybridised for a minimum of 4 hours at 65°C to prevent non-specific binding of the probe. The rate and specificity of renaturation of single stranded DNA molecules depends on the temperature and salt concentration at which the reaction occurs. Therefore, hybridisation buffer contained a selected salt concentration (table 2.4) for use at 65°C. Denhardt's solution (table 2.4), sonicated and boiled salmon sperm and polyadenylic acid (poly A) were included in this buffer to reduce non-specific binding of DNA. After pre-hybridisation, the buffer was discarded and fresh buffer was introduced into the bag together with the denatured DNA probe with a specific activity of approximately $10^6$ counts ml$^{-1}$. The volume of buffer used was equivalent to 1/20th of the area of the membrane. Hybridisation was carried out for 16-24 hours at 65°C in a shaking water bath (Grant Instrument Ltd.). The bag was then cut open and the radioactive solution discarded. The membrane was washed at various stringencies, depending on the probes used. The washing solution contained varying concentrations of SSC from 0.5x (highest stringency used) to 2x (lowest stringency used) with 0.1% SDS. The washing procedure typically consisted of one 15 minute wash at room temperature and one or two 15 minute washes at 65°C. Washed membranes were wrapped in Saran wrap, placed in light proof cassettes and exposed to XAR-5 autoradiography film (Kodak) with intensifying screens (Cronex Quanta 3, Dupont) at -70°C for 16-72 hours or longer depending on the intensity of the signal following hybridisation. Saran wrap kept membranes moist so that the bound probe may be stripped off more easily if the membrane was to be re-used. For stripping, membranes were placed in a boiling hot 0.1% SDS solution and left to cool to room temperature by gentle agitation. They were then wrapped in Saran wrap and kept at 4°C until further use.
2.16 RNase PROTECTION

For RNase protection assays (Winter et al 1985), radio-labelled RNA probes (riboprobes) (see Results section 3.2.1), transcribed from cDNA fragments cloned into appropriate transcription vectors, were hybridised to cellular RNA obtained from lymphoblastoid cell lines (Section 2.7) and the nucleic acid hybrids were digested with RNase A. The products were analysed by denaturing polyacrylamide gel electrophoresis.

2.16.1 Generation of RNA probes

Riboprobes were generated from RB1 cDNA fragments which were cloned into multiple cloning sites (MCS) of either Bluescript M13 or pTZ18 transcription vectors. Immediately adjacent to the MCS in pTZ18 is the bacteriophage T7 promoter. Bluescript M13 contains both T7 and T3 promoters. Both vectors allow in vitro synthesis of RNA. The cDNA fragments were cloned adjacent to these promoters in an orientation that would result in the generation of antisense riboprobes when transcribed from either the T7 or T3 promoters. The plasmid vectors were first linearised by cutting from an appropriate restriction site downstream of the promoter used. After linearisation was checked on an agarose gel, a transcription reaction mixture (table 2.5) was set up using reagents from a riboprobe synthesis kit (Promega). Reagents were added in the order given in table 2.5 and the mixture was incubated at 37°C for one hour. To remove the DNA template, 2 units of RNase free DNase were added together with a further 20 units of RNasin ribonuclease inhibitor to protect the RNA. The mixture was incubated at 37°C for 15 minutes. The volume was then made up to 100 μl with DEPC-treated water (Section 2.1) and the solution was extracted with an equal volume of phenol/chloroform (50:50) mixture. After addition of 20 μg of tRNA (BRL) carrier, the riboprobe was precipitated with 10M ammonium acetate (final concentration 2M) and 3 volumes of ethanol on dry ice for 15 minutes or at -70°C for 30 minutes. The precipitate was recovered by spinning in a microfuge for 15 minutes. The pellet was vacuum dried and resuspended in 20 μl of formamide dye mix (table 2.5).
Table 2.5 Solutions used in the generation and purification of riboprobes.

<table>
<thead>
<tr>
<th>Riboprobe transcription reaction mixture</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5x transcription buffer</td>
<td>4 μl</td>
</tr>
<tr>
<td>0.1M DTT</td>
<td>2 μl</td>
</tr>
<tr>
<td>RNasin ribonuclease inhibitor</td>
<td>0.5 μl (20 units)</td>
</tr>
<tr>
<td>GTP, ATP, UTP (all 10mM)</td>
<td>1 μl from each</td>
</tr>
<tr>
<td>100 μM CTP</td>
<td>2.4 μl</td>
</tr>
<tr>
<td>Linearised plasmid DNA</td>
<td>1 μl (0.2-1 μg)</td>
</tr>
<tr>
<td>α32P CTP</td>
<td>2.5 μl (50 μCi)</td>
</tr>
<tr>
<td>T7 or T3 RNA polymerase</td>
<td>1 μl (10 units)</td>
</tr>
<tr>
<td>DEPC-treated water to 20 μl</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5x transcription buffer</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>200mM Tris-HCl, pH 7.5</td>
</tr>
<tr>
<td></td>
<td>30mM MgCl2</td>
</tr>
<tr>
<td></td>
<td>50mM NaCl</td>
</tr>
<tr>
<td></td>
<td>10mM spermidine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Formamide Dye Mix</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>80% formamide (deionised)</td>
</tr>
<tr>
<td></td>
<td>0.1% Bromophenol blue</td>
</tr>
<tr>
<td></td>
<td>0.1% Xylene Cyanol</td>
</tr>
<tr>
<td></td>
<td>20% TBE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maxam and Gilbert elution buffer</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5M NH₄OAc</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA, pH8</td>
</tr>
<tr>
<td></td>
<td>0.1% SDS</td>
</tr>
</tbody>
</table>
2.16.2 Purification of riboprobes

RNase protection assays require full-length labelled probes. Shorter transcripts, however, are also generated during synthesis. To obtain the correct size full-length transcript, the riboprobes were electrophoresed through 6% denaturing (7M urea) polyacrylamide gel. The riboprobe, in formamide dye mix, was heated to 90°C for 5 minutes and loaded onto the gel, which had been pre-run for 30 minutes, and run at 60 mA for 1-2 hours. The plates were then split apart and the gel was covered with Saran wrap. Labels, which were marked with radioactive ink, were attached to the four corners of the gel for orientation. The gel was exposed to XAR-5 autoradiography film for 30 to 45 seconds. A slot was cut in the exposed film where the desired riboprobe band appeared and the film was aligned with the radioactive markers. A scalpel was used to cut around the band and the Saran wrap was removed with forceps. The piece of acrylamide containing the probe band was lifted on the scalpel blade and placed in an eppendorf tube. The riboprobe was eluted overnight, in 300 μl of Maxam and Gilbert elution buffer (table 2.5), in a 37°C water bath. After spinning in a microfuge for 5 minutes, the supernatant containing the riboprobe was transferred to another eppendorf tube. The acrylamide gel piece was washed with a further 100 μl of elution buffer and the supernatant was added to the previous one. The supernatant was spun for 5 minutes to pellet any residual acrylamide and transferred to a fresh eppendorf. 900 μl of ethanol was then added and the riboprobe was precipitated at -70°C for 20 minutes. After centrifugation in a microfuge for 15 minutes, the pellet obtained was vacuum dried for 5 minutes and resuspended in 30 μl RNA hybridisation buffer (table 2.6) and kept at -20°C until used. The specific activity was typically $2.5 \times 10^6$ cpm μl$^{-1}$ of the probe.
Table 2.6 Solutions used in RNase protection assays.

<table>
<thead>
<tr>
<th>Hybridisation Buffer</th>
<th>RNase Digestion Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>80% (v/v) formamide (deionised)</td>
<td>10mM Tris-HCl, pH 7.5</td>
</tr>
<tr>
<td>40mM PIPES, pH 6.7</td>
<td>5mM EDTA</td>
</tr>
<tr>
<td>400mM NaCl</td>
<td>300mM NaCl</td>
</tr>
<tr>
<td>1mM EDTA, pH 8</td>
<td></td>
</tr>
</tbody>
</table>

2.16.3 Hybridisation and RNase digestion

20 to 40 μg aliquots of total cellular RNA obtained from lymphoblastoid cell lines (Section 2.7) were ethanol precipitated, vacuum dried and resuspended in 28 μl of RNA hybridisation buffer (table 2.6). 1 x 10^6 cpm of gel purified riboprobe was added to each sample and heated to 85°C for 20 minutes. The samples were then incubated at 50°C overnight to hybridise. To each hybridised sample, 300 μl of RNase digestion buffer (table 2.6) containing 20 μg ml⁻¹ RNase A and 1 μg ml⁻¹ RNase T1 was added and the reaction mixtures were incubated at 30°C for one hour. Digestion was terminated by addition of 20 μl of 10% SDS, 50 μg ml⁻¹ proteinase K and incubation at 37°C for 15 minutes. The samples were extracted with an equal volume of phenol/chloroform (50:50). After addition of 5 μg carrier tRNA, the RNA was precipitated with ethanol and left in dry-ice for 15 minutes or at -70°C for 30 minutes. The samples were spun in a microfuge for 15 minutes and washed in 70% ethanol. The pellets were vacuum dried and dissolved in 3 μl of formamide dye mix (table 2.5).

The samples were heated to 90°C for 5 minutes, loaded onto a 6% polyacrylamide/7M urea gel together with the appropriate molecular weight
markers (Section 2.11) and electrophoresed in TBE (table 2.2) for 1.5-3.0 hours at 30 mA. The gel was then fixed, transferred onto 3MM paper and dried (Section 2.10). Exposure to XAR-5 autoradiography film at -70°C was for 6-48 hours depending on the intensity of the bands.

2.17 POLYMERASE CHAIN REACTION (PCR)

Enzymatic amplification of DNA from various sources was carried out using the polymerase chain reaction (Saiki et al 1985, Mullis and Faloona 1987) using the thermostable Thermus aquaticus (Taq) DNA polymerase in a programmable thermocycler (Techne PHC-2). PCR is based on the use of two synthetically made oligonucleotides from opposite strands which bind to the target sequence and promote synthesis of the intervening sequences. Using repeated cycles of denaturation of the template DNA strands, annealing of primers and DNA synthesis, an exponential increase of a specific DNA fragment is achieved. Although for the majority of amplifications, isolated DNA dissolved in water or TE (after extraction from cells) was used as the template DNA, it was also possible to amplify DNA directly from white blood cells. For direct amplification, a small aliquot of cells was first denatured in 150 µl of 0.05M NaOH for 5 minutes at room temperature and then the solution was neutralised by addition of 35 µl of 1M Tris-HCl, pH 7.5. 5-10 µl aliquots were used in PCR. Oligonucleotide primers were synthesised on a phosphoramidite column (ICRF, Central Services Division). The PCR was carried out in a total volume of 50 µl and a typical reaction mixture was as given in table 2.7 except that, when the sequences amplified were G/C rich, 10% dimethyl sulphoxide was added to the PCR mix. The reaction mix was overlaid with mineral oil (Sigma) to prevent evaporation. Taq DNA polymerase (Promega or NBL) was added after an initial 15 minute denaturation step at 96°C. Amplification conditions consisted of 30 cycles of three steps: denaturation, annealing and extension at different temperatures as indicated in table 2.7. After amplification the mineral oil was removed by chloroform extraction. 5-10 µl aliquots were checked on agarose gels to ensure amplification of the appropriate sequence had occurred.
### Table 2.7 Reagents and conditions of PCR.

<table>
<thead>
<tr>
<th>PCR Mix</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
</tr>
<tr>
<td>Primers</td>
</tr>
<tr>
<td>10x Taq polymerase buffer</td>
</tr>
<tr>
<td>dGTP, dATP, dTTP, dCTP</td>
</tr>
<tr>
<td>Taq DNA Polymerase</td>
</tr>
<tr>
<td>Mineral oil to cover</td>
</tr>
</tbody>
</table>

| 10x Taq polymerase buffer                    | 100mM Tris-HCL, pH 8. |
|                                              | 500mM KCl |
|                                              | 15mM MgCl$_2$ |
|                                              | 1% Triton X-100 |

### Conditions

Initial denaturation : 96°C; 15 minutes

30 cycles of:

| Denaturation:          | 94°C; 20-30 seconds |
| Annealing:             | a temperature optimal for the particular primer pair; |
| Extension:             | 72°C; 20-30 seconds |

**Amplification Dilution Solution (ADS)**

| 0.1% SDS |
| 10mM EDTA, pH 8 |

108
2.18 PCR BASED DETECTION OF POLYMORPHIC SITES

Primers designed from sequences flanking the polymorphic sites (see Results, table 3.2) were used to amplify polymorphic regions in the DNA. The PCR amplification procedure was as described in Section 2.17. After amplification was checked on agarose gels, the amplified product was digested (Section 2.8) with the appropriate restriction endonuclease. The resulting DNA fragments were resolved by electrophoresis through 1.4 to 2.0% agarose gels and visualised by ethidium bromide staining (Section 2.9).

To detect the variable number tandem repeat sequence (VNTR-RB1.20, see Results, Section 3.1.2) polymorphism, PCR was as described in Section 2.17 except that 1 μCi α³²P-dCTP (3000 Ci mmol⁻¹) was added to the PCR mix to generate a labelled product and the 'cold' dCTP concentration was reduced to 0.02mM. 1 μl of the PCR product was diluted tenfold in ADS (table 2.7) and 1 to 2 μl of this solution was mixed with 2 μl of stop solution (table 2.2). The sample was heated to 95°C for 5 minutes for denaturation, and resolved by electrophoresis at 60 W constant power on a polyacrylamide 'wedge' gel containing 7M urea. The gel was transferred onto 3MM paper and dried (Section 2.10). Autoexposure to XAR-5 autoradiography film was at -70°C for 16-24 hours with Cronex Quanta 3 intensifying screen.

2.19 CARBODIIMIDE MODIFICATION OF HETERO-DUPLEXES OF PCR PRODUCTS AND PRIMER EXTENSION

Carbodiimide (CDI) (Sigma) modification of heteroduplexes followed by primer extension was performed using the method of Ganguly and Prockop (1990). This is a method to detect single base changes in the DNA and is based on the chemical modification of unpaired Gs and Ts in the DNA by CDI. PCR amplification from DNAs that differ by a single base was as described in PCR Section (2.17). PCR products were extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1) and once with chloroform/isoamyl alcohol (24:1). The DNA was precipitated for one hour at -70°C by addition of 0.5 volume of 7.5M ammonium acetate and 2 volumes ethanol. The pellets were recovered by centrifugation in a
microfuge for 15 minutes, washed in 70% ethanol and dried. They were then resuspended in 50 μl of TE pH 7.5.

2.19.1 Generation of heteroduplexes

20 μl from each PCR product was mixed in an eppendorf tube. 10 μl of hybridisation buffer (table 2.8) was added and the total volume was made up to 120 μl with water. After addition of a drop of mineral oil, the eppendorf tube was placed in a boiling water bath for 10 minutes to denature the DNA. The reaction mix was then placed in a 42°C water bath overnight to allow reannealing. After extraction with chloroform/isoamyl alcohol (24:1), DNA was precipitated as before. The pellet was washed in 70% ethanol, dried and dissolved in 60 μl TE, pH 7.5.

2.19.2 CDI modification of heteroduplexes

A fresh solution of CDI (table 2.8) was prepared just before the modification reaction which was carried out under two different conditions. Firstly sodium borate, pH 8, to 0.1M, was added to 26 μl of the heteroduplex DNA with 10 μl of CDI solution and incubated at 30°C for 3 hours. Secondly sodium borate, to 0.01M, was added and incubated at 37°C for one hour. To remove unreacted CDI, 40 μl of 10mM sodium phosphate, pH 7, and 40 μl of 7.5M ammonium acetate were added and the samples were extracted 3 times with 480 μl of isoamyl alcohol freshly equilibrated with a solution 2.5M ammonium acetate in TE, pH 7.5. The aqueous lower phase was transferred to a siliconised eppendorf tube (CDI-treated DNA adheres to plastic surfaces) and the DNA was precipitated at -20°C overnight with 3 volumes of ethanol. The pellet was washed and dried as before and dissolved in 26 μl of TE, pH 7.5. The samples were then heated in a boiling water bath for 5 minutes to destroy any residual CDI and then cooled on ice for 5 minutes.
Table 2.8 Solutions used in heteroduplex formation, CDI modification and primer extension.

<table>
<thead>
<tr>
<th>Hybridisation Buffer for heteroduplex formation</th>
<th>CDI solution</th>
<th>PCR Cocktail (for 20 reactions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M NaCl</td>
<td>200mM CDI in water (84.7 mg ml(^{-1}))</td>
<td>38 (\mu)l water</td>
</tr>
<tr>
<td>35mM MgCl(_2)</td>
<td></td>
<td>32 (\mu)l of dNTP solution (1.25mM each of dGTP, dATP, dTTP and 0.625mM dCTP)</td>
</tr>
<tr>
<td>30mM Tris HCl, pH 7.5</td>
<td></td>
<td>10 (\mu)l of (\alpha^{32}P) dCTP (0.1mCi, 3000 Ci mmol(^{-1}))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20 (\mu)l of 10x Taq polymerase buffer (table 2.7)</td>
</tr>
</tbody>
</table>

2.19.3 Primer extension

A cocktail for PCR was prepared (table 2.8) and heated at 94°C for 5 minutes to inactivate any proteases. After cooling on ice for 5 minutes, 2.5 units of Taq DNA polymerase was added. PCR reactions were carried out in siliconised eppendorf tubes by adding 4 \(\mu\)l of the CDI modified DNA, 1 \(\mu\)l (5 pmol) of the appropriate primer and 5 \(\mu\)l of the PCR cocktail. The reaction mix was overlaid with mineral oil. The PCR cycle consisted of 3 minutes at 94°C; 1 minute at 56°C and 3 minutes at 72°C. After one cycle, the reaction was stopped by adding stop solution (table 2.2). The samples were then heated at 94°C for 2 minutes and 8 \(\mu\)l aliquots were loaded on a 5% polyacrylamide 7M urea gel together with the appropriate molecular weight markers (Section 2.11). Electrophoresis was at 30 W constant power for 3.5 hours. The gel was fixed, dried and exposed to XAR-5 autoradiography film (Section 2.10), at -70°C, overnight.
2.20 SINGLE STRAND CONFORMATION POLYMORPHISM (SSCP)

For SSCP analysis (Orita et al 1989a), PCR amplified DNAs were denatured and run in non-denaturing polyacrylamide gels. PCR amplification was as described in Section 2.17 except that 1 µCi of α32P-dCTP (3000 Ci mmol⁻¹) was added to the PCR mix to generate a labelled product and the 'cold' dCTP concentration was reduced to 0.02mM. When the amplified PCR product was longer than 300 bp, shorter fragments were generated for SSCP analysis by restriction enzyme digestion (Section 2.8). A 5 µl aliquot of the digested or undigested PCR product was diluted with 40 µl of ADS (table 2.7) to minimise re-annealing after denaturation and 2 µl of this dilution was mixed with 2 µl stop solution (table 2.2). Samples were denatured at 95°C for 3 minutes and placed immediately on ice to prevent renaturation before loading on 6% polyacrylamide non-denaturing gels containing 10% glycerol. The DNA was electrophoresed at 30 W constant power at room temperature for 6 hours. Gels were transferred to 3MM paper, dried and exposed to XAR-5 film (Section 2.10) at -70°C for 12-72 hours without intensifying screens.

2.21 DIRECT SEQUENCING OF PCR PRODUCTS

Sequencing was performed by the dideoxy chain termination method (Sanger et al 1977) directly from PCR products after separation of the individual strands of the DNA.

2.21.1 Separation of DNA strands

For separation of the DNA strands, one of the primers used in the PCR reaction was biotinylated at the 5' end allowing immobilisation of the DNA on streptavidin-coated magnetic 'Dynabeads' (Dynal). The PCR reaction was as described in Section 2.17 except that the amount of primer used was decreased to 15 pmoles to avoid saturation of streptavidin with excess biotinylated primer. Separation of DNA using Dynabeads was carried out according to the manufacturer's instructions (Dynal). The Dynabeads beads were washed in 100 µl of TES (table 2.9), immobilised on a magnetic particle concentrator (MPC, Dynal) and the supernatant removed. The PCR products
(after removal of mineral oil) were incubated with pre-washed Dynabeads for 5 minutes at room temperature to allow immobilisation of DNA on the beads. Supernatant was removed using the MPC and the remaining beads were resuspended in 100 μl of 0.15M NaOH for 5 minutes at room temperature to denature the double stranded DNA. Both single strands of the DNA could then be isolated, the biotinylated strand being bound to the beads and the non-biotinylated strand remaining in the supernatant. The non-biotinylated strands were precipitated by adding 1/3 volume of 3M sodium acetate, pH 5.2, and 2.5 volumes of 100% ethanol at -70°C overnight. DNA was recovered by centrifugation and the pellet was washed in 70% ethanol, dried and resuspended in 5 μl of water. The immobilised biotinylated DNA was washed with 100 μl of TES, followed by 100 μl of water and finally resuspended in 5 μl of water.

### 2.21.2 Dideoxy sequencing

Both strands were sequenced using a Sequenase (version 2.0) kit (USB) according to the manufacturer's instructions except that a 1:15 dilution of labelling mix was used. Sequenase is a genetic variant of bacteriophage T7 DNA polymerase created by *in vitro* genetic manipulation (Tabor and Richardson 1989). To anneal the primer to the template DNA, 0.5 pmol of the appropriate primer, 2 μl of 5x Sequenase buffer (USB) (table 2.9) and the single stranded DNA was incubated at 65°C for 3 minutes and then allowed to cool slowly to room temperature. DNA synthesis was then carried out in 2 steps; a labelling step and the chain termination step. To the annealed template-primer, 1 μl of 0.1M DTT, 2 μl of 1:15 diluted labelling mix (table 2.9), 0.5 μl (5 μCi) α^{35}S-dATP (1000 Ci mmol⁻¹) and 2 μl of 1:8 diluted (table 2.9) Sequenase were added. The reaction mix was incubated for 4 minutes at room temperature allowing the primer to be extended using limited concentrations of dNTPs which included labelled dATP. For chain termination 3.5 μl of this mix was added to each of four eppendorfs containing 2.5 μl of one of four ddNTPs (ddGTP, ddATP, ddTTP, ddCTP) which lack the 3' -OH group necessary for DNA chain elongation. Incubation was at 37°C for 5 minutes. In this step, DNA synthesis occurs until all growing chains are terminated by a ddNTP. The reactions were terminated by adding 4 μl of stop solution (table 2.2). When sequencing was from a biotinylated template immobilised on magnetic beads, the samples
were heated to 65°C for 5 minutes and placed on the MPC to remove the beads. The supernatants containing the sequencing reactions were then transferred to fresh eppendorf tubes and 4 µl aliquots were electrophoresed on 4-6% denaturing polyacrylamide 'wedge' gels (allow greater reading capacity) in TBE buffer (table 2.2) at 60 W constant power for 1.5 to 5 hours depending on the length of the sequence to be read. For runs over 3 hours a long run TBE buffer (table 2.2), which gave better results under these circumstances, was used. The gels were fixed, transferred onto 3MM paper and dried (Section 2.10). Autoradiography was at room temperature for 12-72 hours.

Table 2.9 Solutions used in separation of DNA strands and dideoxy sequencing.

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TES</strong></td>
<td>10mM Tris-HCl, pH 8</td>
</tr>
<tr>
<td></td>
<td>1mM EDTA, pH 8</td>
</tr>
<tr>
<td></td>
<td>0.1M NaCl</td>
</tr>
<tr>
<td><strong>5x Sequenase Buffer (USB)</strong></td>
<td>200mM Tris-HCl, pH 7.5</td>
</tr>
<tr>
<td></td>
<td>100mM MgCl₂</td>
</tr>
<tr>
<td></td>
<td>250mM NaCl</td>
</tr>
<tr>
<td><strong>Labelling Mix (5x concentrate, USB)</strong></td>
<td>7.5 µM each of dGTP, dCTP, dTTP</td>
</tr>
<tr>
<td><strong>Enzyme Dilution Buffer (USB)</strong></td>
<td>10mM Tris-HCl, pH 8</td>
</tr>
<tr>
<td></td>
<td>5mM DTT</td>
</tr>
<tr>
<td></td>
<td>0.5 mg ml⁻¹ BSA</td>
</tr>
</tbody>
</table>
3.0 RESULTS
3.1 APPLICATION OF INTRAGENIC POLYMORPHISMS IN SCREENING FOR MUTANT GENE CARRIER STATUS IN RETINOBLASTOMA FAMILIES

At the outset of this thesis the objectives were to study the nature of the genetic variation in patients with hereditary Rb and to determine whether, using the RB1 gene, pre-natal screening and carrier detection could be provided for the UK population. Although linkage was the main means of achieving these objectives, methods to identify the causative mutations in hereditary cases were investigated so that identification of mutant gene carriers would be unequivocal. In addition, by investigating mutational epidemiology it might be possible to understand more about the function of the RB1 gene.

To carry out a meaningful study of Rb genetics and comment about rare events it is necessary to have a large series of patients. The majority (70-75%) of Rb patients in the UK are referred to the Ophthalmology Departments at Moorfields Eye Hospital and St. Bartholomew's Hospital. Over the past ten years the ICRF unit at the Institute of Child Health (ICH) has provided an ESD quantitation service for Rb patients in this series for the detection of chromosome deletions. As such, blood samples from a large number of families have been processed and were available for analysis. Some families, however, have also been referred to the unit at the ICH directly from other regions in the UK. 57 Rb families from this series were studied using intragenic DNA sequence polymorphisms in this project.

The pedigrees of the 57 Rb families studied are given in Appendix III. In each case the eyes of the affected individuals have been examined by an experienced ophthalmologist as Rb can be confused, phenotypically, with other conditions such as Coat's disease (Section 1.3.1). All tumours from enucleated eyes were confirmed histopathologically as Rb. Blood samples were collected from as many relevant family members as possible, either during their routine visits to the referral centres or by the family General Practitioners. For many of the key family members lymphoblastoid cell lines were also generated. For pre-natal screening, samples were obtained by chorionic villus sampling performed by Dr K Nikolaides at the Harris Birthright Research Centre, King's College, London.
In the linkage analysis described below, seven RB1 intragenic polymorphisms were used, the approximate positions of which are indicated in figure 3.1. The summary of results obtained from the study are given in table 3.3. A more detailed table of results is provided in Appendix IV. The results of the study are presented in two sections, the first using RFLP analysis (Section 3.1.1) and the second using other polymorphisms (Section 3.1.2).

Figure 3.1 A schematic representation of the RB1 gene showing the approximate, relative positions of polymorphic sites. Bam HI, Kpn I, Xba I and Tth III I polymorphisms are detected by probes M1.8, HS0.5, PRO.6 and R0.6 respectively as shown in table 3.1. Probe RS2.0 detects a VNTR in intron 17 (table 3.1). RB1.3 is a single base pair polymorphism and RB1.20 is another VNTR region located in intron 20 (table 3.2 and Section 3.1.2). Vertical bars indicate the position of exons which are numbered 1 to 27. The position of two very big (35 kb and 70 kb) introns (V) are indicated.

3.1.1 Studies using restriction fragment length polymorphisms (RFLPs)

Naturally occurring DNA sequence variation may affect a restriction enzyme site which, in turn, results in variation between individuals in the length of DNA fragments generated by that specific restriction enzyme. Four of the polymorphisms analysed in this study were examples of such RFLPs. The variation in the length of DNA fragments generated in one other polymorphism was due to the presence of a variable number of tandem repeats (VNTR) within the restriction fragment. Details about these recombinant DNA probes (which were a kind gift of Dr T Dryja) from within the RB1 gene are summarised in table 3.1. In all cases purified insert was used as a probe after release from the Bluescribe plasmid vector using the appropriate enzymes (Appendix V). The probes are called M1.8, HS0.5, PRO.6, R0.6 and RS2.0 and they identify polymorphic sites for enzymes
Bam HI, Kpn I, Xba I, Tth III I and the intron 17 VNTR respectively (Wiggs et al 1988, Bookstein et al 1988). These polymorphic sites will be referred to by the name of their respective probes in this thesis for the sake of simplicity and to avoid confusion. In order to assess the usefulness of these RFLPs it was first important to study the allele frequencies in the UK population which is given in table 3.1. These frequencies are comparable with those from the North American population (Wiggs et al 1988, Goddard et al 1990) with the exception of the RS2.0 locus (see later).

**Table 3.1** Characteristics of the DNA probes used in RFLP analysis and their allele frequencies in the UK and North American population. PIC* values are calculated for each polymorphism using the UK allele frequencies alone.

<table>
<thead>
<tr>
<th>DNA probe</th>
<th>Restr. Enzyme</th>
<th>Intron position</th>
<th>Allele size (kb)</th>
<th>Allele freq. UK (North Am)</th>
<th>PIC*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS2.0</td>
<td>Rsa I</td>
<td>17</td>
<td>2.00</td>
<td>0.07(0.13)</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.95</td>
<td>0.12(0.02)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.90</td>
<td>0.27(0.07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.85</td>
<td>0.22(0.07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.80</td>
<td>0.16(0.35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.75</td>
<td>0.04(0.20)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.65</td>
<td>0.06(0.09)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.60</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.50</td>
<td>0.01(0.09)</td>
<td></td>
</tr>
<tr>
<td>PR0.6</td>
<td>Xba I</td>
<td>17</td>
<td>7.8</td>
<td>0.46(0.55)</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.5</td>
<td>0.54(0.45)</td>
<td></td>
</tr>
<tr>
<td>M1.8</td>
<td>Bam HI</td>
<td>1</td>
<td>4.5</td>
<td>0.30(0.30)</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.3/2.2</td>
<td>0.70(0.70)</td>
<td></td>
</tr>
<tr>
<td>R0.6</td>
<td>Tth III I</td>
<td>26</td>
<td>4.95</td>
<td>0.20(0.20)</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.35</td>
<td>0.80(0.80)</td>
<td></td>
</tr>
<tr>
<td>HS0.5</td>
<td>Kpn I</td>
<td>2</td>
<td>12.0</td>
<td>0.86(0.95)</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.0</td>
<td>0.14(0.05)</td>
<td></td>
</tr>
</tbody>
</table>

*PIC (Polymorphism Information Content) calculated using the formula given in Section 1.4.6.1.
Representative autoradiographs, using each of the DNA probes are shown in figure 3.2. The RS2.0 probe recognises a VNTR where the repeat unit is between 50 and 53 bp long (figure 3.4). Eight distinct alleles were reported by Wiggs et al (1988) and Scheffer et al (1989) reported an additional 1.55 kb allele with a frequency of 0.01 in the Dutch population. In this project, the 1.55 kb allele was not detected but instead a 1.60 kb allele was (table 3.1, figure 3.2). All 10 different alleles differ in size by 50 bp and are contained within a single Rsa I fragment. Three of the alleles are present in 60% of individuals, the others being relatively rare (table 3.1). Conventional Southern blot analysis will not usually allow DNA fragments which differ by so little in size to be distinguished. It was necessary, therefore, to separate the DNA fragments on 1.2% agarose gels which were no more than 4 mm thick to prevent diffusion. Overnight electrophoresis was essential to achieve good separation. In practice, the gels were run until the 1.3 kb fragment of the λHHR molecular weight marker (Section 2.11, Appendix II) approached the bottom of a 20 cm gel. It was also important that minimal distortion of the gel occurred during blotting, so minimum weight was placed on the nylon membrane. Although, under these conditions, it was possible to distinguish between alleles differing in size by only 50 bp, variations in gel thickness, conditions which result in non-homogeneity of the gel or local heat gradients during electrophoresis sometimes distorted the migration rate of the DNA restriction fragments. λHHR markers were loaded onto both sides of the gel and they were often observed to have migrated at slightly different rates compared to each other. Although, within a single gel, the Mendelian inheritance could usually be followed easily, it was sometimes difficult to establish the exact sizes of alleles which made comparing allele sizes between gels difficult. This would account for variations of the calculated frequencies of different size alleles (table 3.1) within the population and might possibly explain why different groups report different allele frequencies for this polymorphism.

Using the restriction enzymes Tth III I and Kpn I, achieving complete digestion of genomic DNA was often difficult and, especially with Kpn I digested DNA, the resulting bands in the autoradiograph were weak (figure 3.2e). Given this problem and the low frequencies of the rare alleles involved, only a few families were analysed using these polymorphisms and mostly represented those which were uninformative with the other probes.
Figure 3.2  Examples of the polymorphic banding profiles for (a) the RS2.0, (b) PR0.6, (c) M1.8, (d) R0.6 and (e) HS0.5 polymorphisms as identified by Southern blotting and hybridisation techniques. Allele sizes in kb pairs are given at the sides.
At the time of writing Southern blotting was the only means of identifying the HS0.5, RS2.0 and R0.6 polymorphisms. The DNA sequence of the M1.8 and PR0.6 polymorphisms, however, were available, allowing PCR primers to be constructed (Bookstein et al. 1990c, McGee et al. 1990). Using the method described in Section 2.18 and the primers given in table 3.2, PCR products around the polymorphic sites were generated. To analyse these polymorphisms, amplified products were digested with the appropriate enzyme and resolved in 2-2.5% agarose gels for M1.8 and 1.4% gels for the PR0.6 polymorphism. Representative gels showing the amplified products before and after digestion are shown in figure 3.3. For the M1.8 polymorphism a 188 bp PCR fragment is generated. The polymorphic Bam H1 site is located 53 bp from the 3' end. Thus, in heterozygotes, fragments of 188 (upper allele), 135 and 53 bp (lower allele) are generated after Bam H1 digestion. Curiously, however, compared with the 1 kb ladder markers (Section 2.11, Appendix II) on 2-2.5% agarose gels, the fragments always appeared slightly larger at 200, 140 and 60 bp (figure 3.3a). For the PR0.6 polymorphism, a 945 bp fragment is amplified which, following Xba I digestion identifies two alleles; an upper allele 945 bp long and a lower allele consisting of the two digestion fragments which are 630 bp and 315 bp long (figure 3.3b). There were, however, some difficulties in analysing the M1.8 polymorphism using PCR. The polymorphic Bam H1 site is situated in a highly GC rich region. It was necessary, therefore, to include 10% DMSO in the PCR cocktail to achieve amplification since these GC nucleotides not only predominate in the primers used but also comprise 75% of the amplified product (Bookstein et al. 1990c) (see also Section 3.2.2). DMSO decreases inter- or intrastrand base pairing (secondary structure) in the DNA (Winship 1989, Saiki 1989). In these experiments, amplification of non-specific bands was frequently observed (figure 3.3a).

Clearly, for a particular polymorphism to be informative for linkage in any given family, the individuals transmitting the predisposing mutation must be heterozygous at that locus so that both copies of the RBI gene can be distinguished. As a preliminary screen, therefore, key members from each family were studied first. Where these individuals were found to be heterozygous for a particular polymorphism, the rest of the family was analysed. Most of the families (52 out of 57) were studied using more than one polymorphism. Although the majority were only informative for one RFLP, a few were informative for several (see later and Appendix IV).
Figure 3.3  PCR based detection of RFLPs. The full length PCR products are shown in the left hand side gels for (a) the M1.8 and (b) PRO.6 polymorphisms. Following digestion with Bam HI (a) and Xba I (b) (right) homozygotes and heterozygotes can be identified. The sizes of fragments in base pairs are given at the sides. M = marker lane containing 1Kb DNA ladder (Gibco-BRL) (Section 2.11, Appendix II).
Table 3.2 Description of the primers used for PCR amplification of the regions containing the polymorphic sites within the RBI gene.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Polymorp. Type</th>
<th>Primers: Sense (5'-3')</th>
<th>Antisense (5'-3')</th>
<th>Location</th>
<th>Fragment size (bp)</th>
<th>Anneal. Temp°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1.8</td>
<td>RFLP (Bam HI)</td>
<td>CAGGACACGGCCGGCCGGAG</td>
<td>CTGCAGACGCTGCCCCT</td>
<td>exon 1</td>
<td>188</td>
<td>60</td>
</tr>
<tr>
<td>PR0.6</td>
<td>RFLP (Xba I)</td>
<td>TCCAATGAAGAACAAATGG</td>
<td>GCAATTGCACAAATCCAAGT</td>
<td>intron 17</td>
<td>945</td>
<td>52</td>
</tr>
<tr>
<td>RBI.20</td>
<td>VNTR</td>
<td>GTATGAACCTCATGAGACAGCAT</td>
<td>AATTAACAGGTGTGTTGACACG</td>
<td>exon 20</td>
<td>300-360</td>
<td>60</td>
</tr>
<tr>
<td>RBI.3</td>
<td>Single bp</td>
<td>TTGACCTAGATGAGATGTCGACC</td>
<td>GGCAGTTCACATTTTTGGTCCAAGT</td>
<td>exon 3</td>
<td>200</td>
<td>59</td>
</tr>
</tbody>
</table>

The VNTR polymorphism, detected with probe RS2.0, proved to be the most useful with around 60% of families studied being informative at this locus (table 3.3). A part of the repeated sequence (figure 3.4) has homology to core sequences of some other VNTRs (Nakamura et al 1987). Because such tandemly repeated sequences tend to be genetically unstable, the number of units is highly variable among individuals, giving rise to multiple alleles. The usefulness of this multiallelic polymorphism is reflected in its polymorphism information content (PIC) value which approaches 1 (table 3.1). This value measures the degree of polymorphism (Botstein et al 1980, Section 1.4.6.1). All the other polymorphisms were informative in varying degrees as shown in table 3.3. The percentage informativity of HS0.5, however, does not accurately reflect the usefulness of this polymorphism because, in fact, only a small number of families were studied.
Figure 3.4 DNA sequence of the 50 to 53 bp repeat of the VNTR (Intron 17) detected by the RS2.0 probe. The regions showing variability within the repeat sequence are given in brackets with the variation shown above and below. The 11 bp sequence representing the core sequence of other VNTRs reported by Nakamura et al (1987) are shown above with homologous sites indicated by the vertical bars.

In this study, 45 of the 57 families analysed (79%) were informative for at least one of the RFLPs. 19 families (33.3%) were also informative for at least one other RFLP and 12 families remained uninformative of which 9 were studied using at least four of the RFLPs. The remaining 3 were only studied with the most informative RFLPs (RS2.0, PR0.6, M1.8). The HS0.5 and/or R0.6 polymorphisms were not used in the analysis in some of these families, partly because of the low frequency of the rare allele, but also because of time constraints and the subsequent availability of a more informative polymorphism (Section 3.1.2). Taken together, therefore, approximately 80% of families could be offered screening for mutant gene carrier status using the 5 intragenic RFLPs. It was still important to investigate the other 20%, however, using other polymorphic sites.
Table 3.3  Summary of informativity of the polymorphic sites in Rb family studies.

<table>
<thead>
<tr>
<th>RFLP</th>
<th>Number*</th>
<th>Informative No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS2.0 (Rsa I)</td>
<td>57</td>
<td>35 (61.4)</td>
</tr>
<tr>
<td>PR0.6 (Xba I)</td>
<td>51</td>
<td>18 (35.3)</td>
</tr>
<tr>
<td>M1.8 (Bam HI)</td>
<td>47</td>
<td>13 (27.7)</td>
</tr>
<tr>
<td>R0.6 (Tth III I)</td>
<td>21</td>
<td>4 (19.0)</td>
</tr>
<tr>
<td>HS0.5 (Kpn I)</td>
<td>10</td>
<td>3 (30.0)</td>
</tr>
</tbody>
</table>

Other polymorphisms

<table>
<thead>
<tr>
<th>RFLP</th>
<th>Number*</th>
<th>Informative No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB1.20 (VNTR)</td>
<td>30</td>
<td>26 (86.7)</td>
</tr>
<tr>
<td>RB1.3 (single bp)</td>
<td>9</td>
<td>1 (11.1)</td>
</tr>
</tbody>
</table>

* Number of families studied with the particular polymorphisms.

3.1.2 Analyses of families with the RB1.20 and RB1.3 polymorphisms

During the course of the analysis of the Rb families in our series two other polymorphisms were identified which are not detectable as RFLPs, since neither involved the recognition site of a known restriction enzyme. Thus, the RB1.20 polymorphism (Yandell and Dryja 1989) consists of a variable number (n=14-26) of [CTTT(T)]n repeats and occurs 54 bp from the 3' end of exon 20 (figure 3.1). This polymorphism could only be identified by electrophoresis of PCR-amplified products through denaturing polyacrylamide gels as described in Section 2.18. The primers used for this amplification are given in table 3.2. The RB1.20 polymorphism consists of many alleles which can differ by as little as 1 bp. It was difficult, therefore,
as with RS2.0, to estimate allele sizes between autoradiographs. For this reason the alleles were numbered separately in each family, labelling them sequentially according to size, the largest being '1' and so on. For the same reason, it was not possible to calculate allele frequencies and PIC values for this polymorphism.

The second polymorphism which became available was RB1.3 (Yandell and Dryja 1989), which is a single base pair polymorphism, located 45 bp from the 3' end of exon 3 (figure 3.1); either a C or a T (coding strand) is present at this site. This polymorphism could only be detected unequivocally by sequencing the PCR products directly (Section 2.21) and the primers used for this amplification are given in table 3.2. Representative autoradiographs of both the RB1.20 and RB1.3 polymorphisms are given in figure 3.5.

RB1.20 proved to be the most informative (approx. 87%) of all the polymorphisms studied so far (table 3.3). Of 68 unrelated individuals 71% were heterozygous at this locus and of those 12 families that were uninformative using the other RFLPs, 10 were found to be informative for RB1.20. In the North American population Yandell and Dryja (1989) found that 94% of individuals were heterozygous for RB1.20.

Because of the large number of highly informative polymorphic probes which were already available in RB1, only 9 families were eventually analysed with RB1.3 and only 1 was found to be informative (table 3.3). In another family, although the key affected member (the parent) was homozygous, the affected child was now heterozygous for RB1.3 which will be important for screening his children in the future. From 19 unrelated individuals the frequency of the rarer allele (T) was found to be 0.21 and the PIC value was 0.28. This compares favourably with the frequency of 0.27 for individuals of mixed North American descent (Yandell and Dryja 1989).
Figure 3.5 Representative gels showing the RB1.20 VNTR and the RB1.3 single base pair polymorphism.  (a) The RB1.20 VNTR region was amplified and labelled using PCR as described in Section 2.18. The amplified products were size fractionated on a 6% denaturing polyacrylamide gel. The sizes of the amplified fragments ranged from 300-360 bp depending on the number of "CTTT(T)" repeats. The alleles are numbered sequentially by size, the largest being labelled '1', etc.  (b) The non-coding DNA strand sequence of the RB1.3 locus from three different individuals is shown. Dideoxy sequencing of the PCR products was carried out as described in Section 2.21. The genotypes of individuals are indicated at the top of the gels and the arrow indicates the position of the polymorphic base. GG: homozygous for G; AA: homozygous for A; GA: heterozygous.
When the analyses of polymorphic loci in the RB1 gene are combined, 55 out of 57 (96%) families were found to be informative for at least one of the polymorphisms. The 2 remaining families were universally uninformative although one of these has not been analysed with HS0.5 (Kpn I). Occasionally it is not possible to coordinate the collection of blood samples from patients through GPs, consultants etc. This was the case in a further 7 families where only the key affected members have been analysed of which 6 were found to be heterozygous for at least one of the polymorphisms; the other one was homozygous at all loci examined. Since the transmitting parent is heterozygous, it is highly probable that the family will be informative and could probably be offered genetic screening. In summary, the total number of families that could be offered screening using intragenic polymorphisms represents 95% (61/64).

3.1.3 Lod score analysis of family linkage data

In the early stages of this analysis, although intragenic RFLPs were being used it was not clear whether recombination would be an important factor in interpreting the data. It was important, therefore, to establish lod (logarithm of the odds ratio) scores (Section 1.4.6.1) from RFLP data. The LIPED 6 computer program (Ott 1974) was used in this analysis. Family pedigree data were prepared for lod score analysis using the program LINKSYS (Attwood and Bryant 1988). The frequency of mutant alleles at the RB1 locus was assumed to be 1:10000. The new mutation rate was assumed to be 1.5 x 10^{-5} (Vogel 1979). As no recombinations were recorded, haplotypes were constructed for each individual using the intragenic probes. Lod scores were calculated for linkage between the Rb phenotype and intragenic RB1 probes for 13 of the families studied in this project and combined with 8 other families analysed previously by Dr CD Mitchell, assuming a penetrance of 90% (Onadim et al 1990, Appendix III). The results of lod score analyses are summarised in table 3.4.
Table 3.4 Lod scores (at $\Theta=0$) calculated for linkage between the Rb phenotype and RB1 intragenic probes for 21 Rb families (Appendix III) and the total lod scores at various recombination fractions from the same data.

<table>
<thead>
<tr>
<th>Family</th>
<th>Lod score at $\Theta=0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Penetrance 90%)</td>
</tr>
<tr>
<td>RBF02</td>
<td>0.301</td>
</tr>
<tr>
<td>RBF03</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF04</td>
<td>0.048</td>
</tr>
<tr>
<td>RBF05</td>
<td>0.223</td>
</tr>
<tr>
<td>RBF06</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF07</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF08</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF09</td>
<td>0.301</td>
</tr>
<tr>
<td>RBF10</td>
<td>0.561</td>
</tr>
<tr>
<td>RBF11</td>
<td>0.862</td>
</tr>
<tr>
<td>RBF12</td>
<td>-0.030</td>
</tr>
<tr>
<td>RBF13</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF14</td>
<td>0.422</td>
</tr>
<tr>
<td>RBF15</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF16</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF17</td>
<td>0.125</td>
</tr>
<tr>
<td>RBF18</td>
<td>0.085</td>
</tr>
<tr>
<td>RBF19</td>
<td>0.260</td>
</tr>
<tr>
<td>RBF20</td>
<td>0.204</td>
</tr>
<tr>
<td>RBF21</td>
<td>0.561</td>
</tr>
<tr>
<td>RBF22</td>
<td>0.204</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5.947</td>
</tr>
</tbody>
</table>

Total lod scores at various recombination fractions ($\Theta$)

<table>
<thead>
<tr>
<th>$\Theta$</th>
<th>0.000</th>
<th>0.001</th>
<th>0.005</th>
<th>0.100</th>
<th>0.150</th>
<th>0.200</th>
<th>0.250</th>
<th>0.300</th>
<th>0.350</th>
<th>0.400</th>
<th>0.450</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLS*</td>
<td>5.947</td>
<td>5.844</td>
<td>5.152</td>
<td>4.413</td>
<td>3.666</td>
<td>2.942</td>
<td>2.247</td>
<td>1.598</td>
<td>1.029</td>
<td>0.554</td>
<td>0.204</td>
</tr>
</tbody>
</table>

* Total Lod Score
In most of the families used to calculate the lod score (Appendix III), the inheritance of the Rb phenotype followed the normal dominantly inherited pattern. Two families (one of whom, RBF12, was investigated by Dr CD Mitchell), however, were atypical. Family RBF12 (figure 3.6a) was unusual in that each of the affected members only had either a unifocal tumour or evidence of spontaneously regressed tumours. Phenotypic analysis indicated that there was an apparently unaffected transmitting member (II.7). This subject had died some time ago so was not available for ophthalmological examination; it is possible he also had regressed tumours. As will be discussed later these "mild" phenotypes tend to concentrate in families, a fact reinforced by the demonstration that another member of the family (III.3) showed evidence of incomplete penetrance using molecular probes (figure 3.6a). The lod score for this family was -0.03 at $\Theta=0$ and 90% penetrance (table 3.5) and 1.78 at $\Theta=0$ with a penetrance of 80%. Assuming 90% penetrance the maximum lod score value was 0.96, at a value of $\Theta=0.2$ (table 3.5). This value is misleading since the larger number of unaffected gene carriers are scored as recombinants which is not true.

In RBF04 (figure 3.6b) three affected siblings were born to unaffected parents. The genotype of all affected individuals was the same (1.9/1.65 kb RS2.0). However, III.2 inherited the 1.9 kb allele from her unaffected mother II.5, and the 1.65 kb allele from her affected father II.4, indicating that the mutation cosegregates with the 1.65 kb allele. Individual II.4 died as a result of developing a second cancer, a small cell lung carcinoma. The parents I.1 and I.2 were unavailable for analysis. It was not possible, therefore, to determine whether one of the parents was a true case of incomplete penetrance or had regressed tumours, or whether one or the other was a gonadal mosaic for the mutation. The maximum lod score for this family, assuming a penetrance of 90%, was 0.058 at $\Theta=0.1$ (table 3.5) which, again, is probably misleading because of the unusual nature of the inheritance.
Figure 3.6 Segregation of RS2.0 alleles in families RBF12 and RBF04.

(a) In family RBF12, the Rb phenotype segregates with the 1.75 kb allele. In Family Unit (FU) 1, the affected subject (III.2) is homozygous so it is not possible to determine whether her children (IV.1, IV.2) are mutation carriers or not. In FU 2, both children are unaffected but III.3 (arrow) has received the 1.75 kb allele from his mother and therefore must carry the mutation. In FU 3 both affected members have only retinal scarring (indicated by the hatched symbols) as a consequence of the Rb gene mutation and II.7 was deceased and unavailable for analysis.

(b) In family RBF04, the Rb phenotype segregates with the 1.65 kb allele. The unaffected parents (I.1, I.2) of three affected children (II.2, II.3, II.4) were unavailable for analysis.
Table 3.5 Lod scores at different values of $\Theta$ at 90% penetrance for the two families that showed incomplete penetrance.

<table>
<thead>
<tr>
<th>FAMILY</th>
<th>Lod Score at $\Theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0</td>
</tr>
<tr>
<td>RBF04</td>
<td>0.048</td>
</tr>
<tr>
<td>RBF12</td>
<td>-0.030</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0.018</td>
</tr>
</tbody>
</table>

Despite these few exceptional families, the cumulative lod score in this study from 21 families was 5.947 at $\Theta=0$ assuming 90% penetrance (table 3.4). When data from this series (Onadim et al 1990) is combined with that from Wiggs et al (1988), Scheffer et al (1989) and Goddard et al (1990) the cumulative lod score is 27.091 at $\Theta=0$, assuming a penetrance of 90%. No recombinations have been observed between the Rb phenotype and intragenic polymorphisms to date.

3.1.4 Application of PCR analysis to paraffin-embedded tissue sections

In the linkage analysis presented above key family members were, on occasion, unavailable for analysis because they had died although in many cases archival material from them is available in histopathology departments. The need, in one particular family, to analyse the deceased patient to establish phase led to experiments to try to adapt the PCR analysis to this archival material. Because the PCR process is relatively tolerant to the quality of the DNA, it is potentially possible to analyse DNA from a variety of sources including formalin fixed paraffin-embedded tissue. Ordinarily the processing of this tissue crosslinks the DNA making it unsuitable for Southern blotting. This procedure was used to analyse one Rb family (RBF58) where the key member has died some time previously (Onadim and Cowell 1991).

The pedigree of family RBF58 (RB-29 in Onadim and Cowell 1991) is given in figure 3.7a. The transmitting parent (I.2), who developed Rb at the age of

132
18 months, had had one eye enucleated. There was no previous family history of disease prior to this on either the paternal or maternal side. At the age of 2 years, the first born child (II.1), was treated with radiation for bilateral Rb. The tumour regressed but six months later a rhabdomyosarcoma arose within the radiation field in the right cheek and eventually caused the patient's death (Brookes et al 1990). Although a rare sub-type, this second malignancy was part of a group of tumours frequently seen in Rb mutant gene carriers (Abramson et al 1984, Draper et al 1986). The second child (II.2) was born in 1987 and the family was referred to us for genetic screening.

At the time of referral, II.2 was 3 years old and, although beyond the peak age for hereditary tumours (14 months), could potentially still develop a tumour or be an unaffected gene carrier. To establish linkage phase, unequivocally, it was essential to analyse DNA from the dead child (II.1). No blood samples had been saved and no necropsy was performed (Brookes et al 1990) but, during the treatment of the rhabdomyosarcoma, a tooth was removed. Attached to its base was a small piece of tumour which had been fixed in formalin and embedded in wax. Courtesy of the Histopathology Department at Alder Hey Children's Hospital, it was possible to obtain tissue sections from this tumour material for analysis. DNA was extracted from 5 \( \mu \text{m} \) thick tissue sections as described in Section 2.6.4 and its quality is shown in figure 3.7b.

DNA from the other family members, together with that from the fixed tumour section was amplified for the M1.8 (Bam HI) and PR0.6 (Xba I) polymorphisms (figure 3.7c-d, top) using the primers shown in table 3.2. The PCR products were then digested with the appropriate restriction enzyme (figure 3.7c-d, bottom) to study the polymorphism. For the M1.8 polymorphism, the father (I.1) is homozygous for the lower allele, the mother (I.2) is heterozygous and the (deceased) affected daughter (II.1) is homozygous for the lower allele (figure 3.7c). Since the, as yet, unaffected son (II.2) is also heterozygous at this locus it would be predicted that the mutant gene is segregating with the lower allele (140/60 bp). For the PR0.6 polymorphism, the mother (I.2), the father (I.1) and the surviving child (II.2) are all heterozygous (figure 3.7d). Since II.1 was homozygous for the lower (630/315 bp) allele, the Rb mutation must segregate with this allele.
Figure 3.7  Linkage analysis of family RBF58 using PCR and DNA extracted from formalin fixed paraffin-embedded tissue sections.

(a) Pedigree of family RBF58.

(b) Size range of DNA extracted from the formalin fixed paraffin-embedded tissue. One-tenth of the extracted DNA (lane 1) was run on a 0.6% agarose gel and stained with ethidium bromide. A smear is seen between 200 bp and 3 kb. M= marker lane containing 2 μg of a 1 Kb DNA ladder (Gibco-BRL).

(c) PCR amplification of the polymorphic Bam HI site from DNA from RBF58 family members (top). In all cases a dominant band of the expected size, approximately 200 bp, is seen as well as fainter, non-specific bands in some cases. The amplified DNA from each individual was digested with Bam HI (bottom). The marker lane (M) contains the Gibco-BRL 1 Kb DNA ladder. I.1 is homozygous for the lower allele but I.2 is heterozygous. Although a weak residual upper band is seen in this example, in the DNA extracted from the tissue sections representing the affected dead child (II.1), the dominant bands are those for the lower allele indicating that II.1 is homozygous for the lower allele. The unaffected child (II.2) is heterozygous. The mutant Rb allele is therefore segregating with the lower allele from the mother (I.2).

(d) PCR amplification of the polymorphic Xba I site from family RBF58 (top). The DNA extracted from paraffin sections (II.1) gives a reduced PCR product yield compared to other DNAs. The parents, I.1 and I.2 are both heterozygous as is the second child, II.2 (bottom). II.1, the dead affected child, is homozygous for the lower allele. The mutant allele, therefore, is segregating with the lower allele from the mother. M= marker lane, 1 Kb DNA ladder (Gibco-BRL).
Figure 3.7

(a) RBF58

(b) 3054 - 2036 - 1636 - 220

(c) M 1 200

(d) M 1 945 - 630 - 315 - 135
The DNA extracted from the tissue sections was degraded, in the size range of 200 to 3000 bp (figure 3.7b). Owing to the cross linking of the DNA by the formalin during the fixation process, the smaller sizes are overrepresented. This was not a problem when amplifying the 200 bp sequence needed to identify the M1.8 polymorphism. For the longer 945 bp fragment carrying the polymorphic PR0.6 site, however, yields of the PCR product were somewhat smaller (figure 3.7d, top, lane II.1). This is presumably due to the fact that, because of the highly fragmented nature of the DNA isolated from the formalin fixed tissue, there are fewer intact template DNA molecules of this size. Amplified DNA was digested directly with restriction enzymes but was sometimes incomplete as shown in figure 3.7c (bottom), lane II.1. The same finding has been noted in several cases which were known to be homozygous for the lower allele following conventional Southern blotting and hybridisation. The relative intensity of the undigested band in these cases, however, was significantly weaker than in 'true' heterozygotes (figure 3.7c, bottom). In fact, when this procedure was repeated using different tissue sections it was clear when the upper band was absent. It is probable, therefore, that the degree of enzyme digestion reflects the purity of the extracted DNA.

Because tumour tissue was used as representative of the dead child it is possible that it may not be truly representative of normal tissue from the patient. Loss of heterozygosity for chromosome 13 markers is frequently reported in Rb tumours, osteosarcomas, and soft tissue sarcomas (Cavenee et al 1983, Dryja et al 1986a, Friend et al 1987, Zhu et al 1989). However, consistent with the theory that loss of heterozygosity 'exposes' recessive alleles in these childhood tumours, it is the mutant allele which is retained (Cavenee et al 1985, Dunn et al 1988, Yandell et al 1989). Thus, even if loss of heterozygosity had occurred it would be expected that, in this case, the mutant allele would be retained in the rhabdomyosarcoma tissue. Since both parents are heterozygous for the PR0.6 polymorphism, their daughter, II.1, could have been constitutionally heterozygous, inheriting the upper allele from the father. Whether II.1 is homozygous for the lower allele or it has been reduced to homozygosity in this case in the tumour, the result is still the same; the tumour predisposition gene cosegregates with the lower allele of the mother. Similarly, it would be predicted that the lower allele cosegregates with the disease using the M1.8 polymorphism. Here the loss of heterozygosity is not relevant to the argument since the father is
homozygous and must contribute the lower allele to his children. Since the lower allele from the mother is associated with the mutant gene, II.1 could not have been constitutionally heterozygous.

For future screening purposes, this family is fully informative for the M1.8 polymorphism. Using the PR0.6 polymorphism, however, since both parents are heterozygous, carrier status can only be confirmed or otherwise in 50% of future offspring. The family is also informative for RS2.0 VNTR but not informative for RB1.20 (see next section and Appendix IV).

3.1.5 Pre- and peri-natal screening using intragenic RB1 polymorphisms

Having established that the vast majority of families are informative for the intragenic polymorphisms and that recombination is, empirically, a very rare event, it became possible to offer pre- and peri-natal screening to Rb families. During the 3.5-year period covered by this project, presymptomatic prediction of mutant gene carrier status was carried out for 11 individuals from 8 different Rb families. The intragenic RB1 polymorphisms used for this purpose were the M1.8 and PR0.6 polymorphisms and the RS2.0 and RB1.20 VNTRs (Sections 3.1.1, 3.1.2). The individual characteristics of each of the families are discussed below.

In family RBF06 (figure 3.8a) post-natal screening of II.2, after 16 months and pre-natal screening of II.3 showed that neither had inherited the mutant RB1 gene, which co-segregated with the 1.75 kb, RS2.0 allele from the affected father, I.2. This result was confirmed using the M1.8 polymorphism. Both II.2 and II.3 were homozygous for the lower allele (2.3+2.2 kb when detected by M1.8 probe) whereas the Rb predisposition clearly segregates with the upper (4.5 kb with M1.8) allele of the father in this family (figure 3.8a). Using the RB1.20 VNTR, the Rb predisposition segregates with allele 5 (figure 3.8a). At the time of screening with the RB1.20 polymorphism, no DNA was available from II.2, but II.3 was shown not to have inherited allele 5 (figure 3.8a).
Figure 3.8 Presymptomatic screening in families (a) RBF06, (b) RBF13 in which none of the probands (arrows) inherited the Rb predisposition gene. For each family, the pedigree together with the autoradiographs and/or agarose gels showing the segregation of alleles with informative polymorphisms are presented. In the family pedigrees, for each individual whose DNA was available for analysis, the size (in kb) or the allele number for each informative polymorphism is shown. For the M1.8 and PRO.6 polymorphisms, allele sizes, as determined by Southern blotting only, are given in the family pedigree to avoid confusion. For simplicity, in the case of M1.8 polymorphism, the alleles of heterozygous individuals are given as 4.5/2.3 (instead of 4.5/2.3+2.2); the alleles of individuals homozygous for the lower allele are given as 2.3/2.3 (instead of 2.3+2.2/2.3+2.2).

In (a) the mutant RBI gene in RBF06 segregates with the 1.75 kb, RS2.0 allele, the upper M1.8 allele and allele 5 from RB1.20. Neither II.2 nor II.3 inherited the 1.75 kb RS2.0 or the upper M1.8 allele (DNA from II.2 was not included in the M1.8 agarose gel shown but was analysed previously using Southern blotting). DNA from II.2 was not available for analysis with RB1.20 but II.3 has clearly not inherited allele 5. M= marker lane containing 1 Kb DNA ladder (Gibco-BRL) (Section 2.11, Appendix III).

In (b) the segregation of RS2.0 and RB1.20 alleles in RBF13 is shown. Neither child (II.2, II.3) has inherited either the 1.95 kb, RS2.0 allele or allele 1 for RB1.20 which co-segregate with the Rb predisposition.
Figure 3.8

(a) RBF06

```
I
2.00/1.90
4.5/2.3
2/3

II
2.00/1.75
4.5/4.5
2/5
```

(b) RBF13

```
I
2.00/1.90
4.5/2.3
2/3

II
1.95/1.90
1/2
```

RS2.0

```
2.00 -
2.00 -
I.3
II.3
2.00 -
II.2
II.1
I.1
I.2
I.3
II.3
II.1
I.1
```

RB1.20

```
2 -
3 -
5 -
6 -
1 -
2 -
3 -
4 -
```

M1.8

```
M
II.1
I.1
I.2
I.3
II.3
M
```

```
200bp upper
140bp lower
60bp
```

RS2.0

```
2.00 -
2.00 -
I.3
II.3
2.00 -
II.2
II.1
I.1
I.2
I.3
II.3
II.1
I.1
```

RB1.20

```
2 -
3 -
5 -
6 -
1 -
2 -
3 -
4 -
```
Family RBF13 (figure 3.8b) had not chosen to take advantage of the pre-natal screening option but was informative for RS2.0 and RB1.20. II.2 was screened post-natally when the child was 6 months old (Onadim et al 1990). Subsequently, a second post-natal screen was performed for II.3 after 0.5 months. It was shown that neither child had inherited the allele associated with the Rb predisposition (figure 3.8b).

The segregation of the Rb phenotype in family RBF29 (figure 3.9a) was unusual and is discussed in detail in Section 3.1.6. Informative for RS2.0 and RB1.20, in March 1991 a CV sample was analysed from the fetus (III.5), whose father (II.4) is an unaffected gene carrier, and the fetus was found not to have inherited the predisposing allele (figure 3.9a).

In family RBF32, whose member I.1 is probably an unaffected gene carrier (see Section 3.1.6), the Rb mutation is segregating with the 1.95 kb allele identified by RS2.0 and the lower (6.5 kb) allele of PR0.6. III.2, who was already 10 months old, was excluded from carrying the mutant allele using both of these polymorphisms (figure 3.9b). In August 1991 a second, post-natal screen was carried out for III.3 in this family. This child was found to be heterozygous for the PR0.6 polymorphism which means that, since both parents were also heterozygous at this locus, a decision could not be made using this polymorphism. Using RS2.0, however, III.3 was found not to have inherited the 1.95 kb allele and therefore was predicted to be unaffected (figure 3.9b).

Although family RBF33 (figure 3.10a) were uninformative for three of the polymorphisms used in the screening programme (RS2.0, PR0.6 and M1.8), they were informative using the RB1.20 VNTR. Using DNA from a cord blood sample, III.1 was shown to have inherited the mutant RB1 gene segregating with allele 3 (figure 3.10a) from the affected father. Two weeks later, ophthalmological examination of III.1 identified tumours in both eyes confirming this prediction.
Figure 3.9  Presymptomatic screening in families (a) RBF29, (b) RBF32. For each family, the pedigree together with the autoradiographs and/or agarose gels showing the segregation of alleles with informative polymorphisms are shown. The probands are indicated by an arrow.

In (a) the 1.90 kb RS2.0 allele and allele 5 for RB1.20 segregates with the Rb predisposition. Here III.5 is shown to have inherited the lower 1.85 kb (normal) allele from the father who is an unaffected gene carrier. In the RB1.20 autoradiograph the bands exist as doublets at each position which is probably due to slipped pairing during amplification. However, the inheritance of each allele is still followed easily and III.5 is shown not to have inherited allele 5.

In (b) the mutant RB1 gene co-segregates with 1.95 kb RS2.0 allele and the lower PRO.6 allele in RBF32. III.2 is shown not to have inherited either of these alleles. III.3 is not informative for PRO.6 as both he and his parents are heterozygous at this locus. With RS2.0, however, III.3 is shown not to have inherited the 1.95 kb allele. M = 1 Kb DNA ladder (Gibco-BRL).
Figure 3.9

(a) RBF29

II

III

II

III

(b) RBF32

II

III

PR0.6

upper

lower

M II.1 II.2 II.3 III.1 III.2 III.3

1.95 - 1.75 - 2.00

- 1.90

1.95

1.75 -
Figure 3.10  Presymptomatic screening in families (a) RBF33 and (b) RBF41 in which the probands (arrows) were found to have inherited the Rb predisposition gene.

(a) Linkage analysis in family RBF33 using the RB1.20 polymorphism. III.1 who was screened post-natally, was shown to inherit the mutant Rb predisposition gene segregating with allele 3 from her father.

(b) Linkage analysis in family RBF41 using the RS2.0 and RB1.20 polymorphisms. IV.5 was screened pre-natally and was shown to inherit the 1.90 kb, RS2.0 and allele 3 from RB1.20 which segregates with the Rb predisposition in this family.
Family RBF34 (figure 3.11) was referred to the unit for assessment of the possibility of future pre-natal screening having already had an affected child. Unusually, this family was informative at all the polymorphic sites analysed. Using the RS2.0 probe, the affected father was apparently homozygous for the 1.75 kb allele (figure 3.11b) which was unusual since his father (I.1) did not carry this allele. His mother was not available for analysis. At first, non-paternity was considered as an explanation, until patient III.1 was also shown apparently not to have inherited an allele from his father (figure 3.11b). The same pattern of inheritance was shown using the RB1.20 polymorphism (figure 3.11a). II.1 was heterozygous, however, for the M1.8 (figure 3.11c) and PR0.6 (figure 3.11d) polymorphisms. Thus, it became clear that the predisposing mutation, which originated in I.1, was a deletion which included that part of intron 17 distal to the Xba I site (figure 3.1), and extended at least to the point in intron 20 carrying the RB1.20 polymorphic site. In August 1991, pre-natal screening was carried out for III.2, after 11 weeks of pregnancy, and the fetus was shown to have inherited the normal allele at all loci. The prediction, therefore, is that the child will not be affected.

In family RBF41 tumour predisposition segregates with the 1.90 kb RS2.0 allele and allele 3 for RB1.20. In November 1991, pre-natal screening was carried out for IV.5 using a CV sample obtained after 11 weeks of pregnancy. The fetus was shown to have inherited the mutant Rb predisposition allele and the prediction, therefore, was that it would be affected. After birth, ophthalmological examination of IV.5 identified a unilateral tumour in the left eye confirming this prediction.

In family RBF58 (figure 3.12), which was described in Section 3.1.4, using paraffin embedded tumour material from II.1, it was established that the Rb predisposition segregates with the lower M1.8 and PR0.6 alleles and II.2 was excluded from being a mutant gene carrier. This family was also informative with RS2.0 and the unaffected child, II.2, carried the 2.00 kb RS2.0 allele from the affected mother, I.2 (figure 3.12). In June 1992, pre-natal screening was carried out for II.3 using a CV sample obtained after 11 weeks of pregnancy. The fetus was shown to have inherited the lower M1.8 and PR0.6 alleles and 1.90 kb RS2.0 allele from the affected mother (figure 3.12). The prediction, therefore, was that II.3 would be affected. The family decided to terminate the pregnancy.
Figure 3.11  Linkage analysis in family RBF34. The family pedigree is shown in the centre. Note that II.1 and III.1 have only a single allele for the 1B1.20 VNTR (a) and the RS2.0 (b) polymorphism from their unaffected parents. The M1.8 (c) and the PR0.6 (d) polymorphism analysis confirm that the fetus, II.2, has not inherited the mutant allele. M = marker lane containing 1 Kb DNA ladder (Gibco-BRL).
Figure 3.12 Pre-natal screening in family RBF58 using the RS2.0, M1.8 and PR0.6 polymorphisms. The family pedigree is shown on top. The fetus, II.3, was shown to inherit the mutant Rb predisposition gene segregating with 1.90 kb RS2.0 allele and lower M1.8 and PR0.6 allele from the mother, I.2. M = 1 Kb DNA ladder (Gibco-BRL).
In summary, therefore, five pre-natal and six post-natal tests were carried out in 8 different families. A summary of the results obtained is given in table 3.6. This series of patients represents the first cohort who have been followed ophthalmologically as a result of genetic screening carried out using intragenic RB1 polymorphisms. Because there was no prior experience with the intragenic probes the ophthalmic surgeons felt it inappropriate at the time to discontinue the routine screening, conventionally, under anaesthesia, of the retinae of both eyes every 3 months to the age of 2, 4 months to the age of 3 and 6 months to the age of 5. With genetic screening, of the eleven probands, eight were shown not to have inherited the mutant allele (table 3.6). The majority of familial cases present before 2 years and five patients were followed at least this period. However, since in fact, the mean age of onset in this group is 14 months, seven have reached this age disease free (table 3.6). Three probands were shown to have inherited the mutant allele. The presence of bilateral tumours was confirmed in one of these (III.1, RBF33) and a unilateral tumour in one other (IV.5, RBF41). In family RBF58, the pregnancy was terminated.
Table 3.6  The results of five pre-natal and six post-natal screenings carried out over a period of 3.5 years (Jan 1989 - Jun 1992) involving eight different families.

<table>
<thead>
<tr>
<th>Family</th>
<th>Pro-band</th>
<th>Age at testing (m)</th>
<th>Age at 12 92 (m)</th>
<th>Informative Polymorphisms</th>
<th>Predis- position</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF06</td>
<td>II.2</td>
<td>16</td>
<td>68</td>
<td>RS2.0, M1.8</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF06</td>
<td>II.3</td>
<td>Prenatal</td>
<td>27</td>
<td>RS2.0, M1.8, RB1.20</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF13</td>
<td>II.2</td>
<td>6</td>
<td>53</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF13</td>
<td>II.3</td>
<td>0.5</td>
<td>32</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF29</td>
<td>III.5</td>
<td>Prenatal</td>
<td>14</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF32</td>
<td>III.2</td>
<td>10</td>
<td>47</td>
<td>RS2.0, PR0.6</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF32</td>
<td>III.3</td>
<td>4</td>
<td>22</td>
<td>RS2.0</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF33</td>
<td>III.1</td>
<td>0.25</td>
<td>12</td>
<td>RB1.20</td>
<td>Y</td>
<td>Affected</td>
</tr>
<tr>
<td>RBF34</td>
<td>III.2</td>
<td>Prenatal</td>
<td>11</td>
<td>RS2.0, M1.8, PRO.6, RB1.20</td>
<td>N</td>
<td>Unaffected</td>
</tr>
<tr>
<td>RBF41</td>
<td>IV.5</td>
<td>Prenatal</td>
<td>08</td>
<td>RS2.0, RB1.20</td>
<td>Y</td>
<td>Affected</td>
</tr>
<tr>
<td>RBF58</td>
<td>II.3</td>
<td>Prenatal</td>
<td>-</td>
<td>RS2.0, M1.8, PR0.6</td>
<td>Y</td>
<td>Affected</td>
</tr>
</tbody>
</table>
3.1.6 Summary of phenotypic and genotypic data of Rb families in this series

As a result of this extensive linkage analysis it is possible to draw some general conclusions about the Rb population in this series. A summary of the observed phenotypes of gene carriers in the 57 Rb families studied (see Appendix III) is given in table 3.7. Of the 172 affected individuals 75% (129/172) were bilaterally affected. RFLP analysis detected 10 unaffected gene carriers, of which 8 could not have been predicted as such using pedigree analysis alone. Taken together the non-penetrance value in this series is 5.5%. Spontaneous regression of tumours was noted to have occurred bilaterally in 2.3% (4/172) and unilaterally in 3.5% (6/172). The percentage of cases of regressed tumour among 182 mutant gene carriers was 5.5%. Of the 60 or more apparently unaffected individuals born to Rb families 32 were excluded from carrying Rb predisposition. DNA samples from the remainder were not available. Among 172 affected individuals, 11 (9 bilateral and 2 unilateral) were reported to have developed second cancers (J. Kingston personal communication). In addition, one apparently unaffected transmitting parent also died of cancer (cancer of the colon). The cancers developed were osteosarcoma (1), breast (1), rhabdomyosarcoma (1), oat cell (1), ovary (1), oesophagus (1) and small cell lung carcinoma (4). The type of cancer in one of the individuals was not reported. The percentage of second cancers among gene carriers in this series is, therefore, 6.4%. However, the cause of death is not determined in all the deceased affected individuals. There may have been additional cancer deaths that were not reported to us.

Table 3.7 The Rb phenotype of mutant RB1 gene carriers* in 57 Rb families.

<table>
<thead>
<tr>
<th>Gene Carrier</th>
<th>Number (%)</th>
<th>Spontaneous regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No (%)</td>
</tr>
<tr>
<td>Bilateral</td>
<td>129 (70.9)</td>
<td>4 (2.3)</td>
</tr>
<tr>
<td>Unilateral</td>
<td>43 (23.6)</td>
<td>6 (3.5)</td>
</tr>
<tr>
<td>Total affected</td>
<td>172 (94.5)</td>
<td>10 (5.8)</td>
</tr>
<tr>
<td>Unaffected</td>
<td>10 (5.5)</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>182 (100)</td>
<td>10 (5.5)</td>
</tr>
</tbody>
</table>

* II.3 from RBF58 is not included in this analysis.
3.1.7 Expression of the Rb phenotype

In the majority of the families studied the Rb phenotype was typical of that expected for autosomal dominant inheritance. There were some families, however, which did not conform to this pattern. This group included families where more than one affected child was born to unaffected parents such as RBF04 (Onadim et al 1990), described in Section 3.1.3. There were also families where the majority of gene carriers only developed unilateral disease and/or spontaneously regressed tumours and in some cases no evidence of disease at all (incomplete penetrance). These are referred to as "low penetrance" or "mild phenotype" families.

One particularly interesting family with an unusual pedigree is RBF29 (figure 3.13a) where, with the exception of bilaterally affected, non-identical twins (III3, III4), the expression of the Rb phenotype is extremely "mild". The parents (II.3, II.4) of the affected twins are first cousins and each has a unilaterally affected sister. These observations are consistent with one or both of their parents (I.1, I.3) being gene carriers. In 1984, prior to the birth of twins, family counselling was sought at a local hospital by II.3. The disease was diagnosed as an autosomal recessive disorder by virtue of the fact that both parents had an affected sib and a 1:32 risk of having affected children was given. The retinæ of the parents, I.1 and I.3 were not examined at this time. In February 1987 non-identical twins were born, and in June of that year a diagnosis was made of bilateral Rb in both children at the age of four months following their referral to St Bartholomew's Hospital for treatment. III.3 had a juxtapapillary Rb in the right eye and a posterior pole tumour in the left eye. III.4 had a tumour nasal to the optic disc in the left eye and later developed a tumour temporal to the posterior pole in the right eye. No evidence of extraocular spread was discovered in either.

In July 1987 both grandparents attended the ophthalmology clinic at St Bartholomew's Hospital, where I.1 was found to have a large, grey translucent lesion at the posterior pole of the left eye (figure 3.13a). I.3 was found to have a small posterior lesion of the right eye which was translucent with pigment epithelial disturbance (figure 3.13b), and a small lesion nasal to the optic disc in the left eye. These lesions were typical of spontaneously regressed Rb (retinoma). Close examination of the retinæ of II.3 and II.4 showed no evidence of retinal scarring.
Figure 3.13 Segregation of (a) RS2.0 alleles in RBF29 and (b) R0.6 in RBF18 where most of the individuals have "mild" expression of the disease. The unaffected gene carriers are indicated by an arrow.

(a) In family RBF29 the Rb phenotype co-segregates with the 1.9 kb RS2.0 and number 5 RB1.20 allele derived from I.1 and I.3. The RS2.0 autoradiograph is shown below the pedigree. The RB1.20 autoradiograph is given in figure 3.9a of Section 3.1.5. DNA from III.5 is not included in the RS2.0 autoradiograph but shown separately in figure 3.9a. DNA from I.3, I.4, III.2 and III.6 were not available at the time either but later III.2 was shown to carry the Rb predisposition gene and III.6 was excluded from being a mutant gene carrier (Section 3.3.2).

(b) Family RBF18 where I.1, II.4, III.3 and IV.1 have regressed tumours. The Rb phenotype segregates with 4.95 kb allele of R0.6 polymorphism the autoradiograph of which is shown below.
Figure 3.13

(a)

RBF29

I

II

III

RS2.0

(b)

RBF18

I

II

III

IV

R0.6
Figure 3.14  Spontaneously regressed Rbs (retinomas) discovered in the eyes of the grandparents in family RBF29.  (a) Large spontaneously regressed Rb at the posterior pole of the left eye of I.1. (b) Small regressed Rb at the posterior pole of the right eye of I.3.
Esterase D studies showed all family members to be of genotype 1-1 and chromosomal analysis showed no deletions of the 13q14 locus in any family member. At the time of the initial consultation DNA-based analysis was not generally available.

In 1989, the family was referred to us for genetic screening (Onadim et al 1991). Since both parents, II.3 and II.4 were potentially unaffected mutant gene carriers, there was a possibility that their children could be homozygous for the RB1 gene mutation. The family was first screened using the intragenic probe RS2.0, the details of which are given in figure 3.13a. Analysis of the affected cousins, II.2 and II.5, showed that the common alleles are 1.9 kb which must have been derived from mutant gene carriers I.1 and I.3. Both parents of the twins also carry the 1.9 kb allele which II.4 inherited from I.3 but II.3 received the 1.5 kb allele from the father, I.1, and the 1.9 kb from the unrelated parent, I.2. Therefore II.3 is not a mutant gene carrier, thus excluding the possibility that their children could become homozygous for this mutation. The twins inherited the 1.9 kb allele from their father confirming the co-segregation of the Rb phenotype with this allele. Analysis of other family members also identified II.7 and III.1 as mutant gene carriers (figure 3.13a). At the time when this autoradiograph was prepared, DNA from III.2 was not available but it was subsequently shown that, she, too, was a mutant gene carrier having inherited the 1.9 kb from her affected mother (II.2). In March 1991, a CV sample for III.5 was analysed and was found not to carry the Rb predisposition allele as described in Section 3.1.5. At the same time, the results were confirmed using the RB1.20 VNTR (Section 3.1.5, figure 3.9a).

Another family, RBF18 (figure 3.13b), showed individuals with regressed tumours in four generations. This family, first reported in the literature by Hine (Hine 1937), was unusual in that there were also individuals whose phenotypes cannot be described as mild. However, the presence of regressed tumours in the majority of cases could arguably put this family in the "mild" phenotype category. There are eight affected individuals (figure 3.13b), three of whom (II.2, III.3, IV.1) had unilateral disease. Four generations of males (I.1, II.4, III.3, IV.1) have a mild form of the disease. Although, I.1 had one eye removed for Rb, the tumour in his other eye regressed naturally. II.4 has spontaneously regressed tumours in both eyes and has never been treated for Rb. Likewise III.3 and IV.1 have unilateral regressed tumours.
II.1 and II.3, however, died at the age of 1 year 9 months and 3.5 years respectively, as a result of orbital extension of their Rb (Hine 1937). I.1 and II.2 died as a result of developing second tumours (for I.1 the particular type of cancer was not reported, II.2 died of cancer of the lung) at the age of 61 and 40 respectively. Only DNA from individuals III.2, III.3, III.4 and IV.1 were available for analysis. The family was only informative for the R0.6 polymorphism. The Rb predisposition was linked to the 4.95 kb allele (figure 3.13b).

Regressed tumours are also seen in families RBF02, RBF38 and RBF41 pedigrees of which are given in figure 3.15a, b and c respectively. In these families regressed tumours seem to be a "one-off" phenomenon with the exception of RBF02 (figure 3.15a). In this family the affected father (I.1) has unilateral Rb and the tumours of affected sons II.1 and II.2, unilateral and bilateral respectively, have regressed spontaneously. This family might therefore be considered as a low penetrance family. The Rb phenotype segregates with the 1.60 kb RS2.0 allele from the father (figure 3.15a).

In family RBF38 (figure 3.15b), the unilateral tumour of the affected father (I.1) has regressed spontaneously. His only child, II.1, is, however, bilaterally affected inheriting the 7.8 kb PR0.6 allele from him (figure 3.15b).

There is a great deal of heterogeneity in the Rb phenotype between families. This is illustrated in the analysis of family RBF41 (figure 3.15c) where there are many bilaterally affected individuals. RBF41 has already been described in Section 3.1.5 in the context of pre-natal screening carried out for individual IV.5 who was found to carry the predisposition gene. This family is also one of the families where the first generation parent was apparently unaffected. I.2, himself being unaffected, had three bilaterally affected children (II.2, II.3, II.5) and two apparently unaffected children (II.4, II.6) from two separate marriages. I.2 was, therefore, either an unaffected gene carrier or had regressed tumours which went undetected. Alternatively, I.2 could carry the mutation as a germ-line mosaic. II.5 has spontaneously regressed tumours and stands out as an exception since all the other gene carriers in this family are bilaterally affected with no regressed tumours.
Figure 3.15 Pedigrees of families (a) RBF02, (b) RBF38 and (c) RBF41 showing one or more individuals with regressed tumours (hatched symbols). The number or the sizes of alleles for informative polymorphisms are given for individuals whose DNA were available for analysis. For families RBF02 and RBF 38 autoradiographs are also shown.

(a) In family RBF02, II.1 and II.2 have regressed tumours. The Rb phenotype co-segregates with the 1.60 kb allele of the RS2.0 polymorphism, the autoradiograph of which is shown alongside.

(b) In RBF38, I.1 has a unilateral regressed tumour and passed the 7.8 kb PR0.6 allele to his bilaterally affected daughter, II.1, as shown in the autoradiograph.

(c) The pedigree of RBF41 with an unaffected first generation parent (I.2) and one individual (II.5) with regressed tumours. The sizes (in kb) of the RS2.0 alleles and the numbered alleles for RB1.20 are given for each individual whose DNA was analysed. The relevant autoradiographs are given in figure 3.10b of Section 3.1.5. The Rb phenotype segregates with 1.90 kb RS2.0 allele and allele 3 of RB1.20.
Figure 3.15

(a) RBF02

(b) RBF38

(c) RBF41
In family RBF41 a pre-natal screen was carried out for IV.5, who was found to have inherited the mutant gene (Section 3.1.5). A unilateral tumour was discovered in his left eye soon after birth and he is undergoing radiotherapy to both eyes as there is also a possibility of tumour(s) developing in his right eye due to a strong family history (JE Kingston personal communication). Individual IV.4 (the twin of IV.3) died at the age of 8 from his Rb. Apart from I.2, who died of colon cancer, two other individuals developed second cancers. II.2 died from cancer of the ovary aged 33 and II.3 died from cancer of the lung aged 49.

RBF32 and RBF56 (figure 3.16) are two families with low penetrance phenotypes which contrast with RBF41. One part of the pedigree of RBF32 was shown in figure 3.9b of Section 3.1.5, where two post-natal screenings carried out for this family were described. In figure 3.16a the full pedigree with the RS2.0 and RB1.20 autoradiographs is shown. In this low penetrance family, four individuals in the third generation (III.1, III.7, III.8, III.9) have unilateral Rb. The transmitting parents of these four individuals (II.1, II.2, II.3), who are brothers, and their parents (I.1, I.2) are all apparently unaffected. The parents, except for II.3, were unavailable for analysis. It is extremely unlikely that the four tumours seen in the third generation could have arisen as a result of independent new mutations. The first generation parents, II or I.2, could have been either unaffected gene carriers or have had regressed tumours or have been germ-line mosaics for the Rb mutation. Their children, II.1, II.2 and II.3, are probably unaffected gene carriers. It is not known whether II.1, II.2 and II.3 were examined for retinal scars (regressed tumours) which would have placed them in the affected category rather than the incomplete penetrance one. In many Rb families, however, it is unlikely that the eyes of older generation parents have been examined for regressed tumours since it is only recently that ophthalmologists have become aware of this phenomenon. In the third generation, III.1 was not available for analysis but the common allele between III.7 of family unit (FU) 1 and II.3, III.8 and III.9 of FU2 is identified to be the 1.95 kb RS2.0 and allele 4 of RB1.20 (figure 3.16a). In FU2, the affected sisters III.8 and III.9 share the 1.95 kb allele. IV.3 who was born in 1984, and is unaffected, did not inherit the 1.95 kb allele and neither did IV.4 (born 1989) and IV.5 (born 1991) who are also unaffected (Section 3.1.5). RB1.20 is informative in FU1 but not in FU2 as III.8 and III.9 are homozygous for allele 4. In FU1, unaffected IV.1 who was born in 1987, inherited the 1.95 kb RS2.0 and the
RB1.20 allele 4 from the affected mother III.7 which makes him an unaffected gene carrier. IV.2 (born 1988), on the other hand, did not inherit the mutant Rb allele.

Family RBF56 (figure 3.16b) illustrates a few unique features in Rb pedigrees and highlights some of the problems experienced in counselling. The parent I.2 was apparently unaffected with two unilaterally affected children (II.2, II.4) from separate marriages. The first and second marriages produced one and four apparently unaffected children respectively. Individual II.3 has been confirmed to be unaffected but no information was given about retinal examinations of individuals I.2 and II.5-8. The first generation parent I.2 could have been a germ-line mosaic, unaffected gene carrier or had regressed retinal tumours. In this mild phenotype family, all affected individuals are unilaterals (II.2, II.4, III.2) and there is one possible incomplete penetrance case (IV.1). The genetic make up of II.2, who is deceased, is not known. In the third generation, the affected individual, III.2, and his unaffected sister and brother (III.3, III.4) inherited the same allele with all three polymorphisms (7.8 kb PR0.6, 1.95 kb RS2.0, allele 2 RB1.20) from their affected father, II.2 (figure 3.16b). However, as the genetic make up of II.2 is not known, the question of whether they actually inherited the mutant allele could not be solved. It is possible, however, that II.2 was homozygous for these alleles and passed the mutated copy to III.2 and the normal copy to III.3 and III.4. The 1.95 kb allele of the RS2.0 polymorphism is not one of the rarest alleles in the population (table 3.1, Section 3.1.1) making this possibility likely. In addition, the unaffected sister of II.2 (II.3) also carries the 1.95 kb allele and is, in fact, homozygous for allele 2 using RB1.20. DNA from II.4, II.5-8, III.5 and III.6 were not available for analysis. The affected III.2 passed the allele he inherited from II.2 to his son IV.1 who is unaffected. IV.1 is therefore an unaffected gene carrier.

There were three other families, RBF26, RBF55 and RBF01, in this series with incomplete penetrance. The pedigrees and the examples of linkage analysis in these families are given in Figures 3.17 a, b and 3.18 respectively.
Figure 3.16 Two families, (a) RBF32 and (b) RBF56 with members who are only unilaterally affected or unaffected gene carriers (indicated with an arrow). The autoradiographs and sizes (in kb) or numbers of alleles of informative polymorphisms are given for each family.

(a) In family RBF32, the Rb phenotype co-segregates with 1.95 kb, RS2.0 allele and allele 4 of RB1.20, autoradiographs of which are shown below the pedigree. DNA from IV.4 and IV.5, which were not included in the RS2.0 autoradiograph below, are shown in the one given in figure 3.9b (lanes marked III.2 and III.3). DNA from II.1, II.2 and III.1-5 were not available for analysis. The RB1.20 VNTR is informative in family unit (FU) 1 but not in FU2. The first and second generation parents were apparently unaffected.

(b) Family RBF56 has an apparently unaffected transmitting parent (I.2) in the first generation and a possible case of incomplete penetrance (IV.3). Affected individual III.2 and unaffected individuals III.3 and III.4, inherited the same size alleles from their affected father II.2 (7.8 kb PR0.6, 1.95 kb RS2.0, allele 2 RB1.20). It is possible that II.2 was homozygous for all of these alleles but only passed the mutant allele to III.2. The fact that the unaffected sister, II.3, also carries the 1.95 kb, RS2.0 allele and is homozygous for the RB1.20 2-allele supports this hypothesis.

DNA samples from II.4, II.5-8, and III.5-6 were not available. Since the predisposition associated alleles were transmitted from III.2 to his unaffected son, IV.1 must be an unaffected gene carrier. M = 1 Kb DNA ladder (Gibco-BRL).
Figure 3.16

(a)

RBF32

I

II

III

IV

FU1

FU2

RS2.0

RB1.20

162
Figure 3.16

(b)

RBF56

I

II

III

IV

PRO.6

upper

lower

RS2.0

RB1.20
Figure 3.17  Family pedigrees and the segregation of alleles with informative polymorphisms in (a) RBF26 and (b) RBF55. Unaffected gene carriers are indicated with arrows. The allele sizes for informative polymorphisms are given in kb for each family member analysed. For the M1.8 polymorphism, allele sizes in the pedigree are given as 2.3 kb, instead of 2.3 +2.2 kb, for simplicity.

(a) In family RBF26, the Rb phenotype segregates with the 1.95 kb, RS2.0 allele, 7.8 kb PR0.6 allele and 4.5 kb M1.8 allele as shown in the autoradiographs given below. III.5 is an unaffected gene carrier.

(b) In family RBF55, the only informative polymorphism was RB1.20. Both the affected (II.2) and unaffected (II.3) sisters inherited the same allele from their affected father (I.1). DNA from II.1 was not available for analysis.
In RBF26 (figure 3.17a), where the Rb phenotype segregates with the 4.5 kb M1.8, 7.8 kb PR0.6 and 1.95 kb RS2.0 alleles, unaffected III.5 is an obvious incomplete penetrance case as he inherited the same alleles as his affected brother (III.4) from the transmitting parent (II.5) and he has an affected daughter (IV.1). For the couple, III.3 and III.4, future screening would be possible with RS2.0 and PR0.6. M1.8, however, would be informative for homozygous offspring only.

Family RBF55 (3.17b) was informative only with the RB1.20 VNTR. The affected (II.2) and unaffected (II.3) sisters in this family inherited the same allele (2) from the affected father, I.1. DNA from unilaterally affected II.1, who died aged 9.5 years, was not available. As a recombination event between Rb phenotype and intragenic loci has never been observed and the theoretical frequency of such an event in the Rb gene is very low (1:550), II.3 has to be considered an incomplete penetrance case.

Family RBF01 (figure 3.18a), with two potential incomplete penetrance cases, was initially analysed by Dr C Mitchell who carried out pre-natal screening for II.3. Among available intragenic probes at the time the family was only informative for RS2.0. It was found that the affected (II.1) and the unaffected (II.2) daughters both inherited the 1.6 kb allele from their affected mother, I.2 (figure 3.18a). The result of pre-natal screening (in 1988) showed that II.3, who is so far unaffected, also inherited the same allele from the mother. There are two possible explanations; one is that a recombination event between RS2.0 locus and the actual mutation has occurred in II.1, the alternative being that the unaffected daughter (II.2) represents a case of non-penetrance. The possibility that recombination might have occurred was seriously considered in this family which was not the same as others with a mild phenotype since the mother and daughter were affected multifocally and bilaterally. However, the second, and now the third, child must be examples of incomplete penetrance if no recombination has occurred. No recombination was observed in this family between ESD and RB1 (CD Mitchell personal communication) (figure 3.18a). At the time of this analysis the available distal markers were uninformative in this family. To extend this study, the family was reanalysed with distal 9D11, (13q21); 1E8, (13q22); 9A7, (13q32) and proximal 7F12, (13q12-13); 7D2, (13q14) chromosome 13 markers using Southern blotting/hybridisation as well as the more recent intragenic RB1.20 VNTR which lies distal to RS2.0.
Figure 3.18 Analysis of family RBF01 using a variety of chromosome 13 probes. The pedigree is shown in (a) together with the segregation pattern at defined loci (ESD and RS2.0 results provided by C D Mitchell). The autoradiograph alongside shows the analysis using RB1.20. DNA from II.3 was not available at the time of ESD and RB1.20 analyses, but II.1 and II.2 were shown to have inherited the same allele from the affected mother at all loci (see figure 3.19). In (b) Southern blot analysis using proximal (7F12) and distal (9D11, 1E8, 9A7) chromosome 13 probes are shown. The allelic forms are shown on the left of each panel, "1" is the longer and "2" the shorter fragment. The molecular weight of each fragment (in kb) is shown to the right of each panel. Using 7F12 and 9A7, the affected mother I.2 was uninformative. With 9D11 and 1E8 both the parents and the unaffected daughter, II.2, were heterozygous. The question, therefore, whether II.2 inherited the same allele (allele 2) from the mother can not be resolved at these loci unless combination of alleles at flanking loci are taken into consideration (see figure 3.19). It is clear, however, that II.3 inherited the same allele as II.1 from the mother at loci 9D11 and 1E8.
Figure 3.18

(b)
The autoradiographs from these analyses are shown in figure 3.18a, b except for the result using 7D2 for which all members were homozygous. Using ESD, RS2.0 and RB1.20 it was clear that the affected and the unaffected daughters all inherited the same 'haplotype' (in the mother, it is not a proven haplotype as we do not have the grandparents but rather the 'haplotype reconstruction' with the least need in recombination) from their mother. The possibility of a distal recombination event was also excluded using the distal markers 9D11 and 1E8 for which I.2 is heterozygous. The haplotypes for each family member, at each locus, are represented schematically in figure 3.19. The information from loci between ESD and 1E8 exclude the possibility of meiotic recombination between these markers and the actual predisposing mutation in II.1. However, a meiotic recombination event that occurred distal to 1E8 in the paternal chromosomes, and inherited by II.3, was identified (figure 3.19). Combinations of markers, 7F12, 9A7M (Msp I), 9A7H (Hind III), for which I.2 is homozygous and uninformative are included, therefore, as they show the presence of paternal recombination event. At the 9D11 and 1E8 loci, both parents and the unaffected daughter II.2 are heterozygous. The combination of alleles at loci 9D11 to 9A7H, however, also exclude the possibility of a recombination event at these loci (figure 3.19). We were thus able to show that no recombination event occurred around RB1 in II.1, and confirm the suggestion that II.2 and II.3 had inherited the mutant Rb allele from their mother making them cases of incomplete penetrance.

Another problem in analysing Rb phenotypes is the presence, in some families, of unaffected parents who have more than one affected child. Three such families (RBF41, RBF32, RBF56) were described in this section. In these families it was possible to identify the transmitting parent because they had affected children from separate marriages. In two other families, one of which was described in Section 3.1.3 (RBF04), it was not possible to determine the transmitting parent. In family RBF24 (figure 3.20a) as in RBF04, both parents were unaffected. It was not possible, therefore, to identify the transmitting parent because DNA of I.1 was not available. The Rb phenotype apparently co-segregates with the 7.8 kb PR0.6 allele because the affected children, II.2 and II.3, carry this allele, whereas the unaffected child, II.1, being homozygous for 6.5 kb allele, does not. The 7.8 kb allele could have come from either parent either as an unaffected gene carrier or a germ-line mosaic.
Figure 3.19 Schematic representation of haplotype analysis on chromosome 13 in RBF01. I.2 is only informative between ESD and 1E8. The loci for which the father I.1 is heterozygous are included in the representation to illustrate the meiotic recombination event which occurred and also to show the origin of 9D11 and 1E8 allele 2 in individual II.2 who is not informative at 9D11 and 1E8 loci. The chromosome carrying the mutant Rb allele is designated "rb" and the "+" refers to the wild type. The numbers beside each chromosome represent the allelic form of each locus derived from the data presented in figure 3.18 (a) and (b). Each one of RB1.20 alleles of the mother and father are different in size i.e. there are four different alleles (figure 3.18a) at this locus in the family, numbered 1 to 4 and these uniquely identify each chromosome of the mother and father. The names and the order of loci on the chromosome are shown on the left hand side. ESD and RB1.20 analyses were not performed for II.3 whose genetic make up at these loci is, therefore, not known and this is indicated with a "-" sign at the relevant locus. The 'haplotypic' combinations show that, in the chromosome II.3 inherited from his father, a meiotic cross-over of paternal chromosomes has occurred between 1E8 and 9A7 loci (indicated on the "+" chromosome of II.3) but no recombination event is identified in the chromosomes inherited from the mother as shown by combinations at loci ESD to 1E8. After RB1.20 locus, at loci 9D11 and 1E8, II.2 has the same haplotypic combinations as the mother and father on both of her chromosomes, therefore, this part of her chromosome could have come either from mother's or father's chromosomes. The combination of alleles at loci 9D11 to 9A7H, however, show that II.2 inherited the same allele of 9D11 and 1E8 from the affected mother as II.1 and II.3.
In RBF31 (figure 3.20b), the unaffected mother I.1 is the common link in the two families having one affected child from each marriage. Analysis of this family using RB1.20 showed that neither II.1 nor II.3 inherited an allele from their mother I.1 (figure 3.20b). The affected children of II.3 (III.1 and III.2) do not share an allele with their father either. DNA of unaffected individuals II.2 and III.3 were not available for analysis. The same results were obtained with the RS2.0 polymorphism (figure 3.20b) except that, in this case, II.1 and his mother, I.1, both carry a 1.90 kb allele. II.1, however, must be hemizygous for this allele. These results suggested a predisposing deletion extending at least through the RS2.0 and RB1.20 loci. For M1.8, however, both II.1 and II.3 were heterozygous so the 5' end of the gene is not involved in the deletion. Neither does the deletion involve PR0.6 (Xba I) locus as III.1 and III.2 are heterozygous for this polymorphism (figure 3.20b). The deletion, therefore, starts in intron 17, somewhere between the PR0.6 and RS2.0 loci. As with RBF24, I.1 could either be a true incomplete penetrance case carrying the deleted gene herself or a germ-line mosaic.

In summary, therefore, the analysis of the Rb phenotype in different families indicates a great deal of heterogeneity in the expression of the Rb phenotype. There are families, such as RBF29 and RBF32, where the expression of the phenotype is mild. This contrasts sharply with the Rb phenotype exhibited in other families, such as RBF41, where most of the affected individuals are multifocally and bilaterally affected. The appearance of spontaneously regressed tumours (or retinomas) is also very interesting especially when aggregation of such cases occurs in one family (RBF18). Although linkage analysis is able to identify cases of incomplete penetrance and undetected retinomas, a proper explanation of heterogeneity in expression may not be possible unless the causative mutations are identified and compared.
Figure 3.20  Pedigrees from families (a) RBF24 and (b) RBF31. Autoradiographs and/or agarose gels showing the segregation of alleles at informative polymorphic loci are shown. In (a) the autoradiograph shows that the affected patients II.2 and II.3 both have the 7.8 kb PRO.6 allele not present in II.1. In (b) the segregation of alleles at four intragenic loci are shown as autoradiographs for RS2.0 and RB1.20 and as agarose gels after digestion of the PCR products carrying the M1.8 and PRO.6 polymorphic sites. In the pedigree, for the M1.8 and PRO.6 polymorphisms, the allele sizes are given as defined by Southern blotting for simplicity. The genotype of II.1 is not determined for PRO.6 ("--"). With RB1.20, II.1 and II.3 do not have an allele from their mother I.1. Likewise, III.1 and III.2 do not have an RB1.20 allele from their father II.3. With RS2.0, II.1 and I.1 both carry the 1.90 kb allele but II.3 does not have an allele from I.1 and similarly III.1 and III.2 do not have an allele from II.3. DNA samples from II.2 and III.3 were not available for analysis. M= 1 Kb DNA ladder (Gibco-BRL).
3.2 ASSESSMENT OF TECHNIQUES TO DETECT MUTATIONS

Clearly, linkage analysis is only possible when there is a prior family history of Rb. In sporadic cases, the actual predisposing mutation needs to be identified for predictive screening. Characterisation of individual mutations will also provide an in depth understanding of the disease itself and the mutational mechanisms involved. In the past the identification of subtle mutations in the RB1 gene was limited by the non-availability of fast and accurate techniques for characterising them. During the course of this project, four different approaches to mutation screening have been investigated to determine their usefulness in identifying RB1 gene mutations: RNase protection, PCR amplification and sequencing of individual exons, CDI modification of heteroduplexes and SSCP analysis. The results of these investigations are given below.

3.2.1 RNase protection experiments

RNase protection is one of the techniques and, at the time of its development the only one, for identification and localisation of single base pair mutations in both mRNA and in cloned or genomic DNA. The method which was developed independently by two groups (Winter et al 1985, Myers et al 1985a), is based on the ability of RNase A (bovine pancreatic ribonuclease) to cleave a significant percentage of single base mismatches present in RNA:RNA (Winter et al 1985) and DNA:RNA (Myers et al 1985a) hybrids. For RNase protection experiments total RNA was extracted from lymphoblastoid cell lines from 7 Rb patients and 3 normal controls using the single step acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi 1987) as described in Section 2.7. The resultant RNA was analysed on 1% agarose gels (Section 2.9) to determine the integrity of the 28S and 18S ribosomal RNA bands which is an indirect indication of the quality of the mRNA. A representative of such an RNA checker gel is shown in figure 3.21.

The 4.7 kb of the RB1 mRNA, however, is too large to analyse in a single step using RNase protection so seven overlapping RNA probes (riboprobes) covering the entire 4.7 kb of RB1 were used in these experiments.
Figure 3.21 Typical agarose gel analysis of total cellular RNA. 0.5 μl aliquots (1-2 μg) of total RNA were loaded in each lane. The 28S and 18S bands are seen on a background of heterogeneously sized RNA fragments. The integrity of the ribosomal RNA bands is an indication of the overall integrity of the RNA preparation. The RNA sample in lane 5 does not exhibit these bands and is, therefore, considered degraded.
Individual riboprobes were transcribed from the 4.7 kb cDNA fragments which had been subcloned into transcription vector plasmids (the kind gift of Dr J Dunn in Toronto). The details of these vectors, the size of the riboprobes transcribed and the size of the protected fragments expected are summarised in table 3.8. For the transcription of the correct size riboprobe, each plasmid was linearised with the appropriate enzyme (table 3.8) and radio-labelled antisense riboprobes were synthesised using phage T7 or T3 RNA polymerase as described in Section 2.16.1. These probes were then purified (Section 2.16.2) and hybridised in solution with total RNA from patients and normal controls (Section 2.16.3). After RNase A and T1 digestion, the products were resolved on denaturing polyacrylamide gels (Section 2.16.3). RNase T1 cleaves single stranded RNA 3' to G residues and is added to achieve a more extensive digestion of small fragments of the non-hybridised riboprobes.

Table 3.8 Details of the RB1 subclones used to generate riboprobes and the transcription vectors they were cloned into. For antisense transcription either T7 or T3 promoters were used after linearisation of the plasmids with the enzymes indicated in the table. The size of the transcribed probes and the sizes of the RB1 transcript they hybridise to are shown together with the exons they protect.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Plasmid</th>
<th>Promoter</th>
<th>Enzyme</th>
<th>Probe Size (bp)</th>
<th>Protect. size (bp)</th>
<th>RB1 exons*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PJD 11</td>
<td>PTZ 18</td>
<td>T7</td>
<td>Hind III</td>
<td>899</td>
<td>837</td>
<td>1-8</td>
</tr>
<tr>
<td>PJD 12</td>
<td>PTZ 18</td>
<td>T7</td>
<td>Hind III</td>
<td>523</td>
<td>475</td>
<td>1-5</td>
</tr>
<tr>
<td>PJD 13</td>
<td>Bluescript</td>
<td>T3</td>
<td>Pvu II</td>
<td>827</td>
<td>763</td>
<td>8-16</td>
</tr>
<tr>
<td>PJD 17</td>
<td>Bluescript</td>
<td>T7</td>
<td>Hind III</td>
<td>966</td>
<td>908</td>
<td>9-18</td>
</tr>
<tr>
<td>PJD 18</td>
<td>Bluescript</td>
<td>T7</td>
<td>Bal I</td>
<td>914</td>
<td>912</td>
<td>17-23</td>
</tr>
<tr>
<td>PJD 23</td>
<td>PTZ 18</td>
<td>T7</td>
<td>Hind III</td>
<td>209</td>
<td>163</td>
<td>1-2</td>
</tr>
<tr>
<td>PJD 27</td>
<td>PTZ 18</td>
<td>T7</td>
<td>Bgl II</td>
<td>1068</td>
<td>1047</td>
<td>18-26</td>
</tr>
</tbody>
</table>

* Exon 27 which contains the 3' untranslated region was not included in this analysis.

Tumour tissue was not available for RNA extraction so constitutional DNA from mutant gene carriers was used which, in any case, would have to be the tissue examined if RNase protection was to be used for screening. In all, RNA from lymphoblastoid cell lines from 7 Rb patients were analysed using RNase protection assays as part of a pilot study to assess the usefulness of
this technique. In all cases only the full length protected fragments were observed of which examples are shown in figure 3.22b. A positive control (chicken progesterone receptor RNA, the kind gift of Dr C Dickson), known to have a mutation at 47th base of the protected fragment (with riboprobe Prc 1020) was included as well as a negative control which did not have this mutation (figure 3.22a). In RNase protection, transcription of full-length probes is very important and riboprobes, after transcription, are purified through polyacrylamide gels to obtain just the full-length transcript (Section 2.16.2). If good separation of riboprobe bands are not achieved at this purification step, less than full-length probes present problems in the assay. figure 3.22a illustrates this point. In the negative control sample after RNase digestion (lane II) two bands were observed; one being 180 bp long (the expected fully protected fragment) and the other slightly smaller. The reason for this is the presence of less than a full-length riboprobe. At the purification stage, these two different length probes were extracted from the gel as one band as their sizes were similar to each other. This problem was eliminated later on by running the purification gel longer to achieve better resolution of the riboprobe bands. In the positive control, with the mutation at position 47, two protected bands would be expected, 133 bp and 47 bp long. In figure 3.22a, the 133 bp fragment can be clearly seen but not the 47 bp fragment. Bands approximately 180 bp long were also seen, although fainter than the 133 bp band, in the same lane indicating that RNase digestion of the unprotected fragments was incomplete. This was confirmed since the same bands were present in the control lane (containing riboprobe and no sample RNA) after digestion (lane II). Incomplete digestion was also evident in some of the lanes of the autoradiograph shown in figure 3.22b which results from protection of 4.7R transcripts with probe pJD12. RNA from three Rb patients and one normal individual were hybridised with pJD12 (lanes I) and digested with RNase A and T1 (lanes II). pJD12 is 523 bp long and hybridises to 475 bases at the 5' end of the RB1 mRNA (corresponding to exons 1-5). After digestion, although the most prominent band in all samples, is the fully protected 475 bp fragment, there are also other smaller bands in some lanes. Despite repeated attempts incomplete digestion of unprotected fragments continued to be a problem in these experiments making identification of abnormal cleavage patterns and interpretation of the results difficult. Because of this and other limitations which are discussed in Section 4.2, other techniques were subsequently investigated which would potentially detect mutations more efficiently in RB1.
Figure 3.22 Autoradiographs of two RNase protection gels comparing (a) positive and negative control total RNA samples (chicken progesterone receptor RNA) and (b) total RNA from Rb patients and controls. M = marker lane containing $^{35}$S-dATP labelled Hinf cut pAT153 DNA (Section 2.11, Appendix II).

(a) The chicken progesterone receptor probe (Prc 1020) is shown in lane "P". The products after hybridisation of Prc 1020 to total RNA, but before digestion, are marked "I" and after digestion are marked "II". The control, "C", contained all of the reaction ingredients, including the riboprobe, but no sample RNA. The negative control RNA (-), which does not carry the mutation, produced full length (180 bp) and nearly full length protected fragments. The positive control RNA (+) carrying the mutation produced a 133 bp protected band demonstrating the presence of the mutation at the 47th base position. The faint bands around 180 bp in the same lane is the result of incomplete digestion which is confirmed by the presence of the same faint bands in lane II of C.

(b) RNase digestion patterns of 4.7R transcripts from lymphoblastoid cell lines of three Rb patients, ET27, 139A, ET88 and one normal individual, N. The riboprobe used (lane P) was pJD12 which is 523 bp long and protects a 475 bp fragment. C is the control sample which contained the riboprobe but no sample RNA. Lanes marked "I" and "II" show the products before and after digestion with RNases, respectively. Using pJD12, which covers an area of the 4.7R transcript corresponding to exons 1-5, only the full length protected fragment is seen in the transcripts from the normal individual and Rb patients. No band is observed in the control lane after digestion as might have been expected. High background (due to incomplete digestion), however, is seen in some of the other lanes after digestion.
3.2.2 PCR amplification of individual exons and sequencing

With the publication of the complete coding sequence of RB1 and approximately 200 base pairs of the introns flanking each exon (McGee et al 1989), another approach to mutation detection became possible. The published sequence was used to design oligonucleotide primer pairs for the amplification of the 27 individual exons of the RB1 gene and their flanking regions using PCR amplification. Since it was likely that not all of the mutations in Rb patients would occur within the exons, primers were designed which lay in the intron sequences at least 40 bp away from the splice junctions. Details of primer sequences and their annealing temperatures are given in table 3.9. The length of the PCR products ranges from 212 to 625 bp (table 3.9). The length of the flanking intron region included in the amplified product varies between 42 and 344 bp.

Several factors influenced the selection of sequences used to make PCR primers. Particularly important was the size of the amplified product that could be easily sequenced which is between 500-600 bp. Non-specific DNA amplification was another potential hazard and particularly troublesome areas of the RB1 gene were those showing homology to Alu repeat sequences and GC-rich sequences which were avoided for primer design. The majority of Alu-like sequences are located between exons 2-17 (McGee et al 1989) and, by avoiding these and stretches of polypurines and polypyrimidines, non-specific priming, which might otherwise have produced "background" bands, was reduced. Two exceptions to this principle, however, were primers 11290 (exon 11) and 5535 (exon 17) which are located in areas which show homology to Alu repeats but, in spite of this, generated specific PCR products. The GC:AT ratio is another important factor to be considered when designing primers. Wherever possible, sequences were selected for primers with a random base distribution. High AT content requires low annealing temperatures which, in turn, increases the chances of non-specific amplification and were avoided. Where possible, GC-rich sequences were also avoided to reduce the formation of secondary structures in the template which are known to cause a variety of problems from non-specific priming to lack of product (Innes 1990). All primer pairs were also studied for complementarity. Primers with homologous 3' sequences were avoided to reduce the incidence of "primer dimer" formation which occurs when one primer is extended by the polymerase over the other primer. With these
simple precautions primers were designed which, empirically, produced specific amplification.

The length of the primers ranged from 19-27 bp (table 3.9). Primers which were less than 19 bp generally generated non-specific amplification products. The annealing temperatures were calculated using melting temperatures ($T_m$) of the primers. $T_m$ for each primer was obtained from one of two different equations (depending on primer lengths) which are given at the bottom of table 3.9. For a pair of primers, the average of their $T_m$ was taken (if different) and a temperature which is 2-3°C lower than $T_m$ was used as the annealing temperature. Although the primer pairs were usually the same length (mean of 22 bp), the length of some was altered to obtain a primer pair with a similar annealing temperature. 10% DMSO was added to the PCR reaction to obtain a specific PCR product for regions such as the RBI promoter, exon 1 and flanking intron region which are extremely GC-rich. DMSO is an important reagent which has been shown to reduce secondary structure formation in the DNA (Winship 1989, Saiki 1989). DMSO, however, can also inhibit Taq polymerase thereby decreasing the product yield.

Although the majority of primer pairs were designed from sequences located towards the distal extremes of the sequenced intron regions, there were a few exceptions. For example, the 3' primer for exon 20 is situated upstream to the RBI.20 VNTR region. Also, since exons 15 and 16 are separated by only an 80 bp intron sequence, a single pair of primers was used to amplify both exons. When amplifying particular exons for sequencing, one of the primer pair was biotinylated at the 5' end (Section 2.21). Although, initially, the primers listed in table 3.9 were used, additional internal primers were designed later for exons 17 and 20 (Section 3.3, table 3.10), which lay closer to the exon improving the quality of the sequencing ladders.

All PCR products were analysed on 2% agarose gels to ensure the correct sized band was present and that no background bands were there. Representative gels are shown in figure 3.23.
Table 3.9 Details of the oligonucleotide primers designed for amplification of individual exon and flanking intron sequences of the RB1 gene. The annealing temperatures* for PCR amplification, full size of the amplified fragment, restriction enzymes used to digest PCR products for SSCP analysis and sizes of cut fragments are given. Both the promoter region and exon 1 required 10% DMSO in the PCR reaction (#).

<table>
<thead>
<tr>
<th>Primer</th>
<th>Location</th>
<th><strong>Primer Location</strong> Sequence (5'-3')</th>
<th>Temp °C</th>
<th>Full size (bp)</th>
<th>Enzyme Cut size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6007</td>
<td>RB 5 x PRO GATCCCAAAAGCCAGCAAGTGCTCT</td>
<td>62#</td>
<td>570</td>
<td>Sma I 230</td>
<td></td>
</tr>
<tr>
<td>6008</td>
<td>RB 3 x PRO TCAACGTCCCCCTGAGAAAACCGGA</td>
<td></td>
<td></td>
<td>BspH I 176/164</td>
<td></td>
</tr>
<tr>
<td>9691</td>
<td>RB 5 x 1</td>
<td>CGTGGCGGCCGCGTCGTCTCT</td>
<td>62#</td>
<td>307</td>
<td>Ddel 153</td>
</tr>
<tr>
<td>9692</td>
<td>RB 3 x 1</td>
<td>ACCGGGGCCCTGCGAGACGAC</td>
<td></td>
<td></td>
<td>154</td>
</tr>
<tr>
<td>10143</td>
<td>RB 5 x 2</td>
<td>GTTCTTTTTTACAGTAGTTATATGTCT</td>
<td>60</td>
<td>409</td>
<td>HpaI 214</td>
</tr>
<tr>
<td>10142</td>
<td>RB 3 x 2</td>
<td>CGTGCCCGGCCCTACACATTTTAAA</td>
<td></td>
<td></td>
<td>191</td>
</tr>
<tr>
<td>8202</td>
<td>RB 5 x 3</td>
<td>GCCATCAAGAGATGTGTTCTGCAA</td>
<td>58</td>
<td>477</td>
<td>AluI 243</td>
</tr>
<tr>
<td>8201</td>
<td>RB 3 x 3</td>
<td>GACACAAACTGCATCTCCTTAAG</td>
<td></td>
<td></td>
<td>234</td>
</tr>
<tr>
<td>8203</td>
<td>RB 5 x 4</td>
<td>CTTCCAAAAGATATAGTGTTTGTGG</td>
<td>58</td>
<td>445</td>
<td>RsaI 269</td>
</tr>
<tr>
<td>8204</td>
<td>RB 3 x 4</td>
<td>CCAAGGACCTCAGATAAGCATATT</td>
<td></td>
<td></td>
<td>176</td>
</tr>
<tr>
<td>8206</td>
<td>RB 5 x 5</td>
<td>GAAGACTAATAGGAGAGATTACTGA</td>
<td>58</td>
<td>488</td>
<td>AflIII 218</td>
</tr>
<tr>
<td>8205</td>
<td>RB 3 x 5</td>
<td>TGATCTGAATCAATCCCCTATT</td>
<td></td>
<td>Taq I 142/128</td>
<td></td>
</tr>
<tr>
<td>8208</td>
<td>RB 5 x 6</td>
<td>GAAACACCCCAAGATATATCTGG</td>
<td>58</td>
<td>326</td>
<td>AluI 179</td>
</tr>
<tr>
<td>8207</td>
<td>RB 3 x 6</td>
<td>CCAAGTGGTTTGTTCTAGTACCAG</td>
<td></td>
<td></td>
<td>147</td>
</tr>
<tr>
<td>7090</td>
<td>RB 5 x 7</td>
<td>ACTCTACCTCTCGCATTTTCTCTCAT</td>
<td>60</td>
<td>491</td>
<td>RsaI 193/176</td>
</tr>
<tr>
<td>7091</td>
<td>RB 3 x 7</td>
<td>CTCTCCTCTCCTACAAACAAAGATTAA</td>
<td></td>
<td></td>
<td>61</td>
</tr>
<tr>
<td>20085</td>
<td>RB 5 x 8</td>
<td>GACCTAGTTAGTAGTGAATACCTTC</td>
<td>55</td>
<td>316</td>
<td>TaqI 192</td>
</tr>
<tr>
<td>20084</td>
<td>RB 3 x 8</td>
<td>CATGCTCATAACAAAAACAAATGAAA</td>
<td></td>
<td></td>
<td>124</td>
</tr>
<tr>
<td>7095</td>
<td>RB 5 x 9</td>
<td>TGCAATGGGGAATGGACACCTCTAAC</td>
<td>60</td>
<td>316</td>
<td>EcoRI 171</td>
</tr>
<tr>
<td>7094</td>
<td>RB 3 x 9</td>
<td>CTACTGGCTAGTTCTTCTTGGGC</td>
<td></td>
<td></td>
<td>145</td>
</tr>
<tr>
<td>10145</td>
<td>RB 5 x 10</td>
<td>TCTGTACCTCCTTTAGTTAGATACCC</td>
<td>60</td>
<td>492</td>
<td>BglII 225/218</td>
</tr>
<tr>
<td>10144</td>
<td>RB 3 x 10</td>
<td>CTGTTATAGGACACACAAATTAC</td>
<td></td>
<td>Hinfl 49</td>
<td></td>
</tr>
<tr>
<td>11289</td>
<td>RB 5 x 11</td>
<td>GACACAGAAGCATTATATCTGC</td>
<td>55</td>
<td>294</td>
<td>MboI 151</td>
</tr>
<tr>
<td>11290</td>
<td>RB 3 x 11</td>
<td>CCTGCTCCTAAATATATATTCTT</td>
<td></td>
<td></td>
<td>143</td>
</tr>
<tr>
<td>9987</td>
<td>RB 5 x 12</td>
<td>CCCAGGTTTTTTGAGGGAATG</td>
<td>60</td>
<td>465</td>
<td>BclI 198</td>
</tr>
<tr>
<td>9986</td>
<td>RB 3 x 12</td>
<td>GGTGAGCAAGGCAAATAGGTAA</td>
<td></td>
<td>RsaI 162/105</td>
<td></td>
</tr>
<tr>
<td>5528</td>
<td>RB 5 x 13</td>
<td>TAAATGAGGTGTGGTTAGTACTGT</td>
<td>60</td>
<td>570</td>
<td>EcoRI 225</td>
</tr>
<tr>
<td>5529</td>
<td>RB 3 x 13</td>
<td>AATTCTACATGGCTATGTGTTCC</td>
<td></td>
<td>AseI 166/179</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>13679</td>
<td>RB 5 x 14</td>
<td>CTAAAATAGCAGGCTCTTTATTTTTC</td>
<td>58</td>
<td>212</td>
<td>-</td>
</tr>
<tr>
<td>13680</td>
<td>RB 3 x 14</td>
<td>ATCTTGATGCTTGTGACCTCCTCTGAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11293</td>
<td>RB 5 x 15</td>
<td>ATTCATGCTGACCAAATAAGGTT</td>
<td>55</td>
<td>361</td>
<td>NdeI</td>
</tr>
<tr>
<td>11294</td>
<td>RB 3 x 16</td>
<td>TTCTCCTAAACCTCACACTATCC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5535</td>
<td>RB 5 x 17</td>
<td>ATAAAAATGTTTTAACCCTTCTACT</td>
<td>55</td>
<td>555</td>
<td>Rsal</td>
</tr>
<tr>
<td>5536</td>
<td>RB 3 x 17</td>
<td>GTGCCATGTTTTAAGAAAAGCTATT</td>
<td></td>
<td>NdeI</td>
<td>104/101</td>
</tr>
<tr>
<td>10604</td>
<td>RB 5 x 18</td>
<td>ATGTACCTGGGAAAATTATGCT</td>
<td>58</td>
<td>221</td>
<td>BclI</td>
</tr>
<tr>
<td>10603</td>
<td>RB 3 x 18</td>
<td>CTATTGGCATTTGATGGGCAAC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9439</td>
<td>RB 5 x 19</td>
<td>AGGCAGATACCCCAAGAAAGCCA</td>
<td>62</td>
<td>485</td>
<td>Hinfl</td>
</tr>
<tr>
<td>9440</td>
<td>RB 3 x 19</td>
<td>CACAGAGATATAGTGACTTGGCC</td>
<td></td>
<td>MluI</td>
<td>176/91</td>
</tr>
<tr>
<td>9438</td>
<td>RB 5 x 20</td>
<td>TTCTCCTGGGAAAGAAAAGATGGG</td>
<td>60</td>
<td>350</td>
<td>HpaII</td>
</tr>
<tr>
<td>14928</td>
<td>RB 3 x 20</td>
<td>AGTTAACAAGTAAAGTGGGAGGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9435</td>
<td>RB 5 x 21</td>
<td>GACTTTCAAACCTGAGCCTCAGTATGG</td>
<td>58</td>
<td>517</td>
<td>AseI</td>
</tr>
<tr>
<td>9434</td>
<td>RB 5 x 22</td>
<td>GCAGCTATAATCAACCTAAGAAG</td>
<td>60</td>
<td>363</td>
<td>BclI</td>
</tr>
<tr>
<td>9433</td>
<td>RB 3 x 22</td>
<td>GTTTTGGGAGCCATTACATTAGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9694</td>
<td>RB 5 x 23</td>
<td>TCTAAATGTAATGGGTCCACCA</td>
<td>58</td>
<td>420</td>
<td>BstNI</td>
</tr>
<tr>
<td>9695</td>
<td>RB 3 x 23</td>
<td>CATCTTGCGTGTAACTAGTG</td>
<td></td>
<td>AluI</td>
<td>136/98</td>
</tr>
<tr>
<td>9696</td>
<td>RB 5 x 24</td>
<td>TAAAACTAAGAGACTAGGGTGATGAT</td>
<td>58</td>
<td>579</td>
<td>BstEII</td>
</tr>
<tr>
<td>9697</td>
<td>RB 3 x 24</td>
<td>TAGATTTGCGTTGAGAAAGAACAC</td>
<td></td>
<td>BstNI/HindIII</td>
<td>93/50</td>
</tr>
<tr>
<td>9991</td>
<td>RB 5 x 25</td>
<td>ATTTTGTCGAATGAGCGAAATT</td>
<td>60</td>
<td>625</td>
<td>BstNI</td>
</tr>
<tr>
<td>9990</td>
<td>RB 3 x 25</td>
<td>TGATGCTATGATTTTTCAGTG</td>
<td></td>
<td>Bsp1286</td>
<td>198/161</td>
</tr>
<tr>
<td>9993</td>
<td>RB 5 x 26</td>
<td>AACCCTGTTATTTTGTGAGAACCAC</td>
<td>60</td>
<td>524</td>
<td>PstI</td>
</tr>
<tr>
<td>9992</td>
<td>RB 3 x 26</td>
<td>TGAACTGCTGAACTGAGCATTTTCAC</td>
<td></td>
<td>NsiI</td>
<td>166/143</td>
</tr>
<tr>
<td>10609</td>
<td>RB 5 x 27</td>
<td>AAGGTCTCAGCGCCACATCAGTTGGA</td>
<td>62</td>
<td>218</td>
<td>-</td>
</tr>
<tr>
<td>10608</td>
<td>RB 3 x 27</td>
<td>GAGGTGTACACAGTGGGCCACAGGG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*For primers less than 20 bp; 2 - 3°C below \( T_m (^\circ\)C) = 4 \((G+C)\) + 2 \((A+T)\)

For longer primers; 2 - 3°C below \( T_m (^\circ\)C) = 69.3 + [0.41 \times (G+C)\%] - 650/length of primer

184
Figure 3.23 Agarose gels showing the PCR amplification products for a variety of RBI exons and their flanking intron sequences. The template DNAs were derived from peripheral blood samples. In (a) the amplified products for exons 15-16, 17 and 20 of the RBI gene are shown. The size of the fragments are 361, 555 and 350 bp long respectively. In (b) PCR products of exons 22, 23, 24, 25 and 26 are shown where the sizes (in bp) are 363, 420, 579, 625 and 524 respectively. M = marker lane containing 1 μg of 1 Kb DNA ladder (Gibco-BRL).
In practise, it was necessary to optimise the conditions for PCR amplification and sequencing for each exon. During the course of these sequencing experiments a few minor errors in the published sequence and a new polymorphism was identified (figure 3.24), which are recorded below for the record.

In intron 14 (figure 3.24a), the 17th and 18th base pairs of the intron was published as 'AA' (coding strand) but in the DNA of all individuals sequenced, only one 'A' was observed in this region. McGee et al (1989) had an undetermined base at position 32 which is demonstrated here to be a 'G' (coding strand) (figure 3.24a). In intron 19, bases 78-80 bp 5' to exon 20 were published as 'ACA' (non-coding strand) which has now been shown to be 'TCG' using DNA from Rb patients and normal individuals (figure 3.24b).

In intron 17, bases 81-83, an 'AAA' (coding strand) was reported which, in fact, was shown to be polymorphic in the population with one to three 'A's at this site (figure 3.24c). This polymorphism does not alter a known restriction enzyme site. These variations to the published sequence and the presence of the A(n) polymorphism in intron 17, has since been confirmed by others (A Hogg, D Yandell, personal communication). Since the identification of potential mutations depends on the comparison of patient material with the "normal" sequence it is important to establish these errors, most of which have now been incorporated in the reference databases.

Preliminary application of this technology used asymmetric PCR amplification (see discussion Section 4.2) and sequences were generated but no mutation was found. It proved a very time consuming process and, with the development of new techniques (see later) became redundant. The advantage of pursuing this approach, however, proved to be the development of the sequencing protocols directly from PCR products.
**Figure 3.24** DNA sequencing ladders from those regions in introns 14, 17 and 19, of the RB1 gene which differed from the published sequence.

In (a) sequence from the coding strand of intron 14 from a normal individual is shown. The published sequence and the sequence observed are given above. Where two 'A' residues (the 5th and 6th bases) are seen in the published sequence only one is observed on the gel. The base labelled 'N' in the published sequence was confirmed to be a 'G'.

(b) Sequence from the non-coding strand of intron 19 from the DNA of a normal individual. The 'ACA' given in the published sequence is observed as 'TCG' on the gel. The published sequence and the sequence observed is compared on top.

In (c) the polymorphism $A_n (n=1-3)$, in intron 17 is shown. Sequence from the coding strand of intron 17 from the DNA of two normal individuals and the non-coding strand sequence (bottom) of the same area from the DNA of a third individual is shown. The published sequence is given on top. In the sequence ladder shown on the top, one 'A' is observed whereas in the middle three 'A' residues are seen (arrows) as shown. In the non-coding strand sequence from a third individual (bottom), only one 'T' is observed in the same area.
Figure 3.24

(a) The published sequence
Sequence observed
5'-CAATATAAAAAATTTCA\text{NCC}-3'
5'-CAAT A TAAAAAAATTTCA\text{GCC}-3'

(b) The published sequence
Sequence observed
3'-TTAATCAA ATTAAG-5'
3'-TTAATCGT TAAG-5'
Figure 3.24
(c) The published sequence

5'-CTTAACAAATCTACC-3' (coding)
3'-GAATTGTAGATGG-5' (non-coding)

coding strand sequence

non-coding strand sequence
3.2.3 Detection of single base mismatches by CDI modification of DNA heteroduplexes followed by primer extension

Although, using PCR, the RB1 gene became more accessible for mutation analysis, it was obvious that an exon-by-exon analysis of the 27 exons was not a practical approach for routine screening designed to identify mutations in Rb patients. RNase protection had not proved particularly useful and so it was important to find a more practical means of screening each amplified exon for those most likely to carry mutations. A variety of such techniques were described in the literature during the execution of this thesis one of which was the CDI method (Ganguly and Prockop 1990).

Water soluble CDI [l-cyclohexyl-3-(2-morpholino-ethyl carbodiimide metho-p-toluene sulfonate] modifies unpaired Gs and Ts in DNA which interrupts primer extension of the DNA template heteroduplexes in PCR reactions. To evaluate the potential of this method, the RB1.3 polymorphic locus of the RB1 gene was amplified. The RB1.3 polymorphic site is situated in intron 3, the 45th base of which is either a 'C' or a 'T' (Section 3.1.2). A 477 bp fragment containing the RB1.3 site was amplified using primers 8202 and 8201 (table 3.9) from the DNA of two individuals, one of whom was homozygous CC and the other homozygous TT. When the PCR products were mixed, denatured and renatured (Section 2.19.1) a proportion of DNA formed heteroduplexes containing C-A and T-G mismatches (figure 3.25). This DNA was then treated with CDI at either 30°C or 37°C as described in Section 2.19.2. Primer extension (Section 2.19.3) was then performed with the 5' primer (8202) or the 3' primer (8201) in separate PCR reactions which contained 32P-dCTP to generate labelled products. As CDI only modifies unpaired Gs and Ts, the mismatched bases T-G were modified in heteroduplexes. In the RB1.3 polymorphism, the mismatch is located at position 311 from the 5' end in the 477 bp amplified fragment. When the modified heteroduplexes are extended from the 5' primer (8202), a 310 bp product is generated and when extension is from the 3' primer (8201), a 166 bp product is expected (figure 3.25).
Figure 3.25 Schematic diagram summarising the CDI modification procedure. The detection of T-G mismatches formed by mixing PCR products, 477 bp long, from DNA1 (with a 'C' at position 311 from the 5' end) and DNA2 (with a 'T' at position 311) is given as the specific example.
The autoradiograph produced when the products were resolved is shown in figure 3.26. In addition to test heteroduplex samples, homoduplexes from each DNA sample (lanes I, II) and heteroduplexes not modified with CDI (lane III) were included as controls. Because there is considerable loss of DNA during processing of the CDI-treated samples the unmodified heteroduplex DNA has been diluted two fold before loading onto the gel in order to generate bands of comparable intensity to the CDI-modified DNA bands. The loss of material from CDI-treated DNA is due to the difficulty in precipitating the DNA as CDI interacts non-covalently and alters its solubility (Ganguly and Prockop 1990). Also, CDI-treated DNA adheres to glass and plastic surfaces. Siliconised eppendorf tubes were used routinely, therefore, in this experiment to minimise these losses. In the control lanes (C), only the 477 bp full length fragment is observed. In the lanes containing the 5' primer-extended heteroduplex (H5'), a 310 bp fragment is seen, as expected. A 166 bp fragment is seen in the lanes containing the 3' primer-extended heteroduplexes (H3'), also as expected. However, there is also a high background in these lanes. The difference between 30°C and 37°C lanes seems to be only a loading concentration effect especially in the H5' lanes. There could be a variety of reasons for the presence of these background bands including internal priming of the DNA heteroduplexes and non-specific modification of bases by CDI. The appearance of slight background in the two fold diluted unmodified heteroduplex sample and the difference in the background between H5' and H3' samples point to non-specific extension during PCR especially with the 3' primer. The low intensity of the 310 bp and 166 bp bands points to loss of CDI-modified DNA and also reflects the fact that the modified heteroduplex constitutes only 1/4 of all the DNA duplexes in the starting DNA sample. It is essential, therefore, to optimise the reaction conditions to minimise the loss of CDI-modified DNA further. Concentrating the CDI-treated samples by ultrafiltration instead of ethanol precipitation, for example, is one way. The PCR conditions also need to be optimised to reduce background resulting from non-specific amplification. On the whole, this method looked promising but the large number of steps involved in this method is a disadvantage for a routine mutation screening technique.
Figure 3.26 Detection of T-G mismatches using the RB1.3 polymorphism. PCR products, 477 bp long, were generated. Heteroduplexes containing mismatches were prepared by mixing PCR products containing a 'C' at nucleotide 311 (from the 5' end), with PCR products containing a 'T' at the same site. The heteroduplexes were then CDI-treated under two different conditions (30°C and 37°C) and extended with primer 8202 (H5') and 8201 (H3') in separate PCR reactions which contained 32P-dCTP to generate labelled products. Lanes marked "C" on the gel contain control samples. Homoduplexes of DNA1 (with base C) are present in lane I and DNA2 (with base T) are in lane II after treatment. Lanes marked "III" contain heteroduplexes of DNA1 and DNA2 prepared under the same conditions as the test samples except that they were not modified by CDI. Lanes marked "M" contain 35S-dATP labelled molecular weight markers; from left to right, 1 Kb DNA ladder, Hinf I cut pBR322 fragments, Hinf I cut pAT153 fragments. The 477 bp full-length fragment is seen in every lane as expected. In the H5' lanes, a 310 bp fragment and in the H3' lanes, a 166 bp fragment is also observed indicating the presence of the mismatch at position 311 from the 5' end in the heteroduplexes formed.
3.2.4 PCR-SSCP analysis

The previously described methods had not proved too satisfactory for the large number of patients which need screening at this time so another method, SSCP analysis, was investigated. In this method, which was first described by Orita et al (1989a), the target sequence is radioactively labelled and amplified simultaneously using PCR (Section 2.20). The amplified DNA product is then denatured and the single stranded molecules are resolved in non-denaturing polyacrylamide gels. The migration of a single stranded DNA molecule in a gel under non-denaturing conditions is sequence dependent. Mutations affecting a DNA sequence lead to a conformation change affecting mobility, allowing detection of a mutation as altered mobility of the separated single strands in gels. Single base pair mutations can be detected using this technique in DNA fragments of up to approximately 350 bp. Subtle changes in nucleotide sequence are more readily detectable, however, in smaller molecules (Orita et al 1989a). As can be seen in table 3.9, most of the exon containing regions amplified from the RB1 gene are longer than 350 bp. Restriction enzyme sites within each of the amplified regions, therefore, were identified so that shorter fragments could be generated for SSCP analysis. Digestion with the appropriate restriction enzyme generally produces DNA fragments 100-300 bp long (table 3.9). In some cases, it was necessary to use a combination of up to 3 different enzymes to generate fragments of the required sizes. Complete digestion of PCR products, which is essential to avoid artefact bands on SSCP gels, was confirmed on agarose gels. In some cases, both digested and undigested PCR products were analysed.

In SSCP gels, the detection of mutations depends on the conformational changes of the single stranded DNA molecule induced by the mutation, and therefore, sensitive to the physical environment in the gel. Prevention of ohmic heating is especially important for reproducible results. The mobility patterns of single stranded DNA molecules vary according to the electrophoresis temperature, the presence or absence of glycerol (which has a weak denaturing action on nucleic acids), salt concentration and the extent of cross-linking of the gel. Under one particular set of conditions, some mutations might exhibit more readily detectable mobility changes than others. In this project, most of the SSCP analysis were performed using gels containing 19:1 acrylamide:N,N'-methylenebisacrylamide and 10% glycerol.
Electrophoresis was at 30 W constant power at room temperature for 6 hours. These conditions were reported to be satisfactory for the detection of majority of mutations (Orita et al 1989a) and adopted in this preliminary screening of the RB1 exons.

The analysis of exons 13-22 was given priority because these exons code for the regions of the Rb protein needed for binding to viral oncoproteins and cellular proteins and have been reported to be common sites for mutations (Hu et al 1990, Huang et al 1990). Exons 3, 9 and 13-24 of the RB1 gene from lymphoblastoid cell line DNA of one Rb patient and blood DNA of 14 other Rb patients were analysed on SSCP gels. In 7 of these DNAs, exons 4-7, 10 and 12 were also analysed. Each SSCP analysis takes 3-4 days, so in practice, all patients were analysed for a single exon at one time. Abnormal banding patterns were identified in the DNA of 6 unrelated Rb patients using SSCP analysis of exons 3, 13, 17 and 20 each of which is described in more detail below.

SSCP analysis of exon 13 (figure 3.27a) revealed an abnormal banding pattern in one of the DNAs analysed. This DNA, from individual I.2 of family RBF58 which was described in Sections 3.1.4 and 3.1.5, exhibited two extra bands and a slight upward shift in the fourth band on the gel compared to other samples. SSCP analysis of exon 17 (figure 3.27b) revealed one additional band (just below the first band) in the DNA from individual I.1 of family RBF59 (Appendix III, Section 3.3.1) compared to other samples.

PCR amplification of exon 20 region using primers 9438 and 14928 (table 3.9) produces a 350 bp fragment which was first analysed undigested. In the gel containing the 350 bp PCR products shown in figure 3.28a, abnormal banding patterns were observed in the DNA of three individuals from families RBF18, RBF62 and RBF29. The exon 20 DNA from individual III.3 of family RBF18, which was described in Section 3.1.7, showed an additional upper band compared to the normal mobility pattern (figure 3.28a). Two extra lower bands were also present in the DNA of individual I.1 from RBF62 (Appendix III, Section 3.3.1). There was also a slight upward shift in the two uppermost bands. The DNA amplified from II.4 of RBF29 (Sections 3.1.5, 3.1.7), on the other hand, exhibited one (or two not well resolved) additional bands compared with samples from other patients.
**Figure 3.27** SSCP analysis of (a) exon 13 and (b) exon 17 of the RB1 gene.

(a) SSCP analysis of exon 13. DNA was amplified using primers 5528 and 5529. 570 bp PCR products were digested with Eco RI and Ase I producing 225, 179 and 166 bp fragments. The digestion products were analysed on 6% non-denaturing polyacrylamide gels containing 10% glycerol at 30 W at room temperature for 6 hours. DNA in the lane marked RBF58 exhibits two additional bands (arrows) towards the lower part of the gel compared with other samples.

(b) SSCP analysis of exon 17. DNA was amplified using primer pair 5535/5536 and the 555 bp amplification product was digested with Rsa I, Nde I and Sau3A I. 142, 119, 104, 101 and 89 bp digestion products were analysed on 6% non-denaturing polyacrylamide gels containing 10% glycerol. DNA marked RBF59, when compared to the other samples, shows an extra band (arrow) just below the first band on the gel.
Figure 3.27

(a) RBF58

(b) RBF59
When exon 20 was analysed following Hpa II digestion of the PCR products (figure 3.28b) fragments 177 and 173 bp long were produced. In this case the DNA from RBF18 exhibited an extra band which was lower down the gel. A slight upward shift of all bands in the middle section was also observed (figure 3.28b). In DNA from RBF62, an extra band was seen in between the two lower bands. The second band (from the top) had a slower and a faster migrating component associated with it and the intensity of the third band was much reduced compared with the same band of the other samples. In DNA from RBF29, the intensity of the second band was low and the third band migrated more slowly compared to other samples. The band seen at the fourth position in the other samples was not present in this DNA.

SSCP analysis was also used to detect the RB1.3 single base pair polymorphism described in Section 3.1.2 and is much simpler and quicker than PCR-sequencing. RB1.3 primers 13774 and 13773 (table 3.2) amplify a 200 bp fragment containing 58 bp from exon 3 and 142 bp from intron 3. Although it was difficult to differentiate between CC and TT homozygous individuals in some gels, the difference between heterozygous and homozygous individuals was clear in every gel. More work needs to be done, however, on this type of analysis of RB1.3, to establish the optimal conditions for differentiation between homozygotes in SSCP gels. Time constraints did not allow this work to be completed in this project.

In an SSCP analysis of the RB1.3 locus, an abnormal mobility pattern, superimposed on the RB1.3 banding pattern, was observed in the DNA of two affected individuals from family RBF64. In figure 3.29, extra bands are seen on the gel in the DNA of individuals II.2 and III.1. Although these individuals carry the same mutation, the mobility pattern of their DNAs were slightly different due to their different RB1.3 genotypes. II.2 is a homozygote for base 'C' whereas III.1 is a heterozygote carrying base 'C' in one of his alleles and base 'T' in the other at the 45th position of intron 3 where the polymorphic site is located. Compared to other homozygotes II.2 exhibited five extra bands and compared to other heterozygotes III.1 exhibited four extra DNA bands on the gel (figure 3.29). I.2 and II.3 in this family are unaffected individuals and the difference in their banding pattern reflects the fact that I.2 is a heterozygote and II.3 is a homozygote for base 'T'.

199
**Figure 3.28** SSCP analysis of exon 20 of the RB1 gene. A 350 bp fragment was amplified using primers 9438 and 14928. In (a) the whole PCR products and in (b) Hpa II digested PCR products (177 and 173 bp) were analysed on 6% non-denaturing polyacrylamide gels containing 10% glycerol at 30 W run at room temperature for 6 hours. In the lanes marked *, undenatured samples were run to indicate the position of double-stranded DNA. In (a) DNA in the lane marked RBF18 shows an additional upper band (arrow); DNA marked RBF62 shows two additional bands (arrows) in the lower part of the gel as well as a slight upward shift of the first two bands; DNA marked RBF29 exhibits one (or two not well resolved) extra band (arrow) at the bottom of the gel compared with other samples. In (b) in the lane marked RBF29, the second larger band is reduced in intensity, the third band is more slowly migrating and the fourth band, seen in other samples, is missing; DNA marked RBF18 exhibits one extra band (fourth band from top) and a slight upward shift of the middle section of bands; DNA marked RBF62 shows one extra band (seventh band from top), a slower and a faster migrating component associated with the second band and reduced intensity of the third band.
Figure 3.28

(a)
Figure 3.29 SSCP analysis of the RB1.3 region of the RB1 gene. A 200 bp DNA fragment containing 58 bp from exon 3 and 142 bp from intron 3 was amplified using the RB1.3 primers given in table 3.2. Full-length PCR products were analysed on a 6% polyacrylamide gel containing 10% glycerol. The RB1.3 genotype of each DNA run on the gel is indicated below each lane. The pedigree for RBF64 is shown above and DNA from II.2 and III.1 exhibit extra bands on the gel compared with other samples. The normal heterozygote pattern (I.2 and the first two lanes) has two lower bands whereas the normal homozygote pattern has only one lower band (II.3 and the 3rd lane). In the DNA of a 'T' homozygote (II.3) the third band migrates more slowly compared with the same band of a 'C' homozygote (3rd lane). DNA from II.2 who is a 'C' homozygote, shows five extra bands compared with the DNA from a normal homozygote. DNA from III.1, who is a heterozygote, shows four extra bands compared with the DNA from a normal heterozygote.
3.3 IDENTIFICATION OF ONCOGENIC MUTATIONS IN SSCP ANALYSED DNA BY PCR-SEQUENCING

The relevant exons from the 6 Rb patients, which exhibited abnormal mobility patterns on SSCP gels were sequenced. When exons 17 and 20 were amplified, additional internal primers, designed specifically to make sequencing easier, were also used (table 3.10). In each case, one of the primers used in the PCR reaction was biotinylated at the 5' end for immobilisation of the DNA on streptavidin-coated magnetic beads to allow separation of the individual strands of the DNA (Section 2.21.1). After amplification and separation of DNA strands, dideoxy sequencing of both strands was performed as described in Section 2.21.2. The exact nature of each of these mutations is discussed in detail below.

Table 3.10 Details of internal primers used in PCR-sequencing of exons 17 and 20.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Location</th>
<th>Sequence (5' - 3')</th>
<th>Temp °C</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20877</td>
<td>RB 5 x 17</td>
<td>ACTTCCAAAAAAATACCTAGCTCAAG</td>
<td>58</td>
<td>318</td>
</tr>
<tr>
<td>20876</td>
<td>RB 3 x 17</td>
<td>TTTGTTAGCCATATGCACATGAATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17295</td>
<td>RB 5 x 20</td>
<td>CATGATTTGAAAAAAAATCTACTTG</td>
<td>52</td>
<td>269</td>
</tr>
<tr>
<td>14928</td>
<td>RB 3 X 20</td>
<td>AGTTAACAAAGTAAGTGGGAGGAGA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3.1 Mutations identified in bilateral Rb families

**RBF64**

In family RBF64 (figure 3.30), there are four bilaterally affected individuals (I.1, II.1, II.2, III.1) and those with a cancer predisposition were identified using linkage analysis and RB1.20. SSCP analysis of DNA from constitutional cells from this family identified an abnormal banding pattern in the DNA from individuals II.2 and III.1 in exon 3 (Section 3.2.4, figure 3.29). DNA from individuals I.1 and II.1, who are both deceased, were not
available for analysis. Sequence analysis revealed a 2 bp heterozygous deletion in the DNA from II.1 and III.1 (figure 3.30a). In these DNAs the 106th (A) and 107th (T) nucleotides of exon 3 are deleted in one allele. From the point of the deletion, both the mutant and normal sequencing ladders are superimposed, one lagging behind by 2 bp. DNA from unaffected individuals I.2 and II.3 showed the normal exon 3 sequence (figure 3.30a). When the sequence around the deletion was analysed, an imperfect 8-9 bp direct repeat within which the deleted bases lay was identified (figure 3.30b). Alternatively, the first repeat could be regarded as a quasi repeat (i.e. identical to the mutant sequence) which is reproduced by the deletion. This sequence also forms part of a direct repeat which is repeated once more in the 3' end of the gene. The sequence between 26 bp and 3 bp upstream of the deleted nucleotides and a sequence (11th-32nd nucleotides) in exon 4 are inverted repeats of each other (T'Ang et al 1989). The area around the deletion and the upstream intron contain Alu-like sequences. A sequence, TGAGC[286]T, similar to the deletion consensus sequence (TG A/G A/G G/T[313] A/C) of Krawczak and Cooper (1991) which has also been implicated as the putative arrest sites for DNA polymerase α, is observed 10 bp upstream of the deletion (figure 3.30b). As a result of this deletion, the first two nucleotides of codon 124 (ATA) are altered and a reading frameshift generates a premature stop codon in exon 4 at codon 129. This would be predicted to lead to premature termination of translation and the production of a truncated protein only 128 amino acids long. However, the splice donor site for exon 3 is only 9 bp downstream of the deleted nucleotides and it is not possible to predict how this would affect the use of this site, if at all. If the deletion reduces the efficiency of splicing from this site relative to any existing cryptic splice site nearby, alternative splicing from the cryptic site might result. Alternatively, exon 3 might be spliced out all together from the mature mRNA. This, however, will affect the reading frame and result in a stop codon further down in exon 4.
Figure 3.30 Sequence analysis of exon 3 from family RBF64.

(a) Primer 13774 (RB1.3 sense, table 3.2) was biotinylated and used in conjunction with primer 13773 (RB1.3 antisense, table 3.2) to amplify a 200 bp fragment containing sequences from exon 3 and intron 3. Primer 13773 was used to sequence the non-coding strand from the DNA of each member of RBF64 except I.1 and II.1 whose DNA were not available for analysis. The family pedigree, together with the RB1.3 polymorphic alleles of each individual, is given above. Individuals I.2 and II.3 show the normal exon 3 sequence. Deletion of two bases (TA, closed arrow) in one allele is identified in the DNA from II.2 and III.1. As a result of the deletion, the sequence ladder of the mutant allele is superimposed on the normal, resulting in two bands at most positions on the gel. At the bottom of each sequencing ladder, the polymorphic bases of RB1.3 (open arrows) are seen. Individuals I.2 and III.1 are heterozygous for this polymorphism having both G and A at the same position. II.2 is homozygous for G and II.3 is homozygous for A.

(b) Nucleotide sequence (the coding strand) of exon 3 (upper case) and intron 3 (lower case) flanking the deletion endpoints. The splice donor site at exon3/intron3 boundary is indicated (|). The 8-9 bp imperfect direct repeat sequences are underlined and the deleted bases are enclosed in the box. Nucleotides marked * vary between repeats. The sequence ACAGAAA also represents a quasi repeat. A sequence (TGAGCT) which is similar to the deletion consensus sequence of Krawczak and Cooper (1991) is shown in italics.
Family RBF58 (figure 3.31), with two bilaterally affected individuals, was previously described in Section 3.1.4 and 3.1.5. Linkage analysis in this family included using formalin fixed paraffin-embedded tumour section DNA from II.1 who is deceased. As a result of this analysis, II.2 was excluded from being a mutant gene carrier (Section 3.1.4). SSCP analysis of exon 13 showed two extra bands in the DNA from individual I.2 compared to other samples (figure 3.27a, Section 3.2.4). Sequence analysis revealed a 1 bp heterozygous insertion 50 bp from the 5' end of exon 13 in the DNA from I.2. An extra 'A' was noted in the coding strand (figure 3.31a). From the point of insertion two bands are observed at each position on the gel as the mutant and normal sequencing ladders are superimposed, one lagging behind by 1 bp. The DNA from individual II.2, was found to be free of this mutation confirming the original prediction made with linkage analysis (Section 3.1.4). Unfortunately, at the time of these findings no more tumour tissue sections from II.1 were available and it was not possible, therefore, to confirm the presence of the mutation in II.1. The DNA from normal individual I.1, as expected, exhibited the normal exon 13 sequence (figure 3.31a). In June 1992, pre-natal screening was carried out for II.3 and the fetus was shown to have inherited the Rb predisposition allele through linkage analysis (Section 3.1.5). The CV DNA obtained was also sequenced and exhibited the same mutation identified in the DNA of I.2 (figure 3.31a), again confirming the results of the linkage analysis. When the sequence around the insertion was examined, the insertion was again found to occur within 8-9 bp imperfect direct repeat sequences of which parts are symmetrical (figure 3.31b). This sequence is also found repeated in exons 10, 11 and 27 of the RB1 gene (T'Ang et al 1989). 3 bp upstream of the insertion a TGAAGG sequence is also present (figure 3.31b). The insertion causes a frameshift in exon 13 starting at codon 422 and leads to a premature stop codon nearby at codon 427 (figure 3.31b). This mutation would be expected to result in the production of a truncated protein 426 amino acids long.
Figure 3.31 Sequence analysis of exon 13 in family RBF58.

(a) Primer 5529 was biotinylated and used with primer 5528 to amplify exon 13. Sequencing from primer 5528 is shown for DNA from I.1, I.2, II.2 II.3. The DNA of II.3 was obtained from CV tissue. Insertion of an 'A' is present in the DNA of I.2 and II.3 in one allele (arrow). As a result of the insertion, the sequence ladder of the mutant allele is superimposed on the normal, resulting in two bands at each position on the gel. II.2 who is unaffected does not carry the mutation. The DNA from normal individual I.1 also exhibits the normal sequence as expected. DNA from II.1 was not available at the time of sequencing.

(b) Nucleotide sequence of exon 13 and the amino acids it encodes around the insertion point. The point of insertion of 'A' between the first two bases (A-TA) of codon number 422 is indicated with an arrow. The reading frameshift caused by the insertion converts Ile to Asn (codon 422) and then leads to a stop codon further down at codon number 427. The 8-9 bp imperfect direct repeat sequence, within which the extra 'A' is located, is underlined. Nucleotides marked * vary between repeats. A part of the repeat sequences is symmetrical (AGGATATAGGA). A sequence (TGAAGG) similar to the deletion consensus sequence of Krawczak and Cooper (1991) is shown in italics.
Figure 3.31

(a) 

(b) 

5' AGAGTCAGAGCATAGTACGACATCTTAAAGAGAAA-3'

Codon no: 422 427
In family RBF59 (figure 3.32), there are two bilaterally affected individuals (I.1, II.1). SSCP analysis of exon 17 DNA from I.1 showed an abnormal banding pattern on the gel compared to the other samples (figure 3.27b, Section 3.2.4). Sequence analysis of exon 17 revealed a 1 bp, heterozygous deletion in the DNA from I.1 and II.1. Two 'A's are present at the 131st and 132nd nucleotides of exon 17 in the coding strand and one of these was found to be deleted from one allele in the DNA from I.1 and II.1 (figure 3.32a). Individual I.2, who is unaffected, has the normal sequence (figure 3.32a). Figure 3.32b shows the sequence of the opposite (non-coding) strand from the DNA of I.1. This sequence was obtained using the internal set of primers given in table 3.10 which amplify a shorter fragment from exon 17 and, when used for sequencing, produce a better sequence ladder (figure 3.32b). When the sequence around the deletion was examined, the deleted nucleotide was found to be part of an imperfect 5-7 bp repeat which is present three times in tandem (figure 3.32c). An upstream sequence, TGACAA, similar to the deletion consensus sequence of Krawczak and Cooper (1991) includes the deleted 'A' as its last base and 6 bp downstream of this deleted base, a TGATAA sequence is also present (figure 3.32c). There are two sequences (boxed in figure 3.32c), starting 20 bp upstream and 6 bp downstream, of the deleted nucleotide that are inverted repeats of which parts are also dyad symmetries of each other (T'Ang et al 1989). Exon 17, especially upstream sequences, is rich in Alu repeats (McGee et al 1989). The resulting frameshift leads to a premature stop codon at codon 546. This would be expected to result in the production of a truncated protein 545 amino acids long.
Figure 3.32 Sequence analysis of exon 17 in family RBF59.

(a) Sequence of the coding strand of exon 17 from the DNA of the family members of RBF59. Primers 5535 and 5536 (biotinylated) were used to amplify exon 17. The coding strand was sequenced from primer 5535. In one of the alleles, an 'A' nucleotide is deleted at the position indicated by the arrow in the DNA of I.1 and II.1. The sequence carrying the deletion is superimposed on the normal sequence giving double bands from that point. The sequence from the DNA of I.2 shows the normal exon 17 sequence.

(b) Sequence of the non-coding strand of exon 17 from the DNA of I.1 (left) and a normal individual (right). Internal primers 20877 (biotinylated) and 20876 were used to amplify exon 17. The non-coding strand was sequenced from 20876. The deletion point is indicated with an arrow.

(c) Nucleotide sequence of the coding strand of exon 17 around the deletion area. The deleted nucleotide (A) is boxed and the 5-7 bp imperfect direct repeat sequences are underlined. Nucleotides marked * vary between direct repeats. The boxed sequences are inverted repeats. Sequences (TGACAA and TGATAA) found to be similar to the deletion consensus sequence of Krawczak and Cooper (1991) are shown in italics.
Figure 3.32

(a) RBF59

(b) 1.1 NORMAL

(c) 5′- C[TTTGGATTTTACAAAGTGATCGAAAGTTTTATCA]AAGCAGAAGGCAA

CTTCAACAAGAAAATCAAACATTTGAGATGGAACATCGAAT -3′
Family RBF62 is one of the two families which remained uninformative with all the polymorphisms studied. The transmitting parent, I.1, has one affected (II.1) and one unaffected (II.2) child (figure 3.33a) and SSCP analysis of exon 20 showed an abnormal banding pattern in DNA from I.1 (figure 3.28, Section 3.2.4). Sequence analysis revealed a 1 bp heterozygous insertion 53 bp from the 5' end of exon 20. An extra 'A' was noted in the non-coding strand in the DNA of I.1 compared to the normal sequence (figure 3.33a). The same mutation was also identified in DNA from II.1 who is affected but not in the DNA from II.2 who is unaffected. Alteration of the reading frame results in the generation of a premature stop codon at codon 672. This mutation would, therefore, be expected to result in the production of a truncated protein 671 amino acid long. When the sequence around the insertion was examined, a TGAGCA sequence was found to flank the insertion downstream (figure 3.33b). The insertion occurs within a 4 bp direct repeat sequence and also destroys a Dde I site (CTGAG --> CTTGAG) (figure 3.33b). When a normal 350 bp exon 20 amplification product (primers 9438/14928, table 3.9) is digested with Dde I, fragments 205 and 145 bp long are generated. If the insertion mutation is present the Dde I site is not and an undigested fragment remains following enzyme treatment. This was found to be the case in Dde I digested DNA of I.1 and II.1 confirming the presence of the mutation, whereas the DNA from unaffected II.2 and the normal individual I.2 generated only fully digested products (figure 3.34).
Figure 3.33 Sequence analysis of exon 20 in family RBF62.

(a) Primers 9438 (biotinylated) and 14928 were used to amplify exon 20 and the non-coding strand was sequenced using 14928. The family pedigree is given on top. Non-coding strand sequence of exon 20 from the DNA of each family member is shown. An extra 'A' (arrow) is seen in the DNA sequence from I.1 and II.1 whereas it is not present in the DNA sequence of unaffected II.2 and normal I.2. In the DNA where the insertion is present, the sequence ladder of the mutant allele is superimposed on the normal, resulting in two bands at each position on the gel.

(b) Nucleotide sequence of the coding strand of exon 20 around the site of the insertion. The extra 'T' (arrow) is inserted within a 4 bp direct repeat sequence which is underlined. The Dde I site (CTGAG) which is destroyed by the insertion is highlighted. The TGAGCA sequence flanking the insertion is shown in italics.
Figure 3.33

(a)

(b)

5'-GCCTTCGCTCTCACCCCA-3'

216
Figure 3.34  Restriction enzyme analysis of the mutation in exon 20 from family RBF62 the pedigree of which is shown on top. A 350 bp exon 20 fragment was amplified using primers 9438/14928 from the DNA of RBF62 family members and a normal control. The amplification products were digested with Dde I. The 1 Kb DNA ladder is shown in the lanes marked "M" and the sizes are indicated at the right of the figure. The lane marked "U" contains an uncut control sample (350 bp) and the lane marked "C" contains a Dde I cut normal control sample (205 + 145 bp). Dde I digested DNA from affected I.1 and II.1 exhibited both full-length and digested fragments indicating the presence of the mutant allele. Dde I digested DNA from unaffected II.2 and normal I.2 contained only the fully digested fragments.
3.3.2 Mutations identified in families showing incomplete penetrance and mild expression of the Rb phenotype

RBF29

Linkage analysis for family RBF29 (figure 3.35) was described in Sections 3.1.5 and 3.1.7. In this family there are six affected individuals two of whom have spontaneously regressed tumours (I.1, I.3). There are also four unaffected mutant gene carriers (arrowed in figure 3.35). Extra bands were identified in exon 20 in the DNA from individual II.4 (unaffected gene carrier) in SSCP gels (figure 3.28, Section 3.2.4) and not in any of the other (20 out of 27) exons analysed. Sequence analysis of exon 20 from II.4 revealed a heterozygous C—>T transition in the coding strand (G—>A in the non-coding strand, figure 3.35), 21 bases from the 5' end of exon 20. This mutation converts codon 661 from an arginine (CGG) to a tryptophan (TGG) codon. The same mutation was identified in the DNA of all affected members and unaffected gene carriers (I.1, I.3, II.2, II.4, II.5, II.7, III.1, III.2, III.3, III.4) but not in the unaffected members (I.2, II.1, II.3, II.8, III.5, III.6) of the family. The exon 20 sequence from the DNA of II.4, I.1, II.5, II.7, II.8, III.3 and III.5 are shown in figure 3.35. The C—>T transition occurs within a C_pG dinucleotide pair and affects a Hpa II/Msp I restriction site in exon 20 (CCGG —> CTGG). This produces a simple means to identify mutation carriers. Msp I digestion of the 350 bp PCR product results in two fragments of almost equal size (177 and 173 bp). A smaller PCR product (269 bp) was generated, therefore, using an internal primer, 17295 (table 3.10), in conjunction with primer 14928 so that two clearly distinguished fragments (177 and 92 bp) are produced upon Msp I digestion. The presence of an undigested fragment is indicative of a mutant gene carrier. In individuals heterozygous for the C—>T mutation in RBF29, the normal digested bands (177 and 92 bp) and an undigested 269 bp fragment are expected. DNA from all individuals of this family, from whom blood samples could be obtained, were analysed for the presence of this mutation using Msp I digestion. The undigested 269 bp band was observed in the DNA of all affected members and unaffected mutant gene carriers whereas only the two smaller normal bands were seen in individuals from this family identified not to be mutant gene carriers (figure 3.36). In the gel shown in figure 3.36, the DNA of individual III.5 was obtained from two
sources. DNA in the lane marked CV was obtained from II.3 for pre-natal screening from chorionic villus tissue sampled after 10 weeks of pregnancy. The other DNA sample (in lane marked B) was obtained from cord blood samples taken from III.5 at birth. Sequence analysis of exon 20 from CV DNA was found to be free of the mutation described above (figure 3.35) and confirms the original prediction made with linkage analysis (Section 3.1.5, Onadim et al 1992a). In addition, DNA from both sources was analysed using Msp I digestion (figure 3.36) and showed two fully digested bands 177 and 92 bp long, further confirming that this individual did not carry the mutant gene. Msp I digestion of exon 20 DNA from 34 other unrelated Rb patients and 38 unrelated healthy individuals resulted in full digestion excluding the possibility that this DNA sequence change represented a polymorphic site.
Figure 3.35 Sequence analysis of exon 20 in RBF29. The family pedigree is shown on top. Incomplete penetrance cases are indicated with an arrow. Sequence from the non-coding strand of exon 20 (from primer 14928) showing the heterozygous G→A transition in II.4 (left) is compared with the same sequence from a normal individual (right). Sequence obtained from the DNA of individuals I.1, II.5, II.7, II.8, III.3 and III.5 are also shown (next page). The mutation (arrow) is present in the DNA of affected individuals I.1, II.5, III.3 and also in the DNA of II.7 who is an unaffected gene carrier but not present in the DNA from unaffected individuals, II.8 and III.5. The DNA of III.5 was obtained from CV tissue sampled from II.3 after 10 weeks of pregnancy.
Figure 3.36  Restriction enzyme analysis of the mutation in exon 20 from family RBF29, the pedigree of which is shown on top. In the pedigree incomplete penetrance cases are indicated with an arrow. DNA was amplified using primers 17295/14928 which generated a 269 bp fragment. Msp I digestion of this PCR product results in two fragments, 177 bp and 92 bp long (lanes C, II.1, I.2, II.3, CV, B, II.8). Individuals carrying a C-->T (coding strand) mutation, which destroys the Msp I site, display the uncut 269 bp fragment (lanes III.1, III.2, II.2, I.1, III.3, II.4, II.5, II.7). M=1 Kb molecular weight marker (Gibco-BRL); U=uncut control 269 bp fragment; C=control sample known to be homozygous for the 177 bp and 92 bp fragments; CV=chorionic villus DNA obtained from II.3; B=cord blood DNA from III.5.
Family RBF18 was previously described in Section 3.1.7. Three of the eight affected individuals in this family (II.2, III.3, IV.1) had unilateral disease (figure 3.37). Four generations of males (I.1, II.4, III.3, IV.1) have a 'mild' form of the disease with spontaneously regressed tumours in one or both eyes. SSCP analysis of exon 20 showed abnormal banding patterns in the DNA from patient III.3 compared to other samples (figure 3.28, Section 3.2.4). Sequence analysis of exon 20 from III.3 revealed a heterozygous G→T transversion in the coding strand (C→A in the non-coding strand, figure 3.37), 63 bases from the 5' end of the exon. The same mutation was identified in the DNAs of individuals III.2 and IV.1 who were known to carry the Rb predisposition gene. III.4, who is an unaffected individual, had the normal sequence (figure 3.37). DNA from other family members was not available for analysis. A sequence (TGAGCA) which is similar to the putative DNA polymerase α arrest sites is present 4 bp upstream of the transversion (figure 3.38). The G→T transversion converts codon 675 from a glutamic acid (GAA) to a stop codon (TAA). However, 3 bp downstream of this mutation lies a TAG sequence (figure 3.38) which is compatible with the consensus splice acceptor sequence (Mount 1982). The G→T transversion also removes an AG doublet (converting it to AT) which ordinarily would have prevented the downstream TAG sequence becoming a splice acceptor site. In addition, this transversion increases the pyrimidine:purine ratio in the region immediately preceding the TAG site thereby enhancing its potential to be a cryptic splice site (Hoshijima et al 1991). A branch point sequence exists (figure 3.38) upstream of this cryptic site which, although not as good as the real branch point in intron 19, would nevertheless be adequate if this cryptic site is activated (T. Maniatis, personal communication). If this cryptic site was used, the reading frame would be intact and, as a result, the first 23 amino acids encoded by exon 20, codons 654 to 676, would be lost. Such a deletion would disrupt the leucine zipper motif in exon 20 with 3 out of the 4 leucines being deleted.
Figure 3.37 Sequence analysis of exon 20 in family RBF18. The family pedigree is shown on top. Sequence from the non-coding strand of exon 20 (from primer 14928) showing the heterozygous C-->A transversion in III.3 (left) is compared with the same sequence from a normal individual (right). Sequence of exon 20 from III.2, III.4 and IV.1 is also shown. The mutation is present in the DNA of III.2 and IV.1 (arrow) who are affected. The DNA from the unaffected individual III.4 exhibits the normal sequence.
(a)

Intron 19 ← 5′-aaaaatgactaattttttctttatceccacag|TG TAT CGG

CTA GCC TAT CTC CGG CTA AAT ACA CTT TGT GAA CGC

CTT CTG TCT CAC CCA GAA T|TA|GAA CAT ATC ATC

↑

TGG ACC CTT TTC CAG CAC ACC CTG CAG AAT GAG TAT

GAA CTC ATG AGA GAC AGG CAT TTG GAC CAA|gta-3′→ Intron 20

(b)

Consensus splice sequence

---------- (T) ≥ 11 ---------- N C A G | G

(c)

Real splice acceptor

T T C T T A T T C C C A C A G | T G

Normal Cryptic splice

A G C A C C C A G A A T T A G | A A

Mutated Acceptor

A G C A C C C A T A A T T A G | A A

Figure 3.38 Nucleotide sequence (a) of the coding strand of exon 20 (upper case) and its flanking intron sequences (lower case). A putative DNA polymerase α arrest site (TGAGCA), 4 bp upstream of the transversion, is shown in italics. The normal splice acceptor site in intron 19 is underlined as is the cryptic splice site in the exon sequence containing the G→T transversion. The absolutely required TAG sequence of the splice acceptor (described in (b)) is enclosed in the box. A potential branch site (TTGTGAAC) located 24 bp upstream of cryptic splice site is highlighted. In (c) the real splice acceptor site from intron 19 is compared with the cryptic splice acceptor with and without the G→T mutation.
In summary, therefore, through SSCP analysis and subsequent sequencing, mutations were detected in 6 out of 15 DNAs (40%) analysed. However, because of constraints on time during this project, it was not possible to analyse all exons in all of those patients chosen for analysis. For example, only half the total number of exons (14/27) were analysed in 8 of these patients and in the other 7 an additional 6 exons were studied. Exons 1, 2, 8, 11, 25, 26, 27 and the promoter region of the RB1 gene were not analysed in any of the DNAs. False positives were encountered in two cases where although band shifts were observed on the SSCP gels, no mutation was found upon sequencing the relevant area. The shifts observed in these DNAs were very slight, however, and could easily be artefacts of the gel. Although time constraints did not allow the repeat of these particular gels, in a previous occasion when a gel showing a shift in one of the DNAs was repeated, the same shift was not observed and no mutation was found upon sequencing. It is quite possible that false negatives also occurred as some mutations might not have produced a detectable mobility difference to others under the conditions used.

2 deletions, 2 insertions and 2 point mutations were identified in the DNAs sequenced. These mutations are summarised in table 3.11. The deletions and insertions were found to have occurred in areas of the RB1 gene containing direct and inverted repeat sequences and/or dyad symmetrical elements. All the deletions and insertions give rise to frameshifts leading to premature stop codons, however, the deletion in exon 3 and the transversion in exon 20 can potentially affect the use of nearby splice junction sequences. Four of the mutations (the deletions and insertions) were from bilateral Rb families (RBF64, RBF58, RBF59, RBF62) and the other two (the point mutations) from 'mild phenotype' families (RBF29, RBF18).
<table>
<thead>
<tr>
<th>Family</th>
<th>Mutation</th>
<th>Exon</th>
<th>Location in exon (bp)</th>
<th>Effect</th>
<th>Codon</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF 64</td>
<td>2 bp deletion</td>
<td>3</td>
<td>106/107</td>
<td>Frameshift→Stop codon</td>
<td>129</td>
</tr>
<tr>
<td>RBF 58</td>
<td>1 bp insertion</td>
<td>13</td>
<td>50</td>
<td>Frameshift→Stop codon</td>
<td>427</td>
</tr>
<tr>
<td>RBF 59</td>
<td>1 bp deletion</td>
<td>17</td>
<td>132</td>
<td>Frameshift→Stop codon</td>
<td>546</td>
</tr>
<tr>
<td>RBF 62</td>
<td>1bp insertion</td>
<td>20</td>
<td>53</td>
<td>Frameshift→Stop codon</td>
<td>672</td>
</tr>
<tr>
<td>RBF 29</td>
<td>C→T transition</td>
<td>20</td>
<td>21</td>
<td>Arg→Trp</td>
<td>661</td>
</tr>
<tr>
<td>RBF 18</td>
<td>G→T transversion</td>
<td>20</td>
<td>63</td>
<td>Glu→Stop</td>
<td>675</td>
</tr>
</tbody>
</table>

Table 3.11 Summary of mutations identified in the RB1 gene.
4.0 DISCUSSION
4.1 GENETIC SCREENING IN RETINOBLASTOMA FAMILIES

Prior to the commencement of this thesis there was very little to offer Rb patients in terms of pre-natal screening and identification of mutant gene carriers and genetic counselling was limited to telling the patient what their relative risk of having affected children might be. Although these figures have changed over the past several years, and differ between studies, the best estimates given to families were essentially as described by Draper et al (1992) recently. Thus, in the absence of genetic linkage/mutation analysis the risk to sibs, of inheriting the mutant gene is 50% and because incomplete penetrance occurs in 10% of these individuals, the actual risk is 45%. This risk refers to "old" germline mutations. According to Knudson's hypothesis all bilaterally affected individuals are considered "new" germline mutations and their children have the same risk but it is not definitely known whether those individuals have actual mutations since, with the few exceptions so far whose mutations have been identified, their having affected children is the only sure way of establishing this fact. The problem comes with unilateral sporadic cases of which 2.5% may have new germline mutations. For those that do the risk for tumorigenesis to their children is 45%; for those that do not the risk is negligible. Empirically, families with unilaterally affected individuals are given a 30% risk that their children will develop tumours because of the possibility of incomplete penetrance. The risk to the sibs of bilaterally affected sporadic cases being affected is 1.6% and to unilaterally affected sporadic cases 0.6%. Clearly it was important to develop a method to determine who was carrying a mutation and who was not, if only to relieve the considerable burden of the ophthalmologists who were screening all relatives of affected children. Early detection of Rb is essential as it significantly improves the prospects of therapy to preserve vision and ensure survival. By regular screening, during the first five years of their lives, of all members of a family in which there has been an affected individual, early detection of tumours is usually guarantied although, depending on the exact location of tumour(s) in the eye, a favourable outcome is not. Screening involves regular hospitalisation and full ophthalmological examination under anaesthetic. The vast majority of cases are identified during the first two years but of all the people screened, less than 10% ultimately develop Rb. This form of 'blanket' screening means that resources are not always focused on those who need them. This, not only wastes resources but also this form of screening is unable to identify those 10% of gene carriers who do not
develop tumours due to incomplete penetrance. These individuals are normally identified as such only when they produce affected children. An additional problem, which can not be addressed with this form of screening, is the emotional issue of the dilemma of the Rb patients who would like to have children but who are deterred because of the recurrence risk. Sometimes included in this group are close relatives of Rb patients who have no way of knowing whether they are gene carriers or not. For this reason many Rb patients have opted either not to marry, not to have children and, in some less well informed cases, to have irreversible sterilisation (M Jay, personal communication).

The counselling given to all of these patients has now changed with the development of a predictive genetic test which identifies individuals who carry the predisposition gene. Only those patients are offered ophthalmological examinations for their children which reduces or eliminates the screening load for the ophthalmologists. Such tests could be performed pre-natally or post-natally and could also be applied to all unaffected relatives of Rb patients to identify possible incomplete penetrance cases who have not been identified as such and who can yet produce affected children and are themselves at considerable risk of developing second cancers (Cowell and Onadim 1990). The unaffected subjects shown not to be carriers of the predisposition allele may be reassured that they will not have affected children or develop the tumour themselves. This eliminates the considerable anxiety felt by those who, although reassured that risks are low, nevertheless must subject their children to screening with no guarantee of the outcome.

As explained in the introduction Section 1.4.6.2, every effort has been made to use the knowledge gained through research into this gene in clinical applications. None of the earlier tests developed, namely karyotype analysis, ESD protein and RFLP analysis, were entirely useful for extensive genetic screening for Rb mutant gene carrier status. The new era dawned in 1986 when the Rb predisposition gene was cloned. It was not possible to use the RB1 cDNA itself in linkage analysis, however, and polymorphic DNA probes from within the genomic sequence were needed to make extensive family studies possible. Despite the application of mutation detection to Rb patients there is still an important role for linkage analysis since the sequencing programme is very time consuming and the backlog of patients large.
After the cloning of the Rb gene, the first polymorphism in Rb was reported by Horsthemke et al (1987b). This was a Hind III polymorphism detected by a genomic clone, H3-8 (Lalande et al 1984), which led to the isolation of the RB1 cDNA, and was used by Greger et al (1988) together with other flanking chromosome 13 markers in linkage analysis in 6 Rb families. Only 2 of the 6 families were informative, however, for the H3-8 polymorphism and phase was established using flanking markers in the remaining 4 families. The need for more polymorphic markers from within the gene itself was obvious. Such markers were later isolated (Wiggs et al 1988, Bookstein et al 1988) and formed the basis of linkage studies in the analysis of Rb families worldwide.

To perform any meaningful population studies it was necessary to have sufficient numbers of families available. These had previously been documented, mainly through the considerable effort by Dr Marcelle Jay in the compilation of family pedigrees and the ICRF Oncology group in ICH which had collected DNA samples prospectively from these patients. The application of the DNA polymorphisms to this analysis was firstly designed to establish the optimal experimental conditions for linkage analysis using the intragenic DNA polymorphisms and to evaluate their usefulness in genetic screening of Rb families in UK. This was not altogether straightforward because these analyses had their attendant problems. Among these were the difficulties observed in comparison of DNA fragment sizes between gels for the RS2.0 polymorphism, the problem of incomplete digestion with Kpn I and Tth III I (Section 3.1.1) and optimisation of PCR conditions including the design of the best primers to avoid non-specific amplification (Section 3.1.1, 3.2.2). During the course of this project, linkage analysis was performed in 57 Rb families and it was found that 79% was informative using RFLP probes alone (Section 3.1.1) and that this figure could be improved to 95% using other intragenic polymorphisms (Section 3.1.2). These figures compared favourably with others obtained from the North American and the European populations.

In the first of the studies using RFLP probes Wiggs et al (1988) was able to make predictions in 19 of the 20 (95%) Rb families of mixed North American descent. In 18 of these families, they demonstrated a consistent association of marker RFLPs with the mutation predisposing to Rb. In the 19th family, there was a possibility of lack of segregation of the RFLPs and the site of the mutation predisposing to Rb in a key member. However, there was also
uncertainty about the clinical diagnosis of the retinal lesion in this individual. The most likely explanation was that the child in question did not have Rb. In 3 (16%) of the families, they also detected hereditary deletions involving the Rb locus and identified 2 incomplete penetrance cases. In a similar study, Scheffer *et al.* (1989) analysed 19 European families, 15 (79%) of which were informative with the intragenic markers. They identified an inherited deletion of RB1 in one (5%) family and an unaffected gene carrier in 2 families. Goddard *et al.* (1990), on the other hand, studied 7 North American Rb families and found 6 (86%) to be informative for intragenic probes. 4 of these families had previously been analysed using flanking markers by others (Cavenee *et al.* 1985, Greger *et al.* 1988). In one of the families analysed, one individual previously diagnosed as having Rb did not carry the haplotype associated with Rb in the rest of the family. This was a case similar to the one observed by Wiggs *et al.* (1988) and misdiagnosis because of the presence of Rb-like retinal lesions and a strong family history of Rb was considered to be the most likely explanation. In both of these cases alternative explanations included a new and independent mutation in these individuals; multiple recombination (at least two must have occurred to explain the linkage data in these families) and the involvement of a separate locus all of which were highly unlikely. The frequency of a new independent mutation (1.5 x 10^-5) is quite low, multiple recombination events are very rare and there is no evidence to suggest the existence of a second gene involved in Rb genesis. Misdiagnosis, however, may not be so uncommon as conditions such as Coat's disease can be difficult to differentiate from Rb in the absence of pathological analysis. Especially when there is previous family history of Rb, any Rb-like retinal lesions observed will be assumed to represent retinoma or spontaneous regression of Rb.

In the 3 studies described above 40 out of 46 (87%) families studied were informative for at least one of the intragenic probes used. Also, a direct detection of deletions was shown by the use of these intragenic probes in 4 of 46 (8.7%) pedigrees analysed.

The linkage analysis carried out in this project, using intragenic probes in 57 Rb families, is the largest of its kind. In the studies described above only the RFLP probes (RS2.0, M1.8, PR0.6, HS0.5 and R0.6) were used and a smaller number of families were studied. The total % informativity obtained in these studies (87%) is only slightly higher than the one obtained in this

233
project (79%) using the RFLP probes which is identical to the study by Scheffer et al (1989) although the sample (19 families) was smaller. The two North American studies (Wiggs et al 1988, Goddard et al 1990) yielded 92.6% informativity for the 27 families studied. The difference in these values may reflect a difference between the North American and the European populations. In view of the similar allele frequencies (apart from the RS2.0 alleles) obtained from the North American and the UK populations, such a difference is surprising, however, and alternatively may simply be due to the difference in the numbers of families studied and/or different number of probes used. In this study for example the probes HS0.5 and PR0.6 were not used extensively and more extensive use may have increased the % informativity. The number of families studied is also important because as more families were recruited the % informativity obtained for some of the probes gradually declined from quite high initial values before stabilising as larger numbers of families were included. The difference in frequencies of the RS2.0 alleles between the UK and the North American populations is most likely due to the difficulty in comparing allele sizes between gels as explained in Section 3.1.1. For simplicity, it might be better to number the alleles for the RS2.0 polymorphism in future, as done with RB1.20 polymorphism, instead of estimating their size on each gel.

Although the RB1.3 polymorphism was less helpful, the characterisation of the RB1.20 polymorphism in particular made a big difference in the ability to screen patients. As a result the % of informative families using these polymorphisms increased to 95%. The only other comparable study is that of Yandell and Dryja (1989) who initially identified these polymorphisms by enzymatic amplification and sequencing of the RB1 gene. In their studies, although a large number of unrelated individuals were analysed, only a small number (4-9) of Rb families were examined, all of which were informative. The allele frequencies for RB1.20 were not determined in either population as numerous alleles were found. In the present study approximately 87% of the 30 families analysed using RB1.20 were informative and, of 68 unrelated individuals, 71% were heterozygous at this locus. These 30 families, however, were not chosen randomly from the 57 in the series, rather they tended to represent families uninformative with the other polymorphisms. In the North American population Yandell and Dryja (1989) found 94% of patients were heterozygous at this locus, although the size of the sample from which this figure was drawn was not specified. Ignoring the apparent
variation, the general finding is that the RB1.20 polymorphism on its own is as good as the others put together, although the techniques used to visualise it are more precise and involved.

Polymorphism analysis based on single base pair changes that do not affect restriction enzyme sites are inherently more difficult to use for technical reasons. For the RB1.3 polymorphism, Yandell and Dryja (1989) studied 41 unrelated individuals whereas only 19 were studied in this project. The figures obtained for the frequency of the rare allele was comparable with ours, however, being 0.27 and 0.21 respectively. In practise, constructing the sequencing gels is more involved and more prone to failure and so would only be useful where other polymorphisms were unsuccessful. Using RB1.3, Yandell and Dryja (1989) found all four of the families they studied informative. In this project only 1 out of 9 (11%) families studied were informative.

Although approximately 95% of families can be offered genetic screening it is possible that some of the remaining 5% may still be informative for other intragenic RB1 polymorphisms which have been discovered relatively recently but which have not yet been used generally in family studies. For example, single base pair polymorphisms exist in intron 3, intron 25 (Yandell and Dryja 1989) and intron 5 (Blanquet et al 1992) of the RB1 gene, the rare allele frequencies being 0.05, 0.15 and 0.06 respectively. Although rare we owe it to the uninformative families to test them at these loci. The intron 5 polymorphism is easier to analyse than others as it changes the cleavage site for the restriction endonuclease Mae III. Another new polymorphism was described in Section 3.2.2 involving an A1.3 polymorphism in intron 17. Estimated allele frequencies for this polymorphism are 0.74, 0.008 and 0.243 and the heterozygous frequency was found to be 40% (DW Yandell, personal communication). The A1.3 polymorphism, therefore, is likely to be more useful than the other three described above.

Chromosome deletions have proved to be an important feature of mutation analysis in RB1. Southern blotting using DNA from tumours readily identified large deletions which were often homozygous but analysis of constitutional cells relied on quantification which was less successful. If the deletion lies in the region identified by the polymorphic probe, however, it is possible to identify deletions in family studies. Two such deletions (RBF34,
Section 3.1.5; RBF31, Section 3.1.7) were found in our 57 families (3.5%). Others (Wiggs et al 1988, Scheffer et al 1989, Goddard et al 1990) found 4 deletions from 46 families (8.7%). The slightly higher frequency is probably due to a reporting bias since not all families analysed were included. There is likely to be a referral bias also. For example, using ESD analysis, Cowell et al (1989) found 15 deletions in 500 Rb patients, Turleau et al (1986) found 7 deletions in 66 Rb patients and Dryja et al (1983) found no deletions among 51 Rb patients. Referral bias is also apparent in the number of unaffected gene carriers reported by the North American and Dutch groups. No unaffected gene carrier was reported in the 7 families studied by Goddard et al (1990) whereas 4 unaffected gene carriers were detected in the 39 families studied by Wiggs et al (1988) and Scheffer et al (1989). In contrast, we reported 10 such cases from 57 families.

At the start of this thesis the intragenic probes had only just been reported and whilst evaluating the usefulness of these intragenic probes, it was important at the outset to calculate the lod score values to assess the recombination frequency. This is clearly important for risk assessment. In this project lod scores were calculated using data from 21 Rb families (Onadim et al 1990). Wiggs et al (1988), Scheffer et al (1989) and Goddard et al (1990) also calculated lod scores for the families they analysed. The cumulative lod score from these four studies is 27.091 at θ=0, assuming a penetrance of 90%. Given this already very high score, it was deemed unnecessary to continue adding to this value following analysis of every Rb family recruited in this project. To date there have been no cases of recombination in any of the families reported so far (Wiggs et al 1988, Scheffer et al 1989, Goddard et al 1990, Onadim et al 1990, 1991, Onadim and Cowell 1991, Onadim et al 1992a) which has surveyed over 200 meioses. In addition, over 100 extra meiosis have been analysed but not published bringing the total to well over 300. The RB1 gene spans approximately 180 kb of genomic DNA and, assuming the generally accepted recombination frequency of 1 cross-over per 10^6 base pairs, the theoretical chances of recombination occurring within the RB1 gene is 0.18 or 1:550. Unless the causative mutation is identified, therefore, recombination between marker and mutation can not be totally excluded. However, as mutations seem to occur equally along the length of the gene (Yandell et al 1989, Dunn et al 1989, Hogg et al 1992, Onadim et al 1992b), the possibility of a predictive error is decreased if patients are informative at different loci that cover most of the gene. For example, the
M1.8 probe (intron 1) and the RB1.20 polymorphism (intron 20) cover between them 75% of the RB1 gene. It is important, therefore, where the causative mutation has not been found, to screen families with as many probes as possible to decrease the possibility of a predictive error. This has been the practise throughout this project when performing pre- and post-natal screenings for Rb families. Thus, of the 11 screenings (in 8 families) carried out only 2 predictions was based on the use of a single probe because the individuals concerned were uninformative at the remaining loci.

Although many Rb families were analysed using the RB1 RFLP probes, there have been few reports of the successful application of these probes in pre-natal and peri-natal screening programmes. The very first pre-natal screening in Rb was actually carried out in 1986, before the cloning of the RB1, by Cavenee et al (1986) using flanking DNA markers. This study met with limited success, however, due to recombination events between the markers and the Rb locus. Pre-natal screening was performed in five Rb families and the likelihood of Rb was predicted in two cases and freedom from disease in three. Two of the cases showed evidence of meiotic recombination and the predictive accuracy in one other was only 70% since only loci distal to the Rb locus were informative. Predictions were also made in the past using deletion analysis. Pre-symptomatic prediction of Rb development was made in two patients (J K Cowell, unpublished data) who were carriers of chromosome 13 deletions and who were identified using ESD measurements in the series described by Cowell et al (1989). Both patients eventually developed Rb before 12 months, although not all such cases develop tumours (Wilson et al 1987, Cowell et al 1988).

The first report of a pre-natal screening carried out using intragenic probes came from Mitchell et al (1988). They used DNA obtained from chorionic villus and predicted that the child would be unaffected. This child is now (12/92) 50 months old and, following repeated ophthalmological examination, shows no evidence of a tumour. A second pre-natal test carried out by Dr CD Mitchell was actually reported in Onadim et al (1990). The family involved (RBF01) was described in Section 3.1.7. In this case, although the prediction was "would be affected", the proband did not develop Rb and upon exclusion of recombination (Section 3.1.7), it must be concluded that this case represents incomplete penetrance of the RB1 gene.
In this project, five pre-natal screenings were carried out in five Rb families (Section 3.1.5). In two of these cases, the probands were found to carry the predisposition allele. In the first case, the family involved (RBF41) decided against termination of pregnancy and after birth, the presence of a single tumour was confirmed in the new born child. The second family involved (RBF58) opted for termination of the pregnancy.

There are ethical issues to consider in the pre-natal prediction of a disease such as Rb, which is rarely lethal these days and has no effect on mental ability. Although the vision of some patients who are treated early for single tumours may not be greatly affected, in some cases enucleation (which may be bilateral) is necessary. Another factor that needs to be considered and which also constitutes an argument against using radiation treatment in hereditary cases, is the involvement of second tumours. 5-15% of Rb patients are expected to develop second tumours, mostly osteosarcoma and soft tissue sarcomas, in adolescence (Abramson et al 1984, Sanders et al 1989) and it is now believed that most of the Rb patients would develop other, more common, cancers (e.g. gastrointestinal, lung, breast) when they approach the usual age of onset of these neoplasms (Roarty et al 1988).

It has been observed that decision to terminate pregnancy in cases of fetuses expected to be affected depends very much on the past experience of the families with Rb and other special family circumstances. From the three families (RBF01, RBF41, RBF58) which received a positive result for the pre-natal screenings, only one (RBF58) opted for termination of pregnancy. In this family, the affected members had bilateral disease and the first child had actually died of a rhabdomyosarcoma which developed after she was treated for Rb (Section 3.1.4). The second child was not affected and the family decided for a third pregnancy after they were informed that pre-natal screening would be possible in their case (Section 3.1.4). It was their belief that, in the event of a fetus inheriting the mutant allele, termination of pregnancy was acceptable. In family RBF01, termination was not acceptable, because an unaffected previous child could have been the result of non-penetrance or a recombination between the probe and the predisposing mutation (Section 3.1.7). Family RBF41, on the other hand, decided against termination in spite of the fact that many members had bilateral disease and there had been many cancer deaths in the family (Section 3.1.7). This decision was due to special circumstances in this pregnancy. The mother was
42 years of age and saw this pregnancy as her last chance of having a child and Rb, being curable, was not strong enough reason for them to terminate this pregnancy. In fact the reason for pre-natal screening being carried out rather than post-natal was the fact that chorionic villus sampling was being performed anyway to obtain DNA for Downs Syndrome testing. If it was not for this test, the Rb screening would have been performed after birth using DNA obtained from umbilical cord blood.

The pre-natal test are usually performed using DNA obtained from chorionic villus (CV) sample obtained after 10 weeks of pregnancy (Williamson et al 1981). Earlier sampling, although possible and had been performed in the past, was observed on occasion to cause limb defects in the fetuses (Dr K Nikolaides, personal communication, Firth et al 1991). Chorionic villus sampling, however, is still the method of choice because earlier sampling is possible, although it carries a higher risk for miscarriage than amniotic fluid sampling (MRC working party on the evaluation of chorion villus sampling 1991, Hajianpour et al 1991). In amniocentesis, DNA is obtained from cultured cells of amniotic fluid which may take 2-3 weeks or fail. Post-natal tests are carried out using DNA obtained from either cord blood or whole blood obtained early in the child's life. Sometimes it is difficult to obtain large volumes of blood from newborns. This is not the case with cord blood samples. In cases where families are informative for RB1.20, M1.8 or PRO.6, smaller volumes of blood is not a problem, since these polymorphisms can be identified within 24 hours using PCR which need minute amounts of DNA. However, one of the important probes used in Rb screening, RS2.0, still requires reasonable amounts of DNA (5-10 μg) and the results are only available after 4-7 days. This problem are expected to be overcome in the near future as the sequencing of 180 kb of the RB1 gene has recently been completed (D W Yandell, personal communication) and primers for all the other polymorphisms can now be designed for PCR.

The use of PCR improves the possibility of genetic screening in more ways than one. Accurate genetic screening in family RBF58, for example, was only made possible by the application of PCR to linkage analysis. Linkage phase was established in this family through PCR amplification of polymorphic regions from the paraffin-embedded tumour section DNA from the deceased child (Section 3.1.4). This kind of application has great potential as linkage phase can potentially be established in any family whose
key members are deceased if paraffin blocks have been saved. During this thesis I have tried to analyse DNA from paraffin sections of tumours however, which presented specific problems. The block must be trimmed to avoid contamination by normal cells which, in turn, must be carried out by experienced ophthalmologists. Since these tumours may be necrotic in the centre this can leave very little material for DNA isolation. In my experience either the isolation works and PCR is possible for all exons or the block is of no use at all. For some sporadic cases of Rb this will deny them mutation screening. The philosophy behind screening tumour cells is quite clear. Mutations in RB1 must be present and, if they can be found, it is a simple task to analyse constitutional DNA from the same patient at the mutant locus. If the DNA shows the same mutation then the patient is a carrier of a germ line mutation. If not, then this represents a truly sporadic case and future screening of children is unnecessary. Approximately 70% of tumours (Cavenee et al 1983, Zhu et al 1989) are homozygous for the causative mutation which makes the analysis easier. In fact once the homozygous mutation has been found, despite screening all of the remaining exons, no other abnormalities have been found in tumours (A Hogg, personal communication). When the mutation is heterozygous the same rule applies and the sites of both mutations in these cases must be analysed in constitutional cells. In practise, however, hereditary tumours are rarely removed since post-natal screening usually allows successful treatment. In fact, so good, is the treatment these days that many sporadic tumours are not removed either (Section 1.3.6). This is not the case in the past, however, and identifying mutations from archival material from patients now approaching childbearing age has got to be considered beneficial. Screening constitutional cells from sporadic cases is not only frustrating if no mutations are found but the possibility that a mutation has been overlooked cannot be excluded in these cases.

The versatility of PCR comes from the fact that it is possible to amplify DNA from minute amounts of starting material. It is, therefore, theoretically possible to amplify DNA from a single cell. If the optimum conditions for such amplifications are worked out for the RB1 gene, it might obviate the need for pre-natal screenings in the future. A query has recently been received from an Rb family who is interested in such a possibility. In this family the bilaterally affected mother gave birth to triplets one of whom was born dead. Bilateral tumours were observed in the other two children. Upon
linkage analysis, this family was found to be informative for the RS2.0 probe and it was shown that all three children inherited the same RS2.0 allele from the affected mother. The family is interested in in vitro fertilisation if one of the cells from a fertilised egg can be removed and analysed for the presence of mutant allele. With such analysis it is then possible to implant a fertilised egg that does not carry the mutant allele. This kind of implantation is already being tried for other inherited diseases such as cystic fibrosis. Use of PCR technology in such applications requires utmost care, however, as PCR is prone to error due to the very properties that makes it versatile. Contamination from other sources or replication errors could be a problem especially when the template DNA is present in much less than the usual concentration of genomic DNA (more than $10^4$ molecules per reaction) used in PCR (Krawczak et al 1989). In some cases contamination could be due to carryover of product from previous amplification reactions into unamplified samples. To avoid this type of contamination, the original source DNAs and samples to be amplified must be kept away from the PCR products, reagents must be premixed before dividing them into aliquots and DNA must always be added last. Erroneous amplification is also possible with DNA polymerase enzyme incorporating the wrong base. Such rare errors, between $2 \times 10^{-4}$ to $<1 \times 10^{-5}$ error/base pair incorporated for Taq polymerase (Eckert and Kunkel 1991), nevertheless become significant in PCR applications involving the characterisation of small numbers of DNA molecules. These problems can be largely overcome by keeping the number of PCR cycles to minimum, by optimising the PCR conditions to maximise DNA polymerase fidelity (Ling et al 1991, Eckert and Kunkel 1991), by setting up appropriate controls and by repeating the amplification as long as the source DNA is not contaminated. The possibility of erroneous amplification should always be born in mind, however, when using PCR technology.

All future applications to genetic screening depend on the predictive accuracy of tests involved. Most of the cases involving predictions made during the course of this project have now been followed sufficient time to be sure that the prediction was accurate (Section 3.1.5). Given the low chance of intragenic recombination in the RB1 gene, it must, therefore, be concluded that it is unnecessary to repeatedly screen patients shown not to have inherited the predisposing mutation following linkage analysis. The ophthalmologists also agree with this conclusion and it has been decided that, in St Bartholomew's and Moorfields Eye hospitals at least, examination under
anaesthesia (EAU) will discontinue in the future for patients shown not to have inherited the predisposing mutation (J Hungerford, personal communication). Techniques that are used in routine service laboratories should ideally be rapid, inexpensive, safe and amenable to automation. The use of $^{32}$P-labelled probes presents a potential safety problem and also means that the probe has short shelf life ($t_{1/2}=14.3$ days). A non-radioactively labelled probe would provide a satisfactory solution. Cross-linking with enzymes, biotinylation or fluorescence labelling may be used instead. Devising ways of non-radioactive detection is still a consideration for mutation detection methods, however (see later). As far as the detection of polymorphisms in the RB1 gene is concerned, Southern blotting technique and the use of labelled probes will gradually become redundant as the RB1 gene is sequenced in its entirety and PCR-based detection will soon become possible for all polymorphisms, although this approach is not without its problems. When restriction enzyme digestion is carried out on PCR products, digestion may be taking place in less than ideal conditions. For example, the buffer of the PCR reaction may be suboptimal resulting in partial digestion which give "false" heterozygotes. To avoid such a problem, either the concentration of buffer ingredients are suitably adjusted (if the PCR and enzyme buffers are compatible and only the concentrations of the ingredients are different) or it is necessary to clean (phenol/chloroform extraction and reprecipitation) the PCR product which adds to the work load.

4.1.1 Manifestation of the Rb genotype

St Bartholomew's and Moorfields Eye Hospitals, from where most of the patients analysed in this series were received, are major referral centres for Rb in the UK. Because the most serious cases or cases with unusual inheritance patterns tend to be preferentially referred to these centres, there exists a referral bias in the data presented in this thesis. Also in this project only the hereditary Rb cases were studied. Therefore, it is not possible to arrive at general conclusions about the incidence and phenotypic expression of Rb in the UK using this data. Some parts of the data, however, can be compared with data from other studies of hereditary Rb, although the referral bias should be born in mind in such comparisons.
Although most (75%) of the affected individuals in this series were bilaterally affected (and not surprisingly so because the series contains only hereditary cases) the expression of the Rb phenotype in the remainder ranged from asymptomatic gene carriers, regressed tumours through unifocal to multifocal lesions in unilateral cases. 5.5% of all gene carriers were found to be unaffected in this series. In the inherited form of Rb it is believed that non-penetrance arises because the predisposed retinoblasts do not suffer the second mutation that would result in the development of malignancy. Vogel (1979) reviewed all published pedigrees up to 1979 and calculated that penetrance was between 0.996 to 0.927, with a higher frequency in offspring of bilaterally affected parents. The penetrance value, 0.945, found in the series presented here is consistent with this finding. 5 of the 10 unaffected gene carriers were born to unilateral parents (in 4 different families) and the other 5 to bilateral parents in 3 separate families. In one of these families (RBF29), however, the parent had bilateral retinal scars only. As indicated with the presence of 10 non-penetrance cases in 6 separate families, there is tendency of such cases to aggregate in particular families. This indicates two different types of non-penetrance cases among Rb families. The first kind seems to occur simply due to chance as explained by the Poisson distribution of the second mutation, regardless of the type of predisposing mutation in the families involved. The second kind, on the other hand, seems to be due to an increased chance of non-penetrance because of the nature of the mutation in a particular family (Section 4.3.3). Family RBF29 (Section 3.1.7) is one of the best examples of such a family with 4 non-penetrance cases. The same kind of explanation probably applies to spontaneous regression of Rb. In this series the percentage of regressed tumour cases was 5.5%, a very high figure compared to the figure of 1% suggested by Gallie et al (1977, 1982a), largely due to the presence of a couple of families showing aggregation of such individuals. Again family RBF29 is a good example with two cases but the best example is family RBF18 (Section 3.1.7) with four such cases in four generations: one in every generation. In contrast the spontaneous regressions observed in families RBF38 and RBF41 (Section 3.1.7) are probably one-off chance occurrences and could be explained either by the late occurrence of second mutation in committed retinal cells or the nature of the second mutation itself. The aggregation of such cases in a particular family, however, is difficult to explain by chance occurrences and the nature of the first mutation is probably important (Section 4.3.3). It is difficult to estimate the true incidence of spontaneous regression among Rb cases. As explained
in Section 3.1.7, in the cases of unaffected parents having more than one affected child (RBF04, RBF32, RBF41, RBF56), it is not known whether these parents have been examined for retinal scars. It is only relatively recently that ophthalmologists have become aware of this phenomenon. Examination of apparently unaffected relatives would have altered the advice given to family RBF29 for example. The identification of spontaneously regressed tumours in the grandparents in this family raised the risk to the children to 1:2 compared with the 1:32 risk offered at the time of first consultation. Ocular fundus examination of apparently unaffected relatives is, therefore, an important step in counselling Rb families. Because of the unusually low penetrance of the gene mutation in some families, this needs to be extended as far back in the pedigree as possible. Since the ophthalmoscopic appearance of regressed tumours may fall outside the experience of some ophthalmological practices, it should be routine to refer families or individuals concerned to specialist centres for examination.

RBF29 is one of the best examples of a low penetrance family. As stated in Section 1.4.7, Ellsworth (1969) and Matsunaga (1976) were among the first who suggested that penetrance varies according to the manifestation in the carrier parent. Indeed, a positive correlation was found between expressivity (unilateral versus bilateral involvement) and penetrance in various studies (Matsunaga 1976, Bonaiti-Pellie and Briard-Guillemot 1981). In this series, although all affected parents had more bilateral children than unilateral children, the percentage of bilaterally affected children born to bilateral parents (approx. 82%) was much higher than the percentage of bilaterally affected children born to unilateral parents (approx. 55%). When the families with unaffected parents but more than one affected child were investigated, only 45% of the offspring born to these unaffected parents, who could be unaffected gene carriers or germ-line mosaics, was found to be bilaterally affected. Therefore, the data from this series also points to increased penetrance with increased expressivity (unaffected gene carrier to unilateral to bilateral) in the carrier parent. In one of these families with unaffected parents (RBF04) where two out of three first generation offspring were unilaterally affected (Section 3.1.3), the causative mutation was identified by Weir-Thompson et al (1991) who was also able to obtain DNA from the unaffected mother (I.1) and trace the defective allele to the unaffected father through linkage analysis. It was, therefore, the father who must have been a
carrier either as a non-penetrance case or as a mosaic for a germline mutation.

Matsunaga (1980a) also investigated the expressivity in the carrier parent in relation to the occurrence of second primary tumours. He found no correlation between the two: the risk of second primary tumours did not differ depending on the phenotypes of the gene carriers. The percentage of second cancers among gene carriers in this series was found to be 6.4% although this may be an underestimation as the cause of death is not determined in all the deceased affected individuals. Also there may have been additional cancer incidences that were not reported to us. Although the majority of individuals who developed second cancers were bilaterally affected, this is only to be expected as bilaterally affected individuals comprise 75% of this series. Perhaps the incidence of second cancers have been overstated in the past but the increasingly real observation that Rb mutant gene carriers develop second tumours at an earlier age than non-gene carriers still puts them at high risk.

A great deal of heterogeneity in the Rb phenotype between families has been observed in this series (Section 3.1.7). Such heterogeneity is occasionally observed within a single family as in the case of RBF29. A family similar to RBF29 has also been reported by Connolly et al (1983) in which there were three distinct categories of RB1 gene expression: unilateral and bilateral Rb, retinoma and non-penetrance. In this family, however, in contrast to RBF29, a striking difference has been observed in Rb penetrance between the two principal generations. In the second generation the penetrance was low, non-penetrance and only a single unilateral Rb being observed. In the third generation, the penetrance was very high with many unilaterally and bilaterally affected individuals. A similar family has also been described previously by Bundey and Morten (1981). Although many theories have been put forward to explain the existence of such families (Section 1.4.7), none of them adequately explains why susceptibility has suddenly increased significantly from one generation to the next. If delayed mutation is postulated (Hermann 1976), an unstable intermediate would be expected to persist for several generations before undergoing final transformation to Rb. If, on the other hand, as Matsunaga (1978) suggested, the second mutation involves a host-dependent error in differentiation, the existence of an host-resistance gene segregating independently to Rb has to be postulated. If such
a gene exists, however, it cannot be explained why it is strongly associated with the Rb allele in the low penetrance generation but segregates assortively away from Rb in the high penetrance generation. Although families where there is a striking difference in penetrance between two generations have not been observed in our series, one of the families, RBF01, exhibits two extreme categories of Rb expression in a single generation. In this family where the mother and the first child are bilaterally affected, the two subsequent children are non-penetrance cases and recombination has been ruled out (Section 3.1.7). The existence of these two non-penetrance cases in an otherwise high penetrance family is difficult to explain invoking chance occurrence, and none of the other theories postulated applies in this case, unless the nature of the mutations is important. In this family, DNA from a lymphoblastoid cell line from the affected mother was studied using SSCP analysis for most of the RB1 gene apart from the 5' promoter area and exons 1, 2, 8, 11, and 25-27 and no mutations have been identified so far. It is possible that, under the conditions used, SSCP analysis has missed a mutation since the conditions of analysis were not optimised individually for each exon. The continuation of this project should, therefore, try to identify the predisposing mutation in this family as a priority because the nature of this mutation may be relevant to the differential penetrance observed in families.

4.2 MUTATION DETECTION TECHNIQUES FOR THE IDENTIFICATION OF THE RB1 GENE MUTATIONS

Although many mutation detection techniques have become available recently (Fischer and Lerman 1983, Myers et al 1985a, Cotton et al 1988, Orita et al 1989a, Ganguly and Prockop 1990), finding a technique that is both rapid and sensitive for routine analysis of mutations in a gene which is 180 kb long encoding a 4.7 kb mRNA proved difficult. Unlike cystic fibrosis, for example, where a single mutation, ΔF508, accounts for approximately 70% of the mutant alleles (Kerem et al 1989), Rb mutations are relatively unique for each familial and sporadic case. In addition, the RB1 gene does not seem to have hot spots for mutations, unlike the p53 gene, for example. I chose, within the time available, to study four different approaches as they become available: RNase protection, PCR amplification and sequencing of individuals
exons, primer extension of CDI modified heteroduplexes and SSCP analysis of PCR amplified DNA.

Clearly for such a large gene analysing the mRNA would be much quicker and far more efficient. Because both Rb alleles carry mutations in Rb tumours this would be the appropriate place to start. Also it would be relatively straightforward then to identify new germ-line mutation cases. Only about one-third of tumours, however, are removed the rest being treated in situ. Of those which are removed many have already been treated and the tumours are largely necrotic. It is rare, these days, for tumours from a patient with a hereditary predisposition to be removed since they are usually identified at an early stage and treated. An additional complication is that the RBI RNA is not produced in up to 40% of tumours (Friend et al 1986, Fung et al 1987, Lee et al 1987a, Goddard et al 1988). In the absence of tumour material, it might have been expected that RNA from constitutional cells of patients could be analysed although the presence of the normal allele appears to interfere with production of the abnormal message. At the commencement of this project RNase protection was the only method generally available which was capable of analysing large stretches of RNA or DNA, although this method was certainly not ideal because it was estimated to detect only about half of known mutations (Myers et al 1985a, Winter et al 1985). Lymphoblastoid cell lines were available for a number of Rb patients and the RNA obtained from them were analysed with normal controls. Only the full-length protected fragments were detected (Section 3.2.1). Dunn et al (1988), however, were able to obtain mRNA from previously established Rb cell lines and detected abnormal ribonuclease cleavage patterns in five of 11 cases (45%). All of these cell lines had normal-sized RNA transcripts. The localisation of mutations in these tumours was deduced from the cleavage patterns of overlapping riboprobes. In two of the tumours analysed, the two protected fragments together equalled the length of the fragment normally protected, suggesting that the mutations are single base pair mismatches at the site of cleavage. In three other tumours, both a full-length and a shortened fragment was observed locating the abnormality 52 bp from the 5' end of the probe. The small, 52 bp fragment, however, could not be detected in any of the tumours. A similar situation was encountered in one of the control experiments performed in this project (Section 3.2.1, figure 3.22a) where an expected 47 bp fragment could not be identified on the gel. The presence of the full-length fragment in the tumours described by Dunn et al (1988), on
the other hand, indicated either the presence of transcripts from the other allele or incomplete digestion of the RNA hybrids protected by the mutant allele. Three of the tumours analysed by Dunn et al (1988) came from patients with bilateral Rb but they were unable to detect the same mutation in constitutional cells from these patients. This suggested the possibility that at least some mutant transcripts may be suppressed by the normal transcript and/or the mutant transcript may be unstable in lymphoblastoid cell lines and normal fibroblasts. Later on Dunn et al (1989) cloned and sequenced the PCR-cDNA from the constitutional cells of one of these patients and identified a 10 bp deletion within exon 18 which was the same mutation as that found in the tumour cells. This was the only case where they could detect transcripts from the mutant allele in lymphocytes. To explain these results they suggested that a functional RB1 gene product might regulate its own transcription rate and that the transcripts produced from the mutant allele might be less stable than the normal transcripts. They further postulated that the transcription rate from the normal allele may increase to compensate for the functional loss of RB1 product from the mutant allele in normal cells. Taken together with the proposed unstable nature of the mutant transcript it is unlikely that these would be detected in normal tissues. Because both alleles are mutant in Rb tumours the dominant effect of the normal gene does not come into play and transcription from both mutant alleles can occur which are easily detectable despite their instability. If true, this model, whereby the normal allele expression masks the expression of the mutant allele in constitutional cells, would explain why no mutant transcripts were detected in the assays performed in this project and means that RNase protection assays for RB1 gene mutations have a limited use in diagnostics. Indeed, results of some recent experiments have provided supporting evidence for this model. The RB1 protein has been demonstrated to repress transcription of genes with promoters containing E2F binding sites (Hamel et al 1992) (see also Section 4.3.3). Such a binding site is present also in the promoter of RB1 and it has been shown that expression of RB1 itself would be subject to regulation by pRB (Hamel et al 1992). It has also been observed that cell lines which were produced containing RB1 transgenes under control of heterologous promoters express pRB at low to moderate levels and it has been difficult to produce stable lines containing pRB constitutively expressed from a strong heterologous promoter (Hamel et al 1992). These observations have suggested that the levels of pRB are normally tightly controlled by autoregulation. In tumours, lack of normal pRB and hence autoregulation,
would result in overproduction of mutant message whereas in constitutional
cells, as the normal protein is present, lower levels of mutant transcript would
be produced due to transcriptional repression.

The only other reported RB1 gene mutation identified using RNase protection
was a single base pair mutation detected in a small cell lung cancer cell line
(NCI-H209) by Kaye et al (1990). RNase protection has successfully been
used by others to identify mutations in an estimated 50% of cases in other
genes. Single base mismatches have been detected in RNA:DNA heteroduplexes of β-globin gene from cloned and genomic DNA (Myers et al
1985a) and in DNA:DNA heteroduplexes in c-K-ras genes from human
mutations in the hprt (hypoxanthine-guanine phosphoribosyl transferase)
mRNAs from 5 of 14 Lesch-Nyhan syndrome patients. One interesting
application was from Lopez-Galindez et al (1988) who analysed genetic
variability in RNA viruses. The analysis of RNA:RNA heteroduplexes from
different viral strains led to the identification of single base mismatches and
allowed a qualitative estimation of the genetic relatedness and evolution of
field strains in this study. Mutations have also been identified in the cDNAs
for ornithine transcarbamylase from the sparse fur mouse (Veres et al 1987),
in the sequences from c-myc from Burkitt lymphoma cell lines (Cesarman et
al 1987) and Smith et al (1988) identified multiple RNAs expressed from the
int-2 gene in mouse embryonal carcinoma cell lines. In a recent study by
Miyoshi et al (1992), using RNase protection, germ line mutations of the
APC gene were identified in 67% of FAP patients analysed which represents
a surprisingly high figure. The authors attributed their higher success rate to
the fact so many (30/53) of the mutations in the APC gene were small
deletions or insertions, which are more readily detectable with RNase
protection.

As the evidence accumulated, it soon became clear that only a fraction of all
single base mismatches would be cleaved by RNase A and that the extent of
cleavage varies depending on the mismatch (Winter et al 1985, Myers et al
1985a, Lopez-Galindez et al 1988). The reasons for this variability were not
clear. The recognition of a particular mismatch by the enzyme appeared to
depend not only on the bases forming the mismatch but also on the
nucleotides surrounding the mispaired bases (Winter et al 1985, Myers et al
1985a, Lopez-Galindez et al 1988). It also appeared that mismatches are
more easily cleaved in DNA:RNA hybrids than in RNA:RNA hybrids. The problem with the use of genomic DNA, however, was that the relative concentrations of target sequences when analysing single copy genes are much less abundant than those present in total cellular RNA. This problem could be circumvented, however, by using PCR amplified sequences (Veres et al 1987, Miyoshi et al 1992). This approach could have solved the problem encountered in the identification of mutations from constitutional cells of the Rb patients using RNase protection where the production of the mutant mRNA was suppressed but, it was felt that amplification and RNase protection assay of 27 separate exons was too lengthy a procedure for use with a method which is only capable of detecting a fraction of single base mismatches.

Whichever initial screening technique is used, ultimately the precise identification and characterisation of mutations is achieved by sequencing. Conventional sequencing protocols involve generating overlapping clones, which is a lengthy procedure. PCR technology, on the other hand allows easy access to any desired part of a gene provided primers are available for the amplification step. Using the published RBI sequence (McGee et al 1989), primers were designed to amplify each exon and flanking intron sequences of the RBI gene as described in Section 3.2.2. Even though some predictions can be made about the efficiency of pairs of primers, designing these primers and optimising the conditions for amplification and sequencing of each exon was still done empirically and proved to be a formidable task which took approximately one year to complete. The amplification and sequencing of the 5' region of the RBI gene and the regions containing Alu-like sequences was especially difficult as was reported by others (Yandell et al 1989). The 5' regions were found impossible to amplify because of their very high 'GC' content unless DMSO is added to the reaction mix. As such regions (and also regions containing repeats) tend to give rise to secondary structure formation, they were also difficult to sequence and background bands, stops and compressions were often observed on the sequencing ladders. Stops appear as bands across all four lanes of the autoradiograph and are presumably due to areas of strong secondary structure in the template causing the polymerase to terminate synthesis at that point independent of the incorporation of a dideoxynucleotide. Compressions, on the other hand, appear as inappropriate spacing between bands in one or more lanes on the autoradiograph and are due to strong secondary structure in the labelled
products of the sequencing reactions which cause their mobility in a gel to be influenced by conformation in addition to size. It was sometimes possible to eliminate these artefacts by substituting a nucleotide analog for dGTP (dITP or 7-deaza-dGTP) which forms weaker secondary structure but there were still areas, such as introns 13 and 14 which contain Alu-like repeats and intron 21 which contains long stretches of pyrimidines, which proved very difficult to sequence and if they harbour the mutation they may be difficult to detect.

Although direct sequencing from PCR products is much faster and simpler, several problems were encountered before establishing and optimising a consistent protocol. The difficulty in sequencing PCR products directly arises from the nature of the PCR amplified DNA which consists of short, linear double stranded DNA. The two strands of the DNA can rapidly reanneal during DNA sequencing, blocking or displacing the sequencing primer from the template strand, thereby decreasing the amount of specific termination products formed in the sequencing reactions. One of the first ways of circumventing this problem was to generate a single-stranded template in subsequent rounds of amplification using only a single primer, a variation of the so-called "asymmetric" amplification procedure (Gyllensten and Erlich 1988). This approach is very time consuming. Firstly, two rounds of PCR amplification are required and PCR products need to be extracted with phenol/chloroform and reprecipitated each time to eliminate the first set of primers and free nucleotides. Secondly, it proved difficult to generate sufficient single-stranded product from the second PCR reaction so repeated amplification was necessary. The problem of insufficient DNA for sequencing is in part due to the DNA losses that incur during the cleaning up of the PCR products. Reamplification from PCR product also increases the risk of contamination and amplification from an external source. When the asymmetric method was tried in this project, it was found to be unreliable and only worked in approximately 3 out of 10 cases. Single-stranded DNA, however, can also be obtained using the streptavidin-biotin system and magnetic beads (Hultman et al 1989). After a considerable amount of effort with asymmetric PCR the magnetic beads method became available (Sections 2.21 and 3.3) which gave satisfactory results in at least 95% of experiments. With the advent of this technique it became possible to consider screening large numbers of exons. As mentioned in Section 4.1, there is always the risk of getting false positives and negatives when PCR is involved. To make
sure such errors did not occur, appropriate controls were set up and, where possible, PCR amplification and sequencing from original source DNA was repeated.

Direct sequencing of the entire coding region and limited parts (splice junctions) of the flanking intron regions of a gene is the most sensitive and sure way of identifying mutations. Indeed, Yandell et al. (1989) identified 10 causative point mutations in an exon-by-exon screen of tumour DNA using PCR amplification and direct sequencing analysis. The same approach has identified mutations in other genes such as the rhodopsin gene from patients with autosomal dominant retinitis pigmentosa (Dryja et al. 1990). Mutations could also be identified using cDNA, obtained by reverse transcription (RT) from mRNA, which is then amplified and sequenced. RT-PCR-sequencing has been used to identify mutations of the RB1 and p53 genes in human tumours and tumour cell lines (Shew et al. 1990b, Mori et al. 1990, Scheffner et al. 1991). With this approach the amount of sequencing is reduced but 4.5 kb still needs to be sequenced for each patient. There is also the possibility of the mutant transcript not being produced in some cases. Although direct sequencing is a highly sensitive technique, application of this technique to a large gene with no known hot spots for mutations, is not a practical approach especially for routine screening of patients.

To avoid the need to sequence the whole gene in patients and/or their tumours, several techniques were developed which allowed prescreening of exons for mutations. Among these are denaturing (DGGE) and temperature gradient gel electrophoresis (TGGE) (Fischer and Lerman 1983, Myers et al. 1985b, Myers et al. 1987, Wartell et al. 1990). DGGE and TGGE is based on the melting characteristics of DNA which is affected by its base composition. The separation of sequences altering by only a single base pair is achieved by generating either a denaturing solvent gradient (DGGE) or temperature gradient (TGGE) in a polyacrylamide gel. The sequence affects the melting transition and the electrophoretic mobility of duplex DNA with single-stranded regions. Melting transition takes place as DNA unwinds in a series of cooperatively melting domains. A DNA molecule moving in a polyacrylamide gel with a gradient of denaturant or temperature migrates until it reaches a denaturant concentration or temperature which induces a domain to unwind. At this point the mobility of the DNA decreases, probably because of entanglement of its branched structure in the gel matrix.
DNA fragments 100-1000 bp long generally have 2 to 5 melting domains and DNAs differ in the stability of their first melting domain unwind at different positions in the gel. This technique was estimated to detect 40-50% of single base changes in DNA fragments of up to 1 kb in length (Myers et al. 1985b). However, with the attachment of a GC-rich region (GC clamp) to the test sequence, the sensitivity of the technique was reported to be improved (Myers et al. 1985b, Sheffield et al. 1989) so that virtually all single base changes within a given fragment are theoretically detectable. Mutations have been identified using DGGE in various genes, for example the factor VIII gene (Traystman et al. 1990). Blanquet et al. (1991b) used DGGE to identify mutations in the RB1 gene and detected only 13 abnormalities including a new polymorphism (Blanquet et al. 1992) among the DNA of 110 patients studied. Both DGGE and TGGE require several manipulations, extensive theoretical work on melting profiles and use of GC clamps to increase the sensitivity of detection and analysable area.

There are mutation detection techniques which involve chemical modification of mismatches. One of these techniques is the hydroxylamine / osmium tetroxide (HOT) method (Cotton et al. 1988). This method depends on the reactivity of these chemicals with unpaired pyrimidines in DNA heteroduplexes. The DNA is then cleaved adjacent to modified residues using piperidine. Using this method point mutations have been identified in many genes including the factor IX gene and the large dystrophin gene (Montandon et al. 1989, Roberts et al. 1992). So far, there has been only a single report (Weir-Thompson et al. 1991) on the use of this technique in the RB1, describing a single base mutation identified initially in the liver tumour (SCLC metastasis to liver) from an Rb patient (from family RBF04 also described in this thesis).

In this project, the HOT technique was considered to be unsuitable for a routine screening procedure because of the dangerous nature of the many of the chemicals involved. I, personally, felt uncomfortable using the chemicals involved especially because of their damaging effects on the retina. Instead, another chemical mismatch analysis, CDI-modification of heteroduplexes, was briefly investigated as described in Section 3.2.3. Using the CDI method, Ganguly and Prockop (1990) detected all eight of the possible single base mismatches in the heteroduplexes they created in a model test system from the four variants of phage M13 which were prepared so as to differ by a
single base at the same site. They were also able to detect single base mutations in the coding sequences of the type I (COL1A2) and type III (COL3A1) procollagen genes. In the preliminary experiments performed in this project using the CDI method (Section 3.2.3) it was possible to detect the single base mismatch in the heteroduplexes formed from the RB1.3 polymorphic variants although the background was very high in the autoradiographs produced. This method was only briefly studied because it soon became clear that a great deal of time would have to be spent on the optimisation of the conditions for each of the 27 exons of RB1. This was particularly true because the experimental procedure consisted of many steps. Again, as with RNase protection, it was not clear whether the method was capable of detecting all mutations. It is unlikely that a single chemical reaction with DNA heteroduplexes will detect all single base mutations in all sequence contexts as pointed out by Ganguly and Prockop (1990). In fact, it is unlikely that any of the techniques involving chemical procedures for detecting mismatches will prove totally satisfactory. Despite this, using the combination of hydroxylamine and osmium tetroxide, all possible single base mutations within a series of different sequence contexts were detected (Cotton and Campbell 1989). In other studies, however, single base mismatches in other sequence contexts could not be detected (Bhattacharyya and Lilley 1989). Therefore, it is necessary to test these chemical procedures in a variety of sequence contexts and with several different reaction conditions. Ganguly and Prockop (1990) were able to detect one of the mutations in one procollagen gene sequence only by altering the reaction conditions with CDI. Altering the reaction conditions in this case, however, was found to increase the non-specific interaction of the DNA with CDI, thereby decreasing the yield of the specific mutant product. The need to perform reactions under different conditions in order to detect all possible mismatches makes a procedure, which is already too lengthy, even more laborious especially when analysing a large gene.

It appears from many informal discussions that all techniques for detecting mutations is subject to limitations and, in practise, different laboratories will employ those techniques which they can get to work consistently and which produce results. For this Rb project, the aim was not simply to find techniques that will identify mutations in the RB1 gene, but rather to find a technique that is simple, reliable, reproducible and is able to identify most of the possible mutations that might be encountered. Only then would it be
possible to transfer the technology to the clinic. During the time that many of these techniques were under investigation the SSCP technique was evolving. The most notable feature of SSCP analysis was its relative simplicity. Only a single PCR reaction was required which produces an amplified and labelled product. Although in the case of large fragments enzyme digestion to produce smaller products is an extra step, it is conveniently performed overnight whereafter the analysis of the products is simple involving an electrophoresis run on non-denaturing polyacrylamide gels. This was the technique which proved most successful for us and so far mutations were identified in 40% of the samples analysed. It should be stressed, however, that not all of the exons of the RB1 gene have been analysed in these samples and only a limited set of electrophoresis conditions was used in all cases. As this technique rapidly became the one of choice experience allowed more subtle band shifts to be recognised, which might have been ignored earlier and certainly would not be seen in experiments reported from other laboratories with less experience. Since PCR-SSCP analysis is a relatively new technique, further accumulation of data may be necessary to accurately estimate its sensitivity and success in mutation detection. It has already been shown that SSCP is capable of detecting a variety of mutation types: transitions, transversions, deletions and insertions (Orita et al 1989b, Dryja et al 1991, Onadim et al 1992b, Hogg et al 1992, A. Hogg unpublished results, D W Yandell personal communication). The question is, therefore, whether it is capable of detecting all mutations in all sequence contexts. So far, the figures reported for the rate of mutation detection for this technique ranges from 35% to 99% depending on the size of the fragments analysed and conditions used (Hayashi 1991, Dryja et al 1991, Sarkar et al 1992). Clearly, as with all techniques, with a little effort and using different experimental conditions, it may be possible to identify all mutations. There is one fundamental difference, however, because in SSCP only the size of the analysed fragment and the conditions of the electrophoresis run needs to be varied. Since only approx. 1/100th of the original amplification product is necessary for each analysis, several attempts could be performed using the same amplified product. Thus, different electrophoretic conditions and different combinations of enzymes could be run simultaneously. In this project, had there been time to optimise SSCP conditions, individually, for each exon of the RB1 gene and analyse all exons and the promoter region, the number of mutations detected would certainly have been greater.
SSCP analysis has also been used successfully by others to detect mutations in the RB1 gene (A Hogg, D W Yandell personal communication, Murakami et al 1991, Sakai et al 1991a). Several mutations have been identified in tumours and constitutional DNA of the Rb patients, with or without a family history, in the ICRF laboratory at the ICH (A Hogg, unpublished results). The ophthalmology group in Harvard Medical School have identified a large number of somatic and germinal mutations in the RB1 gene using SSCP analysis (D W Yandell, personal communication). Sakai et al (1991a) detected two mutations (one transition and one transversion) in the 5' promoter area of the RB1 gene from Rb families and Murakami et al (1991) detected a transversion in the RB1 gene from a human lung carcinoma cell line (Lu65). There have also been reports of mutations identified in a variety of other genes. PCR-SSCP was successfully used to detect mutations in the cystic fibrosis gene (Dean et al 1990), p53 gene (Mazars et al 1992, Iavarone et al 1992), WT1 gene (Baird et al 1992), rhodopsin gene (Dryja et al 1991), NFI gene (Cawthon et al 1990), factor IX gene (Demers et al 1990) and the gene for phenylketonuria (Labrune et al 1991) among others.

Interestingly, conformational polymorphism has recently been extended to the study of RNA. Danenberg et al (1992) and Sarkar et al (1992), using RNA-SSCP detected point mutations in p53, dihydrofolate reductase (DHFR) (Danenberg et al 1992) and factor IX (Sarkar et al 1992) genes. They used RNA generated from PCR amplified DNA hypothesising that RNA may have multiple stable conformational states due to a larger repertoire of secondary structures, thus making it more amenable to conformational analysis. In both studies, RNA-SSCP was found to be superior to DNA-SSCP especially in the analysis of larger fragments. Sarkar et al (1992) performed a blind study in which RNA and DNA SSCP and direct genomic sequencing were compared in 28 patients with haemophilia B. A total of 2.6 kb of the factor IX genomic sequence was examined through analysis of fragments 180 to 497 bp in length. Sequence changes at 20 different sites were detected by direct sequencing. It was found that only 35% of these were detected by DNA-SSCP whereas 70% were detected by RNA-SSCP. It is possible, however, that the low incidence of mutations detected in DNA reflected the poor quality of the SSCP autoradiographs in these experiments. Nevertheless, it seems that RNA-SSCP will be an important addition to currently available methods for mutation screening in the future.
Among all the methods available, SSCP analysis seems to carry the highest potential for use in clinical applications. To transfer this research technique to a clinical test system, however, requires that experimental conditions are optimised depending on the application. For the RB1 gene, the optimal fragment sizes and electrophoresis conditions had to be optimised for each exon. The technique can be made time and cost-efficient by using multiplex PCR which involves simultaneous amplification of more than one exon. This approach has already been used for some of the exons of the RB1 gene (Hogg et al 1992) and, although the interpretation of resultant autoradiographs in some cases was complicated, it was still considered to be feasible if PCR primers are carefully designed specifically for this purpose. For routine analysis the use of a non-radioactive detection method would also be very important. Ainsworth et al (1991) demonstrated that the banding profile seen in radioactive SSCP analysis can be detected by silver staining although this is a time-consuming and costly process. Yap and McGee (1992) reported the detection of non-isotopic SSCP products by ethidium bromide staining, but Sarkar et al (1992) were only able to detect the renatured duplex DNA in similar experiments. Sarkar et al (1992), using RNA SSCP, reported limited success using ethidium bromide staining; although mobility shifts were detectable, the resolution was not as good as with autoradiography. Using fluorescence-labelled primers, PCR products could be analysed using a fluorescence-based automated sequencing machine (Chehab and Kan 1989). Using this technique mutations were apparently detected at a sensitivity equal to that of the radioactivity-based PCR-SSCP analysis (Hayashi 1991). Clearly these alternative approaches need more development before they will replace the consistently successful procedures using radioactively labelled compounds.

4.3 MUTATIONS OF THE RB1 GENE

4.3.1 The physical changes in RB1

In this project, 8 different RB1 mutations were detected in the DNA from constitutional cells of patients with familial Rb. Two mutations, found in the DNA of the affected individuals from families RBF34 and RBF31, were large deletions identified by Southern blotting analysis (Sections 3.1.5, 3.1.7).
Although the exact breakpoint sites in these deletions were not identified, linkage analysis indicated that both 5' breakpoints lie beyond intron 17 between the PR0.6 and RS2.0 loci and extend at least the RB1.20 site in intron 20, a distance of at least 50 kb. In family RBF31 (Section 3.1.7) the deletion was inherited from the mother who is either an incomplete penetrance case or a germ-line mosaic whereas in RBF34 (Section 3.1.5), the deletion was paternal in origin. The latter result is consistent with the observation that new germ-line mutations are usually paternal in origin. Ejima et al (1988) showed that cytogenetically visible germ-line mutations are usually in the paternally derived gene. Such a bias would not be expected in sporadic tumours, where both mutations occur in somatic tissue, although there is evidence of a bias towards initial somatic mutation in the paternally derived gene on chromosome 11 in sporadic Wilms' tumour (Schroeder et al 1987, Williams et al 1989, Section 1.5) and in the paternally derived RB1 gene in sporadic osteosarcoma (Toguchida et al 1989). Toguchida et al (1989) examined 13 non-hereditary osteosarcomas and reported that, in 12 cases, the new mutation arose in the paternally derived gene suggesting the involvement of germinal imprinting in producing the differential susceptibility of the two RB1 genes to mutation. Dryja et al (1989) performed a similar analysis of Rb patients and found no such predilection for the initial somatic mutations since, of seven non-hereditary Rb tumours the initial mutation occurred on the chromosome derived from the father in three cases, and from the mother in four. In contrast, when an initial mutation was a new germ-line mutation, it always arose in the father which is consistent with new germ-line mutations arising primarily during spermatogenesis. These results were also confirmed by Zhu et al (1989) who found no evidence that the paternal Rb allele is preferentially retained in nine sporadic tumours. By contrast, tumours from patients with bilateral tumours preferentially retained the paternal Rb allele (Zhu et al 1989, Dryja et al 1989). Several suggestions, such as differences between male and female meiosis, DNA methylation or environmental exposure or the paternal chromosome in the early embryo being more at risk for mutation or deficient in DNA repair (Zhu et al 1989), have been made to explain why new germ-line mutations arise more frequently during spermatogenesis than oogenesis. The most likely explanation, however, is that spermatogenesis, unlike oogenesis, involves many thousands of cell divisions during the lifetime of the male, each of which provides opportunity for mutations which can become fixed in the population.
Large deletions within the RB1 gene have also been reported by others in a variety of tumours including Rb tumours (Friend et al 1987, Fung et al 1987, Lee et al 1988, Bookstein et al 1989, Canning and Dryja 1989, Wunder et al 1991, Venter et al 1991) and in constitutional DNA of Rb patients (Horsthemke et al 1987a, 1987b, Janson et al 1990, Hashimoto et al 1991, Kloss et al 1991, Blanquet et al 1991a). No evidence for the existence of a hot spot for such deletions was found, although exons 13-17 and 21-24 were frequently observed to be included in these deletions (Fung et al 1987, Canning and Dryja 1989, Hong et al 1989, Hashimoto et al 1991). All of these deletions are expected to result in either truncated (presumably non-functional) Rb proteins (pRB) (Shew et al 1990b, Hashimoto et al 1991) or no protein at all. Shew et al (1990b), for example, identified a truncated pRB which was missing regions encoded by exons 21 to 27 in an osteosarcoma cell line (Saos-2) that resulted from a transcriptionally active RB1 allele. Two 'new' germinal RB1 deletions (2 kb and 3.7 kb long) showed mRNAs and truncated proteins of the expected size (Hashimoto et al 1991). Others, however, were unable to detect mRNA or pRB in tumours with various sized RB1 deletions (Friend et al 1986, Fung et al 1987, Lee et al 1987a, 1988, Goddard et al 1988, Horowitz et al 1990).

Deletions or rearrangements that are detectable by Southern or Northern analysis occur only in 10-30% of Rb patients (Fung et al 1987, Goddard et al 1988, Kloss et al 1991, Blanquet et al 1991a), the remainder are more subtle mutations of the kind reported in this thesis (Section 3.3). Four of these mutations were small deletions or insertions and occurred in Rb families where all affected individuals had bilateral tumours (Section 3.3.1). The other two were point mutations and were identified in families showing incomplete penetrance and mild expression of the Rb phenotype (Section 3.3.2).

The 2 bp heterozygous deletion in exon 3 in family RBF64, the 1 bp heterozygous deletion in exon 17 in RBF59 and the 1 bp heterozygous insertion in exons 13 and exon 20 respectively in RBF58 and RBF62 resulted in frameshifts leading to premature stop codons nearby. It would be expected, therefore, that truncated proteins, 128, 545, 426 and 671 amino acids long would be produced. Lack of tumour cells precluded mRNA analysis and it has since been shown (JK Cowell, personal communication) that these abnormal mRNAs are not produced in constitutional cells. Despite
the prediction it is possible that those mutations close to splice recognition sequences affect mRNA processing. This applies particularly to the 2 bp deletion in exon 3 in RBF64 which is only 9 bp upstream of the splice donor site (Section 3.3.1 figure 3.30b) and might reduce the efficiency of splicing from this site or promote the use of any existing cryptic splice site nearby resulting in alternative splicing. Alternatively, exon 3 might be spliced out all together from the mature mRNA -exon-skipping-, although this would again disturb the reading frame and give rise to a stop codon downstream in exon 4. There is evidence from the study of other genes that a change in the local sequence environment can affect splicing patterns (Reed and Maniatis 1986, Steingrimsdottir et al 1992). It has been suggested that the interaction between factor(s) present in a splicing extract and the splice sites is affected by exon sequences and that exon sequences play a key role in distinguishing between normal splice sites and cryptic splice sites located throughout pre-mRNAs (Somasekhar and Mertz 1985, Reed and Maniatis 1986, Kuo et al 1991). It has also been shown that mutations some distance away from the splice sites can affect splicing in the human hypoxanthine-guanine phosphoribosyltransferase (hpri) gene (Steingrimsdottir et al 1992). Mutations affecting splicing in unexpected ways have also been identified in the RB1 gene (Horowitz et al 1990, Mori et al 1990, Murakami et al 1991, Weir-Thompson et al 1991). For example, a G→A transition (table 4.1) at the fifth base of intron 21 resulted in exon 21 being lost from the mRNA (Weir-Thompson et al 1991). Also the C→A transversion in exon 2 of the RB1 gene in the human lung carcinoma cell line, Lu65, although apparently creating a stop codon actually resulted in the excision of exon 2 (Murakami et al 1991).

A summary of small deletions/insertions and point mutations reported in the RB1 gene are given in table 4.1. Small deletions seem to be randomly distributed throughout the gene and constitute 33 % of mutations reported by Dunn et al (1989) and Yandell et al (1989). These tumours, with one exception, were all from bilaterally affected Rb patients. The small deletions reported by Lohmann et al (1992) were also from bilaterally affected hereditary Rb cases. Approximately 35 % of all RB1 mutations given in table 4.1 are small deletions (table 4.1) most of which resulted in frameshifts leading to stop codons. Exceptions were 3 bp deletion in exon 16, which led to an in-frame loss of one amino acid (Lohmann et al 1992), and a 43 bp deletion including the exon 15/intron 15 splice donor site which resulted in
exon 15 being joined directly to exon 17, skipping exon 16, via a cryptic slice donor site (Shew et al 1990a). One other deletion (RB-2, Yandell et al 1989, table 4.1), although apparently resulting in a frameshift leading to a stop codon, was adjacent to a splice signal. This deletion, therefore, was considered to be a potential splice site mutation (Yandell et al 1989) similar to the 2 bp deletion found in exon 3 in this project. The only other insertion reported in an Rb patient, who was bilaterally affected, was a 55 bp duplication in exon 10 of the RB1 gene (Dunn et al 1989). This insertion resulted in a frameshift leading to a stop codon (table 4.1) and the resulting mutant mRNA, although easily detectable in the Rb tumour (RB538), was not detectable in lymphoblastoid cells (Dunn et al 1989). Taken together with the mutations reported in this thesis, roughly 40% of all mutations are small deletions/insertions but the majority (60%) are point mutations.

In this study both point mutations identified were in exon 20 and were from families which exhibited mild expression of the Rb phenotype. The C→T transition identified in family RBF29 (Section 3.3.2) occurred within a C_pG dinucleotide pair and converted an arginine residue to a tryptophan residue. One other C→T transition within a C_pG pair, which converted an arginine to a stop codon, was identified in exon 14 (Hogg et al 1992). This mutation was from a bilaterally affected Rb family (RBF25 in Appendix III, also studied in this thesis using intragenic polymorphisms). C→T transitions are the most frequently observed point mutation in the RB1 gene and often occur within C_pG dinucleotide pairs (Ludeke et al 1991). Ludeke et al (Bi Ludeke, personal communication) who analysed over 70 mutations in Rb patients found 60 to 80% to be point mutations, a large percentage (56%) of which were C→T transitions, mostly at C_pG pairs. This finding is also consistent with the finding that 35% of all single base-pair substitutions causing human genetic disease occur within C_pG dinucleotides (Cooper and Youssoufian 1988, Youssoufian et al 1988). It has been suggested that C_pG dinucleotide pairs represent mutational hot spots in the RB1 gene (Bi Ludeke, personal communication) as was suggested previously for the factor VIII gene associated with haemophilia A (Youssoufian et al 1988, Traystman et al 1990). Out of five C→T transitions (coding strand) summarised in table 4.1, three (RB-74, RB-53, RB-W24) also occurred at C_pG pairs (Yandell et al 1989). Among point mutations transitions are more frequent (65%) than transversions, a finding consistent with observations at other genetic loci (Vogel and Rathenberg 1975, Cooper and Krawczak 1990).
Table 4.1 Summary of some of the reported point mutations, small deletions (Δ) and insertions affecting the RB1 gene and their consequences*.

<table>
<thead>
<tr>
<th>NAME / (SOURCE)</th>
<th>MUTATION</th>
<th>LOCATION</th>
<th>RESULTING CHANGE</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB429 (bilat Rb/tum)</td>
<td>Δ 5bp</td>
<td>exon 8</td>
<td>Frameshift→Stop</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB538 (bilat Rb/tum)</td>
<td>55bp dupl.</td>
<td>exon 10</td>
<td>Frameshift→Stop</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB543 (bilat Rb/tum)</td>
<td>Δ 10 bp</td>
<td>exon 18</td>
<td>Frameshift→Stop</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB570B(bilat Rb/tum)</td>
<td>Δ 9bp</td>
<td>exon 19</td>
<td>Frameshift→Stop</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB570C(bilat Rb/tum)</td>
<td>G→A</td>
<td>intron 21</td>
<td>Loss of SAS</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB571(unilat Rb/tum)</td>
<td>G→A</td>
<td>intron 12</td>
<td>Loss of SDS</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB600(unilat Rb/tum)</td>
<td>G→A</td>
<td>intron 12</td>
<td>Loss of SDS</td>
<td>Dunn et al 1989</td>
</tr>
<tr>
<td>RB-2 (bilat Rb/tum)</td>
<td>Δ 1bp</td>
<td>exon 24</td>
<td>Frameshift→Stop</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-88 (unilat Rb/tum)</td>
<td>T→C</td>
<td>intron 19</td>
<td>Loss of SDS</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-74 (bilat Rb/tum)</td>
<td>C→T</td>
<td>exon 14</td>
<td>Arg→Stop</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-104 (bilat Rb/tum)</td>
<td>C→T</td>
<td>exon 18</td>
<td>Ser→Leu</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-53 (bilat Rb/tum)</td>
<td>C→T</td>
<td>exon 23</td>
<td>Arg→Stop</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-45 (unilat Rb/tum)</td>
<td>Δ 1bp</td>
<td>exon 22</td>
<td>Frameshift→Stop</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-119(unilat Rb/tum)</td>
<td>G→T</td>
<td>intron 10</td>
<td>Loss of SDS</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>RB-W24 (?/Rbtum)</td>
<td>C→T</td>
<td>exon 11</td>
<td>Arg→Stop</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>Rb patient second tum</td>
<td>G→A</td>
<td>intron 21</td>
<td>Loss of SDS</td>
<td>Weir Thompsonetal 1991</td>
</tr>
<tr>
<td>Rb fam. 172 (unilat)</td>
<td>G→T</td>
<td>promoter</td>
<td>effect on transcription</td>
<td>Sakai et al 1991a</td>
</tr>
<tr>
<td>Rbfam 229 (unilat)</td>
<td>G→A</td>
<td>promoter</td>
<td>effect on transcription</td>
<td>Sakai et al 1991a</td>
</tr>
<tr>
<td>Rb fam Ro (bilat)</td>
<td>Δ 3bp</td>
<td>exon 16</td>
<td>loss of one aminoacid</td>
<td>Lohmann et al 1992</td>
</tr>
<tr>
<td>Rb fam Fe (bilat)</td>
<td>Δ 1bp</td>
<td>exon 20</td>
<td>Frameshift→Stop</td>
<td>Lohmann et al 1992</td>
</tr>
<tr>
<td>Rb fam Be (bilat)</td>
<td>Δ 10bp</td>
<td>exon 23</td>
<td>Frameshift→Stop</td>
<td>Lohmann et al 1992</td>
</tr>
<tr>
<td>H69 (sclc)</td>
<td>G→T</td>
<td>exon 22</td>
<td>Glu→Stop/newSDS</td>
<td>Yandell et al 1989</td>
</tr>
<tr>
<td>NCI-H1436 (sclc)</td>
<td>G→T</td>
<td>intron 21</td>
<td>Loss of SDS</td>
<td>Horowitz et al 1990</td>
</tr>
<tr>
<td>NCI-H209 (sclc)</td>
<td>G→T</td>
<td>exon 21</td>
<td>Cys→Phe</td>
<td>Kaye et al 1990</td>
</tr>
<tr>
<td>Lu-24 (sclc)</td>
<td>GC→TT</td>
<td>exon 22</td>
<td>Gln→Stop/abnormal splicing</td>
<td>Mori et al 1990</td>
</tr>
<tr>
<td>Lu-135 (sclc)</td>
<td>Δ 4bp</td>
<td>exon 20</td>
<td>Frameshift→Stop</td>
<td>Mori et al 1990</td>
</tr>
<tr>
<td>Lu-141 (sclc)</td>
<td>Δ 1bp</td>
<td>exon 23</td>
<td>Frameshift→Stop</td>
<td>Mori et al 1990</td>
</tr>
<tr>
<td>Lu-65 (sclc)</td>
<td>C→A</td>
<td>exon 2</td>
<td>Ser→Stop/alternative splicing</td>
<td>Murakami et al 1991</td>
</tr>
<tr>
<td>Primary sclc</td>
<td>Δ 43bp</td>
<td>ex/int 15</td>
<td>Loss of SDS</td>
<td>Shew et al 1990a</td>
</tr>
</tbody>
</table>

*Abbreviations: bilat=bilateral; unilat=unilateral; tum=tumour; sclc=small cell lung cancer; gcc=glial cell lung carcinoma; dupl=duplication; SAS=splice acceptor site; SDS=splice donor site; abn=abnormal; alter=alternative; sp=splicing; fam=family.
The majority of point mutations reported so far in the RB1 gene either directly lead to the generation of stop codons or affect splicing. Missense mutations in RB1 are very rare. In addition to the C-->T transition in exon 20 in RBF29 resulting in an amino acid substitution, only two other such mutations have been reported so far (table 4.1). One of these was from a SCLC cell line, in exon 21 (Kaye et al 1990), and the other was in exon 18 in a bilaterally affected Rb patient (RB-104, Yandell et al 1989). Of six C-->T mutations (table 4.1 and Hogg et al 1992), 4 affected an arginine codon. This is as expected since most of the C-->T mutations occur at CpG pairs and, of the 8 CpG containing codons, 4 encode arginine. The other four encode serine, proline, threonine and alanine. Of course when a 'C' at the third position of a codon is changed to a T and this codon is followed by a codon starting with a 'G' this mutation will also have occurred at a CpG pair. If C-->T mutations at CpG pairs really constitute hot spots for mutations then identification of the position of these pairs in the coding sequence of RB1 and their screening should be considered as a priority.

At first sight the G-->T transversion identified in exon 20 in family RBF18 converted a glutamic acid codon to a stop codon (Section 3.3.2). Such transversions are among the common physical changes observed in the RB1 gene (table 4.1). The expected result in this case is a truncated pRB, missing 254 amino acids at the C-terminal end. This mutation, however, was found in a family which we have classified as having a mild phenotype and, therefore, is apparently inconsistent with the suggestion that these phenotype result from subtle mutations. However, we have argued (Section 3.3.2, figure 3.38), because a TAG sequence lies 3 bp downstream of the mutation, a potential splice acceptor sequence (Mount 1982) is created. It is important to understand how apparently minor sequence changes can activate splice sites. Hoshijima et al (1991), for example, showed that by increasing the pyrimidine:purine ratio in the polypyrimidine tract of female-specific acceptor sequence of dsx (doublesex, which regulates somatic sexual differentiation in Drosophila melanogaster) makes it possible for this normally suboptimal splicing site to be used constitutively. The mutation in RBF18 also increases the pyrimidine:purine ratio in the region immediately preceding the TAG site thereby potentially generating a cryptic splice acceptor site in that region. The existence of a suitable branch point is also important in splicing. The mammalian branch point, which has the loose consensus TNCTA/GAC sequence, is usually located 18-40 nucleotides upstream from the 3' splice
junction and the underlined 'A' residue is involved in the formation of the
lariat intermediate of splicing (Green 1986, Reed and Maniatis 1988). The
lariat formation is the joining of the 2' hydroxyl group of the 'A' of the
branch point to the 5' phosphate residue of the first 'G' at the 5' end of the
intron whereby the upstream exon is separated from the rest of the pre-
mRNA. Splicing then proceeds with the joining of the 3' hydroxyl of the free
upstream exon to the 5' phosphate of the first nucleotide of the downstream
exon freeing the intron. There exists, upstream of the cryptic site in exon 20,
a branch point sequence which could be utilised if the cryptic site is activated
(T. Maniatis, personal communication). Under normal circumstances, in the
presence of cis-competition with the normal site, cryptic sites are never used.
A change in the local sequence environment, however, can change the
splicing pattern (Somasekhar and Mertz 1985, Reed and Maniatis 1986). In a
cis-competition assay for splice site selection, Reed and Maniatis (1986)
showed that sequences located downstream from intron 1 in the human β-
globin gene splice acceptor site, for example, can have a profound effect on
the efficient use of the adjacent splice site. It has also been found that the
interaction between factors present in a splicing extract and the splice sites is
affected by exon sequences, which may play a key role in distinguishing
between normal splice sites and cryptic splice sites (Somasekhar and Mertz
sequence such as the one in RBF18, therefore, might improve the chances of
recognition and/or the affinities of splicing components for the cryptic site,
thus giving rise to a stronger and more stable splice complex. It is possible
that under some circumstances, or in a specific cell type, the cryptic site is
used either exclusively or in combination with the real site. Another
possibility in RBF18 is that the mutation affects the splice sites in such a way
that exon 20 is spliced out completely. However, many factors affect splicing
and, without a functional assay, it is difficult to predict the outcome. For
example, a G-->T transversion in exon 22 of the RB1 gene in a SCLC cell
line (table 4.1) simultaneously created a stop codon and a novel splice donor
site (Yandell et al 1989). However, the mutation must have also influenced
the normal splice acceptor site immediately upstream of exon 22, as it
resulted in the removal of the entire exon (Horowitz et al 1990). Similarly, a
two base mutation GC-->TT in exon 22 of RB1 in the Lu-24 cell line (table
4.1) gave rise to abnormal splicing resulting in the deletion of exon 22 (Mori
et al 1990). As also mentioned earlier, the C-->A mutation in exon 2 of the
RB1 gene in the Lu-65 cell line gave rise to a stop codon (table 4.1).
However, transcripts missing exon 2 sequences were also detected in addition to the ones carrying the mutation (Murakami et al 1991). Other than these unexpected effects on splicing, out of 20 point mutations recorded in table 4.1, more than half directly affected a splicing site. Half of the point mutations in RB1 from Rb patients also occurred at splicing sites. Of the 12 point mutations in Rb cases, 4 were from bilateral cases, 5 were from unilateral cases and 2 were from families that had both bilateral and unilateral members. The laterality in one case was not known. Two of the point mutations occurred in the promoter region presumably affecting transcription and these were from families exhibiting a mild phenotype (see Section 4.3.3).

It seems, therefore, that the majority of physical changes observed in RB1 are point mutations the likes of which were observed also in other tumour suppressor genes: p53 (Hollstein et al 1991, Mazars et al 1992, Iavarone et al 1992) and the APC gene (Miyoshi et al 1992, Horii et al 1992) although the spectrum of point mutations were found to be different in different tumours in p53 (Hollstein et al 1991). In the RB1 gene, the observed physical changes predominantly give rise to formation of premature stop codons and/or splicing abnormalities in Rb and mostly affect splicing in other types of cancer (8/11 non-Rb cancers/cell lines in table 4.1). Missense mutations, however, are very rare. This especially contrasts with the mutations of the p53 tumour suppressor gene (Levine et al 1991) most of which are missense mutations. Although the mutations from exon 13 to 23 region in RB1 constitute over 60% of the mutations compiled in table 4.1 recent, more extensive, studies that have been carried out by the ICRF group and others confirm that there are no hot spots for point mutations in the RB1 gene (A. Hogg, DW Yandell, personal communication) other than CpG dinucleotides (BI Ludeke, personal communication) and that point mutations are distributed randomly throughout the gene. This, again, is in contrast to the p53 gene, where the majority of mutations are clustered in highly conserved regions of the gene and there are hot spots or codons frequently found mutated (Levine et al 1991). Moreover, the frequency and distribution of these hot spots differs among cancers from different tissue types (Levine et al 1991, Levine 1992). The fact that similar physical changes give rise to different functional consequences in the p53 and the RB1 gene reflects differences in mechanisms which lead to impairment of function in the two genes. In p53 most mutations increase the stability of p53 mutant protein and result in a stability of function rather than an absence and the mutant p53, by binding to normal
p53, inhibits its normal function (Section 1.5). This kind of dominant negative affect is not seen in pRB and loss of tumour suppressor ability requires major disruption of the Rb protein. The number of missense mutations capable of inactivating the pRB, however, may be very small.

Identification of causative RB1 mutations provides a direct means of genetic screening for carrier status and is the only means of such screening in sporadic cases. With the available intragenic polymorphisms of RB1 and linkage analysis, although it is now possible to offer screening to the majority of Rb families, there will always be some families uninformative with all the available probes. RBF62 (Section 3.3.1) was one such family but we have now identified the causative mutation and so genetic screening is now possible. Direct genetic screening is also possible for all the other families where the causative mutation is now known. In family RBF58, for example, the causative mutation had already been known at the time of fetal screening for II.3, and by sequencing DNA from the CV sample, the fetus was shown to carry the causative mutation (Section 3.3.1). Direct screening is even simpler to carry out for families where the mutation affects a restriction enzyme site as in family RBF62 and family RBF29 where the mutations affects a Dde I site (Section 3.3.1) and a Hpa II site (Section 3.3.2) respectively. It will be relatively simple in the future to identify mutant gene carriers in these families by restriction enzyme digestion of PCR amplified DNA products.

4.3.2 Possible mechanisms of mutagenesis in RB1

Gene mutations may arise either as a consequence of endogenous error-prone processes such as DNA replication and recombination and repair, or as a result of exposure to exogenous factors such as chemical mutagens or ionising/UV irradiation. Induced mutations are usually many orders of magnitude more frequent than their spontaneous in vivo counterparts and most probably arise by different mechanisms compared with spontaneous mutations. Although the majority of RB1 mutations are thought to arise through endogenous in vivo processes, there is evidence suggesting that radiation might be involved as a causative factor in at least some of the mutations (Morris et al 1990). In particular, treatment of Rb patients with
external beam irradiation is well known to be the cause of second tumours (Section 1.3.8). One of the aims of gene mutation analysis is to search for the underlying mechanisms of mutagenesis. Although the molecular mechanisms involved in spontaneous mutagenesis is still not well understood, there is evidence that some distinctive sequence motifs, such as interspersed repetitive DNA sequences, direct and inverted repeats, and dyad symmetrical elements, are involved in the generation of mutations. This is particularly true for deletions and insertions. When the sequences flanking the small deletions and insertions identified in this project were analysed, all were found to be part of short direct repeats (Section 3.3.1, Figures 3.30b, 3.31b, 3.32c, 3.33b). The breakpoints of the large deletions detected were not characterised, however, and it is not known whether they also contain such sequences at their breakpoints.

Short direct repeats have previously been described at the breakpoints of both large (Canning and Dryja 1989, Hashimoto et al 1991) and small deletions (Dunn et al 1989, Lohmann et al 1992) in the RB1 gene as well as in other genes such as the E. coli lacI gene (Farabaugh and Miller 1978) and the human β-globin gene (Efstratiadis et al 1980). In fact, analysing many eukaryotic deletions, Roth et al (1985) concluded that short direct repeats, of at least 2 bp, occurred more frequently than would be expected from random breakage and reunion. Krawczak and Cooper (1991) reported the presence of 2-8 bp direct repeats in all but one of the 60 small (<20 bp) deletions analysed at many different genetic loci. Efstratiadis et al (1980) argued that such repeats are not long enough to support recombination between chromosomes and postulated that deletions arose instead by 'slipped mispairing' during DNA replication. In this model, originally proposed by Streisinger et al (1966) to explain the generation of frameshift mutations, the repeat sequences mispair at the replication fork leading to the formation and excision of a single-stranded loop containing one repeat unit and the intervening sequences. The extent of the deletions, in this model, depends on the exact excision points within the loop, but one direct repeat sequence is usually lost together with some intervening sequences. Canning and Dryja (1989) found this model consistent with their observations on the large deletions of the RB1 gene and proposed that this mechanism may predominate in the generation of deletions at this locus. Krawczak and Cooper (1991), however, found that deletion of one whole repeat copy occurred rarely (8/59) in small deletions. In both the exon 3 and exon 17 deletions reported here,
short imperfect direct repeats were observed but it was difficult to predict the site of deletion accurately using the slipped mispairing model. Alternative mechanisms involving the repeats exist, however. One possibility is the transient mispairing of two imperfect repeat units at some stage during DNA replication giving rise to the looping out and excision of the non-homologous bases. In exon 3, for example, transient mispairing of the two repeat units, ACAGAAA and ACATAGAAA, could have resulted in the excision of the unpaired 'AT'. Alternatively, the first repeat is a quasi repeat which might have been copied producing the deletion. Another mechanism might involve short runs of the same bases. In exon 17, for example, one of the two 'AA' nucleotides is deleted. At this locus, misalignment of newly synthesised DNA strand by one base pair could have resulted in the excision of the missed, unpaired 'A'.

Short direct repeats were not the only sequence motifs identified around the exon 3 and 17 deletions, however. In exon 17, there are also two sequences (boxed in figure 3.32c), starting 20 bp upstream and 6 bp downstream of the deleted nucleotide, that are inverted repeats of which parts are also dyad symmetries of each other (T'Ang et al 1989). There are similar sequence motifs in exon 3. The sequence between 26 bp and 3 bp upstream of the deleted nucleotides in exon 3 and a sequence (11th-32nd nucleotides) in exon 4 are inverted repeats and a part of this sequence, including the deleted area, is repeated in the 3' end of the Rb1 gene (T'Ang et al 1989). 27 of 60 deletions studied by Krawczak and Cooper (1991) contained inverted repeats which either flanked or span the deletions. Frequent co-occurrence, as in the case of exon 17 deletion reported here, of inverted repeats and direct repeats at sites of deletions were also noted (Krawczak and Cooper 1991). The existence of such sequences, which can form stem and loop structures, tend to suggest that they can potentially be hot spots that promote structural rearrangements. Nalbantoglu et al (1986), for example, showed that inverted repeats were present at, or near, the end points of deletions in the hamster aprt (adenosine phosphoribosyltransferase) locus and proposed that hairpin loops might constitute recognition structures for eukaryotic endonucleases such as topoisomerase I. Shew et al (1990a) observed a sequence which was self-complementary over 12 nucleotides in the 43 bp deletion sequence they identified in the Rb1 gene. They speculated that single-stranded DNA could form a hairpin loop in this region, perhaps during DNA replication, which might be followed by one or more unknown steps leading to its excision.
Structural motifs similar to the ones that were identified around deletions were also identified around the insertions in exon 13 and exon 20 which suggests that similar mutational mechanisms might be involved in insertional mutagenesis. In exon 13, not only was the insertion a part of 8-9 bp direct repeat a part of which is symmetrical (Section 3.3.1, figure 3.31b) but this particular sequence was also found repeated in exon 10, 11 and 27 of the RB1 gene (T'Ang et al 1989). The exon 20 insertion occurred within a 4 bp direct repeat sequence (Section 3.3.1, figure 3.33b). The CTTT(T) repeat of RB1.20 also lies 150 bp downstream to this insertion. The insertion of 'T' to the 'TCT' sequence created a 'TCTT' sequence in exon 20; 5 bp upstream of the insertion position there is another 'CTT' motif. Transient misalignment during DNA synthesis or mispairing at the replication fork with such sequences might have given rise to the incorporation of another 'T'. The exon 13 insertion also involved the incorporation of another 'A' besides an existing 'A' nucleotide creating a tandem repeat of the sequence ATA (ATATA-->ATAATA). Misalignment with the ATA sequence and then realignment could have given rise to this insertion creating the tandem repeat or the symmetrical element flanking the insertion (AGGATAATAGGA) might have given rise to secondary structures such as Moebius loops (Cooper and Krawczak 1991) facilitating an insertion. Cooper and Krawczak (1991) concluded that all 20 randomly occurring insertions they studied could have been caused by slipped mispairing, mediated by direct repeats and/or the templated misincorporation of bases (Kunkel and Soni 1988) by secondary structure intermediates whose formation is facilitated by inverted repeat sequences or symmetrical elements. Most importantly both deletions and insertions occurred when a TG A/G A/G G/T A/C sequence was nearby (Krawczak and Cooper 1991, Cooper and Krawczak 1991). This sequence has also been implicated as the putative arrest sites for DNA polymerase α which often contain a GAG motif (Weaver and DePamphilis 1982). DNA polymerase α is the main eukaryotic polymerase which synthesises DNA complementary to template DNA strands during DNA replication. Its detachment from the synthesis complex at the replication fork will arrest DNA synthesis. It has been proposed that dissociation of polymerase α at arrest sites may, by providing a stable single stranded substrate, lead to deletion/insertion of a DNA sequence either by slipped mispairing via a number of different secondary structure intermediates, or by strand-switching or base misincorporation (Krawczak and Cooper 1991, Cooper and Krawczak 1991). One of the DNA polymerase α arrest site sequences (T/AGGAG) fits
perfectly with the deletion consensus sequence of Krawczak and Cooper (1991). Taking the TG as invariant and allowing a maximum of 2 variations in the other 4 bases, similar sequences were observed around all four deletion/insertion mutations identified in this project (Section 3.3.1, figures 3.30b, 3.31b, 3.32c, 3.33b). Also surveying the deletions/insertions given in table 4.1, this consensus sequence was observed nearby (maximum 10 bp upstream or downstream) in 8 of 12 cases. In the case of 10 bp deletion of exon 18 (table 4.1, Dunn et al 1989), for example, a TGAATC sequence formed part of the 10 bp deletion, the 5' breakpoint being between the TG. In the 3 bp deletion of exon 16 reported by Lohmann et al (1992) (table 4.1), the 5' breakpoint was between the last two bases (AA) of the sequence TGACAA. Similar sequences were also observed around deletions in other genes such as the hamster aprt gene (Nalbantoglu et al 1986).

The exon 3 and 17 deletion regions also contain Alu-like sequences (McGee et al 1989). Intron 17, where the 5' breakpoints of the large deletions identified in families RBF34 and RBF31 lie, is the largest of all RB1 introns being approximately 70 kb in length and it contains the RS2.0 VNTR locus. Another VNTR, RB1.20, is relatively nearby in intron 20. The region of the RB1 gene containing exons 2-17 is also rich in Alu-like sequences and contains Kpn I repeats (Section 1.4.4). It is possible that these sequences are particularly susceptible to recombinational events and have been observed at the breakpoints of deletions in the low-density lipoprotein receptor locus, for example (Lehrman et al 1985). These repeat sequences, however, are less likely to be involved in the mutational mechanisms that gave rise to the exon 3 and 17 deletions, as the loss of one or a few base pairs was found to be more compatible with replication-based models than with recombination-based models (Krawczak and Cooper 1991). The high concentration of Alu and Kpn I repeats in the region covering exons 7-17 of the RB1 gene (Section 1.4.4), however, might provide one explanation for the breakpoint cluster observed in the region for large deletions involving exons 13-17 (Ciccarelli et al 1991). The 3' breakpoint of one deletion (Canning and Dryja 1989) lay within a Kpn I repetitive DNA sequence in intron 17 but not in other deletions (Hashimoto et al 1991). It is too early to determine whether recombination due to repetitive sequences are involved in the large deletions in RB1 and many more deletion breakpoints need to be characterised to resolve this issue.
Mechanisms involved in deletions/insertions might also be involved in point mutations because misalignments and mispairings might lead to base substitution depending on the template sequence (Kunkel and Soni 1988). Base substitutions could arise by other means, however, involving chemical (eg depurination, Loeb and Preston 1986), or enzymatic (eg postreplicative mismatch repair, Modrich 1987) mechanisms. One of the point mutations identified in exon 20, the C—>T transition occurring within C_pG dinucleotide, for example, probably resulted from spontaneous hydrolytic deamination of 5-methylcytosine to thymine (Duncan and Miller 1980). In an investigation of relative mutability of dinucleotides, a 14-fold higher probability of mutation at a C_pG dinucleotide was found compared with the mutability exhibited by any of the other 15 dinucleotides (Cooper and Krawczak 1990). Non-randomness of mutations was also detected among non-C_pG mutations. In one study, in a sample of 95 non-C_pG mutations, G was found to be the most mutable among bases (Cooper and Krawczak 1990). The other point mutation identified in this thesis is a G—>T transversion which occurred in the sequence context CAGAA—>CATAA in exon 20. A large proportion of the observed non-C_pG mutations are thought to arise through misincorporation during the replication/repair process (Cooper and Krawczak 1990). DNA polymerases are involved in all these processes and, therefore, the fidelity of their base incorporation might be an important factor. Kunkel and Alexander (1986) calculated the frequencies of formation of specific mispairs for DNA polymerases α and β. When these mispairing frequencies were compared with the observed frequencies of single base pair (non-C_pG) substitutions, a significant correlation between the two was found (Cooper and Krawczak 1990). Among the models for the polymerase-associated base misincorporation is simple base mispairing (Topal and Fresco 1976) and misalignment/dislocation mutagenesis involving the transient misalignment of the primer template during DNA replication resulting in mispair formation and incorrect base incorporation (Streisinger et al 1966, Kunkel and Alexander 1986, Kunkel and Soni 1988). The G—>T transversion in exon 20 could have evolved by any one of these mechanisms. 8 bp downstream of this mutation there is a CATATA sequence which, by misalignment during DNA replication, could give rise to this base substitution. Interestingly, 4 bp upstream of the mutation there is also a sequence, TGAGCA, (figure 3.38, Section 3.3.2) similar to the arrest site sequences of DNA polymerase α which might facilitate a variety of physical mechanisms. The CTTT(T) repeats of RB1.20 lies 139 bp downstream to
this transversion and such repeats might mediate transient misalignment and/or create a suitable environment whereby mispairing involving a 'T' nucleotide is tolerated and protected from repair. In one study, in the process of analysis of B-DNA crystal structures, a base-pairing shift which was favoured by (CA)_n tracts was observed in the major groove of the B-DNA molecules (Timsit et al 1991). In this conformation the complementary base-pairing was found to be disrupted and the bases formed interactions with the 5' neighbour of their normal complement. It has been suggested that such sequence dependent transient misalignments could stabilise mismatches and could be involved in substitution mutagenesis (Timsit et al 1991). The mutability of a given nucleotide is influenced by neighbouring bases (Mendelman et al 1989) and incorporation of a specific base at a given locus might depend upon the state of the available nucleotide pool and upon the local environment of the DNA sequence (Cooper and Krawczak 1990). It is possible that this 'local environment' is not just the immediate neighbours but extends beyond neighbouring bases as has been demonstrated in prokaryotes (Conkling et al 1980). It has also been suggested that transient misalignment can operate over great distances (Kunkel and Soni 1988). Interestingly, all three mutations identified in exon 20 involved the misincorporation/insertion of a 'T' nucleotide. It is tempting to suggest that the sequence in and around exon 20, perhaps the CTTT(T) VNTR just 3' to the exon, might have a role in the incorporation of this base in all three mutations. Only two small deletions (Mori et al 1990, Lohmann et al 1992) and no point mutations have so far been reported in exon 20. In one of these deletions a 'T' is deleted from the sequence CCCTT(T) (Lohmann et al 1992) which could have arisen by misalignment or by mispairing with similar local motifs. The other deletion involved the removal of CCGG (or CGGC) from CTCCGGCT leaving CTCT (Mori et al 1990) and might have involved similar mechanisms. Interestingly this is the locus of the C-->T transition (CCGG-->CTGG) reported in this thesis. It will be interesting to see whether any other similar mutations will be observed in this exon.

The results of the analysis of sequences around RB1 mutations identified in this project support the involvement of various sequence motifs in the generation of mutations in the RB1 gene. For each mutation alternative mechanisms exist, which are in some cases not mutually exclusive, and it is possible that more than one mechanism acting in combination gave rise to the particular mutations observed. Although the mechanism involved in the
generation of G→T transversion in exon 20, for example, might be replication-dependent as distinct from the probable chemical nature of C→T mutation, both might have occurred under the promotive and protective influence of a strong structural motif provided by the CTTT(T) VNTR. Although mutations seem to be randomly distributed throughout the RB1 gene at the moment, as more mutations are characterised analysis of mutational mechanisms might identify areas of the gene more likely to give rise to mutations which would be particularly useful in the clinical application of mutation screening in the RB1 gene.

4.3.3 The role of pRB in tumorigenesis and possible functional consequences of RB1 mutations

RB1 is a tumour suppressor gene and its protein product, pRB, is present only when both its genomic structure and mRNA transcript appear normal (Lee et al 1987a, Horowitz et al 1990). The phenotypic consequence of the absence of a normally functioning pRB is tumour formation in the eye. Also, the loss or impaired function of pRB is at least a contributory factor in tumour formation in various other soft tissues and in bone tissue (Weichselbaum et al 1988, Harbour et al 1988, Stratton et al 1989, Varley et al 1989, Bookstein et al 1990b, Ishikawa et al 1991, Takahashi et al 1991). Although many RB1 gene mutations have been described in both Rb and other tumours, the exact functional consequences have not been determined in many of them. The ones that have been investigated, however, gave clues about the possible effects of certain gene mutations on pRB and contributed to the understanding of its function.

As explained in Section 1.4.5, pRB seems to be part of a signalling pathway controlling cell proliferation (Weinberg 1991). In this pathway the role of pRB might be that of a transducer conveying extracellular signals into the cell nucleus (Weinberg 1992). As pRB is a tumour suppressor protein, the nature of the signals it responds to are the ones that induce cells to stop proliferation. pRB, in nucleus, might act to process and transduce these affarent anti-proliferation signals by releasing efferent signals which in turn ultimately cause a shutdown of cell growth. The elucidation of exact biochemical mechanisms of this process involved many complex analyses.
pRB was shown to be phosphorylated during progression through the cell cycle; underphosphorylated forms predominating in G1, whereas phosphorylated forms appear as cells enter S phase and persist during the G2 and M phases (Chen et al 1989, DeCaprio et al 1989, Buchkovich et al 1989, Mihara et al 1989) (Section 1.4.5). The earlier discovery that pRB associates with viral oncoproteins (DeCaprio et al 1988, Whyte et al 1988, 1989, Dyson et al 1989) (Section 1.4.5) has provided a powerful experimental tool for studying pRB function. The pRB domains involved in oncoprotein binding (amino acids 393-572 and 646-772) were defined by direct mutagenesis of the RB1 cDNA (Hu et al 1990, Huang et al 1990, Kaelin et al 1990). The link between oncogenes and tumour suppressor genes offered a possible model for the growth suppressing ability of pRB and how viruses make cells grow. In order to replicate their genome viruses need the cellular DNA synthesising machinery. By binding to pRB, early viral proteins may compromise or inactivate pRB function and liberate the cell from the growth constraints imposed by pRB. It has been proposed that the viral oncoproteins mimic the structure of a cellular protein(s) binding of which is critical to the growth suppressor function of pRB. Indeed several cellular pRB-associated proteins were identified (Kaelin et al 1991, Huang et al 1991, Bandara and La Thangue 1991, Bagchi et al 1991, Chittenden et al 1991, Defeo-Jones et al 1991). It was observed that viral oncoproteins and cellular endogenous proteins fail to bind or associate with pRB containing mutations involving the pocket region (protein binding region of pRB formed with appropriate folding of the protein to facilitate binding, Section 1.4.5) and that such mutant versions of pRB also lose the ability to become phosphorylated (Horowitz et al 1989, Shew et al 1990a, Kaelin et al 1991, Huang et al 1991, Chittenden et al 1991). This observation suggested that the oncoprotein binding 'pocket' contributes to the growth regulatory function of pRB which involves direct interaction with cellular proteins. Among the proteins that associated with pRB were the E2F transcription factor (Bagchi et al 1991, Chittenden et al 1991, Chellapappan et al 1991) and a related factor termed DRTFI (Bandara and La Thangue 1991, Bandara et al 1991, Partridge and La Thangue 1991). Mutant versions of pRB and the normal cell cycle-regulated phosphorylation of the protein were shown to result in the loss of the E2F-pRB interaction (Chellapappan et al 1991) indicating this interaction as a functionally important event. In fact, recently, Qin et al (1992) showed that, in addition to a functional pocket domain, sequences extending from the carboxy-terminal boundary of the pocket to the carboxy terminus of the protein are necessary
for growth suppression thereby extending the minimal region necessary for
growth suppression from residue 379 to 928. Moreover, they showed that
this region is also required for E2F binding suggesting that growth
suppression function and E2F binding may be linked. Many genes including
some that encode proteins that are essential for S phase in the cell cycle, such
as dihydrofolate reductase (DHFR), thymidine kinase and DNA polymerase
α, contain E2F binding sites in their promoter regions (Hiebert et al 1992).
Transfection experiments have shown that the interaction of pRB with E2F,
although having no effect on DNA binding, inhibits the ability of E2F to
stimulate transcription (Hiebert et al 1992). These observations suggested
that at least a part of the antiproliferative activity of pRB is mediated through
an interaction with E2F. The function of pRB-E2F interaction may not be
just to sequester E2F and hence inhibit E2F mediated transcription but the
pRB-E2F complex itself may actively repress the transcription of certain
target genes (Weintraub et al 1992). Among the promoters containing E2F
binding sites are the cdc2 and c-myc gene promoters and transfection assays
demonstrated that pRB can repress c-myc transcription as well as cdc2
transcription dependent on E2F sites (Hamel et al 1992, Dalton 1992). As
mentioned in Section 4.2, autoregulation of RB1 also occurs through its own
E2F binding sites. Previously, Robbins et al (1990) demonstrated that pRB
negatively regulates the c-fos gene through a short sequence, which they
termed the retinoblastoma control element (RCE), in the human c-fos
promoter. Subsequently, Kim et al (1991) proposed that pRB may regulate
transcription of c-fos, c-myc, and transforming growth factor β (TGF-β)
genes, acting through RCE elements in the promoters of these genes.
Whether pRB acts through RCE or E2F sites or both is not yet clear but these
observations suggest a role for pRB in transcriptional regulation of certain
genes. The role of pRB does not seem to be just confined to negative
regulation either, as it has also been demonstrated that pRB can positively
regulate Sp1-mediated transcription from the P4 promoter of the insulin-like
growth factor II (IGF II) gene (Kim et al 1992). pRB has also been shown to
physically bind to c-myc and N-myc proteins in vitro (Rustgi et al 1991)
which might indicate post-translational control by pRB. It was recently
demonstrated that co-injection of pRB with c-myc into cells in early G1 phase
of the cell cycle, inhibits the ability of pRB (but not of p53) to arrest the cell
cycle (Goodrich and Lee 1992). This observation indicates that c-myc and
pRB functionally antagonise one another in the cell and might participate in
the same, or overlapping, regulatory networks. Consistent with this line of
thinking is the observation that the ability of TGF-β to shut-off c-myc expression is prevented in the presence of viral oncoproteins that bind pRB. This effect may be an intermediary in the TGF-β:myc signalling pathway (Pietenpol et al 1990, 1991). As explained in Section 1.1, the three viral oncoproteins that bind to pRB are in the category of oncogenes which exert their effect in the nucleus as does myc. Like myc they are also able to collaborate with a ras oncogene in the transformation of embryo cells (Land et al 1983a, Ruley 1983). This might suggest that, by inactivating pRB (a putative negative regulator of myc) these oncoproteins are able to elicit constitutively expressed myc genes, thus mimicking the mutations that create myc oncogene alleles (Weinberg 1992). The same kind of argument can be raised concerning transcription of c-fos which was also shown to be negatively regulated by pRB (Robbins et al 1990).

Recent gene targeting experiments in mice, which were performed by three separate groups (Lee et al 1992, Jacks et al 1992, Clarke et al 1992), shed new light on the function of pRB. In these experiments, mice carrying a mutation in one of their RB1 alleles (heterozygous mutant) were generated and bred to obtain offspring which were homozygous for the mutation. Homozygous mutants were found to be nonviable and died early in gestation, exhibiting neuronal cell death and defective erythropoiesis (Lee et al 1992, Jacks et al 1992, Clarke et al 1992) in stark contrast to the p53 homozygous mutant mice which develop normally but are predisposed to tumours when they are older (Donehower et al 1992). The defects found in both the nerve- and blood tissues of these embryos were indicative of deregulation of cell division and differentiation. These experiments established that RB1, unlike p53, is an essential developmental gene in the mouse, controlling proliferation and differentiation in the CNS and haematopoietic system and that its absence causes developmental arrest. Transfer of the human RB1 transgene into the mutant mice corrected the developmental defects (Lee et al 1992). Here, there is also supporting evidence for a possible interaction of c-myc and c-fos with pRB. As pointed out by Goodrich and Lee (1992), tissues affected by the absence of pRB coincide with the tissues affected by overexpression of c-myc. Overexpression of c-myc also interferes with haematopoietic differentiation in vitro and in vivo (Cole 1986, Bar-Ner et al 1992). Similarly c-fos over- and/or lack of expression was shown to affect development of bone, cartilage and haematopoietic cells in transgenic and chimeric mice (Ruther et al 1988, 1989, Takao et al 1991, Wang et al 1991,
1992). The mice experiments, however, also produced some surprising and unexpected results, and in some ways, raised more questions than they answered. No apparent developmental defects were observed, for example, in the retinal tissues of these mice. Moreover, none of the heterozygous mutant mice developed Rb, although some were observed to develop brain tumours (Lee et al 1992) and, specifically, pituitary tumours (Jacks et al 1992). In summary, therefore, although there is strong evidence that pRB has a key role in signal transduction network, the precise mechanisms by which loss of pRB leads to malignancy still remains to be elucidated.

Structure-function analyses so far indicate that the pocket region of pRB is related to its function. The region 3' to the pocket has recently been implicated in E2F binding and growth suppression (Qin et al 1992). What could be the functional consequences of the RB1 mutations identified in this project? Although by comparing the phenotype and genotype it may be possible to speculate on the functional consequences of mutations, it is difficult to predict the real consequences with any degree of precision as this requires knowledge on the behaviour of mutant pRB in a developing retinal cell. In families with bilaterally affected individuals (RBF34, RBF31, RBF64, RBF58, RBF59 and RBF62) the mutations were all deletions and insertions. The majority of mutations identified to date in bilaterally, multifocally affected individuals cause major disruptions of the gene and its consequent processing resulting, in some cases, in non-production of the Rb protein (Friend et al 1986, Fung et al 1987, Lee et al 1987a, Goddard et al 1988). The extent of deletions in RBF34 and RBF31 were not determined precisely but would certainly include exons 18-20. The deletions/insertions identified in families RBF64, RBF58, RBF59 and RBF62 would all be predicted to result in the production of truncated proteins, 128, 426, 545 and 671 amino acids long respectively. The C→T mutation generating a stop codon in exon 14 from family RBF25 (with bilaterally affected members) is also expected to result in a truncated protein 454 amino acids long (Hogg et al 1992). If produced, these truncated proteins would be missing all or part of the pocket domain and therefore would be expected to be non-functional as shown for two mutations which cause deletion of exon 21 and exon 22 (Kaelin et al 1991, Templeton et al 1991, Mittnacht and Weinberg 1991, Chittenden et al 1991). The pRB from the primary SCLC reported by Shew et al (1990a), which was missing exon 16, was found to be unphosphorylated and unable to bind LT. Similar defects were observed in the pRB produced
by the RB1 allele containing a deletion of exons 21-27 in the Saos-2 cell line (Shew et al 1990b). Since normal pRB can suppress the neoplastic phenotype in Saos-2 (Huang et al 1988), deletion of exons 21-27 suggests that the C-terminal region of pRB is important for its function. In two other SCLC cell lines (Mori et al 1990) with deletions in exons 20 (Δ 4bp) and 23 (Δ 1bp) no abnormal pRB was detected. This observation suggests that either mutant pRB is not produced at all or the amount of pRB in these cells is under the lowest level detectable by the methods used (immunoprecipitation by antibodies and Western blotting). It is also possible that structurally altered pRB is less stable than normal pRB. Some mutations give rise to unexpected transcripts since they affect splicing. The deletion reported here in exon 3 has such a potential being quite near to the splicing site. It is difficult to predict the kind of transcript that would be produced if splicing is affected but production of a less severely mutated transcript is a possibility with alternative splicing. If exon 3 is skipped, however, this disturbs the reading frame and leads to a new stop codon in exon 4.

The severe phenotype observed in the affected members of the families carrying these constitutional deletion/insertion mutations point to loss of function of pRB. The nature of the first mutations identified in these patients point to the same conclusion although there is the possibility that the mutations produce unexpected results due to splicing effects. It is not possible to determine the nature of the second mutation(s) in any of these patients as tumour material is not available. In 70 % of Rb tumours, however, the initial mutation is duplicated in tumour precursor cells (Cavenee et al 1983, Zhu et al 1989). If this is the case for these deletion/insertion mutations, the severe phenotype observed would indicate that the consequence of these mutations is the loss of function of the Rb protein.

Mutations identified in families showing incomplete penetrance and mild expression of the Rb phenotype are point mutations, both in exon 20 of the RB1 gene (Section 3.3.2). The mutation in family RBF29 results in a non-conservative amino acid substitution, Arg→Trp, in codon 661 whereas in family RBF18 the mutation converts codon 675 from a Glu to a stop codon but there is possibility of alternative splicing as discussed in Section 4.3.1. In family RBF29, there are four unaffected gene carriers and the Rb phenotype is extremely mild with the exception of the bilaterally affected twins. In family RBF18, regressed tumours were observed in four generations. Sakai
et al (1991a) reported mutations in Sp1 and ATF sites in the promoter region of the RB1 gene in two similar families with low-penetrance phenotypes. A plausible explanation for this is that promoter mutations may result in reduced levels of pRB. In many cells the production of sufficient pRB protects them against tumorigenesis but, occasionally, a cell produces insufficient pRB, thereby escaping its normal growth control. Another explanation, especially for the occurrence of regressed tumours, has been suggested by Greger et al (1989) who proposed that epigenetic changes such as DNA methylation may contribute to the formation of spontaneous regression of Rb. The maintenance methylase does not replicate methylation patterns during DNA synthesis with absolute fidelity (Holliday 1987) and progressive loss of methylation may eventually lead to the reactivation of the RB1 gene and suppression of tumour growth in cases where inactivation is due to hypermethylation (Greger et al 1989). Sakai et al (1991b) and Greger et al (1989) found evidence for hypermethylation in the RB1 gene in some Rb tumours they studied, all from unilateral patients. Alternative possibilities exist, however, as not all such families contain mutations in the promoter region or have hypermethylated RB1. Subtle changes in pRB amino acid sequence, for example, may reduce only its functional efficiency, and only when a threshold level of activity is not maintained do tumours develop. The mutation in family RBF29 is possibly one such example with only a single amino acid change. Why, then, did the twins in this family develop multifocal disease? Duplication of a "weak" mutation might still result only in a mild phenotype but if, as in 30% of tumours, the second mutation is more serious the combination could result in multifocal tumour formation. Such independent mutations have been identified in tumours from bilaterally affected Rb patients (Dunn et al 1989). The tumours from RBF29 family members, however, were successfully treated and so not available for analysis. Whether subtle changes anywhere in the gene would result in a mild phenotype or whether specific regions, such as exon 20, are more important is not clear. Two other RB1 amino acid substitutions have been reported (table 4.1). Only one of these, however, was from an Rb patient (bilaterally affected) and it was in exon 18 (codon 567) (Yandell et al 1989). The second one, in exon 21 (codon 706) (Kaye et al 1990) was from a SCLC cell line. In this cell line the resulting protein was found to be defective in phosphorylation, oncoprotein binding and nuclear association (Kaye et al 1990, Mittnacht and Weinberg 1991). When the mutation that gave rise to the amino acid substitution in codon 567 was introduced into a wild-type
clone by *in vitro* mutagenesis and expressed, the resulting pRB was also found to be defective in a similar manner (Templeton *et al* 1991). The question is, however, would the effects of amino acid substitutions anywhere in the pocket be the same? There might be regions where the effect on the protein is minimal depending also on the nature of the substituted amino acid. The fact that amino acid substitutions are very rare in the RB1 gene points to this possibility. The number of missense mutations that are capable of inactivating the RB1 gene might be very small. The ones that give rise to tumours and are, therefore, observed as causative mutations might be the rare missense mutations that are capable of causing enough disruption to inactivate the pRB (eg codons 567 and 706 in exon 18 and 21 respectively) or reduce its functional efficiency in such a way that tumours develop occasionally when a threshold level of activity is not maintained (codon 661 in exon 20).

It is interesting that the other mild-phenotype family (RBF18) also has a mutation in exon 20. At first sight the mutation in codon 675 cannot be described as mild because it gives rise to a stop codon and the resulting protein would be missing 254 amino acids. If there is alternative splicing from the cryptic site (figure 3.37, Section 3.3.2), however, only 23 amino acids encoded by exon 20 (codons 654 to 676) would be missing from the protein. Even then this lesion is in the critical pocket domain and it also interrupts the leucine zipper motif with three out of four leucines being lost, although the functional importance of this motif in pRB is not clear. This region of the protein is proline rich which destabilises the α helix formed by the zipper. In addition, leucine zipper motifs are not present, for example, in the pRB binding domains of viral oncoproteins indicating that these associations are formed through interaction with motifs that are distinct from the leucine zipper motifs (Whyte *et al* 1988, DeCaprio *et al* 1988). There have been reports of shorter Rb proteins, resulting in in-frame deletions of exons 20-22, with impaired biochemical properties (Kaelin *et al* 1991, Templeton *et al* 1991, Mittnacht and Weinberg 1991, Chittenden *et al* 1991, Shew *et al* 1990a). Scheffner *et al* (1991) reported that the deletion of 4 amino acids encoded by exon 20 (as a result of splice acceptor site mutation in intron 19, table 4.1) in the C33A cervical carcinoma cell line produced biochemically defective pRB. The precise relationship between the biochemical properties of pRB and its biological activity in tumour suppression is not clear, however. It is possible that not all biological activity of pRB involves binding through the pocket region. There might exist
proteins that bind regions outside the pocket domain or bind to the pocket domain in a manner analogous to that of proteins which bind to the pRB-related factor p107 (Ewen et al 1992, Faha et al 1992, Cao et al 1992). Cyclin A (another protein that functions to regulate the cell cycle probably by regulating kinase activity) binds to p107 through its pocket region (very similar to pRB pocket) and a missense mutation or a peptide introduced into this region of p107, although abolishing binding of viral oncoproteins, does not abolish binding of cyclin A (Ewen et al 1992, Faha et al 1992). Similarly loss or impairment of binding ability in the pocket region for some proteins might not mean that all biological activity of pRB is lost. An example of this notion is the unexpected observation that a mutant pRB (with deletion of amino acids 733 to 768, five encoded by exon 21 and thirty by exon 22) which failed to bind LT antigen and become hyperphosphorylated during the cell cycle, nevertheless activated transcription through E2F sites in the adenovirus early promoter EIIaE and in c-myc promoter (Hamel et al 1992). Normal pRB represses activation of these promoters and this mutant might prevent the normal pRB from repressing promoter activity correctly by competing for normal cellular factors required in the repression complex. What this demonstrates, however, is that a supposedly defective pRB has biological activity. Moreover research has so far been focused only on the functions of pRB involving binding to other proteins. It has been shown, however, that pRB has DNA binding ability which is intrinsic to its carboxy-terminal region (Wang et al 1990) which ties in well with the recent report from Qin et al (1992) who showed that this region is necessary for growth suppression by pRB. The presence of such a DNA binding domain, which is usually not altered by mutations (the mutation in RBF29 would certainly not alter it, neither the one in RBF18 if alternative splicing occurs from the suggested cryptic site), may signal the ability of pRB to interact physiologically with DNA, perhaps even on a sequence-specific basis. Another problem that needs to be addressed is the fact that, although pRB is ubiquitously expressed in all tissues, only a particular set of tissues seem to be predisposed to neoplasms when its activity is abolished. This suggests that pRB plays different roles in different tissues and/or at specific times during their development. It might turn out that RB1 function is related to its protein binding ability which is required in one cell type whereas its DNA binding ability at a certain stage in development might be crucial in another cell type. The effect of most naturally occurring mutations in various different cell types and at different stages of development have not yet been investigated.
Although gene targeting experiments in mice (Lee et al 1992, Jacks et al 1992, Clarke et al 1992) confirmed the crucial role of pRB in development, the fact that there is no damage to the retina in these mice points to the very special role of pRB in human retinal cells. Precisely it is not known what this role is neither are the functional consequences of single amino acid changes and short proteins in a developing retinal cell. The appearance of unaffected gene carriers and regressed tumour cases in otherwise highly penetrant families point to the existence of a short defined "window of susceptibility" which, if missed, tumours fail to result in or they are arrested at a later stage. It is conceivable that, at this crucial time, "weak" mutations rather than abolishing the critical function of RB1 in a developing retinal cell modify it so that it works less efficiently. One consequence of this modification might be that only occasionally is there insufficient pRB to prevent tumorigenesis. Alternatively, once initiated the transformed phenotype might be overcome by subsequent adequate production of pRB, for example, through progressive loss of methylation in cases where inactivation is due to hypermethylation (Greger et al 1989) or through alternative splicing where stop codons are involved as in the case of family RBF18, resulting in regressed/benign tumours. Such models present attractive explanations of the phenotypes observed in low penetrance/regressed tumour families as it is highly unlikely that the occurrence of regressed tumours and the consistent presence of unaffected gene carriers is unrelated to the nature of first predisposing mutation in these families. Confirmation or otherwise of these theories awaits further analysis of mutations in more families with a low-penetrance phenotype and also the elucidation of the pathways through which pRB suppresses tumorigenicity in the developing retinal cells.

A step towards a better understanding of the role of RB1 in development and tumorigenesis has recently been taken by gene targeting experiments in mice (Lee et al 1992, Jacks et al 1992, Clarke et al 1992) but these experiments, as mentioned before, also produced some unexpected results. It was observed, for example, that heterozygous RB1 mutation in mouse also predisposes to malignancy but not in the retina in stark contrast to the situation in humans. Although it is possible that this species difference is due to the relatively small number of target cells in the mouse retina, the fact that no apparent abnormality exists in the retinal tissues of homozygous embryos suggests that even if homozygously mutant retinoblasts were produced in heterozygous mice, they would not be transformed. Rb has been observed in
mice, however, in an earlier set of experiments (Windle et al. 1990) involving generation of transgenic mice carrying SV40 LT antigen. These animals developed tumours in the retina and the phenotype was heritable with complete penetrance in transgenic offspring. In the light of the recent experiments, it must be concluded, therefore, that Rb formation in transgenic mice must have involved functions of LT beyond its effects on pRB, including, for example, sequestration of the p53 protein. The fact that spontaneous Rb is observed exclusively in humans suggests that human retinoblasts are uniquely sensitive to the loss of RB1 function. This difference in requirement for pRB is not so surprising because, even in the same tissue, the effects of pRB might not be the same in every cell lineage. In mice, for example, in some parts of the nervous system absence of pRB causes cell death whereas in others increased cell proliferation is observed (Lee et al. 1992, Jacks et al. 1992). Concerning the apparent lack of abnormal retinal tissue in homozygous mutant mice, it is possible that homozygous mice die before the retina reaches a susceptible stage and that there is a difference between germline and somatic homozygosity. To test this hypothesis, and to observe lack of pRB function in retina cells, mice with inactivated RB1 in their retinal cells only can be generated. Chimeric mice made between embryonic stem (ES) cells carrying homozygous mutant RB1 gene and transgenic ES cells with the human RB1 gene carrying natural RB1 mutations such as the ones identified in families RBF29 and RBF18 would be a better model. Such experiments would provide insight into the specific processing of these mutations in a natural environment and the effect on the retinal cells, even though the role of RB1 in tumour suppression may be quite different in mice and men. The elucidation of the exact relationship between the genotype and phenotype awaits not only more information on the precise role of pRB in tumour formation but also a large study of mutations in families with widely contrasting phenotypes. Most of the RB1 mutations identified to date have been from spontaneous Rb and non-Rb tumours or cell lines but family studies are limited.

Rb is a rare childhood tumour and at the outset it was hoped that research would offer insights into a highly specialised mechanism of tumorigenesis. To everybody's surprise, however, research in the last five years has proven RB1 to be in the centre of a growth controlling machinery that is crucial to many cell types and has also provided insights into the pathogenic mechanisms exploited by DNA tumour viruses. Oncogenic viruses, cellular
oncogenes and tumour suppressor genes were all proved to be parts of a one large and complex picture. Further interactions in this picture, making it even more complex, will no doubt be unravelled. It is possible, for example, that tumour suppressor genes cooperate in growth-regulatory systems and to prevent oncogenesis just as oncogenes cooperate in transformation. One recent finding in support of this view is the report by Shiio et al (1992) which demonstrated negative regulation of RB1 expression by the p53 gene product. They identified a cis-acting element within the RB1 promoter and showed that wild-type p53 (and not mutant p53) down-regulates RB1 transcription in vitro by binding to this element. It is not known whether such an interaction takes place in vivo, and if it does, what the biological significance is. It is possible that p53 and RB1 belong to different growth-regulatory systems and p53 suppresses RB1 transcription through a negative feedback mechanism which operate in one of the complex pathways in the control of the cell cycle and/or differentiation. This might not be a global but rather a specialised interaction operating in certain cell types and/or in response to specific stimuli. Ongoing research into the biochemical pathways involved will no doubt reveal how pRB and other tumour suppressor proteins promote tumour suppression in the future. Once the function of RB1 and its interactions in a variety of cells are established precisely, improvements in treatment and even gene therapy will become a possibility. In the mean time the knowledge gained so far will continue to benefit Rb families worldwide in the form of genetic screening and counselling.
REFERENCES


Attwood J, Bryant S (1988) A computer program to make linkage analysis with LIPED and LINKAGE easier to perform and less prone to input errors. Ann Hum Genet 52:259


Bar-Ner M, Messing LT, Cultraro CM, Birrer MJ, Segal S (1992) Regions within the c-Myc protein that are necessary for transformation are also required for inhibition of differentiation of murine erythroleukemia cells. Cell Growth Differ 3:183-190


Bischoff JR, Friedman PN, Marshak DR, Prives C, Beach D (1990) Human p53 is phosphorylated by p60-cdc2 and cyclin B-cdc2. Proc Natl Acad Sci USA 87:4766-4770


Chigwin JM, Przybyla AE, MacDonald RJ, Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294-5299


Collett MS, Erikson RL (1978) Protein kinase activity associated with the avian sarcoma virus src gene product. Proc Natl Acad Sci 75:2021-2024


Der CJ, Krontiris TG, Cooper GM (1982) Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. Proc Natl Acad Sci USA 79:3637-3640


Dias PLR (1979b) Correlation of aqueous humor lactic acid dehydrogenase activity with intraocular pathology. Br J Ophthalmol 63:574-577


Dryja TP, Rapaport JM, Joyce JM, Petersen RA (1986a) Molecular detection of deletions involving band q14 of chromosome 13 in retinoblastomas. Proc Natl Acad Sci USA 83:7391-7394


Eckert KA, Kunkel TA (1991) DNA polymerase fidelity and the polymerase chain reaction. PCR Methods and Applications 1:17-24


Fakharzadeh SS, Trusko SP, George DL (1991) Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumour cell line. EMBO J 10:1565-1569


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


302


Hu Q, Dyson N, Harlow E (1990) The regions of the retinoblastoma protein needed for binding to adenovirus E1A or SV40 large T antigen are common sites for mutations. EMBO J 9:1147-1155


Knudson AG (1978) Retinoblastoma: A prototypic hereditary neoplasm. Seminars in Oncology 5:57-60


Lane DP, Crawford LV (1979) T-antigen is bound to a host protein in SV40 transformed cells. Nature 278:261-263


319


McGee TL, Cowley GS, Yandell DW, Dryja TP (1990) Detection of the Xbal RFLP within the retinoblastoma locus by PCR. Nucleic Acids Res 18:207


Michalski AJ, Cotter FE, Cowell JK (1992) Isolation of chromosome-specific DNA sequences from an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13)(q22;q12) and ganglioneuroblastoma. Oncogene 7:1595-1602


Onadim Z, Hogg A, Baird PN, Cowell JK (1992b) Oncogenic point mutations in exon 20 of the RB1 gene in families showing incomplete penetrance and mild expression of the retinoblastoma phenotype. Proc Natl Acad Sci USA 89:6177-6181


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


331


Sheffield VC, Cox DR, Lerman LS, Myers RM (1989) Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc Natl Acad Sci USA 86:232-236


Somasekhar MB, Mertz JE (1985) Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. Nucleic Acids Res 13:5591-5609


Stehelin D, Varmus HE, Bishop JM, Vogt PK (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature 260:170-173


Templeton DJ, Park SH, Lanier L, Weinberg RA (1991) Nonfunctional mutants of the retinoblastoma protein are characterized by defects in phosphorylation, viral oncoprotein association, and nuclear tethering. Proc Natl Acad Sci USA 88:3033-3037


Van Kempen C (1966) A case of retinoblastoma, combined with severe mental retardation and a few other congenital anomalies, associated with complex aberrations of the karyotype. Maandschrift Kindergeneeskd 34:90-95


Verhoeff FH (1921) Glioma retinae treated by x-rays with apparent destruction of the tumour and preservation of normal vision. Arch Ophthalmol 50:450-456


Vogel F (1979) The genetics of retinoblastoma. Hum Genet 52:1-54


Von Graefe (1868) (cited by Weller 1941) Zusatze uber intraoculare Tumoren. Arch F Ophthal 14:103-144


## APPENDIX I

### Restriction Endonucleases

<table>
<thead>
<tr>
<th>Restriction Endonuclease</th>
<th>Source</th>
<th>Recognition Sequence</th>
<th>Optimum Temp. °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afl III</td>
<td>Anabeana flosaquae</td>
<td>A↓CPuPyGT</td>
<td>37</td>
</tr>
<tr>
<td>Alu I</td>
<td>Arthrobacter luteus</td>
<td>AG↓CT</td>
<td>37</td>
</tr>
<tr>
<td>Apa I</td>
<td>Acetobacter pasturiensis</td>
<td>GGGCC↓C</td>
<td>37</td>
</tr>
<tr>
<td>Ase I</td>
<td>Aquaspirillum serpens</td>
<td>AT↓TAAT</td>
<td>37</td>
</tr>
<tr>
<td>Bal I</td>
<td>Brevibacterium albidum</td>
<td>TGG↓CCA</td>
<td>37</td>
</tr>
<tr>
<td>Bam HI</td>
<td>Bacillus amyloliquifaciens</td>
<td>G↓GATCC</td>
<td>37</td>
</tr>
<tr>
<td>Bcl I</td>
<td>Bacillus caldolyticus</td>
<td>T↓GATCA</td>
<td>50</td>
</tr>
<tr>
<td>Bgl II</td>
<td>Bacillus globigii</td>
<td>A↓GATCT</td>
<td>37</td>
</tr>
<tr>
<td>Bsp HI</td>
<td>Bacillus species H</td>
<td>T↓CATGA</td>
<td>37</td>
</tr>
<tr>
<td>Bsp 1286 I</td>
<td>Bacillus sphaericus (IAM 1286)</td>
<td>G(G/A/T)GC(C/A/T)↓C</td>
<td>37</td>
</tr>
<tr>
<td>Bst EII</td>
<td>Bacillus steatherophilus H3</td>
<td>G↓GTNACC</td>
<td>60</td>
</tr>
<tr>
<td>Bst NI</td>
<td>Bacillus steatherophilus N1</td>
<td>CC↓(A/T)GG</td>
<td>60</td>
</tr>
<tr>
<td>Dde I</td>
<td>Desulfovibrio desulfuricans</td>
<td>C↓TNAG</td>
<td>37</td>
</tr>
<tr>
<td>Eco RI</td>
<td>Escherichia coli RY13</td>
<td>G↓AATTC</td>
<td>37</td>
</tr>
<tr>
<td>Hind III</td>
<td>Haemophilus influenzae Rd</td>
<td>A↓AGCTTT</td>
<td>37</td>
</tr>
<tr>
<td>Hinf I</td>
<td>Haemophilus influenzae Rf</td>
<td>G↓ANTC</td>
<td>37</td>
</tr>
<tr>
<td>Hpa I</td>
<td>Haemophilus parainfluenzae</td>
<td>GTT↓AAC</td>
<td>37</td>
</tr>
<tr>
<td>Hpa II</td>
<td>Haemophilus parainfluenzae</td>
<td>C↓CGG</td>
<td>37</td>
</tr>
<tr>
<td>Kpn I</td>
<td>Klebsiella pneumonia 0K8</td>
<td>GGTAC↓C</td>
<td>37</td>
</tr>
<tr>
<td>Mbo II</td>
<td>Moraxella bovis</td>
<td>GAAGAN&lt;sub&gt;8/7&lt;/sub&gt;↓</td>
<td>37</td>
</tr>
<tr>
<td>Mlu I</td>
<td>Micrococcus luteus</td>
<td>A↓CGCGT</td>
<td>37</td>
</tr>
<tr>
<td>Mse I</td>
<td>Micrococcus species</td>
<td>T↓TAA</td>
<td>37</td>
</tr>
<tr>
<td>Msp I</td>
<td>Moraxella species</td>
<td>C↓CGG / C↓&lt;sup&gt;m&lt;/sup&gt;CGG</td>
<td>37</td>
</tr>
<tr>
<td>Nde I</td>
<td>Neisseria denitrificans</td>
<td>CA↓TATG</td>
<td>37</td>
</tr>
<tr>
<td>Nsi I</td>
<td>Neisseria sicca</td>
<td>ATGCA↓T</td>
<td>37</td>
</tr>
<tr>
<td>Pst I</td>
<td>Providenci stuartii</td>
<td>CTGCA↓G</td>
<td>37</td>
</tr>
<tr>
<td>Pvu II</td>
<td>Proteus vulgaris</td>
<td>CAG↓CTG</td>
<td>37</td>
</tr>
<tr>
<td>Rsa I</td>
<td>Rhodopsseudomonas sphaeroides</td>
<td>GT↓AC</td>
<td>37</td>
</tr>
<tr>
<td>Sac I</td>
<td>Streptomyces achromogenes</td>
<td>GAGCT↓C</td>
<td>37</td>
</tr>
<tr>
<td>Sal I</td>
<td>Streptomyces albus G</td>
<td>G↓TCGAC</td>
<td>37</td>
</tr>
<tr>
<td>Sau 3AI</td>
<td>Staphylococcus aureus 3A</td>
<td>↓GATC</td>
<td>37</td>
</tr>
<tr>
<td>Sma I</td>
<td>Serratia marcescens</td>
<td>CCC↓GGG</td>
<td>37</td>
</tr>
<tr>
<td>Taq I</td>
<td>Thermus aquaticus</td>
<td>T↓CGA</td>
<td>65</td>
</tr>
<tr>
<td>Tth III I</td>
<td>Thermus thermophilus III</td>
<td>GACN↓NNGTC</td>
<td>65</td>
</tr>
<tr>
<td>Xba I</td>
<td>Xanthomonas badrii/X.campestris</td>
<td>T↓CTAGA</td>
<td>37</td>
</tr>
</tbody>
</table>

Abbreviations: Pu=purine; Py=pyrimidine; N=any one of four bases; m=methylated
APPENDIX II

Molecular Weight Markers

<table>
<thead>
<tr>
<th>λHHR* (kb)</th>
<th>1 Kb Ladder** (kb)</th>
<th>Hinf I cut pBR322 (kb)</th>
<th>Hinf I cut pAT153 (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>23.13</td>
<td>12.216</td>
<td>1.631</td>
<td>1.631</td>
</tr>
<tr>
<td>21.23</td>
<td>11.198</td>
<td>0.517</td>
<td>0.517</td>
</tr>
<tr>
<td>9.42</td>
<td>10.180</td>
<td>0.506</td>
<td>0.396</td>
</tr>
<tr>
<td>6.56</td>
<td>9.162</td>
<td>0.396</td>
<td>0.298</td>
</tr>
<tr>
<td>5.15</td>
<td>8.144</td>
<td>0.344</td>
<td>0.221</td>
</tr>
<tr>
<td>4.98</td>
<td>7.126</td>
<td>0.298</td>
<td>0.220</td>
</tr>
<tr>
<td>4.36</td>
<td>6.108</td>
<td>0.221</td>
<td>0.154</td>
</tr>
<tr>
<td>4.27</td>
<td>5.090</td>
<td>0.220</td>
<td>0.145</td>
</tr>
<tr>
<td>3.52</td>
<td>4.072</td>
<td>0.154</td>
<td>0.075</td>
</tr>
<tr>
<td>2.32</td>
<td>3.054</td>
<td>0.075</td>
<td></td>
</tr>
<tr>
<td>2.02</td>
<td>2.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.91</td>
<td>1.636</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.58</td>
<td>1.018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.38</td>
<td>0.517</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>0.506</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.83</td>
<td>0.396</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.56</td>
<td>0.344</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.13</td>
<td>0.298</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.220</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.201</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.154</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.134</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.075</td>
<td></td>
</tr>
</tbody>
</table>

*λHHR = a mixture of Hind III cut and Hind III + Eco RI cut bacteriophage lambda DNA

**1 Kb DNA ladder from Gibco-BRL
APPENDIX III

Pedigrees of Rb families

* families from the series of Dr CD Mitchell, which were used in lod score analysis only in this project (Section 3.1.3, Onadim et al 1990).
**APPENDIX IV**

Results* of Rb Screening in 57 Rb Families

<table>
<thead>
<tr>
<th>Family</th>
<th>RS2.0</th>
<th>PR0.6</th>
<th>M1.8</th>
<th>R.06</th>
<th>HS0.5</th>
<th>RB1.20</th>
<th>RB1.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF01</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>RBF02</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF03</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>RBF04</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF05</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>RBF06</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>RBF08</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>RBF10</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF13</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>RBF16</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF18</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF21</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF22</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF23</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF24</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF25</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF26</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF27</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF28</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF29</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF30</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF31</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>RBF32</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF34</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF35</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF36</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF37</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF38</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF39</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF40</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF41</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF42</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF43</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF44</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF45</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>353</td>
</tr>
<tr>
<td>RBF46</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF47</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF48</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF49</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF50</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF51</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF52</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF53</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF54</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF55</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF56</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF57</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF58</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF59</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF60</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF61</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF62</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF63</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF64</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBF65</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*For each family where tested for a particular polymorphism a '+' (informative) or a '-' (not informative) indicates the result.
APPENDIX V

DNA Probes Used For Identification of RB1 RFLPs

<table>
<thead>
<tr>
<th>DNA probe in Bluescribe</th>
<th>Enzyme(s) releasing probe</th>
<th>Probe size (kb)</th>
<th>Restriction endonuclease for RFLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>p123 M1.8</td>
<td>EcoRI / Hind III</td>
<td>1.8</td>
<td>Bam HI</td>
</tr>
<tr>
<td>p95 HS0.5</td>
<td>Hind III / Sal I</td>
<td>0.6</td>
<td>Kpn I</td>
</tr>
<tr>
<td>p88 PR0.6</td>
<td>Hind III / Pst I</td>
<td>0.6</td>
<td>Xba I</td>
</tr>
<tr>
<td>p68 RS2.0</td>
<td>EcoRI / Sac I</td>
<td>2.0</td>
<td>Rsa I</td>
</tr>
<tr>
<td>p35 R0.6</td>
<td>EcoRI</td>
<td>0.6</td>
<td>Tth III I</td>
</tr>
</tbody>
</table>
APPENDIX VI

The Genetic Code

<table>
<thead>
<tr>
<th>2nd position</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>UUU Phe</td>
</tr>
<tr>
<td>UUC Phe</td>
</tr>
<tr>
<td>UUA Leu</td>
</tr>
<tr>
<td>UUG Leu</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>CUC Leu</td>
</tr>
<tr>
<td>CUA Leu</td>
</tr>
<tr>
<td>CUG Leu</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>AUC Ile</td>
</tr>
<tr>
<td>AUA Ile</td>
</tr>
<tr>
<td>AUG Met</td>
</tr>
<tr>
<td>G</td>
</tr>
<tr>
<td>GUC Val</td>
</tr>
<tr>
<td>GUA Val</td>
</tr>
<tr>
<td>GUG Val</td>
</tr>
</tbody>
</table>
CARRIER DETECTION AND PRENATAL SCREENING OF THE RETINOBLASTOMA GENE

Retinoblastoma (Rb) is the most common intra-ocular tumour in children with an incidence of 1:20,000 live births and evidence for both sporadic and hereditary forms.\(^1\) In the familial form, the Rb phenotype follows an autosomal dominant pattern of inheritance with high (90 per cent) penetrance. Since there have been few survivors of metastatic Rb, early detection is essential to save life and vision. By regular screening, during the first decade of their lives, of all members of a family in which there has been an affected individual, early detection of tumours is usually guaranteed although, depending on the exact location of tumour(s) in the eye, a favourable outcome is not always guaranteed. Screening involves regular hospitalization and full ophthalmological examination under anaesthetic.

Since less than half of those examined will actually develop the tumour, this form of 'blanket' screening means that resources are not always focused on those who need them. A system for the identification of predisposition gene carriers would have major clinical advantages since:

1. Early diagnosis of gene carriers would be possible;
2. 'Normal' individuals would be identified and excluded from frequent examination;
3. Older, unaffected gene carriers would be identified; and
4. This screening would be available prenatally.

The probable location of the gene on chromosome 13 in band q14 was first suggested by karyotype analysis of constitutional chromosome deletions.\(^2\) Genetic linkage studies using the adjacent esterase-D (ESD) gene\(^3\) confirmed this assignment. The demonstration of close proximity between ESD and Rb allowed detection of predisposed individuals in two groups: first, those rare cases carrying constitutional deletions of 13q14 with reduced enzyme levels\(^4\)—approximately 3–4 per cent of patients fall into this group,\(^5\) and second, families in which the protein polymorphism identified by Hopkinson et al.\(^7\) was used in linkage analysis and prenatal screening.\(^8,9\) Only about 10 per cent of families in the UK benefit from either of these screening procedures.\(^4\)

A new era dawned in 1986, when the Rb predisposition gene was cloned. Lalande et al.\(^8\) had isolated a random DNA sequence which lay within the genomic sequence of the Rb gene. Using this probe, a cDNA was isolated.\(^9\) Circumstantial evidence for the authenticity of this gene (RBI) as the Rb predisposition gene came from the analysis of tumour cells\(^10–13\) where structural abnormalities of the genomic sequence, as well as in the messenger-RNA (m-RNA), were identified. The subsequent demonstration of abnormalities of RBI in individuals predisposed to Rb provided more conclusive evidence.\(^10–14\)

The entire Rb gene has now been sequenced; it is 4757 base pairs long and spans approximately 200 kb of DNA on chromosome 13. There are 27 exons, all but one of which is between 50 and 150 base pairs long. The cDNA clone, however, cannot be used directly in family linkage studies since it does not identify any restriction fragment length polymorphisms (RFLPs) (see ref. 15 for discussion). However, unique sequences from within the introns of the RBI gene,\(^16\) which recognized relatively frequent RFLPs, were isolated soon afterwards and made genetic linkage analysis and prenatal screening possible. No evidence of recombination has been reported and, in our hands, over 90 per cent of families are informative for these probes. The most useful probe, RS2, recognizes a variable number tandem repeat (VNTR) sequence, with eight alleles. Approximately 70 per cent of individuals are heterozygous for this probe, which allows 'gene tracking' in families.\(^15\) We have recently used this approach for prenatal screening\(^17\) using both chorionic villus sampling and cultured amniotic cells. The majority of the 30 per cent of families uninformative with the RS2 probe are informative for one of four other probes. In our recent study, only 10 per cent of the
families could not be offered (prenatal or perinatal) screening.

It will be clear from the preceding discussion that prenatal screening will be possible for virtually all cases of hereditary retinoblastoma (where there is a demonstrable family history) and when DNA samples can be obtained from the majority of family members. The challenge, however, is to determine whether 'new' cases of Rb are truly sporadic or the start of a hereditary line. All bilateral, multifocal cases are considered gene carriers but unilaterally affected individuals can, on rare occasions, have affected children. Because gene tracking requires at least two affected generations, prenatal screening is not possible. Identification of a specific mutation in the RB1 gene in these cases would prove their carrier status and permit prenatal screening.

The most direct method of identifying these specific mutations would be to analyse the whole gene in a single step. Theoretically, this can be achieved using the RNase protection procedure. A sequence complementary to that of the normal RB1 mRNA can be made relatively easily and when mixed with mRNA from a patient forms a stable duplex. Any abnormalities in the patient’s RB1 mRNA causes a breakdown in the homology and renders the duplex susceptible to cleavage with RNase. The products of this reaction can be resolved on acrylamide gels; normal genes produce a single band whereas mutant genes produce two or more. This technique 'looks' at the whole gene but, for reasons that are not clear, does not identify mutations in every case.

The technique known as the polymerase chain reaction (PCR) can selectively amplify small regions of the genome such as a single exon from within the Rb gene. DNA sequencing of the amplified product can define specific mutations (by comparison with the normal sequence). In this way, gene carriers might be identified even in the absence of a family history. Not all mutations, however, will occur within the coding sequence of the gene defined by the exons. Splice junction mutations as well as point mutations generating stop codons in the flanking intron regions will also result in an abnormal mRNA. In addition, mutations in the promoter region of the gene will sometimes affect function without affecting the coding sequence.

The Rb gene has large (> 50 kb) introns and it is not practicable to analyse their sequence entirely; initially, therefore, only the intron sequences flanking the exons will be analysed. These sequence data are now available for all of the 27 exons16 and several mutations have been detected following PCR amplification and sequencing.19 At present, this type of analysis is time-consuming because the site of the mutation is usually unknown and exons are analysed at random.

During the course of the sequencing study by McGee et al.18 another VNTR, based on a 4-base pair repeat, was identified in the 3' region flanking exon 20. In this case, 24 alleles have been recognized so far following PCR amplification. This analysis is complex, since alleles may differ in size by only one base pair and acrylamide sequencing gels are needed to resolve the PCR products. Using this procedure, it has been estimated19 that 99 per cent of individuals are heterozygous at this locus. Theoretically, this type of analysis could replace standard RFLP analysis.16 However, PCR is more time-consuming, expensive, and prone to technical error and, in our laboratory, only those families uninformative with standard RFLP procedures will be analysed in this way.

Tumorigenesis in Rb is due to the 'unmasking' of a predisposing mutation.20 Though tumour cells are homozygous for the mutation,11 gene carriers are constitutionally heterozygous. Thus, analysis of the RB1 gene in somatic cells using either RNase protection or PCR sequencing is always complicated by the presence of the normal allele. It is essential, for this reason, that tumour cells are made available for analysis since only the mutant alleles of interest are present. The demonstration of a mutation in constitutional cells defines the patient as a gene carrier.

As more and more data accumulate and patterns emerge, it may be that this molecular pathology will be able to predict the course of the disease in terms of invasiveness/prognosis, whether patients will be bilaterally or unilaterally affected, and their susceptibility to 'second tumours', a frequent complication in gene carriers.

ACKNOWLEDGEMENTS

We would like to thank Dr J. Pritchard for his critical reading of the manuscript. Z. Onadim is supported by a grant from the David Allen Retinoblastoma Appeal.

J. K. COWELL AND Z. ONADIM
ICRF Laboratory of Molecular Genetics
Institute of Child Health
London WC1 1EH, U.K.
REFERENCES


Application of intragenic DNA probes in prenatal screening for retinoblastoma gene carriers in the United Kingdom

Z O ONADIM, C D MITCHELL, P C RUTLAND, B G BUCKLE, M JAY, J L HUNGERFORD, K HARPER, AND J K COWELL
Application of intragenic DNA probes in prenatal screening for retinoblastoma gene carriers in the United Kingdom


Abstract

Restriction fragment length polymorphisms (RFLPs) in 55 families affected by retinoblastoma have been studied using recombinant DNA probes derived from within the retinoblastoma predisposition gene. Only six families were uninformative for any of the DNA polymorphisms. The remaining 49 families can be offered prenatal screening. No obligate recombination between any of the polymorphic loci and the retinoblastoma phenotype were observed. Four previously unknown cases of non-penetration were identified. Prenatal testing for the inheritance of mutant alleles was performed in two cases and perinatal screening in two additional cases. One fetus inherited the normal allele from the affected parent and is therefore not at risk of retinoblastoma; the second fetus inherited the mutant allele and will require frequent screening for early detection of retinoblastoma. Both perinatal tests showed the absence of the mutant allele.

Retinoblastoma, although rare, is the most common ophthalmic malignancy in childhood. It has been the focus of a large amount of research, because in about one third of cases the predisposition to cancer is dominantly inherited from affected parents. Early detection of retinoblastoma ensures a better prognosis, both for life and for vision. Until recently the only available method of early detection was regular full ophthalmological assessment under anaesthetic. In practice this means that all children at risk of retinoblastoma on clinical grounds are screened. As the retinoblastoma mutation is not fully penetrant (not all subjects carrying the dominant mutant gene show the mutant phenotype) and the disease is able to 'skip' a generation, even children of unaffected subjects from within an affected family must be considered 'at risk' as well as those of apparently sporadic cases. Thus there is considerable investment in a small number of patients of whom less than half ultimately develop tumours. In addition, many patients with retinoblastoma have elected not to have children because of the risk of an offspring being similarly affected. The ability to detect the inheritance of normal or mutant retinoblastoma alleles (one of a series of possible alternative forms of a given gene differing in DNA sequence) would allow clinical screening procedures to focus on those patients at high risk of developing retinoblastoma, and make pregnancy, with prenatal testing and termination of predisposed fetuses, an option for those families not wanting to have an affected child.

The discovery that the esterase-D gene lay in chromosome band 13q14 meant that a protein polymorphism (the existence of two or more genetically different classes in the same inter-breeding population) for that enzyme could be used for 'gene tracking' in families with retinoblastoma. Although close linkage was firmly established, the low incidence of heterozygotes in the population made it impractical to apply this technique widely. Randomly isolated DNA sequences, which had been mapped around the 13q14 region, and which recognised restriction fragment length polymorphisms (RFLPs—variations occurring within a species in the length of DNA fragments generated by a specific endonuclease) were used for prenatal screening. Success was, however, limited because of recombination (the occurrence of progeny with combinations of genes other than those that occurred in the parents as a result of independent assortment or crossing over) between these markers and the retinoblastoma locus. In 1986 Friend et al isolated a candidate retinoblastoma gene that showed frequent structural abnormalities in retinoblastoma tumour cells. Mutations of this gene have now been found in somatic cells of predisposed subjects. It is not possible to recognise RFLPs directly using this complementary DNA (cDNA), but recently Wiggs et al isolated a series of DNA probes from within the genomic sequence that recognise high frequency RFLPs. We have investigated the segregation of alleles detected by these sequences in a large series of families with retinoblastoma and report on their application to family linkage studies.

Subjects and methods

Most families were identified in the ophthalmic oncology clinics at St Bartholomew's Hospital and Moorfields Eye Hospital, although some were submitted from other regional centres. In each case the diagnoses of affected individuals has been confirmed by an experienced ophthalmologist (J L H). Blood samples were collected from as many relevant family members as possible and stored at −80°C until required. For many of the key family members lymphoblastoid cell lines were also generated. Prenatal samples were obtained either by chorionic villus sampling or by amniocentesis.

DNA was prepared from white blood cells, lymphoblastoid cell lines, chorionic villi, or cultured fetal fibroblasts using standard methods. Between 2–5 μg DNA from each...
member of the family was digested with the appropriate restriction endonuclease, size fractionated by electrophoresis through 1% agarose gels, and transferred to nylon membranes (Hybond-N) as described by Southern. All DNA probes were labelled to high specific activity by the oligopriming extension method.

Prehybridisation and hybridisation were carried out in quadruple strength 0-15M sodium chloride and 0-015M sodium citrate (SSC) sonicated boiled salmon sperm, 10 µg/ml polyadenylic acid and 0-1% sodium dodecyl sulphate (SDS) at 65°C for 16 hours. Filters were then washed at 65°C for two washes of 15 minutes each in double strength SSC/0-1% SDS, and then exposed to Kodak XAR-5 autoradiographic film with intensifying screens for 16-72 hours.

Family pedigree data were prepared for analysis using the program LINKSYS, and log of the odds scores were calculated using the program LIPEP. Assuming that the frequency of new gene mutations gamete is 1-5x10^-5, data were analysed for a penetrance (the proportion of subjects in a specified group that show the expected phenotype under a set of environmental conditions) of 0-9.

Results

PATTERN OF INHERITANCE

In our series there were 17 families in which subjects had either an affected parent or sibling, and also an affected child. Three of these 29 subjects were not themselves affected and ophthalmological examination of their retinas failed to detect evidence of regressed tumours. From these observations we assess the penetrance of the retinoblastoma gene as 90% and we have used these figures to calculate the log of the odds scores.

The inheritance of the retinoblastoma phenotype follows the normal pattern of autosomal dominant inheritance in most families (fig 1). Family RBF-12 is unusual in that each of the affected members has only a unifocal tumour or evidence of spontaneous regression (fig 2). Phenotypic analysis indicates that there is an apparently unaffected transmitting member (II.7). Unfortunately this subject was not available for ophthalmological examination; it is possible he also had regressed tumours, especially as evidence of incomplete penetrance in another member of the family was detected using molecular probes (fig 2). The log of the odds score for this family was -0.03 at 0 = 0 and 90% penetrance and 0-178 at 0 = 0 and a penetrance of 80%. 0 Represents the proportion of recombinants out of all the opportunities for recombination. Assessing the log of the odds score, assuming 90% penetrance, the maximum value of 0-96 was obtained at 0 - 0-2 (table 1). In family RBF-04 three affected siblings were

Results

PATTERN OF INHERITANCE

In our series there were 17 families in which subjects had either an affected parent or sibling, and also an affected child. Three of these 29 subjects were not themselves affected and ophthalmological examination of their retinas failed to detect evidence of regressed tumours. From these observations we assess the penetrance of the retinoblastoma gene as 90% and we have used these figures to calculate the log of the odds scores.

The inheritance of the retinoblastoma phenotype follows the normal pattern of autosomal dominant inheritance in most families (fig 1). Family RBF-12 is unusual in that each of the affected members has only a unifocal tumour or evidence of spontaneous regression (fig 2). Phenotypic analysis indicates that there is an apparently unaffected transmitting member (II.7). Unfortunately this subject was not available for ophthalmological examination; it is possible he also had regressed tumours, especially as evidence of incomplete penetrance in another member of the family was detected using molecular probes (fig 2). The log of the odds score for this family was -0.03 at 0 = 0 and 90% penetrance and 0-178 at 0 = 0 and a penetrance of 80%. 0 Represents the proportion of recombinants out of all the opportunities for recombination. Assessing the log of the odds score, assuming 90% penetrance, the maximum value of 0-96 was obtained at 0 - 0-2 (table 1). In family RBF-04 three affected siblings were

<table>
<thead>
<tr>
<th>Families</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF-04</td>
<td>0-048</td>
</tr>
<tr>
<td>RBF-12</td>
<td>0-030</td>
</tr>
<tr>
<td>Total</td>
<td>0-018</td>
</tr>
</tbody>
</table>

Table 1 Log of the odds scores at different values of 0 for the two families that showed evidence of incomplete penetrance
born to unaffected parents. In this family it was not possible to determine whether one of the parents was a true case of incomplete penetrance or whether one or the other was a gonadal mosaic for the mutation. The maximum log of the odds score in this family assuming a penetrance of 90% was 0.058 at $\theta = 0.1$. The log of the odds scores for 20 informative families were calculated and are shown in Table 2. The cumulative scores, including that derived from family RBF-12, was 5.947 at $\theta = 0$ and 90% penetrance.

**GENE TRACKING WITH INTRAGENIC PROBES**

Five unique DNA sequence probes from within the genomic sequence of the 4.7R (RB1) gene were used in this study (Table 3). Each probe recognises RFLPs that occur sufficiently frequently to be of value in linkage analysis.

Fifty five families were analysed, and they form two groups. The first (n=22) contained at least two children and had affected family members in at least two generations, in most cases allowing log of the odds scores to be calculated. The pedigrees of these families are shown in Fig 1. In the second group of families (n=33), either there was only one affected child, or not all family members were available for testing. All the families were selected because prenatal screening could be offered to at least one family member if requested.

Altogether, five informative probes were used (Table 3). The probes were digested with restriction enzymes that did not cleave the probes themselves, and the fragments were separated by electrophoresis in agarose gels. The polymorphisms identified by these probes were found to be informative, with allele frequencies ranging widely. The polymorphism identified by RS2-0 detected a VNTR of roughly 50 base pairs, with eight alleles so far detected. To separate these alleles, DNA (2-3 pg) is run overnight in a 0.4 cm thick, 1.2% agarose gel until lambda markers below 1.0 Kb in size have migrated off the bottom of a 20 cm gel. Under these conditions it is usually possible to distinguish between alleles differing in size by only 50 base pairs. Variations in gel thickness, however, conditions which result in non-homogeneity of the gel, or local heat gradients during electrophoresis can distort the migration rate of the DNA restriction fragments. Thus it is difficult to establish the sizes of alleles and compare them among gels. This almost certainly accounts for the discrepancy between our estimate of allele frequencies and those reported by Wiggs et al. Within a single gel, however, the mendelian inheritance can easily be followed (Fig 3). In some families only key members were analysed with this probe. If they were heterozygotes the analysis was extended to the remaining family members. Homozygotes were not investigated further but were analysed with the other probes, which identify less frequent RFLPs (Table 4). The second most informative probe was PRO-6 and the third was M1-8 (Fig 4), the two other less informative probes being used only if required.

Six families (11%) were uninformative for all the probes used so far and cannot therefore be offered prenatal screening. Of the remainder,
Table 4  Families with retinoblastoma in which a given RFLP was informative in a particular kindred

<table>
<thead>
<tr>
<th>Family</th>
<th>Probes</th>
<th>p68RS2-0</th>
<th>p88RS0-6</th>
<th>p123MI-8</th>
<th>p95HSO-5</th>
<th>p35RS0-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF-01</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td>Not tested</td>
</tr>
<tr>
<td>RBF-02</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-03</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-04</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Not tested</td>
<td>Not tested</td>
</tr>
<tr>
<td>RBF-05</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-06</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-07</td>
<td>Yes</td>
<td>Not tested</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-08</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-09</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-10</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-11</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-12</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-13</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-14</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-15</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-16</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-17</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-18</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-19</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>RBF-20</td>
<td>Yes</td>
<td>Yes</td>
<td>Not tested</td>
<td>No</td>
<td>Not tested</td>
<td>Not tested</td>
</tr>
<tr>
<td>RBF-21</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>RBF-22</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Figure 3  Linkage analysis in family RBF-22 using the RS2-0 intragenic Rb probe. DNA from family members shown in the pedigree (fig I) was digested with RSAl and six different alleles ranging in size between 2.0-1.6 Kb were detected. In this family the mutant gene is segregating with the 1.85 Kb allele. The genotype of the subject represented in lane 9, showing only weak hybridisation, was 1.85/1.60.

Figure 4  Linkage analysis in family RBF-8 using the M1-8 intragenic Rb probe. Individual alleles are either 4.5 Kb long or a doublet, 2.3 and 2.2 Kb long. In this case the disease phenotype is segregating with the lower allele.

73% were informative with RS2-0 alone. Of the 16 families not informative for RS2-0, 11 (69%) were informative for PRO-6. In several cases, though the transmitting subject was heterozygous, the children became homozygous for a particular probe. In these cases the analysis was extended so that informative probes were also available for affected subjects in the second generation.

Prenatal Screening

The possibility of screening for carriers has been discussed with each family. So far two infants have been screened perinatally and in both cases it was possible to say that they were not carriers. Two families have taken advantage of prenatal screening. In the first (RBF-14), DNA was obtained by chorionic villus sampling, and in the second (RBF-01) from cultured amniotic fluid cells obtained for chromosome analysis because of advanced maternal age. The analysis of family RBF-01 with the RS2-0 probe is shown in fig 5. The affected and unaffected children have both inherited the same maternal allele, suggesting that the unaffected daughter apparently carries the mutant allele, as does the fetus. Alternatively, a recombination event within the RBI locus in the affected daughter, proximal to the RS2-0 site, may have occurred. Because available flanking markers were homozygous in the transmitting parent, however, it is not possible to exclude recombination. This family was not included in the log of the odds score calcula-

Figure 5  Segregation of RS2-0 alleles in family RBF-01. The retinoblastoma phenotype is segregating with the 1.6 Kb allele, which is also present in the unaffected 8 year old daughter (arrow) as well as in the amniotic cell from the fetus (○).
Application of intragenic DNA probes in prenatal screening for retinoblastoma gene carriers in the United Kingdom

655

In this family the phenotype may result from a reported by Connolly rather than contributing to that part of the scarring. This family is reminiscent of that one family of phenotypic expression of the retinoblastoma gene. In family RBF-12, however, no subject of phenotypic expression has had bilateral tumours, and in one family of phenotype analysis, the remaining subjects having the full range of malignancy. Vogel reviewed all published pedigrees up to 1979 and calculated that penetrance was between 0·996 to 0·927, with a higher frequency in offspring of bilaterally affected parents. Our calculation of penetrance (clinically defined) is in agreement with these figures, though selection bias is possible as the pedigrees in this series were mostly from a single referral hospital. Using RFLP analysis we have identified apparently unaffected gene carriers. Presumably these subjects remain at risk of other tumours associated with retinoblastoma, and their offspring are at risk of retinoblastoma; they will now receive appropriate genetic counselling. Conversely, the (unaffected) subjects shown not to be carriers of the predisposition allele may be reassured that they do not carry the mutant allele.

Incomplete penetrance can also be the result of inheritance of a balanced translocation where the unbalanced form gives rise to tumour predisposition or missed diagnosis where apparently unaffected gene carriers show evidence of regressed tumours or retinomas.

Excluding family RBF-01, where the issue of penetrance has still to be resolved, two other families showed evidence of unaffected gene carriers. In one family, for whom gene tracking using the intragenic probes was not possible, the incomplete penetrance was completely predictable from phenotype analysis, the remainder of the affected subjects having the full range of penetrance has still to be resolved, two other families showed evidence of unaffected gene carriers. In one family, for whom gene tracking using the intragenic probes was not possible, the incomplete penetrance was completely predictable from phenotype analysis, the remainder of the affected subjects having the full range of penetrance expression of the retinoblastoma gene. In family RBF-12, however, no subject has had bilateral tumours, and in one family nucleus affected members bore only retinal scarring. This family, reported by Connolly et al. It is possible that in this family the phenotype may result from specific mutations resulting in non-penetrance, rather than contributing to that part of the Poisson distribution described by Knudsen, where the second mutation has not occurred.

The availability of the retinoblastoma gene sequence may answer this question in the future. Within this family we have detected unaffected gene carriers who do not express retinal scarring. Cytogenetic analysis (J.K. Cowell, unpublished observations) failed to show any serious structural abnormality of chromosome 13 in affected family members, and their esterase-D gene concentrations were normal, precluding any large deletion in 13q14.

There are ethical issues to consider in the prenatal diagnosis of a disease such as retinoblastoma, which is rarely lethal and has no effect on mental ability. Although the vision of some patients who are treated early for single tumours may not be greatly affected, in some cases enucleation (which may be bilateral) is necessary. This operation is disfiguring, and many parents have expressed an interest in screening. Even though the child will survive the primary tumour, 5–15% will develop second tumours in adolescence, usually soft tissue sarcomas or osteosarcomas. In family RBF-14, the first to opt for prenatal screening, many of the members had bilateral disease, and were blind or partly sighted. It was their belief that in the event of a fetus inheriting a mutant allele, termination of pregnancy was acceptable. In the second family, RBF-01, termination was not acceptable, particularly as an unaffected previous child could have been the result of non-penetrance or a recombination between the probe and the predisposing mutation. In both cases the validation of the results awaits the passage of time.

In linkage studies such as this it is essential that the affected parent is heterozygous for at least one of the polymorphisms. Such is the variety of probes now available that most families are informative. Using standard linkage analysis prenatal screening can be offered to 90% of families, but new probes are still required for the remainder. The discovery of a VNTR, based on a four base pair repeat in the region adjacent to exon 20, should be an important addition to the screening armoury. There are between 10 and 24 different alleles in this system and it is estimated that 97% of subjects will be heterozygous at this locus. Although the techniques used for this analysis are more complex than in conventional agarose gel electrophoresis, specialists and genetic laboratories should be able to apply this procedure to prenatal screening. The ultimate goal is to determine whether each newly diagnosed case of retinoblastoma is a carrier; there is likely to be a whole range of mutations in different families. Though the whole retinoblastoma gene is large, the 27 exons occupy only approximately 5 Kb. Polymerase chain reaction amplification, and sequencing of each exon and their flanking regions offers a realistic possibility in the near future of providing unequivocal evidence for specification of that family history.

Discussion

Using five intragenic, unique sequence DNA probes for the retinoblastoma locus, 90% of families were informative for prenatal screening. As we have used probes from within the retinoblastoma gene itself, recombination among markers is unlikely. Nevertheless in the absence of mutation specific probes, recombination between marker and mutation cannot be excluded. Data from this series and from that of Wiggs et al. give a cumulative log of the odds score of 14·55 at 8 = 0, assuming a penetrance of 0·9.

Excluding family RBF-01, where the issue of penetrance has still to be resolved, two other families showed evidence of unaffected gene carriers show evidence of regressed tumours or retinomas.

There are between 10 and 24 different alleles in this system and it is estimated that 97% of subjects will be heterozygous at this locus. Although the techniques used for this analysis are more complex than in conventional agarose gel electrophoresis, specialists and genetic laboratories should be able to apply this procedure to prenatal screening. The ultimate goal is to determine whether each newly diagnosed case of retinoblastoma is a carrier; there is likely to be a whole range of mutations in different families. Though the whole retinoblastoma gene is large, the 27 exons occupy only approximately 5 Kb. Polymerase chain reaction amplification, and sequencing of each exon and their flanking regions offers a realistic possibility in the near future of providing unequivocal evidence for specification of that family history.
We thank all the those general practitioners in the UK who collected blood and sent it to us, and Dr T Dryja for making the intragenic retinoblastoma probes available to us. We also thank Dr J Pritchard and Dr M Pembrey for their critical reading of the manuscript. Zerrin Onadim was supported by a grant from the David Allen Retinoblastoma Appeal.

1 Vogel W. The genetics of retinoblastoma. Hum Genet 1979; 52:1-34.


20 Attwood J, Bryants S. A computer program to make linkage analysis with LIPED and LINKAGE easier to perform and less prone to input errors. Am Hum Genet 1985;52:259-62.


Genetic counselling in retinoblastoma: importance of ocular fundus examination of first degree relatives and linkage analysis

Z Onadim, P G Hykin, J L Hungerford, J K Cowell
Genetic counselling in retinoblastoma: importance of ocular fundus examination of first degree relatives and linkage analysis

Z Onadin, P G Hykin, J L Hungerford, J K Cowell

Abstract

We report an unusual family pedigree segregating the retinoblastoma predisposition gene. Expression of the phenotype in different individuals in this family ranges from asymptomatic gene carriers, regressed tumours, through unifocal to bilateral multifocal lesions. Because of the unusual pattern of inheritance in this family, initial genetic counselling at a local hospital did not take into account the possibility of incomplete penetrance of the gene, and complete ophthalmological examination of unaffected family members was not undertaken. We have used DNA probes from within the retinoblastoma predisposition gene for unequivocal identification of gene carriers. The subsequent demonstration of regressed tumours in founder members of the family confirmed the diagnosis of a dominantly inherited disease. The circumstances of the management of this family emphasises the need for specialist ophthalmic examination of first degree relatives and detailed genetic analysis of all such families with DNA probes.

Retinoblastoma (RB) is the commonest primary malignant intraocular tumour in children and has both hereditary and sporadic forms. In familial cases the inheritance follows an autosomal dominant pattern, usually with high penetrance. Thus offspring of gene carriers have a 50% risk of inheriting the mutant gene. Although the retinoblastoma phenotype shows a dominant mode of inheritance, it is clear that at the cellular level additional mutations are required, since not all retinal cells produce tumours, and in approximately 10% of gene carriers tumours do not develop at all. This latter phenomenon is referred to as incomplete penetrance. Knudson showed in a statistical analysis of RB that a single additional mutation was sufficient for tumorigenesis. This second mutation occurs sporadically in the homologous normal RB gene in retinal precursor cells. Hereditary cases are usually affected bilaterally, and, since the probability of the second mutation occurring is relatively high, the mean number of tumours in each eye averages four to five.

Retinoblastoma can also regress spontaneously, leaving characteristic retinal scars which often resemble successfully treated tumours. These may be compatible with normal visual acuity and remain undetected. The presence of multiple regressed tumours in a single individual or a single regressed tumour in an individual with relatives with RB implies gene carrier status for those individuals and puts their children at approximately 50% risk of tumour development. It is essential therefore, to examine the retinae of parents of apparently sporadic cases.

The gene for RB predisposition – RB1 – was mapped to chromosome region 13q14 following observations that, in a small percentage of patients, predisposing constitutional chromosome deletions of this region were present. Genetic linkage analysis using the adjacent esterase-D gene confirmed that the hereditary non-deletion form of the disease was due to mutations on the same locus. The RB1 gene was isolated by several groups independently and authenticated following the demonstration of mutations in this gene in tumour cells and predisposed persons. Individual unique DNA probes from within the genomic sequence of RB1, which recognised restriction fragment length polymorphisms (RFLPs), were isolated by Wiggs et al. which now form the basis of linkage analysis in RB families. So far no cases have been reported of genetic recombination between these DNA probes and the disease phenotype, and they have been used successfully in prenatal and perinatal prediction of carrier status.

Despite the now well characterised genetics of retinoblastoma and the availability of genetic probes for screening and counselling, advice to some families is still sometimes incomplete. This is most often the case in families where the phenotype segregates with unusual characteristics. In this report we present the details of one such family as an example of the difficulty in making the correct prediction in the absence of genetic linkage data and ophthalmic examination of first degree relatives.

Materials and methods

DNA was prepared from white blood cells and lymphoblastoid cell lines by standard methods. All procedures for the analysis of DNA samples were as described previously.

Results

The family pedigree is shown in Figure 1. The parents (II.6 and II.7) of the affected twins are first cousins and each has a sister affected by unilateral retinoblastoma. These observations are consistent with one or both of their parents (I.2 and I.3) being unaffected gene carriers. Family counselling was sought at a local hospital by the mother (II.7). In 1984, prior to their birth, a diagnosis of an autosomal recessive disorder was given at the local hospital by virtue...
of the fact that both parents had an affected sib, and a risk of having affected children of 1:32 was given. The retinae of the parents, I.2 and I.3, were not examined at this time.

In February 1987 non-identical twins were born, and in June of that year a diagnosis was made of bilateral retinoblastoma in both children at the age of 4 months following their referral to St Bartholomew's Hospital for treatment. Proband 1 had a juxtapapillary RB in the right eye and a posterior pole tumour in the left eye. Proband 2 had an RB nasal to the optic disc in the left eye and later developed a tumour temporal to the posterior pole in the right eye. No evidence of extraocular spread was discovered in either proband.

In July 1987 both grandparents attended the ophthalmology clinic at St Bartholomew's Hospital, where I.3 was found to have a large, grey translucent lesion at the posterior pole of the left eye (Fig 2). I.2 was found to have a small posterior pole lesion of the right eye which was translucent with pigment epithelial disturbance (Fig 3), and a small lesion nasal to the optic disc in the left eye. These lesions were typical of spontaneously regressed RB. Close examination of the retinae of II.6 and II.7 showed no evidence of retinal scarring. Esterase-D studies showed all family members to be genotype I-I. Chromosome analysis showed no deletions of the 13q14 locus in any family member. At the time of the initial consultation DNA-based analysis was not generally available.

Since both parents are potentially unaffected gene carriers, there is a possibility that their children could be homozygous for the RB gene mutation. To investigate this possibility we used the intragenic RB DNA probe RS2.0, which recognises an 8 allele variable number tandem repeat (VNTR). We were able to follow the segregation of the chromosome 13 in this family, the details of which are given in figures 4 and 5.

Analysis of patients II.5 and II.8, the affected cousins, shows that the common alleles are 1.9, which must have been derived from gene carriers I.2 and I.3. Since both I.2 and I.3 showed evidence of retinal scarring it must be assumed that the 1.9 allele came from a common ancestor carrying the predisposing gene. Both parents II.6 and II.7 also carry the 1.9 allele. For II.6 this allele was inherited from I.2 but II.7 received the 1.5 allele from her father and the 1.9 allele from the unrelated parent I.4. We would predict therefore that II.7 is not a gene carrier, thus excluding the possibility that their children could become homozygous for this mutation. The twins have inherited the 1.9 allele from their father, confirming the cosegregation of the RB phenotype with this allele. Analysis of the other members of the family shows that II.1 is definitely not a gene carrier, since the 1.9 allele is missing, but, being homozygous for 1.9, II.2 must be an asymptomatic gene carrier.

The affected patient II.5 is now homozygous for the 1.9 allele, which means that this probe will not be useful for future prenatal screening. In contrast II.8 is heterozygous and could be offered prenatal screening in the future. The so far unaffected child in this family unit, III.4, is also homozygous for the 1.9Kb allele, which places him at high risk for tumour development.

Discussion

In this unusual pedigree segregating the retino-
Genetic counselling in retinoblastoma: importance of ocular fundus examination of first degree relatives and linkage analysis

The segregation of the alleles identified by the RS2.0 DNA probe. The mutant retinoblastoma gene is segregating with the 1-9 allele. Asymptomatic gene carriers are indicated by the arrows. Symbols as in Figure 1.

Figure 4 Segregation of the alleles identified by the RS2.0 DNA probe. The mutant retinoblastoma gene is segregating with the 1-9 allele. Asymptomatic gene carriers are indicated by the arrows. Symbols as in Figure 1.

Figure 5 Southern blot analysis of key members of family RB-29 using the intragenic VNTR DNA probe RS2.0. In this family five different alleles 2.0 kb, 1.95 kb, 1.90 kb, 1.85 kb, and 1.5 kb are segregating. The RB predisposition gene segregates with the 1-9 kb allele.

Genetic counselling in retinoblastoma: importance of ocular fundus examination of first degree relatives and linkage analysis

The identification of spontaneously regressed tumours in the grandparentes in this case raised the risk to the children to 1:2 compared with the 1:32 risk offered at the time of first consultation. By establishing the phenotypes for the majority of the members of this family it was possible to perform linkage analysis with DNA probes. This analysis established several important points for counselling. The segregation of the RB phenotype with the 1-9 kb allele in the core family allowed us to identify an unaffected gene carrier who has not yet begun her own family. Although this individual is homozygous for the 1-9 kb allele, and so would not qualify for prenatal screening with this probe, new screening procedures are being introduced with other polymorphisms which could be used. For individual II.8, who is heterozygous for the RS2.0 polymorphism, prenatal screening is available. The same service can be offered to the twins. We have also demonstrated that the parents of the twins are not both gene carriers, which excludes the possibility of their having children homozygous for the mutation. Since such homozygous individuals have not been reported, it is not clear what the expected phenotype might be, especially in 'mildly affected' pedigrees.

We thank Dr J Prichard for his critical reading of the manuscript. Zerrin Onadim is supported by a grant from the David Allen Retinoblastoma Appeal.

5. Vogel W. The genetics of retinoblastoma. Hum Genet 1979; 52: 1–54


Application of PCR amplification of DNA from paraffin embedded tissue sections to linkage analysis in familial retinoblastoma

Z ONADIM AND J K COWELL
Application of PCR amplification of DNA from paraffin embedded tissue sections to linkage analysis in familial retinoblastoma

Z Onadim, J K Cowell

Abstract
A family segregating for the retinoblastoma predisposition gene has been analysed using the polymerase chain reaction to exclude their son as being an affected gene carrier. The unusual feature of this family is that the affected child, who would ordinarily have been used to establish phase in a linkage study, died as a result of developing a second tumour some years ago. The only tissue available from this child was a paraffin embedded, formalin fixed histopathological specimen from the second tumour. It was possible to isolate DNA from this tissue and amplify the DNA flanking two polymorphic restriction enzyme sites to establish alleles which cosegregated with tumour predisposition. Archival material can now be used to offer families such as this prenatal screening to provide informed genetic counselling.

Retinoblastoma (Rb) is the most common intraocular tumour of children. In its familial form the tumour phenotype segregates as an autosomal dominant trait.\(^1\) Rb arises as a result of mutations in both copies of the Rb gene.\(^2\) Most cases are apparently sporadic, owing to random key mutations in both genes; in hereditary cases (25 to 40%), however, the first mutation is transmitted through the germline and only the second mutation occurs as a random somatic event. Subjects carrying the Rb predisposition gene are also at a significantly higher risk than the general population for development of a number of other tumours, especially osteosarcoma and soft tissue sarcomas.\(^3\)

Through the analysis of rare constitutional chromosomal deletions\(^4\) and classical linkage studies using the adjacent esterase D gene,\(^7\) the Rb gene was mapped to chromosome region 13q14 and these observations eventually led to its isolation.\(^10-12\) The gene, called RB1, is 4.7 kb long and spans 200 kb of genomic DNA.\(^12-14\) The 4.7 kb mRNA comprises 27 exons and structural rearrangements have been observed within the genomic sequence and mRNA of approximately 30% of tumours. The cDNA cannot be used directly in linkage studies, because it does not identify polymorphic sites for restriction enzymes. Therefore a series of unique sequence, intragenic DNA probes were isolated\(^15\) allowing gene tracking in families and the detection of apparently unaffected Rb gene carriers.\(^16\) The sequence for the entire Rb cDNA is now available\(^17\) and the flanking intron sequences around two of the polymorphic sites have also been determined.\(^18-20\) This means that linkage analysis in Rb families can be performed using the polymerase chain reaction (PCR).

The intragenic DNA probes are used in our laboratory for pre- and postnatal screening for Rb. However, between 10% and 15% are either uninformative for these probes or DNA samples have not been available from key family members. Here we present one such family, whose first child would normally have been used to establish linkage phase but who died before DNA analysis was available.

Materials and methods
DNA preparation
DNA from whole blood samples was prepared using standard phenol/chloroform extraction followed by ethanol precipitation.\(^21\) DNA from 10 μm thick paraffin embedded tumour sections was prepared by a method modified from Shibata et al.\(^22\) Firstly, individual paraffin sections were dissolved in 500 μl of xylene. The tissue was recovered by centrifugation and then washed twice in ethanol and desiccated for 0.5 to 2 hours. The DNA extraction solution consisted of 100 mmol/l Tris–Cl, 4 mmol/l EDTA pH 8–0, and 0.5 mg/ml proteinase K. After 12 to 18 hours at 37°C, this solution was boiled for seven minutes and 1 to 10 μl used for PCR.
PCR BASED DETECTION OF POLYMORPHIC SITES

BamHI site

Two primers (table), one from exon 1 and the other from intron 1 of the human Rb gene, were used to amplify a genomic DNA fragment containing the polymorphic BamHI site. In conventional Southern blotting experiments this polymorphism is identified using the M1.8 DNA probe. PCR was carried out in a total volume of 50 μl containing approximately 1 μg DNA and Promega buffer consisting of 50 pmol of each primer, 0.2 mmol/l each dNTP (dTTP, dCTP, dGTP), 50 mmol/l KCl, 10 mmol/l Tris HCl (pH 9.0 at 25°C), 1.5 mmol/l MgCl₂, 0.01% gelatin (w/v), 0.1% Triton X-100, 10% dimethyl sulphoxide (DMSO), and 2 to 3 units Taq DNA polymerase (Promega). The reaction mix was overlaid with 50 μl of mineral oil to prevent evaporation. Amplification was performed using a programmable thermal cycler (Techne PHC-1). Amplification conditions consisted of three steps (after an initial 15 minute denaturation step at 96°C): denaturation at 96°C for 20 seconds, annealing at 60°C for 20 seconds, followed by an extension step at 72°C for 60 seconds. On completion of 30 cycles the mineral oil was removed by chloroform extraction. The amplified product was then digested overnight with BamHI. DNA fragments were resolved by electrophoresis through 2% agarose gels.

XbaI site

A genomic DNA fragment containing the polymorphic XbaI site (21.8 kb downstream of exon 17) was amplified using primers (table) from intron 17 of the human Rb gene. Using conventional Southern blotting procedures this polymorphism is identified by the PRO.6 DNA probe. The PCR conditions were the same as those used for the BamHI site except that the PCR mix did not contain DMSO and the annealing temperature was 50°C. The amplification product was digested overnight with XbaI. DNA fragments were resolved by electrophoresis through 2% agarose gels.

Results and discussion

Prenatal screening for carriers of the RB predisposition gene is now relatively routine using a panel of unique DNA sequences derived from within the genomic sequence of Rb.15, 23, 24 The polymorphic variants identified using these probes are sufficiently common in the population to offer a prediction of carrier status in 85 to 90% of families (unpublished observations). To date, there have been no reported recombinations between the intragenic probes and the Rb phenotype, allowing predictions to be made with 95% confidence. There are, however, still families for whom this service is not available. Approximately 8 to 10% of families will be uninformative using the available probes because family members transmitting the predisposition to tumour development are homozygous at all of these loci. The application of new technology looking at short variable number tandem repeats (VNTRs) around exon 20 of the Rb gene23, 25 or at single base pair polymorphic sites by direct sequencing26 offers the possibility of prenatal screening in these cases. There are some families, however, where prenatal screening is not an option because, for a variety of reasons, tissue samples from key family members are not available. One such family is the subject of this report.

The pedigree of family RB-29 is shown in fig 1. At the age of 18 months the mother developed a tumour in the right eye which was enucleated. There was no previous history of Rb in this family. There was no history of Rb on the paternal side of the family either. At the age of 2 years the first born child, a girl, was treated for bilateral Rb with radiation. The tumour regressed but six months later a rhabdomyosarcoma arose in the right cheek and eventually caused the patient's death.27 Although a rare subtype of this tumour, this second malignancy was part of the group frequently seen in Rb gene carriers.1, 4 The second child, a boy, was born in 1987 and the family was referred to us for genetic screening. The mother was shown to be heterozygous using the M1.8 and PRO.6 DNA probes described by Wiggs et al.15 The boy had

Figure 1 Pedigree of family RB-29.

<table>
<thead>
<tr>
<th>Polymorphic site</th>
<th>Sense</th>
<th>Antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>BamHI</td>
<td>5'-CAGGAAGCCGGGCGCGAG-3'</td>
<td>5'-CTGCCACGCGCTCCGCCGT-3'</td>
</tr>
<tr>
<td>XbaI</td>
<td>5'-TCCAAATGAGAAAATGCG-3'</td>
<td>5'-GGAAATTTGAGAATTTTTTT-3'</td>
</tr>
</tbody>
</table>
no evidence of Rb at the time of referral but, although the peak age for hereditary tumours is 10 to 14 months, he is still at risk for developing a tumour. In any case he could still be an apparently unaffected gene carrier which is known to occur in approximately 10% of cases. To establish phase unequivocally it was essential, therefore, to analyse DNA from tissue from the dead child. No blood samples had been saved and no necropsy was performed but during the treatment of the rhabdomyosarcoma a tooth was removed. Attached to its base was a small piece of tumour which had been fixed in formalin and embedded in wax. We were able to obtain tissue sections from this tumour material for analysis.

The DNA sequences flanking the polymorphic restriction enzyme sites in RB1 identified by probes M1.820 and PRO.619 have been established. Using these sequences, oligonucleotide primers have been designed (see Materials and methods) to analyse the polymorphic sites using PCR. Using DNA from the tissue section from the dead child's second tumour, PCR products around both polymorphic sites were generated. For the BamHI polymorphism an approximately 200 bp genomic DNA fragment containing the intron 1 splice donor site was amplified (fig 2, top). The polymorphic BamHI site is located within the amplified fragment 50 bp from the 3' end. Thus, fragments of 140 bp and 60 bp are generated after BamHI digestion. Using this polymorphism the results were unequivocal; the mother is heterozygous and the affected daughter homozygous for the lower allele (fig 2, bottom). Since the as yet unaffected son is also heterozygous we would predict that the mutant gene, therefore, is segregating with the lower (140/60 bp) allele. Using the other set of primers (table) an approximately 945 bp genomic DNA fragment containing the polymorphic XbaI site was amplified. XbaI digestion of this fragment identifies two alleles; an upper allele 945 bp long and a lower allele consisting of two fragments 630 bp + 315 bp long. At the XbaI locus (fig 3) the mother, the father, and the surviving child are all heterozygous. The dead child was homozygous for the lower allele indicating that the Rb mutation is segregating with this allele.

The DNA extracted from the tissue sections was degraded, in the size range of 50 to 1500 bp (fig 4). Owing to the cross linking of the DNA with formalin during the fixation process, smaller sizes are over-represented although this did not present a problem in analysing the 180 bp sequence flanking the BamHI polymorphism, yields of PCR product being relatively good. For the larger 945 bp fragment, however, PCR product yields were somewhat smaller. We interpret this to mean that, because of the fragmented nature of the DNA removed from the formalin fixed tissue, there are fewer intact molecules of the appropriate size in the DNA used as template in the PCR reaction. The DNA thus generated was used directly for restriction enzyme digestion where, although left for six to 24 hours, digestion was sometimes incomplete.

![Figure 2](image2.png)  
Figure 2  PCR amplification of the polymorphic BamHI site from DNA samples from the members of family RB-29 (above). In all cases a dominant band of the expected size, ±200 bp long, is seen as well as fainter, non-specific bands in some cases. The amplified DNA from each patient was digested with BamHI (below). The marker lane (M) in each case contains the Gibco BRL 1 kb ladder. The father (lane 1) is homozygous for the lower allele but the mother (lane 2) is heterozygous. Although a weak residual upper band is seen in this example, in DNA extracted from the tissue section representing the affected dead child (lane 3), the dominant bands are those for the lower allele indicating that this patient is homozygous for the allele. The unaffected child (lane 4) is heterozygous indicating that the Rb gene is segregating with the lower allele from the mother.

![Figure 3](image3.png)  
Figure 3  Segregation of the alleles for the polymorphic XbaI site in family RB-29. The father (lane 1) and mother (lane 2) are both heterozygous as is the unaffected child (lane 3). The dead affected child (lane 4) is homozygous for the lower allele. The Rb gene is therefore segregating with the lower allele from the mother. M = marker lane containing 1 kb DNA ladder (Gibco-BRL).
Application of PCR amplification of DNA from paraffin embedded tissue sections to linkage analysis in familial retinoblastoma

315

220

Figure 4  Size range of DNA extraced from the formalin fixed, paraffin embedded tissue. One tenth of the extracted DNA (lane I) was run on a 0.6% agarose gel and stained with ethidium bromide. A smear is seen between 200 bp and 3 kb. The marker lane (M) contains DNA from a 1 kb ladder (Gibco-BRL).

as shown in fig 2. We have noted this same finding in several other cases where subjects are known to be homozygous for the lower allele from conventional Southern blotting/radiolabelled probing. The relative intensity of the band caused by DNA molecules resistant to digestion, however, was significantly weaker than in 'true' heterozygotes (fig 2). Repeated amplification and digestion of DNA isolated from different tissue sections showed that, in some cases, this upper band was absent. It is possible, therefore that this finding varies in relation to the purity of the extracted DNA.

It is possible that the tumour tissue used in this analysis is not truly representative of normal tissue from the patient. Loss of heterozygosity is frequently reported in retinoblastoma tissue, osteosarcomas, and soft tissue sarcomas. However, it is consistent with the theory of 'exposure' of recessive alleles in these childhood tumours that the mutant allele is retained. This must also be expected to be the case in the rhabdomyosarcoma used in this study if loss of heterozygosity had occurred. Using the PRO.6 polymorphisms, however, since both parents are heterozygous their daughter could have been constitutionally heterozygous, inheriting the upper allele from the father. Whether their daughter is homozygous for the lower allele or it has been reduced to homozygosity in this case, in the tumour the result is still the same; the tumour predisposition gene cosegregates with the lower allele. From our linkage analysis using the BamHI polymorphism our prediction is that it is the lower allele which cosegregates with the mutant allele. Here loss of heterozygosity is not relevant to the argument since the father is homozygous and must contribute the lower allele to his children. Since the mother also contributes the lower allele, associated with the mutant gene, her daughter could not have been constitutionally heterozygous.

Since both parents are heterozygous for the PRO.6 polymorphism, carrier status can only be confirmed or otherwise in 50% of future offspring but results will be unequivocal for the BamHI polymorphism. Taken together these tests will allow counselling of this family in the future.

The use of DNA from formalin fixed tissue in PCR has improved our capability to analyse mutations in Rb families. As long as DNA is available from key dead family members linkage phase can be established. The identification of the specific mutations in apparently sporadic cases will soon be possible via direct sequencing of the amplified DNA. Tumour tissue from many patients now approaching child bearing age exists as fixed material in pathology archives. Using this material and the methods described here, accurate genetic counselling is now available for many of these families.

We would like to thank C Brookes for helpful discussions about the dead child and the pathology department of the Royal Liverpool Children's Hospital, Alder Hey, for providing the tissue sections. We are grateful to Dr J Pritchard for critical reading of the manuscript. Zerrin Onadim is supported by a grant from the David Allen Retinoblastoma Appeal.

1 Vogel W. The genetics of retinoblastoma. Hum Genet 1979;52: 1-54.
7 Sparkes RS, Sparkes MG, Wilson MG, et al. Regional assignment
11 Dryja TP, Rapaport JM, Joyce JM, Petersen RA. Molecular detection of deletions involving band q14 of chromosome 13 in retinoblastomas. Proc Nail Acad Sci USA 1986;83:7391-4.
Follow-up of retinoblastoma patients having prenatal and perinatal predictions for mutant gene carrier status using intragenic polymorphic probes from the RB1 gene

Z. Onadim, J. Hungerford & J.K. Cowell

ICRF Oncology Group, Institute of Child Health, 30 Guilford Street, London WC1N 1EH; Department of Clinical Ophthalmology, St Bartholomew's Hospital, West Smithfield, London EC1, UK.

Summary We have carried out presymptomatic prediction of mutant gene carrier status in ten individuals with a family history of retinoblastoma. In all cases standard linkage studies were employed using intragenic DNA probes which recognise restriction fragment length polymorphisms. In four cases foetal DNA samples were obtained by chorionic villus sampling, the remaining six were derived from either cord blood samples or venipuncture of neonates. We demonstrated that the mutant gene was inherited by only one of these patients who has subsequently developed bilateral tumours. Six of the other cases have now reached the age beyond which it might have been expected that tumours would develop and are all disease free. It must be concluded that repeated ophthalmological examination of these and future patients shown not to have inherited the mutant gene, is unnecessary.

The application of reverse genetics procedures has led to the successful isolation of genes responsible for human genetic disease in recent years. It is now commonplace, in the UK at least, for these genes to be used in prenatal screening programmes to determine mutant gene carrier status in families. The advantage of using the causative gene is that the chances of recombination between marker and phenotype are very small. However, despite the relative success in the isolation of genes responsible for biochemical disorders such as PKU, thalassemia, and many other genetic diseases there have been few genes isolated and characterised which are responsible for hereditary predisposition to cancer. The first, and still most extensively studied, cancer predisposition gene to be isolated was the retinoblastoma gene, RB1 (Friend et al., 1986).

Retinoblastoma (Rb) is an intraocular tumour of children which has both hereditary and sporadic forms (Vogel, 1979; Cowell, 1991). Only 15% of patients have a positive family history although all bilateral cases (around 40%) are generally considered to be germ line carriers of a predisposing mutation (Knudson, 1971). The Rb phenotype segregates as an autosomal dominant trait with high penetrance (Vogel, 1979). The mean age of onset in hereditary cases is approximately 14 months (Draper, G.J., pers. comm.) compared with 25-30 months for sporadic cases. Hereditary Rb rarely arises after the age of two (although some cases present in an autosomal dominant trait with high penetrance (Vogel, 1979). The mean age of onset in hereditary cases is approximately 14 months (Draper, G.J., pers. comm.) compared with 25-30 months for sporadic cases. Hereditary Rb rarely arises after the age of two (although some cases present in advanced forms in older children) and hardly ever after the age of five. In patients with bilateral, multifocal disease their tumours may develop sequentially over a period of months. Treatment of small tumours has become very successful in the past two decades with an overall survival of 90%. Early detection of tumours, therefore, is the most critical feature in the clinical management of the disease.

With the cloning of the RB1 gene it was expected that the cDNA probe could be used in standard linkage analysis to identify restriction fragment length polymorphisms (RFLP). This was not the case, however, and it was necessary to isolate intragenic, unique-sequence DNA probes which recognise RFLPs (Wiggs et al., 1988). These probes form the basis of linkage studies in the analysis of RB families worldwide and have been successfully applied in 80-90% of families (Wiggs et al., 1988; Scheffer et al., 1989; Onadim et al., 1990). More recently other RFLPs and DNA sequence polymorphisms (Yandell & Dryja, 1989; McGee et al., 1989) have been added to the armoury of probes available. In cases of familial Rb, it is now possible to establish with which chromosome the mutant allele segregates using standard linkage studies (Cowell & Onadim, 1990). This, in turn, allows for the unequivocal identification of gene carriers as well as excluding those individuals who do not carry the mutant gene. This analysis is particularly important in families where incomplete penetrance occurs. There have been few reports, however, of the successful application of the DNA probes in prenatal and perinatal screening programmes for Rb. Since the first ever report of prenatal prediction by Mitchell et al. (1988) we have undertaken several additional tests in Rb families and have been able to follow them for several years. In this report we present our experience with presymptomatic prediction using standard RFLP analysis and using DNA obtained from foetal chorionic villus sampling and cord and peripheral blood samples from neonates.

Materials and methods

Our Unit at the Institute of Child Health offers an extensive service for family linkage studies and prenatal screening for families from throughout the UK. This programme also includes investigations of esterase-D (ESD) levels in affected patients to detect individuals with 13q14 deletions (Cowell & Onadim, 1986; 1989). Although the majority of Rb patients in the UK are referred to the Ophthalmology Departments at Moorfields Eye Hospital and St Bartholomew's Hospital, several families have been referred to us directly from other regions in the UK.

All of the procedures for the RFLP analysis using the probes RS2.0, PRO.6 (Wiggs et al., 1988) and M1.8 (Bookstein et al., 1988) were as described previously (Onadim et al., 1990). The procedures for PCR based detection of polymorphic sites have been described in Onadim and Cowell (1991).

The Rb1.20 polymorphism consists of a variable number of (CTTT(T))n (n = 14-26) repeats (Yandell & Dryja, 1989) and occurs 54 bp from the 3’ end of exon 20.

The two primers used to amplify the Rb1.20 polymorphism were:

5’ Primer 5’-GTATGAACTCATGAGACAGGCAT-3’
3’ Primer 5’-AATTAAACAGGTGGTGGTGTTACAGC-3’

We tested a series of primers to amplify this region and found this particular pair to give the best results. These are not the same primers originally used by Yandell and Dryja (1989).
The 5' primer is from within exon 20 and the 3' primer from intron 20 of the human RB gene. In a PCR reaction the amplify a genomic DNA fragment (300–350 bp long depending on the number of repeats) containing the RBI.20 VNTR region. PCR was carried out in a total volume of 50 μl containing; approximately 1 μg DNA, 50 pmol of each primer, 5 μl of 10 x Taq-polymerase buffer (Northumbria Biologicals Limited), 0.2 mM each of dATP, dTTP, dGTP and 0.02 mM dCTP, 1 μCi 32P-dCTP and 1–2 units Taq DNA polymerase (NBL). The NBL 10 x buffer consisted of 100 mM Tris-HCl pH 8.8, 500 mM KCl, 15 mM MgCl2, 1% Triton X-100. The reaction mix was overlaid with 50 μl of mineral oil to prevent evaporation. Amplification was performed using a programmable thermal cycle machine (Techne PHC-1). Following an initial 15 min denaturation step at 96°C, after which the Taq polymerase enzyme was added, amplification conditions consisted of three steps: denaturation at 94°C for 20 s, annealing at 59°C for 20 s, followed by an extension step at 72°C for 60 s. On completion of 30 cycles, the mineral oil was removed by chloroform extraction.

The individual PCR products were resolved by electrophoresis at constant power of 60 watts on 6% polyacrylamide gels containing 7 M urea for 6 h. Since allele sizes in this polymorphism can differ by 1 bp, it is important not to use a sample with too much radioactivity otherwise, following autoradiography, bands of similar size tend to merge. We therefore diluted the amplification product in serial dilutions to find the optimum concentration, which was usually 1:10. In the majority of cases, therefore, 1 μl of the PCR product was diluted tenfold in amplification dilution solution (0.1% SDS, 10 mM EDTA) and 1–2 μl of this dilution was mixed with 2 μl of formamide dye mix (95% formamide, 20 mM EDTA, 0.05% Bromophenol blue and 0.05% Xylene cyanol). This solution was heated to 95°C for 5 min for denaturation before loading into the gel. The gel was transferred onto 3 M Whatmann filter paper and dried for 2 h. Dried gels were overlaid with Kodak XAR-5 autoradiographic film and autoradiographed at −70°C for 16–24 h with Cronex Quanta III intensifying screens.

Results

For our genetic linkage analysis in Rb families we have used RFLPs identified by probes RS2.0, PR0.6 and M1.8 which are revealed following DNA digestion using Rsal, Xbal and BamHI respectively (Wigg et al., 1988; Bookstein et al., 1988). The M1.8 probe recognises two variant alleles, the first is 4.5 kb long and the second consists of a doublet 2.3 kb + 2.2 kb. For simplicity heterozygous individuals are described throughout as 4.5/2.3 (instead of 4.5/2.3 -I - 2.2) (Figure 1). Homozygous individuals for the lower allele are described as 2.3/2.3 (instead of 2.3 + 2.2/2.3 + 2.2). The conventional way of identifying RFLPs using PR0.6 and M1.8 is by Southern blotting (Onadim et al., 1990). However, the same polymorphisms can also be detected using PCR techniques (Mc Gee et al., 1990; Bookstein et al., 1990; Onadim & Cowell, 1991) in which heterozygous individuals are described as 4.5/2.3. Later, the predisposing allele identified by PRO.6. In our experience families are usually only informative for a few of the RFLPs. Clearly, the more probes that can be used the more confident our predictions will be. In RB06 we were able to confirm the results with a third polymorphism, the RBI.20 VNTR. In this analysis, Rb predisposition clearly segregates with the 4.5 kb allele in this family.

In family RB06 (Figure 1) post-natal screening of II.2 after 16 months and prenatal screening of II.3 showed that neither inherited the mutant Rb gene, co-segregating with the 1.65 kb, RS2.0 allele from the affected father. This result was confirmed using the M1.8 probe. Both II.2 and II.3 were homozygous for the 2.3 kb allele whereas the RB predisposition clearly segregates with the 4.5 kb allele in this family. In our experience families are usually only informative for a few of the RFLPs. Clearly, the more probes that can be used the more confident our predictions will be. In RB06 we were able to confirm the results with a third polymorphism, the RBI.20 VNTR. In this analysis, Rb predisposition clearly segregates with allele 5 in the family and, although no DNA was available from II.2, II.3 was not shown to inherit this allele in IV.

In family RB13 (Figure 1) a post natal screen of II.2 was undertaken when the child was 6 months old (reported in Onadim et al., 1991) and, subsequently, a second post-natal screen was performed after 0.5 months for II.3. Neither child inherited the allele associated with Rb predisposition (1.95 kb with RS2.0 and allele 2 with RBI.20).

Family RBF29 (Figure 1) has an unusual inheritance pattern which has been described (Onadim et al., 1991) where the possibility of future prenatal screening was indicated. In March 1991 we analysed a CV sample from the fetus (III.5) who was found not to carry the Rb predisposition allele (1.9 kb RS2.0; allele 6 RBI.20). Later, the predisposing mutation itself was identified in this family. All the affected individuals and unaffected gene carriers were found to carry this mutation. The mutation was not present in III.5 (Onadim et al., submitted).

In family RB22 (Figure 1), the Rb mutation is segregating with the 1.95 kb allele identified by RS2.0 and the 6.5 kb allele identified by PR0.6. III.2 who was already 10 months old was excluded from carrying the mutant allele using both probes. More recently (August, 1991) a second, post-natal screen was carried out in this family for III.3. This patient was found to be heterozygous for PR0.6 which means that, since both parents were also heterozygous at this locus, this probe was not informative in his case. Using RS2.0, however...
Figure 1 The pedigrees of the families for whom presymptomatic screening was undertaken. For each individual the size or the number of allele(s) for each informative polymorphism are given. An arrow ✓ indicates the individuals for whom a presumptomatic prediction was given.

Table 1 The results of four prenatal and six post-natal screenings carried out in a period of 4 years (1988-1991) involving seven different families

<table>
<thead>
<tr>
<th>Family</th>
<th>Proband (months)</th>
<th>Age at testing (1/91)</th>
<th>Informative probes</th>
<th>Predisposition</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF06</td>
<td>II.2</td>
<td>16</td>
<td>55</td>
<td>RS2.0, M1.8</td>
<td>N</td>
</tr>
<tr>
<td>RBF06</td>
<td>II.3</td>
<td>Prenatal</td>
<td>14</td>
<td>RS2.0, M1.8, RB1.20</td>
<td>N</td>
</tr>
<tr>
<td>RBF13</td>
<td>II.2</td>
<td>6</td>
<td>40</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
</tr>
<tr>
<td>RBF13</td>
<td>II.3</td>
<td>0.5</td>
<td>19</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
</tr>
<tr>
<td>RBF14</td>
<td>III.2</td>
<td>Prenatal</td>
<td>37</td>
<td>RS2.0</td>
<td>N</td>
</tr>
<tr>
<td>RBF29</td>
<td>III.5</td>
<td>Prenatal</td>
<td>01</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
</tr>
<tr>
<td>RBF32</td>
<td>III.2</td>
<td>10</td>
<td>34</td>
<td>RS2.0, PR0.6</td>
<td>N</td>
</tr>
<tr>
<td>RBF32</td>
<td>III.3</td>
<td>4</td>
<td>9</td>
<td>RS2.0</td>
<td>N</td>
</tr>
<tr>
<td>RBF33</td>
<td>III.1</td>
<td>0.25</td>
<td>3.5</td>
<td>RB1.20</td>
<td>Y</td>
</tr>
<tr>
<td>RBF34</td>
<td>III.2</td>
<td>Prenatal</td>
<td>–</td>
<td>RS2.0, RB1.20</td>
<td>N</td>
</tr>
</tbody>
</table>
III.3 was found not to have inherited the 1.95 allele and therefore is expected to be unaffected.

Family RBF33 (Figure 1) was not informative for three of the probes used in our screening programme, i.e. RS2.0, PR0.6 and M1.8. They were, however, informative for the polymorphism RB1.20. Using DNA from a cord blood sample, III.1 was shown to have inherited the mutant Rb predisposition gene segregating with allele 3 (Figure 2) in this family. Two weeks later, ophthalmological examination of III.1 identified tumours in both eyes.

Family RBF34 (Figure 1) was referred to us for assessment for future prenatal screening having an affected child already. Unusually this family was informative for all the polymorphic restriction enzyme sites analysed (Figure 3). Using the RS2.0 probe the affected father was apparently homozygous for the 1.75 kb allele (Figure 3b) which was unusual in that his father II.1 did not carry this allele. His mother was not available for analysis. At first, we considered non-paternity as an explanation until patient III.1 was shown apparently not to have inherited an allele from his father also (Figure 3b). The same pattern of inheritance was shown for the RB1.20 locus (Figure 3a). II.1 was heterozygous however for BamHI (Figure 3c) and XbaI (Figure 3d) polymorphisms. It was clear that the predisposing mutation, which originated in I.1, is a deletion including a part of intron 17 (after the XbaI site) and extending at least to intron 20.

In August 1991 prenatal screening was carried out for II.2 after 11 weeks of pregnancy. The foetus was shown to inherit the normal allele (allele 2) with RB1.20 and all of the other

---

**Figure 2** Linkage analysis in family RBF33 using the RB1.20 polymorphism. The VNTR region was amplified by using two flanking primers in a reaction containing 3P dCTP. The products were then separated on a denaturing polyacrylamide gel (see Methods). III.1 was shown to have inherited the mutant RB predisposition gene segregating with allele 3 in this family.

**Figure 3** Linkage analysis in family RBF34, the pedigree for which is shown in the centre. Note II.1 and III.1 have only a single allele derived from their unaffected parents for the RB1.20 VNTR and the RS2.0 polymorphism shown in a and b respectively. PCR analysis of polymorphisms revealed by the M1.8 probe c and PR0.6 probe d confirm that the foetus III.2 has not inherited the mutant allele. M = marker lane containing the 1 kb DNA ladder (Gibco-BRL).
probes from the affected father II.1 (Figure 3) and it is therefore expected to be unaffected.

Discussion

We have followed a cohort of patients, with a family history of Rb, who have received pre-symptomatic genetic screening. In those cases where prenatal tests were performed, DNA from chorionic villi were used and were obtained after approximately 10 weeks of pregnancy. Post-natal tests are carried out using DNA obtained from either cord blood or whole blood obtained early in the child's life. Sometimes it is difficult to obtain large volumes of blood from newborns. This is not the case with cord blood samples. In cases where families are informative for RBl.20, Pr0.6 or M1.8, lack of DNA is not a problem, since these polymorphisms can be identified using PCR within 24 h. However, one of the important probes used in Rb screening, Rs2.0, still requires reasonable amounts of DNA (5–10 μg) and the results are only available after 4–7 days.

Prior to our application of intragenic DNA probes to prenatal screening in Rb families the only such report was by Cavenee et al. (1986). In this study markers flanking the RBl gene were used and met with limited success due to recombination events between marker and Rb locus. Prenatal screening was performed in five Rb families and the likelihood of Rb was predicted in two cases and freedom from disease in three. Two of the cases showed evidence of meiotic recombination and the predictive accuracy in one other was only 70% since only loci distal to the Rb locus were informative. We have previously been able to predict, pre-symptomatically, the development of Rb in two patients (J.C., unpublished data) who were carriers of chromosome 13 deletions and who were identified using esterase-D measurements in the series described by Cowell et al. (1989). In both cases the referral for testing was warranted because of the presence of other congenital abnormalities and dysmorphology which are frequently associated with 13q14 deletions. Both patients eventually developed Rb, before 12 months, although it should be noted that not all such cases develop tumours (Cowell et al., 1988; Wilson et al., 1987).

Our current series represents the first reported cases of pre-symptomatic predictions which have been followed for sufficient time to be sure that the prediction was accurate. The majority of familial cases present before 2 years and we have followed four patients for at least this period although, since in fact, the mean age of onset is 14 months six have reached this age disease free. A surprising result was that, to date, all but one of the patients were shown not to have inherited the mutant RBl gene. There is still a small possibility, however, that intragenic recombination might have occurred. The RBl gene consists of 27 exons spanning approximately 200 kb of genomic DNA (Friend et al., 1987; McGee et al., 1989). Assuming the generally accepted recombination frequency of 1 cross-over per 10^6 base pairs, the theoretical chances of recombination occurring within the RBl gene is 0.2, or 1:500. To date there have been no cases of recombination in any of the families reported so far (Wiggs et al., 1988; Scheffer et al., 1989; Onadim et al., 1990) which surveyed approximately 140 meioses. The M1.8 unique sequence DNA probe is located in the first exon of RBl (Bookstein et al., 1988) and a VNTR locus occurring in the 3′ intron adjacent to exon 20 (McGee et al., 1989) which covers most of the gene (75%). If mutations can occur equally along the length of the gene, as appears to be the case at present (Yandell et al., 1989; Dunn et al., 1989), the possibility of a predictive error is decreased accordingly if patients are informative at these loci. Given this low chance of intragenic recombination it must be concluded that it is unnecessary to repeatedly screen patients shown not to have inherited the predisposing mutation following linkage analysis.

In one of the families we described, RBF29, the predisposing mutation itself was identified and III.5 was found not to carry this mutation proving our prediction using polymorphic probes. Identification of the actual mutations require the use of different techniques. Only gross structural rearrangements and large deletions are detected by Southern blot analysis. The majority of Rb mutations, however, are small deletions or point mutations which require the use of techniques such as SSCP (single strand conformation polymorphism) and PCR sequencing. Using a combination of these techniques, it is now theoretically possible to identify predisposing mutations in most Rb families and indeed to search for predisposing mutations in constitutional DNA of sporadic patients to determine whether or not they carry a germ-line mutation. However, this approach is very expensive and time consuming and, at present, it is not practical to analyse every patient with Rb, although this situation might improve in the future with the availability of quicker techniques and automated sequencing.

Zerrin Onadim was supported by a grant from the David Allen Retinoblastoma Appeal.

References


TECHNICAL REPORT

Detection of heterozygous mutations in the RB1 gene in retinoblastoma patients using single-strand conformation polymorphism analysis and polymerase chain reaction sequencing

Annette Hogg, Zerrin Onadim, Paul N. Baird & John K. Cowell

ICRF Oncology Group, Department of Haematology and Oncology, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK

Several families segregating the autosomal dominant form of the hereditary retinoblastoma predisposition gene have been analysed for the causative mutation. We have used the single-strand conformation polymorphism (SSCP) technique to screen for mutations, exon by exon, in the RB1 gene in affected patients from these families. The SSCP technique has proved a rapid and simple technique which relies on the sequence-dependent migration of single-stranded DNA in a non-denaturing polyacrylamide gel. Oligonucleotide primers flanking all 27 exons and the promoter region of the RB1 gene are reported here. The polymerase chain reaction (PCR)-amplified products range in size from 212 to 625 bp and include a flanking intron sequence which allows detection of mutations in these regions. The sensitivity of SSCP is optimal when DNA fragments are approximately 200 bp long. Consequently, restriction enzyme sites for each amplified region were identified, reducing the size of the PCR products analysed to less than 250 bp. Bands with aberrant migration patterns were observed on SSCP gels in the lymphocyte DNA from two patients with bilateral, familial retinoblastoma. Sequence analysis of these DNA fragments revealed the causative mutations. These consisted of a 1-bp insertion of a T in the coding strand of exon 14 and a G→A mutation in the coding strand of exon 20. This approach has proved to be a powerful method for the rapid detection of germline mutations in the RB1 gene, a programme which can be extended to individuals with new mutations.

Many oncogenes require the production of an abnormal gene product to transform cells, but a class of 'recessive cancer genes' whose loss of function is required for initiation of tumorigenesis has been described (Weinberg, 1991). The retinoblastoma (Rb) predisposition gene (RB1) is one of these recessive cancer genes, and both copies must be inactivated for tumorigenesis to occur (Cavenee et al., 1985; Knudson, 1971; Yandell et al., 1989). In tumours this is achieved in 70% of cases by duplicating the initial mutation and losing the chromosome with the normal gene (Cavenee et al., 1983; Zhu et al., 1989). In the remaining cases functional inactivation of both copies of RB1 occurs as a result of independent inactivating mutational events. However, approximately one-half of retinoblastoma patients carry a constitutional predisposing mutation (Vogel, 1979). This means that their children have a 50:50 chance of inheriting the predisposition, and 90% of children who do will develop the tumour (Vogel, 1979). However, only 10–15% of Rb patients have a family history of the disease, the remaining 30–35% representing new germline mutations. For those patients with familial Rb, over 95% can be offered prenatal screening (Onadim & Cowell, 1991) using standard linkage analysis. However, for these with new germline mutations, unless the causative mutation is identified counselling will not be available for their firstborn children.

A detailed analysis of events which play a role in the production of an inactive Rb gene product has, until recently, been limited by the lack of rapid and sensitive detection techniques. Investigations have also been complicated because the gene is approximately 200 kb long, encodes a 4.7-kb mRNA and is composed of 27 exons ranging from 31 bp to 1873 bp in size (McGee et al., 1989).

Although analysis of RNA from tumours would be the most straightforward way of detecting mutations, only about one-third of tumours are removed, the rest being treated in situ. Of those which are removed, many have already been treated and the tumours are largely necrotic. An additional complicating factor is that, in up to 40% of tumours, the RB1 RNA is not produced (Friend et al., 1986; Fung et al., 1987; Lee et al., 1987; Goddard et al., 1988). Although, in most cases, DNA analysis of tumours would avoid complications due to the presence of the normal allele, for the successfully treated hereditary tumours this is not an option. RNAase protection has been used to identify mutations in RB1, but was only successful in an esti-
mated 50% of cases (Dunn et al., 1989). The examination of RNA by RNAase protection or the polymerase chain reaction (PCR) is limited because the normal allele appears to suppress expression of the mutant transcript (Dunn et al., 1988; Zhu et al., 1989; Murakami et al., 1991). Thus, a procedure for identifying heterozygous mutations in constitutional cells is required. Yandell et al. (1989) identified mutations in Rb tumours following exon-by-exon sequencing of the Rb1 gene, which is a very time-consuming process. To avoid the need to sequence the whole gene in patients and/or their tumours, several groups have analysed ways of prescreening exons for mutations. Some of the more recently described methods, following amplification of DNA by PCR, include the chemical cleavage mismatch method (CCM; Cotton et al., 1988), carbodiimide modification (Ganguly & Prockop, 1990) and denaturing (DGGE) and temperature gradient gel electrophoresis (TGGE) (Fischer & Leiman, 1983; Myers et al., 1985; Wartell et al., 1990). These techniques generally require several manipulations and extensive preparation. Consequently, we chose to use the recently described single-strand conformation polymorphism (SSCP) technique (Orita et al., 1989), which relies on the fact that the migration of single-stranded DNA molecules in non-denaturing polyacrylamide gels is sequence dependent.

In this report we present details for the construction of oligonucleotides for the amplification and analysis of all 27 exons of the Rb1 gene and demonstrate the application of SSCP and PCR sequencing to detect constitutional heterozygous mutations in Rb patients.

Results

Design of PCR primers

The full sequence of the coding region of Rb1, together with approximately 200 bp of the introns flanking each exon, has been reported by McGee et al. (1989). This sequence was used to design oligonucleotide primer pairs for the amplification of all 27 exons of the Rb1 gene. Since not all mutations occur within the exons, the sequences used for the oligonucleotide primers are generally situated well into the intron region. Details of primer sequences and their annealing temperatures are given in Table 1. The length of the flanking intron region included in the amplified product varied between 42 and 344 bp depending on the exon. The lengths of PCR products range from 212 to 625 bp (Table 1).

Following PCR amplification, the DNA fragment was analysed on a 2% agarose gel to check both the size and specificity of the product. The location of primers was influenced by several factors. Firstly, the size of product which can be easily sequenced is limited to 500–600 bp. Secondly, in some cases, 'background' bands were seen in addition to those of the expected size. Clearly, these background bands would complicate SSCP and sequence analysis. As a result, several sets of primers were designed for some exons before a single band was seen on agarose gels.

An additional factor which influenced the location of the primers was the GC–AT ratio. The higher the AT content, the lower the annealing temperature and the greater the likelihood of non-homologous pairing which results in background amplification. Generally, the higher the GC content of the primer, the higher the annealing temperature. To reduce background due to very low annealing temperatures, we avoided regions with an AT content of more than 70%. Where possible, GC-rich sequences were also avoided to reduce potential problems occurring as a result of the formation of secondary structures in the template. This can cause a variety of problems from non-specific priming to lack of product (Innes, 1990). Although primer pairs were usually the same size (mean of 22 bp), the length of some primers was altered to obtain a primer pair with a similar annealing temperature; these range from 19 to 27 bp in length. Primers of <19 bp generally generated non-specific amplification products. The promoter region, exon 1 and flanking intron region are extremely GC rich. To obtain a specific PCR amplification product for both the promoter region and exon 1 we found that addition of 10% dimethylsulphoxide to the PCR reaction was essential (Williams, 1989).

Although the majority of primer pairs were designed from sequences located towards the distal extremes of the sequenced intron regions, there were a few exceptions. The 3' primer for exon 20 is situated upstream to the highly polymorphic variable number of tandem repeats (VNTR) region comprised of varying numbers of CTTT(T) repeats (McGee et al., 1989). Since exons 15 and 16 are separated by only an 80-bp intron sequence, a single pair of primers was used covering both exons. All of the primers listed in Table 1 have been tested and their efficiency confirmed using a variety of DNAs from blood cells from normal individuals and patients with bilateral Rb in whom heterozygous mutations might be expected.

SSCP analysis

Single base pair mutations can be detected following SSCP analysis of PCR products up to 350 bp. However, subtle changes in nucleotide sequence are more readily detectable in smaller molecules (Orita et al., 1989). Restriction enzyme sites within each of the 27 PCR-amplifiable regions have been identified and generally result in the generation of DNA fragments between 100 and 300 bp (see Table 1). In some cases, it was necessary to use several enzymes, in combination, to generate fragments of the required length. We found that overnight digestion of the PCR products, under the conditions described in the Materials and methods section, produced complete digestion which is essential to avoid artefact bands on the SSCP gel as a result of incomplete digestion. The expected sizes of digested products were confirmed on agarose gels in all cases.

Analysis of patient DNA

To determine whether the PCR–SSCP analysis could be used in the identification of heterozygous mutations in patients known to carry a predisposing mutation, we performed a limited survey of patients with a prior family history. DNA samples used in this study were from Rb patients who had presented with bilateral tumours and were selected randomly from the large group of Rb families previously described by Onadim et al. (1990).
Table 1  Details of the oligonucleotide primers, the annealing temperatures for PCR amplification, restriction enzymes used to digest PCR products and sizes of cut fragments. Both the promoter region and exon 1 required 10% dimethylsulphoxide in the PCR reaction*

<table>
<thead>
<tr>
<th>Oligo Location</th>
<th>Sequence</th>
<th>Temperature (°C)</th>
<th>Full size (bp)</th>
<th>Restriction enzyme</th>
<th>Cut size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6007 5x PRO</td>
<td>GATCCCAAAGGCAGCAAGTGCTCT</td>
<td>62*</td>
<td>570</td>
<td>Smal</td>
<td>230</td>
</tr>
<tr>
<td>6008 3x PRO</td>
<td>TCAACGTGCCCCCTAGAAAACCGGG</td>
<td></td>
<td></td>
<td>BspHI</td>
<td>176/164</td>
</tr>
<tr>
<td>9691 5x1</td>
<td>CGTGGCGGCGGCTGTGTCCTT</td>
<td>62*</td>
<td>307</td>
<td>Ddel</td>
<td>153</td>
</tr>
<tr>
<td>9692 3x1</td>
<td>ACCGGGCCCCCTGGCAGAG</td>
<td></td>
<td></td>
<td></td>
<td>154</td>
</tr>
<tr>
<td>10143 5x2</td>
<td>GTGTTTTTTCACAGTGTAGAC</td>
<td></td>
<td>409</td>
<td>HpaI</td>
<td>214</td>
</tr>
<tr>
<td>10142 3x2</td>
<td>GTGCGGGGGCTCACAATATTTAA</td>
<td></td>
<td>591</td>
<td></td>
<td>191</td>
</tr>
<tr>
<td>8202 5x3</td>
<td>GCCATCGAGAGATGTGTACAA</td>
<td></td>
<td>477</td>
<td>AluI</td>
<td>243</td>
</tr>
<tr>
<td>8201 3x3</td>
<td>GGCACAACACTGCTACCTCTAAGG</td>
<td></td>
<td>234</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8203 5x4</td>
<td>CTTGACCAAGTTATAGTGTAGGTGG</td>
<td></td>
<td>445</td>
<td>RsaI</td>
<td>269</td>
</tr>
<tr>
<td>8204 3x4</td>
<td>CGGAGGACTTCTGAATCTGCTTTT</td>
<td></td>
<td>176</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8206 5x5</td>
<td>GAGAGCTAATTGGAGGAACTGAC</td>
<td></td>
<td>488</td>
<td>AsfIHI</td>
<td>218</td>
</tr>
<tr>
<td>8205 3x5</td>
<td>GTCCTCTGCCATTCCCTATT</td>
<td></td>
<td>142/128</td>
<td>TaqI</td>
<td>81</td>
</tr>
<tr>
<td>8208 5x6</td>
<td>GAAACCCCAAGGATATATCTGG</td>
<td></td>
<td>326</td>
<td>AluI</td>
<td>179</td>
</tr>
<tr>
<td>8207 3x6</td>
<td>CCAAGGTGTTGGTCTAGACCA</td>
<td></td>
<td>147</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7900 5x7</td>
<td>ACTCTACCTCTGCTTCTCTATC</td>
<td></td>
<td>491</td>
<td>RsaI</td>
<td>193/176</td>
</tr>
<tr>
<td>7901 3x7</td>
<td>CTTCTGTTCTCCCTAACCCATCTTGG</td>
<td></td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20085 5x8</td>
<td>GACCTAAGATTAGTTAAGATACCTCT</td>
<td></td>
<td>316</td>
<td>TaqI</td>
<td>192</td>
</tr>
<tr>
<td>20084 3x8</td>
<td>CAGGACTCTATAACAAAGAAGTAAA</td>
<td></td>
<td>124</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7905 5x9</td>
<td>TGCATGGGGGATGACACCTTAC</td>
<td></td>
<td>316</td>
<td>EcoRI</td>
<td>171</td>
</tr>
<tr>
<td>7904 3x9</td>
<td>CTGTTGATGAGACACAACTTAC</td>
<td></td>
<td>145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10145 5x10</td>
<td>TCGTGACTCTACTTTAGTAGACCC</td>
<td></td>
<td>492</td>
<td>BglII</td>
<td>225/218</td>
</tr>
<tr>
<td>10144 3x10</td>
<td>CTGTTGATGAGACACAACTTAC</td>
<td></td>
<td>117</td>
<td>Hinfl</td>
<td>49</td>
</tr>
<tr>
<td>11289 5x11</td>
<td>GAAACCAAGAAGGATTATACGTG</td>
<td></td>
<td>294</td>
<td>MboII</td>
<td>163</td>
</tr>
<tr>
<td>11290 3x11</td>
<td>CCTGCGCTCTATATATATTCT</td>
<td></td>
<td>131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9987 5x12</td>
<td>CCAAGCTTCTTATTTGAAGATG</td>
<td></td>
<td>645</td>
<td>BclI</td>
<td>198</td>
</tr>
<tr>
<td>9986 3x12</td>
<td>GTGAGAAAGGCAATAAGGAC</td>
<td></td>
<td>162</td>
<td>Ndel</td>
<td>209</td>
</tr>
<tr>
<td>5528 5x13</td>
<td>TAATAGGTTTTTTGATGGTACTG</td>
<td></td>
<td>570</td>
<td>EcoRI</td>
<td>232</td>
</tr>
<tr>
<td>5529 3x13</td>
<td>AATTTTCTAATGACTGATGTTC</td>
<td></td>
<td>213</td>
<td>Hinfl</td>
<td>113</td>
</tr>
<tr>
<td>13679 5x14</td>
<td>CTAAAAATAGCGGCTCTTTATTTTC</td>
<td></td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13680 3x14</td>
<td>ATCTTGCGCGTCTTGGCCTCTCTGT</td>
<td></td>
<td>212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11293 5x15</td>
<td>ATGATCGGCTAGCAGCAAATAAGGGTT</td>
<td></td>
<td>361</td>
<td>Ndel</td>
<td>209</td>
</tr>
<tr>
<td>11294 3x16</td>
<td>TTTCTCTTTAATCTGACACACATCTCC</td>
<td></td>
<td>152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5535 5x17</td>
<td>ATAATAATGTTAATACGCTTACTT</td>
<td></td>
<td>555</td>
<td>RsaI</td>
<td>142/119</td>
</tr>
<tr>
<td>5536 3x17</td>
<td>GTTGCGATTTTAAAGAAGCTATT</td>
<td></td>
<td>555</td>
<td>Ndel</td>
<td>104/101</td>
</tr>
<tr>
<td>10604 5x18</td>
<td>ATGTACCCTGGCAAATATTTG</td>
<td></td>
<td>221</td>
<td>BglI</td>
<td>113</td>
</tr>
<tr>
<td>10603 3x18</td>
<td>CTATTGGCCTGGTATGTTGGAAC</td>
<td></td>
<td>108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9439 5x19</td>
<td>AGGGCGTAAATCCCGAGAAGACCA</td>
<td></td>
<td>485</td>
<td>Hinfl</td>
<td>218</td>
</tr>
<tr>
<td>9440 3x19</td>
<td>CTACAGTGAAGTAACTTTGCTGCC</td>
<td></td>
<td>176/91</td>
<td>MluI</td>
<td>89</td>
</tr>
<tr>
<td>9438 5x20</td>
<td>TTCTCTGGGGGAAAGAAAGGAGTG</td>
<td></td>
<td>350</td>
<td>HpaII</td>
<td>177</td>
</tr>
<tr>
<td>14928 3x20</td>
<td>GACAATACGAGATGACGAGAGA</td>
<td></td>
<td>173</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9436 5x21</td>
<td>GACCTTCAATTCGCAGTTAAGG</td>
<td></td>
<td>518</td>
<td>Asel</td>
<td>180</td>
</tr>
<tr>
<td>9435 3x21</td>
<td>AACAAATCATCTGATATACAGGGAG</td>
<td></td>
<td>173</td>
<td>Ndel</td>
<td>156/164</td>
</tr>
<tr>
<td>9434 5x22</td>
<td>GACCACCTAAATCAGCAGATGAAG</td>
<td></td>
<td>363</td>
<td>BclI</td>
<td>210</td>
</tr>
<tr>
<td>9433 3x22</td>
<td>GATTTTTTGGAGGACCATACATAGGA</td>
<td></td>
<td>153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9694 5x23</td>
<td>TCTAAATGTAAATGCGCTGAAAAC</td>
<td></td>
<td>420</td>
<td>BspNI</td>
<td>186</td>
</tr>
<tr>
<td>9695 3x23</td>
<td>CACTCTGGTTTGGCTTTTGATGTA</td>
<td></td>
<td>420</td>
<td>AluI</td>
<td>136/98</td>
</tr>
<tr>
<td>9696 5x24</td>
<td>TAAACAACTGGAGACTGGTATGAT</td>
<td></td>
<td>579</td>
<td>BstEII</td>
<td>235/200</td>
</tr>
<tr>
<td>9697 3x24</td>
<td>TAGATTTGGGGTGGATGAAAGAAAATCTC</td>
<td></td>
<td>93/50</td>
<td>HindIII/BspNI</td>
<td>89</td>
</tr>
<tr>
<td>9991 5x25</td>
<td>ATTGGTGCTAATGAGCAGAAAATT</td>
<td></td>
<td>625</td>
<td>BspNI</td>
<td>266</td>
</tr>
<tr>
<td>9990 3x25</td>
<td>TGATGCTATGTATTTTTCGCTG</td>
<td></td>
<td>198/161</td>
<td>BspI286I</td>
<td>89</td>
</tr>
<tr>
<td>9993 5x26</td>
<td>AACCACCCAATTTGCTGAAACCAC</td>
<td></td>
<td>524</td>
<td>PstI</td>
<td>215</td>
</tr>
<tr>
<td>9992 3x26</td>
<td>TGACTGTGTGTCACGAAAATTT</td>
<td></td>
<td>166/143</td>
<td>NotI</td>
<td>91</td>
</tr>
<tr>
<td>10609 5x27</td>
<td>AAGGCTCCTGAGGGCTCACATGTTGGA</td>
<td></td>
<td>218</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10608 3x27</td>
<td>GAAGTGTCATGCTAGTGCAACAGCCAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Although we have not yet analysed all exons in all patients, we report here details from two patients, S.T. and A.F., to illustrate the effectiveness of our screening approach. Analysis of exon 14 showed an abnormal SSCP banding pattern (Figure 1) for patient A.F. When compared with DNA from a normal individual, A.F. contained the same number of bands in the upper part of the gel, although two of the four bands were thickened and appeared to contain a slower migrating component which could not be resolved even on longer gel runs. We would predict, therefore, that this patient is heterozygous for the mutation. Sequence analysis of A.F. DNA showed that, in addition to the normal sequence, a band in the A lane was present in the same position as the G seen in the normal control (Figure 2). The intensity of these two bands was 50% of that seen in the normal control, confirming the heterozygous nature of the first mutation. The mutation is 31 bp from the 5' end of exon 14 on the coding strand and converts codon 455 from an arginine (CGA) to a stop codon (TGA).

A 350-bp fragment containing exon 20 was amplified using DNA purified from the blood of several patients and two additional bands were present in the DNA of patient S.T. (Figure 3, lane 3). Sequence analysis revealed a 1-bp insertion 52 bp from the 5' end of exon 20. An extra A was noted in the antisense strand (Figure 4). From the point of insertion both the mutant and normal sequencing ladders are superimposed, one lagging by 1 bp. Alteration of the reading frame results in the generation of a premature stop codon at codon 672. This mutation destroyed a Ddel site. Digestion of DNA from this patient with Ddel shows the typical digested band pattern in addition to the full-length undigested fragment (Figure 5).

**Multiplex SSCP**

Exon-by-exon analysis of the *RBL* gene using SSCP is relatively time-consuming. We tried to improve the efficiency of the process by using combinations of primers with the same annealing temperature in a
DETECTION OF RB1 MUTATIONS USING SSCP

Figure 4 Sequence analysis of exon 20 from patient S.T. Primer 9438 was biotinylated and used with primer 14928 to amplify exon 20. Sequencing of the immobilized strand from primer 14928 is shown for DNA from patient S.T. and normal DNA. Insertion of an 'A' was present in the DNA of S.T. in one allele (arrow). As a result of the insertion, the sequence ladder of the mutant allele is superimposed on the normal, resulting in two bands at each position on the gel.

Figure 5 Confirming the presence of the mutation in patient S.T. A 350-bp DNA fragment from S.T. (lane 1) was amplified and digested with Ddel (lane 3). DNA markers are shown in lane 2 and their sizes are indicated at the right of the figure. Ddel digests of normal DNA (lane 4) gave bands of expected size (145 and 205 bp). Ddel-digested DNA from S.T. contained both full-length and digested fragments, indicating the presence of the mutant allele.

Discussion

One of the major challenges for the clinical management of Rb is the identification of genetically predisposed individuals. These individuals can be screened ophthalmologically and tumours treated as they arise. Identification of the causative mutations in Rb patients will allow unequivocal diagnosis of carriers of germline mutations and unaffected gene carriers in Rb families and allow prenatal screening in these cases. We have demonstrated here that, using a combination of SSCP and PCR sequencing, it is possible to detect heterozygous mutations in constitutional cells, which is essential given that tumour tissue is not often available. At present the siblings of all 'new' cases presenting with Rb will be subjected to regular ophthalmological examinations. Thus, the unequivocal identification of mutant gene carriers will also eliminate the need for the costly and time-consuming procedures for confirmed non-gene carriers.

We have presented mutations in two individuals to illustrate the effectiveness of the SSCP procedure. These mutations consist of a 1-bp insertion in exon 20 in patient S.T. and a C→T single base pair change in the coding strand in exon 14 in patient A.F. Both mutations would be predicted to result in the production of truncated proteins of 672 and 455 amino acids respectively.

Although mutant gene carriers in Rb families can be identified using linkage analysis (Onadim et al., 1990), a few uninformative families still do not qualify for prenatal screening. Patient S.T. was one such example, having one affected and one unaffected child but being uninformative for all of the predictive polymorphisms. Using SSCP analysis, we were able to identify the predisposing mutation in exon 20. Because the 1-bp insertion affects a restriction enzyme site, it will be possible in the future to identify mutant gene carriers in this family using this simple procedure. However, not all mutations affect known restriction enzyme sites and DNA sequencing will still be required.

Although we chose to use the SSCP technique for screening the RB1 exons, because of its relative speed and simplicity, it is still not clear whether this approach will successfully identify all mutations in the RB1 gene. However, in addition to the initial studies on the ras oncogenes (Orita et al., 1989), it has been used successfully to detect mutations in the cystic fibrosis gene (Dean et al., 1990) and the p53 gene (Mazars et al., 1991). Other techniques have been used in the analysis of mutation epidemiology but are limited by the lengthy procedure involved. For large genes, where mRNA is not always produced, exon-by-exon analysis of the genomic DNA is necessary. For these reasons we prefer the SSCP method, and it will not be until sufficient numbers of patients have been analysed that the efficiency of the procedure can be determined. Analysis of SSCP gels depends on the identification of changes in the normal banding pattern but does not
provide any information about the type of mutation. Samples with altered banding patterns are sequenced and so far, in these cases, we have always found a mutation. Multiplex SSCP provides a quicker way to screen a gene. However, digestion of the samples following co-amplification is not feasible because enzyme sites are often present in inappropriate positions in other fragments and an extremely complex gel pattern would be generated. Hence, full-length PCR products must be used and sensitivity may be decreased for fragments > 200 bp in length (K. Hayashi, personal communication).

It is a theoretical possibility that mutations in other parts of the gene may also have occurred. We plan to screen all exons in a subset of patients to investigate this possibility. It is important not to accept base pair changes which do not lead to changes in the amino acid sequence as mutations and to be aware that these changes may result in changes in SSCP banding profiles.

We have described the use of PCR and SSCP for the analysis of mutations in the RB1 gene and have shown that this technique has potential as a rapid and efficient aid to genetic counselling of familial Rb. In addition, it offers the possibility of the further characterization of mutational events in the RB1 gene, which has been shown to be abnormal in a wide range of human cancers. We have also previously described the PCR amplification of DNA from paraffin blocks (Onadim & Cowell, 1991). SSCP analysis of this DNA would open up the possibility of studying archival material.

Materials and methods

PCR amplification

DNA was amplified in a Techne PHC-2 thermocycler. Oligonucleotide primers for PCR were synthesized on a phosphoramidite column (ICRF, Central Services division). The PCR reaction mixture consisted of approximately 100 ng of DNA, 50 pmol of each PCR primer, Taq polymerase reaction buffer (Promega, Madison, WI, USA) and 0.2 mM each of dATP, dGTP, dTTP and dCTP (Pharmacia, Milton Keynes, UK). This cocktail was overlaid with 50 μl of mineral oil to prevent evaporation, and an initial denaturation, the temperature was decreased to 65°C and 1 U of Taq polymerase at 5 U ml⁻¹ (Promega) was added through 0.5 min and an extension step at 72°C for 1.0 min were carried out.

PCR cycles consisting of an initial denaturation step at 96°C for 0.5 min followed by an annealing step at a temperature optimal for the particular primer pair (see Table 1) for 0.5 min and an extension step at 72°C for 1.0 min were carried out.

SSCP

For SSCP analysis, individual exons were amplified as described above except that 1 μCi of [³²P]dCTP (3000 Ci mmol⁻¹) (Amersham, UK) was added to the PCR reaction mix at the outset to generate a labelled product and the ‘cold’ dCTP concentration was reduced to 0.02 mM. A 5-μl aliquot (10%) of the PCR-amplified product was then diluted with 40 μl of 0.1% SDS/10 mM EDTA and 2 μl of this mixture was added to 2 μl of Sequenase stop mix: 95% formamide, 20 mM EDTA, 0.05% bromphenol blue, 0.05% xylene cyanol (United States Biochemicals, OH, USA). Samples were denatured at 95°C for 3 min and placed immediately on ice to prevent renaturation before being run on 0.3 mm × 40 cm × 30 cm, 6% polyacrylamide non-denaturing gels containing 10% glycerol. The DNA was electrophoresed in TBE (0.09 M Tris base, 0.09 M boric acid and 2.5 mM EDTA) running buffer at 30 W at room temperature. Gels were dried and exposed to XAR-5 film (Kodak) for 12–72 h without an intensifying screen, which tends to blur the image.

Restriction enzyme digestion of PCR products

When it was necessary to generate shorter fragments for SSCP analysis, 15 μl of the PCR-amplified product was removed from beneath the oil and incubated overnight with 8 U of each of the appropriate restriction enzymes at 37°C (65°C for TaqI and 50°C for BclI). Enzymes were obtained mainly from Bioexcellence, Essex, UK, but also from New England Biolabs, Beverly, MA, USA, and Northumbrian Biologicals, Northumberland, UK.

Direct sequencing

When the PCR product was to be used for sequencing, amplification was performed with one non-biotinylated primer and one primer biotinylated at the 5' end. This was to allow immobilization of single-stranded DNA on streptavidin-coated magnetic ‘Dynabeads’, (Dynal, Merseyside, UK). The amount of primer used was decreased to 5–15 pmol. This does not affect the quality of the PCR product but was necessary to avoid saturating the streptavidin with excess primer in preference to full-length PCR amplification products. The supernatant was removed from a 30-μl aliquot of resuspended beads after placing them in a magnetic particle concentrator (MPC; Dynal) for 30 s. The MPC immobilizes the beads, allowing removal of supernatant. The beads were washed twice in 100 μl of 10 mM Tris–HCl, pH 8.0, 1 mM EDTA, 0.1 mM sodium chloride (TES), immobilized on the MPC and the supernatant removed, as described above. The amplified PCR product was carefully taken from beneath the mineral oil and added to the washed beads, which were then gently resuspended. Adsorption of the DNA to the beads took place at room temperature for 5 min with occasional gentle mixing. Following adsorption of the DNA to the beads, the supernatant (containing excess unbound DNA) was removed on the MPC and the beads were resuspended in 100 μl of 0.15 M sodium hydroxide and left for 5 min to denature the double-stranded DNA. The non-biotinylated strands were then removed while the beads with the biotinylated strands attached were immobilized on the MPC. The non-biotinylated strands remained precipitated by adding 10 μl of 3 M sodium acetate, pH 5.6, and 250 μl of 100% ethanol and left overnight at −70°C. The DNA was recovered by centrifugation and the pellet was washed in 75% ethanol and resuspended in 5 μl of water. The immobilized biotinylated DNA was washed with 100 μl of TES, followed by 100 μl of water, and finally resuspended in 5 μl of water.

Dideoxy sequencing using a Sequenase kit (United States Biochemicals) was performed on the non-biotinylated fraction and on the immobilized biotinylated strands, thus confirming the presence of the mutation in both directions. To ensure that extension from the primer during the labelling step (when dideoxy nucleotide are not present) was minimal, the labelling mix was diluted 1:15, which allowed the sequence close to the primer to be read.

Acknowledgements

The blood samples used in this study were collected as part of a genetic analysis of Rb patients in the UK, and we are indebted to the Staff of the Ophthalmology departments of St Bartholomew's and Moorfields Eye Hospital, and in par-
ticular Dr Marcelle Jay and John Hungerford. Our special thanks go to Dr D. Yandell for sharing his experience with SSCP and PCR sequencing and Dr I. Goldsmith and his group at the ICRF for supplying the oligonucleotides. A.H. and Z.O. are supported by a grant from the David Allen Retinoblastoma Appeal.

References

Oncogenic point mutations in exon 20 of the \textit{RB1} gene in families showing incomplete penetrance and mild expression of the retinoblastoma phenotype

\textit{Genetics}

\textit{Proc. Natl. Acad. Sci. USA}

\textbf{Vol. 89}, pp. 6177-6181, July 1992

\begin{abstract}

The retinoblastoma-predisposition gene, \textit{RB1}, segregates as an autosomal dominant trait with high (90\%) penetrance. Certain families, however, show an unusual low-penetrance phenotype with many individuals being unaffected, unilaterally affected, or with evidence of spontaneously regressed tumors. We have used single-strand conformation polymorphism analysis and PCR sequencing to study two such families. Mutations were found in exon 20 of \textit{RB1} in both cases. In one family a C \rightarrow T transition in codon 661 converts an arginine (CGG) to a tryptophan (TGG) codon. In this family, incomplete penetrance and mild phenotypic expression were observed in virtually all patients, possibly indicating that single amino acid changes may modify protein structure/function such that tumorigenesis is not inevitable. In the second family the mutation in codon 675 is a G \rightarrow T transversion that converts a glutamine (GAA) to a stop (TAG) codon. However, this mutation also occurs near a potential cryptic splice acceptor site, raising the possibility of alternative splicing resulting in a less severely disrupted protein.

Retinoblastoma (RB) is an intraocular eye tumor with an incidence of 1 in 15,000-25,000 (1). It occurs predominantly in children under the age of 2 years and is rare over the age of 5 years. Approximately 15\% of all RB patients have a prior family history and the tumor phenotype segregates as an autosomal dominant trait in most cases (1). In \sim 10\% of families, unaffected individuals can be identified who can transmit the mutant gene (1). This phenomenon is referred to as "incomplete penetrance."

Knudson (2) demonstrated that, in cases of hereditary RB, a second random mutation must occur for tumor development, and he provided one explanation of the phenomenon of incomplete penetrance by suggesting that these patients formed part of a Poisson distribution where, by chance, the second random mutation did not occur. These patients, however, still carry germ-line mutations and their children have a 50\% chance of inheriting the predisposing mutation. Knudson’s "two-hit" hypothesis also predicted that mutant gene carriers would develop mostly bilateral, unilateral tumors with an earlier age of onset when compared with sporadic cases, which would be mostly unilateral, unifocal, and present later in life. However, the distribution of cases of incomplete penetrance is not entirely random and families have been reported where the majority of gene carriers are either unaffected or only unilaterally affected (3–6). These we refer to as "low-penetrance" families.

Another feature of RB is that it sometimes apparently regresses spontaneously, leaving characteristic scars on the retina (4, 5). An alternative suggestion is that these scars represent a more benign form of the disease, retinoma (7).

Identification of unaffected mutant gene carriers has, until recently, required that they have affected children, but use of classical linkage analysis and restriction fragment length polymorphisms (RFLPs) from within the \textit{RB1} gene (8, 9) has allowed unequivocal identification of cases of incomplete penetrance in RB families (5, 6). In addition, patients with typical retinal scarring and a family history also carry a mutant \textit{RB1} gene (5). This heterogeneity is difficult to interpret in terms of the Knudson two-hit hypothesis, especially when the majority of gene carriers are only unilaterally affected. We suggest that there are alleles of the \textit{RB1} gene that are only partially defective and, as a consequence, give rise to a mild or incompletely penetrant phenotype. With the cloning of \textit{RB1} (10), it should now be possible to determine the nature of mutations in low-penetrance families to establish whether they affect the gene in a distinct way or simply reflect random mutation events throughout the gene.

\textit{RB1} has a complex structure with 27 exons varying in size from 31 to 1873 base pairs (bp) (11) and two very large (>33-kilobase (kb)) introns. The sequences of intron regions immediately adjacent to each exon have been determined (11), allowing amplification of individual exons by the polymerase chain reaction (PCR). This approach has already been used to identify mutations in RB tumors (12). Such a screening program can proceed more rapidly with analysis of each exon of \textit{RB1} by single-strand conformation polymorphism (SSCP) analysis (13). This procedure depends on the sequence dependence of the migration of a single-stranded DNA molecule in a nondenaturing polyacrylamide gel. Hence, mutations affecting a DNA sequence will lead to a conformational change affecting mobility and produce novel bands on gels.

We have used standard linkage analysis in RB families to identify cases of incomplete penetrance and have undertaken an exon-by-exon SSCP analysis combined with PCR sequencing to identify the specific mutation responsible for mild phenotypes in two low-penetrance families.

\textbf{MATERIALS AND METHODS}

\textbf{PCR Amplification.} Two primers were used to amplify a 350-bp fragment including the entire 146 bp of exon 20 and flanking intron sequences. The 5' primer (no. 9438), 5'-TTTCTGTGGGAGAAAGGAGTGG-3', was located in intron 19 and the 3' primer (no. 14928), 5'-AGTTAACAGTAAGTGAGGAGGAGA-3', was located in intron 20. For direct sequencing from PCR products either the biotinylated version of primer 9438 (no. 18322) or a biotinylated internal primer, 5'-CATGATTTGAAAAATCTACTTG-3' (no. 11928), was used. The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. \S1734 solely to indicate this fact.

\begin{thebibliography}

\item \textbf{ABSTRACT}

The retinoblastoma-predisposition gene, \textit{RB1}, segregates as an autosomal dominant trait with high (90\%) penetrance. Certain families, however, show an unusual low-penetrance phenotype with many individuals being unaffected, unilaterally affected, or with evidence of spontaneously regressed tumors. We have used single-strand conformation polymorphism analysis and PCR sequencing to study two such families. Mutations were found in exon 20 of \textit{RB1} in both cases. In one family a C \rightarrow T transition in codon 661 converts an arginine (CGG) to a tryptophan (TGG) codon. In this family, incomplete penetrance and mild phenotypic expression were observed in virtually all patients, possibly indicating that single amino acid changes may modify protein structure/function such that tumorigenesis is not inevitable. In the second family the mutation in codon 675 is a G \rightarrow T transversion that converts a glutamine (GAA) to a stop (TAG) codon. However, this mutation also occurs near a potential cryptic splice acceptor site, raising the possibility of alternative splicing resulting in a less severely disrupted protein.

Retinoblastoma (RB) is an intraocular eye tumor with an incidence of 1 in 15,000–25,000 (1). It occurs predominantly in children under the age of 2 years and is rare over the age of 5 years. Approximately 15\% of all RB patients have a prior family history and the tumor phenotype segregates as an autosomal dominant trait in most cases (1). In \sim 10\% of families, unaffected individuals can be identified who can transmit the mutant gene (1). This phenomenon is referred to as "incomplete penetrance."

Knudson (2) demonstrated that, in cases of hereditary RB, a single, additional, random genetic event was required for tumor development, and he provided one explanation of the phenomenon of incomplete penetrance by suggesting that these patients formed part of a Poisson distribution where, by chance, the second random mutation did not occur. These patients, however, still carry germ-line mutations and their children have a 50\% chance of inheriting the predisposing mutation. Knudson’s "two-hit" hypothesis also predicted that mutant gene carriers would develop mostly bilateral, unilateral tumors with an earlier age of onset when compared with sporadic cases, which would be mostly unilateral, unifocal, and present later in life. However, the distribution of cases of incomplete penetrance is not entirely random and families have been reported where the majority of gene carriers are either unaffected or only unilaterally affected (3–6). These we refer to as "low-penetrance" families.

Another feature of RB is that it sometimes apparently regresses spontaneously, leaving characteristic scars on the retina (4, 5). An alternative suggestion is that these scars represent a more benign form of the disease, retinoma (7).

Identification of unaffected mutant gene carriers has, until recently, required that they have affected children, but use of classical linkage analysis and restriction fragment length polymorphisms (RFLPs) from within the \textit{RB1} gene (8, 9) has allowed unequivocal identification of cases of incomplete penetrance in RB families (5, 6). In addition, patients with typical retinal scarring and a family history also carry a mutant \textit{RB1} gene (5). This heterogeneity is difficult to interpret in terms of the Knudson two-hit hypothesis, especially when the majority of gene carriers are only unilaterally affected. We suggest that there are alleles of the \textit{RB1} gene that are only partially defective and, as a consequence, give rise to a mild or incompletely penetrant phenotype. With the cloning of \textit{RB1} (10), it should now be possible to determine the nature of mutations in low-penetrance families to establish whether they affect the gene in a distinct way or simply reflect random mutation events throughout the gene.

\textit{RB1} has a complex structure with 27 exons varying in size from 31 to 1873 base pairs (bp) (11) and two very large (>33-kilobase (kb)) introns. The sequences of intron regions immediately adjacent to each exon have been determined (11), allowing amplification of individual exons by the polymerase chain reaction (PCR). This approach has already been used to identify mutations in RB tumors (12). Such a screening program can proceed more rapidly with analysis of each exon of \textit{RB1} by single-strand conformation polymorphism (SSCP) analysis (13). This procedure depends on the sequence dependence of the migration of a single-stranded DNA molecule in a nondenaturing polyacrylamide gel. Hence, mutations affecting a DNA sequence will lead to a conformational change affecting mobility and produce novel bands on gels.

We have used standard linkage analysis in RB families to identify cases of incomplete penetrance and have undertaken an exon-by-exon SSCP analysis combined with PCR sequencing to identify the specific mutation responsible for mild phenotypes in two low-penetrance families.

\textbf{MATERIALS AND METHODS}

\textbf{PCR Amplification.} Two primers were used to amplify a 350-bp fragment including the entire 146 bp of exon 20 and flanking intron sequences. The 5' primer (no. 9438), 5'-TTTCTGTGGGAGAAAGGAGTGG-3', was located in intron 19 and the 3' primer (no. 14928), 5'-AGTTAACAGTAAGTGAGGAGGAGA-3', was located in intron 20. For direct sequencing from PCR products either the biotinylated version of primer 9438 (no. 18322) or a biotinylated internal primer, 5'-CATGATTTGAAAAATCTACTTG-3' (no. 11928), was used. The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. \S1734 solely to indicate this fact.

Abbreviations: RB, retinoblastoma; RFLP, restriction fragment length polymorphism; SSCP, single-strand conformation polymorphism.

*To whom reprint requests should be addressed.
was used with primer 14928. The 17295/14928 primer pair amplifies a 269-bp fragment. This particular primer set was also used to amplify exon 20 for Msp I digestion in family RBF29.

PCR was carried out essentially as described by Hogg et al. (14) with 30 cycles of denaturation at 94°C for 20 sec., annealing at 60°C (for the 9438/14928 primer pair) or 52°C (for the 17295/14928 primer pair) for 20 sec., and extension at 72°C for 60 sec.

SSCP and Sequencing. Detailed methods for SSCP and direct sequencing from PCR products have been described (14). For SSCP, exon 20 was amplified using primers 9438 and 14928 in a PCR mixture with 1 μCi of [α-32P]dCTP (3000 Ci/mmol; Amersham; 1 Ci = 37 GBq) added and a nonradioactive dCTP concentration of only 0.02 mM. Denatured samples were electrophoresed in nondenaturing 6% (wt/vol) polyacrylamide/10% (vol/vol) glycerol gels. For direct sequencing, primers 18322/14928 or 17295/14928 were used. Primers 18322 and 17295 were biotinylated at the 5' end to allow immobilization of single-stranded DNA on streptavi-
d-coated magnetic Dynabeads (Dynal; Merseyside, UK), which were used to separate the DNA strands (14). Both single strands were sequenced by using a Sequenase kit (United States Biochemical) according to the manufacturer's instructions.

RESULTS

Included in our SSCP analysis of a large number of patients with hereditary RB were two families that showed an unusual pattern of inheritance in that many individuals had "mild" forms of the disease or were unaffected gene carriers. The pedigrees of these two "low-penetrance" families, RBF29 and RBF18, are given in Fig. 1. In both cases, abnormal banding patterns were seen in the SSCP gel for exon 20.

RBF29. Linkage analysis for family RBF29 (Fig. 1) was reported previously (5). In this family there are six affected individuals (I.1, I.3, II.2, II.5, III.3, III.4), as well as four unaffected gene carriers (II.4, II.7, III.1, III.2; arrows in Fig. 1) who were identified by using the RS2.0 polymorphism (5).
The parents (II.3, II.4) of the bilaterally affected twins (III.3, III.4) are first cousins and each has a unilaterally affected sister (II.2, II.5). The father of the twins (II.4) is an unaffected member of a family showing a Mendelian pattern of inheritance. The eyes of the parents, I.1 and I.3, were examined and spontaneously regressed tumors were identified in both.

Our preliminary screening program involved exon-by-exon analysis of the RB1 gene using SSCP in conjunction with PCR sequencing. Identified in exon 20 in the DNA of individual II.4 from RBF29 in SSCP gels (Fig. 2) and not in any of the other (20 out of 27) exons analyzed. In the SSCP gel shown in Fig. 2 the 350-bp fragment encompassing exon 20 amplified from II.4 exhibited an additional lower band when compared with samples from other patients. Sequence analysis of exon 20 from II.4 in RBF29 (Fig. 3) revealed a heterozygous C → T transition in the coding strand, 21 bases from the 5' end of exon 20. This mutation converts codon 661 from an arginine (CGG) to a tryptophan (TGG) codon. The same mutation was identified in the DNA of all the affected members and unaffected gene carriers in RBF29, but not in the unaffected members of the family. The mutation in RBF29 occurs within an Hpa II/Msp I restriction site in exon 20 (CCGG → CTGG), so that the presence of an undigested 269-bp fragment indicates a mutant gene carrier (Fig. 4). The undigested 269-bp band was observed in the DNA of all affected members and unaffected gene carriers, whereas only the two smaller normal bands were seen in individuals from this family identified not to be gene carriers (Fig. 4). DNA from III.3 was obtained for prenatal screening from chorionic villus (CV) tissue and from cord blood samples (B) taken from the same individual at birth. Exon 20 from II.4 was also sequenced and found to be free of the 661 mutation, confirming our original prediction made with linkage analysis (15). Msp I digestions of exon 20 DNAs from 34 unrelated RB patients and 38 unrelated healthy individuals did not show this mutation.

RBF18. Family RBF18, first reported by Hine (16), shows eight affected individuals in the family (Fig. 1), three of whom (II.2, III.3, IV.1) had unilateral disease. Four generations of males (I.1, II.4, III.3, IV.1) have a mild form of the disease. Although I.1 had one eye removed for RB, the tumor in his other eye regressed naturally. II.4 has spontaneously regressed tumors in both eyes and III.3 and IV.1 have unilateral regressed tumors. II.1 and II.3, however, died at the age of 1 year 9 months and 3.5 years, respectively, as a result of orbital extension (16) of their tumors. I.1 and II.2 died of second tumors at the age of 61 and 40, respectively. Only DNA from individuals III.2, III.3, III.4, and IV.1 was available to us. The RB predisposition is linked to the 4.95-kb allele in the R0.6 polymorphism (6).

SSCP analysis of DNA from III.3 from RBF18 showed an additional upper band (Fig. 2) when compared with other samples. Sequence analysis of exon 20 from III.3 (Fig. 5) revealed a heterozygous G → T transversion in the coding strand, 63 bases from the 5' end of exon 20. This mutation converts codon 675 from a glutamine (GAA) to a stop (TAA) codon. However, 3 bp downstream of this mutation lies a TAG sequence (Fig. 6), which is compatible with a consensus splice acceptor sequence (17). The G → T transversion removes an AG dinucleotide (converting it to AT) that ordinarily would have prevented the downstream TAG sequence from becoming a splice acceptor site. In addition, this transversion increases the pyrimidine/purine ratio in the region immediately preceding the TAG site, thereby enhancing its potential to be a cryptic splice site (18). A branch-point sequence exists (Fig. 6) upstream of this cryptic site, which, although not as good as the real branch point in intron 19, would nevertheless be adequate if this site were activated. In this case, the reading frame would be intact but the first 23 amino acids encoded by exon 20, codons 654–676, would be lost. Such a deletion would disrupt the leucine zipper motif in exon 20, removing 3 of the 4 leucines. The same mutation was identified in the DNA of III.2 and IV.1, who were known to carry the RB-predisposition gene. III.4, who is an unaffected normal individual, has the normal sequence. This G → T mutation does not alter any known restriction enzyme site.

**DISCUSSION**

There have been few reports of mutations within the RB1 gene in patients with RB, and those mutations that have been reported rarely involve amino acid changes that are under intense selective pressure. The data presented here demonstrate that an amino acid substitution of the type found in RB1 does occur, and may therefore be under some selective pressure. Furthermore, the mutations described here involve sites in the protein that could be involved in the interaction of a tumor suppressor protein with a partner protein involved in signal transduction. The RB protein is thought to be involved in the regulation of cell growth and division, and the mutation described here could affect this function. The data presented here demonstrate that an amino acid substitution of the type found in RB1 does occur, and may therefore be under some selective pressure.
Fig. 5. Sequence from the biotinylated (5') strand of exon 20 showing the heterozygous C → A transversion in RBF18 (Left) compared with the same sequence from a normal individual (Right).

described seem to be randomly distributed throughout the gene (12, 19). The majority of these mutations in bilaterally, multifocally affected individuals cause major disruptions of the gene and its consequent processing, resulting in nonproliferating cells. The exact role of RBI in tumorigenesis is still not fully understood, but its product, RBI, seems to be part of a signaling pathway controlling cell proliferation (21). The RBI protein interacts with the transforming oncproteins of DNA tumor viruses (22–24). The tumorigenesis is still not fully understood, but its product, exons 13–17 and 18–22, amino acids 393–572 and 646–772, respectively (25, 26). It has been suggested that, with appropriate folding of RBI, a "pocket" is created that facilitates binding to the viral transforming proteins and endogenous cellular proteins (27). The observation that RBI mutations exist that produce proteins which fail to bind viral proteins or to associate with endogenous cellular proteins has led to the suggestion that this pocket contributes to the growth-regulatory function of RBI (27–30).

Recently, mutations in the promoter region of RBI were detected in two families with low-penetration phenotypes (31). A plausible explanation for this is that promoter mutations may result in reduced levels of RBI. In many cells the production of sufficient RBI protects them against tumorigenesis, but occasionally a cell produces insufficient RBI, thereby escaping its normal growth control. Since promoter mutations were not found in a large number of other families exhibiting a similar phenotype, one would assume that alternative possibilities exist. Subtle changes in the RBI amino acid sequence, for example, may reduce only its functional efficiency, and only when a threshold level of activity is not maintained do tumors develop. The mutation in family RBF29 is possibly one such example with only a single amino acid change. Why, then, did the twins in this family develop multifocal disease? In 70% of tumors the initial mutation is duplicated (32, 33) in tumor precursor cells. Duplication of a 'weak' mutation might still result only in a mild phenotype but, as in 30% of tumors, the second mutation is not serious the combination could result in multifocal tumor formation. Such independent somatic mutations have been identified in tumors from bilaterally affected RB patients (19). The tumors from RBF29 family members, however, were successfully treated and so not available for analysis. Whether subtle changes anywhere in the gene would result in a mild phenotype or whether specific regions, such as exon 20, are more important is not clear. The only other amino acid substitution reported in RB tumors was in exon 18, and that patient was bilaterally affected (12).

It is interesting that the other family in our study showing a low-penetration phenotype also has a mutation in exon 20. At first sight it appears that the mutation in RBF18 resulted in the generation of a stop codon, which could not be described as a mild mutation since the protein would be missing 254 amino acids at the C-terminal end. However, the mutation also alters the DNA sequence, potentially generating a cryptic splice acceptor site in that region. Under normal circumstances, in the presence of cis competition with the normal site, cryptic sites never bind to splicing factors; however, a change in the local sequence environment, however, can change the splicing pattern (34). In a cis-competition assay for splice-site selection, Reed and Maniatis (34) showed that sequences located downstream from intron 1 in the human /3-globin gene splice acceptor site, for example, can have a profound effect on the efficiency of the adjacent splice site. Moreover, they found that the interaction between factor(s) present in a splicing extract and the splice sites is affected by exon sequences, which may play a key role in distinguishing between normal splice sites and cryptic splice sites located throughout pre-mRNAs. A mutation in the exon sequence, therefore, might improve the chances of recognition and/or the affinities of splicing components for the cryptic site, thus giving rise to a stronger and more stable splice complex. It is possible that under some circumstances, or in a specific cell type, the cryptic site is used either exclusively or in combination with the real site. However, many factors affect splicing and, without a functional assay, it is difficult to predict the outcome. For example, a G → T transversion in exon 22 of the RBI gene in the small-cell lung cancer cell line NCI-H69C simultaneously created a stop codon and a novel splice donor site (20). However, the mutation must have also influenced the normal splice acceptor site immediately upstream of exon 22, as it resulted in the removal of the entire exon.

It is possible, therefore, that the RBF18 mutation could have similar consequences. We do not have access to RNA from this family to investigate this possibility. It is also difficult to assess functional properties of protein(s) thus produced. There have been reports of shorter RB proteins, resulting from in-frame deletions of exons 20–22, with impaired biochemical properties (27, 28, 30, 35, 36). Sheffiner et al. (37) reported that a 4-amino acid deletion resulting from a splice acceptor-site mutation in exon 20 in a cervical carcinoma cell line produced defective RBI. It is not clear, however, what the functional consequences of such protein(s) would be in a developing retinal cell.

The RBI gene seems to have a variety of functions, depending on the stage of development and cell type. The effect of naturally occurring RBI mutations on this expanding
reertoire of RBL activities has not been investigated, but it is conceivable that particular mutations—for example, amino acid substitutions in particular regions of the gene—rather than abolishing the function of RBL, modify it so that it works less efficiently. One consequence of this modification might be that only occasionally is there insufficient RBL to prevent tumorigenesis. Alternatively, once initiated the transformed phenotype might be overcome by subsequent adequate production of RBL—for example, through alternative splicing where stop codons are involved—resulting in regressed/benign tumors. To clarify this situation, more families of the phenotype might be overcome by subsequent adequate production of RBL—for example, through alternative splicing where stop codons are involved—resulting in regressed/benign tumors. To clarify this situation, more families of the low-penetrance phenotype need to be analyzed for their mutations.

We thank Sheila Giles for typing the manuscript and preparing the graphics. We thank Dr. I. Goldsmith and his group at the Imperial Cancer Research Fund for preparing the oligonucleotides. We are especially grateful to Dr. Tom Maniatis for helpful discussions in our analysis of the cryptic splice site. Z.O. and A.H. are supported by a grant from the David Allen Retinoblastoma Appeal.