A Molecular Analysis of Genes Involved in the Cell Cycle
in southern African Blacks with Hepatocellular Carcinoma.

Carla Suzana Pinto Martins

A dissertation submitted to the Faculty of Medicine, University of the
Witwatersrand, Johannesburg, for the degree of Master of Science in Medicine.

1999
ABSTRACT

Hepatocellular carcinoma (HCC) is a leading cause of death in both Africa and Asia. It is multifactorial in aetiology and complex in its pathogenesis. Genes that might affect tumour progression, invasion, and metastasis are good candidates to investigate in attempting to understand the transformation process. The \( p53, \ RBI, \ BRCA1, \ BRCA2, \ WT1 \) and \( E\-cadherin \) genes were analysed for allelic imbalance/loss of heterozygosity (LOH), polymorphisms, and mutations. Tumour and non-tumorous liver tissue from 25 southern African blacks were examined, using polymerase chain reaction restriction fragment length polymorphisms (PCR-RFLP) and PCR-single stranded conformational polymorphisms (SSCP), sequencing, and Southern blotting techniques. Allele frequencies for polymorphisms at the \( WT1, \ D13S137, \ D13S120, \ D13S127, \ D17S855, \ D16S301, \) and \( D16S260 \) loci were determined in 20 random African blacks using microsatellite analysis to determine allele frequencies, polymorphism information content (PIC) and diversity (H) values. To our knowledge this has not been done previously for these loci in this population. The chromosomal region 11p13, containing the \( WT1 \) gene, and the gene itself has been reported to be deleted in 4.5\% of HCCs. LOH was detected at the \( WT1 \) locus for 1/13 HCCs (8\%) in this study. The \( RBI \) gene has been described to be mutated in 32.4\% (China), 33.3\% (Korea), 29\% and 50\% (Japan), and 27\% (Australia), of advanced stage HCCs. In our study LOH at this locus was found in 3/19 HCCs (16\%). Our finding of LOH at the \( BRCA2 \) locus in 2/20 HCCs (10\%) supports the previously proposed notion that \( BRCA2 \) may function as a tumour suppressor gene in a hormone-related pathway in the liver, and that it may in some way be involved in HCC. No conclusive findings were made
for any of the other loci. Microsatellite instability was detected in 3/22 (14%) individuals. We propose that microsatellite/genomic instability may play a role in a subset of HCCs only. Of this population, 27% had the specific \( p53 \) codon 249 AGG-AGT mutation in some tumour and non-tumorous liver. This was expected as the great majority of the individuals were from Mozambique, a country where heavy aflatoxin exposure is prevalent. All the loci examined in the allele frequency studies proved to be highly informative, showing high PIC and ED values, and should therefore be useful in population studies.
DECLARATION

I declare that this is my own unaided work. It is being submitted for the degree of Master of Science in Medicine at the Faculty of Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted for any degree at another university.

Carla S. P. Martins

30th day of January 1999
To all those I hold dear, who, because of their love and support, have made this possible.
ACKNOWLEDGEMENTS

I would like to thank:

My parents, my husband, and friends, for their support, encouragement, and interminable patience;

Professor M.C. Kew, and the staff of the Molecular Hepatology Research Unit in the Faculty of Medicine, University of the Witwatersrand, for all their help and advice; and my fellow students for their support.

A special thanks goes to: Dr S. Tavtigian (Myriad Genetics Tnc., Salt Lake City, Utah, USA) who generously provided me with two partial \textit{BRCA2} cDNA clones; Dr A. Lane (SAIMR) for his counsel, and the use of his statistical program; Dr A. Kramvis for her support and guidance.

The Photographic and Illustrations department of the University of the Witwatersrand for the computerised photographs displayed in this work.

Dr M.A. Kedda for her encouragement, input and guidance.
TABLE OF CONTENTS

Abstract i
Declaration iii
Dedication iv
Acknowledgements v
Table of contents vi
List of abbreviations xiv
List of tables xviii
List of figures xix

CHAPTER 1 INTRODUCTION 1

1.1) Rationale of study 2

1.2) HCC 3

1.2.1) The role of HBV in HCC 4

1.2.2) HBV markers 5

1.2.3) Tumour suppressor genes and their role in HCC 6

1.3) The Cell Cycle 7

1.3.1) Cell cycle regulation 9

1.4) Genes Investigated 11

1.4.1) The p53 Gene 11

1.4.1.1) Structure and Function of the p53 Gene and Protein 11

1.4.1.2) p53 Gene Mutations 15
1.4.1.3) p53 and HBx

1.4.2) The Retinoblastoma (RB1) gene
1.4.2.1) The RB1 Gene and Protein
1.4.2.2) pRB1 Function
1.4.2.3) RB1 Gene Mutations
1.4.2.4) RB1 and HCC

1.4.3) The Breast Cancer Susceptibility gene 1 (BRCA1)
1.4.3.1) BRCA1 Gene and Protein Structure
1.4.3.2) BRCA1 Gene Function
1.4.3.3) BRCA1 Mutations

1.4.4) The Breast Cancer Susceptibility Gene 2 (BRCA2)
1.4.4.1) The BRCA2 Gene and the BRCA2 Protein Structure
1.4.4.2) BRCA2 Gene Function
1.4.4.3) BRCA2 Gene Mutations
1.4.4.4) BRCA2 and HCC

1.4.5) The Wilms' tumour (WT1) gene
1.4.5.1) WT1 Gene, Protein Structure and Function
1.4.5.2) Mutations of the WT1 gene

1.4.6) The E-cadherin (Uvomorulin) gene
1.4.6.1) E-cadherin Protein Structure and Function
1.4.6.2) Mutations of the E-cadherin gene

1.5) The Polymorphism Concept
1.5.1) Restriction Fragment Length Polymorphisms (RFLPs)
1.5.2) Variable Number Tandem Repeats (VNTRs) 38

1.6) Allele Frequencies and their Uses 39
  1.6.1) Population Genetics 39
  1.6.2) Genotypic/Chromosomal Markers 40
  1.6.3) Mapping 40

1.6.4) Statistical Studies Involving Allele Frequency 41
  1.6.4.1) Population Size and the Standard Error 41
  1.6.4.2) The Hardy-Weinberg Model (H-W) 41
  1.6.4.3) Polymorphism Information Content (PIC), Diversity (H)
          and Observed Heterozygosity (OH), 42

1.7) Purpose and Importance of the Study 43

CHAPTER 2: MATERIALS AND METHODS 45

2.1) Subjects 45

2.2) DNA extraction 45

2.3) Techniques Used in Gene Analysis 46

  2.3.1) Loss of Heterozygosity and Allele Frequency Studies 47
     2.3.1.1) PCR 47
     2.3.1.2) Gel Electrophoresis of PCR Products 47
       2.3.1.2.1) Polyacrylamide Gels 48
       2.3.1.2.2) Agarose Gels 49
     2.3.1.3) Sequencing - Standard Protocol 49
       2.3.1.3.1) Sequencing Gels 52
2.3.1.4) Repeat Sequence Analysis and Allele Frequency Analysis

2.3.1.4.1) \textit{p53} Gene 52

2.3.1.4.2) \textit{BRCA2} Gene 53

2.3.1.4.3) \textit{BRCA1} Gene 53

2.3.1.4.4) \textit{WTI} Gene 53

2.3.1.4.5) \textit{RB1} Gene 54

2.3.1.4.6) \textit{E-cadherin} Gene 54

2.3.2) SSCP Analysis of the \textit{BRCA2} Gene 55

2.3.2.1) Primer design for the \textit{BRCA2} Gene 55

2.3.2.2) Gel Electrophoresis of PCR Products 56

2.3.3) LOH Analysis of the \textit{BRCA2} Gene 57

2.3.3.1) Restriction Analysis 58

2.3.3.2) Agarose Gel Electrophoresis 59

2.3.3.3) Southern Blotting 59

2.3.3.4) Probe Preparation 59

2.3.3.4.1) Transformation 60

2.3.3.4.1.1) DNA Plasmid Preparation 61

2.3.3.4.1.2) Plasmid Digests 61

2.3.3.4.1.2.1) Restriction Endonuclease Digestion of the pUC18 Plasmid Containing the "AC" Insert 62
2.3.3.4.1.2) Restriction Endonuclease Digestion of
the pUC18 Plasmid Containing the “DG” Insert

2.3.3.4.1.3) Agarose Gel Electrophoresis of Digested Plasmid DNA

2.3.3.4.1.4) DNA Extraction from Agarose Gel

2.3.3.4.1.5) Labelling of Probes

2.3.3.5) Hybridisation

2.3.3.6) Stripping the Membrane

2.3.3.7) Autoradiography

2.3.4) p53 Mutation Analysis

2.3.4.1) PCR Conditions

2.3.4.1.1) Agarose Gel Electrophoresis of PCR Product

2.3.4.2) Restriction Enzyme Digest of PCR Product

2.3.4.2.1) Agarose Gel Electrophoresis of Digested PCR Product

2.3.4.3) Confirmation by sequencing

2.3.5 Statistical Tests for the Allele Studies on the Random Black Population

2.3.5.1 Hardy-Weinberg Equilibrium (HWE)

2.3.5.2 Polymorphism Information Content (PIC), Diversity (H) Observed Heterozygosity (OH) Values

CHAPTER 3 RESULTS

3.1 Gene Studies in HCC
3.1.1 VNTR Analysis

3.1.1.1 WT1 Gene

3.1.1.2 BRCA1 Gene/Chromosome 17

3.1.1.2.1 D17S855

3.1.1.2.2 D17S846

3.1.1.3 BRCA2 Gene/Chromosome 13

3.1.1.3.1 D13S120

3.1.1.4 BRCA2/RB1 Genes/Chromosome 13

3.1.1.4.1 D13S127

3.1.1.5 RB1 Gene/Chromosome 13

3.1.1.5.1 D13S137

3.1.1.5.2 RB(1.20)

3.1.1.6 p53 Gene

3.1.1.7 E-cadherin Gene/Chromosome 16

3.1.1.7.1 D16S301

3.1.1.7.2) D16S260

3.1.1.8) Blurring of Bands on Polyacrylamide Gels for the D17S846, RB(1.20) and p53 Loci

3.1.1.9) Genomic Instability

3.1.2) SSCP Analysis of the BRCA2 Gene

3.1.2.1) Southern Blotting

3.1.3) p53 Gene Codon 249 Analysis

3.1.4) HBV Status

3.1.5) Gene Studies in Random Black Subjects/Polymorphism Frequencies
3.1.5.1) WT1 Internal AC Repeat
3.1.5.2) D13S137
3.1.5.3) D13S120
3.1.5.4) D13S127
3.1.5.5) D17S855
3.1.5.6) D16S301
3.1.5.7) D16S260
3.1.5.8) Statistical Test for Hardy-Weinberg Equilibrium
3.1.5.9) Polymorphism Information Content (PIC), and Diversity (H) Values for All Loci

CHAPTER 4 DISCUSSION

4.1) Molecular aetiology of HCC
4.2) Gene Studies in HCC
   4.2.1) Genomic Instability/LOli
   4.2.2) SSCP Analysis of the BRCA2 Gene
   4.2.3) p53 Codon 249 AGG→AGT Mutation Studies
      4.2.3.1) Gel Artefact
   4.2.4) Techniques
      4.2.4.1) PCR-RFLP Studies
      4.2.4.2) PCR-SSCP Studies
      4.2.4.3) VNTR Studies
4.3) Gene Studies in Random Black Subjects/Polymorphism Frequencies
   4.3.1) Allele Frequencies at the Different Loci
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFB1</td>
<td>Aflatoxin B1</td>
</tr>
<tr>
<td>Anti-HBc</td>
<td>Antibody to the hepatitis B virus core antigen</td>
</tr>
<tr>
<td>Anti-HBs</td>
<td>Antibody to the hepatitis B virus surface antigen</td>
</tr>
<tr>
<td>APS</td>
<td>Ammonium persulphate</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair/s</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>CDK</td>
<td>Cyclin dependent kinase</td>
</tr>
<tr>
<td>CDKI</td>
<td>Cyclin dependent kinase inhibitor</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribo-nucleic acid</td>
</tr>
<tr>
<td>Ci</td>
<td>Curie</td>
</tr>
<tr>
<td>cm</td>
<td>Centimetre</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyribo-adenine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxyribo-cytosine triphosphate</td>
</tr>
<tr>
<td>dH₂O</td>
<td>Distilled water</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Deionised distilled water</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyribo-guanine triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxyribo-thymine triphosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>ddNTPs</td>
<td>2',3'-dideoxynucleoside 5'-triphosphates</td>
</tr>
<tr>
<td>DTE</td>
<td>Dithioerythritol/DTT dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediamine Tetra-Acetic Acid</td>
</tr>
<tr>
<td>E2F</td>
<td>Family of endogenous regulators</td>
</tr>
<tr>
<td>g</td>
<td>Gram/s</td>
</tr>
<tr>
<td>H</td>
<td>Diversity</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HBcAg</td>
<td>Hepatitis B virus core antigen</td>
</tr>
<tr>
<td>HBeAg</td>
<td>Hepatitis B virus e antigen</td>
</tr>
<tr>
<td>HBsAg</td>
<td>Hepatitis B virus surface antigen</td>
</tr>
<tr>
<td>HBV</td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>HBx</td>
<td>Hepatitis B virus x protein</td>
</tr>
<tr>
<td>HCC</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl B-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase/s</td>
</tr>
<tr>
<td>kD</td>
<td>KiloDalton/s</td>
</tr>
<tr>
<td>kV</td>
<td>Kilovolts</td>
</tr>
<tr>
<td>l</td>
<td>Litre/s</td>
</tr>
<tr>
<td>LOH</td>
<td>Loss of Heterozygosity</td>
</tr>
<tr>
<td>M</td>
<td>Molar/Molarity</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram/s</td>
</tr>
<tr>
<td>ml</td>
<td>Millilitre/s</td>
</tr>
<tr>
<td>mm</td>
<td>Millimetre/s</td>
</tr>
<tr>
<td>mM</td>
<td>MilliMolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger ribonucleic acid</td>
</tr>
<tr>
<td>ng</td>
<td>Nanogram/s</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>OH</td>
<td>Observed heterozygosity</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pH</td>
<td>Measure of acidity or alkalinity</td>
</tr>
<tr>
<td>PIC</td>
<td>Polymorphism Information Content</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>pmol</td>
<td>Picomol/s</td>
</tr>
<tr>
<td>pRB</td>
<td>Retinoblastoma protein</td>
</tr>
<tr>
<td><strong>RB/RB1</strong></td>
<td>Retinoblastoma gene</td>
</tr>
<tr>
<td>RFLPs</td>
<td>Restriction fragment length polymorphisms</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SSC</td>
<td>Sodium chloride/sodium citrate</td>
</tr>
<tr>
<td>SSCP</td>
<td>Single stranded conformational polymorphisms</td>
</tr>
<tr>
<td>Std E</td>
<td>Standard error</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-Boric acid-EDTA</td>
</tr>
<tr>
<td>TE</td>
<td>Tris EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N,N-tetramethylethylenediamine</td>
</tr>
<tr>
<td>TGFβ1</td>
<td>Transforming growth factor beta one</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris-(hydroxy)methylamino triethylamine</td>
</tr>
<tr>
<td>U</td>
<td>Units</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>VNTRs</td>
<td>Variable number of tandem repeats</td>
</tr>
<tr>
<td>W</td>
<td>Watt/s</td>
</tr>
<tr>
<td><strong>WT1 gene</strong></td>
<td>Wilm's tumour gene</td>
</tr>
<tr>
<td>Xgal</td>
<td>X galactosidase</td>
</tr>
<tr>
<td>xg</td>
<td>Centrifugal force (gravity)</td>
</tr>
<tr>
<td>μg</td>
<td>Microgram/s</td>
</tr>
<tr>
<td>μl</td>
<td>Microlitre/s</td>
</tr>
<tr>
<td>μmol</td>
<td>Micromolar</td>
</tr>
<tr>
<td>%</td>
<td>Per cent</td>
</tr>
</tbody>
</table>
\( \alpha^{32} \text{P} \text{dCTP} \quad \) Deoxycytidine-5'-[\( \alpha^{32} \text{P} \)]triphosphate, triethyleneammonium salt

\( \Omega \quad \) Ohms

\( \xi \quad \) Extinction coefficient (number of dNTP residues in oligonucleotide constant)

\( X \quad \) Chi
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>PCR primers</td>
<td>50</td>
</tr>
<tr>
<td>2.2</td>
<td>Restriction enzymes used for the digestion of DNA prior to Southern blotting</td>
<td>58</td>
</tr>
<tr>
<td>2.3</td>
<td>Details of the “AC” and “DG” probes</td>
<td>60</td>
</tr>
<tr>
<td>3.1</td>
<td>LOH, SSCP and sequence analysis</td>
<td>105</td>
</tr>
<tr>
<td>3.2</td>
<td>HBV status</td>
<td>108</td>
</tr>
<tr>
<td>3.3</td>
<td>p53 codon 249 analysis</td>
<td>109</td>
</tr>
<tr>
<td>3.4</td>
<td>VNTR analysis for the <em>WT1</em> gene and D13S137, D13S120, D13S127, D17S855, D16S301, D16S260 loci in HCC subjects</td>
<td>113</td>
</tr>
<tr>
<td>3.5</td>
<td>VNTR analysis for the <em>p53</em> gene and the D17S846 and RB(1.20) loci</td>
<td>115</td>
</tr>
<tr>
<td>3.6</td>
<td>VNTR analysis for the <em>WT1</em> gene and D13S137, D13S120, D13S127, D17S855, D16S301, D16S260 loci in random black subjects</td>
<td>117</td>
</tr>
<tr>
<td>3.7</td>
<td>Allele frequencies for the loci D13S137, D13S120, D13S127, D17S855, D16S301 and D16S260 for the random black subjects (RB)</td>
<td>119</td>
</tr>
<tr>
<td>3.8</td>
<td>Observed heterozygosity, Polymorphism information content, Diversity and Hardy-Weinberg Equilibrium for random black subjects</td>
<td>121</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Diagrammatic representation of the cell cycle 9
2. Sequencing of the antisense strand of the AC microsatellite repeat at the D16S301 locus 72
3. Allelic imbalance at the D17S846 locus 75
4. LOH at the D13S127 locus 77
5. LOH/allelic imbalance at the $RB1 [RB(1.20)]$ locus 80
6. Allelic imbalance at the $p53$ locus 81
7. SSCP analysis at position 97 of the $BRCA2$ gene 85
8. SSCP analysis at position 508 of the $BRCA2$ gene 86
9. Restriction enzyme digests of the “AC” and “DG” Probes used for Southern blotting 87
10. Random black DNA samples restricted with four different restriction endonucleases for Southern blotting 88
11. $p53$ codon 249 mutation detected by $HaeIII$ restriction enzyme digest of PCR product 90
12. Confirmation of the presence of the $p53$ codon 249 mutation by sequencing of the anti-sense strand 91
13. Confirmation of the presence of the $p53$ codon 249 mutation by sequencing of the sense strand 92
14. Allele frequency distribution for the $WT1 (ALU)$ internal repeat 98
15. Allele frequency distribution for the D13S137 locus 99
16. Allele frequency distribution for the D13S120 locus 100
17. Allele frequency distribution for the D13S127 locus 101
Allele frequency distribution for the D17S855 locus

Allele frequency distribution for the D16S301 locus

Allele frequency distribution for the D16S260 locus

Diagram depicting mini-hairpin formation, which occurs during electrophoresis, in the region of the p53 gene codon
Allele frequency distribution for the D17S855 locus 102

Allele frequency distribution for the D16S301 locus 103

Allele frequency distribution for the D16S260 locus 104

Diagram depicting mini-hairpin formation, which occurs during electrophoresis, in the region of the p53 gene codon 133
CHAPTER 1
INTRODUCTION

For the normal development of an organism, proliferation, differentiation, and apoptosis of cells must be tightly regulated. Should any of these processes become deregulated, normal development will be compromised (Menke et al., 1998). The evolution of cancer is believed to occur from the stepwise accumulation of a number of genetic aberrations in a cell. Some of the changes in cancer cells contribute to the overall phenotype, whereas others do not. Those changes that cause phenotypic differences will be those that result in loss of growth control, invasion of neighbouring tissue, generation of blood vessels (angiogenesis), or metastasis to distant organs (reviewed by Nicolson 1982). These include loss of function of tumour suppressor genes, activation of proto-oncogenes, faulty DNA mismatch repair, and the integration of viral DNA. (Yee et al., 1994; Fujimoto and Kohgo, 1998).

Hepatocellular carcinoma (HCC) is one of the most common cancers in Africa and Asia and is often associated with chronic hepatitis B or C viral infection and frequent allelic loss of one or more cellular tumour suppressor genes. This functional loss may be secondary to HBV integration events or spontaneous mutations (Wang and Rogler 1988; Slagle et al., 1991; Fujimoto and Kohgo, 1998).
For the normal development of an organism, proliferation, differentiation, and apoptosis of cells must be tightly regulated. Should any of these processes become deregulated, normal development will be compromised (Menke et al., 1998). The evolution of cancer is believed to occur from the stepwise accumulation of a number of genetic aberrations in a cell. Some of the changes in cancer cells contribute to the overall phenotype, whereas others do not. Those changes that cause phenotypic differences will be those that result in loss of growth control, invasion of neighbouring tissue, generation of blood vessels (angiogenesis), or metastasis to distant organs (reviewed by Nicolson 1982). These include loss of function of tumour suppressor genes, activation of proto-oncogenes, faulty DNA mismatch repair, and the integration of viral DNA (Yee et al., 1994; Fujimoto and Kohgo, 1998).

Hepatocellular carcinoma (HCC) is one of the most common cancers in Africa and Asia and is often associated with chronic hepatitis B or C viral infection and frequent allelic loss of one or more cellular tumour suppressor genes. This functional loss may be secondary to HBV integration events or spontaneous mutations (Wang and Rogler 1988; Slagle et al., 1991; Fujimoto and Kohgo, 1998).
1.1) Rationale for this study

Genes, such as tumour suppressors that might affect tumour initiation, progression, invasion, and metastasis, are good candidates to investigate in attempting to understand the transformation process (Slagle et al., 1993). The \textit{p53} and \textit{retinoblastoma (RB1)} genes are known to function in cell cycle regulation (Kim \textit{et al.}, 1994; Schaff \textit{et al.}, 1995). Coincident mutation of the \textit{p53} and \textit{RB1} genes have been observed in 25\% (Murakami \textit{et al.}, 1991), and 12.9\% (Yumoto \textit{et al.}, 1995) of advanced HCCs. Aberrations have been found in \textit{RB1} gene in advanced HCCs (Nishida \textit{et al.}, 1992, 1993, 1994), and abnormalities of the \textit{p53} gene, such as gene mutation, deletion or the nuclear accumulation of mutant p53 protein have also been found to correlate with increased allelic loss at three other loci: breast cancer susceptibility gene 1 (\textit{BRCA1}), D17S588, and D13S267 (Tseng \textit{et al.}, 1997). These observations suggest that \textit{p53} and \textit{RB1} could contribute to carcinogenesis.

The \textit{BRCA1} gene is thought to encode a transcription factor (Smith \textit{et al.}, 1992); and the breast cancer susceptibility gene 2 (\textit{BRCA2}) gene is possibly involved in cellular proliferation and differentiation (Fischer \textit{et al.}, 1996). It has been proposed that the \textit{BRCA1} and \textit{BRCA2} genes may have overlapping functions or act in concert (Rajan \textit{et al.}, 1997). \textit{BRCA2} has been proposed to be involved in HCC development (Katagiri \textit{et al.}, 1996), and loss of heterozygosity (LOH) of the \textit{BRCA1} has recently been reported in HCCs (Piao \textit{et al.}, 1997).

The Wilm's tumour (\textit{WT1}) gene may be implicated in proliferation, differentiation and
apoptosis (Menke et al., 1998), while the E-cadherin protein is expressed by nearly all epithelial cells, and is thought to be the main protein involved in establishing intercellular adhesion (Shimoyama and Hirohasi 1991b). LOH on 11p13 and 16q22, in the region of the WT1 and E-cadherin genes, respectively, has been reported in some HCCs (Wang and Rogler 1988). Moreover, microsatellite instability has been documented in 40% of HCCs from the USA and Korea by Kazachkov et al. (1998) and in 5% of HCCs from China by Sheu et al. (1998).

Thus the objectives of this study were to characterise LOH, polymorphisms and mutations in the p53, RB1, BRC41, BRCA2, WT1 and E-cadherin genes in tumour and non-tumourous liver tissue from southern African blacks with HCC. Microsatellite instability was also analysed in a random sample of southern African blacks.

1.2) HCC

Hepatocellular carcinoma (HCC) is a leading cause of death in both Africa and Asia (reviewed by Tabor 1995), causing at least 310 000 deaths annually worldwide (Parkin et al., 1984). It is multifactorial in aetiology and its pathogenesis is complex. HCC may be monoclonal or polyclonal in origin (Rabes et al., 1982; Williams et al., 1983) and the multistep process of carcinogenesis is characterised by an increasing incidence of HCC with increasing age, following after a long period of chronic liver disease and frequently associated with cirrhosis.
It is not known whether the genetic events which lead to the development of HCC in different populations are the same, similar or distinct (Unsal et al., 1994). The known risk factors of HCC are chronic hepatitis B and C virus infections, non-viral chronic liver disease, and exposure to dietary aflatoxins (Slagle 1995). Several risk factors may jointly contribute to the development of HCC (Yumoto et al., 1995).

1.2.1) The Role of HBV in HCC

Chronic infection with HBV is a major risk factor in the development of HCC (Farshid et al., 1994), and this risk is generally significant only if HBV infection occurs in childhood (reviewed by Tabor 1995).

HBV is a member of the Hepadnaviridae family of viruses (Howard and Melnick, 1991), targeting the liver and in some cases also found in mononuclear blood cells. HBV causes acute and chronic liver disease (Kuhns 1995).

The role of HBV in the development of HCC is complex. Several consequences of HBV infection may contribute to initiation and progression of HCC. The integration of the HBV DNA into host chromosome/s may interrupt growth-regulatory genes or result in the loss of tumour suppressor genes. This integration is not part of the normal replication cycle of the virus, but may occur during chronic HBV infection (reviewed by Butel et al., 1996). Moreover, no unique integration site of HBV within the host genome has been described (Zhou et al., 1987; Harrison et al., 1990). Analyses of integrated HBV DNA cloned from HCCs have shown that conspicuous
chromosomal abnormalities (including deletions, duplications and translocations) frequently occur at the site of viral integration (Hino et al., 1986). The HBV x protein has transactivation activity which can affect host-cell genes, resulting in their inappropriate expression, and possibly, in uncontrolled cell growth. Virus infection and carcinogen exposure may have a synergistic effect in cancer development, and the host's immune response to the chronic viral infection may result in oxidative DNA damage, ultimately contributing to malignant transformation.

1.2.2) HBV Markers

When an individual is infected with HBV, a series of antigens is released into the bloodstream followed later by the production of specific antibody corresponding to each antigen. The antigens appear in the following order: HBsAg (surface antigen) together with HBeAg, approximately two weeks before the onset of symptoms or biochemical evidence of the disease. HBsAg and HBeAg (core antigen - HBcAg) are contained in the infectious viral particle, and HBsAg, in addition, in the small spherical particles and tubular forms. HBeAg which is formed by the post-transcriptional modification of the precore/core fusion protein, is not part of the virus, and is secreted into the serum. When HBeAg is detected in the serum, it indicates that mature virus particles are being produced. The next phase begins when anti-HBc is produced and the first clinical symptoms and biochemical indicators are observed. HBcAg is confined to liver cells. Both IgG and IgM forms of this antibody appear at this time. HBsAg and HBeAg also persist during this period. Once symptoms disappear, HBeAg is cleared from the serum, followed by HBsAg. Phase three is
characterised by termination of HBV infection. HBsAg levels are drastically decreased and anti-HBe appears. Once HBsAg disappears completely it is normally followed a few weeks later by the appearance of anti-HBs, which persists and confers lifelong immunity to the individual (reviewed by Schoub 1994). The anti-HBc IgMs are present for approximately six to nine months and can be used as a good diagnostic marker for acute or recent viral infection (Kuhns et al., 1995). The anti-HBc IgGs are present lifelong, but are not associated with immunity.

When the HBsAg antigen persists for longer than six months and clinical and biochemical markers of liver dysfunction are present, the patient is suffering from chronic hepatitis and is designated a carrier. The incidence of HCC is higher in carriers than in non-carriers (Beasly et al., 1981).

1.2.3) Tumour Suppressor Genes and Their Role in HCC

According to the well-established Knudson model, inactivation of tumour suppressor genes is recessive, requiring mutation or loss of both alleles (Knudson 1971). This may result from two independent events in the case of somatic tumours, and transmission of an altered gene in the germline followed by somatic mutation of the second gene in familial carriers (Haber et al., 1990). The Knudson hypothesis does not preclude the involvement of more than one tumour suppressor gene in the development of a cancer, and although it predicts two rate-limiting steps it cannot exclude additional genetic events (Haber et al., 1990).
A number of tumour suppressor genes is known to be mutated or deleted in human carcinomas. Among them are the \textit{RBI} (retinoblastoma) gene (Harbour \textit{et al.}, 1988; Lee \textit{et al.}, 1988; Bookstein \textit{et al.}, 1990); the \textit{WT1} gene (Gessler \textit{et al.}, 1990) and the \textit{p53} gene (Takahashi \textit{et al.}, 1989; Nigro \textit{et al.}, 1989; Tsai \textit{et al.}, 1990). Other tumour suppressor genes that may play a role during the development of HCC are believed to be located on chromosomes 1p, 4q, 5q, 8p, 11p, 13q, where LOH is frequently observed (Nose \textit{et al.}, 1993; Yeh \textit{et al.}, 1994, 1996; Kuroki \textit{et al.}, 1995; Yumoto \textit{et al.}, 1995; Nagai \textit{et al.}, 1997). Piao \textit{et al.} (1997), have reported LOH of more than one (2-4) tumour suppressor genes in the same HCC. In their study the genes most frequently lost were \textit{p53, RBI, APC}. Furthermore LOH of \textit{p53} and multiple LOH of tumour suppressor genes was most frequently found in poorly differentiated HCCs (Piao \textit{et al.}, 1997). Cumulative LOH presumably reflects the multi-step genetic mechanism of progression of HCC (Tamura \textit{et al.}, 1997).

Inactivation of tumour suppressor genes may result from HBV DNA integration (Slagle \textit{et al.}, 1991), aflatoxin-mediated mutations (Ozturk \textit{et al.}, 1991), or other mutations (Slagle \textit{et al.}, 1993).

1.3) \textbf{The Cell Cycle}

Growth and development require cell multiplication, enlargement, and differentiation. The cell cycle is a term used to describe the systematic sequence of events that guarantee the faithful duplication of all the cellular components in their appropriate order and the
distribution of these components into two daughter cells (Levine et al., 1994).

Many cells undergo a continuous alternation of division and nondivision. The division of nucleate cells requires two distinct integrated activities: karyokinesis (nuclear division) and cytokinesis (cytoplasmic division). In man cytokinesis is initiated after nuclear division is almost complete. Two types of nuclear division are known to occur: mitosis (in somatic cells), and meiosis (for the creation of reproductive cells) (Klug and Cummings 1986).

The cell cycle can be divided into two main stages, interphase and mitosis/meiosis. Interphase can be subdivided into four sections (Fig 1), G0, G1, G2 and S phase. The G1 and G2 phases of the cell cycle are the growth phases, characterised by intense metabolic activity, cell growth and differentiation. By the end of G2 the cell is approximately double the size it was before it entered the cell cycle. The S phase is when the synthesis of DNA and histones takes place. Some cell growth and differentiation also occur. Cells can withdraw from the cell cycle with an unduplicated DNA content and enter a quiescent state where macromolecular synthesis is drastically decreased, this is termed G0 (Klug and Cummings 1986).
1.3.1) Cell Cycle Regulation

The cell cycle is tightly regulated, forty genes or more being required for cell division. In mammalian cells extracellular growth factors, mitogen antagonists, differentiation inducers, and spatial cues are all taken into account before a cell enters the S phase (reviewed by Goodrich and Lee 1993). Cells are sensitive to these stimuli until the restriction point in late G1. Once S phase has commenced, the cell cycle will progress to completion. Thus, if the cell cycle is to be stopped, it has to be before S phase. This G1/S 'checkpoint' is the most crucial cell cycle control signal for the initiation (or not) of DNA synthesis (Hartwell and Weinert 1989).

The key regulators of G1 and G2 in mammals are the cyclins, the cyclin dependent kinases, and various genes (including cell division control [CDC] genes) whose
1.3.1) Cell Cycle Regulation

The cell cycle is tightly regulated, forty genes or more being required for cell division. In mammalian cells extracellular growth factors, mitogen antagonists, differentiation inducers, and spatial cues are all taken into account before a cell enters the S phase (reviewed by Goodrich and Lee 1993). Cells are sensitive to these stimuli until the restriction point in late G1. Once S phase has commenced, the cell cycle will progress to completion. Thus, if the cell cycle is to be stopped, it has to be before S phase. This G1/S 'checkpoint' is the most crucial cell cycle control signal for the initiation (or not) of DNA synthesis (Hartwell and Weinert 1989).

The key regulators of G1 and G2 in mammals are the cyclins, the cyclin dependent kinases, and various genes (including cell division control [CDC] genes) whose
products function in mediating a 'go-no-go' response. Cyclins derive their name from the fact that their concentrations rise and fall in a regular pattern during the cell cycle. Cyclins are a group of proteins which can be modified at cell cycle checkpoints in a manner which induces them to halt progression of the cell cycle. They activate the catalytic subunits of protein kinases by binding to them to form a complex termed a cyclin-dependent kinase (cdk) (Levine et al., 1994). Cdk5 can phosphorylate those proteins that drive the cell cycle through G1 to S to G2 and M phases (Hartwell and Weinert, 1989). Checkpoint pathways are influenced by such factors as availability of nutrients and growth factors, and can then be stimulated to regulate levels or ability of cyclins to activate cdk5 (Levine et al., 1994).

Tumour-suppressor genes and oncogenes can alter the basic genetic mechanism that controls the cell cycle, as they function, at least in part, by mediating responses to external signals. Oncogene products have been shown to be components of signal transduction pathways. Certain tumour-suppressor genes are known to negatively regulate or stop cell division, and may monitor the completion or efficiency of a process in the cell cycle. The products of these genes, however, are known to be unnecessary for cell division (Levine et al., 1994).
1.4) Genes Investigated

1.4.1) The p53 Gene

Mutations in, or deletions of, the p53 gene are found with relatively high frequency in human HCCs (Bressac et al., 1990; Murakami et al., 1991; Scorsonone et al., 1992; Lin et al., 1996; Tabor 1997) and the functional loss of this tumour suppressor gene, as well as its abnormal expression, has been proposed to have a significant role in HCC development (Bressac et al., 1990; Slagle et al., 1991) or at least in the development of a subset of HCCs (Bressac et al., 1990; Slagle et al., 1991; Bressac et al., 1991; Hsu et al., 1991).

Cancer cells often express a high level of mutant p53 protein (Ponchel et al., 1994). Most mutant p53 proteins have been found to have amino acid substitutions at 'mutational hot spots' and these vary according to tumour site (de Fromentel and Soussi 1992). Mutational hot spots are thought to be the target of endogenous and exogenous mutagens (Ponchel et al., 1994), and certain mutations of the p53 gene are seen in one specific type of cancer only.

1.4.1.1) Structure and Function of the p53 Gene and Protein

The p53 gene is located on 17p13.1 (Benchimol et al., 1985) and contains 11 exons. The p53 transcript is spliced producing a mRNA that is 2.2-2.5kb in size and is present in all the cells of the body. The p53 protein is a 53kD phosphorylated protein, with 393 amino acids, that is normally found in the
nucleus, and is divided into four structural and functional domains (reviewed by Levine 1997; Tabor 1997).

p53 appears to be involved in DNA repair, cell division, apoptosis (Lowe et al., 1993), cell senescence (Carson and Lois 199 and centrosome duplication (Fukasawa et al., 1996) at different stages in the cell cycle.

The p53 protein has been shown to be essential to prevent propagation of DNA damage to daughter cells. It appears to have a role as a G₁ to S checkpoint control in the cell cycle for recognising DNA damage (Kuerbitz et al., 1992), and acts by delaying the progress thorough the cell cycle, thereby allowing time for DNA to be repaired (Kastan et al., 1991). p53-induced G₁ arrest is thought to require p53 activated signal pathways. Here p53 acts as a transcription factor enhancing the transcription rate of genes containing p53 responsive elements. p53 may also negatively regulate or inhibit the transcription of genes not containing p53 responsive elements in their regulatory regions (Zambetti and Levine 1993). Besides its function as a transcription factor, p53 binds directly to several cellular proteins (Ko and Prives 1996) such as helicases which form part of the transcription-repair complex TFIIH. Here it acts as a nuclear excision repair factor-binding protein playing an important role in the maintenance of genomic integrity by both controlling the cell cycle and by participating directly in the nuclear excision repair pathway. Furthermore, the p53 protein has a high binding affinity for single-stranded DNA and RNA, catalysing their double-stranded
annealing (Oberosler et al., 1993). It is thus possible that the p53 protein is involved in inhibiting critical helicase activities required for DNA replication and recombination, thus prohibiting single-strand recombination intermediates that could lead to gene duplications, amplifications, and the activation of oncogenes (Levine et al., 1994 review; Wang et al., 1995). Two other possible cell cycle checkpoints involving p53 are G₀-G₁ and G₂/M. G₂/M acts in preventing premature entry into a second S phase. Cells are maintained in G₀ by Gas1, a membrane protein expressed only at this time, and which seems to carry out its function only if wild-type p53 is present. In this case p53 is functioning not as a transcription factor, but possibly by direct protein-protein signalling (Guillouf et al., 1995).

Should DNA damage be too extensive to repair, the p53 protein may induce apoptosis (Lowe et al., 1993); a type of cell death employed in eliminating excessive or unnecessary cells during organ involution, development of embryonic tissue, and regression of tumours (Wyllie et al., 1980). p53-mediated apoptosis is found to prevail in those circumstances in which a cell has limited survival factors (abnormal environment), significant DNA damage (unstable genome), or where an activated oncogene is forcing the cell into a replicative cycle. This is probably the reason why most cancerous cells select against wild-type p53 function (Levine 1997 review).

p53 is thought to play a role in centrosome replication. The centrosome is a microtubule-organising area in eukaryotic cells. In mitosis it establishes spindle
bipolarity and microtubule assembly, constitutes the cleavage furrow plane, and balances the segregation of the chromosomes (Bornens 1992). The p53 protein associates with the centrosome during interphase, but not during mitosis (Blair et al., 1988) and is thought to exert its checkpoint functions through the regulation of centrosome duplication, thereby regulating the number of centrosomes in a cell (Fukasawa et al., 1996). Duplication of the centrosome occurs only once during the cell cycle, commencing at the G1-S boundary and ending in G2 (Robbins et al., 1989). p53 has also been linked to the cell senescence clock (Carson and Lois 1995), where senescence is defined as an antineoplastic mechanism employed to confine the proliferative potential of cell clones in the body (Weinberg 1995). Senescence is a potent defence against cancer.

How the transcription of p53 itself is regulated is unclear. DNA damage may in some way induce a cytoplasmic response that affects p53 concentration or modifies its transcription functions or binding capacity (Wu and Lozano 1994; Milne et al., 1995). The activation of p53 appears to require one of three events: (1) DNA strand breaks are thought to generate the release of restriction enzyme nucleases which in turn induce the increase of p53 levels and its activity (Levine 1997); (2) Hypoxia stimulates the increase of p53 levels and activates the p53 proteins. This is possibly an alternative pathway through which p53 may prevent cells from cycling to carcinogenesis - when tumours reach a critical size their blood supply becomes limiting. Angiogenic factors must then come into play for the tumour to continue growing and this hypoxia triggers p53 (Levine 1997 review).
(3) Ribonucleoside triphosphate pools are a third signal which activates p53. If such pools fall below a certain threshold, DNA replication and progression through the cell cycle is hampered and p53 is activated (Irie et al., 1996). There may be other signals of cellular distress that might induce a response of p53 - for example recognition of birth defects in response to teratogenic agents (Nicol et al., 1995).

Changes in p53 are not a direct cause of cancer. However, those cells that lack p53 have a faulty DNA damage control response and a prolonged lifespan. These cells are at increased risk for malignant transformation (Carson and Lois 1995 review).

1.4.1.2) p53 Gene Mutations

The majority of p53 alterations reported to date have been loss of one allele accompanied by mutations of the second allele (Baker 1989; Nigro 1989). Point mutations are the most commonly found mutation, comprising 40-100% (Nigro et al., 1989; Oda et al., 1992; Nishida et al., 1993; Li et al., 1993; Shi et al., 1995; Hui et al., 1996) and 50% of mutations in HCCs from China and southern Africa respectively (Bressac et al., 1991). The remainder of p53 gene mutations are mainly deletions and insertions of several base pairs, often resulting in frameshifts, and splice site mutations. The greatest number of missense mutations that occur in the p53 gene occur in the DNA-binding domain (amino acids 102-292). Mutations in these amino acid residues usually result in defective contact with DNA, and in the loss of the p53 protein’s ability to act as a transcription factor. Another type of mutation known to occur in this region is one which leads to the physical
disruption of the tertiary structure of the protein, producing a mutant protein which reacts with PAb240 (a monoclonal antibody) (reviewed by Levine 1997).

Point mutations within exons 5-8 of the p53 gene (Nigro 1989) often lead to the production of a mutant p53 protein with increased stability (Jenkins et al., 1985; Finlay et al., 1988; reviewed by Levine 1997). In primary rat cells in culture, mutant p53 protein has been demonstrated to exhibit a phenotype or activity which seems to contribute to abnormal cell growth in some way, and to have the ability to form tetrameric protein complexes with endogenous wild-type p53, thereby inactivating it (Finlay et al., 1988; Martinez et al., 1991). In such cases, this is considered a dominant loss of function phenotype. In the model proposed for HBV-related HCCs (Scorsone et al., 1992), the p53 gene missense point mutation is believed to be the first event followed at a later stage by the loss of the second allele. The fact that p53 gene inactivation in HCCs appears to require alterations of both alleles, indicates that a recessive role is predominant in HCC progression rather than a dominant negative role. This does not seem to hold true for Japanese HCCs, however. Here loss of the one allele may precede mutation of the other allele (Oda et al., 1992). It has not been conclusively established which of these two models applies for HCCs from southern Africa, and the results from this study may aid in establishing this at a later date.

LOH of the p53 gene has been reported with relatively high frequency (49-69%) in HCCs from Japan (Murakami et al., 1991; Oda et al., 1992; Nishida et al.,
1993; Teramoto et al., 1994). and also from southern Africa (60%), China (60%), and Taiwan (39.3%) (Bressac et al., 1991; Slagle et al., 1991; Sheu 1998). In 44-95% of Japanese HCCs, LOH was accompanied by mutation of the second p53 allele (Oda et al., 1992; Nishida et al., 1993; Teramoto et al., 1994).

The direct role of HBV infection in the mutation of the p53 gene in HCCs is unclear. Li et al. (1993), found no correlation between HBV DNA infection and p53 aberration, and Lin et al., (1996) propose that HBV infection does not have a direct correlation with genetic changes in p53. Teramoto et al. (1994), however, claim that HBV infection is hepatocarcinogenic through frequent induction of p53 abnormalities causing G:C to T:A alterations in their Japanese tumours. This alteration did not occur at codon 249, however, which may be explained by the fact that Japan is a low aflatoxin exposure area, and that both aflatoxin and HBV infection appear to be required for the induction of the codon 249 AGG →AGT transversion mutation (Oda et al., 1992; Kirby et al., 1996). Heavy dietary aflatoxin (AFB1) intake is believed to cause an AGG to AGT mutation at the third base of codon 249 of the p53 gene in about 50% of HCCs (Hsu et al., 1991; Bressac et al., 1991; Li et al., 1993) and for this reason clustering of this point mutation occurs in HCCs from China (Hsu et al., 1991; Ozturk et al., 1991) and Africa (Bressac et al., 1991; Hsu et al., 1991; Ozturk et al., 1991).

1.4.1.3) p53 and HBx

There are some cellular proteins and several viral transforming oncoproteins known
to interact with the \( p53 \) gene product (Feitelson et al., 1993; Wang et al., 1994; Ueda et al., 1995), thereby functionally inactivating or degrading it. HBx has been found to be expressed in HCCs from geographic areas of high cancer incidence (Paterlini et al., 1995). Throughout the progression of HCC the \( p53 \) protein may form a complex with HBx (Feitelson et al., 1993; Wang et al., 1994; Ueda et al., 1995; Lin et al., 1996). \( p53 \) protein bound to HBx loses its sequence specific DNA binding properties and its transcriptional activation properties (Wang et al., 1994), thus preventing \( p53 \)-mediated apoptosis. HBx has also been found to decrease the expression of \( p53 \) regulated endogenous genes (Wang et al., 1995).

\( p53 \) mutations may occur late in HCC in certain cases (Murakami et al., 1991; Teramoto et al., 1994; Yumoto et al. (1995), thus appearing to contribute to HCC progression rather than to its formation (Yumoto et al., 1995). Several studies have shown different \( p53 \) mutations in different tumour nodules in the same patient, which has led to the belief that \( p53 \) mutations may develop in a stepwise progression in conjunction with histologic progression of the tumour. This may reflect growth advantage for tumours with \( p53 \) mutations (Nishida et al., 1993). The time when mutations occur, however, could conceivably vary depending on the pathogenesis in different cases (Tabor 1997 review). Bressac et al. (1991), and Carson and Lois (1995), found no apparent association of \( p53 \) mutations with the stage of tumour development (early versus late). It is thus possible that a second tumour suppressor gene on 17p is involved in the development of some HCCs (Nishida et al., 1993; Yumoto et al., 1995).
1.4.2) The Retinoblastoma (RBI) gene

The RBI gene is a tumour suppressor gene (Hsia et al., 1994) or one of a class of cancer genes where the functional loss of both alleles is critical for tumour formation (reviewed by Benedict et al., 1990). The retinoblastoma gene product (pRB) functions as a cell cycle regulator, and disturbance of its function is involved in the pathogenesis of hereditary and sporadic retinoblastoma (Hsia et al., 1994; Wooster et al., 1995).

When the retinoblastoma protein is absent, unrestricted cell growth occurs. An absence of pRB can be attributed to: (i) mutation or deletion of the RBI gene, (ii) binding of certain DNA tumour viral proteins to wild-type pRB thereby inactivating it, (iii) non-functional TGF-β1, resulting in a lack of TGF-β1 suppression of the enzyme that phosphorylates pRB and ultimately in freely phosphorylated pRB (Laiho et al., 1990). Other physiologic growth inhibitory signals such as cyclic AMP and contact inhibition are known to prevent the phosphorylation of pRB. p21 is also known to block pRB phosphorylation in response to DNA damage. In this procedure p53 acts as an intermediary, its steady state level increasing rapidly upon DNA damage. p53 then activates p21 CDKI expression causing p21 to block the activity of CDK2 and CDK4/CDK6. The ultimate goal of this system is to prevent pRB phosphorylation and cell cycle progression beyond G1 (reviewed in Weinberg 1995).
1.4.2.1) The *RB1* Gene and Protein

The *RB1* gene is located on chromosome 13q14.2 (Lee *et al.*, 1987; McGee *et al.*, 1989) and comprises 27 exons, spanning approximately 200kb of DNA. This gene encodes a 4.7kb mRNA (McGee *et al.*, 1989) from which a 298 amino acid nuclear phosphoprotein is translated (Lee *et al.*, 1987; McGee *et al.*, 1989; Hsia *et al.*, 1994).

1.4.2.2) pRB1 Function

The retinoblastoma protein has been shown to suppress neoplastic properties of cells and to form complexes with several viral and cellular proteins involved in growth regulation. It has been suggested that inhibition of pRB function may lead to the immortalisation of cells, and that at least in certain lineages, pRB might be a component of a 'generational clock', which acts to record the number of divisions that a cell undergoes before it differentiates, dies, or irreversibly exits the cell cycle (Weinberg 1991). The retinoblastoma protein is thus believed to act as a signal transducer connecting the cell cycle clock with the transcriptional machinery, and creating a mechanism for braking cell proliferation by modulating gene expression (reviewed by Weinberg, 1995). If a cell sustains genetic damage while in G2, M, or most of G, a pRB imposed cell cycle block in G1 will allow for DNA repair before the cell proceeds to replication. Once damaged DNA has been repaired the pRB-block may be lifted and the cell permitted to advance into S and replicate its
restored genome. If a DNA lesion is impossible to repair the cell may proceed to apoptosis, where p53 acts as an intermediary. Inhibition of pRB has been shown to induce apoptosis in hepatocytes directly from G₁ phase. Loss of pRB can also promote cell cycle progression into S phase from where the cells undergo apoptosis (Fan et al., 1996).

The activity of pRB is believed to be regulated through phosphorylation, where underphosphorylated pRB localises in the nucleus, binds tightly to an unidentified site in the nucleus and prevents cell growth beyond the G₁ phase (Hsia et al., 1994). When phosphorylated, pRB appears to localize in the cytoplasm allowing for cell growth to occur (reviewed by Tabor 1997). pRB can bind E2F transcription factors when hypophosphorylated but not when hyperphosphorylated (Chellappan et al., 1991). E2F transcription factors comprise a group of proteins all of which can interact with variants of the consensus nucleotide sequence TTTCGC'G which is present in the promoters of several genes known, or suspected to be, important in cell growth control (Weinberg 1995). The pRB/E2F complex is found mainly in the G₁ and S phases of the cell cycle (Chellappan et al., 1991; Mudryj et al., 1991). The binding of pRB to E2F is thought to convert E2F from an activator to an inhibitor of transcription. Each E2F appears to form a complex with pRB or one of its cousins in defined periods of the cell cycle. Thus, each may be specialised to control genes at a specific period of G₁ or S. The participation of pRB in inhibiting cell growth while also protecting cells from apoptosis may seem self-contradictory. However, it is thought that E2F provides
a link between these two processes and the abundance and timing of E2F1 determine pRB regulation of the cell cycle by the binding of active pRB to E2F1 and the release of E2F1 from inactive pRB. Simple alternating cycles of binding and release of E2Fs by pRB and its cousins may however, be an oversimplified view of the mechanism by which gene function is regulated. The molar amount of pRB present towards the end of G1 being two orders of magnitude above E2F levels indicates that pRB may regulate several downstream effectors other than E2Fs (reviewed in Goodrich and Lee 1992).

pRB is also thought to play a critical role in cell senescence and in the differentiation of several cell types. The loss of functional pRB may result in cell clones gaining replicative advantage and eventually undergoing immortalisation. In differentiation pRB appears to be a prerequisite to stopping the cell cycle in mid to late G1 and inducing these cells to exit the cell cycle and differentiate further (Weinberg 1995).

In 1971, Knudson proposed a model by which the \textit{RB1} gene may act as a tumour suppressor. The products of a tumour suppressor gene would act to inhibit tumorigeneses by negatively regulating growth or enhancing differentiation. Thus, if the functional loss of both alleles of the \textit{RB1} gene occurred via a two-step model, where one allele of the gene would contain a germline mutation, and a somatic event would later lead to the loss of the other allele, unregulated cell growth and tumour formation would result.
a link between these two processes and the abundance and timing of E2F1
determine pRB regulation of the cell cycle by the binding of active pRB to E2F1
and the release of E2F1 from inactive pRB. Simple alternating cycles of binding
and release of E2Fs by pRB and its cousins may however, be an oversimplistic
view of the mechanism by which gene function is regulated. The molar amount of
pRB present towards the end of G, being two orders of magnitude above E2F levels
indicates that pRB may regulate several downstream effectors other than E2Fs
(reviewed in Goodrich and Lee 1992).

pRB is also thought to play a critical role in cell senescence and in the
differentiation of several cell types. The loss of functional pRB may result in cell
clones gaining replicative advantage and eventually undergoing immortalisation.
In differentiation pRB appears to be a prerequisite to stopping the cell cycle in mid
to late G, and inducing these cells to exit the cell cycle and differentiate further
(Weinberg 1995).

In 1971, Knudson proposed a model by which the RBl gene may act as a tumour
suppressor. The products of a tumour suppressor gene would act to inhibit
tumorigenesis by negatively regulating growth or enhancing differentiation. Thus,
if the functional loss of both alleles of the RBl gene occurred via a two-step model,
where one allele of the gene would contain a germline mutation, and a somatic
event would later lead to the loss of the other allele, unregulated cell growth and
tumour formation would result.
a link between these two processes and the abundance and timing of E2F1 determine pRB regulation of the cell cycle by the binding of active pRB to E2F1 and the release of E2F1 from inactive pRB. Simple alternating cycles of binding and release of E2Fs by pRB and its cousins may however, be an oversimplistic view of the mechanism by which gene function is regulated. The molar amount of pRB present towards the end of G1 being two orders of magnitude above E2F levels indicates that pRB may regulate several downstream effectors other than E2Fs (reviewed in Goodrich and Lee 1992).

pRB is also thought to play a critical role in cell senescence and in the differentiation of several cell types. The loss of functional pRB may result in cell clones gaining replicative advantage and eventually undergoing immortalisation. In differentiation pRB appears to be a prerequisite to stopping the cell cycle in mid to late G1 and inducing these cells to exit the cell cycle and differentiate further (Weinberg 1995).

In 1971, Knudson proposed a model by which the \textit{RBI} gene may act as a tumour suppressor. The products of a tumour suppressor gene would act to inhibit tumorigeneses by negatively regulating growth or enhancing differentiation. Thus, if the functional loss of both alleles of the \textit{RBI} gene occurred via a two-step model, where one allele of the gene would contain a germline mutation, and a somatic event would later lead to the loss of the other allele, unregulated cell growth and tumour formation would result.
a link between these two processes and the abundance and timing of E2F1 determine pRB regulation of the cell cycle by the binding of active pRB to E2F1 and the release of E2F1 from inactive pRB. Simple alternating cycles of binding and release of E2Fs by pRB and its cousins may however, be an oversimplistic view of the mechanism by which gene function is regulated. The molar amount of pRB present towards the end of G1 being two orders of magnitude above E2F levels indicates that pRB may regulate several downstream effectors other than E2Fs (reviewed in Goodrich and Lee 1992).

pRB is also thought to play a critical role in cell senescence and in the differentiation of several cell types. The loss of functional pRB may result in cell clones gaining replicative advantage and eventually undergoing immortalisation. In differentiation pRB appears to be a prerequisite to stopping the cell cycle in mid to late G1 and inducing these cells to exit the cell cycle and differentiate further (Weinberg 1995).

In 1971, Knudson proposed a model by which the RB1 gene may act as a tumour suppressor. The products of a tumour suppressor gene would act to inhibit tumorigenesis by negatively regulating growth or enhancing differentiation. Thus, if the functional loss of both alleles of the RB1 gene occurred via a two-step model, where one allele of the gene would contain a germline mutation, and a somatic event would later lead to the loss of the other allele, unregulated cell growth and tumour formation would result.
1.4.2.3) \textit{RBI} Gene Mutations

Evidence now indicates that pRB is involved in a regulatory pathway that is disrupted during the pathogenesis of multiple human tumours (reviewed by Weinberg 1995). After examination of the frequency of \textit{RBI} gene loss in human tumours, it was concluded that loss of this gene did not occur at analogous stages in all tumours (reviewed by Schubert \textit{et al.}, 1994).

Mutations of the \textit{RBI} gene which result in pRB inactivation have been documented in such malignancies as retinoblastomas, small cell lung carcinomas, bladder carcinomas, prostate cancer, and many sarcomas (reviewed by Weinberg 1995).

1.4.2.4) \textit{RBI} and HCC

There is no direct evidence that mutations of \textit{RBI} are involved in HCCs, although allelic loss at this locus has been frequently observed in these tumours with the use of polymorphic markers (Walker \textit{et al.}, 1991). LOH of the \textit{RBI} gene has been documented in 19\% to 44\% of human HCCs (Murakami \textit{et al.}, 1991; Fujimoto \textit{et al.}, 1994; Kuroki \textit{et al.}, 1995), and in 20\% to 29\% of HCCs pRB has been undetectable by immunohistochemistry (Kawakita \textit{et al.}, 1994; Hsia 1994). The expression of pRB is altered in HCC (Hsia \textit{et al.} 1994), and a significant association between altered pRB expression and tumour grade has been found (histologic grades III and IV).

Inactivation of both alleles of the \textit{RBI} gene has been implicated in the genesis of
HCCs (Zhang et al., 1994), where one allele of the *RBI* gene was lost and mutations (deletions of 1-8bp) were found in the remaining *RBI* allele. No pRB was detected in these tumours. The second allele of the *RBI* gene need not be inactivated by mutations however, but simply by hypermethylation of the promoter region. Normal mutation-screening techniques will not detect this (Ohtani-Fujita et al., 1993).

1.4.3) The Breast Cancer Susceptibility Gene 1 (*BRCA1*)

*BRCA1* has been proposed to regulate proliferation in a negative fashion in adult tissues. The reduced expression of *BRCA1* in vitro results in accelerated growth of breast and ovarian cancer cell lines and overexpression of the gene inhibits the growth of the same cell lines (Thompson et al., 1995; Holt et al., 1996; Rao et al., 1996).

*BRCA1* germline mutations are thought to be responsible for most cases of familial inherited breast and ovarian cancer, and for about 50% of those of inherited breast cancer alone (Easton et al., 1993; Wooster et al., 1994, 1995; Phelan et al., 1996; Tavtigian et al., 1996). No aberrations of this gene have been reported in HCC in the literature, and it is not known to play a role in HCC. However, as it may act in concert with *BRCA2*, which was recently proposed to be involved in HCC development (Katagiri et al., 1996), it was a good candidate to investigate.
1.4.3.1) **BRCA1** Gene and BRCA1 Protein Structure

The **BRCA1** gene has been localised to 17q21 and has 22 exons distributed over approximately 100kb of genomic DNA. The transcript of this gene is 7.8kb in length. Transcripts analysed so far show a complex pattern of alternative splicing near the 5' end, upstream of the presumptive initiation codon. This is similar to the regulatory mechanism operating in the **WTI** gene where alternative splicing of exons that encode parts of the zinc finger domains alters the DNA binding properties of that protein. It remains unclear as to whether this alternative splicing is coordinated with alternative splicing further downstream, and it is also not known if all splice variants produce proteins with an identical NH$_2$-terminus (Miki et al., 1994).

1.4.3.2) **BRCA1** Gene Function

The **BRCA1** gene is believed to encode a protein with DNA binding properties and transcriptional activity (Smith et al., 1992; Miki et al., 1994). Expression of the **BRCA1** gene is regulated in a cell cycle dependent manner, and the BRCA1 protein has also been found to be a substrate for certain cyclin-dependent kinases (Chen et al., 1996). It is thus likely that **BRCA1** plays an important role in cell cycle checkpoint control (Chen et al., 1996), the regulation of cellular proliferation, regulation of recombination and genome integrity (Scully et al., 1997), cellular differentiation (Lane et al., 1995; Marquis et al., 1995), and apoptosis government (Shao et al., 1996). Little is known about the exact role of BRCA1 in any of these processes. It is known that BRCA1 has the ability to form complexes with Rad51.
and p53, either singly or together (Sturzbecher et al., 1996). Rad51 is also known to interact with p53, and it has been postulated that Rad51/BRCA1/p53 may form a tertiary complex involved in sensing DNA damage, as the C-terminal region of p53 (which is also the interaction site with BRCA1) can recognise damaged DNA (Jayaraman and Prives 1995). The interaction of BRCA1 with p53 may provide a BRCA1/p53 complex with an additional activation domain, or BRCA1 may enhance p53 activity by inducing a conformational change in p53 that increases its DNA binding. The BRCA1 protein can also stimulate p53-responsive elements (Ouchi et al., 1998) and can enhance p53-dependent gene expression by acting as a p53 activator. BRCA1 has also been shown to induce p21 (a universal inhibitor of CDKs) thereby regulating the mammalian cell cycle negatively (Somasundaram et al., 1997).

The phosphorylation state of BRCA1 may be important for its function as a cell cycle regulator. The protein undergoes hyperphosphorylation during the G1-S transition and remains in this state throughout mitosis. The levels of BRCA1 are highest during the progression of S phase and do not decrease until M phase (Ruffner and Verma 1997). BRCA1 appears to have different target motifs for activation, the choice of which may vary according to the presence of transcription factors in different cell types (Ouchi et al., 1998).

1.4.3.3) BRCA1 Mutations

The mutations identified thus far in BRCA1 are distributed throughout the gene and
the majority are frameshifts, believed to cause protein truncation and presumably loss of protein function (Miki et al., 1994; Couch and Weber 1996). Breast tumours containing mutations of the BRCA1 gene display loss of the wild-type allele, indicating that loss of functional BRCA1 protein is involved in the genesis of early onset breast cancers (Cornelis et al., 1995). A high frequency of LOH has been found at the D17S856 and D17S855 (flanking the BRCA1 locus) loci in prostate cancer (Gao et al., 1995).

LOH at the BRCA1 locus correlates significantly with a two-to five-fold decrease in BRCA1 protein expression in tumour tissue, as compared with corresponding normal tissue. LOH occurring in a region at the 5' end of the BRCA1 gene also results in decreased protein expression, which may also indicate the deletion of regulatory elements of the promoter resulting in transcription factor failure to regulate BRCA1 expression.

In a study on breast cancer by Tseng et al., (1997), increased loss of heterozygosity at BRCA1 appeared to correlate with abnormal p53 protein. Abnormal p53 protein cannot mediate cell cycle checkpoints, thus leading to genomic instability and neoplastic transformation. There seemed to be a connection between deletions of these loci and abnormal p53 protein. This was not the case for BRCA2. Cells with an unfunctional BRCA1 die unless their checkpoint control has also been lost.
1.4.4) The Breast Cancer Susceptibility Gene 2 (BRCA2)

BRCA2 is a tumour suppressor gene, and most breast tumours in patients with germline mutations of this gene show loss of heterozygosity at the BRCA2 locus, typically the loss of the wild-type allele (Collins et al., 1995; Gudmundsson et al., 1995). Although the incidence of familial breast cancer is lower for germline mutations in the BRCA2 gene than in the BRCA1 gene, breast cancer incidence in men is increased, as well as in women who inherit BRCA2 mutations (Rajan et al., 1997).

The BRCA2 gene is believed to be responsible for 17% of total hereditary breast cancers (Wooster et al., 1995; Phelan et al., 1996).

1.4.4.1) The BRCA2 Gene and BRCA2 Protein Structure

The BRCA2 gene is located on chromosome 13q12-13 (Wooster et al., 1994, 1955; Tavtigian et al., 1996) and consists of 27 exons distributed over approximately 70kb of genomic DNA, from which a mRNA of approximately 11kb is transcribed (Wooster et al., 1995; Couch et al., 1996c; Tavtigian et al., 1996). This transcript is believed to encode a polypeptide comprising 3418 amino acids and has little or no homology to any previously identified protein. Furthermore, the gene appears to be unique, as no close homologue has been found in the human genome (Tavtigian et al., 1996).

1.4.4.2) BRCA2 Gene Function

Presently very little is known about the function of the BRCA2 gene product or the
regulation of its expression. Two functions have been proposed for the BRCA2 protein. The first, that it may have a transcription regulation function (Chapman et al. 1996), stems from the fact that the protein contains a 45-amino acid region with weak similarity to the c-jun protein, a transcription activator in yeast (Milner et al., 1997). The second possible function, that BRCA2 may be involved in the DNA repair process, was proposed when BRCA2 was shown to bind to Rad51, a protein involved in DNA repair and recombination (Sharan et al., 1997). The presence of the BRCA2 protein in the nucleus has also been demonstrated, supporting its involvement in DNA repair (Bertwistle et al., 1997). Furthermore, BRCA2 has been shown to be regulated in a cell cycle dependent manner (Vaughn et al., 1996; Bertwistle et al., 1997). It is induced before DNA synthesis at late G1/early S phase and its presence at this time is thought to reflect a possible role in maintaining genome integrity during the replication of DNA (Bertwistle et al., 1997). The coregulation of both BRCA1 and BRCA2 proteins in a cell cycle dependent manner, suggests that they are possibly regulated by similar stimuli and pathways in multiple cell types. They have however been shown to be differentially regulated during the development of specific endocrine target tissues such as the testis during spermatogenesis and the breast during pregnancy (Rajan et al., 1997). It has been proposed that the BRCA1 and BRCA2 genes may have overlapping functions or act in concert (Rajan et al., 1997). That the functions of these two genes are not entirely redundant is supported by the discovery that mutations of the BRCA2 gene predispose carriers to increased risks of breast cancer despite the presence of wild-type copies of the BRCA1 gene, and vice versa.
Furthermore, homologous regions in the promoter regions of these genes suggest that they may be coordinately controlled by the same transcription factors (Vaughn et al., 1996). The induction of BRCA1 and BRCA2 expression may be a protective response to proliferation. These putative tumour suppressor genes would then participate in a homeostatic regulatory loop in which proliferation would stimulate the expression of the BRCA1 and BRCA2 genes, which in turn would slow the proliferation rate, or exert checkpoint control functions. The normal function of BRCA1 and BRCA2 may well be limited to specific developmental stages (Rajan et al., 1997).

1.4.4.3) BRCA2 Gene Mutations

The mutations thus far identified in BRCA2 are distributed throughout the gene, the majority resulting in protein truncation (Wooster et al., 1995; Couch et al., 1996b; Phelan et al., 1996; Tavtigian et al., 1996). Certain mutations in the coding sequence of the BRCA2 gene may destabilise the transcript as well as disrupt the protein sequence (Tavtigian et al., 1996).

1.4.4.4) BRCA2 and HCC

Katagiri et al. (1996), have proposed that the BRCA2 gene is involved in the development of HCC, and the mRNA of the BRCA2 gene has been shown to be up-regulated in HCCs versus normal livers (Du et al., 1997).

Only three mutations (each in a different tumour) of the BRCA2 gene have thus far
been documented in HCC, in a study involving 60 tumours (3/60 - 5%). Two of these alterations were also present in corresponding normal tissue of the respective patients (Katagiri et al., 1996). Katagiri et al. (1996), also determined that the \textit{BRCA2} gene product is expressed in normal human liver, although the level of expression is lower than in the breast.

In view of these facts, it has been suggested that mutations of the \textit{BRCA2} gene may be involved in hepatocarcinogenesis, and that this gene could function as a tumour suppressor in a hormone-related pathway in the liver and in the breast (Katagiri et al., 1996).

1.4.5) The Wilms' tumour (\textit{WT1}) gene.

The \textit{WT1} gene is a tumour suppressor gene essential for the correct development of the urogenital system, and is localised to 11p13 (Huff et al., 1991). It encodes a potential DNA binding protein with four zinc finger domains (Evans and Hollenberg 1984), which appears to be implicated in processes such as proliferation, differentiation and apoptosis. The exact function of \textit{WT1} is still unknown although it is reported that expression of the gene may allow a cell to respond appropriately to signals from its environment (Menke et al., 1998).

Deregulated activity of the \textit{WT1} gene may lead to tumorigenesis, as it appears to be expressed during normal development in a strict temporal and spatial pattern (Armstrong et al., 1992; Rackley et al., 1993; Kent et al., 1995). A lack of \textit{WT1} can
cause upregulation of apoptosis-inducing factors or downregulation of survival factors; (Miyashita and Reed 1992, 1993; Hewitt et al., 1995).

1.4.5.1) **WT1 Gene, Protein Structure and Function**

The **WT1** gene is composed of 10 exons spanning approximately 50kb (Call et al., 1990; Haber et al., 1991; Gessler et al., 1992; Tadokoro et al., 1992) from which a 3kb mRNA is transcribed (Call et al., 1990; Gessler et al., 1992). The protein's four zinc fingers suggest that it may function as a transcription factor (Rauscher et al., 1990). The transcriptional activity of **WT1** can be affected by its physical interaction with the p53 protein (Maheswaran et al., 1993). Recently it has become apparent that because of alternate RNA splicing and editing (Sharma et al., 1994) mechanisms, the **WT1** gene may encode as many as 16 different **WT1** protein isoforms (Menke et al., 1998).

**WT1** may act as a transcription factor repressor or promoter of the same gene promoters but in different cell types. The ratio of **WT1** protein isoforms in a cell may determine which **WT1** target genes will be expressed. The balance and timing of expression of **WT1** target genes may in turn determine whether a cell will proliferate, differentiate, or undergo apoptosis. **WT1** is therefore important in the cell cycle in different cells at different stages (Miyashita and Reed 1992, 1993; Selvakumaran et al., 1994; Hewitt et al., 1995; Menke et al., 1997). **WT1** has been shown to regulate expression of three genes involved in apoptosis: **TGFβ**, **c-myc** and **bcl-2**. However, whether **WT1** is involved in their transcriptional activation
in vivo has yet to be determined (reviewed by Menke et al., 1998). WT1 proteins may also play a more direct role in apoptosis as WT1 is able to suppress p53-induced apoptosis (Maheswaran et al., 1995) as well as induce apoptosis (Englert et al., 1995; Menke et al., 1997) by mediating signals of both extracellular and internal origin (Evans et al., 1992; Miyashita and Reed 1992, 1993; Belamy et al., 1995). The exact mechanisms are not known. WT1 may also regulate gene expression at the post-transcriptional level and may be involved in splicing (Ward et al., 1995; Larsson et al., 1995).

1.4.5.2) Mutations of the WT1 Gene

Point mutations in the WT1 gene have been found to alter the DNA binding properties of the WT1 protein (Little et al., 1993). Mutations such as deletion of both alleles, partial deletion, and intragenic deletion of the gene have also been reported (Kikuchi et al., 1992). WT1 has been found to be overexpressed in several types of tumours (reviewed by Menke et al., 1998). Wang and Rogler (1988) reported LOH on 11p13 in several HCC tumours and proposed that the tumour suppressor gene lost was WT1. No other studies have reported LOH or polymorphism allele frequencies for this gene to our knowledge.

1.4.6) The E-cadherin (Uvomorulin) Gene

The E-cadherin gene localises to 16q22.1 (Natt et al., 1989) and its product is the primary adhesion molecule in epithelium (Peyrières et al., 1983; Shimoyama et al.,
This molecule is often lost in epithelial cancers (Takeichi 1993; Birchmeier et al., 1994).

Cell-cell adhesion and cell invasion of surrounding tissue are two of the main features that characterise the development of a malignant tumour. In HCC, multicentric development and the formation of intrahepatic metastases is common (Lin et al., 1987; Nagao et al., 1990). Intrahepatic metastasis is known to occur earlier and more frequently than extrahepatic metastasis in these tumours. Furthermore, intra-hepatic recurrences are frequent in patients with HCC treated by liver transplantation (Iwatsuki et al., 1991). These factors contribute to a poor prognosis for HCC patients. Loss of function of E-cadherin may lead to decreased cell-cell adhesion (Takeichi 1990), cellular phenotypic changes, and the development of invasive properties (Behrens et al., 1989; Takeichi 1991). Decreased E-cadherin expression in various human carcinomas in vivo has been found to correlate with invasive and metastatic potential (Shimoyama et al., 1991a, 1991b; Shiozaki et al., 1991), indicating that down-regulation of E-cadherin may aid tumour cell invasion by causing separation of cells from the primary tumour mass.

LOH on chromosome 16q, in the region of the E-cadherin gene, has been previously reported to be important in the initiation or progression of HCC (Tsuda et al., 1990; Zhang et al., 1990). Loss or reduction of E-cadherin expression (Shimoyama et al., 1989; Shimoyama and Hirohashi 1991b) in primary hepatocellular carcinomas has also been observed. Loss of heterozygosity on chromosome 16 has also been found to be
more frequent in poorly differentiated than in well-differentiated liver carcinomas. It has been reported that the ability of carcinomas to invade and to metastasise mainly depends on the degree of epithelial differentiation within the tumours: in other words, poorly differentiated carcinomas are more invasive than well-differentiated ones (Frixen et al., 1991). E-cadherin is thus a good candidate for the tumour suppressor gene at 16q22.1, and may play a particular role in HCC development.

1.4.6.1) E-cadherin Protein Structure and Function

E-cadherin is a 120kD cell surface glycoprotein (Hyafil et al., 1981; Damsky et al., 1983; Gallin et al., 1983) that is involved in a molecular zipper-mediated cell-cell adhesion.

In human liver E-cadherin function and distribution is still not well understood (Ichara et al., 1996). Perl et al. (1998), were the first authors to show that E-cadherin has a causal role in tumour progression in vivo, from benign adenoma to malignant carcinoma when it appears to counteract the transition of the tumour cell to an invasive phenotype by maintaining intercellular adhesion and epithelial organisation. E-cadherin was also shown by Vleminckx et al. (1991), to act as an effective invasion suppressor.

1.4.6.2) Mutations of the E-cadherin Gene

In neoplastic cells, changes in cadherin function may be due to quantitative or qualitative abnormalities. A quantitative abnormality would comprise the reduction
in expression of cadherins at the neoplastic cell surface, while a qualitative abnormality would usually comprise impairment of the normal interaction of cadherin molecules with the actin cytoskeleton of the cell, a necessary cooperation for successful cadherin adhesion of cells (Shimoyama et al., 1989; Behrens et al., 1989; Takeichi et al., 1990, 1991; Frixen et al., 1991; Vleminkx et al., 1991). A number of mechanisms has been proposed to explain E-cadherin dysfunction in cancer, namely, loss or reduction of expression of the E-cadherin gene, and E-cadherin gene mutations (Oda et al., 1994; Kanai et al., 1994).

Loss or reduction of E-cadherin expression in invasive carcinoma cells has coincided with suppression of E-cadherin promoter activity resulting from chromosomal rearrangement, the loss of transcription factor binding, and CpG methylation around the promoter region of this gene. In HCCs aberrant CpG methylation has been found on chromosome 16 near to the promoter region of the E-cadherin gene (Kanai et al., 1997). This may lead to altered gene expression that could contribute to the development of HCC.

The role of cadherins in carcinogenesis is believed to vary depending on the tissue of origin. The E-cadherin system may be involved in intercellular binding and cellular polarity formation of hepatocytes in humans (Ihara et al., 1996). Normal hepatocytes express E-cadherin and possibly N-cadherin and other cadherins, whereas bile duct cells express only E-cadherin. In HCCs decreased or absent expression of one of the two cadherins expressed by normal hepatocytes has been
recorded, as well as abnormal distribution of cadherins at the surface of neoplastic cells, and abnormal cadherin isoforms (Kozyraki et al., 1996). Ihara et al. (1996), have shown that some HCCs represent a novel type of carcinoma that overexpresses E-cadherin.

1.5) The Polymorphism Concept

A polymorphism may be defined as the occurrence together in a population of two or more discontinuous traits in such proportions that the frequency of the rarest could not be maintained only by recurrent mutation (Connor and Ferguson-Smith 1993). The frequency is most often arbitrarily delineated by biologists as 1% or greater.

Polymorphisms may be subdivided into i) protein polymorphisms, where there are different structural forms of the same protein, all genetically determined and ii) DNA polymorphisms, where there are structural variations in the DNA sequence; these changes may be detected at the protein level, or only at the DNA level, depending on the type of mutation.

1.5.1) Restriction Fragment Length Polymorphisms (RFLP)

In this case the term polymorphism refers to variation at a locus and does not imply that there is more than one common allele, although there may be

RFLPs may be employed as genetic markers and they can be used to reveal allelic variation by electrophoretic separation. A polymorphic genetic marker is thus defined
as a locus for which two or more common alleles can be detected (Sutton 1998). Due to a large number of RFLPs in the human genome, markers for almost any chromosomal segment may be obtained, and used in genetic studies.

1.5.2) Variable Number Tandem Repeats (VNTRs)

These are a special case of RFLPs and may also be referred to as microsatellite polymorphisms. These loci are evenly distributed throughout the genome and are extremely useful in studies involving complex genetic traits; they comprise a series of uninterrupted or interrupted dinucleotide (AC) pairs, or other simple repeats. Because of the high number of repeat sequences, such areas of DNA are susceptible to certain types of mutation: mispairing followed by unequal crossing over or enzyme slippage. This can lead to hypervariable numbers of the repeat sequence and the restriction fragments of different lengths (numbers of repeats) are the different alleles (Litt and Luty 1989; Weber and May 1989). The hybridisation of labelled probes to any region of repetitive DNA enables the detection of different alleles at any one chromosomal position. Alternatively, single stranded conformational polymorphisms (SSCPs) enables these VNTRs to be used as genetic markers. SSCP involves PCR of such regions, followed by polyacrylamide gel electrophoresis, where the DNA strands are denatured by heating and coil in on themselves forming different conformational structures based on the variation in the number of repeats in their sequences, thereby changing the mobility of tumour DNA in electrophoresis in relations to non-tumour DNA from the same individual. VNTRs facilitate studies in: DNA fingerprinting, paternity testing, prenatal diagnosis, linkage analysis and the
evolutionary development of, and relationships between, populations and species (Sutton 1998).

VNTRs have been proposed not to be implicated in generating functional differences in gene or polypeptide expression (Sturm et al., 1995).

1.6) Allele frequencies and their Uses

1.6.1) Population Genetics

A population is considered to be a local group of random mating individuals who belong to a single species. Population genetics is the study of populations, and focuses on populations/groups rather than on individuals.

One effective way of studying a population is by the use of allele/gene frequencies, which are the proportions of the different alleles of a gene which exist in a population. Allele frequencies are measured from generation to generation, rather than by focusing on single matings. Populations are dynamic, changing with time as a result of births, deaths, migration, and mergers with other populations. Once the number of alleles and mode of inheritance of a specific gene have been established, the structure of any given population can be studied by measuring the frequency of a given allele controlling a known trait. The problem often arises that allele/gene frequencies cannot be directly determined as in most cases only phenotypes, and not genotypes can be
observed (Klug and Cummings 1986).

1.6.2) Genotypic/chromosomal Markers

When the exact location of a gene is not known, nearby markers can be employed to study or map it. These markers may be morphological, such as abnormal centromeres, knobs, satellites, rearrangements, or deletions of chromosomal region/s, or biochemical markers such as measurable amounts of gene products, enzymes or mRNA molecules, that can be used to assign a gene to a specific chromosome (Stine 1989).

1.6.3) Mapping

Allele/gene frequencies are also vital to establishing genetic maps (the localisation of genes to specific chromosomes). Genetic maps are useful tools for experimental study and manipulation of organisms, as well as for the diagnosis and study of human genetic diseases.

There are two types of genetic maps, genetic linkage maps and physical maps. The genetic linkage map requires that a number of polymorphic markers, located on the same chromosome be found. Genetic recombination between these markers can be used to establish the distance between the loci, and their order along the chromosome. The physical map comprises the distance between genes, which is measured in base pairs. Physical maps are generated using DNA sequencing, recombinant DNA technology, and chromosome banding techniques (Stine 1989).
1.6.4) Statistical Studies Involving Allele Frequency

1.6.4.1) Population Size and the Standard Error

Allele frequency studies generally require large sample sizes. Samples of 20 individuals or less, are considered small and may therefore result only in crude estimates. Confidence in the estimate may however be established by the use of the standard error of the mean calculation. This calculates the chances that the actual frequency in the population, as opposed to the frequency ascertained from a sample of the same population, lies within a certain (but small) range of this observed value (Sutton 1988).

1.6.4.2) The Hardy-Weinberg Model

The Hardy-Weinberg (H-W) model is a mathematical model which was developed by G.H. Hardy and W. Weinberg which can be used to calculate allele frequencies in a population where alleles are not codominant. This law presumes the following conditions: (1) that the size of the population is large enough so as to make the sampling error negligible, (2) that random mating occurs within the population, (3) that no selective advantage exists for any genotype, and (4) that there is an absence of other factors such as mutation, migration, random genetic drift.

The converse is also true. If by molecular techniques the frequency of alleles at a locus can be established, they can be "fitted" to the H-W formula to establish whether that population is in Hardy-Weinberg Equilibrium. If so, the
abovementioned conditions may be assumed with respect to that population. When a population is in equilibrium for a specific allele, the frequency of the allele remains constant from generation to generation. Several alleles may exist for a particular locus. Once the frequencies of these alleles become fixed in a population, they do not change during equilibrium, and a state of genetic variability is conserved. Genetic variability is crucial to the process of evolution (Klug and Cummings 1986).

1.6.4.3) Polymorphism Information Content PIC, Diversity (H), and Observed Heterozygosity (OH)

In order for a locus (and the frequency of its alleles) to be used in population/mapping studies, the locus needs to be informative. The type of informativeness depends on the context of the study. Two measures of informativity are PIC and H. The PIC value gives one an indication of how polymorphic a locus will be, and is often used in studies where pairs of loci are examined for linkage. It reflects how the degree of polymorphism at a marker locus will determine the probability of detecting linkage to the other (index) locus (Botstein et al., 1980).

The Diversity for a locus provides an indication of the degree of expected heterozygosity at that locus, and is calculated by adding together the homozygote genotypic frequencies (calculated from allele frequencies using the H-W model formula) and subtracting these from 1. The closer the PIC or ED values are to 1, the more informative a locus will be (Nei 1987).
Observed Heterozygosity is a value calculated from actual numbers/data collected for a sample of a population. This value indicates the degree of heterozygosity at a locus for the specific population of which the sample is representative.

1.7) Purpose and importance of the study

The genes studied in this work have been investigated in HCCs at one time or another in European, Japanese, and Chinese populations. However, populations differ in environmental, hereditary genetic, and dietary factors (Walker *et al.*, 1991), and it is not yet clear which genetic events contributing to HCC development in different populations are the same, similar or distinct (Unsal *et al.*, 1994). Such information is vital if we are to treat HCC effectively based on its region of origin, especially once gene therapy becomes available.

HCC is a leading cause of death in Africa and China (section 1.2) where the primary risk factors in its induction are HBV infection and aflatoxin exposure. With the exception of the *p53* gene, where only the codon 249 AGG → AGT mutation has been examined to any great extent, little or no work has been done in southern African blacks. Our stored HCC samples originated from a black southern African population mainly from Mozambique, a country with high aflatoxin B1 exposure.
Observed Heterozygosity is a value calculated from actual numbers/data collected for a sample of a population. This value indicates the degree of heterozygosity at a locus for the specific population of which the sample is representative.

1.7) Purpose and importance of the study

The genes studied in this work have been investigated in HCCs at one time or another in European, Japanese, and Chinese populations. However, populations differ in environmental, hereditary genetic, and dietary factors (Walker et al., 1991), and it is not yet clear which genetic events contributing to HCC development in different populations are the same, similar or distinct (Unsal et al., 1994). Such information is vital if we are to treat HCC effectively based on its region of origin, especially once gene therapy becomes available.

HCC is a leading cause of death in Africa and China (section 1.2) where the primary risk factors in its induction are HBV infection and aflatoxin exposure. With the exception of the p53 gene, where only the codon 249 AGG→AGT mutation has been examined to any great extent, little or no work has been done in southern African blacks. Our stored HCC samples originated from a black southern African population mainly from Mozambique, a country with high aflatoxin B1 exposure.
Hypotheses of this study:

1) Loss of heterozygosity is present in the HCC samples studied, and its frequency will be determined.

2) The frequency of loss of heterozygosity differs between the population investigated in this study and the frequencies published for other populations.

3) The \textit{p53} codon 249 AGG→AGT mutation is present in the HCCs studied, and it will be established how it differs from published data on other populations.

4) Coincident mutation/loss of heterozygosity of any of these genes will be present in this population.

5) Microsatellite/genetic instability is an important factor in this population.

6) Allele frequencies at the WT1, D13S137, D13S120, D13S127, D17S855, D16S301, D16S260 microsatellite repeat loci will differ from those previously published on genomic libraries and other populations. They will be established in this study, creating a further reference for future research involving population studies, forensic analyses or paternity testing (sections 1.5 and 1.6).

7) Statistical tests on the allele frequency data will establish how useful the data will prove in such studies (point 6).
SECTION 2
MATERIALS AND METHODS

2.1) SUBJECTS

Tumour and non-tumorous liver tissue, obtained at necropsy or laparotomy and stored at -70°C, was available from 27 southern African blacks with hepatocellular carcinoma, and tumour tissue alone from 29. DNA from 20 random blacks was used to determine allele frequencies for several loci.

2.2) DNA EXTRACTION

DNA from tumour and non-tumorous liver tissue was extracted using a modified "salting-out" procedure (Miller et al. 1988). This method involves the use of high salt concentrations to dehydrate and precipitate the cellular proteins, while the DNA remains soluble. The amount of DNA obtained with this technique was comparable to that obtained using a phenol-chloroform DNA extraction method and was suitable for the procedures employed in this study.

100mg of tumour or non-tumorous tissue was chopped finely with a sterile scalpel blade and suspended in 5ml of buffer A (10mM Tris; 100mM NaCl; 10mM EDTA; pH8), 100μl proteinase K (20mg/ml), 10% SDS (800μl/ml buffer A). The samples were placed on a gently rotating shaker at 37°C, for 2 hours. A final concentration of 1M NaCl was obtained by the addition of 2ml of 6M NaCl to the tubes and these were shaken gently
for 15 minutes at room temperature. Samples were centrifuged at 6 000 xg for 15 minutes at 22°C. The supernatant was removed and equal quantities of chloroform/isoamyl alcohol (24:1) were added and the tubes were mixed gently by inversion for 5 minutes. Samples were centrifuged at 6 000 xg for 10 minutes at 8°C and the supernatant removed to a clean, clear glass tube. Isopropanol was added gradually and the tube gently inverted until the DNA strands precipitated out of solution. DNA was removed using a sterile pipette tip, washed by dipping for 2 minutes in ice cold ethanol, transferred into 200-400 µl of TE (Tris-EDTA) buffer and resuspended on a gently rotating shaker overnight at room temperature.

2.3) TECHNIQUES USED IN GENE ANALYSIS

Loss of heterozygosity (LOH) studies were done for p53, BRCA2, BRCA1, WT1, RB1, and E-cadherin. LOH at the BRCA2 locus was also investigated using Southern blotting and hybridisation with a cDNA probe; and mutations at this locus were detected by SSCP analysis. The p53 codon 249 mutation was detected by PCR-RFLP and confirmed by sequence analysis. Allele frequencies were established by PCR and polyacrylamide gel electrophoresis for the WT1 AC repeat polymorphism and the D13S137, D13S120, D13S127, D17S855, D16S301 and D16S260 loci. The precise number of repeats present in each repeat sequence was established by sequencing several homozygous samples as controls from which the other samples could be sized.
2.3.1) LOSS OF HETEROZYGOSITY AND ALLELE FREQUENCY STUDIES

A standard PCR protocol was followed for all LOH and allele frequency studies. The primers used for PCR are described in Table 2.1

2.3.1.1) PCR

Each PCR reaction consisted of: 100 ng DNA; 1 U Taq DNA polymerase (Promega); 1xNH₄ reaction buffer (16mM(NH₄)₂SO₄); 67mM Tris-HCl (pH 8.8 at 25°C); 0.01% Tween-20; 1 mM MgCl₂; 1 mM dATP, dTTP, dGTP; 0.1 mM dCTP; 0.025 μCi α²³⁵P dCTP; 50 pmol of each primer; in a total volume of 50 μl. Samples were amplified for 30 cycles of: a denaturation step at 94°C for 30 seconds, an annealing step at 55°C for 30 seconds, and a 72°C extension step for 1 minute and a final cycle of 72°C for 10 minutes. Negative water control samples were included in each PCR run to ensure that there was no contamination.

If no PCR product was obtained the procedure was repeated up to three times. A doubling of DNA concentration, increase in MgCl₂ by 1 to 2 mM and a decrease of up to 3°C in primer annealing temperature was attempted in different reactions. If no PCR product was obtained after three attempts the sample was discarded.

2.3.1.2) Gel electrophoresis of PCR products

All PCR products were run on 2 % agarose gels (D1-LE, Hispanagar, Spain) and sized against 50bp DNA ladder (GibcoBRL) or a 100bp DNA ladder (Promega). VNTRs of D13S137 (RB), D13S120 (BRCA1), D13S127 (BRCA2), D17S855
2.3.1) LOSS OF HETEROZYGOSITY AND ALLELE FREQUENCY STUDIES

A standard PCR protocol was followed for all LOH and allele frequency studies. The primers used for PCR are described in Table 2.1.

2.3.1.1) PCR

Each PCR reaction consisted at final volume of: 100ng DNA; 1U *Taq* DNA polymerase (Promega); 1xNH₄ reaction buffer (16mM(NH₄)₂SO₄); 67mM Tris-HCl (pH8.8 at 25°C); 0.01% Tween-20; 1mM MgCl₂; 1mM dATP, dTTP, dGTP; 0.1mM dCTP; 0.025µCi α³²P dCTP; 50pmol of each primer; in a total volume of 50µl. Samples were amplified for 30 cycles of: a denaturation step at 94°C for 30 seconds, an annealing step at 55°C for 30 seconds, and a 72°C extension step for 1 minute and a final cycle of 72°C for 10 minutes. Negative water control samples were included in each PCR run to ensure that there was no contamination.

If no PCR product was obtained the procedure was repeated up to three times. A doubling of DNA concentration, increase in MgCl₂ by 1 to 2 mM and a decrease of up to 3°C in primer annealing temperature was attempted in different reactions. If no PCR product was obtained after three attempts the sample was discarded.

2.3.1.2) Gel electrophoresis of PCR products

All PCR products were run on 2% agarose gels (D1-LE, Hispanagar, Spain) and sized against 50bp DNA ladder (GibcoBRL) or a 100bp DNA ladder (Promega). VNTRs of D13S137 (*RB*), D13S120 (*BRCA2*), D13S127 (*BRCA2*), D17S855
(BRCA1), D16S301 (E-cadherin); D16S260 (E-cadherin) were resolved on polyacrylamide gels, and polymorphic alleles of p53, D17S846 (BRCA1) and RB1.20 (RB1) were resolved on composite agarose gels (see section 3.1.1.8). Tumour and non-tumorous liver DNA samples of subject 2 were included on gels to ensure consistency of scoring.

2.3.1.2.1) Polyacrylamide gels

Radioactively labelled PCR products of the D13S137, D13S120, D13S127, D17S855, D16S301 and D16S260 loci were run on a 6% sequencing gel (Appendix A). A mixture of 5μl of each sample and 5μl of denaturing loading dye (Appendix A) was denatured at 100°C for 3 minutes and loaded from ice onto the gel. Up to 3 loadings were done on each gel, each 2 hours apart. The gel was run at 60W for 2-6 hours, blotted onto Whatman filter paper (3mm) and covered with plastic wrap before being dried in a slab gel drier, at 80°C for 40 minutes. The gel was autoradiographed with Kodak X-ray film (Kodak, Sigma, USA) in an Okamoto S-type cassette with an intensifying screen, for several hours, or overnight. Samples were sized against a pBR322 DNA which had been restricted with HpaII restriction endonuclease and radioactively labelled with Klenow enzyme (Boehringer Mannheim, Mannheim, Germany). A control sample was run on every gel to ensure uniformity of scoring. The number of repeats in each system were established by direct sequencing of the PCR product from a homozygous individual.
(BRCA1), D16S301 (E-cadherin); D16S260 (E-cadherin) were resolved on polyacrylamide gels, and polymorphic alleles of p53, D17S846 (BRCA1) and RB1.20 (RB1) were resolved on composite agarose gels (see section 3.1.1.8). Tumour and non-tumorous liver DNA samples of subject 2 were included on gels to ensure consistency of scoring.

2.3.1.2.1) Polyacrylamide gels

Radioactively labelled PCR products of the D13S137, D13S120, D13S127, D17S855, D16S301 and D16S260 loci were run on a 6% sequencing gel (Appendix A). A mixture of 5μl of each sample and 5μl of denaturing loading dye (Appendix A) was denatured at 100°C for 3 minutes and loaded from ice onto the gel. Up to 3 loadings were done on each gel, each 2 hours apart. The gel was run at 60W for 2-6 hours, blotted onto Whatman filter paper (3mm) and covered with plastic wrap before being dried in a slab gel drier, at 80°C for 40 minutes. The gel was autoradiographed with Kodak X-ray film (Kodak, Sigma, USA) in an Okamoto S-type cassette with an intensifying screen, for several hours, or overnight. Samples were sized against a pBR322 DNA which had been restricted with HpaII restriction endonuclease and radioactively labelled with Klenow enzyme (Boehringer Mannheim, Mannheim, Germany). A control sample was run on every gel to ensure uniformity of scoring. The number of repeats in each system were established by direct sequencing of the PCR product from a homozygous individual.
2.3.1.2.2) Agarose gels

If products could not be resolved on polyacrylamide gels they were run on 4% composite agarose gels (3:1 MS8:D1 LE, Hispanagar). Products were sized against 100bp and 50bp DNA ladders.

2.3.1.3) Sequencing - standard protocol

All sequencing was carried out using Sequenase PCR Product Sequencing Kit (United States Biochemical Corp., Cleveland, Ohio), according to the manufacturer’s instructions. Approximately 15ng of DNA was pre-treated (to remove residual single-stranded primers, extraneous single-stranded DNA and unincorporated dNTPs) with 10U of Exonuclease and 1U of Shrimp Alkaline Phosphatase at 37°C for 15 minutes, followed by 15 minutes at 80°C to inactivate the enzymes. Pre-treated DNA was mixed with 10pmol of appropriate primer and denatured at 100°C for 5 minutes, snap cooled in an ice-water bath for 5 minutes and after centrifuging briefly, placed on ice. To this 7.5μl of labelling mixture was added, comprising: sequenase reaction buffer (1.4x concentrate) (143mMTris-HCl pH7.5; 71mM MgCl₂; 179mM NaCl); 2μl 0.01M labelling mix (dGTP) 5x concentrate diluted 1:10; 7.5μM each dGTP, dATP, dCTP and dTTP; 3.2U Sequenase version 2.0 T7 DNA polymerase (1.6U/μl with inorganic Pyrophosphatase [2U/ml] in 20mM Tris-HCl, pH7.5, 2mM DTT, 0.1mM EDTA and 50% glycerol); 0.01M DTT; 0.5-1μl ^35S dATP (3000 mCi). Samples were labelled at room temperature for 2 minutes and placed on ice for a further 10 minutes. Labelled samples were added in aliquots of 3.5μl to 4 micro-titre plate
### Table 2.1: PCR Primers

<table>
<thead>
<tr>
<th>GENE/LOCUS</th>
<th>PRIMER</th>
<th>PRIMER SEQUENCE</th>
<th>AMPLICON</th>
<th>AMPLICON LENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>p53</em></td>
<td>p5IF1</td>
<td>5'-GTTGGCTCTGACTGT-ACAC-3'</td>
<td>exon 7 spanning codon 249</td>
<td>110 bp</td>
</tr>
<tr>
<td></td>
<td>p5IR1</td>
<td>5'-CTGGAGTCCTCCGAGTTGT-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>p53</em></td>
<td>p5IFS1a</td>
<td>5'-GCACTTTCCTCAACTCTACA</td>
<td>ALU sequence within intron 1 of <em>p53</em> gene</td>
<td>200-300bp</td>
</tr>
<tr>
<td></td>
<td>p5IFS1b</td>
<td>5'-AACAGCTCTTAAATGGGCAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>BRCA2</em></td>
<td>F1</td>
<td>5'-CCATATTTAAACACCTAGG</td>
<td>ALU sequence within intron I of <em>BRCA2</em> gene</td>
<td>153 bp</td>
</tr>
<tr>
<td></td>
<td>R1</td>
<td>5'-ATAGTCACACCTGTGTTTCAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>BRCA2</em></td>
<td>F2</td>
<td>5'-GAGAATATTAGTGGAGGAC</td>
<td>polymorphic site intron 11 encompassing transversion at position 508</td>
<td>190 bp</td>
</tr>
<tr>
<td></td>
<td>R2</td>
<td>5'-CAAAAGTGCCAGTAGTCATT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D13S120</td>
<td>133L</td>
<td>5'-ATGACCTGAAATGATACTCGG</td>
<td>(AC)9, repeat at D13S120</td>
<td>112-136 bp</td>
</tr>
<tr>
<td></td>
<td>133R</td>
<td>5'-CAGACACACAACACAGATT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D17S846</td>
<td>FF</td>
<td>5'-TGCTATTCTGCTACCACTCAG</td>
<td>(GGAA)8, repeat at D17S846</td>
<td>250-300bp</td>
</tr>
<tr>
<td></td>
<td>RF</td>
<td>5'-TCCTTGAGAGGATTTTCACTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D17S551</td>
<td>FS</td>
<td>5'-GGATGCGCTTTTAGAAAGTG</td>
<td>AC repeat at D17S551</td>
<td>146 bp</td>
</tr>
<tr>
<td></td>
<td>RS</td>
<td>5'-CAACAGACTGCTGTGTTGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WTI</td>
<td>400</td>
<td>5'-AATGAGACGATTCTGCG</td>
<td>AC repeat within 3' untranslated sequence of WTI</td>
<td>100-200bp</td>
</tr>
<tr>
<td></td>
<td>401</td>
<td>5'-TTGAGTCTGTAATTTTCAGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>RBI</em></td>
<td>R57</td>
<td>5'-TGATCCTGCTAGCTCTCTG</td>
<td>(CTT)(T), (n=14-26), repeat within intron 20 of <em>RBI</em> gene</td>
<td>400-600bp</td>
</tr>
<tr>
<td></td>
<td>R103</td>
<td>5'-AATGGATCTGCGGTTGGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D13S127</td>
<td>134L</td>
<td>5'-CAGATATGCTACTGCACTAG</td>
<td>(AC)35 repeat at D13S127</td>
<td>130-142bp</td>
</tr>
<tr>
<td></td>
<td>134R</td>
<td>5'-AAACAAATGATGCTGGCTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D13S137</td>
<td>RBa</td>
<td>5'-TTTCTCATTCCTTCCTCAATGG</td>
<td>(GT)22, repeat at D13S137</td>
<td>+/-135 bp</td>
</tr>
<tr>
<td></td>
<td>Rbb</td>
<td>5'-CAGGGAGATGGGTGACCTCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E-cadherin</em></td>
<td>E-cadF1</td>
<td>5'-GATCTCTAGAGGACAAATGCTAGTCCT</td>
<td>D16S301 locus Polymorphic AC repeat region</td>
<td>145 bp</td>
</tr>
<tr>
<td></td>
<td>E-cadR1</td>
<td>5'-AGCCACTTCAGAGGACTGCTTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>E-cadF2</td>
<td>5'-GGTTGAGATGCTGACATGC</td>
<td>D16S260 locus Polymorphic AC repeat region</td>
<td>±234 bp</td>
</tr>
<tr>
<td></td>
<td>E-cadR2</td>
<td>5'-CAGGGTGCTGTTGATATG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Key:

WT1 - Wilm's tumour gene
RB1 - Retinoblastoma gene
bp - base pairs
1 - Bressac et al., 1991
2 - Futreal et al., 1991
3 - Bowcock et al., 1993
4 - Fletjer et al., 1993
5 - Anderson et al., 1993
6 - Haber et al., 1990; Call et al., 1990
7 - Henderson et al., 1994; Yandell et al., 1989
8 - Petrukhin et al., 1993
9 - Thompson et al., 1992
10 - Weber et al., 1990
wells, each containing 2.5μl of each termination mix respectively ddG, ddA, ddT, ddC. Termination mixes comprised 80μM dNTP, 8μM ddNTP and 50mM NaCl. The termination reaction was carried out at 42°C for 5 minutes and the reaction was stopped with the addition of 4μl of sequencing stop solution (95% formamide; 20mM EDTA; 0.05% Bromophenol blue; 0.05% Xylene Cyanol FF).

2.3.1.3.1) Sequencing gels

Samples were heated to 83°C on a Hāgar dry block heater for 2 minutes and loaded from ice onto 8% glycerol tolerant sequencing gels (8% polyacrylamide; 7M urea; 1x glycerol tolerant buffer). Gels were run at 60W for 2 hours and 30 minutes, fixed by soaking in fixing solution (5% acetic acid; 15% methanol) for 10 minutes, blotted onto Whatman filter paper (3mm), dried at 80°C for 90 minutes in a slab gel drier and autoradiographed with Kodak X-ray film in an Okamoto cassette without an intensifying screen. Films were developed after 3 days.

2.3.1.4) Repeat Sequence Analysis and Allele frequency analysis

2.3.1.4.1) p53 gene

This PCR amplifies an ALU sequence which terminates in (AAAAT)₈ in intron 1 of the p53 gene (Futreal et al., 1991). The reaction was altered by the addition of 1mM dCTP rather than 0.1mM dCTP and 0.025μCi α²²P dCTP.
2.3.1.4.2) **BRCA2 gene**

Two polymorphic DNA regions which flank the *BRCA2* gene, D13S127 and D13S120 (Bowcock *et al.*, 1993), were amplified in separate PCR reactions.

2.3.1.4.3) **BRCA1 gene**

Two polymorphic DNA regions which flank the *BRCA1* gene, D17S846 (Fletjer *et al.*, 1993) and D17S855 (Anderson *et al.*, 1993), were amplified in separate PCR reactions.

The PCR for D17S846 amplified a [GGAA]_{15} repeat region and was modified by the addition of 1mM dCTP rather than 0.1mM dCTP; 0.025μCi α^{32}P dCTP and an annealing temperature of 57°C instead of 55°C.

The PCR for D17S855 amplified an AC repeat region and was modified by the addition of 2.5U *Taq* DNA polymerase (Promega) rather than 1U of *Taq* DNA polymerase.

2.3.1.4.4) **Wilm's Tumour gene**

This PCR amplified an AC repeat fragment within the 3' untranslated region of the *WT1* gene (Call *et al.*, 1990) and was modified by the addition of 2.5U *Taq* DNA Polymerase (Promega) rather than 1U of *Taq* DNA Polymerase.
2.3.1.4.5) **RB1 gene**

Two polymorphic DNA regions which flank the RB gene, D13S127 (Bowcock et al., 1993) and D13S137 (Petrukhin et al., 1993), were amplified in separate PCR reactions. A third PCR amplified a polymorphic region within intron 20 of the RB gene, Rb(1.20) (Yandell et al., 1989; Henson et al., 1994).

The PCR for D13S127 amplified an (AC)$_{3j}$ repeat. The PCR for D13S137 amplified a (GT)$_{22}$ repeat and was modified by the addition of 2.5U Taq DNA Polymerase (Promega) rather than 1U of Taq DNA Polymerase. The PCR for Rb(1.20) amplified a variable number of CTTT(T) repeats and was modified by the addition of 1mM dCTP rather than 0.1mM dCTP, and 0.025μCi α$^{32}$P dCTP.

2.3.1.4.6) **E-cadherin gene**

Two polymorphic DNA regions which flank the E-cadherin gene, D16S260 (Weber et al., 1990), and D16S301 (Thompson et al., 1992) were amplified in separate PCR reactions.

The PCR for both D16S301 and D16260 amplified a variable number of AC repeats. The PCR reaction consisted of, at final volume: 100ng DNA; 1U Taq DNA Polymerase; 1xKCl reaction buffer (1.5mM MgCl$_2$); 0.1% gelatin; 1mM each dGTP, dATP, dTTP; 0.1mM dCTP; 0.025μCi α$^{32}$P dCTP; 50pmol of each primer; in a total volume of 25μl. Samples were amplified in a Hybaid thermal cycler for 25 cycles of: denaturation at 94°C for 1 minute, annealing at 55°C
for 2 minutes, 72°C tension for 2 minutes and 30 seconds, and a final cycle of 72°C for 10 minutes. Negative water control samples were included in each PCR run to ensure that there was no contamination.

2.3.2) SSCP ANALYSIS OF THE BRCA2 GENE

PCR-SSCP analysis was used to investigate possible mutations in the BRCA2 gene at two polymorphic sites in the gene. These had previously been described by Tavtigian et al., 1996.

2.3.2.1) Primer design for the BRCA2 gene

The primer pairs described by Tavtigian et al., 1996 were not suitable for our purposes. Primers (Table I) had to encompass the two polymorphic sites located at positions 97(A→G)(PM-3668) and 508(T→C)(PM-3045), (Tavtigian et al., 1996) and were designed accordingly using the Oligo computer program (MWG-Biotech GmbH, Ebersberg, Germany).

The PCR reaction to amplify the polymorphisms at position 97 of the BRCA2 gene consisted at final volume of: 100ng DNA; 2U Taq DNA Polymerase (Promega); 1xNH4 reaction buffer; 1mM MgCl₂; 1mM each dGTP, dATP, dTTP; 0.1mM dCTP; 0.025μCi α³²P dCTP; 50pmol of each primer; in a total volume of 50μl. Samples were amplified in a Hybaid thermal cycler for 30 cycles of: denaturation at 94°C for 30 seconds, annealing at 50°C for 30 seconds, and a 72°C extension
step for 1 minute. Negative water control samples were included in each PCR run to ensure that there was no contamination. Samples that did not work at the first attempt were repeated up to three times. MgCl₂ was increased by 1 to 2 mM and a decrease of up to 3°C in primer annealing temperature was attempted in different reactions. If no PCR product was obtained after three attempts the sample was discarded.

The PCR reaction to amplify the polymorphism at position 508 of the *BRCA2* gene consisted at final volume of: 100ng DNA; 2.5U *Taq* DNA Polymerase (Promega); 1xNH₄ reaction buffer; 1mM MgCl₂; 1mM each dGTP, dATP, dTTP; 0.1mM dCTP; 0.025μCi α³²P dCTP; 50pmol of each primer; in a total volume of 50μl. Samples were amplified in a Hybaid thermal cycler for 30 cycles of: denaturation at 94°C for 30 seconds, annealing at 50°C for 30 seconds, and a 72°C extension step for 1 minute. Negative water control samples were included in each PCR run to ensure that there was no contamination.

Samples that did not work at the first attempt were repeated up to three times. MgCl₂ was increased by 1 to 2 mM and a decrease of up to 3°C in primer annealing temperature was attempted in different reactions. If no PCR product was obtained after three attempts the sample was discarded.

2.3.2.2) *Gel electrophoresis of PCR products*

PCR products were run on an MDE gel (MDE [FMC Bioproducts, Natick,
Massachusetts, USA] containing 1.2xTBE (Appendix A), glycerol (Appendix A) [SAARCHEM-HOLPRO Analytic, Cape Town, SA], APS [BIO RAD Laboratories, Hercules, California, USA], TEMED [Promega]). Four \( \mu l \) of PCR product for each sample was mixed with 6\( \mu l \) of denaturing loading dye, denatured at 95\(^\circ\)C for 2 minutes and 7\( \mu l \) loaded. The gel was run for 12 hours at 6W at room temperature, dried in a slab gel drier at 80\(^\circ\)C for 40 minutes and autoradiographed with Kodak X-ray film in an Okamoto cassette with an intensifying screen, for several hours, or overnight. Samples were scored in terms of differences in band patterns between tumour and non-tumorous tissues of the same individuals. Those samples in which a difference was noted were sequenced.

Direct sequencing of the PCR product revealed that multiple sequences (Section 2.3.1.3) were present, so the SSCP approach was abandoned and polymorphic repeat sequences flanking the \textit{BRCA2} gene were investigated.

### 2.3.3) LOH ANALYSIS OF THE \textit{BRCA2} GENE

LOH at the \textit{BRCA2} gene was established by Southern blotting and hybridisation. DNA from 4 southern African black subjects was digested with 8 different restriction enzymes before Southern blotting and hybridisation were performed in order to expose any \textit{BRCA2} polymorphic sites.
2.3.3.1) Restriction Analysis

DNA from 4 southern African black subjects was digested with 8 different restriction enzymes (Table 2.2) to reveal any polymorphic alleles, and Southern blotting performed in order to reveal RFLPs at the *BRCA2* gene. All digest reactions, at final volume, comprise: 5μg DNA; 1x appropriate reaction buffer; 30U enzyme; in a total volume of 50μl. Digests were carried out at 37°C overnight, except for the *TaqI* digest which was carried out at 65°C for 4 hours. Reactions were stopped by placing the tubes on ice for 10 minutes. Thereafter, 5μl of digest mixture was electrophoresed on a 1% agarose gel to check that digestion did occur. All enzymes were supplied by AEC Amersham except for *TaqI* which was supplied by Boehringer Mannheim.

Table 2.2: Restriction Enzymes used for the digestion of DNA prior to Southern blotting, and their respective buffers.

<table>
<thead>
<tr>
<th>ENZYME</th>
<th>BUFFER</th>
<th>BUFFER COMPOSITION (10x)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>PstI</em></td>
<td>H</td>
<td>500mM Tris-HCl, pH7.5; 100mM MgCl₂; 10mM Dithiothreitol; 1000mM NaCl</td>
</tr>
<tr>
<td><em>EcoRV</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>BglII</em></td>
<td>M</td>
<td>100mM Tris-HCl, pH7.5; 100mM MgCl₂; 10mM Dithiothreitol; 500mM NaCl</td>
</tr>
<tr>
<td><em>SaiI</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>HindIII</em></td>
<td>M</td>
<td>100mM Tris-HCl, pH7.5; 100mM MgCl₂; 10mM Dithiothreitol; 500mM NaCl</td>
</tr>
<tr>
<td><em>EcoRI</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>BamHI</em></td>
<td>K</td>
<td>200mM Tris-HCl, pH8.5; 100mM MgCl₂; 10mM Dithiothreitol; 1000mM KCl</td>
</tr>
<tr>
<td><em>TaqI</em></td>
<td>B</td>
<td>10mM Tris-HCl; 5mM MgCl₂; 100mM NaCl; 1mM 2-Mercaptoethanol</td>
</tr>
</tbody>
</table>
2.3.3.2) **Agarose gel electrophoresis**

DNA restricted with each of the 8 restriction enzymes (section 2.3.3.1) was electrophoresed on 0.8% 22cmx20cm agarose gels in 1xTBE (Appendix A) buffer at 60V overnight. A 1kb DNA ladder (Promega) was loaded on the gels in both the first and last lanes. The gels were photographed and the lanes containing ladder removed prior to Southern blotting.

2.3.3.3) **Southern Blotting**

The DNA was denatured by washing the gels in 1.0M NaCl/0.5M NaOH (Appendix A) twice, for 20 minutes each time. The gels were neutralised by soaking in 0.5M Tris/pH7.5, 1.5M NaCl, twice for 20 minutes each time. The gels were blotted using the Southern blot (Southern, 1975) procedure, according to the manufacturer’s instructions for the MSI nylon membrane (Micron Separations Incorp., Westborough, Massachusetts, USA).

2.3.3.4) **Probe preparation**

Two partial \(BRCA2\) cDNA clones (Table 2.3) were kindly provided by Dr S. Tavtigian (Myriad Genetics Inc., Salt Lake City, Utah, USA). The inserts, cloned into pUC18 with a slightly modified polylinker, were termed “AC” and “DG” and overlapped to make up a complete coding sequence for the \(BRCA2\) gene (S. Tavtigian - personal communication).
Table 2.3: Details of the "AC" and "DG" probes.

<table>
<thead>
<tr>
<th>CLONE</th>
<th>CONTENTS</th>
<th>CLONING SITES</th>
<th>INSERT SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;AC&quot;</td>
<td>exons 1-11</td>
<td>Sal I- Eco RV</td>
<td>5.3 kb</td>
</tr>
<tr>
<td>&quot;DG&quot;</td>
<td>exons 11-27</td>
<td>Sal I - Sac I</td>
<td>4.4 kb</td>
</tr>
</tbody>
</table>

2.3.3.4.1) Transformation

The clones were electroporated into competent JM109 cells which have a lacZΔM15 deletion allowing for blue/white colour screening of recombinant clones by expression of the lacZ α-peptide. The cells were thawed on ice and 40μl of cell suspension was mixed, in a cold polypropylene tube, with 1-2μl of DNA and set on ice. The Gene Pulser apparatus was set at 25μF and 2.5kV with the Pulse Controller at 200Ω. The mixture of cells and DNA was transferred to a cold, 0.2cm electroporation cuvette and placed in a chilled safety chamber slide. The cells were pulsed once at the above settings, which produced a pulse with a time constant of 4-5msec and a filed strength of 12.5kV/cm. Cells were resuspended quickly in 1ml of SOC medium (Appendix A), transferred to a polypropylene tube and incubated at 37°C for 1 hour with shaking, before being plated on selective medium (Luria agar with Ampicillin [0.05mg/ml]; 0.2M IPTG; 5% Xgal). Transformants (white colonies) were selected, transferred onto grid plates containing the same selective medium and incubated overnight at 37°C. From the grid plates 10 transformants for each insert were chosen and grown overnight in 5 ml Luria broth containing 0.05mg/ml Ampicillin (Appendix A) at 37°C with gentle shaking. A 2ml aliquot of culture was taken for each transformant and spun at 40 000 xg in a
microfuge for 10 minutes before proceeding to the plasmid prep.

2.3.3.4.1.1) DNA Plasmid Preparation

The plasmid DNA was isolated from the cells using the Qiagen Spin Miniprep Kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's instructions. The pellet of bacterial cells was resuspended in 250μl of resuspension buffer, to which 250μl of lysis buffer was added. The tube was inverted several times, 350μl of neutralisation buffer was added and the tube inverted again. The samples were centrifuged at 35 000 xg in a microfuge for 10 minutes and the supernatants loaded onto a spin column placed inside a 1.5ml polypropylene tube and centrifuged for 30-60 seconds. The columns were washed by adding 0.5ml of wash buffer to each and centrifuging them for 30-60 seconds. The flow-through was discarded and the columns washed with 0.75ml of wash buffer and centrifuged for 30-60 seconds. The flow-through was discarded and the column was spun for an additional 1 minute spin to remove residual buffer. The DNA was eluted by addition of 50μl of ddH₂O to the columns and allowing them to stand for 1 minute before centrifuging them for 1 minute. A small aliquot of eluate was run on an agarose gel to check the yield.

2.3.3.4.1.2) Plasmid digests

Two batches of pUC18 plasmid were originally provided (section 2.3.3.4) - one containing the cDNA clone termed “AC” and the second the cDNA clone
termed “DG”. Together, the “AC” and “DG” cDNA’s would later be used as probes for hybridisation after Southern blotting. Before labelling of the probes could proceed the “AC” and “DG” inserts were excised from the plasmid by restriction endonuclease digestion and isolated.

2.3.3.4.1.2.1) Restriction endonuclease digestion of the pUC18 plasmid containing the "AC" insert

The “AC” insert was excised from the plasmid in two stages. This was necessary due to the inactivity of SalI in buffer B, the optimal buffer for EcoRV. Initially the vector was digested in a reaction comprising: 40ng DNA (pUC18 plasmid containing “AC” insert); 10U EcoRV (B/M); 1xreaction buffer B; in a total volume of 20μl. The digest was allowed to proceed at 37°C for 1 hour, and the reaction terminated by placing the tube at 65°C for 10 minutes. A 2μl aliquot of this reaction was electrophoresed on a 1% agarose gel to ensure that digestion of the plasmid had occurred. To fully excise the “AC” insert, 10U of SalI (B/M) and 1mM DTE was added to the initial reaction and the Tris-HCl increased to 40mM before the volume was adjusted to 50μl with water. The remainder of the procedure was as stated above for the initial digest.

2.3.3.4.1.2.2) Restriction endonuclease digestion of the pUC18 plasmid containing the “DG” insert

Once again it was necessary to digest the plasmid DNA in two stages, due
to the inactivity of *SalI* in buffer L, the optimal buffer for *SacI*. The procedure was as for the pUC18 plasmid with the "AC" insert with the following modifications: (1) the initial digest comprised 40ng DNA (pUC18 plasmid containing "DG" insert); 10U *SacI* (B/M); 1xreaction buffer L (1mM Tris-HCl; 1mM MgCl₂; 0.1mM DTE) (B/M); (2) to fully excise the "DG" insert, 10U of *Sall* (B/M) and 100mM NaCl was added to the initial reaction and the Tris-HCl concentration was increased to 50mM before the volume was adjusted to 50μl with water.

2.3.3.4.1.3) Agarose gel electrophoresis of digested plasmid DNA

For each insert the remaining 48μl of digested plasmid DNA was separated on a 2% agarose gel and the DNA insert ("AC" insert: 5.3kb; "DG" insert: 4.4kb) excised and placed in a reaction tube.

2.3.3.4.1.4) DNA extraction from agarose gel

The DNA was extracted using an Agarose Gel DNA extraction kit (B/M), as per the manufacturer's instructions. Agarose solubilisation buffer was added to the reaction tube at 300μl of buffer per 100mg of agarose gel. The silica suspension was resuspended until homogenous and 10μl was added to the sample. If the sample contained more than 2.5μg of DNA the silica suspension was increased by 4μl for each additional μg of DNA. The mixture was incubated at 56-60°C for 10 minutes with vortexing every 2-3 minutes and centrifuged at 35 000 xg for 30 seconds. The supernatant was
discarded and the matrix containing the DNA resuspended with 500\mu l of nucleic acid binding buffer on a vortex mixer. The sample was centrifuged and the supernatant discarded as before. The pellet was washed with 500\mu l of washing buffer, centrifuged as before, and the supernatant discarded. This step was repeated once. All liquid was removed with a pipette and the tube inverted on an absorbent tissue before the sample was left to dry at room temperature for 15 minutes. The DNA was eluted with 20-50\mu l of ddH$_2$O. The sample was vortexed and incubated at room temperature for 10 minutes with vortexing every 2-3 minutes. After centrifugation at 35 000 xg for 30 seconds the DNA containing solution was transferred to a clean reaction tube without carrying over any matrix material. Eluted DNA was run on a 2% agarose gel to check the yield.

2.3.3.4.1.5) Labelling of probes

The "AC" and "DG" probes were radioactively labelled using the Megaprime DNA labelling systems kit (United States Biochemical). The DNA was diluted to a concentration of 5ng/\mu l in ddH$_2$O, placed in a microcentrifuge tube, and 5\mu l of primer solution (random nonamer primers in an aqueous solution) was added. The DNA was denatured by heating to 95-100°C for 5 minutes in a Hågar heating block. The tube was briefly spun in a microcentrifuge at 35 000 xg to bring the contents to the bottom before 5\mu l of 10xlabelling buffer (dATP; dGTP; dTTP; Tris-HCl pH5.2; 2-mercaptoethanol; MgCl$_2$) and 2\mu l of enzyme solution (1U/\mu l DNA
Polymerase I Klenow fragment [cloned] in 50mM potassium phosphate pH6.5; 10mM 2-mercaptoethanol; 50% glycerol) were added. The volume was made up to 45μl with water. The tube was spun in a microcentrifuge for a few seconds and 5μl of α32P-dCTP (activity 3000Ci/mol) added. This was incubated at 37°C for 10 minutes and then placed on ice. Spin columns were prepared by placing packing a 1ml syringe with 5% G50-fine sephadex (Sigma Bio Sciences, St. Louis, Missouri, USA). Unincorporated dCTPα32P would be trapped between the sephadex beads within the column. The labelled probes were removed from the ice, loaded onto the columns and spun at 35 000 xg in a microcentrifuge for 5 minutes. The flow-through which was collected in 1.5ml tubes, checked with a Geiger counter, and placed on ice.

2.3.3.5) Hybridisation

The probe was hybridised to the membrane using Quikhyb hybridisation solution (Stratagene Cloning Systems, La Jolla, California, USA). A piece of nylon mesh (21cmx21cm) was rolled with the membrane to prevent the membrane from coming into contact with itself. This was placed inside a hybridisation bottle, rinsed with ddH2O and capped. The Quikhyb hybridisation solution was placed at 68°C for 10 minutes and shaken well. The membrane was pre-hybridised for 20 minutes in 12ml of Quikhyb solution, on a rotating stand (to ensure the membrane was covered in Quikhyb at all times) in a Hybaid micro-oven set at 68°C. "AC" or "DG"
labelled probe was mixed with 100μl of 10mg/ml salmon sperm DNA (Appendix A), boiled for 2 minutes and placed on ice. After pre-hybridisation the probe was mixed into the Quikhyb solution surrounding the membrane. Hybridisation was carried out at 68°C, for 1 hour, with rotation.

After hybridisation the membrane was washed twice for 15 minutes, with gentle shaking at room temperature, with a 2xSSC buffer; 0.1% SDS wash solution (Appendix A). A further two washes were carried out for 30 minutes each at 60°C with a 0.1xSSC buffer; 0.1% SDS wash solution for a high-stringency wash.

2.3.3.6) Stripping the membrane

Two probes were hybridised separately to the same membrane to avoid ambiguity of results. Thus, before each new hybridisation the membrane had to be stripped of any probe DNA still bound to it from the previous hybridisation. Stripping required that a membrane be placed in a glass dish, covered with a 0.1xSSC; 0.1% SDS wash solution, which had been heated to boiling point, and left for 15 minutes.

This wash was repeated once, followed by the pre-hybridisation step before the next hybridisation.

2.3.3.7) Autoradiography

After hybridisation the membrane was sealed in a plastic bag, placed in an Okamoto
cassette with an intensifying screen with Kodak X-ray film and left at -70°C for 4 days before being developed.

2.3.4) *p53 MUTATION ANALYSIS*

PCR was carried out using published primer sequences F3 and R3 (Bressac *et al.*, 1991).

2.3.4.1) **PCR Conditions**

The PCR reaction consisted of, at final volume: 100ng DNA; 2.5U of *Taq* DNA Polymerase (Promega); 1x NH₄ reaction buffer; 1mM MgCl₂; 0.8mM each of dCTP, dATP, dGTP, dTTP; 50pmol of each primer; in a total volume of 50μl. Samples were amplified in a Hybaid thermal reactor for 30 cycles of: denaturation at 94°C for 15 seconds, annealing at 56°C for 15 seconds and an extension step at 72°C for 30 seconds. Negative water controls were used in each PCR run to ensure that there was no contamination of the PCR products.

2.3.4.1.1) **Agarose gel electrophoresis of PCR product**

The 110bp PCR product was sized on 2% ethidium bromide-stained (B/M) agarose gels against a 100bp DNA ladder (Promega).
2.3.4.2) Restriction enzyme digest of PCR product

A G to T transversion at the third base of codon 249 was detected by the presence or absence of a HaeIII restriction site.

The restriction digest consisted at final volume of: 40μl of PCR product; 1U of HaeIII (B/M); 5μl buffer M; in a final volume of 50μl. Digestion was allowed to proceed for 1 hour at 37°C, or until complete, and the enzyme was inactivated by heating to 65°C for 10 minutes. A 5μl aliquot of the digest mix was electrophoresed on a 2% agarose gel to check that digestion had occurred. If the digestion was not complete, samples were re-digested with a further 1U of HaeIII for 2 hours.

2.3.4.2.1) Agarose gel electrophoresis of digested PCR product

Digested PCR products were electrophoresed at 100V on 3% composite gels (1:1; D1 LE:MS8 agarose) and scored as to the presence or absence of the mutation.

2.3.4.3) Confirmation by sequencing

The labelling reaction mix differed from the standard sequencing mix (Section 2.3.1.3) as follows: addition of MN reaction buffer (0.04M Sodium Isocitrate; 0.03M MgCl₂); 2μl 0.01M labelling mix (dGTP) 5xconcentrate diluted 1:15 (7.5μM dGTP, 7.5μM dCTP, 7.5μM dTTP). Samples were labelled at room temperature for 2 minutes and placed on ice for 15 minutes. PCR products were
sequenced bidirectionally both upstream and downstream in separate reactions. All samples shown by digestion to have the codon 249 mutation were sequenced to confirm the presence of the mutation. Samples where the result of the digest was indeterminate were also sequenced along with several control samples not containing the mutation.

2.3.5) Statistical Tests for the Allele Studies on the Random Black Population.

The Hardy Weinberg Equilibrium test (HWE) (Guo and Thompson 1992) was carried out from the allele frequency data obtained for the W77, D13S137, D13S120, D13S127, D17S855, D16S301 and D16S260 loci for the random black population. Polymorphism Information Content (PIC) (Botstein et al., 1980), and Diversity (H) (Nei 1971) were also calculated for each locus.

2.3.5.1) Hardy Weinberg Equilibrium (HWE)

Consistency with Hardy Weinberg expectation was checked by the exact method of Guo and Thompson (1992). The test for Hardy-Weinberg Equilibrium (HWE) for the true allele distribution can be carried out as a goodness of fit test. The Hellinger distance is a measure of the difference between two population samples. Its significance was assessed by bootstrapping in the following manner: two population samples were pooled and two pseudo population samples (of the same size as the actual population samples) were drawn from the pool (with replacement). The Hellinger distance between the pseudo population samples was
then measured and the results compared to the original Hellinger distance. This procedure was repeated 1000 times in order to estimate the probability that the initial Hellinger distance represented nothing more than a sampling error. If less than 5% of the bootstrap Hellinger distances were greater than the original Hellinger distance, the difference between the two population samples was considered to be significant (i.e., values of \( p < 0.05 \) were considered significant). This test is preferable to the \( X^2 \) goodness-of-fit test when sample sizes and/or some genotype frequencies are small (Guo and Thompson 1992).

2.3.5.2) Polymorphism Information Content (PIC), Diversity (H), and Observed Heterozygosity (OH) Values

The PIC of each system was estimated as described by Botstein et al. (1980). A PIC value is an indication of how informative a polymorphism is likely to be in linkage studies (the higher the value of P the more informative a locus).

The diversity (H), which is the expected heterozygosity of each system, was calculated as described by Nei (1987). Loci with high H values are useful in population and forensic studies.

Observed Heterozygosity is a value calculated from the genotypic data collected and is expected to be similar to the Diversity if the population is in Hardy-Weinberg Equilibrium.
CHAPTER 3
RESULTS

3.1) Gene Studies in HCC

3.1.1) VNTR Analyses

Microsatellite loci were amplified by PCR and the radioactively labelled products separated on polyacrylamide gels and viewed by autoradiography. Band mobility shifts between tumour and the matched non-tumorous liver DNA were scored as a change in allele repeat number, and homozygotes were sequenced to determine the exact number of repeats present in each case (Fig. 2). Heterozygotes were scored for number of repeats by comparison to homozygote controls for which exact repeat numbers were known. The microsatellite loci D17S846, RB(1.20) and the ALU repeat within the p53 gene were resolved on agarose gels because polyacrylamide gel electrophoresis generated bands too blurred to score. These loci were scored in terms of a change in allele repeat number only, which was determined by measurable/symmetric mobility shifts between tumour and the matched non-tumorous liver DNA (section 3.1.1.8). LOH was characterised by the disappearance of one band or a considerable (>80%) decrease in band intensity in heterozygotes. LOH excluded expansion and/or contraction of microsatellite sequences in the tumour DNA. Samples appearing to have LOH were duplicated for confirmation.
Figure 2: Autoradiograph of a polyacrylamide gel showing sequencing profile of the antisense strand of tumour tissue (T56) and non-tumorous tissue (NT56) DNA (left panel), at the D16S301 AC microsatellite repeat locus. The repeat consists of (TG)$_8$ interrupted by (TA)$_1$ and a further (TG)$_{12}$ (Composite photograph).
3.1.1.1) *WT1* gene

Alleles were designated according to number of repeats, e.g. allele number 10 consisted of 10 repeats of the sequence AC. Alleles ranged from 8 repeats to 22 repeats with 16 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 19 individuals, 13 of these were informative. No PCR product was obtained for non-tumorous liver in 4 individuals (subjects 6, 17, 38, 54), and for tumour tissue in 3 individuals (subjects 16, 19, 42). No PCR product was obtained in 2 individuals (subjects 7, 41) for either tumour or non-tumorous liver. One out of 19 individuals (subject 8) had an increase in repeat number in one allele at this locus and a decrease in repeat number in the second allele between tumour and non-tumorous liver (Tables 3.1 and 3.4). One out of 19 individuals (subject 18) had either an increase in repeat number in one allele at this locus or a possible LOH.

3.1.1.2) *BRCA1* gene/chromosome 17

3.1.1.2.1) D17S855

Alleles were once again designated according to number of repeats. At this locus, alleles ranged from 10 repeats to 25 repeats with 22 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 18 individuals, 15 of these were informative. No PCR product was obtained for non-tumorous liver in 5 individuals (subjects 4, 6, 17, 18, 38), and for tumour tissue in 3 individuals (subjects 16, 19, 42). No PCR product was obtained in 2 individuals (subjects 7, 41) for either tumour or non-
tumorous liver. Out of 15 individuals, 1 (subject 8) had either an increase in repeat number in both alleles, or an increase in repeat number in one allele coupled with a decrease in repeat number in the second allele (Tables 3.1 and 3.4).

3.1.1.2.2) D17S846

As these samples were run on agarose gels, the exact number of repeats for each allele could not be determined as resolution was insufficient to accurately determine allele sizes. Samples were scored in terms of band shifts between tumour and matched non-tumorous liver DNA (see section 3.1.1.8), and alleles were designated a number from 1-7. The most common allele was allele 3. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 17 individuals, 14 of these were informative. No PCR product was obtained for non-tumorous liver in 1 individual (subject 38), and for tumour tissue in another individual (subject 42). No PCR product was obtained in 7 individuals (subjects 7, 14-19, 41) for either tumour or non-tumorous liver. Out of 16 individuals, 1 individual (subject 8) already had instability at this locus in the non-tumorous tissue, where four alleles were present. Further instability occurred in the tumour tissue, where one of the four alleles was lost, and two of the remaining three suffered changes in repeat number (Tables 3.1 and 3.5; Fig. 3).
Figure 3: Visualisation of PCR products, encompassing the D17S846 locus, on a 4% composite agarose gel. Lane 1: 100 bp molecular weight marker (MWM); lane 2: 50bp MWM; lanes 3 & 4: tumour (T3) and non-tumorous (NT3) tissue DNA; lanes 5 & 6: tumour (T4) and non-tumorous (NT4) tissue DNA; lane 7: 50bp MWM; lanes 8 & 9: tumour (T8) and non-tumorous (NT8) tissue DNA. Allelic imbalance at this locus is evident for individual 8 already in the NT tissue, where four alleles are present. Further instability is evident in the T tissue, where one allele has been lost and two of the remaining three have suffered additional changes in repeat number. (Composite photograph).
3.1.1.3) **BRCA2 gene/chromosome 13**

3.1.1.3.1) **D13S120**

At this locus, alleles ranged from 9 repeats to 29 repeats with 23 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 22 individuals, 20 of these were informative. No PCR product was obtained for non-tumorous liver in 3 individuals (subjects 17, 38, 41), and for tumour tissue in 2 individuals (subjects 19, 42). Out of 22 individuals, 1 individual (subject 8) had a decrease in repeat number in one allele while a second individual (subject 18) had an increase in repeat number in one allele. A third individual (subject 16) either had an increase in repeat number in one allele at this locus or showed LOH (Tables 3.1 and 3.4).

3.1.1.4) **BRCA2/RB1 genes/chromosome 13**

3.1.1.4.1) **D13S127**

At this locus, alleles ranged from 7 repeats to 18 repeats with 15 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 22 individuals, 14 of these were informative. No PCR product was obtained for non-tumorous liver in 2 individuals (subjects 17, 38), and for tumour tissue in 1 individual (subject 42). No PCR product was obtained in 1 individual (subject 41) for either tumour or non-tumorous liver tissue. Out of 21 individuals, 1 individual (subject 8) had a decrease in repeat number in one allele. Another individual (subject 14) had either a decrease in repeat number in one allele at this locus, or LOH, and 1 individual (subject 16)
Figure 4: Autoradiograph showing PCR product, encompassing the D13S127 \((AC)_n\) repeat locus, on a polyacrylamide gel. Lanes left to right: individuals 1-6, 8, 14, and 16 tumour (T) and non-tumorous (NT) tissue DNA. LOH at this locus is evident in individual 14, and LOH/allelic imbalance in individual 8. LOH = loss of heterozygosity. (Composite photograph).
had either an increase in repeat number in one allele at this locus, or LOH (Tables 3.1 and 3.4; Fig. 4).

The \textit{BRCA2} gene lies between markers D13S120 and D13S127. In one individual (subject 16) LOH appeared to have occurred at both these loci. Should this be the case, the \textit{BRCA2} gene is likely to have been lost in this individual. Southern blotting was to be attempted in order to confirm LOH at the \textit{BRCA2} locus for the two individuals (subjects 14, 16) where possible LOH had occurred (VNTR studies) (see section 3.1.2.1).

3.1.1.5) \textit{RBI} gene/chromosome 13  

3.1.1.5.1) D13S137

At this locus, alleles ranged from 15 repeats to 24 repeats with 17 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 17 individuals, 12 of these were informative. No PCR product was obtained for non-tumorous liver in 5 individuals (subjects 6, 14, 40, 41, 51), and for tumour tissue in 2 individuals (subjects 19, 42). No PCR product was obtained in 3 individuals (subjects 7, 17, 18) for either tumour or non-tumorous liver tissue. Out of 17 individuals, 1 individual (subject 16) had an increase in repeat number in one allele and a decrease in repeat number in the other allele (Tables 3.1 and 3.4).
3.1.1.5.2) RB(1.20)

As these samples were run on agarose gels, the exact number of repeats for each allele could not be determined as resolution was insufficient to accurately determine allele sizes. Samples were scored in terms of band shifts between tumour and matched non-tumorous liver DNA (see section 3.1.1.8), and alleles were designated a number from 1-8. The most common allele was allele 4. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 18 individuals, 10 of these were informative. No PCR product was obtained for non-tumorous liver in 3 individuals (subjects 6, 17, 18), and for tumour tissue in 3 individuals (subjects 16, 39, 42). No PCR product was obtained in 4 individuals (subjects 7, 19, 38, 41) for either tumour or non-tumorous liver tissue. Out of 18 individuals, 1 individual (subject 8) had either an increase in repeat number in one allele or LOH at this locus (Tables 3.1 and 3.5; Fig. 5).

3.1.1.6) p53 gene

These samples were run on agarose gels, thus the exact number of repeats for each allele could not be determined. Samples were scored in terms of band shifts between tumour and matched non-tumorous liver DNA (see section 3.1.1.8), and alleles were designated a number from 1-5. There were 5 alleles at this locus. The most common allele was allele 2. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 20 individuals, 11 of these were informative. No PCR product was obtained for tumour tissue in 1 individual (subject 16) and for non-tumorous tissue in another individual (subject 17). No PCR product was
Figure 5: Autoradiograph showing PCR product, encompassing the *RB1* [RB(1.20) CT repeat] locus, on a 4% composite agarose gel. Lanes left to right: individuals 1-4, 8, and 14 non-tumorous (NT) and tumour (T) liver tissue DNA. LOH at this locus is evident in individual 8, where there has been a loss of one allele.
Figure 6: Visualisation of PCR products, encompassing the p53 ALU repeat, on a 4% composite agarose gel. Panels 1-4: tumour (T) and non-tumorous (NT) liver tissue DNA. Panel 5 lanes 1 & 2: tumour (T) and non-tumorous (NT) liver tissue DNA; lane 3: 100 bp molecular weight marker (MWM). A change in repeat number at this locus is visible in individual 18, where there has been an increase in repeat in one allele in the tumour tissue. (Composite photograph).
Figure 6: Visualisation of PCR products, encompassing the $p53$ ALU repeat, on a 4% composite agarose gel. Panels 1-4: tumour (T) and non-tumorous (NT) liver tissue DNA. Panel 5 lanes 1 & 2: tumour (T) and non-tumorous (NT) liver tissue DNA; lane 3: 100 bp molecular weight marker (MWM). A change in repeat number at this locus is visible in individual 18, where there has been an increase in repeat in one allele in the tumour tissue. (Composite photograph).
obtained in 4 individuals (subjects 7, 38, 41, 42) for either tumour or non-tumorous liver tissue. Out of 20 individuals, 2 individuals (subjects 8, 18) had an increase in repeat number in one allele respectively at this locus (Tables 3.1 and 3.5; Fig. 6).

3.1.1.7) E-cadherin gene/chromosome 16

3.1.1.7.1) D16S301

At this locus, alleles ranged from 15 repeats to 29 repeats with 18 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 20 individuals, 11 of these were informative. No PCR product was obtained for non-tumorous liver in 2 individuals (subjects 6, 42), and for tumour tissue in 4 individuals (subjects 14, 17, 19, 48). No PCR product was obtained in 3 individuals (subjects 18, 40, 41) for either tumour or non-tumorous liver tissue. Out of 20 individuals, 1 individual (subject 8) had an increase in repeat number in one allele. Another individual (subject 16) had an increase in repeat number in one allele and a decrease in repeat number in the other allele (Tables 3.1 and 3.4).

3.1.1.7.2) D16S260

At this locus, alleles ranged from 18 repeats to 29 repeats with 20 repeats being the most common. Of the total of 25 individuals, PCR product was obtained for non-tumorous liver in 21 individuals, 13 of these were informative. No PCR product was obtained for non-tumorous liver in 3 individuals (subjects 17, 38,
41), and for tumour tissue in 2 individuals (subjects 19, 48). No PCR product was obtained in 1 individual (subject 42) for either tumour or non-tumorous liver tissue. Out of 21 individuals, 2 individuals (subjects 16, 18) had an increase in repeat number in one allele respectively at this locus (Tables 3.1 and 3.4).

3.1.1.8) Blurring of the Bands on Polyacrylamide Gels for the D17S846, RB(1.20) and p53 Loci.

The blurring of DNA bands on polyacrylamide gels for the loci D17S846, a GGAA repeat region, RB(1.20), a CTTT[T] repeat region, and an ALU repeat region within the p53 gene, was thought to be caused by the secondary structure of the DNA during electrophoresis. The problem was overcome by running the PCR products on 4% composite agarose gels (Appendix A).

3.1.1.9) Genomic instability

The VNTR changes discussed above were confined to 4 of the 25 individuals studied (subjects 8, 14, 16, 18) (Tables 3.1 and 3.4). In 3/22 (14%) of these (subjects 8, 16, 18) this is likely to be indicative of genomic/microsatellite instability, under the definition: changes in microsatellite repeats at two loci or more (Kazachkov et al., 1998). Two of these individuals showed microsatellite instability at 3 loci, and 1 individual at 7 loci.
41), and for tumour tissue in 2 individuals (subjects 19, 48). No PCR product was obtained in 1 individual (subject 42) for either tumour or non-tumorous liver tissue. Out of 21 individuals, 2 individuals (subjects 16, 18) had an increase in repeat number in one allele respectively at this locus (Tables 3.1 and 3.4).

3.1.1.8) Blurring of the Bands on Polyacrylamide Gels for the D17S846, RB(1.20) and p53 Loci.

The blurring of DNA bands on polyacrylamide gels for the loci D17S846, a GGAA repeat region, RB(1.20), a CTTT[T] repeat region, and an ALU repeat region within the p53 gene, was thought to be caused by the secondary structure of the DNA during electrophoresis. The problem was overcome by running the PCR products on 4% composite agarose gels (Appendix A).

3.1.1.9) Genomic instability

The VNTR changes discussed above were confined to 4 of the 25 individuals studied (subjects 8, 14, 16, 18) (Tables 3.1 and 3.4). In 3/22 (14%) of these (subjects 8, 16, 18) this is likely to be indicative of genomic/microsatellite instability, under the definition: changes in microsatellite repeats at two loci or more (Kazachkov et al., 1998). Two of these individuals showed microsatellite instability at 3 loci, and 1 individual at 7 loci.
3.1.2) SSCP Analysis of the BRCA2 Gene

Two polymorphic regions within the BRCA2 gene were amplified by PCR and the products resolved on mutation detection enhancement (MDE) gels (Figs. 7 & 8). Band mobility shifts between tumour and matched non-tumorous liver DNA were scored as changes in the DNA sequence which induced single stranded conformational polymorphisms in the DNA when electrophoresed on the MDE gel. The nature of the changes in SSCP are not associated with specific mutations, and to establish the exact cause of the mobility shifts, the changes have to be analysed by sequencing. Band shifts between T and NT DNA in the BRCA2 gene were detected at position 508 in 3 individuals with the use of SSCP (Table 3.1). In these individuals, the SSCP mobility shifts between the tumour and non-tumorous liver tissues could be indicative of mutations in an allele of the tumour suppressor gene. However the BRCA2 gene was not sequenced and the causes of the mobility shifts are unknown.

3.1.2.1) Southern blotting

It was decided to screen those individuals in whom possible mutations had been detected by SSCP for LOH by this method. Two cDNA probes were obtained (Fig. 9), and radioactively labelled for hybridisation (section 2.3.3.4) to a series of random DNA samples, from black subjects, digested with a number of restriction enzymes (Fig. 10) (section 2.3.3.1). This was done in order to establish which enzyme would reveal a DNA polymorphism within the BRCA2 gene. Upon radiography, however, no probe signal was visible. Blots were stripped and re-probed twice, and a second digest and blotting were performed, but the problem
Figure 7: Autoradiograph showing SSCP analysis at position 97 (A→G)(PM - 3668) of the \textit{BRCA2} gene (PCR product on MDE gel). Lane 1: tumour (T42) tissue DNA PCR positive control; lanes 2-13: tumour (T) and non-tumorous (NT) liver tissue DNA. No mobility changes were detected. SSCP = single stranded conformational polymorphism. (Composite photograph).
Figure 8: Autoradiograph showing SSCP analysis at position 508 (T→C) (PM - 3045 of the BRCA2 gene (PCR product on MDE gel). Lanes 1-14: tumour (T) and non-tumorous (NT) liver tissue DNA. No mobility changes were detected for these individuals. SSCP = single stranded conformational polymorphism.
Figure 9: Visualisation of pUC18 vector (with modified polylinker) containing two partial BRCA2 cDNA clones (S. Tavtigian - Myriad Genetics), termed “AC” and “DG”; restricted with endonucleases; and electrophoresed on a 0.8% agarose gel. These clones were later used for hybridisation after Southern blotting. Lanes 1 & 2: pUC18 vector containing “AC” clone - lane 1: restricted with endonucleases SalI and EcoRV, “AC” clone 5.3 kilobases (kb) - lane 2: unrestricted; lane 3: 1kb “DNA Plus” molecular weight marker (MWM); lanes 4 & 5: pUC18 vector containing “DG” clone - lane 4: restricted with SacI and SalI endonucleases, “DG” clone 4.4kb - lane 5: unrestricted.
Figure 10: DNA from a random black population restricted with endonucleases, and visualised on a 0.8% agarose gel, later used in Southern blotting. Lanes 1 & 22: 1 kilobase (kb) molecular weight marker (MWM); lanes 3-6, 8-11, 13-16, 18-21: DNA of individuals ZY 43, PED33, TSO14, and NDE6 restricted with the endonucleases BamHI, EcoRI, SalI, and TaqI respectively.
persisted. It was established that DNA digests were successful, transfer of DNA to the membrane was successful, and denaturation of both probe and sample DNA had been carried out. The final signal was however, not visible (see discussion).

The exact nature of the problem was not established. However, stringency of washes may have been too high, or there may have been a problem with the DNA sequence of the probe, which was cultured before use. Time constraints impeded further optimisation and sequencing was attempted instead.

3.1.3) p53 gene codon 249 analysis

The p53 codon 249 AGG > AGT mutation was detected using PCR-RFLP analysis, and confirmed by sequencing. In certain tumour samples partial HaeIII digestion was observed, even after 3 independent digests with different PCR experiments of the same individual. This appeared to be indicative of contaminating normal DNA in the tumour samples of these individuals. The presence of the mutation in these samples was then confirmed by sequencing, where it was established that the 'partial digest' was a gel artefact in all cases except in 2 individuals. The mutation was found in the non-tumorous liver of 1 individual (individual 24), and in both the tumour and non-tumorous liver tissues of 2 individuals (individuals 6 & 42) (Table 3.3). Excluding these two individuals, the p53 codon 249 AGG > AGT mutation was detected in the tumour tissues of 4 individuals (Tables 3.1 and 3.3; Figs. 11, 12, 13).
Figure 11: PCR product generated from the codon 249 region of the p53 gene, digested with the endonuclease HaeIII, and electrophoresed on a 2% agarose gel. PCR product from chromosomes with the AGG→AGT mutation at codon 249 is 110 base pairs (bp) in size. PCR product generated from wild-type chromosomes yields, when restricted with HaeIII, two fragments of 75bp and 35bp respectively. Individuals 3 and 6, lanes 1-4, appear to have the mutation in both their tumour (T) and non-tumorous (NT) tissues. Individual 3 was later found by sequencing, not to have the mutation in the non-tumorous tissue. This 110bp band was thus due to the digestion of the PCR product with HaeIII having been terminated before completion. Individual 6 was shown by sequencing to have the mutation in both the tumour and non-tumorous tissue. This was thought to be due to contamination (sections 3.1.3 and 4.1.3).
Figure 12: Autoradiograph of polyacrylamide gels showing sequencing profile of the codon 249 region of the *p53* gene. Panels left to right: sequence of the anti-sense (A) strand of individuals 6, 8 and 24 tumour (T) and non-tumorous (NT) tissue DNA. The AGG → AGT mutation at codon 249 is visible in individuals 6 and 8 - tumour tissue (T6A; T8A). (Composite photograph).
Figure 12: Autoradiograph of polyacrylamide gels showing sequencing profile of the codon 249 region of the \textit{p53} gene. Panels left to right: sequence of the anti-sense (A) strand of individuals 6, 8 and 24 tumour (T) and non-tumorous (NT) tissue DNA. The AGG$\rightarrow$AGT mutation at codon 249 is visible in individuals 6 and 8 - tumour tissue (T6A; T8A). (Composite photograph).
Figure 13: Autoradiograph of polyacrylamide gels showing sequencing profile of the codon 249 region of the p53 gene. Panels left to right: sequence of the sense (S) strand of tumour (T) and non-tumorous (NT) tissue DNA. The AGG→AGT mutation at codon 249 is visible in the tumour tissue of individuals 6 and 8 (T6S; T8S). These individuals are heterozygous for the mutation and thus the wild-type AGG sequence is also visible. The two G bands have, however, been compressed into one as a result of a gel artefact (section 4.1.3.1). The compression results in the uneven spacing of the bands (CCC) directly ahead causing the T band to lie adjacent to the C band. (Composite photograph).
Upon gel electrophoresis of sequenced PCR products for the \textit{p53} codon 249 AGG \textit{\Leftrightarrow} AGT mutation, a curious gel artefact was observed. Due to the formation of a mini-hairpin secondary structure in this region in wild type chromosomes, the sequence of the sense strand appeared to be missing a G at the third base of codon 249 (Fig. 13).

The complementary C base was present when the antisense strand of the same individual was sequenced (Fig. 12). Furthermore, when the PCR product was digested with the endonuclease \textit{HaeIII}, which recognises the cutting site GGCC, the DNA was cut, indicating the presence of the "missing" G (Fig. 13).

3.1.4) HBV Status

Six individuals were found to be super carriers, being HBsAg-positive and HBeAg-negative, while 2 individuals were infected with HBV at the time the tissue was removed, being HBsAg positive. Eight individuals showed evidence of recent infection and acquired immunity, being anti-HBs positive and anti-HBc positive. The HBV status of the remaining individuals was unknown (Table 3.2).

3.1.5) Gene Studies in Random Black Subjects/Polymorphism Frequenc
cies

Allele frequencies were determined for the \textit{WT1} (AC internal repeat), \textit{RB(1.20)}, D13S137, D13S127, D17S855, D16S301 and D16S260 loci. Only alleles smaller than
10 or larger than 29 simple repeats were observed. Allele distribution patterns are reflected in histograms/graphs Figs 14 - 20.

3.1.5.1) WT1 internal AC repeat

Alleles clustered between 13 and 17 repeats with allele 14 being the most common (0.45), followed by alleles 15 (0.21), 13 (0.16) and 16 (0.13), with allele 17 being the least common (0.05). The larger alleles (> 17) were entirely absent. Allele frequencies followed a distribution skewed to the right (Tables 3.6 and 3.7 and Fig. 14). The observed heterozygosity was 0.58, the PIC value 0.67, and the diversity 0.73 (Table 3.8).

3.1.5.2) D13S137

In the literature 12 alleles have been reported at this locus (Petrukhin et al., 1993). We found 9 alleles (Tables 3.6 and 3.7) which followed a bimodal distribution with alleles 17 (0.2) and 24 (0.2) being the most common alleles (Fig 15), followed by alleles 16 (0.18), 18 (0.15), 23 (0.1), 19 (0.08), 20 (0.05), 21 and 22 (both at 0.03). Alleles below 15, and above 24 repeats were absent. The observed heterozygosity was 1, the PIC value 0.83, and the diversity 0.89 (Table 3.8).

3.1.5.3) D13S120

Bowcock et al. (1993) have reported 10 alleles at his locus. We found 11 alleles at this locus (Table 3.6 and 3.7), and these were randomly distributed (Fig 16). The most common allele was allele 25 (0.29) followed by alleles 23 (0.21), 17
(0.12), 11 (0.09), 21 and 22 (0.06), 13, 24, 26 and 29 (all at 0.03). The observed heterozygosity was 0.94, the PIC value 0.82, and the diversity 0.86 (Table 3.8).

3.1.5.4) D13S127

Of the 6 alleles reported in the literature (Bowcock et al., 1993) at this locus we found 5. These followed an almost normal distribution, clustering together mainly between alleles 15 and 18. Alleles below 12 and above 19 repeats were absent (Tables 3.6 and 3.7 and Fig. 17). The most common allele was allele 16 (0.53) followed by alleles 17 (0.21), 15 (0.16), and 13 (0.08), with allele 18 (0.03) being the least common. The observed heterozygosity was 0.68, the PIC value 0.6, and the diversity 0.66 (Table 3.8).

3.1.5.5) D17S855

We found 9 alleles clustering mainly between 16 and 24 repeats (Table 3.6 and 3.7), following an almost bimodal distribution. In the literature 7 alleles have been reported (Anderson et al., 1993). Allele 20 (0.28) was the most common allele, followed by alleles 18 (0.2), 23 (0.1), 21 and 22 (0.13), 24 (0.08), and 16 (0.05) (Fig 18). The least common allele was allele 19 (0.03). Allele 10 was rarely present. The observed heterozygosity was 0.95, the PIC value 0.81, and the diversity 0.88 (Table 3.8).
3.1.5.6) D16S301

This locus was reported by Thompson et al. (1992) to have 6 alleles. We found all 6 in our population and they clustered mainly between 18 and 21 repeats (Tables 3.6 and 3.7 and Fig. 19). The most common allele was allele 19 (0.44) followed by alleles 18 (0.28), 20 (0.16), and 15 (0.06). The least common alleles were alleles 14 and 21 (0.03). Alleles below 13 and above 22 repeats were entirely absent. The observed heterozygosity was 0.69, the PIC value 0.65, and the diversity 0.72 (Table 3.8).

3.1.5.7) D16S260

Weber et al. (1990) reported 4 alleles at this locus. We found 6 alleles with frequencies following a distribution skewed to the right. Allele 19 (0.58) was the most common, followed by alleles 20 (0.14), 21 and 22 (0.11), 18 and 23 (0.03) (Tables 3.6 and 3.7 and Fig. 20). Alleles below 18 repeats and above 23 were absent. The observed heterozygosity was 0.67, the PIC value 0.58, and the diversity 0.63 (Table 3.8).

3.1.5.8) Statistical Test for Hardy Weinberg equilibrium (HWE)

This population was found to be in HWE for all loci, as none of the p values indicated significant departures from expected Hardy Weinberg proportions (ie. all were greater than 0.05) (Table 3.8).
3.1.5.9) Polymorphism information content (PIC), and Diversity (H) Values for All Loci

Very high PIC and H values are evident in this population for all loci studied (Table 3.8). The closer these values are to 1, the more informative the locus will be.
Figure 14: Histogram representing allele frequency distribution for the AC microsatellite repeat within the 3' untranslated region of the WT1 gene. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. Alleles clustered between 13 and 17 repeats with allele 14 (0.45) being the most common. There were 5 alleles in total.
Figure 15: Histogram representing allele frequency distribution for the D13S137 AC microsatellite repeat locus. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. At this locus 9 alleles followed a bimodal distribution with alleles 17 (0.2) and 24 (0.2) being the most common.
Figure 16: Histogram representing allele frequency distribution for the D13S120 AC microsatellite repeat locus. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. At this locus 11 alleles are randomly distributed with allele 25 (0.29) being the most common.
Figure 17: Histogram representing allele frequency distribution for the D13S127 AC microsatellite repeat locus. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. The 5 alleles at this locus followed an almost normal distribution with allele 16 (0.53) being the most common.
Figure 18: Histogram representing allele frequency distribution for the D17S355 AC microsatellite repeat locus. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. The 9 alleles at this locus clustered between 16 and 24 with allele 20 (0.28) being the most common allele.
Figure 19: Histogram representing allele frequency distribution for the D16S301 AC microsatellite repeat locus. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. The 6 alleles at this locus clustered mainly between 18 and 21 repeats with allele 19 (0.44) being the most common.
Figure 20: Histogram representing allele frequency distribution for the D16S260 AC microsatellite repeat locus. Allele frequencies were established with the use of PCR and polyacrylamide gel electrophoresis from a random sample of a southern African black population. The 6 alleles at this locus followed a distribution skewed to the right with allele 19 (0.58) being the most common.
Table 3.1: LOH, SSCP and sequence analysis.

<table>
<thead>
<tr>
<th>Subject number</th>
<th>p53 (AGU)</th>
<th>WT1 (AC)</th>
<th>RB1</th>
<th>VNRs</th>
<th>BRCA2</th>
<th>BRCA1</th>
<th>E-cadherin</th>
<th>BRCA2 SSCP analysis</th>
<th>p53 codon 249</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>NI</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>NI</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>NI</td>
<td>?</td>
<td>NI</td>
<td>-</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>-/-</td>
</tr>
<tr>
<td>16</td>
<td>?</td>
<td>?</td>
<td>Δ</td>
<td>?</td>
<td>+/!</td>
<td>+</td>
<td>?</td>
<td>Δ</td>
<td>-/-</td>
</tr>
<tr>
<td>Sample number</td>
<td>p53 (ALU)</td>
<td>WT1 (AC)</td>
<td>VNTRs</td>
<td>RB1</td>
<td>BRCA2</td>
<td>BRCA1</td>
<td>E-cadherin</td>
<td>BRCA2 SCCP</td>
<td>p53 codon 249</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>24</td>
<td>NI</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>40</td>
<td>NI</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>48</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>?</td>
<td>?</td>
<td>-</td>
</tr>
<tr>
<td>50</td>
<td>NI</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>51</td>
<td>NI</td>
<td>-</td>
<td>?</td>
<td>NI</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>52</td>
<td>NI</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>53</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>54</td>
<td>-</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-/-</td>
</tr>
<tr>
<td>56</td>
<td>NI</td>
<td>-</td>
<td>NI</td>
<td>-</td>
<td>-</td>
<td>NI</td>
<td>NI</td>
<td>NI</td>
<td>-/-</td>
</tr>
</tbody>
</table>
KEY:

LOH : loss of heterozygosity

- : no mutation
  -- > LOH studies: no LOH
  -- > BRCA2 SSCP analysis: no difference in SSCP pattern between
tumour (T) and non-tumorous liver (NT) tissues
  -- > p53 codon 249 AGG → AGT mutation analysis: G → T transversion absent

+ : mutation
  -- > LOH studies: LOH
  -- > BRCA2 SSCP analysis: difference in SSCP pattern between
tumour (T) and non-tumorous liver tissues
  -- > p53 codon 249 AGG → AGT mutation analysis: G → T transversion present

NI : not informative

? : results not obtained (see discussion)

Δ : a change in repeat number between tumour (T) and non-tumorous liver (NT)
tissues in both chromosomes

1/1 : an increase/decrease in repeat number between tumour (T) and non-tumorous liver
(NT) tissues in one chromosome

WT1 : Wilm's tumour gene

RB1 : Retinoblastoma gene

97 and 508 : polymorphic regions in the BRCA2 gene which were analysed using
SSCPs
<table>
<thead>
<tr>
<th>Sample number</th>
<th>HBV status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HBsAg-positive; HBeAg-negative</td>
</tr>
<tr>
<td>2</td>
<td>HBsAg-positive</td>
</tr>
<tr>
<td>3</td>
<td>HBsAg-positive; HBeAg-negative</td>
</tr>
<tr>
<td>4</td>
<td>Anti-HBc positive; Anti-HBs positive</td>
</tr>
<tr>
<td>5</td>
<td>HBsAg-positive; HBeAg-negative</td>
</tr>
<tr>
<td>6</td>
<td>HBsAg-positive; HBeAg-negative</td>
</tr>
<tr>
<td>7</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>8</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>14</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>16</td>
<td>HBsAg-positive</td>
</tr>
<tr>
<td>17</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>18</td>
<td>HBsAg-positive, HBeAg-negative</td>
</tr>
<tr>
<td>19</td>
<td>HBsAg-positive; HBeAg-negative</td>
</tr>
<tr>
<td>24</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>38</td>
<td>?</td>
</tr>
<tr>
<td>39</td>
<td>HBsAg-positive; HBeAg-negative</td>
</tr>
<tr>
<td>40</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>41</td>
<td>Anti-HBs positive; Anti-HBc positive</td>
</tr>
<tr>
<td>42</td>
<td>?</td>
</tr>
<tr>
<td>48</td>
<td>?</td>
</tr>
<tr>
<td>50</td>
<td>?</td>
</tr>
<tr>
<td>51</td>
<td>?</td>
</tr>
<tr>
<td>52</td>
<td>?</td>
</tr>
<tr>
<td>53</td>
<td>?</td>
</tr>
<tr>
<td>54</td>
<td>?</td>
</tr>
<tr>
<td>56</td>
<td>?</td>
</tr>
</tbody>
</table>

**KEY:**
- **HBV status unknown**
- **HBsAg** = hepatitis B virus S antigen
- **HBeAg** = hepatitis B virus E antigen
- **Anti HBs** = antibody to hepatitis B virus S antigen
- **Anti HBc** = antibody to hepatitis B virus C antigen
<table>
<thead>
<tr>
<th>Sample number</th>
<th>HaeIII Digest</th>
<th>Sequencing result</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sense</td>
</tr>
<tr>
<td>T1</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT1</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T2</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT2</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T3</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>NT3</td>
<td>+/-?</td>
<td>CCGGAGG/CC</td>
</tr>
<tr>
<td>T4</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT4</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T5</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT5</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T6</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>NT6</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>T7</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT7</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T8</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>NT8</td>
<td>+/-?</td>
<td>CCGGAGG/CC</td>
</tr>
<tr>
<td>T9</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T10</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T11</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T12</td>
<td>-/-</td>
<td>CCGGAGG/CC</td>
</tr>
<tr>
<td>T13</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>T14</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT14</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T15</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT15</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T16</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>Sample number</td>
<td>HaeIII digest</td>
<td>Sequencing result</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sense</td>
</tr>
<tr>
<td>NT16</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T17</td>
<td>+/-</td>
<td>CCGGAGGG/TCC</td>
</tr>
<tr>
<td>NT17</td>
<td>-/-</td>
<td>CCGGAGGCC</td>
</tr>
<tr>
<td>T18</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT18</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T19</td>
<td>-/-</td>
<td>CCGGAGGCC</td>
</tr>
<tr>
<td>NT19</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T20</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT21</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T22</td>
<td>+/-</td>
<td>CCGGAGGG/TCC</td>
</tr>
<tr>
<td>T23</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T24</td>
<td>-/-</td>
<td>CCGGAGGCC</td>
</tr>
<tr>
<td>NT24</td>
<td>+/-</td>
<td>CCGGAGGG/TCC</td>
</tr>
<tr>
<td>T26</td>
<td>+/-</td>
<td>CCGGAGGG/TCC</td>
</tr>
<tr>
<td>T27</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T28</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T29</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T30</td>
<td>+/-</td>
<td>CCGGAGGCC</td>
</tr>
<tr>
<td>T31</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T32</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T33</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T34</td>
<td>+/-</td>
<td>CCGGAGGG/TCC</td>
</tr>
<tr>
<td>T35</td>
<td>+/-</td>
<td>CCGGAGGCC</td>
</tr>
<tr>
<td>T36</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T37</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T38</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>Sample number</td>
<td><em>HaeIII</em> digest</td>
<td>Sequencing result</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sense strand</td>
</tr>
<tr>
<td>NT38</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T39</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT39</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T40</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT40</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T41</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT41</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T42</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>NT42</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>T43</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T44</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>T46</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T48</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT48</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T50</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT50</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T51</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT51</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T52</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>NT52</td>
<td>-/-</td>
<td>CCGGAGG/CCTCC</td>
</tr>
<tr>
<td>T53</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT53</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T54</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT54</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>T55</td>
<td>+/-</td>
<td>CCGGAGG/TCC</td>
</tr>
<tr>
<td>T56</td>
<td>-/-</td>
<td>NS</td>
</tr>
<tr>
<td>NT56</td>
<td>-/-</td>
<td>NS</td>
</tr>
</tbody>
</table>
KEY:

+  : $G\rightarrow T$ transversion present (PCR product digested with the restriction enzyme $HaeIII$)

-  : $G\rightarrow T$ transeversion absent (PCR product not digested with the restriction enzyme $HaeIII$)

T  : tumour tissue

NT : non-tumorous liver

NS : not sequenced

AGG : sequence of the codon 249 of the $p53$ gene (sense strand)

TCC : sequence of codon 249 of the $p53$ gene (anti-sense strand)

G  : the third base G (in the sense strand) of codon 249 of the $p53$ gene was not visible on sequencing gels due to a gel artefact (section 4.1.3.1), the complementary $T$ of the anti-sense strand was visible
Table 3.4: VNTR analysis for the WT1 gene and D13S137, D13S120, D13S127, D17S855, D16S301, D16S260 loci in HCC subjects.

<table>
<thead>
<tr>
<th>Sample number</th>
<th>WT1</th>
<th>RB1</th>
<th>BRCA2</th>
<th>BRCA1</th>
<th>E-cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>NT</td>
<td>T</td>
<td>NT</td>
<td>T</td>
</tr>
<tr>
<td>1</td>
<td>14/14</td>
<td>14/14</td>
<td>20/23</td>
<td>20/23</td>
<td>25/29</td>
</tr>
<tr>
<td>2</td>
<td>13/17</td>
<td>13/17</td>
<td>17/18</td>
<td>17/18</td>
<td>23/25</td>
</tr>
<tr>
<td>3</td>
<td>14/16</td>
<td>14/16</td>
<td>16/20</td>
<td>16/20</td>
<td>23/27</td>
</tr>
<tr>
<td>4</td>
<td>14/16</td>
<td>14/16</td>
<td>17/19</td>
<td>17/19</td>
<td>18/26</td>
</tr>
<tr>
<td>8</td>
<td>15/15</td>
<td>14/16</td>
<td>17/17</td>
<td>17/17</td>
<td>19/23</td>
</tr>
<tr>
<td>24</td>
<td>16/18</td>
<td>16/18</td>
<td>18/18</td>
<td>18/18</td>
<td>23/26</td>
</tr>
</tbody>
</table>

113
<table>
<thead>
<tr>
<th>Sample number</th>
<th>(WT1)</th>
<th>(RB1)</th>
<th>(BRCA2)</th>
<th>(BRCA1)</th>
<th>(E\text{-cadherin})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>NT</td>
<td>T</td>
<td>NT</td>
<td>T</td>
</tr>
<tr>
<td>38</td>
<td>14/15</td>
<td>?</td>
<td>17/23</td>
<td>17/23</td>
<td>11/13</td>
</tr>
<tr>
<td>40</td>
<td>14/16</td>
<td>?</td>
<td>18/18</td>
<td>?</td>
<td>17/19</td>
</tr>
<tr>
<td>48</td>
<td>15/16</td>
<td>15/16</td>
<td>15/23</td>
<td>15/23</td>
<td>21/23</td>
</tr>
<tr>
<td>50</td>
<td>16/19</td>
<td>16/19</td>
<td>22/24</td>
<td>22/24</td>
<td>9/27</td>
</tr>
<tr>
<td>51</td>
<td>16/18</td>
<td>16/18</td>
<td>15/15</td>
<td>?</td>
<td>21/23</td>
</tr>
<tr>
<td>52</td>
<td>16/17</td>
<td>16/17</td>
<td>15/15</td>
<td>15/15</td>
<td>17/21</td>
</tr>
<tr>
<td>53</td>
<td>16/19</td>
<td>16/19</td>
<td>16/22</td>
<td>16/22</td>
<td>17/23</td>
</tr>
<tr>
<td>54</td>
<td>19/22</td>
<td>?</td>
<td>15/23</td>
<td>15/23</td>
<td>10/25</td>
</tr>
<tr>
<td>56</td>
<td>14/16</td>
<td>14/16</td>
<td>15/23</td>
<td>15/23</td>
<td>17/23</td>
</tr>
</tbody>
</table>

**KEY:**
- ? : PCR was repeatedly unsuccessful (even after three attempts) and no result was obtained.
- \(WT1\) : Wilm's tumour gene.
- \(RB1\) : Retinoblastoma gene.
- Differences are indicated in bold typescript
Table 3.5: VNTR analysis for the *p53* gene and the D17S846, and RB(1.20) loci.

| Sample Number |  |  |  |  |
|---------------|---------------|---------------|---------------|
|               | *p53* ALU repeat | D17S846 (GGAA)<sub>25</sub> | RB (1.20) [CTTT(T)<sub>n</sub>,
<sup>n</sup> = 14-26] |
<p>|               | T  | NT | T  | NT |
| 1             | 2/2 | 2/2 | 1/4 | 1/4 | 4/8 | 4/8 |
| 2             | 2/4 | 2/4 | 3/6 | 3/6 | 3/5 | 3/5 |
| 3             | 2/4 | 2/4 | 4/5 | 4/5 | 3/4 | 3/4 |
| 4             | 2/4 | 2/4 | 3/3 | 3/3 | 4/4 | 4/4 |
| 8             | 3/5 | 3/4 | 2.5/5/6 | 3/4 &amp; 6/8 | 4/4 | 2/4 |
| 17            | ½  | ?  | ?  | ?  | 4/4 | ?  |
| 18            | 2/4 | 2/3 | ?  | ?  | 3/3 | ?  |
| 19            | 2/4 | 2/4 | ?  | ?  | ?  | ?  |</p>
<table>
<thead>
<tr>
<th>Sample number</th>
<th>p53 ALU repeat</th>
<th>D17S846 (GGGA)25</th>
<th>RB (1.20) [CTTT(T)]_n&lt;sup&gt;??&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>NT</td>
<td>T</td>
</tr>
<tr>
<td>24</td>
<td>2/2</td>
<td>2/2</td>
<td>3/4</td>
</tr>
<tr>
<td>38</td>
<td>?</td>
<td>?</td>
<td>3/7</td>
</tr>
<tr>
<td>39</td>
<td>2/4</td>
<td>2/4</td>
<td>2/2</td>
</tr>
<tr>
<td>40</td>
<td>2/2</td>
<td>2/2</td>
<td>3/6</td>
</tr>
<tr>
<td>48</td>
<td>3/3</td>
<td>3/3</td>
<td>1/5</td>
</tr>
<tr>
<td>50</td>
<td>2/2</td>
<td>2/2</td>
<td>3/5</td>
</tr>
<tr>
<td>51</td>
<td>2/2</td>
<td>2/2</td>
<td>3/5</td>
</tr>
<tr>
<td>52</td>
<td>2/2</td>
<td>2/2</td>
<td>3/6</td>
</tr>
<tr>
<td>53</td>
<td>2/4</td>
<td>2/4</td>
<td>2/5</td>
</tr>
<tr>
<td>54</td>
<td>2/4</td>
<td>2/4</td>
<td>4/7</td>
</tr>
<tr>
<td>56</td>
<td>3/3</td>
<td>3/3</td>
<td>2/5</td>
</tr>
</tbody>
</table>

**KEY:**
- ? : results not obtained (see discussion)
Table 3.6: VNTR analysis for the WT1 gene and D13S137, D13S120, D13S127, D17S855, D16S301, D16S260 Loci in Random Black Subjects.

<table>
<thead>
<tr>
<th>RB Sample number</th>
<th>WT1</th>
<th>RB1</th>
<th>BRCA2</th>
<th>BRCA1</th>
<th>E-cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D13S137</td>
<td>D13S120</td>
<td>D13S127</td>
<td>D17S855</td>
</tr>
<tr>
<td>NDE 6</td>
<td>14/14</td>
<td>18/23</td>
<td>17/25</td>
<td>?</td>
<td>19/21</td>
</tr>
<tr>
<td>PED 20</td>
<td>14/14</td>
<td>16/17</td>
<td>?</td>
<td>16/16</td>
<td>18/20</td>
</tr>
<tr>
<td>SOT 54</td>
<td>15/15</td>
<td>17/24</td>
<td>23/25</td>
<td>17/18</td>
<td>18/22</td>
</tr>
<tr>
<td>TSO 14</td>
<td>14/14</td>
<td>17/24</td>
<td>17/25</td>
<td>13/16</td>
<td>10/24</td>
</tr>
<tr>
<td>TSO 8</td>
<td>13/17</td>
<td>16/18</td>
<td>17/25</td>
<td>15/17</td>
<td>18/20</td>
</tr>
<tr>
<td>TSO 41</td>
<td>15/16</td>
<td>18/23</td>
<td>13/23</td>
<td>17/17</td>
<td>20/22</td>
</tr>
<tr>
<td>TSO 40</td>
<td>15/16</td>
<td>16/24</td>
<td>23/25</td>
<td>16/16</td>
<td>20/22</td>
</tr>
<tr>
<td>ZU 13</td>
<td>14/14</td>
<td>19/22</td>
<td>?</td>
<td>15/16</td>
<td>23/20</td>
</tr>
<tr>
<td>SW 50</td>
<td>13/14</td>
<td>17/18</td>
<td>23/25</td>
<td>13/16</td>
<td>16/22</td>
</tr>
<tr>
<td>TSO 13</td>
<td>13/14</td>
<td>16/17</td>
<td>23/25</td>
<td>16/16</td>
<td>21/24</td>
</tr>
<tr>
<td>TSO 6</td>
<td>14/15</td>
<td>17/24</td>
<td>11/27</td>
<td>16/16</td>
<td>20/23</td>
</tr>
<tr>
<td>RB Sample number</td>
<td>WT1</td>
<td>RB1</td>
<td>BRCA2</td>
<td>BRCA1</td>
<td>E-cadherin</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>D13S137</td>
<td>D13S120</td>
<td>D13S127</td>
<td>D17S855</td>
<td>D16S301</td>
</tr>
<tr>
<td>PED 19</td>
<td>?</td>
<td>18/24</td>
<td>?</td>
<td>15/16</td>
<td>18/20</td>
</tr>
<tr>
<td>ZU 63</td>
<td>14/16</td>
<td>16/20</td>
<td>23/25</td>
<td>15/16</td>
<td>16/20</td>
</tr>
<tr>
<td>ZU 62</td>
<td>13/15</td>
<td>17/19</td>
<td>21/22</td>
<td>16/17</td>
<td>18/21</td>
</tr>
<tr>
<td>ZU 61</td>
<td>13/16</td>
<td>17/19</td>
<td>25/26</td>
<td>16/17</td>
<td>18/24</td>
</tr>
<tr>
<td>ZU 52</td>
<td>14/14</td>
<td>21/24</td>
<td>25/27</td>
<td>16/17</td>
<td>20/23</td>
</tr>
<tr>
<td>PED 33</td>
<td>15/15</td>
<td>16/23</td>
<td>11/11</td>
<td>16/16</td>
<td>20/22</td>
</tr>
<tr>
<td>ZU 43</td>
<td>14/14</td>
<td>18/24</td>
<td>21/23</td>
<td>13/17</td>
<td>20/21</td>
</tr>
<tr>
<td>PED 21</td>
<td>13/17</td>
<td>20/23</td>
<td>22/24</td>
<td>15/16</td>
<td>18/18</td>
</tr>
<tr>
<td>TSW 30</td>
<td>14/16</td>
<td>16/24</td>
<td>17/29</td>
<td>15/16</td>
<td>21/23</td>
</tr>
</tbody>
</table>

**KEY:**

- ? : PCR was repeatedly unsuccessful (even after three attempts) and no result was obtained.
- WT1 : Wilm's tumour gene.
- RB1 : Retinoblastoma gene.
- Differences are indicated in bold typescript.
- Nde : Ndebele
- Ped : Pedi
- Soh : Sotho
- Tso : Tsonga
- Zu : Zulu
- Tsw : Tswana

| PED 33 | 15/15 | 16/23 | 11/11 | 16/16 | 20/22 | 18/18 | 19/22 |
| PED 21 | 13/17 | 20/23 | 22/24 | 15/16 | 18/18 | ? | 19/19 |
| TSW 30 | 14/16 | 16/24 | 17/29 | 15/16 | 21/23 | 15/19 | 19/22 |
Table 3.7: Allele frequencies for the loci D13S137, D13S120, D13S127, D17S855, D16S301 and D16S260 for the random black subjects (RB).

<table>
<thead>
<tr>
<th>Allele (No. of repeats)</th>
<th>WT1</th>
<th>D13S137 (RB1)</th>
<th>BRCA2</th>
<th>BRCA1</th>
<th>E-cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D13S120</td>
<td></td>
<td>D17S855</td>
<td>D16S301</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0.09</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0.16</td>
<td>0</td>
<td>0.03</td>
<td>0.08</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0.45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0.21</td>
<td>0</td>
<td>0</td>
<td>0.16</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0.13</td>
<td>0.18</td>
<td>0</td>
<td>0.53</td>
<td>0.05</td>
</tr>
<tr>
<td>17</td>
<td>0.05</td>
<td>0.2</td>
<td>0.12</td>
<td>0.21</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>0.15</td>
<td>0</td>
<td>0.03</td>
<td>0.2</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0.28</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0.03</td>
<td>0.06</td>
<td>0</td>
<td>0.13</td>
</tr>
</tbody>
</table>
## Allele Frequencies

<table>
<thead>
<tr>
<th>Allele (No. of repeats)</th>
<th>WTI</th>
<th>D13S137 (RBI)</th>
<th>BRCA2</th>
<th>BRCA1</th>
<th>E-cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D13S120</td>
<td>D13S127</td>
<td>D17S855</td>
<td>D16S301</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>0.03</td>
<td>0.06</td>
<td>0.13</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>0.1</td>
<td>0.21</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0.2</td>
<td>0.03</td>
<td>0.08</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0.29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0.29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0.06</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>1</td>
<td>1.02</td>
<td>1.01</td>
<td>1.01</td>
<td>1.03</td>
</tr>
</tbody>
</table>

### Key:

- **No. of repeats**: begins at 10
- **WTI**: Wilm's tumour gene.
- **RBI**: Retinoblastoma gene
- **RB**: random blacks.

**Allele frequency** = \( \frac{\text{No. alleles} \times X}{\text{total No. alleles}} \) (where \( X = \text{variable number of repeats} \) )
Table 3.8: Observed heterozygosity, Polymorphism Information Content, Diversity and Hardy Weinberg Equilibrium for Random black subjects.

<table>
<thead>
<tr>
<th>Calculated values</th>
<th>WT1</th>
<th>D13S137 (RB1)</th>
<th>BRCA2</th>
<th>BRCA1</th>
<th>E-cadherin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D13S120</td>
<td>D13S127</td>
<td>D17S855</td>
<td>D16S301</td>
</tr>
<tr>
<td>HWE (p)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.061</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.558</td>
<td>0.579</td>
<td>1</td>
<td>0.941</td>
<td>0.684</td>
<td>0.950</td>
</tr>
<tr>
<td>OH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.670</td>
<td>0.829</td>
<td>0.818</td>
<td>0.601</td>
<td>0.814</td>
<td>0.653</td>
</tr>
<tr>
<td>PIC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.700</td>
<td>0.892</td>
<td>0.861</td>
<td>0.664</td>
<td>0.878</td>
<td>0.722</td>
</tr>
<tr>
<td>H/STD E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.140</td>
<td>0.030</td>
<td>0.144</td>
<td>0.145</td>
<td>0.041</td>
<td>0.152</td>
</tr>
</tbody>
</table>

KEY:

HWE        : Hardy Weinberg equilibrium p value
OH         : observed heterozygosity
PIC        : polymorphism information content
H          : diversity
STD E      : standard error for diversity
Genes which would be good candidates to investigate in the transformation process in human HCC (Slagle et al., 1993), were characterised in this study. Rapid and simple PCR-SSCP, PCR-RFLP techniques were used to screen for mutants, before the use of effective but more time-consuming sequencing and Southern blotting techniques for confirmation of mutations.

4.1) Molecular aetiology of HCC

Chronic infection with the HBV virus is the foremost aetiologic association in southern Africa, and south-eastern Asia. The role of HBV in HCC development is thought to involve both indirect and direct carcinogenic mechanisms. Indirect mechanisms are the contribution of HBV to chronic necroinflammatory parenchymal disease, and to hepatocyte injury over a prolonged period of time - leading ultimately to uncontrolled cell proliferation. Inflammation may also lead to the production of oxygen reactive species which may disrupt the expression of transcriptional regulators such as p53, and to the formation of DNA adducts and consequently mutations. Directly, HBV may influence HCC development by HBV-integration mediated chromosomal alterations such as amplification, translocation, or inverted repeat sequences (Hino et al., 1986). Such alterations may occur in the DNA flanking a tumour suppressor gene resulting in its deletion. If the integration/alteration occurs at a distant site to a tumour suppressor or proto-oncogene, it may still disrupt the relationship between the gene and its far-off regulatory element/s, causing it to be abnormally expressed. It is reckoned that products
of two HBV genes (X gene and preS2/S) may trans-activate cellular genes, resulting in increased cellular proliferation and differentiation or apoptosis (reviewed in Kew 1996a). The HBx protein has also been documented to bind to wild-type p53 protein, thereby inactivating it (Feitelson et al., 1993; Wang et al., 1994). Chori HCV infection is a major risk factor in industrialised countries, and an important factor in southern Africa, and south-eastern Asia, albeit not as important as chronic HBV infection. How HCV contributes to the pathogenes's of HCC is poorly understood. It is almost always associated with necroinflammatory hepatic disease, and thus probably has an indirect carcinogenic effect. HCV does not integrate into the genome of the patient, and its direct effect remains a matter of speculation. It has been proposed that the protein products of the virus may contribute to malignant transformation via proteinase, helicase, and gene regulatory activities. HBV and HCV have a synergistic effect in those patients infected with both viruses. Cirrhosis frequently results from persistent infection with HBV and HCV, and from alcohol abuse, and plays a central role in the pathogenesis of HCC. Increased hepatocyte replication in necroinflammatory hepatic disease may promote tumour development by increasing the likelihood of spontaneous mutations, and HBV DNA integration into the cellular genome (Kew 1996b). Iron overload in liver cells can lead to the increased production of oxygen reactive species which will cause DNA adduct formation, and membrane damage of intracellular organelles (Mandishona et al., 1998). Frequent dietary consumption of the fungal oxin aflatoxin B1 is believed to cause a G>T transversion at the third base of codon 249 of the p53 gene in about 50% of HCCs, and for this reason clustering of this point mutation occurs in HCCs from China and Africa (Hsu et al., 1991; Bressac et al., 1991).
4.2) Gene Studies in HCC

All tumours were at an advanced stage, and because of the paucity of available records (most samples having been obtained at autopsy) no attempt was made to correlate mutations with clinical or other features of the individuals. A great portion of the southern African black population is rural, patients are only seen by a medical practitioner at very advanced stages of the disease, and follow-up on any of them is impossible. Furthermore attempting to draw any conclusions requiring information dependent on stage of HCC development was not possible due to the advanced stage of the tumours.

4.2.1) Genomic instability/LOH

LOH appears to have occurred in 3 individuals at the RB1 (13q14.2), BRCA2 (13q12-13) and WT1 (11p13) loci.

LOH of the RB gene has been documented in HCCs from China (11/34 - 32.4%) (Sheu et al., 1998), Korea (7/21 - 33.3%) (Piao et al., 1997), and Japan (4/21 - 19% and 6/24 25%; 16/20 - 80%) (Murakami et al., 1991; Kuroki et al., 1995), and in an Australian study 4/15 (27%) (Walker et al., 1991). One copy of the RB1 gene was lost in 3/19 (16%) HCCs in this study. This differs from the higher percentages found thus far. Walker et al., also reported LOH of the RB1 gene in their population in HBV negative tumours only. Our tumours showing LOH for RB1 were HBV positive. We presume these to be population differences, as our HCCs were from southern Africa, whereas the samples in the other studies were from Japan and Australia. The RB1 gene may play a
role in some HCCs, although somatic mutation of this gene appears to be rare in HCC (Zhang et al., 1994). As RBI is a tumour suppressor gene, loss of both function of both alleles is required for total gene inactivation and tumour formation. The techniques employed in this study did not allow for the detection of mutations such as small deletions, inversions, or point mutations. The remaining allele in those patients in which LOH was detected may have been wild-type. If however, we assume this allele to have been inactivated, we may propose that RBI had an effect in HCCs studied. When pRB is inactivated unrestricted cell growth occurs (Laiho et al., 1990), which is certainly one of the features of HCC development. The second allele may also have been inactivated by hypermethylation. It can be speculated that the low percentage found with respect to LOH of this gene in our population, may also be due to cells lacking functional pRB having undergone apoptosis at an early stage. Cells in which pRB is non-functional may re-route via p53 for apoptosis (Fan et al., 1996). In contrast to others, no coincident abnormalities of the p53 and RBI genes were found in this study.

The BRCA2 gene has been reported to be a tumour suppressor involved in HCC progression, where mutations were detected in 2/60 (3%) HCCs (Katagiri et al., 1996). LOH in the region of this locus (13q12/13q12.1) has also been reported in 6/9 (67%) HCCs by Walker et al., (1991), and in 5/20 (25%) HCCs by Sheu et al., (1998). BRCA2 is believed to regulate proliferation by functioning at cell cycle checkpoints. In the present study one copy of the BRCA2 gene was lost in 2/20 (10%) HCCs. Our results indicate that BRCA2 may be involved in HCC progression, in accordance with a study by Katagiri et al., (1996). Elimination of the function of the BRCA2 gene may certainly
lead to increased cell division, and the possible unconstrained growth of neoplastic liver cells.

LOH at 11p13 has been reported in 1/14 (7%) HCCs (Wang and Rogler 1988), although it was not confirmed whether the deletions included the WT1 gene. Piao et al., (1997) reported LOH at the WT1 locus in 1/22 (4.5%) HCCs. Our result of 1/13 (8%) HCCs is concordant with these findings. The WT1 gene thus appears to be the tumour suppressor gene targeted in the 11p13 deletions. Whether WT1 plays a significant role in HCC has yet to be established, although as a transcription factor possibly involved in proliferation/differentiation/apoptosis, it could do so.

LOH cases reported could be indicative of 'background' loss rate, which may be expected in any tumour cell by chance, as LOH at each individual marker was <20% (Fearon and Vogelstein 1990). We believe this to be unlikely in the cases of LOH at the BRCA2 gene in individual 16, where both markers flanking the gene were lost, and in individual 18 (LOH WT1 and RB1) where genetic changes appeared to be associated with HCC development (p117). Furthermore, LOH of RB1 and BRCA2 genes, as stated previously, has been documented at percentages above 20 in HCC's.

LOH has been reported in the literature for the p53 gene in 49-69% of HCCs from Japan (Murakami et al., 1991; Oda et al., 1992; Nishida et al., 1993; Teramoto et al., 1994), and in 3/5 (60%) of HCCs from southern Africa (Bressac et al., 1991; Slagle et al., 1991). No LOH was detected for p53 in this study. Inactivation/reduction of the p53 gene

...
gene expression or of its product, by means other than LOH, may have occurred in our population. This was not investigated here. Walker et al., (1991) reported that p53 allele losses occurred only in HBV negative tumours in their study, and no p53 mutations were present in HBV-positive HCCs. They proposed that the inverse correlation between viral infection and p53 mutation implied a mechanism whereby the p53 protein was inactivated by complex association with virus-encoded protein products. Such associations have been documented in the literature (Hosono et al., 1993). In our study for all those samples for which HBV status was known, most were HBV positive, or anti-HBs positive. Our p53 protein may thus have been inactivated by association with virus-encoded proteins or the gene itself could be mutated in ways other than LOH, such as point mutations and small deletions (<50bp). Mutations such as these could have been left undetected in this study, as the entire p53 gene was not sequenced, but only a small (110bp) PCR product generated from a region of the p53 gene encompassing codon 249.

Cells with no p53 exhibit a prolonged lifespan and selection advantage over non-neoplastic cells, thus increasing their risk for malignant transformation. We were not able to determine whether in our population the LOH of the wild-type allele was the primary or secondary event as we did not have tumours of different stages for any individual. LOH has been proposed to preceed mutation of the remaining allele in Chinese tumours, but the reverse has been proposed for Japanese tumours (Oda et al., 1992).

No LOH or changes in repeat number were documented for the BRCA1 gene in this study. To our knowledge, LOH of this gene has been reported only once, in a study from Korea by Piao et al. (1997), where 3.26 (11.5%) of HCCs showed LOH at this
locus. \textit{BRCA1} is proposed to be a tumour suppressor whose decreased expression leads to increased cell growth in breast cancer and which is believed to possibly regulate proliferation by functioning at cell cycle checkpoints (Chen \textit{et al}., 1996). As LOH of this gene has been reported only in one study, and this at 11.5\%, it may be that this may only reflect a 'background' loss rate, which may be expected by chance in any tumour cell. This has been reported for colorectal cancers (Fearon and Vogelstein 1990).

LOH in the region of the \textit{E-cadherin} locus (16q22.1) has been reported in 31/36 (86\%) HCCs in a Japanese study (Tsuda \textit{et al}., 1990), and in 18/28 (64\%) HCCs in a Chinese study (Slagle \textit{et al}., 1993). No LOH was found for this locus in our study. \textit{E-cadherin} may be involved in HCC multicentric development and formation of intrahepatic metastases (Lin \textit{et al}., 1987). As we had only one tumour sample available for each patient, and no information as to whether they had exhibited any intra or extra hepatic metastases, our speculations are limited. Furthermore, although all the HCCs used in this study were in advanced stages, it was not established whether they were highly undifferentiated. There may have been retention of \textit{E-cadherin} expression in these samples, and no loss of intercellular adhesion. Alternately, \textit{E-cadherin} may have been inactivated in our tumours via promoter methylation or chromosomal rearrangements (Kanai \textit{et al}., 1997), which would not have been detected by the methods used in this study.

Changes in repeat number but not LOH was noted at the \textit{p53} and D13S120 loci in 2/25 individuals, at the \textit{W7I}, D13S137, D16S120, D13S127, D17S855, D17S849, and
D16S301 loci in 1/25 individuals. Genetic instability was noted in 3 individuals. In these samples, it was not possible to determine whether joint changes at the loci were required for tumorigenesis, or whether they represented independent events in tumour initiation and/or progression. Microsatellite/genomic instability is believed to occur at random and may reflect alteration of the entire genome of the cancer cell (Li et al., 1998). Furthermore, the cumulative effect of the changes, rather than their order of occurrence is believed to be significant (Tabor 1994; Butel et al., 1996). Sheu et al. (1998), reported genetic instability in 9/34 (26%) of HCC patients. Since instability was detected in only 12/232 of their markers, they concluded microsatellite instability to be insignificant in HCC. In a Korean study by Kazachkov et al., (1998) microsatellite instability was detected in 4/10 (40%) HCCs, where each individual showed instability at two or more loci. Of the 9 markers used in their study, 3 showed genetic instability in one or more individuals. In our study 3/25 (12%) HCCs showed instability at two or more loci. Of the 10 loci investigated, all 10 showed genetic instability in one or more individuals. We propose that microsatellite instability may play a role only in a small subset of HCC in our population. Microsatellite/genetic instability is thought to be caused by mispairing followed by unequal crossing over or enzyme slippage, and has been proposed to be due to a lack of function of certain genes such as $p53$. The three individuals who showed microsatellite instability had no LOH of the $p53$ gene. In one individual there was a codon 249AGG—> AGT mutation in $p53$ which may have contributed to genetic instability in this individual. In a previous study (Kazachkov et al., 1998) found no alteration of the $p53$ gene in any of the tumours displaying genetic instability. However, as other mutations in $p53$ may have been missed in our study (see discussion for $p53$
gene) and there are many genes involved in cell cycle checkpoint pathways which were not analysed here, the occurrence of unequal crossing over or enzyme slippage causing genetic instability in our samples is possible.

In Sheu et al. (1998), tumours in those patients positive for HBsAg tended to have higher frequency of LOH. This was not the case in our samples. In one individual genomic instability was noted along with the \( p53 \) codon 249 AGC \( \rightarrow \) AGT mutation and anti-HBs and anti-HBc positivity, indicating that genetic changes are associated with the presence of HCC development in this individual.

The model proposed for recessive mutations of tumour suppressor genes in Retinoblastoma (Knudson et al., 1971) is almost certainly too basic to explain the events occurring in the development of tumours like HCC. Here certainly, more than one tumour suppressor may be lost, and ethnic differences such as predisposing hereditary, environmental and dietary factors between populations are interfacing to produce the final tumour.

4.2.2) SSCP Analysis of the \textit{BRCA2} Gene

Mobility shifts between tumour and matched non-tumorous DNA were detected for 3 individuals. The \textit{BRCA2} gene could not be sequenced (results not shown) and the causes of the mobility shifts are unknown.
4.2.3) p53 Codon 249 AGG→AGT Mutation Studies

The p53 codon 249 AGG→AGT mutation was detected in (14/52) 27% of the patients. This mutation was established in both tumour and non-tumorous tissues of one individual. Contamination of the non-tumorous tissue with tumour tissue, was suspected. However, insufficient tissue was available for histopathological analysis. In another individual the mutation was detected in the non-tumorous tissue only. The presence of this mutation has been documented before at a higher level in non-tumorous tissue than in the corresponding tumour tissue (Kirby et al., 1996). Kirby et al. (1996), proposed that normal liver subjected to prolonged aflatoxin exposure could gradually accumulate high levels of p53 codon 249 AGT mutations (ie. AGG→AGT transversion), whereas neoplastic populations that were cloned from single progenitor cells resistant to aflatoxin would not necessarily acquire this mutation. We did not confirm that the tissue was indeed non-tumorous liver tissue by histopathological examination so we cannot exclude contamination.

Most of our subjects with a p53 codon 249 AGG→AGT mutation were HBV positive, which concurred with a study by Hollstein et al. (1991) where mutation at codon 249 was not identified in non-HBV-related HCC's. As many of our subjects were from Mozambique, a country with a high aflatoxin B1 exposure (section 1.4.1.3), an occurrence of the p53 codon 249 AGG→AGT mutation in just over one quatre of them was not surprising. Both aflatoxin exposure and HBV infection appear to be required for the induction of this mutation (Kirby et al., 1996; Oda et al., 1992).
4.2.3.1) Gel Artefact

The suggestion of a mini-hairpin structure forming at codon 249 of the \textit{p53} gene position and generating a "missing" G was put forward by Kapelner \textit{et al.} (1994). In their study a G was missing at codon 249 of exon seven of the \textit{p53} gene in all wild type samples. At first it had appeared to be a frameshift mutation AGG$\rightarrow$AG. In all wild-type samples there was a single G band which was both crisp and clean (as in our wild-type samples - Figure 13) with no evidence of a compression. As with our samples \textit{HaeIII} digest confirmed the presence of the recognition sequence GGCC (Figure 11). The possibility of Taq DNA polymerase infidelity was ruled out (Lee-Jackson \textit{et al.}, 1993). When electrophoresed on a formamide gel Kapelner \textit{et al.} (1994) discovered that the "missing" G reappeared and concluded that the artefact was due to a compression.

Our sequence is a variant of the d(GCGA-AGC) structure which is thought to behave as follows: the structure bends and folds back between A4 and A5, and G1 lines up with A4, while A5 lines up with C7. Stacking and hybridisation interactions take place between the GC pairs, and the hairpin loop is formed (Figure 21). The G3A5 pair is non-Watson Crick and this bond helps to form a very stable hairpin structure (Hirao \textit{et al.}, 1994). During sequencing, in individuals with the codon 249 AGG$\rightarrow$AGT mutation, the T is substituted by an A. The replacement of a hybridising stem nucleotide destabilises the loop structure in electrophoresis. The DNA becomes linear, enhanced mobility (due to the loop structure) is lost, and gel compressions are prevented.
Sometimes the compressed G bands appear darker, and the following CCC bands are spaced more widely. This was observed both in this study and in Kapelner's, as well as the fact that the band compression was always reproducible. In Kapelner's study the band was always resolved by formamide.

This problem may also have been experienced by others (Kapelner et al., 1994). However, the absence of other reports of this "G deletion" in such a well-sequenced region led Kapelner et al. (1994) to suggest that it is not a common finding. This is easily explained as sequencing protocols differ between laboratories, and various factors and their interactions in electrophoresis can affect the structure of migrating species.
4.2.4) Techniques

4.2.4.1) PCR-RFLP Studies

This method was shown to be both rapid and simple. Direct sequencing is a viable method for the detection of useful DNA polymorphisms in well-characterised loci and could thus be used to confirm the G→T transversion in the \( p53 \) gene in those samples which had lost the \( HaeIII \) restriction site (Saiki et al., 1986).

4.2.4.2) PCR-SSCP Studies

Detection of mutations in DNA can be rapidly and easily done by the use of PCR-SSCP. PCR can amplify specific DNA sequences thousands of times by repeated cycles of denaturation, primer annealing, and extension by \( \alpha \) DNA polymerase (Saiki et al., 1985). Only small quantities of DNA are required as PCR is capable of amplifying DNA from a single template molecule. Genetic changes can then be identified through analysis of mobility changes of tumour DNA in electrophoresis resulting from the single stranded conformational polymorphisms (Peinado et al., 1992; Ionov et al., 1993). The drawbacks of this technique are that due to its sensitivity it may yield false positives from samples containing a trace amount of contaminant which would not be detected by less sensitive methods. This is overcome by including a negative control (Lo et al., 1988). The exact nature of the sequence changes must be determined by further analysis.

The possibility also exists that the SSCP method may underestimate a small portion
of mutations, (<10%) (Hayashi 1991). PCR-SSCP detects only about 80% of mutations present (Zhang et al., 1994).

To overcome the latter two drawbacks, direct sequencing or Southern blotting may be performed on those samples where possible mutations were detected by PCR-SSCP analysis.

4.2.4.3) VNTR studies

Restriction fragment length polymorphisms can be used to study LOH, but this technique is tedious, labourious and requires a considerable amount of DNA. Microsatellites, which are short tandem sequences, are widely distributed throughout the whole human genome and can be amplified by PCR. Genetic changes in tumours can be detected by using microsatellites as markers, thus providing a rapid and easy way to study LOH (Weissenbach et al., 1992; Cawkwell 1993; Thibodeau et al., 1993; Peltomaki et al., 1993; Cunningham et al., 1993).

The polymerase chain reaction is ideal to type polymorphic DNA as it consumes less DNA (10 nanograms is sufficient) and is faster than standard blotting and hybridisation (Saiki et al., 1988). The only requirement for utilisation of these polymorphic markers is knowledge of a unique set of amplification primer sequences and of the polymorphisms itself (Saiki et al., 1986). PCR products can then be electrophoresed by constant denaturing gel electrophoresis and separated into the different alleles. This approach is very rapid for screening areas where
some genomic sequence is already known, and is traditionally used to screen for mutations in tumour suppressor genes.

Direct sequencing can then be used to characterise the mutations previously detected by the mobility shift analysis (Garcia-Delgado 1997).

4.3) Gene Studies in Random Black Subjects/Polymorphism Frequencies

4.3.1) Allele Frequencies at the Different Loci.
At the D13S137 (Petrukhin et al., 1993), D13S120 (Bowcock et al., 1993), D13S127 (Bowcock et al., 1993), D17S855 (Anderson et al., 1993), and D16S260 (Weber et al., 1990) loci, the number of alleles found by us was not consistent with the number reported in the literature. These discrepancies may be explained as follows: the data for the D13S137, D13S120, D17S855, and D13S127 loci was obtained from DNA libraries, while this study employed a random sample of an actual population. The frequencies reported for the D16S260 locus were generated from a Caucasoid population, while our population was black. As the sample size for this study was small, the allele frequency values may only be crude estimates of the actual frequencies in the population.

4.3.2) Hardy Weinberg Equilibrium (HWE)
This population was found to be in Hardy Weinberg equilibrium which would suggest a number of factors, namely that the population was random mating, that...

...
advantage was evident for any one genotype and that mutation, migration and random genetic drift appear not to be operating. Sample size was however on the small side.

4.3.3) Polymorphism Information Content, Diversity and Observed Heterozygosity.

All the loci investigated proved to be highly informative (high PIC and H values). These high values were expected as these loci are highly polymorphic repeat loci. Furthermore, the D13S137, D13S120 and D17S855 loci had the highest PIC and H values, which was to be expected as they had the greatest number of alleles. These loci should thus be useful in population studies.
CHAPTER 5

CONCLUSION

It is not known whether the genetic events which lead to the development of HCC in different populations are the same, similar or distinct (Unsal et al., 1994).

Genes which would be good candidates to investigate in the transformation process (Slagle et al., 1993), and particularly in HCC, were characterised in this study. The *p53*, *RB*, *BRCA1*, *BRCA2*, *WT1* and *E-cadherin* genes were analysed for LOH, polymorphisms, and mutations, in tumour and non-tumour liver tissue from southern African blacks with HCC, using molecular techniques. PCR-RFLP and PCR-SSCP analyses which are both rapid and simple, were employed to detect mutations. Both these procedures require small amounts of starting tissue and DNA. Direct sequencing, or Southern blotting, was performed to establish the nature of the mutations, on those samples where possible mutations were detected by PCR-SSCP. Microsatellites markers were employed to investigate LOH, as they are widely distributed throughout the whole human genome, and can be amplified by PCR, providing a rapid and easy way to study LOH (Weissenbach et al., 1992; Cawkwell 1993; Thibodeau et al., 1993; Peltomaki et al., 1993; Cunningham et al., 1993).

LOH was detected at the *WT1* locus (on 11p13) in 1 individual, at the *RB1* locus in 3 individuals, and at the *BRCA2* locus (on 13q12-13) in 2 individuals.

Chromosome 11p13, a region containing the *WT* gene, has been reported to be deleted in
some HCCs. However, it was not certain whether WT1 was also deleted (Wang and Rowler 1988). Piao et al. (1997) reported LOH at the WT1 locus in 4.5% of HCCs. Our findings of 8% LOH at this locus in HCCs are concordant with those of Piao et al., (1997) and we suggest that the WT1 gene may be involved in HCC. Further studies will be required before this can be conclusively established.

Mutations of the BRCA2 gene have recently been reported in a subset of HCCs (Katagiri et al., 1996). Our finding of LOH at this locus in 10% of HCCs further supports the notion that BRCA2 may function as a tumour suppressor gene in the liver (Katagiri et al., 1996), and that it may in some way be involved in a subset of HCC.

The p53 and RB1 genes have been reported to be frequently mutated in various human cancers including HCCs. Coincident mutation, in HCC, of these two genes has also been documented (Murakami et al., 1991; Walker et al., 1991; Konishi et al., 1993; Farshid et al., 1994; Kasai et al., 1996; Sheu et al., 1997). Mutations and LOH in these genes is most frequently observed in advanced stage HCCs, like those investigated here. LOH for the RB gene was detected in 3 individuals only. LOH for the RB1 gene was comparatively low in our population. We speculate that this may be due to cells lacking functional pRB having undergone apoptosis at an early stage. It has been documented that cells whose pRB is non-functional may re-route via p53 for apoptosis, and no co-incident mutation of these two genes was found in this study. No LOH was found for the p53 gene. In our study for all those samples for which HBV status was known, most were HBV positive, or anti-HBs positive. We thus propose that a mechanism such as p53 inactivation by complex association
with virus encoded protein products (eg. HBx protein), rather than p53 gene inactivation by physical mutation or LOH, may have been operating in our tumours to eliminate the function of the p53 protein. Alternatively, the p53 and RB1 genes could be mutated in ways other than LOH, such as point mutations and small deletions (<50bp). Mutations such as these could have been left undetected in this study, as the entire p53 gene was not sequenced, but only a small (110bp) PCR product generated from a region of the p53 gene encompassing codon 249. No other genes were sequenced.

Whether HBV is involved in the chromosome losses reported here is unknown (Scorsone et al., 1992). However, as HBV infection increases the incidence of HCC, further studies will be required in order to determine if HBV DNA has been incorporated into the genome of the case subjects. Furthermore, the functional significance of HBV infection will have to be established in these cases.

Loss of E-cadherin expression has been reported to play an important role in HCCs, and LOH of the region where it is located (16q22) has been reported (Tsuda et al., 1990; Vleminckx et al., 1991; Sheu et al., 1998). No LOH was found for this gene in our study. Although all the HCCs used in this study were in advanced stages, it was not established whether they were highly undifferentiated. There may have been retention of E-cadherin expression in these samples, and no loss of intercellular adhesion.

LOH at the BRCA1 locus in HCC has been reported only once as far as we are aware (Piao et al., 1997). No LOH was found for this gene in our sample population, and we propose
that it may not be significant in HCC development or initiation. However, further investigation of the gene will be necessary for confirmation.

Expansion/contraction of microsatellite sequences is indicative of positive microsatellite instability which may indicate genomic instability and replication errors (Gao et al., 1995). Genetic instability was determined by mobility shifts in tumour DNA when compared to the matched non-tumorous DNA on polyacrylamide gels. This was confirmed by sequencing. Changes in repeat number were noted at all loci, however, out of the 25 individuals studied, in only 3 did changes occur at 2 or more loci. In the remaining individuals only one locus showed changes in repeat number. In these individuals it was therefore not possible to claim that the changes were due to genetic instability under its definition as: changes in microsatellite repeats at two or more loci (Kazachkov et al., 1998). The order of these changes is most likely insignificant. Their cumulative effect, however, may be important. We propose that microsatellite/genomic instability may play a role in a subset of HCCs only. This finding is in agreement with a recent study by Sheu et al. (1998)

More than one quarter of this population had the p53 gene codon 249 AGG → AGT mutation. This was expected as many of our subjects were from Mozambique, a country with a high aflatoxin B1 exposure. Dietary aflatoxin (AFB1) intake is believed to result in a G to T transversion at the third base of codon 249 of the p53 gene in 30-50% of HCCs (Hsu et al., 1991; Bressac et al., 1991; Ozturk et al., 1991; Li et al., 1993).

We determined crude allele frequencies for polymorphisms at the W71, D13S137, D13S120,
D13S127, D17S855, D16S301, and D16S260 loci for a random population of southern African blacks in order to determine how informative they would be for population studies, and studies such as this. To our knowledge this has not been done previously for these loci in this population. As expected for microsatellite loci, all the loci proved to be highly informative with high Polymorphism Information Content and Diversity values, thus making them useful in population studies.

In conclusion, our hypotheses proved mostly to be correct. Loss of heterozygosity is present in our population, and the frequency differs from that published for other populations, with WT1 gene being the only exception. The p53 codon 249 AGG→AGT mutation is present in approximately 30% of the HCCs studied, in accordance with previous studies on this population, and this percentage differs from other populations. Microsatellite/genetic instability is an important factor in a subset of this population. Allele frequencies established for microsatellite repeat loci did differ from those previously published, and statistical tests established the usefulness of the allele frequency data for future studies.

Coincident mutation of any of these genes was not observed in this study, in conflict with our hypothesis.

In brief, our observations support a possible role for p53, WT1 and BRCA2 genes in the pathogenesis of HCC, and microsatellite instability appears to be an important factor contributing to HCC development in a subset of our HCCs.
APPENDIX A - RECIPES

2% AGAROSE G\(^1\) L (150ml)
- 3 g Agarose D1 LE
- 150 ml 1X TBE

AMPICILLIN FOR AGAR PLATES (50 mg/ml) (20 ml)
- 1 g ampicillin
- make up to volume with ddH\(_2\)O
- filter sterilise into 1 ml aliquots
- store at -20°C
- use 1 ml/l

4% COMPOSITE AGAROSE GEL (3:1) (150 ml)
- 4.5 g Agarose D1 LE
- 1.5 g MS-8 Agarose
- 150 ml 1X TBE

DENATURED LOADING DYE
- 95% formamide
- 20 mM EDTA
- 0.05% bromophenol blue
- 0.05% xylene cyanol FF
FIXING SOLUTION FOR SEQUENCING GELS

- 150ml methanol
- 50ml acetic acid
- 800ml dH₂O

GLYCEROL TOLERANT BUFFER (500ml)

- 108g Tris base
- 36g Taurine
- 2g Na₂EDTA.2H₂O

8% GLYCEROL TOLERANT BUFFER SEQUENCING GEL (100ml)

- 42g Urea
- 5ml Glycerol tolerant buffer
- 20ml 40% Polyacrylamide
- 35ml distilled water
- 600μl 10% APS
- 25μl TEMED

0.2mM IPTG

- made up in ddH₂O
- use 20μl/plate
LURIA AGAR (100ml)

- 1g bacto-tryptone (1%)
- 0.5g yeast extract (0.5%)
- 0.5g NaCl (0.5%)
- 1.5 agar (1.5%)
- make up to volume with ddH₂O
- stir
- autoclave
- pour plates (±40ml/plate) before sets using sterile technique

LURIA BROTH (1l)

- 10g bacto-tryptone
- 5g bacto-yeast extract
- 5g NaCl
- make up to 1l with ddH₂O
- autoclave
- when cool add 0.05mg/ml ampicillin
MDE HYDROLINK GEL (100ml)

- 25ml MDE
- 52.5ml distilled water
- 12ml 10XTBE
- 10ml Glycerol
- 400µl 10% APS
- 40µl TEMED

NON-DENATURATING DYE FOR SSCP CONTROLS (100ml)

- 0.05g bromophenol blue
- 0.5g xylene cyanol

6% POLYACRYLAMIDE GEL (50ml)

- 7.5ml 40% polyacrylamide
- 15ml dH₂O
- 5ml 10x TBE
- 21g urea

- stir at room temperature until urea has dissolved

Just before pouring add:

- 160µl (10%) ammonium persulphate
- 45µl TEMED
- pour and allow to polymerise
SEQUENCING GEL (100ml)

- 42g Urea
- 30 ml dH₂O
- 15ml 40% Polyacrylamide
- 10ml 10XTBE

Just before pouring:
- 320μl 10% APS
- 90μl TEMED

- pour and allow to polymerise

SOB MEDIUM

- 20g bacto-tryptone
- 5g bacto-yeast extract
- 0.5g NaCl
- 950ml dH₂O

- stir until dissolved
- 10ml 250mM KCL

Just before use add 5ml 2M MgCl₂

SOC MEDIUM

- 1l SOB medium cooled to <60°C
- 20ml 1M glucose

- stir
40% STOCK ACRYLAMIDE (100ml)

- 38g Acrylamide
- 2g N,N'-Methylene-bis-acrylamide
- 5g Amberlite
- make up to volume with dH₂O

0.5M Tris(pH7.5)/1.5M NaCl (1l)

- 60.55g Tris base
- 87.6g NaCl
- make up to volume with dH₂O
- autoclave

5% XGal

- made up in dimethyl formamide
- use 20μl per plate
APPENDIX B

CALCULATIONS: PRIMER DILUTIONS

Example: Primers used to amplify D13S120 (from Bowcock et al., 1993)

GENBANK No. Utsw 1353 (M99152)

1) 1353L (ATGACCTAGAAATGATACTGGC)

(22mer); OD = 15.7

\[
\begin{array}{ccc}
A & 8(15.4) & = 123.2 \\
C & 4(7.3) & = 29.2 \\
G & 5(8.8) & = 44 \\
T & 5(11.7) & = 58.5 \\
\end{array}
\]

\[\text{concentration} = \frac{OD}{\xi} = \frac{15.7}{254.9} = 0.061592781... \mu\text{mol/}\mu\text{l}\]

For a 100pmol/\mu l stock:

\[0.061592781... \times 10000 = 615.9 \ \mu\text{l} \text{dH}_2\text{O}\]

Thus, to your lyophilized primers add 615.9 \ \mu\text{l} \text{dH}_2\text{O}.

But for PCR a 10pmol/\mu l stock is required, thus dilute 1:10

For 1000\mu l:

- 100\mu l primer
- 900\mu l \text{dH}_2\text{O}

Aliquot into tubes of 200\mu l each, and store at -20˚C.

All other primer calculations were done by this method.

The primers used to amplify D13S127 (Bowcock et al., 1993) (1341L and 1341R) also had a Genbank number, namely Utsw 1341 (M99150).
REFERENCE LIST


171

241 V. ooster R., Bignell G., Lancaster J., Swift S., Seal S., Magion J., Collins N.,
Gregory S., Gumbs C., Micklem G., Barfoot R., Hamoudi R., Patel S., Rice C.,
R., Peto J., Devilee P., Cornélise C., Lynch H., Narod S., Lenoir G., Egilsson V.,
Barkadottir R.B., Easton D.F., Bentley D.R., Futreal P.A., Ashworth A. and
Nature (Lond) 378:739-792.


244) Yandell D.W. and Drija T.P. (1989) Detection of DNA sequence polymorphisms
by enzymatic amplification and direct genomic sequencing. Am J Hum Genet
45:547-555.


Genetic alterations at the distal region of chromosome 1p in human hepatocellular


248) Yumoto Y., Hanafusa T. Hada H., Morita T., Ooguchi S., Shinji N., Mitani T.,

wild-type and mutant p53. FASEB J 7:855-865.

250) Zhang W., Hirohashi S. Tsuda H., Shimosato Y., Yokota J., Terada M. and
Sugimura T. (1990) Frequent loss of heterozygosity on chromosomes 16 and 4 in

Benedict W.F. and Sekiya T. (1994) Deletion of chromosome 13q, mutations in
retinoblastoma 1, and retinoblastoma protein state in human hepatocellular