ASSOCIATION OF GENE EXPRESSION AND GENOMIC CHANGE, ANALYSED USING MICROARRAYS, WITH PHENOTYPE IN BREAST CARCINOMA

by

Nicola Turton
MRC Toxicology Unit
University of Leicester

Thesis submitted for the degree of Doctor of Philosophy (PhD)
2003
Abstract

Association of Gene Expression and Genomic Change, Analysed Using Microarrays, with Phenotype in Breast Carcinoma.
Nicola J. Turton, MRC Toxicology Unit, University of Leicester

Breast cancer is the commonest malignancy to affect women and approximately 1 in 9 women (12 %) will develop the disease in developed countries (Burstein and Winer 2000). A major problem for treatment is the development of a drug resistant phenotype, whereby the tumour fails to respond to chemotherapy. This phenotype arises from altered gene expression in the resistant cells, which can occur by changes in transcription and/or genomic alteration. Other phenotypic properties of breast carcinoma cells may also arise in this way, and gene expression profiles can be linked to different breast tumour phenotypes, such as oestrogen receptor (ER) status, clinical tumour stage and tumour size (Martin et al., 2000). In this thesis cDNA microarrays were utilised to study both genomic amplification and RNA expression changes occurring in human breast carcinoma. These changes were related to phenotypic characteristics including a doxorubicin (Dox) resistant phenotype, hormone receptor status, tumour grade and type.

Several gene clusters involving the development of resistance and the eventual Dox resistant phenotype in breast cancer cell lines were elucidated, associated with both these was the multi drug resistance 1 gene (ABCB1). Potential therapeutic targets in these cells e.g. the oxytocin receptor gene (OXTR) were also indicated.

Regions of genomic amplification and specific genes were elucidated; some of these have previously been described while others are novel. Some of the genomic changes were associated with tumour phenotype, for example gene amplification at chromosome 2p25 and 22q11 were specific to lymph node positive tumours and 2 genes were consistently found amplified in these regions, LPIN1 and DGSI respectively.

The findings of this study have contributed to the general understanding of genetic events occurring in breast cancer and associations between these changes and phenotype were suggested.
Acknowledgements

This thesis and the work within it would not have been possible without the support, aid and patience of many people. So it is with pleasure that I can use this opportunity to express my gratitude. Firstly I thank the MRC for funding my PhD and my two supervisors, Tim Gant and Andy Smith for setting this project up. Special thanks go to Tim for his much appreciated support, ideas and assistance throughout this project. I am sure he’ll have many more happy PhD students coming through the door.

I should also thank the Protein and Nucleic Acid Laboratory (PNACL) for their rapid sequencing, plasmid preparations and oligo nucleotide production. Thanks also to Jacqui Shaw and Gavin at the Glenfield hospital for supplying some clinical tumour tissue samples.

During my time at the MRC I have been lucky enough to work with some great people, who not only have provided me with a good scientific environment in which to work, but more importantly a fun, productive one. Things like “Cake” will always make me chuckle! Amongst these people are some who have moved on since I started my PhD especially Pam, Petra and Damon. Some have been constants through out my time, such as the irrepressible Reg Davies, who brings a jubilant atmosphere wherever he goes, David Judah who has been responsible for spotting the microarrays. Whilst others have been more recent arrivals, including Sonia and Katie.

There is of course someone in the lab that I have yet to mention, Joan Riley, well what can I say? If I needed a protocol – Joan had one. If I needed someone to talk to – Joan listened. If I fancied a drink – Joan was more than happy to join me! Joan is a person of many talents: she has played in orchestras with the best, starred in the Commonwealth Games 2002 and will always be a dear friend. Keep in touch!

There are a number of people I wish to thank away from the lab. Especially my family: my brother Dave and Lynsey, Dad and Sharon and of course my Mum for all her hard work and love over the years. And last but not least, my husband and fellow Oz adventurer Nick. I’m fighting back tears of happiness as I’m sitting here trying to put into words what you have bought to my life over these past few years. But stop tickling me and give me back my nose!

Thanks everyone.
CONTENTS

Abstract ......................................................................................................................................1
Acknowledgements ...................................................................................................................2
Abbreviations ..........................................................................................................................10
List of Figures .........................................................................................................................12
List of Tables ..........................................................................................................................14

CHAPTER 1 ....................................................................................................................................16
GENERAL INTRODUCTION ...................................................................................................16

1. General Introduction ......................................................................................................17
1.1. Cancer General...............................................................................................................17
1.2. Breast-Cancer General Outline ..................................................................................18
     1.2.1. Incidence and Risk Factors ....................................................................................18
     1.2.2. Classification of Breast Cancer ..............................................................................19
     1.2.3. Staging of Breast Cancer ......................................................................................21
1.3. Breast Cancer Treatment .............................................................................................21
     1.3.1. Chemotherapy ...........................................................................................................22
     1.3.2. Hormone therapy .......................................................................................................23
     1.3.3. Novel therapies ..........................................................................................................25
1.4. Drug Resistance ..............................................................................................................25
     1.4.1. P-glycoprotein Mediated Multidrug Resistance .......................................................26
     1.4.2. Multi Drug Resistance Associated Protein and Other ABC Members ..................27
     1.4.3. Glutathione ................................................................................................................28
     1.4.4. Metallothionein .........................................................................................................28
     1.4.5. Topoisomerases .........................................................................................................29
     1.4.6. DNA Repair of Damage Caused by Anti Cancer Drugs ........................................29
     1.4.7. Changes in Membrane Properties ..........................................................................29
     1.4.8. Decreased Apoptosis .................................................................................................30
1.5. Genetics of Breast Cancer ............................................................................................30
     1.5.1. Genetic Progression of Breast Cancer .....................................................................30
     1.5.2. Chromosomal Aberrations in Breast Cancer ............................................................33
     1.5.3. Mechanisms of Chromosomal Aberrations ..............................................................35
1.6. Techniques for Studying Genomic Alterations .......................................................35
     1.6.1. G-bandling ..................................................................................................................35
     1.6.2. Fluorescent in situ Hybridisation (FISH) .................................................................36
1.6.3. Comparative Genomic Hybridisation (CGH) .......................................................... 36
1.6.4. Real-Time Quantitative PCR (RT-PCR) ................................................................. 37
1.6.5. Microarray Comparative Genomic Hybridisation .................................................. 38
1.7. DNA Microarrays ...........................................................................................................38
  1.7.1. Practical Applications of Microarrays ....................................................................39
  1.7.2. Different Options Available for DNA Microarrays ................................................40
    1.7.2.1. Targets Printed onto the Array .......................................................................40
    1.7.2.2. The Array Solid Support ...............................................................................41
    1.7.2.3. The Labelled Probe ......................................................................................42
  1.7.3. The Microarray Facility at MRC Toxicology Unit, Leicester ..................................44
1.8. Thesis Aims .....................................................................................................................46
  1.8.1 Analysis of the Drug Resistance Phenotype ............................................................46
  1.8.2 Genomic Amplifications in Breast Cancer .............................................................46

CHAPTER 2 ...................................................................................................................................48
MATERIALS AND METHODS ................................................................................................48
2. Materials and Methods ..................................................................................................49
  2.1. Buffers and Solutions .................................................................................................49
    2.1.1. Agarose Gel Electrophoresis .............................................................................49
    2.1.2. Poly-L Lysine Coating of Slides ......................................................................49
    2.1.3. Printing of Slides ..............................................................................................49
    2.1.4. Cell Culture ........................................................................................................50
    2.1.5. RNA Extraction from Cell Lines ......................................................................50
    2.1.6. DNA Extraction from Cell Lines ......................................................................50
    2.1.7. RNA Labelling for Microarray Analysis ...........................................................51
    2.1.8. Genomic DNA labelling .................................................................................52
    2.1.9. Griding Nylon Membranes .............................................................................52
    2.1.10. Nylon Array and Southern Blotting .................................................................53
    2.1.11. Protein Extraction ............................................................................................53
    2.1.12. Western Blotting .............................................................................................53
  2.2. Method Protocols ........................................................................................................56
    2.2.1. Preparation of Clones for Printing on the Array ................................................56
      2.2.1.1. Streaking and Growing clones ....................................................................57
      2.2.1.2. Creating bacterial Glycerol Socks and Plasmid Preparations ......................58
      2.2.1.3. PCR Amplification of Plasmid preparations .................................................58
      2.2.1.4. Gel Electrophoresis of PCR Product ...........................................................59
2.2.1.5. Precipitation of Amplified Clones ................................................................. 59
2.2.1.6. Sequencing of Clones .................................................................................. 60
2.2.1.7. Transfer of Clones to 384-well Plates for Printing ........................................ 60

2.2.2. Preparation of Slides .......................................................................................... 62
   2.2.2.1. Poly-L lysine Coating of Slides ..................................................................... 62
   2.2.2.2. Printing of Slides ......................................................................................... 63
   2.2.2.3. Post Processing of Printed Slides ................................................................. 63

2.2.3. Preparation of Samples for Microarray Analysis .................................................... 64
   2.2.3.1. Cell Culture and Maintenance ..................................................................... 64
   2.2.3.2. Treatment of Cells with Drug ........................................................................ 66
   2.2.3.3. Mutagenising Cells and Selecting for Drug Resistant Clones ...................... 66

2.2.4. RNA and DNA Extraction from Cell Lines .......................................................... 66
   2.2.4.1. Total RNA Extraction from Cell Culture ....................................................... 66
   2.2.4.2. DNA Extraction from Cell Culture ............................................................... 67
   2.3.4.3. Quantification of RNA/DNA ....................................................................... 68

2.2.5. RNA amplification ............................................................................................... 69
   2.2.5.1. Smart cDNA First Strand Synthesis .............................................................. 69
   2.2.5.2. SMART cDNA Amplification ..................................................................... 70

2.2.6. RNA/DNA/Amplified RNA Labelling and Hybridising ........................................ 70
   2.2.6.1 Total RNA Labelling ...................................................................................... 70
   2.2.6.2. Genomic DNA and Amplified RNA Labelling ............................................. 71
   2.2.6.3. Purification of the Labelled cDNA ............................................................... 72
   2.2.6.4. Hybridisation of Labelled cDNA to Glass Slide ........................................... 73
   2.2.6.5. Washing and Scanning of Hybridised Slides ............................................... 73

2.2.7. Data Analysis ....................................................................................................... 73

2.2.8. Post Microarray Experiments .............................................................................. 75

2.2.9. Nylon Membrane Microarrays ............................................................................ 75
   2.2.9.1. Gridding Clones onto Nylon Membranes ...................................................... 75
   2.2.9.2. Labelling RNA for Nylon Membranes ......................................................... 76

2.2.10. Cytotoxicity Assay in Cell Lines Exposed to a Compound .................................. 77

2.2.11. Southern Blot Analysis ....................................................................................... 78
   2.2.11.1. Southern Blot Membrane Preparation ......................................................... 78
   2.2.11.2. Probe Labelling and Hybridisation ............................................................. 79

2.2.12. Protein Extractions from Cells .......................................................................... 80

2.2.13. Western Blot Analysis ......................................................................................... 80
2.2.14. Real Time PCR Analysis

2.2.14.1. Reverse Transcription Reaction

2.2.14.2. SYBR Green Real Time PCR Reaction

2.2.15. Summary of Microarray Methods

CHAPTER 3

GENE EXPRESSION CHANGE AND THE DRUG RESISTANT PHENOTYPE

3.1. Introduction

3.1.1. Doxorubicin Mechanisms of Action and Toxicity

3.1.2. Mechanisms of Resistance to Chemotherapeutic Agents

3.1.3. Chapter Objectives

3.2. Results

3.2.1 Relative Toxicity of Various Breast Carcinoma Cell Lines to Doxorubicin

3.2.2 Gene Expression Changes and the Doxorubicin Resistant Phenotype

3.2.3. Genomic Amplification and the Doxorubicin Resistant Phenotype

3.2.4. Gene Expression in the Breast Carcinoma Cell Lines Compared to HBL100

3.2.5. Confirmation of Microarray Results

3.3. Discussion

3.3.1. Gene Expression Changes and the Doxorubicin Resistant Phenotype

3.3.2. Genomic Amplification and the Doxorubicin Resistant Phenotype

3.3.3. Gene Expression in the Breast Carcinoma Cell Lines Compared to HBL100

3.3.4. Summary

CHAPTER 4

GENE CHANGES AND DEVELOPMENT OF THE RESISTANCE PHENOTYPE

4.1. Introduction

4.1.1. Selection of Drug Resistant Cells

4.1.2. Mechanisms in Drug Resistance Selection

4.1.3. Chapter Objectives

4.2. Results

4.2.1. Cytotoxicity Assay for MCF7/wt and MDA468 Cells to Doxorubicin

4.2.2. Gene Changes After a 7 Day Time Course Exposure to 20 nM Doxorubicin

4.2.3. Cytotoxicity Data from the 14 Week Doxorubicin Dosing Study

4.2.4. Microarray Analysis from the 14 Week Doxorubicin Dosing Study

4.2.5. Gene Amplification and Resistance Development

4.3. Discussion
### 4.3.1. Gene Changes During 7 Day Time Course Exposure to 20 nM Doxorubicin

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.1.1. Genes Overexpressed After 1 Day Dox Exposure</td>
<td>124</td>
</tr>
<tr>
<td>4.3.1.2. Genes Overexpressed After 4 Days Dox Exposure</td>
<td>126</td>
</tr>
<tr>
<td>4.3.2. Resistance Acquisition by MCF7 Cells</td>
<td>126</td>
</tr>
<tr>
<td>4.3.3. Gene Changes in All Cell Groups After 14 Weeks Exposure to Doxorubicin</td>
<td>126</td>
</tr>
<tr>
<td>4.3.4. Genes Changes in Cells Prior to Doxorubicin Resistance Development</td>
<td>127</td>
</tr>
<tr>
<td>4.3.5. Gene Changes Associated with Doxorubicin Resistance Development</td>
<td>130</td>
</tr>
<tr>
<td>4.3.6. Gene Amplification and Doxorubicin Resistance Development</td>
<td>133</td>
</tr>
<tr>
<td>4.3.7. Summary</td>
<td>133</td>
</tr>
</tbody>
</table>

### 4.3.2. Resistance Acquisition by MCF7 Cells

### 4.3.3. Gene Changes in All Cell Groups After 14 Weeks Exposure to Doxorubicin

### 4.3.4. Genes Changes in Cells Prior to Doxorubicin Resistance Development

### 4.3.5. Gene Changes Associated with Doxorubicin Resistance Development

### 4.3.6. Gene Amplification and Doxorubicin Resistance Development

### 4.3.7. Summary

###CHAPTER 5

GENOMIC MICROARRAY ANALYSIS OF BREAST TUMOUR CELL LINES

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1. Introduction</td>
<td>136</td>
</tr>
<tr>
<td>5.1.1. Genomic Alterations in Breast Cancer Cell Lines</td>
<td>136</td>
</tr>
<tr>
<td>5.1.2. Specific Genes Associated with Genomic Alterations</td>
<td>138</td>
</tr>
<tr>
<td>5.1.3. Chapter objectives</td>
<td>139</td>
</tr>
<tr>
<td>5.2. Results</td>
<td>140</td>
</tr>
<tr>
<td>5.2.1. Overview of Gene Amplification</td>
<td>140</td>
</tr>
<tr>
<td>5.2.2. Specific Regions of Gene Amplification in Tumour Cell Lines</td>
<td>143</td>
</tr>
<tr>
<td>5.2.3. Frequent Regions of Gene Amplifications in Cell Lines</td>
<td>145</td>
</tr>
<tr>
<td>5.2.4. Specific Genes Involved in Frequent Regions of Amplification</td>
<td>146</td>
</tr>
<tr>
<td>5.2.5. Chromosome 20 Amplifications</td>
<td>147</td>
</tr>
<tr>
<td>5.2.5.1. Specific Genes Associated with Chromosome 20 Amplifications</td>
<td>149</td>
</tr>
<tr>
<td>5.3. Discussion</td>
<td>151</td>
</tr>
<tr>
<td>5.3.1. Frequent Regions of Gene Amplification in Cell Lines</td>
<td>151</td>
</tr>
<tr>
<td>5.3.2. Chromosome 20 Amplification</td>
<td>154</td>
</tr>
<tr>
<td>5.3.3. Chromosomal Aberrations in Specific Cell Lines</td>
<td>156</td>
</tr>
<tr>
<td>5.3.4. Problems Associated with Microarray Comparative Genomic Hybridisation</td>
<td>158</td>
</tr>
<tr>
<td>5.3.5. Summary</td>
<td>159</td>
</tr>
</tbody>
</table>

###CHAPTER 6

GENOMIC MICROARRAY ANALYSIS OF CLINICAL BREAST TUMOURS

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1. Introduction</td>
<td>161</td>
</tr>
<tr>
<td>6.1.1. Genomic Alterations in Infiltrating Breast Carcinoma</td>
<td>161</td>
</tr>
<tr>
<td>6.1.2. Genomic Alterations in Ductal Breast Carcinoma</td>
<td>162</td>
</tr>
<tr>
<td>6.1.3. Genomic Alterations Associated with Poor Prognosis and Metastasis</td>
<td>165</td>
</tr>
</tbody>
</table>
Chapter 6.1.4. Chapter Objectives ................................................................. 165

6.2. Results ........................................................................................................... 167
6.2.1. Overview of Gene Amplification .............................................................. 167
6.2.2. Specific Regions of Gene Amplification ................................................. 170
6.2.3. Frequent Regions of Gene Amplification in Clinical Tumours and Phenotype 172
6.2.4. Specific Genes Involved in Frequent Regions of Amplification .......... 174

6.4. Discussion ................................................................................................. 176
6.3.1. Frequent Regions of Gene Amplification in Clinical Breast Tumours .... 176
6.3.2. Region of Gene Amplification and Tumour Phenotype ....................... 178
   6.3.2.1. Region of Amplification Detected and Tumour Type .................. 178
   6.3.2.2. Region of Amplification and Lymph Node Status .................... 180
   6.3.2.3. Region of Amplification, ERBB2 and Menopausal Status .......... 180
   6.3.2.4. Region of Amplification and Tumour Grade .......................... 181
   6.3.2.5. Region of Amplification and ER Status ................................. 182
   6.3.3. Problems Associated With Clinical Tumour Samples .................... 182
   6.3.4. Summary ......................................................................................... 183

CHAPTER 7 ............................................................................................................. 184
CHROMOSOME 17 GENOMIC CHANGES ............................................................ 184

7.1. Introduction ............................................................................................... 185
   7.1.1. Regions of Gain on Chromosome 17q ............................................ 185
   7.1.2. The ERBB2 Oncogene .................................................................. 186
   7.1.3. Chapter Objectives ...................................................................... 188

7.2. Results ........................................................................................................ 189
   7.2.1. Specific Regions Containing Amplified Genes in Cell Lines .......... 189
   7.2.2. Specific Regions Containing Amplified Genes in Clinical Tumours 192
   7.2.3. The 17q11.2-q12 Region of Amplification .................................. 192
   7.2.4. Confirmation of Microarray Data in BT474 Cells ....................... 195

7.3. Discussion .................................................................................................. 200
   7.3.1. Amplification at 17q23 .................................................................. 200
   7.3.2. The ERBB2 region of Amplification at 17q11.2-q12 .................... 201
   7.3.3. Summary ...................................................................................... 203

CHAPTER 8 ............................................................................................................. 204
GENERAL DISCUSSION ....................................................................................... 204

8. General Discussion ....................................................................................... 205
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS</td>
<td>Ammonium Persulphate</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complement Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon Dioxide</td>
</tr>
<tr>
<td>Cy3</td>
<td>Cyanine-3</td>
</tr>
<tr>
<td>Cy5</td>
<td>Cyanine-5</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenosine Triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytidine Triphosphate</td>
</tr>
<tr>
<td>Dox</td>
<td>Doxorubicin</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethylpyrocarbonate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanine Triphosphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide Triphosphate</td>
</tr>
<tr>
<td>dpm</td>
<td>Disintegrations per minute</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiotritol</td>
</tr>
<tr>
<td>dTTP</td>
<td>Deoxythymine Triphosphate</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced Chemiluminesence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene Diamine Tetraacetic Acid</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium Bromide</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>EMS</td>
<td>Ethyl Methane Sulfonate</td>
</tr>
<tr>
<td>g</td>
<td>Gram(s)</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric Acid</td>
</tr>
<tr>
<td>HEPES</td>
<td>(N-[2-Hydroxyethyl]piperazine-N-[2-ethanesulfonic acid])</td>
</tr>
<tr>
<td>HGMP</td>
<td>Human Genome Mapping Project</td>
</tr>
<tr>
<td>H₂O</td>
<td>Water</td>
</tr>
<tr>
<td>hr</td>
<td>Hour(s)</td>
</tr>
<tr>
<td>IMAGE</td>
<td>Integrated Molecular Analysis of Genomes and their Expression consortium</td>
</tr>
<tr>
<td>J</td>
<td>Joules</td>
</tr>
<tr>
<td>K</td>
<td>Thousand</td>
</tr>
<tr>
<td>L</td>
<td>Litre(s)</td>
</tr>
<tr>
<td>LCM</td>
<td>Laser Capture Microdissection</td>
</tr>
<tr>
<td>m</td>
<td>Metre</td>
</tr>
<tr>
<td>M</td>
<td>Molar</td>
</tr>
<tr>
<td>MDR</td>
<td>Multi Drug Resistance</td>
</tr>
<tr>
<td>Min</td>
<td>Minute(s)</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium Chloride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribonucleic Acid</td>
</tr>
<tr>
<td>NaAc</td>
<td>Sodium Acetate</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NaH$_2$PO$_4$</td>
<td>Sodium phosphate monobasic</td>
</tr>
<tr>
<td>NaOCl</td>
<td>Sodium Hypochlorite</td>
</tr>
<tr>
<td>NaOH</td>
<td>Sodium Hydroxide</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>OD</td>
<td>Optical Density</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PMT</td>
<td>Photo Multiplier Tube</td>
</tr>
<tr>
<td>PNACL</td>
<td>Protein and Nucleic Acid Laboratory</td>
</tr>
<tr>
<td>PVP</td>
<td>Polyvinyl pyrrolidone</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions Per Minute</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulphate</td>
</tr>
<tr>
<td>sec</td>
<td>second(s)</td>
</tr>
<tr>
<td>SSC</td>
<td>Salt Sodium Citrate</td>
</tr>
<tr>
<td>SSPE</td>
<td>Sodium chloride/sodium phosphate monobasic/EDTA</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/Borate/EDTA</td>
</tr>
<tr>
<td>TE</td>
<td>Tris/EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N,N-Tetramethylethlenediamine</td>
</tr>
<tr>
<td>U</td>
<td>Units</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra Violet</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.1. The Development of the Drug Resistant Phenotype ...................................... 23
Figure 1.2. Mechanisms in Drug Resistance ...................................................................... 26
Figure 1.3. The Structure of P-glycoprotein ..................................................................... 27
Figure 1.4. A Hypothetical Model for the Progression of DCIS ...................................... 32
Figure 1.5. The Cyanine Labelled Nucleotides ................................................................ 43
Figure 1.6. Diagrammatical Representation of the Hybridisation Process ...................... 45

Figure 2.1. Overview of the Microarray Experimental Stages ........................................... 56
Figure 2.2 Overview of Clone Preparation ......................................................................... 57
Figure 2.3. An Example of an Output File from BLAST .................................................. 61
Figure 2.4. Overview of Slide Preparation ......................................................................... 62
Figure 2.5. Overview of Sample Preparation ...................................................................... 64
Figure 2.6. Overview of Data Analysis .............................................................................. 74
Figure 2.7. Overview of Post Microarray Work ................................................................. 75
Figure 2.8. Membrane Blot Transfer Set up for Southern Blotting ................................... 79
Figure 2.9. A Schematic Summary of the Microarray Process ......................................... 84

Figure 3.1. The Molecular Structure of Doxorubicin ......................................................... 86
Figure 3.2. Relative Toxicity of Breast Carcinoma Cell Lines to Doxorubicin ................... 90
Figure 3.3. Clusters of Gene Expression and Gene Amplification Associated With the Intrinsic and Acquired MDR Phenotype ..................................................................... 92
Figure 3.4. Clusters of Gene Overexpression in the Breast Carcinoma Cell Lines Relative to HBL100 ......................................................................................................................... 96
Figure 3.5. Confirmation of Microarray Data with Nylon Membranes ............................. 97
Figure 3.6. Confirmation of Microarray Data with Real Time PCR .................................. 98
Figure 3.7. The Relationship Between $MMP1$ and $TIMP2$ Expression .......................... 100

Figure 4.1. Cytotoxicity Assay of MCF7/wt and MDA468 Cells to Doxorubicin .......... 109
Figure 4.2. Dox Dosing Scheme for the MCF7/wt Cells ..................................................... 111
Figure 4.3. Gene Changes in MCF7 Cells After 1 Day with 20 nM Doxorubicin ......... 113
Figure 4.4. Gene Changes in MCF7 Cells After 4 Days with 20 nM Doxorubicin ......... 114
Figure 4.5. Cytotoxicity Assays in the 14 Week Study: Scheme 1 .................................... 116
Figure 4.6. Cytotoxicity Assays in the 14 Week Study: Scheme 2 ................................. 116
Figure 4.7. Cytotoxicity Assays in the 14 Week Study: Scheme 3 ................................. 117
Figure 4.8. Cluster of Genes Overexpressed in All Doxorubicin Treated Cells........... 120
Figure 4.9. Cluster of Genes Overexpressed in Cells Dosed Continuously with
Doxorubicin but not in Cells with a Slight Resistant to Doxorubicin ......................... 121
Figure 4.10. Genes Overexpressed in Cells with Slight Resistance to Doxorubicin... 122

Figure 5.1. Regions of Gene Amplification Detected in Breast Cancer Cell Lines .... 141
Figure 5.2. Chromosome 20 Regions Containing Amplified Genes ......................... 148

Figure 6.1. Regions of Gene Amplification Detected in Clinical Tumour Samples ... 168
Figure 6.2. Regions of Gene Amplification Associated with Tumour Phenotype ..... 173

Figure 7.1. Regions of Chromosome 17 Amplification in Tumour Samples .............. 190
Figure 7.2. Chromosome 17 Regions Containing Amplified Genes in Cell Lines...... 191
Figure 7.3. Chromosome 17 Amplification in Clinical Tumour Samples............... 193
Figure 7.4. Confirmation of Microarray Amplification Data by Southern Blot...... 196
Figure 7.5. Confirmation of Microarray Expression Data by Real Time PCR........ 197
Figure 7.6. Western Blot of the ERBB2 Protein................................................. 199

Figure 8.1. Genomic Amplification in Drug Resistant Cells .................................. 206
Figure 8.2. Gene Changes Associated with the Dox Resistant Phenotype .............. 207
Chapter 1

General Introduction
1. General Introduction

To date more than thirty organisms have had their entire genomes completely sequenced, with approximately another one hundred in progress (www.tigr.org). The first draft of the human genome has been released, with the project completion date estimated as 2003. All this sequencing data is an impressive leap forward, although it gives no indication to the function of the genes, how they interact with each other or how the cells work and differentiate. One method to analyse functionality of the genome is to study gene expression, as it is differences in gene expression that are responsible for most morphological and phenotypic differences in cells.

The aim of this project was to analyse gene expression of breast cancer and so further our understanding of the disease. There are over one million new cases of breast cancer in the world each year, it is the most common malignancy in women and constitutes 18% of all female cancers (McPherson et al., 2000). Several aspects of the disease were analysed, including genetic aberrations occurring in breast cancer and the ability of the tumours to develop a resistant phenotype when exposed to drugs.

The first part of this introduction focuses on breast cancer and genetic changes in breast tumours. The latter part will give some explanatory background into the techniques that have been utilised in this thesis to explore these genetic alterations in breast cancer.

1.1. Cancer General

The mechanism by which a normal cell turns neoplastic is a multistage process involving the accumulation of mutations in genes vital for the regulation of normal cell growth and maintenance. These genetic alterations are passed onto daughter cells, which may undergo further mutations. Some believe that normal rates of mutation coupled with clonal expansion are sufficient for this process to occur (Tomlinson et al., 1996), whilst others suggest that an underlying genetic instability is required for the generation of multiple mutations and cancer progression (Loeb, 1991).

The initial mutations may be caused by a variety of factors, including carcinogen binding to DNA, or via spontaneous mutation during DNA replication, which occurs with less than 100% accuracy (Cohen and Ellwein, 1991). The mutations may be reversed by error free DNA repair or may disappear when the cell dies. However if the cell with the altered DNA undergoes mitosis, then the alteration will be retained. Subsequent alterations may occur and
the mutations may drive a wave of cellular multiplication associated with a gradual increase in tumour size, disorganisation and malignancy.

Knudson et al. (1985) suggested one of the first molecular mechanisms of tumour development; they hypothesised that at least two mutation events are necessary for cell transformation from normal to neoplastic. Thus the cell containing the first mutation event has to survive in the tissue long enough to sustain a second. Later it was suggested that three to six mutations are necessary to complete the process (Vogelstein and Kinzler, 1993). During this period the neoplasm may convert from a benign (adenoma) to a carcinoma. Further accumulation of mutations enables the carcinoma to develop from benign to malignant, which may be invasive and metastatic (Vogelstein and Kinzler 1993). Specific genes are associated with these mutations, some of which are activated and are known as oncogenes, whilst others undergo a loss of function and are commonly referred to as tumour suppressor genes.

1.2. Breast-Cancer General Outline

1.2.1. Incidence and Risk Factors

The incidence of breast cancer is increasing and it is thought that 1 in 9 women (12 %) will develop the disease in developed countries and approximately a third of these women will succumb to the disease (Burstein and Winer 2000; Harris et al. 1992a). The incidence of disease increases with age, although the rate of increase slows after menopause.

There are many risk factors for breast cancer, some of these are summarised in Table 1.1. (Harris et al., 1992a). Some of these factors may contribute to sporadic cases (the occurrence of disease with no apparent family history of breast cancer), whilst others are thought to be genetically linked, as approximately 10-15 % of cases of breast cancer have a family history of breast or ovarian cancer (Arver et al., 2000).

Two breast cancer related tumour suppressor genes that have recently been identified and cloned are \textit{BRCA1} and \textit{BRCA2} on chromosome 17 and 13 respectively (Wooster et al., 1995; Miki et al., 1994). Germ line mutations in these two genes are associated with an increase in the overall risk of developing breast, ovarian cancer or both and \textit{BRCA2} has also been linked to the rarer form of male breast cancer (Thorlacius et al., 1995). However many women with a family history of breast cancer do not carry mutations in the BRCA genes and it is known that other candidate genes exist. One such example is the ATM gene, which is mutated in ataxia-telangiectasia (A-T), a rare autosomal recessive disorder associated with a high risk of breast
cancer in multiple-case breast cancer families (Chenevix-Trench et al., 2002). Another example associated with a high incidence of breast cancer is Li-Fraumeni familial cancer syndrome, where germline p53 mutations is one of the underlying defects (Malkin et al., 1990).

Table 1.1. Risk Factors in Breast Cancer

<table>
<thead>
<tr>
<th>High Risk (&gt; 4*)</th>
<th>Moderate Risk (2-4*)</th>
<th>Slight Risk (1-2*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal history of breast cancer</td>
<td>Personal history of benign proliferation breast disease</td>
<td>Personal history of benign breast disease</td>
</tr>
<tr>
<td>&gt;1 first degree relative affected</td>
<td>Mother affected before the age of 60 yrs</td>
<td>Mother affected after the age of 60 yrs</td>
</tr>
<tr>
<td>Presence of hereditary germ line mutations</td>
<td>Postmenopausal obesity</td>
<td>Onset menarche age &lt;12 yrs</td>
</tr>
<tr>
<td>Living in a developed country</td>
<td>High alcohol intake</td>
<td>Moderate alcohol intake</td>
</tr>
<tr>
<td>Elderly</td>
<td>Personal history of carcinoma of ovary or endometrium</td>
<td>Age at birth of first child &gt;25 yrs</td>
</tr>
<tr>
<td></td>
<td>Prolonged uninterrupted menses</td>
<td>Current use of oral contraceptive</td>
</tr>
</tbody>
</table>

Table 1.1. Legend: Some of the established and probable risk factors associated with breast cancer (McPherson et al., 2000; Harris et al., 1992). * Fold greater risk than general population

1.2.2. Classification of Breast Cancer

Once diagnosed breast tumours are classified by histopathological examination. The three main classes of breast disease are benign lesions, in situ carcinomas and invasive carcinomas.

Benign lesions, which can mimic the clinical presentation of cancer, occur far more frequently than breast cancers. There are many types of benign lesions several of which, e.g. fibroadenomas, sclerosing adenosis, papillomas or radical scars have been shown to lead to an increase in long-term risk of subsequent breast cancer (Jacobs et al., 1999; Simmons and Osborne, 1999; Dupont et al., 1994). For example the relative risk of invasive breast cancer in patients with fibroadenomas when compared to control groups is 2.17-3.1 times higher, and this high-risk level may remain elevated for decades (Dupont et al., 1994). Another benign tumour example is atypical hyperplasia, which is found in approximately 4 % of benign biopsies (Page et al., 1985). These lesions show some architectural features of carcinoma in situ but lack the complete criteria for diagnosis and have also been shown to lead to an
increase risk of subsequent invasive carcinoma (Simmons and Osborne, 1999; London et al., 1992; Page et al., 1985).

*In situ* carcinoma (CIS) of the breast is typically recognised by the proliferation of malignant-appearing cells within the ducts (ductal carcinoma *in situ*, DCIS) and lobules (lobular carcinoma *in situ*, LCIS), without evidence of light microscopic invasion of the surrounding stromal tissue (Zaugg and Bodis, 2000). The risk of subsequent breast cancer development for most women diagnosed with CIS has been calculated at 16% after 10 years, and in most cases CIS and subsequent invasive cancer shows the same morphological features (i.e. ductal or lobular) (Franceshi et al., 1998).

DCIS is typically detected with screening mammography and an increase in screening strategies has seen the incidence of DCIS rise rapidly, thus in 1983-1992 the incidence of DCIS increased by 200% in the USA (Emster et al., 1996). DCIS can be divided into two main subtypes, comedo or noncomedo. Comedo DCIS, which is characterised by higher nuclear grade morphology, is associated with a higher risk of subsequent breast cancer as opposed to non comedo (Badve et al., 1998; Silverstein et al., 1990; Lagios et al., 1989), and it has been shown that there is increased rate of local recurrence with higher nuclear grade comedo DCIS (Solin et al., 1993). Comedo DCIS when compared to noncomedo also has a higher expression of the oncogene *ERBB2* (Claus et al., 2001), whose overexpression has also been associated with a worse prognosis in breast cancer (McCann et al., 1991). Non comedo DCIS can be further divided into a number subtypes that include cribriform, micropapillary, papillary and solid (Simmons and Osborne, 1999). An alternative method classifies DCIS into well, intermediate or poorly differentiated. Poorly differentiated DCIS have large atypical nuclei, are more aggressive and have a higher risk of recurrence and invasiveness than well-differentiated DCIS, while intermediate DCIS show a mixed pattern (Van Diest, 1999) and treatment of DCIS will depend on diagnosis of the tumour (Schwartz et al., 1989).

LCIS is multicentric and predominantly occurs in women aged 44-49 years, which is approximately 10-15 years younger than the average age for developing invasive breast carcinoma, and of those women diagnosed with LCIS approximately 2/3 are premenopausal (Hutter, 1984). LCIS is characterised by masses of loosely arranged cells that contain small, uniform, round to oval nuclei and variably distinct borders (Stone et al., 2001). Little is known about its true incidence, as it is not usually grossly evident and lacks clinical or mammographic signs (Harris et al., 1992b), consequently it is often an incidental finding in
breast biopsy samples removed for other reasons (Fryckberg et al., 1987). LCIS also carries an increased risk of invasive breast cancer of 7-10 fold (Page et al., 1991). And the subsequent invasive cancers are equally likely to occur in both the ipsilateral and contralateral breast, unlike DCIS, which normally develops in the same quadrant (Harris, 1992b).

Invasive breast carcinoma is a tumour where there is dissemination of cancer cells outside the basement membrane of the ducts and lobules into surrounding adjacent normal tissue (Sainsbury et al., 2000). Some invasive cancers show distinct patterns of cell morphology and growth and are subdivided into tubular, cribriform, medullary, mucoid, papillary or classic lobular (Sainsbury et al., 2000). While others, which are characterised by the absence of any special histological features, are called invasive carcinoma, not otherwise specified (NOS), and account for approximately 75% of carcinomas (Harris, 1992b).

1.2.3. Staging of Breast Cancer

Several staging systems have been proposed for tumours. For example analysis of histological tumour grade whereby nuclear morphology and architectural differentiation are used to grade tumours can divide DCIS into poor, intermediate and well-differentiated groups. Similarly invasive tumours are often graded I - III depending on degree of glandular formation, nuclear pleomorphism and frequency of mitosis (Sainsbury et al., 2000). An internationally recognised staging system, the TNM system (Tumour, Nodes and Metastasis), is most commonly used to classify cancers (Beahrs et al., 1988). This system stages breast tumours by analyses of the extent of disease, the involvement of lymph nodes and the presence of metastasis. Clinical in vivo tumour samples used in this project have been staged in this way. The information obtained from clinical staging is useful in choosing treatments and estimating prognosis for individual patients.

1.3. Breast Cancer Treatment

The actual treatment received by a patient will depend on factors such as clinical stage and classification of tumour, age of patient, current health of the patient or even where the patient is being treated. There is evidence to suggest that in the UK there are variations in treatment given to women with breast cancer and this treatment variation has been implicated in patient survival (Purushotham, 2001). Other countries in Europe also show strong variation in survival between regions (Quinn et al., 1998); for instance, Spain and Italy show differences of 16% and 22% respectively in 5-year survival between the affluent northern regions and the poorer southern parts of these countries.
Primary breast cancer in some women is a local disease without distant spread, these cancers if diagnosed quickly are highly curable with local or regional treatment alone, although some women with primary breast cancer have subclinical metastases or develop distant metastases after surgery (Tabar et al., 1992). In metastatic breast cancer there are large variations in growth rate and therapy response. The majority of women with metastatic breast carcinoma will succumb to the disease (Greenberg et al., 1996), thus prolongation of life with minimal disruption quality of life is often the treatment goal.

The optimal treatment for women with breast cancer involves multiple methods (Wood, 1998), so combinations of surgery, radiotherapy, hormone therapy, chemotherapy or chemoprevention are normally used. Two of these treatments, chemotherapy and hormone therapy involve the administration of drugs to which breast carcinoma cells can develop a resistant phenotype. A large part of this project is directed towards understanding alterations in gene expression that take place during the development of drug resistance to these treatments.

1.3.1. Chemotherapy

Chemotherapy is one of the more effective treatments used for combating malignant tumours, especially if the lesions have spread or metastasised. Common methods by which administered drugs exert their effects is by crosslinking or intercalating DNA, inhibiting topoisomerases or rectifying the cell cytoskeleton thus limiting cell division.

Women with breast cancer who are treated with adjuvant therapy suffer from many side effects that result from the action of the drug on normal cells. As well as serious side effects another major obstacle in the successful treatment with chemotherapy is the development of a drug resistant phenotype. Resistant cells may arise from sensitive cells and undergo clonal expansion resulting in a resistant tumour (Figure 1.1.). The mechanisms involved in the development of resistance are discussed in Section 1.4. One method to combat the development of resistance is combination chemotherapy. Combination chemotherapy is more effective than single-drug therapy, reducing the annual risk of death by about 20 % (Wood, 1998), and involves the administration of several drugs, which exert their anticancer effect via different mechanisms, thus hindering the development of a resistant phenotype. This is summarised in Figure 1.1., which shows the advantage of administering a combination of drugs rather than repeat administration of the same drug.
Figure 1.1. The Development of the Drug Resistant Phenotype

Figure 1.1. Legend: The development of the resistant phenotype can occur after several rounds of chemotherapy, and resistance to one drug often confers resistance to other drugs not previously used. The use of combination chemotherapy involving drugs that exert their chemotherapeutic affect in different ways (Rx I + Rx II) may result in ablation of the tumour.

The choice of drug regime depends on recurrence risk, morbidity and patient preference. Frequent combinations of chemotherapy include cyclophosphamide, methotrexate and 5-fluorouracil (CMF); fluorouracil, adriamycin (doxorubicin) and cyclophosphamide (FAC); fluorouracil, epirubicin and cyclophosphamide (FEC); and adriamycin and cyclophosphamide (AC). Another approach to combat resistance is the use of chemosensitisers, some of which are undergoing clinical evaluation for the treatment of resistant tumours (Teodori et al., 2002). However, the development of resistance to anti cancer drugs is still a major impediment in the treatment of the disease.

1.3.2. Hormone therapy

Resistance to a single anti cancer drug is often associated with resistance to several unrelated compounds, a process known as multi drug resistance. A major advantage of hormone therapy is that women, who respond to one method of hormonal treatment, will often respond well to a second type of hormone when the first becomes ineffective. Hormone therapy is often used in chemoprevention, a process that involves the use of xenobiotics in cancer therapy to arrest or reverse the process of carcinogenesis before invasion and metastasis occur (Hong and Sporn, 1997).
Tamoxifen, a synthetic antiestrogen synthesised in 1966, has become a main approach to hormone therapy for all stages of breast cancer (Jaiyesimi et al., 1995). Adjuvant tamoxifen is recommended for all women with hormone receptor (HR) positive breast cancer, and the reduction in 10-year mortality in women with HR positive breast cancer treated with 5 years tamoxifen versus no tamoxifen is 47% (Early Breast Cancer Trialists Collaborative Group, 1998). Tamoxifen has additional benefits, as it can decrease cholesterol levels by 12% and low-density lipoprotein cholesterol by about 20% (Love et al., 1991) and it also gives protection against osteoporosis by preserving bone mineral density in postmenopausal women (Powels et al., 1996). However there are a number of risks associated with tamoxifen, firstly there is a low risk of the development of endometrial cancer, (Fisher et al., 1994). Secondly there is a slight decrease in plasma concentrations of antithrombin III, which may contribute to a very small increase in risk of deep vein thrombosis (Jordan et al., 1987). Other symptoms linked to tamoxifen administration include hot flashes and irregular menses (Wood, 2001). Even so the net benefit of tamoxifen therapy is thought to greatly outweigh risk (Fisher et al., 1994). Other oestrogen analogues, such as raloxifene are in developmental stages and show the potential for chemoprevention of breast cancer without some of the side effects of tamoxifen (Cummings et al., 1999; Black et al., 1994).

Ovarian ablation has also been shown to improve survival (Early Breast Cancer Trialists’ Collaborative Group, 1996), although more recently ovarian suppression using a luteinising hormone-releasing hormone analogue is an alternative to surgical or irradiation-induced castration with the added advantage of being reversible. Other hormonal therapies include megestrol acetate, progestin (Brufman and Biran, 1990) or anastrozole, an inhibitor of aromatase, which is the enzyme responsible for catalysing the conversion of androgens to oestrogen (Anderson et al., 2002).

It has been shown that the combination of tamoxifen or ovarian suppression with chemotherapy is more effective than either alone (Early Breast Cancer Trialists’ Collaborative Group, 1996 and 1992). There is some agreement regarding choice of optimal adjuvant therapy once it is determined a women will benefit from this intervention, which is summarised in Table 1.2. (Wood, 1998).
Table 1.2. Adjuvant Therapy for Women with Operable Primary Breast Cancer

<table>
<thead>
<tr>
<th>ER* status</th>
<th>Level of risk</th>
<th>Systemic therapy for women &lt; 50 years</th>
<th>Systemic therapy for women &gt; 50 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>Any</td>
<td>CT</td>
<td>CT</td>
</tr>
<tr>
<td>+</td>
<td>Low</td>
<td>HT, CT or HT + CT</td>
<td>Tamoxifen or HT + CT</td>
</tr>
<tr>
<td>+</td>
<td>Moderate/ high</td>
<td>HT + CT or novel</td>
<td>HT + CT or novel</td>
</tr>
<tr>
<td>Unknown</td>
<td>Any</td>
<td>HT + CT</td>
<td>HT + CT</td>
</tr>
</tbody>
</table>

Table 1.2. Legend: ER* = Oestrogen Receptor; CT = Chemotherapy; HT = Hormone Therapy; novel = novel therapeutic strategies

1.3.3. Novel therapies

Increased understanding of the biology of breast cancer is leading to the development of novel therapies. For example the anti-HER2/neu antibody, directed against the HER2 oncogene encoded by the ERBB2 gene located on chromosome 17, which is an indicator of poor prognosis and is overexpressed in approximately 10-34 % of breast cancers (Revillion et al., 1998). A randomised trial in which chemotherapy was combined with the HER2 antibody reported an improved response rate and an increase in overall survival, when compared to patients receiving chemotherapy alone (Pietras et al., 1998). This region in chromosome 17 is discussed in greater detail in Chapter 7. Another novel therapy is the use of telomerase inhibitors (Keith et al., 2001). Telomerase is expressed in a variety of cancers with a high degree of specificity and inhibition of telomerase has been shown to result in cell death of immortalised breast epithelial cells (Herbert et al., 1999).

1.4. Drug Resistance

In some cases after first round treatment with chemotherapy a reduction in tumour size or complete elimination of the tumour can be achieved. However, in most cases successful treatment is relatively short lived and relapse follows, this temporary period of remission was first described in 1948 (Farber et al.). Thus the major obstacle for successful treatment with chemotherapy is the development a resistant phenotype by breast cancer cells.

Resistance can either be intrinsic, whereby the tumour cells do not respond to treatment from the outset, or acquired where tumour cells develop certain biological mechanisms, which render them resistant to treatment (Osmak, 1998). In many cases when a cell no longer responds after treatment with a single drug, it is also resistant to various structurally and
mechanistically unrelated compounds; this phenomenon is known as multi-drug resistance (MDR). There are several molecular mechanisms that can contribute to the drug resistance phenotype some of which are summarised in Figure 1.2.

**Figure 1.2. Mechanisms in Drug Resistance**

![Figure 1.2. Mechanisms in Drug Resistance](image)

Figure 1.2. Legend: Mechanisms thought to be involved with cellular acquisition of a drug resistance phenotype include increased drug clearance, inactivation of the drug and increased DNA damage repair.

**1.4.1. P-glycoprotein Mediated Multidrug Resistance**

P-glycoprotein (P-gp) is a 170 kDa plasma membrane protein whose structure is shown in Figure 1.3. It is a member of the ATP-binding cassette (ABC) superfamily of transport proteins (Dean et al., 2001; Higgins 1992) and is found in normal tissue, especially in organs that have an excretory function (Thiebaut et al., 1987), and is also present at pharmacological barriers of the body like the blood-brain barrier (Cordon-Cardo et al., 1989).

P-gp exerts its drug resistance effect by actively extruding drugs from the cell, thus reducing the drugs intracellular concentrations to below the threshold of their chemotherapeutic effect (Gottesman and Pastan, 1993). P-gp is able to transport a wide variety of structurally and mechanistically unrelated agents out of the cell. Substrates include anthracyclines (doxorubicin), alkaloids (vincristine, vinblastine), specific peptides (cyclosporin A), local anaesthetics (dibucaine) or dye molecules (rhodamine 123) (Seelig, 1998). P-gp is encoded by the multi drug resistance 1 (*MDR1* or *ABCB1*) gene (Ueda et al., 1987), in multidrug resistant cells increased gene expression has been recorded which can be followed by amplification of the *ABCB1* gene (Shen et al., 1986).
Figure 1.3. Legend: A schematic representation of the structure of P-gp. It is made up of two halves sharing 43% homology; each half contains six α-helical hydrophobic transmembrane domains and a nucleotide binding sequence; an intracellular linker region of 60 amino acids joins the two halves (Juranka et al. 1989).

1.4.2. Multi Drug Resistance Associated Protein and Other ABC Members

Multidrug resistance associated protein (MRP or ABCC1), which is ubiquitously expressed in normal tissue (Zaman et al. 1993), is another member of the ATP-binding cassette family. It was isolated from non P-gp MDR lung carcinoma cells (Cole et al., 1992) and has also been identified in a variety of non P-gp MDR tumour cell lines. ABCC1 is a 190 kDa protein, which has only a 15% sequence similarity to P-gp (Loe et al 1996). It generally has a pattern of drug resistance similar to P-gp (Grant et al 1994; Loe et al 1996), although there are differences, as it confers a low resistance to paclitaxel and colchicines, which are reported to be good substrates for P-gp (Osmak, 1998).

Another member of the ABC family, ABCG2 previously know as the breast cancer resistance protein (BCRP) is also involved in the MDR phenotype (Diestra et al., 2002). Resistance associated with this protein includes cross-resistance to topoisomerase inhibitors, doxorubicin and daunorubicin (Diestra et al., 2002). This protein is mainly expressed at the plasma
membrane in MDR cells and is thought to be widely present in human tumours, thus it may also have clinical significance (Diestra et al., 2002).

1.4.3. Glutathione

Glutathione (GSH) and its related enzymes also play a role in drug resistance. GSH is involved in cellular metabolism and protection against free radical induced oxidant injury (Meister, 1994). GSH is involved in drug resistance to a number of drugs, such as cisplatin (Mitchell and Russo, 1987). It is thought that levels of GSH may contribute to the resistant phenotype in several ways, firstly by binding electrophilic compounds thus rendering them less toxic and more readily extruded from the cell by e.g. the \textit{ABCC1} pump, secondly GSH may assist in the repair of damage induced DNA in the nucleus (Lai et al., 1989).

A group of GSH related enzymes, the glutathione-S-transferases (GSTs) play a role in drug clearance, as they are part of the Phase II detoxification process that metabolise many lipophilic xenobiotics, including anti-cancer drugs (Tew et al., 1994; Waxman, 1990). This results in the conversion of these lipophilic chemicals to more water-soluble derivatives that are more readily eliminated from the body. GSTs have been linked to resistance, as elevated GST levels have been recorded in some cells resistant to anti-cancer drugs e.g. melphalan, cisplatin, and to a lesser extent doxorubicin (Caffrey et al., 1999). Transfection experiments have also shown that GST causes a slight but significant increase in resistance to a number of chemotherapeutic agents (Puchalski and Fahl, 1990).

1.4.4. Metallothionein

Metallothioneins are cysteine rich proteins that are involved in heavy metal detoxification (Kelly et al., 1988). They are known to have a role in drug resistance, as increased levels of metallothioneins have been noted in tumour cells displaying a drug resistant phenotype to anti-cancer drugs, such as cisplatin and chlorambucil (Kelly at al. 1988). However increases in metallothioneins do not always result in a cellular phenotype that is more resistant to the toxic effects of antineoplastic drugs (Schilder et al. 1990). More recent studies have appeared to confirm a role for metallothioneins in drug resistance, as embryonic cells of metallothionein negative transgenic animals have an enhanced sensitivity to a number of anti-cancer drugs (Kondo et al. 1995).
1.4.5. Topoisomerases

A form of drug resistance known as atypical MDR is associated with DNA topoisomerase IIα (Topo IIα). Topo IIα creates double stranded breaks and aids in the unwinding of DNA in mammalian cells (Withoff et al., 1996) and its overexpression has been reported in several human cancers including breast cancer (Rudolph et al., 1999; Lynch et al., 1997). Consequently Topo IIα is a target for anticancer agents, which stabilise Topo IIα-DNA complexes thus inhibiting DNA replication and transcription (Wilson et al., 2001). Drug resistance is thought to occur by a reduction in expression or function of Topo IIα, since etoposide resistance in glioma cells can be achieved by a stepwise increase in exposure to the drug, which was associated with decreased Topo IIα activity (Taki et al., 1998). Another study demonstrated that movement of Topo IIα from the nucleus to the cytoplasm, caused by increased cell to cell contact, also contributes to resistance (Valkov et al., 2000) and it has been suggested that Topo IIα associated MDR is an early event in the selection process for resistance (Matsumoto et al., 2001). Although Topo IIα may play a role in the MDR phenotype, it has been proposed that the primary role of Topo IIα is cell proliferation and consequently highly proliferative tumours are still likely to be sensitive to treatment regimes using anti- Topo IIα agents (Wilson et al., 2001).

1.4.6. DNA Repair of Damage Caused by Anti Cancer Drugs

There is evidence to suggest that enhanced ability of the cell to repair DNA damage caused by a variety of cytotoxic drugs may lead to a resistant phenotype (Osmak 1998). For example, cisplatin causes DNA adducts which contribute to cellular toxicity and an enhancement in repair of this DNA damage have been observed in a number of cisplatin resistant cells (Scanlon et al., 1991). The adducts are also removed much more rapidly in the resistant cell lines as opposed to their sensitive counterparts (Eastman and Schulte 1988). However the exact mechanisms of this enhancement of DNA repair are not well understood.

1.4.7. Changes in Membrane Properties

Changes in membrane properties may play a role in drug resistance, and decreases in membrane fluidity have been observed to occur in resistant cells exposed to Dox (doxorubicin) when compared to sensitive cells (Siegfried et al., 1983). However the resistant cells do not show any differences in ability to take up or efflux drug over sensitive cells, leading to the hypothesis that changes in membrane properties are involved in Dox resistance by mechanisms unrelated to drug efflux or uptake (Siegfried et al., 1983).
1.4.8. Decreased Apoptosis

Drug resistance may also arise as cells become resistant to apoptosis induced by antitumour agents. For example a leukaemia cell line resistance to various chemotherapeutic agents including Dox, and its sensitive parental cell line from which it was selected, were shown to have similar expression levels of the \textit{ABCB1} gene product and DNA topoisomerase (Sakamoto et al., 2000). However drug induced activation of caspases and subsequent apoptosis was significantly suppressed in the resistant cell line. Glyoxlase 1 (\textit{GLO1}), an enzyme that detoxifies methylglyoxal, was found overexpressed in the resistant cells and co-treatment with an inhibitor of \textit{GLO1} enhanced apoptosis in the resistant cells (Sakamoto et al., 2000).

1.4.9. Drug Resistance: A Multifactor Process

Drug resistance is thought to be a multifactorial process. As in cells resistant to cisplatin, decreased drug accumulation, increased levels of glutathiones and metallothioneins and an increase in DNA repair have all been observed (Scanlon et al. 1991). It is thought that due to induction of several molecular mechanisms after treatment with a single anti cancer drug, resistant cells may become multi drug resistant. Not all the mechanisms of the drug resistant phenotype are clearly elucidated and so the need to clarify the events involved in drug resistance and how they affect each other is of paramount importance in the successful treatment of cancer.

1.5. Genetics of Breast Cancer

1.5.1. Genetic Progression of Breast Cancer

There is evidence to suggest that genetic instability increases as the tumour develops from a precursor lesion to an invasive carcinoma and continues to increase with tumour stage (Ried et al., 1999). In some cancers, such as cholorectal, there is a widely excepted model whereby certain genetic aberrations occur in a stepwise manner, and these changes correlate well with morphological change of the tumour from benign to carcinoma (Kinzler and Vogelstein, 1996). In breast cancer there is not such a clear transition, and relatively little is known about the sequence of molecular events in progression of the disease (Rennstam et al., 2001).

It has been hypothesised that different classifications of breast cancer such as LCIS and well-differentiated DCIS, which are closely related lesions, could evolve from one neoplastic cell (Buerger et al., 2000). There is also evidence suggesting that breast cancers grow and develop
from well-differentiated grade I tumours to poorly differentiated grade III tumours and that the malignant capacity of many cancers increases as the tumour grows (Rajakariar and Walker, 1995; Duffy et al., 1991), thus explaining why improved breast screening programs, which are thought to detect the disease at earlier stages, have reduced mortality (Peto et al., 2000). Although there is conflicting evidence regarding the true benefits of screening, likewise there is also conflicting evidence for tumour progression; as a number of studies suggest that a majority of histologically well differentiated grade I tumours do not progress to grade III tumours (Roylance et al., 1999; Millis et al., 1998).

The study by Roylance et al. (1999), which looked at DCIS, found a loss of 16q in 65% of grade I well differentiated tumours, compared to only 16% in grade III poorly differentiated tumours suggesting that different breast tumour grades may have distinct molecular origins. However to test the significance of this numeric finding it would be necessary to determine if any of the 65% of patients with the grade I tumours containing the 16q deletion later develop grade III tumours also containing the deletion.

The findings of Roylance et al. (1999) are supported by a hypothetical model of DCIS progression from well, intermediate and poorly differentiated to invasive carcinoma. The model, which is illustrated in Figure 1.4., was proposed by Buerger et al. (1999) and contains at least 3 genetic pathways. In general they found that well differentiated tumours had less genetic aberrations than those that were poorly differentiated, suggesting a more genetically advanced tumour. One of the pathways proposed follows a route of well-differentiated DCIS to invasive carcinoma and involves an initial loss of 16q and does not include a poorly differentiated DCIS step, which may thus explain the above findings of Roylance et al. (1999), who found a greater number of losses of 16q in well differentiated tumours compared to poorly differentiated tumours.

The findings of a study by Moore et al. (1999), which analysed poorly differentiated DCIS also agrees with this model, as 22 of the 23 tumours studied contained amplifications at 11q13. However, although the proposed routes of progression are generally accepted, there are still problems with the theory, as only 50% of 38 DCIS cases analysed in the study by Buerger et al. (1999) seemed to follow the model. Also the model cannot clearly explain intermediate differentiated DCIS, which exhibits a mixed pattern of both well and poorly differentiated DCIS (Van Diest, 1999) and more information is required on the progression routes of DCIS to invasive carcinoma.
Figure 1.4. A Hypothetical Model for the Progression of DCIS

Figure 1.4. Legend: A hypothetical model of the pathogenesis of DCIS (Buerger et al., 1999). Three main genetic routes of DCIS progression are proposed, the first involves a loss of chromosome 16q and results in well-differentiate DCIS that may progress to invasive carcinoma with or without an intermediate differentiated step associated with a gain at 1q. The second has an amplification at 11q13 resulting in intermediate DCIS that may develop into invasive carcinoma with or without a poorly differentiated step. The final pathway is associated with amplifications at 11q13 and 17q12 resulting in poorly differentiated DCIS which may progress to invasive carcinoma.
1.5.2. Chromosomal Aberrations in Breast Cancer

Breast carcinoma arises due to changes in the genome of normal cells. These different genomic alterations or mutations may be divided into four main types:

1. Small sequence changes

Small sequence changes involve base pair substitutions or deletions, or the insertion of a few nucleotides. Genetic aberrations of this type are thought to be rare in human cancers, but if present they may cause major problems (Lengauer et al., 1998). Two good examples of small sequence changes in breast cancer are mutations in $p53$ or the $BRCA$ genes. The $p53$ tumour suppressor gene is mutated in approximately 50% of all tumours and 15-34% of breast cancers (Hartmann et al., 1997). The $p53$ protein is a transcription factor involved in the arrest of DNA replication when DNA is damaged, allowing either repair or directing the cell to apoptosis (Phillips et al., 1999). The most common form of mutations in the gene are missense mutations resulting in altered protein residues which may cause its inactivation (Levine, 1997). Germ-line mutations in the $BRCA1$ and $BRCA2$ genes, as previously discussed, predispose to breast cancer. Mutations within the $BRCA1$ gene are distributed throughout the coding region and may be of different kinds including small deletions of 1 or 2 base pairs causing a missense mutation which results in an altered protein product (Arver et al., 2000).

2. Chromosomal translocation

Chromosomal translocation results in rearrangement of segments of chromosomes. These translocations are thought to be either random, thus unrelated to translocations observed in other tumours of the same histological subtype, or specific whereby the same chromosomal segments are rearranged in the same tumour types. Specific chromosome translocations have been described in certain breast tumour cells, e.g. ZR75-1 and T-47D cells have break points at 8p12, which may be followed by subsequent rearrangements (Courtay-Cahen et al., 2000). Such rearrangements may lead to the activation of an oncogene via positioning of the gene near a strong promoter or its fusion with another gene (Lengauer et al., 1998) or could result in inactivation of tumour suppressor genes. Neighbouring genes may also be affected and can be switched on or off, or made unstable by the presence or absence of the translocated gene (Cairns, 1981).
3. Aneuploidy

Aneuploidy involves alterations to chromosome number involving losses or gains of whole chromosomes and is thought to be a common event in tumours. Aneuploidy in breast cancer is most commonly associated with an overall DNA content gain, whilst hypodiploidy, where cellular DNA is content is less than normal cells, has an incidence of 1-2% in breast tumours (Tanner et al., 1998), but may be associated with a more aggressive disease and poorer prognosis (Ferno et al., 1992). Tumours with high a degree of aneuploidy compared to those that are diploid or weakly aneuploid may also be more aggressive (Beerman et al., 1990). Thus aneuploidy is thought to be a late event in tumour development that is preceded by other chromosomal arrangements and oncogene amplification (Rennstam et al., 2001). However there is some confusion as aneuploidy has been recorded in early breast lesions, such as DCIS (Ottosen et al., 1995) and some believe that aneuploidy is an early somatic mutation that precedes cancer (Li et al., 2000).

4. Gene amplification or deletion

Amplifications of single genes or a region of the chromosome up to 10 Mb in length known as an amplicon can occur late in tumourgenesis (Ingvarsson, 1999; Lengauer et al. 1998). Some gene amplifications are associated with tumour progression or aggression (Ried et al., 1999) while others may have prognostic significance (Ried et al., 1995). For example amplification of the ERBB2 proto-oncogene, which lies on chromosome 17, occurs in approximately 10-34% of breast tumours (Revillion et al., 1998) and is associated with a poor prognosis (Slamon et al., 1987). Other amplifications in tumour cells occur in drug resistance, whereby upon exposure to a drug cells amplify genes such as the ABCB1 transporter, which confer a resistance to the cell (Gottesman and Pastan, 1993). However in 5 X 10^8 normal non tumour cells that were selected in tissue culture with 3 drugs no amplification events were observed (Wright et al., 1990). This is supported by Tlsty et al. (1989), who demonstrated that the rate of spontaneous amplification is significantly lower in non tumourigenic cell lines than in tumourigenic cell lines. When cells were exposed to the cytotoxic drug N-(phosphonoacetyl)-L-aspartate, resistance occurred much more readily in the tumourigenic cell lines (Tlsty et al., 1989) and resistance to this drug has only been reported to occur through amplification of the CAD gene (Stark and Wahl, 1984).

Gene amplification plays an important role in the progression and initiation of breast cancer (Barlund et al., 2000). There are many specific amplifications involved in breast cancer and these are discussed in greater detail in Chapter 5-7. In contrast DNA amplifications are
thought to be rare events in ‘normal’ cells (Seshadri et al., 1987). This is supported by measures of mutation rates in malignant human lymphocytes, which are much greater than those in normal human lymphocytes (Seshadri et al., 1987).

1.5.3. Mechanisms of Chromosomal Aberrations

The exact mechanisms that give rise to chromosomal aberrations are unknown, although there are a number of theories. Mutations may be caused by external factors such as carcinogens binding to the DNA, or via spontaneous mutation during DNA replication (Cohen and Ellwein, 1991). To prevent replication of cells with damaged DNA, cell cycle progression is regulated at many checkpoints, such as DNA damage or cell cycle progression checkpoints (Paulovich et al., 1997). However these may fail and DNA damage checkpoints are thought to be involved in chromosomal aberrations such as translocations. Although the exact molecular basis of chromosomal translocations is unknown, it has been suggested that they occur in cells, which enter mitosis before recombination-promoting double strand breaks are repaired (Paulovich et al., 1997). Candidate genes that may also be responsible for chromosome translocations are genes involved in double strand break repair, such as p53, ATM or the BRCA genes. The tumour suppressor gene p53, is also thought to be involved in other chromosomal aberrations, since gene amplifications occur more frequently in cells where p53 is inactivated (Livingstone et al., 1992). An example of a cell cycle checkpoint is the spindle assembly checkpoint that ensures chromosomes do not separate until they are aligned correctly along the mitotic spindle (Elledge, 1996), and failure of this may result in aneuploidy and abnormal chromosome separation.

1.6. Techniques for Studying Genomic Alterations

Detection and mapping of chromosomal alterations is an important step for associating aberrations with disease phenotype and for localising critical genes involved in carcinogenesis. There are several methods available to analyse genomic alterations (Popescu and Zimonjic, 1997) some of which are explained here.

1.6.1. G-banding

Banding procedures allow identification of individual chromosomes within a species by producing a pattern of longitudinal banding along the length of each chromosome. Thus individual chromosomes can be recognised and large aberration events may be detected. Giesma (G-) banding is one of the most useful and popular banding procedures (Trent and Thompson, 1987). However chromosome banding has limits, since due to poor sensitivity
many abnormalities cannot be resolved by this technique, as small regions that appear abnormal may be difficult to identify (Trent and Thompson, 1987). The technique is also labour intensive and can be accomplished and interpreted only by skilled observers (Pinkel et al., 1986).

1.6.2. Fluorescent in situ Hybridisation (FISH)

Fluorescent in situ hybridisation (FISH) is a more sensitive approach than G-banding and deletions or amplifications of only a few kilobases may be detected (Ommen et al., 1995). FISH can either be mono labelled, whereby tumour DNA probes are labelled and hybridised to unlabelled normal DNA, or more recently dual labelled, which involves simultaneous in situ hybridisation of tumour and normal DNA probes to normal metaphase or interphase chromosomes. The tumour DNA is labelled with biotin and the normal DNA with digoxygenin, after hybridisation both probes are visualised with fluorochromes which in this case render the tumour green and normal DNA red. The intensity of each fluorochrome is calculated and the relative ratios correspond to relative copy number of each sample probe (Liang et al., 1995). A primary problem with FISH is that it is restricted to the use of probes targeted to specific chromosomes or regions, thus it is not a global approach and its main use is as a tool to characterise previously identified aberrations (James, 1999).

1.6.3. Comparative Genomic Hybridisation (CGH)

Comparative genomic hybridisation (CGH), which was first described by Kallioniemi et al. (1992), is a more global approach to genome analysis and has the ability to give an overview of chromosomal amplifications and deletions throughout the entire genome of a sample in a single hybridisation experiment. Differentially fluorescent-labelled test and control genome samples are simultaneously hybridised for 3-4 days to normal metaphase chromosome spreads. The ratio of the two bound fluorescent samples, which gives an indication of gene amplifications or deletions present, is then calculated. However, due to the highly condensed structure of metaphase chromosomes the technique has limitations, as CGH is unable to detect copy number changes in highly condensed regions of chromosomes. Thus deletions of less than 10 - 12 Mb are difficult to resolve (Bentz et al. 1998) and up to 2 Mb is thought to be the mathematical limit for amplification detection (Piper et al. 1995), so resolution of small or closely spaced aberrations is difficult. Although amplifications of small regions of chromosome may be detected if they have undergone high-level amplification, e.g. 5 fold amplification of a 1Mb region (Forozan et al., 1997). Also when analysing tumour samples CGH will be able to identify chromosomal regions, which may contain putative tumour
suppressor genes or oncogenes involved in carcinogenesis, but it will not be able to identify the actual genes themselves (Takeo et al. 2001). Another flaw with the technique is that it is not able to provide quantitative information about copy number change, and it is insensitive to structural aberrations that do not result in a DNA copy number change (Gray and Collins, 2000).

1.6.4. Real-Time Quantitative PCR (RT-PCR)

Traditionally, Southern blotting was utilised for quantitative analysis of gene copy number, whereby several genomic DNA populations are initially digested, then electrophoresed to resolve the fragments on the basis of size. The double stranded DNA fragments are then denatured and transferred by capillary action to a membrane. A specific gene or labelled probe is then hybridised to the membrane and the relative abundance of the gene can be calculated. However Southern blotting is labour intensive and an alternative approach is to use real-time quantitative PCR (Real Time PCR) (Heid et al., 1996; Yoshida et al., 2001). Real Time PCR can be used to measure either genomic or RNA expression changes of a gene.

In Real Time PCR gene specific primer pairs are designed and an oligonucleotide probe is created. The probe contains two fluorescent dyes, one of which is a reporter dye, and is annealed to the target DNA, when hybridised the two fluorescent dyes are in close proximity to one another and a quenching effect is observed. A PCR reaction using the specific primers is carried out that displaces the probe and the subsequent release of the reporter fluorescence is measured (Bustin, 2000). The technique is very sensitive for the detection of low abundance mRNA (Bustin, 2000), although the disadvantage is that genes are analysed individually and primers and probes have to be designed for each gene. In addition the cost associated with the production of TaqMan probes is high (Gray and Collins, 2000). Alternatively a SYBR green incorporation reaction that does not require a TaqMan probe may be carried out. SYBR green is a dye that binds to the minor groove of double-stranded DNA in a sequence independent way, and upon binding its fluorescence increases over 100 fold (Lekanne Deprez et al., 2002).

Southern blotting and Real Time PCR with SYBR green have been carried out in this report to confirm microarray findings, Real Time PCR has been utilised to confirm RNA expression data and Southern blot to confirm genomic changes.
1.6.5. Microarray Comparative Genomic Hybridisation

The most recent method for the analysis of changes in DNA copy number involves comparative genomic hybridisation coupled to cDNA microarrays (Muller, 2001; Pollack et al. 1999; Pinkel et al. 1998; Solinas-Toldo et al., 1997). This technique follows the same principles as the utilisation of microarrays to analyse gene expression, discussed in Section 1.7. Basically two genomic DNA populations for comparison are uniformly labelled with different Cyanine (Cy) dyes and hybridised to EST clones of open reading frames arrayed on a cDNA microarray. The microarrays are then scanned and ratio of the two fluorescent dyes represents the relative abundance of the gene. This technique has the potential to study global genomic changes although the limiting factor is the material printed on the arrays, which is probed by the two genomic DNA populations. Another disadvantage is that it only compares the presence of genetic material between populations and so does not have the ability to detect other aberrations such as chromosomal translocations or rearrangements. However high copy number amplifications of as little as 40 kb and losses from 75-130 kb have been resolved using this technique (Solinas-Toldo et al., 1997).

Gene based approaches such as these have been successful in breast cancer therapy and have led to the development of several classes of drugs (Bange et al., 2001). Microarray technology offers high throughput screening to aid identification of such targets and may in the future lead to tailored therapies for the treatment of a clinically diverse disease. This project utilises microarray technology to delineate genetic changes occurring in breast carcinoma with the aim to relate specific changes with phenotype.

1.7. DNA Microarrays

Gene microarrays are more usually associated with the determination of gene expression, however as previously discussed a number of studies have now successfully employed them in the determination of genomic change (Pollack et al. 1999; Solinas-Toldo et al., 1997). Gene expression is analysed by monitoring messenger RNA (mRNA) levels, as the genomic DNA is transcribed into RNA and later translated into functional protein products. Thus RNA is an intermediate step in the development of proteins, and the analysis of RNA expression provides information on the activity of genes, cell state and in general for most genes, changes in levels of RNA are related to changes in levels of protein products (Lockhart and Winzeler, 2000). There are also techniques available to measure the functional protein products including western blots, two-dimensional gels and mass spectrometric detection. Proteomics is a good approach as it is a measurement of the final expression product, and
some of the techniques can reveal information on post translation protein modifications, protein – protein interactions and localisation of the proteins within the cells (Pandey and Mann, 2000), although these protein based approaches are generally thought to be more difficult to perform and less sensitive (Lockhart and Winzeler, 2000).

There are a number of ways to monitor gene expression or mRNA abundance. These include differential display (Liang and Pardee, 1998), subtractive hybridisation cloning (Duguid and Dinauer, 1990), representational difference anlaysis (RDA) (Lisitsyn et al., 1993), suppression subtractive hybridisation (SSH) (Diatchenko et al., 1996), reverse transcription polymerase chain reaction (RT-PCR), northern blots and serial analysis of gene expression (SAGE) (Velculescu et al., 1995). However arguably the most valuable tools for studying gene expression are the nucleic acid arrays, as the potential mass of data and information that can be generated within a short period is immense.

1.7.1. Practical Applications of Microarrays

Most genes have multiple effects; the ability of the arrays to monitor many genes simultaneously makes it a powerful tool that enhances the understanding of the working cell. In drug discovery and development, the comparison of expression of thousands of genes between 'disease' and 'normal' tissue and cells allows the identification of multiple protein targets (Debouck and Goodfellow, 1999). In medicine microarrays also have the potential to assist in diagnosis by linking particular gene expression patterns with pathological disease states (Martin et al., 2000; Perou et al., 1999). In drug treatment microarrays have potential use for the investigation of mechanisms of drug action and may be used to screen drugs and determine the potential toxicity of a compound (Waring and Ulrich, 2000), as well as provide an ongoing picture of phenotypic alterations during treatment.

Microarrays have been used to study various cancers and have identified genes related to carcinomas of the ovary (Ono et al., 2000), colon (Backert et al., 1999), skin (Bitter et al., 2000), breast and prostate (Mousses et al., 2001), although the picture is far from complete. However it is thought that the results from microarray experiments may be able to aid in the classification of tumours, delineate the progression of cancer based on the gene expression patterns, or identify new molecular pathways and areas that may be involved in tumourgenesis (Wang et al., 1999). In turn this information will potentially lead to the identification of new targets for tumour diagnosis and treatment in the future.
1.7.2. Different Options Available for DNA Microarrays

DNA arrays are not a new concept and have been crudely used for a number of years (Gillespie and Spiegelman, 1965). However the availability of expressed sequence tag (EST) collections and recent technological and robotic advances have seen the technique flourish, and it is now possible to determine the relative expression ratios of thousands of genes in a single experiment. The basis of DNA arrays is that DNA forms a duplex between sequence complimenting strands, which allows the hybridisation of labelled RNA or DNA in solution to corresponding DNA molecules attached to a support. There are a number of different arrays currently in use, both available commercially or constructed ‘in house’. These arrays may differ from the material printed onto the support, to the type of support itself and sample preparation (Bowtell, 1999).

1.7.2.1. Targets Printed onto the Array

There are two main types of target printed onto the array, synthesised oligonucleotides or cDNA clones. The oligonucleotide arrays may be divided into two groups depending on how they are produced; firstly the oligos may be synthesised on the arrays themselves (Lipshutz et al., 1999). For example Affymetrix use photolithography to fabricate their arrays, whereby a mercury lamp is shone through a photolithographic mask onto the array surface, oligos which are not protected by the mask are activated resulting in a 5’hydroxy group. The array is then incubated with the desired nucleotide and chemical coupling occurs at the activated positions. This process is repeated with different masks until the desired set of oligonucleotides, up to 30 bases in length, are obtained on the array (Lipshutz, 1997). Secondly oligonucleotides may first be synthesised using standard controlled pore glass (CPG) synthesis (Damha et al., 1990) and then subsequently printed onto the array.

Oligonucleotide arrays have the advantage that there is no need for clones, PCR products or sequencing of cDNAs. Additionally oligonucleotides are short chains with single points of constraint at each chain end, and thus they may be far more accessible for the labelled probes to hybridise to. However there is a reduction in sensitivity due to limiting the number of nucleotides from hundreds to 20 – 70 (Duggan et al., 1999). As well as monitoring gene expression, practical uses for oligonucleotides arrays include detection of sequence variations in genomic DNA including single nucleotide polymorphisms (SNPs) (Lindblad-Toh et al., 2000; Hacia, 1999), which are point mutations that constitute the most common type of genetic variation, and can be contributory factors to human disease and may be used as genetic markers (Gilles et al., 1999).
CDNAS may also be printed onto the array as either full length cDNAs, partial sequences of cDNA fragments commonly called expressed sequence tags (ESTs), or randomly chosen unknown cDNAs from a library of interest. The EST system has been widely used for arrays based on higher eukaryotes, where as for yeast and prokaryotes DNA targets are usually generated by amplification of genomic DNA with gene specific primers (Duggan et al., 1999). There are a number of advantages of cDNA arrays, as opposed to oligo arrays. The library approach has the ability of being able to detect novel changes in gene expression, since unknown cDNAs are also incorporated on the array. The technique has the potential of being able to spot whole sequenced genomes onto an array, thus allowing genome wide expression analysis (Brown and Botstein, 1999).

There are different options available for spotting the cDNA onto the microarray, but in general a few nano litres of the purified fragments or oligos are deposited onto the array surface, creating spots in the region of 100-150 μm in diameter (Cheung et al., 1999). Printing can either be carried out by contact printing or non contact printing. Contact printing generally involves a pin with a narrow split at the pointed end. The cDNA solution is taken up into the split by capillary action and when the pins are briefly touched to the array surface a small amount is deposited. Non-contact printing utilises either piezo or inkjet devices which spray the cDNA onto the slide (Marshall and Hodgson, 1998).

1.7.2.2. The Array Solid Support

There are two main types of solid support for arrays, either glass microscope slides or porous membranes. Membrane arrays can be produced in house or there are a variety of filter arrays commercially available e.g. the Clontech filters, which include only known genes and are grouped depending on their involvement in different processes. The types of membranes commonly used for these arrays include nitrocellulose and positive charged nylon commercial varieties (Duggen et al., 1999). The two RNA populations are labelled with radioisotopes such as $^{33}\text{P}$ or $^{32}\text{P}$ and hybridised to separate duplicate membranes, which are subsequently normalised and compared. Colorimetric nylon arrays that do not use hazardous radioisotopes have also been described (Chen et al., 1998).

Filters are useful as smaller quantities of RNA are required, about 50 ng of total RNA per experiment (Rihn et al., 2000). The sensitivity of nylon arrays are thought to be limited to high and medium abundance genes only (Pietu et al., 1996), though this has been disputed, as
nylon arrays are thought to bind larger amounts of target material as opposed to glass, which would increase sensitivity (Bertucci et al., 1998).

Glass microarrays assay differential gene expression or genomic change by hybridisation of fluorescently labelled probes prepared from different RNA or DNA sources. Different fluorochromes are used and labelled RNA or DNA is co-hybridised to the same slide. The relative ratio of hybridisation can be calculated from the distribution of the fluorescent change. Glass is an impermeable and rigid support, which has a number of advantages over porous membranes (Southern et al., 1999). As it is non-porous, liquid is unable to penetrate the surface, thus labelled nucleic acids can hybridise easily to the target without diffusing into the pores. Similarly the washing steps are not impeded by diffusion; also as it is non porous the hybridisation volume can be kept to a minimum, enhancing the kinetics of annealing probes to the targets (Cheung et al. 1999). The printed nucleotide location on glass is also better defined than on flexible membranes, due to the flatness and rigidity of the support, this is crucial for the small ‘spotting’ capability of microarrays. However, the main advantage is the competitive nature of the hybridisation, as both probes may be co-hybridised to the same target array. The main disadvantage of the glass slide and fluorescent probe approach is that for adequate fluorescence the amount of total RNA required is about 50 μg.

1.7.2.3. The Labelled Probe

There are two main approaches to labelling the probes, either with fluorescent dyes, mainly used in dual labelling with glass microarrays, or with radiolabels mainly used with nylon arrays. For expression detection the starting material is normally RNA, which can be either total RNA or polyA RNA. As with many RNA based assays, the purity and quality of the starting RNA has a significant effect on the experiment. For detection of amplifications and deletions genomic DNA is the starting material as in the study by Pollack et al., (1999) who hybridised DNA from a number of tumour breast cell lines and were able to characterize some of the DNA copy number changes occurring in cancer. The RNA labelling reaction generates cDNA most commonly using a polyT primer from the mRNA polyA tail, while genomic DNA is primed randomly with hexamers. During the reaction fluorescent or radio labelled dNTPs are incorporated. The fluorescent labelled markers most commonly used are nucleotides labelled with either cyanine 3 (Cy3) or cyanine 5 (Cy5) (Figure 1.5.), as these dyes have high incorporation efficiencies, good photostability and are widely separated in their excitation and emission spectra (Duggan et al., 1999).
Figure 1.5. Legend: Structures of the cyanine labelled nucleotides. RNA populations are labelled with either Cy3 dUTP or Cy5 dUTP. DNA populations are labelled with either Cy3 dCTP or Cy5 dCTP. Cy3, which fluoresces green is normally used to label the control, whilst Cy5, which fluoresces red is used to label the ‘test’. However reverse labelling to confirm results is recommended.
1.7.3. The Microarray Facility at MRC Toxicology Unit, Leicester

The microarrays used here at Leicester are cDNA clone arrays printed onto glass microscope slides. The microarrayer itself was constructed on site following the ‘Stanford’ design (http://cmgm.stanford.edu/pbrown/mguide/index.html). It has a 16 pin print head and has the capacity to print 137 glass slides at one time. The pins have slots in them at the pointed end and rely on capillary action to take up the cDNA solution from 384 well plates and to spot a few nano litres of the solution onto each slide (Figure 1.6.). The approximate number of clones currently in our human collection is 6000, 4000 of which are from a Research Genetics sequenced verified set, the others were picked from the I.M.A.G.E collection at the Sanger Centre in Cambridge, UK.

Once cDNA clones are robotically printed onto poly-L lysine glass coated microscope slides, as first described by Schena et al., (1995), two RNA populations are labelled with either 5-amino-propargyl-2’-deoxyuridine 5’-triphosphate coupled to Cy3 fluorescent dye (Cy3 dUTP) or 5-amino-propargyl-2’-deoxyuridine 5’-triphosphate coupled to Cy5 fluorescent dye (Cy5 dUTP) (Figure 1.5.). Or two DNA populations are labelled with 5-amino-propargyl-2’-deoxycytidine 5’-triphosphate coupled to Cy3 fluorescent dye (Cy3 dCTP) or 5-amino-propargyl-2’-deoxycytidine 5’-triphosphate coupled to Cy5 fluorescent dye (Cy5 dCTP) (Figure 1.5.). A dual hybridisation method is carried out on the array, with both RNA or DNA test and control being hybridised together on the same slide. The hybridised slides are scanned with lasers and Cy3 fluoresces green when excited at a wavelength of 532 nm and Cy5 fluoresces red at 635 nm. Therefore if both incorporated dyes and so both populations of RNA hybridise to the array at equal concentrations a yellow colour will be observed, if one is greater than the other varying ratios of green or red are seen (Figure 1.6.).
Figure 1.6. Diagrammatical Representation of the Hybridisation Process

RNA 1

Both RNA populations are labelled with different Cy dyes

RNA 2

Labelled RNA populations are hybridised to the array

cDNA clones printed onto the array

Glass slide the cDNA array is printed on

RNA hybridises to corresponding cDNA on the array

The array is scanned, differentially expressed RNAs are visualised

Raw data is normalised and analysed

RNA is more abundant in the Control
RNA is more abundant in the Test
RNA present in equal amounts in the Test and Control: Red + Green = Yellow

Figure 1.6. Legend: Overview of the microarray procedure. Firstly the two populations of RNA to be compared are labelled separately with the different fluorescent Cy labelled nucleotides. Both RNA populations are then hybridised together to the cDNA array. The array is washed and scanned. Differentially expressed genes are identified, by analysing the abundance of Cy3 and Cy5 in each hybridisation feature on the microarray. If RNA from both populations is present in equal abundance then simple colour chemistry dictates that: red + green = yellow.
1.8. Thesis Aims
This project aimed to utilise microarray technology and bioinformatics to gain genetic information on breast carcinoma in the hope of increasing our understanding of the disease by linking genetic changes with phenotype. The 2 main aspects examined were genetic alterations associated with drug resistance (Chapters 3 and 4) and genomic alterations occurring in breast cancer (Chapters 5-7), both of which may affect the susceptibility of the tumour to a variety of xenobiotics.

1.8.1 Analysis of the Drug Resistance Phenotype
A variety of drug sensitive and drug resistant breast carcinoma cell lines were exposed to the anti cancer drug doxorubicin and their relative sensitivity to the drug was calculated. RNA was extracted from these different breast carcinoma cell lines that showed varying resistance to Dox, and was analysed by microarray hybridisation. The aim was to build up a profile of gene expression patterns associated with the drug resistant phenotype, which could be mined for mechanisms and gene interactions that may lead to further elucidation of drug resistance. Both intrinsic and acquired multi drug resistance was studied.

Sensitive cell lines were also continuously exposed to Dox and in one cell line resistance to Dox developed. RNA and DNA were extracted and microarray analysis was carried out at various stages of resistance development with the aim to discern the spatial relationship of genes during the development of resistance.

1.8.2 Genomic Amplifications in Breast Cancer
Various genomic aberrations including amplifications occur during the development of cancer. Tumour DNA and RNA from a panel of human breast cancer cell lines was labelled and hybridised to the microarrays against non tumour DNA and RNA control with the aim to discover common and frequent chromosomal areas or specific genes that are amplified in breast cancer. The RNA hybridisations were carried out to see if these genomic amplifications are also significant at the expression level.

A number of in vivo clinical tumour samples exhibiting different phenotypes were also analysed for genomic amplifications, to discern specific gene amplifications or regions of gene amplification common to the different tumours, or specific for certain breast tumour grades or type.
A specific region of amplification was detected on chromosome 17, which was present in both a cell line and clinical breast tumour samples. The region contains the ERBB2 oncogene gene that is a known indicator of poor prognosis. This region was analysed in greater detail in the attempt to gain information on the surrounding genes involved.

The data generated within this thesis provides further insight into genetic changes that occur in phenotypically different breast cancer types, as well as elucidating some of the changes that lead to the development of the drug resistant phenotype.
Chapter 2

Materials and Methods
2. Materials and Methods

2.1. Buffers and Solutions

2.1.1. Agarose Gel Electrophoresis

5 x TBE Buffer
446 mM Tris base
445 mM Boric acid
12 mM EDTA
Made up to the appropriate volume with distilled H$_2$O and stored at room temperature.

Loading Dye
15 % Ficoll-400
0.25 % xylene cyanol FF
0.25 % bromophenol blue

1 kb ladder
4 ml 1x TE buffer
1 ml 1 kb ladder
1 ml 6 x loading dye

2.1.2. Poly-L Lysine Coating of Slides

Cleaning solution
70 g NaOH dissolved in 280 ml H$_2$O
420 ml 95 % EtOH

Poly-L Lysine solution
70 ml poly-L Lysine
70 ml PBS
560 ml H$_2$O

2.1.3. Printing of Slides

Printing Solution 1
3 X SSC
4.5 µl of 20 X SSC added to 25.5 µl distilled H$_2$O per well. Printing solution used for print runs labelled: Hs990803

Printing Solution 2
3 X SSC
0.1 % N-lauryl sarcosine (sarcosyl)
For print runs Hs000315 – Hs010509. Sarcosyl was added to each well, as this is a detergent that increases the homogeneity of the cDNA spots and aids in printing.

**Printing Solution 3**

3 X SSC
1.5 M Betaine Monohydrate
For latter print runs. Betaine was added to the print solution as it increases the viscosity of the solution and reduces the rate of evaporation, which increases spot homogeneity (Diehl et al. 2001).

### 2.1.4. Cell Culture

**Trypsin/EDTA (10 x)**

- 0.5 % Trypsin
- 0.2 % EDTA
- 0.85 % NaCl

Made up to the appropriate volume with PBS and stored at -20 °C.

**Freezing Down Media**

- 70 % (v/v) Appropriate cell medium
- 22.5 % (v/v) Heat inactivated fetal calf serum
- 7.5 % (v/v) ml DMSO

### 2.1.5. RNA Extraction from Cell Lines

All buffers supplied with the Qiagen RNA extraction kit.

**Buffer RLT (Cell Lysis Buffer)**

**Buffer RW1**

**Buffer RPE**

### 2.1.6. DNA Extraction from Cell Lines

All buffers supplied with the Qiagen Genomic DNA extraction kit.

**Buffer C1 (Cell Lysis Buffer)**

- 1.28 M Sucrose
- 40 mM Tris/HCl pH 7.5
- 20 mM MgCl₂
- 4 % Triton X-100

Buffer C1 should be stored at 4 °C.
**Buffer G2 (General Lysis Buffer)**
- 800 mM Guanidine HCl
- 30 mM Tris/HCl pH 8.0
- 5 % Tween-20
- 0.5 % Triton X-100

**Buffer QBT (Equilibration Buffer)**
- 750 mM NaCl
- 50 mM MOPS pH 7.0
- 15 % Isopropanol
- 0.15 % Triton X-100

**Buffer QC (Wash Buffer)**
- 1 M NaCl
- 50 mM MOPS pH 7.0
- 15 % Isopropanol

**Buffer QF (Elution Buffer)**
- 1.25 M NaCl
- 50 mM Tris/HCl pH 8.5
- 15 % Isopropanol

**2.1.7. RNA Labelling for Microarray Analysis**

**dNTP Master Mix**
- 20 mM dGTP
- 20 mM dCTP
- 20 mM dATP
- 8 mM dTTP
  
  Made up to the appropriate volume with distilled H$_2$O and stored at –20 °C.

**20 x SSPE**
- 3 M NaCl
- 1 mM NaH$_2$PO$_4$
- 20 mM EDTA
  
  The pH was adjusted to 7.4 with NaOH, made up to the appropriate volume with H$_2$O and stored at room temperature.

**50 x Denhardt's solution**
- 1 % BSA
- 1 % PVP
1 % Ficoll
Made up to the appropriate volume with distilled H₂O and stored at − 20 °C.

**Prehybridisation Buffer**
- 5 ml Deionized formamide
- 3 ml 20 X SSPE
- 0.5 ml 10 % SDS
- 1 ml 50 X Denhardts
- 100 µl Sigma sonicated Salmon Sperm DNA (10mg/ml)
- 0.4 ml distilled H₂O

The solution was filtered through a 0.45µ syringe filter and stored at − 20 °C.

**Hybridisation Buffer**
- 5 ml Deionised formamide
- 500 µl 50 x Denhardts solution
- 1 ml distilled H₂O
- 500 µl 10 % SDS

The solution was filtered through a 0.45µ syringe filter and stored at − 20 °C.

### 2.1.8. Genomic DNA labelling

**OLB Master Mix (Feinberg and Vogelstein (1983))**

Also used for labelling DNA for Southern and Northern Blot analysis

A:B:C mixed at ratios 100:250:150

**Solution A**
- 1 ml 1.25 M Tris pH 8.0, 0.125 M MgCl₂
- 18 µl β-mercaptoethanol
- 5µl dATP 0.1 M
- 5µl dGTP 0.1 M
- 5 µl dTTP 0.1 M

**Solution B**
- 2 M HEPES pH 6.6 (bought to pH with NaOH)

**Solution C**
- Hexadeoxyribonucleotides dissolved in TE to 90 OD U/ml

### 2.1.9. Gridding Nylon Membranes

**Denaturing solution**
1.5 M NaCl
0.5 M NaOH
Made up to the appropriate volume with distilled H$_2$O and stored at – 20 °C.

Neutralising solution
1.5 M g Na Cl
121.1 g/L Trizma Base
Adjusted to pH 7.0 with concenratd HCl and Made up to the appropriate volume with distilled H$_2$O and stored at – 20 °C.

2.1.10. Nylon Array and Southern Blotting

Hybridisation solution
5 x Denhardts
6 x SSC
50 % Deionised formamide
1 % SDS
Prepared fresh. SDS was the last addition to the mix from a 10 % stock.

2.1.11. Protein Extraction

Cell Lysis Buffer
50 mM Tris (pH 7.5)
% (w/v) β-mercaptoethanol
5 mM EDTA
10 mM EGTA
1 mM PMSF
10 mM Benzamidine
Made up to the appropriate volume with distilled H$_2$O and stored at – 4 °C. Protease inhibitors leupeptin, aprotinin and pepstatin A were added prior to use, to give a final concentration of 5 μg/ml each.

2.1.12. Western Blotting

Separating Gel (8 %)
H$_2$O 2.6 ml
0.75 M Tris pH 8.8 7.5 ml
30 % Acrylamide 4 ml
2 % SDS 750 μl
10 % APS 105 μl
100 % TEMED 4.5 μl

The APS and TEMED are added last and the gel is poured immediately.
The % of the gel used depends on the size of protein, the larger the protein the lower the %.
To adjust the %, the water and acrylamide volumes were altered.

Stacking Gel
H₂O 6.2 ml
0.5 M Tris/0.4 % SDS 2.5 ml
30 % Acrylamide 1.5 ml
10 % APS 30 μl
100% TEMED 7.5 μl

The APS and TEMED are added last and the gel is poured immediately.
To adjust the percentage, the water and acrylamide volumes were altered.

2 x Sample Buffer
125 mM Tris/HCl pH 6.8
20 % (v/v) Glycerol
4 % (v/v) SDS
50 μg Bromophenol Blue
200 mM DTT
Made up to the appropriate volume with distilled H₂O and stored at room temperature.
The DTT was added prior to use.

10 x Running Buffer
250 mM Tris base
1.92 M Glycine
1 % SDS
Made up to the appropriate volume with distilled H₂O and stored at 4 °C. Diluted to 1X with distilled H₂O prior to use.

Transfer Buffer
50 mM Tris base
40 mM Glycine
20 % (v/v) Methanol
Made up to the appropriate volume with distilled H₂O and stored at 4 °C.

0.1 % TBS-T
20 mM Tris/HCl (ph 7.6)
137 mM NaCl
0.1 % (v/v) Tween 20

Made up to the appropriate volume with distilled H₂O and stored at 4 °C.
2.2. Method Protocols

A simplified schematic overview of all the stages involved in the microarray experimental work carried out in this thesis is shown in Figure 2.1. The process is split into 6 main areas, clone preparation, slide preparation, sample preparation, hybridisation, data analysis and post microarray experiments.

**Figure 2.1. Overview of the Microarray Experimental Stages**

Figure 2.1. Legend: cDNA clones are prepared then spotted onto microscope slides that have been coated with poly-L lysine. Samples for labelling, either RNA or DNA are extracted and hybridised to the slides. The fluorescent data for each spotted clone is acquired and analysed and subsequent post microarray work is then carried out.

2.2.1. Preparation of Clones for Printing on the Array

The clones were selected and processed for printing on the array essentially as described by Eisen and Brown (1999) and Hedge et al. (2000). The clones initially printed on the array were from a 4K sequenced verified collection Research Genetics (http://www.resgen.com), which covered a broad spectra of genes. Additional clones were individually selected from the NCBI website via the Unigene database (http://www.ncbi.nlm.nih.gov/UniGene) and obtained from the IMAGE consortium at the HGMP, Cambridge, UK (http://www.hgmp.mrc.ac.uk). New clones are being continuously added to the array; at the time of writing this report there were a total of 6482 clones currently on the array, 5411 of these were unique, with an extra 1071 duplicate clones representing genes already present at least once on the microarray. The
process for preparation of the clones is summarised in Figure 2.2. and protocols for each stage are in subsequent sections.

Figure 2.2 Overview of Clone Preparation

![Clone Preparation Diagram](image)

Figure 2.2. Legend: Bacterial cDNA clones are grown overnight in liquid broth. Cultures are plasmid prepped and the cDNA fragments amplified by PCR. The products are precipitated, sequenced and resuspended in spotting solution prior to printing onto glass microscope slides.

2.2.1.1. Streaking and Growing clones

The clone collection from Research genetics arrived as glycerol stocks in 96-well plates, which were stored at -70 °C. 1 µl of these stocks was aliquoted into a 96-well block containing 1.8 ml liquid broth with ampicillin and grown overnight at 37 °C in a Gallenkamp incubator being continuously shaken at 250 rpm. The IMAGE clones arrived as bacterial plugs on agar. To obtain individual colonies they were streaked onto agar plates with ampicillin (50 µg/ml) and grown overnight at 37 °C in a Sanyo incubator. Individual colonies were picked for each clone, placed into 96-well blocks and grown overnight containing 1.8 ml liquid broth with ampicillin. The 96-well blocks contained clones of the same vector, thus
facilitating subsequent PCR amplification in 96-well plate format. The 96-well location assigned at this point was maintained throughout and served as the clone’s unique identification. 96-well location nomenclature was: species, plate number, 96-well location e.g. Hs056g09: Homo sapiens, plate 56 location g09.

2.2.1.2. Creating bacterial Glycerol Socks and Plasmid Preparations

Bacterial glycerol stocks of the IMAGE cultures were prepared. 200 µl of the overnight culture was aliquoted into the corresponding well of an appropriately labelled flat bottomed 96-well plate (Genetix). Glycerol was added to each well to a final concentration of 15 %. The glycerol bacterial stocks were stored at -80°C. The remaining overnight culture of the research genetics and IMAGE clones was pelleted for 10 min at 5000 rpm, and all excess liquid broth was removed. The bacterial pellets were given to the Protein and Nucleic Acid Laboratory Group (PNACL) Leicester University to be plasmid prepped. Plasmid preparations were carried out using a Qiagen 9600 Biorobot following the Qiaprep protocol.

2.2.1.3. PCR Amplification of Plasmid preparations

The plasmid inserts were amplified by PCR in their 96-well plate format. Primers used depended on the vectors the inserts were cloned into and were either Universal primers, pT7T3-Pac, Bluescript-SK, pcmvSPORT or pcmvSPORT6 primers, the sequences for which are show in Table 2.1. PCR taq master mix (2.0 mM MgCl) from ABgene was used. 5 ml of the AB taq master mix was diluted 1:1 with 5 ml distilled H₂O. 5000 pmol of the appropriate sense and antisense primers were added. The PCR mix was aliquoted into a 96-well plate, 80 µl /well, therefore primers at a final concentration of 40 pmols /reaction. 1 µl of each plasmid preparation was aliquoted into the appropriate well of the plate via an 8 multi channel pipette. The PCR amplification reactions were carried out in an MJ Research DNA engine Tetrad. The primer annealing temperature was 52 °C for universal primers and 56 °C for pT7T3-pac, Bluescript-SK, pcmvSPORT and pcmvSPORT6 and the PCR parameters used were:

<table>
<thead>
<tr>
<th>Step</th>
<th>Conditions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>92 °C for 5 min</td>
<td>Initial denaturation</td>
</tr>
<tr>
<td>Step 2</td>
<td>72 °C for 3 min 30 sec</td>
<td>Denaturation</td>
</tr>
<tr>
<td>Step 3</td>
<td>92 °C for 30 sec</td>
<td>Primer annealing</td>
</tr>
<tr>
<td>Step 4</td>
<td>52/56 °C for 45 sec</td>
<td>Extension</td>
</tr>
<tr>
<td>Step 5</td>
<td>72 °C for 3 min 30 sec</td>
<td></td>
</tr>
<tr>
<td>Step 6</td>
<td>steps 3 – 5 repeated 28 times</td>
<td></td>
</tr>
</tbody>
</table>

All amplified clones were stored at -20 °C.
Table 2.1. Primer Nucleotide Sequence

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal</td>
<td>Sense GTAAAACGACGGCCAGTG</td>
</tr>
<tr>
<td></td>
<td>Anti sense ACACAGGAAACAGCTATGACC</td>
</tr>
<tr>
<td>pT7T3-pac</td>
<td>Sense GGCCAGTGCCAAAGCTAAAAAT</td>
</tr>
<tr>
<td></td>
<td>Anti sense TCACTATAGGGAATTGGGCC</td>
</tr>
<tr>
<td>Bluescript-SK</td>
<td>Sense GACTCACTATAGGGCGAATTTG</td>
</tr>
<tr>
<td></td>
<td>Anti sense CTAGAACTAGTGATCCCCCG</td>
</tr>
<tr>
<td>PcmvSPORT</td>
<td>Sense AGAGCTATGACGTCGATGC</td>
</tr>
<tr>
<td></td>
<td>Anti sense AGGTACCGGTCCGGAATTC</td>
</tr>
<tr>
<td>PcmvSPORT6</td>
<td>Sense AGAAAGCTGGGTACGCGTAA</td>
</tr>
<tr>
<td></td>
<td>Anti sense AAAAGCAGGCTGGTGAC</td>
</tr>
</tbody>
</table>

2.2.1.4. Gel Electrophoresis of PCR Product

A 300 ml, 96-well 1% agarose gel made with 1 x TBE (Section 2.1.1) and 30 µl EtBr (10 mg/ml) was poured and allowed to set. 3 µl of each PCR reaction was placed into corresponding wells of a new 96-well plate with an 8 multi-channel pipetter and 3 µl of loading dye (section 2.1.1.) was added to each well. Samples were loaded onto the gel and 3 µl of ladder (section 2.1.1.) was loaded into the first well of each row. The gels were run at 150 V for approximately 1 hr 45 min in 1 x TBE running buffer. An image of the gel was created using a Flour-S™ MultiImager (BioRad), and images were stored on a central file server. Filenames used were species and plate number e.g. Hs045. The size of the insert sequence was determined from the agarose gel by comparison with the ladder using MultiAnalyst software (BioRad).

2.2.1.5. Precipitation of Amplified Clones

The PCR products were precipitated with 1/10 volume 3 M NaAc (pH 5.2) and a 1 volume of isopropanol. Reactions were incubated at -20°C for 1 hr to overnight to precipitate the cDNA. Reactions were then centrifuged at 2500 rpm for 30 min to pellet the amplified cDNA. Pelleted products were washed twice with 70 % EtOH, centrifuged at 2500 rpm for 5 min and left to air dry. Pellets could be stored at – 70 °C. In preparation for spotting onto the slides pellets were resuspended overnight on a shaker in printing solution (section 2.1.3.).
2.2.1.6. Sequencing of Clones

Sequencing of clones was carried out to confirm their identity. Sequencing reactions were performed with 1 µl of the PCR product using the Big Dye reaction (Perkin Elmer Applied Biosystems) with 3.2 pm of the appropriate anti sense primer. Sequencing reactions were carried out in an MJ Research DNA engine using the following parameters:

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>94 °C</td>
<td>30 sec</td>
<td>Initial denaturation</td>
</tr>
<tr>
<td>2</td>
<td>96 °C</td>
<td>10 sec</td>
<td>Denaturation</td>
</tr>
<tr>
<td>3</td>
<td>50 °C</td>
<td>05 sec</td>
<td>Primer annealing</td>
</tr>
<tr>
<td>4</td>
<td>60 °C</td>
<td>4 min</td>
<td>Extension</td>
</tr>
<tr>
<td>5</td>
<td>steps 2 – 4</td>
<td>repeated 30 times</td>
<td></td>
</tr>
</tbody>
</table>

Reactions were then purified by dyeEx spin columns (Qiagen) and purified reaction products were dried down and given to the Protein and Nucleic Acid Laboratory (PNACL) Leicester University. A sequencing gel and sequence analysis, were performed by PNACL using ABI PRISM 377 instruments.

Once clones were sequenced their identities were confirmed by running a sequence similarity search in BLAST (basic local alignment search tool) found on the ncbi website (http://www.ncbi.nlm.nih.gov/BLAST/). BLAST, is a set of similarity search programs, which explores all of the available sequence databases with your query sequence. BLAST utilises a heuristic algorithm, which seeks local as opposed to global alignment and is therefore able to detect relationships among sequences which share only isolated regions of similarity (Altschul et al. 1990). An example of the output file from BLAST is shown in Figure 2.3.

2.2.1.7. Transfer of Clones to 384-well Plates for Printing

Prior to printing the clones were sequentially transferred from 96-well plates into 384-well plates via a Hydra96 (Robins Scientific). 10 µl of the cDNAs from the 96-well plates were dispensed into Costar polypropylene 384-square well plates (Fisher Scientific) via the Hydra in a clockwise order around the plate. To ensure there was no cross contamination of cDNA the syringes were washed between each transfer with EtOH and nuclease free water. Plates were sealed and stored at – 20 °C between different print runs.
Figure 2.3 Legend: An output file of BLAST. BLAST generates firstly a diagrammatical view of the sequence alignment, the colour key for alignment scores is shown – with red being the best alignment. Also generated are gene identification numbers and the gene names of ‘hits’ found in the database. The blue score on the left gives an indication of the homology of match, since the higher the score the better the match. As seen for this search of clone HS043h02, it matches very well on a number of occasions with ribosomal protein L11.
2.2.2. Preparation of Slides

Prior to printing the DNA clones onto the surface, the slides were coated with poly-L lysine. In preparation for hybridisation the printed slides were then cross linked, blocked and denatured as described by Eisen and Brown (1999), a schematic diagram of these stages is shown in Figure 2.4.

Figure 2.4. Overview of Slide Preparation

![Slide Preparation Diagram](image)

Figure 2.4. Legend: Prior to printing the cDNA clones slides are firstly coated with poly-L lysine. The cDNA clones are then cross-linked to the slides, blocked then denatured prior to dual hybridisation of labelled samples.

2.2.2.1. Poly-L lysine Coating of Slides

Gold Star glass microscope slides were obtained from Chance Proper ltd. Slides were washed in cleaning solution (section 2.1.2.) on a rocker for 2 hr with the slides horizontal to the direction of rocking. Once slides were clean they were exposed to air as little as possible, since dust particles would interfere with the poly-L lysine coating and subsequent printing. The slides were rinsed vigorously by plunging up and down in distilled H₂O. The rinsing step was repeated 4 times with fresh distilled H₂O each time and slides were dried by centrifugation at 500 rpm for 5 min. The slides were transferred to polypropylene boxes and the poly-L lysine solution (section 2.1.2.) was added. Boxes were rocked on a shaker for 15 – 60 min with slides lying horizontal to the direction of rocking. The poly-L lysine coated slides were rinsed with distilled H₂O 5 times and dried by centrifugation. Slides were then transferred to covered slide boxes and left to age for at least two weeks to allow the surface to become hydrophobic prior to printing. The hydrophobicity of the coated slide surface is important in maintaining the small size of the printed DNA spots for the high-density microarrays.
2.2.2.2. Printing of Slides
The amplified cDNAs were robotically printed onto the poly-L lysine coated glass microscope slides, as first described by Brown (Schena et al. 1995). A 16 split pin arrayer, based on the Stanford design (http://www.le.ac.uk/cmht/twg1/array-images.html) was used (earlier prints were carried out using the 16 pin head) and 137 slides were printed at one time. Printing was achieved by capillary action whereby the split pin tips were placed in wells of the 384-well source plates; approximately 1 μl of cDNA printing solution was drawn up into the reservoir in the split of the tips via capillary flow. The tips were then lightly tapped at identical positions on the poly-L lysine surface of each slide, depositing approximately 0.5 nl of cDNA solution (Eisen and Brown, 1999). The pins were washed with distilled H₂O and blown dry between each sample. The ArrayMaker software version 1.8 from the DeRisi lab (UCSF) controlled the printing (http://derisilab.ucsf.edu/arraymaker.shtml). The average centre-to-centre distance between the cDNA spots or features was 200 μm. The date of the print determined the print name e.g. print Hs000315 was printed on 15th March 2000 and the slides for each print were numbered with the print date and slide number 1 - 137.

2.2.2.3. Post Processing of Printed Slides
To increase the amount of hybridisable DNA stably attached to the arrays, the DNA was firstly cross-linked to the to the slides by UV irradiation with a Stratagene Stradlinker set at 65 mJ. The cross-linked slides were then blocked to ensure that the remaining free lysine groups on the arrays are modified to minimise their ability to bind labelled probes. Finally the printed DNA is denatured to enhance hybridisation availability.

The boundary of the array for a given print was marked onto a blank slide prior to blocking to act as a template, as the array becomes invisible post-blocking. For slides printed after Sept 2001, blocking and denaturing was achieved by baking, which involved heating the slides for 2 min at 100 °C, followed by rinsing in distilled H₂O and drying by centrifugation for 4 min at 500 rpm, these slides were hybridised within 2 hr. Slides used before September 2001 were blocked using a method described by Eisen and Brown (1999) (http://cmgm.stanford.edu/pbrown/protocols/3_post_process.html). The procedure was performed in a fume cupboard and had to be carried out rapidly for blocking to be successful: 6.0 g of succinic anhydride was weighed out, 335 ml of 1-methyl-2-pyrrolidinone was added and stirred with a magnetic stirrer. Immediately after the succinic anhydride dissolved 15 ml of sodium borate was added and stirred for 5 sec. The solution was rapidly transferred to a glass chamber and the slides were plunged up and down 5 times and then left completely immersed in the solution for 15 – 20 min, during which time the chamber was gently rocked.
on a shaker with the slides horizontal to the direction of rocking. The printed DNA was
denatured to enhance hybridisation availability by plunging slides 3 times in 95 °C+ water,
followed by an incubation in the water for 2 min. Slides were then plunged 3 – 5 times in 95%
EtOH and dried via centrifugation for 5 min at 500 rpm. Prior to hybridisation blocked slides
were stored in the dark in plastic slide boxes at room temperature. These slides were
prehybridised with 15 µl of Prehybridisation buffer (section 2.1.5.) for 45 min to decrease
background signal and used within 24 hr.

2.2.3. Preparation of Samples for Microarray Analysis

The source of DNA or RNA for microarray analysis was either human breast tumour cell
lines or clinical breast tumour samples. The DNA and RNA was extracted and quantified
prior to labelling and hybridisation to the arrays. A schematic diagram of these stages is
shown in Figure 2.5.

Figure 2.5. Overview of Sample Preparation

Figure 2.5. Legend: Prior to labelling and hybridisation the RNA and DNA was extracted and
quantified. Clinical tumour samples were limited, therefore laser capture micro dissection and
amplification of the RNA could be carried out.

2.2.3.1. Cell Culture and Maintenance

All cell culture work was done in a class II lamina flow hood under asceptic conditions. Cells
were maintained in a Sanyo CO2 incubator at 37 °C with 5 % CO2. Long-term storage of cells
was in liquid Nitrogen.
The human breast cell lines used are summarised in Table 2.2. Breast tumour cell lines MCF7/wt, T47D, ZR75-1 and BT474, were originally obtained from the American Type Culture Collection and were maintained in phenol red-free RPMI 1640 media supplemented with glutamine. Breast tumour cell lines MDA16 and MDA468 were maintained in phenol red RPMI 1640 media supplemented with glutamine. The breast epithelial cell line HBL100 was maintained in Dulbecco Minimal Essential Medium (DMEM). All media were supplemented with 10% heat-inactivated fetal calf serum. Additionally HPGP cells derived from the MCF7/Dox cells were supplemented with 0.5 μM doxorubicin. MCF7/wt and the drug resistant MCF7/Dox (Batist et al. 1986) cell lines were originally obtained from the National Cancer Institute, MD, while the drug resistant MDA16 line was obtained from the progenitor cell line MDA468 via selection of MDR by cell sorting, using a magnetic bead containing the MRK-16 antibody.

Table 2.2. Human Breast Cell Lines

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Resistant</th>
<th>Media</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBL100</td>
<td>No</td>
<td>DMEM</td>
<td>Non tumour immortalised breast cell line</td>
</tr>
<tr>
<td>BT474</td>
<td>No</td>
<td>RPMI 1640</td>
<td>Isolated from a solid, invasive ductal carcinoma of the breast (Lasfargues et al., 1978)</td>
</tr>
<tr>
<td>T47D</td>
<td>No</td>
<td>RPMI 1640</td>
<td>Isolated from a patient with an infiltrating ductal carcinoma of the breast (Keydar et al., 1979)</td>
</tr>
<tr>
<td>ZR75-1</td>
<td>No</td>
<td>RPMI 1640</td>
<td>Isolated from a patient with ductal carcinoma of the breast (Engel et al., 1978)</td>
</tr>
<tr>
<td>MCF7/wt</td>
<td>No</td>
<td>RPMI 1640</td>
<td>Isolated in 1973 from a patient with metastatic adenocarcinoma of the breast (Soule et al., 1973)</td>
</tr>
<tr>
<td>HPGP</td>
<td>Yes</td>
<td>RPMI 1640</td>
<td>Derived from Dox sensitive MCF7/wt (Davies et al., 1996; Batist et al., 1986)</td>
</tr>
<tr>
<td>MDA468</td>
<td>No</td>
<td>RPMI 1640</td>
<td>Isolated in 1977 from a patient with metastatic adenocarcinoma of the breast (Cailleau et al., 1978)</td>
</tr>
<tr>
<td>MDA16</td>
<td>Yes</td>
<td>RPMI 1640</td>
<td>Derived from Dox sensitive MDA 468 cells using the MRK-16 antibody</td>
</tr>
</tbody>
</table>

Cells were allowed to reach confluency (70 – 80 %) whereupon they were subcultured: The monolayer cells were washed twice with PBS solution and were trypsinised with 1 x Trypsin/EDTA (section 2.1.4.). Cells were pelleted at 200 x g for 3 min and reseeded at approximately 2 x 10⁴ cells/ml. For long term storage in liquid nitrogen, cells were detached and pelleted at 200 x g for 3 min. The pellet was resuspended at 1 x 10⁶ cells/ml in freezing down medium (section 2.1.4.) and stored at -80°C for 48 hrs before being transferred to liquid nitrogen. All cells were only used for experimental work up to 35 passages.
2.2.3.2. Treatment of Cells with Drug

Drug dosing regimes of cells varied for individual experiments and are explained where necessary in each separate results chapter. Cells were grown as described in section 2.2.3.1. Prior to dosing the cells were cultured and left to reseed for 4 hrs to overnight. Once attached the cells were dosed with varying concentrations of drug in appropriate vehicle, for Doxorubicin DMSO was used ensuring that the final concentration of DMSO in contact with cells did not exceed 0.1%.

In Chapter 4 the cells were left to grow for 7 days in the presence of the compound. All cells were then trypsinised and total RNA was extracted and analysed on the microarray to monitor changes in gene expression of the cells in the presence of drug. Varying time points and concentrations of drug were carried out.

2.2.3.3. Mutagenising Cells and Selecting for Drug Resistant Clones

Chapter 4 analyses the expression patterns for the onset of the drug resistant phenotype. 1 x 10^6 cells were first mutagenised with 250 µg of ethylmethane sulfonate (EMS) for 18 hr at 37°C as described by Kataoka et al., (1994) (EMS from Sigma as 1.25 g/ml). The cells were then washed and allowed to recover for 4 days. 3 x 10^6 of mutagenised cells were selected for resistance by exposing the cells for 7 days to a drug at a concentration that gave a lethal dose to 90% (LD_{90}). Single clones were isolated, grown and recloned. The single clones were then grown and altered resistance to the drug was determined with the MTS cytotoxicity assay and RNA was extracted and analysed on the array. Alternatively cells were exposed to drug without prior mutagenisation and the drug selection procedure was repeated several times with increasing concentrations of drug. At each stage MTS assays were carried out (Section 2.2.10.) and RNA was extracted to analyse altered gene expression as drug resistance arose.

2.2.4. RNA and DNA Extraction from Cell Lines

Total RNA or genomic DNA was extracted from cells using either the Qiagen RNAeasy mini kit or Qiagen genomic DNA isolation kit for cells. All buffers were supplied with the kits (Section 2.1.5. and 2.1.6.) and the manufacturers protocols were followed.

2.2.4.1. Total RNA Extraction from Cell Culture

Cell lines were seeded into 3 x 150 cm² flasks, on approaching confluence cells were washed twice with PBS and enzymatically detached with 1x trypsin/EDTA. Cells were pelleted at 200
x g for 3 min and resuspended in 10 ml media, subsequent steps were carried out on ice. Cells were counted using a Casyl cell counter (Schärfe Systems), repelleted at 200 x g for 3 min and 2 ml of RLT buffer (Section 2.1.5.) was added per 3 x 10^7 cells to lyse cells. The cell suspensions were sonicated for 6 x 5 sec at 10 µ power with 10 sec intervals. From here the steps were carried out at room temperature. 2 ml 70 % EtOH was added to each cell suspension and the 4 ml solution was added to an extraction column supplied with the Qiagen kit. The columns were centrifuged at 3000 x g for 5 min, eluates were discarded from the collecting tube. The column was washed by adding 3.8 ml of RW1 (Section 2.1.5.) to the column and was centrifuged as before. Eluates were discarded and 2.5 ml RPE (Section 2.1.5.) was added to the columns, which were centrifuged as before, this RPE step was repeated twice. The columns were transferred to fresh collecting tubes and 150 µl RNase free water was added to the column, which was centrifuged to elute the RNA. This step was repeated and the eluates combined in the collecting tube. 1 ml of Sigma Tri reagent and 200 µl of chloroform was added. Tubes were vortexed and centrifuged. The top aqueous layer was recovered and transferred to a 1.5 ml eppendorf tube; to precipitate the RNA 0.5 volumes of isopropanol was added, mixed by inversion and incubated for 10 min at room temperature. The precipitated RNA was pelleted by centrifugation at 10,000 x g for 10 min. The pellet was washed twice with 70 % EtOH and resuspended in 200 µl RNase free water. The RNA was reprecipitated overnight with 1/10 volume NaAc (3 M pH 5.2) and 3 volumes of EtOH at - 80 °C. The precipitated RNA was pelleted by centrifugation at 10,000 x g for 30 min. The RNA pellet was washed twice with 70 % EtOH, dried and resuspended in 20 µl RNase free water. RNA was quantified and stored at - 80 °C.

2.2.4.2. DNA Extraction from Cell Culture

Cell lines were seeded into 3 x 150 cm² flasks, on approaching confluence cells were washed twice with PBS and were enzymatically detached with 1x trypsin/EDTA. Cells were pelleted at 200 x g for 3 min and resuspended in 10 ml PBS, from here steps were carried out on ice. Cells were counted using a Shaeffer cell counter, repelleted at 200 x g for 3 min, and resuspended in 2 ml PBS per 2 x 10^7 cells (number of cells for genomic 100 tip columns). 6 ml of ice cold distilled H₂O and 2 ml of Buffer C1 (Section 2.1.6.) was added; Buffer C1 lyases the cells, but stabilises and preserves the nuclei. The suspensions were incubated for 10 min on ice, then lysed cells were pelleted by centrifugation for 5 min at 4 °C, 1300 x g. A wash step was carried out to remove residual cell debris from the nuclear pellet: Pellets were resuspended in 1 ml buffer C1 and 3 ml ice cold distilled H₂O and repelleted via centrifugation for 15 min at 4 °C, 1300 x g. Pellets were completely resuspended in 5 ml
buffer G2 (Section 2.1.6.) and 95 µl of the Qiagen protease solution (buffer G2 lyses the nuclei and denatures proteins e.g. nucleases and histones, while proteases digests the denatured proteins into smaller fragments, thus stripping the genomic DNA of bound proteins). The nuclear suspensions were incubated at 50 °C for 60 min. The DNA lysate was vortexed for 10 sec and applied to a genomic 100 tip column, which was prepared by the addition of 4 ml buffer QBT (Section 2.1.6.) that was allowed to drain through via gravity flow. The DNA lysate was ran through the column by gravity flow; gentle pressure with a disposable syringe was applied ensuring that 10 – 20 drops /per min was not exceeded. The column was then washed by the addition of 7.5 ml buffer QC (Section 2.1.6.), this wash step was repeated. The genomic DNA was then eluted by the addition of 5 ml buffer QF (Section 2.1.6.). The eluted DNA was precipitated with 0.7 volumes of isopropanol. Percipitated DNA was centrifuged at 4 °C for 15 min at 5000 x g. The resulting pellet was washed twice with 2 ml 70 % EtOH via centrifuation at 4 °C for 10 min at 5000 x g. The pellet was left to air dry and resuspended in 50 µl TE pH 8.0. Extracted DNA was quantified and stored at 4 °C.

2.3.4.3. Quantification of RNA/DNA
Extracted RNA and DNA were quantified using a Lambda II mass spectrophotometer. A 1 in 500 dilution was made of the RNA or a 1 in 250 dilution of the DNA in nuclease free water. The RNA or DNA were measured at wavelengths of 260, 280 and 320 with nuclease free water as the blank. The 260, 280 and 320 wavelength measurements were used to calculate the ratio and concentration of the RNA or DNA:

\[
\text{Ratio calculated: } \frac{260 \lambda}{280 \lambda} = \text{ratio}
\]

The ratio gives an indication of the quality of the RNA or DNA and should be in the region of 1.8 – 2.0.

\[
\text{Concentration for ds DNA calculated: } 260 \lambda \times \text{dilution} \times 50
\]
\[
\text{Concentration for RNA calculated: } 260 \lambda \times \text{dilution} \times 40
\]

The concentration of the RNA or DNA is initially calculated as µg/ml and was divided by 1000 to acquire µg/µl. Concentration of RNA required for the labelling reaction and subsequent hybridisation is 50 µg/10 µl, while the concentration of DNA required for the labelling and hybridising is 2 µg/10 µl.
2.2.5. RNA amplification

Where only small amounts of RNA were available (0.05 ng – 1 μg), i.e. limited in vivo samples or those from Laser Capture Microdissection (LCM), it was necessary to amplify the RNA prior to microarray analysis. Amplification was achieved using the Smart PCR cDNA synthesis and Advantage 2 PCR amplification kits (Clontech) as described by Chenchik et al. (1998), whereby RNA is first translated into cDNA, which undergoes amplification via long distance PCR. All enzymes and primers were supplied in the kit.

2.2.5.1. Smart cDNA First Strand Synthesis

The first strand reaction is primed with the modified SMART cDNA synthesis oligo dT primer (3' SMART CDS Primer II A). When the reverse transcriptase reaches the 5' end of the RNA the enzymes terminal transfer activity adds a few deoxycytidine bases to the 3' end of the cDNA. A second primer, the SMART Oligonucleotide, which has an oligo(G) sequence at its 3’ end base pairs with the deoxycytidine bases. The reverse transcriptase then switches templates and continues back to the end of the oligonucleotide. The SMART anchor sequence then serves as a priming site for cDNA amplification with the Advantage 2 PCR kit.

10 μM of cDNA synthesis SMART CDS Primer II A and 10 μM of Smart II Oligonucleotide was mixed with 500 ng of RNA in 1.5 μl of H₂O and incubated at 70 °C for 2 min to allow annealing. The 1st strand cDNA synthesis reaction in Table 2.3. was then set up.

<table>
<thead>
<tr>
<th>Reagent (concentration)</th>
<th>Volume (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annealing reaction</td>
<td>2.5</td>
<td>500 ng RNA</td>
</tr>
<tr>
<td>5x first strand buffer</td>
<td>1</td>
<td>1 x</td>
</tr>
<tr>
<td>DTT (20 mM)</td>
<td>0.5</td>
<td>2 mM</td>
</tr>
<tr>
<td>50X dNTP mix (10mM)</td>
<td>0.5</td>
<td>1 mM</td>
</tr>
<tr>
<td>MMLV reverse transcriptase (200 U/μl)</td>
<td>0.5</td>
<td>100 U</td>
</tr>
</tbody>
</table>

The 1st strand cDNA synthesis reaction was incubated at 42 °C for 1 hr. 20 μl of TE buffer was added and tubes were incubated at 72 °C for 7 min. Samples could be stored at – 20 °C for up to 3 months prior to amplification.
2.2.5.2. SMART cDNA Amplification

cDNA was amplified via the 5’ PCR primer II A primer, reactions were prepared as shown in Table 2.4.

The reactions were placed into a preheated MJ Research DNA engine (95 °C) and PCR amplification reactions were carried out using the following parameters:

- **Step 1:** 95 °C for 1 min  
  Initial denaturation
- **Step 2:** 95 °C for 15 sec  
  Denaturation
- **Step 3:** 65 °C for 30 sec  
  Primer annealing
- **Step 4:** 68 °C for 6 min  
  Extension
- **Step 6:** steps 2 – 4 repeated 19 times

Table 2.4. SMART cDNA Amplification Reaction

<table>
<thead>
<tr>
<th>Reagent (concentration)</th>
<th>Volume (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA synthesis reaction</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>10X Advantage 2 PCR buffer</td>
<td>10</td>
<td>1 X</td>
</tr>
<tr>
<td>50X dNTP mix (10mM)</td>
<td>2</td>
<td>0.2 mM</td>
</tr>
<tr>
<td>5’ PCR primer II A (10 μM)</td>
<td>2</td>
<td>0.2 μM</td>
</tr>
<tr>
<td>50X Advantage 2 polymerase mix</td>
<td>2</td>
<td>1 X</td>
</tr>
<tr>
<td>H2O</td>
<td>74</td>
<td>-</td>
</tr>
</tbody>
</table>

For each cDNA synthesis reaction 3 separate cDNA amplification reactions were carried out and the resulting products were pooled. The amplified reactions were precipitated with an equal volume of isopropanol for 30 min at – 20 °C and washed with 70 % EtOH. The appropriate reactions were resuspended, pooled together in 20 μl H2O and stored at – 70 °C prior to labelling and hybridising to the arrays.

2.2.6. RNA/DNA/Amplified RNA Labelling and Hybridising

2.2.6.1 Total RNA Labelling

The main principle of the RNA labelling and hybridisation reactions are described by Duggan et al (1999), whereby two different RNA populations, which are to be compared for gene expression, are firstly reverse transcribed into cDNA via an oligo dT primer. During the reaction cyanine-3 labelled dNTPs are incorporated into one population and cyanaine-5 labelled dNTPs into the other. The two labelled cDNA populations are pooled and co-hybridised to a single microarray.
4 μg of oligo dT<sub>25</sub> acting as a primer, was added to 50 μg of total RNA in 10 μl nuclease free H<sub>2</sub>O. After denaturation at 70 °C for 8 min, annealing occurred as the temperature was allowed to fall to 42 °C over 30 min. At 42 °C the components of the labelling reaction were added (Table 2.5.). Labelling reactions were incubated for 1 hr at 42 °C at which point 0.5 μl of Superscript II Reverse transcriptase was added, and the reaction was allowed to continue at 42 °C for a further 1 hr.

### Table 2.5. Components of the RNA Labelling Reaction

<table>
<thead>
<tr>
<th>Reagent (concentration)</th>
<th>Volume (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x first strand buffer*</td>
<td>4</td>
<td>1 x</td>
</tr>
<tr>
<td>0.1 M DTT*</td>
<td>2</td>
<td>0.01 M</td>
</tr>
<tr>
<td>dNTP mix (dATP, dCTP, dGTP 20 mM, dTTP 8mM)*</td>
<td>0.5</td>
<td>0.5 mM and 0.2 mM</td>
</tr>
<tr>
<td>Cy3 or Cy5 fluor-dUTP (1 mM)</td>
<td>2</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>RNAsin</td>
<td>0.5</td>
<td>20 U/reaction</td>
</tr>
<tr>
<td>Superscript II Reverse transcriptase</td>
<td>0.5</td>
<td>100 U/reaction</td>
</tr>
</tbody>
</table>

* Made as a Master mix, added as a volume of 6.5 μl.

20.5μl of nuclease free water was added to bring the volume to 41μl. RNA was hydrolysed, so removed from the synthesised cDNA, by the addition of 1 μl of 0.5 M EDTA, 1 μl of 10% SDS, 3 μl of 3 M NaOH and was incubated at 70 °C for 10 min. The reaction was neutralised by the addition of 3 μl of 2 M HCl and 10 μl of 1 M Tris/HCl pH 7.5. Labelled cDNA was purified as in section 2.2.6.3.

#### 2.2.6.2. Genomic DNA and Amplified RNA Labelling

Prior to labelling 2 μg of the genomic DNA was digested with DpnII (New England Biolabs) for 1 hr – overnight at 37 °C (see Table 2.6.).

### Table 2.6. The DNA Digestion Reaction

<table>
<thead>
<tr>
<th>DNA 2 μg (μl)</th>
<th>DpnII Buffer 10 x (μl)</th>
<th>Enzyme DpnII (μl)</th>
<th>Water (μl) to 50 μl</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
<td>1 (50 U)</td>
<td>34</td>
</tr>
</tbody>
</table>

Protein was removed from the digested DNA by extracting twice against an equal volume of 1:1 tris buffered phenol:chloroform, followed by one extraction against chloroform. The DNA
was then precipitated overnight at -20 °C with the addition of 1/10 volume of NaAc (3 M pH 5.2) and 2 volumes of EtOH. Percipitated DNA was pelleted by centrifugation at 10,000 x g for 20 min. The DNA pellet was washed twice with 70 % EtOH, air dried and resuspended in 32 μl RNAse/DNAse free water.

Labelling of the amplified RNA or digested DNA was achieved using random hexamers, essentially as described by Feinberg and Vogelstein (1983). The labelling reaction components are shown in Table 2.7.

Table 2.7. Components of the Genomic DNA Labelling Reaction

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplified RNA or Digested DNA (heat denatured)</td>
<td>32</td>
</tr>
<tr>
<td>OLB (Section 2.1.8.)</td>
<td>10</td>
</tr>
<tr>
<td>Bovine Serum Albumin (BSA) 10 μg/μl</td>
<td>2</td>
</tr>
<tr>
<td>Cy3 dCTP OR Cy5 dCTP</td>
<td>5</td>
</tr>
<tr>
<td>Klenow Fragment (high activity 50,000 U/ml)</td>
<td>1</td>
</tr>
</tbody>
</table>

The labelling reaction was left for 5 hr – overnight at room temperature and terminated with the addition of 5 μl EDTA. The labelled cDNA was purified as in Section 2.2.6.3.

2.2.6.3. Purification of the Labelled cDNA

1 μl of 4 μg/ml yeast tRNA was added to each labelled sample to act as a carrier. Labelled cDNA was purified using Centrisep gel filtration columns (Princeton Separations), which removed unincorporated nucleotides by absorption onto the gel surface. Columns were prepared with the addition of 0.8 ml nuclease free water, briefly vortexed and allowed to settle for at least 30 min prior to use. The column caps and bottoms were removed and the columns were left to drain by gravity flow in a 1.5 ml eppendorf. Columns were centrifuged for 2 min at 800 x g to remove the remaining fluid. Columns were then transferred to fresh 1.5 ml eppendorf tubes and ensuring the gel bed was not disturbed, the labelled cDNA was transferred to the centre of the column. The columns were centrifuged in the same orientation as before for 6 min at 800 x g.

To avoid non-specific binding during hybridisation 1μl of COT 1 Human DNA and 1 μl of polyA was added. All labelled samples were then dried to a pellet using a DNA speedvac (Savant).
2.2.6.4. Hybridisation of Labelled cDNA to Glass Slide

Corresponding Cy3 and Cy5 dried, labelled samples were redissolved together in 10.5 µl of hybridisation buffer (Section 2.1.7.) and 4.5 µl 20 x SSPE (Section 2.1.7.). The probes were denatured at 100 °C for 2 min, and incubated at 42 °C for 30 min to allow annealing of the polyT tails. The labelled cDNA in the hybridisation solution were then placed onto the array and a coverslip was laid over the top, ensuring no air bubbles were present. The slides were transferred to watertight hybridisation chambers from GeneMachines and left to hybridise overnight – 48 hr in a 42 °C water bath with a lid.

2.2.6.5. Washing and Scanning of Hybridised Slides

Hybridised slides were placed in a horizontal slide rack and submerged in the first wash solution: 1 x SSC, 0.03 % SDS, slides were left in the solution until the cover slips came loose, followed by a further 10 min with occasional plunging before transfer to the second wash solution: 0.2 x SSC. Slides were washed for 5 min, then transferred to the third wash solution: 0.05 x SSC and washed for a further 5 min. Slides were centrifuged at 500 x g for 4 min to dry.

Laser excitation of the incorporated Cy dyes of the hybridised cDNA populations yields an emission, which is measured using a PMT (Photomultiplier tube) detection system in the Axon 4000A scanner which was controlled by the GenePix software (version 2 or 3) from Axon Instruments. The incorporated Cy dyes were excited at wavelengths of 532 nm (Cy3) or 635 nm (Cy5) and the resulting fluorescence was measured. cDNA spots or feature sizes were determined using the inbuilt automated parameters in GenePix, then adjusted by hand where required.

2.2.7. Data Analysis

After washing and scanning of hybridised microarrays the GenePix software was used to acquire the data for all features or cDNA spots. The data was then normalised prior to further analysis. Cluster programs were used to analyse the data, which was visualised with treeview. A schematic diagram of these stages is shown in Figure 2.6.
The GenePix software (Axon instruments) was used to scan the hybridised microarrays and acquire the data. A single feature or cDNA spot on the array once scanned consists of a number of pixels. The fluorescence of each pixel within the features were determined and the median fluorescence of these pixel measurements was taken as the measure of fluorescence for the whole feature. The default parameters of GenePix were used to measure background fluorescence.

The raw feature data for each channel (Cy3 or Cy5) was globally normalized by reference to the median fluorescence of the whole feature set for that channel. The mean ± SD of the log\textsubscript{2} ratio data was calculated, and used to calculate a significance weighting value using the following equation:

\[ f(x) = \frac{1}{1 + \exp(-a*(x-b))} \]

where \( x = \log\textsubscript{2} \) ratio; \( a = 5/(\text{conf99}-\text{conf95}) \); and \( b = (\text{conf95}+\text{conf99})/2 \).

The signs of the function and confidence values were reversed for the two tails of the distribution. Values falling between 0.05 and -0.05 were rounded to 0. Significance weighting values greater than 0.92, or less than -0.92, were associated with differential expression ratios falling in the greater than 99 % confidence levels in both tails of the distribution. Thus, the middle significance weighting value of 0.5 was associated with the values falling on the 97.5 % confidence value and the whole data set was hence normalized to a scale of -1 to 1. Clustering analysis was performed using the significance weighting.

For clustering analysis data from replicate experiments was entered separately. Data for which there was a fluorescence value in one channel and not the other were treated as a special case.
These values were assigned a significance value of +1 or -1 depending on in which channel the fluorescence was recorded and only if the intensity of that fluorescence was greater than a multiple of 1.5 to 3 (dependent on the array quality) times the average fluorescent intensity for the channel as a whole. Data processing was carried out using the Convert Data program version 3.2.3a. The resulting values were fed into the cluster program of Eisen et al (1998) and clustered using the hierarchically, un-centred, with complete linkage option. The clustering data were displayed using the treeview program of Eisen et al. (http://www-genome.stanford.edu/). Prior to viewing the weighted significance data was converted back into ratio data. The density of colour in the tree view diagrams indicates the ratio as indicated.

2.2.8. Post Microarray Experiments

Post microarray experiments were carried out to confirm data or to provide additional data to the microarray work. Southern blots, nylon arrays and Real Time PCR were carried out to confirm array data. A schematic diagram of these stages is shown in Figure 2.7.

Figure 2.7. Overview of Post Microarray Work

![Diagram of Post Microarray Work]

Figure 2.7. Legend: Experiments such as Southern blots or Real Time PCR were carried out to confirm microarray data and other experiments were carried out to provide additional data.

2.2.9. Nylon Membrane Microarrays

2.2.9.1. Gridding Clones onto Nylon Membranes

Some microarray data was confirmed by gridding PCR products in duplicate onto nylon membranes that were probed with $^{33}$P labelled RNA samples. 0.6 pm of PCR product were prepared in a 96-well plate. A piece of blotting paper was immersed in denaturing solution (Section 2.1.9.) and carefully laid onto a filter block (Genetix). Using foreps a MagnaGraph nylon transfer membrane, 0.45 micron (Osmonics Inc.) was briefly soaked in denaturing
solution and overlaid on top of the blotting membrane. A 10 ml disposable pipette was rolled over the surface of the membrane to remove air bubbles. The filter block was placed into the bed of a Qpix Robot (Genetix). A 96 pin head with 0.25 mm pins was used to grid the PCR products in duplicate. Between each transfer the pins were washed in 1 % NaOCl, distilled H₂O and 80 % EtOH. Using forceps a fresh piece of blotting paper was immersed in the denaturing solution and placed onto a clean filter block. The gridded membrane was placed nucleic acids side down onto the soaked blotting paper and left to denature for 10 min. Using forceps a fresh piece of blotting paper was immersed in neutralising solution (section 2.1.6.) and placed onto a filter block. The membrane was placed nucleic acids face down onto the blotting paper and left to neutralise for 5 min. The membrane was transferred to a dry piece of blotting paper and left to air dry. The DNA was covalently linked to the membrane by UV irradiation, DNA side up, on auto mode at 1200 μJ/cm² (Stratagene Stradlinker). Membranes were stored at 4 °C for up to two weeks prior to use.

2.2.9.2. Labelling RNA for Nylon Membranes

Total RNA extracted from cell lines was labelled with α³²P-dCTP and hybridised to the membranes. 15 ml of hybridisation solution (Section 2.1.10.) was added to the membrane, and allowed to pre hybridise for 3 – 7 hr in a rotating Techne Hybriser HB-1D oven at 42 °C. The hybridisation tube contained sufficient solution to provide cover for the rotating membrane, but not an excess (approximately 10 ml).

The labelling reaction was primed with 8 μg of oligo dT₂₅ (8 g/μl), which was added to 25 μg of total RNA in 15 μl nuclease free H₂O. After denaturation at 70 °C for 8 min, annealing occurred as the temperature fell to 42°C over 30 min. At 42 °C the components of the labelling reaction were added (Table 2.8.).

### Table 2.8. Components of the RNA Labelling Reaction for Nylon Membranes

<table>
<thead>
<tr>
<th>Reagent (concentration)</th>
<th>Volume (μl)</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x first strand buffer*</td>
<td>6</td>
<td>1 x</td>
</tr>
<tr>
<td>0.1 M DTT*</td>
<td>2</td>
<td>0.006 M</td>
</tr>
<tr>
<td>dNTP mix (dATP, dTTP, dGTP 20 mM, dCTP 8 mM)*</td>
<td>2</td>
<td>0.6 mM and 0.2 mM</td>
</tr>
<tr>
<td>³²P-dCTP</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>RNAsin</td>
<td>1</td>
<td>40 U /reaction</td>
</tr>
<tr>
<td>Superscript II Reverse transcriptase</td>
<td>1</td>
<td>200 U /reaction</td>
</tr>
</tbody>
</table>

* Made as a Master mix, added as a volume of 10 μl.
Labelling reactions were incubated for 1 hr at 42 °C at which point 1 µl of Superscript II Reverse transcriptase was added to the reaction, which was allowed to continue at 42 °C for a further 1 hr.

RNA was hydrolysed, so removed from the synthesised cDNA, by the addition of 1 µl of 0.5 M EDTA, 1 µl of 10% SDS, 3 µl of 3 M NaOH and was incubated at 70 °C for 10 min. The reaction was neutralised by the addition of 3 µl of 2 M HCl and 10 µl of 1 M Tris/HCl pH 7.5. 98 µl of nuclease free H₂O was added to bring the volume up to 150 µl. Labelled cDNA was purified using a sephadex G50 column and 1 µl of purified probe was counted with a Liquid Scintillation Counter 1410 (Wallac). A maximum of 10 x 10⁷ dpm of the probe was used for each membrane. 2 µl of dA₈₀ (1 µg/µl) was added to the probe, which was then heat denatured at 100 °C for 5 min, rapidly cooled on ice, then added to the hybridisation solution in the hybridisation chamber and left to hybridise for 16 – 72 hr in the rotating oven.

The membrane was washed for 10 – 15 min at 42 °C with 0.1 X SSC/ 0.1 % SDS. The membrane was then wrapped in cling film and exposed to a phosphor screen (Molecular Dynamics) for 2 hr – 24 hr and scanned using a PhosphorImager (Molecular Dynamics). Subsequent analysis was carried out using Image Quant 3.3 software (Molecular Dynamics).

2.2.10. Cytotoxicity Assay in Cell Lines Exposed to a Compound

The Cell titre 96° AQueous one solution cell proliferation assay from Promega was used to monitor the sensitivity of various cell lines to various compounds. This is a colorimetric assay used for determining the number of viable cells. The AQueous one solution contains the compound [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt (MTS) and the electron coupling agent phenazine ethosulfate (PES). The AQueous one solution is incubated with the cells for 1 – 4 hr, during which time the MTS is bioreduced by the viable cells into a coloured formazan product, which can be measured at a wavelength of 490 nm. The quantity of formazan measured at 490 nm is directly proportional to the number of living cells in culture.

Cell lines were seeded into 1 x 150 cm² flasks, on approaching confluencey cells were washed twice with PBS and were enzymatically detached with 1x trypsin/ EDTA. Cells were pelleted at 200 x g for 3 min and resuspended in 10 ml media. Cells were counted using a Shaeffer cell
counter. Cells were reseeded into 96-well flat bottomed plates, 100,000 in 100 μl per well, media and dosed with varying concentrations of drug in less than 0.1 % DMSO. ‘Blanks’ to measure media background readings were also set up, whereby 100 μl of each media alone was added to three wells. The cells were allowed an attachment period of 4 hours after seeding prior to exposure to the compound. Cells were incubated in a Sanyo CO₂ incubator at 37°C with 5% CO₂ for 72 hr.

20 μl of the cell proliferation assay solution was added to each well. The cells were incubated for a further 1 hr. Colour changes for viable cells were clearly visible and were measured with a plate reader at 490 nm. Percentage of viable cells, from the untreated dose, was calculated and plotted against drug concentration.

2.2.11. Southern Blot Analysis

Southern blotting detects the presence of a specific DNA sequence. Digested genomic DNA is separated via gel electrophoresis and the resolved DNA is transferred to a nylon membrane and a labelled probe is hybridised to the membrane allowing detection of the DNA.

2.2.11.1 Southern Blot Membrane Preparation

10 μg of DNA was digested with 10 U of EcoRI, in a total reaction volume of 20 μl. The DNA was digested at 37 °C for 2 – 3 hr. After digestion 6 μl of 5 x TBE and 4 μl of loading buffer were added. The samples were electrophoresed in a 1 % agarose gel overnight at 22 V. The gel was stained with EtBr (1 μg/ml) for 10 min and the image was scanned and saved using a Flour-S™ MultImager (Biorad). The gel was soaked in 0.2 M HCl for 10 min to further fragment the DNA and assist in membrane transfer. The gel was briefly washed in distilled H₂O then denatured in 1.5 M NaCl/ 0.5 M NaOH for 30 min with a change of solution after 10 min. The gel was again washed in distilled H₂O and neutralised in 1M Tris/HCl pH 7.4/ 1.5 M NaCl for 60 min with a change of solution after 30 min. The gel was then set up for an overnight capillary transfer using 10 x SSC to a MagnaGraph nylon transfer membrane, 0.45 micron (Osmonics Inc.), as illustrated in Figure 2.8. After transfer the DNA was UV crosslinked to the membrane using a Stratagene Stradlinker on auto mode (1200 μJ/cm²). Membranes were probed instantly or wrapped in cling film and stored at 4 °C for up to two weeks prior to use.
Figure 2.8. Legend: The apparatus used in the Southern blot transfer. Capillary transfer is achieved at room temperature overnight with 10 x SSC. Filter paper and paper towels aid in the transfer and a weight is placed on the top to compress the unit and facilitate transfer. Transferred membranes can be stored at 4 °C for up to two weeks prior to use.

2.2.11.2. Probe Labelling and Hybridisation

15 ml of hybridisation solution (Section 2.1.10.) was added to the membrane, and allowed to prehybridise for 3 – 7 hr in a rotating Techne Hybriser HB-1D oven at 42 °C. The hybridisation tube contained sufficient solution to provide cover for the rotating membrane, but not an excess. 100 ng of DNA probe in 32 µl H₂O was heat denatured at 95 °C for 5 min and a labelling reaction as shown in Table 2.9. was set up.

Table 2.9. Components of the Probe Labelling Reaction

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 ng Probe DNA (heat denatured)</td>
<td>32</td>
</tr>
<tr>
<td>OLB (Section 2.1.8.)</td>
<td>10</td>
</tr>
<tr>
<td>Bovine Serum Albumin (BSA) 10 µg/µl</td>
<td>2</td>
</tr>
<tr>
<td>³²P-dCTP</td>
<td>5</td>
</tr>
<tr>
<td>Klenow Fragment (high activity 50,000 U/ml)</td>
<td>1</td>
</tr>
</tbody>
</table>
The reaction was left for 3 – 4 hr at 37 °C, then purified with a Chromospin 10 column (Clontech). The purified labelled probe was counted with a Liquid Scintillation Counter 1410 (Wallac). A maximum of 10 x 10⁷ dpm of the probe was heat denatured at 95 °C for 5 min, rapidly cooled on ice, then added to the hybridisation solution in the hybridisation chamber and left to hybridise for 48 – 72 hr in the rotating oven.

After hybridisation the membrane was washed firstly with 2 x SSC/ 0.1 % SDS for 30 min at 42 °C, then with 0.1 x SSC/ 0.1 % SDS for 1 hr at 55 °C. The membrane was then exposed to a phosphor screen (Molecular Dynamics) for 2 hr – 24 hr and scanned using a PhosphorImager (Molecular Dynamics). Subsequent analysis was carried out using Image Quant 3.3 software (Molecular Dynamics).

### 2.2.12. Protein Extractions from Cells

Protein was extracted from cell lines for Western blot analysis. Cells were seeded into 90 mm petri dishes. When cells were 70 – 80 % confluent the media was removed and the cells were washed twice with PBS. The PBS was removed and 200 µl of cell lysis buffer (Section 2.1.11.) was added to each dish. The dishes were left on ice for 10 min, after which time cells were scrapped and transferred to eppendorf tubes. The cells were incubated for 30 min on ice to allow further lysis. Tubes were then centrifuged at 10,000 rpm for 3 min to pellet the cell debris. The supernatant containing the proteins was collected and the protein concentrations were determined by the Bradford assay (Bradford, 1976), proteins were then stored at −20 °C until use.

### 2.2.13. Western Blot Analysis

Some microarray gene expression data was supported with protein expression data via Western blotting. Proteins were firstly separated from total protein extracts via sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS/PAGE). The resolved proteins are then transferred to a nitrocellulose membrane, which is hybridised with an antibody to detect for the presence of a specific protein.

30 µg of protein extracted from cells was prepared in 15 µl H₂O and 4.5 µl sample buffer (Section 2.1.12.). Samples were incubated for 5 min at 100 °C prior to loading onto a precasted gel comprising of a lower separating gel (Section 2.1.12.) and an upper stacking gel (Section 2.1.12.) (Mighty small gel apparatus, Hoefer Instruments). The gel was
electrophoresed in 1 x running buffer (section 2.1.10.) at 43 V and 47 – 208 kDa high molecular weight standard markers (BioRad) were used to measure electrophoretic progression. The protein samples were transferred onto a hybond nitrocellulose membrane for 3 hr at 60 V, 4 °C. To minimise non specific hybridisation, membranes were then blocked over night at 4 °C in 10 % non-fat dried milk/0.1 % T-BST (section 2.1.10.).

After blocking all steps were carried out at room temperature. Membranes were washed briefly on a rocker in 0.1 % T-BST and then exposed to the appropriate monoclonal mouse derived antibody in 0.5 % non-fat dried milk/0.1 % T-BST for 1 hr (diluted to manufacturers recommendations). Membranes were washed in 0.1 % T-BST 3 x for 5 min, exposed to secondary antibody for 1 hr and then washed as before. The specific proteins were detected with an Enhanced Chemiluminesence (ECL) kit (Amersham). The ECL solutions were mixed 1:1 and exposed to the membranes for 1 min. Chemiluminesence is generated from the horseradish peroxidase catalysed oxidation reaction of a lumiol substrate. The reaction is amplified for 10 sec – 10 min for detection on an X-ray film (Hyperfilm, Amersham), after which films were developed on an X-ograph compact x-2 developer. Developing and fixing solutions were supplied by Kodak.

2.2.14. Real Time PCR Analysis

Real Time PCR was used to confirm microarray expression data. Primers were first designed using PrimerExpress (Applied Biosystems), ensuring that the amplicon spanned an exon-exon boundary to minimise DNA contamination. A reverse transcription (RT) reaction to convert mRNA to cDNA, followed by a SYBR green Real Time PCR was performed. The TATA box-binding protein (TBP) mRNA was used to normalise genes, as due to its unchanging expression is a good endogenous control (Bieche et al., 1999), and each breast cancer cell line was calibrated to the non tumour HBL100 cell line which was thus assigned an expression value of 1.

2.2.14.1. Reverse Transcription Reaction

A master mix of the RT buffer for 100 reactions was made up (Table 2.10), this could be stored at – 20 °C. 20 μl RT reactions were performed using 14.5 μl of the RT master mix buffer, with 0.5 μl RNAsin, 1 μl Reverse transcriptase, 2 μl of RNA (at 100 ng/μl) and 2 μl of water.
Table 2.10. Components of the Real Time Reverse Transcription Master Mix

<table>
<thead>
<tr>
<th>Reagent (concentration)</th>
<th>Volume (µl)</th>
<th>Concentration in RT reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 x TaqMan RT PCR Buffer</td>
<td>200</td>
<td>1 x</td>
</tr>
<tr>
<td>MgCl₂ (50 mM)</td>
<td>100</td>
<td>2.5 mM</td>
</tr>
<tr>
<td>DNTP mix (dATP, dCTP, dGTP)</td>
<td>80</td>
<td>100 µM</td>
</tr>
<tr>
<td>dTTP 20 µl of 100 mM each)</td>
<td>19.8</td>
<td>2.5 µM</td>
</tr>
<tr>
<td>Random Hexamers</td>
<td>20</td>
<td>0.2 mM</td>
</tr>
<tr>
<td>DTT (20 mM)</td>
<td>1030.2</td>
<td>to 1450 µl</td>
</tr>
</tbody>
</table>

RT reactions were placed into a MJ Research DNA engine and reactions were carried out using the following parameters:

Step 1: 23 °C for 10 min  
Step 2: 42 °C for 45 min  
Step 3: 99 °C for 10 min

2.2.14.2. SYBR Green Real Time PCR Reaction

SYBR green Real Time PCR reactions were performed (Section 1.6.4.). Duplicate reactions were set up for each gene (Table 2.11.), placed into a MJ Research DNA engine and carried out using the following parameters:

Step 1: 50 °C for 2 min  
Step 2: 95 °C for 10 min  
Step 3: 95 °C for 15 sec  
Step 4: 60 °C for 1 min  
Step 5: steps 3 – 4 repeated 40 cycles

Table 2.11. Components of the Real Time PCR Reaction

<table>
<thead>
<tr>
<th>Reagent (concentration)</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYBR Master Mix (ABI)</td>
<td>12.5</td>
</tr>
<tr>
<td>Forward Primer (5 pmols/µl)</td>
<td>0.25</td>
</tr>
<tr>
<td>Reverse Primer (5 pmols/µl)</td>
<td>0.25</td>
</tr>
<tr>
<td>CDNA from 2.2.14.1.</td>
<td>1.0</td>
</tr>
<tr>
<td>Water</td>
<td>11</td>
</tr>
</tbody>
</table>

The threshold cycle (Cₜ) for each gene was calculated. Cₜ indicates the fractional cycle number at which the fluorescent signal is first recorded as statistically significant above background.
(Gibson et al., 1996). The mean $C_t$ for each gene was calculated and the TBP was used to normalise genes, and each breast cancer cell line was calibrated to the non tumour HBL100 cell line to gain relative expression data:

\[
\Delta C_t \text{ sample} = \text{target gene mean } C_t - \text{Endogenous gene mean } C_t \text{ (TBP)} \\
\Delta \Delta C_t = \Delta C_t \text{ sample} - \Delta C_t \text{ of calibrator (HBL100)} \\
2^{-\Delta \Delta C_t} = \text{normalised gene relative to HBL100}
\]

2.2.15. Summary of Microarray Methods

The microarray processes carried out in this thesis are summarised in Figure 2.9. As seen the stages are clearly defined into clone preparation, slide preparation and subsequent hybridisation followed by data acquisition, analysis and post microarray experiments.
Figure 2.9. A Schematic Summary of the Microarray Process

Figure 2.9. Legend: A schematic diagram of the microarray process; clones and slides are prepared prior to printing of the arrays. RNA or DNA to be analysed for altered gene expression are labelled and hybridised to the arrays. The resulting data is analysed and further work is carried out.
Chapter 3

Gene Expression Change and The Drug Resistant Phenotype
3.1. Introduction
Chemotherapy is commonly used in the treatment of breast cancer. It has been shown that in chemotherapy naïve patients with advanced breast cancer, doxorubicin (Dox or adriamycin) containing regimes are superior, and the first choice of chemotherapy in many patients with invasive metastatic cancers (Andersson et al., 1999; Nemoto et al., 1982). Consequently in this project Dox with a focus on breast carcinoma was used in several experiments.

The development of resistance to chemotherapeutics is a major obstacle in the successful treatment of patients diagnosed with cancer. Resistance to Dox is known to develop during breast cancer treatment (Kroger et al., 1999) and so it is a good candidate to use to elucidate some of the gene alterations that may occur during the development of drug resistance. Dox is a multi-ring fungal anthracycline whose structure is shown in Figure 3.1.

Figure 3.1. The Molecular Structure of Doxorubicin

![Figure 3.1 The Molecular Structure of Doxorubicin](image)

Figure 3.1.Legend: The structural formula of the anthracycline doxorubicin.

3.1.1. Doxorubicin Mechanisms of Action and Toxicity
Dox exerts its main chemotherapeutic effect on tumour cells through several mechanisms (Gewirtz, 1999). The main effects of Dox are at the DNA level, where it can intercalate with DNA and form cross-links. In human tumour biopsies after a dose of 25 mg/m² administered preoperatively over 80 % of the Dox is present in the DNA-bound form (Cummings et al., 1986) and this intercalation is thought to account for the observed dose mediated inhibition of DNA replication and RNA transcription (Cummings et al., 1991). Dox also interferes with DNA unwinding (Fornari et al., 1994), an observation which is thought to be caused by DNA
intercalation across two strands preventing unwinding or by the inhibition of helicase activity (Bachur et al., 1992).

Dox causes specific single or double strand breaks in DNA (Goldenberg et al., 1986). The single and double strand breaks are primarily caused by topoisomerase II. Topoisomerase II is involved in ATP dependant DNA strand passing activity important for replication fork movement and segregation of daughter cell chromosomes. It binds to both strands of the DNA and breaks open each by forming a transient covalent bond with the 5'-phosphoryl end of the broken strand and a tyrosine residue of each subunit of the protein, thus allowing another DNA molecule to pass through (Cummings, 1991). Dox stabilises this cleavage complex and at higher concentrations inhibits the strand passing activity. Although once the drug is removed from the DNA, the inhibition of topoisomerase II is reversed (Cummings, 1991).

Dox has also been shown to generate reactive oxygen species (Doroshow, 1986), which may induce free radical injury to DNA (Gutteridge et al., 1985). Whether free radicals are generated in vivo at clinically relevant concentrations and whether such free radicals could play a role in Dox toxicity to the tumour is not known (Gewirtz, 1999). However the main limitation in the treatment of breast cancer with Dox is cardiotoxicity, as it has been shown that there is a 2.2% increase risk of congestive heart failure in patients receiving Dox (Van Hoff et al., 1979). The exact mechanism of Dox cardiotoxicity is not known, although it has been suggested that disruption to calcium homeostasis and the generation of free radicals is involved (Olsen and Mushlin, 1990). To a lesser extent Dox has also been shown to induce apoptosis (Skaldanowski and Konopa, 1993), cause growth arrest (Gerwitz, 1999) and influence the properties of cellular membranes (Siegfried et al., 1983).

3.1.2. Mechanisms of Resistance to Chemotherapeutic Agents

Resistance occurs in cancer by various mechanisms and it can be either acquired or intrinsic. In acquired resistance the phenotype is an induced or selected population of tumour cells exposed to drug, whilst intrinsic resistant cells show resistance without prior contact with the drug. Two cell lines exhibiting an MDR phenotype and resistance to Dox were used in this chapter, MDA16 and HPGP. The intrinsically resistant MDA16 cell line was derived from sensitive MDA468 breast cancer cells by selecting with a monoclonal antibody, MRK-16. MRK-16 was first generated by Hamada et al. (1986) and has an affinity for the first and fourth extracellular loops of P-gp (Vasudevedevan et al., 1998). Dox is a substrate of the P-glycoprotein transporter encoded by the ABCB1 gene, which is known to be involved in Dox
resistance and the MDR phenotype (Seelig, 1998). Therefore this resistant cell line was selected for without the use of drug selection, and so has an intrinsic resistance phenotype.

The HPGP cell line was derived from the Dox resistant MCF7/Dox cell line (Davies et al., 1996). The HPGP-progenitor MCF7/Dox cells were initially produced from the sensitive MCF7/WT cells by exposure of these cells to an increasing concentration of Dox (Batist et al., 1986), therefore these cells are considered to have an acquired resistance to Dox. The MCF7/Dox cells also express GSTs, overexpress P-gp, have an amplified \( ABCB1 \) gene and consequently the cells exhibit a 192 fold increase in resistance to Dox (Batist et al., 1986). They also show the MDR phenotype, as they are cross-resistance to a variety of drugs, such as actinomycin D, vinblastine and colchicine. The resistant MCF7/Dox cells were sorted by fluorescence activated cell sorting (FACS) into two separate populations on the basis of rhodamine 123 (R123) accumulation (Davies et al., 1996). R123 is a substrate of P-gp hence cells with a high concentration of R123 were designated low P-gp cells, whilst cells with a high rate of R123 efflux are the high P-gp (HPGP) cells.

3.1.3 Chapter Objectives

In this chapter, genetic changes associated with the intrinsic and acquired Dox resistant phenotype are analysed with the aim to increase our understanding of drug resistance. cDNA microarrays were used to study differential gene expression and genomic alterations between two Dox resistant cell lines (HPGP and MDA 16) and their cell line progenitors (MCF7/wt and MDA468). For comparison with the Dox resistant cell lines a panel of breast carcinoma cells showing a varying intrinsic resistance to Dox was also included in the analysis and the putative normal mammary cell line HBL100, as a common comparison for all cell lines.
3.2. Results

3.2.1 Relative Toxicity of Various Breast Carcinoma Cell Lines to Doxorubicin

Doxorubicin resistant and sensitive progenitor cell lines MDA16, MDA468, HPGP and MCF7/wt and a panel of Dox sensitive breast carcinoma cell lines: BT474, T47D, ZR75-1 and HBL100 were assayed for relative sensitivity to Dox. The cells were grown in 96 well plates in the presence of varying concentrations of Dox for 4 days, with a change of media and Dox at day 2. The viable cells were measured by the Cell titre 96® aqueous one solution cell proliferation assay from Promega. Viable cells were expressed as a percentage of control cells exposed to the Dox vehicle alone, DMSO at 0.1%. The results of this assay are represented by the two graphs in Figure 3.2.

The difference in response to Dox between the resistant cell lines MDA16 and HPGP when compared to their sensitive counterparts MDA468 and MCF7/wt is shown in Figure 3.2.A. The IC50 for the sensitive MDA468 and MCF7/wt cells was 65 nM and 650 nM respectively, whereas the cell viability as a percentage of the control at the top dose (5000 nM) for the resistant MDA16 and HPGP was approximately 80 % and 70 %, so at this top dose the cells did not reach their IC50.

The response of the panel of breast cell lines (BT474, T47D, ZR75-1 and HBL100) to Dox is shown in Figure 3.2.B. These data show a rapid decline in cell viability at 100 nM Dox in the HBL100 and T47D cells, which had an IC50 value of 50 nM and 70 nM respectively. The ZR75-1 cells have an IC50 of 560 nM, whilst the BT474 had an IC50 of approximately 630 nM.
Figure 3.2. Relative Toxicity of Breast Carcinoma Cell Lines to Doxorubicin

A. The resistant and sensitive progenitor cell lines

B. The sensitive cell line panel

Figure 3.2. Legend: Relative toxicity data of the breast carcinoma cell lines to Dox. Cells were grown with Dox for 4 days with a change of media and drug after 48 hr. Viable cell numbers were assessed by the one solution proliferation assay at day 4 and expressed as a percentage of controls, which were exposed to 0.1% DMSO vehicle alone. The IC50 for each cell line is indicated.

Graph A. Shows the resistant cells and their sensitive progenitor cells

Graph B. Shows the sensitive cell line panel with varying response to Dox
3.2.2 Gene Expression Changes and the Doxorubicin Resistant Phenotype

The resistant cell lines (HPGP and MDA16) were analysed by microarray for gene expression and genomic changes against their sensitive progenitor cell line (MCF7/wt and MDA468). All breast cancer cell lines were also analysed against HBL100 cells.

Clustering analysis of the microarray data using the expression scores ascribed to each gene, as described in Materials and Methods (Section 2.2.7.), was performed using the hierarchical clustering algorithms selected by Eisen et al., (1998). Figure 3.3. contains clusters of differential gene expression associated with Dox resistance for both the intrinsically resistant MDA16 cells and the acquired resistant HPGP cells. To the right of the cluster is listed the ‘gene names-96 well location of the clone-chromosome location of the gene’ (e.g. GJA1-10b06-6). Across the top are the cell types arrayed and array number. In each case the RNA population labelled with Cy3 (green) is listed first and the Cy5 (red) labelled population is second (i.e. closest to the top of page). The cluster illustrates that across the different experiments the Dox resistant cell lines clustered together when compared against the unrelated normal mammary cell line HBL100 (green box) and when analysed against their drug sensitive progenitors (red box). The analysis of genomic deletions and amplifications between the resistant cells and their sensitive counter parts also clustered together (blue box).

Clusters of differential gene expression associated with Dox resistance for both the intrinsic and acquired Dox resistant cell lines are shown (Figure 3.3). Figure 3.3.A. is a cluster containing a group of 22 genes that were overexpressed in the resistant cell lines compared to their sensitive progenitor cell lines and also to the putatively normal HBL100 cell line. One of the genes seen overexpressed was the \textit{ABCB1} (ATP-binding cassette, sub-family B (MDR/TAP) member 1), which was expected, as both cell lines are known to express the \textit{ABCB1} gene product P-gp. Other genes in this cluster associated with the resistance phenotype included \textit{MMP1} (matrix metalloproteinase 1), \textit{MSLN} (pre-pro-megakaryocyte potentiating factor), \textit{LSP1} (lymphocyte specific protein 1) and \textit{OXTR} (oxytocin receptor).

Genes that were over expressed in the drug resistant lines as compared to their sensitive counterparts and also overexpressed in the immortalised HBL100 against MCF7/wt cells are shown in Figure 3.3.B. Conversely these genes were down regulated in the panel of breast cell lines against HBL100. Amongst the genes in this cluster were \textit{GST\textpi} (glutathione-S-transferase pi), \textit{SPARC} (secreted protein, acidic, cysteine-rich (osteonectin)) and \textit{SCYA19} (small inducible cytokine subfamily A (Cys-Cys), member 19).
Figure 3.3 Clusters of Gene Expression and Gene Amplification Associated With the Intrinsic and Acquired MDR Phenotype
Figure 3.3. Legend: Clusters of gene expression and gene amplification associated with the intrinsic and acquired MDR phenotypes. Cell pairs are shown on the top with the first cell type of each pair being the reference (Cy3 labelled) and the latter the test (Cy5 labelled). The data was clustered along both the vertical and horizontal axis. Thus the clustering of the replicate experiments across the top reflects their similarity and is not an artificial grouping. These experiment groups include Dox resistant cells compared with their drug sensitive progenitors (red box), genomic alterations in the Dox resistance cells when compared with their drug sensitive progenitors (blue box), and Dox resistance cells compared with the putative normal sensitive HBL100 cell line (green box). The scale of fold change is shown in the bar representing the ratio score between the pairs of RNA populations.

Cluster A. Genes overexpressed in both the Dox resistant cell lines relative to their drug sensitive progenitors and HBL100 cells

Cluster B. Two clusters of genes overexpressed in both Dox resistant cells relative to their progenitors but not relative to HBL100. These genes were thus over expressed in HBL100 cells relative to MCF7/wt and also down regulated in the panel of breast cell lines compared to HBL100

Cluster C. Gene amplifications present in both the intrinsic and acquired resistant phenotype cells relative to their drug sensitive progenitor cells.
3.2.3. Genomic Amplification and the Doxorubicin Resistant Phenotype

Figure 3.3C is a small cluster of amplified genes specific to the Dox resistant cell lines compared to their sensitive progenitors. There are other genes amplified that do not fall within this cluster diagram, for example the ABCB1 gene that is clustered in Figure 3.3.A. All genes amplified by greater than 2-fold in both the resistant cell lines are summarised in Table 3.1. There is one exception, KPNA2, which clusters with the amplified genes (Figure 3.2C), but its average ratio is only 1.69 in MDA16 cells. 1.69 is the ratio from one experiment, as for the other experiments a ratio could not be calculated, as the gene was only present in MDA16 cells and no signal was observed in MDA468 cells. Consequently this gene is still considered to be amplified in the resistant cell line compared its progenitor cells. The table also shows the chromosome location and RNA expression of each gene in the resistant cell lines.

Table 3.1. Genomic Amplifications in Resistant Cell Lines

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Chr</th>
<th>DNA</th>
<th>RNA expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MDA16</td>
<td>HPGP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ratio</td>
<td>Stdev</td>
</tr>
<tr>
<td>ABCB4</td>
<td>7q21.1</td>
<td>12.19</td>
<td>1.89</td>
</tr>
<tr>
<td>ABCB1</td>
<td>7q21.1</td>
<td>7.20</td>
<td>0.14</td>
</tr>
<tr>
<td>BAK1</td>
<td>6p21.3</td>
<td>7.85</td>
<td>3.65</td>
</tr>
<tr>
<td>HNRPA1</td>
<td>12q13.1</td>
<td>2.93</td>
<td>0.88</td>
</tr>
<tr>
<td>ORM2</td>
<td>9q32</td>
<td>2.98</td>
<td>0.22</td>
</tr>
<tr>
<td>RBMX</td>
<td>Xq26</td>
<td>2.60</td>
<td>0.14</td>
</tr>
<tr>
<td>NGFR</td>
<td>17q21-22</td>
<td>2.53</td>
<td>0.56</td>
</tr>
<tr>
<td>BAD</td>
<td>11q13.1</td>
<td>1.99</td>
<td>0.28</td>
</tr>
<tr>
<td>KPNA2</td>
<td>17q23.1</td>
<td>1.69</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 3.1. Legend: A summary of genomic changes that are present in both resistant cell lines when compared to their sensitive progenitor cell lines. The gene abbreviation, chromosome location, DNA amplification and RNA expression data are shown. All fold change or ratio values are an average of \( n = 2 \) or 4 experiments standard deviations are shown. Ratio values highlighted in red are considered amplified or overexpressed in the resistant cells relative to their Dox sensitive progenitor cells.
Two of the genes amplified in the resistant cell lines are members of the ABC family: \textit{ABCB1} and \textit{ABCB4}. Amplification of the \textit{ABCB1} gene has previously been recorded in cells exhibiting an MDR phenotype (Fairchild et al., 1987; Shen et al., 1986) and \textit{in vitro}, amplification of the \textit{ABCB1} genes precedes an MDR phenotype (Shen et al., 1986). The two ABC genes, \textit{ABCB1} and \textit{ABCB4} both lie next to each other on chromosome 7q21.1 and have previously been shown to amplify together as a cassette, however only the \textit{ABCB1} gene is also overexpressed at the RNA level. Two other genes that are amplified and slightly overexpressed in the resistant cells are \textit{HNRPA1} and \textit{RBMX}. The remaining 6 genes, \textit{ABCB4}, \textit{Bak-1}, \textit{ORM2}, \textit{NGFR}, \textit{BAD} and \textit{KPNA2} are amplified but not overexpressed.

\textbf{3.2.4. Gene Expression in the Breast Carcinoma Cell Lines Compared to HBL100}

Genes over expressed in the breast carcinoma cell line panel relative to HBL100 are shown in Figure 3.4. Figure 3.4 (A) are clusters of genes whose overexpression relative to the HBL100 are also maintained in the Dox resistant cell populations. Within these clusters are five different keratin genes.

The cluster represented in figure 3.4.B. contains 2 genes, \textit{FPRL1} and \textit{ESR} (oestrogen receptor), which are also overexpressed in the breast carcinoma cell line panel compared to HBL100. Also both genes are expressed at higher levels in MCF7/wt compared to its drug resistant counterpart HPGP, whilst there are no differences in expression of these genes between the MDA468 cells and their resistant counterpart MDA16. The MDA468 cells are oestrogen negative, thus the resistant MDA16 cells have maintained this phenotype. The HBL100 cells also appear to be \textit{ESR} negative and this is confirmed in literature (Bamberger et al., 1999).
Figure 3.4. Clusters of Gene Overexpression in the Breast Carcinoma Cell Lines Relative to HBL100

Cluster A. Genes over expressed in the carcinoma and DOX resistant cells relative to HBL100 cells, but not (in the case of the Dox resistant cells) relative to their progenitor cells (MCF7/wt and MDA468). This also includes two additional keratin genes, which clustered separately.

Cluster B. A gene expression cluster containing the ESRα gene indicating the ESRα negative receptor status of the putative normal mammary epithelial cell line HBL100.
3.2.5. Confirmation of Microarray Results

To confirm the fluorescent detection microarray results a selection of 48 genes were analysed using nylon microarrays probed with $^{33}$P labelled RNA. The gene expression direction for the majority of genes were confirmed e.g. overexpressed, underexpressed or not differentially expressed. A selection of these genes are represented in Figure 3.5., which is a graph of ratio values from the nylon arrays, and the corresponding microarray data is shown below. The $ABCB1$ gene overexpression detected with microarray analysis was confirmed with the nylon membranes, although the degree of expression varies in both sets of data. For example HPGP cells were shown to over express $ABCB1$ by approximately 5-16 fold across 3 microarray experiments, whilst the nylon data gave fold change values of approximately 3.8 fold increase.

**Figure 3.5. Confirmation of Microarray Data with Nylon Membranes**

![Graph showing confirmation of microarray data with nylon membranes](image)

Figure 3.5. Legend: Confirmation of microarray data by nylon arrays. The nylon data are shown graphically and the microarray results are underneath. The y-axis represents the ratio or fold change figure (resistant/sensitive). $ABCB1$ and $RBMX$ were overexpressed in the 2 resistant cells in both techniques. Gene 20g08 was only overexpressed in the HPGP cells in both techniques and not MDA16. $BAD$, $ORM2$ and $ABCB4$, which were all amplified in the resistant cells, were not expressed in these cells in both techniques.
Overexpression of the \textit{RBMX} gene in both resistant cells compared to their sensitive progenitor cells was also confirmed. Another gene differentially expressed is 20c08, whose overexpression was specific to HPGP cells but not MDA16 cells, which was also confirmed with the nylon membranes. Three genes \textit{BAD}, \textit{ORM2} and \textit{ABCB4}, that were amplified but not overexpressed in the resistant cell lines by microarrays, was also confirmed.

Gene expression of several genes was also confirmed using Real Time-PCR. Figure 3.6. contains 3 of these genes \textit{OXTR}, Bak-1 and \textit{ABCB4}. The \textit{OXTR} gene was found overexpressed in the resistant cells, HPGP and MDA16 only and all other genes were not overexpressed in any breast cell lines compared to HBL100 in both techniques.

**Figure 3.6. Confirmation of Microarray Data with Real Time PCR**

![Graph](image)

Figure 3.6. Legend: Figure 3.5. Legend: Confirmation of microarray data by Real Time PCR. The Real Time PCR data is shown graphically. The y-axis represents the relative expression of the genes in each cell line when compared to HBL100. In agreement with the microarray data, \textit{OXTR} was overexpressed by Real Time PCR in both the resistant MDA16 and HPGP cells and \textit{Bak-1} and \textit{ABCB4}, which were both amplified but not overexpressed on the microarray were also not overexpressed by Real Time PCR.
3.3. Discussion

3.3.1. Gene Expression Changes and the Doxorubicin Resistant Phenotype

Both the intrinsic and acquired MDR resistance cells clustered together and showed a good correlation of gene expression with each other. This suggests that a tumour will develop a MDR phenotype, not as a result of cells acquiring such a phenotype, but rather through the cytotoxic drug selection of cells that already have such a phenotype. Therefore a tumour that has acquired drug resistance has only done so by the existence of intrinsically resistant cells already present as a sub-population within the tumour.

The genes that were differentially expressed in the drug resistant cell lines fell into 3 main categories. Those which were are known to be involved in the resistant phenotype and so expected, those which could be related to other known phenotypic properties of the cells in addition to resistance and those whose reason for differential expression is unclear at the present time.

One of the expected genes seen over expressed in the drug resistant cell lines compared to both their sensitive progenitor cells and HBL100 was **ABCB1** (ATP-binding cassette, sub-family B (MDR/TAP), member 1), which was both over expressed and amplified. Overexpression of this gene was expected, as it codes for the P-gp transporter and HPGP are known to highly express P-gp, while MDA16 cells were selected for with an antibody directed to the P-gp. Expression of **GSTπ** has also been associated with **ABCB1** gene expression and with the acquired resistance phenotype to a number of drugs including Dox (Batist et al. 1986). Therefore as expected, overexpression of **GSTπ** was detected in the resistant HPGP and MDA16 cells relative to their drug sensitive progenitors (Figure 3.3.B.). Clustered next to **GSTπ** was the glutathione peroxidase gene (**GPX1**) which catalyses the reduction of hydroperoxides by glutathione and is an important enzymic defence against these toxic oxidant species (Shen et al., 1994). Interestingly the putative normal cell breast carcinoma cell line HBL100 also had an elevated **GSTπ** and **GPX1** expression compared with drug sensitive carcinoma cells.

Another gene previously associated with a drug resistant phenotype **GLO1** (Section 1.4.8.), was not overexpressed in these cell lines. This gene contributes to a drug resistant phenotype independent of **ABCB1** (Sakamoto et al., 2000), and so its lack of overexpression in the resistant cells, which overexpress **ABCB1** in this study, confirms that **GLO1** resistance is mediated differently.
Included in the resistant cluster was *MMP1* (Matrix metalloproteinase 1), a member of a proteolytic enzyme family, which contain a zinc ion at their active sites and can degrade collagens and other components of the extracellular matrix (Nutt and Lunec, 1996). Overexpression of *MMP1* is associated with many stages of tumour progression, including metastasis, growth and invasion of tumour cells into surrounding tissue (Benbow et al., 1999). Cells with an MDR phenotype have been reported to have an increased invasive ability (Weinstein et al. 1991), thus the increased expression of *MMP1* in the drug resistant cells is consistent with this characteristic. Also chelation of heavy metals by metallothionein is thought to be involved in MDR (Cherian et al., 1993). The matrix metalloproteinases require activation from their latent forms, a process that can be regulated by several proteins called tissue inhibitors of matrix metalloproteinases (*TIMPS*) (Matrisian, 1990). *TIMP2* was also contained in the cluster associated with Dox resistance as it was found overexpressed in the resistant cells compared to their sensitive progenitors, but not in the resistant cells when compared to HBL100 (Figure 3.3.B.). A schematic diagram displaying the relationship between *MMP1* and *TIMP2* gene expression in the different cell types is shown in Figure 3.6, which may suggest that in response to the elevated levels of *MMP1*, the resistant cells are responding by overexpressing an inhibitor.

**Figure 3.6.** Legend: The relationship of gene expression of *MMP1* (A) and *TIMP2* (B) in the different cell lines is summarised. The arrows indicate the direction of comparison. *MMP1* was overexpressed in resistant cells compared to both sensitive progenitor cells and HBL100. *TIMP2* overexpression was only seen in resistant cells when compared to their sensitive cells.
Other genes associated with cancer metastasis that were present within the clusters associated with the resistant phenotype include PLAU, SPARC and VCAM1. Like MMP1 increased production of urokinase is associated with cancer metastases (Stein et al., 1993) and the PLAU (urokinase-plasminogen activator) is within the cluster of genes overexpressed in the resistant cell lines compared to their progenitor cell line. SPARC is a glycoprotein that belongs to a group of matrix-associated factors, which mediate cell-matrix interactions but do not serve in primarily structural roles (Brekken and Sage, 2000). SPARC expression is increased in invasive malignant tumours (Porter et al., 1995), and it has been shown that SPARC is associated with members of the MMP family and increases the activity of MMP2 in invasive human breast cell lines (Gilles et al., 1998). Thus SPARC may also have an effect on other members of the matrix metalloproteinase family, such as MMP1, which was also overexpressed here. There are a number of potential regulators of SPARC mRNA production including the transforming growth factor TGFB1, which has been to shown to increase SPARC expression by 3-4 fold in human fibroblasts after 24 hr and TGFB1 was also overexpressed in the resistant cells. VCAM1 (Vascular cell adhesion molecule 1), which was also overexpressed in the resistant cell lines, is a cell surface P-glycoprotein and is also thought to play a role in tumour metastases (Rice and Bevilacqua, 1989).

Two genes with suggested potential roles in cancer treatment that were present in the resistant cluster are PRAME (preferentially expressed antigen in melanoma) and NK4. Studies have previously reported PRAME over expression in tumours (Steinbach et al., 2002; Pellat-Deceunynck et al., 2000), while it is only expressed at low levels in a few normal tissues, thus making it a good candidate for tumour immunotherapy (Watari et al., 2000). However recently PRAME overexpression has been associated with significantly higher rates of disease free survival and overall survival in childhood acute myeloid leukemia (Steinbach et al., 2002), which may hamper its success as a candidate for tumour immunotherapy. Even though it is over expressed in tumours, PRAME has not previously been linked to the Dox resistant phenotype, although it is confusing to see a gene that has been linked to a favourable prognosis being involved in drug resistance. The presence of NK4 within this cluster is also controversial. NK4 is an antagonist of HGF (hepatocyte growth factor), which is involved in promoting invasion and metastasis, thus NK4 has been suggested as a tool in cancer treatment (Maemondo et al., 2002). Therefore it is apparently contradictory to see both MMP1 and PLAU, which are associated with increased invasion and metastasis and NK4, an invasion and metastatic antagonist within the same cluster. However increased NK4 expression in the resistant cells may suggest that its proposed role in cancer treatment should be followed with
caution, as if resistant cells are over expressing NK4, then treatment of these cells with NK4 may have little effect.

A gene associated with the resistant phenotype that may be a potential target for treatment of resistant tumours is OXTR (Oxytocin receptor). This gene is expressed in the uterus prior to the onset of labour and in myo-epithelium where it allows responsiveness to oxytocin for milk ejection during lactation (Maggi et al., 1994). Expression of the OXTR has previously been reported in endometrial adenocarcinomas (Cassoni et al., 2000) and breast tumours, and oxytocin is thought to have a different biological role in tumours derived from the same organ (Cassoni et al., 2000). In the human endometrial OXTR positive cell line (COLO684) and the MDA-MB231 breast cancer cell line growth and proliferation are retarded in the presence of oxytocin via OXTR (Cassoni et al., 1997). OXTR may thus have a greater effect on retarding growth of the multi drug resistant cells that express high levels of OXTR compared to their drug sensitive counterparts. Due to the specific expression of oxytocin and OXTR in normal physiology, overexpression of OXTR in resistant cells may offer an option in oxytocin drug development.

Other genes that were over expressed in the drug resistant cell lines include, MSLN (pre-pro-megakaryocyte potentiating factor), otherwise known as mesothelin and LSP1. Over expression of MSLN has previously been described in ovarian cancers and mesothelioma cells (Chang and Pastan, 1996). The LSP1 gene was originally isolated from T-lymphocytes, and although its function is unclear it is thought to be involved in cytoskeleton mediated cellular processes in leukocytes (Miyoshi et al., 2001). LSP1 is down regulated in cells undergoing apoptosis (Brockstedt et al. 1998). Thus over expression of LSP1 in the resistant cells could indicate that these cells may less readily undergo apoptosis and resistance to apoptosis has also been associated with increased ABCB1 gene expression (Robinson et al. 1997).

Genes within the resistance-associated cluster whose presence is confusing and difficult to understand or suggest include the Coch gene (hereditary deafness gene), which is involved in hereditary hearing and vestibular disorders (Ikezono et al., 2001). However the Coch gene is expressed in secretory glands, which may explain the presence of the gene in breast cells.

3.3.2. Genomic Amplification and the Doxorubicin Resistant Phenotype

There were 8 genes amplified at the genomic level in both the resistant cell lines. Three of which, the ABCB1, HNRPA1 (heterogeneous nuclear ribonucleoprotein A1) and RBMX (RNA
binding motive protein, X chromosome) genes were also over expressed at the RNA level. Amplification of the $ABCB1$ gene in multi drug resistant lines has been previously described in the literature (Fairchild et al. 1987; Shen et al. 1986). Another member of the ABC family, $ABCB4$ was also amplified but not overexpressed. $ABCB4$ encodes a protein, which does not confer an MDR phenotype when transfected into cells (Abulrob and Gumbleton, 1999), although its expression is detected in some cancers and it has been suggested to play a possible role in resistance in leukaemias (Arai et al., 1997). $ABCB4$ is located alongside $ABCB1$ at chromosome 7q21.1 and its amplification in a cassette form with the $ABCB1$ gene is recognised. Since the gene is not overexpressed in the Dox resistant cells, it is thus not contributing to the resistance phenotype; it is probably amplified due to its close proximity to the $ABCB1$ gene. Its failure to be overexpressed when amplified may also suggest that the gene has not been amplified in a complete form.

$HNRPA1$ and $RBMX$, the other two genes that were both amplified and overexpressed in the resistant cells are both involved in RNA production. $HNRPA1$ is involved in the packaging of pre-mRNA and may also modulate splice site selection. $RBMX$ encodes a RNA-binding protein that is also involved in splicing (Lingenfelter et al., 2001).

$Bak-1$ and $BAD$ were also amplified, but not overexpressed. These two genes are located at chromosome 6p21 and 11q13.1 respectively and are members of a subgroup of the $Bcl-2$ (B cell leukemia-2) family (Adams and Cory 1998). The genes are both $Bcl-2$ antagonists and so are thought to have the ability to induce apoptosis. Thus it seems contradictory to see these genes amplified in the Dox tumour cells, as cells with increased $ABCB1$ gene expression are associated with a resistance to apoptosis (Robinson et al., 1997). However these genes are again not overexpressed and so not contributing to the resistant phenotype. Perhaps the tumour cells are attempting to undergo apoptosis by amplifying these genes, but an up or downstream event in the resistant tumour cells is preventing this gene amplification being carried through to the RNA level.

$Bak-1$ has not previously been associated with a drug resistance phenotype. Whilst expression of $BAD$ has been reported as lower in Dox resistant cells (Watts et al., 2000). However $Bak-1$ is located on chromosome 6p21.3 close to several members of the ABC family: $ABCB2$, $ABCB3$, $ABCC10$ and $ABCF1$. $ABCB2$ and $ABCB3$ are involved with the major histocompatibility complex (MHC) in the process of peptide presentation (Klein et al., 1999). The $ABCB2$ and $ABCB3$ proteins form a heterodimer that actively transports peptides, which
have been degraded by peptidase from the cytosol into the endoplasmic reticulum lumen, where they associate with MHC class I molecules. The two ABC genes are located in the MHC class II region in chromosome 6p21.3, close to Bak-1 (Beck et al., 1992), which may be co-amplified with these genes.

Other genes that were amplified include ORM2 and NGFR. Since these genes are not overexpressed their involvement in the resistance phenotype is questionable. These amplifications may be spontaneous, as DNA replication occurs with less than 100% accuracy (Cohen and Ellwein, 1991). However as the amplifications were detected specifically in both the resistant cell lines it suggests that they are indirectly involved, similar to the ABCB4 gene which was co-amplified with ABCB1 but not overexpressed; or that they may be a prerequisite for the development of a MDR phenotype.

3.3.3. Gene Expression in the Breast Carcinoma Cell Lines Compared to HBL100

Clusters of genes that were overexpressed in the breast carcinoma cell lines relative to the putative normal HBL100 cell line are shown in Figure 3.3. There is an abundance of keratin genes included in the cluster relative to HBL100, including keratin 18. Overexpression of keratin 18 has been previously recorded in other studies of breast carcinoma cells using microarrays (Perou et al., 2000; Perou et al., 1999). The keratin 18 protein dimerises with keratin 8, which was also overexpressed here in the resistant cells. Overexpression of this gene has also been noted in a recent microarray study of breast cancer (Martin et al. 2000) and is thought to be associated with tumour size. Overexpression of these genes in the carcinoma breast cell lines relative to HBL100 was maintained in the MDR phenotype.

CD9 (cluster of differentiation 9) is involved in osteoclastogenesis (Tanio et al., 1999) and was also present in the cluster of genes that are overexpressed in both the resistant and sensitive carcinoma cells relative to HBL100. A more recent study has confirmed the overexpression of CD9 in Dox resistant cell lines (Watts et al., 2001).

The ESR (oestrogen receptor) was also overexpressed in the panel of drug carcinoma cell lines relative to HBL100. This was expected as the HBL100 cells are known to be oestrogen receptor negative (Bamberger et al., 1999), whilst the panel of sensitive cell lines that includes ZR75.1, T47D, MCF7/wt and BT474 are all known to be oestrogen receptor positive. The MDA468 cell lines are oestrogen receptor negative and the MDA16 resistant cells have maintained this phenotype (Figure 3.4.B.). In contrast the HPGP cells derived from
oestrogen receptor positive MCF7/wt cells have acquired an ESR negative phenotype during the process of developing the resistant phenotype.

3.3.4. Summary

In this study microarrays were used to analyse differential gene expression in breast carcinoma cells both resistant and sensitive to Dox. Genes known to be involved in the MDR phenotype such as ABCB1 and glutathiones were overexpressed in the resistant cells compared to their sensitive progenitor cells. A group of genes involved in metastases were also associated with the resistant phenotype, as well as genes that have been proposed as targets in cancer treatment. Interactions between genes that were overexpressed or amplified in the resistant phenotype have been suggested e.g. the interactions between TIMP2 and MMP1.

Nine genes were amplified in the resistant cell lines when compared to their sensitive pairs. Three of these genes were also overexpressed, one of which was the ABCB1 gene. The ABCB4 gene was co-amplified with ABCB1, but not overexpressed.

Expressed genes in both the intrinsic and acquired resistant phenotypes caused these cells to be clustered together and a differential gene expression pattern, associated with both types of resistant breast cancer cell lines was identified. This could imply that the HPGP cells that are thought to have acquired resistance on exposure to Dox may have had an intrinsic resistance at the outset and may have been derived from a mixed cell population that contained both sensitive cells and cells with an intrinsic MDR phenotype, the latter were then selected for via Dox exposure and survived.

The gene expression and DNA amplification patterns shown here could provide a useful diagnostic marker of the Dox resistance phenotype and some of the genes expressed specifically in the resistant cell lines have the potential to be useful chemotherapeutic targets in the treatment of Dox resistant carcinomas.
Chapter 4

Gene Changes and Development of the Resistance Phenotype
4.1. Introduction

Resistance to the chemotherapeutic agent Dox is known to develop in breast cancer treatment (Kroger et al., 1999) and in the previous chapter microarrays were used to compare cells with an intrinsic and acquired resistance with those exhibiting a varying sensitivity to Dox. Although this work elucidated some common patterns of gene amplification and overexpression in the Dox resistant cells, it gave no information of the gene changes that proceeds and are perhaps instrumental to the development of a resistant phenotype.

4.1.1. Selection of Drug Resistant Cells

A widely used technique for studying mechanisms of acquiring resistance is to utilise the phenomenon that cells may acquire resistance and can be selected for resistance by a stepwise increase in drug concentration (Matsumoto et al., 2001; Shen et al., 1995; Gottesman, 1987; Akiyama et al., 1985). In some cases it may be necessary to utilise mutagens, such as ethyl methanesulfate (EMS), to initially mutate cells prior to selection with drug to elevate the mutation frequency to a detectable level (Thompson, 1987; Shen et al., 1986). Although there are potential complications with the use of a mutagen, as DNA damage will occur at many other sites not just the target genes, and such secondary mutations may affect other properties of the cell.

4.1.2. Mechanisms in Drug Resistance Selection

As previously discussed drug resistance is a multifactorial process and several mechanisms are thought to contribute to the phenotype. While the contribution of factors to the overall phenotype can be analysed in the resistant cells which represent in essence the end stage, it is important both for a complete understanding of the mechanisms of resistance and for therapeutic interaction to discern the processes that give rise to the drug resistant phenotype. That is those factors that initiate and maintain the progression prior to the full development of the phenotype. One study that tried to address this was by Shen et al., (1986), who selected human KB carcinoma cells for resistance with drug exposure and looked at expression of the \(ABCB1\) gene. In these cells there was first overexpression of the \(ABCB1\) gene, followed by further increases in \(ABCB1\) gene expression accompanied with genomic amplification of the \(ABCB1\) gene at higher selection levels of Dox (Shen et al., 1986). Such overexpression of the \(ABCB1\) gene in cells selected by drug for resistance is often thought to be associated with the expression of a single allele (Mickley et al., 1998) and the \(ABCB1\) gene is to be a known a common contributor to MDR.
Other members of the ABC family may be involved in resistance and a study by Yoshida et al. (2001) that analysed several ABC transporter family members by Real Time PCR found that after 48 hrs incubation with Dox, induction of the \textit{ABCB1}, \textit{MRP5} and a splice variant of \textit{MRP5}, \textit{SMRP} occurred in human lung resistant cancer cells (Yoshida et al., 2001). As these genes are induced after 48 hrs it was suggested that these genes are unlikely to be involved in early response of the cell to the drug and that they may be mediated by events after initial exposure (Yoshida et al., 2001). Early response genes may play a role in the later development of a resistant phenotype, while others may occur as a result of the initial cell response to the drug.

A study, which analysed protein changes in cells selected for resistance to different drugs, including Dox, vinblastine, mezerein and colchicine suggested that only a limited number of protein changes occur in the development of the MDR phenotype (Shen et al., 1985). These proteins included an increase in a 170 kDa protein, which is probably P-glycoprotein encoded by the \textit{ABCB1} gene. Other changes seen included decreased levels of members of a family of proteins of 70-80 kDa and an increase in protein synthesis of a protein with a mass of 21 kDa, the latter change was only observed in cells resistant to colchicine (Shen et al., 1985).

Other genes are known to be down regulated upon exposure to Dox, such as the down regulation of the heat shock protein 27 (\textit{HSP27}) after treatment with 0.1 \mu M Dox for 2 days in MCF7 cells (Chen et al., 2002), although whether this gene is likely to contribute to a resistant phenotype or is a response of the cells to the drug is not known.

\textit{4.1.3. Chapter Objectives}

There is extensive evidence of \textit{ABCB1} and \textit{ABCC1} activity \textit{in vitro} and their involvement in the multi drug resistant phenotype in certain tumour types. However the underlying mechanisms contributing to the development of resistance remains unclear.

In this chapter MCF7/wt cells were initially exposed to Dox over a 7 day period to elucidate early gene changes occurring in response to the drug. The drug sensitive cell lines MCF7/wt and MDA468 were then exposed to Dox by a continuous or stepwise increase in drug concentration to select cells for resistance. Microarray analysis was carried out at various stages with the aim to elucidate some of the genetic events occurring during the development of resistance and to detect the genes associated with this process.
4.2. Results

4.2.1. Cytotoxicity Assay for MCF7/wt and MDA468 Cells to Doxorubicin

A cytotoxicity assay after 1 week exposure to Dox was initially carried out with the sensitive MCF7/wt and MDA468 cells to choose a suitable concentration to dose to the cells (Figure 4.1.).

Figure 4.1. Cytotoxicity Assay of MCF7/wt and MDA468 Cells to Doxorubicin

![Cytotoxicity Assay of MCF7/wt and MDA468 Cells to Doxorubicin](image)

Figure 4.1. Legend: Cytotoxicity data for MCF7/wt and MDA468 cells exposed to Dox. The percentage of viable cells when compared against the non treated cells dosed with DMSO vehicle only is on the y axis and the log 2 dose is shown on the x-axis. The experiment was carried out in triplicate and standard deviations are represented by the error bars. The IC$_{50}$ for MCF7/wt cells dosed for 1 week with Dox was 20 nM and approximately 30 nM for MDA468 cells.

It has been suggested that to achieve a resistant cell population the cells should be dosed at a concentration that is lethal to 90% of the population, which in this case was approximately 50 nM. When the MCF7/wt and MDA468 cells were dosed with 50 nM the majority of cells died and the remaining cells went into a period of growth arrest followed by cell death. An alternative approach, as discussed earlier, is to firstly mutagenise the cells then select for resistance with drug, but this method has complications, as the mutagen is affecting the DNA
of the cells indiscriminately and secondary effects may occur. Also when exposed to the mutagen EMS (Section 2.2.3.3.), the majority of cells died or went into a period of growth arrest followed by cell death. Consequently the cells were dosed at their IC_{50}, followed by continuous dosing at the same dose or a stepwise increase in drug to encourage the development of resistance.

When the MDA468 cells were exposed to a step increase in Dox concentration the majority of cells died and the remaining cells went into a period of growth arrest followed by cell death. After a period of 7 weeks continuous exposure of the MDA468 cells the cells began to die, consequently resistance to Dox did not develop in these cells, thus all experiments in this chapter were carried out in the MCF7/wt cells alone. When MCF7/wt cells were dosed with Dox they also went into a period of growth arrest, but subsequently recovered and continued proliferation.

MCF7/wt cells were dosed with Dox as illustrated in the dosing scheme in Figure 4.2. Three separate dosing schemes were carried out, the first was a continuous weekly dosing of cells with 20 nM Dox (the IC_{50}), and each week sample was represented by the week or dose number. The second was a stepwise increase in drug concentration, with the amount of drug doubling at each step increase, and the week number and a letter corresponding to drug dose represent each sample. The final scheme was initially a stepwise increase, and at week 9A the cells had become slightly resistant to Dox, therefore dosing was stopped in order to see if the resistant phenotype was stable in the cells without the presence of drug. These samples were labelled 9A followed by a number corresponding to number of weeks free from drug.

At several stages RNA was isolated from both dosed and control unexposed MCF7/wt cells where no drug was administered for comparison on the array. Cytotoxicity assays were carried out at the end of each week to monitor changes in sensitivity of the cells to Dox. At several stages DNA was also extracted to analyse cells for gene amplification. Prior to the 14 week study a 1 week time course of MCF7/wt cells dosed with 20 nM was carried out, this is discussed in the next section 4.2.2.
Figure 4.2. Legend: Doxorubicin dosing scheme for MCF7/wt cells over 14 weeks. The 3 dosing schemes are described. The number at each stage represents the week number and the dose concentration for each week is shown in italics. For all stages control MCF7/wt cells with no drug were grown alongside dosed cells for comparison on the microarray.

Scheme 1 - Continuous
Continuous 14 week Dox dosing at 20 nm

Scheme 2 - Stepwise
Stepwise increase in Dox dosing over 14 weeks from 20 nm to 120 nm

Scheme 3 - Stepwise then 0 dose
Stepwise increase in Dos dosing over 9 weeks from 20 nm to 40 nm when cells became resistant they were left for a period of 5 weeks in the absence of Dox to ascertain the stability of the phenotype.

Figure 4.2. Dox Dosing Scheme for the MCF7/wt Cells
4.2.2. Gene Changes After a 7 Day Time Course Exposure to 20 nM Doxorubicin

To examine possible early changes that are occurring at the initial IC$_{50}$ dose a 7 day time course was set up. Cells were dosed with 20 nM Dox and RNA isolations were performed at 2 hr, 6 hr, 1, 2, 3, 4, 5, 6 and 7 day after exposure. At all time points control MCF7/wt cells with no drug were grown alongside dosed cells and RNA isolated for comparison on the microarray. Two experimental repeats were carried out and the ratio data was used for hierarchical clustering.

Gene expression changes over the 7 day period with 20 nM Dox were limited with small fold changes of around 2. There were no significant gene expression changes at the 2 hr or 6 hr time points. After 1 day exposure to Dox several genes were overexpressed and overexpression of these genes was maintained through to 7 days exposure (Figure 4.3.). There are 2 separate clusters in the figure, cluster A contains several CDKN genes and ATP1B1, all of which were overexpressed by 2 - >5 fold after a 1 day exposure and this expression was maintained in the cells after 7 days exposure. CDKN1A is represented on the array several times, and the replicates clustered together in this experiment validating the microarray data. The second cluster, B represents genes that were slightly overexpressed after day1 and this expression was maintained or increased after 7 days exposure of the cells with 20 nM Dox. A gene that was in this cluster, FN1 was not overexpressed until 4 days exposure of cells Dox, after which time its expression rises to >5 fold.

There were several clusters of genes that were overexpressed at later time points, these are shown in 2 clusters (Figure 4.4). The first cluster contains genes that were overexpressed after a 4 day exposure, but after 6 days the genes were no longer differentially expressed in the cells dosed with 20 nM Dox. The genes in the second cluster were overexpressed in MCF7/wt cells after 4 days, and this overexpression was maintained until 7 days exposure with drug.
Figure 4.3. Gene Changes in MCF7 Cells After 1 Day with 20 nM Doxorubicin

Figure 4.3. Legend: Genes overexpressed after a 1 day exposure to 20 nM Dox, which remain overexpressed in all subsequent time points of Dox exposure. Each experiment has been carried out as a biological duplicate, Cy5 labelled RNA is the dosed MCF7/wt cells (red) and Cy3 is the unexposed reference MCF7/wt cells (green). The experiment details are across the top and gene names are listed down the right. The first number in the experiment details corresponds to the time point, and the latter is the array chip number. The relative fold change is shown on the left. There are 2 separate clusters:

Cluster A: Represents genes that were rapidly overexpressed, reaching peak expression after a 1 day exposure to Dox.

Cluster B: Contains genes that were slightly overexpressed after a 1 day Dox exposure, and whose expression was maintained or partially increased to day 7.
Figure 4.4. Legend: Genes that were overexpressed after a 4 day exposure to 20 nM Dox. Each experiment has been carried out in duplicate, Cy5 labelled RNA is the dosed MCF7/wt cells (red) and Cy3 is the unexposed reference MCF7/wt cells (green). The experiment details are across the top and gene names are listed down the right. The relative fold change is shown on the left. There are 2 separate clusters:

Cluster A: Represents genes that were highly overexpressed after 4 days exposure with Dox and this over expression was maintained until after 5 days Dox exposure, then expression of the genes decreased

Cluster B: Contains genes that were overexpressed after 4 days exposure with Dox and this expression was maintained in cells dosed up to 7 days with Dox.
4.2.3. Cytotoxicity Data from the 14 Week Doxorubicin Dosing Study

At each stage in the scheme in Figure 4.2. a cytotoxicity assay was carried out on the MCF7/wt cells dosed with Dox. The results of these cytotoxicity graphs for Scheme 1 cells (weeks 1-14) are shown in Figure 4.5. In general there was little difference in the response of cells to Dox after 14 weeks of continuously dosing the cells at 20 nM, and the IC\textsubscript{50} varies from 10-50 nM. Weeks 9 and 11 had the highest IC\textsubscript{50} values, although there appears to be no correlation to relative resistance and length of time dosing, as week 4 had the lowest IC\textsubscript{50} and week 13 was in the middle.

The cells that have undergone a stepwise increase to Dox (Scheme 2) became more resistant to Dox, with a marked difference in response to the drug between stages 6A and 7A (Figure 4.6.). The IC\textsubscript{50} of subsequent resistant samples rose to greater than 100 nM, although there appeared to be no further increase in the level of resistance as the Dox concentration was increased further.

Figure 4.7. summarises the cytotoxicity data for the cells grown in scheme 3, whereby once resistance was achieved the cells were grown in the absence of Dox. Once resistance was achieved at 9A (Figure 4.6.) some of the cells were grown without drug for 5 weeks (9A01-9A05) and after 3 weeks the cells maintained a slight resistance to Dox, although there appeared to be a decrease in resistance in the cells grown for 4 and 5 weeks without Dox (94-04 and 9A-05 respectively).
Figure 4.5. Cytotoxicity Assays in the 14 Week Study: Scheme 1

Figure 4.5. Legend: Cytotoxicity results of the cells dosed in Scheme 1 weeks 1-14 (20 nM) (Figure 4.2.). The week number is shown in the legend on the right. The IC$_{50}$ in the cells varied from approximately 10-50 nM, in general all the cells show a similar trend.

Figure 4.6. Cytotoxicity Assays in the 14 Week Study: Scheme 2

Figure 4.6. Legend: Cytotoxicity results of the cells dosed in Scheme 2 weeks 1-14C (Figure 4.2.). The week number is shown in the legend on the right. Weeks 1-5 at 20 nM are shown in blue, 6A and 7A at 40 nM are in green and 8A-14C at 40 to 120 nM are shown in red. It is clear that between weeks 6A and 7A there was an increase in the level of resistance to Dox.
Figure 4.7. Cytotoxicity Assays in the 14 Week Study: Scheme 3

Figure 4.7. Legend: Cytotoxicity results of the cells dosed in scheme 3 weeks 1-9A-05 (Figure 4.2.). The week number is shown in the legend on the right and the dose is indicated down the right. Weeks 1-5 at 20 nM are shown in blue, 6A and 7A at 40 nM are in green, the 2 resistant 8A and 9A populations are in red and the 9A-05 – 9A-05 grown in the absence of Dox are in purple. The cells that were grown in the absence of Dox appeared to maintain a slight resistance to Dox, although cells 9A-04 and 9A-05 were losing their resistance towards the drug.
The approximate IC$_{50}$ of the cells and the percentage of viable cells at 20 nM and 50 nM at each stage is summarised in Table 4.1. It is clear that there was a rise in % of viable cells at the 20 nM dose and the IC$_{50}$ of the cells from stage 7A, as the IC$_{50}$ rises to 90 nM, and the % of viable cells at 20 nM did not drop below 90 % after week 7A. The cells that became slightly resistant and then grown without drug maintained their slight resistance to Dox after 5 weeks growth in drug free media, although the cells were starting to revert to a Dox sensitive phenotype.

Table 4.1. IC$_{50}$ Values in the 14 Week Study

<table>
<thead>
<tr>
<th>Week</th>
<th>Scheme 1 % at 20 nM</th>
<th>% at 50 nM</th>
<th>IC$_{50}$</th>
<th>Scheme 2 % at 20 nM</th>
<th>% at 50 nM</th>
<th>IC$_{50}$</th>
<th>Scheme 3 % at 20 nM</th>
<th>% at 50 nM</th>
<th>IC$_{50}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>53.02</td>
<td>13.96</td>
<td>20</td>
<td>0</td>
<td>53.02</td>
<td>13.96</td>
<td>20</td>
<td>0</td>
<td>53.02</td>
</tr>
<tr>
<td>1</td>
<td>37.20</td>
<td>20.80</td>
<td>15</td>
<td>1</td>
<td>37.20</td>
<td>20.80</td>
<td>15</td>
<td>1</td>
<td>37.20</td>
</tr>
<tr>
<td>2</td>
<td>47.44</td>
<td>25.78</td>
<td>20</td>
<td>2</td>
<td>47.44</td>
<td>25.78</td>
<td>20</td>
<td>2</td>
<td>47.44</td>
</tr>
<tr>
<td>4</td>
<td>24.79</td>
<td>16.21</td>
<td>8</td>
<td>4</td>
<td>24.79</td>
<td>16.21</td>
<td>8</td>
<td>4</td>
<td>24.79</td>
</tr>
<tr>
<td>5</td>
<td>42.60</td>
<td>16.45</td>
<td>17.5</td>
<td>5</td>
<td>42.60</td>
<td>16.45</td>
<td>17.5</td>
<td>5</td>
<td>42.60</td>
</tr>
<tr>
<td>6</td>
<td>75.65</td>
<td>27.21</td>
<td>25</td>
<td>6A</td>
<td>73.58</td>
<td>32.99</td>
<td>28</td>
<td>6A</td>
<td>73.58</td>
</tr>
<tr>
<td>7</td>
<td>68.47</td>
<td>27.09</td>
<td>30</td>
<td>7A</td>
<td>93.06</td>
<td>71.91</td>
<td>90</td>
<td>7A</td>
<td>93.06</td>
</tr>
<tr>
<td>8</td>
<td>55.37</td>
<td>26.52</td>
<td>30</td>
<td>8A</td>
<td>96.80</td>
<td>71.28</td>
<td>100</td>
<td>8A</td>
<td>96.80</td>
</tr>
<tr>
<td>9</td>
<td>95.44</td>
<td>50.86</td>
<td>50</td>
<td>9A</td>
<td>90.54</td>
<td>95.49</td>
<td>300</td>
<td>9A</td>
<td>90.54</td>
</tr>
<tr>
<td>10</td>
<td>74.27</td>
<td>43.17</td>
<td>35</td>
<td>10B</td>
<td>98.45</td>
<td>80.60</td>
<td>100</td>
<td>9A-01</td>
<td>101.26</td>
</tr>
<tr>
<td>11</td>
<td>86.75</td>
<td>48.76</td>
<td>50</td>
<td>11B</td>
<td>98.45</td>
<td>87.62</td>
<td>200</td>
<td>9A-02</td>
<td>98.17</td>
</tr>
<tr>
<td>12</td>
<td>71.95</td>
<td>36.52</td>
<td>30</td>
<td>12B</td>
<td>101.93</td>
<td>83.80</td>
<td>250</td>
<td>9A-03</td>
<td>106.99</td>
</tr>
<tr>
<td>13</td>
<td>39.18</td>
<td>16.75</td>
<td>15</td>
<td>13C</td>
<td>-</td>
<td>63.34</td>
<td>110</td>
<td>9A-04</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>54.41</td>
<td>21.68</td>
<td>22.5</td>
<td>14C</td>
<td>-</td>
<td>64.20</td>
<td>85</td>
<td>9A-05</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.1.Legend: The table shows the IC$_{50}$ value and the % of viable cells at 20 nM and 50 nM for each of the weekly dose points. The 3 different dosing schemes from Figure 4.2. are labelled across the top.

It is clear in Table 4.1. that the cells in scheme 2 after week 6A gained substantial resistance with IC$_{50}$s of < 85 and at 20 nM there was little cell death. In scheme 3, the resistance in the cells grown without drug appeared to be lessoning, as the IC50 drops over time until it reached approximately 30 nM. There appeared to be a slight increase in the IC50 from stages 7-12 in scheme 1, which decrease again in subsequent changes.
4.2.4. Microarray Analysis from the 14 Week Doxorubicin Dosing Study

RNA extractions were performed at most stages for microarray analysis. For each case non exposed MCF7/wt cells that were grown continuously alongside the dosed cells were used as the reference comparison on the microarray. The log ratios were calculated as a measure of gene expression change and the data was clustered to identify possible co-regulated genes.

A number of discrete gene clusters were associated with the 14 week Dox dosing experiment (Schemes 1-3, Figure 4.2.). The first set of clusters contained genes overexpressed in all MCF7 Dox dosed cells whether resistant or not (Figure 4.8.). There are 3 clusters in this figure, cluster A is a set of genes overexpressed in all the Dox dosed cells compared to unexposed MCF7/wt. However although overexpressed in scheme 3 cells (9A-05), which were initially made resistant by a stepwise increase to Dox then grown for 5 weeks without Dox (Figure 4.2.), the genes appeared to be less expressed compared to those cells continuously maintained with Dox.

Cluster B in Figure 4.8. contains a number of genes that were overexpressed in all samples compared to non dosed MCF7 cells. Included in this cluster were 2 genes that have been spotted in duplicate on the microarray, TNSFR11B and AGRN both replicates clustered together, thus validating the microarray data.

The final cluster, C (Figure 4.8.), contains a small set of genes that were weakly overexpressed in all cell samples from all stages. Included in this cluster were several CDKN genes and the ATP1B1 gene, these genes were also overexpressed in MCF7 cells dosed with 20 nM in the 1 week time course (Section 4.2.3.), where overexpression of these genes was seen from 1 to 7 days in MCF7/wt cells continuously exposed to Dox.
Figure 4.8. Cluster of Genes Overexpressed in All Doxorubicin Treated Cells

Figure 4.8. Legend: Experiments were carried out in duplicate, Cy5 labelled RNA is the dosed MCF7 cells (red) and Cy3 is the unexposed reference MCF7/wt cells (green). The experiment details are across the top and gene names are listed down the right. The first number in the experiment details corresponds to the stage number in the dosing scheme (Figure 4.2.), and the latter number is the array chip number. The experiments from the 3 different schemes are labelled. The relative fold change is shown on the top left. There are 3 separate clusters:

Cluster A: Represents genes overexpressed in all experiments, although the genes appeared to be less expressed in the scheme 3 experiments: those which were made resistant and then grown in the absence of Dox (Figure 4.2.)

Cluster B: Contains genes that were overexpressed in all schemes

Cluster C: Contains genes that were weakly overexpressed in all schemes.
There was also a cluster of genes that was overexpressed in the scheme 1 cells: those dosed continuously for 14 weeks with 20 nM Dox, that were not overexpressed in the cells that had acquired low level resistance by a stepwise exposure to Dox (Figure 4.9.).

**Figure 4.9. Cluster of Genes Overexpressed in Cells Dosed Continuously with Doxorubicin but not in Cells with a Slight Resistant to Doxorubicin**

![Gene expression heatmap](image)

Figure 4.9. Legend: A cluster of genes, which were overexpressed in MCF7 cells continuously exposed to 20 nM Dox for 14 weeks (scheme 1), which were not overexpressed in the cells that had acquired partial resistance to Dox (schemes 2 and 3). The experiment details are across the top and gene names are listed down the right. The first number in the experiment details corresponds to the stage number in the dosing scheme (Figure 4.2.), and the latter number is the array chip number. The relative fold change is shown on the left.

The genes included within this cluster were overexpressed in cells exposed continuously to 20 nM Dox (Scheme 1), but once partial resistance to the drug was acquired the genes were no longer overexpressed (Scheme 2 and 3). Included within the cluster were 8 different genes. Several of these genes were also present in the clusters associated with overexpression in the 1 week time course experiment (Section 4.2.2.).

A final group of genes, which clustered into 2 different groups are those that were overexpressed in the MCF7 cells which had gained a partial resistance to Dox compared to those that remained sensitive (Figure 4.10.).
Figure 4.10. Genes Overexpressed in Cells with Slight Resistance to Doxorubicin

Figure 4.10. Legend: Clusters of genes associated with the MCF7 cells that had acquired resistance to Dox after a stepwise increase in exposure to the drug. The experiment details are across the top and gene names are listed down the right. The first number in the experiment details corresponds to the stage number in the dosing scheme (Figure 4.2.), and the latter number is the array chip number. The relative fold change is shown on the left. There were 2 separate clusters:

Cluster A: Represents genes that were highly overexpressed in cells selected for resistance by a stepwise Dox exposure. These genes were also partially overexpressed in cells that were continuously exposed to 20 nM and that did not acquire resistance to Dox in scheme 1 (Figure 4.2.)

Cluster B: Contains a cluster of genes that were overexpressed only in the cells that have acquired Dox resistance.
There are 2 different clusters of genes in Figure 4.10. Cluster A contains a group of genes that were overexpressed in the MCF7 cells that had been selected for resistance by a stepwise exposure to Dox. The genes within this cluster also appeared to be partially overexpressed in the cells with no resistance to Dox that were continuously exposed at 20 nM for 14 weeks (scheme 1). The genes associated with cluster B were only overexpressed in the resistant cells. All genes in these clusters were also overexpressed in the 9A-05 cells, which had gained a slight resistant to Dox and were subsequently grown in the absence of Dox.

Included in cluster B was the $ABCB1$ gene, which encodes a gene known to be involved in the MDR phenotype, P-glycoprotein. The other genes in this cluster included $BTN3A3$, $ERBB2$, $TLE2$, $ARP3$, $SRP9$ and $TACSTD2$.

4.2.5. Gene Amplification and Resistance Development

At several later time points after resistance development, DNA was also extracted and analysed using CGH microarray. The time points analysed were 14 (scheme 1), 11B and 14C (scheme 2) and 9A-05 (scheme 3). Over this period no significant DNA amplification events had occurred in the cells dosed with Dox.
4.3. Discussion
Initially experiments within this chapter were to be carried out on both the sensitive MCF7 and MDA468 cell lines. However, the selection of cells for resistance by Dox exposure was unsuccessful in the MDA468 cells, consequently, only MCF7 cells were studied. The MCF7 cells are known to be able to acquire resistance to Dox, since they are the progenitor cell line of the resistant HPGP cell line, which was selected from the MCF7 cells by drug exposure. Similarly, the resistant MDA16 cell line was derived from the sensitive MDA468 cells, however, these cells were selected for resistance by an antibody that recognizes P-glycoprotein, and thus are thought to have an intrinsic resistance to Dox. Therefore, this may explain why the MCF7 cells readily developed resistance to Dox drug exposure and not the MDA468 cells. An alternative approach to select for resistance is to firstly mutate cells prior to selection with drug to elevate the mutation frequency to a detectable level, then select for resistance with drug, but this method has complications, as the mutagen is affecting the DNA of the cells indiscriminately and secondary effects may occur (Thompson, 1987; Shen et al., 1986).

4.3.1. Gene Changes During 7 Day Time Course Exposure to 20 nM Doxorubicin
MCF7 cells were exposed to 20 nM Dox and RNA was isolated at several time points over a 1 week period and analysed using cDNA microarray. MCF7 cells grown without Dox alongside the dosed cells were used as the reference population. After 2 hr and 6 hr Dox exposure there were no significant changes detected in gene expression, whereas at later time points several genes were overexpressed in the dosed cells. Some of these genes were first expressed at days 4 and 5 and not expressed at day 6 or 7 (Figure 4.4.), while others were expressed after 1 day Dox exposure and this expression was maintained until day 7 (Figure 4.3 and 4.4).

4.3.1.1. Genes Overexpressed After 1 Day Dox Exposure
Several clusters of genes were overexpressed in the dosed cells after 1 day exposure, and these genes maintained their overexpression after 7 days exposure to Dox (Figure 4.3). These genes fell into several clusters, the first (Figure 4.3.A) included a number of CDKN genes, CDKN1A, CDKN2A and CDKN2D as well as the ATP1B1 gene. CDKN1A (p21) is also overexpressed in radioresistant MCF7 cells following irradiation (Li et al., 2001). Li et al. (2001) found that after 12 subsequent passages the level of CDKN1A and resistance dropped so it was proposed that CDKN1A may play a role in resistance, however transfection of the resistant cells with the gene did not alter survival when irradiated and so CDKN1A is not thought to contribute to the resistance (Li et al., 2001). What’s more CDKN1A and DNA
repair proteins rapidly accumulate at sites of DNA damage induced by heavy ions or irradiation and thus are thought to have a role in detecting or processing DNA lesions (Jakob et al., 2002; Jakob et al., 2000), therefore overexpression of CDKN1A here is likely to be a response to the DNA damage effects of Dox. Irradiation of cells has also been shown to induce CDKN2A as well as CDKN1A (Suzuki et al., 2001), a gene that also overexpressed in MCF7 cells exposed to Dox in this study.

The other clusters of genes overexpressed during days 1 to 7 of Dox exposure contained 14 genes (Figure 4.3.B). CD9 was overexpressed in these clusters and CD9 was also overexpressed in Chapter 3 in all breast cancer cell lines when compared to HBL100, thus exposure of MCF7/wt cells to Dox is increasing this expression. A recent study has shown the overexpression of CD9 in Dox resistant cell lines (Watts et al., 2001). Expression of CD9 has an inverse correlation with metastases in breast cancer (Miyake et al., 1995) and reduced CD9 expression is associated with poor prognosis in lung cancer (Higashiyama et al., 1995), thus overexpression of this gene in response to Dox is an indicator of the positive effect of Dox treatment on cancers. Another possible indicator of the beneficial effects of Dox in tumour cell treatment was the overexpression of the TNFRSF6 gene after 1 day exposure, as TNFRSF6 encodes the FAS protein, which is a cell-surface receptor involved in apoptosis initiation (Theuns et al., 2001). Thus cells exposed to Dox may be predisposed to apoptosis.

A gene involved in drug metabolism, the glutathione-S-transferase GSTM3 was also in this cluster and was slightly increased in expression after 2 days exposure to Dox and this overexpression was maintained after 7 days continuous exposure to Dox (Figure 4.3.B). The glutathione transferases are a superfamily of enzymes involved in the metabolism of reactive electrophilic intermediates to less reactive more water-soluble glutathione conjugates, which are more readily extruded from the body (Hayes et al., 1995). They are thought to contribute to drug resistance in chemotherapy, and glutathione overexpression in cells exposed to Dox here may be a cellular response, attempting to increase Dox clearance and thus may contribute to later resistance development.

Several other genes within the cluster in Figure 4.3.B. have previously been seen overexpressed in various tumours. The FN1 gene (fibronectin 1) located within the cluster was highly expressed after 5 days (< 5 fold), this gene can be overexpressed in papillary thyroid carcinoma (Huang et al., 2001). S100A6 and S100A11 are members of the S100 family, which are known to be overexpressed in some tumour cells (Otterbein et al., 2002).
The *S100A6* gene has been found overexpressed in leukemia cells (Otterbein et al., 2002) and is also overexpressed cholangiocarcinoma where it is thought to have a possible role in the differentiation of cholangiocarcinoma from hepatocellular carcinoma (Kim et al., 2002). Although the reasons why and how these genes are induced in cells exposed to Dox is unclear.

**4.3.1.2. Genes Overexpressed After 4 Days Dox Exposure**

There were several clusters of genes that were overexpressed after 4 days exposure to Dox (Figure 4.4.). The first gene cluster contained genes whose expression returned to normal after 5 days continuous exposure to Dox (Figure 4.4.A), suggesting they are not involved in the resistant phenotype but are a late temporary cellular response to the drug. The second cluster also contained late expressed genes as they were overexpressed in cells continuously exposed to 20 nM Dox from 4 days through to 7 days (Figure 4.4.B).

**4.3.2. Resistance Acquisition by MCF7 Cells**

After a stepwise increase in Dox concentration in the 14 week study, the MCF7/wt cells acquired a drug resistant phenotype (Schemes 2 and 3 in Figure 4.6 and 4.7. respectively). In contrast cells continuously exposed for 14 weeks with 20 nM Dox did not appear to alter their relative sensitivity to Dox (Scheme 1 in Figure 4.5.) and the cell populations all had IC50s between 10 and 50 nM. The resistant phenotype acquired by the MCF7 cells appeared to be relatively stable for a number of weeks, although after culture for 5 weeks free of Dox, the level of resistance had decreased.

**4.3.3. Gene Changes in All Cell Groups After 14 Weeks Exposure to Doxorubicin**

There were 3 gene clusters that were overexpressed in all MCF7 cells exposed to Dox in the different schemes (Figure 4.2.) regardless of sensitivity to the drug (Figure 4.8.). The first cluster although overexpressed in all cells appeared to be less expressed in the resistant cells that had not been treated for Dox for 5 weeks (9A-05). This decrease in expression may thus be associated with the decrease in resistance seen over time in these cells or the absence of exposure of the cells to Dox. The first cluster contained 10 different genes, 4 are closely related *IFITM1, IFITM2, IFITM3* and *IFI27*, all of which are interferon (IFN) inducible proteins. *IFNs* have several roles in cellular antiviral and antiproliferative effects, by working through several pathways to activate target genes (Darnell et al., 1994) and their effect in the Dox dosed cells is discussed in further in the next section (Section 4.3.4.).
PORI was also within this cluster and was highly overexpressed in all the dosed cells over 14 weeks (<5 fold). PORI is thought to be involved in cytoskeleton rearrangement at the cell periphery induced by ARF6 and Rac1, which may be involved in endocytosis, proliferation and cell motility (D’Souza-Schorey et al., 1997). When MCF7/wt cells were treated with Dox they underwent a slight morphological change and become slightly elongated, which may have been a consequence of the induction of the PORI gene.

The other 2 gene clusters in Figure 4.8., which were overexpressed in all cell populations exposed to Dox contained several genes that also clustered together in the 7 day time course study in cells continuously exposed to Dox (Figure 4.3.A.), discussed in Section 4.3.1. These included the CDKN genes and ATP1B1 all of which clustered together again, suggesting they may be activated similarly. Likewise AGRN, BMP7, SLBP, CD9 and S100A11 which were overexpressed in the 7 day time course were also overexpressed in all MCF7 cells, both resistant and sensitive, when exposed to Dox over 14 weeks, and again clustered together.

As the above genes were overexpressed in cells exposed to Dox regardless of dose concentration or sensitivity of the cells to the drug, they may merely be a response of the cells to the drug and may not be associated with the resistant phenotype.

4.3.4. Genes Changes in Cells Prior to Doxorubicin Resistance Development

There was an interesting cluster of 8 genes that were overexpressed in all MCF7/wt cell populations exposed continuously to 20 nM Dox, although after resistance development these genes were no longer overexpressed (Figure 4.9.). The decreases in expression may be associated with the development of a slight resistance phenotype or may be a result of higher Dox concentrations used in the stepwise increase scheme.

S100A2, a Ca²⁺ binding protein and candidate tumour suppressor gene was present in this gene cluster. Expression of S100A2 is down regulated in tumours compared to normal cells (Nagy et al., 2002; Liu et al., 2000), and cell growth and proliferation is halted in hepatocellular carcinoma cells transfected with S100A2 (Wang et al., 2001). It has also been suggested that the loss of S100A2 expression may be associated with poor survival outcome (Kyriazanos et al., 2002) and with the development of malignant tumours, and is associated with late tumour development (Liu et al., 2000). The mechanism of S100A2 protein down regulation in tumour cells is thought to be hypermethylation of the promoter region (Wicki et al., 1997). Therefore as S100A2 was overexpressed in sensitive Dox treated cells compared to
untreated MCF7/wt breast cancer cells, it seems that Dox maybe able to cause demethylation
of the promotor region resulting in expression of \textit{S100A2} in tumour cells, although once cells
have gained the resistant phenotype this possible action of Dox is prevented. This expression
of \textit{S100A2} could be another mechanism by which Dox exerts its chemotherapeutic effect.
This is not a new concept in the treatment of cancer, as clinical trials are currently ongoing
with a drug called 5-azadeoxcytidine, which is also thought to result in the induction of
methylation-regulated tumour suppressive pathways (Kanai et al., 2001). Also two genes,
\textit{CDKN2A} and \textit{CDKN2D}, the first of which is methylated in human lung cancer cells are
induced by 5-azadeoxcytidine (Zhu et al., 2001) and both these genes were overexpressed in
this study upon Dox exposure.

The \textit{MEIS1} gene encoding a mouse homolog homeodomain protein was also within this
cluster in Figure 4.9 and so is overexpressed in cells treated with Dox, but once a resistant
phenotype developed the genes are no longer overexpressed. The \textit{MEIS1} gene is involved in
development and is known to interact with several \textit{HOX} proteins, including \textit{HOXA9} in the
induction of acute myeloid leukaemia (Thorsteindottir et al., 2001). There is evidence to
suggest that \textit{HOX} genes and \textit{MEIS1} are also involved in hematopoiesis, the generation of a
large spectrum of stem cells with multi differentiation potential, and \textit{MEIS1} and some \textit{HOX}
genes show similar expression profiles whereby they are preferentially expressed in
hematopoietic stem cells and down regulated following differentiation and maturation
(Pineault et al., 2002). There are a number of \textit{HOX} genes, including \textit{HOXA9} on the Leicester
microarray, none of which were overexpressed during this study, suggesting a role for \textit{MEIS1}
independent of \textit{HOX} genes.

Another gene in this cluster (Figure 4.9.), which was overexpressed in sensitive cells treated
with Dox, but not in resistant cells exposed to Dox, is \textit{IFI41} (interferon inducible protein 41),
also known as \textit{Sp110}. \textit{Sp110} is known to localise to nuclear bodies, cellular structures whose
exact function is unknown, but are thought to have roles in cancer, apoptosis and viral
infections (Regad and Chelbi-Alix, 2001). The level of \textit{Sp110} mRNA increases in \textit{NB4} cells
treated with all-trans retinoic acid (Bloch et al., 2000), although this is not surprising as all-
trans retinoic acid is known to induce a high level of interferons (\textit{IFNs}) in \textit{NB4} cells, which in
turn is known to increase the number of nuclear bodies (Regad and Chelbi-Alix, 2000). Two
other genes in the cluster, which were overexpressed in the Dox treated sensitive cells, but not
the resistance cells were \textit{STAT1} (Signal transducer and activators of transcription 1) and
\textit{IFIT1}, and both are also induced by \textit{IFNs}. \textit{STAT1} is activated by phosphorylation when \textit{IFNs}
bind to their cognate cell surface receptors, which activates different signal cascades known as the Jak/STAT pathways (Regad and Chelbi-Alix, 2000). Phosphorylated STAT1 can form homodimers or heterodimers with STAT2 and these dimers lead to the transcription of a set of target genes (Regad and Chelbi-Alix, 2000). Therefore IFNs may play a role in the sensitive MCF7/wt cell response to Dox, but not in the resistant population, as 3 genes induced by IFNs, STAT1, IFI41 and IFIT1 were overexpressed in the sensitive cells after Dox exposure. However, against this hypothesis is the overexpression in this study of several IFN genes in both the sensitive and resistant cells (discussed in section 4.3.3.).

Interestingly the IFN regulated pathway involving STAT1 and interferon inducible genes has been implicated in cells undergoing senescence (Perou et al., 1999). Senescence is a physiological process that limits cell proliferation and is accompanied by morphological change (enlarged and flattened shape) and senescence has been induced in cells exposed to anticancer agents (Chang et al., 1999). The MCF7 cells when exposed to Dox appeared to enter senescence as cells underwent a morphological change and initial growth arrest prior to resistance development. When MCF7 cells gained partial resistance to Dox, STAT1, IFI41 and IFIT1 were no longer expressed and cells no longer went into senescence after Dox exposure, and thus the lack of expression of these genes may have contributed to the development of resistance.

Further evidence of IFN involvement in resistance is that IFN-gamma can decrease levels of a gene in the same family as ABCB1, the ATP-binding cassette transporter 1 (ABCI) whose protein level is also decreased by 3-4 fold in foam cells (Panousis and Zuckerman, 2000). In STAT1 knockout mice IFN-gamma treatment does not cause a decrease in ABC1 gene expression (Wang et al., 2002), indicating that STAT1 signalling is necessary for the IFN effects on ABC1 gene expression. Thus although STAT1 appeared to be involved in the early cellular response to Dox, it is no surprise to see STAT1 expression decreased in the resistant cells.

Insulin-like growth factor 1 (IGF1) was also overexpressed in sensitive but not resistant Dox treated cells. IGF1 is a mediator of the effects of growth hormone and can promote cell proliferation, differentiation and transformation (Furstenberger and Senn, 2002). IGF1 may also play a role in the inhibition of apoptosis (Hurbin et al., 2002), and the IGF1 receptor is overexpressed in many tumour cell lines (Werner et al., 1996). Thus the sensitive cells dosed with Dox may be attempting to overcome the effects of Dox by overexpression of IGF1 to
increase cell proliferation and decrease apoptosis. In concordance with this hypothesis in the cells resistant to Dox,  \textit{IGF1} was not overexpressed as Dox is having little effect on cell survival.

Therefore, within the cluster of genes overexpressed in sensitive cells and not resistant cells exposed to Dox were a number of \textit{IFN} involved genes, \textit{IFI41} (\textit{Sp110}), \textit{STAT1} and \textit{IFIT1} suggesting a role of \textit{IFN} in cells exposed to Dox. \textit{IGF1}, which is involved in the promotion of cell proliferation and the inhibition of apoptosis is also overexpressed in the sensitive cells and may be a mechanism by which cells are trying to combat the effects of Dox. \textit{SI00A2}, a candidate tumour suppressor gene was overexpressed in sensitive cells treated with Dox, and therefore its activation may be another mechanism by which Dox exerts its chemotherapeutic effect. Alternatively the genes within this cluster may not be overexpressed in the partially resistant cells as a downstream mechanism may prevent their overexpression in the cells exposed to Dox, which could in turn lead to a resistant phenotype.

4.3.5. Gene Changes Associated with Doxorubicin Resistance Development

Only a few genes were found to be overexpressed in the cells made partially resistant to Dox (Figure 4.10.), which is in agreement with a study of Shen et al., (1985) that analysed protein changes in cells selected for resistance to different drugs, including Dox, vinblastine, mezerein and colchicine. In this study it was found that only a limited number of protein changes occurred in the development of the MDR phenotype (Shen et al., 1985). Evidence against the involvement of these genes in the development of a resistant phenotype was that all the genes were overexpressed in the 9A-05 population that had been grown without Dox for 5 weeks and appeared to be reverting back to a Dox sensitive phenotype.

Figure 4.10 represents the 2 gene clusters overexpressed in the cells that had developed a resistance to Dox (Scheme 2 and 3) but not in the cells that remained sensitive after Dox exposure (Scheme 1). Whether these genes were overexpressed due to the resistant phenotype or the increase in Dox concentration is unclear. However the cells that had gained a slight resistance to Dox, which were subsequently maintained for 5 weeks in the absence of Dox also expressed these genes, indicating these changes are related to the drug resistant phenotype and not concentration of Dox exposure.

The first cluster contains 9 genes that were highly overexpressed (2 - <5 fold) in the resistant cell populations and only weakly overexpressed in the sensitive population compared to
MCF7/wt cells with no drug exposure (Figure 4.10.A.). While the genes in the second cluster are associated with the Dox resistant phenotype as they were overexpressed in resistant cells exclusively (Figure 4.10.B.). This cluster contains 7 different genes, including the ABCB1 gene. The ABCB1 gene is located several times on the microarray, all of which were present in this cluster, validating the microarray data. As discussed in the previous chapter the ABCB1 gene encodes the 170 kDa P-glycoprotein known to be involved in the multi drug resistant phenotype (Seelig, 1998), therefore it is not surprising to see the genes overexpressed. In later stages of the experiment (Scheme 2: 14C) this gene was very highly overexpressed with fold changes of greater than 90 occurring.

BTN3A3 was another gene in this cluster associated specifically with the resistant cells (Figure 4.10.B.). It is a member of the human butyrophillin family (BTN) and shares some features of the ABCB1 gene product. BTN are one of the major proteins of the milk fat globule membrane, it is a type I membrane glycoprotein expressed on the apical surface of secretory cells in the lactating mammary gland (Cavaletto et al., 2002). BTN consists of 2 extracellular immunoglobulin domains, a transmembrane domain and a large intracellular domain containing a B30.2 domain or ret finger protein of 170 amino acids (Taylor et al., 1996). It has been suggested that the domain structure of BTN may have some unknown receptor function (Taylor et al., 1996) and thus it may be playing a role in the development of resistance.

The ERBB2 oncogene was present in the cluster associated with resistant MCF7 cells (Figure 4.10.B.). ERBB2 encodes a 185 kDa protein and is a member of the EGFR family. ERBB2 is overexpressed in approximately 10-34% of breast cancers (Revillion et al., 1998) and its overexpression is associated with a poorer response to chemotherapy (Hengstler et al., 1999). Overexpression of both ERBB2 and the Ep-Cam protein, which is encoded by the GA733-2 or TACSTD2 (also overexpressed in this gene cluster) complement each other for predictors of poor prognosis in patients with invasive cancer (Spizzo et al., 2002). The protein product of TACSTD2 is a cell surface protein widely expressed in human cancers and at specific stages of epithelial cell differentiation (Mangino et al., 2002). Both ERBB2 and TACSTD2 are targets for treatment with monoclonal antibodies in cancer (Schwartzberg, 2001; Pietras et al., 1998), and their overexpression in cells that had developed a slight resistance to Dox suggests a therapeutic potential.
*TLE2*, a member of a family of genes that are the human homologues of Drosophila groucho, was also overexpressed in the resistant MCF7 cells (Schemes 2 and 3), but not the sensitive cells (Figure 4.10.B.). The human homologues of Drosophila groucho are members of the NOTCH signalling pathway (Liu et al., 1996) and the Drosophila groucho protein, a widely expressed nuclear protein, is a transcriptional corepressor that lacks DNA binding activity, but can act together with other proteins to regulate embryological developmental processes in the fly (Dang et al., 2001). The possible roles and functions of *TLE2* in these resistant cells or in the development of resistance is unclear.

The *ARPC3* gene was also overexpressed in cells that have acquired some resistance to Dox (Figure 4.10.B.). The *ARPC3* protein is part of an assembly of 7 proteins that make up the *Arp2/3* complex, which plays a role in the control of actin polymerisation in eukaryotic cells (Robinson et al., 2001). Actin polymerisation is involved in the protrusion of cell membrane, which is necessary in cell shape change and locomotion (Welch et al., 1997). Interestingly overexpression of the *ARPC3* gene can be explained in this study because when the cells were exposed to Dox their morphology changed slightly from a round to slightly elongated shape similar to the resistant HPGP cells that were derived from the MCF7 cell line. The gene was not overexpressed in the HPGP resistant cells (Chapter 3), thus it is not associated with the end resistant phenotype, but could be involved in the acquisition of resistance as the cell undergoes a change in morphology.

The final gene in Figure 4.10.B. associated with the MCF7 cells that exhibit some resistance was a signal recognition particle (*SRP9*). *SRP9* is a subunit of the SRP complex, which is a ribonucleoprotein particle involved in targeting of secretory and membrane proteins to the rough endoplasmic reticulum (Mason et al., 2000), it recognises the signal sequence of the nascent polypeptide chain emerging from the ribosome and targets the ribosome-nascent chain-SRP complex to the rough endoplasmic reticulum. Thus it could be overexpressed in cells acquiring resistance as it may have a role in aiding transport of the increased levels of the *ABCB1* gene product P-glycoprotein to the rough endoplasmic reticulum prior to transport to the cell membrane. However other members of the SRP complex, including *SRP54* and *SRP72* are present on the microarray but were not overexpressed.

The genes in this section were all overexpressed in cells that had acquired a resistant phenotype (Scheme 2 and 3) but not in cells, which remained sensitive (Scheme 1) when exposed to Dox (Figure 4.10.B.). It is possible that several of these genes associated with an
acquisition of Dox resistance are expressed as part of common pathways. In some ovarian tumours, *STAT1*, *Ep-CAM* (*TACSTD2*) and *S100A2* (Hough et al., 2001) are overexpressed, and it has been suggested that due to similar expression patterns of *STAT1* and *Ep-CAM* that they may be targets or part of common signalling pathways (Hough et al., 2001). As mentioned earlier *STAT1* is induced by *IFNs*, thus *Ep-CAM* could also be involved in this pathway, although *STAT1* was overexpressed in sensitive cells while *Ep-CAM* was overexpressed in cells that acquired some resistance to Dox.

Glyoxlase 1 (*GLO1*), an enzyme that detoxifies methylglyoxal, has previously been found overexpressed in resistant cells when treated with Dox (Sakamoto et al., 2000), however the expression of *GLO1* was unchanged in the MCF7 cells that exhibited partial resistance in this study (Scheme 2 and 3). However *GLO1* was associated with resistance that was not mediated by the *ABCB1* gene, and thus may be mediated by a different mechanism of resistance. Other genes previously associated with resistance as seen in this study (Chapter 3) including *OXTR*, *LSP1*, *TITF1* etc. were not overexpressed here. Therefore these genes may be late events in resistance development, and may not play a role in cells that are partially or becoming resistant to Dox.

### 4.3.6. Gene Amplification and Doxorubicin Resistance Development

Over the 14 week study there were no changes in gene amplification. The *ABCB1* gene is sometimes amplified and overexpressed in resistant cells, and the gene was shown to be amplified in the highly resistant HPGP cells, which are derived from the MCF7 cells. As no amplification events were seen here it suggests that amplification of the *ABCB1* gene is a late event and is not involved in the initial development of drug resistance. This data therefore confirm the earlier study of Shen et al., (1986).

### 4.3.7. Summary

The chapter has elucidated several genes that are induced by cell exposure to Dox over a 7 day period, and many of these genes continue to be expressed over a total of 14 weeks exposure to Dox. Some of these genes may be responding to DNA damage or are early response genes such as *CDKN1A*, while others are associated with the *IFNs*. There are many genes represented in Figure 4.8. that were overexpressed in both cells that had acquired a partial resistance to Dox (Scheme 2 and 3) and in cells that remained sensitive (Scheme 1), consequently these genes are thought to be part of the cellular response to Dox and may not play a role in the development of a resistant phenotype.
There were several genes associated with the Dox treatment in the cells selected for resistance to Dox when compared to those that remained sensitive (Figure 4.10.). Amongst these was the $ABCB1$ gene that encodes P-glycoprotein known to be involved in an MDR phenotype and 2 genes whose overexpression are predictors of poor prognosis in patients with invasive cancer (Spizzo et al., 2002), $ERBB2$ and $TACSTD2$. As well as genes overexpressed in resistant cells, there was a population that were only induced in the sensitive cells when exposed to Dox. These included several genes associated with $IFNs$, and a potential tumour suppressor $S100A2$.

This chapter has elucidated some possible mechanisms of Dox action or cell response to Dox exposure. And it suggested some likely candidates of genes associated with the initial stages of development of a Dox resistant phenotype in MCF7 cells selected with drug. The $ABCB1$ gene is known to be a major figure in this process, but it is clear there are other contributing factors in the early stages of resistance development.
Chapter 5

Genomic Microarray Analysis of Breast Tumour Cell Lines
5.1. Introduction

Cancer is a multistage process thought to require the accumulation of at least 6 individual mutations in a single cell for transformation to occur (Vogelstein and Kinzler, 1993). Studies using a variety of techniques have been carried out on tumour samples to try and find associations between specific aberrations and the clinical characteristics of cancer. Many of these studies have analysed different aspects of breast cancer including DCIS (Buerger et al., 1999; James et al., 1997), high grade DCIS (Moore et al., 1999), LCIS (Nishizaki et al., 1997), primary breast tumours (Larramendy et al., 2000; Tirkkonen et al., 1998), breast cell lines (Forozan et al., 2000; Larramendy et al., 2000), lymph node negative tumours (Hermsen et al., 1998; Isola et al., 1995), metastasising breast tumours (Adeyinka et al., 1999; Kuukasjarvi et al., 1997; Nishizaki et al., 1997), aneuploidy in breast cancer (Ried et al., 1999; Tanner et al., 1998) and gene alterations involved in breast cancer progression (Roylance et al., 1999).

In this chapter, 7 human breast cancer cell lines, described in the Materials and Methods (Section 2.2.3.1.), have been analysed with microarray GCH to detect areas of amplification. Each cell line has been compared against the putative normal HBL100 breast cell line. Table 5.1. summarises data from CGH studies in the literature regarding chromosomal gains and losses previously detected in human breast cell lines and primary tumours. In general it was found that the mean number of aberrations detected in cell lines (19 and 16.3) appears to be greater than in primary tumours (7.7). This finding is also confirmed in other studies (Larramendy et al., 2000; Kallioniemi et al., 1994). It has been suggested that the greater number of gains or losses detected illustrates that cell lines are more genetically advanced (Kallioniemi et al., 1994). However cell lines are a good resource for the detection of chromosomal losses and gains in breast tumours, since most aberrations detected in primary cells are present in their corresponding cell line (Larramendy et al., 2000), and the highly prevalent changes are frequently found in both primary tumours and cell lines (Kallioniemi et al., 1994).

5.1.1. Genomic Alterations in Breast Cancer Cell Lines

The studies in Table 5.1., which analysed genomic gains and losses in breast cell lines by CGH both studied different cell lines (Forozan et al., 2000; Larramendy et al., 2000). The regions of frequent gain and loss are shown and the studies are in agreement for the most common regions of chromosomal gain found at: 8q, 1q, 20q, 3q, 5p, 17q and 1p and chromosomal loss: 8p, 18q, Xp, Xq and 4p.
Table 5.1. Chromosome Changes in Breast Cell Lines and Primary Tumours

### BREAST CANCER CELL LINES

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>8q</td>
<td>28</td>
<td>75</td>
<td>8p</td>
<td>22</td>
<td>58</td>
</tr>
<tr>
<td>1q</td>
<td>23</td>
<td>61</td>
<td>18q</td>
<td>22</td>
<td>58</td>
</tr>
<tr>
<td>20q</td>
<td>21</td>
<td>55</td>
<td>1p</td>
<td>16</td>
<td>42</td>
</tr>
<tr>
<td>7p</td>
<td>17</td>
<td>44</td>
<td>Xp</td>
<td>16</td>
<td>42</td>
</tr>
<tr>
<td>3q</td>
<td>15</td>
<td>39</td>
<td>Xq</td>
<td>16</td>
<td>42</td>
</tr>
<tr>
<td>5p</td>
<td>15</td>
<td>39</td>
<td>4p</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>7q</td>
<td>15</td>
<td>39</td>
<td>11q</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>17q</td>
<td>12</td>
<td>33</td>
<td>18p</td>
<td>12</td>
<td>33</td>
</tr>
<tr>
<td>1p</td>
<td>11</td>
<td>30</td>
<td>10q</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>20p</td>
<td>11</td>
<td>30</td>
<td>19p</td>
<td>11</td>
<td>28</td>
</tr>
</tbody>
</table>

Mean number aberrations 19

- Forozan et al., 2000
  - n = 38

### PRIMARY BREAST TUMOURS

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>8q22-qter</td>
<td>18</td>
<td>100</td>
<td>Xq11-q12</td>
<td>10</td>
<td>56</td>
</tr>
<tr>
<td>1q31-q32</td>
<td>16</td>
<td>89</td>
<td>Xp11.2-p1</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td>20q12-q13.2</td>
<td>15</td>
<td>83</td>
<td>13q21</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td>8q13</td>
<td>13</td>
<td>72</td>
<td>8p12-p1</td>
<td>8</td>
<td>44</td>
</tr>
<tr>
<td>3q26.1-qter</td>
<td>12</td>
<td>67</td>
<td>4p13-p14</td>
<td>7</td>
<td>39</td>
</tr>
<tr>
<td>17q21-qter</td>
<td>12</td>
<td>67</td>
<td>6q15-q22</td>
<td>7</td>
<td>39</td>
</tr>
<tr>
<td>5p14</td>
<td>11</td>
<td>61</td>
<td>18q11.2-qter</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>6p22</td>
<td>10</td>
<td>56</td>
<td>3p12-p14</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>22qter-qter</td>
<td>9</td>
<td>50</td>
<td>4q23-q24</td>
<td>5</td>
<td>28</td>
</tr>
<tr>
<td>1p31-pter</td>
<td>8</td>
<td>44</td>
<td>4q26-q28</td>
<td>5</td>
<td>28</td>
</tr>
</tbody>
</table>

Mean number aberrations 16.3

- Larramendy et al., 2000
  - n = 18

### PRIMARY BREAST TUMOURS

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q</td>
<td>37</td>
<td>67</td>
<td>16q</td>
<td>21</td>
<td>29</td>
</tr>
<tr>
<td>8q</td>
<td>27</td>
<td>49</td>
<td>8p</td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td>16p</td>
<td>21</td>
<td>38</td>
<td>13q</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>5p</td>
<td>13</td>
<td>24</td>
<td>17q</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>19q</td>
<td>11</td>
<td>20</td>
<td>9p</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>11q</td>
<td>10</td>
<td>18</td>
<td>Xq</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>17q</td>
<td>10</td>
<td>18</td>
<td>6q</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>19p</td>
<td>10</td>
<td>18</td>
<td>11q</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>20q</td>
<td>10</td>
<td>18</td>
<td>18q</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>3q</td>
<td>9</td>
<td>16</td>
<td>1p</td>
<td>9</td>
<td>16</td>
</tr>
</tbody>
</table>

Mean number aberrations 7.7

- Tirkkonen et al., 1998
  - n = 55

Table 5.1. Legend: A summary of data from 3 CGH studies, 2 of which were carried out on breast cancer cell lines and 1 on primary breast tumours. The 10 most frequent regions of amplification and losses detected in the studies and the number and percentage of tumours samples with the aberration are shown. The reference, total number of tumour samples analysed and the mean number of aberrations detected for each study is on the right.
Also included in Table 5.1 are data from a study detecting chromosomal gains and losses in primary tumours. Interestingly 6 of the most common chromosome gains and 6 of the most frequent losses detected in the cell lines are also present in the primary cells. Data such as this substantially validates the use of cell lines to analyse frequent regions of chromosome alteration in breast cancer.

5.1.2. Specific Genes Associated with Genomic Alterations

Some of the regions of chromosomal aberration detected are associated with specific genes thought to be involved in cancer. Several well established oncogenes or tumour suppressor genes have been mapped to regions of high aberration. Chromosome 8q24, which is amplified in some breast tumours is associated with lymph node involvement, advanced stage and an increased rate of relapse (Borg et al., 1992; Borg et al., 1991), contains the *myc* oncogene. *ERBB2*, an oncogene and an indicator of poor prognosis is overexpressed in approximately 20% of breast cancers (Revillion et al., 1998) and is located at chromosome 17q21.2. Chromosome 17 and its involvement in breast cancer is discussed in greater detail in chapter 7. The breast cancer susceptibility genes *BRCA1* and *BRCA2*, located at chromosome 17q21 and 13q12 respectively and the tumour suppressor gene p53 located at 17p13 are frequently mutated in breast cancer.

Based on their location in regions of frequent chromosomal alterations, a number of cancer-related genes have also been detected. Amplification at 20q13 occurs in a variety of tumour types and is associated with aggressive tumour behaviour (Collins et al., 1998). 20q13 contains a number of genes that have been amplified including the steroid co activator *AIB1* and the zinc finger transcription factor *ZNF217*, a proposed oncogene that is associated with increased metastasis (Hadika et al., 2000). In multidrug resistance, amplification occurs of 2 genes located on the long arm of chromosome 7, the *ABCB1* gene, which contributes to the MDR phenotype, and the neighbouring *ABCB4* (Chapter 3; Shen et al., 1986).

In contrast some of the detected regions containing genomic alterations are not currently associated with known genes that play a role in carcinogenesis. Some of the amplified genes such as *ERBB2* and *ABCB1* are also overexpressed and so are thought to play a direct role in tumour development or phenotype. Other genes amplified or deleted in areas of frequent aberration are not thought to contribute to the tumour phenotype, as they are not overexpressed. These gene amplifications or losses could be an incidental event or may be
occurring as a result of neighbouring genes, such as the \textit{ABCB4} gene which is not overexpressed in some resistant tumours but is co-amplified with the \textit{ABCB1} gene.

\textbf{5.1.3. Chapter objectives}

CGH has been used extensively to study regions of possible chromosomal amplification or deletion, however CGH does not have the ability to identify specific genes involved or elucidate small regions of amplification (Takeo et al., 2001). Therefore in this study I have used cDNA microarray to analyse specific gene amplification or deletion events in several breast cancer cell lines. Cell lines are a model as described above that is frequently used for the detection of aberrations in breast cancer. It is hoped the microarray CGH findings will elucidate chromosomal areas of high aberration activity and will identify single genes or specific regions involved in carcinogenesis. The relative RNA expression of the genes in these areas of chromosomal aberration is also elucidated, thus genes that are differentially expressed and so contributing to the cancer may be detected. The data acquired here is to be compared with that from clinical tumour samples (Chapter 6), in order to relate the amplifications occurring to disease state.
5.2. Results

All of the amplified genes, which were detected in the breast tumour cell lines with CGH microarray have been mapped to specific regions on individual chromosomes. However one gene (NAP4), which was amplified on chromosome 1 in the BT474 cells, does not yet have a precise cytogenic position. The chromosomal positions for each gene were retrieved from the Unigene and Locuslink NCBI databases (http://www.ncbi.nlm.nih.gov/). All experiments were in duplicate or triplicate and genes were only considered amplified if they were amplified greater than 2 fold in all experiments and if they were significantly amplified as judged by a T-test. A 2-fold cut off was chosen, as these experiment are analysing chromosomal amplification therefore unlike RNA, where there can be 1.3 or 1.8 fold changes amplifications of both alleles must be whole number changes or single alleles may be half or whole number changes.

5.2.1. Overview of Gene Amplification

The regions of amplification detected in the 7 in vitro breast cancer cell lines as compared to the non tumour HBL100 cell line are summarized in Figure 5.1, the BT474 cell line was also compared against pooled DNA from in vivo clinical breast tumour surround tissue. Each chromosome is represented by its ideogram and the coloured bars to the right of each chromosome represent the regions that contained amplified genes in the different tumour cell lines compared to the non tumour controls. The chromosome regions containing more than one amplified gene are represented by thicker bars. Two chromosomes, 21 and Y, contained no regions of amplification, this was expected in chromosome Y as all DNA tumour samples studied are derived from women and thus are lacking the Y chromosome. Also there was no difference in genomic or RNA expression between the breast tumour cell lines in genes located on chromosome Y when compared to the HBL100. This indicates that the HBL100 cells used in this study are likely to be genuine and not the male chromosome Y containing cells that have been detected by American Type Tissue Culture Collection (ATCC) (http://www.atcc.org).

The BT474 cell line was compared against the HBL100 cell line (red bars) and a pooled DNA sample from in vivo surround clinical breast tissue (pink bars). In general the regions containing amplification in the BT474 cells when compared to HBL100 were also present when compared to the pooled in vivo sample, indicating that the non tumour HBL100 cell line is a good control to use in this case. Although there were more changes detected in the BT474 cell line versus the pooled clinical sample.
Figure 5.1. Regions of Gene Amplification Detected in Breast Cancer Cell Lines

Continued over page
Figure 5.1. Legend: Regions of genomic amplification detected in the 7 different breast cancer cell lines. The ideograms of each chromosome are shown and the number corresponds to the chromosome. Each cell line is represented by a different colour as indicated and the corresponding bars to the right of the chromosomes represent regions that contain amplified genes. The BT474 cell line was compared against DNA from both the non-tumour breast cell line HBL100 and a pooled sample of in vitro surround breast tissue, whilst all other tumour cell lines were compared to HBL100 only. Regions containing a single amplified gene are represented by thin bars, whereas regions containing more than one amplified gene are represented by thick bars.
The ideograms in Figure 5.1 illustrate that there were many areas of common chromosomal amplification occurring in the different human breast cancer cell lines, although some of the areas also contained amplifications specific to one cell line. The 2 Dox resistant cells lines HPGP (light blue) and MDA16 (light green) had some areas of common amplification that may thus be related to the resistant phenotype. Included amongst these was a region at 7q21, which contains the $ABCB1$ and $ABCB4$ genes, as well as regions on chromosome 6p ($Bak1$) and 12q ($HNRPA1$) all of which have previously been associated with the resistant phenotype (Chapter 3). Interestingly the resistant cell lines had different regions of amplifications when compared to their sensitive progenitor cell lines, although there were some common regions of amplification such as in chromosome Xq, which was a region exclusive to the HPGP cells and their progenitor MCF7/wt cells.

There were some chromosomes in the tumour samples that contained multiple genes amplified within a single region. Amongst these regions was chromosome 20q that has a large area of multiple gene amplification in BT474, MCF7/wt and ZR75.1 cell lines and is analysed in greater detail in section 5.2.5. Another chromosome containing several amplified genes was chromosome 17. Due to the high frequency of amplification on chromosome 17 it is analysed in greater detail in chapter 7. Other chromosomes that contain regions with more than one amplification event included 2q, 7q and 12q.

**5.2.2. Specific Regions of Gene Amplification in Tumour Cell Lines**

Specific chromosome regions that contained amplifications are summarised in Table 5.2. There was a greater frequency of amplification in the BT474, MCF7 and ZR75.1 cell lines. The number of regions containing amplified genes were greater in BT474 when compared to the pooled clinical DNA (54), as opposed to BT474 versus HBL100 (23), although most regions detected against HBL100 were also present in the comparison against the in vivo pooled sample. T47D and MDA16 had the least regions containing amplifications, 5 and 4 respectively.

Regions containing several amplified genes are shown in bold. The majority of amplifications detected by the CGH microarray were of a single gene in a separate region, thus suggesting single gene amplifications. But it should be remembered that only about one sixth of the estimated 30,000 human genes are present on the microarrays, therefore some of these apparent single gene amplifications events could be part of a larger region of amplification along the chromosomes that is not detected with the clones currently on the microarray.
Table 5.2. Regions of Amplification Detected in the Breast Cancer Cell Lines

<table>
<thead>
<tr>
<th>Cell Line</th>
<th>Chromosome regions where genes are amplified</th>
<th>No. Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT474 vs. HBL100</td>
<td>1, lp33, 1q22, 1q31.1, 2p25.1, 3p21.3, 8p21.3-p21.2, 11q12, 14q12, 16q21-22.3, 17p13.1, <strong>17q11-q12</strong>, <strong>17q12</strong>, 17q21.1, 17q21-q22, 17q22, <strong>17q23.1</strong>, 17q23.2, 18q11.2, <strong>20q11.2</strong>, <strong>20q12</strong>, <strong>20q13</strong>, 22q11.21</td>
<td>23 (6)</td>
</tr>
<tr>
<td>BT474 vs. pooled in vivo DNA</td>
<td>1, lp33, 1q22, 1q23, 1q31.1, <strong>1q32.1</strong>, 1q32-q31, 2p25.1, 3p21.3, 3q27, 4p16.3, 4q21-q25, 5q31.1-q33.1, <strong>5q32-33</strong>, 5q35, 6p21.3, 6q15, 7p12.2, 8p21.3-p21.2, 8q24.3, 9p13, 9p24.1, 9q21-q22, <strong>9q34</strong>, 11p13, 11q13, 11q22, 12p12.1, <strong>12q13-14</strong>, 12q13-15, 12q23, 13q12, 13q14.1-q14.2, 14q12, 16q21-22.3, 17p13.1, 17q11.2-q12, <strong>17q12</strong>, 17q21-q22, <strong>17q22</strong>, <strong>17q23.1</strong>, <strong>17q23.3</strong>, 17q23-q24, 17q25.3, 18q11.1-q11.2, <strong>18q11.2</strong>, 18q21.3, 19p13.3, <strong>20q11.2</strong>, <strong>20q12</strong>, <strong>20q13</strong>, 22q11.21, Xq28</td>
<td>54 (12)</td>
</tr>
<tr>
<td>MCF7 vs. HBL100</td>
<td>1p22, <strong>2q33</strong>, 2q33-q34, 3p21.1-p12, 4q31.3-q33, 6p21.3, 7q11.22, 8q12-q13, 8q24.3, 9q22, 10p14-p13, <strong>12q13</strong>, <strong>12q23</strong>, 12q23-q24.1, 14q12, 14q23.3, 15q15.2, 15q21, 17q22-q23, <strong>17q23</strong>, 17q23-q24, 18q21.3, 19p13.3, <strong>20q11.2</strong>, <strong>20q12</strong>, <strong>20q13</strong>, Xq22.1</td>
<td>26 (7)</td>
</tr>
<tr>
<td>HPGP vs. HBL100</td>
<td>2q12.3, 3q27, 6p21.3, 7p15, <strong>7q21.1</strong>, 8p22-p21.3, 8p21.1, 8q12-q13, 8q22, 8q24, 11q13.1, 11q23.3-q25, 12q13.1, 17p13, 17q23.1-q23.3, 22q13.31, Xq22</td>
<td>17 (1)</td>
</tr>
<tr>
<td>MDA468 vs. HBL100</td>
<td>7p12, 7q11.2, 7q21.3, 8p22-p21, 8q13, 12p13, 12q12, 12q13-q15, 14q32, 16p13, 17q25.3, 20p13, 22q11.23, Xp11.3-p11.23</td>
<td>14 (0)</td>
</tr>
<tr>
<td>MDA16 vs. HBL100</td>
<td>6p21.3, <strong>7q21.1</strong>, 11q13.1, 12q13.1,</td>
<td>4 (1)</td>
</tr>
<tr>
<td>T47D vs. HBL100</td>
<td>1q22, 8q24.3, 12q23, 13q14.1-q14.2, 19p13.3</td>
<td>5 (0)</td>
</tr>
<tr>
<td>ZR75.1 vs. HBL100</td>
<td>1p22.1, 1p32.1-p31.3, 1p34.2, 1q32.2, 2q21.1, <strong>2q33-34</strong>, 3p21.1-p12, 4q12-q13.3, <strong>6p12</strong>, 6q21, 6q22, 7q21-q22.1, 7q22, 7q31, 8q22.3-q23, 9q34, 11q13, 12p12.1, 12q13.1, 12q23, 15q21, 15q23-q24, 16q22.1, 17p13.1, 17q21-q23, <strong>17q23</strong>, 18q11.2, 18q21.3, 19q12, 19q13.1-q13.2, <strong>20q11</strong>, 20q11.2-q12, <strong>20q13</strong>.</td>
<td>33 (5)</td>
</tr>
</tbody>
</table>

Table 5.2. Legend: A summary of all the regions of amplification detected in 7 breast cancer cell lines. The cell line and its relative control are indicated in the left column, whilst the amplification regions are listed in the middle. Those written in bold are regions which contain more than one amplified gene for that specific breast cell line. The total number of regions containing single or multiple amplifications for each cell line are shown in the right column.
5.2.3. Frequent Regions of Gene Amplifications in Cell Lines

Table 5.3 summarises the most frequent areas of gene amplification that were detected across all the breast tumour cell lines when compared to the putative HBL100 breast cell line only. The first part of the table shows the chromosome arm that contains the amplification and the number of cell lines involved, whilst the second section shows the precise region containing the gains within the chromosome arm. All chromosomes that contained amplification in greater than 43 % (3 samples) are shown and where there are two regions of chromosome gain in one arm they have been separated by parentheses.

Table 5.3. Most Frequent Chromosomal Gene Amplification Regions

<table>
<thead>
<tr>
<th>TUMOUR BREAST CELL LINES</th>
<th>Chromosome Arm Gains</th>
<th>Chromosome Region Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr. region</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>12q</td>
<td>6</td>
<td>86</td>
</tr>
<tr>
<td>17q</td>
<td>5</td>
<td>71</td>
</tr>
<tr>
<td>7q</td>
<td>5</td>
<td>71</td>
</tr>
<tr>
<td>8q</td>
<td>5</td>
<td>71</td>
</tr>
<tr>
<td>6p</td>
<td>4</td>
<td>57</td>
</tr>
<tr>
<td>11q</td>
<td>4</td>
<td>57</td>
</tr>
<tr>
<td>17p</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>20q</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>18q</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>1p</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>8p</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>14q</td>
<td>3</td>
<td>43</td>
</tr>
<tr>
<td>22q</td>
<td>3</td>
<td>43</td>
</tr>
</tbody>
</table>

Amplifications in chromosome 12 were the most frequent, with two specific regions involved 12q13 (71 %) and 12q23 (26 %). The long arm of chromosome 17 also had 2 regions of frequent amplification in the breast cancer cell lines at 17q23 and 17q22. The short arm of chromosome 17 also contained a region of frequent amplification at 17p13. Eight other chromosomes contained regions of common amplification including a large section of chromosome 20 at 20q11-q13.3. The mean number of regions detected, which contain amplifications in the cancer cell lines analysed was 17.
5.2.4. Specific Genes Involved in Frequent Regions of Amplification

There were a number of genes associated with the frequent regions of amplification that were amplified in more than one cell line sample when compared to HBL100, these are summarised in Table 5.4. These included 3 genes located on chromosome 20q, which were amplified in MCF7, BT474 or ZR75.1. Two of these genes CYP24 and BMP7 are both located in the q13 region. Another gene located close to BMP7 is RAE1 and was amplified in BT474. As this region is thought to contribute to a more aggressive tumour phenotype (Collins et al., 1998), and due to the frequent amplification of genes in this region, chromosome 20 is analysed in greater detail in the next section.

Table 5.4. Specific Genes Amplified in the Breast Cell Lines

<table>
<thead>
<tr>
<th>Gene Information</th>
<th>Breast Cancer Cell Line</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BT474</td>
</tr>
<tr>
<td>CASP8</td>
<td></td>
</tr>
<tr>
<td>SCA7</td>
<td></td>
</tr>
<tr>
<td>BAK1</td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td></td>
</tr>
<tr>
<td>ABCB4</td>
<td></td>
</tr>
<tr>
<td>MCM4</td>
<td></td>
</tr>
<tr>
<td>KIAA0014</td>
<td></td>
</tr>
<tr>
<td>BAD</td>
<td></td>
</tr>
<tr>
<td>HNRPA1</td>
<td></td>
</tr>
<tr>
<td>SLC25A3</td>
<td></td>
</tr>
<tr>
<td>ARHGAP5</td>
<td></td>
</tr>
<tr>
<td>MAPK6</td>
<td></td>
</tr>
<tr>
<td>PPM1D</td>
<td></td>
</tr>
<tr>
<td>CDC34</td>
<td></td>
</tr>
<tr>
<td>NCOA3</td>
<td></td>
</tr>
<tr>
<td>CYP24</td>
<td></td>
</tr>
<tr>
<td>BMP7</td>
<td></td>
</tr>
</tbody>
</table>

A = the gene was amplified in this breast tumour cell line

There are several genes, BAK1, ABCB1, ABCB4, BAD and HNRPA1 that were amplified in the resistant cell lines when compared to HBL100, and as seen in chapter 3 may be involved in resistance, although HNRPA1 was also amplified here in the relatively sensitive ZR75.1 cells. MCM4 is a gene that was amplified in both the resistant HPGP cell line and its sensitive
MCF7 progenitor. Four other genes were amplified in both MCF7 and ZR75.1, and 2 were amplified in BT474 and MCF7, and a further 3 genes were amplified in MCF7 and T47D. The MDA468 cells did not contain any genes that were amplified in the other breast cell lines.

5.2.5. Chromosome 20 Amplifications

The long arm of chromosome 20 contained multiple gene amplifications in 3 different cell lines, BT474, MCF7 and ZR75.1 consequently in this section the region is studied further. In order to represent this region graphically relative locations of genes located on chromosome 20 were determined. GeneMap '98 (http://www.ncbi.nlm.nih.gov/genemap98) was used to retrieve the cR3000 of genes, although many genes have not been currently been assigned a cR3000, thus to obtain a more complete representation of all the genes from this region, the genomic sequence generated by the Human Genome Project was used to plot relative cR3000 positions of other genes (http://www.ncbi.nlm.nih.gov/Locuslink). Ratios of the genes present on chromosome 20 were plotted against the cR3000 values for the 3 cell lines BT474, MCF7 and ZR75.1 (Figure 5.2). T47D, a breast cell line that did not show any amplification on chromosome 20 is also shown.

There were 8 regions of gene amplification within the 3 tumour cell lines, some of these regions are represented by a single cell line, while others are common to several. Genes at region 1 and 5 were found amplified in MCF7 cells only, genes at 2 and 4 in ZR75.1 cells only, region 3 was found in BT474 cells and genes at regions 6, 7 and 8 were amplified in all cell lines. The T47D cell line is plotted in graph D, where it can be seen there were no amplified genes in this sample on chromosome 20, when compared with the plots for the other cell lines.
Figure 5.2. Legend: Chromosome 20 plots for 4 different breast cancer cell lines. Three of the cell lines BT474, MCF7 and ZR75.1 had genes that were amplified, while T47D did not. In each case the cR3000 value of the genes along the chromosome is plotted against the gene average ratio of the breast cancer cell line vs HBL100. There were 7 different areas of gene amplification within the 3 cell lines, which are represented by the grey bars.
5.2.5.1. Specific Genes Associated with Chromosome 20 Amplifications

Specific genes that were amplified at the 8 regions indicated in Figure 5.2 are shown in Table 5.5. The fold change or ratio and standard deviations are shown for each gene and where available RNA expression data is present. Genes in red were 2-fold over in the breast tumour cell lines compared to HBL100 and so were considered amplified or overexpressed. Region 1 was situated on the short arm of chromosome 20, and regions 2-8 were located on the 20q long arm.

The gene associated with region 1 on 20p in the MCF7 cells was CST3, although amplified in MCF7 only, the gene was also overexpressed in all 3 breast tumour cell lines. ID1 in region 2 and was amplified in ZR75.1 only. Region 3 contained 5 genes all of which were amplified in the BT474 cells only, CEP2 was also overexpressed at the RNA level in these cells. Two genes were present in region 4, which were amplified in ZR75.1 cells, including TGIF1, which was also overexpressed. Region 5 contained 2 genes that were amplified but not overexpressed in the MCF7 cells only. Regions 6, 7 and 8 contained genes that were amplified in all breast cancer cell lines. Included in these regions are several genes, PRKCBP1, NCOA3, PFDN4, BMP7, CSTF1 and RAE1 that are also overexpressed in 1 or more breast cancer cell lines. CYP24 is the only gene amplified in all breast cancer cell lines, but not overexpressed.
### Table 5.5. Specific Genes Associated with the Chromosome 20 Amplifications

<table>
<thead>
<tr>
<th>Region location</th>
<th>Gene</th>
<th>BT474 DNA</th>
<th>BT474 RNA</th>
<th>MCF7 DNA</th>
<th>MCF7 RNA</th>
<th>ZR75.1 DNA</th>
<th>ZR75.1 RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 20p11</td>
<td>CST3</td>
<td>1.5 (0.4)</td>
<td>2.2 (0.1)</td>
<td>2.1 (0.9)</td>
<td>4.5 (2.2)</td>
<td>0.8 (0.4)</td>
<td>2.2 (0.0)</td>
</tr>
<tr>
<td>2 20q11</td>
<td>ID1</td>
<td>0.5 (0.8)</td>
<td>2.6 (0.8)</td>
<td>1.5 (0.2)</td>
<td>1.4 (0.7)</td>
<td>2.3 (0.8)</td>
<td>1.5 (0.0)</td>
</tr>
<tr>
<td>3</td>
<td>ITCH</td>
<td>3.0 (0.6)</td>
<td>1.2 (0.5)</td>
<td>0.7 (0.1)</td>
<td>NA</td>
<td>0.7 (0.1)</td>
<td>0.1 (0.2)</td>
</tr>
<tr>
<td>20q11.2-q12</td>
<td>GSS</td>
<td>2.1 (0.1)</td>
<td>1.7 (1.0)</td>
<td>1.0 (0.0)</td>
<td>0.6 (0.0)</td>
<td>0.9 (0.1)</td>
<td>0.9 (0.1)</td>
</tr>
<tr>
<td>4 20q12-q13</td>
<td>PROCR</td>
<td>2.4 (0.0)</td>
<td>0.3 (0.4)</td>
<td>1.0 (0.0)</td>
<td>0.4 (0.7)</td>
<td>0.6 (0.5)</td>
<td>0.9 (0.1)</td>
</tr>
<tr>
<td>5</td>
<td>MMP24</td>
<td>1.9 (0.7)</td>
<td>0.3 (0.5)</td>
<td>1.0 (0.0)</td>
<td>NA</td>
<td>0.8 (0.1)</td>
<td>0.5 (0.0)</td>
</tr>
<tr>
<td>6 20q13</td>
<td>CEP2</td>
<td>2.3 (0.4)</td>
<td>2.6 (0.9)</td>
<td>0.8 (0.0)</td>
<td>0.2 (0.3)</td>
<td>1.0 (0.0)</td>
<td>0.6 (0.0)</td>
</tr>
<tr>
<td>7 20q13</td>
<td>TGIF2</td>
<td>1.1 (0.1)</td>
<td>1.1 (0.3)</td>
<td>1.0 (0.0)</td>
<td>NA</td>
<td>2.6 (0.1)</td>
<td>2.4 (0.1)</td>
</tr>
<tr>
<td>20q11-q12</td>
<td>RBL1</td>
<td>1.3 (0.4)</td>
<td>1.6 (1.3)</td>
<td>1.3 (0.1)</td>
<td>0.6 (0.1)</td>
<td>2.2 (0.4)</td>
<td>0.5 (0.0)</td>
</tr>
<tr>
<td>5</td>
<td>BPI</td>
<td>1.0 (0.3)</td>
<td>0.3 (0.4)</td>
<td>2.1 (1.6)</td>
<td>0.8 (0.7)</td>
<td>1.0 (0.0)</td>
<td>0.3 (0.5)</td>
</tr>
<tr>
<td>20q11-q12</td>
<td>TOP1</td>
<td>1.2 (0.1)</td>
<td>1.1 (0.3)</td>
<td>2.3 (0.0)</td>
<td>1.5 (1.0)</td>
<td>1.4 (0.3)</td>
<td>1.2 (0.1)</td>
</tr>
<tr>
<td>6 20q12-q13</td>
<td>PRKCBP1</td>
<td>1.0 (0.0)</td>
<td>3.2 (0.7)</td>
<td>5.0 (0.9)</td>
<td>4.7 (1.1)</td>
<td>2.5 (1.4)</td>
<td>1.0 (0.2)</td>
</tr>
<tr>
<td>7</td>
<td>NCOA3</td>
<td>3.4 (1.5)</td>
<td>6.3 (3.6)</td>
<td>2.9 (0.1)</td>
<td>5.4 (0.0)</td>
<td>1.9 (0.1)</td>
<td>1.4 (0.1)</td>
</tr>
<tr>
<td>8 20q13</td>
<td>BMP7</td>
<td>2.2 (0.5)</td>
<td>5.5 (2.1)</td>
<td>3.0 (0.0)</td>
<td>7.2 (4.4)</td>
<td>2.7 (0.4)</td>
<td>1.3 (0.0)</td>
</tr>
<tr>
<td>20q13</td>
<td>CSTF1</td>
<td>2.3 (0.0)</td>
<td>2.2 (0.0)</td>
<td>0.7 (0.0)</td>
<td>0.3 (0.6)</td>
<td>0.8 (0.2)</td>
<td>1.0 (0.0)</td>
</tr>
<tr>
<td>8</td>
<td>RAE1</td>
<td>4.0 (0.7)</td>
<td>5.9 (0.9)</td>
<td>0.9 (0.0)</td>
<td>2.3 (0.2)</td>
<td>0.5 (0.5)</td>
<td>1.7 (0.1)</td>
</tr>
<tr>
<td>20q13</td>
<td>HSRNASEB</td>
<td>2.3 (0.3)</td>
<td>1.6 (0.3)</td>
<td>1.4 (0.2)</td>
<td>0.5 (0.5)</td>
<td>1.4 (0.3)</td>
<td>0.8 (0.3)</td>
</tr>
<tr>
<td>8</td>
<td>PCK1</td>
<td>2.7 (1.0)</td>
<td>0.9 (0.0)</td>
<td>0.7 (1.0)</td>
<td>0.8 (0.4)</td>
<td>1.3 (0.3)</td>
<td>0.5 (0.0)</td>
</tr>
</tbody>
</table>

Table 5.5. Legend: Specific genes that were amplified within the 8 regions on chromosome 20. Red indicates genes had a fold change (ratio) of greater than 2 and so are considered amplified or overexpressed. NA = no data available. The fold change or ratio and (standard deviation) for each gene are shown.
5.3. Discussion

There was a high incidence of gene amplification recorded in some chromosome regions of the cell lines when compared to the putative HBL100, with the highest incidence occurring in BT474, MCF7 and ZR75.1, which contained 23, 26 and 33 regions respectively. The mean number of aberrations seen across all experiments was 17, this agrees well with previous studies, which recorded 19 (Forozan et al., 2000) and 16.3 (Larramendy et al., 2000) mean aberrations in cell lines. The high number of chromosome aberrations seen in breast cancer cell lines may be explained by the hypothesis that cell lines are a genetically advanced population (Kallioniemi et al., 1994), as the cells in culture are free from some growth restraints found in surrounding cells \textit{in vivo} and so can undergo clonal expansion at an increased rate, and thus may undergo increased chromosomal aberrations.

BT474 cells were also compared against a pooled sample of \textit{in vivo} clinical DNA extracted from surrounding non tumour breast tissue. There was good correlation between amplifications detected using the HBL100 DNA control and pooled \textit{in vivo} control, as generally the regions containing amplified genes in BT474 when compared to HBL100 were also present when compared to the pooled \textit{in vivo} sample. However there were many amplifications detected in the BT474 cell line against the \textit{in vivo} pooled clinical DNA that were not present when compared to HBL100 cells. The HBL100 and BT474 cells are immortalised cells and may be a genetically advanced population as it is possible that subsequent genetic aberrations have occurred in both these cell populations. Thus when compared against each other there will be less differential amplification between the two cell populations as opposed to the cell line against \textit{in vivo} cells.

5.3.1. Frequent Regions of Gene Amplification in Cell Lines

Many of the amplification events only occurred in a single breast cancer cell line when compared to HBL100, however some chromosome arms containing amplifications were detected in several tumours. The most frequent of which are summarised in Table 5.3. Some of these regions of amplification have been observed previously (Table 5.1.). Findings from this study and literature are compared in Table 5.6. The first column contains data from this chapter and the subsequent columns are data from literature. The regions containing chromosome gains are shown and the relative frequency of amplification occurring in these chromosomes in each study are shown.
Table 5.6. Comparison of Frequent Regions of Amplification with Literature

<table>
<thead>
<tr>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>7q</td>
<td>5</td>
<td>71</td>
<td>7q</td>
<td>7</td>
<td>39</td>
<td>7q</td>
<td>15</td>
<td>39</td>
</tr>
<tr>
<td>8q</td>
<td>5</td>
<td>71</td>
<td>8q</td>
<td>18</td>
<td>100</td>
<td>8q</td>
<td>28</td>
<td>75</td>
</tr>
<tr>
<td>17q</td>
<td>5</td>
<td>71</td>
<td>17q</td>
<td>12</td>
<td>67</td>
<td>17q</td>
<td>12</td>
<td>33</td>
</tr>
<tr>
<td>20q</td>
<td>3</td>
<td>43</td>
<td>20q</td>
<td>15</td>
<td>83</td>
<td>20q</td>
<td>21</td>
<td>55</td>
</tr>
<tr>
<td>1p</td>
<td>3</td>
<td>43</td>
<td>1p</td>
<td>8</td>
<td>44</td>
<td>1p</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>6p</td>
<td>4</td>
<td>57</td>
<td>6p</td>
<td>10</td>
<td>56</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11q</td>
<td>4</td>
<td>57</td>
<td>11q</td>
<td>7</td>
<td>39</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14q</td>
<td>3</td>
<td>43</td>
<td>14q</td>
<td>4</td>
<td>22</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1q</td>
<td>16</td>
<td>89</td>
<td>1q</td>
<td>23</td>
<td>61</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3q</td>
<td>12</td>
<td>67</td>
<td>3q</td>
<td>15</td>
<td>39</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5p</td>
<td>11</td>
<td>61</td>
<td>5p</td>
<td>15</td>
<td>39</td>
</tr>
<tr>
<td>12q</td>
<td>6</td>
<td>86</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8p</td>
<td>3</td>
<td>43</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17p</td>
<td>3</td>
<td>43</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18q</td>
<td>3</td>
<td>43</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22q</td>
<td>3</td>
<td>43</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16q</td>
<td>5</td>
<td>28</td>
<td>-</td>
<td></td>
<td></td>
<td>22p</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7p</td>
<td>17</td>
<td>44</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20p</td>
<td>11</td>
<td>30</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Five of the frequent chromosome arms detected in this study, 7q, 8q, 17q, 20q and 1p that contained amplified genes have also been reported in the literature. Chromosomal gains at 6p, 11q and 14q were detected in this study and by Larramendy et al. (2000), and frequent genetic gains at 8q, 11q, 17q and 20q have also been reported in pancreatic cancer (Mahlamaki et al., 1997). The high incidence of chromosomal aberrations in these chromosome arms suggests that they are important in cancer development and may therefore contain specific genes involved in breast cancer.

Gains at 8q were detected with the highest frequency in all studies (Table 5.6). Amplifications at 8q have been associated with an aggressive phenotype (Isola et al., 1995). 8q gains also occur frequently in primary breast tumours (Tirkkonen et al., 1998), prostate (Steiner et al., 2002), pancreatic (Mahlamaki et al., 1997) and endometrial cancer (Suzuki et al., 1997). The well-established oncogene \(myc\), maps within this region to 8q24, although it was not amplified in the breast cell lines in this study so is not apparently involved in the...
amplifications in these breast cancer cell lines. Genes that were frequently amplified at 8q in the breast cell lines here included MCM4 and KIAA0014, the latter was both amplified and overexpressed in MCF7 and T47D cells. Little information is available about KIAA0014 as it is a poorly characterised protein, although it is thought to be similar to PRAME, which is located on chromosome 22. However when the Unigene nucleotide sequence of KIAA0014 is queried using BLAST no sequence similarity with PRAME is found and PRAME itself is not amplified, thus is it unlikely that there is any cross hybridisation between these genes. As KIAA0014 was amplified and overexpressed in cell lines it may play a direct role in breast cancer, although the overexpression of this gene may merely be due the extra copies of the coding region or promoter sequences.

The long arm of chromosome 17, 17q is known to be involved in breast cancer, and individual genes along this chromosome have previously been associated with breast cancer. The breast cancer related tumour suppressor gene (BRCA1) is located at 17q, and when mutated the BRCA1 gene is associated with an increase in the overall risk of developing breast and ovarian cancer or both (Thorlacius et al., 1995). Another gene located on chromosome 17q is the ERBB2 oncogene. ERBB2 is amplified in approximately 20% of breast tumours and its presence is an indicator of poor prognosis (Revillion et al., 1998). The ERBB2 gene was amplified with some neighbouring genes on chromosome 17q21.1 in the BT474 cells, which have previously been shown to amplify the ERBB2 gene (Pollack et al., 1999). Due to the high frequency of amplifications occurring in chromosome 17 its possible role in breast cancer has been analysed and discussed in greater detail in chapter 7.

The ZR75.1, MDA16 and HPGP cells contained amplifications on the long arm of chromosome 11 at 11q13. This region is frequently amplified in human breast cancer (Larramenddy et al., 2000), esophageal carcinomas (Tsuda et al., 1988) and lung carcinoma (Berenson et al., 1990), but specific genes responsible for the emergence of amplifications in this area are largely unknown (Szepetowski et al., 1992). The PRAD1 gene located at 11q13 is both amplified and overexpressed in some breast cancers and so it has been suggested as a candidate for breast cancer amplification of 11q13 (Szepetowski et al., 1992). In this chapter the FAU gene located at 11q13 was also amplified and overexpressed in the ZR75.1 tumour cell line, suggesting a possible role in breast cancer for this gene.

The frequent areas of chromosome amplification detected in this study not previously reported were 8p, 12q, 17p, 18q and 22q. These may still be involved in the development of
breast cancer, since the types of aberrations occurring in cancer can be diverse and often vary according to tumour histology and type or genetic and ethnic background. For example HNRAP1, a gene associated with the resistant phenotype located at 12q (Chapter 3), was in a region that had not previously been detected however, the inclusion of 2 cell lines displaying a resistant phenotype in this analysis has resulted in its detection here.

5.3.2. Chromosome 20 Amplification

Chromosome 20 is approximately 72 Mb in length, more than 80 % of which has been sequenced (http://www.ncbi.nlm.nih.gov/genome/guide/HsChr20.shtml). There appeared to be several discrete regions containing amplified genes on chromosome 20 (Figure 5.2.), some have been recorded in the literature and possible candidate genes for these amplifications have been proposed, however some have not previously been seen. Seven of these regions were located on the long arm of chromosome 20 and some of these regions contained specific genes common to more than 1 tumour (Table 5.5). Amplification of genes on 20q is thought to contribute to neoplastic transformation and/or progression in breast cancer and other tumours (Collins et al., 1998).

Region 3 on chromosome 20 contained 5 amplified genes and was detected at 20q11.2-q12 in the BT474 cell lines (Figure 5.2; Table 5.5.). These genes included ITCH which interacts with the DNA binding transcription factor p45/NF-E3 and whose mouse homologue has been implicated in regulation of growth and differentiation of erythroid and lymphoid cells (Chen et al., 2001), GSS a glutathione synthetase, PROCR, MMP24 and CEP2. CEP2 were also overexpressed in the BT474 cells. Two other regions at 20q11 also harboured amplified genes. The first was found in the ZR75.1 breast cells and contained TGIF2 and RBL1, the former gene was also overexpressed in these cells (Region 4). Amplification and overexpression of TGIF2 also occurs in some ovarian cancer cell lines, thus it has been suggested it plays a role in the development and/or progression of some ovarian tumours (Imoto et al., 2000). The other region in 20q11 was found in MCF7 cells and included 2 genes BPI and TOP1, both of which were amplified but not overexpressed (Region 5).

The results in this chapter and Larramendy et al., (2000) have shown another region of amplification on the long arm of chromosome 20 in breast cancer cell lines, 20q12-q13.2 (Region 6; Table 5.5.). Amplification at 20q13 occurs in a variety of tumour types and is associated with aggressive tumour behaviour (Collins et al., 1998). The first CGH study showed an increase in copy number at 20q13 in 40 % of breast cancer cell lines and 18 % of
primary breast tumours (Kallioniemi et al., 1994). Subsequent CGH studies have also revealed copy number gains at 20q12 in 25% and 20q13 in more than 29% of cancers in the ovary (Tanner et al., 2000), colon and pancreas (Schlegel et al., 1995). Several candidate genes have been suggested in this region including AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer (Tanner et al., 2000; Anzick et al., 1997). BTAK, which is amplified in several cancers including breast, ovarian and colon (Tanner et al., 2000; Sen et al., 1997), CAS, a cellular apoptosis susceptibility gene previously seen amplified in BT474 cells (Brinkmann et al., 1996) and TFAP2C (Williamson et al., 1996). Unfortunately some of these genes are not found on the current MRC Leicester microarray, thus data is not available for them. However AIB1, which is also referred to as NCOA3 is found on the microarray, and was amplified and overexpressed in BT474 and MCF7 cells and marginally amplified in ZR75.1 cells in (Region 6), thus confirming that this gene may have an active role in breast cancer. Another gene PRKCBP1, a protein kinase C binding protein, which is in close proximity to NCOA3 in region 6 was also amplified in this study in MCF7 and ZR75.1 cells, and overexpressed in BT474 and MCF7 cells and thus may have a role in breast cancer. PTPN1 and BTAK (STK15) are 2 genes located at 20q13 that have also been associated with amplification on 20q (Tanner et al., 2000), however neither was amplified nor overexpressed in the breast cancer cell lines here.

Another separate area of increased copy number at 20q13.2, which does not involve the above genes has been described (Collins et al., 1998). High level of amplification at this region is associated with poor prognosis in lymph node negative breast tumours (Isola et al., 1995) and has been correlated to metastatic potential and tumour progression in colorectal cancers (Hidaka et al., 2000). Collins et al., (1998) found 4 previously unseen genes, PIC1L, ZNF217, ZNF218 and NABC1 and a known gene CYP24 located in a 1Mb region studied within this area. Three of the genes CYP24, ZNF218 and PIC1L were excluded as candidate tumour genes, as they were not overexpressed at the RNA level in cancer cell lines, which contained the 20q13.2 amplification. Of these genes only CYP24 is found on the MRC Leicester array chip and the CYP24 (Region 7) was amplified in BT474, MCF7 and ZR75.1 cell lines but not overexpressed, thus confirming literature data. The ZNF217 and NABC1 genes have been proposed as oncogenes in breast cancer (Collins et al., 1998) and the ZNF217 gene is also associated with increased metastatic potential in colorectal cancers (Hidaka et al., 2000) and has the ability to immortalise cultured mammary epithelial cells (Nonet et al., 2001). Neither of these genes are located on the microarray, however a neighbouring gene PFDN4, was also
amplified and overexpressed in the BT474 breast cells in this study (Region 7) and recently it has been suggested as a driver gene for gene amplification in this region (Collins et al., 2001).

The last discrete region detected (region 8) contained 5 genes, which have not previously been seen in breast cancer. Three of these genes, HSRNASEB, PCK1 and CSTF1 were amplified in BT474 cells and the latter was also overexpressed. There is evidence to suggest genetic linkage between PCK1 and other genes in this chromosome region with the development of non-insulin dependant diabetes mellitus (Hani et al., 1996). RAE1, a gene involved in the nuclear export of polyA mRNA (Bharathi et al., 1997) was also amplified in BT474 cells and overexpressed in BT474 and MCF7 cells, which could suggest an increase in mRNA synthesis in these cells. The final gene in this region BMP7 was amplified in all 3 breast cancer cell lines and was also overexpressed in BT474 and MCF7 cells, thus it may play an important role in breast cancer. BMP7 is a bone morphogenic protein that can be used as gene therapy to stimulate bone regeneration (Franceschi et al., 2000). It is also thought to be involved in collecting tubule cell proliferation and apoptosis, at low doses BMP7 increases tubular number and cell proliferation, while high doses inhibit cell proliferation and stimulate apoptosis (Piscione et al., 2001). Therefore although the breast cell lines both amplify and overexpress BMP7 compared to HBL100 cells, this may be at an optimum low dose that results in increased cell proliferation.

The ability of GCH microarray to identify single genes allows the detection of small discrete regions of amplification, which when analysed by CGH may not be resolved and may be seen as one amplified area. However radiation hybrid mapping, which has been used to derive cR3000 values retrieved from GeneMap, cannot determine the relative position of adjacent markers or ESTs, consequently the number as well as the position of amplification peaks identified here may change as this region is sequenced or physically mapped. It is expected that the complete human genome sequence will replace the radiation hybrid map in the near future and as the sequence information from the human genome mapping project becomes available construction of microarrays representing specific regions of chromosomes will be available.

5.3.3. Chromosomal Aberrations in Specific Cell Lines

Some of the cell lines that I have analysed with CGH microarray here have previously been examined for chromosomal aberrations. Kallioniemi et al. (1994), analysed the BT474 cell line for chromosomal regions containing areas of increased DNA copy number. They used
CGH and found DNA copy number increases at 17q12, 17q22-q24 and 20q13. Other regions of chromosomal gain detected in the BT474 breast cell line are 20p11.1-20q11.2, 20q13.1-20q13.2, 9q32-9q34 (Guan et al., 1994). Most of these regions were confirmed in this study, although amplifications on chromosome 9 were only detected at 9q21-22, 9q34 only when the BT474 cells were compared against the in vivo non tumour pooled DNA.

Previous regions of amplification recorded in ZR75.1 cells include 11q13, 12q14-q15, 17q22-qter (Kallioniemi et al., 1994). These regions, apart from 12q14-q15 were also shown to contain amplifications in the data described here. Although amplification was not detected at 12q14-q15 regions surrounding this, 12q13 and 12q23 did contain regions of amplification. Whole arm gains at 1q, 7p, 12p, 16p, 20q and 22 have also been recorded in ZR75.1 cells with CGH (Kallioniemi et al., 1994). Microarray CGH does not have the ability to detect whole chromosome arm gains, however there were many different gene amplifications detected in 20q, suggesting a large region or high activity of amplification.

Forozan et al. (2000) included the MCF7 and ZR75.1 cell lines in their study of 38 cell lines, which coupled the CGH technique with RNA expression analysis of 1236 genes. As previously seen the frequent regions containing increased copy number detected in this study and in this chapter correlate well (Table 5.5.). Forozan et al. (2000) detected some genes that were overexpressed in MCF7 cells in these regions of chromosomal gains including RCH1 (17q23), which is not present on the Leicester Toxicology microarray and thus data is not available, however a gene in the same region SMARCD2 (17q23-24) was overexpressed and amplified here, and other neighbouring genes were amplified in this area but not overexpressed (PPM1D). Good correlation is seen in my data and the Forozan et al. data in expression of the TMP21 gene, which was located in a region that did not contain frequent amplifications (14q24.3), and was not amplified in my experiments but is overexpressed in both studies. The study by Forozan et al., detected regions that had undergone amplification by CGH and corresponding gene expression of a small number of genes was analysed. However discrete or small regions of amplification may not be detected by CGH, so whether the genes are truly amplified as well as overexpressed is not certain, whereas CGH microarray coupled with expression microarray can determine the amplification and expression of each gene. Although unlike CGH, regions of whole arm chromosomal gain will not detected unless all genes are spotted on the microarray.
Some data from the Forozan study contradicted my findings, particularly for MYBL2 and PRKACA genes, which are overexpressed in the Forozan study but not here. However different control samples were used in the 2 studies, the putative non tumour HBL100 cell line was used here and normal mammary gland RNA in the other, which may account for some differences detected in RNA expression. Interestingly Forozan et al. included the HBL100 cell line in their study as part of the 38 breast carcinoma cell lines analysed for chromosomal gains by CGH and so presumed it was a breast cancer cell line. However the HBL100 cell line is derived from non tumour breast cells and they exhibit some non tumour properties eg when transfected into animals do not cause tumours, thus although they have been immortalised they exhibit some properties that are not true tumour.

Regions of chromosome DNA copy number increase at 1cen-q23, 3p14, 8q21-qter, 15q21-qter, 16q23-q24, 17q22-q24, 20q13 have also been detected in MCF7 cells (Kallioniemi et al., 1994). Data from this chapter confirmed that genes located at 1q, 8q21, 15q21, 17q22-24 and 20q13 were amplified. However no genes at 13p14 or 6q23-24 were found amplified in the MCF7 cells, although nearby locations at 3p21.1-pl2 also contained genes that were amplified. Whole chromosome arm gains in MCF7 cells of 5p, 12q and 14q have also been detected (Kallioniemi et al., 1994). The chromosome arms of 12q and 14q both contain many genes that were amplified in my work. Good correlation was also found by a FISH study that confirms the amplification events at 17q22 and 20q12-q13.2 in MCF7 cells (Guan et al., 1994). This good correlation with literature indicates that the cell lines used in this study are similar to those in the literature.

5.3.4. Problems Associated with Microarray Comparative Genomic Hybridisation

There are some problems associated with the use of microarrays to analyse genomic aberrations. The problems associated with background noise are increased in CGH microarrays when compared to RNA microarrays. Consequently it is necessary to ensure good genomic DNA is used as a starting material prior to labelling, otherwise the background noise is too great. Also detection of deletions or chromosomal losses is difficult and maybe at the microarray CGH detection limits, as if a gene is deleted the copy number may be going from 1 to 0 in the tumour cell line, in contrast DNA amplification could be from 1 copy number to many.

Microarray CGH allows for the identification of specific genes present in areas of chromosomal gain that are both amplified and overexpressed and so maybe involved in
carcinogenesis, and thus aids in our understanding of breast cancer and could suggest possible cancer-related genes or therapeutic targets. Although like CGH it does not have the ability to detect other chromosomal aberrations such as translocations or gene mutations, so it is not able to elucidate all the events occurring. Also microarray does not discriminate between alleles, therefore loss of one allele followed by duplication of the other would remain undetected.

5.3.5. Summary

The data validates the use of cell lines as appropriate tools in breast cancer research, as DNA copy number changes detected in cell lines will provide data for investigation of tumour progression in vitro and will aid in more detailed mapping and isolation of genes implicated in cancer. Many regions containing amplified genes were elucidated in the panel of cell lines studied and some of these regions were present in more than one breast cell line. In general there was a good correlation between regions containing chromosome gains, with data obtained in this study using CGH microarray, and those found in literature that used FISH or CGH. The corresponding RNA expression data illustrated that several genes located in frequent regions of amplification were amplified and overexpressed, whilst others were amplified but not overexpressed. These included 2 amplified genes KIAA0014 and MCM4 the former of which was also over expressed. The region of amplification at chromosome 20 previously detected by CGH was found to contain several discrete regions of amplification and some genes such as NCOA3 and BMP7 were also overexpressed. It is thought that genes such as these may play a role in breast cancer, as their amplification and overexpression may be giving the tumour cells increased survival and a more aggressive phenotype.

Some regions of chromosome loss or deletion were thought to be outside the detection limits of microarray. However despite this, microarray CGH has the ability to resolve small and discrete regions of amplification and can identify specific genes within these regions. Thus it can aid in our understanding of breast cancer and could suggest possible novel therapeutic targets. The ability of microarrays to provide information not only about RNA expression but also genomic amplification is an immense benefit, as together this information combined may elucidate specific genes that may be involved in the development or phenotype of breast cancers, and this may be used to select appropriate treatment.
Chapter 6

Genomic Microarray Analysis of Clinical Breast Tumours
6.1. Introduction

Many recurrent aberrations have been seen in different types of tumour, although some specific patterns of genomic change depend on tumour histology, genetic or ethnic background. Several studies have carried out genetic analysis on different aspects of breast carcinoma, including type of tumour (IDC, ILC, DCIS), tumour grade, lymph node status and metastatic potential.

Here 13 human clinical breast tumour samples have been analysed with microarray GCH to detect areas of amplification. The tumours were of different grade or hormone receptor status, and are summarised in Table 6.1. The majority of tumours were IDC, and there was one of each tubular carcinoma, ILC and DCIS. Of the IDC tumours, 8 were node status positive, 2 were negative and 5 were grade II and grade III tumours. The menopausal status, \( ER \) and \( ERBB2 \) (\( HER-2 \)) status if available is also shown.

### Table 6.1. Clinical Breast Tumour Information

<table>
<thead>
<tr>
<th>Reference Number</th>
<th>Menopausal Status</th>
<th>Lymph node status</th>
<th>Type</th>
<th>Grade</th>
<th>( ER ) Status</th>
<th>( ERBB2 ) status</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Post</td>
<td>-ve</td>
<td>Tub</td>
<td>I</td>
<td>+ve</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>+ve</td>
<td>IDC</td>
<td>II</td>
<td>+ve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>Post</td>
<td>-ve</td>
<td>ILC</td>
<td>II</td>
<td>-ve</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>+ve</td>
<td>IDC</td>
<td>III</td>
<td>-ve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>05</td>
<td>Pre</td>
<td>-ve</td>
<td>DCIS</td>
<td></td>
<td>+ve</td>
<td></td>
</tr>
<tr>
<td>06</td>
<td>Post</td>
<td>-ve</td>
<td>IDC</td>
<td>II</td>
<td>+ve</td>
<td></td>
</tr>
<tr>
<td>07</td>
<td>+ve</td>
<td>IDC</td>
<td>II</td>
<td>-ve</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08</td>
<td>Post</td>
<td>+ve</td>
<td>IDC</td>
<td>III</td>
<td>-ve</td>
<td></td>
</tr>
<tr>
<td>09</td>
<td>Pre</td>
<td>+ve</td>
<td>IDC</td>
<td>III</td>
<td>-ve</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Pre</td>
<td>+ve</td>
<td>IDC</td>
<td>II</td>
<td>-ve</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Pre</td>
<td>+ve</td>
<td>IDC</td>
<td>II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Pre</td>
<td>-ve</td>
<td>IDC</td>
<td>III</td>
<td>-ve</td>
<td>+ve</td>
</tr>
<tr>
<td>13</td>
<td>Pre</td>
<td>+ve</td>
<td>IDC</td>
<td>III</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Post: post menopausal (>50 yrs); Pre: pre menopausal (<50 yrs)
Blank: not determined; -ve: negative; +ve: positive
Tub: Tubular carcinoma; IDC: Infiltrating ductal carcinoma; ILC: Infiltrating lobular carcinoma; DCIS: Ductal carcinoma in situ

6.1.1. Genomic Alterations in Infiltrating Breast Carcinoma

Several previous studies have analysed different grades and type of breast tumours and these are summarised in Table 6.2. and 6.3. Table 6.2. summarises data from two published studies regarding frequent chromosomal gains and losses detected in infiltrating ductal and lobular...
human breast tumours. The first study analyses changes occurring in both ILC and IDC tumours (Nishizaki et al., 1997), and the second study looked at IDC tumours that are been divided into poorly differentiated and well differentiated (Roylance et al., 1999).

There are some common regions of gain in the ILCs and IDCs, as chromosomal gains at 1q and 8q were detected at a high frequency in all the studies. Gains at 17q were detected at a high frequency in all infiltrating carcinomas, apart from in the study which analysed IDC tumour that were well differentiated (Roylance et al., 1999). Other frequent chromosomal gains detected in IDC but not ILC were found at 20q and 16p. In general ILC compared to IDC show a greater frequency of recurrent losses at 16q, where as gains at 8q and 20q occur more frequently in IDC (Nishizaki et al., 1997).

There were a high number of regions containing chromosomal losses in both the studies of Nishizaki et al., (1997) which analysed ILC and IDC and many of the regions detected in the studies were common to both. The other study had fewer regions of chromosomal loss, but all regions detected were present in the study by Nishizaki et al., (1997). There was one region of loss at 4q that was only frequently detected in the ILC tumours and 3 regions, which were specific to the IDCs 1p, 3p and 5q. The region of chromosomal loss at 1p was also detected in the poorly differentiated IDC tumours (Roylance et al., 1999).

6.1.2. Genomic Alterations in Ductal Breast Carcinoma

Table 6.3. summarises 3 studies from literature that examined DCIS. Two of the studies looked at DCIS tumours of various grades (Buerger et al., 1999; James et al., 1997), whilst the latter focused specifically on high grade DCIS (Moore et al., 1999). As in the infiltrating tumours gains on chromosome 17 and 1q were frequently detected. Gains on chromosome 19 were also detected in 2 of the DCIS studies and one study had a frequent gain at 8q, which was seen in all infiltrating breast tumour studies. In general there were more regions of frequent amplification detected in the high grade DCIS as opposed to the DCIS tumours of various grade. Three regions of frequent amplification were exclusive to the high grade DCIS, 6p, 12q and 22q, whilst the other regions detected were also found in at least one of the studies that analysed infiltrating carcinomas. 13q contained regions of chromosomal loss in all DCIS studies and losses were also frequently found in 2 of the studies of DCIS at 9p, 14q and 16q.
### Table 6.2. Chromosomal Changes in Different Types of Invasive Carcinoma

#### INFILTRATING LOBULAR CARCINOMA

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q</td>
<td>15</td>
<td>79</td>
<td>16q</td>
<td>12</td>
<td>63</td>
</tr>
<tr>
<td>17q</td>
<td>5</td>
<td>26</td>
<td>17p</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>8q</td>
<td>4</td>
<td>21</td>
<td>6q</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13q</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Xq</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18q</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Xp</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4q</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4p</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11q</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8p</td>
<td>3</td>
<td>16</td>
</tr>
</tbody>
</table>

Nishizaki et al., 1997  
N = 19

#### INFILTRATING DUCTAL CARCINOMA

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q</td>
<td>26</td>
<td>57</td>
<td>13q</td>
<td>18</td>
<td>39</td>
</tr>
<tr>
<td>8q</td>
<td>23</td>
<td>50</td>
<td>17p</td>
<td>16</td>
<td>35</td>
</tr>
<tr>
<td>17q</td>
<td>17</td>
<td>37</td>
<td>16q</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>20q</td>
<td>13</td>
<td>28</td>
<td>Xq</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>11q</td>
<td>10</td>
<td>22</td>
<td>18q</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Xp</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6q</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11q</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8p</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3p</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5q</td>
<td>12</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1p</td>
<td>11</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4p</td>
<td>7</td>
<td>15</td>
</tr>
</tbody>
</table>

Nishizaki et al., 1997  
N = 46

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q</td>
<td>28</td>
<td>70</td>
<td>16q</td>
<td>26</td>
<td>65</td>
</tr>
<tr>
<td>16p</td>
<td>18</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q</td>
<td>12</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Roylance et al., 1999  
N = 40  
Grade I well differentiated

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q</td>
<td>35</td>
<td>70</td>
<td>8p</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>8q</td>
<td>34</td>
<td>68</td>
<td>11q</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>17q</td>
<td>23</td>
<td>46</td>
<td>13q</td>
<td>21</td>
<td>42</td>
</tr>
<tr>
<td>20q</td>
<td>19</td>
<td>38</td>
<td>1p</td>
<td>18</td>
<td>36</td>
</tr>
<tr>
<td>16p</td>
<td>17</td>
<td>34</td>
<td>18q</td>
<td>16</td>
<td>32</td>
</tr>
</tbody>
</table>

Roylance et al., 1999  
N = 50  
Grade III poorly differentiated

---

Table 6.2. Legend: Frequent regions of amplification and loss in different infiltrating breast cancers, which are present in literature.
Table 6.3. Regions of Chromosomal Alterations in DCIS

<table>
<thead>
<tr>
<th>Chromosome gains</th>
<th>Chromosome losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr. Region</td>
<td>n</td>
</tr>
</tbody>
</table>

---

**DUCTAL CARCINOMA in SITU**

- **Chromosome gains**
  - 17q21: 2, 22% (James et al., 1997, n = 9)
  - 19q: 2, 22%
  - 20p: 2, 22%

- **Chromosome losses**
  - 16q: 2, 22%
  - 13q: 2, 22%

---

**Chromosome gains**

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
</table>
| 1q          | 23| 60%
| 17q         | 11| 29%
| 8q          | 8 | 21% |

**Chromosome losses**

- 16q: 24, 62%
- 8q: 10, 26%
- 13q: 10, 26%
- 14q: 10, 26%
- 9p: 6, 15%
- 17p: 6, 15%

(Buerger et al., 1999, n = 38)

**Mean number aberrations 5.6**

**Different grades DCIS**

---

**Chromosome gains**

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
</table>
| 11q13       | 22| 96%
| 1p          | 18| 78%
| 19          | 18| 78%
| 17          | 12| 52%
| 22q         | 12| 52%
| 16p         | 13| 57%
| 6p21.3      | 11| 48%
| 12q24       | 10| 44%
| 20q         | 10| 44%
| 1q31        | 6 | 26% |

**Chromosome losses**

- 13q: 7, 30%
- 9p21: 4, 17%
- 6q16.1: 4, 17%

(Moore et al., 1999, n = 23)

**All high grade DCIS**

---

Table 6.3. Legend: Regions of chromosomal alteration in DCIS detected in literature. The chromosome region, number and percentage of tumours samples with the aberration are shown. Only regions that were present in more than 15% of tumour samples are shown. The reference, total number of tumour samples analysed and the mean number of aberrations detected in the study is on the right.
6.1.3. Genomic Alterations Associated with Poor Prognosis and Metastasis

Chromosome gains at 8q and 20q12-q13 have been associated with recurrence of disease and poor clinical outcome (Isola et al., 1995). Other gains associated with poor prognosis include 11q13 and 17q and it has been suggested that these gains could be late, progression related events and are associated with an aggressive clinical phenotype (Hermsen et al., 1998).

An important prognostic marker in breast cancer is the presence or absence of axillary lymph node metastases. Several studies have analysed genetic changes in these metastasising tumours. There were some frequent chromosome gains common to both metastasising and their primary tumours including 1q, 8q and 17q (Adeyinka et al., 1999; Kuukasjarvi et al., 1997; Nishizaki et al., 1997). However chromosome losses at 6q and 16q are more frequent in lymph node negative tumours (Adeyinka et al., 1999), and so have been suggested as early changes in breast carcinoma. Furthermore allelic losses at 16q23.2-q24.2 have been identified as an independent marker of good prognosis in primary breast cancers (Hansen et al., 1998) and tumours with losses at 16q24.3 have a lower potential to invade and migrate (Driouch et al., 1997), so it has been postulated that loss of a growth inducing or metastasis gene might have occurred on 16q. In contrast losses on chromosome 18 were detected more frequently in lymph node positive tumours (Adeyinka et al., 1999) or in lymph node metastasis compared to their corresponding primary tumour (Nishizaki et al., 1997). A number of genes located on chromosome 18 including DCC and members of the SMAD gene family have been implicated as tumour suppressor genes (Adeyinka et al., 1999). A region of significant chromosome gain in metastasising tumours has been located in the Xq12-q22 region (Kuukasjarvi et al., 1997).

Thus there appear to some regions of chromosomal gain and loss that are common between different types of tumours with different phenotypic properties and there are other regions, which seem to occur more specifically.

6.1.4. Chapter Objectives

In the previous chapter cell lines were analysed for chromosomal amplification and this chapter focuses on clinical in vivo breast tumour samples. CGH has been used extensively to study regions of possible chromosomal amplifications or deletions. However CGH does not have the ability to identify specific genes involved or elucidate small regions of amplification (Section 1.6.). Therefore in this study I have used cDNA microarray to analyse specific gene amplification events.
Genomic alterations in tumour cells are the basis of phenotypic characteristics and of the biological behaviour of cells. Thus genetic analysis may aid in our understanding of the genetic background associated with these phenotypic changes and offers prognostic information. A variety of different grade and types and hormone status breast tumours were analysed, with the objective being to link chromosomal aberrations with phenotypic characteristics.
6.2. Results
The chromosomal position of all amplified genes was retrieved from the Unigene or Locuslink NCBI databases (http://www.ncbi.nlm.nih.gov). All experiments were in duplicate or triplicate and genes were only considered amplified if they had a fold change of greater than 2 fold in all experiments and if they were significantly amplified by means of a T-test. A 2-fold cut off was chosen, as these experiments were analysing chromosomal amplification therefore unlike RNA where there can be 1.3 or 1.8 fold changes, amplifications of both alleles must be whole number changes or single alleles may be half or whole number changes.

Where available the non tumour control was a pooled sample of DNA extracted from surround non tumour tissues from the patients, although non tumour clinical material was not always available, as during breast conservation surgery the removal of surround tissue is kept to a minimum. Therefore the non tumour HBL100 cell line was also used, as seen in the previous chapter HBL100 was a good non tumour control to use, as most amplifications detected against HBL100 were also seen against the pooled in vivo DNA sample.

6.2.1. Overview of Gene Amplification
The regions of amplification detected in the 13 clinical breast tumours as compared to a non tumour control are summarised Figure 6.1. Each chromosome is represented by its ideogram and the coloured bars to the right of each chromosome represent the regions of amplification in the different clinical breast tumours compared to the non tumour controls. The chromosome regions containing more than one amplified gene are represented by thicker bars. Two chromosomes, 21 and Y contained no regions of amplification, which is in concurrence with the breast cancer cell lines in the previous chapter (Chapter 5).

There were some areas of common chromosomal amplification that occurred in the different clinical tumour samples, the most frequent of which were 1q, 8q 12q and 13q. Areas containing more than one amplified gene were seen in different tumour samples in several chromosomes. One such region was on the long arm of chromosome 17, which due to the high frequency of amplification in both the cell lines (Chapter 5) and clinical tumour samples, is discussed in greater depth in chapter 7.

It is clear from the figure that the clinical breast tumour sample 13, which is a pre menopausal lymph node positive IDC grade III had a high frequency of genomic amplification, as regions in many of the chromosomes contain single or multiple amplified genes.
Figure 6.1. Regions of Gene Amplification Detected in Clinical Tumour Samples

Continued over page
<table>
<thead>
<tr>
<th>Tumour Sample Number</th>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Single gene amplified in region</td>
</tr>
<tr>
<td>02</td>
<td>Multiple genes amplified in region</td>
</tr>
<tr>
<td>03</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td></td>
</tr>
<tr>
<td>05</td>
<td></td>
</tr>
<tr>
<td>06</td>
<td></td>
</tr>
<tr>
<td>07</td>
<td></td>
</tr>
<tr>
<td>08</td>
<td></td>
</tr>
</tbody>
</table>

**Figure:**

[Image of chromosome bands with color-coding for Tumour Sample Numbers 1 to 21, and X and Y chromosomes marked with green and purple bands, respectively.]
Figure 6.1. Legend: Regions of amplification detected in the 13 clinical breast tumour samples. The ideograms of each chromosome are shown and the number corresponds to the chromosome. Each in vivo clinical sample (01 - 13) is represented by a different colour as indicated and the corresponding bars to the right of the chromosomes represent regions containing amplified genes. All the clinical samples were compared against either a pooled DNA from surround tissue or the HBL100 breast cell line. Regions containing a single gene that was amplified are represented by thin bars, whereas regions containing more than one amplified gene are represented by thick bars.

6.2.2. Specific Regions of Gene Amplification

Specific chromosome regions that contained amplifications are summarised in Table 6.4. Regions containing several amplified genes are shown in bold, although the majority of amplifications detected by the CGH microarray were of a single gene in a separate region, thus suggesting single gene amplifications, however it should be remembered that genes neighbouring the amplified genes may not be present on the current microarray.

Tumour samples 02, 05 11, 12 and 13 contained the most regions of chromosomal amplification. Tumour 13 had substantially more regions containing amplified genes than all other samples, with 82 different regions 32 of which contained more than one gene. The average number of regions detected across all samples is 15, although if tumour 13 containing many regions is discounted the number drops to 8. Six of the tumour samples contained regions that have more than one amplified gene.
### Table 6.4. Specific Regions Containing Gene Amplification in Clinical Tumours

<table>
<thead>
<tr>
<th>Tumour</th>
<th>Chromosome regions where genes are amplified</th>
<th>No. Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>lq41.1, 7q11.2, 8q24.3, 10q26.11, 12q14, 13q14.1-q14.2, 16p13.3</td>
<td>8 (0)</td>
</tr>
<tr>
<td>02</td>
<td>1p33, 1q25.3, 1q42.1, 2p25.1, 8q24.3, 12q23, 13q14.1-q14.2, 16p13.3, 17q11.2-q12, <strong>17q12</strong>, 22q11.21</td>
<td>11 (1)</td>
</tr>
<tr>
<td>03</td>
<td>lq44.4, 8p21, 8q24.3, 12q23, 13q14.1-q14.2, 14q32.33, 16p13.3</td>
<td>7 (0)</td>
</tr>
<tr>
<td>04</td>
<td>1q42.1, 8q24.3, 12q23, 13q14.1-q14.2, 16p13.3</td>
<td>5 (0)</td>
</tr>
<tr>
<td>05</td>
<td><strong>1q21</strong>, 1q42.1, 3q26.31, 4p14, 5q23, 8q24.3, 9q34.11, 10q11.21-q11.22, 12q23, 13q14.1-q14.2, 14q32.31, 15q21-q22.2, 19p13.3, 19q13.1-q13.2, 20q13</td>
<td>16 (1)</td>
</tr>
<tr>
<td>06</td>
<td>8q24.3, <strong>12q23</strong>, 12q23-q24, 13q14.1-q14.2, 19p13.1, 19p13.3</td>
<td>6 (1)</td>
</tr>
<tr>
<td>07</td>
<td>8q24.3, 12q23, 13q14.1-q14.2, 19p13.3</td>
<td>4 (0)</td>
</tr>
<tr>
<td>08</td>
<td>8q24.3, 12q23, 13q14.1-q14.2</td>
<td>6 (1)</td>
</tr>
<tr>
<td>09</td>
<td>8q24.3, 12q23, 13q14.1-q14.2</td>
<td>2 (0)</td>
</tr>
<tr>
<td>01</td>
<td>2p25.1, 9q31.2, 19q13.1-q13.4, 22q11.21</td>
<td>4 (0)</td>
</tr>
<tr>
<td>02</td>
<td>1q31-q32, 2p25, 2q33-q34, <strong>3p21</strong>, 4q25-q31, 5q12.1, 6p21.1, <strong>7q26</strong>, 8p22-p21, 8q22.1, 9q34, 10q21.1, 11pter-15.5, 11q22.2-q22.3, 12q11-q12, 12q13, 15q24, <strong>17q11-q21</strong>, 20p13, Xp11.3-p11.23</td>
<td>20 (3)</td>
</tr>
<tr>
<td>03</td>
<td>1q25.3, 3p21, <strong>5q13</strong>, 5q13.1-q15, 7p14-p13, 7q36, 8p22-p21, 11p15.5-p15.3, <strong>12q24</strong>, 17q11-q21.3, 17q11.2-q12, <strong>17q12</strong>, 17q23, 20p13</td>
<td>14 (3)</td>
</tr>
<tr>
<td>04</td>
<td><strong>1p31.2-p31.1</strong>, <strong>1p33</strong>, <strong>1p34.1</strong>, 1p36.3-p34.1, 1q21, <strong>1q31-q32</strong>, 1q32-q41, <strong>1q42.1</strong>, 1q42-q43, <strong>2p21</strong>, 2p23-p21, <strong>2p25</strong>, 2q11.2, 2q33-q34, 2q35, 2q36, 2q37.3, <strong>3p21</strong>, 3p21-p14, <strong>3q21.3</strong>, 3q21-q23, 4p16, <strong>4q24</strong>, 4q25-q31, 5q12.1, 5q14, 5q22-q23, <strong>5q31</strong>, 5q33.1, 5q35.3, <strong>6p12.3</strong>, 6p21.2-p12, 6p21.3, 6p24.3, 6q16.1, 6q22, 7p12, 7p13, <strong>7p15.1</strong>, 7p15-p14, 7p22.2, 17q21-q22, <strong>17q32</strong>, <strong>17q36</strong>, <strong>8p21</strong>, 8p22-p21, 8q21, 8q22.1, <strong>9q22</strong>, 9q22-32-q31.3, <strong>9q34</strong>, 10p13, 10q21.1, <strong>10q23-q26</strong>, 11p11.2, <strong>11p15.5</strong>, 11q13, <strong>11q22</strong>, 11q23.2-q23.3, 12p12.1, 12q11-q12, <strong>12q13-q14</strong>, <strong>12q22</strong>, 12q22-q24.2, 13q12-q14, <strong>13q14</strong>, 13q32, 14q32.33, 15q24, <strong>16p13.2</strong>, 16q13, <strong>16q22</strong>, <strong>17p13</strong>, <strong>17q11-q12</strong>, 17q21, 17q25, 18q12.3, <strong>18q21</strong>, <strong>19p13</strong>, 19q13.2, 20p13, 20q11.2-q13.1, <strong>20q13</strong>, 22q11.21, 22q13.1, <strong>Xp11.23</strong>, Xp11.3-p11.23, Xp22.2, <strong>Xq28</strong></td>
<td>89 (32)</td>
</tr>
</tbody>
</table>

Table 6.4. Legend: A summary of all the regions of amplification detected in the 13 clinical breast tumour samples. The tumour sample is indicated in the left column, whilst the amplification regions are listed in the middle. Those written in bold are regions which contain more than one amplified gene for that specific tumour. The total number of regions containing single or multiple amplifications (bold) for each cell line are shown in the right column.
6.2.3. Frequent Regions of Gene Amplification in Clinical Tumours and Phenotype

Table 6.5 summarises the most frequent areas of gene amplification that were detected across the breast tumour samples, all chromosome arms that contain amplification in greater than 23% (3) of samples are shown. The table summarises the regions of gain that were associated with lymph node status, whether the patients are pre or post menopausal, the different grade of tumour and the oestrogen receptor status.

Table 6.5. The Most Frequent Gene Amplification Regions

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>Total No.</th>
<th>Lymph</th>
<th>Menopausal</th>
<th>Grade</th>
<th>ER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+</td>
<td>-</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>8q24.3</td>
<td>9 69</td>
<td>5 4</td>
<td>2 4 3</td>
<td>1 4 3 1</td>
<td>4 5 0</td>
</tr>
<tr>
<td>12q23</td>
<td>9 69</td>
<td>6 3</td>
<td>3 3 6</td>
<td>0 4 4 1</td>
<td>8 1 0</td>
</tr>
<tr>
<td>13q14</td>
<td>9 69</td>
<td>5 4</td>
<td>3 3 6</td>
<td>1 4 3 1</td>
<td>4 4 1</td>
</tr>
<tr>
<td>1q42</td>
<td>5 39</td>
<td>3 2</td>
<td>2 1 2</td>
<td>1 1 2 1</td>
<td>3 1 1</td>
</tr>
<tr>
<td>16p13</td>
<td>5 39</td>
<td>3 2</td>
<td>1 2 2</td>
<td>1 2 2 0</td>
<td>2 2 1</td>
</tr>
<tr>
<td>2p25</td>
<td>4 31</td>
<td>4 0</td>
<td>3 0 1</td>
<td>0 3 1 0</td>
<td>1 1 2</td>
</tr>
<tr>
<td>17q11-q12</td>
<td>4 31</td>
<td>3 1</td>
<td>3 0 1</td>
<td>0 2 2 0</td>
<td>1 0 3</td>
</tr>
<tr>
<td>19p13</td>
<td>4 31</td>
<td>2 2</td>
<td>2 1 1</td>
<td>0 2 1 1</td>
<td>2 1 1</td>
</tr>
<tr>
<td>3p21</td>
<td>3 23</td>
<td>2 1</td>
<td>3 0 0</td>
<td>0 1 2 0</td>
<td>0 1 2</td>
</tr>
<tr>
<td>7q36</td>
<td>3 23</td>
<td>2 1</td>
<td>3 0 0</td>
<td>0 1 2 0</td>
<td>0 1 2</td>
</tr>
<tr>
<td>8p21</td>
<td>3 23</td>
<td>1 2</td>
<td>2 1 0</td>
<td>0 1 2 0</td>
<td>0 2 1</td>
</tr>
<tr>
<td>9q34</td>
<td>3 23</td>
<td>2 1</td>
<td>3 0 0</td>
<td>0 1 1 1</td>
<td>1 0 2</td>
</tr>
<tr>
<td>11p15</td>
<td>3 23</td>
<td>2 1</td>
<td>3 0 0</td>
<td>0 1 2 0</td>
<td>0 1 2</td>
</tr>
<tr>
<td>14q32</td>
<td>3 23</td>
<td>1 2</td>
<td>2 1 0</td>
<td>0 1 1 1</td>
<td>2 0 1</td>
</tr>
<tr>
<td>19q13</td>
<td>3 23</td>
<td>2 1</td>
<td>3 0 0</td>
<td>0 1 1 1</td>
<td>1 1 1</td>
</tr>
<tr>
<td>20p13</td>
<td>3 23</td>
<td>2 1</td>
<td>3 0 0</td>
<td>0 1 2 0</td>
<td>0 1 2</td>
</tr>
<tr>
<td>22q11</td>
<td>3 23</td>
<td>3 0</td>
<td>3 0 0</td>
<td>0 2 1 0</td>
<td>1 1 1</td>
</tr>
</tbody>
</table>

Lymph = lymph node tumours; Grade = tumour grade; ER = oestrogen receptor
+ = positive; - = negative; U = undetermined

There were some regions containing amplified genes, which did not appear to be related to specific phenotypic properties of the tumours, however there were others that may be related to phenotype. The 3 most frequent regions of amplification, 8q24.3, 12q23 and 13q14 were present in 9 tumours regardless of lymph node, menopause, grade or ER status; although the 12q23 region was found predominately in the ER+ tumours. These regions and the different phenotypes are clearly seen in the graphs in Figure 6.2.
Figure 6.2. Legend: Summarises the number of chromosome regions containing amplifications for the different phenotypes. The x-axis is the chromosome region and the y-axis is the number of tumours. The proportion of the bar represents the number of tumour. Each different characteristic is summarised in separate graphs: Graph A represents the lymph node negative and positive tumours. Graph B represents pre and post menopausal patients. Graph C is data for the grade of tumour and Graph D shows the regions of amplification in relation to the tumour oestrogen receptor status. If the tumour characteristic is undetermined it is represented by grey bars.
It is clear from Figure 6.2.A that there were 2 regions containing amplification specific to lymph node positive tumours, 2p25 and 22q11, whereas all frequent regions seen in lymph node negative tumours were also in lymph node positive tumours. There were several regions containing amplified genes that were only seen in the premenopausal patients, 3p21, 7q36, 9q34, 11p15, 17q11-q12, 19q13, 20p13, 22q11 and 2p25, the latter 2 regions were also found in lymph node positive tumours only. There were some tumours included that are of an undetermined menopausal status, thus care should be taken in interpreting these data. Figure 6.2.C represents the regions of frequent amplification when compared to the grade of the tumour. Four of the regions were found in all tumour grades, while 13 were found in Grade II and III tumours only, although for some of the tumours the grade of the tumour was undetermined. The ER status and regions of common amplification are shown in Figure 6.2.D. Some of the regions containing frequent amplification were detected in tumours regardless of ER status, and some were seen only in ER+ or ER- tumours, although many tumours in this category were also undetermined.

6.2.4. Specific Genes Involved in Frequent Regions of Amplification

There were 15 genes amplified in at least 3 of the clinical tumour samples (23 %), which are summarised in Table 4.2. These included 3 genes, ERBB2, GRB7 and PPARBP that are located at 17q12, which were amplified in tumour samples 02, 11, 12 and 13. Another gene located at 17q12 is CrkS, which was also amplified in samples 11, 12 and 13, whether this gene was amplified in tumour 02 is not known, as the Crks gene was a latter addition to the MRC toxicology microarrays and tumour 02 was processed on earlier arrays that did not contain this gene. Due to the high frequency of amplification detected in this region of chromosome 17 in these clinical tumour samples and in the breast cell line BT474 it is studied in greater detail in chapter 7.

Interestingly 2 of the genes commonly amplified in the clinical tumour samples were also frequently amplified in the breast cancer cell lines (Chapter 5). The gene most commonly amplified was the KIAA0014 gene located at 8q24.1, which was amplified in 9 clinical tumour samples, and it was also amplified in 2 of the breast cell lines, MCF7/wt and T47D. SLC25A3 was amplified in 8 clinical tumours and was also amplified in the breast cancer cell lines MCF7/wt and T47D.
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Chr. Location</th>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSNAX</td>
<td>1q42.1</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>LPIN</td>
<td>2p25.1</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
</tr>
<tr>
<td>ABCB8</td>
<td>7q36</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>KIAA0014</td>
<td>8q24.3</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CDK4</td>
<td>12q14</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SLC25A3</td>
<td>12q23</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ESD</td>
<td>13q14.1-q14.2</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>UBE21</td>
<td>16p13.3</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ERBB2</td>
<td>17q12</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GRB7</td>
<td>17q12</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PPARBP</td>
<td>17q12</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CrkS</td>
<td>17q12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CDC34</td>
<td>19p13.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CDC25B</td>
<td>20p13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DGS1</td>
<td>22q11.21</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

A = the gene was amplified in this clinical tumour sample
6.4. Discussion

The mean number of regions containing amplified genes was 15, although if sample 13, which contained a substantially higher number of regions than the other tumours is discounted then the average number of regions containing amplified genes is 8, which is in agreement to numbers detected in literature (Tirkkonen et al., 1998). In contrast the mean number of aberrations detected in the breast cancer cell lines was 17 (Chapter 5). The higher number of chromosome aberrations seen in cell lines have been explained by the suggestion that cell lines are a genetically advanced population (Kallioniemi et al., 1994). Cells in culture are also free from some growth restraints found in surrounding cells in vivo and so can undergo clonal expansion at an increased rate and thus may undergo increased chromosomal aberrations.

Tumour 13 is a premenopausal IDC grade III with positive lymph node metastasis, thus it is considered an advanced tumour and so one would expect to see a high frequency of chromosome aberrations. However tumour 13 contained 89 separate regions of amplification, which is also substantially higher than other clinical samples and the breast cell lines. The regions of amplification were present in most arms of the chromosomes; one explanation for this is that the tumour may be aneuploid and a disadvantage of microarray CGH is that whole chromosome gains or chromosome arm gains are not easily detected. Although if duplication of whole chromosomes had occurred one would expect to see a greater number of individual genes amplified. In general amplified genes that were detected by microarray CGH in this study had a fold change in the range of 2 - 6, whereas in tumour 13 fold changes of 6 - 23 were detected in 25 genes. The high numbers of genes amplified and the high fold changes detected in this grade III invasive tumour suggests that it is advanced with a high degree of genetic instability.

6.3.1 Frequent Regions of Gene Amplification in Clinical Breast Tumours

Many gene amplifications occurred in a single tumour sample. However there were 17 regions that were present in greater then 3 (23 %) tumour samples (Table 6.5.). Some of these chromosome arms are also frequently detected in breast cancer in the literature. Findings from this study and literature studies that have analysed primary tumours are summarised in Table 6.7.
Table 6.7. Comparison of Frequent Regions of Amplification with Literature

<table>
<thead>
<tr>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
<th>Chr. region</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>8q</td>
<td>9</td>
<td>69</td>
<td>8q</td>
<td>27</td>
<td>49</td>
<td>8q</td>
<td>17</td>
<td>59</td>
</tr>
<tr>
<td>1q</td>
<td>5</td>
<td>39</td>
<td>1q</td>
<td>37</td>
<td>67</td>
<td>1q</td>
<td>19</td>
<td>66</td>
</tr>
<tr>
<td>17q</td>
<td>4</td>
<td>31</td>
<td>17q</td>
<td>10</td>
<td>18</td>
<td>17q</td>
<td>9</td>
<td>31</td>
</tr>
<tr>
<td>16p</td>
<td>5</td>
<td>39</td>
<td>16p</td>
<td>21</td>
<td>38</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19p</td>
<td>4</td>
<td>31</td>
<td>19p</td>
<td>10</td>
<td>18</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19q</td>
<td>3</td>
<td>23</td>
<td>19q</td>
<td>11</td>
<td>20</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3q</td>
<td>9</td>
<td>16</td>
<td>3q26-qter</td>
<td>6</td>
<td>21</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12q</td>
<td>9</td>
<td>69</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13q</td>
<td>9</td>
<td>69</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2p</td>
<td>4</td>
<td>31</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3p</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7q</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8p</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9q</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11p</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14q</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20p</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22q</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5p</td>
<td>13</td>
<td>24</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11q</td>
<td>10</td>
<td>18</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20q</td>
<td>10</td>
<td>18</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6q22</td>
<td>6</td>
<td>21</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7p21-pter</td>
<td>6</td>
<td>21</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Three of the chromosome arms 8q, 1q and 17q that contained regions of gain are detected in all studies. Chromosomal gains at 8q and 17q are also associated with pancreatic cancer (Mahlamaki et al., 1997), and are frequently detected in human breast cancer cell lines (Chapter 5; Forozan et al., 2000; Larramendy et al., 2000) and gains of regions on chromosome 8q and 1q are known to occur in endometrial cancers (Suzuki et al., 1997). The gene primarily amplified in the 8q region in this study was KIAA0014, which was also amplified in the breast cell lines MCF7 and T47D. KIAA0014 is a hypothetical protein and is not well characterised, but interestingly this gene was also overexpressed in the cell lines and so it could have a direct role in breast cancer. The cytogenic position of the KIAA0014 gene is 8q24.3, and an oncogene commonly amplified in some cancers located at 8q24 is myc (Gray and Collins, 2000), although this gene is not considered significantly amplified in these
experiments. The frequent gains at chromosome 17q detected in this study in both the cell lines and clinical tumour samples involved the amplification of many genes, consequently this chromosome and its possible role in breast cancer is analysed in greater depth in chapter 7. Since chromosomal gains along 8q, 1q and 17q occur at high frequency in different tumours, it is likely that genes within these regions are providing a critical advantage for tumour survival and progression.

Chromosomal gains at 16p, 19p and 19q were also detected frequently in this study and in that by Tirkkonen et al., 1998 and gains at 3q were detected in the two literature studies but not here. Other frequent gains detected were specific to individual studies, although the regions of 13q amplification detected in this study have been previously detected in breast cancer, and are associated with poor prognosis (Hermsen et al., 1998).

Regions containing amplified genes detected in this study at 7q, 12q and 14q were also seen frequently in the human breast cell lines in chapter 5 and to a lesser extent (11%) in literature (Tirkkonen et al., 1998). The SLC25A3 gene located at chromosome 12q23 was amplified in the MCF7/wt and T47D cells and in 8 clinical samples in this study. This gene encodes a phosphate carrier, which is involved in transport of phosphate groups from the cytosol to the mitochondrial matrix (Jabs et al., 1994). Although it was amplified with high frequency the gene was not overexpressed in vitro or in vivo, thus its involvement in breast cancer is questionable and its frequent amplification could be due to neighbouring genes involved in breast cancer that have undergone amplification.

6.3.2. Region of Gene Amplification and Tumour Phenotype

The associations between the regions of amplification and tumour phenotype were shown in Table 6.5 and Figure 6.2. (Section 6.2.3.).

6.3.2.1. Region of Amplification Detected and Tumour Type

Four different tumour types were included in this study, DCIS, Tubular, ILC and IDC. DCIS tumours of different grade have been studied for regions of chromosomal amplification (Buerger et al., 1999; James et al., 1997), whilst other studies have focused on high grade DCIS (Moore et al., 1999). Only 1 DCIS tumour was included in this study (sample 05), the grade of which was undetermined, thus making comparison difficult. Although frequent regions of amplification detected in DCIS tumours in literature include 1q and 8q (Buerger et al., 1999) and both of these were detected in the DCIS tumour. There is data to suggest that
DCIS is a precursor legion to invasive breast carcinoma (James et al., 1997), however this could not be determined in this study due to the lack of DCIS samples.

Regions commonly detected in ILC tumours in literature include 1q, 8q and 17q (Nishizaki et al., 1997). In the study here only 1 ILC sample was available, thus making a comparison difficult, although amplification was detected on the long arm of chromosome 8. Other regions detected in the ILC tumour in this study include 1p34.3, 12q23, 13q14 and 16p13.3.

The majority of tumours in this study were IDC grade II or III. Several studies have been carried out to determine regions of chromosomal amplification in IDC (Roylance et al., 1999; Nishizaki et al., 1997). The common regions of frequent chromosomal gain found in these literature studies and in this report are shown Table 6.8.

**Table 6.8. Common Regions of Frequent Gene Amplifications for IDC**

<table>
<thead>
<tr>
<th>Chr. Region</th>
<th>This Study</th>
<th>Roylance et al., 1999 (grade III)</th>
<th>Nishizaki et al., 1997</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>Chr. Region</td>
</tr>
<tr>
<td>8q24.3</td>
<td>6</td>
<td>60</td>
<td>8q</td>
</tr>
<tr>
<td>17q11-q12</td>
<td>4</td>
<td>40</td>
<td>17q</td>
</tr>
<tr>
<td>1q42</td>
<td>3</td>
<td>30</td>
<td>1q</td>
</tr>
<tr>
<td>16p13</td>
<td>3</td>
<td>30</td>
<td>16p</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20q</td>
</tr>
</tbody>
</table>

The common regions containing frequent chromosome amplification in IDC include the 8q, 17q and 1q regions that have been previously discussed. 16p also contains genes that were frequently amplified in this study and also by Roylance et al. (1999). The latter of which is a study that analysed grade III IDC tumours, in this study all IDC tumours were grade II or grade III and so amplifications on chromosome 16p may be linked to less well differentiated breast tumours. However the clinical tumour sample 01, a tubulin grade I breast tumour and grade I well differentiated IDC tumours (Roylance et al., 1999) also contained amplification at this site. 20q amplifications are frequently detected in the literature studies, but not in the clinical samples here, although they were frequently detected in the breast cancer cell lines (Chapter 5).
6.3.2.2. Region of Amplification and Lymph Node Status

The lymph node status of all tumours studied was determined, 5 were lymph node negative and 8 were lymph node positive. A CGH study on 53 lymph node negative tumours found frequent regions of chromosome gain at 8q, 1q, 4q, 5q and Xq and in all cases that had 1q gains 8q gains were also present, but gains of 8q could be found without 1q (Hermsen et al., 1998). This was also true in findings here, which suggests that gains at 8q occur before, and are necessary, for 1q gains in lymph node negative tumours. In DCIS development the model suggested by Buerger et al., (1999) (Figure 1.2) hypothesised that chromosomal amplification of 1q is a step involved in the transition of DCIS well-differentiated to DCIS intermediately-differentiated, thus if 8q gains occur prior to 1q they may be an early change involved in the development of well-differentiated DCIS. However some CGH studies have suggested that 1q gain is an early event, as it has been reported in genetically less complex tumours and has also been recorded as the sole genetic event (Tirkkonen et al., 1998) and in some studies it is the most frequent amplification event (Roylance et al., 1999; Nishizaki et al., 1997). However these studies were not restricted to lymph node negative tumours and did not differentiate the amplification events on the basis of lymph node status.

Two regions of frequent amplification, 2p25 and 22q11 occurred in lymph node positive tumours but not lymph node negative tumours. Two genes were consistently found amplified in these regions, LPIN1 at 2p25 and DGSI at 22q11. 22q11 chromosomal rearrangements are known to occur, and are known to be responsible for a number of genomic disorders (McDermid and Morrow, 2002); the DGSI gene is deleted in DiGeorge syndrome, which is associated with learning disability and cleft palate appearance (McDermid and Morrow, 2002). However amplifications of 2p25 or 22q11 have not previously been recorded in breast cancer. Interestingly although LPIN1 and DGSI are located on different chromosomes neither gene was amplified alone (Table 6.6.). Both these genes were also amplified and overexpressed in the BT474 breast cell line. BT474 cells are highly invasive, and as these genes are amplified and overexpressed in BT474 and amplified in the clinical samples that are lymph node positive these genes may play a role in breast cancer metastasis.

6.3.2.3. Region of Amplification, ERBB2 and Menopausal Status

There also appeared to be an association with LPIN1 and DGSI amplification with ERBB2 (or HER-2) and menopausal status. Clinical samples 12 and BT474 breast cells were known to be ERBB2 positive, while the other clinical samples were undetermined. Both samples 12 and BT474 were found to have amplification of the ERBB2 gene at 17q11.2-q12, thus confirming
their ERBB2 positive status. In three other clinical tumours 02, 11 and 13 ERBB2 was also amplified and so were also considered ERBB2 positive, while tumour 10 appeared negative (although it was assigned undetermined as an average fold change of 1.8 was detected, thus it may be amplified a single allele in the tumour sample, but it is not over the 2-fold change used here). This ability to potentially screen breast tumours for hormonal positive status by microarray CGH is a good advantage. Clinical samples 10, 11, 12 and 13 were all premenopausal and 02 is undetermined. Amplification of LPIN1 and DGSI was only found in clinical samples 02, 10, 13 and BT474, this data is summarised in Table 6.9.

Table 6.9. Lymph Node and ERBB2 Status with DGSI and LPIN1 Amplification

<table>
<thead>
<tr>
<th>Tumour sample</th>
<th>ERBB2 Status</th>
<th>Lymph Node</th>
<th>Menopause</th>
<th>DGSI</th>
<th>LPIN1</th>
</tr>
</thead>
<tbody>
<tr>
<td>02</td>
<td>+</td>
<td>+</td>
<td>U</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>U</td>
<td>+</td>
<td>Pre</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>+</td>
<td>+</td>
<td>Pre</td>
<td>nc</td>
<td>nc</td>
</tr>
<tr>
<td>12</td>
<td>+</td>
<td>-</td>
<td>Pre</td>
<td>nc</td>
<td>nc</td>
</tr>
<tr>
<td>13</td>
<td>+</td>
<td>+</td>
<td>Pre</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BT474</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

+ = positive or amplified; - = negative; ? = unknown; U = undetermined; nc = no change

Therefore DGSI and LPIN amplification occurred predominantly in lymph node and ERBB2 positive tumours. And depending on the menopausal status of tumour 02 and BT474 the amplification of ERBB2, DGSI and LPIN1 could be associated with premenopausal tumours.

As well as 2p25, 17q11-q12 and 22q11, other regions containing gene amplification were found only in premenopausal women, including 3p21, 9q34, 11p15, 19q13 and 20p13. Interestingly regions 11p15 and 20p13 were also only observed in premenopausal ERBB2 positive tumours, although not all ERBB2 positive premenopausal women contained these amplifications.

6.3.2.4. Region of Amplification and Tumour Grade

The regions of gene amplification detected appeared to be independent of tumour grade, as there were no regions specific to tumour grade. Some regions were detected in all grade tumours (I – III), 8q24.3, 13q14, 1q42 and 16p13, while others were only detected in tumour grades II – III or undetermined. However it cannot be concluded that the amplifications that
occurred only in grade II and III tumours are late changes in tumour development, as only 1 grade I tumour was present in the study making significant conclusions difficult. Although, as high-grade tumours generally have large pleomorphic nuclei, it has been suggested that such morphological characteristics reflect the presence of profound underlying genomic DNA abnormalities (Moore et al., 1999).

6.3.2.5. Region of Amplification and ER Status
The data suggested that some regions of amplification are related to ER status, however for some tumours ER status was undetermined. Although the region at 12q23 occurs in a high frequency in the ER positive tumours (8/9), the gene involved at this site is the SLC25A3 gene, which as seen before was also amplified in MCF7/wt and T47D, both of which are also ER positive. Thus the high frequency of amplification of this gene in ER positive cells could be significant.

6.3.3. Problems Associated With Clinical Tumour Samples
Several problems were encountered in the acquisition of clinical samples. Due to ethical or sheer lack of numbers clinical samples were not readily available. Where possible a broad selection of tumours was attempted, but only a single DCIS, ILC and grade I tumour was available, which made comparison of these tumour types difficult. Some of the clinical tumour samples were paraffin embedded, this made RNA extraction difficult, and thus RNA was only available for a small number of the samples. Breast conservation therapy is the preferred surgical option for many patients thus surround non tumour tissue is not readily available, and HBL100 had to be used in some cases as the non tumour control, although as seen in chapter 5 most changes present against HBL100 were also detected against pooled in vivo surround tissue.

The tumours samples were also heterogenous, therefore when cross sectioned they contained some tumour and some non tumour cells. Other cell types such as infiltrating lymphocytes may also be present, thus the signal from the tumour cells will be slightly diluted by the presence of normal cells. A way to overcome this problem is to carry out laser capture microdissection (LCM), as this technique will enable the removal of tumour cells exclusively from a tumour mass. In order to achieve this RNA and DNA PCR amplification is carried out on the microdissected tissue to ensure there is enough material for subsequent microarray analysis. LCM was attempted on a number of samples, however the clinical tumours were old
and in some cases not stored correctly for LCM e.g. they were paraffin embedded. Therefore the tumours did not section well and lasered cells could not be captured onto the cap.

6.3.4. Summary

There is a high degree of karyotypic complexity in epithelial tumours such as breast cancer, which is one reason why our understanding of chromosomal aberrations and their chronological order in carcinogenesis is unclear (Hoglund et al., 2002). However this study has elucidated some interesting amplifications that are occurring in breast carcinoma.

A number of regions harbouring amplified genes appeared to be linked to certain phenotypic properties of the tumour, including lymph node status, whether the tumour is pre or post menopausal and ER and ERBB2 hormone status. Individual genes not previously linked to the ERBB2 or lymph node positive phenotype have been suggested, some of which were amplified in clinical samples and breast cancer cell lines. Some regions containing frequently amplified genes were confirmed in literature or were also present in breast cell lines in chapter 5, whilst others were regions not previously detected. Specific genes, such as KIAA0014 were associated with these regions of common amplification between the clinical samples and cell lines. Thus the use of microarray CGH to detect specific gene amplification events occurring in breast tumours is a good approach, although as seen with tumour sample 13, it is insensitive to other possible aberrations that may occur such as aneuploidy.
Chapter 7

Chromosome 17 Genomic Changes
7.1. Introduction

Chromosome 17 is approximately 92 Mb in size, 50% of which has been sequenced (http://www.ncbi.nlm.nih.gov/genome/guide/HsChr17.shtml). The long arm of chromosome 17 is a frequent site of cancer associated genetic aberrations, and studies have shown that both deletions or allele loss and amplification occur (Orsetti et al., 1999; Plummer et al., 1997). There are at least two distinct regions of frequent loss on 17q (Cornelis et al., 1993). The first region at 17q12-q21, found in 38% of breast tumours (Cornelis et al., 1993), is associated with the breast cancer susceptibility gene BRCA1, which maps to 17q21 (Hall et al., 1992). Germ line mutations in the BRCA1 gene are associated with an increase in the overall risk of developing breast, ovarian cancer or both and mutations of the BRCA1 gene are thought to account for approximately 45% of families with significantly higher incidence of breast cancer (Miki et al., 1994). Another region of chromosome loss has been localised to 17q25 (Orsetti et al., 1999).

7.1.1. Regions of Gain on Chromosome 17q

There are several sites of gene amplification in breast cancer on chromosome 17. CGH analysis has shown frequent gains on chromosome 17 within the 17q22-q24 region (Isola et al., 1995; Kallioniemi et al., 1994). In sporadic tumours the reported frequency of gain ranges from 8-20%, although in tumours positive for BRCA2 mutation up to 87% of breast cancers showed gain at 17q22-q24 (Tirkkonen et al., 1997). This region is thought to be important in the pathogenesis of tumours as many other tumour types including neuroblastomas (Lastowska et al., 1997), mesotheliomas (Kivipensas et al., 1996) and dermatofibrosarcomas (Pedeutour et al., 1995) also show gains at 17q22-q24. A FISH study elucidated 2 separate regions of amplification that lie 5 Mb apart within this region (Barlund et al., 1997). The first region may be co-amplified with the ERBB2 oncogene located at 17q11.2-q12, while the second was amplified in cells that did not show ERBB2 amplification. The co-amplification of ERBB2 and genes located at 17q22-q24, such as the PL genes has been confirmed (Letham et al., 2001).

Specific gene target genes within the 17q22-q24 region are not known, although several candidates have been suggested, amongst these are several genes that are amplified and overexpressed thus suggesting a possible functional role in breast cancer. These include RAD51C, PS6K, PAT1 and TBX2, which are located at 17q23 and are amplified and overexpressed in some breast tumours (Barlund et al., 2000). PS6K is also overexpressed at the protein level in some human breast tumours (Couch et al., 1999). Other candidate genes
also amplified and overexpressed are the PL genes at 17q23 (Latham et al., 2001) and SIGMA1B (Wu et al., 2000). The later study suggested that there were several major regions of amplification within the 17q23 region. A more comprehensive analysis of the 17q23 region, whereby a genomic contig map was constructed of the region, which allowed mapping of the amplicon boundaries also revealed the presence of at least 2 common regions of amplification in breast cancer cell lines (Monni et al., 2001). In the study of Monni et al (2001) a number of genes located in these regions that were also overexpressed and thus may have a role in breast cancer progression were suggested including PSK6, MUL, APPBP2, TRAP240 and an uncharacterised EST (Monni et al., 2001). Amplification at 17q11.2-q21 is well documented in breast cancer and is thought to contribute to poor clinical outcome (Kauraniemi et al., 2001). There are thought to be 2 distinct regions of amplification at 17q11-q21 involved in breast cancer (Bieche et al., 1996). The ERBB2 oncogene, which is located at 17q12, is amplified and overexpressed in 10-34 % of breast cancers (Revillion et al., 1998). However the ERBB2 gene is not the only gene involved in this region, as microanalysis has been performed on a number of genes in this region and specific clusters associated with overexpression of ERBB2 and other genes located at 17q12 have been detected (Perou et al., 2000; Pollack et al., 1999). Frequent co-amplification of ERBB2 and several other genes including THRA1, RARA, TOP2A, GRB7, MLN51, MLN64 and PPARBP have also been detected (Zhu et al., 1999; Bieche et al., 1996; Tomasetto et al., 1995; Stein et al., 1994; Keith et al., 1993). And a recent study combining microarray CGH and microarray expression analysis of the 17q21 region found several genes that were also overexpressed and amplified in a selection of human breast cancer cell lines, these include genes previously detected ERBB2, GRB7, MLN51 and MLN64 as well as CDC6 and ZNF144 and some novel ESTs that were not previously reported (Kauraniemi et al., 2001). This region is not unique to breast cancer as amplification of several genes at 17q11-q12 also occurs in gastric cancers, including ERBB2, TOP2A, GRB7, MLN51 and MLN64 the first two are also overexpressed in some gastric tumours (Varis et al., 2002), thus amplification at 17q11-q12 may play an important role in tumourgenesis and progression.

7.1.2. The ERBB2 Oncogene

The c-ERBB-2 gene is a v-erbB related gene that was originally identified as an oncogene in chemically induced rat neuroglioblastomas, where a single point mutation in the transmembrane domain of the molecule conferred oncogenic activation (Bargmann et al., 1986; Shih et al., 1981), although in humans activation of ERBB2 is achieved by
amplification and overexpression. The ERBB2 gene encodes a 185 kDa transmembrane receptor protein that has intrinsic tyrosine kinase activity and is a member of the epidermal growth factor receptor (EGFR) class I family (Reese and Slamon, 1999), members of which are considered important mediators of growth, differentiation and survival (Meric et al., 2002). The precise role and mechanisms of ERBB2 in breast cancer are not known, although overexpression of ERBB2 in breast cell lines has been shown to increase DNA synthesis and promote cell growth (Pietras et al., 1995) and in ‘normal’ cells activation of growth factor receptors leads to the phosphorylation and activation of multiple second messengers.

Amplification and overexpression of ERBB2 is found in several cancers including ovarian (Hengstler et al., 1999), gastric (Varis et al., 2002) and approximately 20% of breast cancers (Tsuda et al., 2001; Revillion et al., 1998). ERBB2 amplification has been associated with poor prognosis and poor response to chemotherapy in breast cancer (Hengstler et al., 1999), however the true clinical significance of ERBB2 amplification is debatable (Revillion et al., 1999). ERBB2 overexpression is not a significant predictive marker for patients treated with preoperative high dose anthracycline based chemotherapy (Vincent-Salomon et al., 2000), although overexpression of ERBB2 has been associated with the development of distant metastases in patients treated with high doses of anthracycline chemotherapy (Vargas-Roig et al., 1999) and ERBB2 overexpression has also been shown to contribute to the development of the resistance phenotype against chemotherapy (Hengstler et al., 1999). Others have suggested that amplification of ERBB2 alone is not a prognostic marker, but co-amplification of ERBB2 and other genes such as myc, is associated with a reduction of patient survival (Cuny et al., 2000).

The apparent involvement of ERBB2 in the pathogenesis of breast cancer led to its use as a therapeutic target. Potent and selective ATP-competitive inhibitors of ERBB2 kinases have been developed, including Iressa (AstraZeneca), which is currently in Phase III clinical trials (Wissner et al., 2002). Several monoclonal antibodies against the intracellular domain of the ERBB2 protein have been also found to inhibit proliferation in human cancer cells that overexpress ERBB2 (Shepard et al., 1991; Pietras et al., 1998). One such antibody, trastuzumab, which inhibits tumour growth alone (Greenberg et al., 1996), also has synergistic or increased effects when administered with some chemotherapeutic agents (Pegram et al., 1999; Pietras et al., 1998) or ionising radiation (Pietras et al., 1999). Phase I and II trials of trastuzumab have shown the antibody to be safe and confined to the tumour, additionally women with ERBB2 positive metastatic tumours who had relapsed after
chemotherapy have had a clinical response to the antibody (Baselga et al., 1996). Phase III trials have shown that the combination of trastuzumab and chemotherapy was associated with a longer time to disease progression, a higher rate and longer duration of response and longer survival with a 20% reduction in the risk of death (Slamon et al., 2001). Thus amplified and overexpressed genes such as ERBB2 are potential targets for breast cancer treatment, and antibodies such as trastuzumab have been shown to increase the clinical benefit of chemotherapy in some tumours.

7.1.3. Chapter Objectives

Chromosome 17 is known to harbour several regions of amplification that may be indicators of poor prognosis. In some of the in vivo clinical tumours and human breast tumour cell lines studied here chromosome 17 gains were common. Thus in this chapter the regions of chromosomal gain especially 17q11-q12 have been analysed in greater detail in the hope to further define the extent of amplification in this region.
7.2. Results

There was a high frequency of regions containing gene amplification on chromosome 17 in the breast cells and clinical tumour samples (Chapter 5 and 6). Figure 7.1. displays the ideogram of chromosome 17 and summarises the regions containing amplification, the breast cell lines are represented by bars to the right and clinical samples are shown to the left.

Five clinical samples 02 and 11-13 all contained regions of amplification on the long arm of chromosome 17 and one sample, 13, also had a region with more than one amplified gene at 17p. Five different breast cancer cell lines also contained regions harbouring amplified genes on the long arm of chromosome 17, BT474, MCF7, HPGP, MDA468 and ZR75.1, 3 of which also had an amplification in 17p. The most frequent regions of amplification were between 17q11.2-q23.

7.2.1. Specific Regions Containing Amplified Genes in Cell Lines

Breast tumour cell line samples that contained regions of amplification containing more than one gene (represented by thick bars) are represented graphically in Figure 7.7. The ratio for each gene is plotted against the cR3000 value from GeneMap '98, if there is no cR3000 value available the genes were assigned to relative positions along the chromosome using the Locuslink NCBI database that contains data from the genomic sequence generated by the Human Genome Mapping Project (http://www.ncbi.nlm.nih.gov/).

Within the 3 breast cell line samples there were 7 separate regions that contain amplified genes, the majority of these regions were specific to certain tumours, thus contained genes that were only amplified in a single tumour. However amplifications within region 6 were found in all 3 tumours and in BT474 and MCF7 cells, and the same gene was amplified, PPMID. This gene was also slightly overexpressed in both BT474 and MCF7 cells with a ratio of 2.0 (+/- 1.1) and 3.3 (+/- 0.9) respectively. Region 4 in the BT474 tumour cell line contained a high abundance of amplified genes and corresponded to 17q11.2-q12, this region also contained more than one amplified gene in 4 clinical tumour samples, 02, 11, 12 and 13 (Figure 7.1.) and consequently is studied in greater detail in section 7.2.3.
Figure 7.1 Legend: Regions of amplification on chromosome 17 detected in the *in vivo* clinical (left of ideogram) and breast cell lines (right of ideogram). The ideogram of chromosome 17 with labelled regions is shown. Thin bars indicate regions that contained a single amplified gene and thick bars represent regions that contained more than one amplified gene.
Figure 7.2. Chromosome 17 Regions Containing Amplified Genes in Cell Lines

A. BT474 Cell Line vs HBL100

B. MCF7 Cell Line vs HBL100

C. ZR75.1 Cell Line vs HBL100

Figure 7.2. Legend: Chromosome plots for 3 different breast cancer cell lines that contained regions detected with microarray CGH, which contained more than one amplified gene. In each case the cR3000 value of the genes along the chromosome is plotted against the gene average ratio of the breast cell line vs HBL100. Genes were considered amplified if their fold change was greater than 2. The cell lines contained 7 areas of amplification along the chromosome, each represented by a grey bar.
7.2.2. Specific Regions Containing Amplified Genes in Clinical Tumours

Four clinical tumour samples, 02 and 11-13 also contained regions harbouring more than one amplified gene (Figure 7.1.), thus amplifications within chromosome 17 in the 4 clinical tumour samples are graphically represented in Figure 7.3.

There were several regions of amplification detected in the different clinical tumours along chromosome 17, 2 of which were common to several samples and are highlighted by grey bars (Figure 7.3.). The first region, A contained the *CDK5R1* gene, which was amplified in clinical tumour samples 11-13. In tumour 11 the gene had a fold change of greater than 6 and in tumour 13 a fold change of greater than 10. The other region is region 4, which was detected in BT474 cells (Figure 7.2.) and is located within 17q11.2-q12. This region also harboured a high number of amplified genes in the 4 clinical samples and is discussed in the next section.

7.2.3. The 17q11.2-q12 Region of Amplification

Microarray data for genes located together at a region in 17q11-q12 (Region 4) for 4 clinical breast tumour samples and BT474 cells are shown in Table 7.1. For each gene the average expression ratio and standard deviation are shown. The order of the genes in the table corresponds to the order in which they are found along chromosome 17 in the Locuslink NCBI database (http://www.ncbi.nlm.nih.gov/locuslink; Pruitt and Maglott, 2001). Also located within this region are a number of hypothetical proteins, however as little information is available on these hypothetical genes and no name has yet been assigned they have not been included.

There are several genes that were amplified in all samples including *PPARBP*, the oncogene *ERBB2* and *GRB7*. Amplification of the *NR1D1* gene located at the end of the list did not occur in any tumour sample. RNA expression data is also shown for the BT474 cell line, and 8 of these genes were also overexpressed. Four genes in the region *LASP1*, *RPL23*, *TCAP* and *ZNFIA3* are not present on the current MRC Leicester microarray and so no data is available. Data is not available for 10 genes in the clinical tumour 02, as this tumour was analysed on an early batch of microarrays, prior to the addition of other clones to the microarray.
Figure 7.3. Legend: Chromosome plots of 4 clinical tumours. The cR3000 value of the genes is plotted against the gene average ratio vs HBL100. Two of these regions are highlighted with grey bars. Region A contains genes amplified in 3 samples and region 4 (from BT474 cells corresponding to 17q11.2-q12, Figure 7.2.) was present in the 4 clinical samples.
Table 7.1. Genes Associated with the 17q11.2-q12 Region of Amplification

<table>
<thead>
<tr>
<th>Gene</th>
<th>BT474 DNA</th>
<th>BT474 RNA</th>
<th>Clinical Tumour Sample DNA 02</th>
<th>Clinical Tumour Sample RNA 11</th>
<th>Clinical Tumour Sample RNA 12</th>
<th>Clinical Tumour Sample RNA 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZNF144</td>
<td>2.5 (0.8)</td>
<td>6.5 (2.2)</td>
<td>0.6 (0.6)</td>
<td>1.1 (0.2)</td>
<td>2.2 (0.4)</td>
<td>1.1 (0.0)</td>
</tr>
<tr>
<td>PIP5K2B</td>
<td>1.0 (0.1)</td>
<td>1.2 (0.5)</td>
<td>0.3 (0.5)</td>
<td>1.1 (0.3)</td>
<td>0.8 (0.0)</td>
<td>1.2 (0.3)</td>
</tr>
<tr>
<td>RPL23</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>LASP1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>TEM7</td>
<td>4.1 (1.9)</td>
<td>0.5 (0.7)</td>
<td>NA</td>
<td>1.2 (0.0)</td>
<td>1.7 (0.1)</td>
<td>0.9 (0.0)</td>
</tr>
<tr>
<td>CACNB1</td>
<td>4.4 (0.4)</td>
<td>0.8 (0.1)</td>
<td>1.0 (0.3)</td>
<td>1.2 (0.4)</td>
<td>1.9 (0.9)</td>
<td>4.2 (1.5)</td>
</tr>
<tr>
<td>RPL19</td>
<td>2.2 (0.8)</td>
<td>12.3 (5.6)</td>
<td>1.2 (0.4)</td>
<td>1.3 (0.1)</td>
<td>1.0 (0.1)</td>
<td>0.9 (0.2)</td>
</tr>
<tr>
<td>PPARBP</td>
<td>6.8 (4.3)</td>
<td>10.8 (2.6)</td>
<td>3.6 (0.3)</td>
<td>2.3 (0.2)</td>
<td>3.2 (0.3)</td>
<td>4.2 (0.6)</td>
</tr>
<tr>
<td>CrkRS</td>
<td>5.1 (0.6)</td>
<td>2.6 (1.1)</td>
<td>NA</td>
<td>2.7 (1.3)</td>
<td>3.3 (0.4)</td>
<td>5.8 (1.1)</td>
</tr>
<tr>
<td>NEUROD2</td>
<td>2.5 (0.4)</td>
<td>0.6 (0.1)</td>
<td>NA</td>
<td>1.7 (0.4)</td>
<td>1.7 (0.3)</td>
<td>5.0 (1.0)</td>
</tr>
<tr>
<td>MLN64</td>
<td>2.9 (0.3)</td>
<td>7.2 (4.8)</td>
<td>0.7 (0.6)</td>
<td>1.7 (0.7)</td>
<td>1.8 (0.3)</td>
<td>4.2 (1.1)</td>
</tr>
<tr>
<td>TCAP</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>PNMT</td>
<td>1.0 (0.1)</td>
<td>0.6 (0.8)</td>
<td>1.2 (0.1)</td>
<td>0.7 (0.3)</td>
<td>0.8 (0.2)</td>
<td>1.1 (0.4)</td>
</tr>
<tr>
<td>ERBB2</td>
<td>3.8 (0.4)</td>
<td>12.8 (7.6)</td>
<td>3.0 (1.3)</td>
<td>3.1 (0.3)</td>
<td>2.6 (0.6)</td>
<td>4.8 (1.3)</td>
</tr>
<tr>
<td>ZNFN1A3</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>GRB7</td>
<td>4.6 (2.0)</td>
<td>13.7 (8.0)</td>
<td>2.4 (0.6)</td>
<td>3.3 (1.6)</td>
<td>2.0 (0.2)</td>
<td>4.8 (1.1)</td>
</tr>
<tr>
<td>CSF3</td>
<td>2.5 (1.5)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.4 (0.6)</td>
</tr>
<tr>
<td>PSMD3</td>
<td>4.4 (0.3)</td>
<td>10.1 (8.3)</td>
<td>NA</td>
<td>2.4 (0.9)</td>
<td>1.8 (0.3)</td>
<td>1.2 (0.0)</td>
</tr>
<tr>
<td>MLN51</td>
<td>1.1 (0.2)</td>
<td>1.4 (1.2)</td>
<td>NA</td>
<td>3.3 (1.2)</td>
<td>2.2 (0.3)</td>
<td>0.8 (0.1)</td>
</tr>
<tr>
<td>NR1D1</td>
<td>0.6 (0.8)</td>
<td>0.7 (0.3)</td>
<td>1.0 (0.1)</td>
<td>0.9 (0.0)</td>
<td>1.1 (0.1)</td>
<td>0.8 (0.0)</td>
</tr>
</tbody>
</table>

Table 7.1. Legend: A summary of microarray data for known genes present in the region at 17q11-q12 in 4 clinical tumour samples and 1 breast cancer cell line. There are a number of hypothetical proteins and genes that also lie in this region, but as little information is available for them and as they are only hypothetical they have not been included in the table. The order of the genes in the table corresponds to the order in which they are found along chromosome 17 in the Locuslink NCBI database (http://www.ncbi.nlm.nih.gov/locuslink; Pruitt and Maglott, 2001). For the BT474 breast cell line RNA expression data is also shown. Red indicates genes had a fold change (ratio) of greater than 2 and so were considered amplified or overexpressed. NA = no data available. The fold change or ratio and (standard deviation) for each gene are shown.
7.2.4. Confirmation of Microarray Data in BT474 Cells

Gene amplification and expression microarray data for the genes located within 17q11.2-q12 was confirmed with Southern blotting (Figure 7.4.) and Real Time PCR respectively (Figure 7.5.). Protein expression of the ERBB2 gene was also analysed by Western blotting (Figure 7.6.). Confirmation was only carried out in breast cancer cell lines as there was insufficient clinical in vivo material.

A selection of genes recognised as amplified in the BT474 cells by microarray CGH analysis, CACNB1, PPARBP, ERBB2 and GRB7 were confirmed by Southern blotting (Figure 7.4.). For each gene amplification was confirmed in the BT474 cells when compared to HBL100 and the other breast cell lines. The ERBB2 gene was also amplified in the MCF7 cell line when compared to HBL100 and T47D cells. A southern blot was also carried out for NR1D1, which confirmed the microarray CGH finding that amplification of this gene did not occur in any cell line compared to HBL100.

Microarray RNA expression data for a number of genes located in the 17q11-q12 region (Locuslink) were confirmed by Real Time PCR in the 7 breast cancer cell lines (Figure 7.5.). In each case the relative abundance of the gene was calculated by comparing all cancer cell lines against the HBL100 non tumour cell line that was thus assigned a value of 1. For RPL23, LASP1, TCAP and ZNF1A3 that were not present on the microarray and thus no microarray data was available, were also analysed with Real Time PCR for relative RNA expression in the breast cell lines and all of these genes were overexpressed in the BT474 cell line compared to HBL100 by Real Time PCR. Overexpression of genes by microarray analysis in the BT474 cell line compared to HBL100 were confirmed by Real Time PCR. For the majority of genes overexpression was specific to the BT474 cells, although RPL23 and CACNB1 were overexpressed in T47D cells and ZNF1A3 was also overexpressed in MDA16 and HPGP cells when compared to HBL100. Microarray data of CACNB1 and NR1D1, indicated their expression did not change in BT474 cells compared to HBL100 these findings were also confirmed by Real Time PCR.

Changes in RNA expression or DNA amplification may not necessarily correlate to protein abundance. Therefore a Western blot was carried out on the known oncogene within this region ERBB2 (Figure 7.6.), which showed the ERBB2 protein was also overexpressed in BT474 cells compared to other cell lines.
Figure 7.4. Legend: Southern blots confirming microarray CGH data for 5 genes located in the 17q11.2-q12 region. The gene names are shown on the right and red boxes indicates the gene was amplified on the microarray, while yellow boxes indicate there was no change. Four cell lines were analysed for each gene BT474, HBL100, MCF7, and T47D. Amplification of *CACNB1*, *PPARBP*, *ERBB2*, and *GRB7* is clearly shown in the BT474 cells compared to the other cell lines. The *NR1D1* gene was not amplified in any cell line compared to the HBL100 cells.
Figure 7.5. Confirmation of Microarray Expression Data by Real Time PCR

Key

<table>
<thead>
<tr>
<th>Microarray data</th>
<th>Real Time PCR data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overexpressed</td>
<td>HBL100</td>
</tr>
<tr>
<td>No change</td>
<td>BT1474</td>
</tr>
<tr>
<td>No data</td>
<td>ZR75.1</td>
</tr>
<tr>
<td></td>
<td>MDA468</td>
</tr>
<tr>
<td></td>
<td>MDA16</td>
</tr>
<tr>
<td></td>
<td>MCF7</td>
</tr>
<tr>
<td></td>
<td>HGP</td>
</tr>
</tbody>
</table>

Continued over page
Figure 7.5. Legend: Confirmation of microarray expression data for 12 genes located in the 17q11.2-q12 region by Real Time PCR. The gene names are shown. Red boxes indicates the gene was amplified on the microarray, yellow boxes indicate there was no significant change in expression and grey boxes indicate there was no microarray data available. The genes are positioned in order as they appear on the chromosome in LocusLink (www.ncbi.nlm.nih.gov/LocusLink). All 7 human breast carcinoma cell lines and the non tumourgenic HBL100 cell line were analysed for each gene. Cell lines are indicated on the x axis in the key.
Figure 7.6. Legend: Western blot of the protein product of the ERBB2 gene. The 209 kDa molecular weight marker is shown and the cell lines are indicated across the top. The BT474 cell line contained an abundance of the ERBB2 protein and MCF7 and T47D cells contained small amounts. The protein was absent in the non tumourgenic HBL100 cells.
7.3. Discussion

Amplification of genes on the long arm of chromosome 17 was the most common amplification observed in this tumour cohort, as 71% of breast cancer cell lines and 31% of clinical samples analysed here contained amplified genes on 17q. Several genes located on chromosome 17 including TP53, ERBB2 and BRCA1 are known to be involved in breast cancer, and may be contributing to the high frequency of genetic change.

7.3.1. Amplification at 17q23

There are thought to be several areas within chromosome 17 containing amplified genes. One such area is 17q23 that has been associated with a more aggressive clinical course and poor prognosis in breast cancer (Monni et al., 2001). Two separate regions of amplification within the 17q23 area have previously been recorded in human breast cell lines in a FISH study and one of these regions was found frequently co-amplified with the ERBB2 region located at 17q11.2-q12 (Barlund et al., 1997). Amongst the cell lines in the study by Barlund et al. (1997) was the BT474 cell line that contained both amplifications at the ERBB2 regions and 17q23. In this study co-amplification of the ERBB2 region and genes located on 17q23 was also detected in BT474 cells and a clinical tumour sample, 11 (Figure 7.1.). The BT474, MCF7 and ZR75.1 cell lines all contained amplifications at 17q23 and copy number gain in these cell lines at chromosome 17q22-q24 has previously been shown by CGH (Barlund et al., 1997).

A study of 372 primary tumours associated several genes with regions of amplification in 17q23 including RAD51C a gene amplified in 3% breast tumours, S6K, PAT1 and TBX2, which were co-amplified in 10% of breast tumours (Barlund et al., 2000). A later study also showed amplification of PAT1, PS6K, RAD51C and SIGMA1B in breast cancers (Wu et al., 2000), and in most cases these genes were overexpressed when amplified (Wu et al., 2000; Couch et al., 1999). Of these genes only RAD51C is currently located on the MRC Toxicology microarray, and this gene was not significantly amplified in tumours studied here, although this gene has previously been recorded as amplified in BT474 and MCF7 cells by Southern blotting and Real Time PCR (Wu et al., 2000). This discrepancy may be due to different non tumour controls or the different techniques used, as in some cases microarrays are thought to underestimate changes occurring or there may be a difference in the cells themselves, as variation of the same cell line is known to occur between laboratories, as cells may be cultured differently or may be at a different passage number, although as seen in chapter 5 the MCF7 cell used in this study are genetically similar to those in literature.
Another source of variation can be found in the clinical tumours, as the samples studied here are a heterogeneous population of breast tumour and non tumour cells, therefore amplification events in the tumour cells may be diluted by the presence of normal cells and other cell types, and thus not detected. In this study only 7 breast tumour cell lines and 13 clinical samples were studied, thus amplifications present in few tumours, such as RAD51C (3 %) may also not be detected due to the small population studied. However it could be suggested that rare amplifications, although they may play a role in tumour progression and development may not be as important as those genes which are amplified at higher frequencies.

A gene in close proximity to RAD51C, PPM1D was both amplified and overexpressed in BT474 and MCF7 cell lines in this study. PPM1D encodes a serine/threonine protein phosphatase and is known to be amplified and overexpressed in MCF7 breast tumour cell lines (Li et al., 2002). PPM1D is thought to have weak oncogenic properties and when transfected into cells has the ability to cooperate with RAS to transform primary mouse fibroblasts, however though transformed the cells were not immortalised and did not survive beyond 7 passages (Li et al., 2002). Amplification of PPM1D in human tumours is also known to contribute to the development of cancer by suppressing activation of the p53 tumour suppressor gene (Bulavin et al., 2002; Takekawa et al., 2000), as activation of p53 may be achieved by phosphorylation at the Ser33 and Ser46 by the p38 mitogen activated protein kinase (MAPK), and PPM1D overexpression can inhibit this phosphorylation (Bulavin et al., 2002). Thus the PPM1D gene is thought to be oncogenic due to its negative effect on p53 function (Li et al., 2002).

7.3.2. The ERBB2 region of Amplification at 17q11.2-q12

Amplification of ERBB2 and several neighbouring genes was detected in the BT474 cell line and 4 clinical tumour samples (31 %), in agreement with previous studies where amplification and overexpression of the ERBB2 oncogene has been detected in 10-34 % of breast cancers (Revillion et al., 1998). Several other genes are sometimes co-amplified in this region including THRA1, RARA, TOP2A, GRB7, MLN51, MLN64, PPARBP, CDC6 and ZNF144 (Kauraniemi et al., 2001; Zhu et al., 1999; Bieche et al., 1996; Tomasetto et al., 1995; Stein et al., 1994; Keith et al., 1993). In this chapter the 17q11.2-q12 region harbouring ERBB2 and other amplified genes was elucidated in greater detail.
The number of genes amplified within the region varied and BT474 cells contained the greatest number of amplified genes as a cluster of 12 genes located in close proximity to ERBB2 was amplified in BT474 cells. Amplification of MLN51 was not detected in BT474 cells, although it was amplified in clinical tumour samples 11 and 12. These data suggests that this area is susceptible to gene amplification but the size of the amplicon varies from tumour to tumour. Interestingly 4 genes were amplified in all 5 samples that contained amplifications in this region, PPARBP, CrkRS, ERBB2 and GRB7. Thus these genes may lie at the centre of the amplicon. In all cases the NR1D1 gene was not amplified and so marked the upper boundary of amplification. CD6, RARA and TOP2A amplification was also not detected in the samples studied here, and in Locuslink these genes lie further away from the amplification than NR1D1, suggesting that regions of amplification larger than those detected here may occur, or that a discrete separate region of amplification may exist.

Several genes amplified in this region have also been shown to be overexpressed, ZNF144, MLN64, ERBB2, GRB7, MLN51 and CDC6 (Kauraniemi et al., 2001). Eight of the amplified genes in BT474 cells, ZNF144, RPL19, PPARBP, CrkRS, MLN64, ERBB2, GRB7 and PSMD3 were also overexpressed. Data was confirmed by Real Time PCR and 4 additional genes also located in this region, RPL23, LASP1, TCAP and ZNF1A3 were also found to be overexpressed in BT474 cells in this study. Thus these genes as well as ERBB2 may have an active role in breast cancer development and progression, as they are differently amplified and overexpressed in breast tumour cells. The 3 genes, PPARBP, CrkRS and GRB7 which were co-amplified with the ERBB2 oncogene in all 5 samples were also overexpressed in BT474 cells and thus may be of greater importance. Amplifications within this region have also been detected in other cancers including gastric and ovarian cancers (Varis et al., 2002; Hengstler et al., 1999), thus genes involved may play a more general role in tumourgenesis and progression.

ERBB2 amplification has been associated with poor prognosis and poor response to chemotherapy in breast cancer (Hengstler et al., 1999), however the clinical significance of ERBB2 amplification is debatable (Revillion et al., 1999). The reason for this discrepancy may depend on the other genes which are co-amplified and overexpressed with ERBB2, since it is clear the region of amplification varies between tumours and the increased biological activity of several genes, not just ERBB2 alone, may jointly contribute to the tumour phenotype and more aggressive clinical behaviour of some tumours. There is some evidence to support this as co-amplification of the PL genes situated at 17q22-q24 and ERBB2 is seen
mainly in lymph node positive tumours, suggesting a possible association with high malignancy (Latham et al., 2001).

There were several genes that lay in the region of amplification (as determined by relative position in LocusLink and GeneMap98) including \textit{PIP5KB2} and \textit{PNMT} which were not amplified in any breast tumour sample. Radiation hybrid mapping used to generate the cR3000 values in GeneMap98 can not determine the relative position of adjacent markers or ESTs, consequently the number as well as the position of amplification peaks here may change as this region is sequenced or physically mapped. It is expected that the complete human genome sequence will replace the radiation hybrid map in the near future and as the sequence information from the human genome mapping project becomes available construction of microarrays representing specific regions of chromosomes can be constructed. Currently LocusLink provides information on gene location along chromosomes, but again the position of certain genes within a region may change as the genome is fully sequenced or physically mapped (Pruitt and Maglott, 2001). Thus the 2 genes \textit{PIP5K2B} and \textit{PNMT}, which were not in BT474 cells may not actually be located within this region.

\textit{7.3.3. Summary}

Due to the high incidence of chromosome amplification detected in the breast tumour samples along chromosome 17, microarrays were used in this chapter to analyse the regions of chromosomal gain especially at 17q11.2-q12 in greater detail in the hope to further elucidate these regions. A number of genes were amplified in various tumour samples including \textit{PPM1D}, \textit{CDK5RI} and several genes located at 17q11.2-q12. The amplification region at 17q11.2-q12 that was mapped out in greater detail in this chapter appeared to vary between tumours, and whether the oncogene \textit{ERBB2} is a predictor of clinical outcome or prognosis may depend on the other genes involved. A number of genes that were more frequently amplified in the tumours were also overexpressed in BT474, and may play an active role in breast cancer. Due to the role that \textit{ERBB2} has in breast cancer and its differential expression, it has become a successful target for breast cancer treatment. In a similar manner other genes also amplified and differentially expressed may also have an important role in breast cancer and are potential targets for treatment.
Chapter 8

General Discussion
8. General Discussion
Despite the large number of studies that have been conducted into the different aspects of breast cancer, the underlying mechanisms of development, progression and treatment failure remain elusive. The overall aim of this thesis was to relate breast carcinoma phenotype to genetic changes using cDNA microarrays to cover a broad spectrum of genes.

8.1. Gene Changes and the Drug Resistance Phenotype

Human breast cancer cell lines with both an acquired and intrinsic resistance were initially compared for altered gene expression against their sensitive progenitor cells and a panel of breast cells with varying sensitivity to Dox (Chapter 3). Subsequent experiments analysed genetic changes that were occurring in the initial stages of resistance acquisition in MCF7 cells that were selected for resistance by increasing concentrations of Dox (Chapter 4).

The acquired (HPGP) and intrinsic (MDA16) resistant cells both had a similar pattern of gene amplification and expression and thus clustered together (Chapter 3), suggesting that a tumour will develop a MDR phenotype, not as a result of cells acquiring such a phenotype, but rather through the cytotoxic drug selection of cells that already have such a phenotype.

8.1.1. Genomic Amplification and the Drug Resistant Phenotype

There were 8 genomic amplifications that were common to both HPGP and MDA16 cells which are summarised in Figure 8.1., several of these genes were also overexpressed (shown in red). \(ABCB1\), was amplified and overexpressed, and this gene is involved in the drug resistant phenotype as it actively extrudes Dox from the cell and genomic amplification of the \(ABCB1\) in resistant cells has previously been described (Fairchild et al. 1987; Shen et al. 1986). \(ABCB4\), which was also amplified, is located alongside \(ABCB1\) at chromosome 7q21.1 and its amplification in a cassette form with the \(ABCB1\) gene is recognised. However the gene was not overexpressed in the Dox resistant cells, thus it is not contributing to the resistance phenotype. Its failure to be overexpressed when amplified may also suggest that the gene has not been amplified in a complete form.

The other genes amplified in both Dox resistant cells included 2 genes involved in RNA production, which were also overexpressed and 4 genes that were not overexpressed and so are thought not to be contributing to the resistant phenotype.
8.1.2. Gene Overexpression and the Drug Resistant Phenotype

The majority of gene changes observed in both the resistant cells (Chapter 3) and the early gene changes in cells during the development of a minor resistant phenotype by a stepwise exposure to Dox (Chapter 4) are summarised in Figure 8.2. The diagram is split into 3 main sections, the first summarises genes that were overexpressed in the resistant cells compared to sensitive cancer cell lines (Chapter 3). The second section represents gene changes that occurred in the sensitive cell lines in the stepwise increase in Dox after the development of a minor resistant phenotype and the final section summarises changes that occurred in cells exposed to 20 nM Dox for one week, after which time resistance did not develop (Chapter 4).
Figure 8.2. Gene Changes Associated with the Dox Resistant Phenotype

Figure 8.2. Legend: Summarises genes that were firstly associated with the resistant cells (Chapter 3) and secondly the genes that were involved in the development of the resistant phenotype (Chapter 4). All genes in red are present on the array and were overexpressed in the resistant cells. The effect of genes on each other are shown by \(\text{"-\" inhibition or \"+\" induction. Chapter 4 genes are split into two groups, those that were expressed in the 7 day 20 nM Dox study that did not result in resistance development and those that were expressed during the stepwise increase in drug concentration during resistance development. The genes in the blue boxes were no longer expressed when the cells became resistant.
Among the genes associated with the Dox resistant phenotype in Figure 8.2. were MMP1, PLAU and VCAM1, which are associated with increased metastases or invasiveness (Stein et al., 1993; Rice and Bevilacqua, 1989). Cells with a MDR phenotype have previously been reported to have an increased invasive ability (Weinstein et al. 1991). However in contrast NK4 was also overexpressed, this gene is an antagonist of HGF (not included on the Leicester microarray), which promotes invasion and metastases, thus expression of NK4 is conflicting. MMP1 is also involved with several other overexpressed genes, as TGFB1 stimulates SPARC, which in turn can induce MMP2, and as discussed earlier may also regulate MMP1 (Chapter 3). TIMP2 is an inhibitor of MMP1 (Matrisian, 1990), so to see it overexpressed is strange, although it may be overexpressed due to the high concentration of MMP1, however its inhibition appears not significant enough to lower MMP1 levels. Other genes included 2 potential therapeutic targets, a gene down regulated in cells undergoing apoptosis and genes involved in the drug resistant phenotype, including ABCB1 and GST\(\pi\).

ABCB1 was also overexpressed in cells that have gained a slight resistance to Dox (Chapter 4) and ABCB1 appeared proportional to degree of resistance attained. Other genes expressed in these cells include predictors of poor prognosis in invasive cancers, another membrane glycoprotein, and a gene involved in cell shape change. The latter is ARPC3 whose overexpression can be explained, as when sensitive MCF7/wt developed a slight resistance they underwent a morphological change and became more elongated.

There were several genes overexpressed at various time points when sensitive MCF7/wt cells were exposed to 20 nM Dox over a 7 day period, which did not result in resistance. Early genes overexpressed included TNFRSF6, which is involved in the stimulation of apoptosis (Theuns et al., 2001) and genes associated with the repair of DNA lesions caused by Dox. Thus these genes could to be a cellular response to Dox and not involved in resistance. After 7 days exposure to Dox S100A2, a candidate tumour suppressor, STAT1 and 2 IFN inducible genes were overexpressed. Interestingly once resistance developed these genes were no longer overexpressed. Induction of STAT1 and IFN inducible genes are known to be involved in senescence (Perou et al., 1999). When MCF7 cells were exposed to Dox they went through a period of growth arrest prior to resistance development, thus when these genes are no longer expressed the cell may come out of senescence and begin to undergo proliferation in the presence of Dox thus aiding in the development of resistance.
8.2. Analysis of Genomic Amplification

There are many genomic aberrations that have been detected in malignant cells compared to normal cells, although only 6 mutations are thought to be necessary for carcinogenesis (Vogelstein and Kinzler, 1993). Therefore other changes that occur may be the result of downstream or cascading effects from the initial changes, and work still needs to be done to determine which genetic changes are the causing events in the transformation of normal cells to tumour cells. There is also a high degree of karyotypic complexity in epithelial cells such as breast cancer, which is one reason why our understanding of chromosomal aberrations and their chronological order in breast carcinogenesis is unclear (Hoglund et al., 2002).

In this study genomic amplifications were analysed by microarray, as gene amplification is a mechanism by which important genes may become up regulated in genetically unstable tumours cells (Pollack et al., 2002). Initially a set of human breast cancer cell lines were analysed for both genomic amplification and expression (Chapter 5), then genomic analysis was carried out on 13 clinical breast tumour samples (Chapter 6). In both these chapters chromosome 17 exhibited a high number of amplification events and was analysed in greater detail (Chapter 7).

8.2.1. Regions of Frequent Genomic Amplification

The breast cancer cell lines exhibited a high number of genetic aberrations (mean of 17), this high number may be explained by the hypothesis that cell lines are a genetically advanced population (Kallioniemi et al., 1994), which occurs because the cells in culture are free from some growth restraints found in surrounding cells in vivo and so can undergo clonal expansion at an increased rate, leading to an increased number of chromosomal aberrations. In general there was a good correlation of amplification regions in the breast cancer cell lines between data obtained in this study using CGH microarray and those found in literature that used FISH or CGH.

The clinical breast tumour samples used here (Chapter 6) were a selection of different grade and type, which allowed the detection of regions of genetic change specific to certain phenotypes. There were 17 regions of amplification present in greater than 3 (23 %) tumour samples. These included 6 regions previously seen frequently amplified (greater than 15 %) in literature (Tirkkonen et al., 1998; Kuukasjarvi et al., 1997) and 8q and 17q that were also seen in 71 % of the breast cancer cell lines (Chapter 5). Chromosomal gains that occur at high frequency in different tumours are likely to harbour genes, which are providing a critical
advantage for tumour survival and progression. For example the gene primarily amplified in the 8q region in this study was KIAA0014, which was amplified in 9 clinical tumours and 2 breast cancer cell lines. KIAA0014 is not well characterised, but interestingly this gene was also overexpressed in the cell lines, therefore it could have an important role in breast cancer.

8.2.2. Genomic Amplification and Tumour Phenotype

Gene expression patterns have previously been linked to different types of breast tumour, as clusters of genes were associated with ER status and clinical tumour stage (Martin et al., 2000). In this study associations between genomic amplification and tumour phenotype were detected.

Five of the clinical tumours used here were lymph node negative and 8 were lymph node positive and several trends of amplification were detected. Previously Hermsen et al. (1998) in a CGH study of 53 lymph node negative tumours found frequent regions of chromosome gain at 8q, 1q, 4q, 5q and Xq and in all cases that had 1q gains 8q gains were also present, and gains of 8q could be found without 1q. The same was also true in findings here, which suggests that gains at 8q occur before, and are necessary, for 1q gains in lymph node negative tumours. However some CGH studies have suggested that 1q gain is an early event, as it has been reported in genetically less complex tumours and has also been recorded as the sole genetic event (Tirkkonen et al., 1998), and in other studies it is the most frequent amplification event (Roylance et al., 1999; Nishizaki et al., 1997). However these studies were not restricted to lymph node negative tumours and did not differentiate the amplification events on the basis of lymph node status.

Here, 2 regions of frequent amplification, 2p25 and 22q11 were found to be specific to lymph node positive tumours. Two genes were consistently found amplified in these regions, LPIN1 at 2p25 and DGSI at 22q11. Amplifications of 2p25 or 22q11 have not previously been recorded in breast cancer and interestingly although LPIN1 and DGSI are located on different chromosomes neither gene was amplified alone. There was also an association with LPIN1 and DGSI amplification with ERBB2 and menopausal status, as DGSI and LPIN amplification occurred predominantly in premenopausal lymph node and ERBB2 positive tumours. As well as 2p25 (LPIN) and 22q11 (DGSI), other regions containing gene amplification were found associated with premenopausal women only including 3p21, 9q34, 11p15, 17q11-q12, 19q13 and 20p13. Interestingly regions 11p15 and 20p13 were also only detected in premenopausal
ERBB2 positive tumours, although not all ERBB2 positive and premenopausal women contained these amplifications.

The regions of gene amplification detected generally appeared to be independent of tumour grade. However some regions were detected in all grade tumours (I – III), 8q24.3, 13q14, 1q42 and 16p13, while others were only detected in tumour grades II – III or undetermined. However it cannot be concluded that the amplifications occurring only in grade II and III tumours are related to tumour development, as only 1 grade I tumour was present in the study making significant conclusions difficult.

8.2.3. Gene Changes and Chromosome 17

Amplification of genes on the long arm of chromosome 17 was the most common amplification observed in this tumour cohort, as 71 % of breast cancer cell lines and 31 % of clinical samples analysed contained amplified genes on 17q (Chapter 7).

There are several areas within chromosome 17 that contain amplified genes, including 17q23, which is associated with an aggressive clinical phenotype and poor prognosis in breast cancer (Monni et al., 2001). Two separate regions of amplification within the 17q23 area have previously been recorded in human breast cell lines by FISH and one of these regions was frequently co-amplified with the ERBB2 region located at 17q11.2-q12 (Barlund et al., 1997). In this study co-amplification of the ERBB2 region and genes located on 17q23 was detected in BT474 cells and clinical tumour sample, 11 (Figure 7.1.). Some specific genes were associated with the 17q23 region in this study e.g. PPM1D, which is thought to have weak oncogenic properties (Li et al., 2002), was both amplified and overexpressed in BT474 and MCF7 cells.

Amplification of ERBB2 and several neighbouring genes at 17q11.2-q12 was detected in the BT474 cell line and 4 clinical tumour samples (31 %), in agreement with previous studies where amplification of the ERBB2 oncogene has been detected in 10-34 % of breast cancers (Revillion et al., 1998). Amplifications within this region have also been detected in other cancers (Varis et al., 2002; Hengstler et al., 1999), thus these genes may play a more general role in tumourgenesis and progression. The numbers of gene amplified within the region varied in each of the samples with BT474 containing the most (12), and so this area is susceptible to amplification though the size of the amplicon may vary. Interestingly 4 genes, PPARBP, CrkRS, ERBB2 and GRB7 were amplified in all 4 clinical tumour samples and
BT474. Thus these genes may lie at the centre of the amplicon. In all cases the NR1D1 gene was not amplified and so marked the upper boundary of amplification.

Several genes amplified in this region can also be overexpressed (Kauraniemi et al., 2001), unfortunately RNA data is not available for clinical tumours but 12 of the amplified genes in BT474 cells were also overexpressed here by either cDNA microarray or Real Time PCR. These genes as well as ERBB2 may have an active role in breast cancer development and progression. PPARBP, CrkRS and GRB7, which are co-amplified with ERBB2 in BT474 and the 4 clinical tumour samples were amongst those overexpressed and so may be of greater importance. ERBB2 amplification has been associated with poor prognosis and poor response to chemo therapy in breast cancer (Hengstler et al., 1999), however the clinical significance of ERBB2 amplification is debatable (Revillion et al., 1999). The reason for this discrepancy may depend on the other genes which are co-amplified and overexpressed with ERBB2, since it is clear the region of amplification varies between tumours and the increased biological activity of several genes not just ERBB2, may jointly contribute to the tumour phenotype and more aggressive clinical behaviour.

8.3. Future Work

It may be necessary to validate more of the microarray data by Southern blot or Real Time PCR as this was not carried out on all genes. However to date our microarrays have yielded a low incidence of false positive results, thus it may not be necessary to confirm every change for the results to be valid and trustworthy; especially if conclusions are based on changes in sets of genes rather than individual genes (Lockhart and Winzeler, 2000) and the microarrays are conducted as experimental replicates, as they are in this thesis.

As Dox is frequently used in chemotherapy the analysis of the drug resistant phenotype was carried out in Dox resistant cells that overexpressed the ABCB1 transporter. However to further increase our understanding of the genetic events occurring in a drug resistant phenotype other cells, which display a resistance to different compounds or via different mechanisms should be assessed. For example Glyoxlase 1 (GLO1), an enzyme that detoxifies methylglyoxal, has previously been found overexpressed in resistant cells when treated with Dox (Sakamoto et al., 2000), and this GLO1 overexpression was associated with resistance independent of ABCB1 gene expression, indicating that there are other mechanisms of resistance are involved.
Only 13 clinical breast tumour samples were analysed, so to increase the statistical significance of the data and its association with phenotype more samples should be assessed. LCM should be used to overcome the problem of heterogeneous cell populations and this coupled with PCR amplification protocols will allow the analysis of both genomic amplification and RNA expression with microarrays.

This thesis has also suggested specific genes that warrant follow up work. For example the \textit{OXTR} gene, which was overexpressed in both the Dox resistant cell lines HPGP and MDA16 (Chapter 3) may be a therapeutic target. In the human endometrial \textit{OXTR} positive cell line (COLO684) and the MDA-MB231 breast cancer cell line growth and proliferation is retarded in the presence of oxytocin via \textit{OXTR} (Cassoni et al., 1997) and due to the specific expression of oxytocin and \textit{OXTR} in normal physiology, overexpression of \textit{OXTR} in resistant cells may offer an option in drug development.

8.4. \textit{Summary}

This thesis has focused on the study of genetic and gene expression changes that accompany the development of phenotypic properties of breast cancer including drug resistance, tumour grade, type and hormone status. Clusters of genes have been elucidated in cells that are resistant (Chapter 3) or gained some resistance by a stepwise increase in Dox concentration (Chapter 4). The \textit{ABCB1} gene was overexpressed in both Dox resistant cell populations and it was also amplified in the resistant HPGP and MDA16 cells. Both HPGP and MDA16 cells showed similar patterns of gene expression compared to their progenitor cell lines. A number of genes that were overexpressed or suppressed as cells initially gain a Dox resistant phenotype were also identified.

Common regions of amplification associated with various phenotypic properties of tumours have been elucidated and specific genes have been implicated in the phenotype (Chapter 5-7). For example, 2 regions of frequent amplification specific to lymph node positive tumours were detected, 2p25 and 22q11. These have not previously been detected in breast cancer, and 2 genes were consistently found amplified in these regions, \textit{LPIN1} and \textit{DGSl}. Associations between \textit{ERBB2}, \textit{ER} and menopausal status were also found and amplifications in chromosome 17 linked with \textit{ERBB2} were elucidated. The chronological order of amplifications in breast cancer is unclear, but there is evidence both in this study and literature that in lymph node negative tumours gains at 8q occur before, and are necessary, for 1q gains.
This project has answered some of the asked questions that led to the study and in doing so has contributed to our general knowledge of the genetics of breast cancer and relationships of the genes to phenotype. It is clear that microarrays have the potential to elucidate regions of amplification and simultaneously analyse gene expression and that as the human genome is sequenced and fully mapped this technique has great potential to assist in the understanding and treatment of breast cancer.
References


Bamberger A. N., Methner C., Lisbola B. W., Stadtler C., Schulte H. M., Loning T. and Milde-Langosch K. "Expression pattern of the AP-1 family in breast cancer: association of...


Barlund M., Tirkkonen M., Forozan F., Tanner M. M., Kallioniemi O. and Kallioniemi A. "Increased copy number at 17q22-q24 by CGH in breast cancer is due to high level amplification of two separate regions." Genes, Chromosomes and Cancer 20: 372-376.(1997).


225


Suzuki A., Fukushige S., Nagase S., Ohuchi N., Satomi S. and Horii A. "Frequent gains on chromosome arms 1q and/or 8q in human endometrial cancer." Human Genetics 100(5-6): 629-636.(1997).


