Identification of Mesenchymal Genes Involved in Prostate Organogenesis by LongSAGE Analysis.

Brigid Orr
BSc. (Hons.), MSc.

Doctor of Philosophy
University of Edinburgh
2006
Declaration

Except where acknowledgement has been made by reference, the experiments detailed in this thesis were the unaided work of the author. No part of this work has previously been accepted for any other degree, nor is part of it being concurrently submitted in candidature for any other degree or qualification at this university or any other institute of learning.

Brigid Orr,  
2006
Abstract

Growth and development of the prostate are regulated by androgens and mesenchymal-epithelial interactions. Prostate growth starts in the embryo, continues throughout neonatal life, and is completed at puberty. The male and female embryonic urogenital sinus (UGS) both contain a prostate precursor, but due to the action of testosterone only the male will develop a prostate. Testosterone acts in concert with paracrine signalling from inductive mesenchyme to regulate epithelial growth and branching of the prostate. The ventral mesenchymal pad (VMP) is an area of inductive mesenchyme that is rich in regulatory growth factors of which few have been identified at present. Recently, the role of mesenchyme/stroma in cancer has been established and developmental factors may be involved in mediating the effects of cancer stroma. The aim of this thesis was to identify mesenchymal factors and characterise their expression and function in the developing prostate, and to investigate their expression in prostate cancer stroma.

The LongSAGE technique was used to generate a comprehensive transcriptional profile of the neonatal rat prostate rudiment. LongSAGE is an open-ended and unbiased gene-profiling method. Two libraries were made from tissues in the female UGS; one library was prepared from the whole female prostate rudiment and the other library was prepared from the inductive mesenchyme (VMP) isolated from within the rudiment. Several essential factors known to be involved in prostate development were identified in the libraries, including low abundance cDNAs such as AR and Fgf10. This suggested that the approach had sufficient sensitivity to identify key mesenchymal factors. The two libraries were compared, and the comparison was statistically analysed, highlighting genes that were VMP-enriched (P<0.05). Candidate mesenchymal transcripts were selected from the VSU and VMP Libraries by two different approaches, either because of their status as VMP-enriched (P<0.05) transcripts, or by an 'intuitive' approach, because the transcripts were associated with genes that are known members of developmental pathways and/or have been associated with prostate cancer. The expression and abundance of candidate transcripts were quantified by qRT-PCR in the male and female neonatal UGS. Subsequently, candidate transcripts were verified as VMP-enriched and were quantified during prostate development by Northern blot analysis. The protein distribution of selected candidates were localised within the neonatal rat UGS by immunocytochemistry, and the effect of testosterone treatment on the protein distribution was studied. To test the function of one of the candidates on prostate growth and development, recombinant protein was added to prostate rudiments grown in vitro. Finally, candidate transcripts were investigated in human prostate cancer associated fibroblast cells (CAFs) and normal prostate fibroblast cells (NPFs) by RT-PCR and Northern blot analysis.

SAGE analysis, qRT-PCR and Northern blot analysis identified six candidate transcripts as VMP-enriched; Dlk1, Notch2, Ptn Nell2, MMP2 and MMP14. The transcript expression of each candidate was most abundant in the developing rat prostate during the perinatal period. Proteins for Dlk1, Ptn and Notch2 were localised to mesenchymal cells of the neonatal VMP and ventral prostate (VP). Ptn expression was also associated with the basement membrane and cell-surface of the epithelial duct cells of the VP. Treatment of VP organs with recombinant DLK1 in vitro increased the organ size and
epithelial branching. Also, \textit{PTN}, \textit{NOTCH2}, \textit{MMP2} and \textit{MMP14} transcript expression was observed in CAFs and NPFs. \textit{PTN} and \textit{NOTCH2} showed a decrease in CAFs compared to NPFs suggesting a tumour-suppressive role.

In summary, a comprehensive gene profiling technique was used to identify mesenchyme specific/enriched transcripts in the developing prostate. The expression, distribution and function of candidate transcripts and proteins were investigated in the developing prostate and in prostate cancer stroma.
Acknowledgements

I would first of all like to express my gratitude to Dr. Axel Thomson, my supervisor, for giving me the opportunity to undertake this Ph.D., and for being a constant source of guidance, enthusiasm and motivation. Also, a huge thank-you to the other members of the ‘Prostate Group’. I am grateful to Dr. Griet Vanpoucke, for her patient teaching and supervision as the all-knowing post-doc, her friendly support and good humour, and the data acknowledged in this thesis. A priceless source of information and motivation! Also, many thanks to Cathal Grace for all his help in the lab, the never-ending organ dissections, organ culture experiments, and helping me to “take it easy”, and to Raymond Chan for the endless Mini-Preps needed to make the LongSAGE libraries.

Further individual thank-you’s to Frances Collins for her being a consistent source of molecular biology tips when I first tackled LongSAGE, to Mike Millar and Sheila Macpherson for all the hours on the Confocal, to Arantza Esnal for embedding those tiny organ cultures, and to Nina Hallmark for answering the wealth of “little thesis questions”, and cheering me on. Also, thank-you to the HGU Sequencing Facility, Edinburgh, for the superhuman amount of Sequencing and to Dr. Simon Hayward, University of Vanderbilt, USA, for supplying the CAF and NPF cells.

A collective thank-you to everyone I’ve worked with at the MRC and the University of Edinburgh, too numerous to mention, who have contributed to this project, and who have made completing a Ph.D., a happy and productive time.

Finally, a special thank-you to my family, who now know a bit too much about the prostate. I couldn’t have done it without you!

I dedicate this thesis to my aunt, Anne-Marie Andrews, to my grandmother, Laura Wight, and to my great aunt, Mary Orr.
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC/HRP</td>
<td>avidin biotin complex/horseradish peroxidase</td>
</tr>
<tr>
<td>AIS</td>
<td>androgen insufficiency syndrome</td>
</tr>
<tr>
<td>AMACR</td>
<td>α-methyl-acyl-coA racemase</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine 5′-triphosphatase</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>AS</td>
<td>anti-sense</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone morphogenetic protein</td>
</tr>
<tr>
<td>BLE</td>
<td>bladder epithelium</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BPH</td>
<td>Benign prostatic hyperplasia</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>CAF</td>
<td>cancer-associated fibroblast</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CPM</td>
<td>counts per minute</td>
</tr>
<tr>
<td>CO₂</td>
<td>carbon dioxide</td>
</tr>
<tr>
<td>DAB</td>
<td>deoxyadenosine triphosphatase</td>
</tr>
<tr>
<td>dATP</td>
<td>deoxyadenosine triphosphatase</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>double distilled water</td>
</tr>
<tr>
<td>dGTP</td>
<td>deoxyguanosine triphosphatase</td>
</tr>
<tr>
<td>DHT</td>
<td>Dihydrotestosterone</td>
</tr>
<tr>
<td>Dlk1</td>
<td>Delta-like 1</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's Modern Eagle Media</td>
</tr>
<tr>
<td>DNase</td>
<td>deoxyribonuclease</td>
</tr>
<tr>
<td>dNTPs</td>
<td>2′-deoxynucleoside-5′-triphosphates</td>
</tr>
<tr>
<td>dTTP</td>
<td>deoxytymidine triphosphate</td>
</tr>
<tr>
<td>32P- dCTP</td>
<td>deoxycytidine 5′-[32P] triphosphate triethylammonium salt</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diaminetetra-acetic acid</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>EF1α</td>
<td>Elongation factor 1 alpha</td>
</tr>
<tr>
<td>ERα</td>
<td>Oestrogen receptor alpha</td>
</tr>
<tr>
<td>ERβ</td>
<td>Oestrogen receptor beta</td>
</tr>
<tr>
<td>FGF</td>
<td>Fibroblast growth factor</td>
</tr>
<tr>
<td>FGFR</td>
<td>Fibroblast growth factor receptor</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>Frzd</td>
<td>Frizzled</td>
</tr>
<tr>
<td>FSH</td>
<td>Follicle stimulating hormone</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphatase dehydrogenase</td>
</tr>
<tr>
<td>GDNFRα1</td>
<td>Glial derived neurorophic factor receptor alpha 1</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>GLGI</td>
<td>Generation of Longer cDNA fragments from SAGE tags for Gene Identification</td>
</tr>
<tr>
<td>GnRH</td>
<td>Gonadotrophin hormone releasing hormone</td>
</tr>
<tr>
<td>Hgf</td>
<td>Hepatocyte growth factor</td>
</tr>
<tr>
<td>HSPG</td>
<td>Heparin sulfate proteoglycans</td>
</tr>
<tr>
<td>Igf</td>
<td>Insulin-like growth factor</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobases</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>Lgl1</td>
<td>Late gestation lung factor 1</td>
</tr>
<tr>
<td>LH</td>
<td>Luteinising hormone</td>
</tr>
<tr>
<td>Ltbp</td>
<td>Latent transforming growth factor beta binding protein</td>
</tr>
<tr>
<td>MgCl2</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>MMP</td>
<td>Matrix metalloproteinase</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>MPSS</td>
<td>Massively parallel signature sequencing</td>
</tr>
<tr>
<td>mRNA</td>
<td>Message ribonucleic acid</td>
</tr>
<tr>
<td>Nell2</td>
<td>Nel-like 2</td>
</tr>
<tr>
<td>NTP</td>
<td>Nucleoside 5'-Triphosphate</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Pdgf</td>
<td>Platelet derived growth factor</td>
</tr>
<tr>
<td>PMP</td>
<td>Peripheral myelin protein</td>
</tr>
<tr>
<td>Ptn</td>
<td>Pleiotrophin</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>RACE</td>
<td>Reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>S</td>
<td>Sense</td>
</tr>
<tr>
<td>SAGE</td>
<td>Serial analysis of gene expression</td>
</tr>
<tr>
<td>SAM</td>
<td>s-adenosylmethionine</td>
</tr>
<tr>
<td>Sfrp</td>
<td>Secreted frizzled protein</td>
</tr>
<tr>
<td>Shh</td>
<td>Sonic hedgehog</td>
</tr>
<tr>
<td>SM</td>
<td>Smooth muscle</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SM α actin</td>
<td>Smooth muscle alpha actin</td>
</tr>
<tr>
<td>SSC</td>
<td>sodium chloride/sodium citrate</td>
</tr>
<tr>
<td>T</td>
<td>Testosterone</td>
</tr>
<tr>
<td>Ta</td>
<td>annealing temperature</td>
</tr>
<tr>
<td>TBE</td>
<td>tris-borate EDTA</td>
</tr>
<tr>
<td>TBS</td>
<td>tris buffer saline</td>
</tr>
<tr>
<td>TBST</td>
<td>tris -buffer saline tween 20</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA -box binding protein</td>
</tr>
<tr>
<td>Tfms</td>
<td>Testicular feminized mouse strain</td>
</tr>
<tr>
<td>TF8</td>
<td>Transcription factor 8</td>
</tr>
<tr>
<td>TGFβ</td>
<td>Transforming growth factor beta</td>
</tr>
<tr>
<td>TGFβR</td>
<td>Transforming growth factor beta receptor</td>
</tr>
<tr>
<td>TGFi</td>
<td>Transforming growth factor beta induced</td>
</tr>
<tr>
<td>TIMPS</td>
<td>Tissue inhibitors of matrix metalloproteinases</td>
</tr>
<tr>
<td>Tm</td>
<td>melting temperature</td>
</tr>
<tr>
<td>TNFα</td>
<td>Tumour necrosis factor alpha</td>
</tr>
<tr>
<td>TRAMP</td>
<td>Transgenic adenocarcinoma mouse prostate</td>
</tr>
<tr>
<td>UGE</td>
<td>urogenital epithelium</td>
</tr>
<tr>
<td>UGM</td>
<td>urogenital mesenchyme</td>
</tr>
<tr>
<td>UGS</td>
<td>urogenital sinus</td>
</tr>
<tr>
<td>UGT</td>
<td>urogenital tract</td>
</tr>
<tr>
<td>UV</td>
<td>ultra violet</td>
</tr>
<tr>
<td>VAP1</td>
<td>vascular adhesion protein 1</td>
</tr>
<tr>
<td>VMP</td>
<td>ventral mesenchymal pad</td>
</tr>
<tr>
<td>VP</td>
<td>ventral prostate</td>
</tr>
<tr>
<td>VSU</td>
<td>Ventral mesenchymal pad/Smooth muscle/Urethral epithelium</td>
</tr>
</tbody>
</table>
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declaration</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables and Figures</td>
<td>xiv</td>
</tr>
</tbody>
</table>

## Chapter 1 Introduction

1.1 The Prostate
   - 1.1.1 The Adult Prostate | 2 |
   - 1.1.2 Development of the rat prostate | 3 |

1.2 Endocrinology of Prostate Induction and Growth | 5 |

1.3 Mesenchymal –Epithelial Interactions During Development
   - 1.3.1 The andromedin hypothesis and androgen action | 12 |
   - 1.3.2 Sexually dimorphic smooth muscle expression regulates prostatic organogenesis. | 12 |

1.4 Stromal -Epithelial Interactions in the Adult Prostate and Prostate Cancer. | 15 |

1.5 Molecular Control of Normal and Neoplastic Prostate Development
   - 1.5.1 Normal prostate development | 21 |
   - 1.5.2 Neoplastic prostate development | 25 |

1.6 Summary | 27 |

1.7 Project Aims | 28 |

## Chapter 2 Materials and Methods

2.1 Materials
   - 2.1.1 Equipment | 28 |
   - 2.1.2 Computer Software | 30 |
   - 2.1.3 Enzymes | 31 |
   - 2.1.4 Nucleic Acids | 31 |
   - 2.1.5 Chemicals | 32 |
   - 2.1.6 Bacterial Strain Genotypes | 32 |
   - 2.1.7 Bacterial Culture Media | 32 |
   - 2.1.8 Northern Blotting and Hybridisation Equipment | 32 |
2.1.9 Immunohistochemical Reagents 33
2.1.10 Organ Culture Media, Equipment 34
2.1.12 Standard Solutions (A to Z) 36

2.2 Methods
2.2.1 Total RNA Extraction 37
2.2.2 DNase Treatment of Total RNA 37
2.2.3 cDNA Synthesis 37
2.2.4 Polymerase Chain Reaction 38
2.2.5 qRT-PCR – LightCycler Analysis 38
2.2.6 Cloning of PCR Products 42
  2.2.6.1 Preparation of Insert DNA 42
  2.2.6.2 Ligation 42
2.2.7 Plasmid DNA Preparation 42
  2.2.7.1 Plasmid Mini-Preparation 42
  2.2.7.2 Restriction Endonuclease Digestion 42
2.2.8 Agarose Gel Electrophoresis 43
  2.2.8.1 DNA 43
  2.2.8.2 RNA- Formaldehyde gel 43
2.2.9 Northern Blotting and Hybridisation 43
  2.2.9.1 Northern Blotting 43
  2.2.9.2 In vitro transcription of $^{32}$P labelled DNA probes 44
  2.2.9.3 Hybridisation of Labelled DNA Probes to RNA Blot 44
2.2.10 Immunocytochemistry 45
  2.2.10.1 Paraffin Embedding of Tissue 45
  2.2.10.2 Tissue Immunohistochemistry 45
  2.2.10.3 Fluorescent Tissue Immunocytochemistry 46
2.2.11 In Vitro Organ Culture 46

2.3. LongSAGE 46
2.3.1. Materials 47
  2.3.1.1 Tissue Samples and RNA Preparation 47

2.3.2 Method
  2.3.2.1 SAGE Library Preparation 48
  2.3.3 Library Analysis 52
  2.3.3.1 Tag Quantification 52
  2.3.3.2 Statistical analysis 53
2.3.3.3 Tag-to-Gene Annotation 53
2.3.3.4 Literature Mining and Candidate Selection 53
2.3.3.5 Gene Ontology Assignment 54

### Chapter 3. Gene Expression Profiling and Serial Analysis of Gene Expression (SAGE).

#### 3.1 Introduction 55
- 3.1.1 Gene Expression Profiling Techniques 55
- 3.1.2 SAGE and LongSAGE 56
- 3.1.2.1 SAGE Bioinformatics 58
- 3.1.2.2 Databases of SAGE data 58
- 3.1.2.2.1 SAGEmap 60

#### 3.2 Results 61
- 3.2.1 Library Preparation 61
- 3.2.2 The VSU and VMP LongSAGE Libraries 62
- 3.2.3 Bioinformatics - Library Analysis 65
  - 3.2.3.1 Extracting the SAGE tag data 65
  - 3.2.3.2 Comparison of the VSU and VMP Libraries 66
  - 3.2.3.3 Statistical Analysis 67
  - 3.2.3.4 Tag-to-Gene Annotation 67
  - 3.2.3.5 Assignment to Gene Ontology 69
  - 3.2.3.6 Developmental Pathways and Regulatory Genes 71
- 3.2.4 Summary of Results 71

#### 3.3 Discussion of Library Preparation and Analysis 72
- 3.3.1 VSU and VMP Library Preparation 72
- 3.3.2 Bioinformatics - Library Analysis 73
  - 3.3.2.1 Extracting the Tag Data 75
  - 3.3.2.2 Tag-to-Gene Annotation 76
  - 3.3.2.3 Differential Tag Expression in the LongSAGE libraries 77
  - 3.3.2.4 Gene Ontology of VMP-enriched Tags 78
- 3.3.2.5 Selection of Genes of Interest 79

### Chapter 4. LightCycler Analysis and Selection of Candidate Genes

#### 4.1 Introduction 80

#### 4.2 Results 80
- 4.2.1 Selecting Genes for Further Analysis 80
- 4.2.2 Experimental Validation of LongSAGE Libraries 81
  - 4.2.2.1 VMP-Enriched Genes Verified by LightCycler Analysis 81
  - 4.2.2.2 LightCycler Analysis Identifies VSU-enriched Genes 82
4.2.2.3 LightCycler Analysis Identifies No Difference in Transcript Expression 83
4.2.2.4 Selection of Candidate Genes 84
4.2.3 Summary of Results 85

Chapter 5  Delta-Like 1 homolog

5.1 Introduction 87

5.2 Results 92
5.2.1 Delta-Like 1 Transcript Expression in the VMP and VSU 92
5.2.2 Dlk1 transcript levels in the prostate and selected organs 93
5.2.3 Dlk1 Protein Localisation in the UGS 93
5.2.4 The effect of Testosterone on Dlk1 distribution 95
5.2.5 Functional Studies on Dlk1 100
5.2.6 Summary of Dlk1 Results 102

Chapter 6.  Notch 2

6.1 Introduction 103

6.2 Results 107
6.2.1 Notch2 Transcript Expression in the VSU and VMP 107
6.2.2 Notch2 transcript levels in the prostate and selected organs 108
6.2.3 Notch 2 Protein Localisation in the UGS. 108
6.2.4 The effect of Testosterone on Notch2 distribution 110
6.2.5 Summary of Notch2 Results 113

Chapter 7.  Pleiotrophin

7.1 Introduction 114

7.2 Results 118
7.2.1 Ptn Transcript Expression in the VSU and VMP 118
7.2.2 Ptn transcript levels in the prostate and selected organs 119
7.2.3 Pleiotrophin Protein Localisation in the UGS 121
7.2.4 The effect of Testosterone on Ptn distribution 125
7.2.5 Functional studies on Ptn 127
7.2.6 Summary of Pleiotrophin Results 127
Chapter 8. Other Candidate Genes; Nell2, MMP2 and MMP14

8.1 Introduction 128
8.1.1 Nell2 128
8.1.2 MMP2 and MMP14 131

8.2 Results 135
8.2.1 Nell2, MMP2 and MMP14 transcript expression in the VSU and VMP 135
8.2.2 Nell2 transcript levels in the prostate and selected organs 136
8.2.3 MMP2 transcript levels in the prostate and selected organs 138
8.2.4 MMP14 transcript levels in the prostate and selected organs 138
8.2.5 Summary of results 142

Chapter 9. Candidate Expression in Prostate Cancer Stroma

9.1 Introduction 143

9.2 Results 145
9.2.1 Expression of candidate transcripts in the human prostate 145
9.2.2 PTN protein localisation in the human prostate 148
9.2.3 Summary of results 149

Chapter 10. Discussion

10.1 LongSAGE 150
10.2 Experimental Validation of the LongSAGE Libraries 151
10.3 Delta-Like 1 (Homolog) 153
10.3.1 Dlk1 transcript expression 153
10.3.2 Dlk1 Protein Expression in the Neonatal UGS 155
10.3.2.1 The Effect of Testosterone on Dlk1 Protein Distribution 156
10.3.3 Dlk1 Functional Studies 157

10.4 Notch2 158
10.4.1 Notch2 Transcript Expression 158
10.4.2 Notch2 Protein Expression 160
10.4.2.1 The Effect of Testosterone on Notch2 Protein Distribution 160
10.4.3 Notch2 Function in the Developing Prostate 161
10.5 Pleiotrophin
10.5.1 Pleiotrophin Transcript Expression 163
10.5.2 Pleiotrophin Protein Expression 164
10.5.2.1 Effect of Testosterone on Pleiotrophin Protein Distribution 165
10.5.3 Pleiotrophin Functional Studies 165

10.6 Nell2 167

10.7 MMP2 and MMP14 168

10.8 Candidate Expression in Prostate Cancer Stroma 170

10.9 Project Conclusions 172
10.9.1 A Novel Approach to Studying Prostate Development 172
10.9.2 How successful was the use of LongSAGE analysis? 173
10.9.3 How could the LongSAGE data be further mined? 174
10.9.4 LongSAGE analysis is a robust, reliable and interesting approach to gene expression profiling? 175
10.9.5 What does the LongSAGE data tell us about prostate development and prostate cancer? 176

10.10 Final Summary 178

Chapter 11 Bibliography 180

Appendix 219
List of Tables and Figures

Fig. 1.1 Schematic illustration of the anatomy of the prostate 3
Fig 1.2 Stages of early development of the ventral prostate lobe 5
Fig 1.3 The hypothalamic-pituitary-gonadal axis regulates prostate development 7
Fig 1.4 Summary of tissue recombination experiments between urogenital sinus mesenchyme and epithelium from Tfm and wild-type embryos 10
Fig 1.5 A model of ventral prostate induction and development. 14
Fig 1.6 Interactions between the epithelial and the stromal/mesenchymal components of the prostate during development and adulthood. 17

Table 2.1 Antibodies used for Immunohistochemistry 34
Table 2.2: Recombinant proteins used in organ culture experiments 35
Table 2.3 Table of commercial kits used 35
Table2.4: Rat Primers for RT-PCR, LightCycler Analysis and DNA probes 39
Table 2.5 Human Primers for RT-PCR, LightCycler Analysis and DNA probes 41
Fig: 2.1 The Prostate Rudiment 47
Table 2.6: Tissue collected and RNA used for LongSAGE libraries 48
Fig 2.2: The LongSAGE method 49

Fig 3.1 Steps for data analysis of SAGE Libraries 59
Table 3.1 LongSAGE protocol modifications 62
Table 3.2: Features of the VSU and VMP LongSAGE Libraries 62
Fig 3.2: Flow -Diagram of the LongSAGE protocol 63
Table 3.3: Summary of the VSU and VMP SAGE Library Data 66
Fig 3.3: Flow-diagram of the tag-to-gene annotation process 67
Table 3.4 Full VSU Library-saved to CD and appended at end of thesis 69
Table 3.5 Full VMP Library-saved to CD and appended at end of thesis 69
Table 3.6 Comparison of rat LongSAGEmap vs. BLAST for tag-to-gene annotation of VMP-enriched tags 69
Graph 3.1 Functional classification of VMP-enriched genes (P<0.05) from the LongSAGE Libraries 70
Table3.7 Regulatory and developmental genes from the LongSAGE libraries 71

Graph 4.1 SAGE Tags Verified to be VMP-enriched 82
Table 4.1 Genes identified as VSU-enriched by LightCycler analysis 83
Table 4.2 Genes that show no difference in expression between the VSU and VMP by LightCycler analysis 84
Table 4.3 Genes associated with candidates identified in the VSU and VMP Libraries 85
Fig 5.1: Structure of the Dlk1 and Delta Proteins

Fig 5.2: Dlk1 mRNA expression in the male and female rat UGS

Fig 5.3: Dlk1 mRNA expression during prostatic ontogeny and expression in other selected organs (P0)

Fig 5.4: Dlk1 expression in female and male rat UGS

Fig 5.5: Dlk1 and smooth muscle alpha actin expression in the female UGS

Fig 5.6: Dlk1 and smooth muscle alpha actin expression in the neonatal male UGS

Fig 5.7: The effect of testosterone on Dlk1 expression

Fig 5.8: The effect of Dlk1 on VP growth

Fig 6.1: Schematic diagram of the Notch2 protein

Fig 6.2: Notch2 mRNA expression in the male and female UGS

Fig 6.3: Notch2 mRNA expression during prostatic ontogeny and expression in other selected organs (P0)

Fig 6.4: Notch2 expression in Male and Female UGS and the Ovary

Fig 6.5: The effect of testosterone on Notch2 expression

Fig 7.1: Ptn mRNA expression in the male and female UGS

Fig 7.2: Ptn mRNA expression during prostatic ontogeny and expression in other selected organs (P0)

Fig 7.3: Pleiotrophin expression in the male and female UGS

Fig 7.4: Pleiotrophin and β-catenin expression in the male UGS

Fig 7.5: Ptn and p63 expression in the male UGS

Fig 7.6: The effect of testosterone on Ptn localisation

Fig 8.1: Schematic of NELL2 protein

Fig 8.2: Schematic of MMP2 and MMP14 Proteins

Fig 8.3: Nell2, MMP2, MMP14 and Fgf10 mRNA expression in the male and female UGS

Fig. 8.4: Nell2 mRNA expression during prostatic ontogeny and expression in other organs (P0)

Fig. 8.5: MMP2 mRNA expression during prostatic ontogeny and expression in other organs (P0)

Fig. 8.6: MMP14 mRNA expression during prostatic ontogeny and expression in other organs (P0)

Fig 9.1: Summary of CAF/NPF and BPH tissue recombination Experiments

Fig 9.2: Gene Expression in Human Foetal Tissues and Cells

Graph 9.1: PTN and NOTCH2 mRNA Expression in CAFs and NPFs.

Fig 9.4: Pleiotrophin protein expression in the human prostate

Appendix: VMP-upregulated genes (P<0.05) identified by LongSAGE analysis
1. Introduction

Growth and development of the prostate is mediated by androgens and mesenchymal interactions, and mesenchymal factors are essential for resulting growth and development (Cunha and Chung, 1981). The mesenchyme produces essential paracrine factors that act on the epithelium to stimulate growth in the developing prostate. The mesenchymal factors, fibroblast growth factor 10 (Fgf 10) (Thomson and Cunha, 1999) and fibroblast growth factor 7 (Fgf 7) (Sugimura et al., 1996) have been identified in prostate development, and several other possible candidates have been proposed. A highly localised area of condensed mesenchyme at the base of the bladder, called the ventral mesenchymal pad (VMP) has been identified as a source of the instructive signalling, and is a tissue rich in regulatory factors (Timms et al., 1995). The mesenchymal signals may be directly controlled by androgens, but to date, there is little experimental evidence to support this (Peehl and Rubin, 1995). Thomson et al., (2002) has proposed a novel mechanism by which the sexually dimorphic expression of the smooth muscle layer in the urogenital tract of the embryo, mediates the androgen control of signalling from the mesenchyme. This suggests that prostatic inductive factors are constitutively expressed in the VMP.

Mesenchymal factors from the adult prostate stroma act to maintain normal prostatic epithelial differentiation and homeostasis (Gao et al., 2001), and it has been established that the mis-expression of factors by the stroma may be responsible for directly epithelial mitogenesis during the early stage of prostate cancer (Tuxhorn et al., 2002). The identification of mesenchymal factors in the growing prostate would help to characterise a developmental pathway that may be reactivated in the disease state. This project aimed to identify developmental mesenchymal factors and to determine if they were expressed in neoplastic prostate development. Expression of these factors during prostate cancer may represent markers of disease or potential therapeutic targets.
Chapter 1

Introduction

1.1 The Prostate

1.1.1 The adult prostate

The prostate is a male accessory gland found only in mammals and functions by contributing secretory proteins and other factors to seminal fluid. Interest in studying the biology of the prostate is driven both by the fascinating nature of the developmental processes that give rise to the prostate, and by the high incidence in humans of prostatic diseases, including prostatic adenocarcinoma and benign prostatic hyperplasia. In the UK, BPH is found in approximately 1 in 3 men over the age of 60 and 65,000 BPH patients per year will undergo surgical treatment for this condition. Also, there are approximately 50,000 men living with prostate cancer, while approximately 10,000 men per year will die of prostate cancer, and about 27,000 new cases of prostate cancer will be diagnosed.

The prostate surrounds the urethra close to the base of the bladder and is composed of a tubuloalveolar gland arranged in lobules surrounded by stroma. Located deep within the pelvis, the gland is rich in nerves smooth muscle, collagen and lymphatics. In rodents the prostate (Fig 1.1, Panel B) is divided into morphologically and biochemically distinct lobes that are named according to their position: anterior (also called coagulating gland), dorsal, dorso-lateral and ventral. In contrast, the adult human prostate (Fig 1.1, Panel A) is a compact gland without distinct lobes. It is roughly the size and shape of a walnut and it exhibits three distinct zones: central, peripheral and transitional, reflecting three distinct sets of ducts (McNeal, 1980, 1983). Comparative observations of prostatic development in rodents and humans demonstrated that prostatic morphogenesis occurs in an analogous manner in both humans and rodents, with several distinct sets of epithelial buds growing out of the urethra into the UGM (Price, 1963, Timms et al., 1994). Nevertheless, compelling molecular evidence for homology between specific rodent prostatic lobes and human prostatic zones has yet to be identified.

Mature prostatic ducts contain three major cell types, luminal secretory epithelial cells, basal epithelial cells and stromal smooth muscle cells. Other less common cell types include fibroblasts, neuroendocrine cells as well as rare basal epithelial cells.
with unique marker expression profiles that are candidates for epithelial stem cells (Collins et al., 2001, Wang et al., 2001). In both the mouse and human, the ductal lumen is lined by tall columnar secretory epithelial cells, and the stromal layer, although largely composed of smooth muscle, also contains fibroblastic neuronal, lymphatic and vascular cell types (as reviewed by Marker et al., 2003).

Fig. 1.1 Schematic illustration of the anatomy of the prostate
The human prostate (Panel A) and rat perinatal prostate and associated reproductive organs (Panel B) (adapted from Abate–Shen, 2000 and Thomson et al., 2001, respectively).

1.1.2 Development of the rat prostate
Growth and development of the prostate begins in foetal life and are complete at sexual maturity. In all species the prostate develops from the endodermal urogenital sinus (UGS), which is derived from the cloaca. The endodermal UGS is an
ambisexual embryonic rudiment, which develops into the prostate, prostatic urethra and bulbourethral glands in males, and the lower vagina and urethra in females and, and into the bladder in both sexes (Staack et al., 2003). The endodermal UGS is surrounded by embryonic connective tissue called urogenital sinus mesenchyme (UGM). Before sexual differentiation of the UGS, UGM expresses androgen receptor in both sexes and thus acquires the capacity to undergo masculine development (Cooke et al., 1991).

In response to foetal testicular androgens, epithelial buds emerge from the wall of the UGS (~E18.5) in the rat, and grow into an area of the surrounding UGM, termed the ventral mesenchymal pad (VMP) (Fig 1.2, Panels A and B). The prostate contains highly branched epithelial ducts that grow at their tips during development (Sugimura et al., 1986). Branching morphogenesis starts in the perinatal period (rats and mice) (Fig 1.2 Panel C), and is completed by the end of puberty (Kinbara and Cunha, 1995, Sugimura et al., 1986). It has been demonstrated that 80% of branching morphogenesis is complete by Day 10 of neonatal life in mice (Donjacour and Cunha, 1988). Initially, the emerging prostatic buds and ducts are solid but undergo canalization in the neonatal period. Epithelial cells of the developing, solid prostatic cords are characterized by co-expression of cytokeratins 5, 8, 14, and 18 and p63 (Wang et al., 2001) Ductal canalization begins at the urethra and progresses distally towards the ductal tips. During ductal canalization the epithelial cells differentiate into luminal and basal cells. Basal epithelial cells become localized along the basement membrane to form a discontinuous layer of cells expressing cytokeratins 5 and 14 and p63 in the rat and mouse. Concomitantly, tall columnar luminal cells, which express cytokeratins 8 and 18, differentiate and line the ductal lumina. Secretory cytodifferentiation of the epithelium occurs postnatally in rodents, and prostate specific secretory proteins are initially detected in rats and mice at 12-20 days postnatal (Lopes et al., 1996). Prostatic epithelial differentiation is accompanied by differentiation of the mesenchyme into smooth muscle and fibroblasts (Hayward et al., 1996a, 1997) (reviewed by Marker et al., 2003). Growth of the epithelium is highly regulated forming a precise spatial pattern, and thus establish in rodents the four lobar subdivisions of the prostate. The different prostatic
lobes each have a characteristic ductal branching pattern (Kinbara and Cunha, 1995, Marker et al., 2003, Price, 1963, Sugimura et al., 1986), and also distinct secretions, which may be specified due to differences in the mesenchyme.

![Diagram of prostate development](image)

**Fig 1.2 Stages of early development of the ventral prostate lobe**
The rat prostate rudiment is similar in the embryonic male and female UGS (E17.5) (Panel A). In the male, epithelial buds emerge into the VMP in response to androgens (~E18.5) (Panel B). The buds continue to grow and branching occurs during early postnatal development of the rodent prostate (Panel C); 1) formation of the epithelial duct lumen (proximal to distal, relative to the urethra) 2) epithelial cell differentiation into basal and luminal cells (proximal to distal), 3) differentiation of smooth muscle from mesenchyme (proximal to distal), and 4) ductal growth localised to highly proliferative ductal tips. Androgen receptor (AR) expression is localised to the mesenchyme until the formation of growing epithelial ducts, when AR is also found expressed the epithelia (modified from Marker et al., 2003).

### 1.2 Endocrinology of Prostate Induction and Growth

The major steroids involved in induction and growth of the developing prostate, and the adult prostate, are androgens. Androgens masculinize the reproductive tract during the ambisexual stage to development and lead to the formation of the prostate.
Androgens are required to initiate prostatic development, to continue embryonic and neonatal prostatic growth, and subsequently to begin secretory activity at puberty. The role of androgens in the prostate changes over the course of a lifetime, and during adult life, androgens maintain the production of secretory proteins as well as maintaining the tissue architecture of the prostate. The levels of circulating androgens found in rats correlate with prostate development and production of prostatic fluid during adulthood; androgen levels are relatively high towards the end of gestation, decrease one day after birth and increase again during puberty (Corpechot et al., 1981), and are high in adult life.

Androgen production is under the control of the hypothalamic–pituitary–gonadal axis (Fig 1.3). GnRH from the hypothalamus acts on the pituitary gland to induce the release of the gonadotrophins follicle stimulating hormone (FSH) and luteinising hormone (LH). In the male, LH acts on the Leydig cells of the testis to stimulate synthesis of testosterone, the principal androgen of the testis, while FSH acts on the Sertoli cells of the testis to induce inhibin production in the presence of testosterone. Testosterone and inhibin exert a negative feedback regulation of LH and FSH production at the levels of the pituitary and hypothalamus (as reviewed in Coffey, 1988).

In addition to androgens, prostatic development is sensitive to levels of oestrogenic compounds. Most of the oestrogen in the male is derived from peripheral conversion of oestrogens through an enzymatic aromatization reaction. During neonatal life oestrogen levels exert many effects upon the growth and differentiation of the prostate (Jarred et al., 2002, Prins, 1992, vom Saal, 1997). These effects might be due to the alteration of androgen concentrations via the hypothalamic–pituitary–gonadal axis or via direct effects upon the prostate, although inactivating mutations in the ERα and ERβ receptors have no affect on normal prostate development (Jarred et al., 2002).

Testosterone functions as a prohormone in the prostate, and is converted into 5α–dihydrotestosterone (DHT) (the active form) by 5α–reductase. DHT becomes the
major androgen regulating the cellular events of growth differentiation and function in the prostate. DHT is bound more efficiently by the androgen receptor and is thus a more potent androgen; impairment of 5α-reductase reduces but does not completely inhibit prostatic development (Imperato-McGinley et al., 1985, Marks et al., 1999, Uygur, et al., 1998). Additionally, in a study of prostate specification in marsupials a second testicular androgen 5α-androstane-3α, 17β-diol (5α-adiol) was implicated as the circulating androgen that acts to specify the UGS, but further work is required to verify this finding. In this study testosterone was also found to contribute to epithelial growth through an action triggered by androgen receptor binding (Foster and Cunha, 1999).

Fig 1.3 The hypothalamic-pituitary-gonadal axis regulates prostate development

The systemic androgens produced by the foetal testis act to specify development of the UGS into prostate. Cellular response to systemic androgens is mediated by nuclear androgen receptors (AR), activated by testosterone (T) or dihydrotestosterone (DHT). The requirements for and sufficiency of androgens in
establishing prostate identity in the UGS is shown by the absence of a prostate in mice and human that lack functional androgen receptors due to inactivating mutations (Brown et al., 1988, Charest et al., 1991, Gaspar et al., 1991, He et al., 1991, Lubahn et al., 1989), and by the development of a prostate in female urogenital sinuses exposed to androgens (Takeda et al., 1986).

During prostatic induction, AR is expressed in mesenchymal cells but is absent initially from epithelial cells (Takeda et al., 1985). AR is first observed in mesenchymal cells, preceding the appearance of prostatic epithelial buds (see Fig 1.2, Panel A, B). Once prostatic epithelial buds have formed and begun to develop, AR is observed in these growing epithelia (Hayward et al., 1996a)(see Fig 1.2, Panel C). Androgen signalling is both necessary and sufficient for prostatic organogenesis and epithelial androgen receptor is not required for development of the prostate (Cunha and Chung, 1981).

1.3 Mesenchymal–Epithelial Interactions During Development

In addition to androgens, mesenchymal–epithelial interactions play a key role in directing the growth and development of the prostate, with morphogenesis and differentiation of both the epithelium and mesenchyme aborted if the epithelium and mesenchyme are grown separately.

Paracrine signalling from the mesenchyme to the epithelium is essential for prostatic development. During prostatic development the UGM; (a) specifies prostatic epithelial identity, (b) induces epithelial bud formation, (c) elicits prostatic bud growth and regulates ductal branching, (d) promotes differentiation of a secretory epithelium and, (e) specifies the types of secretory proteins which are expressed (Hayashi et al., 1993, Marker et al., 2003). Also, although epithelial development is regulated by the mesenchyme, the interaction is reciprocal, and it is likely that epithelial signalling regulates differentiation of the mesenchyme into ‘stroma’, composed of smooth muscle and fibroblasts (Hayward et al., 1996a). This has been
Chapter 1

Introduction

demonstrated by the ability of different prostatic epithelia to elicit species-specific patterns of stromal development (Hayward et al., 1998b). Tissue recombinants composed of adult human prostatic epithelium and rat urogenital sinus mesenchyme were grafted beneath the renal capsule of athymic rodent hosts. The rat mesenchymal component differentiated into thick sheets of smooth muscle characteristic of the human but not the rat prostate.

The action of androgen on mesenchymal cells is known to regulate epithelial mitogenesis and morphogenesis and which is suggested to be mediated by paracrine factors made by the mesenchyme, via an unknown mechanism. The role of mesenchymal–epithelial interactions in urogenital development has been studied extensively using grafting and tissue recombination methods in Tfm mice (Cunha, 1996, Cunha and Chung, 1981, Cunha et al., 1987, Cunha and Lung, 1978, Hayward et al., 1997). Tfm mice lack a functional androgen receptor this results in loss of prostatic development in mice carrying the mutation (Quigley et al., 1995). In tissue recombinants made with epithelium (UGE) lacking androgen receptor function (Tfm) combined with mesenchyme (UGM) containing “wild-type” androgen receptor in the presence of androgens, the AR-deficient epithelium underwent androgen–dependent ductal morphogenesis, epithelial proliferation and columnar cytodifferentiation thus forming glandular epithelium resembling prostate (Fig 1.4). However, tissue recombinants made with mesenchyme lacking androgen receptor function (Tfm) and epithelium containing wild-type androgen receptor did not form prostatic tissue. These experiments demonstrated that many androgenic effects expressed in epithelium do not require epithelial AR, and are in fact elicited by the paracrine action of AR-positive mesenchyme. Also, androgen receptor signalling in mesenchymal cells is both necessary and sufficient for prostatic development (Cunha and Chung, 1981). Further analysis of Tfm/wild–type recombinants revealed that epithelial ARs are required for expression of AR-dependent secretory proteins (Cunha and Young, 1991, Donjacour and Cunha, 1993)
Fig 1.4 Summary of tissue recombination experiments between urogenital sinus mesenchyme and epithelium from Tfm and wild-type embryos.

A positive androgenic response (prostatic morphogenesis) occurs when wild-type mesenchyme is grown in association with either wild-type or Tfm epithelium. Conversely, no prostate structures are formed when either wild-type or Tfm epithelium is grown in association with Tfm mesenchyme. These results demonstrate that androgens elicit many of their effects on epithelial development via mesenchymal androgen receptors (redrawn from Cunha et al 1983).
The UGM, in the presence of androgens, determines epithelial identity so that subsequent cell–cell interactions can elicit prostatic bud outgrowth, ductal branching and prostatic differentiation. The ability of UGM to specify prostatic epithelial identity was discovered by analysis of a tissue recombinants composed of UGM plus epithelium of embryonic or fully differentiated adult urinary bladder, in which bladder epithelium was induced by UGM to undergo prostatic differentiation. Formation of prostatic tissue can be explained by two possible options; either the adult BLE was terminally differentiated and was maintained or reprogrammed by paracrine cues from the mesenchyme/stroma, or the BLE contained uncommitted ‘stem cells’ that were induced to form the prostate. Studies so far have been unable to identify molecules produced by the UGM that induce prostatic epithelial differentiation (reviewed by Cunha et al., 2004).

The mesenchyme involved in prostatic induction includes the peri –urethral mesenchyme and ventral mesenchymal pad (VMP), a condensed pad of mesenchyme peripheral to the urethral epithelium that is found in both males and females. The VMP was initially characterized by Timms et al (1995) and has been shown to function as an inducer of prostatic organogenesis in vivo, and is located adjacent to the site of epithelial bud outgrowth and branching. The VMP appears to constitutively express mesenchymal factors (i.e. not androgen-regulated) important for prostatic growth. For example, fibroblast growth factor (Fgfl0), a key mesenchymal regulator of prostate development, is expressed in the VMP of males and females and is not directly regulated by testosterone in vitro (Thomson and Cunha, 1999).

The finding that prostatic development involves epithelial proliferation and differentiation has led to the hypothesis that there are mesenchymal paracrine –acting factors that are regulated by androgens. The androgen–regulated factors are not made in epithelial cells, despite the presence of the androgen receptor. In epithelial cells, the function of the androgen receptor is most likely the regulator of secretory proteins (Donjacour and Cunha, 1993) and perhaps epithelial differentiation. Furthermore, AR expression in the mesenchyme of the early prostate prior to bud
induction and the later appearance of AR in the growing epithelial cells supports the idea that prostatic induction and early prostate growth are mediated by mesenchymal signals.

1.3.1 The andromedin hypothesis and androgen action
As androgens have been proposed to regulate prostatic growth and development by action on the mesenchyme, research has been focused on the identification of paracrine factors that could mediate androgen action. The precise mechanism of paracrine regulation of the prostate is not known, but a current hypothesis is that androgens might regulate the activity of paracrine-acting factors made by the mesenchyme, which regulates epithelial development. At present, mesenchymal paracrine regulators of prostatic growth have been identified (e.g. Fgf7, Fgf10 and Igf1) but androgens do not appear to directly regulate the genes for Fgf7 or Fgf10 (Thomson and Cunha, 1999, Thomson et al., 1997), though other studies have suggested that these factors may be androgen–regulated (Lu et al., 1999, Yan et al., 1992). At present no growth factors, expressed in the mesenchyme, have been shown unequivocally to be directly regulated by androgens (as reviewed by Thomson, 2001). Therefore, at present it appears that the expression of paracrine factors may not be androgen-regulated and alternative mechanisms by which androgens might control development have been investigated.

1.3.2 Sexually dimorphic smooth muscle expression regulates prostatic organogenesis.
An alternative mechanism for androgen regulation of prostatic induction proposed by Thomson et al (2002) has examined the role that differentiation of smooth muscle might play in regulating prostatic induction. They proposed that androgens control prostatic induction by regulating differentiation of smooth muscle within the UGM, and consequently the signalling between mesenchyme and epithelium.

Smooth muscle appears in the rat UGS at ~R15 is formed by the differentiation of mesenchymal cells, and appears to be regulated by epithelial signals. The nature of the epithelial to mesenchymal signalling involved in SM differentiation is not fully
understood, although members of the Tgfβ family (Peehl and Sellars, 1998), Sonic hedgehog (Shh) (Ramahlo -Santos et al., 2000, Sukegawa et al., 2000) and Pod1 (Hidai et al., 1998, Lu et al., 1998) have been implicated.

Thomson et al., (2002) have proposed that differentiation of the smooth muscle layer is androgen–regulated and comprises part of the mechanism regulating prostatic induction. The hypothesis is supported by several observations. Firstly, they demonstrated that in the developing UGS, a layer of smooth muscle differentiates between the VMP and urethral epithelium, and there is a sexually dimorphic difference in the development of this layer. By serial reconstruction of the developing UGS, they showed that the SM layer forms at approximately E20.5 in female but did not form in males. Also, testosterone was shown to reduce the thickness of the SM layer in cultures of female UGS. Plus, rare prostatic buds present in some female UGS were stimulated to start the process of prostatic organogenesis by testosterone treatment, depending upon the bud position relative to the smooth muscle layer.

A diagram of their hypothesis is outlined in Figure 1.5 and will be discussed more fully here. During embryonic development (E17.5), the VMP is separated from the urethral epithelia by a layer of smooth muscle in males and females (Fig 1.5, Panel A). This smooth muscle layer is less well developed in males, and prostatic buds can pass through a gap in the SM layer to enter the VMP (Fig 1.5, Panel B). However, in females, after the SM layer has formed rare prostatic epithelial buds formed cannot enter the VMP to grow and develop in response to androgens (Fig 1.5, Panel C). Androgens have been found to partially inhibit development of the smooth muscle layer, modulate smooth muscle differentiation and lead to sexual dimorphism of this SM layer.
Fig 1.5 A model of ventral prostate induction and development. Prostate development from embryonic to adult life, outlining the smooth muscle hypothesis (Thomson et al., 2002) for prostatic induction and highlighting specific regulatory growth molecules involved in prostate organogenesis (as reviewed by Marker et al., 2003). During the initial stages of prostatic induction (A), the UGS is similar in the male and female, and the SM layer is discontinuous with signalling between the VMP and urethra occurs. At E18.5, prostatic buds become visible, in males. In the female at E20.5 (C), a SM layer has formed and interaction between the VMP and urethra is prevented. Residual buds may be present, but these do not enter the VMP and eventually regress. In the male at E20.5 the UGS has a discontinuous SM layer, prostatic buds have emerged from the urethra and entered the VMP, where subsequent growth and branching morphogenesis takes place (D). Developmental factors localised to the mesenchyme and/or epithelial compartments of the P0 prostate are important in prostate organogenesis and may be mis-expressed during prostate carcinogenesis (modified from Thomson et al., 2002). The smooth muscle may act as a barrier to the passage of inductive factors, or to inhibit contact between mesenchyme and epithelium. These data suggest that smooth muscle can act as a regulator of ductal elongation and branching morphogenesis, and that this peri-urethral smooth muscle is a barrier of prostatic ductal elongation and branching morphogenesis (Thomson et al., 2002).
1.4 Stromal-Epithelial Interactions in the Adult Prostate and Prostate Cancer.

Stromal-epithelial interactions are important, not only during prostatic growth and development (Fig 1.6, Panel A), but also in adult life. In the adult prostate (Fig 1.6, Panel B), under the direct or indirect action of androgens, mesenchymal factors from the stroma/smooth muscle (which express AR) act to maintain normal prostatic epithelial differentiation and homeostasis (Cunha et al., 1996). Homeostasis of the adult prostate is maintained through reciprocal interactions between the stroma and epithelium, with minimal proliferation of either cell type (Cunha et al., 2001). However, it is proposed that disruption of this homeostatic state, by alteration of the stromal and/or epithelial compartments, could affect their reciprocal relationship, to initiate and promote carcinogenesis. Also, it is likely that prostatic carcinogenesis is initiated by prostatic damage to prostatic epithelium.

It is hypothesized that prostate carcinogenesis is associated with changes in the local interactions between the stromal microenvironment and genetically initiated incipient tumour cells, and these changes are suggested to promote malignant tumour progression resulting in tumorigenesis (Fig 1.6, Panels B, C and D) (Cunha et al., 2004, Grossfield et al., 1998, Hayward et al., 1996, 1997, 1998). A model of prostate tumorigenesis proposed (reviewed in Cunha et al., 2004) is that following a genetic insult to the prostatic epithelium, the epithelium fails to signal appropriately to the adjacent stroma (smooth muscle), which in turn begins to differentiate towards a more fibroblastic phenotype (Wong and Tam, 2002a). As smooth muscle begins to de-differentiate, signalling from prostatic smooth muscle to prostatic epithelium becomes anamolous, resulting in progressive loss of control over epithelial differentiation and proliferation. Accordingly, a "vicious cycle" is established during progression in which both prostatic epithelium and smooth muscle de-differentiate and proliferate. This hypothesis is supported by many observations. Firstly, the highly differentiated growth quiescent state of the normal adult prostate is maintained by intimate stromal-epithelial cell interactions. Also, prostatic smooth muscle undergoes de-differentiation during carcinogenesis (Cunha et al., 2002). Furthermore, as prostate epithelium undergoes neoplastic transformation it loses its ability to maintain and induce smooth muscle differentiation (Wong and Tam,
2002b). This was demonstrated in experimental tissue recombinants in which various normal or neoplastic prostatic epithelia were grown in combination with UGM. It was found that only the normal (non-neoplastic) epithelia were fully capable of inducing differentiation of smooth muscle in the UGM (Hayward et al., 1996). This observation is consistent with the idea that one of aspect of the carcinogenic process is a loss of the ability of the epithelium to induce and maintain smooth muscle differentiation. It is now evident that smooth muscle–epithelial interactions play key roles in regulating epithelial differentiation, proliferation and carcinogenesis (reviewed by Cunha et al., 2004)

Prostate carcinogenesis is regulated not only by stromal–epithelial interactions but also by the action of steroid hormones (Fig 1.6, Panels C and D). The pivotal role of androgens and oestrogens will not be discussed in detail, and is included only to provide a more comprehensive picture of the factors involved. The importance of androgens in prostate carcinogenesis has been determined conclusively by the fact that prostate tumours do not develop in the absence of testosterone, as in eunuchs castrated early in life. Also, with relevance to the action of stromal factors in prostate tumorigenesis, there is strong supportive evidence that some of the stromal signals are likely to be mediated by androgens, and the presence of stromal AR has been verified (Cunha et al., 2003, Wang et al., 2001b).

Therefore, prostatic stroma plays a key role in the regulation of prostatic growth and differentiation, and also prostate cancer (Cunha et al., 1996, Cunha et al., 2002, Cunha et al., 2003). The importance of the stroma in human prostatic tumorigenesis has been clearly established, and human and rodent models of stromal effects upon prostatic neoplasia have been developed (reviewed in Hayward, 2002, Matrisian et al., 2001). However, the mechanisms by which stromal cells influence tumorigenesis are poorly understood, although it is proposed that prostatic stroma may affect prostatic carcinogenesis in many different ways, for example, by adopting a wound–like phenotype or by stimulation of angiogenesis (reviewed in Tuxhorn et al., 2001).
Fig 1.6 Interactions between the epithelial and the stromal/mesenchymal components of the prostate during development and adulthood.

During prostatic development (A) low levels of androgens, acting through the mesenchymal androgen receptor (AR), stimulate proliferation and differentiation of prostatic epithelium. Concurrently, the epithelium induces differentiation of the mesenchyme into smooth muscle. In the normal adult (B) high levels of circulating androgens, acting through AR located in the smooth muscle, maintain the morphology of growth-quiescent adult epithelium. Secretory function is elicited by direct androgenic stimulation of AR in the differentiated columnar epithelium. Epithelial and smooth muscle differentiation is maintained by reciprocal paracrine-acting homeostatic factors. In early tumors it is hypothesized that following genetic insult to the epithelium (C), signaling between the epithelium and the adjacent smooth muscle begins to become abnormal, leading (D) to the formation of a fibroblastic "tumor stroma", which responds to androgenic stimulation by producing paracrine-acting mitogens, fueling a cycle of tumor proliferation and stromal de-differentiation. (redrawn from Cunha et al., 2002).
Tumour stroma is histologically different from normal stroma (Ronnov et al., 1996, Tisty and Hein, 2001) and shows enhanced proliferation and migratory behaviour in vitro (Schor et al., 1985, 1988). Also, the constituents of the ECM and the vasculature in tumour stroma are also different from that associated with normal epithelia (Carmeliet and Jain, 2000). The tumour stromal cells are known to produce a variety of MMPs, which may affect tumour initiation, growth, migration, angiogenesis, apoptosis, invasion and metastasis (Lynch and Matrisian, 2002). In comparison to normal stromal cells, tumour stromal cells exhibit alterations in their cell–surface molecules (Chaudhuri et al., 1975, Oishi et al, 1981), altered expression of growth factors (Ellis et al., 1994, Frazier et al., 1997, Nakamura et al., 1997, Ponten et al., 1994, Uan et al., 1993, Yee et al., 1989), expression of prostaglandin synthase enzyme (Shattuck-Brandt et al., 1999, 2000) and alterations in the ECM (Pupa et al., 2002, Werb et al., 1996).

The first studies to demonstrate that stroma could regulate neoplastic prostate epithelia were conducted over a decade ago, and showed that embryonic mesenchyme could regulate epithelial proliferation and differentiation of rat Dunning tumour (Hayashi and Cunha, 1991, Hayashi et al., 1990). Tissue recombination experiments combined perinatal UGM or neonatal seminal vesicle mesenchyme (SVM) or neonatal bulbourethral mesenchyme (BLM) or with androgen-dependent Dunning prostate carcinoma (DT). The resultant issue combinations were grown as subcapsular renal grafts. Grafts of the slow growing DT increased only marginally in size in their hosts and comprised tubules with simple undifferentiated epithelium. When grown in association with BLM, the DT continued to show its characteristic differentiation and showed little growth. By contrast, tissue combinations of UGM + DT, SVM + DT and BUG-M + DT enlarged many fold in size. DT carcinoma cells in combination with UGM, SVM and BUG-M, were induced to undergo morphogenetic changes by forming larger more regularly organized ducts, b) differentiated in tall polarized columnar epithelial cells, and c) produced secretion which accumulated in the lumen when grown in male hosts. These results demonstrated a continued responsiveness of carcinoma cells to mesenchymal
inductors that can induce secretory cytodifferentiation and elicit reduction in growth rate and loss of tumorigenesis (Hayashi et al., 1990).

More recently, data indicate that stromal signals play a key role in directing epithelial mitogenesis during early (Olumi et al., 1999, Tuxhorn et al., 2002) and late stages of prostate cancer (as reviewed in Bhowmick et al 2004b and Cunha et al., 2004). To demonstrate this, cultured human fibroblast cells from normal or tumour prostate were cultured together with immortalized epithelial cells derived from benign prostatic hyperplasia. Tumour fibroblasts but not normal prostatic fibroblasts were found to stimulate epithelial proliferation and malignant transformation. However, tumour fibroblasts were found not to stimulate the growth of normal epithelial cells under identical conditions, suggesting that cancer associated fibroblasts express ECM proteins and growth factors that influence the incipient tumour cells and promote the angiogenesis necessary to maintain epithelial transformation (Olumi et al., 1999, Tuxhorn et al., 2002).

Factors produced by the cancer associated stroma may be involved in prostate cancer by stimulating epithelial mitogenesis in at least two ways; either by aberrant re-expression in the stroma (paracrine action (eg Tuxhorn et al., 2002)) or by mis-expression in epithelia (autocrine action (eg Gao et al., 2001)). Gao et al., (2001), showed that mis-expression of mesenchymal genes in epithelia may contribute to a switch from paracrine (normal) growth control to autocrine (cancer) growth control. They proposed that normal adult epithelial proliferation is mediated by stroma and that at some point during tumorigenesis, there is a direct control of epithelial proliferation by androgens (Gao et al., 2001). Additionally, the stromal cells may influence the incidence and growth rate of prostatic tumour cells through direct interaction and recruitment of endothelial cells, regulating angiogenesis (Tuxhorn et al., 2002).

Therefore, an important avenue in prostate cancer research has been the identification of cancer-associated stromal factors. For example, in a recent study to identify differential gene expression in breast tumour in comparison to normal breast,
two chemokines, CXCL12 and 14, were identified as over-expressed in myofibroblasts and myoepithelial cells, respectively (Allinen et al., 2004). CXCL12 and CXCL14 were reported to bind to receptors, acting on adjacent epithelial cells and causing enhanced proliferation and invasion. More recently, CXCL-12 was identified in prostatic stroma fibroblasts, with higher expression levels in aging men (>63 years old) in comparison to younger men, and this induced cellular proliferation (Begley et al., 2005). This provides a more mechanistic elucidation of the role of ageing in the etiology of prostate cancer.

In the most recent studies, it has been shown that normal epithelial cells without genetically engineered mutations can be induced to form carcinomas by genetically altered fibroblasts (Bhowmick et al., 2004, Kuperwasser et al., 2004). In previous studies, epithelial cells induced to be more carcinogenic by association with genetically altered fibroblasts (by irradiation) were already fully transformed into carcinomas (Ohuchida et al., 2004), or had a known p53 gene mutations (Barcello-Hoff and Ravani, 2004). This new finding of the capacity of genetically altered fibroblasts to induce carcinomas in normal epithelial cells was demonstrated by genetic modification of a member of the TGFβ signalling pathway. TGFβ has been reported to exert both tumour suppressive and tumour promoting functions. However, recent studies have provided further supportive evidence for a major role for TGFβ signalling in fibroblasts in the initiation of carcinomas. In one study, mice were generated in which the TGFβ type II receptor (Tgfr2) was inactivated specifically in stromal fibroblasts, and the mice showed prostatic intraepithelial neoplasia, a presumed forerunner of prostatic carcinoma, as well as invasive squamous cell carcinomas of the forestomach (Bhowmick et al., 2004b). The Tgfr2 knock-out (Tgfrβfopko) mouse model demonstrates that the TGF-β signalling pathway known to suppress cell-cycle progression and tumour formation when acting on epithelial cells can also indirectly inhibit epithelial proliferation when acting on adjacent stromal fibroblasts in vivo. Consequently, loss of this pathway in fibroblasts results in increased epithelial proliferation and may also promote invasive carcinoma in some tissues. This also suggests that pre-malignant and malignant epithelial tumours observed in the Tgfr2 knock-out (Tgfrβfopko) model can develop in tumours
owing to mutations in fibroblasts preceding any subsequent tumorigenic changes in the epithelial cells. Additionally, it was noted that the apart from the prostate and forestromach, other tissues in the Tgfbr2f-in mouse showed no signs of carcinogenic transformation. Therefore, from these results it was proposed that not all fibroblasts are created equal, and there may be differences in fibroblasts within tissues and between tissues, which have distinct functions (as reviewed by Bhowmick et al., 2004b).

Finally, with regard to the accumulating evidence that stromal– epithelial interactions affects the behaviour of tumour cells, stromal therapy has been proposed as a viable approach to cancer prevention and intervention. Targeting these mechanisms offer a fundamental new direction and potential for future cancer therapies. In principle three main points of attack are conceivable: 1) targeting the tumour signals that are responsible for the development of “tumour stroma”; 2) targeting CAF signals that initiate and promote tumour growth and facilitate invasion and metastases; and 3) eliminating the CAFs themselves to abolish their interaction with other fibroblast and epithelial cells. However, with the heterogeneity of tumour-stroma interaction, it is questionable whether one single factor is the “magic bullet” to halt tumorigenesis (as reviewed by Micke and Ostman, 2004).

1.5 Molecular Control of Normal and Neoplastic Prostate Development

1.5.1 Normal prostate development
Several genes and gene families have been suggested to play a role in prostatic development (reviewed by Marker et al., 2003). Functional studies of both spontaneous and engineered mutations in mice and humans that have provided direct evidence for the roles of specific genes in prostatic development. These genes include those associated with epithelium–specific factors Sonic hedgehog (Shh) (Podalesk et al., 1999 a), Nkx3.1 (Bhatia–Gair et al., 1999, Schneider et al., 2000, Tanaka et al., 2000) and p63 (Signoretti et al., 2000), factors expressed in both the
mesenchyme and epithelium, HoxA10 (Podalesk et al., 1999b), Hox A13 (Podalesk et al., 1999c, Warot et al., 1997) and Hox D13 (Podalesk et al., 1997, Warot et al., 1997), and mesenchyme-specific factors AR (Brown et al., 1988, Lubahn et al., 1989, He et al., 1991, Charest et al., 1991 and Gaspar et al., 1991) and BMP4 (Lamm et al., 2001).

The epithelial factor Nkx3.1 is expressed during embryogenesis and appears to demarcate prospective prostate epithelium prior to prostate formation, and continues to mark prostate epithelium during neonatal development. Nkx3.1 null mutants display defects in ductal morphogenesis and secretory protein production. Regulation of Shh expression in the epithelium is proposed to be upstream of Nkx3.1 since Shh-/- mice fail to express Nk3.1 transcripts (Schneider et al., 2000). Shh is synthesised in epithelial cells and in many situations, acts as a paracrine factor through its receptor Patched1 (Ptc), that is expressed in adjacent mesenchymal cells. The exact role of Shh on prostate development remains somewhat controversial as to whether Shh is androgen-regulated and is required for prostatic induction (Podalesk et al., 1999a) or whether Shh expression is not androgen-dependent and is more important in later stages of prostate development (Freestone et al., 2003). Both studies report that inhibition of SHH-signalling reduced growth of the prostate, however, it has been reported that SHH is not critical for prostatic induction, but plays important roles in prostatic growth, branching morphogenesis and epithelial differentiation (Berman et al., 2004, Freestone et al 2003). Inhibition of SHH-signalling in vitro led to the precocious and aberrant differentiation of epithelial cells, resulting in lesions similar to human cribriform prostatic intraepithelial neoplasia (PIN). The third epithelial-derived factor highlighted here, p63, has been discussed previously, since its expression becomes restricted to the basal epithelial cells during prostate development (reviewed in Marker et al., 2003). p63 expression has been reported to be essential for differentiation of prostatic basal epithelial cells, but not for differentiation of the prostatic luminal epithelial cells or neuroendocrine cells (Kurita et al., 2004). Additionally, grafts of p63-/- mouse embryo UGS transplanted into adult male mice developed into prostate, demonstrating that p63 is not essential for development and regeneration of the prostate (Kurita et al., 2004).
Chapter 1

Introduction

The homeobox containing transcription factors HOXA10, HOXA13 and HOXD13 are expressed in both the prostatic mesenchyme and epithelium. HOXA13 and HOXD13 are both required for normal morphogenesis of the prostate and act in a partially redundant fashion. Hoxa13+/− mice have small prostates with reduced branching morphogenesis of the VP and DLP (Podalesk et al., 1999c). Hoxd13 −/− mice also have smaller VP and DLP with reduced DLP branching morphogenesis (Podalesk et al., 1997). Compound Hoxa13 +/−Hoxd13 −/− mutant mice have more severe failure of prostate development including the absence of the AP (Warot et al., 1997). Hoxa10 −/− mice have reduced AP branching and partial AP to DLP transformation (Podalesk et al., 1999).

As for the mesenchyme, the VMP as previously mentioned, appears to constitutively express mesenchymal factors important for prostatic growth. The VMP is present in the embryonic UGS of both sexes but few factors specific to the mesenchyme have been identified so far. The importance of functional AR expression in the mesenchyme of the developing prostate has already been discussed. No prostate is formed in the absence of functional AR expression in the Tfm mouse (Charest et al., 1991, Gaspar et al., 1991, He et al., 1991) or in the human male with androgen insufficiency syndrome (AIS) (Brown et al., 1988, Lubahn et al., 1989).

Additionally, signals derived from the mesenchyme have been reported to instruct of the epithelium to grow and branch. Branching morphogenesis is a common process in mammalian organogenesis, and gives rise not only to the excretory epithelia of the prostate gland, but also to the salivary and mammary glands, the airways of the lungs and urine-collecting ducts of the kidney. As branching organs exhibit many anatomical features, matrix requirements and signalling pathways, it has been suggested that some of the branching mechanisms are conserved. A ‘consensus’ model of the mechanisms and molecules regulating branching epithelium has been proposed (Davies, 2002), although it is recognised that it is possible that important features of branching remain to be discovered that are quite different in different systems.
Some of the growth factors reported to be active in branching morphogenesis, and in prostate development include fibroblast growth factors (FGFs) (as reviewed by Thomson, 2001), epidermal growth factors (EGFs) (McKeehan et al., 1984), transforming growth factors (TGFα and β) (Lucia et al., 1998, Peehl and Sellars, 1997) and hepatocyte growth factor (HGF). Most of these factors (except for TGFβs including BMPs and activins) are predominantly stimulators of proliferation, and also have tumour-promoting functions.

Members of the Fgf family play key roles in the in development of many branched organs and generally act to promote branching. Mammals have a multitude of FGFs and several FGF receptors, in general branching epithelial use either FGF7 or FGF10, and their receptor FGFR2-IIIb. The degree to which different organs depends on these FGFs varies. Lungs for example show an absolute requirement for FGF but in the kidney, the branching or urinary collecting duct is only reduced, not eliminated by the loss of renal FGF (as reviewed by Davies, 2002).

During development of the prostate, Fgf7 and Fgf10 are expressed in the mesenchyme, while their receptors are predominantly found on adjacent epithelial cells (Finch et al., 1995, Mason et al., 1994, Orr-Utreger et al., 1993). They both play significant roles as paracrine regulators of prostate development. Fgf7 and Fgf10 both regulate the growth of the ventral prostate (Sugimura et al., 1986, Thomson and Cunha, 1999) and seminal vesicle (Alardi et al 1994, Thomson and Cunha, 1999). However, deletion of Fgf7 by gene knock-out does not appear to affect the formation of the male reproductive tract (Guo et al, 1996), while the Fgf10 knock-out (Fgf10−/−) has absent or atrophic male secondary sex organs, including the prostate and seminal vesicle (Donjacour et al., 2003, Min et al., 1998, Sekine et al., 1999, Suzuki et al., 2000). In the prostate, as in the lung, and other branched organs, Fgf10 is restricted to the distal aspects of the organ where it is believed to act as a chemoattractant for elongating ducts and an inducer of branching through stimulation of cell proliferation (Bellusci et al., 1997, Thomson and Cunha, 1999). Fgf10 is the first mesenchymal specifically expressed gene known to be required for prostatic organogenesis.
HGF, and EGF/TGFα have also been shown to promote branching. Generally, HGF is produced by the mesenchyme and its receptor is expressed in the epithelium. Addition of exogenous HGF or EGF/TGFα was found to stimulate the production of branching tubules by cells derived from prostate, kidneys, mammary glands, salivary glands and pancreas (as reviewed by Davies, 2002).

On the other hand members of the TGFβ family, TGFβ, BMPs and activins are proposed to inhibit branching. TGFβ1 is expressed in the mesenchyme of the prostate, lungs, and mammary gland. In the prostate and the lungs, it is a potent inhibitor of epithelial proliferation and branching (Itoh et al., 1998, Zhao et al., 1996) and inhibits production of FGF10 (Tomlinson et al., 2004).

BMPs are involved in regulation of epithelial branching in several organs. In the prostate, BMP4 is expressed by the mesenchyme and is suggested to restrict prostatic ductal branching. Increased branching morphogenesis is observed in the anterior and posterior prostatic lobes of a heterozygous BMP4 knock-out mouse (Lamm et al., 2001). However, in the lungs, BMP4 is proposed to set up differences between differentiation characteristic of early and late generation branching (Bellusci et al., 1996).

An awareness of the features of different branching systems and their regulation is very useful for hypothesising on possible similarities with the poorly characterised prostate.

1.5.2 Neoplastic prostate development

Some molecules that regulate normal prostate development are also active during neoplastic prostate growth. Virtually all properties of neoplasms have a counterpart in normal prostate development. Accordingly proliferation, differentiation, invasion and apoptosis are events especially relevant to neoplasms as well as to the developing embryo. In the prostate these events are cardinal features of both normal development and carcinogenesis (as reviewed by Cunha et al., 2004). Therefore, it is
not surprising that molecules common to both processes have been reported, and of particular interest are factors expressed by cancer associated fibroblast cells (reviewed in Bhowmick et al., 2004). These include developmental growth factors already identified above with relation to prostatic disease, plus members of the developmental Wnt and MMP gene families (reviewed in Bhowmick et al, 2004b).

An important technique now for the analysis of prostate tumours is gene expression profiling. It is hoped that the biological heterogeneity of the disease might be reflected in expression profiles, yielding classifications for prognosis and treatment and new diagnostic markers and therapeutic targets. Another hope from using this technique is the identification of primary genetic changes in prostate carcinomas, by analysis of altered gene expression patterns. Specifically, it may be able to pinpoint transcriptional profiles for invasion, for androgen independence, or indicating the activation of a particular "cancer pathway". More than a dozen expression profiling studies have been published on clinical specimens and pre-clinical models, prior to and during the course of the Ph.D. project (Bera et al., 2002, Bubendorf et al., 1999, Xu et al., 2000, Dhansakeran et al., 2001, Luo et al., 2001a, b, Mousses et al., 2001, Stamey et al., 2001, Waghray et al., 2001b, Welsh et al., 2001, DePrimo et al., 2002, Ernst et al., 2002, LaTulippe et al., 2002, Rhodes et al., 2002, Shou et al., 2002, Sing et al., 2002, Varambally et al., 2002).

Studies using clinical samples have focused on the classification of prostate carcinoma progression and on identification of marker genes. Two single gene markers found reproducibly altered in prostate cancer samples; are hepsin (Luo et al., 2001b, Magee et al., 2001) and α-methyl-acyl-coA racemase (AMACR)(Jiang et al., 2001, Rubin et al., 2002) although at present it is uncertain if they represent good therapeutic targets.

An inherent problem associated with gene profiling of cancer is the tissue heterogeneity of the samples. A tumour is a complex tissue containing different cells, and as such it has a complex transcriptome. Therefore, preparation of gene expression profiles from whole tumour samples makes it impossible to assign the
genes identified to a specific cell type. Furthermore, without knowledge of the specific expression of a gene, it makes it difficult to pin-point changes in gene expression between normal prostate tissue and prostate cancer, that are not merely the result of alterations in the proportions of the different cell types, as the tissue becomes tumorigenic.

Finally, gene expression profiling has also been used to identify genes that are active during development, to investigate the regulatory mechanisms occurring during organ formation. Several studies have been performed, including the production of a gene expression profile of the embryonic mouse (Wagner et al., 2005) and more specifically, profiles of individual organs such as the embryonic mouse colon (Park et al., 2005), and testis (Shima et al., 2004), developing rat and human kidney (Dekel et al., 2002, Stuart et al., 2001), and human foetal liver (Yu et al., 2001). Of particular relevance to this thesis, are published transcriptome profiles, generated by SAGE, of the developing mouse prostate (E16.5, P1 and Adult) (Zhang et al. 2006), available at http://www.mouseatlas.org. In addition to contributing to the understanding of developmental processes, identifying genes involved in morphogenetic events taking place during development of individual tissues may give insights into genes that function during the pathogenesis of disease in these tissues. In cancer for instance, it is proposed that key developmental genes may be dysregulated. Also, relating patterns of gene expression between different organs during the development, will determine the presence of conserved developmental mechanisms in tissues and also between species.

1.6 Summary
Development of the prostate is regulated by mesenchymal-epithelial interactions, and is dependent on the presence of androgens, plus androgen receptor action in the mesenchyme. Numerous molecules are involved in mesenchymal-epithelial interactions, and have been shown to regulate prostate organogenesis. More specifically, the ventral mesenchymal pad, an area of condensed mesenchyme has been identified to express mesenchymal factors that are important in prostate growth. Few mesenchymal factors have been identified so far. One hypothesis for the
mechanism of regulation of prostate organogenesis has proposed that the production and secretion of some mesenchymal factors are androgen-regulated, and are termed ‘andromedins’. At present no growth factors, expressed in the mesenchyme have been shown unequivocally to be directly regulated by androgens, but it is possible that factors are constitutively expressed and their activity is indirectly regulated by androgens. More recently, the importance of the stroma in cancer has been established, and it has been proposed that developmental factors may be involved in mediating the effect of cancer stroma.

1.7 Project Aims
The aims of this project were to identify mesenchymal factors involved in normal and neoplastic prostate growth, and also to determine if these factors were androgen-regulated. An open-ended approach was used so that the findings would not be biased by previous assumptions.

A gene expression profile of the prostate precursor was documented, using LongSAGE analysis. The Libraries were prepared from tissues of the neonatal female UGS. The VMP (P0) was proposed to provide an ideal source of homogenous mesenchymal tissue that constitutively expresses factors capable of prostate organogenesis. The LongSAGE experiment was designed with the knowledge that some VMP-expressed factors may be androgen-regulated in the developing male UGS, and also that mesenchymal factors ‘activated’ by invading epithelial cells were likely to be absent from the Libraries generated.

Also, the importance the role of mesenchyme/stroma in cancer has become more established, and it was hoped to determine if regulatory developmental genes in the developing prostate are expressed during neoplastic growth in prostate cancer, and may represent potential molecular therapeutic targets.
2. Materials and Methods

2.1 Materials

2.1.1 Equipment

Agarose gel electrophoresis was performed using Bio-Rad equipment.

Automated sequencing gels were run using an ABI 310 Genetic Analyser by the Human Genetics Unit, MRC, Edinburgh, UK.

Confocal Microscopy was performed on a Zeiss LSM 510 Laser scanning microscope.

Images of immunohistochemistry were captured using a Provis microscope (Olympus Optical Co. London, UK) equipped with a DCS330 camera (Eastman Kodak Co., Rochester, NY).

Images of organ cultures were captured on a Leica MZ6 dissecting microscope with attached Leica ICA camera.

Quantitative RT-PCR analysis was performed on the LightCycler from Roche Applied Systems.

Minipreps for the SAGE libraries were performed on the Biorobot® 9600 from Qiagen.

Paraffin embedding was performed on a Leica TP1050, the tissues embedding centre was purchased from Reichert-Jung and tissue sections were cut using the Leica Microtome RM2135.

PCR was performed using a Peltier Thermal Cycler PTC 200 from MJ Research.
Chapter 2 Method and Materials

Phospho Screen and Phosphoimager Storm 860 were purchased from Molecular Dynamics.

RNA concentration and purity was measured using the Agilent 2100 Bioanalyser from Agilent Biotechnologies

Vertical gel electrophoresis apparatus for large format gels (15 x 17cm) used was from BioRad

Wallac 1450 Microbeta® Trilux, Liquid Scintillation and Luminescence Counter

2.1.2 Computer Software

SAGE2000 Version 4.5 was downloaded from http://www.invitrogen.com/sage and used to extract the LongSAGE tags from the sequence files and generate a tag abundance report.


Microsoft®Access was used to annotate the tags by comparison with the Rat LongSAGE map, and compare the tag abundances between the LongSAGE libraries.

Images were captured on a Macintosh G3/G4 and Apple computer and manipulated with Adobe Photoshop 7.0 Power PC.

Organ culture images were captured using the Scion Java Package 1.1 for Image J.

Phosphorimages were visualised using the Typhoon 9400 Variable Mode Imager from Amersham Biosciences, and quantified using Image Quant Version 5.2 from Amersham Biosciences.
Plasmid preparations on the Qiagen Biorobot® 9600 were performed using the Qiaprep® 96 Turbo protocol run on Qiasoft 3.0 software.

Sequence analysis was performed using GeneJockey II Sequence Processor (P.Taylor, BIOSOFT). Sequence searches were performed using BLAST from NCBI (http://www.ncbi.nlm.nih.gov/) and Ensembl (http://www.ensembl.org).

### 2.1.3 Enzymes

*EcoR*I (Cat. No. R6011), *Sph*1 (Cat. No. R6261), *BamH*I (Cat. No. R6021) and *Xba*I (Cat. No. 6181) were purchased from Promega

*Mme* I type II restriction endonuclease (Cat. No. B9003S) purchased from NEB Biolabs Inc. was used with 50 μM S-adenosylmethionine and the supplied buffer.

SUPERSCRIPT™ II RNase H Reverse Transcriptase was purchased from Invitrogen™ Life Technologies and used with the supplied buffers.

T4 Polynucleotide Kinase (Cat. No. M0201S) purchased from NEB Biolabs Inc. was supplied with 10x kinase buffer and used with 2mM ATP (Cat. No. 27-2056-01) purchased from Pharmacia Biotech Ltd.

T4 DNA Ligase (Cat. No. and Platinum *Tag* Polymerase (Cat. No. 10966-018) were purchased from Invitrogen™ Life Technologies and used with supplied buffers.

### 2.1.4 Nucleic Acids

100bp (Cat. No. G2101) and 1kb (Cat. No. G5711) DNA ladder were purchased from Promega Inc.

dNTP's (Cat. No. U1330) were purchased from Promega Inc.
Radioisotope redivue deoxycytidine 5'-[^32P] triphosphate triethylammonium salt (Cat. No. AA0005) was purchased from Amersham Biosciences Inc.

Synthetic oligonucleotides were synthesised either by Sigma-Genosys Ltd or MWG-Biotech AG.

2.1.5 Chemicals
All chemicals for standard solutions were purchased from BDH, Sigma-Aldrich or Invitrogen™ Life Technologies and were of analytical grade.

2.1.6 Bacterial Strain Genotypes
DH5α; F- 80dlac ΔZM15 Δ(lacZYA-argF) U169 recA1 endA1
hsdR17(rl-, mk+) phoA supE44 - thi-1 gyrA96 relA1
Purchased from Invitrogen™ Life Technologies.

2.1.7 Bacterial Culture Media
Luria Broth (LB): 10g/L baco-tryptone, 5g/L bacto-yeast extract, 5g/L NaCl, adjusted to pH7.0 with 5M NaOH

SOC Media: Cat. No. 46-0821 was purchased from Invitrogen™ Life Technologies.

SOB Media: 20% baco-tryptone, 0.5% baco-yeast extract, 0.05% NaCl, 2.5%mM KCl, 10mM MgCl₂

2.1.8 Northern Blotting and Hybridisation Equipment
Hybond™-N+ paper (Cat. No. RPN-303B), Hybond blotting paper (Cat. No. RPN6101M), and MicroSpin™ S—400 HR Columns (Cat. No. 27-5140-01) were purchased from Amersham Biosciences UK Ltd.
Hybaid boroscilicate tubes (Cat. No, HBOVBS) were purchased from Fisher Scientific

ULTRAhyb™ Ultrasensitive Hybridisation Buffer (Cat. No. 8670) was purchased from Ambion®

2.1.9 Immunohistochemical Reagents

ABComplex/HRP (Cat.No. K0355) and DAB+Substrate Chromagen System (Cat. No. K3468) purchased from DakoCytomation

Table 2.1: Antibodies used for Immunohistochemistry

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Source</th>
<th>Species</th>
<th>Epitope</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti Delta-like1 (A-17)</td>
<td>Santa Cruz</td>
<td>Goat</td>
<td>Polyclonal</td>
<td>1:25 IHC, 1:10 IFC</td>
</tr>
<tr>
<td>Anti-Dlk 1</td>
<td>Gift Dr. S. Morley</td>
<td>Rabbit</td>
<td>Polyclonal</td>
<td>1:300 IHC</td>
</tr>
<tr>
<td>Anti-Pleiotrophin (C-19)</td>
<td>Santa Cruz</td>
<td>Goat</td>
<td>Polyclonal</td>
<td>1:50 IHC</td>
</tr>
<tr>
<td>Anti-Pleiotrophin</td>
<td>RnD Systems</td>
<td>Goat</td>
<td>Polyclonal</td>
<td>1:400 IHC, 1:50 IFC</td>
</tr>
<tr>
<td>Anti-Notch 2 (M20)</td>
<td>Santa Cruz</td>
<td>Goat</td>
<td>Polyclonal</td>
<td>1:200 IHC</td>
</tr>
<tr>
<td>Anti –alpha actin smooth muscle</td>
<td>Sigma -A2547</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>1:5,000 IFC</td>
</tr>
<tr>
<td>Anti-beta catenin (E5)</td>
<td>Santa Cruz</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>1:100 IFC</td>
</tr>
<tr>
<td>Anti-P63</td>
<td>Santa Cruz</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>1:40 IFC</td>
</tr>
<tr>
<td>Anti-rabbit biotinylated</td>
<td>Vector Laboratories</td>
<td>Swine</td>
<td></td>
<td>1:500</td>
</tr>
</tbody>
</table>
Table 2.1 (cont):  

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Source</th>
<th>Species</th>
<th>Epitope</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-goat biotinylated</td>
<td>Vector Laboratories</td>
<td>Rabbit</td>
<td></td>
<td>1:500</td>
</tr>
<tr>
<td>Anti-mouse Cy3</td>
<td>Sigma–Aldrich Inc</td>
<td>Goat</td>
<td></td>
<td>1:200</td>
</tr>
<tr>
<td>Anti-Mouse Cy2</td>
<td>Sigma–Aldrich Inc</td>
<td>Goat</td>
<td></td>
<td>1:200</td>
</tr>
<tr>
<td>Streptavidin Alexafluor 488</td>
<td>Molecular Probes</td>
<td></td>
<td></td>
<td>1:200</td>
</tr>
<tr>
<td>Streptavidin Alexafluor 546</td>
<td>Molecular Probes</td>
<td></td>
<td></td>
<td>1:200</td>
</tr>
<tr>
<td>Topro 3</td>
<td>Molecular Probes</td>
<td></td>
<td></td>
<td>1:2,000</td>
</tr>
</tbody>
</table>

Santa Cruz also supplied blocking peptides to their antibodies delta-like 1 (SC-8625P), pleiotrophin (SC-1394P), and notch 2 (SC-7423P).

2.1.10 Organ Culture Media, Equipment

All incubation dishes and tubes were purchased from Falcon, NUNC™ or Corning Inc. (COSTAR).

Dulbecco’s Modern Eagle Media (DMEM) with glutamax I (Cat. No. 31966-021), Penicillin-Streptomycin (Cat no. 15240-062), Trypsin-EDTA (Cat. No. 25300-062), Bovine Transferrin (Cat. No. 11107) were purchased from Invitrogen™, Life Technologies. Bovine Insulin (Cat. No. 1-0516) was purchased from Sigma.

Manipulations were carried out in statutory Class II tissue culture hoods.

Millicell CM Culture Plate Inserts (Cat. No. PICM03050) were purchased from Millipore.
### Table 2.2: Recombinant proteins used in organ culture experiments

<table>
<thead>
<tr>
<th>Recombinant Protein</th>
<th>Source</th>
<th>Cat. No.</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Delta-like 1 homolog (Drosophila)</td>
<td>Wheat germ cell free in vitro expression system</td>
<td>H00008788-P01</td>
<td>Abnova Corporation</td>
</tr>
<tr>
<td>Human Pleiotrophin</td>
<td>Sf21 insect cell line</td>
<td>252-PL</td>
<td>R&amp;D Systems</td>
</tr>
</tbody>
</table>

### 2.1.11 Kits

### Table 2.3

<table>
<thead>
<tr>
<th>Kit</th>
<th>Cat. No.</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-SAGE kit</td>
<td>T5001-01</td>
<td>Invitrogen™ Life Technologies</td>
</tr>
<tr>
<td>Rneasy Mini Kit</td>
<td>74106</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Turbo DNA-free™ kit</td>
<td>1907</td>
<td>Ambion</td>
</tr>
<tr>
<td>Rneasy Midi Kit</td>
<td>75144</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Platinum Taq Polymerase PCR Kit</td>
<td>10966-034</td>
<td>Invitrogen™ Life Technologies</td>
</tr>
<tr>
<td>LightCycler FastStart DNA Master SYBR Green 1</td>
<td>3 003 230</td>
<td>Roche Applied Science</td>
</tr>
<tr>
<td>Cloned AMV First Strand Synthesis Kit</td>
<td>12328-032</td>
<td>Invitrogen™ Life Technologies</td>
</tr>
<tr>
<td>Qiaquick® PCR Purification Kit</td>
<td>28106</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Qiaquick® Spin Miniprep Kit</td>
<td>27106</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Qiaquick® Gel Extraction Kit</td>
<td>28704</td>
<td>Qiagen</td>
</tr>
<tr>
<td>pGEM T easy Vector System 1</td>
<td>A1360</td>
<td>Promega Inc.</td>
</tr>
</tbody>
</table>
### Table 2.3 (cont):

<table>
<thead>
<tr>
<th>Kit</th>
<th>Cat. No.</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RadPrime DNA Labelling System</td>
<td>18428-011</td>
<td>Invitrogen™ Life Technologies</td>
</tr>
<tr>
<td>Wizard™ SV9600 Plasmid Purification System</td>
<td>A2258</td>
<td>Promega Inc.</td>
</tr>
</tbody>
</table>

#### 2.1.12 Standard Solutions (A to Z)

- **Blue/Orange Loading Dye (Cat. No. G190A) purchased from Promega Inc.**

- **Bouins Fixative**
  
  75ml picric acid, 25ml 40% (v/v) Formaldehyde, 5ml glacial acetic acid.

- **Citrate buffer (0.1M)**
  
  0.1M Citric Acid (monohydrate), pH 5.5.

- **Formaldehyde-gel running buffer (5x)**
  
  0.1M MOPS (pH 7.0), 40mM sodium acetate, 5mM EDTA (pH 8.0)

- **MOPS (10x)**
  
  200μM 3-(N-morpholino)propanesulfonic acid (MOPS), 50μM Sodium acetate, 10μM EDTA.

- **PBS**
  
  Phosphate buffer saline tablets (Cat. No. P4417) supplied by Sigma®

  RNA Sample Loading Buffer: Cat. No. R-1386 purchased from Sigma® without ethidium bromide
SSC (20x) (Cat. No. 51233) purchased from Accugene<sup>TM</sup>, Biowhittaker Molecular Applications

TBS 0.15M NaCl in Tris pH 7.4

TBST 0.15M NaCl, 0.1% Polyoxyethylenesorbitan Monolaurate (Tween 20)

2.2 Methods

2.2.1 Total RNA Extraction

Total RNA was extracted from mammalian tissue using RNeasy™ Mini and Midi kits (Qiagen) as per the manufacturer’s instructions. RNA was eluted ddH₂O in a final volume of 30μl (Mini kit) or 120μl (Midi kit), and the RNA concentration and purity was measured on the Agilent 2100 Bioanalyzer.

2.2.2 DNase Treatment of Total RNA

Total RNA was treated with the Turbo DNA-free kit (Ambion) as per the manufacturer’s instructions to eliminate genomic DNA contamination. 4U TurboDNase and 1x DNase buffer was used in a final reaction volume of 25μl. The reaction was incubated at 37°C for 30 minutes, and then mixed with 0.2V of DNase Inactivation Reaction solution. The mixture was centrifuged at 10,000rpm for 1.5 minutes and the DNase-treated RNA was removed to new eppendorf tube.

2.2.3 cDNA Synthesis

500ng of DNase treated RNA was reverse transcribed in the presence of absence of reverse transcriptase (SUPERSCRIPT™ II RNase H; Reverse Transcriptase; Invitrogen™, Life Technologies) as per the manufacturer’s instructions. The completed reactions (20 μl) were diluted with 80μl of ddH₂O. The quality of cDNA was always checked by PCR using TATA-Binding Protein (TBP) primers, and also with GAPDH primers to ensure the complete elimination of genomic DNA by DNase treatment.
2.2.4 Polymerase Chain Reaction

In a 50μl reaction volume 5μl of cDNA, made from 500ng DNase treated RNA, was mixed with 1x PCR buffer, 1.25 or 2.5 mM MgCl2 and, 0.2mM dNTP's, 0.5 μM of each primer and 1.5U Platinum® Taq DNA Polymerase. The primers used are listed in Table 2.4 and 2.5. Typically 35 cycles of PCR were performed using the following conditions: samples were preheated to 95°C for 2 minutes, then a standard 35 cycles of 95°C for 45 seconds, Ta for 30 seconds and 72°C for 45 seconds. A final extension of at 72°C for 15 minutes was performed and samples were incubated at 15 °C. Ta refers to the annealing temperature in °C used for the oligonucleotides based upon the formula; Ta= 4x (No. G/C bases in the oligonucleotide) + 2x (No. A/T bases in the oligonucleotide). PCR products were visualised by agarose gel electrophoresis and ethidium bromide staining.

2.2.5 qRT-PCR – LightCycler Analysis

The LightCycler FastStart DNA master SYBR Green 1 Kit (Roche Applied Systems) was used for qRT-PCR. 3μl of cDNA (synthesised from 500ng DNase Treated RNA) was added to a reaction mixture of 1x Lightcycler FastStart DNA Master SYBR Green 1, 2mM or 3mM MgCl2 and 0.625 μM of each primer in a final volume of 10μl. The PCR products were measured against a standard curve of cDNA diluted to 10%, 2%, 0.4%, 0.08%. The PCR was performed on a LightCycler machine (Roche Applied Systems). Typically 45 cycles were performed using the following conditions: samples were preheated to 95°C for 10 minutes, then a standard 45 cycles of 95°C for 15 seconds, Ta for 5 seconds, 72°C for 22 seconds and Tm for 1 second. Ta refers to the annealing temperature as described under Polymerase Chain Reaction (Section 2.2.4). Tm refers to the measuring temperature at which the amount of PCR product is recorded. The PCR products were quantified using the LightCycler software and normalised against GAPDH control primers. The PCR products were checked by agarose gel electrophoresis and ethidium bromide staining.
Table 2.4: Rat Primers for RT-PCR, LightCycler Analysis and DNA probes.

| Probe/construct | Primers (5’-3’) | Reverse | PCR Ta | MgCl₂ Ta | LightCycler Tm
|----------------|-----------------|---------|--------|----------|----------------
| Basigin        | CAGAGCTGATCATCACGCGAC | CCTTGTCAATTCAGGTGAGAC | 60    | 1.25     | 60 3 86        |
| BMP3           | GGATCTTTGTTGCCATGGCAG | GGACCACATCTTGTTCTGC | 60    | 1.25     | 60 2 84        |
| BMP7           | CTGAAAGCTATGCTGCTTAC | GCTTATACACACAGGTGTCG | 60    | 1.25     | 60 3 86        |
| Decorin        | CAGTCAACTGCATCTGGATG | CGAAGACACATCTGAAAGTG | 60    | 1.25     | 60 3 80        |
| Delta-like 1   | AGGACTGTGACCACAAAGGT | TGGACAGGTAGCTCAGTCAC | 60    | 1.25     | 60 3 88        |
| FGF10          | CAGTGGAAATCGGGATTGGT | ATGACGGCAATGGACAGAGCT | 60    | 1.25     | 60 3 82        |
| FGFR2          | GTCATAGTGCATAAGCACATG | CCAAGTCTTCCACCAACTGC | 60    | 1.25     | 60 3 84        |
| Fibronectin    | GACCTACACAGTAGGAGAAC | GCTGAAAGCTGAGACTTGCAC | 60    | 1.25     | 60 3 86        |
| GAPDH          | TTAGCACCCCCCTGGGCAAGG | CTTACTCCTGGAGCCATG | 60    | 2.5      |                |
| GDNFR alpha 1  | CAGACACCTTTGCTGTCATC | GAGCTTTCCTGCCAACAGTAC | 60    | 1.25     | 60 2 84        |
| Hepsin         | GCAAGGTCTGTAAGCTGAC | CCAAATCTGGACTCTGATGC | 57    | 2.5      | 60 3 85        |
| Hgf            | GCTTGGAAATCCATGATGCC | CTTGTATCCATGGATGCTTC | 60    | 1.25     | 60 2 81        |
| Jagged 2       | GAGAGGAAACAGCTCAGCTG | GGTAGAGTACGTCTTGGTG | 60    | 2.5      | 60 3 86        |
| Lgl1           | CGTCATTGTGACAGAGGTGG | GCTGTCCACTTGTGTCATTC | 60    | 2.5      | 60 3 86        |
| Midkine        | CCTGAAAGAGGCTCGGTAC | CCAGAAAAGCTCTGGGCCTC | 60    | 1.25     | 60 2 82        |
| MMP14          | GGACAAAGATGGATGCAAGTC | GAGGAGGCAAAAGTCTTGGT | 60    | 1.25     | 60 3 88        |

Ta = Annealing Temperature (°C)  Tm = Measuring Temperature (°C)  MgCl₂ = magnesium chloride (mM)
<table>
<thead>
<tr>
<th>Probe/construct</th>
<th>Forward</th>
<th>Reverse</th>
<th>Ta</th>
<th>MgCl₂ Ta</th>
<th>MgCl₂ Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMP2</td>
<td>CCTGATGTCAGCAAGTATGGC</td>
<td>CTGAAGTCACCAGGTTGAAGG</td>
<td>60</td>
<td>1.25</td>
<td>65</td>
</tr>
<tr>
<td>Nel-like 2</td>
<td>CACAGTTGACCTGCTGCTGC</td>
<td>GACAAGGTACGGTGACACAG</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
<tr>
<td>Notch 2</td>
<td>CTGAGGACTTCTGCTGCAAC</td>
<td>CAGCTGATCAGCTCTGCCAC</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
<tr>
<td>Notch 1</td>
<td>CTATCAGGGTGGCTGCTGGTC</td>
<td>CCAGCAACACTTGGCAGTCA</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
<tr>
<td>Paralemmin</td>
<td>GTGGAGAAGGAGAACAGGTGAC</td>
<td>CAGCTGATCCTGCTGCTTCT</td>
<td>60</td>
<td>1.25</td>
<td>65</td>
</tr>
<tr>
<td>Pdgf</td>
<td>GCTGTGACTCAACAGCGGCTG</td>
<td>GCCACTTGGAATAAGGACCC</td>
<td>60</td>
<td>1.25</td>
<td>65</td>
</tr>
<tr>
<td>Pleiotrophin</td>
<td>GCAGCTGACTGAGGAGCTCTG</td>
<td>CCAGCTGACATGAGCAGACATG</td>
<td>60</td>
<td>1.25</td>
<td>65</td>
</tr>
<tr>
<td>PMP 22</td>
<td>GTCTGTCATTCCCAGCCTGC</td>
<td>GCAGCTGACTGAGCAGACATG</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
<tr>
<td>Syntenin</td>
<td>CACTGATCAACACATCTGTC</td>
<td>CTGATAGTAGAGGAGTTCC</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
<tr>
<td>TBP</td>
<td>GCCTAGTTGAGGTAGGTCTG</td>
<td>CCTAGAGCATCCTCTATCTCT</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
<tr>
<td>TGF beta 2</td>
<td>CCTCCACATATGCCAGGCGG</td>
<td>CAGTGGTTCCAGATCTTGG</td>
<td>60</td>
<td>2.5</td>
<td>60</td>
</tr>
<tr>
<td>VAP1</td>
<td>CCTGGAGGTGCTGTAGAGCA</td>
<td>CTGAGCACAACCCTTGAGGAC</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
</tr>
</tbody>
</table>

Ta = Annealing Temperature (°C) Tm = Measuring Temperature (°C) MgCl₂ = magnesium chloride (mM)
Table 2.5 Human Primers for RT-PCR, LightCycler Analysis and DNA probes.

<table>
<thead>
<tr>
<th>Probe/construct</th>
<th>Primers (5’ to 3’ Forward)</th>
<th>Reverse</th>
<th>PCR Ta</th>
<th>MgCl₂ Ta</th>
<th>Light Cycler Ta</th>
<th>MgCl₂ Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta-like 1</td>
<td>CTGCCCTTACGGCACTCTGT</td>
<td>GCAGGTCTTTGTCGATGACGC</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>GAPDH</td>
<td>TAGCACCCTTGGGCAAGG</td>
<td>CTTACTCTTGGAGCCATG</td>
<td>60</td>
<td>2.5</td>
<td>60</td>
<td>3</td>
</tr>
<tr>
<td>MMP14</td>
<td>GGACAGATTGTGAGTGCACTC</td>
<td>GAGGAGGCAAAGTCTTTGTG</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
<td>2 87</td>
</tr>
<tr>
<td>MMP2</td>
<td>CCAAGCTCATCGAGATGC</td>
<td>CTGAAGTCCACCAAGGTAAGG</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
<td>3 87</td>
</tr>
<tr>
<td>Nel-like 2</td>
<td>GACATCAACAGACTGTGC</td>
<td>GGTGATGTGAGACACAGTAG</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
<td>3 89</td>
</tr>
<tr>
<td>Notch 2</td>
<td>CAGAGGACTCTCTTGTGTAAACG</td>
<td>GAGAGAAGCAAGTAAACTAAC</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
<td>3 89</td>
</tr>
<tr>
<td>Pleiotrophin</td>
<td>CAGCTGGATAGCTCACTG</td>
<td>GACAGTCTCTTGGCATTCG</td>
<td>60</td>
<td>1.25</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>TBP</td>
<td>AGTTAGAAGCGCTTGTGCTGGAG</td>
<td>GGGAGGCAAGGGTACATGAG</td>
<td>60</td>
<td>2.5</td>
<td>60</td>
<td>3.5 78</td>
</tr>
</tbody>
</table>

Ta = Annealing Temperature (°C)  Tm = Measuring Temperature (°C)  MgCl₂ = magnesium chloride (mM)
2.2.6 Cloning of PCR Products

2.2.6.1 Preparation of Insert DNA
100 μl (2 x 50μl reactions) of PCR product was prepared according to the PCR conditions outlined. 10μl of PCR product was checked by agarose gel electrophoresis and the remaining product was purified using the Qiagen™ PCR Purification Kit, eluting in a final volume of 30 μl. The concentration of the final eluate was estimated by running a 5μl aliquot on a 1% agarose gel against a known amount of DNA ladder.

2.2.6.2 Ligation
The pGEMTeasy kit from Promega was used for ligation of insert DNA into the pGEMTeasy Vector supplied as per the manufacturer’s instructions, using the appropriate insert: vector ratios, 1XT4 DNA Ligase buffer and 0.05 U/μl in a 10μl volume. Blunt-end ligation was performed at room temperature for 1 hour or overnight at 16°C. 2μl of ligation mix was transformed into DH5α E.Coli as per the manufacturer’s instructions.

2.2.7 Plasmid DNA Preparation

2.2.7.1 Plasmid Mini-Preparation
A 3ml. culture of E.Coli was grown in LB, with 1X ampicillin at 37°C in a shaking incubator overnight. The culture was pelleted in a 1.5ml eppendorf tube and plasmid DNA was prepared as described in the Qiaprep™ Mini-prep protocol.

2.2.7.2 Restriction Endonuclease Digestion
Plasmid DNA was digested using the recommended buffer and approximately 2U enzyme per 1μg DNA for 2 hours. Restriction endonuclease digestion was performed in a water bath at 37°C.
2.2.8 Agarose Gel Electrophoresis

2.2.8.1 DNA
Plasmid DNA was run on agarose gels of 1% at 100V, determined to give the best possible resolution of fragments. DNA was loaded into the gels using 1X Blue/Orange Loading Dye (Cat. No. G190A) purchased from Promega, and run using 1xTBE. DNA was visualised by ethidium bromide staining and UV-light.

2.2.8.2 RNA- Formaldehyde gel
RNA was run on a horizontal 0.1% (w/v) Agarose gel containing, 1x MOPS and 6% formaldehyde, using 1xFormaldehyde-gel running buffer. RNA was loaded in a ratio of (2:1) with RNA Sample Loading Buffer without ethidium bromide (Sigma), heated to 65°C for 15 minutes and cooled on ice prior to loading. Gels were run at 4-5V/cm, typically 75V for a 16 cm gel. 5μg of total RNA was loaded in each gel lane. A duplicate was performed for each gel with independent RNA samples, except where stated.

2.2.9 Northern Blotting and Hybridisation

2.2.9.1 Northern Blotting
RNA Gels were blotted using Hybond™-N+ paper (Amersham Biosciences) and cut to match the size of gel. 20x SSC and Hybond™ blotting paper (Amersham Biosciences) were used for the capillary transport of RNA from the gel to the Hybond™ N+ paper, carried out overnight. The transfer of RNA was checked by separating and staining a lane of Hybond™ N+ paper containing transferred RNA, with a solution of 0.5M sodium acetate (pH 5.2) and 0.04% methylene blue for 10 minutes at room temperature.

2.2.9.2 In vitro transcription of 32P labelled DNA probes
A DNA template of the probe was prepared using the relevant primers in Table, and inserted into the pGEMTeasy vector as per the manufacturer’s instructions (see
Cloning of PCR Products. The plasmid DNA template was cut out of the pGEMTeasy Vector by digestion with EcoRI supplied by Promega to create a 5' overhang. Typically, a 30µl plasmid mini-prep was digested in a reaction volume of 100µl. Complete digestion was confirmed by gel electrophoresis. At least 90µl of digestion mixture was run on a 1% agarose gel and the digested DNA insert was cut out and gel-purified using the Qiagen™ Gel Extraction Kit. The insert was eluted in a final volume of 30µl, and 5µl of product was checked by gel electrophoresis.

In a reaction volume of 50µl ~25ng of insert DNA, typically 1-3µl, was heated to 95°C for 5 minutes and cooled on ice prior to mixing with a final concentration of 1x buffer, 10µM of each dATP, dGTP, and dTTP, 40U Klenow enzyme from the RadPrime DNA Labelling System Kit (Invitrogen™ Life Technologies) and 50pCi 32P dCTP (Amersham Biosciences). The reaction mixture was incubated for 30 minutes at 37°C. The labelled DNA probe was recovered by column purification with MicroSpin™ S—400 HR Columns (Amersham Biosciences). The probe was measured for 32P incorporation in a Scintillation counter (Wallac 1450).

2.2.9.3 Hybridisation of Labelled DNA Probes to RNA Blot
The Northern blot was placed in a Hybaid borosilicate tube (Fisher Scientific) with 10mls of ULTRAhyb™ Ultrasensitive Hybridisation Buffer (Ambion®), and incubated for at least 1 hour at 42°C in a rotating oven. The 32P DNA Labelled Probe was heated to 95°C for 5 minutes and cooled on ice prior to adding 10⁶CPM of probe per ml to the hybridization buffer. The probe was hybridized to the blot overnight at 42°C. The non-hybridized probe was then discarded and the blot was washed with solutions 2x SSC, 0.1%SDS followed by 0.1x SSC, 0.1%SDS at 48°C. Hybridised probes were visualised using a phosphor-screen (Molecular Dynamics), and quantified using a Phosphoimager (Amersham) and normalised to a GAPDH control probe.
2.2.10 Immunocytochemistry

2.2.10.1 Paraffin Embedding of Tissue

Microdissected tissue or cultured VP’s were fixed in Bouins fixative for 2 hours on ice. Tissue was stored in 70% ethanol until processed. Dehydration and further processing was performed on a Leica TP1050. Paraffin embedding was performed on the Tissue Embedding Center and 5μm sections were cut using a microtome (Leica RM2135).

2.2.10.2 Tissue Immunohistochemistry

Sections were dewaxed in xylene for 10 minutes and rehydrated through an ethanol series. Sections were pressure cooked for 5 minutes in citrate buffer (0.01M) and cooled for 20 minutes. Sections were washed two times for 5 minutes in TBS solution, and blocked by washing in a 3% hydrogen peroxide, methanol solution for 30 minutes at room temperature. The sections were blocked in 20% normal serum (from host species the secondary antibody was raised in) diluted in TBST with 5% BSA, for 45 minutes. The sections were then incubated overnight at 4°C with the primary antibody (see table) diluted in the normal blocking serum. Negative control sections were incubated with primary antibody that had been preabsorbed with its relevant peptide, or with normal blocking serum.

Slides were given two times 5 minute washes of TBST, followed by incubation with the required IgG-biotinylated secondary antibody at 1:500 diluted in 20% normal blocking serum. Sections were then washed two times 5 minutes in TBST and incubated with ABCComplex/HRP (DakoCytomation). The slides were washed again in TBST and the antigen was finally localised using the Liquid DAB+Substrate Chromagen System (DakoCytomation). Sections were counterstained with haematoxylin. Images were visualised and captured using the Provis (Olympus Optical Co.)
2.2.10.3 Fluorescent Tissue Immunocytochemistry
Sections are prepared similarly to the protocol outlined for Tissue Immunocytochemistry, except for the omission of the blocking step (3% hydrogen peroxide, methanol solution) and the washes are made with PBS. Also, the secondary antibody is either the required biotinylated IgG as described, or the required direct conjugate and streptavidin fluorophore complex (diluted 1:200 in 20% normal blocking serum). The nuclei were counterstained with Topro 3. Sections were visualised and imaged using confocal microscopy (Zeiss LSM 510).

2.2.11 In Vitro Organ Culture
Male VP lobes (VP) were microdissected from P0 Wistar rats and grown in serum-free organ culture (DMEM/NUT.MIX F-12 supplemented with transferrin (20μg/ml), insulin (20μg/ml) and 1 x penicillin/streptomycin) (Foster and Cunha, 1999). VP’s were positioned in a drop of media on Millicell® CM filters that were floated on 1ml of culture media in 4-well NUNCLONM™ plates. Organ rudiments were cultured at 37°C in 5% CO₂ for up to 6 days. The VP’s were imaged under lightfield illumination using the Image J software prior to fixation. To study the effects of recombinant proteins on the growth of the VP’s, organ rudiments were cultured with varying concentrations of each recombinant protein in the presence or absence of testosterone (10⁻8M).

2.3. LongSAGE
The LongSAGE technique (Saha et al., 2002, Velculescu et al., 1995) is a genome profiling method based on the acquisition of a 17bp sequence tag from the 3’ end of a transcript, which uniquely identifies that transcript. The method is similar to the original SAGE approach, but uses a different type IIS restriction endonuclease (MmeI) and incorporates other modifications to produce longer transcript tags. The resulting 21bp tag consists of a constant 4bp sequence representing the restriction site at which the transcript was cleaved, followed by an adjacent sequence in each transcript.
2.3.1. Materials
The LongSAGE™ libraries were made using the I-SAGE™ kit purchased from Invitrogen™ Life Technologies and an adapted form of the I-SAGE™ protocol. An I-LongSAGE™ kit is now supplied by Invitrogen™ to make LongSAGE libraries, but was not available at the time of making these libraries.

2.3.1.1 Tissue Samples and RNA Preparation
Two LongSAGE libraries were made with starting tissue microdissected from the urogenital tract of P0 female Wistar rats. One library was made from the whole region comprising of the prostatic precursor, and a second library was made from only the mesenchyme involved in prostatic induction and growth (the ventral mesenchymal pad, VMP). The whole prostate rudiment was referred to as the VSU as it is composed of VMP, Smooth muscle and Urethra (urethral epithelium). The VMP library was made from microdissected VMP mesenchyme only (Fig 2.1). RNA was prepared according to the RNA extraction method using the RNeasy Mini Kit from Qiagen. The RNA concentration and purity was measured on the Agilent 2100 Bioanalyzer.

![Diagram of the female rat (P0) UGT showing the relative position of the bladder (BL), and the two regions microdissected, and used to prepare the LongSAGE libraries; the VSU and the condensed pad of mesenchyme (VMP).](image)

**Fig: 2.1 The Prostate Rudiment**
A schematic of the female rat (P0) UGT showing the relative position of the bladder (BL), and the two regions microdissected, and used to prepare the LongSAGE libraries; the VSU and the condensed pad of mesenchyme (VMP).
2.3.2 Method

2.3.2.1 Library Preparation

A flow diagram of the LongSAGE protocol is included as Figure 2.2. The following modifications were made to the I-SAGE™ protocol (http://www.invitrogen.com), based on recommendations from an EMBO SAGE practical course (Jan 2003), and a LongSAGE protocol released by Genzyme Molecular Oncology in January 2003:

1) Primers for Rat EF-1α were used in the PCR checks of cDNA. The EF-1α primers supplied with the kit were recommended only for human samples.

   EF-1S       CATGAGGCAGACAGTTGCT
   EF-1AS      AGCCTTCTGAGCTTTCTGG

2) Linkers containing the Mme1 recognition site were ligated to 3’ cDNA ends after NlaIII digestion. To prepare the Linkers for use, 3.15 μg of Linker1B and Linker2B were kinased using 100 U T4 Polynucleotide Kinase, 1X T4 DNA Kinase buffer, 1mM ATP in a 20 μl reaction volume, at 37°C for 30 mins. The reaction was heat inactivated at 65°C for 10 minutes 1x Linker 1A and 1x Linker B was mixed with 2x Kinased Linker 1B and 2B, respectively, and hybridized at 95°C for 2 minutes, followed by 20 minutes at room temperature. The hybridised linkers were checked by agarose gel electrophoresis and diluted to a concentration of 20ng/μl for use. Linker 1A/B was ligated to half the cDNA population and Linker 2A/2B was ligated to the remaining half of the cDNA as per the I-SAGE kit instructions.

### Table 2.6

<table>
<thead>
<tr>
<th></th>
<th>VSU Library</th>
<th>VMP Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tissue collected</td>
<td>~ 100 VSUs</td>
<td>~90 VMPs</td>
</tr>
<tr>
<td>Starting Tissue Weight</td>
<td>~180 mg</td>
<td>~65 mg</td>
</tr>
<tr>
<td>Total RNA Used</td>
<td>50 μg</td>
<td>21 μg</td>
</tr>
</tbody>
</table>
Mix sample RNA with oligo (dT) magnetic beads

cDNA synthesis

Digest cDNA with a 4-base cutter "anchoring enzyme" (NlaIII).

Divide pool in half & ligate to different linkers (1 or 2), both of which have a restriction site for the "tagging enzyme" (Mme1).

Digest with a "type 2" enzyme (Mme1) which recognizes the linker sequences and cuts 17bp downstream in cDNA, leaving a 2bp overhang.

Fig 2.2: The LongSAGE method (Part 1)
Synthesis of double stranded cDNA, restriction enzyme digestion of cDNA and isolation of 3'-cDNA fragments, adaptor modification and excision of tags (redrawn from LongSAGE protocol supplied by Invitrogen™)
Fig 2.2 (cont): LongSAGE method (Part 2)
Ditag formation followed by amplification of ditags and adaptor removal. Formation of concatemers, cloning into pZero1 vector and Sequencing.
Chapter 2  Method and Materials

Linker 1A:
5'TTTGGATTTGCTGGTGCAGTACA ACTAGCTTTAATATCCGACATG3'

Linker 1B:
5'TCGGATATTAAGCTAGTTGTACTGCACCAGC CAAATCC
Amino modified C7-3'

Linker 2A:
5'TTTCTGCTGATCAGCTTTAAGCTTTAATATCCGACATG3'

Linker 2B:
5' TCGGACGTACATCGTTAGAAGCTTGATACGTCCGACATG3' Amino Modified C7-3'

3) The Linker molecules were released from the cDNA (attached to the magnetic beads) using 20U MmeI type IIS restriction endonuclease, 1X MmeI buffer, and 50mM S-adenosylmethionine for each Linker-tag in a 100 µl reaction volume. Digestion was performed at 37°C for 2.5 hours.

4) The LongSAGE tags created by digestion with the enzyme MmeI leaves a 2bp 3' overhang which was not polished with a Klenow reaction. The Linker 1A/B-Tag and Linker 2A/B -Tag molecules were ligated together overnight at 16°C as per the I-SAGE kit protocol, to create a ~130bp ditag.

5) The ditag molecules were amplified with LongSAGE primers diluted to a stock concentration of 175ng/µl:

LS-1P 5'GTGCTCGTGATTTGTGCTTGCAGTACA-3
LS-2P 5' GAGCTCGTGCTTGATTCAGCTTT3

7) Prior to ligation into pZero®-1 Vector, the concatemer fragments were digested with 20U of restriction endonuclease SphI and 1x required buffer purchased from
Promega in a final reaction volume of 20μl. The reaction was incubated at 37°C for 30 minutes, giving the concatemer fragments a 5’ overhang to improve their cloning efficiency. The digest mixture was extracted with phenol/chloroform and ethanol precipitated, as outlined in the 1-SAGE protocol. Ligation of the SphI digested-concatemer into the pZero®-1 vector was performed overnight at 16°C to improve the efficiency of cloning longer concatemer fragments.

8) The plasmid minipreps for the LongSAGE libraries were performed by Research Support Technician Raymond Chan on a Biorobot® 9600 from Qiagen using the Wizard™ SV9600 Plasmid Purification System (Promega). The plasmid minipreps were screened for inserts by double enzyme restriction digestion using recommended enzymes Bam HI and Xba I (Promega) with the supplied buffer, as per the manufacturer’s instructions. The sequencing was performed by the Human Genetics Unit (HGU) sequencing facility, Edinburgh.

2.3.3 Library Analysis

2.3.3.1 Tag Quantification

SAGE2000 Version 4.5 software, (http://www.invitrogen.com/sage) was used to prepare a report of the tag sequences, their abundance and associated transcripts from the raw Sequence files.

Within the sequence files each tag was ligated to another tag as a “ditag”, one tag with 5’ to 3’ orientation and the other in reverse. The sequence file contained a number of the ditags ligated as a concatemer, and present as an insert in a vector. The software was able to recognise the ditag sequences within the concatemer by the Nla III restriction enzyme sites (CATG) at each end. The software then separated the ditags as two individual tags and the orientation of the second tag (3’ to 5’) was reversed. The tags were then listed with their frequency. Each tag should represent a unique transcript and their quantification was a measure of the transcript abundance in the tissue.
2.3.3.2 Statistical analysis
The VSU and VMP libraries were compared according to the tag abundances of each transcript using Microsoft® Access, and Microsoft® Excel. Statistical analysis of VMP up-regulated and VSU up-regulated genes (P<0.05) was performed by Z-test using StatView® 5.0.1 Software from http://www.statview.com, and the Bayesian method (http://igs-server.cnrs-mrs.fr/~audic/winflat.cgi)( Audic and Claverie, 1997).

2.3.3.3 Tag-to-Gene Annotation
The tags were identified according to their transcript using a reference database, the Rat LongSAGEmap located at the NCBI web-site: http://www.ncbi.nlm.nih.gov.uk. The tags were compared to the Rat LongSAGE map in Microsoft® Access, and a file was created listing the LongSAGE tags, their abundance, and the tag-to-gene annotation. Tag-to-gene annotations were verified by BLAST using Ensembl Genome Browser (http://www.ensembl.org).

2.3.3.4 Literature Mining and Candidate Selection
The named genes in the VSU and VMP libraries were processed, to select genes of interest that should be studied further. The aim was to validate the SAGE libraries, and identify mesenchymal genes that may be important during prostate development. To fully interrogate the libraries four routes of research were pursued.

1) Genes that were found to be upregulated/enriched in the VMP Library by statistical analysis (P<0.05) (Appendix). The function of their associated proteins had been determined previously, and secretory proteins, transcription factors and intracellular signalling molecules were considered to be of particular interest because of their potential activity in regulating developmental processes.

2) Genes that were found through Literature Mining using PubMed (NCBI), to be of importance during development of branched organs such as the lung, kidney or breast, or to be members of known developmental pathways. These genes could also be acting in the prostate to regulate growth and development.

3) Genes that were previously found upregulated in prostate cancer by microarray (Lapointe et al., 2004, Rhodes et al., 2002, Stuart et al., 2004). Genes robustly over-expressed in more than one study were identified by R. Chan (research support
technician). It is thought that genes that are present during prostate organogenesis may also be re-expressed during disease state prostate, cancer or benign prostatic hyperplasic. 4) Interesting development associated molecules not previously studied in the prostate. The accuracy of the tag-to-gene annotation for genes of interest was verified by BLAST searches using Ensembl (http://www.ensembl.org) before further experiments.

2.3.3.5 Gene Ontology Assignment
Tags verified by BLAST were grouped according to their function. The assignments were based on searches of Ensembl (http://www.ensembl.org) and GeneCards (http://www.bioinfo.weizmann.ac.il/cards/) provided by the Weizman Institute. The tag assignments were performed by Dr. G. Vanpoucke.

3.1 Introduction

3.1.1 Gene Expression Profiling Techniques

Nearly all the biological events such as cell division and differentiation, responsiveness to hormones or growth factors and ultimately cell death are associated with changes in expression. By comparing gene expression profiles of different tissues or tissues under different conditions, individual genes or groups of genes can be identified that play an important role in a particular signalling cascade or disease process. Methods for assaying gene expression can be grouped into two major types: (1) closed methods, which measure expression levels of already collected clones or sequences, and (2) open methods, which do not require prior knowledge of genes being measured.

The sequencing of cDNA clones provides the first glimpse of gene expression patterns in different human tissues (Adams et al., 1995). Techniques based on subtractive hybridisation (Hedrick et al., 1984) and differential display (Liang and Pardee, 1992) have been useful in identifying differences in amounts of transcript populations. However, these techniques provide only a partial picture and may miss transcripts expressed at low levels. Other methods such as oligonucleotide arrays (Lockhart et al., 1996) have been used to compare the expression of thousands of genes in a variety of tissues including small cell populations (Sgroi et al., 1999, Mahadevappa and Warrington, 1999) but they are limited to analysing only previously identified transcripts. Also, the methods discussed cannot be used for accurately monitoring quantitative differences in genes expression. However, despite the possible caveats to research using micro-arrays, they have proved a popular method for investigation into gene markers of prostate cancer (Dhanasekaran et al., 2001, Lapointe et al., 2004, Ernst et al., 2002, Stuart et al., 2004) and most recently, the Institute of Cancer Research has developed a tissue microarray (TMA)
method to be carried out on needle biopsy samples to look for multiple makers of various genes associated with prostate cancer (Jhavar et al., 2005).

A recent technique, called massively parallel signature sequencing (MPSS) (Brenner et al., 2000), is able to surmount some deficiencies of the previously discussed techniques. MPSS can perform quantitative global analysis of gene expression, and is provided as a service by Lynx Therapeutics, Inc., (California, USA) (http://www.lynxgen.com). This sequencing approach combines non-gel-based signature sequencing with an *in vitro* cloning of millions of templates on separate microbeads. Like other clone-and-count techniques this approach provides a digital picture of the gene expression profile. However, the cost of MPSS can be prohibitively expensive. In this thesis, long serial analysis of gene expression (LongSAGE), a modified version of the gene-profiling method, Serial Analysis of Gene Expression (SAGE) was used (Saha et al., 2002, Velculescu et al., 1995) to investigate the “transcriptome” of the neonatal rat UGS.

### 3.1.2 SAGE and LongSAGE

Serial analysis of gene expression (SAGE) is a high-throughput, high efficiency method to simultaneously detect and measure the expression levels of genes expressed in a cell at a given time, including novel and low abundance genes (Velculescu et al., 1995). It does not depend on the prior availability of transcript information. The technique is comprehensive, unbiased, quantitative, sensitive and reliable. The SAGE technique can be used in a wide variety of applications including the definition of transcriptomes, analysis of the effect of drugs on tissues, identification of disease-related genes and elucidation of development and disease pathways (Tuteja and Tuteja, 2004a).

A major use of SAGE has been the identification of genetic markers of cancer (as reviewed in Tuteja and Tuteja, 2004b). Genetic events including mutations, chromosomal gains, losses and rearrangements, along with epigenetic alterations, lead to significant transcriptional changes in cancer cells. Changes in the expression of many genes associated with the onset and progression of cancer likely contribute
to the cancerous phenotype. By characterising and comparing normal and malignant gene expression patterns, this could yield many new diagnostic and prognostic markers as well as therapeutic targets in cancer. The SAGE technique has been used in prostate cancer research (Xu et al., 2001, Untergasser et al., 2002, Waghray et al., 2001a, Waghray et al., 2001b), and more recently in a study of mouse prostate development (Zhang et al., 2006).

The SAGE technique produces 14bp sequences or "tags" from the 3' end of each transcript that identify one (or more) mRNAs. More recently, a modified version of the SAGE method called long serial analysis of gene expression (LongSAGE) has been described (Saha et al., 2002). The LongSAGE method generates 21bp tags derived from the 3' ends of transcripts that can rapidly be analysed and matched to genomic sequence data (Saha et al., 2002). This method is similar to the original SAGE approach but uses different type IIIS restriction endonuclease (Mmel) and incorporates other modifications to produced longer transcript tags. The resulting 21bp tag consists of a constant 4bp sequence representing the restriction site at which the transcript was cleaved followed by a unique 17bp sequence derived from an adjacent sequence in each transcript. Like SAGE, these sequence tags are concatenated serially into long DNA molecules for sequencing. The frequency of each tag in the sequence reflects the cellular abundance of the corresponding transcripts. This serial analysis of many thousands of gene-specific tags allows the simultaneous accumulation of information for genes expressed in the tissue of interest and gives rise to an expression profile for that tissue. These sequencing data are then analysed to identify each gene expressed in the cell and the levels at which each gene is expressed.

The benefit of using LongSAGE over SAGE is in the accuracy of tag-to-gene annotation. LongSAGE can be used for the rapid identification of novel genes and exons. Theoretical calculations show that >99.8% of 21bp tags are expected to occur only once in the genomes the size of the human genome. Likewise, similar analyses based on actual sequence information from ~ 16,000 known genes suggest that >75% of 21bp tags would expected to occur only once in the human genome, with the
remaining tags matching duplicated genes or repeated sequences (Saha et al., 2002). In contrast, conventional SAGE tags of 14bp do not allow unique assignment of tags to genomic sequences, though they do allow assignment to the much less complex compendium of expressed sequence tags and previously characterized mRNAs (Velculescu et al., 1995). An I-SAGE and I-LongSAGE kit is supplied by Invitrogen™ Life Technologies to produce SAGE and LongSAGE libraries.

Since the invention of SAGE technology, a number of modifications have been reported in the technique. Some of the important modifications include SAGE-lite (Peters et al., 1999), micro-SAGE (Datson et al., 1999), mini-SAGE (Ye et al., 2000), a SAGE adaptation for downsized extracts (SADE) (Cheval et al., 2002), SuperSAGE (Matsumara et al., 2005), Robust LongSAGE (Gowda et al., 2004), aRNA-LongSAGE (Heidenblut et al., 2004) and generation of longer cDNA fragments from SAGE tags for gene identification (GLGI) (Chen et al., 2000).

### 3.1.2.1 SAGE Bioinformatics

The sequence data generated by a typical SAGE experiment is analysed following a typical procedure (Fig 3.1). Invitrogen™ Life Technologies supplies SAGE2000 Version 4.5 software available for download from http://www.invitrogen.com/sage to purchasers of the I-SAGE or I-LongSAGE kits. The core functions of the SAGE software are: (1) extraction and tabulation of tag sequences and counts from raw sequence files, (2) comparison of tag abundances between projects and (3) matching tags to reference sequences in other databases.

### 3.1.2.2 Databases of SAGE data

SAGE is a powerful technology that has been adopted as a standard for gene expression. SAGE has become increasingly popular and as the number of users has increased, the size and number of SAGE databases has grown. There are presently a number of SAGE databases available, some commercial and some public.
Fig 3.1 Steps for data analysis of SAGE
The SAGE clones are sequenced and the tags are extracted after removal of duplicated ditags by SAGE software. The software also quantifies the transcript abundance of the SAGE clones. The observed tags are then matched to known sequences from various sources and databases. Following this, the sorting and filtering of tags based on differential expression reveals candidates for further investigation.

The Genzyme SAGE database is an example of a commercial database. This database is an integral part of Genzyme’s therapeutic discovery efforts, and includes over 4 million SAGE tags, believed to represent over 100,000 genes from major cancers and normal human tissues. Other non-commercial methods are also available for determining gene identification from a SAGE tag, namely SAGE Genie (http://cgap.nci.nih.gov/SAGE) and USAGE (http://www.cmbi.kun.nl/usage)(van Kampen et al., 2000). In this thesis, a publicly available gene expression database was used, LongSAGEmap (http://www.ncbi.nlm.nih.gov/SAGE/), for high-throughput tag-to-gene annotation.
3.1.2.2.1 SAGEmap

A valuable and commonly used resource for SAGE tag-to-gene mapping is NCBI’s SAGEmap (http://www.ncbi.nlm.nih.gov/SAGE/). LongSAGEmap is a variant of SAGEmap based on 17bp (21bp-CATG) transcript sequences as opposed to 10bp sequences. Tag mapping refers to the unambiguous determination of the gene represented by a SAGE tag. As described by Lash et al, (2000), this tag mapping is done on the basis of matching SAGE data to tags extracted from Expressed Sequence Tag (EST) and cDNA sequences found in the NCBI Unigene clusters (Schuler et al, 1997) (http://www.ncbi.nlm.nih.gov/Unigene/). The Unigene database is composed of expressed sequences (ESTs and cDNAs) clustered on the basis of sequence similarity to gene sequences found in Genbank. The 3’ most predicted SAGE tags are extracted from each expressed sequence, and experimentally derived tags can be compared with these predicted tags to determine the Unigene cluster that most likely represents the gene from which the experimental SAGE tag was derived.

Unfortunately, there are several drawbacks to this method of tag mapping, primarily due to the underlying data. The same gene may be represented in multiple clusters, resulting in a single tag mapping to multiple clusters, and thus being erroneously assigned as ambiguous. The sequences in the clusters are primarily ESTs, which have an error rate estimated at ~1% (Lash et al., 2000). This error rate will result in tags not matching the appropriate cluster and/ or matching an incorrect cluster. As this approach is based primarily on expressed sequences, any genes that are not represented by ESTs or cDNAs in Genbank, will not be represented by tags in SAGEmap, resulting in tag-to-gene assignments that are missing or potentially erroneous. As a result, tag assignments can be ambiguous, incorrect or unavailable, leading to potentially incomplete or erroneous gene expression interpretation (Pleasance et al., 2003).

In summary, the most reliable are tag-to-gene matches derived from well-characterised mRNA/cDNA sequences from Genbank. Of less reliability are tags extracted from EST sequences, which can be further partitioned into classes depending on the presence of absence of a polyadenylation signal and if they have a
3' or 5' orientation. The most reliable among the EST sequences are those with a polyadenylation signal and/or polyadenylation tail and annotated as 3' sequences. Fortunately, it is possible to check the accuracy of tag-to gene assignments using BLAST, which searches genomic sequences, although of course this is not feasible for a whole Library. Originally started as a database of gene expression, SAGEmap now accepts SAGE sequence data from any source. This database presently contains over 3 million SAGE tags for expressed sequences from over 50 SAGE tag libraries. By accessing SAGEmap the user can compare transcript populations between any of the posted libraries (Lash et al., 2000).

In summary, the LongSAGE method was used in this thesis to generate a comprehensive and quantitative gene expression profile of the prostate precursor. LongSAGE was used because it was an open-ended method of gene profiling, does not require prior knowledge of transcript expression, and had the potential for identification of novel genes. LongSAGE was used in preference to SAGE because of the improvement in the accuracy tag-to gene-annotation process.

### 3.2 Results

#### 3.2.1 Library Preparation

The LongSAGE Libraries were prepared according to a modified protocol (Fig 2.2) presented in Materials and Methods. The modifications made were recommended on an EMBO SAGE practical course (Jan, 2003), and were also included in a LongSAGE protocol released by Genzyme Molecular Oncology (Jan 2003). The aims of the modifications were to increase the efficiency of the SAGE technique, and to produce Libraries of 21bp instead of 14bp tags.

A summary of the changes made to the conventional I-SAGE protocol is presented in Table 3.1. The LongSAGE method starts with more total starting RNA than SAGE, and isolates a tag of 21bp (17bp +CATG) rather than 14bp (10bp +CATG) by changing the tagging enzyme, from *BsmF1* to *MmeI*. The action of the tagging enzyme (*MmeI*) was such that no Klenow reaction was required to ligate the tags,
and there was an additional SphI digestion step of the concatemer fragments to improve their cloning efficiency.

### Table 3.1: LongSAGE protocol modifications

<table>
<thead>
<tr>
<th>Library Features</th>
<th>LongSAGE modifications made</th>
<th>Conventional SAGE (Velculescu et al., 1995)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total RNA quantity</td>
<td>20- 50 μg</td>
<td>5-10μg</td>
</tr>
<tr>
<td>Tagging enzyme</td>
<td>Mme I</td>
<td>BsmF I</td>
</tr>
<tr>
<td>Klenow Reaction</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SphI digestion of concatemer fragments</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

### 3.2.2 The VSU and VMP LongSAGE Libraries

To identify mesenchymal factors involved in prostate development, two LongSAGE libraries were made from female neonatal rat prostate rudiment (Fig 2.1). A flow-diagram of the step-by-step preparation of the two libraries is presented in Fig 2.2. The VSU library was generated first, followed by the VMP library. Preparation and sequencing of each Library was completed in ~9 months. The main features of the libraries are summarised in Table 3.2.

### Table 3.2: Features of the VSU and VMP LongSAGE Libraries

<table>
<thead>
<tr>
<th>Library Features</th>
<th>VSU Library</th>
<th>VMP Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tissue collected</td>
<td>~100 VSU's</td>
<td>~90 VMP's</td>
</tr>
<tr>
<td>Starting Tissue Weight</td>
<td>~180 mg</td>
<td>~65 mg</td>
</tr>
<tr>
<td>Total RNA</td>
<td>50 μg</td>
<td>21 μg</td>
</tr>
<tr>
<td>130bp Template Dilution per Reaction</td>
<td>1:40</td>
<td>1:40</td>
</tr>
<tr>
<td>No. PCR Cycles</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>No. PCR Reactions</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>Concatemer fragment size insert</td>
<td>700-1kb</td>
<td>500-800bp</td>
</tr>
<tr>
<td>Number of minipreps</td>
<td>2,747</td>
<td>3,788</td>
</tr>
<tr>
<td>Number of tags per concatemer insert</td>
<td>24</td>
<td>21</td>
</tr>
</tbody>
</table>
Fig 3.2: Flow Diagram of the LongSAGE protocol.

**Gel 1:** PCR check for cDNA synthesis from RNA (Lanes 1; dH2O, 2; +ve (GAPDH), 3; +ve (EFα1)), Nla III restriction enzyme digestion of cDNA (lanes 4; +ve, 5; no enzyme (EFα1)), and ligation of Linkers LS1A/B and LS2A/B (40bp) to cDNA pool A or B (Lanes 6; LS1A/B + pool A cDNA, +ve, 7, dH2O (EFα1)), (Lanes 8; LS2A/2B + pool B cDNA, +ve, 9, dH2O (EFα1))

**Gel 2:** PCR check for digestion of each pool of cDNA + linkers with Mme 1 to isolate tags (resulting in loss of EFα sites) (Lanes 1,2 (EFα1)) and ligation of the two pools of tags plus linkers (~65bp) to form 130bp ditags (lane 3).

**Gel 3:** PCR amplified ditags (130bp) separated from ligated Linkers (~80bp).
Fig 3.2 (cont): Flow-Diagram of the LongSAGE protocol

Gel 4: Separation of 34bp ditags, from Linkers (~40bp) and ligated Linkers (~80bp).

Gel 5: Ditags were ligated to form concatemers

Gel 6. The size of concatemer fragment insert was checked by restriction enzyme digestion.
The differences in the features of the library preparations will be explained more fully here. Firstly, a greater amount of total RNA used to make the VSU (50μg) than VMP library (21μg). Consequently, in the PCR amplification step, which was performed to generate a minimal amount of concatemer cDNA for cloning, 27 PCR cycles were required for the VMP library and only 23 PCR cycles were required for the VSU library. In other words, by maximising the RNA input, this minimised the amount of PCR amplification needed. Furthermore, minimal PCR amplification reduced the possibility creating a transcript bias in the Libraries, where the high abundance tags could overwhelm the low abundance tags making them more difficult to find.

Additionally, a minor change was made to the size of concatemer fragments insert cloned into the pZero1 vector, from 700bp-1kb (VSU) to 500-800bp (VMP). Initially it was thought that this might make the Sequencing process more efficient because it usually returns only ~700-750bp of sequence, and therefore, any extra sequence is lost. However, as small inserts are preferentially cloned, it was the 500bp concatemer fragments from the VMP Library that were cloned more often. These obviously contain fewer ditags than the 700bp minimum concatemer fragment of the VSU Library, and meant it was necessary to increase the number of minipreps and amount of Sequencing in order to generate a VMP Library of similar size.

3.2.3 Bioinformatics - Library Analysis

3.2.3.1 Extracting the SAGE tag data

The raw sequence data from the two SAGE libraries was compiled and analysed using SAGE 2000 version 4.5 software provided by K. W. Kinzler (http://www.sagenet.org). The SAGE 2000 software tabulated the tags using the raw sequence files, and quantified the tags according to their abundance. During this process, any Linkers and duplicate dimers were removed from the sequence data by the software. The Linkers are the ~40bp sequences containing the Mme 1 restriction site, that are linked to the ditags during the LongSAGE protocol, and may not have been removed from the cDNA pool by Nla III digestion and PAGE gel purification. These sequences are not part of the ditag sequence information and were therefore
discarded. Duplicate dimers are when two identical tags form a ditag. This occurrence is so statistically improbable that it is considered to be a sequencing error or a PCR artefact, and so should be discarded.

A summary of the VSU and VMP library data is presented in Table 3.3, and the full VSU and VMP library tag lists have been saved to CD (Tables 3.4 and 3.5, respectively), and have been appended in a pocket at the back of the thesis. The tag lists were annotated according to the process of Tag-to-Gene Annotation (2.3.3.3). The lists include the tag sequences, the tag abundance, tag-to-gene annotation if found, and the relevant reference Unigene ID number from Rat LongSAGEmap (NCBI).

Table 3.3: Summary of the VSU and VMP SAGE Library Data.

<table>
<thead>
<tr>
<th></th>
<th>VSU Library</th>
<th>VMP Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Tags</td>
<td>67,833</td>
<td>80,008</td>
</tr>
<tr>
<td>Unique Tags</td>
<td>22,755</td>
<td>26,932</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tag Categories</th>
<th>No. Tags</th>
<th>% of Library</th>
<th>No. Tags</th>
<th>% of Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>High 200-500+</td>
<td>21</td>
<td>0.1</td>
<td>24</td>
<td>0.1</td>
</tr>
<tr>
<td>Med 30-199</td>
<td>234</td>
<td>1</td>
<td>257</td>
<td>1</td>
</tr>
<tr>
<td>Low 2-29</td>
<td>6,867</td>
<td>30.2</td>
<td>7,824</td>
<td>29</td>
</tr>
<tr>
<td>Singletons</td>
<td>15,633</td>
<td>68.7</td>
<td>18,827</td>
<td>69.9</td>
</tr>
</tbody>
</table>

3.2.3.2 Comparison of the VSU and VMP Libraries
The tag sequences and their abundances were tabulated from each library and imported into Microsoft® Access for further analysis. Using the Microsoft® Access software, a comparison was made between all tags in the VMP library with their matching tags in the VSU Library. This meant that all VMP tags, but not all VSU tags were represented, as some transcripts from the VSU library would not be present in the VMP library. The VMP Library was found to have 2,044 unique tags (>1 tag count), specific to the VMP but not present in the VSU Library. The converse of this
process, comparing “all tags” in the VSU library with only matching tags in the VMP library was also performed (data not shown).

3.2.3.3 Statistical Analysis
The VSU and VMP library comparisons described in Section 3.2.3.2, were analysed for statistically significant differences in the tag abundances. The comparison between “all tags” in the VMP and matching VSU tags as described in Section 2.3.3.2 was imported into Microsoft® Excel and manipulated to perform two statistical tests, a Z-test and Bayesian analysis. The level of statistical significance for both tests was set at P < 0.05. Both tests produced the same list of VMP up-regulated genes (Appendix). Similarly, a comparison made between “all tags” in the VSU and matching tags in the VMP was also subjected to the same statistical analyses (data not shown).

3.2.3.4 Tag-to-Gene Annotation
The tags in the VSU and VMP libraries were processed to determine the identity of their associated genes. By using the process described in Section 2.3.3.3 it was attempted to match all tag sequences in each library to their associated genes using Rat LongSAGE map (Lash et al., 2000). A modified version of rat SAGEmap (10bp sequences) was requested from NCBI. LongSAGE map was compiled with 17bp reference sequences to improve the tag-to-gene annotation of the LongSAGE tags. Accurate gene assignments are essential if the named genes will be the basis of further studies. Therefore, a more detailed analysis of the VMP-enriched genes using BLAST was performed (Fig 3.3), to ensure the tag –to –gene annotations were accurate. The results of the Library comparison, statistical analysis and tag-to-gene annotation is included in the Appendix.

Of the 218 VMP-enriched tags (P<0.05) (Appendix), 143 tags were identified as “known”, that is they were associated with named rat transcripts. The accuracy of gene assignment was verified by BLAST searches of the genomic sequences (Ensembl). The remaining 75 tags (35%) were “unknown” according to the LongSAGE map, and could not be associated with a named transcript (Table 3.4).
To identify the “unknown” tags, BLAST searches (Ensembl) were performed. The majority of BLAST searches were performed by Dr. G. Vanpoucke.

![Flow-diagram of the tag-to-gene annotation process](image)

**Fig 3.3: Flow-diagram of the tag-to-gene annotation process**

After performing BLAST searches on the 75 ‘unknown’ tags, 42 tags were found to give a unique hit in the rat genome, and to lie with 5kb of a gene locus and could be identified. Using BLAST, increased the number of tag-to-gene assignments from 143 to 185, and is useful for accurately identifying smaller numbers of tags, but is not practical for use on LongSAGE libraries. A comparison of the tag-to gene annotation process by LongSAGEmap and BLAST is presented in Table 3.6

Finally, to name the remaining 19 ‘unknown’ tags, it would be necessary to perform GLGI also known as 3’RACE (Chen et al., 2002), a technique that allows the acquisition of 50-250bp of further sequence data using the original 21bp tag. With this extra sequence information it would be easier to identify the associated transcripts by BLAST searches. GLGI was not undertaken as part of my research.

The process of tag-to-gene annotation was similarly performed for the VSU-upregulated genes (P<0.05) (data not shown) and several genes of interest were identified by genomic BLAST searches (Ensembl). Tags associated with epithelial markers, cytokeratins 8, 19 smooth muscle markers and transgelin were found in the list of VSU up-regulated genes (data not shown).
Table 3.6: Comparison of rat LongSAGEmap vs. BLAST for tag-to-gene annotation of VMP-enriched tags.

<table>
<thead>
<tr>
<th>Genes Identified by LongSAGEmap</th>
<th>VMP upregulated P&lt;0.05 (gene no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Known genes</td>
<td>143</td>
</tr>
<tr>
<td>EST’s</td>
<td>63</td>
</tr>
<tr>
<td>No match</td>
<td>12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>218</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BLAST searches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Known genes</td>
</tr>
<tr>
<td>EST’s</td>
</tr>
<tr>
<td>No match</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

Known genes = named genes and Rat EST’s with similarity to a named gene
EST’s = rat EST’s with no similarity to a named gene
No Match = No hit or multiple hits in the rat genome, or a hit in the wrong direction.

3.2.3.5 Assignment to Gene Ontology

The genes identified as VMP-enriched (P<0.05), were assigned a molecular function to their proteins. The assignments were based on searches of Ensembl (http://www.ensembl.org) and GeneCards (http://www.bionfo.weizmann.ac.il/cards/) provided by the Weizman Institute. The assignments were performed by Dr. G. Vanpoucke, and are presented in Graph 3.1. The proteins were divided into categories according to their function, and 60% of the proteins were assigned a known function. For the remaining 40% of the tags (~ 88), half were classified as ‘hypothetical proteins’ (20%) and half as ‘unknown’ function (20%). The majority of the tags classified as “unknown” proteins (38 tags) had not been assigned to a gene. The “hypothetical proteins”, were assigned from “known” identified genes, but had no known function, and the majority of these were identified as novel transcripts.
Graph 3.1: Functional classification of VMP-enriched genes from the SAGE Libraries.
A pie chart depicting the percentage of distinct genes encoding proteins in various functional categories in the list of 218 VMP-enriched genes (P<0.05) identified by statistical analysis of the SAGE library data.
### 3.2.3.6 Developmental Pathways and Regulatory Genes

The LongSAGE libraries were interrogated for ‘genes of interest’ (Table 3.7). Three routes of investigation were pursued:

1) genes that are already known as being important during prostate development,
2) genes that are members of a known developmental pathway
3) cDNAs/transcripts associated with prostate cancer

<table>
<thead>
<tr>
<th>Regulatory Genes</th>
<th>Development of Prostate Cancer</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate Development</td>
<td>Upregulated in Prostate Cancer</td>
<td>androgen receptor, FGF10, smoothened, BMP4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MMP2, Paralemmin, PMP22, VAP1, Calnexin, TF8, Four and a half LIM domains, Hepsin</td>
</tr>
<tr>
<td>Notch</td>
<td>Notch1, Notch2, (Dlk1?) lunatic and radical fringe</td>
<td></td>
</tr>
<tr>
<td>Wnt</td>
<td>Wnt4, Wnt11, Sfrp2, Frzld1, Dkk3, Disabled2, Dishevelled 1,</td>
<td></td>
</tr>
<tr>
<td>TGFβ /BMP</td>
<td>TGFβ2, TGFβ1, Ltpb4, BMP4 and 6, SMAD 1, SMAD4 and SMAD5.</td>
<td></td>
</tr>
<tr>
<td>FGF</td>
<td>FGF10, FGF13, Sprouty1, FGFR1 and 2.</td>
<td></td>
</tr>
<tr>
<td>Hedgehog</td>
<td>smoothened</td>
<td></td>
</tr>
</tbody>
</table>

### 3.2.4 Summary of Results

The LongSAGE technique was used to prepare two libraries from the female prostate rudiment. The VSU library was made from the whole UGS, containing the ventral mesenchymal pad (VMP), smooth muscle and urethral epithelium, and the VMP Library was made from the microdissected VMP only. The LongSAGE tags were extracted from the raw sequence data, and quantified according to their tag count. Tag-to-gene annotation was performed in a high-throughput manner on both
Libraries, using LongSAGEmap, and the accuracy of the gene assignment was verified by BLAST searches. The contents of the Libraries were validated by the presence of established developmental and regulatory pathways, including Fgf10, AR and BMP4, mesenchyme-specific factors, known to be active during prostate development. To identify genes that demonstrated VMP enriched/specific expression, a back-comparison was made of the VSU and VMP libraries, based on their tag counts. Statistical analysis (Z-test) of the library comparison highlighted 218 VMP-enriched (P<0.05) genes (Appendix).

3.3 Discussion of the Library Preparation and Analysis

LongSAGE is an expensive, time-consuming and multi-step technique, which is not in common usage. Therefore, it was thought to be important to provide the reader with a detailed insight into the preparation and analysis of the Libraries. This section aims to provide further background on the LongSAGE method, the choice of tissue used, and analysis process undertaken.

LongSAGE was chosen as a genome profiling method over other techniques, such as differential display PCR, microarrays and MPSS, because they were either too time-consuming, relied on prior knowledge of transcript expression eliminating the possibility of identifying novel genes, or were too expensive, respectively. Also, when this project was initiated microarrays provided poor coverage of the rat transcriptome and there were concerns over their sensitivity and reproducibility between different laboratories. The most efficient but expensive method, MPSS (Jongeneel et al., 2003, Brenner et al, 2000) generates millions of signature sequences from cells and tissues would have provided the most in-depth analysis.

The LongSAGE method was used instead of SAGE because it generated 21bp tags instead of 14bp tags, allowing the unique identification of the associated transcripts (Saha et al., 2002). This increased tag size helped to identify rat transcripts with improved accuracy. However, this also created a trade-off between the need for accurate tag-to-gene annotation, and the depth of library sampling. It is more
expensive to generate a LongSAGE than a SAGE library of equivalent size because the tag concatemers sequenced comprise fewer 21bp than 14bp tags.

3.3.1 The VSU and VMP LongSAGE Libraries
The LongSAGE libraries were generated to profile the gene expression of the mesenchyme/stroma of the prostate rudiment. The Libraries were prepared from VSU and VMP tissue carefully microdissected from the female UGS (P0). The quality of the tissue was paramount. Using female tissue to make the Libraries was proposed to be valid, based on the hypothesis that females produce prostatic inducing molecules constitutively (but that these are not ‘activated’ unless androgens are present). Also, studies have shown that it is clear that females have the capacity to form a prostate but they do not do so because of a lack of androgens (as reviewed by Thomson, 2001). Female UGS (E16.5) can be induced to produce epithelial buds by treatment with Testosterone in vitro (Takeda et al., 1986). Consequently, as the ‘andromedin hypothesis’ of prostatic induction (regulation of mesenchymal factors at the level of gene transcription by androgens) remains controversial, it was argued that studying the female UGS would not be a disadvantage in finding ‘inductive’ molecules.

Neonatal tissue was used as opposed to embryonic tissue due to its larger size, which was beneficial for preparing large amounts of RNA. Also, the female rat UGS had the advantage of a homogeneous mesenchyme, the VMP (P0), with no contaminating epithelial and smooth muscle tissue, unlike the male prostate at the same age.

An additional concern about the experimental design of this project was that in using the VMP, any factors that are ‘triggered’ as a consequence of invading epithelium, would not be expressed. A common mechanism of branching organs is for epithelial tips to induce a local change in the mesenchyme as it invades (as reviewed by Davies, 2002). However, alternative experimental designs that considered this problem were also found to contain caveats.
Two possible approaches suggested, were to compare LongSAGE libraries prepared from male UGM and UGS (E17.5), or male UGS and female UGS (E17.5 or P0). However, preparing libraries from embryonic tissues would have required sacrificing a considerable number of dams, and it would be extremely technically difficult to isolate the UGM. Zhang et al., (2006) has reported isolating E16.5 UGM in the mouse using enzymatic and mechanical separation, but it is suggested that this may compromise the quality of the RNA. Furthermore, it is proposed that if LongSAGE Libraries of both male and female UGS (P0) had been generated and compared, the heterogeneous nature of the tissues, and the developmental changes in cell proportions between the male and female UGS, would make it very difficult to determine transcript expression in distinct cell populations. It would not have been a useful approach for identifying the mesenchyme-enriched genes.

It is concluded that the approach used, comparing neonatal female UGS and VMP, has proved to be relatively successful so far, in identifying mesenchyme-enriched molecules. However, if the project had not been aimed at identifying primarily mesenchymal factors, and also was awarded unlimited time, funding and manpower, it is proposed that several Libraries would have been prepared, from embryonic to adult prostate, including Libraries of the isolated male and female UGM, and prostate cancer stroma. It is noted that Zhang et al., (2006) has reported preparing several, bigger SAGE libraries to study gene expression in the developing mouse prostate. Also, a more comprehensive picture of changes in gene expression during prostate development could be studied with the more powerful gene profiling technique, MPSS.

The VSU and VMP LongSAGE libraries were prepared from microdissected neonatal female UGS, using a modified SAGE protocol (Invitrogen). It would have been ideal to generate two VMP Libraries in addition to the VSU Library, as this would have ensured that the data was reproducible, but limitations in tissue collection and time available for experiments were prohibitive. It was understood that tissue dissections are not always accurate and by preparing two VMP Libraries
Chapter 3 SAGE

this would have acted as a 'control', and the Libraries could have been compared to confirm if the LongSAGE technique could produce reproducible, reliable results. Also, the 'cDNA amplification step' was highlighted as a stage in the technique that could introduce significant errors into quantification of transcript abundance. It was recognised that the amplification of cDNA by PCR would favour the high abundance transcripts, exponentially increasing their expression more rapidly than the low abundance transcripts, and could create a transcript bias within the Libraries. Consequently, a greater amount of starting RNA was used to prepare the Libraries than recommended, in order to minimise the number of PCR cycles necessary at the amplification step. Also, a greater amount of RNA was used for the VSU Library than the VMP Library because it was hypothesized that this would ensure the low abundance tags in the more complex VSU tissue, would be adequately represented, and because of the difficulty of obtaining a large amount of VMP tissue due to its size. Subsequently, it was noted that the VMP Library contained more unique tags than the VSU library even though they were sampled to a similar depth (70-80,000 tags). It was concluded that the VMP Library (26,932 unique tags) gene profile was likely to be more representative of its homogenous source tissue than the VSU Library (22,755 unique tags) prepared from the more complex VSU, and that further sequencing would be required to generate an accurate VSU transcriptome profile.

3.3.2 Bioinformatics - Library Analysis

3.3.2.1 Extracting the Tag Data

The VSU and VMP libraries were sequenced to a depth of ~70-80,000 tags. The VSU library had 67,833 tags, comprising 22,755 unique tags. The VMP library had 80,008 tags, comprising 26,932 unique tags. The size of the libraries was carefully considered based on a report by Velculescu et al., (1999), who stated that a range of 50-100,000 tags in a SAGE library (of a single cell line) was optimal, and that relatively few new tags were found with more in-depth sampling. They proposed that ~90% of the unique tags in a SAGE library of up to 650,000 tags, were present in a sampling of 50,000 tags. This thesis has reported that the VSU Library was not sequenced adequately at ~80,000 tags, and recently Zhang et al., (2006) concluded
that SAGE libraries of the developing mouse prostate required deeper sequencing than 100,000 tags to reveal the entire transcriptome.

Fortunately, identifying the entire transcriptome of the VSU was not critical to finding VMP-enriched genes, and it also highlighted the benefit of using the homogeneous VMP, to prepare the second Library. Additionally, expression of some low abundance VMP-enriched transcripts i.e. Fgf10, AR and BMP4, was confirmed in both Libraries, and therefore the LongSAGE technique and the sampling of 80,000 tags was verified to be sensitive enough to identify similarly expressed transcripts. Further sequencing would have been expensive and could not be justified based on the research by Velculescu et al., (1999), that few further unique tags would be found.

### 3.3.2.2 Tag-to-Gene Annotation

The tags isolated in the VSU and VMP LongSAGE libraries were used to identify their associated transcripts/genes. Using 21bp LongSAGE tags instead of 14bp SAGE tags improves the accuracy of the tag-to-gene annotation process (Saha et al., 2002). However, even with the benefit of having longer tags, the data in the VSU and VMP Libraries demonstrated that several in individual tags had been assigned to different multiple genes. Also, conversely, in several instances, different tags had been assigned to the same gene. The proportion of the whole Libraries that were poorly annotated was not determined, although it was demonstrated that only 65% of the tags were assigned in the list of 218 VMP-enriched tags (P<0.05) using the LongSAGEmap method.

This poor rate of accurate tag-to-gene annotation can be explained by some potential problems that had been highlighted in previous studies. 1) Rat sequences are under-represented (particularly at the 3' end) in the Unigene database, in comparison to the human and mouse. Therefore, a single rat tag can be wrongly matched (by LongSAGEmap), to a sequence that is present in different transcripts, but which does not represent the most 3' end. The tag will be assigned several gene names. 2) Partial digestion of the cDNA by the tagging enzyme *NlaIII* results in heterogenous cleavage at the poladenylation site, producing multiple tags for one transcript. 3) One
gene can have multiple transcripts because of gene splicing, alternatively spliced or polyadenylated mRNA sequences (Pauws et al., 2001), and therefore different tags may represent different isoforms of the same gene. 4) 10% of SAGE tags potentially contain a sequencing error, therefore cannot be correctly assigned to a gene (Hillier et al., 1996, Lash et al, 2000). It is suggested here that sequencing of the longer LongSAGE tag increases the possibility of inclusion of a sequence error. 5) Genes that do not have a CATG (Nla III digest site) will not be included in the LongSAGE protocol if NlaIII was used. To address this situation it is possible to use a different “tagging” enzyme such as Sau3A.

Unfortunately, there is little that can be done to solve the several drawbacks of this method of tag mapping, which are primarily due to the underlying data. Therefore, until the rat genome is more fully characterised, genes that may be of interest could unfortunately remain hidden within rat Libraries. It could be worthwhile re-examining the Libraries as the rat transcriptome becomes better characterised

Consequently, the tag-to-gene annotation resulting from matches to the rat LongSAGEmap was not wholly reliable, the transcripts identified as VMP-enriched (P<0.05) were verified by BLAST (Ensembl). BLAST uses rat genomic data, which is more fully characterized than rat transcript data available (used to compile the rat LongSAGEmap). BLAST was able to annotate a further 42 tags (~19%) which had previously been classified as “unknown” by the LongSAGEmap method, and also gave additional confidence to the accuracy of the gene assignments. BLAST was not suitable for the high-throughput process required for annotation of whole SAGE libraries, although it was recommended that BLAST should be used to verify the transcript identities prior to further experimental analysis.

10.2.3 Differential Tag Expression in the LongSAGE libraries
A process of ‘back-comparison’ was performed between the VMP and VSU Libraries, which highlighted 218 VMP-enriched tags (P<0.05) (Appendix A) and 103 VSU-enriched tags (P<0.05) (data not shown).
Performing a ‘back-comparison’ between the VSU and VMP Libraries was proposed because it would help to highlight the VMP-enriched low abundance transcripts. As the VMP is only a subset of the VSU, it was thought that genes which are VMP-specific/enriched would be likely to be represented as either very low or absent in a VSU Library, and so will be more apparent by comparison of the Libraries.

Additionally, the list of VSU-enriched tags (P<0.05), was found to contain known epithelial cell markers, i.e. cytokeratin 8 and 19, and smooth muscle markers, i.e. transgelin and myosin light chain. This helped to support the validity of the data in the Libraries. Specifically it showed that there were significant differences in gene expression depending on the tissues that had been used to make the Libraries, and supported analysis process used. In other words, if expected gene markers could be identified in the VSU-enriched data (P<0.05), there could be more confidence in the accuracy of the VMP-enriched/specific transcript data (P<0.05).

However, one significant problem noted by statistical analysis of the Library comparisons was that for a tag to be identified as demonstrating statistically significantly different expression (P<0.05) between the Libraries, the minimum difference in tag counts had to be at least 5 vs. 0 tags. This meant that all tags present at ≤ 4 counts in both Libraries would be discarded by statistical analysis, even though they would be likely to contain some interesting but low abundance genes. Consequently, the low abundance transcripts (≤ 4) were analysed separately and interrogated for possible “genes of interest”. Even though candidates from this group have not been selected in this project, some have been included other studies.

### 3.3.2.4 Gene ontology of VMP-enriched Tags

The proteins associated with the VMP-enriched tags (P<0.05) were assigned a function, to assist in the selection of candidate genes for further studies. Unfortunately, it was only possible to assign a molecular function to 60% of the tags, while the remaining tags (40%, or ~88 tags) were evenly split between ‘hypothetical’ proteins, and proteins of ‘unknown’ function. The poor level of assignment is suggested to be the result of the inadequate characterisation of the rat genome. The
groups that were highlighted as most likely to contain regulatory developmental genes included the secreted factors (5%), transcription factors (6%) and intracellular signalling molecules (7%). In particular, it was proposed that secreted factors would be ideal candidates that could be active in a paracrine signalling pathway, from the mesenchyme to the epithelium.

3.3.2.5 Selection of Genes of Interest
The Libraries were interrogated for genes that are members of known signalling pathways expressed during prostate development, in order to validate the SAGE data, and identify possible mesenchymal candidate genes. Also highlighted were genes that have been reported upregulated during prostate cancer.

It was proposed that some of these genes, although not identified by the bioinformatic and statistical analysis approach, might also represent possible mesenchymal candidate genes. This process, described as intuitive selection, is a recognised method of selecting candidate genes from gene profiles, which has been extensively reported. By using an additional method of candidate selection it was hoped to reduce the chance of ‘missing’ candidate genes through problems with tag quantification in the Libraries and the subsequent bioinformatic analysis. For example, the low abundance genes (tags counts of ≤ 4) that were excluded as possible mesenchymal candidates by statistical analysis of the library comparison, but still may represent important factors.

Genes that were intuitively selected from the Libraries, included members of known developmental and regulatory pathways i.e Notch, TGFβ1/BMP, FGF, Hedgehog and Wnt and genes associated with prostate cancer including, MMP2, Plm, PMP22 and VAP1. This provided a “snap-shot” of the likely molecular processes taking place in these tissues, while also validating the data in the Libraries. Additionally, identification of cancer–associated genes in a developing tissue, supports the hypothesis that regulatory developmental genes may be active during carcinogenesis (McNeal et al., 1969).
4. LightCycler Analysis and Selection of Candidate Genes

4.1 Introduction

After generating the LongSAGE Libraries, the aim of this project was to characterise the expression and function of a small number of candidate mesenchymal genes such that a gene identified by SAGE could be demonstrated to have a function in the developing prostate.

Quantitative RT-PCR was used as a preliminary means of validating the LongSAGE data by an independent technique. In other words, qRT-PCR was used as a screening process for confirming the accuracy and reliability of the SAGE data. Unfortunately, due to the expense of LightCycler analysis, and also insufficient time, it was not possible to study and categorise a greater proportion of the gene expression in the LongSAGE Libraries, as has been reported in some other cells/tissues (as reviewed by Tuteja and Tuteja, 2004(a)). Therefore, only a few genes were selected for further analysis, as likely mesenchymal candidates.

4.2 Results

4.2.1 Selecting Genes for Further Analysis

Genes were selected from the VMP and VSU SAGE libraries for further analysis. The categories considered can be summarised as follows;
1) VMP-enriched transcripts (P<0.05),
2) genes found in the SAGE libraries from known developmental pathways
3) genes associated with identified developmental pathways, but not identified in the libraries,
4) transcripts /cDNAs present in the libraries and upregulated in prostate cancer studies (Lapointe et al., 2004, Rhodes et al., 2002, Stuart et al., 2004).
A group of 26 transcripts were selected for further experimental studies. The accuracy of the tag-to-gene annotation process by the rat LongSAGEmap was verified by BLAST searches of the tags against the rat genome (Ensembl).

4.2.2 Experimental Validation of LongSAGE Libraries

Initial analysis of the selected genes was performed by the LightCycler (qRT-PCR), to determine if they were VMP-enriched, by an independent technique. Levels of transcript expression in VSU and VMP tissue were normalised by housekeeping gene levels (TBP), and then compared. The experiment was repeated on at least two individual batches of VSU and VMP cDNA, synthesised from their own freshly prepared RNA samples. The primers designed for normal PCR were also used for the LightCycler reaction (Table 2.4). When LightCycler analysis was performed on genes that had not been identified in the SAGE libraries, transcript expression in the VSU and VMP was initially tested with general RT-PCR. LightCycler analysis was performed on the relative transcript expression levels of 26 genes in the VSU and VMP. The results divided the transcripts into three groups; 1) VMP-enriched transcripts (6), 2) VSU-enriched transcripts (9) and 3) transcripts with no difference in expression levels in the VSU and VMP (11).

4.2.2.1 VMP-Enriched Genes Verified by LightCycler Analysis.

LightCycler analysis identified 6 of the 26 genes tested as 'VMP-enriched', that is showed at least a 1.5 fold higher level of expression in the VMP in comparison to the VSU. The fold-difference in VMP/VSU transcript expression for the six 'VMP enriched' genes, according to the LongSAGE and LightCycler analysis results are presented in Graph 4.1. Three of the genes were identified as VMP-enriched; Dlk1 (2.4, 1.5-fold), Notch2 (6, 1.8-fold), Nell2 (5, 1.7-fold) by SAGE tag counts (P<0.05) and LightCycler analysis, respectively. The 3 remaining genes (MMP14, Fgf 10 and Ptn) were not identified as VMP-enriched (P<0.05) by SAGE analysis and were analysed by LightCycler after intuitive selection from the Libraries. MMP14 (1.5-fold), Fgf10 (2.4-fold) and Ptn (1.6-fold) were identified as VMP-enriched according to LightCycler analysis. Fgf10 was included as an internal control, a low abundance VMP-specific gene, known to have an important role in
prostate organogenesis (Thomson and Cunha, 1999). Therefore, identification of Fgf10 as VMP-enriched gene validated the LightCycler as a reliable technique to quantify even low abundance transcripts.

![Graph 4.1: SAGE Tags Verified as VMP-enriched by LightCycler Analysis](image)

**Tag count**

<table>
<thead>
<tr>
<th></th>
<th>VSU</th>
<th>VMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta-like 1</td>
<td>17</td>
<td>41</td>
</tr>
<tr>
<td>Notch 2</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Nel-like 2</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>MMP 14</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>FGF10</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Plectrophin</td>
<td>63</td>
<td>74</td>
</tr>
</tbody>
</table>

Graph 4.1: SAGE Tags Verified as VMP-enriched by LightCycler Analysis

A graph of the SAGE and LightCycler analysis data, showing the fold-difference of transcript expression in the VMP in comparison to the VSU. LightCycler analysis classified a difference in transcript expression of ≥1.5-fold in the VMP, as VMP-enriched.

**4.2.2.2 LightCycler Analysis Identifies VSU-enriched Genes.**

Of the remaining genes tested, some were found to be VSU-upregulated (>1.5 fold) in comparison to the VMP (9 of 26) (Table 4.1). These transcripts were likely to be specific to the epithelial or smooth muscle cells, present only in the VSU tissue.

‘Control’ genes, Notch1 (Shou et al., 2001, Wang et al., 2004), Hepsin (Dhansekaran et al., 2001, Ernst et al., 2002, Klezovitch et al., 2004) and FgfR2 (Igarashi et al., 1998) were included in the analysis. They have each been previously localised to the
epithelial cells of the developing and/or disease prostate. Additionally, Jagged2 was included as a ligand of Notch1 (Valsecchi et al., 1997). Verifying their predicted gene expression as VSU-enriched helped to confirm the robustness of the LightCycler screening process.

Table 4.1: Genes identified as VSU-enriched by LightCycler analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>LongSAGE</th>
<th>VMP tags</th>
<th>VSU tags</th>
<th>VSU enriched genes (P&lt;0.05)</th>
<th>VSU enriched gene (≥1.5 fold-difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepsin</td>
<td>27</td>
<td>43</td>
<td>No</td>
<td></td>
<td>2.7</td>
</tr>
<tr>
<td>FGFR2</td>
<td>2</td>
<td>6</td>
<td>No</td>
<td></td>
<td>2.2</td>
</tr>
<tr>
<td>TGFβ2</td>
<td>6</td>
<td>1</td>
<td>No</td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>Notch 1</td>
<td>4</td>
<td>0</td>
<td>No</td>
<td></td>
<td>1.9</td>
</tr>
<tr>
<td>GDNFRα1</td>
<td>3</td>
<td>4</td>
<td>No</td>
<td></td>
<td>1.9</td>
</tr>
<tr>
<td>BMP 7</td>
<td>3</td>
<td>0</td>
<td>No</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Jagged 2</td>
<td>0</td>
<td>0</td>
<td>No</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Hgf</td>
<td>0</td>
<td>0</td>
<td>No</td>
<td></td>
<td>1.6</td>
</tr>
<tr>
<td>Pdgf</td>
<td>0</td>
<td>0</td>
<td>No</td>
<td></td>
<td>2.5</td>
</tr>
</tbody>
</table>

4.2.2.3 LightCycler Analysis Identifies No Difference in Transcript Expression

The remaining genes investigated (11 of 26) showed no difference in transcript expression levels in the VSU and VMP (Table 4.2). Several of these genes had been selected from the list of VMP-enriched genes identified by SAGE analysis (P<0.05)(Appendix), and several of them had also been identified as cancer-associated genes (Lapointe et al., 2004, Rhodes et al., 2002, Stuart et al., 2004). In particular, MMP2 is highlighted within this group as a VMP-enriched gene (P<0.05), identified by SAGE and is included in further analyses (Chapter 8).
### Table 4.2: Genes that show no difference in expression between the VSU and VMP by LightCycler analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>VMP tags</th>
<th>VSU tags</th>
<th>VMP-enriched genes by SAGE tag counts (P&lt;0.05)</th>
<th>Identified as a cancer-associated by literature mining.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decorin</td>
<td>384</td>
<td>270</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>MMP2</td>
<td>69</td>
<td>33</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fibronectin</td>
<td>21</td>
<td>10</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Basigin</td>
<td>18</td>
<td>6</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>PMP22</td>
<td>16</td>
<td>5</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Paralemmin</td>
<td>12</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Syntenin</td>
<td>11</td>
<td>2</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Midkine</td>
<td>10</td>
<td>11</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Lgl</td>
<td>5</td>
<td>2</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>VAP 1</td>
<td>5</td>
<td>0</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>BMP3</td>
<td>0</td>
<td>0</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

#### 4.2.2.4 Selection of Candidate Genes

Six genes were selected after LongSAGE and LightCycler analysis as "candidate mesenchymal genes"; Dlk1, Ptn, Notch2, Nell2, MMP2 and MMP14. The associated transcripts had each been identified by SAGE and/or LightCycler analysis as VMP-enriched (Appendix and Graph 4.1). Additionally, in support of the choice of candidate selection, the VSU and VMP libraries were interrogated for candidate-associated genes (Table 4.3), to determine if other member of their developmental and/or signalling pathways were present.

Characterisation of the candidates will be discussed in the following chapters of this thesis. In particular, Dlk1 (Chapter 5), Notch2 (Chapter 6) and Ptn (Chapter 7), were studied in most detail.
Table 4.3:

<table>
<thead>
<tr>
<th>Candidate Gene</th>
<th>Associated Genes in the VSU identified by SAGE analysis and/ or RT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notch 2</td>
<td>Notch1, radical and lunatic fringe, Jagged 2, numb, nicastrin</td>
</tr>
<tr>
<td>Dlk1</td>
<td>Notch associated genes?</td>
</tr>
<tr>
<td>Pleiotrophin</td>
<td>midkine, syndecan1, 3 and 4, syntenin and nucleolin.</td>
</tr>
<tr>
<td>MMP2 and MMP14</td>
<td>MMP11, 16, 23, TIMP 1 and 2, ADAM 15 and 17, ADAMTS-1 and 12.</td>
</tr>
<tr>
<td>Nell2</td>
<td>Unknown signalling pathway</td>
</tr>
</tbody>
</table>

4.2.3 Summary of Results

LightCycler analysis (qRT-PCR) was used as a screening process to verify the VMP-enriched expression of 26 selected genes. Some genes were selected either by bioinformatic analysis of the i.e VMP-enriched genes (P<0.05) (Appendix) (10), or by intuitive selection, i.e those from established developmental pathways and/or prostate cancer studies (16).

LightCycler identified six genes as confirmed as VMP-enriched (≥1.5 fold) in comparison to the VSU; Dlk1, Notch2, Ptn, Nell2, MMP14 and Fgf10 (a known mesenchyme-specific factor). Dlk1, Notch2 and Nell2 had been selected by bioinformatic analysis of the SAGE data, and had been identified as VMP-enriched (P<0.05). However, these were only 3 out of 10 (30%) VMP-enriched genes (P<0.05), determined by SAGE, to be confirmed as VMP-enriched (>1.5-fold) by LightCycler analysis. The other 7 VMP-enriched genes (P<0.05) showed no difference in expression between the VMP and VSU tissues, by LightCycler analysis (Tables 4.1 and 4.2) although they did exhibit a trend towards VMP-enriched expression according to their tag counts. This group included MMP2, a gene commonly expressed with MMP14.

The intuitively selected genes, MMP14 and Ptn exhibited noticeable but not statistically significant VMP-enriched expression by SAGE analysis, however, subsequent LightCycler analysis confirmed that they demonstrated VMP-enriched
expression (≥1.5-fold). Finally, 9 genes were confirmed as VSU-enriched (≥1.5 fold) by LightCycler analysis and were therefore, likely to contain factors from the epithelium and smooth muscle. This group of genes included Notch1, Hepsin and Fgfr2, each previously identified in prostate epithelia.

Based on the SAGE and LightCycler data, six candidates were selected for further characterisation studies: Dlk1, Notch2, Ptn, Nell2, MMP2 and MMP14. To support the selected candidates the SAGE Libraries were interrogated for associated genes and for other members of their developmental pathways.
5. Delta-Like 1 homolog

5.1 Introduction

The mammalian protein Delta-like 1 (Dlk1) is a member of the epidermal growth factor (EGF)-like family of homeotic genes that includes Notch receptors and their ligands (Laborda et al., 1993, Laborda, 2000). Notch ligands are transmembrane proteins and belong to the DSL (Delta, Serrate, Lag-2) family of proteins. The ligands are characterized by two conserved motifs: the DSL domain, important for Notch binding and a series of EGF-like repeats. They are grouped into two subfamilies based on the presence of a cysteine-rich region in the extracellular region: those with the cysteine-rich region belong to the Serrate/Jagged family, whereas those without belong to the Delta family. The delta-like 1 protein was so named because of its similarity in structure and amino acid sequence in the EGF-like repeats to Delta proteins (see Fig 5.1). Four Delta proteins (Delta1, Delta2, Delta 3, Delta4) have been identified as Notch ligands in vertebrates, to date (Gray et al., 1999, Joutel and Tournier-Lasserve, 1998, Shutter, et al., 2000).

Fig 5.1: Structure of the Dlk1 and Delta Proteins
Comparison of the Dlk1 (Panel A) and Delta (Panel B) protein structures. Dlk1 lacks the DSL (Delta, Serrate, Lag-2) domain. Signal peptide (SP), epidermal growth factor repeats (EGF), protease target region (PRO) transmembrane domain (TM), intracellular region (INT) (Baladron et al, 2005)
The Dlk1 protein comprises six EGF-like repeats, a signal sequence in the extracellular domain, a unique transmembrane domain, and a short intracellular region (Laborda et al., 1993). The EGF-like repeats, are 35- to 40-amino acid domains with conserved spacing of six cysteine residues, first identified in epidermal growth factor. However, there is one important difference between Notch ligands and dlk, in the lack of a DSL domain at the dlk N-terminus. The cysteine-rich DSL motif is thought to be important for binding to the Notch receptor in the extracellular region (Artavanis-Tsakonas et al., 1995) therefore dlk has been reported to be an unlikely ligand for Notch. However, a controversy still remains as to a possible interaction between Dlk and Notch1. Recently Baladron et al., 2005 reported that despite lacking a DSL domain, dlk does interacts with Notch1, competing with its ligands, and inhibiting Notch 1 activation.

Dlk1 (Pref-1) was first cloned in a cDNA library screen for differentially expressed during adipocyte differentiation using the mouse preadipocyte cell line 3T3-L1(Smas and Sul, 1993). A clone for the rat homolog was more recently isolated (Carlsson et al., 1997, Halder et al., 1998). Extensive sequence similarity has been reported between reported between rat and mouse and rat and human FA1 (Carlsson et al., 2000). Dlk orthologues have been identified in human, rat, bovine as well as mouse independently and so many names have been given to the same molecule; pG2 (pheochromocytoma-specific protein)(a human adrenal specific mRNA), Pref-1 (Preadipocyte factor 1) (Fahrenkurg et al., 1999, Smas and Sul, 1993), SCP-1 (stromal cell-derived protein 1) (GenBank/D16847), ZOG, (zona glomerulosa specific mRNA)(Halder et al., 1998) as well as FA1 (foetal antigen 1), (Jensen et al., 1994).

The Dlk protein, not only has several names, but different forms of the protein have also identified. The Dlk mRNA undergoes differential splicing which encodes for two main species of polymorphic proteins with or without, an extracellular amino acid target sequence which is a target for an unknown protease capable of releasing the EGF-like region to the external medium. The alternate mRNA splicing and protease activity are thought to modulate the ratio of expression of the two variants; the soluble and membrane-bound forms. Several spliced mRNA variants have been reported (Lee et al., 1995, Smas et al., 1994, Smas et al. 1997). Four major forms of the transcript, Pref-1A-D were identified during the initial study in the mouse 3T3-L1
cells (Smas and Sul, 1993) and more recently in the mouse colon (Dong et al., 2004). Of the four splicing products only the -1A and -1B forms, released from the extracellular dlk region to the medium as soluble proteins, were found to be biologically active (Mei et al., 2002). The soluble, Pref-1A was found to be the most abundant form. Similar to the findings in the mouse, multiple cleaved soluble proteins (48-60 kDa) have reported in the rat adrenal gland leaving a residual cell-associated protein (25kDa), after cleavage of the extracellular domain (Halder et al., 1998, Smas et al., 1997). The extracellular soluble part of the dlk protein often referred to as foetal antigen 1(FA1), was originally isolated from second trimester normal human amniotic fluid, and is highly expressed during pregnancy in maternal and foetal circulation in the urine and amniotic fluid (Fay et al., 1988, Jensen et al., 1994).

The Dlk/FA1 protein is widely expressed in embryonic tissues, including liver, tongue, vertebrae, skeletal myotubes, chodroblasts and pancreas (Floridon et al., 2000, Smas and Sul, 1993, Tornehave et al., 1989, Tornehave et al., 1996), where its role remains unknown. However, in the adult Dlk/FA1 expression is confined to zona glomerulosa of the adrenal gland (Halder et al., 1998, Jensen et al., 1994), the somatotrophs in the pituitary gland (Larsen et al., 1996), the monoaminergic neurons in the central nervous system (Jensen et al., 2001), the Leydig cells, and theca interna and hilus cells of the testes and ovaries (Jensen et al., 1999), and the β cells in the endocrine pancreas (Tornehave et al., 1996). Dlk1/FA1 has also been demonstrated in endocrine and neuroendocrine tumours (Jensen et al., 1994, Tornehave et al., 1996) and increased serum concentrations have been demonstrated in patients with small cell lung carcinoma (Jensen et al, 1999). Therefore, the distribution in embryonic and adult tissues suggest that Dlk1/FA1 may be involved in the development and maintenance of endocrine tissues.

More generally, Dlk is reported as a member of the EGF-like family of proteins that mediate protein-protein interactions through their EGF-like repeats and mediate cell fate and differentiation decisions during development in many organisms (Artavanis-Tsakonas, et al., 1999, Artavanis-Tsakonas, et al., 1999, Laborda et al., 2000). Dlk1 is reported to be critical in modulating extracellular signals leading to either cell differentiation or proliferation (Baladron et al., 2002, Garces et al., 1999, Ruiz-Hidalgo, et al., 2002). Originally identified during a differential screening designed to
isolate genes that regulate adipogenesis (Smas et al., 1993), Dlk1/FA1 is known to participate in several differentiation processes, including adipogenesis (Smas et al., 1997, Garces et al., 1999, Bauer et al., 1998, Boney et al., 1996), haematopoiesis (Bauer et al., 1998, Kaneta et al., 2000), Miyazoto et al., 2001, Moore et al., 1997, Ohno et al., 2001), and adrenal gland and neuroendocrine cell differentiation (Laborda et al., 1993, Halder et al., 1998, Okamoto et al., 1998, Raza et al., 1998, Van Limpt et al., 2003). Dlk1 is also reported as involved in peripheral and central nervous system differentiation (Costaglioli et al., 2001, Jensen et al., 2001), growth arrest and increased malignance of undifferentiated tumours (Baladron et al., 2001, Laborda et al., 1993, Van Limpt et al., 2000) and in wound repair (Samulewicz et al., 2002). In addition mice lacking Dlk display pre- and postnatal growth retardation and accelerated adiposity, whereas transgenic mice expressing the complete ectodomain of Dlk show decreased adiposity and glucose tolerance. Dlk1 null mice also displayed eyelid and skeletal abnormalities smaller kidneys and lungs plus pulmonary defects, and a dramatic increase in perinatal mortality. This study therefore concluded that Dlk1 was important for perinatal survival, normal growth and fat deposition. In adulthood the deletion of Dlk1 led to accelerated adiposity (Moon et al., 2002).

The role of Dlk remains elusive, but it is generally reported as a likely de-differentiation factor, maintaining cells in an undifferentiated state and/or localised to cells that are capable of regeneration. For example, in adipogenesis, on differentiation of preadipocytes the expression of Dlk is downregulated, and cannot by detected in mature adipocytes (Smas et al., 1993). Also, Dlk is expressed in foetal liver stromal cells that are able to support haematopoiesis (Moore et al., 1997). Further, Dlk modulates proliferation of thymocytes (Kaneta et al., 2000), foetal liver haematopoietic cells (Ohno et al., 2001) and pre-B cells in the pancreas (Bauer et al., 1998). It is not fully understood which forms of Dlk are regulating these processes, however, it has been reported that only the secreted Dlk variants inhibit adipocyte differentiation but membrane-bound variants do not; in fact they appear to play a positive effect in differentiation (Garces et al., 1999, Mei et al., 2002). The mechanism of Dlk regulation as a differentiation factor is unknown. There is an hypothesis that up-regulation of Dlk may lead to suppression of certain differentiated functions that may facilitate responsiveness to mitogenic stimuli in adjacent cells thereby maintaining a population of stem cell/progenitor cells. Such a mechanism
was suggested for haematopoietic stem cells, in which stromal cells expressing Dlk, were found to support self-renewal of the stem cells (Moore et al., 1997).

_Dlk_ gene expression and protein levels are subject to genomic imprinting, via a differential methylation mechanism, and regulation of this process is important for normal development. _Dlk1_ is a paternally expressed imprinted gene on human chromosome 14q32. Studies on the human chromosome 14 deletions and maternal uniparental disomy (mUPD)14 have suggested that expression of unknown imprinted genes on chromosome 14 cause congenital disorders. Further studies on _dlk1_ null mice, where Dlk is located on chromosome 12, found similar effects, and therefore it was suggested that _dlk_ may be responsible for most of the imprinting related phenotypes of mUPD12 and syntenic human mUPD14 (Moon et al., 2002).

The signal transduction pathway mediated by Dlk1 is not well understood. It has been hypothesized that _dlk1_ exerts its inhibitory effect on cell differentiation through an unidentified receptor via its EGF-like domains, either through a paracrine or juxtacrine mechanism. The presence of the full form Dlk1 molecule, or the cleaved smaller forms, would be as the result of alternate splicing, and this would determine its mode of action (Laborda, 2000). However, there is evidence that Dlk1 participates in cell-to-cell interactions capable of modulating the ERK/MAPK pathway and Notch 1 has been proposed as a good candidate for a _dlk_ receptor (Baladron et al., 2005, Ruiz-Hidalgo et al., 2002). In addition, a potential cross-talking between the Ras-ERK/MAPK signalling pathways has already been reported (Ruiz-Hidalgo et al., 1999, Weijzen et al. 2002) and that there is a functional relationship between Dlk1 and HES-1, a transcription factor whose expression is activated by Notch1, although this relationship is controversial (Baladron et al., 2005, Kaneta et al., 2000, Ohno et al., 2001). Therefore, if Dlk1 is a ligand for Notch1, it could be involved in Notch signalling in many different systems, including, but not limited to adipogenesis, haematopoiesis, nervous system differentiation and cancer. This is only a suggestion at present, and I chose to study the Dlk1 molecule independently of its possible relationship with Notch1.
5.2 Results

5.2.1 Delta-Like 1 Transcript Expression in the VMP and VSU

Bioinformatic analysis of the SAGE libraries (P<0.05) (Appendix) and LightCycler analysis (Graph 4.1) identified Dlk1 expression was VMP-enriched. Additionally, Northern blot analysis was also used to measure the relative levels of Dlk1 transcript expression in the VMP and VSU (Fig 5.2). To verify the correct RNA loading, the Northern blot was also hybridised with Fgf10 (Fig 8.3), a known VMP-specific gene (Thomson and Cunha, 1999). The radioactively labelled probes were prepared using primers in Table 2.4. A Northern blot comprising neonatal VP, VMP and VSU RNA determined that Dlk1 transcript abundance was ~2-fold higher in the VMP, in comparison to the VSU, and also ~1.5-fold higher in the VMP than in the VP. Therefore, Dlk1 was identified as a VMP-enriched gene by three independent techniques.

![Figure 5.2](image_url)  
**Fig 5.2** Dlk1 mRNA expression in the male and female rat UGS. Analysis of Dlk1 transcript levels in P0 female and male prostate organ rudiments (n=2). The VSU is the whole female prostate rudiment, the VMP is the microdissected mesenchymal pad, and the VP is the microdissected male ventral prostate lobe. Transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represent the average transcript abundance of Dlk1 transcript relative to the VP.

5.2.2 Dlk1 transcript levels in the prostate and selected organs

Further experiments were performed to investigate the expression of Dlk1 transcript during embryonic and postnatal prostate development and also expression in selected neonatal organs (Fig 5.3). Dlk1 transcript expression was found to be most abundant...
on the day of birth (PO) (100%), upregulated from embryonic life (e17.5) (23%) and then downregulated after PO, most dramatically between P10 (61%) and P28 (3%), and remaining low in adult life. Also, no difference in transcript abundance was found between the male (E17.5m) and female embryonic UGS (E17.5f) (Fig 5.3, Panel A). In the selection of PO organs, Dlk1 expression was most abundant in the VP (100%), liver (173%), ovary (147%) and testes (117%). Dlk1 levels by comparison were very low or negligible in the other organs tested (2-11%)(Fig 5.3, Panel B).

5.2.3 Dlk1 Protein Localisation in the UGS
Localisation of Dlk1 protein in the neonatal rat male and female UGS was investigated by immunocytochemistry. Two independent antibodies (see Table 2.1) were used to verify the staining pattern found and P3 adrenal gland was included as the control tissue. Dlk1 was localised to the medulla and zona glomerulosa of the adrenal gland, where downregulation of Dlk1 expression is thought to be an important step in adrenal zone differentiation (Hadler et al., 1998). In addition, Dlk1 protein expression was found in the kidney (PO), a branched organ, with similarities in its developmental processes to the prostate, and Dlk1 protein was localised to the renal tubules (data not shown). This is consistent with a report of DLK1 expression in the developing tubules of human foetal kidneys (8-21wks)(Fukuzawa et al., 2005).

The Dlk1 antibody supplied by Santa Cruz gave the lowest background staining, and was therefore used on the sections displayed in Figures 5.4-5.7. In the male and female PO UGS, Dlk1 protein expression was found in the mesenchyme of the VMP and VP and also in the smooth muscle (SM) layer surrounding the urethra. The urethral epithelium was mostly negative for Dlk1, and occasional weak staining was attributed as background. Additionally, in the female UGS, it was noted that the Dlk1 staining in the SM layer was not uniform. Staining for Dlk1 was weak or absent in the SM layer that lies adjacent to the VMP, but positive in other areas of SM. This was interesting as the SM layer is in general considered to be homogeneous, but here displayed differential distribution of Dlk1, to a subset cells in the SM layer.
Fig 5.3: Dlk1 mRNA expression during prostatic ontogeny and expression in other organs (P0).
Analysis of Dlk1 transcript levels in the prostate; E17.5 until adulthood (Panel A), and in branching and non-branching organs collected from P0 rats (Panel B) (n=2). Ventral prostate (VP), bladder (Bld), brain (Brn), gut (Gt), heart (Ht), kidney (Kid), liver (Lvr), ovary (Ov), testes (Test), lung (Lg). The transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represents the average Dlk1 transcript abundance relative to the P0 VP.
To study this interesting pattern of Dlk1 further, fluorescent immunocytochemistry was used to co-localise Dlk1 and smooth muscle α-actin, a marker of smooth muscle cells in the male and female neonatal UGS. In the female UGS, Dlk1 expression was confirmed in the SM layer, except in the area adjacent to the VMP, where only staining for SM α-actin was observed (Fig 5.5). Additionally, the Dlk1 staining where present in the SM layer, was most strongly expressed towards its outer edge. The SM cells most adjacent to the UGM had low or absent Dlk1 expression. Some SM cells appeared to co-express Dlk1 and SM α-actin, while some cells expressed the proteins individually. In the male P0 UGS, Dlk1 protein expression was localised to the mesenchymal cells of the VP and the smooth muscle layer (Fig 5.6). Similar to staining in the SM layer of the female UGS, Dlk1 expression was most evident in cells towards the outside of the SM layer. A narrow band of the SM layer that lies most adjacent to the UGM was absent of Dlk1 staining.

Finally, in the sections of P0 male UGS stained by fluorescent immunocytochemistry, it was noted that strong staining for Dlk1 protein was observed in the epithelial duct cells of the VP. This was considered background staining due to the high concentration of Dlk1 antibody used, and no similar staining had been observed previously in the male UGS by general immunocytochemistry (Fig 5.4).

5.2.4 The effect of testosterone on Dlk1 distribution

Androgens may directly or indirectly induce the release of regulatory factors from the mesenchyme during development. Consequently, it was investigated whether the distribution of Dlk1, as a mesenchyme-expressed factor, in the VP, would be affected by exposure to Testosterone. VP organ rudiments were cultured in the presence (+T) or absence of testosterone (-T), and stained for Dlk1 protein (Fig 5.7). The VPs cultured with testosterone were found to grow larger in size and show more epithelial bud branching than VPs grown without. Dlk1 protein was found similarly expressed in the mesenchymal cells of the VPs under both culture conditions, and therefore Dlk1 distribution was not affected by testosterone treatment in vitro. Preliminary studies were also performed to investigate changes in the Dlk1 transcript levels in treated and untreated organs by Northern blot analysis (data not shown, personal communication, Dr. G. Vanpoucke), but no differences were found.
Fig 5.4: Dlk 1 expression in Neonatal Female and Male Rat UGS
Immunohistochemical localisation of the Dlk1 protein, in P0 female (Panel A and B) and male (Panels C-E) rat UGT. Ventral mesenchymal pad (VMP), smooth muscle (SM), urethral epithelium (URE), mesenchyme (M), Epithelium (E). P3 Adrenal gland (Panel F) is the positive control tissue. Zona glomerulosa (ZG), Medulla (Md). In panel A arrows point to areas of high Dlk1 expression and * marks an area of smooth muscle that has very low or absent Dlk1 expression. Preabsorbed primary antibody were added to Panels B, E and F (insert). The scale bar in panel A represents 200μm in B, C, E, and F, and 100 μm in panel D.
Fig 5.5 Dlk1 and smooth muscle alpha actin expression in the female UGS. Immunohistochemical co-localisation of Dlk1 and SM $\alpha$-actin in P0 female rat UGS (Panels A-D) (n=3). Dlk1 is shown in red, SM $\alpha$-actin in green, and nuclear staining in blue. Co-expression of Dlk1 and SM $\alpha$-actin is shown in yellow. The plane of tissue section in Panel A does not include urogenital epithelium. Ventral mesenchymal pad (VMP), smooth muscle (SM), urogenital mesenchyme (URM). Panel B is a higher power magnification of the SM region marked * in Panel A, and Panel C is a higher power magnification of the SM layer adjacent to the VMP. In Panel B, note that Dlk1 staining in the SM layer is not uniform, and is strongest towards the outer edge of the tract. Panel C shows Dlk1 staining is absent in the SM layer adjacent to the VMP. Primary antibodies were not added to Panel D. The scale bar in Panel A represents 100 $\mu$m. The scale bar in panel B represents 50 $\mu$m in panels C and D.
**Fig 5.6 Dlk1 and smooth muscle alpha actin expression in the male UGS.**

Immunohistochemical co-localisation of dlk1 and SM α-actin protein in P0 male rat UGS, including the VP lobe (Panels A-D) (n=3). dlk1 is shown in red, SM α-actin in green, and nuclear staining in blue. Co-expression of Dlk1 and SM α-actin is shown in yellow. mesenchyme (M), epithelium (E), smooth muscle (SM), urogenital mesenchyme (UGM)). In panel A, * marks an area of the smooth muscle layer that lacks Dlk1 staining. Also, note that strong Dlk1 staining in the epithelial cells of the ducts is background, due to the high antibody concentration used. Primary antibodies were not added to Panel D. The scale bar in Panel A represents 100μm and 50μm in panel C. The scale bar in Panel B represents 100μm in Panels B and D.
Fig 5.7 The effect of testosterone on Dlk 1 expression.
Immunohistochemical localisation Dlk-1 protein in P0 VP lobes cultured for 6 days in the presence (+T) (Panels C, D and F) or absence of testosterone (-T)(10^{-8} M)(Panels A, B and E). Preabsorbed primary antibody was added to Panels E and F. Epithelium (E), mesenchyme (M) and urethral ends of ducts (U). The scale bars in Panel A represent 200 \mu m in Panels C. and in Panel B the scale bar represents 100 \mu m in Panel D-F.
5.2.5 Functional Studies on Dlk1

*Dlk1* transcript and protein expression in the mesenchyme of the male and female UGS (P0) has been confirmed, in this thesis. Next, a possible functional role for Dlk1 as a mesenchymal regulator of prostate organogenesis was investigated.

VPs were cultured in the presence (+T) or absence of testosterone (-T) (10^8M), plus recombinant human DLK1 protein 0.1–2.5μg/ml (Table 2.2). DLK1 treatment was found to increase the size of VP organs and amount of epithelial budding and branching observed, in a concentration-dependent manner (-T plus 0.1-0.75 μg/ml DLK1) (Fig 5.8). No increase in the effect was observed when organs were treated with the highest concentration of DLK1 (-T plus 2.5 μg/ml DLK1) (data not shown). The greatest increase in VP growth and epithelial branching observed was induced by +T treatment only. No incremental effect was observed when organs were treated +T plus 0.1-0.75 μg/ml DLK1 (data not shown). Unfortunately, due to insufficient time, it was not possible to make measurements of the effect of DLK1 treatment on VP growth, and so the conclusions are based on observed changes in gross morphology of the VPs and the evident increase in number of duct tips after DLK1 treatment, indicating stimulation of branching morphogenesis.
Fig 5.8: The effect of Dlk1 on VP growth.
VP's from P0 rats were grown for six days in the presence (+T) or absence of testosterone (-T)(10^{-8}M) and/or recombinant human DLK1 (0.1, 0.25, and 0.75 μg/ml). The scale bar represents 500μm.
5.2.6 Summary of Results

*Dlk1* was identified as a VMP-enriched transcript by LongSAGE, LightCycler and Northern blot analysis. During development and growth of the prostate *Dlk1* transcript expression, was most abundant in the P0 VP and during the perinatal period, when epithelial growth and branching occurs. In other selected neonatal organs, *Dlk1* expression was most abundant in the liver, ovary, testes and VP. The Dlk1 protein was localised to the mesenchyme of the VMP and VP and also the smooth muscle layer surrounding the UGS. Notably, Dlk1 expression was not uniform in the SM layer, and was most strongly expressed towards the outside of the layer. Also, in the female UGS, Dlk1 expression was low or absent in an area of SM adjacent to the VMP. The SM layer was previously thought to be homogenous, but here it exhibits differential protein distribution. Additionally, Dlk1 was localised to the mesenchyme of VP organs cultured *in vitro* in the presence or absence of Testosterone, and the protein distribution was not affected by treatment with Testosterone. Finally, VP organ cultures treated with recombinant DLK1 were found to show an increase in overall size and epithelial branching. No additional effect was found when Testosterone was also added.
6. Notch 2

6.1 Introduction

Notch2 is a member of the evolutionary conserved Notch family of genes that help to determine cell fate during development. Notch was originally reported in Drosophila, which has only one copy of the Notch gene, while mammals possess at least four such genes, Notch1, Notch2, Notch3, and Notch4 (Lardelli et al., 1994, Reaume et al., 1992, Robbins et al., 1992, Weinmaster et al., 1992). Genes of the Notch family encode large cell-surface receptors with several conserved domains, which interact with membrane bound ligands encoded by the Delta, Serrate and Jagged family genes. Six proteins have been recognized as Notch ligands in vertebrates (Delta1, Delta2, Delta3, Delta4, Jagged1 and Jagged2), to date (Gray et al., 1999, Joutel and Tournier-Lasserve, 1998, Lindsell et al., 1995, Shutter, et al., 2000, Valsecchi et al., 1997). Proteins of the Notch family share several repeated peptide motifs, including extracellular tandemly repeated copies of an epidermal growth factor (EGF-) like motif, involved in ligand binding, repeats of the Lin-12/Notch motif (LNR) with unknown function in the extracellular domain and CDC10/ankyrin-like repeats in the cytoplasmic domain (Hamada et al., 1999).

In vertebrates, Notch2 was originally cloned in the rat brain (Weinmaster et al., 1992) and since then expression of Notch2 transcript and/or protein has been reported in diverse tissues particularly during development with more restricted expression during adult life, and has also been associated with tumorigenesis.
Notch2 expression has been reported in the developing brain (Tanaka and Marunouchi, 2003), eye and heart (McCright et al., 2001), liver (Tanimizu and Miyajima, 2004), lung (Ito et al., 2000), kidney (Leimeister et al., 2003, McCright et al., 2001), pancreas (Lammert et al., 2000, Norgaard et al., 2003), small intestine (Schröder and Gossler, 2002) and tooth (Mitsiadis et al., 2003). Kitamoto et al., (2005) has reported a comprehensive study of Notch1, Notch2 and Notch3 transcript expression in the foetal mouse. Reports of Notch expression in the developing prostate have so far been confined to Notch1, which is expressed in the basal epithelial cells of prostatic ducts (Wang et al., 2004). Whereas Notch2 expression has been reported elsewhere in the male reproductive tract, in the developing rat Wolffian duct (Bowman et al., 2005) and epididymis (Turner et al., 2003). In the adult, Notch2 expression is reported in developing follicles of the ovary (Johnson et al., 2001), in the brain (Tanaka and Marunouchi, 2003) and in adult wounded teeth (Mitsiadis et al., 2003). Notch2 is also reported to be involved in embryonal brain tumour growth (Fan et al., 2004) and breast (Parr et al., 2004) and lung cancer (Garnis et al., 2005)

As already mentioned, the only member of the Notch family studied in the prostate, so far has been Notch1. Notch1 expression was reported to be essential for prostatic branching morphogenesis, and also for re-growth of the prostate following castration and androgen replacement. Notch1 expression in the basal epithelial cells may define the progenitor cells in the epithelial cell lineage, which are essential and indispensable for prostate development and re-growth (Wang et al., 2004). Additionally, Notch1 was found upregulated in malignant prostatic epithelial cells of TRAMP mice (transgenic adenocarcinoma of the mouse prostate) (Shou et al., 2001), and dysregulation of Jagged1, a Notch ligand has been associated with prostate cancer progression and metastasis (Satagata et al., 2004). Therefore the Notch signalling pathway is well represented during normal and neoplastic prostate growth.

Notch2, on the other hand, has been reported in the developing epididymis (Turner et al., 2003) and Wolffian duct (Bowman et al., 2005), where it is suggested to be
involved in the mesenchymal-epithelial interactions that regulate development. These studies also reported that \textit{Notch2} transcript expression levels in these tissues were altered by anti-androgen treatment \textit{in utero}, and that in the case of the epididymis at least, this was not due to changes in the mesenchyme/epithelial cell ratio. Therefore, \textit{Notch2} may represent an androgen-regulated molecule.

In general, \textit{Notch} genes have been reported as essential for embryonic development and adult homeostasis. During development \textit{Notch} signalling is thought to be involved in regulation of tissue patterning and morphogenesis via direct cell-cell interactions. Expression of mammalian \textit{Notch} genes were reported to show complementary and combinatorial patterns during development, suggesting distinct functions for different members of the Notch family, with the possibility of some functional redundancy (Kitamoto et al., 2005). For example, Notch1 and -2 have been reported as differentially regulated by modulated interaction with their ligands, caused by signalling factors upstream in the Notch pathway. Therefore, the differential expression pattern of Notch receptors in the embryo, combined with differential modulation of their ligands may act in concert to control tissue patterning and morphogenesis during development (Hicks et al., 2000). Additionally, \textit{Notch} expression has been found throughout development in uncommitted proliferative cells (Artavanis-Tsakonas et al., 1995, Lardelli et al., 1995, Reaume et al., 1992, Williams et al., 1995), and later in development and in adult life, expression of Notch continues in the proliferative layer of mature tissues (Kopan and Weintraub, 1993, Milner et al., 1994, Zagouras et al., 1995). Therefore, these patterns of expression might suggest that Notch proteins function in maintenance of the proliferative capacity of immature cells.

Analysis of both inherited human diseases and mutant mice has shown that mutations in Notch receptors result in developmental abnormalities and neoplasias (Gridley, 1997, Joutel and Tournier-Lasserve, 1998). Notch malfunction can disrupt aspects of neurogenesis, somite formation, angiogenesis and kidney and lymphoid development (Anderson et al., 2001, Conlon et al., 1995, de la Pompa et al., 1997, Lawson et al., 2001, McCright et al., 2001). For example, studies of the \textit{Notch2} null mouse found
that Notch2 expression is necessary for perinatal mortality because of its role in the developing kidney (Hamada et al., 1999, McCright et al., 2001). Also, in humans, constitutively active aberrant forms of Notch have been associated with leukaemia and solid tumours of the cervix, colon and pancreas (Allenspach et al., 2001, Capobianco et al., 1997, Ellisen et al., 1991, Gray et al., 1999, Nickoloff et al., 2003, Radtke and Raj, 2003, Zagouras et al., 1995). However, the role of Notch in cancer is further complicated by reports that Notch1 and 2 have opposing roles in tumorigenesis, and this might parallel their functions in normal development (Fan et al., 2004). In breast cancer for example Notch2 has been proposed to have a tumour-suppressive role, whereas Notch1 possesses tumour-promoting functions (Parr et al., 2004).

Notch signalling is an evolutionary conserved mechanism that is known to be involved in a variety of cellular processes, including the maintenance of stem cells, specification of cell fate, differentiation, proliferation and apoptosis at all stages of development, thus controlling organ formation and organogenesis (Artavanis-Tsakonas et al., 1999, Frisen and Lendahl, 2001). As both Notch and its ligands are transmembrane proteins, signalling only occurs between closely apposed cells. Activation of Notch by ligand binding triggers cleavage of the receptor, releasing the intracellular domain, which then translocates to the nucleus (Struhl and Adachi, 1998). In the nucleus, Notch forms transcriptional complexes with transcription factors of the CSL family in mammals (Greenwald, 1998), and transactivates genes targets such as those in the Hes and Hey families (Iso et al., 2003).

Additionally, in a study of Notch activity during development, it was proposed that communication between the fibroblast growth factor receptor (FGFR) pathway and Notch pathway might represent an important reciprocal autoregulatory mechanism for the regulation of normal cell growth (Norgaard et al, 2003, Small et al., 2003). A possible Notch/Fgf pathway interaction is interesting with relevance to the developing prostate, as members of the FGF family are known to be important regulatory factors in prostate growth and development (Thomson, 2001).
In conclusion, the importance of the Notch pathway in development and cancer, and in particular the proposed role of Notch1 in the prostate made Notch2 an excellent candidate gene to study in more detail during prostate development.

6.2 Results

6.2.1 Notch2 Transcript Expression in the VSU and VMP

Bioinformatic analysis of the SAGE libraries (Appendix) and LightCycler analysis (Graph 4.1) identified Notch2 expression was VMP-enriched. Additionally, Northern blot analysis was used to measure the relative levels of Notch2 transcript expression in the VMP and VSU (Fig 6.2). To verify the correct RNA loading, the Northern blot was also hybridised with Fgf10 (Fig 8.3), a known VMP-specific gene (Thomson and Cunha, 1999). The radioactively labelled probes were prepared using primers in Table 2.4. A Northern blot comprising neonatal VP, VMP and VSU RNA demonstrated that Notch2 transcript abundance was ~2-fold higher in the VMP, in comparison to the VSU, and expression levels were similar in the VP. Therefore, Notch2 was identified as a VMP-enriched gene by three independent techniques.

![Notch2 and GAPDH expression](image)

**Fig 6.2: Notch2 mRNA expression in the male and female UGS**

Analysis of Notch2 transcript levels in P0 male and female rat UGS (n=2). The VSU is the whole female rudiment, the VMP is the microdissected mesenchymal pad, and the VP is the microdissected male ventral prostate lobe. Transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represent the average Notch2 transcript abundance relative to the VP.
6.2.2 Notch2 transcript levels in the prostate and selected organs.

Further Northern blot analysis experiments were performed to study Notch2 transcript expression during embryonic and postnatal prostate development, and in other selected neonatal organs (Fig 6.3). Notch2 transcript expression was found to be most abundant in the prostate on the day of birth (P0) (100%), upregulated from embryonic life (e17.5) (54%), and then downregulated towards adulthood, most dramatically between P4 (81%) and P10 (39%) (Fig 6.3 Panel A). Also, it was noted that Notch2 expression in embryonic life was upregulated immediately prior to Dlk1 expression (Fig 5.3). Dlk1 has been proposed as a possible Notch ligand (Baladron et al., 2005). No difference in transcript abundance was found between the male (E17.5m) and female embryonic UGS (E17.5f). However, it should be noted that with the increasing age of VP organs, the increasing proportion of epithelial duct tissue would act to dilute the expression of the Notch2 transcript in the mesenchyme, and so the measurement of Notch2 represents a decreasing trend of mRNA expression rather than definitive values.

In the other neonatal organs (Fig 6.3, Panel B), Notch2 was found expressed highly selectively, and was most abundant only in the VP (100%) and brain (152%), with much lower levels measured in the remaining tissues (1-5%).

6.2.3 Notch 2 Protein Localisation in the UGS.

Localisation of Notch2 protein in the neonatal rat male and female UGS was investigated by immunocytochemistry (Fig 6.4). The antibody used, also localised Notch2 to the granulosa cells of P21 ovary (control tissue) (Fig 6.4, Panel D), as previously reported by Johnson et al., (2001). The Notch signalling pathway has been implicated in regulation of mammalian folliculogenesis. In the UGS, Notch2 protein was localised to the mesenchymal cells of the VMP and VP and also the SM layer (Fig 6.4). However, notably Notch2 expression was not uniform throughout the SM layer. A thin ribbon of brown staining representing Notch 2 expression was observed only in SM cells most proximal to the peri-urethral mesenchyme (Fig 6.4, Panels A, C and D). This differential protein expression in the SM layer suggests it is not a homogeneous tissue, as previously thought.
Fig 6.3 Notch2 mRNA expression during prostatic ontogeny and expression in other selected organs (P0).
Analysis of Notch2 transcript levels in the prostate; E17.5 until adulthood (Panel A), and in branching and non-branching organs collected from P0 rats (Panel B) (n=2). Ventral prostate (VP), bladder (Bld), brain (Brn), gut (Gt), heart (Ht), kidney (Kid), liver (Lvr), ovary (Ov), testes (Test), lung (Lg). The transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represents the average transcript abundance of Notch2 transcript relative to the P0 VP.
Furthermore, it was noted that the Notch2 expression in the SM layer was the converse of Dlk1 expression previously reported in this thesis (Fig 5.5; Panels A and B and Fig 5.6; Panel A). Comparison of Dlk1 and Notch2 expression in the male UGS found expression of both proteins in the mesenchyme of the VP, but expression in the SM layer only in adjacent cells. Dlk1 was found in cells at the outer edge of the SM layer, while Notch2 as localised to cells most proximal to the peri-urethral mesenchyme. As Dlk1 has been reported as a possible Notch ligand (Baladron et al., 2005), this distinctive protein expression pattern may provide supportive evidence of a Notch2-Dlk1 ligand relationship. Unfortunately, due to insufficient time, co-localisation of Dlk1, Notch2 and SM α-actin was not performed to confirm this expression pattern.

6.2.4 The effect of Testosterone on Notch2 distribution.
Androgens may directly or indirectly induce the release of regulatory factors from the mesenchyme during development. Consequently, it was investigated whether the distribution of Notch2, as a mesenchyme-expressed factor, in the VP, would be affected by exposure to Testosterone. VP organ rudiments were cultured in the presence (+T) or absence of testosterone (-T), and the distribution of Notch2 was localised by Immunocytochemistry (Fig 6.5). The VPs cultured with testosterone were found to grow larger in size and show more epithelial bud branching than VPs grown without. Notch2 protein expression was localised to the mesenchymal cells of the VPs under both culture conditions, with no discernable differences, and was therefore apparently unaffected by testosterone treatment.
Fig 6.4 Notch 2 expression in Male and Female UGS and the Ovary
Immunolocalisation of Notch2 protein in the P0 female (Panel A and B), male UGS (Panels C -E), and P21 Ovary (Panel F) (n=4). Ventral mesenchymal pad (VMP), smooth muscle (SM), urethral epithelium (URE), urogenital mesenchyme (UGM), epithelium (E), mesenchyme (M), follicle (F), granulosa cells (G). Arrows in Panels A point to areas of high Notch2 expression. Arrows in Panels C and D, point in particular to areas of Notch2 expression in a subset of cells in the SM layer, adjacent to the peri-urethral mesenchyme. Preabsorbed primary antibody was added to Panels B, E and F inset. The scale bar in Panel A represents 200μm in Panels B, C and E, and 100μm in Panels D and F.
Fig 6.5 The effect of testosterone on Notch2 expression.
Immunohistochemical localisation Notch2 protein in P0 VP lobes cultured for 6 days in the presence (Panels C, D and F) or absence (Panels A, B and E) of testosterone (10^{-8}M). Preabsorbed primary antibody was added to Panels E and F. Epithelium (E), mesenchyme (M) and end of VP organ proximal to urethra (U). The scale bars in Panel A represent 200\mu m in Panels C. and in Panel B the scale bar represents 100\mu m in Panel D-F.
6.2.5 Summary of Notch2 Results

Notch2 was identified as a VMP- enriched transcript by LongSAGE, LightCycler and Northern blot analysis. During development and growth of the prostate, Notch2 transcript expression, was most abundant during late embryogenesis and the perinatal period, when bud induction, epithelial growth and branching occurs. In other selected neonatal organs, Notch2 expression was most abundant in the VP and brain. In the P0 UGS, Notch2 protein was localised to the mesenchyme of the VMP and VP and the smooth muscle layer. Notch2 expression was not uniform in the SM layer, and was most strongly expressed in a narrow ribbon of cells lying adjacent to the periurethral mesenchyme. Comparison of Notch2 and Dlk1 expression in the UGS, supported a possible Notch2 –Dlk1 ligand relationship, but this requires further studies. Additionally, Notch2 was localised to mesenchyme of VP organ cultures, and the protein distribution was unaltered by treatment with Testosterone. Finally, due to insufficient time, this thesis lacks functional data on the activity of Notch2 in the developing prostate, and this will be the subject of further studies.
Chapter 7

Pleiotrophin

7. Pleiotrophin

7.1 Introduction

Pleiotrophin (Ptn) belongs to a novel two-member family of heparin binding molecules, with the structurally related (50%) protein midkine (Mk), also called retinoic acid-induced heparin-binding factor (Kadomatsu et al., 1997). Besides amino acid homology, Ptn and Mk also use the same receptors and share many biological activities. Pleiotrophin (Li et al., 1990) is an 18kDa protein that has been independently identified by different approaches leading to several names including heparin affin regulatory peptide (HARP) (Courty et al., 1991) and heparin-binding growth-associated molecule (HB-GAM)(Rauvala, 1989). The amino acid sequence of pleiotrophin is highly conserved among human, rat, mouse and bovine species, with 98% similarity (Li et al., 1990, Tezuka et al., 1990). Also, as its alternative names suggest, Ptn protein interacts with heparin sulphate proteoglycans (HSPGs) at the extracellular matrix (Vacherot et al., 1999b). Therefore, Ptn expression is most often localised to the basement membrane at the surface of differentiating cells (Mitsiadis et al., 1995).

The Ptn protein contains 24% of basic residues (18% of lysines), mainly arranged in two clusters at the N- and C- terminal regions, and five intrachain disulfide bonds (Kilpelainen et al., 2000). The molecule is organised in two β-sheet domains linked by a flexible linker, and each of these two domains includes one heparin -binding site (as reviewed in Papdimitriou et al., 2004).

Pleiotrophin was first isolated from neonatal rat brain as a molecule that induced neurite outgrowth of embryonic neurons (Rauvala, 1989), but further studies have found that it is present in a variety of non-neuronal tissues including bone (Imai et al., 1998), cartilage (Dreyfus et al., 1998), heart, (Chen et al., 2004), liver (Asahina et al., 2002) limb bud (Szabat and Rauvala, 1996), uterus (Milhiet et al., 1998) and testes (Zhang et al., 1999). In general however, the most frequent reports of Ptn expression are in tissues derived from mesoderm (Li et al., 1992) and in organs in which mesenchymal-epithelial interactions are thought to play an important role,
such as the salivary gland, lung, pancreas, kidney and mammary gland (LeDoux et al., 1997, Mitsiadis et al., 1995, Vanderwinden et al., 1992).

Ptn expression has been reported as a highly regulated temporal and cell-type specific process. Ptn expression is usually only present in tissues during late embryogenesis and perinatal growth, and expression usually fades around the time of birth (Li et al., 1990, Rauvala, 1989 Rauvala et al., 1994). In the adult, Ptn is expressed only at minimal levels in a very few tissues. The highest levels of Ptn expression reported in the adult are in the brain (Vanderwinden et al., 1992), but Ptn has also been found in the adult pancreas (Klomp et al., 2002), uterus (Milhiet et al., 1998), testes (Zhang et al., 1999) and mammary gland (Bernard-Pierrot et al., 2004). Additionally, Ptn expression has been associated with prostate (Vacherot et al., 1999), ovarian (Nakanishi et al., 1997), testicular (Aigner et al., 2003), pancreatic (Klomp et al., 2002) and breast cancer (Zhang et al., 1997), solid gliomas (Mentlein and Held-Feindt, 2002), neuroblastomas (Nakagawara, et al., 1995), melanomas (Souttou et al., 1998), and in several malignant cell lines of different origin (as reviewed in Papadimitriou et al., 2004).

Ptn and the closely related Mk have both been implicated as important proteins during development. Furthermore, Ptn and Mk share a similar pattern of expression during a period of foetal development. However, Mk expression peaks around midgestation, whereas Ptn expression peaks around the day of birth. In the mouse urogenital tract, for example, Ptn and Mk was localised to the surface of the epithelial cells and mesenchyme of the genital tubercle during early embryonic development, but only Ptn expression persisted at 16.5 dpc (Mitsiadis et al., 1995). It has been proposed therefore, that development involving these factors may be regulated by the difference in their temporal expression (Mitsiadis et al., 1995). Additionally, it has been reported that Ptn-deficient mice (Arnet et al., 2000) and Mk-deficient mice (Nakamura et al., 1998)) have no gross anatomical abnormalities. This does not necessarily indicate that Ptn or Mk is not required for normal development. Instead it has been proposed, that because they are closely related and found co-expressed during development, it is likely that Ptn might compensate for
Mk and vice versa, when there is an overlap in temporal expression (Mitsiadis et al., 1995).

Pleiotrophin is generally described as a secreted growth factor that plays a key role in cellular growth and differentiation. Ptn has also been demonstrated as involved in angiogenesis and tumour growth (as reviewed in Papadimitriou et al., 2004). The angiogenic properties of Ptn have been associated with inflammatory diseases such as endometriosis (Chug et al., 2002), rheumatoid (Pufe et al., 2003) oesteoarthritis (Pufe et al., 1993) and renal fibrosis (Henger et al., 2004), and in tumour growth and progression, Ptn may act to recruit stromal tissue and blood to expanding tumours (Czubayko, et al., 1996, Choudhuri et al., 1997, Jager et al., 1997, Sugino et al., 2002). Ptn has also been linked to cancer because of its activity in vitro, stimulating the proliferation of a wide range of cells including epithelial, endothelial and fibroblastic cells (Fang et al, 1992), and preventing apoptosis. The highly restricted expression of Ptn protein in the adult, unlike the many other growth factors implicated in cancer, make it a possible target for tumour therapy. Additionally, elevated serum levels of Ptn have been measured in patients with pancreatic (Klomp et al., 2002) or testicular cancer (Aigner et al., 2003), and so monitoring Ptn serum levels has been suggested as possible diagnostic marker of these diseases (Papadimitriou et al., 2004).

The most compelling evidence for the mechanism of Ptn activity has been provided by the expression pattern of Ptn mRNA and its protein. Ptn mRNA is mainly found in the neuroectoderm and mesoderm of the rat and mouse foetus, whereas its protein is found in endoderm and nerve tissues (Vanderwinden et al., 1992, Mitsiadis et al., 1995). Therefore, this suggests that Ptn act as a mesenchymal factor on epithelial tissues during development. With relevance to branching organs like the prostate, Ptn mRNA and protein expression patterns have implicated Ptn in regulating development of the kidney (Sakurai et al., 2001) and mammary gland (Bernard-Pierrot, et al., 2004, LeDoux et al., 1997).
In the mammary gland, for example, Ptn mRNA and protein were found expressed in adjacent myoepithelial and glandular epithelial cells, during development, pregnancy and lactation. This expression pattern was reported to suggest that Ptn is acting in a paracrine and autocrine growth factor in the regulation of mammary gland development and its homeostatic regulation during pregnancy and lactation (Bernard-Pierrot, et al., 2004, LeDoux et al., 1997).

Additionally, Ptn mRNA is expressed in the metanephric mesenchyme at the onset of kidney development (Vanderwinden et al., 1992), and has been reported as regulating branching morphogenesis of the uterine bud. Ptn protein isolated from the conditioned medium of metanephric mesenchyme cells was found to induce morphogenetic changes in the uterine bud in a concentration-dependent manner. Ptn was proposed to be acting like a classic "morphogen", inducing phenotypic changes in tissues according to its concentration (Sakurai et al., 2001).

Studies of Ptn in branched organs have not included reports of Ptn expression during prostate organogenesis, however Ptn activity has been associated with prostate cancer. Vacherot et al., (1999) reported Ptn mRNA was expressed in the stromal cells of normal adult prostate, prostate cancer and benign prostatic hyperplasia (BPH). The Ptn protein was similarly expressed in the stromal cells, but in addition, Ptn protein was also found in the epithelial cells of prostate cancer tissue, but not in the normal prostate or BPH tissues. They suggested that Ptn was therefore, acting in a paracrine manner from mesenchymal to epithelial cells, involved in tumour growth (Vacherot et al., 1999).

Finally, although there is considerable evidence for the paracrine manner of Ptn activity, the signal transduction pathways used by Ptn are not clearly established. The biological actions of Ptn are reported to be mediated by three different receptors; syndecan-3 (N-syndecan)(Raulo et al., 1994, Kinnunen et al, 1998), receptor protein tyrosine phosphatase, RPTB (Maeda et al., 1996) and anaplastic lymphoma kinase (ALK) (Stoica et al., 2001). Interaction of Ptn with syndecan-1 has also been proposed (Mitsiadis et al., 1995). Studies of Ptn signalling however have so far only
been able to conclude that a signal transduction pathway induced by Ptn includes phosphatidylinositol 3-kinase (PI3K) and mitogen activated protein kinase (MAPK), with MAPK downstream of PI3K (Qi et al., 2001, Stoica et al., 2001, Souttou, et al., 1997, Souttou et al., 2001).

7.2 Results

7.2.1 Ptn Transcript Expression in the VSU and VMP
Bioinformatic analysis of the LongSAGE libraries did not identify Ptn transcript expression as VMP-enriched (P<0.05), but did show a trend toward VMP-enriched expression (Graph 4.1). Ptn was selected for LightCycler analysis after mining the published research literature on Ptn activity. Ptn has been proposed to have a role in growth and development of other branching organs, like the kidney (Sakurai et al., 2001) and mammary gland (Bernard-Pierrot, et al., 2004, LeDoux et al., 1997), and Ptn expression has also been associated with prostate cancer (Vacherot et al., 1999). Subsequent LightCycler analysis identified Ptn as VMP-enriched by (Graph 4.1). Additionally, Northern blot analysis was used to compare the relative levels of Ptn transcript expression in the VMP and VSU (Fig 7.1). To verify correct RNA loading, the Northern blot was also hybridised with Fgf10 (Fig 8.3), a known VMP-specific gene (Thomson and Cunha, 1999). The radioactively labelled probes were prepared using primers in Table 2.4. A Northern blot comprising neonatal VP, VMP and VSU RNA demonstrated that Ptn transcript levels were ~2.6-fold higher in the VMP, in comparison to the VSU, and ~1.5 fold higher than in the VP. Therefore, two independent techniques identified Ptn as a VMP-enriched transcript.
Chapter 7  Pleiotrohin

Fig 7.1 *Ptn* mRNA expression in the male and female UGS
Analysis of *Ptn* transcript levels in P0 male and female rat UGS (n=2). The VSU is the whole female prostate rudiment, the VMP is the microdissected mesenchymal pad, and the VP is the microdissected male ventral prostate lobe. Transcript levels were quantified using a phosphoimager and normalised to *GAPDH*. The numbers below the figure represent the average transcript abundance of *Ptn* transcript relative to the VP

7.2.2 *Ptn* transcript levels in the prostate and selected organs.
Further experiments were performed by Northern blot analysis, to study *Ptn* transcript expression during embryonic and postnatal prostate development and also expression in selected tissues (P0) (Fig 7.2). *Ptn* transcript expression was most abundant in the prostate on the day of birth (P0)(100%), upregulated from embryonic life (E17.5) (59%), and downregulated after P0, most dramatically between P10 (49%)and P28 (8%), and of low abundance in adult life (3%)(Fig 7.2, Panel A). Also, no difference in transcript abundance was found between the male (E17.5m) and female embryonic UGS (E17.5f).

In the selected neonatal organs *Ptn* expression was most abundant was in the VP (100%) and ovary (94%). Similar *Ptn* expression levels were found in the brain (50%), liver (36%) and testes (39%), and low levels of *Ptn* expression were measured in the remaining tissues (3-22%) (Fig 7.2, Panel B).
Fig 7.2: *Ptn* mRNA expression during prostatic ontogeny and expression in other organs (P0).

Analysis of *Ptn* transcript levels in the prostate; E17.5 until adulthood (Panel A), and in branching and non-branching organs collected from P0 rats (Panel B) (n=2). Ventral prostate (VP), bladder (Bld), brain (Brn), gut (Gt), heart (Ht), kidney (Kid), liver (Lvr), ovary (Ov), testes (Test), lung (Lg). The transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represents the average transcript abundance of ptn transcript relative to the P0 VP.
7.2.3 Pleiotrophin Protein Localisation in the UGS

Localisation of Ptn protein expression was investigated in the male and female neonatal rat UGS by immunocytochemistry. Two antibodies against Ptn (Table 2.1) were used to localise Ptn and generated the same results (data not shown), therefore there was confidence in the staining pattern. Ptn protein expression was also immunolocalised in the neonatal kidney, (data not shown), where expression was most evident in the renal tubules. This result is consistent with previous reports of Ptn in the early kidney and a role for Ptn in renal branching morphogenesis (Sakurai et al., 2001, Vanderwinden et al., 1992). The Ptn antibody supplied by RnD Systems was used on the sections displayed in Figures 7.3-7.6.

In the neonatal female UGS, Ptn was localised to the VMP and UGM, and weak staining for Ptn was also present in the smooth muscle layer (Fig 7.3, Panel A). In the neonatal male UGS (Fig 7.3, Panels C and D), Ptn was expressed in the mesenchymal cells of the VP, plus there was strong staining for Ptn associated with the epithelial cells of the VP, and weak Ptn staining present at the basement membrane of the URE, and in the SM layer. Furthermore, the Ptn associated with the epithelial cells in the VP displayed a gradient of expression, with the strongest staining at the tips of the ducts, and decreasing towards the urethra. At this stage, it was unclear if the Ptn protein was expressed within the cytoplasm or at the surface of the epithelial cells. It was speculated that the Ptn expressed by the mesenchyme, might diffuse into the epithelial cells, to give this staining pattern. This was investigated further.

To study the precise location of Ptn expression in the epithelial duct cells, fluorescent probes were used to co-localise Ptn with β-catenin, a cell-surface protein marker (Fig 7.4). Ptn expression was as previously reported (Fig 7.3). The strongest Ptn staining in the mesenchyme was in the peripheral VP, and in the mesenchyme cells juxtaposed to the epithelial cells. Plus, Ptn was found co-expressed with β-catenin, and therefore located at the surface, of the epithelial duct cells.
Fig 7.3: Pleiotrophin expression in Female and Male UGS

Immunohistochemical localisation of the Ptn protein in female (Panel A and B) and male rat UGS (P0) (Panel C-E) (n=5). Panels B and E are the negative control sections. ventral mesenchymal pad (VMP), smooth muscle (SM), urogenital mesenchyme (UGM), urethral epithelium (URE), BM (basement membrane), ventral prostate (VP), anterior prostate (AP). Black arrows in Panels A and D point to areas of high Ptn expression. Panel D shows the concentration of Ptn at the epithelial duct tips. The scale bar in Panel A represents 200μm in Panels B, C, E and 50μm in Panel D.
Also confirmed was the gradient of Ptn expression, with the strongest Ptn staining associated with the epithelial cells most distal from the urethra (the epithelial duct tips). From previous studies on Ptn expression, it is known that Ptn protein interacts with heparin sulphate proteoglycans (HSPGs) at the extracellular matrix (Vacherot et al., 1999b), and its expression is most often localised to the basement membranes on the surface of differentiating cells (Mitsiadis et al., 1995). Therefore, the findings of this thesis correlate well with the literature reports.

The epithelial cells at the tips of the branching ducts in the prostatic lobes are characterized as undifferentiated and undergoing rapid proliferation (Marker et al., 2003). In this thesis, it was hypothesized that the association of strong Ptn staining with the epithelial tip cells might implicate Ptn as having a role in the differentiation process that occurs in these cells during development. To investigate this theory, Ptn was co-localised with p63, (a differentiation marker protein) in the neonatal male UGS (Fig 7.5). The expression of p63 becomes restricted to basal epithelial cells of the developing prostate, during the perinatal period. These basal cells are undifferentiated and still have the ability to proliferate (Marker et al., 2003). In the male P0 UGS, (Fig 7.5) p63 expression was localised to the undifferentiated basal cells of the VP ducts, and Ptn expression was localised, as described previously. Ptn and p63 expression showed some similarity in distribution i.e. at the epithelial duct tips and Ptn associated with the basement membrane of the VP ducts correlated with some p63 expression in the basal epithelial cells. However, replication of the co-localisation on different sections, determined no clear relationship between Ptn and p63 (undifferentiated cells) expression, and therefore the results are inconclusive, so far.
Fig 7.4: Pleiotrophin and β-catenin expression in Male UGS
Analysis of the co-localisation of Ptn and β-catenin protein in P0 male rat UGS by fluorescent immunocytochemistry. (n=4) The Ptn protein is shown in green, the β-catenin protein is shown in red and nuclear staining is shown in blue. Co-expression of Ptn and β-catenin is shown in yellow. Primary antibody was not added in Panel H. Mesenchyme (M), epithelium (E), urethra (UR), basement membrane (BM). The scale bar in Panel A represents 100μm. The scale bars in Panels B and C represent 50μm. Panel E is the same magnification as Panel C. The scale bar in Panel D represents 20μm.
Fig 7.5: Ptn and p63 expression in the Male UGS
Analysis of Ptn and p63 protein co-localisation in P0 male rat UGS by fluorescent immunocytochemistry (n=5). Ptn is shown in green, p63 is shown in red and the nuclear staining is shown in blue. No primary antibodies were added to Panel B. epithelial cells (E), mesenchyme (M), basement membrane (BM), basal cells (B), luminal cells (L). The scale bar in Panel A represents 50µm in both panels.

7.2.4 Effect of Testosterone on Ptn distribution
Androgens may directly or indirectly induce the release of regulatory factors from the mesenchyme during development. Consequently, it was investigated whether the distribution of Ptn, as a mesenchyme- expressed factor in the VP, would be affected by exposure to Testosterone. VP organ rudiments were cultured in the presence (+T) or absence (-T) of Testosterone, and Ptn expression was investigated by immunocytochemistry (Fig 7.6). The VPs cultured with testosterone were found to grow larger in size and show more epithelial bud branching than VP grown without. Ptn protein was found associated with the basement membrane and cell-surface of the epithelial duct cells, in a similar pattern to that reported previously in vivo (Fig 7.3-7.5). Additionally, it was noted Ptn expression was not uniform in all epithelial ducts within an organ. The ducts most proximal to the urethra had Ptn expression associated only with the basement membrane, while in more distally located ducts (at the edge of the organ), Ptn expression was also associated with the cell -surface of epithelial cells within the ducts. Overall however, there was no noticeable difference in Ptn expression between the treated and untreated organs, and Ptn distribution was not affected by Testosterone treatment.
Fig 7.6: The effect of testosterone on Ptn localisation
Immunohistochemical localisation of the Ptn protein in P0 VP organs cultured for 6 days in the presence (+T) (Panels E-H) or absence (Panels A-D) of testosterone (-T)(10^{-8}M). Panels D and H are the negative control sections. Epithelial cells (E), mesenchyme (M), basement membrane (BM), end of organ proximal to urethra (U). Peripheral ducts are marked * and ducts proximal to the urethra are marked with a black dot. The scale bars in Panel A and E represent 200\mu m. The scale bar in Panel B represents 100\mu m in Panels F and G, and 200\mu m (Panels A-D) in Panels C and G.
7.2.5 Functional Studies on Ptn

_Ptn_ transcript and protein expression in the mesenchyme of the male and female UGS (PO) has been confirmed, in this thesis. Next, a possible functional role for Ptn as a mesenchymal regulator of prostate organogenesis was investigated. VP organs were cultured in the presence or absence Testosterone plus recombinant human PTN (RnD Systems) (0.25 -2.5 µg/ml) (Table 2.2). Treatment with Testosterone induced an increase in VP growth and branching morphogenesis, but treatment with PTN only, had no effect on organ growth (data not shown).

7.2.6 Summary of Pleiotrophin Results

_Ptn_ was intuitively selected from the LongSAGE libraries as a candidate gene, and also showed a trend towards VMP- enriched expression by tag counts. LightCycler analysis and Northern blot analysis confirmed _Ptn_ was VMP –enriched. During development and growth of the prostate, _Ptn_ transcript expression, was most abundant in the P0 VP and late embryogenesis and the perinatal period, when bud induction, epithelial growth and branching occurs. In other selected neonatal organs, _Ptn_ transcript expression was most abundant in the VP and ovary.

_Ptn_ protein was localised to the mesenchyme of the VMP and VP and weakly in the SM layer in the UGS (PO). Additionally, in the male UGS, _Ptn_ protein was associated with the basement membrane and surface of the epithelial duct cells, and also exhibited a gradient of expression, where with most abundant expression was associated with the undifferentiated tip cells. This was thought to suggest a role for _Ptn_ in regulating differentiation of these epithelial duct cells. However, co-localisation of _Ptn_ and p63 (differentiation marker) showed no clear relationship between _Ptn_ expression and p63 (undifferentiated cells), and at present the results remain inconclusive. Additionally, _Ptn_ was localised to mesenchyme of VP organ cultures, and the protein distribution was not affected by treatment of the growing organs with Testosterone. Finally, due to insufficient time, this thesis lacks data on the function of _Ptn_ in the developing prostate, and will be the subject of further studies.
8. Other Candidate Genes; *Nell2, MMP2 and MMP14*

8.1 Introduction

The three remaining candidate genes *Nell2, MMP2 and MMP14* were selected for further studies after LightCycler analysis (Chapter 4). Unfortunately, there was insufficient time to characterise their expression fully.

8.1.1 Nell2

Neural epidermal growth factor –like like 2, (*Nell2*), is a member of the mammalian *Nell* gene family, which has only two members (*Nell1* and *Nell2*). The *nel* (neural epidermal growth factor) gene was originally identified as a gene encoding a protein containing six epidermal growth factor (EGF-) like domains from a chick-embryo cDNA library (Masuhashi et al., 1995). NELL1 and NELL2 were first isolated from a human foetal brain cDNA library as part of the Human Genome Project (Watanabe et al., 1996), and then from a rat brain cDNA library (Kuroda et al., 1999). The *NELL2* gene is more closely related to the *nel* gene than *NELL1* (Masuhashi et al., 1995). Chicken NEL is approximately 40-50% homologous to mammalian NELL1 and approximately 80-90% homologous to mammalian NELL2. The *NELL* genes code for large peptides of approximately 90kDa and are heavily glycosylated with an additional approximately 50kDa of N-linked carbohydrate moieties for total monomeric size of approximately 140kDa. The secreted forms of the NELL peptides are processed to approximately 130 kDa. monomers and form approximately 400 kDa. trimers in solution (Kuroda et al., 1999).

NELL2 is a secreted glycoprotein, and belongs to a class of proteins that contain epidermal growth factor domains. This group includes amongst others, Notch and its ligands (Delta/Serrate/Lag (DSL), the Neuregulin (NRG) family of ligands and the SLIT family of ligands (Kuroda et al., 1999, Kuroda and Tanizawa, 1999, Oyasu et al., 2000). The EGF-like domains are likely to serve as potential binding sites for novel protein interactions. NELL2 also belongs to the laminin G/N-terminal thrombospondin-1 (N-TSP1)/Pentraxin supergene family, a large class of multi-domain adhesive proteins in the extracellular matrix that act as molecular bridges between cells and the matrix, and participate in cell-cell communication and contact
Lastly, NELL2 also belongs to the chordin-like domain family.

NELL2 contains a cleavable signal peptide and N-TSP1 domain, six EGF-like and five cysteine-rich chordin-like/von Willebrand Factor C (C/vWc) domains. Also, NELL2 has also been identified as a protein that interacts with protein kinase C (PKC), by which it is phosphorylated, and as a member of the EGF-like family it has been shown to have three Ca^{2+} binding sites in its EGF-like repeat domains and suggested to play a role in the regulation of cellular events related to Ca^{2+} (Kuroda et al., 1999).

*NELL1 and NELL2 expression is tightly regulated in several vertebrate tissues both during development and in the adult (Kim et al., 2002, Kuroda et al., 1999, Luce and Burrows, 1999, Matsuhashi et al., 1995, Nelson et al., 2002, Oyasu et al, 2000, Watanabe et al., 1996). During development, NELL2 expression is temporally and spatially restricted to a period of differentiation in discrete regions of the nervous system, the pharyngeal arches, and the forelimb and trunk skeletal muscle (Nelson et al., 2002). NELL2 expression has also been reported in the early kidney and heart (Matsuhashi et al., 1995, Watanabe et al., 1996). In the adult human, NELL2 is reported to show brain specific expression (Oyasu et al., 2002, Watanabe et al., 1996). However, NELL2 expression has also been found in the adult prostate, where an overexpression of NELL2, localised to the basal epithelial cells was associated with benign prostatic hyperplasia (DiLella et al., 2001). Additionally, NELL2 and NELL1 have been detected in several transformed cell lines, hyperplasias and malignancies (DiLella et al., 2001, Kuroda et al., 1999, Luce and Burrows, 1999, Maeda et al.,

*Fig 8.1: Schematic of NELL2 protein
N-terminal thrombospondin 1 motif (TSP-1) with N-terminal heparin-binding domain (HB) and coiled-coiled (cc) domain; chordin-like/von Willebrand factor c motif (C/vWc); epidermal growth factor-like motif (EGF-like) (Nelson et al., 2004).*
The expression of NELL2 has been well documented but the role of NELL2 in vivo has not been fully studied.

The focus of Nell2 studies have been on its expression in the developing and adult nervous system, where it is thought to play an important role during neuronal development (Nelson et al., 2004). The likely role of Nell2 in other developing tissues has not been studied. However, based on homologies to the Notch/DSL, NRG, SLIT and TSP-1 proteins, Nell2 has been proposed as a signalling ligand in similar key developmental events. A summary of Notch function during development is included in Chapter 4. Also, the ligand Slit and its receptor ROBO are known to acts as target-derived factors that attract or repel extending neurites during neuronal development (reviewed in Nguyen-Ba-Charvet and Chedotal, 2002). Finally, TSP-1 is known to regulate cell proliferation, migration and apoptosis in a variety of physiological and pathological conditions such as wound healing, inflammation, angiogenesis and neoplasia (Chen et al., 2000). The association of NELL with some hyperplasias and malignancies indicates that mis -regulation of NELL genes is correlated with oncogenic potential in certain cancerous pathologies (DiLella et al., 2001, Kuroda et al., 1999, Luce and Burrows, 1999, Maeda et al., 2001).

The mechanism of Nell2 activity has not been elucidated, as yet, but it has been suggested that NELL2 activity is reminiscent of the classical lateral inhibition mediated by Notch and Delta (Morrison et al., 2000, Wakamatsu et al., 2000). Recently, NELL2 was shown to demonstrate dual cell autonomous and paracrine functions, in the brain. NELL2 promoted differentiation of neuronal progenitor cells that express NELL2 while simultaneously stimulating neighbouring cells to proliferate (Nelson et al., 2004). How NELL2, a secreted glycoprotein could simultaneously exert both cell autonomous and paracrine activities are not yet understood. Possible direct or indirect interactions with the Notch pathway and Wnt/β-catenin signalling pathway have been suggested (Hari et al., 2002, Nelson et al., 2004). Also, interactions with cell surface receptors such heparin sulfate proteoglycans or FGF receptor-heparin complexes or indirectly with the heparin sulfate -rich extracellular matrix (HS-ECM) (Kuroda et al., 1999). Finally, NELL2 could also bind other secreted factors since it contains cysteine-rich chordin-like/von Willebrand factor-C and N-TSP1 motifs, which can interact with TGFβ/ BMP family
members and a myriad of cell-surface and ECM molecules, respectively (as reviewed in Adams and Tucker, 2000, Bornstein and Sage, 2002, Garcia Abreu et al., 2002).

8.1.2 MMP2 and MMP14
Matrix metalloproteinases (MMPs) are a large family of Zn$^{2+}$ and Ca$^{2+}$ dependent enzymes with different substrate specificities. MMPs share many structural and functional properties and are known to take part in extracellular matrix (ECM) remodelling. MMPs are divided into groups according to their substrate specificities and exist as both secreted and membrane bound forms. MMPs are often secreted as inactive pro-forms and their activity is carefully regulated by an activation process (Docherty et al., 1992, Sommerville et al., 2003, Werb, 1997). At the level of the gene, different signalling pathways will activate different MMPs in response to cytokines, hormones or growth factors such as IL-1β, IL-6, TNFα, Platelet derived growth factor (PDGF) and basic fibroblast growth factor (Jones et al., 2003). Also, following activation, the MMPs can be inhibited by tissue inhibitors of metalloproteases (TIMPs) (Brew et al., 1999).

Structurally, nearly all MMPs have a basic minimal structure consisting of a terminal peptide signal sequence (enables their secretion), a propeptide domain (removed during activation) and a catalytic domain. MMP2 and MMP14 have an additional hemopexin like domain (involved in substrate recognition and the binding of endogeneous inhibitors-TIMPS). MMP2 also contains fibronectin like domains (that bind gelatin), and MMP14 contains a recognition motif for furin-like serine proteases (Fig 8.2) (Nuttal et al., 2004).

MMP14 (MT1-MMP) is the first member of a more recently established group of MMPs containing a membrane-spanning sequence, and is the most extensively studied of the cell-surface MMPs. MMP14 can degrade molecules such as collagen I, laminin, fibrin and fibronectin (Hotary et al., 2002, Ohuchi et al., 1997), but it is better known as the cell-surface activator of pro-MMP2 (Sato et al., 1994).
MMP2

Fibronectin-like domains

SP PRO C Zn²⁺ C H Hemopexin domain

MMP14

Furin recognition sequence MT-loop

SP PRO C Zn²⁺ C H Hemopexin domain TM

Fig 8.2: Schematic of MMP2 and MMP14 Proteins
Signal peptide (SP), pro-domain (PRO), catalytic domain (C), Zinc binding site (Zn²⁺), Hinge region (H), membrane-type loop (MT-loop), transmembrane domain (TM) (modified from Apte et al., 1997).

MMP14 also activates MMP13 (Knauper et al., 2002). MMP14 is subject to a high degree of transcriptional regulation during development. The pro-domain of MMP14 is constitutively removed by a proprotein convertase (furin) in the secretory pathway, resulting in delivery of catalytically active enzyme at the cell-surface (Sato et al., 1996).

MMP2, also called gelatinase A, is a member of the gelatinase subgroup of MMPs, with the capacity to degrade basement membrane Type IV collagen (Massova and Kotra, 1998). MMP2 is synthesised and released in the inactive form (pro-MMP2) (72kDa). It requires extracellular activation for biological function and the activation process is mediated in part by MMP14. Activation of pro-MMP2 occurs most efficiently at the cell surface and requires the formation of a trimolecular process in which TIMP2 links MMP14 to MMP2. The active form of MMP2 is 62kDa (Sato et al., 1994, Strongin et al., 1995).

The ability of MMPs to remodel the ECM has implicated their involvement in embryonic development and morphogenesis, as well as in the invasion and metastasis of cancer, including the process of angiogenesis. During development and early
postnatal life in mice, MMP2 is most highly expressed in the kidney, spleen, lung, skeletal muscle and heart, which partially mirrors the reported most abundant expression of MMP14 in placenta, kidney, heart and lung tissue (Nuttal et al., 2004). The destructive nature of these enzymes means that in normal tissue they tend to be expressed at low levels. High expression occurs as part of inflammation tissue remodelling (wound healing in adults) in response to growth factors, cytokines and interactions between cells and components of the ECM (Dyer and Hupe, 1998). The actions of MMP2 and MMP14 are too extensive to report fully, therefore only those relevant to prostate development and prostate cancer will be discussed here.

MMPs are often investigated for their involvement in development of branched organs, such as the kidney and lung and submandibular and mammary gland, where the key to successful branching morphogenesis is controlled turnover of the ECM and thus a balance between the activities of MMPs and TIMPs. The process of branching involves the thinning of the basement membrane of epithelial cells followed by metalloproteinase – mediated penetration into the mesenchyme. MMPs are described as mesenchymal factors that are ideally positioned to regulate such stromal-epithelial interaction (Sternlicht and Werb, 2001), although their expression is not always restricted to the mesenchyme. MMP2 expression has been localized to the mesenchyme of branching organs like the lung (Reponen et al., 1992), kidney (Stuart et al., 2003), and mammary gland (Wiseman et al., 2003) while MMP14 expression is often reported in the mesenchyme and epithelium (Kheradmand et al., 2002, Wiseman et al., 2003).

In the embryonic kidney, both MMP2 and MMP14 have both been reported to show increased expression at the uterine bud tips (Kanwar et al., 1999) and MMP14 is associated with the branching uterine bud tip in vitro (Meyer et al., 2004). Also, MMP14 is reported to be essential for tubular formation in vitro (Kadono et al., 1998, Miyamori et al., 2000), but blocking MMP2 activity has no effect on kidney morphogenesis in vitro (Lelongt et al., 1997). The roles of MMP2 and -14 in the developing kidney are not fully understood. Additionally, MMP14 null mice have been reported to have defects in the lung and submandibular gland. Lack of MMP14 substantially affected branching morphogenesis in the submandibular gland, and impaired angiogenesis in the lung affecting postnatal development. Also, it was
proposed that the fact that MMP2 null mice did not have comparable anomalies suggests that MMP14 acted in this context via mechanism(s) independent of pro-MMP2 activation (Oblander et al., 2005). However, it should be noted that a previous study had reported that lung branching and morphogenesis were severely affected by inactive MMP2 caused by low levels of MMP14 (Kheradmand et al., 2002). Finally, in development of the mammary gland MMP2 has been shown to regulate ductal invasion and repress lateral budding (Wiseman et al., 2003), while overexpression of MMP14 in mammary epithelium leads to development of roughly twice the number of branch sites (Ha et al., 2001).

Therefore, MMP2 and MMP14 are both important during development, but MMP14 null mice have been reported to have the most distinctive phenotype of all the null MMP alleles. MMP14 null mice usually die within 3 weeks of age, with postnatal dwarfism and musculoskeletal defects (Holmbeck et al., 1999, Zhou et al., 2000). Whereas, MMP2 null mice do not have major developmental anomalies, but have been found to show defective angiogenesis (Itoh et al., 1997, Itoh et al., 1998) and impaired postpubertal mammary gland duct development (Wiseman et al., 2003). Also, mutations in human MMP2 cause a multicentric osteolysis and arthritis syndrome (Martigentti et al., 2001).

MMP2 and MMP14 are also involved in pathological processes including cancer. MMP2, as a gelatinase, has the ability to degrade basement membrane Type IV collagen a major structural barrier for malignant tumor cells. Therefore, gelatinase expression in tumors has been associated with tumor invasion and metastasis. MMP 2 overexpression has been reported in many neoplasms (Nelson et al., 2000) including ovarian (Davidson et al, 1999, Sakata et al, 2000), urothelial (Davies et al., 1993, Gohiji et al., 1996, Kanayama et al, 1998), cutaneous (Dumas et al., 1999 Ikebe et al., 1999), gastric (Mori et al., 1997), breast (Pacheco et al., 1998, Talvensaari-Mattila et al. 2001), cervical (Nuovo et al., 1995) and prostate cancers (Kuniyasu et al, 2000, Riddick et al, 2005, Ross et al, 2003, Still et al, 2000, Stearns and Stearns, 1996, Trudel et al, 2003). Additionally, in prostate cancer, reports of inconsistent results of MMP2 expression suggest a discrepancy exists between MMP2 mRNA and protein expression levels. Riddick et al (2005) and Lightingaghen et al (2002) reported a decrease in MMP2 mRNA in prostate tumours, which differs from the increase in
MMP2 protein observed by Brehmer et al., (2003). Also, in contrast, a study identified an increase in MMP2 RNA in the highest grades of cancer (Still et al., 2000) and high expression of MMP2 has been associated with aggressive behaviour and metastasis (Wood et al., 1996). Finally, MMP2 and MMP14 co-localisation has been reported in the epithelial compartments of benign prostate epithelium, high-grade prostatic intraepithelial neoplasia and prostate cancer (Upadhyay et al., 1999).

MMP14 expression has also been reported to correlate with the malignancy of different tumour types such as the lung (Kettunen et al., 2004, Nawrocki et al., 1995, Sato et al., 1994, Tokuraki et al., 1995), gastric (El-Rifai, et al., 2001, Martinez et al., 2005, Nomura, et al., 1995, Mori et al., 1997), colon (Okada et al., 1995, Ohtani et al., 1996), breast (Okada et al., 1995, Polette et al., 1996, Ueno et al., 1997), cervical (Gilles et al., 1996), prostate carcinomas (Cao et al., 2005, Upadhyay et al., 1999), gliomas (Beliën et al., 1999) and melanomas (Hofman et al., 2000). Cao et al., (2005) recently reported that MMP14 promotes human prostate cancer invasion and metastasis, and suggest MMP14 as a suitable target to prevent cancer metastasis. Finally, one of the key roles that MMP2 and MMP14 play in promoting tumor growth is their function with TIMP-2 to promote tumoral angiogenesis (Sounni et al., 2002).

8.2 Results

8.2.1 Nell2, MMP2 and MMP14 Transcript Expression in the VSU and VMP

*Nell2* was identified by bioinformatic analysis of the SAGE libraries (P<0.05) (Appendix) and LightCycler analysis (Graph 4.1) as VMP-enriched. *MMP2* was identified by SAGE analysis as VMP-enriched, and although this was not confirmed by LightCycler analysis, it did show a trend towards VMP-enriched expression (1.2 fold). MMP14 was not identified as VMP-enriched by SAGE analysis but was included in LightCycler analysis after finding reports of MMP14 activity in other branching organs, the lung and submandibular gland (Oblander et al., 2005) plus, MMP14 has an important role as a pro-MMP2 activator (Sato et al., 1994).
Additionally, *Nell2, MMP2* and *MMP14* transcript expression was investigated in the male and female UGS (P0) by Northern blot analysis (Fig 8.3). The radioactively labelled probes were prepared using primers in Table 2.4. To verify the correct RNA loading, the Northern blot was also hybridised with *Fgf10*, a known VMP-specific gene (Thomson and Cunha, 1999). In a Northern blot containing P0 VP, VSU and VMP, *Fgf10* transcript abundance was shown to be ~4-fold higher in the VMP than the VSU, and ~2-fold higher than the VP. The high expression of *Fgf10* in the VMP in comparison to the VSU, confirmed that the blot had been correctly loaded. *Nell2* transcript expression was ~3-fold more abundant in the VMP than the VSU, and ~1.9 fold more abundant than in the VP. *MMP2* transcript expression was ~1.4 fold more abundant in the VMP than in the VSU, and was expressed at a similar level in the VP. The *MMP14* transcript expression was ~3-fold more abundant in the VMP than the VSU, and ~1.9 fold higher than the VP. Therefore, *Nell2, MMP2, MMP14* and *Fgf10* all showed VMP-enriched expression, to varying degrees and they were also expressed in the VP. Unfortunately, due to insufficient time and lack of working antibodies, no analysis of the their associated protein distribution was performed.

### 8.2.2 *Nell2* Transcript Levels in the Prostate and Selected Organs.

Further experiments were performed to investigate the expression of *Nell2* transcript during embryonic and postnatal prostate development and also expression in other selected neonatal organs. *Nell2* transcript expression (Fig 8.4, Panel A) was most abundant on the day of birth (P0), upregulated from embryonic life (31%) and downregulated dramatically by P4 (40%), with subsequent reduced expression levels until adult life (11%). Also, no difference in transcript abundance was found between *Nell2* expression in the male (E17.5m) and female embryonic UGT (E17.5f). The peak of *Nell2* expression coincided with an important period of prostate development, epithelial growth and branching morphogenesis.

In other selected neonatal tissues (Fig 8.4, Panel B), levels of *Nell2* expression were most abundant in the brain (152%) and VP (100%), and were measured at much lower levels (1-15%) in the remaining tissues.
Fig 8.3: Nell2, MMP2, MMP14 and Fgf10 mRNA expression in the male and female UGS.

Analysis of the transcript levels for the named genes in P0 male and female UGS. The VSU is the whole female rudiment, the VMP is the microdissected ventral mesenchymal pad, and the VP is the microdissected male ventral prostate. Transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represent the average transcript abundance of each transcript relative to the VP.
8.2.3 MMP2 Transcript Levels in the Prostate and Selected Organs.
MMP2 transcript expression was investigated during embryonic and postnatal development of the prostate (Fig 8.5, Panel A). MMP2 expression in the prostate was most abundant on the day of birth (P0) (100%, and downregulated dramatically by P10 (44%), and subsequently until adulthood (4%). Also, no difference in transcript abundance was found between the male (E17.5m) and female embryonic UGS (E17.5f). The most abundant levels of MMP2 expression coincide with the period of bud induction and epithelial growth and branching in prostate organogenesis. In the selected P0 organs (Fig 8.5, Panel B), MMP2 transcript expression was most abundant in the VP (100%) and ovary (85%), also with significant levels in the liver (58%), testes (42%) and bladder (29%). Low levels of MMP2 were measured in the remaining tissues (1-12%).

8.2.4 MMP14 Transcript Levels in the Prostate and Selected Organs.
MMP14 transcript expression was also investigated during prostatic ontogeny (Fig 8.6, Panel A). MMP14 expression was most abundant on the day of birth (P0) (100%), and downregulated dramatically by P10 (53%), and subsequently until adulthood (5%). Also, no difference in transcript abundance was found between the male (E17.5m) and female embryonic UGT (E17.5f). The most abundant levels of MMP14 expression coincide with the period of bud induction and epithelial growth and branching in prostate organogenesis. In the selected P0 organs (Fig 8.6, Panel B) MMP14 transcript expression was most abundant in the VP (100%), and similar levels were found in the ovary (87%), liver (78%) and testes (70%). Lower levels of MMP14 were measured in the remaining tissues (6-40%).

Also, it was noted that MMP2 and MMP14 expression levels were similar at each stage of prostate development, and both were most abundantly expressed in the same P0 organs.
Fig. 8.4: *Nell2* mRNA expression during prostatic ontogeny and expression in other organs (P0). Analysis of *Nell2* transcript levels in the prostate; embryonic (E17.5) until adult life (Panel A), and in branching and non-branching organs collected from P0 rats (Panel B) (n=2). Ventral prostate (VP), bladder (Bld), brain (Brn), gut (Gt), heart (Ht), kidney (Kid), liver (Lvr), ovary (Ov), testes (Test), lung (Lg). The transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represents the average *Nell2* transcript abundance relative to the P0 VP.
Fig. 8.5: MMP2 mRNA expression during prostatic ontogeny and expression in other organs (P0). Analysis of MMP2 transcript levels in the prostate; E17.5 until adulthood (Panel A), and in branching and non-branching organs collected from P0 rats (Panel B) (n=2). Ventral prostate (VP), bladder (Bld), brain (Brn), gut (Gt), heart (Ht), kidney (Kid), liver (Lvr), ovary (Ov), testes (Test), lung (Lg). The transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represents the average MMP2 transcript abundance relative to the P0 VP.
Fig. 8.6: *MMP14* mRNA expression during prostatic ontogeny and expression in other organs (P0).
Analysis of *MMP14* transcript levels in the prostate; E17.5 until adulthood (Panel A), and in branching and non-branching organs collected from P0 rats (Panel B) (n=2). Ventral prostate (VP), bladder (Bld), brain (Brn), gut (Gt), heart (Ht), kidney (Kid), liver (Lvr), ovary (Ov), testes (Test), lung (Lg). The transcript levels were quantified using a phosphorimager and normalised to GAPDH. The numbers below the figure represents the average *MMP14* transcript abundance relative to the P0 VP.
8.2.5 Summary of Results

*Nell2, MMP2 and MMP14* were identified as VMP-enriched genes by bioinformatic analysis of the LongSAGE libraries and/or LightCycler analysis. They were each further verified as VMP-enriched by Northern blot analysis. During prostatic prostate development, *Nell2, MMP2* and *MMP14* were all highly expressed around the day of birth (P0) and during the perinatal period, but with low expression in adult life. *Nell2* was expressed highly selectively in the panel of neonatal organs, with significant transcript levels present in only the brain and VP. *MMP2* and *MMP14* expression was most abundant in the VP, and with lower but significant expression in the liver, ovary and testes. Also, it was noted that *MMP2* and *MMP14* showed synchronicity in their levels of expression during prostate development and in other P0 organs.

These candidates support the use of the experimental approach used in this thesis to identify VMP-enriched genes. Unfortunately, due to insufficient time it was not possible to characterise their protein distribution and function.
9. Candidate Expression in Prostate Cancer Stroma

9.1 Introduction

Developmental mechanisms involved in prostate organogenesis may have relevance to the genesis and biology of prostate cancer. Prostate carcinogenesis is associated with perturbation of the reciprocal homeostatic mesenchymal-epithelial interactions that in the normal prostate maintains epithelial and stromal differentiation and growth quiescence. Changes in the stromal microenvironment, have been linked to disturbance of these homeostatic stromal-epithelial interactions, which promotes tumorogenesis (Grossfeld et al., 1998, Hayward et al., 1996, 1997, 1998). A process possibly mediated by developmental factors.

The role of stroma in the initiation and promotion of carcinogenesis has been considered for many years and is based upon the observation that “tumour stroma” frequently exhibits a variety of phenotypic differences relative to normal stroma (Bosman et al., 1993, Selijelid et al., 1999). Tumour stromal cells have been shown to exhibit abnormal migratory behaviour in vitro (Schor et al., 1988), alterations in cell-surface molecules (Choudhuri et al., 1975, Oishi, et al., 1981), altered expression of growth factors (Ellis, et al., 1994, Frazier and Grotendorst, 1997, Nakamura et al., 1997, Ponten et al., 1994, Yan et al., 1993, Yee et al., 1989), expression of prostaglandin synthase enzymes (Shattuck-Brandt et al., 1999, 2000) and alterations in the ECM (Pupa et al., 2002, Werb et al., 1996). They also show increased expression of MMPs, which may affect tumour initiation, growth, migration, angiogenesis, apoptosis, invasion and metastasis (Lynch and Matrisian, 2002).

The mechanisms by which stromal cells influence tumorogenesis are poorly understood, although it is suggested that it involves the differential expression of factors by cancer-associated fibroblasts (CAFs) versus normal prostate fibroblasts (NPFs), which act to modulate the local micro-environment. Studies have reported that CAFs have the ability to promote carcinogenesis in initiated, but non-tumorigenic human prostate epithelial cells, also increasing angiogenesis (Tuxhorn et al., 2002), while NPFs, restrain epithelial proliferation (Fig 9.1) (Grossfeld et al., 1998, Olumi et al., 1999).
Fig 9.1 Summary of tissue recombination experiments in which normal human prostatic fibroblasts (NAF) or carcinoma-associated fibroblasts (CAF) were grown in vivo in association with either normal human prostatic epithelial cells or initiated human prostatic epithelial cells (BPH-1). As indicated in the four-way grid, tumors only develop when initiated human prostatic epithelial cells (BPH-1) are grown in association with CAF. The bottom depicts the gross size differences between CAF + BPH-1 recombinants, which form tumors, versus NAF + BPH-1 recombinants, which do not. (modified from Olumi et al., 1999, in Cunha et al., 2004)

Therefore, cancer associated stroma and the molecules they express are proposed to be involved in regulating prostatic carcinogenesis (as reviewed in Cunha et al., 2002, 2004). In this thesis, it was hypothesized that mesenchyme-expressed molecules, abundantly expressed during prostate organogenesis, may be re-expressed in prostate cancer stroma cells (CAFs) to regulate neoplastic growth.
9.2 Results

9.2.1 Transcript Expression in the Human Prostate

Samples of human CAFs and NPFs were tested for the expression of the human orthologues of the developmental candidate genes; DLK1, PTN, NELL2, NOTCH2, MMP2 and MMP14. Initial studies were performed by general RT-PCR (Fig 9.1), using the relevant primers in Table 2.5. CAFs and NPFs were supplied as a gift from Dr. Simon Hayward, University of Vanderbilt, USA. The CAF and NPF samples were supplied as matched pairs, derived from the same patient. The CAFs and NPFs had been tested for their ability and inability, respectively, to stimulate tumorigenesis. PTN, NOTCH2, MMP2 and MMP14 expression was found in both CAFs and NPFs, but DLK1 and NELL2 expression was absent. Immunocytochemistry of tissue arrays with normal adult prostate and prostate cancer also verified DLK1 protein was not expressed (data not shown, personal communication, Mr. O. C. Grace).

PTN and NOTCH2 expression levels were then measured by LightCycler analysis in CAFs and NPFs, using the relevant primers in Table 2.5. The results were normalised by GAPDH. PTN and NOTCH2 were both found to show a downregulation of transcript expression in the CAFs in comparison to their matched NPFs (Graph 9.2). PTN mRNA expression was downregulated by 1.7 and 1.4-fold, and NOTCH2 mRNA expression was downregulated by 1.7 and 2.6-fold in CAF1 and CAF2, respectively.

PTN and NOTCH2 transcript expression in the CAF and NPF samples was then investigated by Northern blotting analysis. Radioactively labelled probes were prepared using the primers in Table 2.5. The results were normalised by GAPDH. Similar to the LightCycler results, PTN and NOTCH2 transcript expression was found to be downregulated in the CAFs in comparison to their matched NPFs, in both patients (see Fig 9.3). PTN mRNA was ~1.6-fold and downregulated ~2-fold, and NOTCH2 mRNA was downregulated ~1.4-fold and ~1.5-fold in CAF1 and CAF2, respectively.

However, these results are based CAF and NPF cells from only two patients, and therefore are not conclusive. It will be necessary to repeat these experiments on further CAF and NPF samples, to make robust conclusions.
Candidate Expression in Prostate Cancer Stroma

**Fig 9.2: Gene Expression in Human Foetal Tissues and Cells**
PCR analysis of DLK1 (429bp), PTN (346bp), NELL2 (377bp), NOTCH2 (692bp), MMP2 (294bp), and MMP14 (354bp) expression in cancer associated fibroblasts (CAF) and normal prostate fibroblasts (NPF) of human origin. Foetal liver (Lvr) and brain (Brn) were included as positive controls, and a cDNA template made with no reverse transcriptase enzyme (-RT) was the negative control.

**Graph 9.1: PTN and NOTCH2 mRNA Expression in CAFs and NPFs.**
LightCycler analysis of transcript expression in cancer-associated fibroblasts (CAFs) and normal prostate fibroblasts (NPFs) (n=2). Transcript expression was normalised to TBP. The samples were analysed as matched pairs; NPF1, CAF1 and NPF2, CAF2, where the paired samples were derived from the same patient. The graph displays the downregulation of PTN and NOTCH2 transcript expression in each CAF sample in comparison to its matched NPF.
**Fig 9.3: PTN and NOTCH2 mRNA Expression in CAFs and NPFs.**

Analysis of PTN and NOTCH2 transcript levels in cancer associated fibroblasts (CAF) and normal prostate fibroblasts (NPF). Foetal prostate (Prost), liver (Lvr) and brain (Brn) were included as positive control tissues (n=1). The samples were analysed as matched pairs (CAF1, NPF1 and CAF2, NPF2), where each pair was derived from the same patient. The transcript levels were quantified using a phosphoimager and normalised to GAPDH. The numbers below the figure represents the average transcript abundance relative to the NPF within each sample group.
9.2.2 Pleiotrophin Protein Localisation in the Human Prostate

The distribution of PTN protein in the human foetal (15wk) prostate (Fig 9.3) was investigated by immunocytochemistry using a Ptn primary antibody supplied by RnD Systems (Table 2.1). PTN protein was localised to the epithelial cells of the prostate ducts and more diffusely in the surrounding mesenchyme. The precise localisation of the PTN with relation to the epithelial cells was not clear. A further co-localisation experiment would be required to verify if PTN was associated with the cell-surface and basement membrane of the human epithelial cells, as was found in the neonatal male rat UGS (Fig 7.4). Due to insufficient time, localisation of the other candidate proteins was not performed in the human foetal prostate.

Fig 9.4: Pleiotrophin protein expression in the human prostate
Immunohistochemical localisation of PTN protein expression in human foetal (wk15) prostate (n=3). Epithelium (E), mesenchyme (M), wolffian ducts (WD). No primary antibody was added to Panel D. The scale bar represents 200μm in Panels A and D, 40μm in Panel B and 20μm in Panel C.
9.2.3 Summary of Results

Human cancer associated fibroblast cells (CAFs) and normal prostate fibroblast cells (NPFs) were tested for expression of candidate mesenchyme genes: DLK1, NOTCH2, PTN, NELL2, MMP2 and MMP14. The CAF and NPF cells expressed PTN, NOTCH2, MMP2 and MMP14 transcripts but DLK1 and NELL2 expression was absent. Further analysis of PTN and NOTCH2 expression, by LightCycler and Northern blot analysis, determined transcript levels were downregulated in the CAFs in comparison to their matched NPFs. However, these results are based on cells isolated from only two patients and are therefore are not considered as conclusive. Finally, PTN protein expression was immunolocalised in the human foetal prostate. Strong staining for PTN was associated with the epithelial duct cells and weak PTN associated staining was present in the surrounding mesenchyme.
10. Discussion

10.1 LongSAGE

The LongSAGE technique was used to investigate the gene expression profile of the prostate rudiment, and in particular, genes that are expressed in the mesenchyme/stromal cells. The LongSAGE method was updated from the published SAGE protocol (Invitrogen), and used to generate Libraries of the VSU and VMP from the neonatal female UGS. The tissue quality was recognised as being essential to the validity of the Libraries they would generate. Also, the number of PCR cycles in the cDNA amplification step was minimised, to reduce the possibility of creating a transcript bias in the gene expression profiles.

The Libraries generated were analysed and presented by the software according to the tag sequences and the associated tag abundance. A ‘back-comparison’ between the VMP and the VSU Libraries, and statistical analysis identified 218 VMP-enriched tags (P<0.05), where the minimum tag count in the VMP Library was 5 tags. The Libraries of tags were annotated to their associated genes using the rat LongSAGEmap, although this was not a comprehensive procedure. BLAST was able to increase the number of genes named, using the more fully characterised rat genomic data. BLAST was used to verify the accuracy of the gene assignments for the VMP-enriched transcripts (P<0.05), and further genes were intuitively selected from the Libraries. Genes expressed in the VSU and/or VMP Libraries were found to include those from known prostate development signalling pathways, such as the low abundance gene, Fgf10, and also genes known to be upregulated in prostate cancer e.g. hepsin. Therefore, there was confidence in the validity of the Library data, and in the sensitivity of the technique to identify transcripts with low levels of expression. A more detailed discussion of the preparation and analysis of the Libraries have been included in Section 3.3.

After completion of the Libraries, the aim of the project was to characterise the expression and function of a few of the candidate mesenchymal genes in more detail. It was hoped that the results of this study would help to validate the data in the
Libraries and support the main hypothesis of the project, that mesenchymal/stromal factors play a role in prostate development and cancer.

10.2 Experimental Validation of the LongSAGE Libraries
Completion of the LongSAGE Libraries had taken ~18 months, but ultimately planned to characterise some of the VMP-enriched genes identified. In the time remaining, only a few selected genes could be studied in any depth. QRT-PCR was used as the initial screening process to confirm the VMP-enriched expression of candidate mesenchymal transcripts, by an independent quantitative technique.

LightCycler analysis was able to confirm the preferential VMP expression (≥1.5 fold) of the majority of genes (18 out of 26) that had been identified by SAGE as VMP-enriched (P<0.05) or had demonstrated a trend towards VMP-enriched expression. The LightCycler analysis process was further validated by confirmation of expression of the ‘control’ genes. Fgf10 was confirmed as a VMP-enriched gene (Thomson and Cunha), while FgfR2 (Giri et al, 1999, Igarashi et al, 1998, Leung et al, 1997), Notch1 (Shou et al., 2001, Wang et al., 2004) and Hepsin (Dhanasekaran et al., 2001, Ernst et al., 2002, Klezovitch et al., 2004) each of which has been localised to the epithelial cells of the developing prostate and/or prostate cancer, were confirmed as VSU-enriched. Also, it was noted that on no occasion did the SAGE data and LightCycler data directly contradict each other i.e a VMP-enriched tag (P<0.05) identified by the SAGE data was never identified as VSU-enriched (≥1.5 fold) by qRT-PCR or vice versa.

However, only 3 of the 10 genes (30%) identified as VMP-enriched (P<0.05) by SAGE were confirmed by qRT-PCR (≥1.5 fold). It is proposed that this is not a direct measurement of the accuracy of the SAGE method, as the trends in gene expression in the VSU/VMP were generally upheld by qRT-PCR. It is suggested that any problems with the SAGE and/or the LightCycler analysis would contribute to poor rates of verified gene expression levels. For example, in the SAGE analysis, the cDNA amplification step and/or the depth of sampling in the Libraries, could
have introduced inaccuracies in the transcript quantification. While, on the other hand, LightCycler analysis was found to produce variable results in transcript quantification, particularly with the low abundance transcripts. Also, as both the SAGE and LightCycler analysis contain cDNA amplification steps, it was concluded that any variability between starting material used would result in different relative transcript expression levels being measured by each technique. It was hoped to minimise these differences during the LightCycler analysis by taking an average of the results from at least two individually prepared batches of VSU and VMP RNA. However, it was ultimately concluded that the LongSAGE or LightCycler analysis could not be wholly relied on to identify VMP-enriched tags, and Northern blotting analysis was used as a final confirmation.

Additionally, it was noted that the bioinformatic analysis of the SAGE data was more effective at identifying VMP-enriched genes than intuitive selection. Only 2 out of the 13 (~15%) intuitively selected possible VMP-enriched genes (i.e not including the VSU ‘control’ genes) were confirmed by LightCycler analysis. Therefore, SAGE analysis was twice as effective as intuitive selection in identifying VMP-enriched genes i.e 30% vs. 15%.

Further studies were only performed on transcripts that had demonstrated VMP-enriched expression, either in this project and/or in the Literature. Other considerations for the choice of transcripts to be studied in more depth, included the availability of relevant commercial antibodies and recombinant proteins that would be needed for a full characterisation of protein localisation and protein function. In the time remaining it would not have been possible to produce the relevant factors, by other means.

The six final candidates, Dlk1, Notch2, Ptn, Nell2, MMP2 and MMP14, were selected based on the results of the SAGE, qRT-PCR data, and/or previously reported studies. Specifically, four of the candidates (Dlk1, Notch2, Nell2 and MMP2) had been initially identified as VMP-enriched (P<0.05) by SAGE, whereas Ptn and MMP14 showed a trend towards VMP-enriched expression by SAGE, with
both factors also previously reported in association with roles in branching morphogenesis in other developing organs. Additionally, Midkine (Mk), a molecule closely related to Ptn, which also play an important role during early development, was identified in the LongSAGE libraries but was not included in the experimental analysis, as both the SAGE and qRT-PCR data did not suggest VMP-enriched expression.

The following Sections will discuss the results of more detailed experimental analysis of the six candidates, their transcript expression, and protein distribution in the developing rat prostate, and also in human cancer stroma.

10.3 Delta-Like 1 (Homolog)

10.3.1 Dlk1 Transcript Expression

*Dlk1* was identified as a VMP–enriched gene, and its expression was also confirmed in the male UGS during prostatic development. *Dlk1* expression was found to be most abundant during the processes of ductal branching morphogenesis, epithelial and mesenchymal/stromal cytodifferentiation (reviewed in Marker et al., 2003), which would suggest *Dlk1* has a role in one or more of these processes.

*Dlk1* has previously been reported as being highly expressed during embryonic life, and its expression becomes more restricted in adulthood, and the levels of Dlk1 expression are proposed to be critical to modulate the extracellular signals leading to either cell differentiation or proliferation (Baladron et al., 2002, Garces et al, 1999, Ruiz-Hidalgo, et al., 2002). As both of these processes are occurring during prostate development, it is speculated that the *Dlk1* expression is carefully regulated during the perinatal period to regulate normal organogenesis. Also, based on preliminary data, it is suggested that *Dlk1* mRNA is not an androgen–regulated molecule, as no difference was found between expression in the embryonic male and female UGS (E17.5), when levels of Testosterone would be high in the males in comparison to the
female. *Dlk1* expression is sufficiently high at E17.5 that a relatively small difference in expression between the sexes would have been noticeable.

*Dlk1* mRNA levels were found to be most abundant in the neonatal VP, liver, ovary and testes. High levels of *Dlk1* expression have been reported previously in each of these tissues, except the prostate (Floridon et al., 2000, Jensen et al., 1999, Ohno et al., 2001, Rockett et al., 2004, Tanimizu et al., 2003). Unfortunately, they provide no additional evidence of a likely role for Dlk1 in the prostate, as there are no obvious similarities between the morphology of their development.

Furthermore, it was noted that it would have been possible to test for molecules common to the branching process if actively branching embryonic lung and kidney RNA had been included in the tissue samples tested for candidate transcript expression. Low levels of transcript expression were measured in the neonatal lung and kidney for each of the candidate genes tested in this project. However, as the embryonic kidney and lung are more actively branching than their functioning neonatal tissues, this does not preclude the genes studied from a role in branching. Conversely, it is proposed that if high levels of a candidate gene transcript were found expressed in branching organs, other than the developing prostate, then this would suggest the candidate plays a common role in regulating branching morphogenesis.

It has been reported that *Dlk1* is expressed in branching tissues and may have a role in branching morphogenesis. Studies of *Dlk1* null mice have reported smaller lungs and kidneys than wild-type mice at birth, and that they did not catch up in size, such that the null mice had resulting pulmonary defects (Moon et al., 2002). However, a possible role for Dlk1 in branching in the developing prostate has not been studied.

**10.3.2 Dlk1 Protein Expression in the Neonatal UGS**

Dlk1 protein was localised to the mesenchymal cells of the VMP and VP, and also in the smooth muscle layer surrounding the urethra in the neonatal male and female rat UGS. The pattern of protein expression is supported by the transcript data presented
and by *in situ* hybridisation (data not shown, G. Vanpoucke, personal communication).

The distribution of the Dlk1 protein suggests it may be acting in a paracrine and/or autocrine manner, or a combination of both in the developing prostate. Here it is suggested that the Dlk1 protein expressed in the mesenchyme of the ventral prostate, could act in a paracrine manner on adjacent epithelial cells, either as the full transmembrane protein or as a secreted soluble factor, regulating epithelial proliferation/differentiation, and/or acting in an autocrine manner on the mesenchyme.

However, the story is further complicated by reports that the secreted and membrane-bound variants of Dlk1 may have opposite effects on differentiation (Garces et al., 1999, Mei et al., 2002). In this thesis, Dlk1 immunolocalisation has recognised both the soluble and membrane-bound forms, and without knowledge of definite Dlk1 receptor and its localisation, it is difficult to suggest a likely signalling pathway.

A possible role for Dlk1 is as a member of the Notch signalling pathway. Notch1 is expressed in the epithelial duct cells, and is reported to be essential for normal prostatic branching morphogenesis (Wang et al., 2004). Dlk1 is proposed to be able to bind to the Notch1 receptor (Baladron, et al, 2005, Laborada, 2000), even though it lacks the CSL domain reported to be important for Notch binding (Artavanis-Tsakonas et al., 1995), and could therefore act as a Notch ligand. Additionally, in this thesis it has been speculated that Dlk1 could also act as a ligand for Notch2.

The expression of Dlk1 in the smooth muscle layer was of interest because it contrasts with a report that Dlk1 was found absent in embryonic and foetal human smooth muscle (Floridon et al., 2000). Also, because Dlk1 demonstrated differential expression in subset of cells in the SM layer, a tissue generally thought of as homogeneous. Dlk1 was found co-localised with some of the SM α-actin expressing cells, particularly in the female UGS, but the exact nature of the Dlk1 only, expressing cells has not been confirmed. Additionally, in the female UGS (P0),
Dlk1 expression was notably absent from the section of the SM layer that lies adjacent to the VMP.

As a known differentiation/de-differentiation factor, it is speculated that Dlk1 may be active in the differentiating SM cells, as they form from the mesenchyme. The differential distribution of Dlk1 in the SM layer could implicate it in a role in SM differentiation. The direct/indirect control of SM distribution in the male and female UGS has been proposed as one hypothesis for the regulation of prostate organogenesis (Thomson et al., 2002) and if Dlk1 was involved in this process, would make it an interesting avenue for further studies. More specifically, it is speculated there may be some relationship between the absence of Dlk1 expression in the SM cells adjacent to the VMP, and the maintenance of these SM cells during development of the female UGS, while this area of SM layer thins and disappears in the male UGS.

10.3.3 The Effect of Testosterone on Dlk1 Protein Distribution.

Normal development of the prostate is dependent on androgen action, and may directly/indirectly act to regulate the release of growth factors from the mesenchyme that are also necessary for prostate organogenesis. It was hypothesised that the Testosterone could be acting indirectly through its control of Dlk1 distribution to regulate the differentiated or de-differentiated state of the epithelial cells and to control normal growth. According to this proposal, the Dlk1 expression would become restricted in its distribution either to the mesenchyme most proximal to the urethra i.e. adjacent to epithelial ducts (differentiating) or to the mesenchyme cells at the edge of the VP organ i.e. adjacent to the proliferating epithelial tip cells (undifferentiated). However, cultures of VP organs treated with Testosterone demonstrated the Dlk1 protein was immunolocalised to the mesenchymal cells of both the treated and untreated organs, and no obvious difference was noted in the distribution of the Dlk1 protein. Unfortunately, it was not possible to use Western blotting analysis to measure changes to protein levels, as the Dlk1 antibody was not compatible with this technique.
Therefore, the data presented here has not fully investigated the question of androgen-regulation of Dlk1, and so it is interesting that two previous studies have reported Dlk1 expression modified by Testosterone levels. Rats treated in utero with anti-androgens Linuron or DBP were found to show significantly altered Dlk1 expression in the foetal (GD21) epididymis (Turner et al. 2003) and Wolffian duct (Bowman et al., 2005), respectively. Dlk1 was significantly upregulated in the epididymis and slightly upregulated in the Wolffian duct. This was confirmed to not be the result of changes to the mesenchyme–epithelial cell ratio in the epididymis only (Turner et al. 2003). However, here it is speculated that the observed changes to Dlk1 expression may not be the result of androgen-regulation but are an indirect result of the differentiation process occurring in these developing tissues. At present androgen-regulation of Dlk1 remains inconclusive and further studies are required.

10.3.4 Dlk1 Functional Studies
The functional role of Dlk1 in normal prostate growth was studied using a VP assay system. Addition of DLK1 to VP organs cultured in the absence of Testosterone was found to induce an increase in their growth and the amount of epithelial branching observed, apparent by increased number of duct tips. This effect was found to be concentration-dependent (0.1 –0.75μg/ml), although no increase in the effect was found at the highest dose of 2.5mg/ul. On the other hand, VPs cultured in the presence of Testosterone showed an increase in organ size and amount of epithelial branching, but no additional effect was found with increasing DLK1 treatment. It is suggested that the Testosterone in this system was driving the growth process maximally, and overwhelming any effect of DLK1 treatment.

Additionally, it is noted that as Dlk1 is already expressed the VP organ culture system, addition of excess DLK1 creates an artificial overloading of the growth system. Therefore, an interesting and necessary further experiment could be the addition of Dlk1 antibodies to block the effect of Dlk1 activity on VP growth. If the converse results were observed in the VPs, with an inhibition of growth and branching, this would verify the positive effects recorded with DLK1 treatment. Based on the findings in vitro it is speculated that Dlk1 may have a similar role in
vivo, and may be involved in regulation of VP growth and branching morphogenesis, although whether this is via a paracrine and/or autocrine mechanism not known.

In conclusion, Dlk1 was identified as a mesenchyme-enriched gene. As yet its role remains unknown although it increases growth and branching in a VP organ culture system, and may have a similar effect in vivo. Also, it is speculated that the differential distribution of Dlk1 in the SM layer of the male and female UGS may play a role in the earliest stages of prostate induction, although it remains unknown if Dlk1 expression may be affected directly/indirectly by Testosterone. Finally, it is difficult to speculate on the mechanism of Dlk1 activity without knowledge of a definite receptor and its localisation, therefore further studies are required to understand the role of Dlk1 in the developing prostate.

10.4 Notch2

10.4.1 Notch2 Transcript Expression

*Notch*2 was identified as a VMP-enriched gene, and its expression was also confirmed in the male UGS (P0). Therefore, as a member of the Notch family of receptors with important roles during organ development, *Notch*2 was selected as an interesting candidate mesenchymal gene. *Notch*2 transcript and/or protein expression has been previously reported in other organs where development is regulated by epithelial-mesenchymal interactions, such as the lung (Ito et al., 2000, Weinmaster et al., 1992), kidney (Liemaster et al., 2003, McCright et al., 2001, Weinmaster et al, 1992), teeth (Mitsiadis et al., 1995) and breast (Parr et al., 2004) but not the prostate.

Other members of the Notch signalling family, *Notch1* and *Jagged2* (a Notch ligand) were also found expressed in the female UGS (P0), and identified as VSU-enriched genes, thereby localising their expression to the epithelial or SM cells. *Jagged2* expression has not previously been localised in the prostate but *Notch1* expression has been reported in the epithelial cells of the prostatic ducts (Wang et al., 2004).
Confirmation of Notch1 expression as a VSU-enriched gene supports the validity of the LightCycler data.

The differential distribution of Notch1 and Notch2 in epithelial and mesenchymal cells, respectively, is not unusual as their individual expression in different cell types has been reported in other developing organs (Johnson et al., 2001, Leimeister et al., 2003, Schroder and Gossler, 2002, Tanaka and Marunouchi, 2003, Tanimizu and Miyahjima, 2004, Weinmaster et al., 1992), and with different members of the Notch family having distinct functions (Kitamoto et al., 2005).

The most abundant Notch2 expression in the developing prostate was found to coincide with the processes of prostatic bud induction (E18.5), ductal branching morphogenesis, epithelial and mesenchymal/stromal cytodifferentiation (reviewed in Marker et al., 2003), and suggested Notch2 has a role in one or more of these processes. Also, the Notch2 levels in the embryonic UGS were shown to be upregulated immediately prior to the upregulation of Dlk1 expression, and it was speculated that this is further evidence of a link between Notch2 and Dlk1 activity.

The Notch family of receptors are generally known for their role in tissue patterning, during embryonic development, but the specific role of Notch2 has not been elucidated. Among the selected neonatal organs, expression of Notch2 was most abundant in the VP ovary, liver and testes. Notch2 expression has been previously reported in each of these tissues (as reviewed in Kitamoto et al., 2005), except the prostate. Additionally, it is suggested that the high levels of Dlk1 found in these tissues, would support the presence of a Notch2–Dlk1 relationship. Unfortunately, the abundant Notch2 expression in this selection of developing tissues does not provide additional evidence of a likely role for Notch2 activity in the prostate.

10.4.2 Notch2 Protein Expression

Notch2 was found strongly expressed in the mesenchymal cells of the VMP and VP, and was also present in the peri–urethral mesenchyme and smooth muscle layer in the male and female UGS (P0). The protein expression in the mesenchyme was
supported by the transcript expression data. The Notch2 expression in the SM was of particular interest because the SM layer is generally considered to be homogenous, but here demonstrated differential protein distribution. Also, the Notch2 localisation in the SM was noted to be the converse of the Dlk1 localisation, where the two factors were apparently expressed in adjacent SM cells. The significance of this distribution is not understood. Unfortunately, due to insufficient time, this complementary expression pattern of Dlk1/Notch2 has not been verified by co-localisation of Dlk1, Notch2, and SM α–actin. If confirmed this would support a hypothesis that Notch2 and Dlk1 are may represent a ligand-receptor relationship, acting in a paracrine manner between adjacent SM cells, while the expression of both molecules in the VP mesenchyme would act via an autocrine mechanism.

10.4.2.1 The Effect of Testosterone on Notch2 Protein Distribution

The Notch2 protein was localised to the mesenchyme of cultured VP organs, treated +/- Testosterone, but no difference was observed in the distribution of protein between the treatment groups. This experiment was not considered to be a full investigation of whether Notch2 expression was modulated by Testosterone and unfortunately the Notch2 the antibody tested was not compatible with Western blotting analysis to measure protein levels. However, this experiment did confirm that Notch2 was expressed in an in vitro VP, and could therefore, be used as a model system for studying Notch2 function. Also, it was suggested that in the absence of a working antibody, the levels of Notch2 transcript expression could be compared between treated and untreated organs, but unfortunately there was insufficient time to complete this.

Two previous studies, have proposed that Notch2 expression may modulated by androgens. These studies have also been referred to in this thesis with relation to Dlk1. Rats treated in utero with anti-androgens Linuron or DBP showed altered Notch2 expression levels in the foetal (GD21) epididymis (Turner et al. 2003) and Wolffian duct respectively (Bowman et al., 2005). Notch2 was significantly downregulated in the epididymis, and significantly upregulated in the Wolffian duct. As with Dlk1, these modified transcript levels were confirmed in the epididymis to
not be the result of changes in the mesenchyme/epithelial cell ratio, a possibility that was ignored in the Wolffian duct. It is speculated that the altered Notch2 expression may not be an example of androgen-regulation, but may be the indirect result of differentiation taking place in these in the developing tissues, or due to altered expression of its ligand. As yet, there are no conclusive studies as to the possible effect of Testosterone on Notch2 expression.

10.4.3 Notch2 Function in the Developing Prostate
As member of the family of Notch receptors, Notch2 could have a wide variety of roles, including determination of cell fate and tissue patterning (Kitamoto et al., 2005). Unfortunately, due to insufficient time, the role of Notch2 during prostate development was not studied in this thesis. The effect of a loss of Notch2 activity on VP growth and branching morphogenesis, could have been studied using an in vitro culture system. By adding recombinant Notch2 protein to VP cultures, this would have acted as a ligand trap, binding the endogenous Notch2 ligands, and inhibiting normal Notch2 activity. Alternatively, it was proposed to add an inhibitor of a member of the Notch signalling pathway, such as Presenilin, (reported to cleave the transmembrane Notch receptor prior to translocation to the nucleus) (as reviewed by Baron, 2003), to cultured VPs, and thus block Notch activity. Although in this case, the activity of all the Notch1-4 receptors could be affected, and the specific role of Notch2 would not be characterised.

Notch1, the only member of the Notch family previously studied in the prostate, has been reported as essential for normal prostatic branching morphogenesis (Wang et al., 2004) and, it is tempting to speculate that Notch2 may be of similar importance. For example, high expression of Notch2 has been reported in the branching embryonic lung (Ito et al., 2000) and kidney. More specifically, a study on Notch2 null mice reported perinatal mortality as a result of kidney malformation, specifically defects in glomerular development (McCright et al., 2001). As Notch2 expression is found during development of branching organs, and plays a critical role in the developing kidney, it is suggested this could be indicative of a role for Notch2 in the developing prostate.
Also, while Notch1 has been localised to the epithelial duct cells of the VP (Wang et al., 2004), Notch2 expression has been demonstrated in the mesenchyme, and therefore it is likely that each receptor has distinct roles. Also, studies have reported that Notch1 and Notch2 may have opposing effects. For example, in the mammary gland NOTCH2 is proposed to have a tumour-suppressive role while NOTCH1 has a tumour-promoting role (Parr et al., 2004). Consequently, it is speculated that this may be suggest opposing roles for the receptors during development? Further work is required to investigate Notch2 function in the prostate.

In conclusion, the Notch2 transcript was identified as a mesenchyme–enriched gene and from its expression pattern is speculated to be involved in prostatic bud induction and/or epithelial branching, although no functional studies were performed. The Notch2 protein localisation confirmed the transcript data, and also highlighted a possible receptor–ligand relationship between Notch2 and Dlk1. The possible modulation of Notch2 expression by Testosterone remains unknown. Further work is required to determine the role of Notch2 in prostate organogenesis, and whether Dlk1 could be acting as a receptor.

10.5 Pleiotrophin

Ptn has been associated with development of organs that are regulated by epithelial-mesenchymal interactions. Ptn expression has been reported in the developing kidney (Sakurai et al., 2001) and lung (Lu et al., 2003) and in the adult mammary gland (Bernard-Pierrot et al., 2004), the adult prostate, and prostate cancer (Vacherot et al., 1999), but has not previously been studied in the developing prostate.

10.5.1 Pleiotrophin Transcript Expression

Ptn was identified as a VMP-enriched molecule, and its expression was also confirmed in the VP of the male UGS (P0). During development, Ptn was most abundant in the VP during the processes of prostatic bud induction, ductal branching morphogenesis, epithelial and mesenchymal/stromal cytodifferentiation (reviewed in Marker et al., 2003), suggesting it has a role in one or more of these processes.
Among selected neonatal organs, *Ptn* expression most abundant in the VP and ovary, with lower levels measured in the brain, liver and testes. In support of these findings *Ptn* expression has been reported previously in each of these tissues during development (Asahina et al., 2002, Rauvala, 1989, Vanderwinden et al., 1992, Zhang et al., 1999), and has also been implicated as a growth factor in neuroblastomas (Nakagawara et al., 1995), testicular (Aigner et al., 2003), ovarian (Nakanishi et al, 1997) and prostate cancer (Vacherot et al., 1999). Confirmation of a growth factor expressed during development and in cancer is further evidence in support of the hypothesis that mis-expression of developmental factors may regulate neoplastic growth. However, unfortunately, the organs that demonstrated high *Ptn* expression also provided no clues of a likely common role in organogenesis.

### 10.5.2 Pleiotrophin Protein Expression

*Ptn* protein was immunolocalised to the mesenchymal cells of the VMP and VP, and also to the UGM with weak expression in the smooth muscle layer. The pattern of protein expression supports the transcript data already discussed, and also *in situ* hybridisation data (not shown, G. Vanpoucke, personal communication). Of particular interest, were the association of *Ptn* with the epithelial duct cells and its gradient of expression, with the most abundant expression at the ductal tips. Co-localisation of *Ptn* with β-catenin, (a cell-surface marker) confirmed that *Ptn* was associated with the cell-surface, and was not expressed in the cytoplasm of the epithelial cells. The distribution of *Ptn* found here is supported by previous studies which have reported that *Ptn* protein interacts with heparin sulphate proteoglycans (HSPGs) at the extracellular matrix (Vacherot et al., 1999b), and therefore, that *Ptn* is often found localised at the cell-surface and basement membrane of cells (Mitsidias et al., 1995).

Additionally, the gradient of *Ptn* expression exhibited in the ductal tips is reminiscent of a study in the developing kidney. Sakurai et al., (2001), demonstrated that *Ptn* was expressed in the metanephric mesenchyme of the developing mouse kidney (E13) *in vitro* and was also localised to the basement membrane of the branching uteric bud, where a concentration gradient of protein exhibited the most abundant expression at
the uterine bud tip. Furthermore, they treated kidney cultures with recombinant Ptn and caused an increase in the branching morphogenesis observed. Therefore, here it is highlighted that there are similarities between the distribution of Ptn in the developing kidney and prostate, with protein expression in the mesenchyme and also at the sites of proliferation and growth, the uterine bud and ductal tips, respectively. It is speculated that as Ptn has been reported to increase branching morphogenesis in the kidney it may have a similar role in the prostate. This will be discussed in more detail later.

The epithelial ducts tips that exhibited the most abundant Ptn expression in the developing prostate contain undifferentiated cells, and consequently it was proposed that Ptn might be involved in epithelial cell differentiation. Co-localisation of Ptn with p63 (a differentiation marker) (Wang et al., 2001) in the male UGS (P0) demonstrated no consistently positive correlation between the distributions of each protein. At present the results are inconclusive and further work is required to establish if Ptn plays a role in differentiation.

10.5.2.1 Effect of Testosterone on Pleiotrophin Protein Distribution

VPs cultured +/- Testosterone demonstrated that the distribution of the Ptn protein was not affected by exposure to androgens. Unfortunately, as the Ptn antibody was not compatible with Western blotting analysis, it was not possible to confirm if the levels of protein expression were affected. The localisation of Ptn to the mesenchymal cells and also to the cell-surface and basement membrane and of the epithelial cells was similar to that that observed in vivo. Additionally, it was suggested that the concentration of Ptn in the ducts located most distally from the urethra, was the result of the high Ptn concentration found at these growing tips. Also, it may be indicative of an association between Ptn and the undifferentiated cells at the ductal tips, as opposed to the more differentiated cells proximal to the urethra. Further work is required to study this conclusively.
10.5.3 Pleiotrophin Functional Studies

Ptn is a secreted factor that has been proposed to play a role in branching morphogenesis, including the developing kidney (Sakurai et al., 2001) and mammary gland (Bernard-Pierrot, et al., 2004, LeDoux et al., 1997). This was of particular interest because the developing prostate is a branching organ. As already mentioned, Sakurai et al., (2001), was able to induce an increase in branching morphogenesis, in a concentration-dependent manner, in kidney cultures by adding conditioned medium isolated from metanephric mesenchymal cells (containing secreted Ptn).

This thesis has highlighted the parallels between Ptn in the developing kidney and the prostate, however, when recombinant PTN (RnD Systems) was added to VP organ cultures, it was found to have no effect on growth or branching morphogenesis. As a similar result had been found in the kidney with this particular source of PTN protein (Sakurai et al., 2001) it was speculated that the recombinant protein was inactive, as opposed to concluding that Ptn does not play a role in prostate growth. No alternative source of PTN protein was available at the time. Other sources of recombinant PTN, such as conditioned medium isolated from prostate mesenchyme cells or metanephric mesenchymal cells may be more effective, but unfortunately, due to insufficient time, these experiments were not undertaken.

A model has been reported of the possible mechanism of Ptn activity in the developing kidney, and here it is speculated that it may also describe a likely model for Ptn activity in the developing prostate. Sakurai et al., (2001) proposed that during early kidney development, the high concentration of Ptn associated with the uterlic bud tip may act as a protein “reservoir”, and from here the Ptn could act in a paracrine manner to regulate branching morphogenesis. The concentration gradient of Ptn expression observed along the elongating uterlic stalk could be generated by matrix degrading proteases. The most abundant Ptn is found at the proliferating tips with decreasing amounts along the length of the stalk, which may induce elongation of the forming tubule. Also, the concentration gradient of Ptn would provide a basis
for modulating the shape and directionality of the developing uterine bud, and help to regulate branching morphogenesis.

Consequently, it was proposed that Ptn in the developing prostate is secreted by the mesenchyme of the VP and becomes associated with the epithelial cells in the ducts. Here it is bound to HSPGs and its release and activity could then be regulated by MMPs. Several members of the MMP family have been identified in the LongSAGE VSU library, e.g. MMP-2, -11, -14, -16 and -23, TIMP 1 and -2, and therefore could play a role in this mechanism. Then the direction and shape of the growing epithelial ducts would be modulated by the concentration gradient of the Ptn, with a high protein concentration at the duct tips, and decreasing amounts along the epithelial duct, a pattern that has already been confirmed. However, further studies are required to investigate this hypothesis more fully. Additionally, if the mechanism of Ptn activity is similar in two branching organs, this may be also related to other organs with branching epithelia.

In conclusion, Ptn was identified as a mesenchyme-enriched molecule that is abundantly expressed during prostate development. A possible association between Ptn and regulation of differentiation remains inconclusive but it has been suggested that similarities between the Ptn expression in the developing kidney and prostate are indicative of a role for Ptn in regulation of branching morphogenesis during prostate organogenesis.

10.6 Nell2

Nell2 was identified as a VMP-enriched gene and was also found in the VP of the male UGS (P0), with the most abundant expression on the day of birth. Nell2 was unique amongst the mesenchymal candidates, in having a very brief "window" of expression, but the significance of this is not yet understood. Therefore, it was concluded that Nell2 activity is most active during epithelial growth and branching morphogenesis taking place in the VP at this time. There have been no previous reports of Nell2 expression in the developing prostate, although Nell2 has been found
upregulated in BPH (DiLella et al., 2001), but its role is unknown. Additionally, as no commercial Nell2 antibody was available, it was not been possible to confirm the distribution of the Nell2 protein.

Most research on Nell2 has concentrated on its role in neuronal development (Aihra et al., 2003, Kuroda et al., 1999, Matsuhashi et al., 1995, Nelson et al., 2004). Therefore, it was not surprising that among the selected neonatal organs, Nell2 expression was most abundant in the VP and brain, where its highly selective expression suggested that there could be similarities in Nell2 activity in both tissues. Unfortunately, little is known in general about the activity of Nell2.

Previous studies have suggested that because of its homology to the Notch/DSL, NRG, SLIT and TSP-1 proteins, Nell2 may be acting as a signal ligand in similar key developmental events, including; cell proliferation, apoptosis and/or as a possible trophic factor (Aihra et al., 2003, Garcia Abreu et al., 2002, Kuroda et al., 1999, Matsuhashi et al., 1995, Nelson et al., 2004, Watanabe et al., 1996). During neuronal development NELL2 is proposed to act autonomously on progenitor cells and promotes differentiation into neuronal cells, and also acts in a paracrine manner to stimulate mitogenesis in adjacent cells (Nelson et al., 2004). Consequently, it was speculated that Nell2 could regulate proliferation/differentiation in the prostate mesenchyme and/or epithelial cells via an autocrine and/or paracrine mechanism. Additionally, as a factor with a possible role in branching morphogenesis in the developing prostate, it is interesting that Watanabe et al., (1996) has reported NELL2 expression in the early, developing kidney, although they made no suggestion of its likely roles.

Finally, a report of NELL2 over-expression in BPH (DiLella et al., 2001) supports the speculation that Nell2 is involved in prostate organogenesis, and more specifically can regulate epithelial growth. Also, it supports the hypothesis of developmental “awakening” proposed by McNeal, (1969).
In conclusion, Nell2 may be acting as proliferation and/or differentiation factor during prostate organogenesis with a role in epithelial growth and/or branching. The mechanism of Nell2 activity may be similar to that reported during neuronal development.

10.7 MMP2 and MMP14

MMP2 and MMP14 were identified as VMP-enriched genes that were also expressed in the male VP (P0). There are no previous reports of MMP2 and MMP14 transcript or protein expression in the developing prostate, although both have been associated with BPH and prostate cancer (Cao et al., 2005, Kuniyasu et al, 2000, Riddick et al, 2005, Ross et al, 2003, Still et al, 2000, Stearns and Stearns, 1996, Trudel et al, 2003, Upadhyay et al., 1999). The most abundant MMP2 and-14expression was measured in the developing VP during late embryogenesis and the perinatal period, suggesting that both genes could be involved in prostatic bud induction and/or epithelial growth and/or branching. Furthermore, it was demonstrated that both MMPs were expressed at similar levels in the developing prostate at each time-point measured, a finding which is supported by reports of the co-expression of MMP2 and MMP14 (Nuttal et al., 2004), in a variety of tissues and the role of MMP14 as a pro-MMP2 activator (Sato et al., 1994).

Unfortunately, in the absence of working antibodies it was not possible to confirm the distribution of the MMP2 and MMP14 proteins in the UGS (P0) although Karl Sternad (MSc. student), in collaboration with this project, used zymography to confirm the expression of pro- and active MMP2 in the male and female UGS (P0)(personal communication, data not shown).

The discovery of MMP2 and MMP14 expression in the developing prostate was not surprising as MMPs in general, are described as mesenchymal factors that are ideally positioned to regulate stromal-epithelial interactions (Sternlicht and Werb, 2001). Also, MMP2 and MMP14 have both been associated with branching morphogenesis in other organs, such as the kidney (Kadono et al., 1998, Kanwar et al., 1999, Meyer
et al., 2004 Miyamori et al., 2000), lung, submandibular gland (Kheradmand et al., 2002, Oblander et al., 2005) and the mammary gland (Ha et al., 2001, Wiseman et al., 2003). Their exact roles remain controversial, and whether they act independently or in concert. MMPs act with TIMPS in the controlled turnover of the ECM and thinning of the basement membrane of epithelial cells to facilitate penetration into the mesenchyme (as reviewed in Nuttal et al., 2004). Consequently, it was speculated that MMP2 and MMP14 may be either acting in concert or independently, in the developing prostate to regulate branching morphogenesis, and may also play additional roles.

Additionally, among the other selected neonatal organs tested, MMP2 and MMP14 expression was similar in each tissue, which is further support for their co-expression with the most abundant expression measured in the VP, ovary, liver and testes. A high expression of MMP2 and MMP14 has been previously reported in all of these tissues (Mazaud et al., 2005, Robinson et al., 2001, Zhou et al., 2004) except the developing prostate. Also, expression of MMP2 and MMP14 in the developing prostate, BPH and prostate cancer may be further support for the developmental ‘awakening’ hypothesis proposed by McNeal, (1969).

In conclusion, MMP2 and MMP14 were identified as mesenchyme-enriched molecules and may be involved in the processes of prostatic bud induction, epithelial growth and branching. The strongest case is for their involvement in branching morphogenesis, either acting in concert or independently. However further characterisation of the distribution and function of the MMP2 and MMP14 proteins in the developing male UGS is required.

10.8 Candidate Expression in Prostate Cancer Stroma

The mesenchyme is important not only during development, and recently, a role for mesenchyme/stroma in cancer has been established where developmental factors may be involved in mediating the effects of the stroma. Studies have demonstrated that cancer associated fibroblast cells (CAF) cells have the ability to promote
tumorigenesis in initiated, but non-tumorigenic human prostate epithelial cells, whereas normal prostate fibroblast cells (NPFs) restrain epithelial proliferation (Cunha et al., 2002, 2004, Grossfeld et al., 1998, Olumi et al., 1999). Therefore, CAF cells and the molecules they express have the ability to regulate neoplastic development (Barclay et al., 2005, Cunha et al., 2002, 2004, Orimo et al., 2005).

Research has suggested that there is a relationship between development and carcinogenesis, based on the observations that virtually all properties of neoplasms have a normal counterpart in embryonic development (as reviewed by Cunha et al., 2004). In this thesis, it was hypothesized that developmentally expressed mesenchymal factors may be mis-expressed in prostate cancer stroma, and regulate neoplastic epithelial development. This theory was supported by the identification of NOTCH2, PTN, MMP2 and MMP14 expression in CAF and NPF cells, although DLK1 and NELL2 expression were absent. PTN, MMP2 and MMP14 expression have previously been reported in association with prostate cancer (Cao et al., 2005, Kuniyasu et al., 2000, Riddick et al., 2005, Ross et al., 2003, Still et al., 2000, Stearns and Stearns, 1996, Trudel et al., 2003, Upadhyay et al., 1999, Vacherot et al., 1999), but NOTCH2 has not, although the Notch signalling pathway is known to be active (Shou et al., 2001, Satagata et al., 2004).

Preliminary data indicated that NOTCH2 and PTN expression was modestly downregulated in CAFs in comparison to their matched NPFs, but these results remain inconclusive as the cells were derived from only two patients. However, it is noted that this data supports a report by Parr et al (2004) that proposed NOTCH2 plays a tumour-suppressive role in breast cancer differentiation in contrast to a tumour-promoting role of NOTCH1. Consequently, as Notch1 upregulation is associated with prostate cancer, and is proposed to have a tumour-promoting role in TRAMP (transgenic adenocarcinoma of the mouse prostate) mice (Shou et al., 2001), it is suggested Notch2 could have a tumour suppressive role in prostate cancer, where downregulation in expression would be expected in the CAFs in comparison to their NPFs.
Conversely, Ptn has been reported as playing a significant role in promoting tumour growth and angiogenesis (as reviewed in Papadimitriou et al., 2004), and can stimulate the proliferation of a wide range of cells including epithelial, endothelial and fibroblastic cells, but also prevents apoptosis *in vitro* (Fang et al, 1992). In a study on prostate cancer, Vacherot et al., (1999), reported that PTN mRNA and protein was expressed in the stroma, of normal prostate, BPH and prostate cancer, but additional PTN protein expression was also found only in the epithelial cells of the prostate cancer. They proposed PTN was acting in a paracrine manner, secreted by the stroma, and promoting neoplastic epithelial cell growth. Although this data would at first appear to contradict the downregulation of PTN observed in the CAF cells, it is suggested that as Vacherot et al., (1999) did not measure PTN levels, their findings cannot be directly compared with the data reported here. Also, it is suggested that the modest PTN downregulation observed might not be significant, and levels of other transcript expression (data not shown) were found to vary considerably between samples. Additionally, it is speculated that PTN expression may have been altered when CAF and NPF cells were isolated from their associated epithelial cells and cultured *in vitro*. The presence of PTN binding receptors on epithelial cells undergoing neoplasia, but which are absent from normal epithelial cells and BPH (Vacherot et al., 1999), may affect PTN expression in prostate cancer. Therefore, levels of PTN expression in the isolated CAFs tested may have been affected by the absence of epithelial cells and their associated PTN receptors.

Alternatively, another proposal for PTN activity in the adult prostate is in the maintenance of homeostasis. This proposal is supported by two main facts; 1) PTN has been associated with the regulation of homeostasis in the glandular epithelial and myoepithelial cells of the mammary gland, during pregnancy and lactation (Bernard-Pierrot et al., 2004, LeDoux et al., 1997), and 2) PTN expression in the adult is normally highly restricted but is expressed in the normal adult prostate (Vacherot et al., 1999). Therefore, perhaps the downregulation of PTN observed in the CAF cells may be indicative of a loss of PTN activity and homeostatic regulation which precedes neoplastic growth.
Finally, PTN protein distribution in the human foetal prostate (15wks) was localised to the epithelial cells and weakly in the surrounding stroma, a pattern which was moderately similar to that observed in the neonatal male rat prostate, although Ptn was more evident in the mesenchyme of the rat. However, further work is required to determine if PTN is localised to the cytoplasm or the cell-surface of the epithelial ducts, and if there is a gradient of PTN expression, as identified in the rat epithelial ducts.

In conclusion, verification of the expression of four developmental genes; NOTCH2, PTN, MMP2 and MMP14, in CAF and NPF cells, supports the proposal that the mis-expression of developmental genes can result in neoplastic growth in adult tissues McNeal (1969). Further studies are needed to confirm the modified transcript expression measured in the CAF in comparison to the NPFs. These studies could include the addition of recombinant PTN or NOTCH2 proteins to CAF cells recombined with ‘initiated’ BPH cells in vitro.

10.9 Project Conclusions
10.9.1 A Novel Approach to Studying Prostate Development
Prostate development is a complex process involving androgen-regulation and mesenchymal-epithelial interactions. Most studies of gene expression in prostate development have focused on individual signalling pathways, or use techniques such as micro-arrays that preclude the possibility of identifying novel genes. However, as it is likely that a group of different factors regulates prostate development, LongSAGE was considered the most comprehensive approach to studying prostate organogenesis, providing a ‘snap-shot’ of the gene expression and signalling pathways.

This project focused on the importance of mesenchyme/stroma in prostate development and cancer. The mesenchyme/stroma produces molecules that can regulate prostate organogenesis, and may also play a role in prostate cancer. To date, only a few developmental mesenchymal factors had been identified and studied in
any detail e.g. Fgf7, Fgf10 and BMP4 (Lamm et al., 2001, Sugimura et al., 1986, Thomson and Cunha, 1999). At the time of preparing the VSU and VMP LongSAGE Libraries, no other rodent libraries of the developing prostate had been published. Recently, Zhang et al., (2006) has published SAGE Libraries of embryonic, (including UGM) neonatal, and P12 mouse prostate.

Also, the majority of research on gene expression in the prostate has concentrated on the study of prostate cancer because of its increasing prevalence and the associated healthcare costs. The importance of the stroma in regulation of neoplastic growth was first highlighted over a decade ago, but the mechanisms are poorly understood (Cunha et al., 1996, Cunha et al., 2002, Cunha et al., 2003). Prostate cancer is a heterogeneous tissue, making it difficult to study at the cellular level. The VMP provided a homogenous source of mesenchyme/stroma tissue, which is technically difficult to isolate in prostate cancer. Also, identifying VMP-enriched factors in the SAGE libraries highlighted developmental mesenchymal factors that might be active in prostate cancer stroma.

10.9.2 How successful was the use of LongSAGE analysis?

The LongSAGE approach was found to be moderately successful in the identification of mesenchyme–enriched specific transcripts. The VSU and VMP Libraries were found to contain signalling pathways known to be active during prostate development and factors upregulated during prostate cancer. Also, LongSAGE and the subsequent bioinformatics analysis of the data, was confirmed by qRT-PCR analysis, to be twice as successful as ‘intuitive’ selection in identifying VMP-enriched genes.

Also, it is proposed that SAGE analysis was successful in highlighting genes that have not previously been studied in prostate development and prostate cancer. Six candidate genes, previously not reported in prostate organogenesis, have been confirmed as demonstrating VMP-enriched and VP expression, and one gene (Dlk1) demonstrated a functional role in prostate development in vitro. Additional VMP-enriched/specific genes are currently under investigation by other members of Dr.
Thomson's research group, and their mesenchymal expression has also been confirmed.

Furthermore, one of the major benefits of using the LongSAGE technique was the potential identification of novel genes. The VSU and VMP Libraries, and more specifically the list of VMP-enriched ($P<0.05$) tags were found to contain many ‘unknown’ and novel tags. It is proposed that some of these previously unrecognised tags may represent important developmental regulatory factors that may have not have been highlighted if SAGE had not been used. It is hoped that improvements in characterisation of the rat genome will increase the number of genes able to be identified.

10.9.3 How could the LongSAGE data be further mined?
The LongSAGE libraries were interrogated for candidate genes by bioinformatic analysis and by intuitive selection. Two possible further avenues of investigation are; 1) the very low abundance transcripts (tag counts $\leq 4$) and 2) using other developmental SAGE data sets.

Statistical analysis of the SAGE Library comparison excluded genes with tag counts $\leq 4$. It was noted that this group of low abundance transcripts contained mesenchyme-expressed genes that are known to be important during prostatic development, including $BMP4$, $Fgf7$ and $Fgf10$. Therefore, it is proposed that this group of low abundance genes has not been adequately investigated and may also contain other important regulators of prostate organogenesis.

Additionally, comparison of the VSU and VMP Libraries with other SAGE libraries, such as those of the developing rat and human kidney (Dekel et al., 2002, Stuart et al., 2001) may highlight genes that are commonly expressed in branching organs. It is proposed that similarities in gene expression between developing organs may indicate similar regulatory mechanisms. These mechanisms have been studied more comprehensively in organs such as the kidney (as reviewed by Davies, 2002). Also, comparison of the rat VSU/VMP Libraries with the SAGE Libraries of the
developing mouse prostate (Zhang et al., 2006) may provide mutually supportive evidence of the genes and signalling pathways identified.

10.9.4 is LongSAGE analysis a robust, reliable and interesting approach to gene expression profiling?
The LongSAGE method was shown to be robust but not infallible. The contents of the Libraries were verified based on expected gene expression and the data was further verified using independent quantitative techniques. The ‘control’ transcripts identified supported the validity of the Library data. However, preparation of the Libraries was too expensive in terms of tissue, time and cost to repeat, and so the reproducibility of the Libraries has not been tested. The results of the LightCycler analysis screening suggest that quantification of the transcripts by LongSAGE was not always accurate, although it was noted that the trends in expression were upheld by qRT-PCR and Northern blotting analysis. It is recommended that in the future, a more cost-efficient method than LightCycler analysis be used to screen the candidate genes, in a high throughput manner. This would help to characterise the Library data more fully, ensuring that quantification of the transcripts were verified. Also, if gene expression levels were estimated by an independent technique, this could help to indicate when a SAGE library has been adequately sampled.

As an open-ended technique, the LongSAGE approach could be considered of high risk. In particular, the wealth of data generated has been cited as a possible weakness when searching for candidate genes. This was overcome by comparison of the VSU and VMP libraries and statistical analysis, which highlighted a manageable number of VMP-enriched candidate genes (P<0.05) (218). A similar process is commonly used to identify disease-related genes by comparison of normal and cancer cell/tissue SAGE profiles. The additional intuitive selection process in this project is similar to the candidate-based approach that has been used with success by multiple research groups.

It was concluded that any problems associated with the LongSAGE technique are outweighed by the enormous potential that an open-ended approach provides.
Furthermore, undertaking this project has been extremely interesting and informative, and the limited success presented in this thesis, has not fully explored the several avenues of possible further research that the VSU/VMP SAGE libraries represent.

10.10.5 What does the LongSAGE data tell us about prostate development and prostate cancer?

The mesenchyme/stroma is known to produce factors that regulate prostate development and neoplastic growth, but the mechanisms of action are not fully understood. The VSU/VMP Libraries were shown to contain developmental mesenchymal factors that are expressed in both the male and the female UGS (P0). It is proposed that this data is supportive of the hypothesis that constitutive expression of molecules in the VMP, are capable of regulating prostate development (as reviewed by Thomson et al., 2001). However, it also raises the question of how these factors are differentially regulated during development so that prostate growth is induced in the male and not the female. One mechanism proposed is the direct/indirect regulation of mesenchymal factors by androgens, but this project did not identify any factors as androgen-regulated. Consequently, it is suggested that the findings of the SAGE data may be supportive of an alternative mechanism proposed for regulating prostate induction, via sexually dimorphic smooth muscle expression (Thomson et al., 2002). In the absence of androgen-regulation, perhaps the physical barrier of the smooth muscle layer in the female UGS is responsible for preventing inductive factors signalling from the VMP to the epithelium. Furthermore, it is suggested that the possible importance of the smooth muscle layer during prostate development may be supported by the unusual distribution of Dlk1 and Notch2 in the SM, reported in this thesis.

The LongSAGE data has also provided evidence that the developing prostate expresses signalling pathways that are active in other branched organs i.e Fgfs, Tgfβ, MMPs, BMPs (as reviewed by Davies, 2002). Branching epithelia in different organs share a number of anatomical features, matrix requirements and signalling pathways, but they are by no means identical. A ‘consensus model’ of mammalian branching
epithelium has been reported by Davies, (2002), but not all branching systems make used of all the pathways in the ‘consensus set’. Epithelial branching in organs such as the kidney and lung are essential for viability, and attract greater research interest, while the factors and mechanisms that regulate branching morphogenesis in the prostate are poorly understood. The VSU and VMP libraries could be a valuable data source to investigate the factors that control branching in the prostate.

However, it is also recognised that as a result of using VSU and VMP tissues to make the SAGE Libraries, it is speculated that the genes identified in this project may only represent factors expressed during prostate induction, and that any changes to gene expression that take place epithelial branching could have been missed. A common mechanism for controlling branching epithelia has been proposed for the developing kidney and lung (as reviewed by Davies, 2002), where the tips of a branching tubule renders the mesenchyme incapable of supporting further branching after it has passed through it. Therefore, it is possible that the prostate mesenchyme may only act to regulate epithelial branching once the epithelium has invaded the virgin mesenchyme, and the epithelia may induce changes in the local mesenchyme as it passes through it. These local changes in gene expression in the mesenchyme would be absent from the SAGE libraries prepared.

Finally, the LongSAGE libraries and the subsequent characterisation of genes expressed in the developing rat prostate and the human CAF cells has supported the use of a developmental model when looking at neoplastic growth in the prostate, and the theory of developmental ‘awakening’ proposed by McNeal (1969). The LongSAGE libraries provided a useful source of data for comparison with prostate cancer, with the ability to highlight the mesenchymal/stromal compartment. Identifying the genes that are mis-expressed during prostate cancer, and understanding the factors that regulate their expression, could represent useful therapeutic targets.
Chapter 10

Discussion

10.10 Final Summary

The aim of this thesis was to identify mesenchyme specific/enriched genes that regulate normal and neoplastic growth. The LongSAGE technique was used to generate comprehensive transcriptome profiles of the VSU and VMP tissues micro-dissected from the female rat UGS (P0). The expression of ‘control’ genes within the Libraries confirmed the validity of the data. Candidate gene transcripts were selected from the Libraries for further studies either by bioinformatic analysis or intuitive selection. The genes selected had not previously been reported in association with the developing prostate.

Six candidate transcripts; Dlk1, Notch2, Ptn, Nell2, MMP2 and MMP14 were confirmed as showing VMP–enriched transcript expression in the female UGS (P0). Candidate transcript expression was also identified in the VP of the male UGS (P0), and the levels of expression were studied from embryonic to adult life. Each candidate demonstrated the most abundant expression during late embryogenesis and the perinatal period. Investigation of the candidate protein distribution confirmed the localisation of Dlk1, Notch2 and Ptn to the mesenchymal cells of the VMP and VP. Each protein also exhibited an individual distribution pattern in the UGM and SM layer of the UGS. An unusual and complementary pattern of Dlk1 and Notch2 expression in the smooth muscle layer, apparently in adjacent cells, suggested they represented a possible receptor–ligand relationship in the Notch signalling pathway. The significance of this is not yet understood. Ptn, on the other hand, demonstrated a strong association with the cell-surface of the epithelial cells in the VP, and a gradient of expression, which has been likened to a model of Ptn activity in the developing kidney. Consequently, it was speculated that the prostate and kidney might share similar regulatory mechanisms involving Ptn, during branching morphogenesis. Furthermore, the function of one of the candidates was successfully tested in on VP cultures in vitro. Addition of rhDLK1 to the cultures caused an increase in the size of the organ and the amount of epithelial branching in the VP. It is proposed that Dlk1 may also play a similar role in vivo. Finally, the possibility that developmental mesenchymal factors may be active during neoplastic development was studied using isolated human CAF and NPF cells. Four of the six
candidates were expressed in both cell types, but levels of NOTCH2 and PTN were found moderately downregulated in CAFs in comparison to their matched NPFs. PTN protein expression was also confirmed in the human foetal prostate.

In conclusion, this project has been successful in generating two LongSAGE Libraries, which were used for the identification of six mesenchymal genes, at least one of which has demonstrated the ability to regulate prostate growth in vitro. Furthermore, some of the developmental mesenchymal factors were also found expressed in prostate cancer stroma cells, and therefore may represent therapeutic targets, although this remains to be investigated in the long-term.
11. Bibliography


epithelial cells to express androgen-dependent seminal vesicle secretory proteins in chimeric tissue recombinants. *Endocrinology* 128: 3293-98.


of prostatic epithelial cells. *Cancer Res* 61: 5038-44.


mCytochem 45: 1-7.


cell lines corresponds to a differentiated chromaffin cell type. *Int J Cancer* **105**: 61-9.


## APPENDIX

### VMP-enriched genes (P<0.05) identified by LongSAGE analysis

**Genes selected for LightCycler analysis**

**Candidate genes selected after LightCycler analysis**

<table>
<thead>
<tr>
<th>Tag</th>
<th>VMP</th>
<th>VSU</th>
<th>fold</th>
<th>P-value</th>
<th>Unigene</th>
<th>Description</th>
<th>genomic BLAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATAACACAT AAAAAAAA</td>
<td>492</td>
<td>363</td>
<td>1.4</td>
<td>4.38E-02</td>
<td>11317</td>
<td>protein tyrosine phosphatase, non-receptor type 1</td>
<td>no hits, not 3' tag</td>
</tr>
<tr>
<td>GACTTTTGA AAACATTIT</td>
<td>370</td>
<td>139</td>
<td>2.7</td>
<td>4.67E-05</td>
<td>2953</td>
<td>collagen, type 1, alpha 1</td>
<td>unique hit</td>
</tr>
<tr>
<td>GGCTATGTGA AAATTTTC</td>
<td>362</td>
<td>258</td>
<td>1.4</td>
<td>3.25E-02</td>
<td>106103</td>
<td>decorin</td>
<td>unique hit</td>
</tr>
<tr>
<td>TTTGGTTTTT CAAAAAG</td>
<td>252</td>
<td>142</td>
<td>1.8</td>
<td>9.76E-05</td>
<td>107239</td>
<td>procollagen, type 1, alpha 2</td>
<td>unique hit</td>
</tr>
<tr>
<td>CGAACCTCTC ACGTGAT</td>
<td>236</td>
<td>142</td>
<td>1.7</td>
<td>1.16E-03</td>
<td>98989</td>
<td>secreted acidic cysteine rich glycoprotein</td>
<td>unique hit, also called osteonectin</td>
</tr>
<tr>
<td>TGTTCTCTCT GCCTTAA</td>
<td>158</td>
<td>83</td>
<td>1.9</td>
<td>3.67E-04</td>
<td>3247</td>
<td>collagen, type III, alpha 1</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGCTACCCT TCTAAAAA</td>
<td>151</td>
<td>91</td>
<td>1.7</td>
<td>9.69E-03</td>
<td>101464</td>
<td>Rattus norvegicus similar to gamma actin-like protein (LOC295810), mRNA</td>
<td>multiple hits</td>
</tr>
<tr>
<td>TGCTACCCT TCTAAAAA</td>
<td>151</td>
<td>91</td>
<td>1.7</td>
<td>9.69E-03</td>
<td>98160</td>
<td>smooth muscle alpha-actin</td>
<td>multiple hits</td>
</tr>
<tr>
<td>CATTTTCTG GCAAAATC</td>
<td>124</td>
<td>34</td>
<td>3.6</td>
<td>2.35E-05</td>
<td>964</td>
<td>insulin-like growth factor 2</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGGAATAATGA AAAAAAAA</td>
<td>108</td>
<td>51</td>
<td>2.1</td>
<td>4.85E-04</td>
<td>2953</td>
<td>collagen, type 1, alpha 1</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>TGGAATAATGA CCCCCAAA</td>
<td>105</td>
<td>52</td>
<td>2.0</td>
<td>1.34E-03</td>
<td>2953</td>
<td>collagen, type 1, alpha 1</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>GATCACCTG TTTTATAC</td>
<td>99</td>
<td>39</td>
<td>2.5</td>
<td>5.41E-05</td>
<td>17012</td>
<td>Rattus norvegicus similar to microfibril-associated glycoprotein 1 (LOC313662), mRNA</td>
<td>similar to microfibril-associated glycoprotein 1</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>-------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>TAGAATTCA</td>
<td>97</td>
<td>57</td>
<td>1.7</td>
<td>2.71E-02</td>
<td>1046</td>
<td>transforming growth factor, beta induced, 68 kDa</td>
<td>unique hit, close to novel transcript with homology to TGF beta induced 68kDa (betaIG-H3)</td>
</tr>
<tr>
<td>GATTCCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGAGAGGCA</td>
<td>92</td>
<td>55</td>
<td>1.7</td>
<td>3.93E-02</td>
<td>12328</td>
<td>Rattus norvegicus similar to RIKEN cDNA 1110007F23 (LOC287382), mRNA</td>
<td>unique hit close to novel transcript with similarity to mRIKEN cDNA 1110007F23 (= MFAP4: microfibrillar associated protein 4)</td>
</tr>
<tr>
<td>GCAGAGAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGGAAATGA</td>
<td>88</td>
<td>40</td>
<td>2.2</td>
<td>8.99E-04</td>
<td>2953</td>
<td>collagen, type 1, alpha 1</td>
<td>unique hit - collagen, type 1, alpha 1</td>
</tr>
<tr>
<td>CCAAACGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCCAATGGC</td>
<td>85</td>
<td>44</td>
<td>1.9</td>
<td>7.27E-03</td>
<td>11889</td>
<td>Rattus norvegicus similar to procollagen, type VI, alpha 2 (LOC361821), mRNA</td>
<td>multiple hits</td>
</tr>
<tr>
<td>CCAATAAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACATTTC</td>
<td>84</td>
<td>43</td>
<td>2.0</td>
<td>6.52E-03</td>
<td>8527</td>
<td>Rattus norvegicus transcribed sequence with strong similarity to protein pir:S55054 (H.sapiens) S55054 Sm protein G (small nuclear ribonucleoprotein polypeptide G)</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>ATATGAAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCCTGAGTC</td>
<td>78</td>
<td>34</td>
<td>2.3</td>
<td>9.81E-04</td>
<td>94978</td>
<td>actin, beta</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>CACACCGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCGTGGTC</td>
<td>78</td>
<td>34</td>
<td>2.3</td>
<td>9.81E-04</td>
<td>115044</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pir:S55054 (H.sapiens) S55054 Sm protein G (small nuclear ribonucleoprotein polypeptide G)</td>
<td>unique hit no assigned rat locus within 5kb</td>
</tr>
<tr>
<td>CACACCAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACAAAAAAAAA</td>
<td>77</td>
<td>39</td>
<td>2.0</td>
<td>8.03E-03</td>
<td>47029</td>
<td>SEC61, alpha subunit (S. cerevisiae)</td>
<td>unique hit - Sec61 alpha-1</td>
</tr>
<tr>
<td>CCAACCTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGAATGGGC</td>
<td>69</td>
<td>33</td>
<td>2.1</td>
<td>6.09E-03</td>
<td>6422</td>
<td>matrix metalloproteinase 2 (72 KDa type IV collagenase)</td>
<td>unique hit</td>
</tr>
<tr>
<td>AAAAAAGTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCCGCGG</td>
<td>61</td>
<td>33</td>
<td>1.8</td>
<td>3.60E-02</td>
<td></td>
<td></td>
<td>multiple random hits</td>
</tr>
<tr>
<td>GCAGCTTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCTCACCAA</td>
<td>55</td>
<td>24</td>
<td>2.3</td>
<td>5.68E-03</td>
<td>54384</td>
<td>elastin</td>
<td>hit - elastin</td>
</tr>
<tr>
<td>CCCACCTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>TCTCACAA CCCCCACCT</td>
<td>55</td>
<td>24</td>
<td>2.3</td>
<td>5.68E-03</td>
<td>11372</td>
<td>lysyl oxidase</td>
<td></td>
</tr>
<tr>
<td>ATGTAGTAG TGTCITTAC</td>
<td>54</td>
<td>20</td>
<td>2.7</td>
<td>1.14E-03</td>
<td>33807</td>
<td>Rattus norvegicus similar to CCT (chaperonin containing TCP-1) zeta subunit (LOC288620), mRNA</td>
<td>unique hit, close to heterogenous ribonucoprotein D0 (HNRNP D0)</td>
</tr>
<tr>
<td>ATGTAGTAG TGTCITTAC</td>
<td>54</td>
<td>20</td>
<td>2.7</td>
<td>1.14E-03</td>
<td>94022</td>
<td>RNA binding protein p45AUFI</td>
<td></td>
</tr>
<tr>
<td>CTAATAAAT GTTTTTGG</td>
<td>53</td>
<td>26</td>
<td>2.0</td>
<td>2.07E-02</td>
<td>102247</td>
<td>Rattus norvegicus similar to DREV protein (LOC300672), mRNA</td>
<td>similar to DREV protein</td>
</tr>
<tr>
<td>CTAATAAAT GTTTTTGG</td>
<td>53</td>
<td>26</td>
<td>2.0</td>
<td>2.07E-02</td>
<td>109804</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, part of mDREV protein</td>
</tr>
<tr>
<td>TGTATAAAA ATAAAAAA</td>
<td>51</td>
<td>19</td>
<td>2.7</td>
<td>1.66E-03</td>
<td>101146</td>
<td>Rattus norvegicus similar to Endoplasmin precursor (Endoplasmic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorphic tumor rejection antigen 1) (Tumor rejection antigen gp96) (LOC362862), mRNA</td>
<td>unique hit- wrong direction- close to named gene</td>
</tr>
<tr>
<td>CCTAGCCCC CCCCCACCA</td>
<td>49</td>
<td>15</td>
<td>3.3</td>
<td>3.22E-04</td>
<td>7961</td>
<td>Rattus norvegicus similar to latent transforming growth factor-beta binding protein 4S (LOC292734), mRNA</td>
<td>unique hit, close to novel transcript with homology to mLTbp4 short splice variant</td>
</tr>
<tr>
<td>TTTGGTTTTTC CAAAAAAA</td>
<td>49</td>
<td>20</td>
<td>2.5</td>
<td>4.85E-03</td>
<td>107239</td>
<td>procollagen, type I, alpha 2</td>
<td>multiple hits</td>
</tr>
<tr>
<td>ATATAATGA ATAATAAT</td>
<td>38</td>
<td>13</td>
<td>2.9</td>
<td>3.47E-03</td>
<td>14547</td>
<td>delta-like 1 homolog (Drosophila)</td>
<td>unique hit</td>
</tr>
<tr>
<td>TTCAATTAA AATCAGTT</td>
<td>37</td>
<td>14</td>
<td>2.6</td>
<td>8.25E-03</td>
<td>19642</td>
<td>Rattus norvegicus similar to Iκ protein (LOC291659), mRNA</td>
<td>unique hit, close to novel transcript with homology to Iκ protein (RER)</td>
</tr>
<tr>
<td>CACTCTATT AAAATAAA</td>
<td>34</td>
<td>13</td>
<td>2.6</td>
<td>1.22E-02</td>
<td>2862</td>
<td>calpain 6</td>
<td>unique hit</td>
</tr>
<tr>
<td>TATTCATGT GAAATATT</td>
<td>33</td>
<td>13</td>
<td>2.5</td>
<td>1.65E-02</td>
<td>3313</td>
<td>adenyl cyclase 6</td>
<td>unique hit</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>AAATAAAGT CTTGAAAAC</td>
<td>29</td>
<td>10</td>
<td>2.9</td>
<td>1.12E-02</td>
<td>2163</td>
<td>agrin</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGGAAAAACA CCAAATAA</td>
<td>28</td>
<td>7</td>
<td>4</td>
<td>2.12E-03</td>
<td>22352</td>
<td>Rattus norvegicus similar to gamma-filamin (LOC362332), mRNA</td>
<td>unique hit- similar to gamma-filamin</td>
</tr>
<tr>
<td>TTTTAAAAT AAACAAATT</td>
<td>27</td>
<td>7</td>
<td>3.9</td>
<td>3.08E-03</td>
<td>37666</td>
<td>Rattus norvegicus similar to tweety homolog 2 (LOC304315), mRNA, RIKENcDNA 2900029G13</td>
<td>unique hit close to novel transcript with tweety homology domain</td>
</tr>
<tr>
<td>GTAATCCCTG CTCAGTAC</td>
<td>25</td>
<td>9</td>
<td>2.8</td>
<td>2.31E-02</td>
<td></td>
<td></td>
<td>no hit in rat genome, PRESENT IN DIFF PATENT SEQ</td>
</tr>
<tr>
<td>AATAAACAT TTATTATT</td>
<td>25</td>
<td>10</td>
<td>2.5</td>
<td>3.98E-02</td>
<td>2816</td>
<td>Ras-related GTP-binding protein ragA</td>
<td>unique hit</td>
</tr>
<tr>
<td>GCCCTCCAG TAATTATT</td>
<td>23</td>
<td>6</td>
<td>3.8</td>
<td>6.48E-03</td>
<td></td>
<td></td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>TAACAGAAG TCCAGAAG</td>
<td>23</td>
<td>6</td>
<td>3.8</td>
<td>6.48E-03</td>
<td>2875</td>
<td>collagen, type V, alpha 2</td>
<td>unique hit- wrong direction - close to collagen, type V, alpha 2 (Col5a2), mRNA</td>
</tr>
<tr>
<td>TATAACGAG CCAAATTC</td>
<td>22</td>
<td>8</td>
<td>2.8</td>
<td>3.47E-02</td>
<td>54261</td>
<td>four and a half LIM domains 1</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>TGAAGGCCC CACTGAAC</td>
<td>22</td>
<td>8</td>
<td>2.8</td>
<td>3.47E-02</td>
<td>9073</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pdb:1LBG (E. coli) B Chain B, Lactose Operon Repressor Bound To 21-Base Pair Symmetric Operator Dna, Alpha Carbons Only</td>
<td>unique hit, close to novel transcript with similarity to NIPSNAP2 protein (glioblastoma amplified sequence</td>
</tr>
<tr>
<td>CCTCGGGGG CACCTAGA</td>
<td>21</td>
<td>5</td>
<td>4.2</td>
<td>6.39E-03</td>
<td></td>
<td></td>
<td>unique hit, close to mouse fibulin gene</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>AATAAATGA</td>
<td>21</td>
<td>7</td>
<td>3</td>
<td>2.66E-02</td>
<td>8685</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein ref:NP_001488.2 (H.sapiens) UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase 1, membrane-bound form; beta-N-acetylglucosaminyl-glycopeptide beta-1,4-galactosyltransferase</td>
<td>unique hit, 2.5kb from N-acetylglucosamine galactosyltransferase</td>
</tr>
<tr>
<td>CCCTATCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAGCTGGTT</td>
<td>20</td>
<td>6</td>
<td>3.3</td>
<td>1.96E-02</td>
<td>3437</td>
<td>small nuclear ribonucleoprotein polypeptides B and B1</td>
<td>Small nuclear ribonucleoprotein associated protein B (snRNP-B) (Sm protein B) (Sm-B) (SmB) (SM11) (Fragment).</td>
</tr>
<tr>
<td>TATAATAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit</td>
</tr>
<tr>
<td>CATCAAGCC</td>
<td>19</td>
<td>6</td>
<td>3.2</td>
<td>2.81E-02</td>
<td>106103</td>
<td>decorin</td>
<td>unique hit</td>
</tr>
<tr>
<td>AAATATAAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGATGTACT</td>
<td>18</td>
<td>3</td>
<td>6</td>
<td>3.67E-03</td>
<td>7055</td>
<td>Rattus norvegicus similar to VE-cadherin (LOC307618), mRNA</td>
<td>unique hit- no assigned gene locus within 5kb</td>
</tr>
<tr>
<td>GTACCTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCTGACTT</td>
<td>18</td>
<td>5</td>
<td>3.6</td>
<td>2.01E-02</td>
<td>2269</td>
<td>basigin</td>
<td>unique hit</td>
</tr>
<tr>
<td>ACCCTGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCCTAGCCA</td>
<td>18</td>
<td>5</td>
<td>3.6</td>
<td>2.01E-02</td>
<td>91417</td>
<td>Rattus norvegicus similar to microtubule-associated protein 4 (LOC367171), mRNA</td>
<td>unique hit, close to novel transcript with homology to MAP4</td>
</tr>
<tr>
<td>TTTCTGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAAAAAGTT</td>
<td>18</td>
<td>6</td>
<td>3</td>
<td>4.01E-02</td>
<td>2517</td>
<td>Rattus norvegicus similar to LIM domain only 4; LIM only 4; ethanol induced 4 (LOC362051), mRNA</td>
<td>no hits</td>
</tr>
<tr>
<td>AAAAAAATAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTGTTGATC</td>
<td>17</td>
<td>3</td>
<td>5.7</td>
<td>5.58E-03</td>
<td>774</td>
<td>laminin, beta 2</td>
<td>unique hit</td>
</tr>
<tr>
<td>GCACATTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCCCTGGGT</td>
<td>16</td>
<td>3</td>
<td>5.3</td>
<td>8.48E-03</td>
<td>102346</td>
<td>Rattus norvegicus similar to hypothetical protein 5730443G10 (LOC303260), mRNA</td>
<td>unique hit, close to novel transcript with homology to sperm associated antigen</td>
</tr>
<tr>
<td>TACTTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTGGGGCAG</td>
<td>16</td>
<td>5</td>
<td>3.2</td>
<td>4.24E-02</td>
<td>8450</td>
<td>Rattus norvegicus similar to integral membrane protein 2C (LOC301575), mRNA</td>
<td>unique hit, close to novel transcript with similarity to Mus musculus integral membrane protein 2C (Itm2c)</td>
</tr>
<tr>
<td>CCGTATAGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTTTACCTGA</td>
<td>16</td>
<td>5</td>
<td>3.2</td>
<td>4.24E-02</td>
<td>8450</td>
<td>Rattus norvegicus similar to integral membrane protein 2C (LOC301575), mRNA</td>
<td>unique hit, close to novel transcript with similarity to Mus musculus integral membrane protein 2C (Itm2c)</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ATGACTGAG CCCCTAAT</td>
<td>15</td>
<td>2</td>
<td>7.5</td>
<td>4.76E-03</td>
<td>2595</td>
<td>Rattus norvegicus similar to divalent cation tolerant protein CUTA (LOC294288), mRNA</td>
<td>unique hit- in wrong direction</td>
</tr>
<tr>
<td>CACTTCCCT CGGGGTTT</td>
<td>15</td>
<td>3</td>
<td>5</td>
<td>1.29E-02</td>
<td>106701</td>
<td>Rattus norvegicus similar to Ierepo4-pending protein (LOC304839), mRNA</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>GATGAAGAC TGCCCTGTT</td>
<td>15</td>
<td>3</td>
<td>5</td>
<td>1.29E-02</td>
<td>117705</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>GATGAAGAC TGCCCTGTT</td>
<td>15</td>
<td>3</td>
<td>5</td>
<td>1.29E-02</td>
<td>8538</td>
<td>splicing factor, arginine/serine-rich (transformer 2 Drosophila homolog) 10</td>
<td>unique hit- no loci within 5kb</td>
</tr>
<tr>
<td>GCCCGTGCC ACCCTTC</td>
<td>15</td>
<td>3</td>
<td>5</td>
<td>1.29E-02</td>
<td>1226</td>
<td>Rattus norvegicus similar to GAMM1 protein (LOC300258), mRNA</td>
<td>unique hit, close to novel transcript with homology to GAMM1 (MYG1)</td>
</tr>
<tr>
<td>TGTATTACA GTGGATTT</td>
<td>15</td>
<td>3</td>
<td>5</td>
<td>1.29E-02</td>
<td>7967</td>
<td>Rattus norvegicus similar to SON protein (LOC304092), mRNA</td>
<td>Unique hit, close to novel transcript belonging to the SON protein family</td>
</tr>
<tr>
<td>TGTTAGAGT CAATCTCG</td>
<td>15</td>
<td>3</td>
<td>5</td>
<td>1.29E-02</td>
<td>16968</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1.5kb from retinol dehydrogenase10</td>
</tr>
<tr>
<td>TAAATATCT GTGTCCCC</td>
<td>15</td>
<td>4</td>
<td>3.75</td>
<td>2.98E-02</td>
<td>29232</td>
<td>step II splicing factor SLU7</td>
<td>no exact hit in rat genome</td>
</tr>
<tr>
<td>GGGATTTGT GGGGGCGAG</td>
<td>14</td>
<td>3</td>
<td>4.67</td>
<td>1.95E-02</td>
<td>2458</td>
<td>tubulin, beta 5</td>
<td>unique hit</td>
</tr>
<tr>
<td>GATAAAAGCA AGATTTCC</td>
<td>14</td>
<td>4</td>
<td>3.5</td>
<td>4.40E-02</td>
<td>120904</td>
<td>Rattus norvegicus transcribed sequence with weak similarity to protein ref:NP_116255.1 (H.sapiens) hypothetical protein FLJ14957 [Homo sapiens]</td>
<td>unique hit, close to novel transcript with homology to hFLJ14957</td>
</tr>
<tr>
<td>GGTACCTTA CTTCTCTC</td>
<td>14</td>
<td>4</td>
<td>3.5</td>
<td>4.40E-02</td>
<td>3552</td>
<td>Rattus norvegicus similar to CG14967-PA (LOC303280), mRNA</td>
<td>unique hit, close to novel transcript with homology to mRIKEN cDNA2610507B11 (= KIAA0100 gene product)</td>
</tr>
<tr>
<td>GTCATTTTCT AGTTTTT</td>
<td>14</td>
<td>4</td>
<td>3.5</td>
<td>4.40E-02</td>
<td>10600</td>
<td>transcription factor 8</td>
<td>unique hit</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------------</td>
<td>---------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
</tr>
<tr>
<td>GTTTGTACA AATAATAC</td>
<td>14</td>
<td>4</td>
<td>3.5</td>
<td>4.40E-02</td>
<td>37338</td>
<td>latent transforming growth factor beta binding protein 3</td>
<td>unique hit</td>
</tr>
<tr>
<td>TTTCCAGCAG TGTAAGATC</td>
<td>14</td>
<td>4</td>
<td>3.5</td>
<td>4.40E-02</td>
<td>47</td>
<td>hypoxanthine guanine phosphoribosyl transferase</td>
<td>unique hit</td>
</tr>
<tr>
<td>AACCACAGA GATGGTTG</td>
<td>13</td>
<td>1</td>
<td>13</td>
<td>3.63E-03</td>
<td>23810</td>
<td>casein kinase 1, alpha 1</td>
<td>unique hit</td>
</tr>
<tr>
<td>TATTTCAG CATTTCG</td>
<td>13</td>
<td>2</td>
<td>6.5</td>
<td>1.14E-02</td>
<td>2744</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>TACACTCCC CCTACTCC</td>
<td>13</td>
<td>3</td>
<td>4.33</td>
<td>2.94E-02</td>
<td>8628</td>
<td>protocadherin gamma subfamily C, 3</td>
<td>only 1 EST seq to confirm =&gt; more likely to be gluco reg kinase tag</td>
</tr>
<tr>
<td>TACACTCCC CCTACTCC</td>
<td>13</td>
<td>3</td>
<td>4.33</td>
<td>2.94E-02</td>
<td>4636</td>
<td>serum/glucocorticoid regulated kinase</td>
<td>unique hit</td>
</tr>
<tr>
<td>GATGAATGT TATATGTT</td>
<td>12</td>
<td>2</td>
<td>6</td>
<td>1.77E-02</td>
<td>18947</td>
<td>fibronectin</td>
<td>unique hit, 2kb from mRIKEN cDNA1200009022, fibronectin like containing structures</td>
</tr>
<tr>
<td>TTCTTCAGCTT TCCAGTA</td>
<td>12</td>
<td>2</td>
<td>6</td>
<td>1.77E-02</td>
<td>18947</td>
<td>paralemmin</td>
<td>unique hit</td>
</tr>
<tr>
<td>TAAACGTIT TCAGGGAA</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>4.42E-02</td>
<td>69085</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pir:139175 (H.sapiens) I39175 SH2-domain protein Grb-IR</td>
<td>unique hit, 2.5kb from mGrb10 locus (growth factor receptor binding protein)</td>
</tr>
<tr>
<td>TGACAGTGT TCATTAAC</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>4.42E-02</td>
<td>3882</td>
<td>Rattus norvegicus similar to cleavage and polyadenylation specific factor 5; cleavage and polyadenylation specific factor 5, 25 kD subunit (LOC291877), mRNA</td>
<td>unique hit-similar to cleavage and polyadenylation specific factor 5</td>
</tr>
<tr>
<td>TTAATGTAT CGTCTTAA</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>4.42E-02</td>
<td>3622</td>
<td>hypothetical protein CDA08</td>
<td>unique hit-close to T-cell immunomodulatory protein (TIP protein) (CDA08-like protein)</td>
</tr>
<tr>
<td>TTAATGTAT CGTCTTAA</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>4.42E-02</td>
<td>115615</td>
<td>Rattus norvegicus transcribed sequence</td>
<td>unique hit, part of RIKEN cDNA 2310047C21 (= CDA08 protein, TIP protein, predicted membrane protein)</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>----------</td>
<td>----------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>AAGCTAAAC</td>
<td>11</td>
<td>1</td>
<td>11</td>
<td>9.05E-03</td>
<td>106046</td>
<td>Rattus norvegicus similar to RIKEN cDNA 2310047013 (LOC291320), mRNA</td>
<td>unique hit, close to novel transcript with homology to mRIKEN cDNA 2310047013 (BRAIN MY042, FLJ133397)?</td>
</tr>
<tr>
<td>GCTGGTAT</td>
<td>11</td>
<td>1</td>
<td>11</td>
<td>9.05E-03</td>
<td>79423</td>
<td>Rattus norvegicus similar to Wbscr1 (LOC288599), mRNA</td>
<td>unique hit, close to novel transcript with similarity to Wbscr1 gen</td>
</tr>
<tr>
<td>CATTGCGTG</td>
<td>11</td>
<td>1</td>
<td>11</td>
<td>9.05E-03</td>
<td>8257</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2kb from novel transcript (mSemaphorin6D)</td>
</tr>
<tr>
<td>TTAGGTAAT</td>
<td>11</td>
<td>1</td>
<td>11</td>
<td>9.05E-03</td>
<td></td>
<td></td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>GGTGAAAC</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>96446</td>
<td>Rattus norvegicus similar to RIKEN cDNA 5330414D10 (LOC311726), mRNA</td>
<td>unique hit, 1kb from novel transcript</td>
</tr>
<tr>
<td>AATAAAGT</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>4309</td>
<td>syntenin</td>
<td>Syntenin 1 (Syndecan binding protein 1)</td>
</tr>
<tr>
<td>ATACACGTTG</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>23941</td>
<td>dynactin 4</td>
<td>only 1 EST seq to confirm =&gt; more likely to be calnexin tag</td>
</tr>
<tr>
<td>GTCTTTAA</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>19707</td>
<td>Rattus norvegicus similar to importin 9 (LOC304817), mRNA</td>
<td>unique hit- wrong direction - similar to importin 9</td>
</tr>
<tr>
<td>TATTTTTCTA</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>29900</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1kb from locus with similarity to Cbx1 (heterochromatin protein P25)</td>
</tr>
<tr>
<td>GGGGACAGT</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>16536</td>
<td>Rattus norvegicus transcribed sequence with strong similarity to protein pdb:1LBG (E. coli) B Chain B, Lactose Operon Repressor Bound To 21-Base Pair Symmetric Operator Dna, Alpha Carbons Only</td>
<td>unique hit, no assigned loci within 5kb</td>
</tr>
<tr>
<td>TTTCAGCAT</td>
<td>11</td>
<td>2</td>
<td>5.5</td>
<td>2.73E-02</td>
<td>101054</td>
<td>Rattus norvegicus similar to RAN binding protein 3 isoform RANBP3-d; RAN-binding protein-3 (LOC316126), mRNA</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>ACATTTTAT</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>3.60E-03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>Fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>TGTTTGGA ACATTTAT</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>3.60E-03</td>
<td>110951</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pir:T08778 (H.sapiens) T08778 hypothetical protein DKFZp58611520.1</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>CATCATTCC GTGGCTGA</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>1.43E-02</td>
<td>22555</td>
<td>Rattus norvegicus similar to KIAA0368 (LOC313196), mRNA</td>
<td>unique hit, close to novel transcript with homology to m AI314180 expressed sequence</td>
</tr>
<tr>
<td>TAAGAGCTC TITCCATC</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>1.43E-02</td>
<td>8672</td>
<td>Rattus norvegicus similar to hypothetical protein, estradiol-induced (LOC308843), mRNA</td>
<td>unique hit, close to novel transcript, with similarity to h estradiol-induced (E2IG4) - NM_015516</td>
</tr>
<tr>
<td>TAGACTGCC TCACAATA</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>1.43E-02</td>
<td>108352</td>
<td>Rattus norvegicus similar to Multifunctional protein ADE2 (LOC288115), mRNA</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>TTAGACACG TGTCCTTT</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>1.43E-02</td>
<td>55771</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pir:160307 (E. coli) 160307 beta-galactosidase, alpha peptide</td>
<td>unique hit, novel transcript in neighbourhood but not clear if tag is part of it...</td>
</tr>
<tr>
<td>AGTTAAGCA TACTGTTT</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>24871</td>
<td>phosphatidylinositol binding clathrin assembly protein</td>
<td>multiple hits</td>
</tr>
<tr>
<td>ATGTCTGTT CTATATA</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>101845</td>
<td>Rattus norvegicus similar to entrophrin; epsin 4; clathrin interacting protein localized in the trans-Golgi region (LOC317516), mRNA</td>
<td>unique hit- no assigned gene locus within 5kb</td>
</tr>
<tr>
<td>GTGTGTGGGT GCCCCCAG</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>3222</td>
<td>Rattus norvegicus transcribed sequence with weak similarity to protein pir:T46432 (H.sapiens) T46432 hypothetical protein DKFZp434D2426.1</td>
<td>unique hit close to novel transcript with similarity to prion protein interacting protein 1</td>
</tr>
<tr>
<td>TACATAGGA TTGTTTGC</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>1476</td>
<td>peripheral myelin protein 22</td>
<td>unique hit</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----------</td>
<td>---------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>TCAGAATTI</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>94979</td>
<td>Rattus norvegicus similar to Hypothetical protein FLJ25422 (LOC294789), mRNA</td>
<td>unique hit, close to novel transcript with similarity to human hyp prot FLJ25422</td>
</tr>
<tr>
<td>AGCCCATAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCTGGCAGT</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>79893</td>
<td>Rattus norvegicus similar to karyopherin beta 3; Ran_GTP binding protein 5; importin beta-3 subunit (LOC306182), mRNA</td>
<td>unique hit- close to named transcript</td>
</tr>
<tr>
<td>CTCCCTTCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCTGGCTCC</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>17033</td>
<td>Rattus norvegicus similar to RNA binding motif protein 5 (LOC300996), mRNA</td>
<td>unique hit- wrong direction - similar to RNA binding motif protein 5</td>
</tr>
<tr>
<td>TTTACAAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGGGCATAG</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>41133</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2kb from novel transcript with similarity to mSorC2 precursor (VPS10 domain containing receptor)</td>
</tr>
<tr>
<td>CTGAGGTTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTATTTGTATG</td>
<td>10</td>
<td>2</td>
<td>5</td>
<td>4.22E-02</td>
<td>10622</td>
<td>Rattus norvegicus similar to Murine homolog of human ftp-3 (LOC308650), mRNA</td>
<td>unique hit, close to novel transcript with homology to murine heterogeneous nuclear ribonucleoprotein H2 (ftp3)</td>
</tr>
<tr>
<td>GGAATTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATGTCACAA</td>
<td>9</td>
<td>0</td>
<td></td>
<td>5.74E-03</td>
<td>12759</td>
<td>fibrillin-1</td>
<td>no hit, no poly A</td>
</tr>
<tr>
<td>TAAAGTGGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACGCTGTAC</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>20981</td>
<td>phosphodiesterase 9A</td>
<td>unique hit</td>
</tr>
<tr>
<td>AGAATTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTGTGTGTTG</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>101908</td>
<td>Rattus norvegicus similar to hypothetical protein DJ667H12.2 (LOC360899), mRNA</td>
<td>unique hit, close to novel transcript with similarity to hyp protein DJ667H12.2</td>
</tr>
<tr>
<td>TAATATAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATGTCATG</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>115909</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>TTITAAACT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATGTGTATG</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>2560</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>2 genomic hits, no transcripts in neighbourhood</td>
</tr>
<tr>
<td>TTITAAACT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGTGTGTTG</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>95170</td>
<td>Rattus norvegicus similar to EIB-55kDa associated protein 5 (LOC361522), mRNA</td>
<td>unique hit, close to novel transcript with homology to E1B-55KDA-associate protueb 5 isoform A</td>
</tr>
<tr>
<td>GTAAATTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGAGCTGGC</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>41227</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, close to novel transcript with homology to ADAMTS-1 precursor</td>
</tr>
<tr>
<td>TGCTGAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>GTCAAGGTTCTGTGCT</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>9517</td>
<td>nucleolar phosphoprotein p130</td>
<td>unique hit</td>
</tr>
<tr>
<td>TACATACGTACATAAAT</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>70655</td>
<td>Rattus norvegicus similar to Probable ATP-dependent RNA helicase DDX27 (DEAD-box protein 27) (LOC362274), mRNA</td>
<td>unique hit, close to DDX27</td>
</tr>
<tr>
<td>TATTAACCTCCTCCGGA</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>22808</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1.5kb from novel transcript with homology to hARHGEF17</td>
</tr>
<tr>
<td>TCTGTCTTTATTACCTT</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>28161</td>
<td>fatty acid desaturase 1</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGACTGCTGCTTCAATA</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td></td>
<td></td>
<td>unique hit, no assigned locus in same direction</td>
</tr>
<tr>
<td>TTTAGTGACTTTACAT</td>
<td>9</td>
<td>1</td>
<td>9</td>
<td>2.28E-02</td>
<td>94887</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 5kb from novel transcript with similarity with KLF13</td>
</tr>
<tr>
<td>AGTTGGAAAAGGGGAAT</td>
<td>8</td>
<td>0</td>
<td></td>
<td>9.21E-03</td>
<td>119131</td>
<td>Rattus norvegicus transcribed sequence with strong similarity to protein sp: P00722 (E. coli) BGAL_ECOLI Beta-galactosidase (Lactase)</td>
<td>unique hit, no assigned locus in neighbourhhood</td>
</tr>
<tr>
<td>GTGTGGTCCCAGGGCAT</td>
<td>8</td>
<td>0</td>
<td></td>
<td>9.21E-03</td>
<td>6171</td>
<td>H19 fetal liver mRNA</td>
<td>unique hit</td>
</tr>
<tr>
<td>AATAAACTTGCTTTCC</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>28050</td>
<td>Rattus norvegicus similar to pre B-cell leukemia transcription factor 3 (LOC311876), mRNA</td>
<td>unique hit, close to Pbx1 (or 3?)</td>
</tr>
<tr>
<td>CACACACAAACAACACAC</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>24541</td>
<td>Rattus norvegicus similar to old astrocyte specifically induced substance (LOC362165), mRNA</td>
<td>100s of genomic hits</td>
</tr>
<tr>
<td>CATACGCTCAAAAGAA</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td></td>
<td></td>
<td>unique hit, novel transcript at 5kb, based on mouse data, tag comes from ribophorinI</td>
</tr>
<tr>
<td>CTTAGGTGCTGTCTGTA</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>30223</td>
<td>Rattus norvegicus similar to RIKEN cDNA 0610027018 (LOC363171), mRNA</td>
<td>Unique hit close to rat transcript -XP_343511.1</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>CTGCAGCAA</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>10293</td>
<td>low density lipoprotein receptor-related protein associated protein 1</td>
<td></td>
</tr>
<tr>
<td>AAGTTTAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit</td>
<td></td>
</tr>
<tr>
<td>GAATCTCAT</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>1878</td>
<td>Rattus norvegicus transcribed sequences</td>
<td></td>
</tr>
<tr>
<td>ACCTAAACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, no assigned locus within 5kb</td>
<td></td>
</tr>
<tr>
<td>GCAATCATT</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>1712</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein prf:2018199A (E. coli) 2018199A beta lactamase IRT-4 [Escherichia coli]</td>
<td></td>
</tr>
<tr>
<td>TCATACTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, 3.5kb from novel transcript with homology to hZimp10</td>
<td></td>
</tr>
<tr>
<td>TATACAGAG</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>9547</td>
<td>Rattus norvegicus transcribed sequences</td>
<td></td>
</tr>
<tr>
<td>CGTCATCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, 5kb from novel transcript with similarity to putative ankyrin repeat containing prot.3526402J09Rik</td>
<td></td>
</tr>
<tr>
<td>TCTGAATAT</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>22787</td>
<td>Rattus norvegicus similar to sprouty 1 (LOC294981), mRNA</td>
<td></td>
</tr>
<tr>
<td>AACATATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, chr2, close to novel transcript with homology msprouty1</td>
<td></td>
</tr>
<tr>
<td>TCTGTCCTG</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>34792</td>
<td>degenerative spermatocyte homolog (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>CTGCAAGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit</td>
<td></td>
</tr>
<tr>
<td>TGGGGAGAG</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>101799</td>
<td>Rattus norvegicus Ab2-450 mRNA, complete cds</td>
<td></td>
</tr>
<tr>
<td>GAATAGCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, close to MAN2B2 locus (Mannosidase alpha class 2B member 2)</td>
<td></td>
</tr>
<tr>
<td>GAAATAGCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, overlaps with mouse dip3 locus, dip13)</td>
<td></td>
</tr>
<tr>
<td>TATAATCACAAAG</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>1100</td>
<td>Rattus norvegicus similar to Pon2 protein (LOC296851), mRNA</td>
<td></td>
</tr>
<tr>
<td>TTCTGCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, close to novel transcript with similarity to Pon2 9 (serum paraxonase/arylesterase 2)</td>
<td></td>
</tr>
<tr>
<td>TTAATTTCTC</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>3.63E-02</td>
<td>17950</td>
<td>degenerative spermatocyte homolog (Drosophila)</td>
<td></td>
</tr>
<tr>
<td>ATAAAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit</td>
<td></td>
</tr>
<tr>
<td>ATTGTATCTGT</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>1.48E-02</td>
<td>50150</td>
<td>Rattus norvegicus similar to coat protein gamma-cop (LOC297428), mRNA</td>
<td></td>
</tr>
<tr>
<td>TTCTGCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>multiple genomic hits</td>
<td></td>
</tr>
<tr>
<td>ATTGTTCT</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>1.48E-02</td>
<td>36822</td>
<td>Rattus norvegicus transcribed sequences</td>
<td></td>
</tr>
<tr>
<td>TTCTGCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, no assigned locus in neigbourhood</td>
<td></td>
</tr>
<tr>
<td>CAGCATAAA</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>1.48E-02</td>
<td>7015</td>
<td>Rattus norvegicus similar to DIP13 beta (LOC362860), mRNA</td>
<td></td>
</tr>
<tr>
<td>TTAATTTGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit, overlaps with mouse dip3 locus, dip13)</td>
<td></td>
</tr>
<tr>
<td>CCCCCAAAATT</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>1.48E-02</td>
<td>12939</td>
<td>vesicle-associated membrane protein 2</td>
<td></td>
</tr>
<tr>
<td>CTGTGCGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>unique hit</td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CCCCCAAATT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>114545</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pdb:1LBG (E. coli) B Chain B, Lactose Operon Repressor Bound To 21-Base Pair Symmetric Operator Dna, Alpha Carbons Only</td>
<td>unique hit, close to VAM-2 locus (Vesicle associated membrane protein)</td>
</tr>
<tr>
<td>CTGTGGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCCTATGTTT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>84878</td>
<td>Rattus norvegicus similar to vascular endothelial zinc finger 1 (LOC287615), mRNA</td>
<td>unique hit, close to novel transcript with homology to vascular endothelial zinc finger 1 (= ZnF 161)</td>
</tr>
<tr>
<td>GTGCCCTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCCTATGTTT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>23175</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein pir:RGECDW (E. coli) RGECDW transcription activator of D-serine dehydratase</td>
<td>unique hit, close to novel transcript with similarity to mATP sulfurylase (PAPssl)</td>
</tr>
<tr>
<td>GTGCCCTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTGAGGTAT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>3507</td>
<td>Rattus norvegicus similar to ATP sulfurylase/APS kinase (LOC295443), mRNA</td>
<td>unique hit, close to novel transcript with homology to vascular endothelial zinc finger 1 (= ZnF 161)</td>
</tr>
<tr>
<td>TCTTGCAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCCTAATGT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td></td>
<td>no unique hit</td>
<td></td>
</tr>
<tr>
<td>ATACNAAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGAAGAGGC</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>2924</td>
<td>Rattus norvegicus transcribed sequence with weak similarity to protein prf:2018199A (E. coli) 2018199A beta lactamase IRT-4 [Escherichia coli]</td>
<td>unique hit, close to novel transcript with homology to Cbx6</td>
</tr>
<tr>
<td>AGGACCAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGTACITTT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>10266</td>
<td>polypeptide GalNAc transferase T1</td>
<td>unique hit</td>
</tr>
<tr>
<td>AGAACTCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCTTCATCT</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1.48E-02</td>
<td>3247</td>
<td>collagen, type III, alpha 1</td>
<td>unique hit</td>
</tr>
<tr>
<td>AAGTCTGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>--------------</td>
<td>---------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>TTTCTAGTG AATAAAT</td>
<td>7</td>
<td>0</td>
<td>1.48E-02</td>
<td>7870</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein ref:NP_055680.1 (H.sapiens) chromosome condensation-related SMC-associated protein 1; chromosome condensation-related SMC-associated protein 1; KIAA0159 gene product</td>
<td>unique hit, close to novel transcript with homology to chromosome condensation-related SMC-associated protein 1 (structural maintainance of chromosomes)</td>
<td></td>
</tr>
<tr>
<td>AAAAAATAA ATATATT</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>40381</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, no assigned locus within 5kb</td>
<td></td>
</tr>
<tr>
<td>AAATAAAAC AGGCAACT</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>22244</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, &gt; 5kb from closest assigned transcript</td>
<td></td>
</tr>
<tr>
<td>AATGCGAAC ACCATCAC</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>25144</td>
<td>squamous cell carcinoma antigen recognized by T-cells 1</td>
<td>no hits in same direction</td>
<td></td>
</tr>
<tr>
<td>ATACTGACA CTTNGTAG</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td></td>
<td></td>
<td>No exact hit in rat genome</td>
<td></td>
</tr>
<tr>
<td>CACGCAACAG TCAGCCGA</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>11305</td>
<td>dentatorubral pallidolysian atrophy (=atrophin 1)</td>
<td>unique hit, part of DRPL gene, atrophin 1</td>
<td></td>
</tr>
<tr>
<td>CACTGTCGCC TTCAAGAT</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>35353</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 3.5kb from carboxypeptidaseD locus</td>
<td></td>
</tr>
<tr>
<td>CAGTTTGTCA CATTACT</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>32253</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein ref:NP_054738.1 (H.sapiens) HSPC009 protein [Homo sapiens]</td>
<td>2 genomic hits</td>
<td></td>
</tr>
<tr>
<td>CAGTTTGTCA CATTACT</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>101891</td>
<td>zinc finger protein Y1 (RLZF-Y)</td>
<td>multiple genomic hits</td>
<td></td>
</tr>
<tr>
<td>CGCGCTTGA ACTCCCTC</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>25124</td>
<td>Rat insulin-like growth factor I mRNA, 3' end of mRNA</td>
<td>chr7, 3.5kb from Igfl locus</td>
<td></td>
</tr>
<tr>
<td>CCTACACCTA ACTCTTAA</td>
<td>6</td>
<td>0</td>
<td>2.41E-02</td>
<td>12859</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>multiple genomic hits</td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CCTCTGTTTG</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>100120</td>
<td>cysteine string protein (chaperone)</td>
<td>unique hit, no assigned rat locus within 5kb</td>
</tr>
<tr>
<td>GTGGTAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTTTTATCCC AATTCCT</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td></td>
<td></td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>CTAGCTCAC TAAGGCA</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>37737</td>
<td>iGb3 synthase</td>
<td>unique hit</td>
</tr>
<tr>
<td>CTCCAATAA TTGATATT</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>94808</td>
<td>protein kinase inhibitor, alpha</td>
<td>unique hit</td>
</tr>
<tr>
<td>CTGGCCTCT GAGCTTGG</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>20730</td>
<td>Rattus norvegicus transcribed sequence with weak similarity to protein pir:S10889 (H.sapiens) S10889 proline-rich protein</td>
<td>unique hit, close to unovel transcript with homology to hKIAA0720 protein</td>
</tr>
<tr>
<td>CTTGCATAAA AATTTAAG</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>16704</td>
<td>Rattus norvegicus transcribed sequence with weak similarity to protein pir:S15433 (M.musculus) S15433 hypothetical protein</td>
<td>unique hit, chrX, no assigned transcript within 5kb</td>
</tr>
<tr>
<td>GAGACGGAG TCCAAAGG</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>63247</td>
<td>Rattus norvegicus similar to chromodomain helicase DNA binding protein 4; Mi-2b (LOC312712), mRNA</td>
<td>hit close to named transcript</td>
</tr>
<tr>
<td>GCAGATGTT ATGAATTG</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>13832</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1.5kb from novel locus, similar to hypothetical protein KCTD8)</td>
</tr>
<tr>
<td>GCATAGTCT GAGATGCA</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>40510</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>2kb from Wnt4 locus??</td>
</tr>
<tr>
<td>GGTAGGAGC ACCGCACT</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>72442</td>
<td>Rattus norvegicus similar to TBC1 domain family, member 8; BUB2-like protein 1; vascular Rab-GAP/TBC-containing (LOC360520), mRNA</td>
<td>unique hit, close to novel transcript with homology to KIAA clone of the TBC1 domain family</td>
</tr>
<tr>
<td>GTGAAAGAA TAAATGCA</td>
<td>6</td>
<td>0</td>
<td></td>
<td>2.41E-02</td>
<td>98686</td>
<td>Rattus norvegicus similar to KDEL-containing protein 1; endoplasmic reticulum-resident KDEL protein (LOC315664), mRNA</td>
<td>&lt; 5kb - similar to KDEL containing protein 1; endoplasmic reticulum-resident KDEL protein</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene Description</td>
<td>genomic BLAST</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>GTGAAAGAAA TAAATGCA</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>14013 Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2kb from mRIKEN cDNA 2010004J24</td>
<td></td>
</tr>
<tr>
<td>GTGTACATT GTTCTTAT</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>34113 Rattus norvegicus similar to hypothetical protein MGC10067 (LOC360514), mRNA</td>
<td>unique hit, 1kb from novel locus, with similarity to hypothetical protein MGC10067</td>
<td></td>
</tr>
<tr>
<td>GTTTCCCCCTT TATTCAA</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>24229 Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1kb from mBTB and Kelch domain protein 2 = Kelch-like 13 (Klh13)</td>
<td></td>
</tr>
<tr>
<td>TAAATAACT TTTCTCCA</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>8311 Rattus norvegicus transcribed sequence with weak similarity to protein sp:Q64707 (M.musculus) U2R1_MOUSE U2 small nuclear ribonucleoprotein auxiliary factor 35 kDa subunit related-protein 1 (SP2)</td>
<td>unique hit- wrong direction</td>
<td></td>
</tr>
<tr>
<td>TAAATATAA TTTATGTT</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>3989 Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1kb from mCyclinG2 locus</td>
<td></td>
</tr>
<tr>
<td>TCCAATAAAA GAGCTGAG</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td></td>
<td>No exact hit in rat genome</td>
<td></td>
</tr>
<tr>
<td>TGCACGACT ATGTATT</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>8266 Rattus norvegicus transcribed sequences</td>
<td>unique hit, 3kb from novel transcript with similarity to RING1 and YY1 binding protein (Rybp)</td>
<td></td>
</tr>
<tr>
<td>TGGTTCTCTT AATCCAGT</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>98823 Rattus norvegicus similar to KIAA1064 protein (LOC308385), mRNA</td>
<td>unique hit- no assigned gene locus within 5kb</td>
<td></td>
</tr>
<tr>
<td>TTCCGTGACT AATGTTAG</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>65930 notch gene homolog 2, (Drosophila)</td>
<td>unique hit, 2.5 kb from Notch2 locus</td>
<td></td>
</tr>
<tr>
<td>TTGATTTGTCT ATGTGT</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>12498 Rattus norvegicus transcribed sequences</td>
<td>unique hit, 5kb from novel transcript (similarity to human TU12B1-TY protein)</td>
<td></td>
</tr>
<tr>
<td>TTTTATTGCA AAGATGT</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>2.41E-02</td>
<td>9376 Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1.5kb from hNeurobeachin locus</td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>AAAATTTCT TGGCAAAA</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>6532</td>
<td>Rattus norvegicus transcribed sequence with weak similarity to protein ref:NP_037370.1 (H.sapiens) DNAJ domain-containing [Homo sapiens]</td>
<td>unique hit, close to novel transcript, similar to hriken cDNA 1810055D05</td>
</tr>
<tr>
<td>AAAGTCAAA AAAAAAAA</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>53915</td>
<td>PV-1</td>
<td>no hits</td>
</tr>
<tr>
<td>AACTACACC TGCTGTAA</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>9609</td>
<td>cystatin N</td>
<td>maps to cystatin E/M locus</td>
</tr>
<tr>
<td>AACTCCCGG TGCTGTAC</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>27227</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 5kb from bcl-2 locus</td>
</tr>
<tr>
<td>AAGTAGTTG GACTCGGG</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>6570</td>
<td>Rattus norvegicus similar to vaccinia related kinase 3 (LOC361565), mRNA</td>
<td>unique hit, close to hVRK3</td>
</tr>
<tr>
<td>AATAAAGAT AAATAAAC</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>16544</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>AATAATGAT ATATCGGT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>61277</td>
<td>Rattus norvegicus similar to hypothetical protein D930024B17 (LOC303552), mRNA</td>
<td>unique hit, close to novel transcript</td>
</tr>
<tr>
<td>ACAGACAGA ACATTCCC</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td></td>
<td></td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>ACCCTGGTG GGAGACA</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>16016</td>
<td>Rattus norvegicus putative eps protein (MGEPS) mRNA, partial cds</td>
<td>unique hit, 2kb from transcript homolog to EHD2</td>
</tr>
<tr>
<td>ACTGGGATG TTCTTGGT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td></td>
<td></td>
<td>unique hit 2kb from polyA polymerase alpha</td>
</tr>
<tr>
<td>AGAATGAAG TTCAACAG</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>2918</td>
<td>Rattus norvegicus similar to MUM2 protein (LOC287427), mRNA</td>
<td>unique hit, close to mVoltage gated K channel</td>
</tr>
<tr>
<td>AGATAAGCA GGAGTTC</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>9560</td>
<td>Rat mRNA</td>
<td>unique hit, no assigned loci in neighbourhood</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>AGGAGATGG AGAACCTTT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>93304</td>
<td>Rattus norvegicus similar to TRAF-binding protein (LOC293571), mRNA</td>
<td>unique hit, close to novel transcript with homology to TRAF binding protein (TRABID)</td>
</tr>
<tr>
<td>AGTGTGTAC AAGCTTAG</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td></td>
<td>unique hit, close to novel transcript similar to mEphB3 receptor (MDK5)</td>
<td></td>
</tr>
<tr>
<td>ATACATAC ACTGATAGA</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td></td>
<td>No exact hit in rat genome</td>
<td></td>
</tr>
<tr>
<td>ATGCTGAGA GGAAGCTT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>52627</td>
<td>Rattus norvegicus similar to hypothetical protein (LOC366980), mRNA</td>
<td>unique hit, close to novel transcript with homology to mRIKEN cDNA 4432409D24 = Arid rfx like</td>
</tr>
<tr>
<td>ATTAGTTAA CAACATC</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>3111</td>
<td>Rattus norvegicus similar to semaphorin cytoplasmic domain-associated protein 3A (LOC312607), mRNA = PDZ domain containing RING finger 3</td>
<td>unique hit, close to mRIKEN cDNA 1110020C07, SemaF cytopl dom ass protein 3A</td>
</tr>
<tr>
<td>ATTCCTTTGG AATGTAT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>2342</td>
<td>mannosyl (alpha-1,6)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase</td>
<td>unique hit- close to named gene</td>
</tr>
<tr>
<td>ATTAAGAC AGTATTTTT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>29900</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, no assigned locus in neighbourhood</td>
</tr>
<tr>
<td>CAAAAGGCT CTCCTAAT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>100904</td>
<td>Rattus norvegicus similar to Pmcf-pending protein (LOC297342), mRNA = potassium channel modulatory factor</td>
<td>unique hit- in wrong direction</td>
</tr>
<tr>
<td>CAAGGCCCT CCTAAAAG</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td></td>
<td>unique hit, no assigned locus in neigbourhood</td>
<td></td>
</tr>
<tr>
<td>CACTCCTGC TGAAGAAT</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>14510</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2kb from novel transcript (similar to neurofilament heavy polypeptide)</td>
</tr>
<tr>
<td>CAGATTTTT GTAATAAA</td>
<td>5</td>
<td>0</td>
<td></td>
<td>3.95E-02</td>
<td>11189</td>
<td>Ras-related associated with diabetes</td>
<td>unique hit (GTP binding protein RAD 1)</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CAGCAGCCT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>103390</td>
<td>Rattus norvegicus similar to RIKEN cDNA 0610008A10 (LOC300802), mRNA</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>CTGGAAGA</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>116313</td>
<td>Rattus norvegicus transcribed sequence with moderate similarity to protein ref:NP_112591.1 (H.sapiens) hypothetical protein DKFZp564D0372 [Homo sapiens]</td>
<td>multiple genomic hits</td>
</tr>
<tr>
<td>CATAATTGT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>8133</td>
<td>Rattus norvegicus similar to RIKEN cDNA 1110038M16 (LOC313529), mRNA</td>
<td>unique hit close to novel transcript with homology to mRIKEN cDNA 1700012A05 (=1110038M16)</td>
</tr>
<tr>
<td>GTCCAGTG</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>8458</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>CCTTGGGGG</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>91975</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2.5kb from locus with homology to diff proteins</td>
</tr>
<tr>
<td>AGGGTGGGG</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>98420</td>
<td>Rattus norvegicus similar to RIKEN cDNA 1700027N10 (LOC316238), mRNA</td>
<td>unique hit, part of novel transcript (similar to RIKEN cDNA 1700027N10)</td>
</tr>
<tr>
<td>CCTTTGTGTA</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>17732</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1kb from mSOX7 locus</td>
</tr>
<tr>
<td>GAGGATCAT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>118492</td>
<td>Rattus norvegicus transcribed sequence with strong similarity to protein sp:Pl1174 (H.sapiens) RS15_HUMAN 40S ribosomal protein S15 (RIG protein)</td>
<td>multiple hits</td>
</tr>
<tr>
<td>TGAATATC</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>118492</td>
<td>Rattus norvegicus transcribed sequence with strong similarity to protein sp:Pl1174 (H.sapiens) RS15_HUMAN 40S ribosomal protein S15 (RIG protein)</td>
<td>multiple hits</td>
</tr>
<tr>
<td>GAGGAGCGG</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>23241</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1.5kb from ZF13 locus (Zfp110, Zfp274)</td>
</tr>
<tr>
<td>GAGGAGAGT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>23241</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 1.5kb from ZF13 locus (Zfp110, Zfp274)</td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>--------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>GCCGAAATG</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>1078</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 3kb from novel transcript similar to mSpindlin</td>
</tr>
<tr>
<td>TCCTGTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCTAACTCA</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>28360</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 3.5kb from mADAMTS12 locus</td>
</tr>
<tr>
<td>TGCCGGAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCTGCGG</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>33523</td>
<td>Rattus norvegicus similar to A301 protein (LOC308592), mRNA</td>
<td>unique hit, close to novel transcript with homology to glutamate rich WD containing 1 (Grwd1 = A301)</td>
</tr>
<tr>
<td>GCTTCCTCCA</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>8046</td>
<td>casein kinase 1, delta</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGTTCCCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGAGATACC</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>11567</td>
<td>nel-like 2 homolog (chicken)</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGTTCAAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGCAAGCCCC</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td></td>
<td></td>
<td>multiple genomic hits (for 19bp)</td>
</tr>
<tr>
<td>CAGCGCTNT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTCTATTAT</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>41830</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2Kb from novel transcript similar to ZFP236</td>
</tr>
<tr>
<td>ACATTGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTGATTTCA</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>121231</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 2kb from novel transcript similar to cyclin M3</td>
</tr>
<tr>
<td>TTGTCTCTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAAAAAATG</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>36521</td>
<td>Rattus norvegicus similar to FGFR1 oncogene partner (LOC365103), mRNA</td>
<td>unique hit, 1.5kb from novel transcript similar to FGFR1 oncogene partner</td>
</tr>
<tr>
<td>TCTTTTGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAAAAGTACT</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>1868</td>
<td>beta-2 microglobulin</td>
<td>unique hit, beta 2 microglobulin</td>
</tr>
<tr>
<td>CACTTCTGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAAAGTACT</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>117190</td>
<td>Rattus norvegicus transcribed sequence</td>
<td>unique hit, 4kb from novel transcript with homology to mSCUBE1 (contains CUB domain, mostly found in developmentally reg proteins)</td>
</tr>
<tr>
<td>CACTTCCGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATTAAC</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>103231</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit</td>
</tr>
<tr>
<td>ACTTGTGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAATTTGAT</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>8202</td>
<td>ubiquitin-conjugating enzyme UBC7</td>
<td>unique hit</td>
</tr>
<tr>
<td>TTTGGGAAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACCTAAAAA</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td>66545</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, 4.5kb from novel transcript similar to PLAG1 (pleiomorphic adenoma gene)</td>
</tr>
<tr>
<td>TGATTTTCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TATTTTACTCT</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3.95E-02</td>
<td></td>
<td></td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>GCCAAGGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tag</td>
<td>VMP</td>
<td>VSU</td>
<td>fold</td>
<td>P-value</td>
<td>Unigene</td>
<td>Description</td>
<td>genomic BLAST</td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-----------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>TGACCATAT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>56603</td>
<td>amine oxidase, copper containing 3 (=vascular adhesion protein 1)</td>
<td>unique hit</td>
</tr>
<tr>
<td>TGTATATGA</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>99017</td>
<td>Pr2 protein</td>
<td>unique hit -orthologue of thyroid receptor interacting protein</td>
</tr>
<tr>
<td>TGTGGCATATA</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>16844</td>
<td>Rattus norvegicus similar to amyotrophic lateral sclerosis 2 chromosome region candidate 9; putative GTPase regulator (LOC363239), mRNA</td>
<td>unique hit, no assigned loci in same direction</td>
</tr>
<tr>
<td>TGTGGCACT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>62699</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>unique hit, no assigned locus within 5kb</td>
</tr>
<tr>
<td>TTCACAATG</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>24283</td>
<td>Rattus norvegicus transcribed sequences</td>
<td>No exact hit in rat genome</td>
</tr>
<tr>
<td>TTCATTAATT</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td>103261</td>
<td>Rattus norvegicus similar to retinoid x receptor interacting protein (LOC290997), mRNA</td>
<td>unique hit, close to novel transcript with similarity to retinoid x receptor interacting protein (=RAP 80)</td>
</tr>
<tr>
<td>TTCCCTGCAA</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>3.95E-02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>