The role of Inhibitor of DNA binding 4 (Id4) in mammary gland development and breast cancer

Simon Junankar

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

UNSW

Garvan Institute of Medical Research

St. Vincent's Hospital Clinical School

Faculty of Medicine

November 2012

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: Junankar

First name: Simon | Other name/s: Rajesh

Abbreviation for degree as given in the University calendar: PhD

School: St Vincent's Hospital Clinical School | Faculty: Medicine

Title: The role of Inhibitor of DNA binding 4 (Id4) in mammary gland development and breast cancer

Abstract 350 words maximum: (PLEASE TYPE)

The inhibitor of DNA binding (Id) proteins Id1-4 are transcriptional regulators that control many cell fate and developmental processes and are often deregulated in cancer. In this dissertation I examine the role of Id proteins in mammary development and neoplasia. Initially the role of Id1 in regulating the immune response to senescent tumour cells was examined, before a more thorough investigation into the role of Id4 during mammary gland development and breast cancer. Prior to the studies described in this thesis, the role for Id4 during mammary gland development had not been investigated and its role in breast cancer was controversial. Id4 expression patterns in the mammary gland were analysed throughout development by immunohistochemistry and the phenotype of Id4 loss was determined using the Id4 knockout mouse. The function of Id4 was further analysed using the normal mouse mammary epithelial cell line Comma-Dβ. Id4 expression was examined in breast cancer cell lines and a number of cohorts of breast cancer patients. The results presented here show that Id4 is a critical regulator of mammary gland development through its control of differentiation, proliferation and extracellular matrix (ECM) remodelling pathways. Id4 expression is restricted to the myoepithelial cells of the mammary gland and its loss leads to impaired mammary gland development. Overexpression and knockdown studies utilising the Comma-Dβ cells demonstrated that Id4 inhibited luminal differentiation and was required for cell proliferation. Furthermore Id4 overexpression promoted neoplastic transformation of the normal Comma-Dβ cell line *in vivo*. Transcript profiling experiments further demonstrated Id4 regulated a number of ECM remodelling genes. Our results from examining patient cohorts demonstrate that Id4 expression associates with the Her2 and basal-like subtypes of breast cancer and that high expression significantly correlates with improved patient survival. Despite high Id4 expression associating with improved survival, our combined results demonstrate that Id4 nonetheless has the capacity to promote tumourigenesis.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

Signature | Witness | 19/11/12 Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

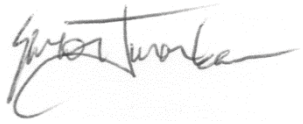
FOR OFFICE USE ONLY | Date of completion of requirements for Award:

THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS

Originality Statement

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that the assistance from others in project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed:



Date: 14/11/12

Acknowledgements

Firstly I would like to thank Alex Swarbrick for being a great supervisor, friend and mentor. You gave me the freedom to follow my own ideas and also gave direction when I needed it. I would also like to thank the other members of the Swarbrick group for their help, advice and friendship. In particular I would like to thank Radhika for conceptualising the Id4 project and for so much advice and help along the way.

I would like to thank my family and friends for their support and forbearance during my PhD studies. In particular my wife Marcel Batten for all of her love, help and support over the last four years and also our son Perry who has made the last 20 months all the more exciting and enjoyable. I would also like to acknowledge the support of both my parents and Marcel's parents over the last four years. I hope now to be able to spend some more time with my friends and family.

Conferences, Publications and Awards

Conferences

Simon Junankar, Sandra O'Toole, Ewan Millar, Radhika Nair, Andrea McFarland, Kyuson Yun, Robert L. Sutherland, Alexander Swarbrick. 2012, A critical role for Id4 in breast cancer and mammary gland development. *Poster presentation. 24th Lorne Cancer Conference, Lorne, Victoria, Australia.*

Simon Junankar, Sandra O'Toole, Ewan Miller, Radhika Nair, Andrea McFarland, Kyuson Yun, Robert L. Sutherland, Alexander Swarbrick. 2011. Inhibitor of DNA binding 4 plays a crucial role in breast cancer and mammary gland development. *Poster presentation. 19th NSW Scientific Meeting, ASMR, The University of Sydney, NSW, Australia.*

Simon Junankar, Sandra O'Toole, Ewan Miller, Radhika Nair, Andrea McFarland, Robert L. Sutherland, Alexander Swarbrick. 2011. A crucial role for Id4 in breast cancer and mammary gland development. *Poster presentation. Stem Cells, Cancer and Metastasis Keystone Symposia, Keystone, Colorado, USA.*

Simon Junankar, Sandra O'Toole, Ewan Miller, Radhika Nair, Andrea McFarland, Robert L. Sutherland, Alexander Swarbrick. 2011. A crucial role for Id4 in breast cancer and mammary gland development. *Poster presentation. 23rd Lorne Cancer Conference, Lorne, Victoria, Australia.*

Simon Junankar, Emie Roy, Alex Swarbrick. 2009. Id1 overexpression protects mammary tumours from senescence and immune system recognition. *Poster presentation. 39th Annual Scientific Meeting Australasian Society for Immunology, Gold Coast, Queensland, Australia.*

Simon Junankar, Emie Roy, Alex Swarbrick. 2009. Id1 overexpression plays a role in protecting from senescence and immune cell recognition in Ras overexpressing mammary tumours. *Poster presentation. 21st Lorne Cancer Conference, Lorne, Victoria, Australia*

Publications

Junankar S, Baker L, Yang J, O'Toole S, McFarland A, Phua Y, Yun K, Nair R, Swarbrick A. Id4 regulates stem cell maintenance and luminal differentiation during mammary gland development, (manuscript in preparation).

Nair R, Junankar S, O'Toole S, Shah J, Borowsky AD, Bishop JM, Swarbrick A. Redefining the expression and function of the inhibitor of differentiation 1 in mammary gland development. PLoS One. 2010; 5 (8):e11947

Awards

2012 – 24th Lorne Cancer Conference – Poster award.
2009 – Cancer Institute NSW – Research Scholars award.

Abstract

The inhibitor of DNA binding (Id) proteins Id1-4 are transcriptional regulators that control many cell fate and developmental processes and are often deregulated in cancer. In this dissertation I examine the role of Id proteins in mammary development and neoplasia. Initially the role of Id1 in regulating the immune response to senescent tumour cells was examined, before a more thorough investigation into the role of Id4 during mammary gland development and breast cancer. Prior to the studies described in this thesis, the role for Id4 during mammary gland development had not been investigated and its role in breast cancer was controversial. Id4 expression patterns in the mammary gland were analysed throughout development by immunohistochemistry and the phenotype of Id4 loss was determined using the Id4 knockout mouse. The function of Id4 was further analysed using the normal mouse mammary epithelial cell line Comma-Dβ. Id4 expression was examined in breast cancer cell lines and a number of cohorts of breast cancer patients. The results presented here show that Id4 is a critical regulator of mammary gland development through its control of differentiation, proliferation and extracellular matrix (ECM) remodelling pathways. Id4 expression is restricted to the myoepithelial cells of the mammary gland and its loss leads to impaired mammary gland development. Overexpression and knockdown studies utilising the Comma-Dβ cells demonstrated that Id4 inhibited luminal differentiation and was required for cell proliferation. Furthermore Id4 overexpression promoted neoplastic transformation of the normal Comma-Dβ cell line *in vivo*. Transcript profiling experiments further demonstrated Id4 regulated a number of ECM remodelling genes. Our results from examining patient cohorts demonstrate that Id4 expression associates with the Her2 and basal-like subtypes of breast cancer and that high expression significantly correlates with improved patient survival. Despite high Id4 expression associating with improved survival, our combined results demonstrate that Id4 nonetheless has the capacity to promote tumourigenesis.

Frequently used abbreviations

ABC transporter	ATP binding cassette transporter
ABCTB	Australian breast cancer tissue bank
αSMA	alpha smooth muscle actin
AREG	Amphiregulin
ATCC	American type tissue collection
BaA	Basal A
BaB	Basal B
bHLH	basic helix-loop-helix
BMP	Bone morphogenic protein
BRCA1	Breast cancer associated 1
C3TAg	C3(1)/SV40 T-antigen
CD	Cluster of differentiation
CFC	Colony forming capacity
CK14	Cytokeratin 14 (also known as Keratin 14/K14)
CK5	Cytokeratin 5 (also known as Keratin 5/K5)
CK6	Cytokeratin 6 (also known as Keratin 6/K6)
CK8	Cytokeratin 8 (also known as Keratin 8/K8)
CpG	Cystein-phosphate-guanine
CSC	Cancer stem cell
d.n.s.	Data not shown
D18	Day 18
DAPI	4',6-diamidino-2-phenylindole
DCIS	Ductal carcinoma *in situ*
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
Dox	Doxycycline
E16	Embryonic day 16
E2	Oestrogen
ECM	Extracellular matrix
EcoR	Ecotropic receptor
EGF	Epidermal growth factor
EMT	Epithelial to mesenchymal transition
ER	Oestrogen receptor
ES	Embryonic stem
FACS	Fluorescence activated cell sorter
FBS	Foetal bovine serum
FCS	Foetal calf serum
FFPE	Formalin fixed paraffin embedded
FGF	Fibroblast growth factor
FISH	Fluorescent *in situ* hybridisation
FSC	Forward scatter

GFP	Green fluorescent protein
GSEA	Gene set enrichment analysis
H&E	Haematoxylin and eosin
H2B	Histone-2B
HA	Haemagglutanin
Her2	Human epidermal growth factor receptor 2
Het	Heterozygous
HR	Hazard ratio
IC	Immunocytochemistry
Id1	Inhibitor of differentiation/DNA binding 1
Id2	Inhibitor of differentiation/DNA binding 2
Id3	Inhibitor of differentiation/DNA binding 3
Id4	Inhibitor of differentiation/DNA binding 4
IF	Immunofluorescence
IHC	Immunohistochemistry
IL2Rγ	Interleukin 2 receptor gamma
IP	Intra-peritoneal
IPA	Ingenuity pathway analysis
K14	Keratin 14
K8	Keratin 8
KD	Knockdown
KO	Knockout
LB	Luria broth
MHC	Major Histocompatability complex
mIR	MicroRNA
mMEC	Mouse mammary epithelial cell
MMP	Matrix metalloproteinase
MMTV	Mouse mammary tumour virus
mRNA	Messenger RNA
MRU	Mammary repopulating unit
MTS	(3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium)
n	Number
NA	Not applicable
ND	Not done
NK cell	Natural killer cell
NOD	Non-obese diabetic
NOG	NOD/SCID/IL2Rγ$^{-/-}$
NS	Not significant
OE	Over expression
OIS	Oncogene induced senescence
OS	Overall survival
OVX	Ovariectomized
p value	Probability value
P4	Progesterone

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PR	Progesterone receptor
Prl	Prolactin
PyMT	Polyoma middle T-antigen
RAG	Recombination activating gene
Rb	Retinoblastoma protein
RFS	Recurrence free survival
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SASP	Senescence associated secretion phenotype
SA-β-gal	Senescence associated β-galactosidase
SCID	Severe combined immunodeficiency
SD	Standard deviation
SEM	Standard error of the mean
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
SOC	Super optimal broth
SSC	Side scatter
TBS	Tris buffered saline
TDLU	Terminal ductal lobular unit
TEB	Terminal end bud
Tet	Tetracycline
TetR/TRE	Tetracycline response element
TGF-β	Transforming growth factor beta
tTA	Tetracycline transactivator
UTR	Untranslated region
WAP	Whey acidic protein
WT	Wild type

Table of Contents

Copyright and Authenticity statement I
Originality Statement II
Acknowledgements III
Conferences, Publications and Awards IV
Abstract V
Frequently used abbreviations VI
Table of Contents IX
Chapter 1. Introduction 1
1.1 General Introduction 1
1.2 Mammary Gland Development 1
1.2.1 Mammary gland structure 1
1.2.2 Normal mammary development 3
1.2.3 Mammary epithelial cell hierarchy 7
1.2.4 Markers that delineate different mammary epithelial cell populations 9
1.2.5 Factors that regulate mammary gland development 10
1.3 Breast Cancer 17
1.3.1 Introduction 17
1.3.2 Subtypes of Breast Cancer 18
1.3.3 The cancer stem cell theory and breast cancer 21
1.4 Id proteins 21
1.4.1 Overview 22
1.4.2 Regulation of Id proteins 24
1.4.3 Id proteins in development 25
1.4.4 Id proteins in cancer 25
1.5 Inhibitor of DNA binding 4 (Id4) 26
1.5.1 Gene structure, expression, localisation 26
1.5.2 The role of Id4 in stem and developmental pathways 27
1.5.3 The role of Id4 in cancer 29
1.6 Outline of thesis 31
Chapter 2. Materials and Methods 34
2.1 Media and Buffers 34
2.2 Mice 35

2.2.1 Oestrus staging, Whittening, and timed mating 35
2.2.2 Surgical procedures 36
2.2.3 Mammary epithelial cell transplants 36
2.2.4 Mammary piece/mammary tumour transplants 37
2.2.5 *In vivo* imaging 37
2.2.6 Tissue collection 37
2.3 Plasmids and siRNAs 38
2.3.1 Transformation of competent bacteria 40
2.3.2 Plasmid DNA preparations 41
2.4 Cell culture 41
2.4.1 Cell growth conditions 41
2.4.2 Cryopreservation of mammalian cells 41
2.4.3 Proliferation assay 42
2.4.4 *In vitro* doxycycline treatment 42
2.4.5 *In vitro* differentiation of Comma-Dβ cells 42
2.4.6 Non-lethal irradiation of NIH-3T3 cells 42
2.4.7 Transfection of Comma-Dβ cells with siRNA 43
2.5 Production of viral supernatants and viral transduction of primary mMECs and cell lines 43
2.5.1 Retroviral production 43
2.5.2 Lentiviral production 43
2.5.3 Retroviral transduction of primary mammary epithelial cells 44
2.5.4 Retroviral transduction of Comma-Dβ cells 44
2.5.5 Lentiviral transduction of Comma-Dβ cells 44
2.6 Preparation of single cell suspensions 44
2.6.1 Mouse mammary epithelial cell preparations 44
2.6.2 Mouse mammary tumour cell preparation for the analysis of immune cell infiltrates 45
2.7 Colony forming capacity (CFC) assay 45
2.8 Flow cytometry and FACS 46
2.8.1 Flow cytometry for immune cells infiltrating mammary tumours 46
2.8.2 Flow cytometry for mammary epithelial stem cell populations 46
2.8.3 Fluorescence activated cell sorting (FACS) 47
2.9 Histological methods 47

2.9.1 Wholemount histology 47
2.9.2 Tissue processing for IHC 47
2.9.3 Immunostaining 47
2.9.4 Senescence associated β-galactosidase staining 50
2.10 Protein methods 50
2.10.1 Protein extraction 50
2.10.2 Quantifying protein concentration 50
2.10.3 Western blotting 50
2.11 RNA methods 51
2.11.1 RNA extraction 51
2.11.2 RNA analysis 52
2.11.3 RNA Microarray analysis 52
2.11.4 cDNA synthesis 52
2.11.5 Quantitative RT-PCR/Microarray analysis 52
2.12 DNA methylation analysis 53
Chapter 3. The role of Id1 in modulating the immune response to Ras driven mammary tumourigenesis 54
3.1 Introduction: 54
3.1.1 Oncogene induced senescence 54
3.1.2 Role of Id1 in breast cancer 55
3.1.3 Senescence and immune recognition 57
3.1.4 Hypothesis: 59
3.1.5 Aims: 59
3.2 Results: 59
3.2.1 Analysing immune infiltrates into previously generated Id1 Ras mammary tumours by IHC 59
3.2.2 Determining the best vector combination for inducible Id1 overexpression 61
3.2.3 Generating mammary tumours overexpressing inducible Id1 and Ras 62
3.2.4 Switching Id1 off after mammary tumours are established 64
3.2.5 Repeating Id1 switch off with tumours with very high Id1 overexpression 66
3.2.6 Comparing Id1/Ras mammary tumour growth between immune compromised mice and controls 69
3.2.7 Repeat of switching Id1 off after mammary tumours are established with additional doxycycline 71
3.2.8 Immune infiltrate into Id1 Ras mammary tumours 72

3.3 Discussion: 75
Chapter 4. The role of Id4 in mammary gland development 78
4.1 Introduction: 78
4.1.1 Id4 and mammary gland biology 78
4.1.2 Hypothesis: 79
4.1.3 Aims: 79
4.2 Results: 79
4.2.1 Id4 is expressed by cap and myoepithelial cells of the mouse mammary gland and is regulated by the oestrus cycle 79
4.2.2 Loss of Id4 leads to a delay in pubertal mammary gland development 87
4.2.3 Loss of Id4 does not affect levels of myoepithelial cell markers 97
4.2.4 Loss of Id4 leads to an increase in luminal progenitor cells and reduction in stem cells 99
4.3 Discussion: 106
Chapter 5. Understanding the function of Id4 using the mouse normal mammary epithelial cell line Comma-Dβ 111
5.1 Introduction: 111
5.1.1 The Comma-Dβ cell line model 111
5.1.2 Hypothesis: 112
5.1.3 Aims: 112
5.2 Results: 113
5.2.1 Comma-Dβ cells endogenously express Id4 113
5.2.2 Overexpression of Id4 in a normal mouse mammary cell line 114
5.2.3 Id4 overexpression promotes tumour formation in the Comma-Dβ cells 122
5.2.4 Knockdown of Id4 in a normal mouse mammary cell line 123
5.2.5 Overexpression and knockdown of Id4 in a normal mouse mammary cell line does not regulate p38MAPK phosphorylation 130
5.2.6 Genome-wide determination of Id4 target genes and pathways 131
5.2.7 Analysing profiling results 132
5.2.8 Validation of transcript profiling results 140
5.3 Discussion: 141
Chapter 6. The role of Id4 in breast cancer 146
6.1 Introduction: 146
6.1.1 Id4 as a tumour suppressor in breast cancer 146
6.1.2 Id4 as a tumour promoter in breast cancer 147

6.1.3 Hypothesis: 148
6.1.4 Aims: 148
6.2 Results: 149
6.2.1 Analysis of Id4 promoter methylation, mRNA expression and protein expression in a panel of breast cancer cell lines 149
6.2.2 Id4 is highly expressed by a subset of triple-negative and Her2 amplified breast cancer samples 156
6.2.3 High Id4 mRNA level correlates with better survival within the basal-like and Her2 subtypes of breast cancer 158
6.2.4 High Id4 protein levels correlate with better survival in basal-like breast cancer patients 161
6.2.5 Association of better survival with chemotherapy response 164
6.2.6 Id4 protein expression levels in a number of mouse models of breast cancer 164
6.3 Discussion: 167
Chapter 7. Discussion 173
Chapter 8. References 179
Appendix 1 200
Appendix 2 217

Chapter 1. Introduction

1.1 General Introduction

Developmental processes and cancer are deeply intertwined. Genes and pathways that regulate developmental and differentiation pathways are frequently disrupted in cancer (Visvader, 2009). Cancer pathologists have long recognised that less differentiated tumours are more aggressive and are associated with poorer patient outcome (Bloom and Richardson, 1957). It is thus not surprising that genes that regulate differentiation are often deregulated during carcinogenesis. This link has been clearly demonstrated between mammary gland development and breast cancer, which is the main focus of this thesis (Brisken and O'Malley, 2010). Understanding the function and mechanism of action of factors controlling differentiation during normal development will thus increase our understanding of the aetiology of the cancers in which they are deregulated. This understanding can lead to the discovery of novel therapeutic targets as well as improving the targeting of current therapies.

1.2 Mammary Gland Development

The mammary gland is the defining organ of mammals. Lactation provides newborns with all the nutrients they need to continue their development and also provides them with protection from diseases in the form of maternal antibodies (Khokha and Werb, 2011). Due to its crucial and highly dynamic nature, the development and homeostasis of the mammary gland are tightly regulated processes.

1.2.1 Mammary gland structure

The mammary gland is composed of a mixture of cell types including epithelial cells, adipocytes, fibroblasts, endothelial cells, and a variety of immune cells (Polyak and Kalluri, 2011). Adipocytes make up the bulk of the mammary gland which is thus known as the mammary fat pad. The epithelial cells form a bilayered ductal system within the fat pad that is composed of an inner ring of luminal cells surrounded by an outer layer of myoepithelial cells (Figure 1-1A). A network of fibroblasts, immune cells, and extracellular matrix surrounds the ducts. The ductal network extends from the nipple to fill the entire mammary fat pad. Studies in the mouse have led to many crucial discoveries about mammary gland development that have gone on to improve our understanding and treatment of breast cancer. However, there are a few differences in structure between the mammary glands of humans and mice. Firstly, whilst in humans there are several ducts extending from the nipple into the fat pad that terminate in terminal ductal lobular units (TDLUs) in mice there is one duct leading from the

nipple that branches numerous times to fill the mammary fat pad (Figure 1-1B) (Visvader, 2009). Secondly, human mammary glands tend to have a lower percentage of adipocytes and an increased level of extracellular matrix deposition around the mammary epithelium (Simian et al., 2001). Aside from these modest differences, the processes that regulate the mammary gland are very similar between the two species, so for the rest of this introduction I will focus on mammary gland development in the mouse.

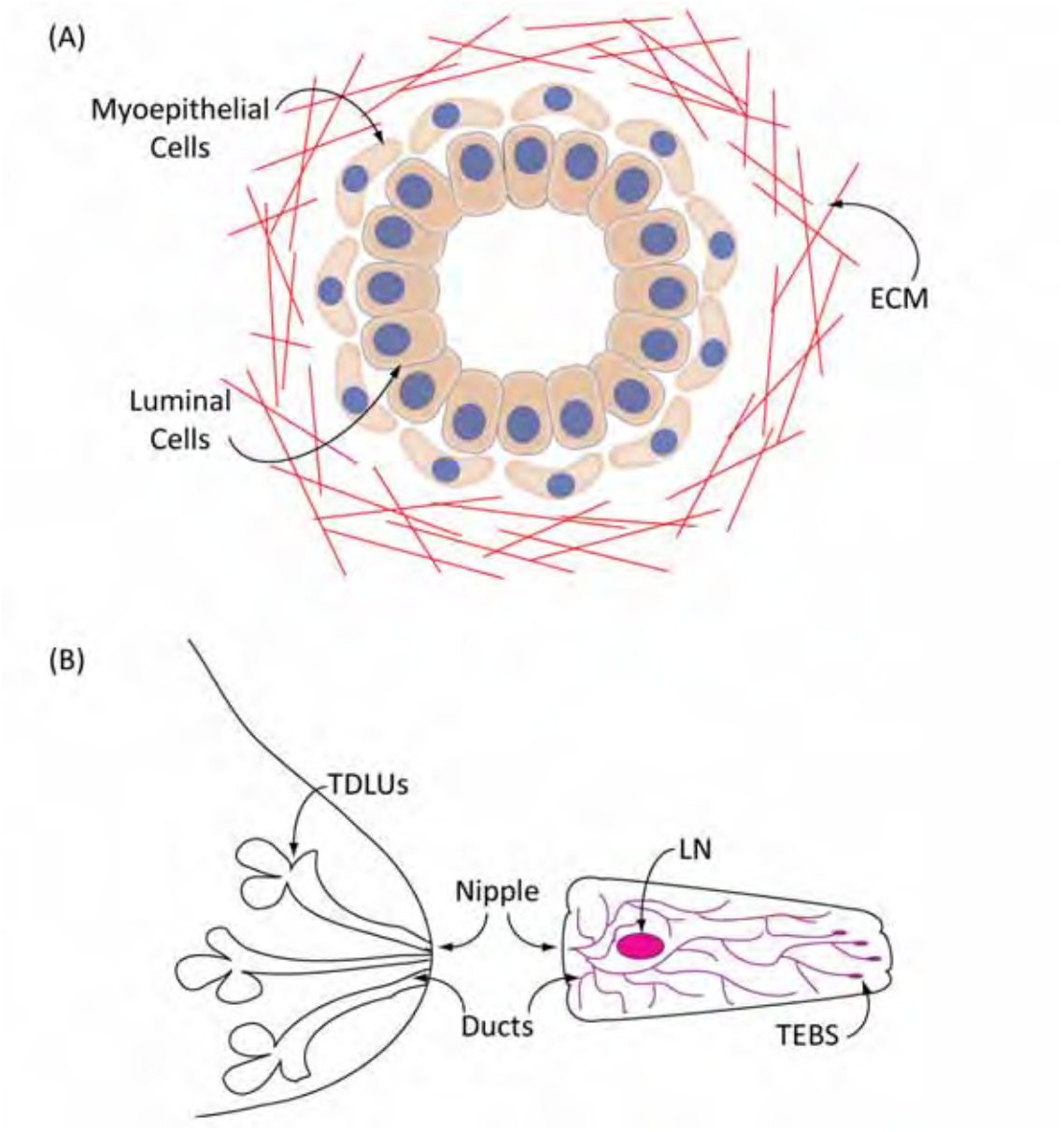


Figure 1-1. Structure of the mammary gland.
(A) Cross section of a mammary gland duct indicating the luminal epithelial, myoepithelial, and ECM. (B) Organisation of the ducts in humans (left) and mice (right) with the terminal ductal lobular units (TDLUs), nipple, ducts, terminal end buds (TEBS), and lymph node (LN) indicated.

1.2.2 Normal mammary development

The mammary gland as an organ is unique in that the majority of its development occurs after birth (Figure 1-2). I will now give a brief overview, followed by a detailed description of the stages of mammary gland development. During embryogenesis the mammary gland placode develops into a rudimentary ductal structure that is present at birth and that does not undergo significant further development until puberty. At puberty the levels of the hormones oestrogen and progesterone increase and drive a rapid growth in the mammary ductal tree (Watson and Khaled, 2008). In the mouse the epithelial ducts in the mammary gland go from a rudimental ductal structure to completely filling the entire fat pad with an elaborate ductal tree within about 3-5 weeks (McNally and Martin, 2011). After the development of the mature ductal tree the mammary epithelium continues to undergo cyclical proliferative and apoptosis phases driven by the oestrus cycle (Brisken and O'Malley, 2010). Upon pregnancy there is a second rapid burst of proliferation and differentiation driven this time by prolactin and progesterone in addition to oestrogen. These hormones lead to the expansion and differentiation of the mammary epithelial cells into the milk producing alveolar structures that will feed the newborn pups. Following weaning of the pups there is a rapid reduction in the alveolar structures which is called involution (Watson and Khaled, 2008). During involution there is extensive apoptosis and cell clearance by the immune system. After involution the mammary gland is returned to a ductal morphology similar to that present pre-pregnancy (Watson and Kreuzaler, 2011). I will now describe the different stages of mammary gland development in more detail.

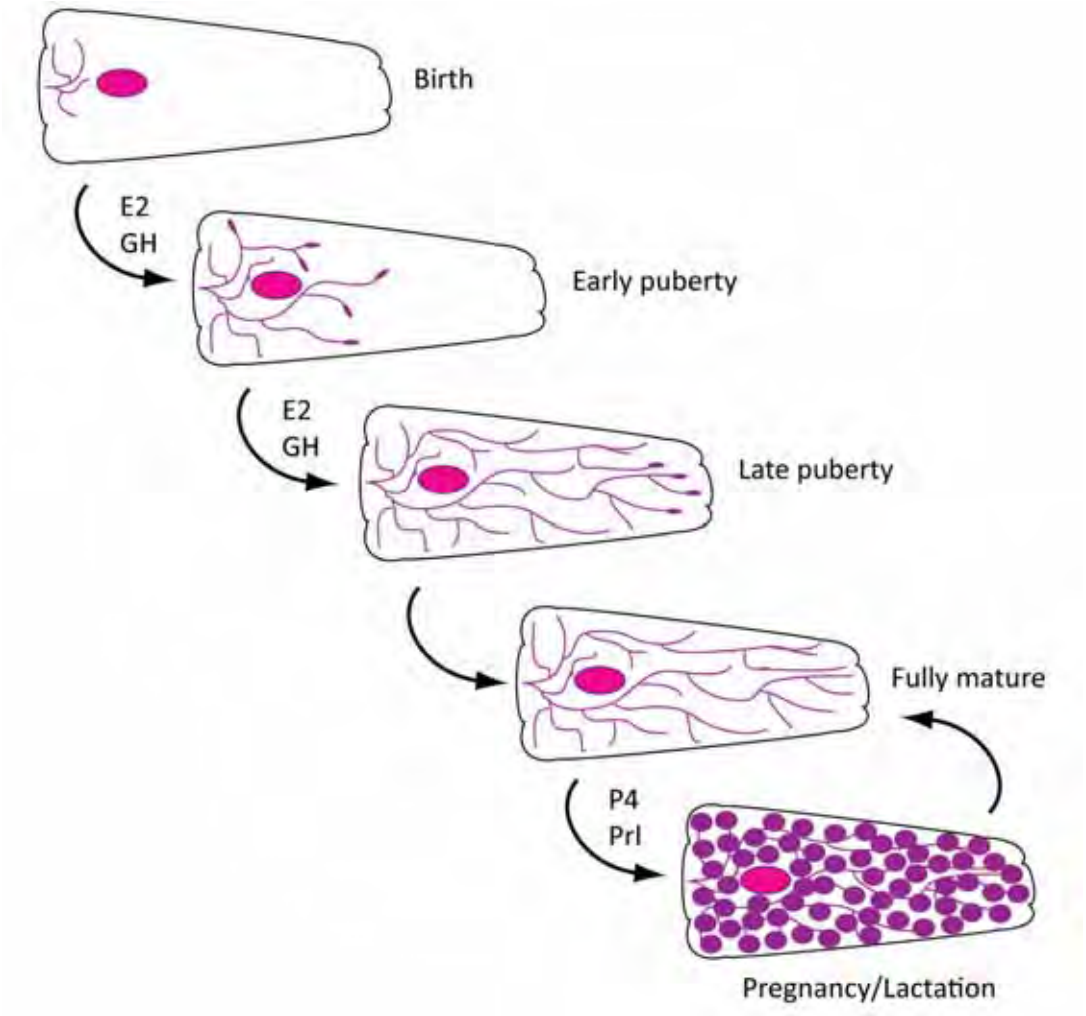


Figure 1-2. Mammary gland development in the mouse and hormonal regulation.
Changes in mammary gland structure from birth, through early and late puberty to a fully mature gland and pregnancy. Following lactation the gland returns to the fully mature state via involution. Oestrogen (E2) and growth hormone (GH) regulate pubertal development while progesterone (P4) and prolactin (Prl) regulate pregnancy induced development.

1.2.2.1 Embryonic development

Mammary gland development starts during embryogenesis. At embryonic day 10.5 (E10.5) structures known as milk lines form, there are two of these and they run in an anteroposterior direction from the forelimb to hindlimb buds (Veltmaat et al., 2004). The milk lines are ridges composed of multilayered ectoderm that form from the skin. At E11.5 five pairs of mammary placodes form along these milk lines. The mammary buds form from the placodes and invade into the fat pad precursor, with ductal branching starting at around E16.5, which subsequently leads to the formation of a rudimentary ductal structure by E18.5. Development of the mammary gland then halts at E18.5. A number of factors are important during embryonic development, these include parathyroid hormone related protein (PTHrP), bone morphogenic protein 4 (BMP4), Wnt signalling, fibroblast growth factor 10 (FGF10), and the inhibition of

hedgehog signalling pathway via Gli3 expression (Chu et al., 2004; Hatsell and Cowin, 2006; Hens et al., 2007; Veltmaat et al., 2006). The development of each pair of mammary placodes is regulated independently with different signalling molecules being more important for the development of some over others, for example mice lacking *Fgf10* or its receptor *Fgfr2b* only form placodes at position 4 (Mailleux et al., 2002). The exact timing of development is also variable with placodes 3 and 4 being the first to develop, followed by 1 and 5, and finally placode pair 2 (Watson and Khaled, 2008).

1.2.2.2 Pubertal development

Mammary gland development is limited between E18.5 and puberty, with the rudimentary ductal structure expanding only in a manner commensurate with the overall growth of the animal, this is also known as allosteric growth (Watson and Khaled, 2008). The next stage of mammary gland development is initiated at puberty, in the mouse this occurs from around 3-4 weeks of age (Watson and Khaled, 2008). At puberty the levels of the ovarian hormone oestrogen rise dramatically and stimulate a burst of growth in the mammary gland that takes the mammary epithelium from the rudimentary ductal structure to a highly branched ductal network that fills the entire mammary fat pad (Mueller et al., 2002). This development is a highly regulated process that requires the interaction and crosstalk of all the cell types within the mammary gland including the epithelial cells, fibroblasts, adipocytes and immune cells (Coussens and Pollard, 2011; Mueller et al., 2002). The proliferation of the epithelial cells and the invasion into the fat pad by the ductal tree is driven by bulbous structures at the end of the ducts termed terminal end buds (TEBs) (Figure 1-3) (McNally and Martin, 2011).

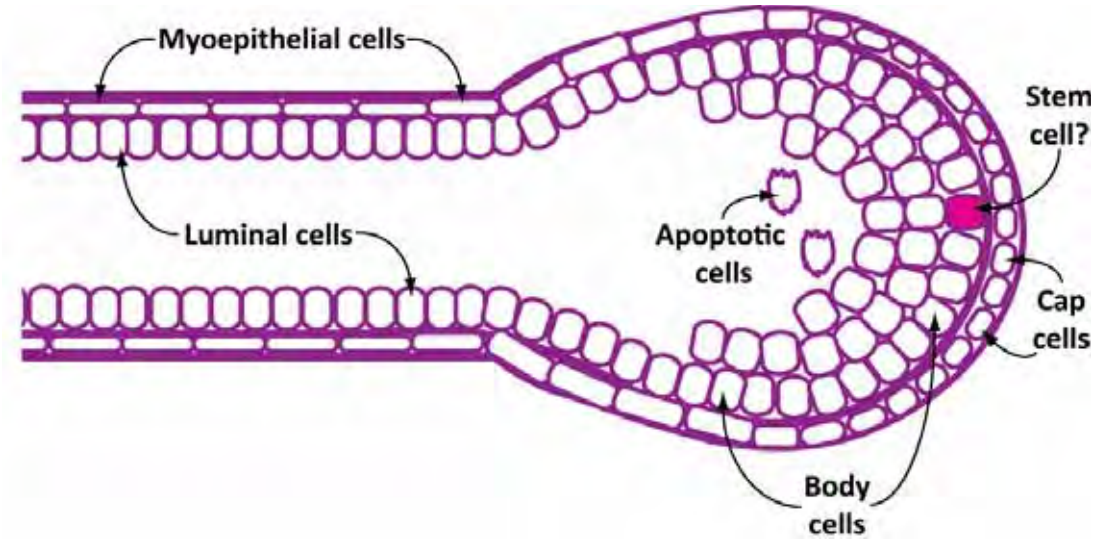


Figure 1-3. Structure of the terminal end bud (TEB).
With the localisation of the putative stem cell, the cap cells, body cells, and apoptotic cells. The left of the diagram showing the formation of the mature duct with mature luminal, and myoepithelial cells indicated.

TEBS are composed of an outer layer of cap cells and inner multilayered structure composed of body cells (Sternlicht, 2006). The highly proliferative cap and body cells are thought to differentiate into the other cells of the ducts, with the cap cells differentiating to form the outer myoepithelial cell layer and the body cells differentiating into the luminal cell layer (Woodward et al., 2005). Lumen formation occurs through the apoptosis of some of the inner body cells in a BIM dependent process (Mailleux et al., 2007). Once the TEBS reach the edge of the fat pad at roughly 8-10 weeks of age they stop proliferating and differentiate into terminal ducts (McNally and Martin, 2011).

After the ductal tree is completely formed the mammary gland is maintained in a dynamic state in the nulliparous mouse by the oestrus cycle. Initiated by every oestrus there is a round of epithelial proliferation that leads to the formation of tertiary side branching along the ducts in a TGFβ1 and MMP3 regulated manner (Pierce et al., 1993; Wiseman et al., 2003). This side branching reaches a maximum at diestrus, when in the absence of pregnancy there is then a round of organised apoptosis that returns the mammary gland to its original pre-oestrus state (Sternlicht et al., 2006).

1.2.2.3 Pregnancy induced development and lactation

The next major developmental stage of the mammary gland is during pregnancy, which in the mouse takes 21 days. A burst of proliferation leads to tertiary branching from the primary

ductal tree and lobular alveolar development. This is initiated by increased systemic levels of the hormones progesterone and prolactin. The induction of the transcription factor Elf-5 by prolactin is critical for the differentiation of luminal epithelial cells into the milk producing alveolar cells (Harris et al., 2006; Oakes et al., 2008). By the end of pregnancy the mammary gland is completely filled with a large number of milk secreting alveoli, these are composed of luminal alveolar cells surrounded by a network of myoepithelial cells able to constrict and pump the milk down the ducts to the nipple (Watson and Khaled, 2008). Along with the expansion of the epithelial cells during pregnancy there needs to be a commensurate reduction in adiposity and this is achieved by the de-differentiation of the adipocytes (Watson and Kreuzaler, 2011).

1.2.2.4 Involution

The final developmental stage of the mammary gland is involution, which occurs following weaning. During involution, there is extensive epithelial cell apoptosis and clearance by the immune system. Involution occurs in two distinct phases, the first is reversible whereas the second is not (Lund et al., 1996). The first phase in the mouse lasts only 48 hours and is not dependent on external growth factors or hormones, it is characterised by the appearance of dying cells within the lumen of the alveoli and the infiltration of neutrophils (Watson and Kreuzaler, 2011). The second phase is dependent on circulating factors, and the extensive alveolar epithelial cell death is associated with massive ECM remodelling, adipocyte expansion, an influx of macrophages, and plasma kallikrein produced by mast cells (Lilla et al., 2009; Watson and Kreuzaler, 2011). These processes are tightly controlled at many levels by cytokines, growth factors, proteases, and transcription factors. Apoptosis of the epithelial cells is thought to be mediated by a caspase dependant mechanism initiated by the loss of cellular adhesion and anoikis (Watson and Kreuzaler, 2011).

1.2.3 Mammary epithelial cell hierarchy

The mammary gland is a highly dynamic organ that undergoes large proliferative bursts during puberty and pregnancy. It thus requires the ability to rapidly produce multiple cell lineages in a coordinated manner. To accomplish this it is proposed that there exists a mammary stem cell at the top of a hierarchical tree extending to multipotent progenitors, then committed progenitors, and finally into fully differentiated luminal and myoepithelial cells (Figure 1-4). The presence of mammary stem cells in the adult mouse was first demonstrated by transplantation studies by DeOme and colleagues. These showed that small pieces of mammary gland could grow and recapitulate an entire mammary gland when transplanted

into the fat pad of a recipient mouse cleared of endogenous epithelium (Smith and Medina, 2008). The fact that any section of mammary gland could be transplanted suggested that stem and progenitor cells are located throughout the mammary gland. More recently this has been confirmed by using single cell mammary epithelial cell transplants that demonstrated that a single cell has the ability to recapitulate an entire mammary epithelial ductal tree (Shackleton et al., 2006; Stingl et al., 2006). These studies demonstrated that mammary stem cells made up roughly 1:3000-1:5000 of the epithelial cells within a mammary gland and that this could be enriched by fluorescence activated cell sorting (FACS) on specific cell surface marker to 1:60-1:500 (Shackleton et al., 2006; Stingl et al., 2006). The existence of luminal progenitors has also been confirmed using FACS combined with *in vitro* and *in vivo* studies (Asselin-Labat et al., 2007). The study of the mammary stem and progenitor cells revolves around FACS studies as described above often combined with *in vivo* studies involving limiting dilution transplantation to determine enrichment of stem cell activity in different fractions of mammary epithelial cells, and *in vitro* studies examining colony forming capacity (CFC) (Stingl, 2009). There are several variations of the CFC assay and these include epithelial cells plated onto collagen coated plates; onto plates with a feeder layer of irradiated fibroblasts; and into soft agar.

A variant on the stem cell hierarchy model adds another layer of complexity to our understanding of mammary epithelial cells. It suggests that there is a division between ductal luminal and myoepithelial progenitors, and alveolar luminal and myoepithelial progenitors (Smith and Medina, 2008). This was based on transplantation studies where ductal restricted or alveolar restricted outgrowths were formed (Smith and Medina, 2008). Our understanding of alveolar restricted progenitors is still limited as they have only been described by a couple of groups and have not been extensively analysed (Britt et al., 2009).

A recent study has suggested that the endogenous proliferation and differentiation of cells during puberty and pregnancy is limited to expansion of committed progenitor cells and does not involve the multipotent progenitors or stem cells that are necessary for mammary epithelial outgrowth following transplantation (Van Keymeulen et al., 2011). This study used elegant lineage tracking experiment utilizing CK14 and CK8 reporter mice, however one caveat of interpreting these experiments is the assumption that all stem cells express sufficient levels of CK14 to induce Cre expression to allow for lineage tracing. At least one group has shown that CK14 expression is low or non-existent in the cap cells of the developing mammary gland, where the putative stem cell is thought to be enriched during development (McCaffrey and

Macara, 2009). The evolutionary advantage of retaining a mammary stem cell population after embryogenesis is unclear if these cells serve no function later in life.

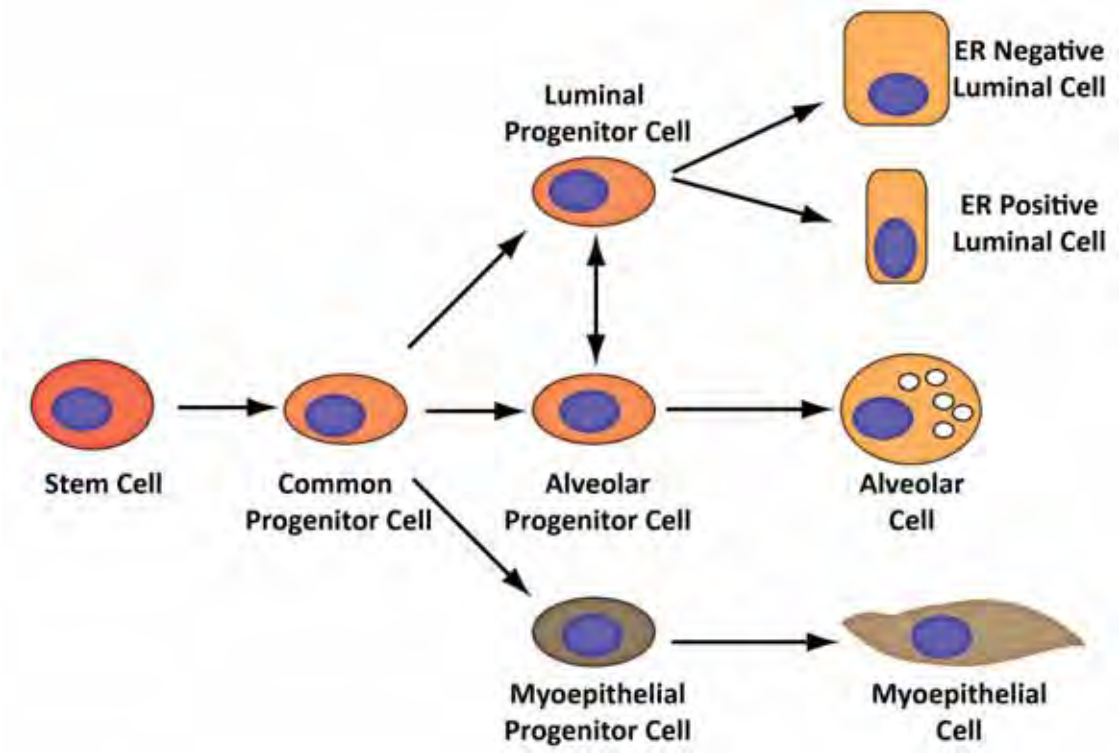


Figure 1-4. The mammary stem cell hierarchy.
Adaption of the stem cell hierarchy models of Jane Visvader and Gilbert Smith showing the differentiation of mammary epithelial cells down the luminal, alveolar, and myoepithelial lineages (Smith and Medina, 2008; Visvader, 2009).

1.2.4 Markers that delineate different mammary epithelial cell populations

The mature luminal and myoepithelial cells can be identified by their expression of various protein markers by immunohistochemistry (IHC) (Table 1-1). Some of the lower frequency stem and progenitor cell populations can be identified and enriched using various cell surface markers by flow cytometry (Table 1-2). Using cell surface markers to sort the different populations of mammary epithelial cells and subsequent analysis using *in vivo* transplantation experiments has demonstrated that only cells within the stem/myoepithelial fraction can repopulate a mammary gland and form new mammary epithelial outgrowths (Shackleton et al., 2006; Stingl et al., 2006). Currently used cell surface markers can distinguish mature luminal cells, luminal progenitor cells and stem/myoepithelial cells (Asselin-Labat et al., 2007; Shackleton et al., 2006; Stingl et al., 2006). There are currently no cell surface markers that can readily distinguish mammary stem cells from myoepithelial progenitor cells or mature myoepithelial cells.

IHC Marker	Cell type
CK5	Myoepithelial (Taddei et al., 2008)
CK6	ER/PR positive body and luminal cells (Grimm et al., 2006)
CK8	Luminal Cells (Guo et al., 2012)
CK14	Myoepithelial cells(Stingl et al., 2006)
CK18	Luminal cells(Stingl et al., 2006)
ER	Sensor luminal cells (Zeps et al., 1998)
PR	Sensor luminal cells (Beleut et al., 2010)
p63	Myoepithelial cells (Yalcin-Ozuysal et al., 2010)
αSMA	Myoepithelial cells (Guo et al., 2012)
ZO1	Luminal cells (Lee et al., 2011a)

Table 1-1. IHC markers of mammary epithelial cell subpopulations

Flow Marker	Cell type
CD24	Epithelial cells (Shackleton et al., 2006)
CD29	Stem/myoepithelial cells (Shackleton et al., 2006)
CD49f	Stem/myoepithelial cells (Stingl et al., 2006)
CD61	Stem/myoepithelial cells, Luminal progenitors (Asselin-Labat et al., 2007)
Kit	Luminal progenitors (Smart et al., 2010)
Sca1	ER positive luminal cells (Regan et al., 2011)
CD326/EpCAM	Epithelial cells (Stingl Personal Communication)

Table 1-2. Flow cytometric markers of mammary epithelial cell subpopulations

1.2.5 Factors that regulate mammary gland development

Mammary gland development is regulated at many levels, firstly at the whole body level by systemic hormones, secondly at an organ level by crosstalk between the epithelial and stromal cell types, thirdly at an epithelial level by crosstalk between different epithelial cell populations, and finally at the individual epithelial cell level by transcription factors, intracellular signalling proteins, and receptor expression. The focus of this introduction will be the global hormones, paracrine signalling molecules, stromal factors, and the transcriptional regulators. While intracellular signalling events are also important they are not as well elucidated. The factors that regulate mammary gland development are described below and some key ones are also summarised in Figure 1-5.

1.2.5.1 Hormonal regulation

Mammary gland development is primarily driven by the ovarian hormones oestrogen and progesterone, and the pituitary hormones prolactin and growth hormone. The oestrogen, progesterone, and prolactin receptors are only expressed by a subset of luminal cells known as sensor cells. The steroid hormones oestrogen and progesterone can freely diffuse across cell membranes, they are then bound by their receptor that then oligomerise, localise to the nucleus and regulate target gene expression (Htun et al., 1999; Mangelsdorf et al., 1995). Upon hormone signalling these cells then secrete a variety of soluble signalling molecules to induce proliferation of the surrounding luminal and myoepithelial cells (Brisken and O'Malley, 2010). Signalling by these soluble factors in conjunction with signals from stromal cells regulate proliferation and differentiation of the surrounding epithelial cells in a highly organised manner. Oestrogen is a critical regulator of mammary gland development and is the main hormone regulating pubertal development (Mallepell et al., 2006; McNally and Martin, 2011). Growth hormone (a pituitary hormone) also plays an important role in pubertal mammary gland development, and is thought to regulate insulin like growth factor (IGF) and fibroblast growth factor (FGF) levels in the mammary gland (McNally and Martin, 2011). Growth hormone receptor is not however required in the epithelium indicating its function is mediated through stromal cells (Gallego et al., 2001). Progesterone receptor signalling is also important although not necessary for pubertal development, however, progesterone along with prolactin control development during pregnancy, largely through the paracrine factors RANKL and Wnt ligands (Fernandez-Valdivia et al., 2009a; McNally and Martin, 2011; Watson and Khaled, 2008). During pregnancy levels of prolactin rise dramatically and induce the differentiation of mammary epithelial cells into milk secreting alveolar cells (Brisken et al., 1999). Prolactin receptor expression in the epithelium is necessary for this development (Ormandy et al., 1997). Recent studies have also suggested that oestrogen and progesterone do not just regulate growth of the mammary gland through proliferation signals but actively regulate both growth and differentiation signals in the stem and progenitor cells (Asselin-Labat et al., 2010; Joshi et al., 2010). While RANKL was proposed as the main candidate for this regulation, Wnt ligands are also likely to play a key role. This suggests that progesterone mediates a number of its effects by regulating the proliferation and differentiation of stem and progenitor cells via paracrine signalling molecules (Asselin-Labat et al., 2010; Joshi et al., 2010).

1.2.5.2 Secreted and Stromal Factors

Hormonal signalling to the sensor cells induces the production of a number of secreted factors that then regulate the proliferation and differentiation pathways of the surrounding cells

within the mammary gland in a paracrine manner. Hormones can also signal directly to the stromal cells (Mueller et al., 2002). These secreted factors include amphiregulin, RANKL, Wnt ligands, bone morphogenic proteins (BMPs) and TGFβ.

Amphiregulin (AREG) is an epidermal growth factor-like (EGF-like) growth factor that signals through the EGF receptor (EGFR), with the receptor composed of either an EGFR homodimer or heterodimer with ErbB2, ErbB3, or ErbB4 (McBryan et al., 2008). AREG is a crucial modulator of oestrogen signalling in the developing mammary gland, its absence leading to a failure of ductal elongation during puberty where the epithelial cells showed no proliferative response to oestrogen (Ciarloni et al., 2007). AREG is synthesised as a transmembrane protein that requires proteolytic cleavage to release the active protein by a member of the disintegrin and metalloproteinase family 17 (ADAM17 also known as TACE) (McBryan et al., 2008). Intriguingly it appears that the role of AREG is to signal to the surrounding stromal cells that then secrete factors that promote epithelial cell proliferation and regulate mammary gland development. This was shown using transplantation studies where loss of stromal EGFR recapitulated the AREG null phenotype but loss of epithelial EGFR had little effect on development (Wiesen et al., 1999).

Receptor of Activated Nuclear Factor-κβ Ligand (RANKL) is another important signalling molecule in the mammary gland downstream of hormone signalling. Unlike AREG though, RANKL is required for ductal side branching and alveolar development rather than pubertal development (Fata et al., 2000; Fernandez-Valdivia et al., 2009b). RANKL has been shown to be a progesterone target gene that is expressed by the ER/PR positive sensor cells in response to hormone signalling. RANKL then signals to the surrounding RANK positive epithelial cells (and possibly to the mammary stem/progenitor cells) to elicit its effects (Asselin-Labat et al., 2010; Fata et al., 2000; Joshi et al., 2010). Furthermore RANKL has been shown to be the key driver of progesterone signalling and via targeted epithelial cell overexpression was shown to compensate for the loss of progesterone signalling in the mammary gland (Mukherjee et al., 2010).

Wnt proteins also play an important role in paracrine signalling during mammary gland development and in particular regulate mammary stem cell proliferation and survival (Zeng and Nusse, 2010). Wnt-5a was shown to be crucial for pubertal development. Its expression is enriched in the TEBs where it acts to reduce lateral side branching (Roarty and Serra, 2007). Wnt-4 on the other hand functions downstream of progesterone signalling during alveolar development (Brisken et al., 2000). Wnt signalling is very complex with multiple receptors and

co-receptors leading to either canonical signalling through the transcriptional activity of β-catenin or through non-canonical signalling pathways that are β-catenin independent (Grumolato et al., 2010). Non-canonical Wnt signalling is modulated through one of three known pathways: the polarity pathway, the calcium flux pathway, and the atypical protein kinase C pathway (Grumolato et al., 2010). Wnt-4 is known to signal through the canonical pathway whereas Wnt-5a signals through a non-canonical pathway (Roarty and Rosen, 2010). The Wnt co-receptor LDL-receptor-related Protein 5 (Lrp5) is crucial for mammary stem cell and myoepithelial cell proliferation maintenance and its absence leads to a stunted ductal tree that cannot be serially transplanted (Badders et al., 2009; Lindvall et al., 2006).

BMPs and TGF-β are cytokines that have the ability to modulate growth and differentiation signals in the mammary gland. TGF-β is secreted in an inactive complex that is then activated by protease digestion, in the mammary gland it was shown that this activation is very tightly regulated only around a subpopulation of luminal cells during puberty, oestrus and pregnancy (Ewan et al., 2002). Furthermore BMP-4 but not BMP-2 or BMP-7 has been shown to be a key promoter of the proliferation of mammary epithelial cells *in vitro* in combination with EGF, fibroblast growth factors (FGF), or hepatocyte growth factor (HGF) (Montesano et al., 2008; Perk et al., 2005).

As well as the soluble factors discussed above stromal cells such as fibroblasts and immune cells as well as the extracellular matrix (ECM) are important regulators of mammary gland development (Coussens and Pollard, 2011; Maller et al., 2010; Wiseman and Werb, 2002).

The extracellular maxtix (ECM) plays a crucial role in regulating mammary development and function. The ECM gives structure to the tissue surrounding the mammary epithelial ducts and thus needs to be remodelled to allow for the growth of the ducts and alveoli. The ECM is composed of many different proteins including laminins, collagens, nidogens, perlecan, fibronectin, tenascins, SPARC, small leucine-rich proteoglycan, decorin, biglycan, and elastic fibres (Maller et al., 2010). The myoepithelial layer which is located adjacent to the ECM thus plays a crucial role in the interactions including growth and differentiation signals between the ECM and the luminal epithelial cells during mammary gland development (Faraldo et al., 2005). The ECM also plays a crucial role in binding and storing growth factors and cytokines such as TGFβ that can then be released in a controlled manner by specific proteases (Maller et al., 2010). The remodelling of the ECM by proteases such as the matrix metalloproteinases (MMPs) and ADAMs (a disintegrin and a metalloproteinase) and inhibited by TIMPs (tissue inhibitors of metalloproteinases) play a crucial role in mammary branching and ductal

outgrowth (Fata et al., 2004). Via integrins epithelial cells can sense changes in ECM composition such as changes in laminin concentration and ECM stiffness and this can have a dramatic impact on the cells phenotype (Naylor et al., 2005). For example loss of β1 integrin (CD29) by mammary epithelial cells led to a significant reduction in stem cells (Klinowska et al., 1999; Taddei et al., 2008). The mammary ECM is composed of several layers. Closest to the myoepithelial cells is the basal lamina which is about 20-100nm thick, next there is the intra- and interlobular stroma which is 10-100um thick, then there is the fibrous connective tissue which in humans can be up to centimetres think but is not as abundant in the mouse.

The stromal cells within the mammary gland are a mixture of immune cells, fibroblasts and adipocytes, all of which are required for mammary gland development and homeostasis. Macrophages, eosinophils and mast cells all play an important role in pubertal mammary gland development with the loss of any of these cell types leading to a marked deficit in ductal elongation (Coussens and Pollard, 2011; Gouon-Evans et al., 2002). Immune cells also play a role in mammary gland development during pregnancy, lactation, and involution (Coussens and Pollard, 2011). Immune cells are localised in a specific manner around the TEBs of developing ducts: macrophages are located around the base of the TEB near the mature ducts, eosinophils around the TEB and are further enriched where TEBs are starting to bifurcate, and mast cells are localised around the leading edge of the TEB (Coussens and Pollard, 2011; Gouon-Evans et al., 2000; Lilla and Werb, 2010). These immune cells can modulate ductal growth through the secretion of matrix remodelling agents and growth factors. Macrophages have also been demonstrated to play an important role in the turn-over of tertiary branches during the normal oestrus cycle, with a role both in promoting tertiary branching and phagocytosis of the apoptotic cells as the tertiary branches regress (Chua et al., 2010). It was also recently shown that macrophages express the oestrogen receptor and can respond directly to oestrogen at least in a breast cancer context (Iyer et al., 2012).

The exact role of fibroblasts is less well understood although they are known to play a role in producing the ECM surrounding the ducts and are also likely play a role in mediating growth factor signalling in response to amphiregulin (Maller et al., 2010). In particular fibroblasts are likely sources of factors from the BMP, IGF, FGF, and HGF families that are paracrine regulators of mammary gland development (McNally and Martin, 2011).

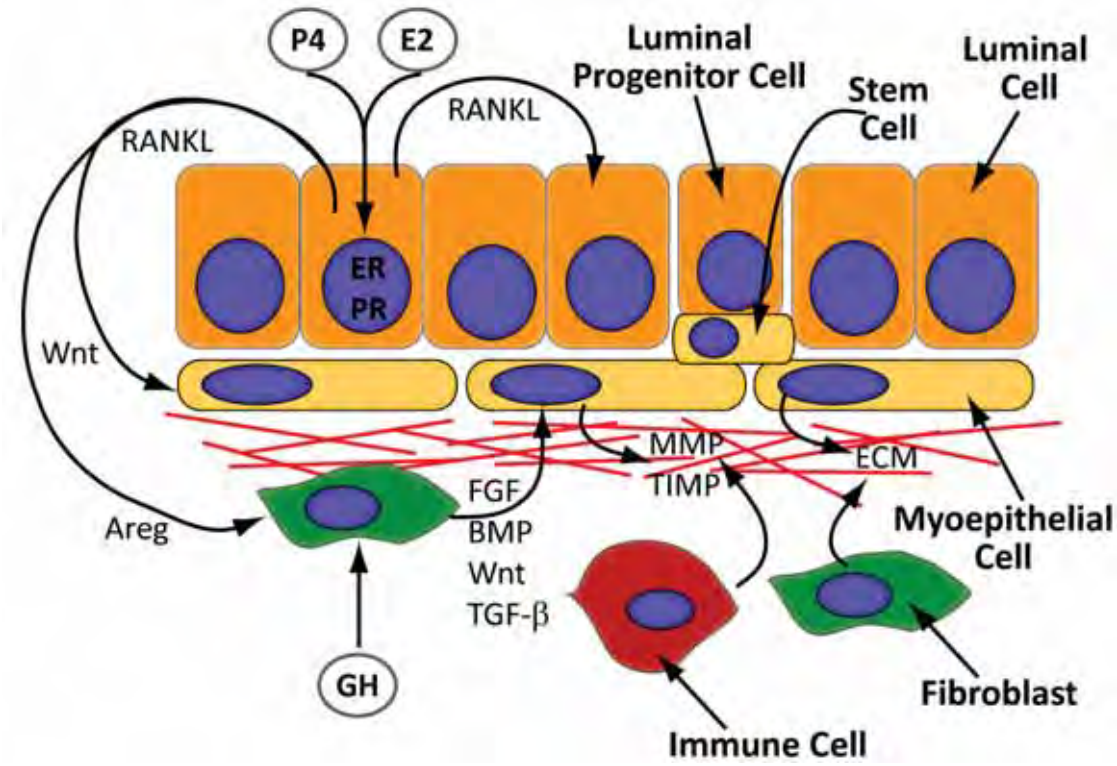


Figure 1-5. Factors and cell types that regulate mammary gland development.
Systemic hormones such as oestrogen (E2), progesterone (P4), and growth hormone (GH) regulate mammary epithelial differentiation and growth through a number of paracrine mechanisms. Paracrine factors include amphiregulin (Areg), RANKL, Wnt ligands (Wnt), fibroblast growth factor (FGF), bone morphogenic proteins (BMP), and transforming growth factor-beta (TGF-β). In addition the extracellular matrix (ECM), the ECM remodelling enzymes metalloproteinases (MMP) and their regulators the tissue inhibitor of metalloproteinases (TIMP).

1.2.5.3 Transcription factors and epigenetic regulators

All of these signalling pathways from global hormone levels, to the local milieu of ECM factors, growth factors and cytokines converge on the mammary epithelial cells and induce them to up regulate specific transcription factors and down regulate others. This then leads to proliferation, migration, or differentiation of the mammary epithelial cells. A number of transcription factors regulate the differentiation of mammary stem cells into the different mature mammary epithelial cell subtypes: ductal luminal, alveolar luminal, and myoepithelial. Most recently it was shown that by the ectopic expression of the transcription factors Slug and Sox9 that mature luminal cells could be forced to dedifferentiate back into multipotent mammary stem cells (Guo et al., 2012). This suggests a level of plasticity within mammary epithelial cell differentiation pathways that is not fully appreciated by the linear hierarchical structure that is normally ascribed to it.

A number of factors are known to regulate the mammary stem cell phenotype while others are known to promote luminal differentiation. Less is known about the factors that regulate myoepithelial cell differentiation. The polycomb group protein Bmi-1 is critical for the self renewal of multiple stem cell types. However, in the mammary gland despite the loss of Bmi-1 leading to a dramatic decrease in stem cell activity, its expression is actually highest in the luminal epithelial cell population (Pietersen et al., 2008). Pygo2 is a homeo domain containing factor that has been shown to be crucial for the expansion of mammary progenitor cells from stem cells, it does this through the epigenetic regulation of histone H3 (Gu et al., 2009). Other factors important in the early fate decisions from stem cells into either the myoepithelial or luminal progenitor cells are p63, along with the Hedgehog and Notch pathways. Hedgehog signalling appears to regulate the expression of the p63 isoforms ΔNp63 and TA-p63. ΔNp63 expression is crucial for the stem/myoepithelial cells and Indian hedgehog signalling induces a switch from this isoform to TA-p63 thus promoting luminal differentiation (Li et al., 2008). Notch signalling also leads to a down regulation of ΔNp63, however it is not currently known how the Notch and Hedgehog pathways interact to regulate mammary epithelial cell differentiation (Yalcin-Ozuysal et al., 2010). Further evidence for the important role of Notch in luminal differentiation was provided by Bouras and colleagues who demonstrated that blockade of the Notch pathway via knockdown of the Cbf-1 transcription factor caused a dramatic increase in stem cell activity, and conversely overexpression of a constitutively active form of Notch triggered precocious luminal development (Bouras et al., 2008). Interestingly p53 was demonstrated to promote stem cell differentiation through a mechanism that increased asymmetric division suggesting an important role for p53 in mammary stem cell homeostasis (Tao et al., 2010).

Factors specifically required for luminal cell differentiation and alveologenesis have been further characterised, these include FoxM1, BRCA1, GATA3, and Elf-5. BRCA1 and GATA-3 are key regulators of mammary epithelial differentiation and in their absence luminal differentiation is abrogated leading to an accumulation of luminal progenitor cells (Kouros-Mehr et al., 2006; Lim et al., 2009). FoxM1 expression is enriched in the luminal progenitor cells and it was recently demonstrated to inhibit luminal differentiation by repressing GATA-3 expression (Carr et al., 2012). Stat5a is a key controller of the luminal differentiation pathway and its absence causes a dramatic reduction of luminal progenitor cells and a large defect in alveologenesis (Yamaji et al., 2009). Whether this is due to increased differentiation down the ductal luminal pathway or reduced differentiation of stem to luminal progenitor cells is not clear (Yamaji et al., 2009). Interestingly SOCS2 up regulation by prolactin signalling functions

by inhibiting Stat5a activity suggesting it needs to be activated and then suppressed for normal alveologenesis (Harris et al., 2006). Elf-5 is a key target of prolactin signalling, is up regulated by an epigenetic mechanism during pregnancy and in its absence mammary glands are unable to form alveolar structures (Harris et al., 2006; Lee et al., 2011a; Oakes et al., 2008). Id2 is another key regulator of alveologenesis that acts downstream of RANKL signalling to promote the normal formation of alveoli (Kim et al., 2011).

The highly dynamic nature of the mammary gland with regular proliferative bursts probably contributes to the relative frequency of neoplastic conversion in this organ. I will now go on to discuss breast cancer.

1.3 Breast Cancer

1.3.1 Introduction

When proliferation and differentiation pathways in the mammary epithelial cells go awry whether it is due to specific gene mutations, aberrant gene silencing, or gene rearrangement or a combination of these, breast cancer may potentially develop. Breast cancer is a very common malignancy in Australia with more than 18 000 new cases diagnosed in Australia in 2008 (Ferlay et al., 2012). Screening programs have dramatically helped to reduce mortality from breast cancer (Blanks et al., 2000; Tabar et al., 2000). Furthermore the survival of people diagnosed with oestrogen receptor positive (ER+) breast cancer has improved dramatically over recent decades due to the success of targeted therapies, most notably the anti-oestrogen Tamoxifen and more recently the aromatase inhibitors such as letrozole. A percentage of these patients however will go onto relapse and patients whose cancer is oestrogen receptor negative (ER-) have a much poorer outlook. It is thus imperative to better understand these hard to treat breast cancers so that we can develop new therapeutics for them.

Breast cancer progresses in a step wise manner from atypical hyperplasia, to ductal carcinoma in situ, through to invasive breast cancer, and finally to metastatic disease (Allred et al., 2001). Breast cancer patients are treated with various combinations of surgery, endocrine therapy, chemotherapy and radiotherapy. Surgery can involve removal of the tumour and axillary lymph nodes or complete mastectomy. Endocrine therapy is used to treat patients that have oestrogen receptor positive breast cancers as described previously. Chemotherapy may be provided pre-surgery (neo-adjuvant chemotherapy) or post-surgery (adjuvant chemotherapy), and can comprise the following drugs: cyclophosphamide, methotrexate, fluorouracil,

doxorubicin, and paclitaxel. Radiotherapy is usually provided following surgery for several weeks (Brennan et al., 2005).

It is now well understood that the developmental processes that allow for the rapid growth of the mammary gland during puberty and pregnancy are corrupted during carcinogenesis. The most common of these developmental pathways to be corrupted is oestrogen signalling. This is demonstrated by the majority of breast cancers being ER positive and responding to anti-oestrogen therapy. Another corrupted pathway is related to the mutation of breast cancer associated gene 1 (BRCA1). Mutations in BRCA1 are the most commonly inherited mutations predisposing patients to breast and ovarian cancer (Wooster and Weber, 2003). At first the reason for the association of BRCA1 mutation with these specific types of cancer was unknown. This was because the function of BRCA1 was thought to be primarily in homologous DNA repair (Turner et al., 2004). However it is now known that BRCA1 also plays a key role in the mammary epithelial cell differentiation processes during development (Lim et al., 2009). Furthermore patients with mutations in BRCA1 have increased percentages of luminal progenitor cells (Lim et al., 2009). A number of other regulators of mammary gland development are also deregulated during tumour progression in the breast, these include GATA3 and FoxM1 (Carr et al., 2012; Kouros-Mehr et al., 2008).

1.3.2 Subtypes of Breast Cancer

Breast cancer is a very heterogeneous disease that has been divided into a number of subtypes. At first the subtypes were classified based on expression of the oestrogen receptor, progesterone receptor and the epidermal growth factor like receptor ErbB2 using immunohistochemistry. However in recent years the number of breast cancer subtypes has increased as we have shifted from only using an immunohistochemical approach to a whole genome mRNA expression approach using microarray profiling. These developments have led to diagnostic tests such as Oncotype DX that allow us to better select treatments for patients and also better understand a patient's response to current therapies (Kelly et al., 2012). Findings from these whole genome studies are leading to the better management of breast cancer and helping to determine novel targets for therapy. Based on this, breast cancer can be divided into a number of clinically relevant categories with the most common classifications as follows: luminal A, luminal B, Her2 amplified, basal-like, claudin-low, and normal-like (Prat and Perou, 2010; Sorlie et al., 2001). The relevance of these different subtypes of breast cancer is shown by the marked difference in survival seen between the subtypes (Figure 1-6) (Prat and Perou, 2010; Sorlie et al., 2001). Very recently a number of other systems for determining

breast cancer subtypes have been proposed, most recently based on gene copy number aberrations, leading to 10 subtypes of breast cancer (Curtis et al., 2012). While this new research may reshape the way breast cancer is classified in the future I will go on to discuss in more detail the established subtypes of breast cancer based on transcript profiling. It is also becoming clearer that the different subtypes of breast cancer are likely derived from different cells of origin within the mammary epithelial stem cell hierarchy (Lim et al., 2009; Visvader, 2009, 2011).

1.3.2.1 Luminal A subtype

The luminal A subtype is the most common type of breast cancer making up roughly 30% of all breast cancers, while the intrinsic subtype is based on a wide range of gene expression levels it is most often categorised based on ER and PR positivity and negativity for the up regulation of the Her2 pathway (Prat and Perou, 2010). This subtype of cancer is dependent on oestrogen signalling for its proliferation and these cancers initially respond well to anti-oestrogen therapy. However, unfortunately after some time these cancers often develop resistance to anti-oestrogen therapy (Cook et al., 2011). Luminal A tumours are thought to be derived from relatively differentiated ER positive luminal cells and generally have a more differentiated histopathological phenotype (Visvader, 2009).

1.3.2.2 Luminal B subtype

The luminal B subtype of breast cancer is categorised based on being ER and PR positive and also positive for a highly proliferative transcriptional signature that may include Her2 pathway activation (Prat and Perou, 2010). Luminal B cancers make up roughly 20% of all breast cancers. Breast cancers of the luminal B subtype are much more aggressive than cancers of the luminal A subtype and do not respond as well to anti-oestrogen therapy, however in cases where the Her2 pathway is activated they can also be treated with anti-Her2 antibody therapy (Trastuzumab) and/or a tyrosine kinase inhibitor (Lapatinib) (Higa and Abraham, 2007; Hudis, 2007). The luminal B subtype of breast cancer is thought to be derived from a less differentiated ER positive luminal cell compared to the luminal A subtype (Visvader, 2009).

1.3.2.3 Her2 Amplified subtype

The Her2 amplified subtype of breast cancer is highly aggressive and is characterised by overexpression or amplification of the ErbB2/Her2 gene and makes up roughly 15% of all breast cancers (Prat and Perou, 2010). A proportion of Her2 positive cancers respond well to Trastuzumab or Lapatinib (Higa and Abraham, 2007; Hudis, 2007). The exact cell of origin for

Her2 amplified breast cancers is not known but is presumed to be derived from either a luminal or alveolar progenitor (Visvader, 2009).

1.3.2.4 Basal-like subtype

The basal-like subtype is generally negative for ER, PR expression, and Her2 amplification, but has some characteristics of normal basal/myoepithelial cells such as basal cytokeratin expression. This subtype makes up roughly 15% of all breast cancers, is frequently associated with mutation of p53, it has no clinically validated targeted therapy and the poorest outcome of any breast cancer subtype (Sorlie et al., 2001). Despite no current targeted therapies, evidence from recent studies shows that patients with basal-like breast cancers tend to respond better to generic chemotherapy particularly when platinum based compounds are used (Leong et al., 2007; Liedtke et al., 2008; Rottenberg et al., 2007). It was recently shown that patients with mutations in the BRCA1 gene have an increased number of luminal progenitors and that these cells are likely candidates for cell of origin of breast cancers from these patients (Lim et al., 2009). Tumours from patients with BRCA1 mutations have a high degree of similarity to the basal-like subtype of breast cancer and thus it has been proposed that all basal-like breast cancers are derived from luminal progenitors (Lim et al., 2009).

1.3.2.5 Claudin-low subtype

Similar to the basal-like subtype the claudin-low subtype is negative for ER, PR and Her2, however it is further defined based on low expression of a number of claudin genes. Claudin-low tumours tend to be slower growing than the basal-like subtype, but also have no clinically approved targeted therapy and poor patient outcome compared to patients with Luminal A type breast cancer (Prat and Perou, 2010). The claudin-low subtype makes up roughly 10% of breast cancer patients (Prat and Perou, 2010). The transcript profile of claudin-low tumours has a significant overlap with the mammary stem cell enriched fraction of normal mammary epithelial cells and they have a number of characteristics of stem cells, they are thus thought to be derived from the mammary stem cell population (Visvader, 2009).

1.3.2.6 Normal-like subtype

The normal-like subtype of breast cancer appears most similar in gene expression to normal breast tissue, and often contain a low percentage of cancer cells. It is thus controversial as to whether it represents a true breast cancer subtype (Prat and Perou, 2010). Normal-like breast cancers make up roughly 5-10% of all breast cancers (Prat and Perou, 2010).

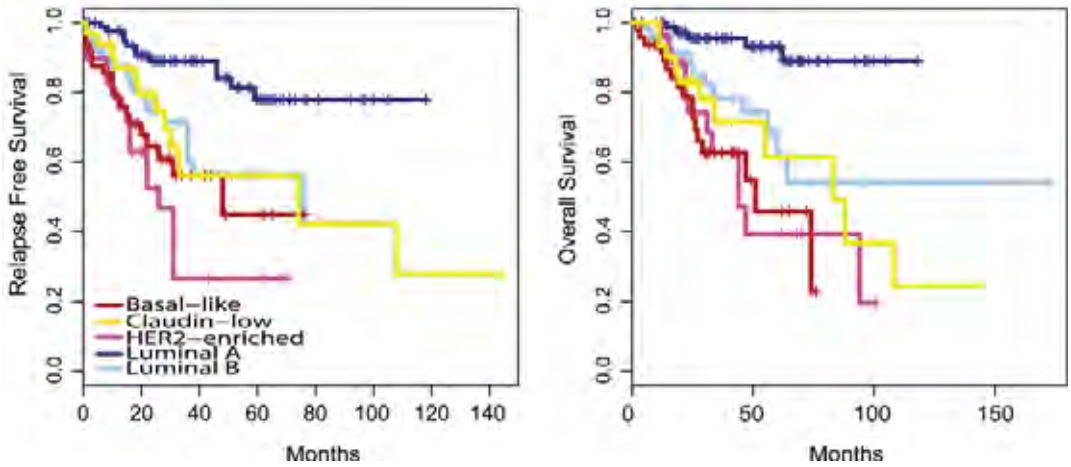


Figure 1-6. Kaplan-Meyer analysis of relapse free survival and overall survival l in a cohort of 337 breast cancer patients divided based on their molecular subtype. Modified from Prat & Perou 2010.

1.3.3 The cancer stem cell theory and breast cancer

The cancer stem cell (CSC) theory proposes that only a proportion of cells within a tumour can regrow a tumour following surgery, have the ability to metastasise, and are resistant to chemotherapy and these are known as CSC (Clevers, 2011). CSCs could be derived from normal stem or progenitor cells, however they could also be derived from more differentiated cells that have acquired stem like characteristics due to specific mutations. One reason for thinking that CSCs are derived from normal stem cells is twofold; firstly stem cells already have many characteristics of tumour cells, and secondly the long lived nature of normal stem cells allows them to acquire the multiple mutations required for malignant transformation. There is growing evidence that certain cancer types follow the cancer stem cell model while others do not (Gupta et al., 2009). Understanding whether CSCs are driving a particular tumour can be important for devising therapy. CSCs are considered resistant to current chemotherapy regimes and thus they can promote relapse after the bulk of the more differentiated cancer cells are killed or removed by surgery. Devising strategies to kill CSCs in combination with conventional therapies that remove the bulk of the tumour would allow for complete treatment of the cancer. Breast cancer is one type of cancer where there is accumulating evidence for the existence of a CSC (Korkaya et al., 2011).

1.4 Id proteins

The Id proteins are important for many developmental processes and are often deregulated in cancer. These proteins are the focus of this thesis and are discussed in detail below.

1.4.1 Overview

The inhibitor of DNA binding/differentiation (Id) proteins are a family of four proteins, namely Id1, Id2, Id3, and Id4. Id proteins are characterised by a helix-loop-helix (HLH) motif found in all basic HLH (bHLH) transcription factors. However Id proteins lack the basic DNA binding domain and are thought to act primarily by dimerizing with bHLH factors and preventing them from binding to their target genes (Figure 1-7) (Norton, 2000). Id proteins can also block the function of some helix-turn-helix factors such as the Ets family of transcription factors (Norton, 2000). Outside of the highly conserved HLH domain the different Id genes contain significant sequence divergence. However due to their overlapping expression pattern it is thought that one Id protein can compensate for the loss of another Id protein, in particular this appears to be the case with Id1 and Id3 (Norton, 2000). Id proteins are not completely redundant though as they have different affinities for the different bHLH factors (Liu and Harland, 2003).

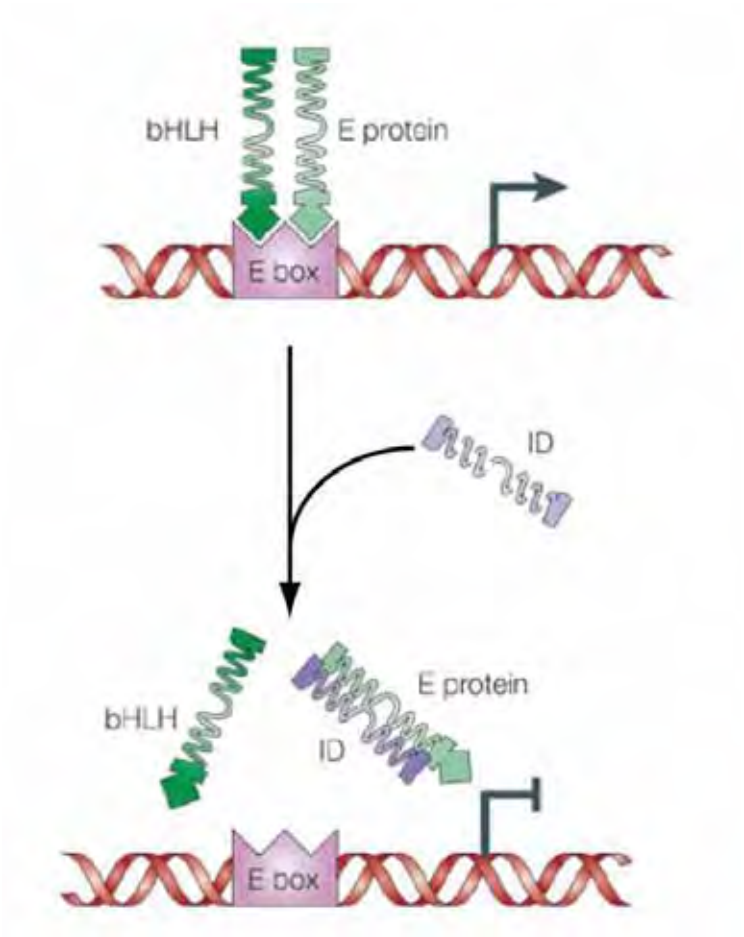


Figure 1-7. Schematic of the mode of action of Id proteins.
Id protein expression leads to heterodimerisation with a class of bHLH protein called the E proteins and the subsequent inhibition of target gene expression. Modified from Perk et al. 2005 (Perk et al., 2005).

1.4.1.1 Basic HLH Factors

The bHLH family is composed of 107 members in mice and 118 in humans, in mammals these have recently been divided into 5 clades based on their full amino acid sequence (Figure 1-8) (Skinner et al., 2010). bHLH factors are distinguished by two α-helices linked by a loop and a DNA binding domain that contains several basic amino acids (Norton, 2000). To function, they either homo or hetero-dimerize and bind to specific target DNA sequences known as the E box sequences (CANNTG), there they regulate transcription of their target genes (Desprez et al., 2003). All members of the bHLH family contain the highly conserved HLH dimerization domain. While some bHLH proteins are ubiquitously expressed such as the E-proteins E12 and E47, others are expressed in a tissue specific manner such as MyoD and NeuroD (de Candia et al., 2004; Ruzinova and Benezra, 2003). Basic HLH proteins are known to play an important role in stem cell homeostasis and development and some such as MyoD have the ability to specify cell fate (Davis et al., 1987).

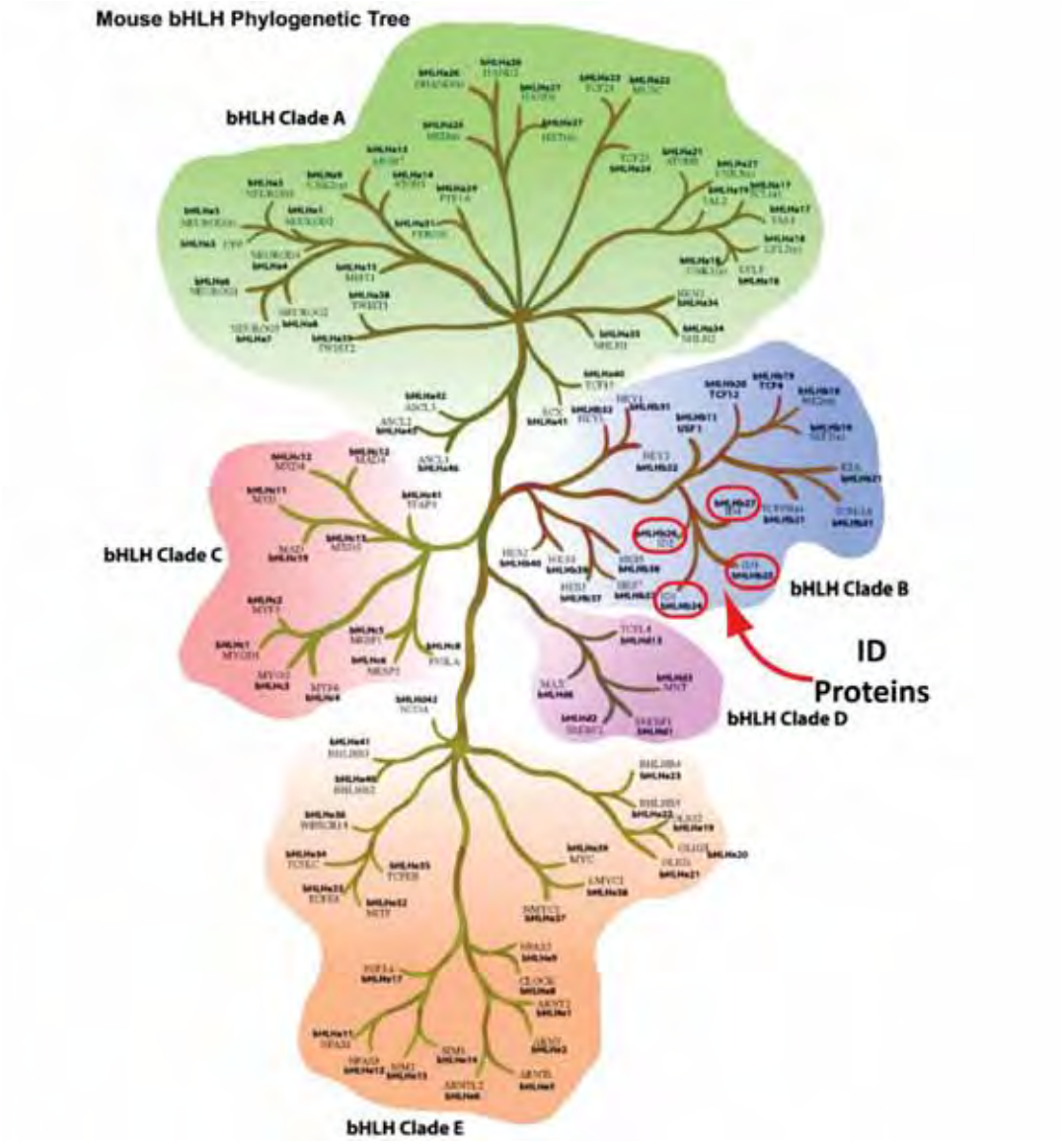


Figure 1-8. Phylogenic tree of mouse basic HLH Factors.
The Id proteins located in Clade B indicated in red. Modified from Skinner et al 2010 (Skinner et al., 2010).

1.4.2 Regulation of Id proteins

Numerous proteins including members of the transforming growth factor beta (TGFβ) superfamily, epidermal growth factor (EGF), and insulin-like growth factor-I (IGF-I) have been shown to induce Id protein expression *in vitro*, however which of these factors are important for Id activation *in vivo* is less certain (Yokota and Mori, 2002). Stability of the Id proteins seems to be promoted by dimerisation induced nuclear localisation, and some family members are also regulated by the relative levels of ubiquitinase and deubiquitinase enzymes such as USP1 (Williams et al., 2011; Yokota and Mori, 2002). Id4 the focus of my thesis however does not appear to be regulated by ubiquitin mediated degradation (Williams et al., 2011).

1.4.3 Id proteins in development

The Id proteins play an important role in a wide diversity of cell differentiation and developmental processes. The Id proteins are particularly important in the development of the various cell types of the brain and the immune system, but they have also been shown to have a role in mammary epithelial biology.

In the brain Id proteins regulate the generation of neurons and glia from common progenitor cells. Ectopic overexpression of Id1 in the embryonic cerebral cortex led to exclusive formation of glia cells at the expense of neurons (Cai et al., 2000). Overexpression of Id1 and Id3 *in vitro* also blocked neurogenesis, and in the Id1/Id3 double knockout mouse premature neuronal differentiation was observed (Ruzinova and Benezra, 2003). Id2 plays a similar role in blocking oligodendrocyte differentiation as shown by both *in vitro* overexpression and *in vivo* knockout studies (Wang et al., 2001). Like the other Id proteins Id4 also plays a role in neuronal differentiation pathways, this will be discussed in more detail in section 1.5.2.1 (Yun et al., 2004).

In the immune system Id proteins are expressed by hematopoietic stem cells and some progenitor cells (Murre, 2005). For example Id2 is crucial for NK cell differentiation from a common T cell/NK cell progenitor. Loss of Id2 also slows the differentiation of some dendritic cell subsets including splenic dendritic and Langerhans cells (Kee, 2009; Ruzinova and Benezra, 2003). Furthermore Id3 plays a role in both preBCR (preB cell receptor) and preTCR (pre T cell receptor) signalling in the developing B and T cells (Murre, 2005).

1.4.4 Id proteins in cancer

The Id proteins have been suggested as potential oncogenes due to their ability to inhibit differentiation and promote proliferation, however, it appears that their role in cancer is more complex (Perk et al., 2005). The Id proteins are often aberrantly expressed in cancer both at the mRNA and protein level, and Id expression often correlates with poor outcome (Lasorella et al., 2001). Elevated Id levels have been reported in a wide variety of both solid (Lin et al., 2000; Lyden et al., 1999) and haematological carcinomas (Bellido et al., 2003), as well as in sarcomas (Lasorella et al., 2001; Ruzinova and Benezra, 2003). However, analysing Id proteins at the whole tissue level can be complicated by their expression by multiple cell types including blood vessels and immune cells. This is particularly true for Id1 (Perk et al., 2005).

The functions of Id proteins in cancer are wide, varied and complex, for example in mouse models of glioma high Id1 expression was associated with high stem/self renewal capacity *in*

vitro but with lower proliferation and tumour formation capacity *in vivo* (Barrett et al., 2011). Up until recently the role for Id proteins in cancer had been determined by correlating *in vitro* studies with clinical expression levels and patient outcome data. It is however becoming clear that the role of the Id proteins in cancer need further investigation utilising both *in vitro* and *in vivo* systems to truly appreciate their mechanism of action.

1.5 Inhibitor of DNA binding 4 (Id4)

Id4 is the focus of the majority of my thesis. Id4 was the last Id protein to be identified, and has a expression pattern distinct from the other Id proteins (Riechmann et al., 1994).

1.5.1 Gene structure, expression, localisation

ID4 is localised to chr6:19,837,617-19,840,915 in humans and chr13:48,356,796-48,359,405 in mice and has a high degree of conservation across species particularly in the coding region (Figure 1-9) (Fujita et al., 2010). The *Id4* gene was initially cloned from mouse tissue followed by human tissue. The initial amino acid sequence analysis showed a 91% homology between the two species (Pagliuca et al., 1995; Riechmann et al., 1994). Subsequent analysis shows a higher degree of homology between human and mouse Id4 with 92.9% identity in the translated sequence and 82.4% identity in the 3' untranslated region, and that this led to a 98% identity at the amino acid sequence (Rigolet et al., 1998). The *Id4* gene is composed of three exons and two introns with the entire coding region localised within exons 1 and2 (Figure 1-9) (van Cruchten et al., 1998). *Id4* transcription is initiated within a 30bp region 300bp upstream of the translation start site (ATG) (van Cruchten et al., 1998). The *Id4* promoter contains an E box and two Sp1 binding sites, and these factors were shown to regulate the rate of *Id4* transcription (Pagliuca et al., 1998). Using nuclear extracts from HELA cells it was shown that upstream factor (USF) and bHLH-zip proteins could bind to the E box of Id4 (Pagliuca et al., 1998). The *Id4* gene is highly GC rich around the promoter and most of the open reading frame suggesting it could be regulated epigenetically by methylation (Rigolet et al., 1998). The *ID4* gene can produce four different transcripts in humans (1.1, 1.7, 2.4 and 3.8kb) and mice (1.1, 1.8, 2.7, and 4.2kb) either by differential splicing or more likely alternate polyadenylation start sites (Pagliuca et al., 1995; Riechmann et al., 1994; van Cruchten et al., 1998). Unlike the multiple transcripts of *Id4*, only one major transcript is seen for the other three Id proteins (van Cruchten et al., 1998). The alternative splicing or polyadenylation of *Id4* mRNA changes the length of the 3' UTR rather than the protein coding region (Riechmann et al., 1994; van Cruchten et al., 1998). Different organs, and tissues at different developmental stages express the four *Id4* transcripts to varying extents, this suggests a novel level of

regulation of the length of the 3' UTR. Furthermore, the ratio of the different transcripts was shown to change in Sertoli cells following different treatments, with cAMP treated cells having mainly the 2.7 kb and 1.8kb form while serum treated cells had mainly the 1.1kb form (Chaudhary et al., 2001). The purpose of the regulation of the length of the 3'UTR is not known but could indicate an additional level of regulation of Id4 translation or mRNA stability. ID4 has a distinct expression pattern during development compared to the other Id proteins suggesting that it has a unique and non-redundant role (Jen et al., 1996; Riechmann et al., 1994).

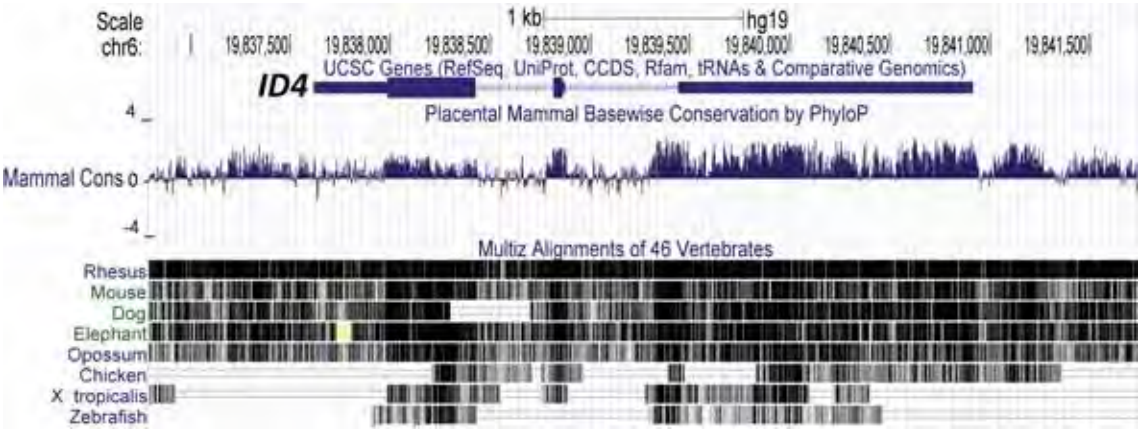


Figure 1-9. Structure and mammalian conservation of the human Id4 gene.
Id4 is made up of 3 exons separated by 2 introns. Exons are indicated by dark blue bars with the thickened sections indicating the translated region of the transcript. From the UCSC genome browser (Fujita et al., 2010).

1.5.2 The role of Id4 in stem and developmental pathways

Id4 has been demonstrated to play a role in stem and progenitor cell pathways during development in numerous tissues including the brain, testes, skin, bone, cartilage, fat, and in embryogenesis (Bedford et al., 2005; Huang et al., 2011; Jen et al., 1996; Murad et al., 2010; Oatley et al., 2011; Samanta and Kessler, 2004; Tokuzawa et al., 2010; Yun et al., 2004). This is discussed in detail below. The role of Id4 in mammary gland development is not well understood and is discussed in more detail in Chapter 4 "The role of Id4 in mammary gland development".

1.5.2.1 The role of Id4 in neural development

When examining Id4 null mice it was observed that they had a smaller brain size due to defects in neural differentiation and proliferation. Neurons in Id4 null mice differentiated precociously and displayed a reduced rate of proliferation at E12.5, which increased by E15.5 (Yun et al., 2004). Precocious differentiation of oligodendrocytes was also observed in the Id4 null mice,

associated with increased apoptosis (Marin-Husstege et al., 2006). During neurogenesis Id4 is induced by BMP4 signalling and interacts with the bHLH proteins OLIG1 and OLIG2 to inhibit oligodendrocyte differentiation from neural precursors, which is thought to allow for the normal differentiation of astrocytes (Cheng et al., 2007; Marin-Husstege et al., 2006; Samanta and Kessler, 2004). As oligodendrocytes differentiate, the levels of Id4 go down, and this can be enhanced *in vitro* by PDGF withdrawal (Kondo and Raff, 2000). Id protein and in particular Id4 expression thus regulates both oligodendrocyte and astrocyte development, with Id4 promoting astrocyte differentiation from neural progenitor cells whilst blocking oligodedrocyte differentiation. It is thus necessary for Id4 expression to be down regulated to allow for normal oligodendrocyte differentiation. Three mechanisms have been proposed for this down regulation during neural development in the brain. Firstly a role for the methyltransferase PRMT5 in regulating the methylation of a CpG island in the Id4 gene has been proposed (Huang et al., 2011). Secondly the transcription factor Yin Yang 1 represses the transcription of Id4 by recruiting histone deacetylase-1 to the Id4 promoter (He et al., 2007). Furthermore the transcriptional repressor Rp58 was shown to regulate all four Id proteins and that in its absence there was increased differentiation down the astrocyte pathway, that could be rescued by gene knockdown of all four Id proteins (Hirai et al., 2012).

1.5.2.2 The role of Id4 in the development of other tissues

As well as playing a crucial role in neural development Id4 plays an important role in other tissues. Firstly Id4 was shown to play a crucial role in adipocyte differentiation and adipose tissue formation, with loss of Id4 leading to impaired adipocyte differentiation and reduced adipose tissue formation as well as reduced weight gain when fed an high fat diet (Murad et al., 2010). However the converse has been found when mesenchymal stem cells differentiate either into bone marrow adipocytes or osteoblasts, in this case Id4 promoted adipocyte formation. This study showed loss of Id4 led to an increase in PPARγ and a switch from osteoblast differentiation towards adipocyte differentiation through the release of Hes1 from Hes1/ Hey2 complexes (Tokuzawa et al., 2010). This suggests the role of Id4 even in a particular cell type can be highly dependent on the microenvironment or other factors. Id4 also plays a role in normal homeostasis of the skin epidermis, with abnormal Id4 promoter methylation leading to an inhibition of differentiation in the epidermal cells and parakeratosis (Ruchusatsawat et al.). Finally the Id proteins including Id4 regulate the differentiation of the reproductive organs of both males and females, in particular they play a role in the function of Sertoli cells and spermatogonial stem cells (SSC) in the testes, and steroidogenic cells of the ovary, in response to GDNF, cAMP and TGFβ signalling respectively (Chaudhary et al., 2001;

Hogg et al., 2010; Oatley et al., 2011). Knockdown of Id4 in the SSCs using siRNAs led to an inhibition of proliferation, and the importance of Id4 in SSCs was further demonstrated by a rapid drop in fertility in Id4 null mice when compared to controls (Oatley et al., 2011).

1.5.3 The role of Id4 in cancer

Deregulation of Id4 has been associated with many different types of cancer. This dysregulation can be divided into several categories. Firstly up regulation of Id4 is associated with the cancer stem cell phenotype (CSC) particularly in certain types of brain tumours. Secondly in a small number of cases, Id4 is dysregulated at the genomic level by translocation or gene amplification. Thirdly Id4 silencing by promoter methylation is associated with poor outcome in numerous solid and haematological tumour types. The role of Id4 in breast cancer is controversial and will be discussed in more detail in Chapter 6 "The Role of Id4 in Breast Cancer".

1.5.3.1 The association of Id4 with the CSC phenotype and differentiation pathway

The enhanced ability to metastasise is one of the core phenotypes of CSCs. In a cohort of high grade prostate cancer patients, high nuclear Id4 expression in the primary tumour was associated with a higher risk of metastasis (Hazard ratio 3.215, P=0.026) (Yuen et al., 2006). Furthermore Id4 was the only Id protein to associate with outcome in prostate cancer despite Id1 and Id2 nuclear protein levels increasing significantly from nodular-hyperplasia to cancer (Yuen et al., 2006).

Elevation of Id4 levels has in particular been associated with the CSC or tumour initiating cell from glioblastoma multiforme (GBM). Id4 was upregulated in samples from GBM patients, and overexpression of Id4 in primary mouse Ink4a/Arf-/- astrocytes was sufficient to drive transformation (Jeon et al., 2008). In this mouse model it was shown that Id4 could dedifferentiate Ink4a/Arf astrocytes into neural stem-like cells as shown by the expression of Nestin, Tuj1 and NG2 markers. Furthermore Id4 was shown to activate Notch signalling and cyclin E expression, with Notch signalling being necessary for CSC development in this mouse model. Both Notch and cyclin E were confirmed to be crucial for neurosphere formation by human glioma cell lines (Jeon et al., 2008). Separate lines of evidence suggest that the increased stem cell phenotype in Id4 overexpressing GBM was due to a down regulation of the bHLH factors oligodendroglial lineage-associated factors Olig1 and Olig2 (Wu et al., 2012).

Another important characteristic of CSCs is increased resistance to chemotherapy, and it has recently been shown that Id4 promotes chemotherapy resistance through a multistep

mechanism that leads to the up regulation of the ATP-binding cassette (ABC) drug efflux transporters (Jeon et al., 2011). In this study Id4 was shown to down regulate miR-9* and this in turn allowed for an up regulation of SOX2 expression. SOX2 then promoted ABCC3 and ABCC6 expression and that this was necessary for resistance to apoptosis induced by the chemotherapeutics 1,3-bis(2-choroethyl)-1-nitrosourea (BCNU), etoposide, and staurosporine (Jeon et al., 2011).

As well as having a cell intrinsic role to confer a CSC like state to GBM cells Id4 also has an effect on the extracellular matrix and in promoting angiogenesis (Kuzontkoski et al., 2010). In this study GBM cell lines overexpressing Id4 grew much faster than controls when xenografted into mice but actually grew more slowly or at the same rate as controls *in vitro*. It was noted that the xenografts from GBM cells overexpressing Id4 were better vascularised and that a factor or factors secreted by GBM-Id4 cells *in vitro* could stimulate endothelial cell growth. Matrix GLA protein (MGP) was identified as a candidate of Id4 overexpression that plays a role in angiogenesis, and knocking down MGP suppressed some of the Id4 induced growth of GBM xenografts (Kuzontkoski et al., 2010). However, in a different murine glioma model PDGF-B driven oligodendroglioma, Id4 was suggested to be a tumour suppressor as Olig2 was shown to be necessary for malignancy and Id4 down regulates Olig2 (Appolloni et al., 2012). This study went on to show using publically available datasets that while Id4 was overexpressed in all glioma subtypes that higher Id4 levels were associated with increased survival (Appolloni et al., 2012).

It thus appears that overexpression of Id4 modulates tumourigenesis in a number of cancers and in several examples this is through the deregulation of the normal differentiation pathways that it normally regulates during development.

1.5.3.2 Genetic and epigenetic deregulation of Id4 expression in cancer

There are three reports of genetic alterations of Id4 being associated with cancer. The first example is from a patient with B-cell lineage acute lymphoblastic leukemia (ALL) where Id4 was translocated from chromosome 6 to chromosome 14 and abnormally expressed (Bellido et al., 2003). The second study finds that the chromosomal translocation identified by Bellido and colleagues is not an isolated case with 13 more cases found in ALL patients (Russell et al., 2008). In the third case a region of chromosome 6 where Id4 lies is commonly amplified in bladder cancer, however, the functional relevance of Id4 amplification in bladder cancer has not been demonstrated (Wu et al., 2012).

In addition to the relatively rare genetic deregulation of Id4 in cancer, Id4 appears to be epigenetically silenced in a wide variety of cancers. When the Id4 gene was characterised it was noticed that its promoter and most of the coding region was covered by a CpG island. CpG islands allow for the epigenetic regulation of gene expression through the methylation of the CpG dinucleotides which reduce gene expression (Momparler and Bovenzi, 2000). In cancer there is a global decrease in DNA methylation in conjunction with a gene specific increase in methylation of CpG islands, and a number of the genes silenced by hypermethylation turn out to be tumour suppressor genes (Hon et al., 2012; Momparler and Bovenzi, 2000). Id4 is aberrantly hypermethylated in numerous types of cancer including breast cancer (Noetzel et al., 2008; Umetani et al., 2005), gastric cancer (Chan et al., 2003), prostate cancer, leukemia (Chen et al., 2011; Wang et al., 2010; Yu et al., 2005), lung cancer (Castro et al., 2010), and colorectal cancer (Gomez Del Pulgar et al., 2008; Umetani et al., 2004). In leukemia the tumour suppressor function of Id4 has been tested by overexpressing Id4 in cell lines where Id4 is methylated and showing a decrease in growth both *in vitro* and *in vivo*, whereas most other studies have only correlated Id4 methylation with poor clinical outcome (Yu et al., 2005). In colorectal cancer it was shown that a Rho GTPase Cdc42 enhanced Id4 promoter methylation, however, no specific mechanism was identified (Gomez Del Pulgar et al., 2008).

These data indicate that depending on the context Id4 can either promote tumourigenesis through enhancing the stem cell phenotype or that it can act as a tumour suppressor gene and is epigenetically silenced to promote tumourigenesis. It also appears that Id4 can have different effects on different subtypes of the same cancer. The role of Id4 in breast cancer is the topic of Chapter 6 and is discussed in detail there.

1.6 Outline of thesis

The Id proteins are key regulators of developmental pathways that are often deregulated in cancer. The specific role of Id proteins in breast cancer and mammary gland development however is not well understood. During my doctoral studies I investigated the role of Id proteins in mammary gland development and breast cancer. In my first results chapter I describe a project that aimed to better understand the role of Id1 in promoting breast cancer. In particular I aimed to understand the role of Id1 in blocking senescence and the immune recognition of tumour cells utilising a mouse model of breast cancer. This research followed on from previously published results from our group, in which Id1 was shown to promote tumourigenesis and that its continued high expression was required to block senescence and tumour regression (Swarbrick et al., 2008). Increasing the ability of the immune system to

recognise tumour cells is a key goal of cancer immunotherapy, thus we sought to better understand the role Id1 played in this process. To do this we attempted to examine the changes in immune regulatory molecules and immune cells following the inducible down regulation of Id1. Unfortunately the model system established at the University of California San Francisco (UCSF) in the USA could not be re-established in our lab in Sydney. This unfortunately prevented any further investigation into this project.

I continued to investigate the role of Id proteins in breast cancer and mammary gland development, however I shifted my focus to investigate the role of Id4 in these processes. The role of Id4 in these processes is not well understood, however, there is significant evidence that Id4 does have an important role in their regulation (Beger et al., 2001; Fernandez-Valdivia et al., 2008; Fontemaggi et al., 2009; Noetzel et al., 2008; Umetani et al., 2005). My second results chapter describes my studies characterising Id4 protein expression in the normal mouse mammary gland and investigating the effect of Id4 loss on normal mammary development. The third results chapter looks to further investigate the role of Id4 in mammary epithelial cell function using the normal Comma-Dβ cell line as a model system for overexpression and knockdown studies. The final results chapter investigates the role of Id4 in breast cancer. The role of Id4 in breast cancer is controversial with some studies suggesting that it promotes tumourigenesis while others suggest that it acts as a tumour suppressor. Firstly we sought to better understand the association of Id4 with the different subtypes of breast cancer by correlating Id4 mRNA and protein expression in a panel of breast cancer cell lines. We then went on to investigate Id4 protein expression in cohorts of breast cancer patients and correlated this with patient outcome and other clinical parameters.

The overall aim of this thesis is to better understand the role of Id proteins in breast cancer. To do this we also need to understand the function of Id proteins in a developmental context. Understanding their role during development gives us insights into their function during tumour progression. Furthermore as breast cancer is not just one disease it is important to understand the subtypes of this disease that the individual Id proteins are most relevant to.

From these studies I have discovered that Id4 is necessary for normal pubertal mammary development where I show that it inhibits luminal differentiation and promotes proliferation. In breast cancer I show that Id4 overexpression is associated with the basal-like and Her2 amplified subtypes of breast cancer. I go onto show that high Id4 expression is also associated with improved patient survival within these subtypes of breast cancer. We also note that Id4 mRNA expression cannot always be used as a surrogate marker for Id4 protein expression.

Our studies of Id4 during development have increased our understanding of how Id4 functions both during mammary gland development as well as during breast carcinogenesis, and further conclude that deregulated Id4 expression can promote tumour progression.

The overall hypothesis of this thesis is that the Id proteins control networks of genes that regulate both cell intrinsic and extrinsic pathways during normal development and that these pathways can be corrupted during carcinogenesis.

The specific aims of the four results chapters are:

1. Determine the role of Id1 in regulating oncogene induced senescence and subsequent tumour clearance by the immune system.
2. Examine the role of Id4 in normal murine mammary gland development.
3. Determine the pathways regulated by Id4 and also it's transcriptional targets in mammary epithelial cells.
4. Determine whether Id4 protein expression has any clinical prognostic significance in breast cancer patients.

Chapter 2. Materials and Methods

2.1 Media and Buffers

Solution	Composition	Preparation and Storage
mMEC media	DMEM/F12 media (Gibco, Grand Island, NY, USA), 2%FBS (Thermo-Scientific), 10mM HEPES (Gibco), Penicillin/Streptomycin (Gibco), 1.25ml human insulin (Novo Nordisk, Bagsvaerd, Denmark), 5ng/ml mEGF (Sigma, St Luis, MO, USA)	4°C
Comma-Dβ media	DMEM/F12 media (Gibco), 2%FBS (Thermo-Scientific), 5ml 10mM HEPES (Gibco), 5ml Penicillin/Streptomycin (Gibco), 1.25ml human insulin (Novo Nordisk), 5ng/ml mEGF (Sigma)	4°C
293T/Phoenix Eco media	DMEM (Gibco), 10% FBS (Thermo-Scientific), 5ml Non-essential amino acids (NEAA), 5ml L-glutamine (Gibco), 5ml Sodium Pyruvate (Gibco)	4°C
NIH-3T3 media	DMEM (Gibco), 5% FBS (Thermo-Scientific)	4°C
CFC Assay media	Epicult-B (Stem Cell Technologies) media, 5% FBS (Stem Cell Technologies), 4μg/ml heparin (Sigma), 10ng/ml, β–FGF (Sigma), 10ng/ml mEGF (Sigma).	4°C
Eukaryotic cell freezing media	90% FBS (Thermo-Scientific) 10% Dimethyl Sulphoxide (DMSO, Sigma)	Made fresh.
Collagenase buffer	RPMI (Gibco), 2.5% FBS (Thermo-Scientific) 10mM HEPES (Gibco)	4°C
FACS buffer	PBS plus salts (Gibco), 2%FBS (Thermo-Scientific), 2%HEPES (Gibco)	4°C

Carmine staining solution	0.2% Carmine (Sigma), 0.5% aluminium sulphate (Sigma), 0.1% Thymol (Sigma)	Add 1g Carmine and 2.5g of aluminium sulphate to 500ml of distilled water, boil for 20 minutes, and allow to cool. Filter using vacuum filtration unit or filter papers. Add 50mg thymol and store at 4°C.
RIPA buffer	50mM Tris-HCl pH7.4, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 137.5mM NaCl, 1% glycerol, 0.5mM EDTA	Filter sterilise and store at 4°C
Tris buffered saline (TBS)	100mM Tris-Cl pH7.8, 1mM EDTA	Stored at room temperature.
TBS-Tween (TBST)	TBS 0.1% Tween 20 (USB Corp., Cleveland, OH, USA)	Stored at room temperature.

2.2 Mice

All experiments involving mice were performed in accordance with the regulations of the Garvan Institute Animal Ethics Committee. The $Id4^{GFP/GFP}$ mice were a gift from Kyuson Yun (Jackson Laboratory, Bar Harbor, Maine, USA) and were generated as previously described on the C57BL/6 background (Yun et al., 2004). Briefly, a construct containing GFP with flanking sequence from 5' and 3' of the Id4 gene was introduced into JM-1 embryonic stem (ES) cells by homologous recombination. The mice generated from these ES cells were thus null for Id4 and expressed GFP from the endogenous Id4 promoter. These mice were maintained on the C57BL/6 background and were also backcrossed five generations onto the FVB/N strain.

Wildtype C57BL/6, FVB/N, and BALB/C mice were sourced from the Australian BioResources Ltd. (Moss Vale, NSW, Australia). The NOD SCID $Il2\gamma^{-/-}$ (NOG) mice were a gift from Andrew Biankin (Garvan Institute, Darlinghurst, NSW, Australia). $Rag1^{-/-}$ mice were obtained from the Animal Resource Centre (Canning Vale, WA, Australia). Doxycycline food 700mg/kg (Gordon's Speciality Stock Feed, Yanderra, NSW, Australia) was given to the mice ad libitum during studies involving doxycycline induced switch-off of Id1.

2.2.1 Oestrus staging, Whittening, and timed mating

Oestrus staging was performed using vaginal swab (sterile 80mm FLOQSwabs from Copan Flock Technologies, Brescia, Italy) cytology. The vaginal swabs were transferred to Superfrost slides and stained using the DiffQuik kit (Polysciences Inc., Warrington, PA, USA) a modified

Giemsa stain with five dips into solutions A (fixative), B (stain 1) and C (stain 2). Slides were rinsed in tap water, dried and analysed by microscopy (Joshi et al., 2010).

To synchronise the oestrus stage the mice were whittened for 3 days. Whittening is achieved by the addition of dirty bedding from a male cage. On the following day the majority of mice should be in oestrus.

Timed matings were performed after the females had been whittened for 3 days. One male was housed with one female, and each morning the female was checked for a vaginal plug to determine the date of pregnancy.

2.2.2 Surgical procedures

Mice were anaesthetised using isoflourane (Baxter, Deerfield, IL, USA) anaesthetic machine, set on 5 for induction and 3 for maintenance. Anaesthesia was checked by a foot pad pinch check. Topical bupivacaine (AstraZenica, London, UK) 8mg/kg, and subcutaneous ketoprofen (Parnell, Sydney, Australia) 5mg/kg were used for analgesia. A small 1-2cm incision is made up the midline from in between the hind legs up to just below the ribcage cutting through the skin but not the peritoneum, a second ~1cm incision from the bottom of the midline towards the left or right hind leg was then made. The skin flap is then peeled back and pinned down to expose the 4th mammary gland. Clearing was performed by removing the portion of the mammary fat pad from the lymph node to the nipple. At 3 weeks of age the endogenous epithelium has not travelled pass the lymph node thus all endogenous epithelium is removed. Transplants were performed as described below sections 2.2.3 and 2.2.4. For double sided surgery this procedure was then repeated on the contralateral 4th mammary gland. The skin flaps were then closed with Autoclip 9mm wound clips (BD Primary Care Diagnostics, Sparks, MD, USA). Mice were weighed and transferred to a heat pad until they had recovered.

2.2.3 Mammary epithelial cell transplants

Single cell suspensions of primary mouse mammary epithelial cells (mMECs) (see section 2.6.1 for how these cells were prepared) or Comma-Dβ cells (see section 2.4.1) were resuspended in PBS containing magnesium and calcium salts (Gibco/Life Technologies) and then injected into the cleared 4th mammary gland of 3 week old recipient mice in a 10μl volume using a Hamilton syringe (Reno, NV, USA). In the case of mMECs retrovirally transduced with oncogenes, cells were injected into the 4th mammary gland without clearing (unless otherwise stated) of 3-8 week old recipient mice. Normal outgrowths were allowed to form for 5-8 weeks before the glands were harvested for wholemount histology. For the pregnancy studies after allowing the

transplants to engraft for 8 weeks the mice were time mated, and the glands were collected at D18 of pregnancy. For tumour studies: once tumours were palpable tumour growth was monitored using Vernier callipers and mice were sacrificed when mice reached ethical endpoint. Ethical endpoint was determined based on deterioration in body condition (hunched body position, ruffled coat), a greater than 20% loss of body weight, or a tumour of greater than $1cm^3$.

2.2.4 Mammary piece/mammary tumour transplants

Viably frozen ~$1mm^3$ mammary/tumour pieces were quickly thawed in a 37°C water bath, rinsed in PBS, and placed on ice. Mammary pieces were then transplanted into the cleared 4th mammary gland of 3 week old recipient mice. Tumour pieces were then transplanted into 4th mammary gland of 8-12 week old recipient mice. To perform these transplants a small pocket was created in the mammary fat pad with a sharp pair of scissors and then the mammary/tumour piece was placed into the pocket. Tumour growth was monitored using Vernier callipers and mice were sacrificed when they reached ethical endpoint.

2.2.5 *In vivo* imaging

Mice bearing tumours transduced with GFP expressing vectors were monitored using the IVIS *in vivo* imager (Xenogen) using Living Image 3.1 software. GFP fluorescence was captured using a 465nm excitation wavelength and a GFP emission filter, background fluorescence was determined using a 430nm excitation wavelength and GFP emission filter. GFP fluorescence intensity was analysed and quantified using the Image Math feature in Living Image 3.1 software after changing the units from counts to efficiency and where a background region of interest (ROI) is subtracted from the GFP signal. The area covering the tumour was first shaved to allow for better visualisation of the GFP signal.

2.2.6 Tissue collection

Depending on the experiment the following sample types were collected from the mice.

For tumours:

A cross section of the tumour was placed in 10% neutral buffered formalin (Australian biostain, Traralgon, VIC, Australia).

A cross section of the tumour was placed in optimal cutting temperature (OCT) media (Sakura Finetek, Alphen aan den Rijn, Netherlands) on dry ice. Then stored at -80°C

2-3 cryo vials containing 5-10 small 1mm^3 pieces of tumour were snap frozen in liquid nitrogen. Then stored at -80^{o}C

2-3 cryo vials containing 5-10 small 1mm^3 pieces of tumour and 1ml of viable freezing media (7% DMSO in Collagenase buffer). Placed into a 'Mr Frosty' and stored at -80^{o}C before being transferred to liquid nitrogen.

For flow analysis tumour pieces were kept in PBS plus salts (Gibco) on ice and processed as described in section 2.6.2.

For mammary glands:

One 4th mammary gland was wholemounted as described in section 2.9.1.

One 4th mammary gland was placed in 10% neutral buffered formalin (Australian biostain).

One 2/3rd mammary gland was placed into OCT (Sakura Finetek) on dry ice and stored at -80^{o}C.

One 2/3rd mammary gland was cut into 10-20 small 1mm^3 pieces of tumour and transferred into 2-3 cryo vials containing were snap frozen in liquid nitrogen. Then stored at -80^{o}C

Mammary glands for transplants:

From a 4th mammary gland 5-10 small 1mm^3 pieces of mammary gland and 1ml of viable freezing media (7% DMSO in Collagenase buffer) were transferred into 3 cryo vials. Placed into a 'Mr Frosty' and stored at -80^{o}C before being transferred to liquid nitrogen.

For single cell preparations the 3rd, 4th, and 5th mammary glands were collected from 3-8 mice and were processed as described in section 2.6.1.

2.3 Plasmids and siRNAs

The plasmids used in these studies are described in Table 2-1, and the shRNA sequences contained in them are summarised in Table 2-2. Id4 and control siRNAs were purchased commercially (Dharmacon, Lafayette, CO, USA).

Construct	Antibiotic selection	Insert	Backbone	Obtained from
pMig-Ras	Ampicillin	Oncogenic H-Ras carrying an Alanine to Valine substitution at codon 12	pMig – MSCV IRES GFP	pMig was obtained from Yosef Refaeli (National Jewish Medical and Research Center, Denver, CO, USA).
pRQ-Id1	Ampicillin	C-terminal HA-tagged murine Id1	Id1 was cloned downstream of the TREtight promoter (Clontech). pRQ also contains a PGK-PuroR cassette for selection.	The pRQ vector was a kind gift of Drs Alana and Bryan Welm.
pMig-tTA	Ampicillin	Tetracycline transactivator (tTA)	pMig – MSCV IRES GFP	pMig was obtained from Yosef Refaeli (National Jewish Medical and Research Center, Denver, CO, USA).
pMSCV-Id4-DSRed	Ampicillin	Murine Id4	pMSCV IRES DSRed	From Kyuson Yun (Jackson Laboratory, Bar Harbor, ME, USA)
pMSCV-DSRed	Ampicillin	DS Red control	pMSCV IRES DSRed	From Kyuson Yun (Jackson Laboratory, Bar Harbor, ME, USA)
pSLIK-shId4-venus	Ampicillin	Id4 shRNA sequence was determined using the RNAi codex website (http://cancan.cshl.edu/cgi-bin/Codex/Codex.cgi)	pSLIK –Tet inducible mIR30 based shRNA backbone, upstream of a constitutive rtTA IRES and venus fluorescent protein cassette.	pSLIK backbone from ATCC (**Manassas, VA,USA). pSLIK-shId4-venus from Wee Teo** (Garvan Institute, Darlinghurst, NSW, Australia)
pLV4311-IRES-Thy1.1	Ampicillin	IRES-murine Thy1.1 control vector	pLV 3rd generation lentiviral vector	From Brian Rabinovich (The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA)
pLV4301-IRES-EGFP	Ampicillin	IRES-EGFP control vector	pLV 3rd generation lentiviral vector	From Brian Rabinovich (The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA)
pCL-Eco	Ampicillin		ecotropic retroviral packaging plasmid	From Bob Navaiux (The Salk Institute, San Diego, CA, USA).
pMDLg/pRRE, pRSV-Rev, pMD2.G	Ampicillin		3rd generation lentiviral packaging vectors	From Addgene (Cambridge, MA, USA).

Table 2-1. Summary of plasmids used in this study.

Vector/siRNA	shRNA/siRNA sequence
pMISSION shId4 1443	CCGGGTGCGATATGAACGACTGCTACTCGAGTAGCAGTCGTTCATATCGCACTTTTTG
pMISSION shId4 1444	CCGGGCACGTTATCGACTACATCCTCTCGAGAGGATGTAGTCGATAACGTGCTTTTTG
pMISSION shId4 1445	CCGGCGCCGTGAACAAGCAGGGTGACTCGAGTCACCCTGCTTGTTCACGGCGTTTTTG
pMISSION shId4 1446	CCGGGTGCCTGCAGTGCGATATGAACTCGAGTTCATATCGCACTGCAGGCACTTTTTG
pMISSION shId4 1447	CCGGGCTTTGCTGAGACAGCCGCCACTCGAGTGGCGGCTGTCTCAGCAAAGCTTTTTG
pMISSION shCont	CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTT
pMISSION shGFP	CCGGTACAACAGCCACAACGTCTATCTCGAGATAGACGTTGTGGCTGTTGTATTTTT
pSLIK-shId4	Sense: AGCTGTGCCTGCAGTGCGATA Antisense: TCGACACGGACGTCACGCTAT
siId45, #J-04687-05	GCAUUCACCGUACGUAUUC
siId4 #6, J-04687-06	GUAGAAGAGCAAUCGUGAA
siId4 #7, J-04687-07	GUACUGGUCUUGCAUGAUU
siId4 #8, J-04687-08	GAAACGUCGUCUCGUCUUG
siCont #2, D-001810-02	UGGUUUACAUGUUGUGUGA
siCont #4, D-001810-04	UGGUUUACAUGUUUUCCUA

Table 2-2. siRNA and shRNA sequences used in the Id4 knockdown studies.

2.3.1 Transformation of competent bacteria

The following competent cells were used for the different types of vectors: DH5α E.coli cells for regular vectors, Max Efficiency Stbl2 (Invitrogen) E.coli cells for retroviral vectors, and Max Efficiency Stbl3 (Invitrogen) E.coli cells for lentiviral vectors. DH5α competent cells were transformed by heat shock treatment. Briefly, 100μl of cells was thawed on ice in 15ml polypropylene tubes then mixed with 50ng of plasmid and placed back on ice for 30 minutes. Tubes were then incubated in a 42°C waterbath for 45 seconds and cooled on ice for 2 minutes. 900μl of SOC media (Invitrogen) was added and cells incubated for 1 hour at 37°C shaking at 225rpm. 20μl and 200μl of cells were then plated onto LB Agar plates containing the appropriate antibiotic, and incubated overnight at 37°C. Single colonies were then picked the next day. Max Efficiency Stbl2 cells (Invitrogen) and Max Efficiency Stbl3 cells (Invitrogen) were transformed as per the manufacturer's instructions. Briefly Stbl2 cells were transformed

as DH5α cells except cells were heat shocked for only 25 seconds and were incubated at 30°C instead of 37°C. Stbl3 cells were transformed as DH5α cells except only 250µl of SOC media is added following the heat shock.

2.3.2 Plasmid DNA preparations

Wizard plus mini-prep (Promega, Madison, WI, USA) and Maxi-prep (QIAGEN, Hilden, Germany) DNA extractions were performed as per manufacturer's instructions. DNA was resuspended in nuclease free water (Promega) and quantitated using a Nanodrop ND1000 spectrophotometer (Thermo-Scientific, Waltham, MA, USA)

2.4 Cell culture

2.4.1 Cell growth conditions

The following cell lines were maintained in culture: Comma-Dβ, Phoenix-Eco, 293T, NIH-3T3 fibroblasts, and primary mouse mammary epithelial cells (mMECs). All cells were maintained as described here with a few modifications for specific cell lines described below: cells were maintained in a 37°C humidified 5%CO_2 incubator, they were routinely passaged while subconfluent every 3-4 days. To passage cells, cells were washed once with PBS (Gibco), the incubated with 0.05% Trysin/EDTA (Gibco) for 10-15 minutes at 37°C. Trypsin was inhibited with growth media and cells were transferred to a new flask at appropriate split ratios.

The mouse mammary epithelial cell line Comma-Dβ were a gift from Joseph Jeffery (University of Massachusetts, Amherst, MA, USA). Comma-Dβ cells retain a more normal phenotype when grown at high density and were passaged so that they would reach confluence within 2 days of splitting. Cells grown at low density readily generate abnormal clones of this cell line. Cells were transferred to a new flask at a dilution of between 1:6 and 1:8.

Phoenix-Eco and 293T cells are very loosely adherent and thus were not rinsed with PBS prior to trypsinisation and only required 5 minutes of trypsinisation at room temperature.

Following trypsinisation primary mMECs were transplanted back into mice and not subcultured.

2.4.2 Cryopreservation of mammalian cells

Following centrifugation, the supernatant was removed and the cells were resuspended gently in freezing media to a final concentration of $1x10^6$-$1x10^7$ cells/ml. 1ml aliquots were transferred to Cryo vials (Corning, NY, USA) labelled with cell line name and passage number. Cryovials were transferred to a "Mr Frosty" (Nalgene/Thermo Scientific) that controls the rate

of cooling to 1°C/minute which is then placed at -80°C. Cells are transferred to liquid nitrogen for storage after 24 hours.

2.4.3 Proliferation assay

Cellular proliferation was measured using the Cell Titre 96 AQ kit (Promega). 500 Comma-Id4 and Comma-DSRed cells were seeded (replicates of 6) into 96-well plates along with 6 wells of media alone as a negative control. On days 1, 2, 4, 7 and 9 days following seeding, the 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) and PMS reagents were mixed (20:1) and 20μl added to each well of a plate. The plate was incubated for 3 hours at 37°C and then absorbance at 490nm was measured using the FLUOstar Optima plate reader with Optima software (BMG LabTech, Ortenberg, Germany).

2.4.4 *In vitro* doxycycline treatment

293T cells were plated at $6x10^5$cells/well of a 6-well plate and the next day 2 wells were transiently transfected using Effectine (QIAGEN) as per manufacturer's instructions with 0.6μg of each of the following combinations of plasmids: pRQ-Id1 and pMig-tTA, pRQ-Id1 and pMir-tTA, or pMIG-Id1 alone. One well of each transfection combination was treated with doxycycline 1μg/ml. The next day the media was changed, and the following day protein lysates were made from the cells.

Comma-Dβ-pSLIK-shId4-venus cells were plated at $1x10^5$ cells/well of a 6-well plate, the next day 1mg/ml of doxycycline was added to half of the wells. Protein lysates were then collected 24hrs, 48hrs, and 72hrs after doxycycline addition.

2.4.5 *In vitro* differentiation of Comma-Dβ cells

Comma-Dβ cells were seeded at $0.8x10^5$ cells /well (or for acute lentiviral knockdown studies at $1.1x10^5$ as the infection process slows the cells growth) into 6-well plates, 72 hours later the media was removed and the cells were washed once with PBS and then media without mEGF was added to the cells. 24 hours later (Day 0 of the assay) the media was replaced with mEGF-free Comma-Dβ media containing 0.5μg/ml prolactin (Sigma) and 1nM dexamethasone (Sigma) or fresh mEGF-free media as a control. Media was replaced daily with fresh media containing prolactin and dexamethasone. RNA and protein lysates were collected at Day 0, Day 2, and Day 4.

2.4.6 Non-lethal irradiation of NIH-3T3 cells

NIH 3T3 cells were trypsinised while 50-70% confluent and counted. Cells were resuspended at 10^6 cells/mL in FACS buffer and irradiate with 50 Grays using an X-ray machine. After

irradiation the cells were centrifuged and resuspended in freezing medium and frozen at a concentration of 1.7 x10^6cells/vial at -80^{o}C.

2.4.7 Transfection of Comma-Dβ cells with siRNA

Comma-Dβ cells were seeded at 0.8x10^5 cells/well into 6-well plates with 2ml of Comma-Dβ media without penicillin and streptomycin. On the same day as seeding siRNA constructs were transfected into the cell using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions at between 5-50nM. The media was changed the next day and cells were harvested between 48 and 120 hours after transfection for protein lysates.

2.5 Production of viral supernatants and viral transduction of primary mMECs and cell lines

2.5.1 Retroviral production

Low passage number Phoenix-Eco cells were plated out at a density of 1.8x10^6 cells into 6cm dishes. 16-24 hours later the cells were transfected with the retroviral vectors (2μg) and the retroviral packaging plasmid pCL-Eco (1μg) using Effectine transfection reagent (QIAGEN) in OPTIMEM (Invitrogen) following the products standard transfection protocol (QIAGEN). 16-24 hours later the media was changed. 48 hours post transfection virus containing media was collected from the cells and fresh media was put back on the cells. At 72 hours post transfection the virus containing media was again collected. The viral supernatant was purified by passing it through a 0.45μm filter and the virus was either used immediately or frozen in aliquots at -80^{o}C.

2.5.2 Lentiviral production

Low passage number 293T cells were plated out at a density of 1.5x10^6 cells into 6cm dishes. 16-24 hours later the cells were transfected with the lentiviral vectors (1μg) and the lentiviral packaging plasmids pMD2.G (1.5μg), pRSV-REV (2.25μg), and pMDLg/pRRE (2.25μg) using Lipofectamine 2000 (Invitrogen) in OPTIMEM (Invitrogen) following the products standard transfection protocol (Invitrogen). Media was changed 16-24 hours post transfection. Virus containing media was then collected 48 hours post transfection. The viral supernatant was purified by passing it through a 0.45μm filter and the virus was then concentrated using Amicon-Ultra 15 concentrators (Millipore, Billerica, MA, USA). 8mls of viral supernatant was concentrated to ~1.5ml. Concentrated virus was frozen in aliquots at -80^{o}C.

2.5.3 Retroviral transduction of primary mammary epithelial cells

Mammary epithelial organoids that had been prepared as per section 2.6.1 were infected 24 hours after plating into 6-well plates. Infection was performed using a spinfection protocol as follows: media was removed from the mMECs and replaced with 2ml of fresh viral supernatant and 8μg/ml polybrene (Sigma). The plates were then spun at 1400rpm for 1 hour. The viral supernatant was then removed and replaced with mMEC media. This spinfection procedure was repeated the next day. Infection efficiency was determined by fluorescence microscopy for GFP.

2.5.4 Retroviral transduction of Comma-Dβ cells

Comma-Dβ cells ($1.1x10^5$) were seeded into a 6-well plate. 16-24 hours later the cells were infected with pMSCV-Id4-DSred or pMSCV-DSred retrovirus diluted 1:10 in Comma-Dβ media with 8μg/ml polybrene. 24 hours later the media was changed. DSred positive cells were then FACS enriched using the BD FACSAria fluorescence activated cell sorter and BD FACSDIVA software.

2.5.5 Lentiviral transduction of Comma-Dβ cells

Comma-Dβ cells ($1.1x10^5$) were seeded into a 6-well plate. 16-24 hours later the cells were infected with the various pMISSION viruses or pLV4301 virus diluted 1:10 into Comma-Dβ media with 8μg/ml polybrene. 24 hours later the media on the cells was replaced with fresh media. Infection efficiency was determined by fluorescence microscopy on the pLV4301 infected cells. pMISSION shRNA infected cells could be selected for using 4mg/ml puromycin (Sigma) for one week.

2.6 Preparation of single cell suspensions

2.6.1 Mouse mammary epithelial cell preparations

The 3rd, 4th and 5th mammary glands were removed from 3-8 euthanized mice and pooled. Glands were minced mechanically using scalpels followed by 5 straight razors taped together. Minced glands were digested in Collagenase buffer (RPMI, 2.5%FBS, 10mM HEPES) 5ml/g of mammary gland with 1mg/ml Collagenase blend L (Sigma) for 90mins in a 37°C shaking incubator. Cells were pelleted for 10 minutes at 1200rpm and the supernatant removed. Cells and organoids were resuspended in PBS containing 50ug/ml DNase (Sigma) for 3 minutes, pelleted again (5 minutes at 1200rpm) and resuspended in PBS+salts. To enrich for epithelial organoids, cells were subjected to 3 rounds of differential centrifugation by centrifuging the tubes up to 1000rpm and then rapidly stopping and discarding the supernatant. For studies

involving the retroviral transduction of mMECs the organoids were plated out onto 6-well plates with organoids from 2-3 donor mice/6-well plate. For studies requiring single cell suspensions the following steps were performed. The pelleted organoids were then resuspended in 0.017% Trypsin/EDTA (Gibco) for 15 minutes with pipetting every 3 minutes. Trypsin was then neutralised using PBS containing 2% FBS. Cells were then incubated with 1-3ml of Dispase (Stem Cell technologies, Vancouver, BC, USA) for 5 minutes at 37°C. Samples were diluted in PBS+salts (Gibco) and then passed through a 40μm cell strainer, pelleted, resuspended in 1-5ml of FACS buffer and counted with a haemocytometer using trypan blue (Gibco) to exclude dead cells.

2.6.2 Mouse mammary tumour cell preparation for the analysis of immune cell infiltrates

Tumour bearing mice were euthanized and then intra-cardiac PBS perfused using a blunt 20 gauge feeding needle (Fine Science Tools, North Vancouver, BC, Canada) into the left ventricle while also cutting the right atrium with scissors. Mice were perfused with 10ml of PBS using a syringe to remove circulating blood leukocytes from the tumour. The tumour was then removed, minced using scalpels and scissors, and then digested for 30 minutes in a shaking incubator at 37°C in 2mg/ml collagenase A (Roche, Basel, Switzerland) and 2U/ml DNase (Sigma) in 30mls of serum free DMEM. FCS (Thermo Scientific) 5ml was added to stop the digestion and the digested tumour was passed though a 70μm cell strainer. Cells were pelleted by centrifugation for 5 minutes at 1200rpm. Cells were then passed through a second 70μm cell strainer, pelletted (5 minutes at 1200rpm) and resuspended in FACS buffer.

2.7 Colony forming capacity (CFC) assay

Single cell suspensions of mammary epithelial cells (5000 cells) were resuspended in 4ml of Epicult-B (Stem Cell Technologies) media containing 5% FCS (Stem Cell Technologies), 4μg/ml heparin (Sigma), 10ng/ml, β–FGF (Millipore), and 10ng/ml mEGF (Sigma). Irradiated NIH-3T3 fibroblasts $1.6x10^6$ were resuspended in 4mls of the same media. The mammary epithelial cells were mixed with the irradiated fibroblasts and plated into two 6cm dishes. 16-24 hours later the media was replaced with the same media except containing 1% FCS. Five to six days later the media was removed then the cells washed with PBS and fixed and stained using a modified Geimsa stain the DiffQuik kit (Polysciences Inc.). Colonies were visualised and counted using the Leica MZ12 dissecting microscope.

2.8 Flow cytometry and FACS

2.8.1 Flow cytometry for immune cells infiltrating mammary tumours

Primary mouse mammary tumour cells were incubated for 10 minutes with anti-CD16/CD32 antibody (1:200 BD Biosciences, San Jose, CA, USA) in FACS buffer (PBS plus salts 2%FBS, 2%Hepes) to block non-specific antibody binding by immune cells. Cells were then pelleted and resuspended in FACS buffer containing the following combinations of antibodies. All antibodies were from eBioscience (San Diego, CA, USA), clones are listed in brackets:

1. CD45-APC-AF750 (30-F11), and CD31-PE (390)
2. CD11b-APC-AF750 (M1/70), CD11c-PE (N418), and F4/80-APC (BM8)
3. CD11b-APC-AF750 (M1/70), Gr1-PE (RB6-8C5), and F4/80-APC (BM8)
4. CD3-APC-AF750 (17A2), CD4-PE (CK1.5), and CD8-APC (53-6.7)
5. CD3-APC-AF750 (17A2), and CD49b-PE (DX5)
6. CD45-APC-AF750 (30-F11), B220-PE (RA3-6B2), and CD19-APC (MB19-1)
7. CD4-APC-AF750 (CK1.5), CD8-APC (53-6.7), CD62L-PerCP (MEL-14), and CD44-PE (IM7)

All antibodies were diluted 1:200 in FACS buffer. Cells were incubated with the antibodies for 20 minutes on ice, pelleted, washed twice with FACS buffer then resuspended in FACS buffer containing 7AAD (1:10) except for the sample containing CD62L-PerCP. Flow cytometry was then performed on a BD FACSCANTO II using BD FACS DIVA software, and the results were analysed using Flowjo software (Treestar, Ashland, OR, USA).

2.8.2 Flow cytometry for mammary epithelial stem cell populations

Single cell suspensions of primary mouse mammary epithelial cells were incubated for 10 minutes with anti-CD16/CD32 antibody (1:200 BD Biosciences) in FACS buffer (PBS plus salts 2%FBS, 2%Hepes) to block non-specific antibody binding by immune cells. Cells were then pelleted and resuspended in FACS buffer containing the following lineage markers: anti-CD31-biotin (1:40 BD Biosciences, Clone: 390), anti-CD45-biotin (1:100 BD Biosciences, Clone:30-F11), anti-TER119-biotin (1:80 BD Biosciences, Clone: TER119), and anti-BP- biotin (1:50 eBiosciences, Clone: 6C3) for 20 minutes on ice. Cells were then pelleted and resuspended in FACS buffer containing streptavidin-APC-Cy7 (1:400 BD Biosciences) and the following epithelial stem cell markers anti-CD24-PE-Cy7 (1:400, BD Biosciences, Clone: M1/69), anti-CD29-Pacific Blue (1:100 Biolegend (San Diego, CA, USA), Clone: HMβ1-1),and anti-CD61-APC (1:100 Invitrogen, Clone: HMβ1-1), and incubated for 20 minutes on ice. Cells were then washed twice in FACS buffer before being resuspended in FACS buffer containing DAPI (1:1000

Invitrogen). Flow cytometry was then performed on a LSRII SORP using BD FACS DIVA software, and the results were analysed using Flowjo software (Treestar).

2.8.3 Fluorescence activated cell sorting (FACS)

Cells were prepared as described above and resuspended at $1x10^7$ cells/ml in FACS buffer. FACS was performed in the Garvan Institute flow cytometry facility on a BD FACS Vantage sorter with BD FACS DIVA software.

2.9 Histological methods

2.9.1 Wholemount histology

Mammary glands were dissected and whole mounted onto glass slides at the indicated ages. These were then fixed in 10% neutral buffered formalin overnight, the fat was removed with several changes of acetone and the ductal network was stained with Carmine alum overnight. Glands were then dehydrated through graded alcohols (70%, 95% and 100%), cleared using Slidebrite, then stored and imaged using a Leica MZ12 dissecting microscope in methyl salicylate.

2.9.2 Tissue processing for IHC

Tissue was fixed overnight at 4°C in neutral buffered formalin (Australian Biostain) rinsed in PBS then transferred to Unisette tissue processing/embedding cassettes (Simport, Beloeil, QC, Canada) and stored in 70% ethanol at 4oC. Samples were then transferred to the Garvan Institute Histology facility for tissue processing and embedding.

2.9.3 Immunostaining

Immunohistochemistry (IHC) and immunofluorescence (IF) were performed on 4μm sections of formalin fixed paraffin embedded tissue. Antigen retrieval was performed using DAKO (Glostrup, Denmark) antigen retrieval solutions s1699 or s2368 either for 20 minutes in a boiling waterbath (WB), or in a pressure cooker (PC). IHC was performed on the DAKO Autostainer. Briefly, slides were rinsed with DAKO wash buffer, incubated with DAKO peroxide block for 5 minutes, rinsed again, then incubated with DAKO protein block for 30 minutes. Primary and secondary antibody conditions are summarised in Table 2-3. Slides were then incubated with DAKO DAB+ reagent for 10 minutes and then rinsed in water. Slides were then counterstained with haematoxylin for 2 minutes for human tissue and for 20-30 seconds for mouse tissue. Slides were then dehydrated through graded alcohols (70%, 95% and 100%), cleared in xylene, and mounted using Ultramount #4 (Fronine, Riverstone, NSW, Australia).

Staining was analysed and imaged using a Leica DMRB microscope, with DCF420C camera and Leica Aplication Suite V3 software.

IF was performed manually following the same procedure as IHC except fluorescent secondary antibodies were used and slides were then incubated with DAPI 1:100 in PBS for 5-10minutes. All fluorescent secondary antibodies were obtained from Jackson ImmunoResearch (West Grove, PA, USA). Slides were then mounted with Vectorshield (Vector Labs, Burlingame, CA, USA) mounting media. Slides were analysed using either a Ziess upright fluorescent microscope or a Leica inverted confocal microscope.

Immunocytochemistry was performed on freshly fixed cells grown in 4- or 8-well chamber slides (BD Biosciences). Briefly, cells were washed once with PBS (Gibco), fixed with 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA, USA) in PBS for 15 minutes at room temperature, and followed by 4% paraformaldehyde/0.02% Tween 20 for another 15 min to permeablise the cells. Cells were washed with PBS 3 times for 5 minutes each. Non-specific binding was blocked by incubating the cells in 5% BSA/PBS/0.1% Tween or Protein block (DAKO) for 1 hour. After blocking cells were incubated with primary antibodies (diluted in blocking solution) for 1-2 hours. The concentration of antibodies was the same as for IHC/IF as described in Table 2-3. Cells were washed as before and incubated for 1 hr with the fluorescently labelled secondary antibodies (1:100) in the dark. Cells were again washed and then incubated with DAPI 1:100 in PBS for 5-10minutes. The chambers were then removed and the slides mounted with Vectorshield (Vector Labs) mounting media.

Antigen	Manufacturer and clone/catalogue number	Antigen retrieval	Antibody dilution and incubation time	Secondary detection	Fluorescent secondary detection
Human/ Mouse Id1	Biocheck (Foster City, CA, USA) BCH 1/195-14	PC 1699 5min	1:50, 60 min	Envision Rabbit 30 min	NA
Mouse CD45	BD Biosciences 30-F11	WB 1699 20min	1:500, 90 min	Rabbit anti-Rat-HRP 30 min/Envision Rabbit 15 min	NA
Mouse F4/80	Serotec (Oxford, UK) CI:A3-1	WB 1699 20min	1:200, 90 min	Rabbit anti-Rat-HRP 30 min/Envision Rabbit 15 min	NA
Mouse Neutrophils	Cedarlane Labs (Burlington, Ontario, Canada) 7/4	WB 1699 20min	1:500, 90 min	Rabbit anti-Rat-HRP 30 min/Envision Rabbit 15 min	NA
Human/ Mouse CD3	DAKO A0452	WB s2368 20min	1:100, 90 min	Envision Rabbit 30 min	NA
GFP	Abcam (Cambridge, UK) ab6556-25	WB 1699 20min	1:50, 60 min	Envision Rabbit 30 min	NA
Human/ Mouse Id4	Biocheck 9/82-12	PC 1699 1min	1:400, 90 min	Envision Rabbit 30 min	α-Rabbit-Cy3
Mouse CK5	Covance (Princeton, NJ, USA) PRB-160P	WB 1699 20min	1:1000, 60 min	Envision Rabbit 30 min	NA
Mouse CK6	Covance PRB-169P	WB 1699 20min	1:200, 60 min	Envision Rabbit 30 min	NA
Mouse CK8	DSHB (Iowa City, IA, USA), TROMA1	WB 1699 20min	1:500, 60 min	Rabbit anti-Rat-HRP 30 min/Envision Rabbit 15 min	α-Rat-Cy2
Mouse CK14	Covance PRB-155P	WB 1699 20min	1:2000, 60 min	Envision Rabbit 30 min	NA
Mouse SMA	Abcam ab5694	WB 1699 20min	1:50, 60 min	Envision Rabbit 30 min	NA
Mouse p63	Novus (Littleton, CO, USA), NB100-691	WB 1699 20min	1:100, 60 min	Envision Rabbit 30 min	α-mouse-Cy2
Mouse αSMA-FITC	Sigma F3777	PC 1699 1min	1:200, 60 min		

Table 2-3. IHC/IF conditions

2.9.4 Senescence associated β-galactosidase staining

Staining was performed on 10μm frozen OCT tumour sections using the senescence associated β-galactosidase kit (Cell Signalling Technologies, Beverley, MA, USA) following the manufacturer's instructions.

2.10 Protein methods

2.10.1 Protein extraction

Cells grown in monolayer were lysed using RadioImmune Precipitation Assay (RIPA) buffer containing Complete mini EDTA-free protease inhibitor cocktail (Roche), 100μM vanadate, and 20μM MG132. 120ul of ice cold lysis buffer was added to each well of a 6-well plate, cells were scraped down on ice and the lysates were transferred to a microcentrifuge tube, incubated on ice for 5 – 10 minutes and then spun at 14 000 rpm at 4°C for 5 minutes. The supernatant from the lysate was then aliquotted and stored at -20°C until analysed.

50mg of snap frozen small tumour pieces were transferred to a microcentrifuge tube on dry ice. These were ground on dry ice using a 1.5ml pellet pestle (Lomb Scientific/Thermo Scientific) for roughly 10 seconds until all large pieces are broken up. 400μl of RIPA buffer with protease inhibitors as described above was added then the sample was sonicated on ice for a total of 20 seconds with 2 second pulses and 0.5 second pauses. If sample had not fully broken down sonication was repeated. Samples were incubated on ice for 5 – 10 minutes and then spun at 14 000 rpm at 4°C for 5 minutes. The supernatant from the lysate was then aliquotted and stored at -20°C until analysed.

2.10.2 Quantifying protein concentration

Protein concentration was determined using the micro bicinchoninic acid (BCA) kit (Pierce, Rockford, IL, USA) following the manufacturer's instructions. Briefly, the BSA standard (2mg/ml) was serially diluted in PBS to generate the following dilution range (2μg/ml-0.06125μg/ml), a 10μl aliquot of lysate was thawed and diluted 1:6 with PBS, and the samples and standards were transferred in duplicate 25μl volumes into a 96-well plate. The BCA reagents were then mixed (50:1 Part A:Part:B) and 200ul was added to each well. Absorbance at 560nm was analysed using the FLUOstar Optima plate reader with Optima software (BMG LabTech).

2.10.3 Western blotting

Protein lysates (5-20μg) were prepared with 4x loading buffer (Invitrogen) and sample reducing agent (Invitrogen), denatured by heating to 70°C for 10 minutes and loaded onto 4-

12% bis/tris gels (Invitrogen). Gels were run at 200V for approximately 40 minutes using MES running buffer (Invitrogen). Proteins from the gels were then transferred to Hybond ECL nitrocellulose membranes (Amersham Biosciences, Germany) using the Invitrogen western blotting module and transfer buffer for 1-1.5 hours at 30V. Membranes were then rinsed in TBST and blocked for 1hr at room temperature (or overnight at 4°C) in either TBST 5% skim milk powder (or TBST 2%BSA for the anti-milk and anti-phospho protein western blots). Antibody dilutions for western blotting are summarised in Table 2-4. Western blot bands were visualised using Western lightning Plus ECL reagent (Perkin Elmer, Waltham, MA, USA), and Fuji SuperRX film (Tokyo, Japan).

Protein detected	Primary antibody and concentration	Secondary antibody and concentration
Human/Mouse Id4	Biocheck #BCH 9/82-12 1:25 000	Cell Signalling α-Rabbit #7074 1:2 000
Human/Mouse Id1	Biocheck #BCH 1/195-14 1:500	Cell Signalling α-Rabbit #7074 1:2 000
Human Ras	Calbiochem #OP38 1:100	GE Healthcare α-Mouse #NA931V, 1:5000
HA-tag	Roche #12CA5 1:1000	GE Healthcare α-Mouse #NA931V, 1:5000
Mouse Milk proteins	Accurate Chemical & Scientific Corp #YNRMTM 1:10 000	Cell Signalling α-Rabbit #7074 1:2 000
Mouse p38 MAPK	Cell Signalling #9212 1:1000	Cell Signalling α-Rabbit #7074 1:2 000
Mouse phospho-p38 MAPK	Cell Signalling #9211 1:1000	Cell Signalling α-Rabbit #7074 1:2 000
Human/Mouse β-Actin	Sigma #A1978 1:200 000	GE Healthcare α-Mouse #NA931V, 1:5000

Table 2-4. Antibody dilutions and secondary detection antibody for western blotting.

2.11 RNA methods

2.11.1 RNA extraction

RNA was extracted from cells using either Trizol (Ambion/Life Technologies) or the RNAeasy Minikit (QIAGEN) following the manufacturer's instructions. RNA was eluted or resuspended in nuclease free water (Promega). RNA was extracted from mammary glands using the RNAeasy Minikit as follows. 20-30mg of snap frozen mammary gland pieces were ground in microcentrifuge tubes on dry ice using a 1.5ml pellet pestle (Lomb Scientific/Thermo Scientific) for roughly 10 seconds until all large pieces are broken up. 600μl of buffer RLT was added then

the sample was sonicated on ice for a total of 20 seconds with 2 second pulses and 0.5 second pauses. RNA extraction was then continued as per the QIAGEN RNAeasy minikit protocol.

2.11.2 RNA analysis

All RNA preparations were checked for quality and concentration using the Nanodrop spectrophotometer. For routine RT-PCR RNA was considered of sufficient quality if the 260/280 was >1.8 and the 260/230 was >1.8. For transcript profiling the RNA was further analysed using the Agilent 2100 Bioanalyser using a RNA Nano chip (Agilent Technologies) and following the manufacturer's instructions. The Bioanalyser generates a RNA integrity score with a maximum of 10. Samples with a score of >9 were considered of sufficient quality to be used for microarray analysis.

2.11.3 RNA Microarray analysis

RNA (500ng) was sent to the Ramaciotti Centre (UNSW, Kensington NSW, Australia) for hybridisation to Affymetrix gene expression mouse 1.0 ST microarray chips (Affymetrix, Santa Clara, CA, USA). Array data was analysed using GenePattern software (Broad Institute, Cambridge, MA, USA) utilizing the following modules. Data was normalised using the NormalizeaffymetrixST module. Changes in gene expression were determined using the LimmaGP module, these were then visualised using the ComparativeMarkerSelectionViewer module, the HeatmapViewer module, and the HierarchicalClusteringViewer module. Gene set enrichment was performed using the GSEAPreranked module. Overlapping gene sets were determined using the VennDiagram module. For all GenePattern analyses the default settings were used.

2.11.4 cDNA synthesis

cDNA was synthesised from 0.5-1μg of RNA using the Superscript III RT-PCR kit (Invitrogen) using oligo-dT primers and following the manufacturer's instructions.

2.11.5 Quantitative RT-PCR/Microarray analysis

Gene expression and the Taqman probe (Applied Biosystems/Life Technologies) used to analyse it are listed in Table 2-5. All assays were set up using an EPmotion 5070 robot and run on an ABI PRISM 7900 HT machine. PCR cycling follows the standard Taqman assay protocol and is as follows: 1 cycle at 95^{o}C for 10 minutes, followed by 40 cycles of 95^{o}C for 15 seconds and 60^{o}C for1 minute. Data is removed from the machine in a tab delineated format and analysed using Microsoft Excel.

Gene	Taqman assay
Mouse Id4	Mm00499701_m1
Mouse CK8	Mm00835759_m1
Mouse Elf-5	Mm00468732_m1
Mouse β-Casein	Mm00839664_m1
Mouse WAP	Mm00839913_m1
Mouse PTPRN	Mm01258989_m1
Mouse MMP9	Mm00442991_m1
Mouse β-Actin	4352341E
Mouse GAPDH	4352339E
Human Id4 (designated Id4 M1)	Hs00155465_m1
Human Id4 (designated Id4 G1)	Hs02912975_g1
Human GAPDH	4326317E

Table 2-5. Genes analysed and the corresponding Taqman assay used to analyse their expression level.

2.12 DNA methylation analysis

DNA from breast cancer cell lines for methylation analysis was obtained from the Garvan Institute Cancer Program Cell line DNA bank. Methylation analysis was performed by Dr Shalima Nair from the Epigenetics group within the Cancer Program (Garvan Institute, Darlinghurst, Australia). DNA was bisufite converted and analysed using a PCR and the Sequenom MassArray technology as previously published (CIark et al., 1994; Coolen et al., 2007). Primers used for this analysis are outlined below Table 2-6. The regions of the Id4 gene analysed are described in Chapter 6, Figure 6-1A. As a positive control serological DNA (Millipore) was fully methylated after treatment with Sss1 (Millipore). As a negative control whole genome amplified DNA (Millipore), prepared using the Sigma WGA amplification kit, where human genomic DNA is fragmented and amplified leaving it unmethylated.

Primer set	Forward sequence	Reverse sequence
Id4 #1	GAGAGCGTAGTGGAGGAGG	CGAATATCCTAATCACTCCCTT
Id4 #2	GTTCGGAGTTTGTTTGTTTTTTTC	GACTACGAAACCACCCAAACTATAAC
Id4 #3	TGGTGGTTGTTTTAGTAGGGT	AAAACTAATACCCACCATCCC
Id4 #4	GGGAATTAGGGAAGGTAGAAGAGTT	CTCTCCCTACCACCTAAATAACAAA

Table 2-6. Primer sequences for Sequenom DNA methylation analysis.

Chapter 3. The role of Id1 in modulating the immune response to Ras driven mammary tumourigenesis

3.1 Introduction:

For cancers to develop, several mutations or epigenetic changes are required. Some will result in activation of growth promoting genes (oncogenes) and others will involve the suppression of growth control genes (tumour suppressors). As described in the introduction, there are several oncogenes and tumour suppressor genes, however, they function by affecting a smaller number of pathways. In normal cells aberrant activation of an oncogene leads to growth arrest, and then either senescence or apoptosis. Other mutations or epigenetic changes are then required to overcome these growth suppressive pathways. The focus of this chapter will be on oncogene-induced senescence and the recognition of senescent tumour cells by the immune system.

3.1.1 Oncogene induced senescence

Oncogene induced senescence (OIS) is a tumour suppressor pathway that is activated by a number of oncogenes and can block tumourigenesis. Unlike cellular quiescence a reversible cell cycle arrest, senescence is considered an irreversible growth arrest in the G1 phase of the cell cycle that involves chromatin remodelling and changes in gene expression, however, the cells remain metabolically active (Larsson, 2011). Senescence is also associated with changes in cellular morphology, senescent cells tend to have a larger flattened morphology and contain a large number of vacuoles within the cytoplasm (Larsson, 2011). Gene expression changes following senescence include the increased expression of senescence-associated β-galactosidase (SA-β-gal) and an increase in secretion of certain cytokines known as the senescence associated secretion phenotype (SASP) (Larsson, 2011; Lee et al., 2011b). OIS was first described in cultured fibroblasts overexpressing oncogenic K-Ras^{V12} but there is accumulating evidence that it plays a crucial role in blocking tumour formation *in vivo* (Reddy and Li, 2011; Serrano et al., 1997). Human melanocytic nevi (moles), which commonly express oncogenic mutant BRAF, are growth arrested and express markers associated with senescence such as SA-β-gal and p16INK4A (Michaloglou et al., 2005), suggesting that senescence acts a barrier to progression of nevi to malignant melanoma. A number of mouse models of cancer, particularly those driven by oncogenic Ras, also show signs of premalignant lesions that express senescence markers (Braig et al., 2005; Guerra et al., 2003). Other studies have shown the importance of OIS in prostate tumours driven by PTEN loss and aberrant PI3K-Akt signalling

(Chen et al., 2005; Reddy and Li). OIS is induced by three main signalling pathways, the first being the p16/Cyclin D/pRb pathway, the second being the Arf/Mdm2/p53/p21 pathway, and the third being the DNA damage response pathway that also converges on p53/p21 (Larsson, 2011). These pathways include two of the classic tumour suppressor genes, namely p53 and pRb. Mutations in p53 and pRb are a common mechanism for evading OIS, however, there are numerous other regulators of these senescence pathways that can also be mutated or suppressed to overcome OIS and these have not all been well characterised (Larsson, 2011).

3.1.2 Role of Id1 in breast cancer

Id1 is a member of the Inhibitor of Differentiation/DNA binding (Id) family of proteins that act as dominant negative regulators of other helix-loop-helix transcriptional factors, as discussed in the introduction. Id1 plays an important role in regulating embryonic development where it is expressed in complex spatiotemporal patterns; however, its expression is undetectable in most mature epithelial tissues. In contrast, Id1 is re-expressed in a number of advanced cancers, including in breast, endometrial, ovarian and prostate cancers (Perk et al., 2005). In breast cancer, Id1 is overexpressed in 54% of high grade basal-like tumours (Jang et al., 2006), and its expression correlates negatively with disease-free survival of patients within the ER-negative subtypes (Schoppmann et al., 2003). Unfortunately, the study of the role of Id1 in tumourigenesis has been hampered by its expression in multiple cell types. In the breast Id1 is expressed by endothelial cells, a rare subset of epithelial cells and the majority of immune cells. Thus if Id1 levels are analysed at the whole tissue level rather than the individual cell level it is uncertain which cell types are contributing. Furthermore the relatively high expression levels of Id1 in the endothelial and immune compartments can mask any potentially clinically relevant changes in the epithelial cells. A number of mechanisms have been proposed for the role of Id1 in promoting tumourigenesis. However, these have primarily been related to its roles in angiogenesis and immune cell function. More recently Id1 has been associated with stem cell and cancer stem cell maintenance (Barrett et al., 2011; Romero-Lanman et al., 2011; Williams et al.2011). It is therefore important to gain insight into how Id1 is promoting tumourigenesis in mammary epithelial cell.

To better understand the role of Id1 in epithelial cells during mammary carcinogenesis our lab established an inducible Tet-off Id1 overexpression system where doxycycline administration turns off Id1 overexpression by binding the tetracycline response element (TRE) and preventing transactivation of the Id1 overexpression cassette (Figure 3-1). The Id1 in this system is also HA tagged so it can be differentiated from endogenous Id1 expression. This

system allows us to see how Id1 cooperates with certain oncogenes at different stages of tumourigenesis. Initially, how Id1 cooperated with oncogenic Ras was examined. The Ras pathway is commonly up regulated in breast cancer, as suggested by a quarter of breast cancer cell lines having this pathway activated either by direct oncogenic Ras mutations or downstream targets such as BRAF (Hollestelle et al., 2007). Furthermore high Ras expression is associated with a subset of basal-like breast cancers (Herschkowitz et al., 2007). Constitutive overexpression of activated Ras by itself induces a proliferative burst in primary mouse mammary epithelial cells followed by the activation of senescence pathways in a dose dependant manner, with senescence pathways being avoided by low level overexpression (Sarkisian et al., 2007; Swarbrick et al., 2008). More recently, an association between functional mutations in the Ras 3' UTR and BRCA1 mutant breast cancers has been shown, these mutations block let7 binding and lead to an increase in Ras protein expression (Hollestelle et al., 2011). Using the Id1 inducible system mentioned above, it was recently shown that Id1 overexpression cooperated with oncogenic Ras to promote aggressive metastatic disease. Furthermore when Id1 overexpression was switched off in established mammary tumours the tumours regressed. This was not due to any increase in apoptosis but was in fact due to the induction of senescence and what appeared to be clearance by the immune system, similar to the findings of Xue et al and Ventura et al described above (Swarbrick et al., 2008; Ventura et al., 2007a; Xue et al., 2007). This suggested that part of the role of Id1 in epithelial cancers was to block OIS and subsequent immune recognition. The induction of senescence following Id1 switch-off is not completely surprising as Ras is known to activate senescence pathways and Id1 has previously been shown to negatively regulate replicative senescence in fibroblasts (Alani et al., 2001). What is less well characterised is the clearance of senescent cells by the immune system.

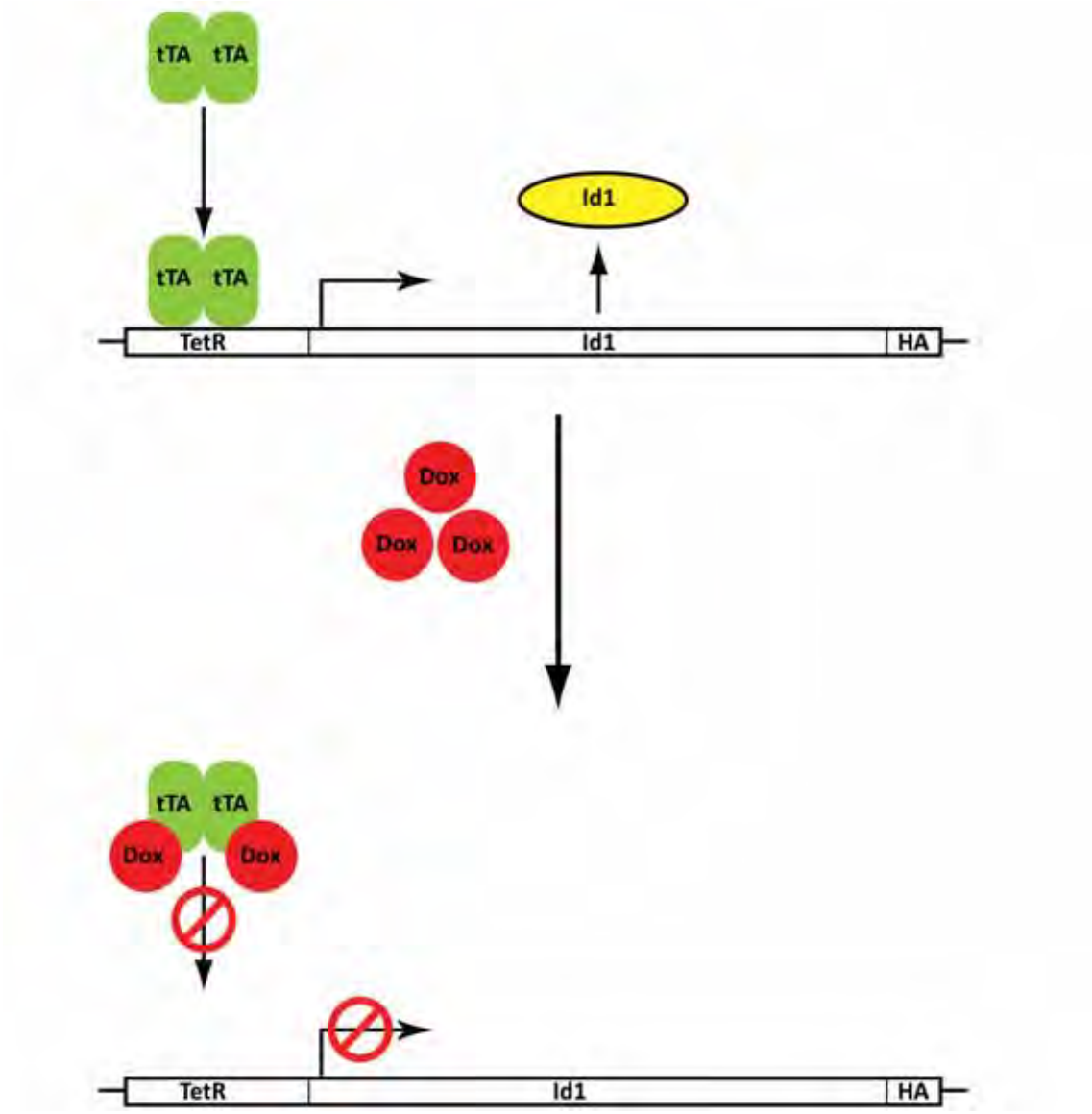


Figure 3-1. Model of doxycycline-inducible Id1 overexpression using the Tet-Off system.
The tetracycline transactivator (tTA) protein binds to the tetracycline response element (TetR) to induce haemagglutanin antigen (HA) tagged Id1 transgene expression in the absence of doxycycline (Dox) leading to Id1 protein expression. With the addition of Dox the tTA can no longer bind to the TetR and the Id1 transgene is not transcribed and no Id1 protein is expressed.

3.1.3 Senescence and immune recognition

Senescence was initially thought to be just a cell autonomous mechanism to block cell proliferation following DNA damage or cell stress, however, it is now becoming clear that senescent cells interact with the surrounding cells via the secretion of various factors and probably via changes to cell surface protein expression (Lee et al., 2011b). A small number of studies have found that when the tumour suppressor gene p53 is switched back on after tumour formation that the cells undergo senescence and the tumours regress. In a liver tumour model, Xue et al showed that upon p53 reactivation tumour cells were cleared by

innate immune cells, however, they were using athymic nude mice which lack functional T and B cells (Xue et al., 2007). Thus these mice do not reflect what would happen in the context of an immune competent host. Another study by Ventura et al 2007, also showed sarcoma regression upon reactivation of p53 and induction of senescence, however, they did not fully investigate the mechanism by which this occurred other than showing that apoptosis was not the cause (Ventura et al., 2007a). These studies suggest that initiating senescence in established tumours may be a potential treatment for cancer by initiating an anti-tumour response by the immune system. Immune evasion by the tumour is now understood to be a hallmark of cancer, and it is now becoming clear that the immune system is often corrupted by cancer cells to produce factors that promote their growth (Hanahan and Weinberg, 2011; Mantovani et al., 2008). It is exciting to think that by initiating senescence or a component of the senescence pathway one could convert a tumour promoting immune response into an anti-tumour immune response. However the pathways by which senescent cells might initiate clearance by the immune system are not well understood. Senescent cells are known to secrete cytokines, however, these have been generally thought to induce a pro-tumour inflammatory response (Rodier and Campisi, 2011). The study by Xue et al showed that innate immune cells, in the absence of the adaptive immune system, were able to clear senescent tumour cells. However the mechanism by which the senescent cells trigger an immune response is not known. Possibilities include an up regulation of activating NK cell ligands, an up regulation of cytokines that promote an anti-tumour immune response, or the down regulation of cytokines that promote tumourigenesis. We seek to utilise the inducible Id1 model of Ras driven mammary carcinogenesis to further understand the role of Id1 in blocking senescence and to understand the changes within the senescent cells that make them more immunogenic.

3.1.4 Hypothesis:

Id1 blocks oncogene induced senescence and the associated clearance by the immune system.

3.1.5 Aims:

The specific aims of this chapter are to:

1. Determine whether transient inhibition of Id1 expression leads to irreversible senescence.
2. Determine what role immune cells play in tumour regression following senescence.
3. Identify senescence-induced tumour cell factors that promote immune mediated tumour regression.

3.2 Results:

3.2.1 Analysing immune infiltrates into previously generated Id1 Ras mammary tumours by IHC

Based on pathological examination of H&E stained sections, Dr. Swarbrick and colleagues reported an increase in immune cell infiltrates into the mammary tumours expressing inducible Id1 with oncogenic Ras following Id1 switch off (Swarbrick et al., 2008). To better understand the immune cell infiltrates into these tumours, paraffin sections were obtained from tumours generated during the Swarbrick et al 2008 study. IHC was performed for macrophages (F4/80) (FigureA), and T cells (CD3) (FigureB) after performing optimisation experiments to determine the optimal antibody concentration and antigen retrieval strategy. Considerable numbers of both macrophages and T cells were observed in these tumours, showing that immune cells make up a significant component of mammary tumours driven by oncogenic Ras and Id1 overexpression. However, their relative proportion did not change dramatically following 72 hours of doxycycline treatment. Unfortunately there was no tissue available to analyse later time points following the switch off of Id1. Furthermore, histological analysis is limited in the level of detail about the immune cell profile. To better understand the immune response to Id1 switch-off, we went onto use flow cytometry on freshly prepared tumour samples, as it allows for the analysis of many more immune cell subsets than is possible by IHC (see section 3.2.8).

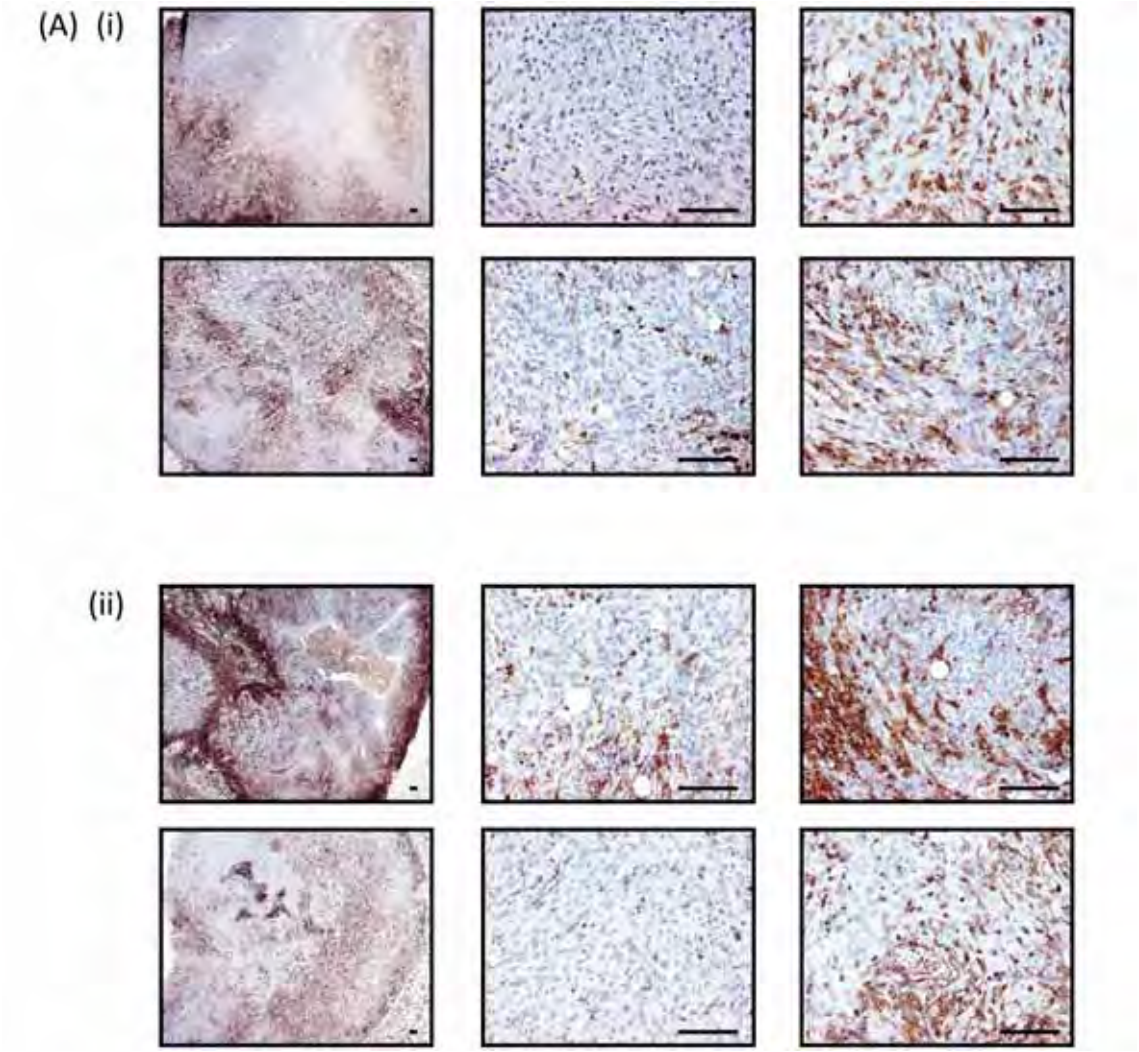


Figure 3-2. Immune cell infiltrate into tumours overexpressing Ras before and after Id1 switch-off.
(A) IHC staining for F4/80 positive macrophages (brown staining) before (i) and after (ii) 72 hours of doxycycline treatment. Two representative tumours are shown in each case with two representative higher power image of each tumour. Scale bar=80μm. (Figure continues on next page)

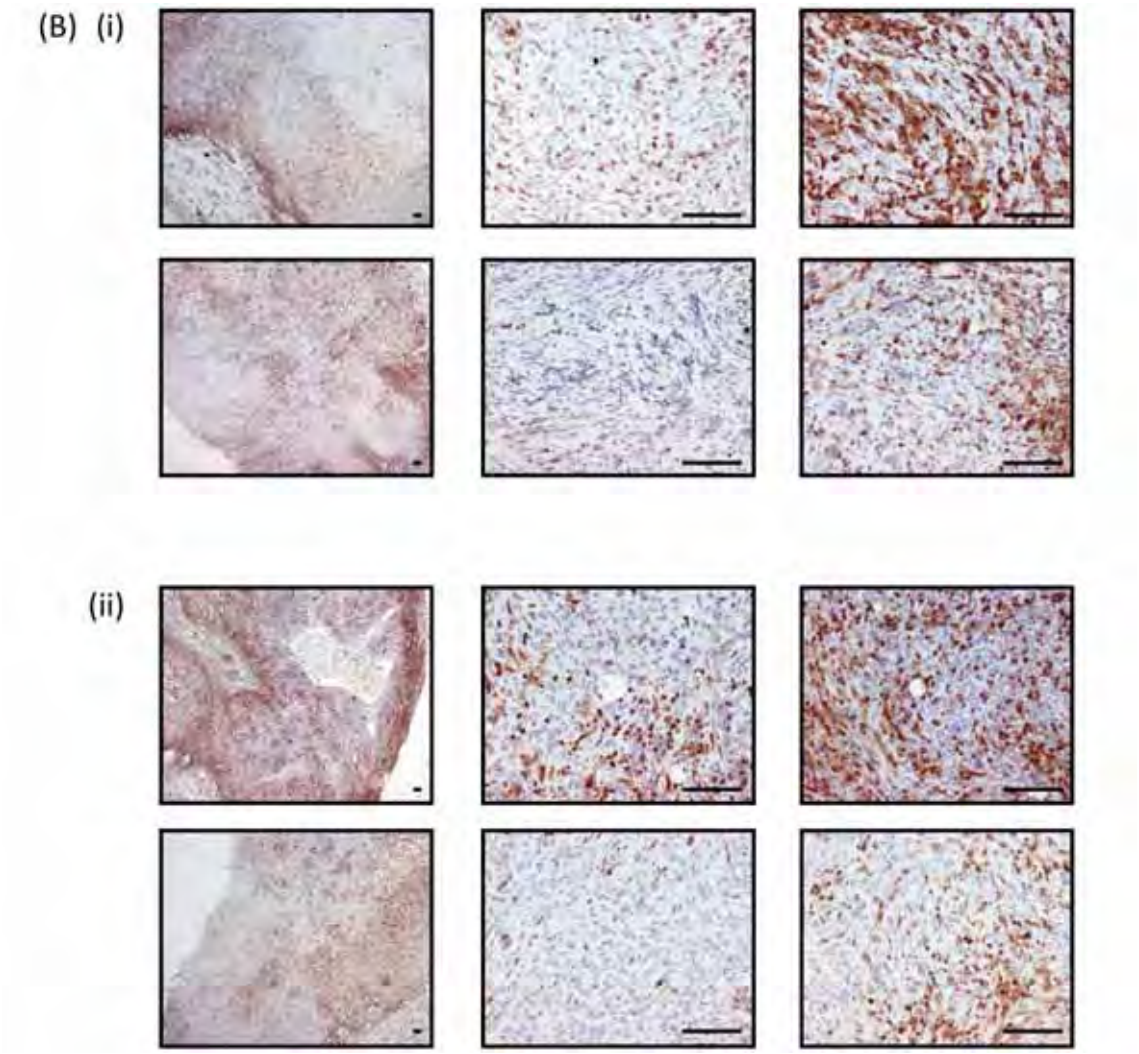


Figure 3-2. Continued. Immune cell infiltrate into tumours overexpressing Ras before and after Id1 switch-off.

(B) IHC staining for CD3 positive T-cells (brown staining) before (i) and after (ii) 72 hours of doxycycline treatment. . Two representative tumours are shown in each case with two representative higher power image of each tumour. Scale bar=80μm.

3.2.2 Determining the best vector combination for inducible Id1 overexpression

Following on from the original study that used pMIG-tTA to switch off Id1 in doxycycline dependent manner we had obtained a second tTA vector pMIL-tTA. The pMIL-tTA construct includes a luciferase reporter that would allow for improved *in vivo* imaging of the tumours generated when compared to the GFP in the pMIG vector. To determine which combination of vectors generated the best inducible overexpression of Id1 we compared the ability of pMIL-tTA with pMIG-tTA to induce Id1 expression. As a control the non inducible construct pMIG-Id1 was also analysed. 293T cells were transiently transfected with the above constructs and the doxycycline response was tested in these cells *in vitro*. Following administration of doxycycline (1μg/ml) for 48 hours, Id1 levels were significantly reduced as measured by anti-

HA western blot (Figure 3-3A). The combination of pMIG-tTA and pRQ-Id1 produced 2.7x more Id1 overexpression than the pMIG-Id1 alone and upon doxycycline treatment this overexpression was reduced by more than 90% (Figure 3-3B). The pMIG-tTA vector provided more robust overexpression of Id1 and greater levels of switch off upon doxycycline administration so the combination of pRQ-Id1 with pMIG-tTA was chosen for further studies.

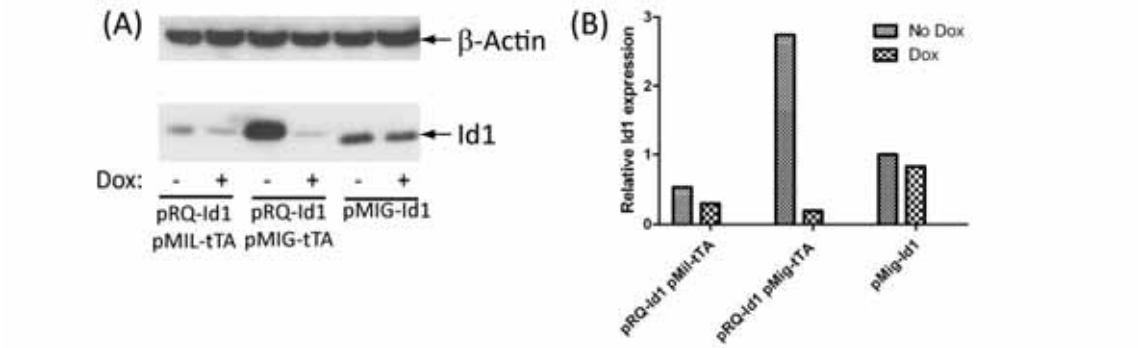


Figure 3-3. Analysis of inducible Id1 systems *in vitro* by anti-Id1 western blot.
293T cells were infected with pRQ-Id1 along with pMIL-tTA or pMIG-tTA or with pMIG-Id1 as indicated, then cultured in the presence or absence of doxycycline for 48 hours. (A) Western blot analysis for HA tagged Id1 with β-actin loading control. (B) Densitometry analysis of Id1 regulation by doxycycline.

3.2.3 Generating mammary tumours overexpressing inducible Id1 and Ras

To further analyse the role of Id1 in tumourigenesis, senescence and the immune response it was first necessary to generate a number of tumours overexpressing oncogenic Ras and inducible Id1. These tumours could then be propagated by transplanting small pieces of the tumours into larger cohorts of mice to allow for the thorough analysis of the interaction between senescence induction and the immune response.

The generation of mouse mammary tumours overexpressing oncogenic Ras and inducible Id1 required the retroviral transduction of primary mouse mammary epithelial cells (Figure 3-4). Retroviruses that contained oncogenic Ras, inducible Id1, and the doxycycline response element tTA were freshly prepared in Phoenix-EcoR cells. For optimal infection of primary cells frozen virus cannot be used. Primary mammary epithelial cells were prepared from wildtype FVB/N mice by collagenase digestion and differential centrifugation to remove stromal cells. Differential centrifugation involves short low speed centrifugation steps that pellet the mammary epithelial organoids but not the stromal fibroblasts and immune cells. The epithelial cell enriched preparations were then plated into 6-well plates. The following day, the cells were retrovirally transduced using a spinfection protocol with viruses containing either Ras (pMig-RasV12), Id1 (pRQ-Id1) and tTA or as a control Ras, an H2B control (pRQ-H2B-RFP) and tTA. H2B is a core histone protein that is normally ubiquitously expressed so

overexpression of this control protein should have minimal effect on primary mammary epithelial cells and will provide an appropriate control for infection with the three different retroviruses. All three retroviruses are required to infect one cell to cause oncogenic transformation. The infection process was repeated for a second day, the next day the cells were observed for GFP expression by microscopy to determine success of viral transduction. Viral transduction efficiency could only be determined for the Ras construct as it was GFP labelled and also led to morphological changes in the infected cells. The tTA construct was also GFP labelled however it alone would not lead to any morphological changes in the invected cells.

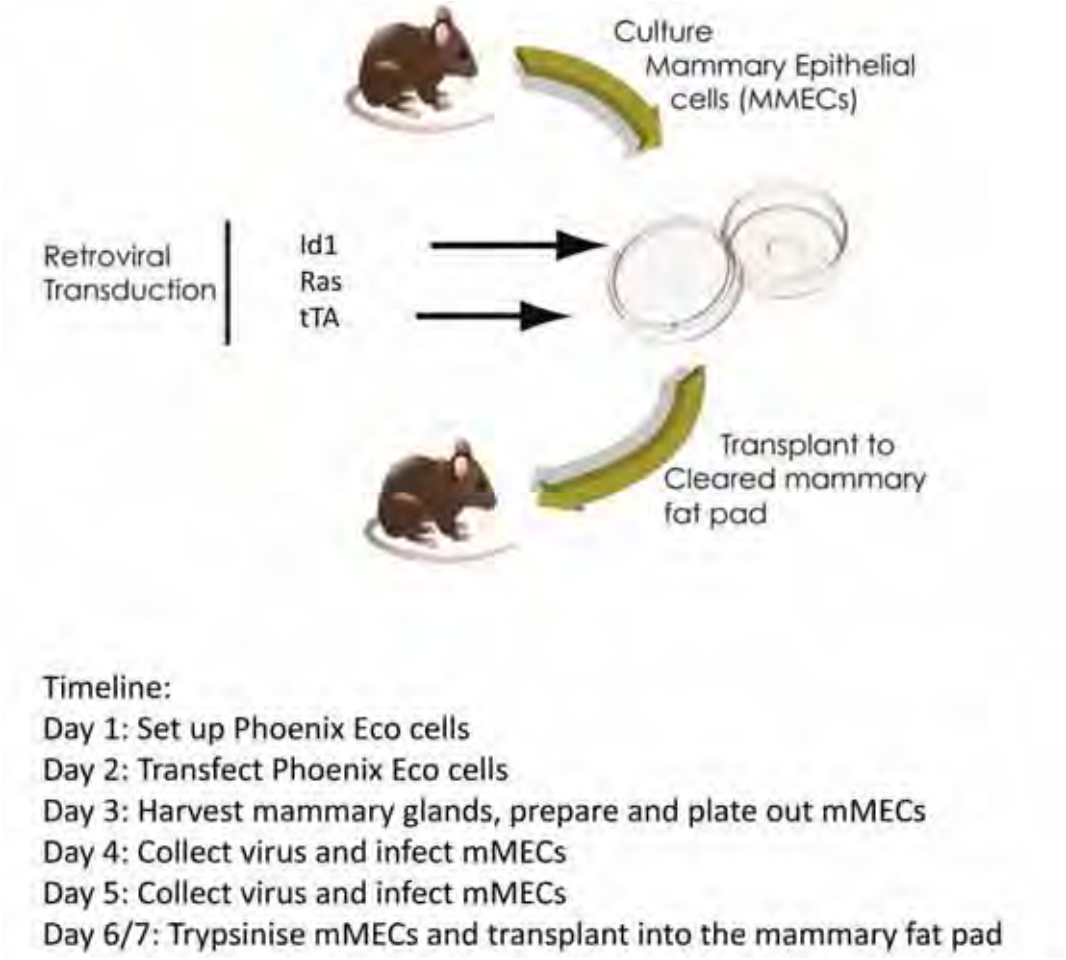


Figure 3-4. Protocol outline and timeline for the retroviral transduction of primary mammary epithelial cells (mMECs).

250 000 of the virally transduced cells were then transplanted per mouse into the 4th mammary gland of recipient FVB/N mice and monitored for tumour development. Mammary glands were not cleared of endogenous epithelium in these studies to minimise the impact of surgery on the immune response. As transformation of mammary epithelial cells requires all

three constructs there is a strong selective pressure for the outgrowth of cells that received all three constructs. 11 mice received the Id1 and Ras transduced cells and 4 mice the H2B and Ras control cells. Others in our group had previously determined that overexpressing Ras alone would not generate any tumours when 250 000 cells were transplanted. The growth of tumours that formed was monitored using callipers. Six tumours formed in the Id1 and Ras mice but none formed in the H2B and Ras control mice (Figure 3-5). Tumours were palpable from around 10 days after transplant and reached ethical endpoint (when the average diameter reached 10mm) at around 25 days following transplant. Tumours were collected, cut into small pieces and viably frozen to be used in subsequent experiments.

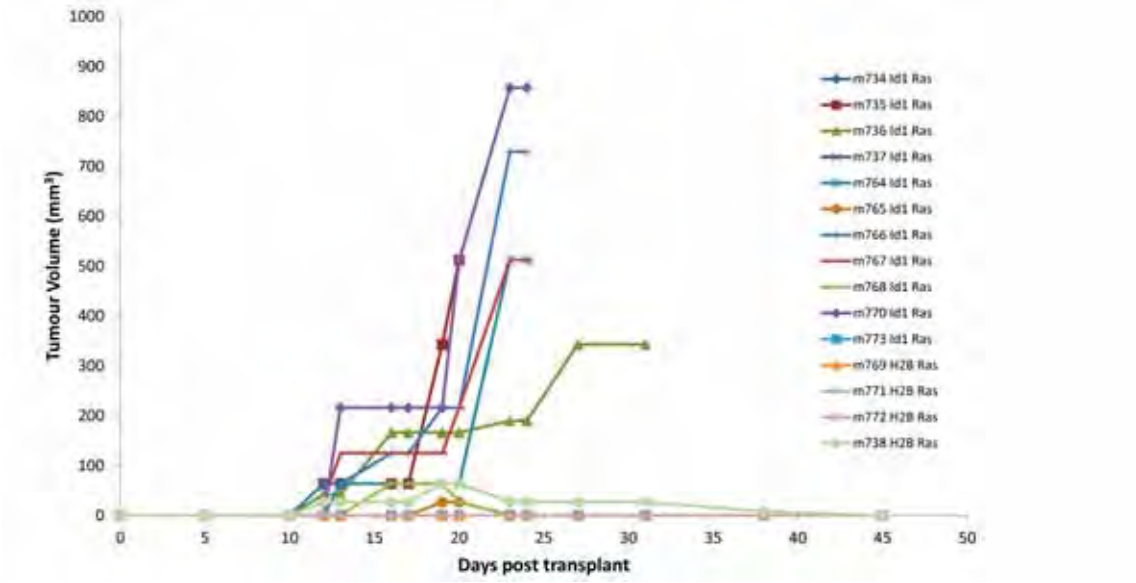


Figure 3-5. Tumour growth following the transplant of primary mammary epithelial cells transduced with oncogenic Ras and inducible Id1.

Tumour growth (mm^3) is given for individual mice transplanted with primary mMECS transduced either with Id1 and Ras, or H2B and Ras, as indicated in the legend. Mice were euthanased when tumours grew beyond 10mm in one dimension or at 45 days pot transplant.

3.2.4 Switching Id1 off after mammary tumours are established

10 mice were transplanted with cryopreserved pieces from one of two tumours overexpressing Id1+Ras from the previous experiment (n=6 and n=4). When the tumours had reached 6-8mm diameter, half of the mice were put on a doxycycline diet to switch-off the Id1 expression. To confirm that tumour regression was dependent on the immune system we also transplanted tumour pieces into 10 highly immuno-deficient NOD/SCID/IL2R$\gamma^{-/-}$ mice (designated NOG mice). The NOG mice lack all cells of the adaptive immune system (T and B cells), as well as having dramatically reduced macrophage and NK cell activity and thus should be unable to clear the tumour cells following induction of senescence. Tumour growth in the

FVB/N mice was monitored until tumours had reached ethical endpoint using both callipers and *in vivo* imaging using the IVIS Illumina II system. The *in vivo* imager could detect GFP fluorescence from the transduced mammary cells, however due to the low penetration of the GFP signal the area over the tumour had to be shaved to allow for imaging and as the tumour grew the level of fluorescence measured by the imager did not correlate with the increased size of the tumour. Surprisingly, although the growth of tumours on the doxycycline diet was slightly but insignificantly reduced, they did not regress over the course of the 22 day experiment (Figure 3-6A), whereas experiments reported in Swarbrick et al showed immediate growth arrest followed by regression (Swarbrick et al., 2008). The relative levels of senescent cells between the doxycycline treated and untreated tumours was analysed by senescence-associated β-galactosidase staining. Small numbers of senescent tumour cells could be observed in both groups of mice with no enrichment seen in the doxycycline treated mice (Figure 3-6B). Thus the slight reduction in tumour growth in the doxycycline treated mice does not appear to be due to increased oncogene induced senescence. To determine why the tumours failed to regress, the level of Id1 overexpression was analysed and it was noted that the level of Id1 overexpression was relatively low in these tumours (Figure 3-6C). To overcome this, we aimed to generate new Id1+Ras tumours in which the level of Id1 overexpression was screened prior to subsequent experiments (Section 3.2.5). Interestingly, although NOG mice were monitored for tumour growth for 7.5 weeks, no tumours developed. Further examination of this phenomenon is examined in section 3.2.6.

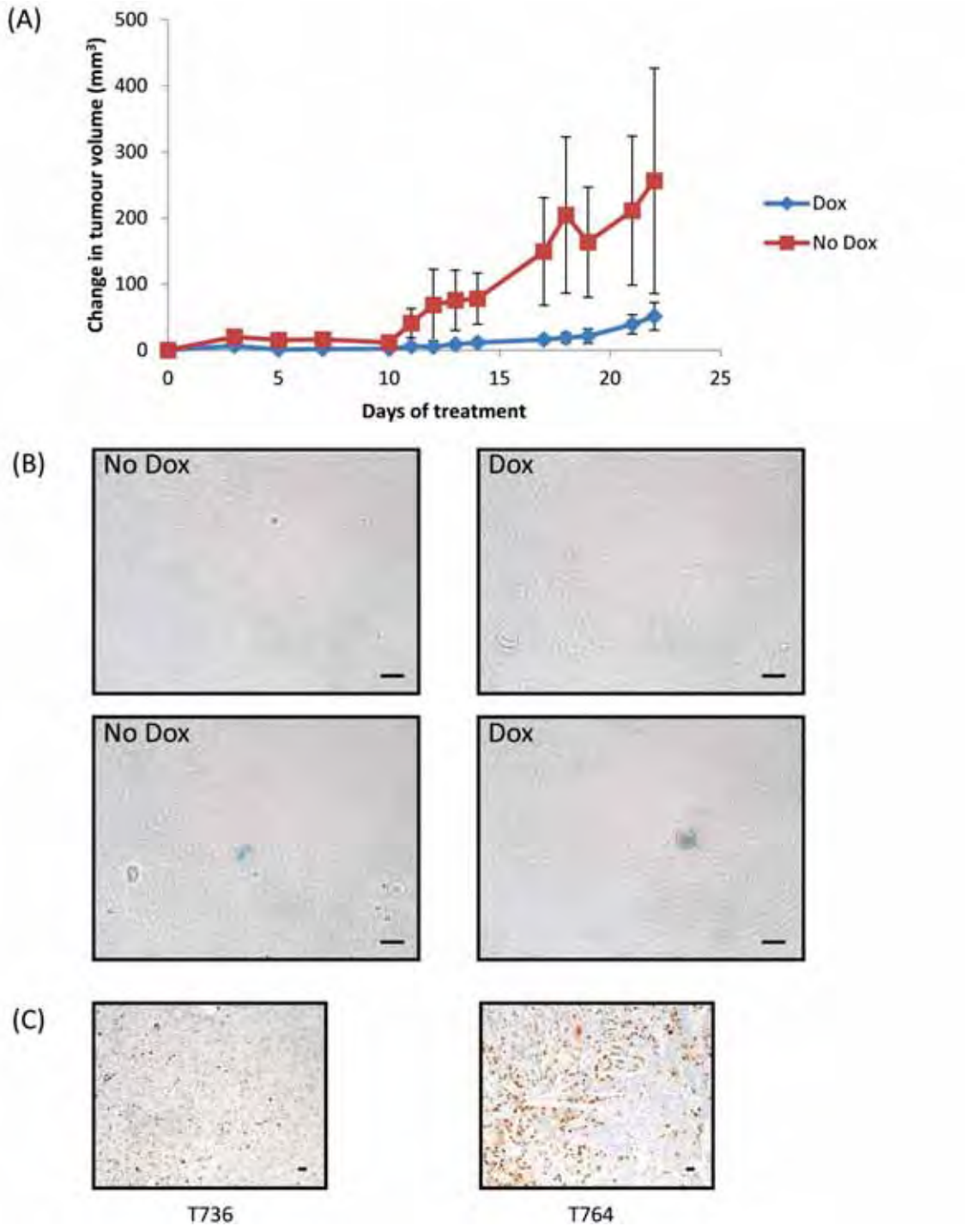


Figure 3-6. Tumours fail to regress following Id1 switch-off.
(A) Average tumour growth (±SEM) following Id1 switch-off. (B) No change in number of senescent cells upon Id1 switch-off as measured by the number of senescence-associated β-galactosidase (SA-βGal) positive cells (blue staining). (C) Analysis of Id1 expression (brown staining) in the tumours used for the analysis of Id1 switch-off. Scale bars=25μm.

3.2.5 Repeating Id1 switch off with tumours with very high Id1 overexpression

To test whether the failure to regress in the established tumours following Id1 switch off was due to relatively low initial overexpression of Id1, fresh mammary tumours were generated by

retroviral transduction with Id1 and Ras. These tumours were then screened for Id1 expression levels and a tumour with high Id1 expression was selected for further analysis (Figure 3-7A). 20 FVB/N mice were transplanted with tumour pieces from this tumour. When the tumours reached 6-8mm in diameter eight of the mice were put on a doxycycline diet. Again no regression was observed with only a partial reduction in tumour growth (Figure 3-7B). Interestingly, despite the tumours from both groups being of similar diameter at endpoint, the tumours from the doxycycline treated mice contained a greater amount of fluid. When the dry weight of the tumours was analysed the doxycycline treated tumours were significantly lighter than the untreated controls 1.4±0.41g verses 2.5±0.26g ($p=0.0136$ Mann-Whitney t-test) (Figure 3-7C). This suggested that while Id1 switch-off was reducing tumour growth it was not leading to regression. To determine if the failure of these tumours to regress was due to failure to switch off Id1 overexpression, the tumours were analysed for HA tagged Id1 expression levels by western blot. This showed that Id1 switch-off was not complete upon doxycycline treatment (Figure 3-7D). This suggested that incomplete shut off of the Id1 transgene was leading to the attenuated response in these tumours.

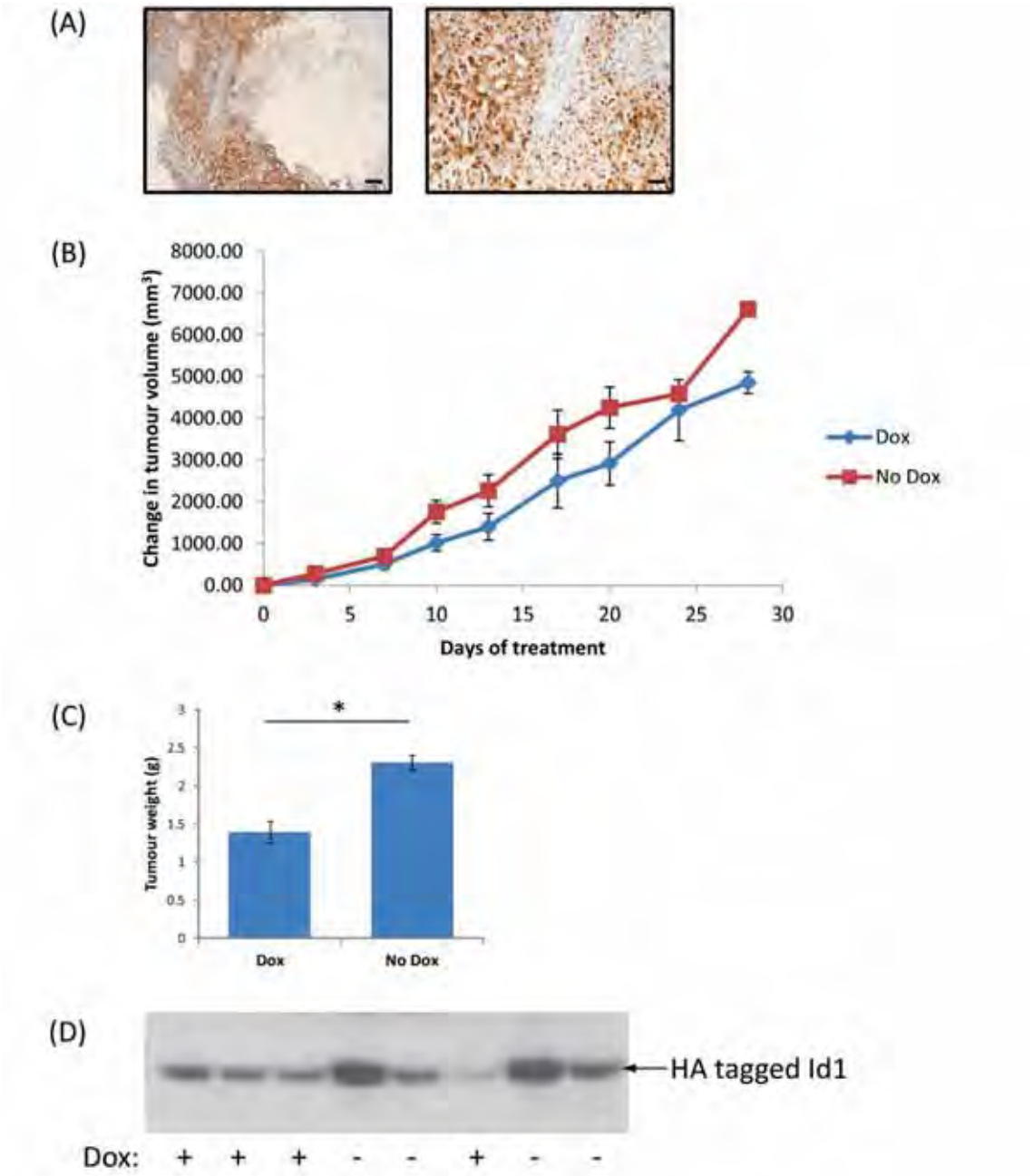


Figure 3-7. Tumours fail to regress following Id1 switch-off in tumours highly overexpressing Id1.
(A) High Id1 expression level (brown staining) in tumour selected for switch-off experiment. Scale bars=200μm. (B) Tumour growth upon Id1 switch off. Change in tumour volume following doxycycline treatment is plotted as an average per group. (C) Tumours were weighed upon dissection when they reached the ethical endpoint in size. The average weight per group is shown (B and C) Error bars indicate SEM* p=0.136 Mann-Whitney t-test. (D) Analysis of Id1 transgene expression in the tumours as measured by western blot for the HA-tagged Id1 with or without doxycycline treatment.

3.2.6 Comparing Id1/Ras mammary tumour growth between immune compromised mice and controls

To rule out that failure of tumours to grow in NOG mice was due to experimental error, we repeated the experiment described in 3.2.4 using tumour pieces from a tumour with higher Id1 expression (Figure 3-7A) and this time the mice were also monitored for GFP signal using *in vivo* imaging. Tumour pieces were transplanted into 6 NOG mice and 15 FVB/N mice. GFP positive cells were detectable in all of the NOG mice by *in vivo* imager over the full time-course of the experiment (FigureA) but tumour growth was only observed in two of the mice, whereas tumours grew in all FVB/N mice. Furthermore the tumours that did grow in NOG mice did so at a significantly reduced rate as compared to those tumours growing in the FVB/N controls (FigureB). In contrast all of the tumours grew in the FVB/N control mice. To determine if the poor tumour growth was specific to the NOG strain of mice we compared the rate of tumour growth in the NOG mice to that in recombination activating gene 1 (Rag-1) deficient mice. Rag-1 encodes an enzyme critical for the rearrangement and recombination of the genes of immunoglobulin and T cell receptor molecules. Mice deficient in Rag-1 therefore lack mature T and B lymphocytes but retain more normal macrophage and NK cell activity. Tumour growth was similarly reduced both in NOG and Rag mice when compared to FVB/N controls (FigureC). While these results are interesting and show that the immune system is necessary to promote tumourigenesis in this model they also mean that it is not possible to analyse the role of the immune system in tumour regression using these immune compromised mouse strains.

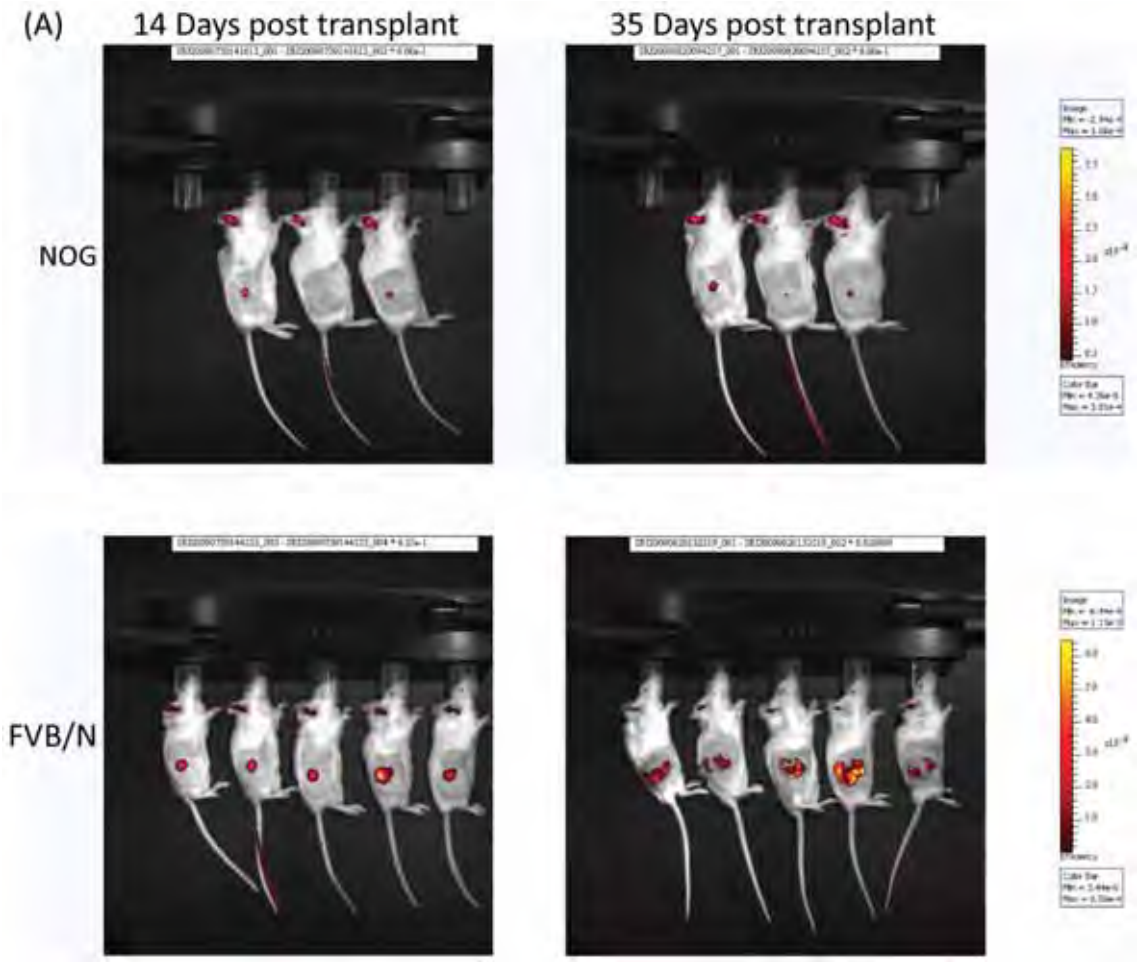


Figure 3-8. Mammary tumours driven by oncogenic Ras and Id1 overexpression do not grow in immuno-deficient mice.

(A) Representative images illustrating the detection of GFP positive transplants in FVB/N and NOG mice at days 15 and 35 following transplant using the *in vivo* imager. (Figure continues on next page)

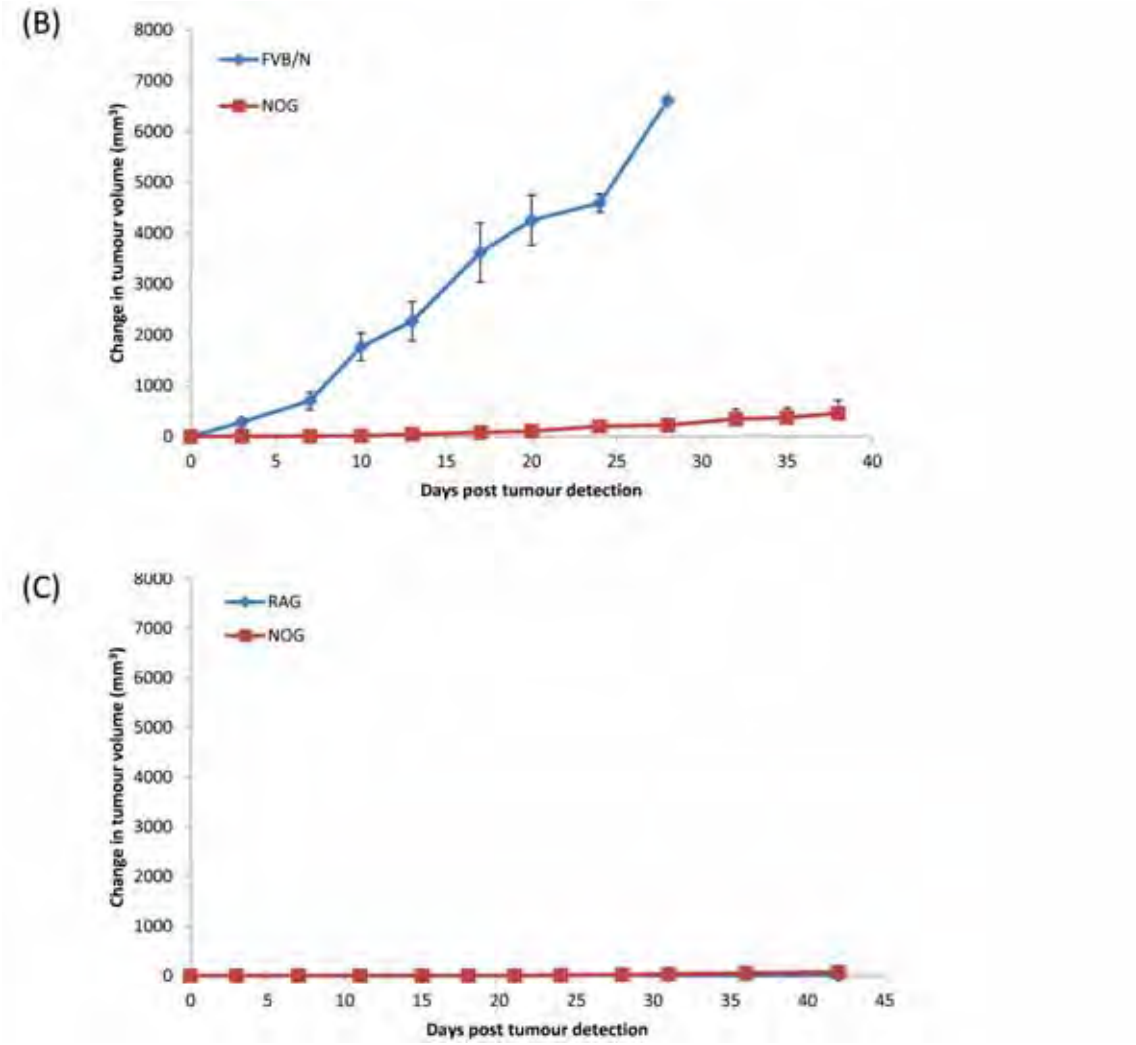


Figure 3–8. Continued. Mammary tumours driven by oncogenic Ras and Id1 overexpression do not grow in immuno-deficient mice.

(B) Average tumour growth (mm^3) in NOG mice compared with FVB/N control mice (±SEM). (C) Average tumour growth (mm^3) in NOG compared to Rag mice (±SEM).

3.2.7 Repeat of switching Id1 off after mammary tumours are established with additional doxycycline

In attempt to overcome the incomplete deactivation of Id1 expression in response to the doxycycline in the previous Id1 switch-off experiment (section 3.2.5), the experiment was repeated wherein mice were treated with doxycycline diet as well as an additional IP injection of doxycycline (500μg) once the tumours had reached 6-8mm. Tumours were harvested 7 days following doxycycline treatment, however, again the tumours did not regress (Figure 3-9). In the original study they had transplanted the tumourigenic cells into 3 week old mice and had cleared the endogenous epithelium. I tested whether this could be causing the difference between my results and the previously published results however this also did not result in tumour regression upon doxycycline treatment (d.n.s).

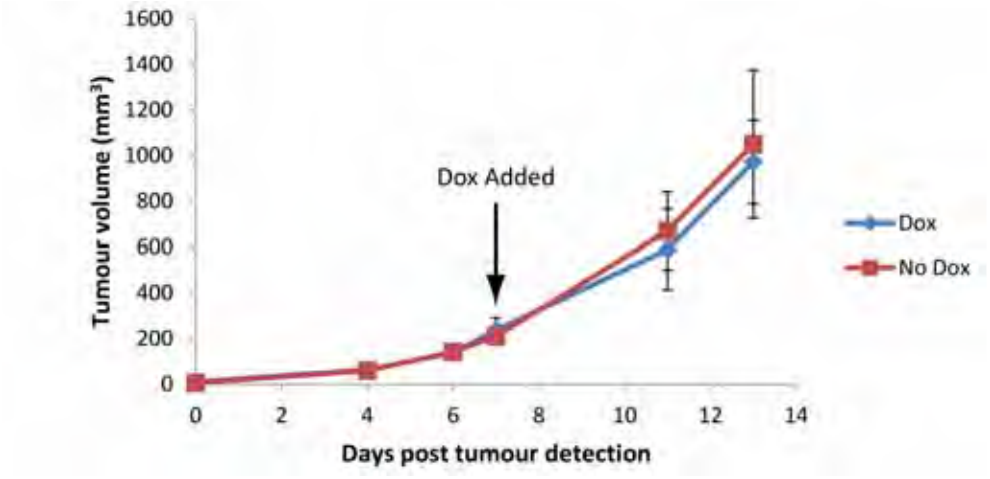


Figure 3-9. Additional injection of doxycycline fails to induce tumour regression.
Plot of average tumour growth of mice treated with or without doxycycline (Dox) diet from day 7, Dox group mice were also injected with 500 µg Dox i.p. on day 7 (±SEM).

3.2.8 Immune infiltrate into Id1 Ras mammary tumours

While establishing the tumour regression model, it was also necessary to establish the flow cytometric analysis of immune cell infiltrates into the mammary tumours for use in later analyses. A panel of immune cell markers was determined to characterise the immune cells infiltrating the tumours (see Table 3-1). This panel was determined by consulting the literature, personal experience, and discussions with immunology experts at the Garvan Institute, Australia, and the University of California San Francisco, USA (DeNardo et al., 2009; Junankar et al., 2006). To determine the cell types present in the mammary tumours generated by Id1 and Ras overexpression two tumours (from 3.2.3) were PBS perfused to remove the blood leukocytes from the tumour-associated cells, the tumours were then excised, collagenase digested and single cells prepared and then analysed by flow cytometry. Roughly a third of the tumour cells were found to be $CD45^+$, and the majority of these were $CD11b^+$ myeloid cells (~80%), the next largest group were $CD3^+$ T cells (~15%), and the remainder were $CD49b^+$ NK cells (~5%) (Figure 3-10A). The $CD11b^+$ cells were further subdivided based on surface marker expression and found to be mostly either $F4/80^+$ macrophages or $Gr1^+$ neutrophils (Figure 3-10A). The $CD3^+$ cells were approximately 60% $CD4^+$ helper T cells and 20% cytotoxic CD8+ T cells. To determine what effect switching off Id1 had on the immune cell infiltrate, three doxycycline treated and three untreated tumours (from 3.2.4) were analysed by flow cytometry. The mice had either been fed doxycycline food once the tumours reached 6-8mm in size or were untreated. The mice were PBS perfused and single cells prepared from the tumours once they had reached >10mm diameter. Cells were stained for cell surface markers and analysed by flow cytometry. Despite the lack of tumour regression

and in line with the reduced tumour mass after Id1 switch-off (Figure 3-7C), upon doxycycline treatment NK cell percentage was significantly increased from 6.0 % to 9.4 % of the CD3 negative population of cells in the doxycycline treated tumours, suggesting an increase in immune surveillance by NK cells (Figure 3-10B). No other cell types were significantly altered.

Cell Type	Markers
All leukocytes	CD45
Granulocytes/Neutrophils	Gr1 or 7/4
Macrophages	CD11b and F4/80
NK cells	NK 1.1 or CD49b positive and CD3negative
Immature myeloid cells	CD11b and Gr1
Dendritic Cells	CD11c
T cells	CD3
T helper cells	CD3 and CD4
Cytotoxic T cells	CD3 and CD8
NK T Cells	CD3 and NK1.1 or CD49f
B cells	CD19 and B220

Table 3-1. Panel of antibodies used to determine the immune cell infiltrate into the mammary tumours.

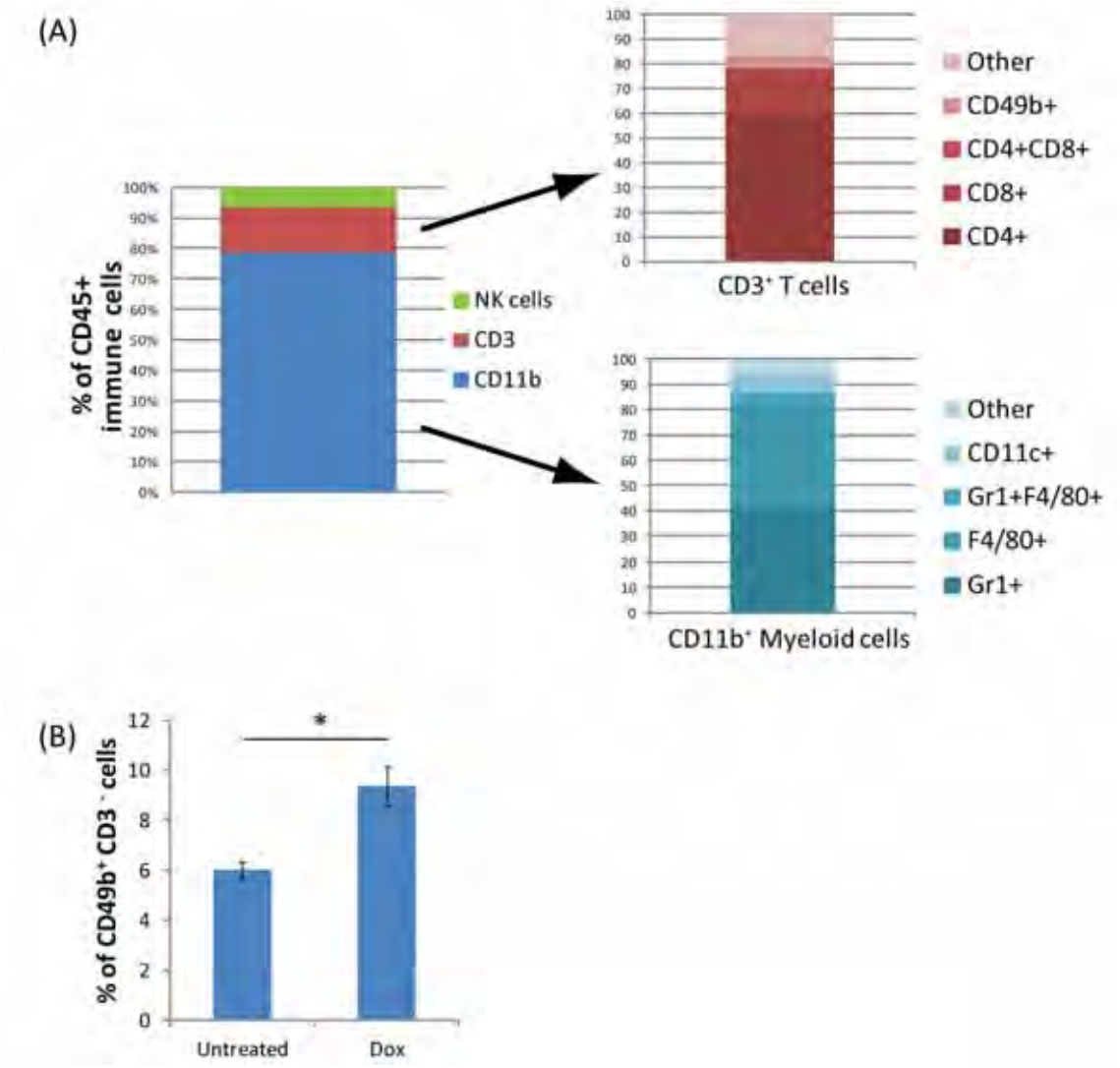


Figure 3-10. Immune cell infiltrate into mammary tumours driven by oncogenic Ras and overexpression of Id1.

(A) Immune cell composition of PBS perfused Id1+Ras tumours was determined by flow cytometry as outlined in Table 3-1. The average percentage of each cell population from 2 mice is depicted (B) Mice bearing Id1+Ras tumours of 6-8 mm were divided into two groups and either left untreated or given Doxycycline diet (Dox). When tumours were greater than 10mm diameter, 3 tumours per group were perfused, harvested and immune cell composition determined as outlined in Table 3-1. The average percentage of $CD3^{-}CD49b^{+}$ NK cells in tumours increases upon Id1 switch-off (±SEM), n=3 per group *p<0.05 Mann-Whitney t test.

Immune infiltrates were also examined by IHC, however the panel of markers available is restricted and thus fewer immune cell sub-types could be analysed. IHC was performed for total leukocytes (CD45), macrophages (F4/80), neutrophils (7/4), and T cells (CD3). No difference was observed in immune cell infiltrate between doxycycline treated and untreated tumours by IHC (d.n.s). Unfortunately there is not currently a good antibody to study natural killer cells (NK cells) by IHC so the flow cytometry results could not be confirmed by IHC.

The immune cell infiltrates into the tumours treated with additional doxycycline i.p. (Section 3.2.7) were also analysed by flow cytometry. Further, to determine if there were changes in the global immune cell composition, the spleens were also analysed. The markers used in Table 3-1 were used, along with additional markers CD62L and CD44, which enable determination of the activation status of the $CD4^+$ T cells. $CD4^+$ T cells that express CD62L but not CD44 are considered naive, double positive cells are considered central memory T cells and CD62L negative CD44 positive cells are considered effector T cells. However, in this experiment no significant differences were observed with any of the immune cell markers upon doxycycline treatment (d.n.s.).

3.3 Discussion:

The aim of this work was to examine the effect of Id1 down regulation and senescence on the recognition of tumour cells by the immune system. Exciting data from Swarbrick et al and others had suggested that the induction of senescence could switch a pro-tumour immune response into an anti-tumour immune response (Swarbrick et al., 2008; Ventura et al., 2007a; Xue et al., 2007). In these studies, I discovered that there is a large immune cell composition in mammary tumours driven by oncogenic Ras and Id1 overexpression. Interestingly it appears that the growth of mammary tumours driven by oncogenic Ras and Id1 overexpression requires an intact immune system in the host animal, as tumour growth was significantly reduced in both the NOG and RAG strains of immuno-deficient mice. Of particular interest was that the tumour cells in the immuno-deficient mice remained viable and continued to express GFP as measured by the *in vivo* imager but failed to grow. Oncogenic Ras is known to induce inflammation which, in the absence of a senescence response, can promote tumourigenesis (Borrello et al., 2008). Interestingly it is generally thought that it is the innate immune system that promotes tumourigenesis, however, since tumours did not grow in RAG-$1^{-/-}$ mice with intact innate immunity, in this case it appears that the adaptive immune system is required. There have been a small number of studies that have shown that the adaptive immune system can promote tumourigenesis. Firstly it has been demonstrated that the adaptive immune system is required for the full metastatic potential of the transgenic mouse mammary tumour model MMTV-PyMT, although there was no affect on the primary tumour growth when it was crossed onto the Rag background (DeNardo et al., 2009). In a transgenic model of skin cancer, tumour incidence was reduced in the absence of the adaptive immune system and that this was due to reduced inflammation. In this case the inflammation was being driven by activated B cells producing antibodies to extracellular components in the stroma surrounding the hyperplastic skin (de Visser et al., 2005).

Despite not seeing complete tumour regression in this model I did see a reduction in tumour growth and, in one experiment, an increase in tumour infiltrating NK cells. Mechanisms by which the tumour cells could become more immunogenic are through the modulation of cell surface receptors and secreted cytokines. For example, the down regulation of immunogenic cell surface markers such as MHC class I molecules and changes in the ratio of NK cell activating and inhibiting receptor ligands can increase NK cell recognition and killing of cancer cells (Langers et al., 2012). Furthermore, changes in the profile of secreted cytokines and chemokines can affect the strength of the immune response. Subsequent to my studies, a recent publication demonstrated in a premalignant liver cancer model that N-Ras driven oncogene induced senescence elicits a $CD4^+$ T cell response that directs clearance of the senescent cells by macrophages (Kang et al., 2011). This immune clearance was dependent on MHC class II up regulation by the senescent cells and was associated with increased levels of a number of inflammatory chemokines. This study also showed that in mice that lacked T cells, there was a failure by the immune system to recognise oncogene induced senescence and an increased incidence of cancer. This is in sharp contrast to my results which demonstrated a reduced incidence of cancer in the absence of the adaptive immune system. Considering the aforementioned potential importance of B cells in tumour growth, the lack of both B and T cells in our $RAG1^{-/-}$ studies may contribute to the discrepancy.

Id1 has been shown to be directly regulated by oncogene Myc (Swarbrick et al., 2004). Interestingly, Myc has been demonstrated to both suppress senescence pathways and also to down regulate components of the cellular stress response that promote immune recognition (Larsson, 2011; Schlee et al., 2007). In particular, in Burkitt's lymphoma Myc was shown to impair the NF-κB and interferon response by blocking Stat1 transcription and activation (Schlee et al., 2007). Both Myc and Id1 are known to positively regulate Bmi1, a member of the polycomb repressor complex that plays a role in the stem cell phenotype and is also known to suppress senescence induction (Larsson, 2011; Qian et al., 2010). Together, these data suggests that the senescence blocking and possibly the immune regulatory roles of Myc may be mediated through Id1 (Swarbrick et al., 2004). It would be interesting to see if the immune suppressive functions of Myc are also mediated via Id1. Recent results from another student in our lab support this proposition, showing that Id1 regulates the NF-κB pathway in the MDA-MB-231 breast cancer cell line (W. Teo, personal communication), further suggesting that Id1 may play a role in suppressing the immune response to cancer cells.

Unfortunately the tumour regression upon Id1 switch off that formed the basis for this project could not be replicated and thus the experiments planned to follow on from these findings could not be pursued. As well as the failure of these tumours to regress it was also of note that none of the mice that grew the Id1+Ras tumours in any of the experiments had any macro-metastasis in their lungs. Metastasis to the lung was a common feature of this model in the previous studies that had been performed in the USA. After trouble shooting Id1 expression levels and the doxycycline administration procedure there was no improvement in tumour regression. With no obvious way of rectifying this problem this study was abandoned. Setting up new systems such as those utilising p53 reactivation as described by other groups would be another way to investigate the role of senescence and the immune response, however this would not allow for the investigation of the role of Id1 in this process.

Chapter 4. The role of Id4 in mammary gland development

4.1 Introduction:

Id4 is known to play a significant role in the developmental processes of a number of organs including the brain, the testes, and the ovaries, and also plays an important role during embryogenesis (Bedford et al., 2005; Jen et al., 1996; Johnson et al., 2008; Sablitzky et al., 1998; Yun et al., 2004). Id4 regulates development in these tissues through the modulation of differentiation pathways, and thus in its absence abnormal cell differentiation and organ dysfunction occur. This suggests that the most common mechanism of action for Id4 is to regulate differentiation pathways in stem and progenitor cells, and that it is required for the correct formation and function of these organs. It is known that Id4 is expressed in the breast and is associated with breast cancer. In particular studies have shown that aberrant Id4 promoter methylation with correlates with poor outcome (de Candia et al., 2006; Noetzel et al., 2008; Umetani et al., 2005). However the role of Id4 in mammary gland development has not been examined.

4.1.1 Id4 and mammary gland biology

Id4 expression is up regulated by progesterone in the mouse mammary gland (Soyal et al., 2002). Furthermore Soyal et al showed evidence that suggested that Id4 was a direct target of the progesterone receptor. However, Id4 mRNA expression negatively correlates with oestrogen receptor (ER) expression in normal human breast epithelium suggesting that Id4 is not expressed in the ER and PR positive cells (de Candia et al., 2006). Furthermore mining of transcript profiling data published by Visvader and colleagues shows that Id4 mRNA is more highly expressed in the stem/myoepithelial population when compared to the luminal progenitor and mature luminal cell populations in both humans and mice (Lim et al., 2010). Despite this, no one has examined the protein expression pattern of Id4 during development or the functional role of Id4 in the mammary gland until very recently. Thus we aim to determine the role of Id4 during mammary gland development and hypothesise that it might act through stem and differentiation pathways similar to what is seen in the brain.

4.1.2 Hypothesis:

Id4 regulates the development of the epithelial ductal tree of the mammary gland through the modulation of stem and progenitor cell differentiation pathways.

4.1.3 Aims:

The specific aims of this chapter are to:

1. Characterise the expression pattern of Id4 in the normal mouse mammary gland during development.
2. Characterise the mammary gland phenotype of the Id4 null mouse during development.
3. Determine the function of Id4 in regulating mammary gland development.

4.2 Results:

4.2.1 Id4 is expressed by cap and myoepithelial cells of the mouse mammary gland and is regulated by the oestrus cycle

To understand the role of Id4 in mammary gland development we first wanted to examine the cell types within the mammary gland that express Id4 protein and then to determine if its level of expression was regulated during the different stages of mammary gland development.

4.2.1.1 Id4 protein is expressed by myoepithelial cells throughout mammary gland development

To examine the level and cellular location of Id4 expression in the mouse mammary gland we first optimised immunohistochemical (IHC) staining for Id4. We used a rabbit monoclonal antibody that recognises both human and mouse Id4. Mammary tissue from the Id4 null mouse served as a negative control. Our group and others have previously shown that polyclonal antibodies against Id proteins are often not completely specific, thus careful antibody selection and optimisation is necessary to generate accurate results (Nair et al., 2010; Perk et al., 2006). Using this monoclonal antibody Id4 protein expression was seen to be localised to the myoepithelial cell layer of the wild type mouse mammary gland and no staining was observed in the Id4 null control mammary gland (Figure 4-1A).

We went on to examine the expression of Id4 throughout several post-embryonic developmental stages in the mouse mammary gland. During pubertal development high Id4 levels could be observed in the cap cells and in a small number of the body cells of the terminal end buds (TEBs), while lower levels of Id4 were observed in the myoepithelial cells of the maturing ducts (Figure 4-1B i-ii). In the mature virgin mammary gland varying levels of Id4

could be detected in all cells in the myoepithelial cell layer, however the expression levels varied between individual mice from very strong to very feint myoepithelial staining. Id4 expression was maintained during early pregnancy however from late pregnancy through to lactation and involution the number of Id4 positive cells appeared to be reduced. In particular around the alveolar structures there was only the occasional Id4 positive cell, however, Id4 expression was maintained around the mature ducts (Figure 4-1B iii-viii). Further investigation is needed to determine if the reduction in Id4 expression around the alveolar structures is due to reduced expression by the myoepithelial cells or just a consequence of less observable myoepithelial cells as they stretch out around the alveolar structures. Notably no nuclear Id4 protein expression was observed in any mature luminal cells of the mammary gland, suggesting that Id4 expression is restricted to the myoepithelial cells and a small subset of body cells.

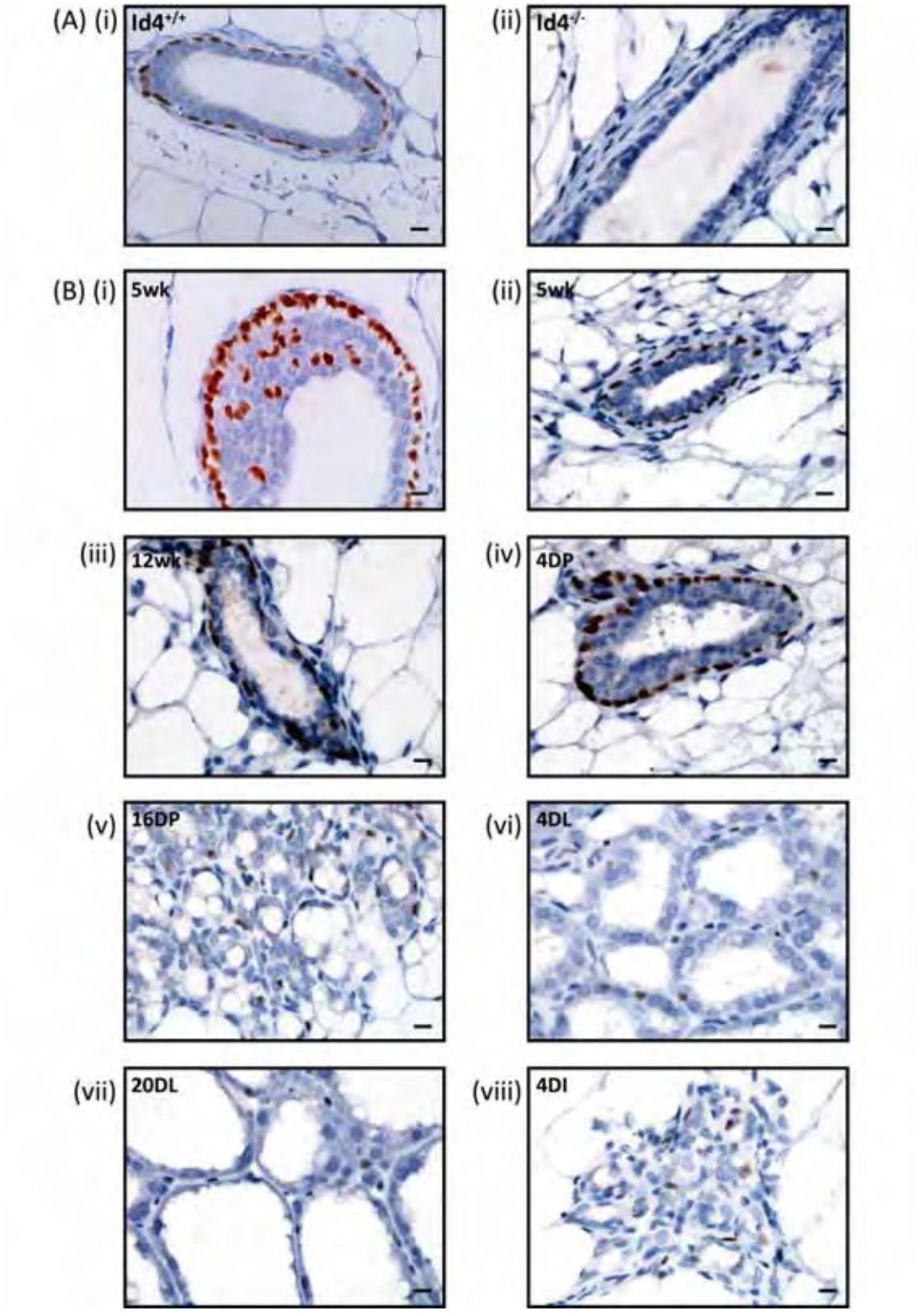


Figure 4-1. Id4 is expressed by the myoepithelial cells of the mammary gland throughout development.

(A) Id4 IHC (brown staining) on wild type $Id4^{+/+}$ (i) and $Id4^{-/-}$ (ii) mammary glands. (B) Id4 IHC (brown staining) on mammary glands from the following developmental stages: pubertal terminal end buds (TEBs 5wk) (i), pubertal ducts (5wk) (ii), 12 week old mature virgin (12wk) (iii), 4 days pregnancy (4DP) (iv), 16 days pregnancy (16DP) (v), 4 days lactation (4DL) (vi), 20 days lactation (20DL) (vii),and 4 days involution (4DI) (viii). Scale bar=10μm.

4.2.1.2 Id4 co-localises with myoepithelial cell but not luminal cell markers

To confirm that Id4 is expressed by myoepithelial cells and not by any of the surrounding stromal or luminal cells I performed co-immunofluorescence for Id4 with a number of myoepithelial and luminal cell markers. This analysis was performed on formalin fixed paraffin embedded (FFPE) sections of mammary glands from mature virgin 12 week old mice. Most importantly, Id4 co-localised with the myoepithelial cell markers p63 and alpha-smooth muscle actin (αSMA), but not the luminal cell marker cytokeratin 8 (CK8) (Figure 4-2). This confirmed that Id4 is expressed by myoepithelial cells but is not expressed by luminal epithelial cells in the mature mammary gland.

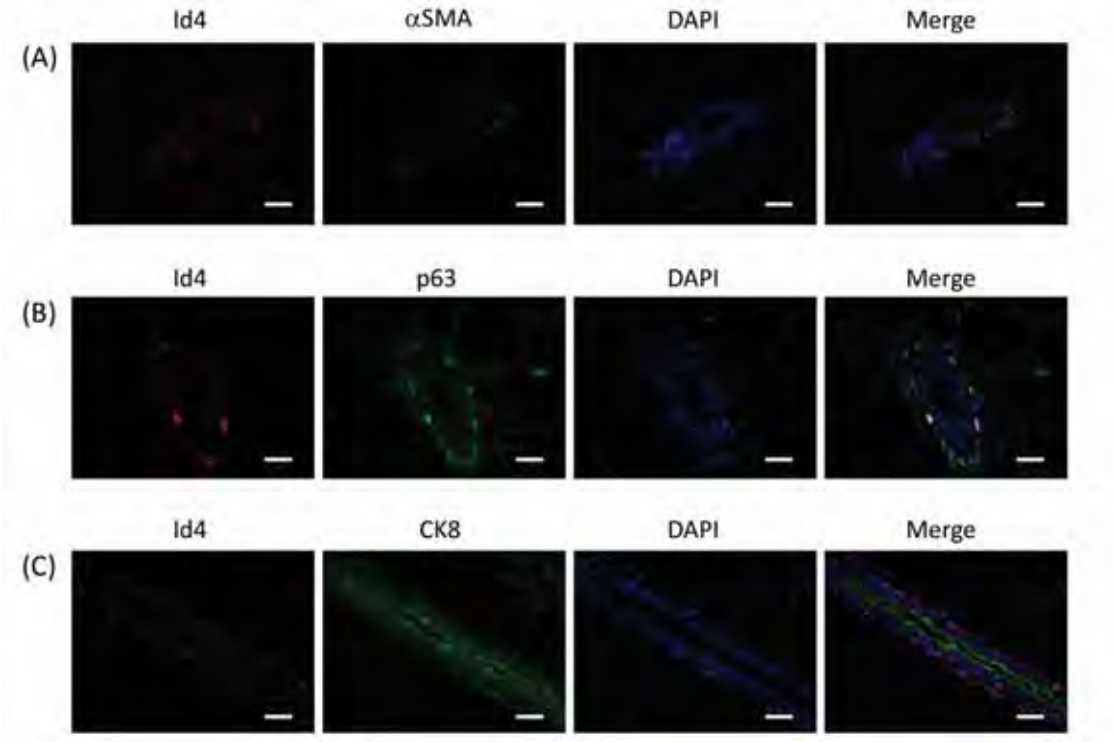


Figure 4-2. Id4 co-localises with myoepithelial cell markers but not with a luminal cell marker in mature 12 week old mammary glands.

(A) Id4 (red), p63 (green), nuclei (blue) and merged image. (B) Id4 (red), αSMA (green), nuclei (blue), and merged image. (C) Id4 (red), CK8 (green), nuclei (blue), and merged image. Scale bar = 20μm.

4.2.1.3 Id4 levels are highest in the cap cells of the developing mammary gland

As Id4 levels appeared to be very high in the TEBs of developing mammary glands by IHC and that a small number of body cells were also positive for Id4 we also used co-immunofluorescence to determine Id4 co-localisation with myoepithelial and luminal markers in the TEBs. This analysis was performed on mammary glands from 8 week old virgin mice. Id4 was highly expressed in the cap cells of the TEBs, however it was almost undetectable in the myoepithelial cells of the ducts. This suggests that Id4 expression is down regulated in the

maturing ducts when compared to the TEBs and the mature ducts of a 12 week old mouse, as described in section 4.2.1.2. In contrast, the myoepithelial marker αSMA was almost undetectable in the cap cells of the TEBs but was strongly expressed by the myoepithelial cells of the ducts (Figure 4-3A). The vast majority of Id4 positive cells in the TEBS were CK8 negative, however a small number of body cells that had lower levels of Id4 expression were double positive for CK8 and Id4 (Figure 4-3B). Taken together, this confirmed that Id4 is highly expressed in the highly proliferative cap cells of the developing mammary gland and that its expression is not present or is down regulated in body cells expressing luminal cell markers.

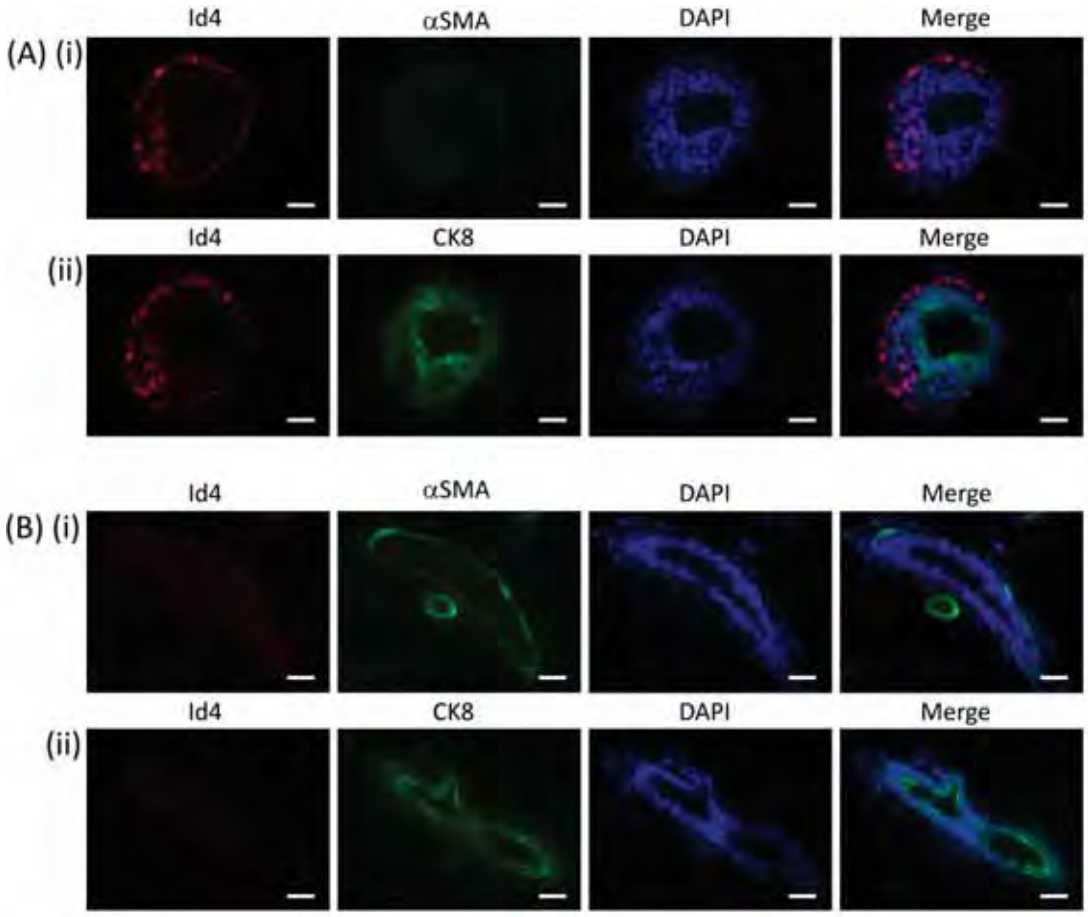


Figure 4-3. Id4 protein levels are highest in the cap cells of the terminal end buds (TEBs) whereas αSMA expression is highest in the ducts of the developing mammary gland.
Co-immunofluorescence analysis of mammary glands from mid puberty (8 weeks of age) examining the TEBs (A) and the maturing ducts (B). (A) (i) and (B) (i)Id4 (red), αSMA (green), nuclei (blue), and merged image. (A) (ii) and (B) (ii) Id4 (red), CK8 (green), nuclei (blue), and merged image. Scale bar = 20μm.

4.2.1.4 Id4 expression in normal human breast tissue

We were interested to see if the strong Id4 staining observed in the myoepithelial cells of the mouse was phenocopied in human mammary glands. Id4 IHC was performed on normal breast tissue from 17 patients. Analysis showed that Id4 expression in human mammary glands was

much more variable than observed in the mouse. Again there was a tendency for higher levels of staining in the myoepithelial cells, however there were also samples that had relatively high levels of Id4 in luminal cells or no Id4 staining at all (Figure 4-4). Thus it appears that the regulation of Id4 in humans is somewhat different to inbred mouse strains. Also the patients in this study were not controlled for age, oestrus stage, or hormonal treatments such as the birth control pill, anti-oestrogens, or hormone replacement therapy and this may influence the expression of Id4 in the mammary gland.

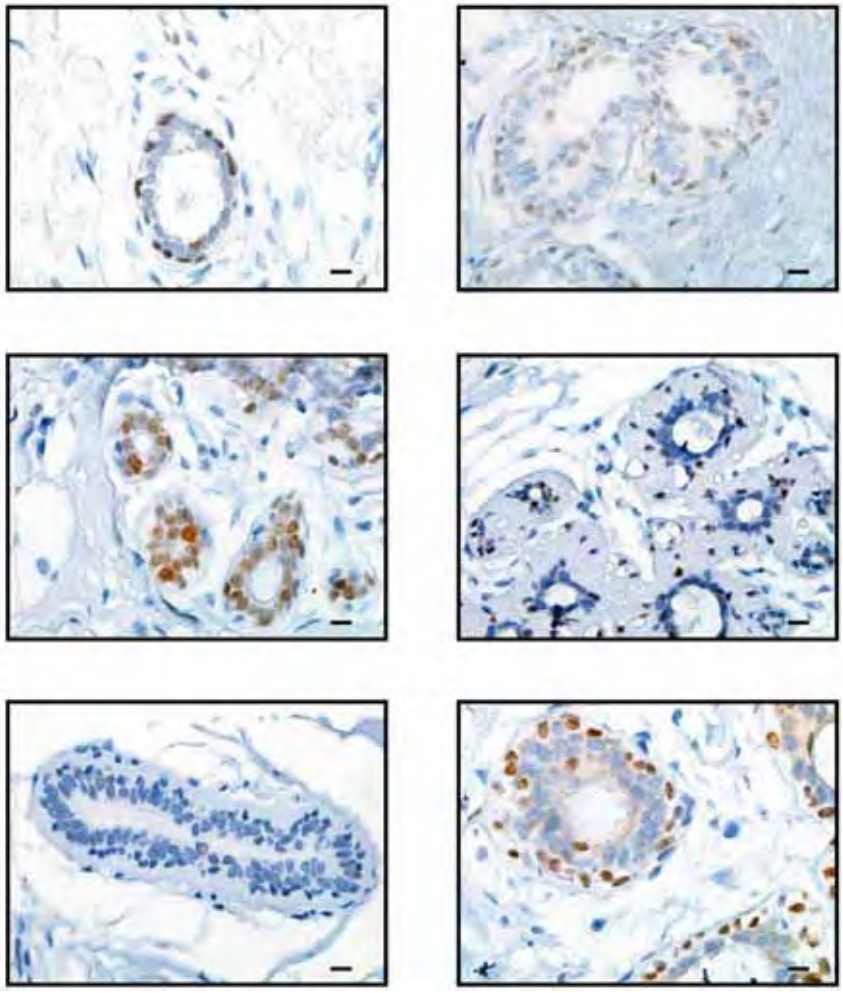


Figure 4-4. Id4 protein expression in normal human breast tissue is enriched but not restricted to the myoepithelial cell layer.
Id4 IHC (brown staining) was performed on samples of normal human breast tissue. Images are from 6 independent patient samples and represent the range of staining patterns observed in samples of human breast tissue. Scale bar=10μm.

4.2.1.5 Id4 levels are regulated through the oestrus cycle

It had been shown by Lydon and colleagues that Id4 mRNA was up regulated in the mammary gland following progesterone treatment (Fernandez-Valdivia et al., 2008). Furthermore I saw variation in mammary gland Id4 expression levels between individual mature 12 week old

virgin mice suggesting that Id4 could be regulated by the oestrus cycle. An outline of how ovarian hormones fluctuate during the oestrus cycle in the mouse is shown in Figure 4-5A. To examine whether Id4 was indeed regulated by the oestrus cycle a cohort of sixteen 10-12 week old virgin female mice were analysed for oestrus stage by vaginal smear cytometry and divided into one of the four stages of the oestrus cycle: proestrus, oestrus, metestrus, or diestrus. FFPE mammary gland sections were then analysed by IHC for Id4 protein expression. Importantly, Id4 was consistently highly expressed at the oestrus stage of the cycle and barely detectable at diestrus (Figure 4-5B). In addition, low levels of cytoplasmic Id4 staining in the luminal cells were also seen at oestrus. This clearly demonstrates that Id4 is regulated by the oestrus cycle, however further experiments are needed to determine if the light cytoplasmic staining in luminal cells at oestrus has any functional significance in mammary gland biology.

4.2.1.6 Ovarian hormones and the progesterone receptor are not necessary for Id4 expression

To determine if the ovarian hormones are necessary for Id4 expression we examined the mammary glands from ovariectomized and progesterone receptor null mice. Ovariectomized mice have had their ovaries removed and do not produce the ovarian hormones oestrogen and progesterone, and thus do not undergo oestrus cycles. In contrast, progesterone receptor null mice still produce oestrogen and progesterone but cannot respond to progesterone signalling. FFPE mammary sections from these mice were obtained from Chris Ormandy (Garvan Institute, Darlinghurst, NSW, Australia). As can be seen in Figure 4-5C Id4 protein was expressed in the mammary glands of these mice at comparable levels to the control mice. This suggests that while the oestrus cycle can regulate the level of Id4 expression the ovarian hormones are not necessary for its basal-level of expression.

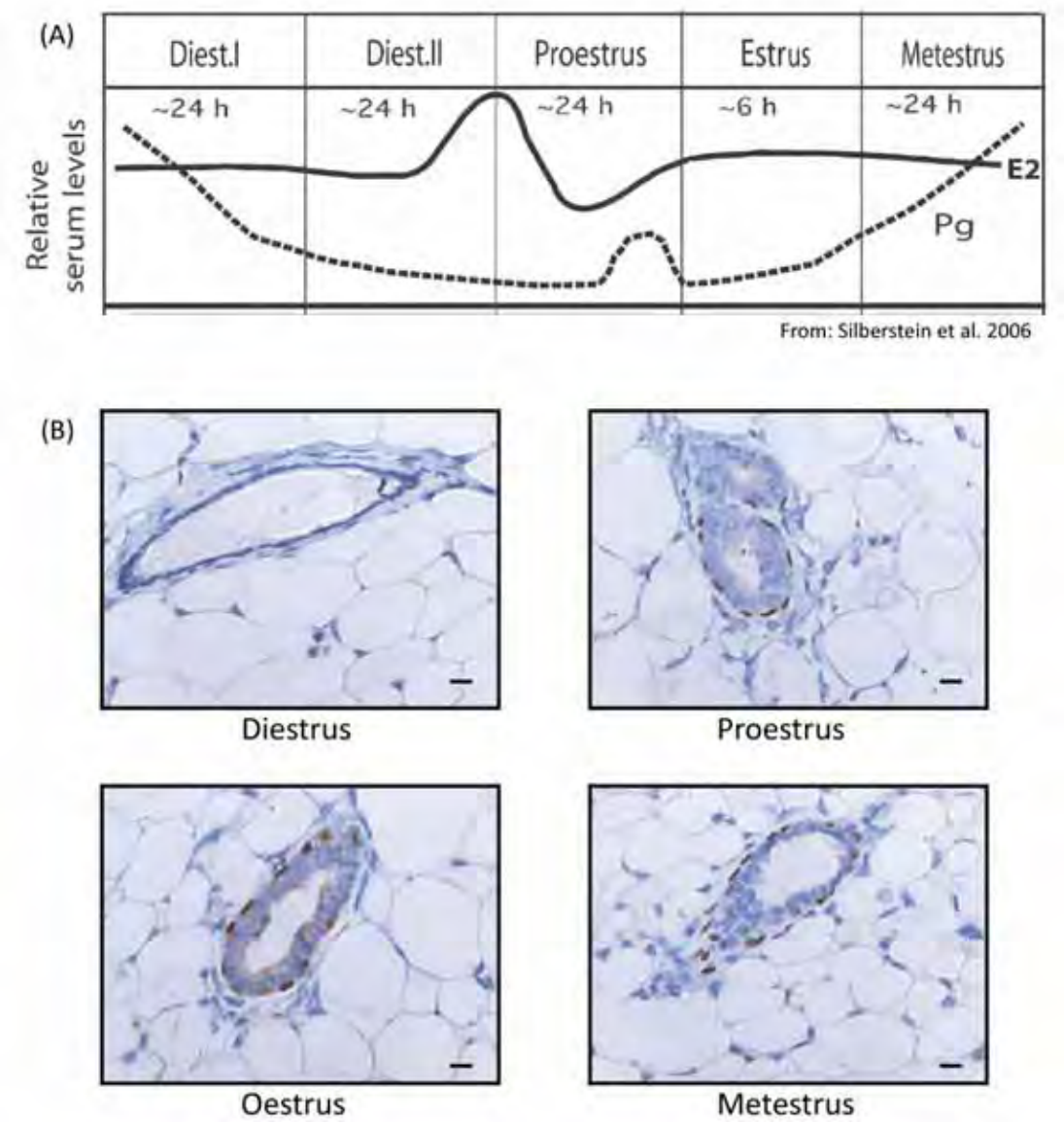


Figure 4-5. Id4 protein levels are regulated through the mouse oestrus cycle, however, Id4 expression does not require ovarian hormones or the progesterone receptor.

(A) Levels of ovarian hormones during the different stages of the oestrus cycle (Silberstein et al., 2006). (B) Id4 IHC (brown staining) on mammary glands from mice at proestrus, oestrus, metestrus, and diestrus, as indicated. (Figure continued on the following page).

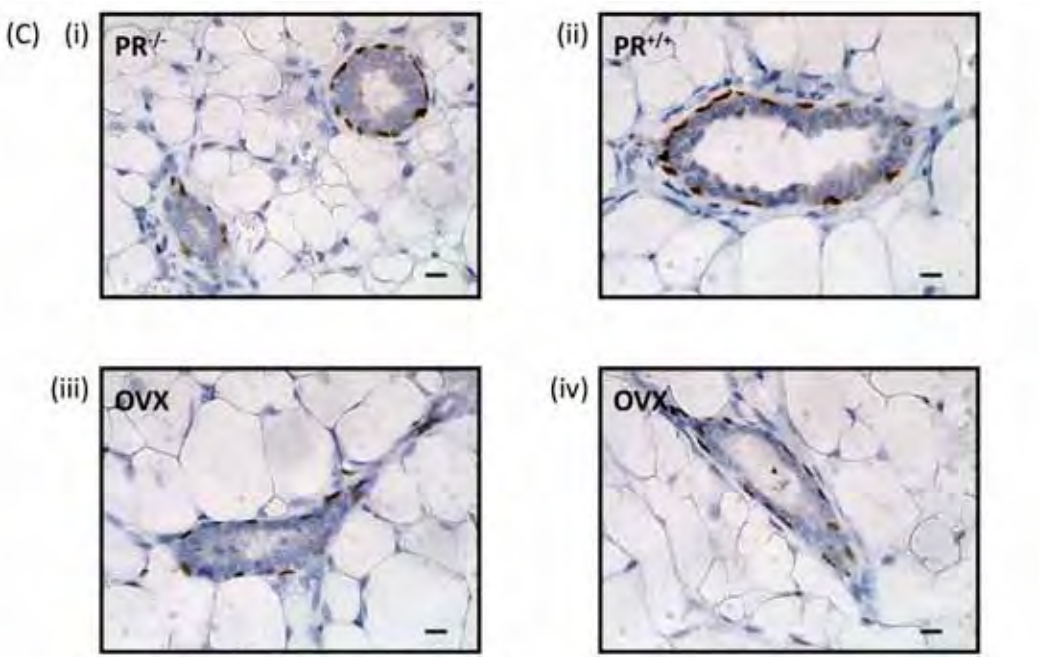


Figure 4-5. Continued. Id4 protein levels are regulated through the mouse oestrus cycle, however, Id4 expression does not require ovarian hormones or the progesterone receptor.
(C) (i) Id4 IHC (brown staining) on mammary glands from progesterone receptor null ($PR^{-/-}$) and (ii) control mice ($PR^{+/+}$). (iii-iv) Id4 IHC (brown staining) on mammary glands from ovariectomized mice (OVX). Scale bars = 10μm.

4.2.2 Loss of Id4 leads to a delay in pubertal mammary gland development

To investigate the role of Id4 in mammary gland development we examined mammary glands from Id4 null mice at different stages of development. These mice had been generated by knocking GFP into the Id4 locus and been previously used to analyse the role of Id4 in the brain (Yun et al., 2004). In these mice GFP is expressed from the endogenous Id4 promoter instead of Id4 protein. The Id4 heterozygous mouse expresses both Id4 protein and GFP. We initially examined the mammary glands of Id4 null mice on the C57BL/6 background, however we also backcrossed these mice onto the FVB/N background and performed further analyses on these mice.

4.2.2.1 Wholemount analysis of Id4 null mammary glands during pubertal development

To determine the role of Id4 in mammary gland development we examined mammary glands from Id4 null mice at various stages of pubertal development by wholemount histology. Wholemount histology is a technique that allows for the clear visualisation of the whole epithelial ductal tree of the mammary gland, and thus is used for the detailed analysis of mammary glands to detect any defects in development. The stages of mammary gland development that were analysed were early puberty when the mice were at 15g which is

approximately 5 weeks of age, 8 weeks of age (late puberty), and 12 weeks of age (fully mature). The Id4 null glands were compared to their heterozygous and wildtype littermates. During early puberty a large defect in mammary gland development was observed (Figure 4-6**Error! Reference source not found.**), this was quantified to show that there was a significant 49% decrease in area of the mammary gland filled. There was also a 58% decrease in the number of ducts formed indicating a branching defect in the Id4 null mammary glands (Figure 4-6**Error! Reference source not found.**). Additionally there were fewer TEBs (46% eduction) in the Id4 null glands although this did not reach significance. In the more severely affected glands the TEBs had very unusual morphologies with the ducts from one mouse displaying a hook-like morphology (Figure 4-6**Error! Reference source not found.** v). By late uberty the mammary gland defect observed in the Id4 null mice was more variable with some glands still having a severe defect while others appearing normal, however, there was still a significant decrease in the area of gland filled (47% reduction) and in the number of ducts (50% reduction) (Figure 4-7). By 12 weeks of age in the mature mammary gland no defect was observed (Figure 4-8). No difference was detected between the wildtype and heterozygous littermates at any stage during mammary gland development, thus heterozygous littermates were used as controls for all subsequent experiments.

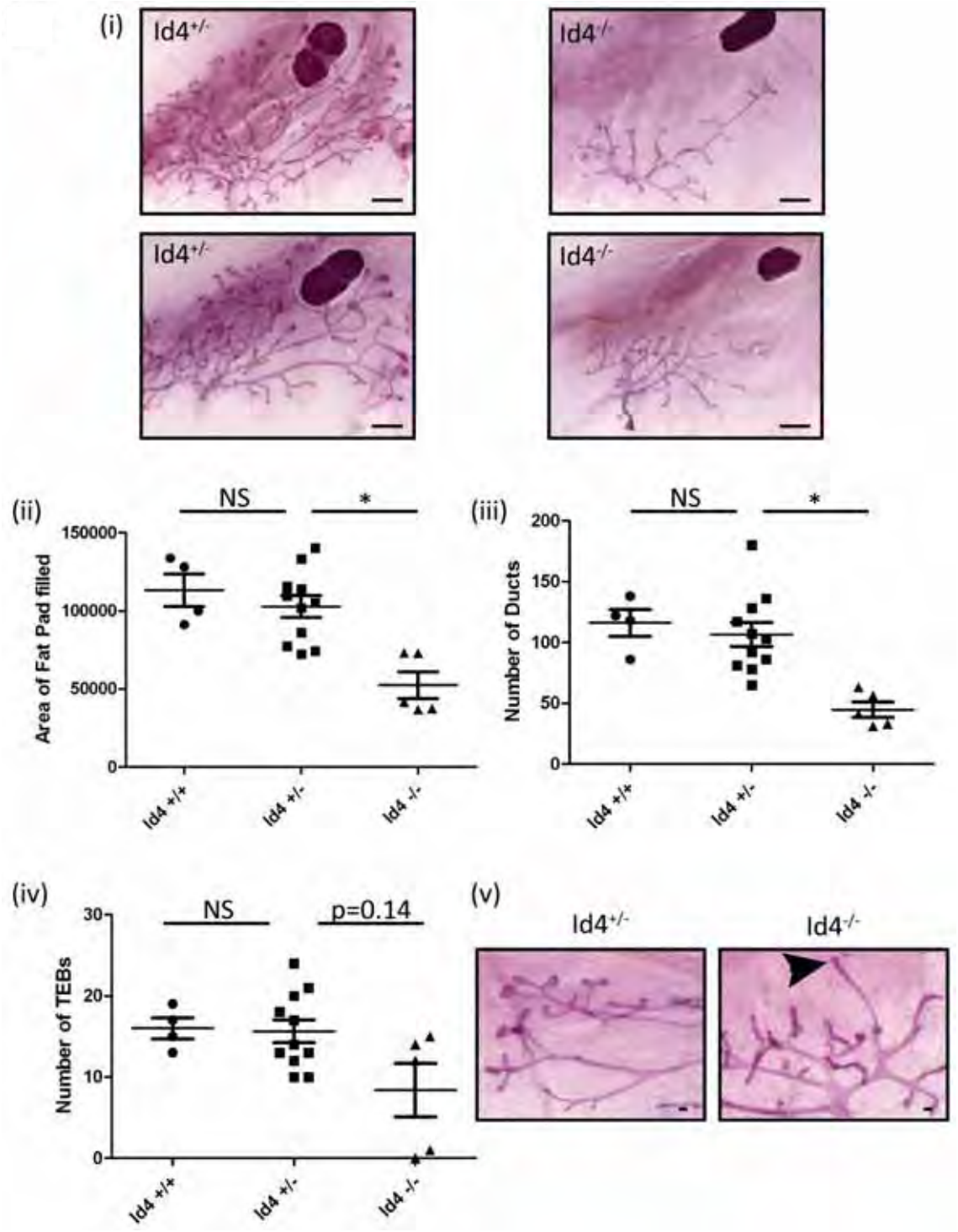


Figure 4-6. Mammary glands from Id4 null mice show a severe defect in pubertal development.
Wholemount histology of mammary glands from 15g C57BL/6 Id4 null ($Id4^{-/-}$), and Id4 heterozygous ($Id4^{+/-}$) mice, scale bar=1mm. (ii)-(iv) mammary tissue from Id4 wild type ($Id4^{+/+}$), Id4 heterozygous and Id4 knockout mice was prepared as in (i) and examined for the following: (ii) area of fat pad filled by epithelium was calculated using ImageJ software in (arbitrary units), (iii) number of ducts per field of view were counted, and (iv) number of TEBs per mammary gland are quantitated from 5-10 mice per group. Graph points indicate individual mice. Bars indicate average +/- SEM. *p<0.05 unpaired t-test. (v) TEBs from some Id4 null mice show a very unusual morphology, scale bar=100μm.

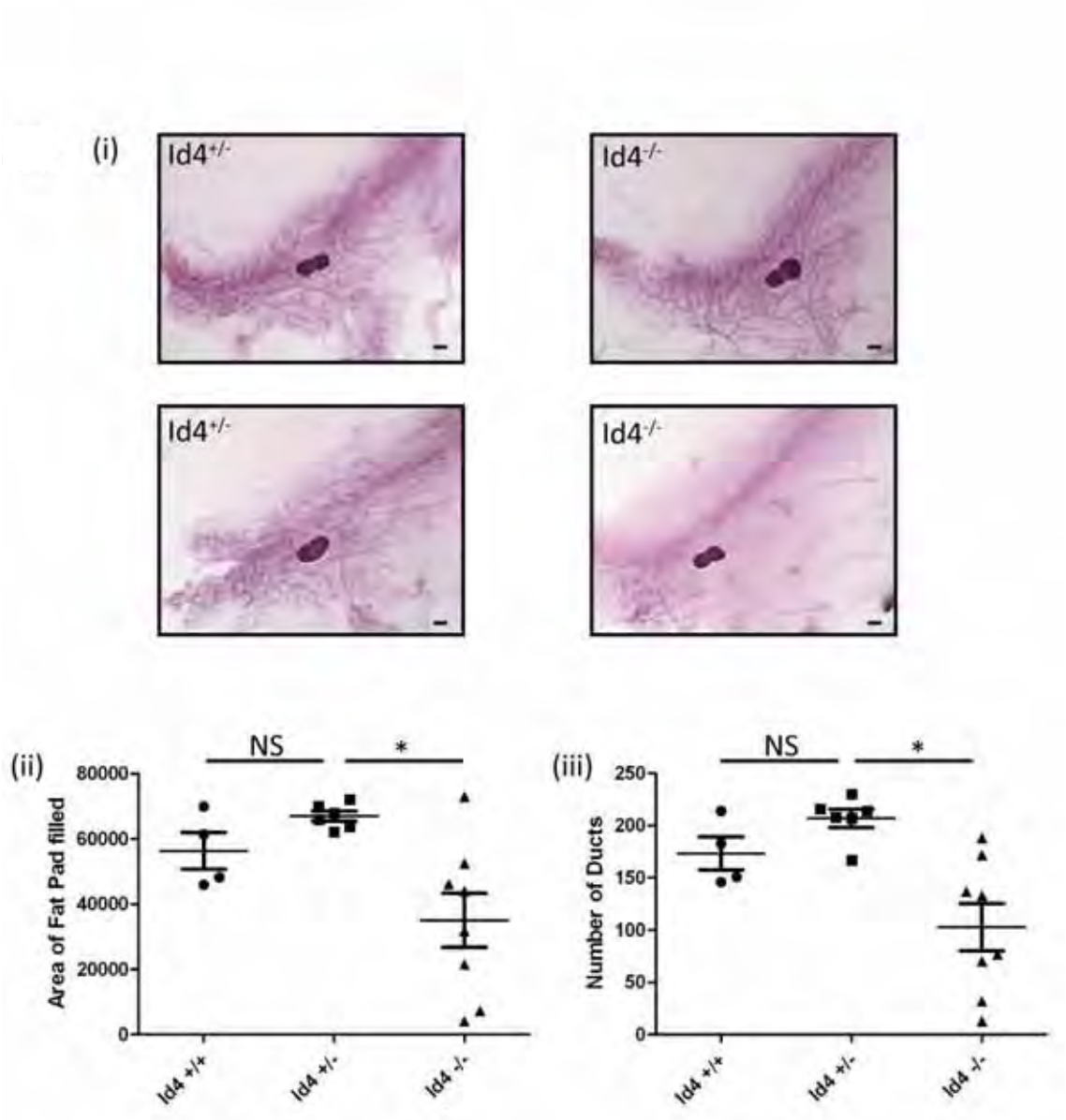


Figure 4-7. Mammary glands from Id4 null mice show a severe defect in pubertal development
Wholemounted mammary glands from late pubertal 8 week old C57BL/6 Id4 null ($Id4^{-/-}$), and Id4 heterozygous ($Id4^{+/-}$) (Scale bar=1mm), (ii) area of fat pad filled (arbitrary units) as determined by ImageJ software, and (iii) number of ducts per gland in Id4 null, Id4 heterozygous and Id4 wild type ($Id4^{+/+}$) C57BL/6 mice. Graph points indicate individual mice. Bars indicate average +/- SEM. *p<0.05 unpaired t-test.

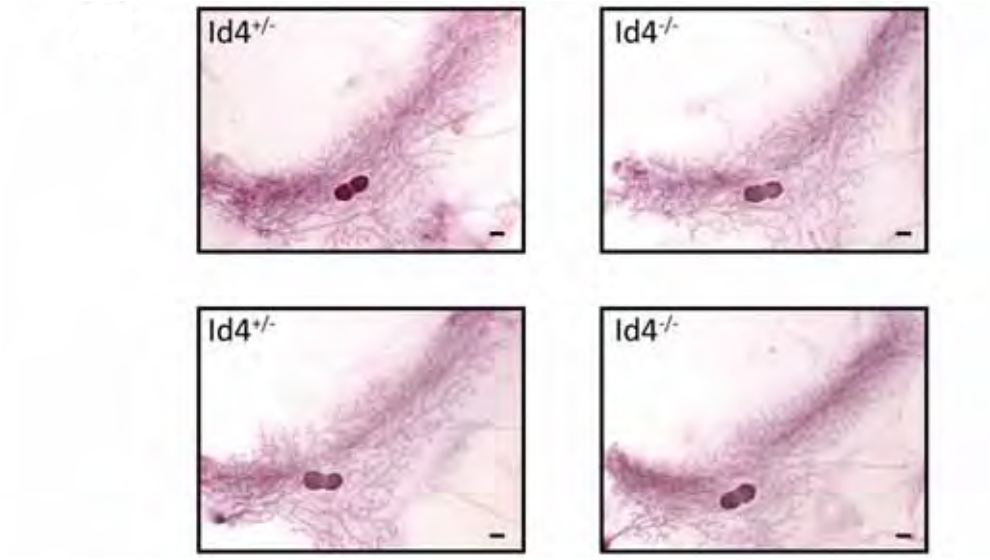


Figure 4-8. Mammary gland development defect is overcome by 12 weeks of age in C57BL/6 mice. Wholemounted mammary glands from mature 12 week old Id4 null ($Id4^{-/-}$) and heterozygous control ($Id4^{+/-}$) mice on a C57BL/6 genetic background.

The Id4GFP mice were also backcrossed 5 generations onto the FVB/N strain primarily for breast cancer studies, however, it was noted that the mammary gland defect in these mice was more consistent than on the C57BL/6 background. Wholemount analysis was initially performed at 8 weeks and 12 weeks of age. The mammary glands from the 8 week old mice showed a more significant defect than the mice on the C57BL/6 strain at this age with a dramatic 72% reduction of area of fat pad filled (Figure 4-9**Error! Reference source not found.**). This defect was maintained to a smaller extent at 12 weeks of age (20% reduction) (Figure 4-10) and was even observed in some mice at 16 weeks of age (data not shown), although the defects at these stages did not reach significance. This confirms that the defect in mammary gland development due to loss of Id4 is not dependent on the background strain of the mouse. As the phenotype was more consistent in mice from the FVB/N background all subsequent experiments were performed on these mice unless specifically indicated.

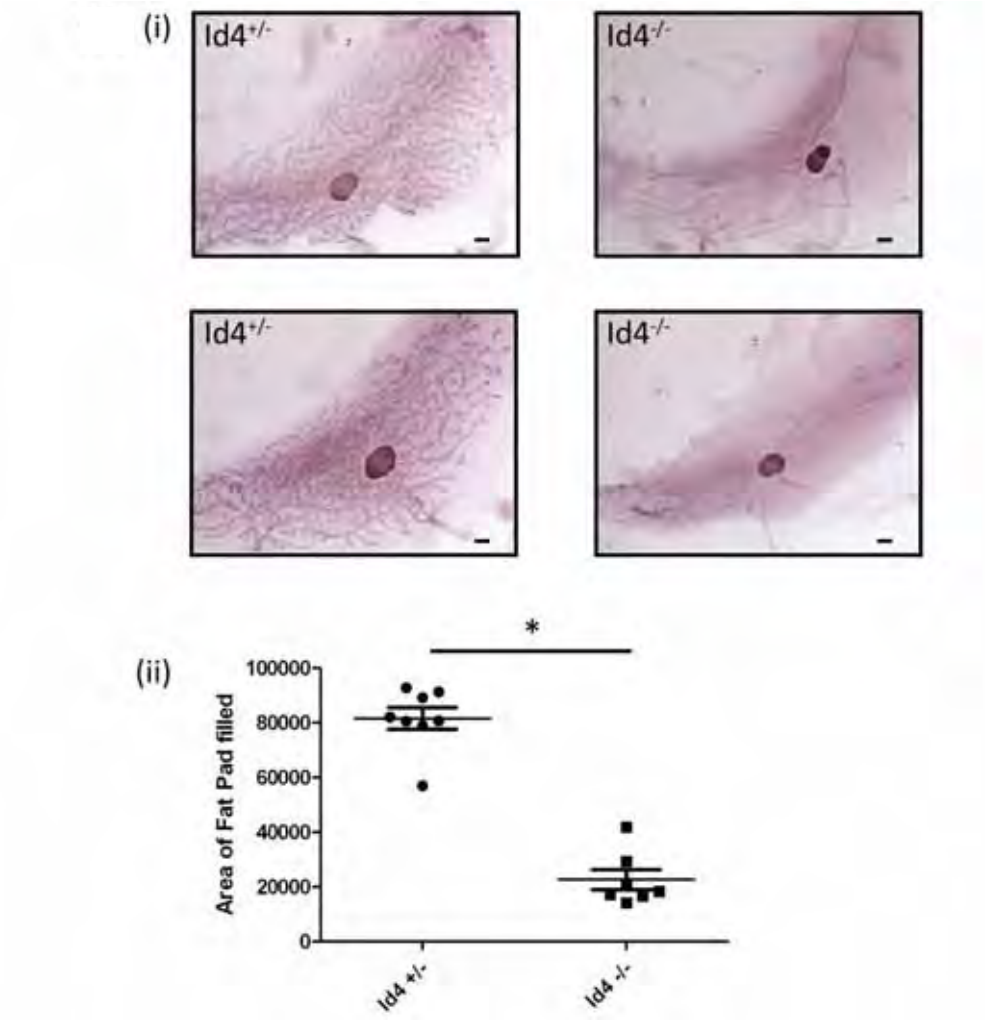


Figure 4-9. Mammary glands from Id4 null mice show a severe defect in pubertal development.
(i) Late pubertal 8 week old Id4 null ($Id4^{-/-}$), and Id4 heterozygous ($Id4^{+/-}$) mice on an FVB/N genetic background. (Scale bar=1mm). (ii) Area of fat pad filled in was determined in for tissues from mice described in (i) (arbitrary units). Bars indicate average +/- SEM. *$p<0.05$ unpaired t-test.

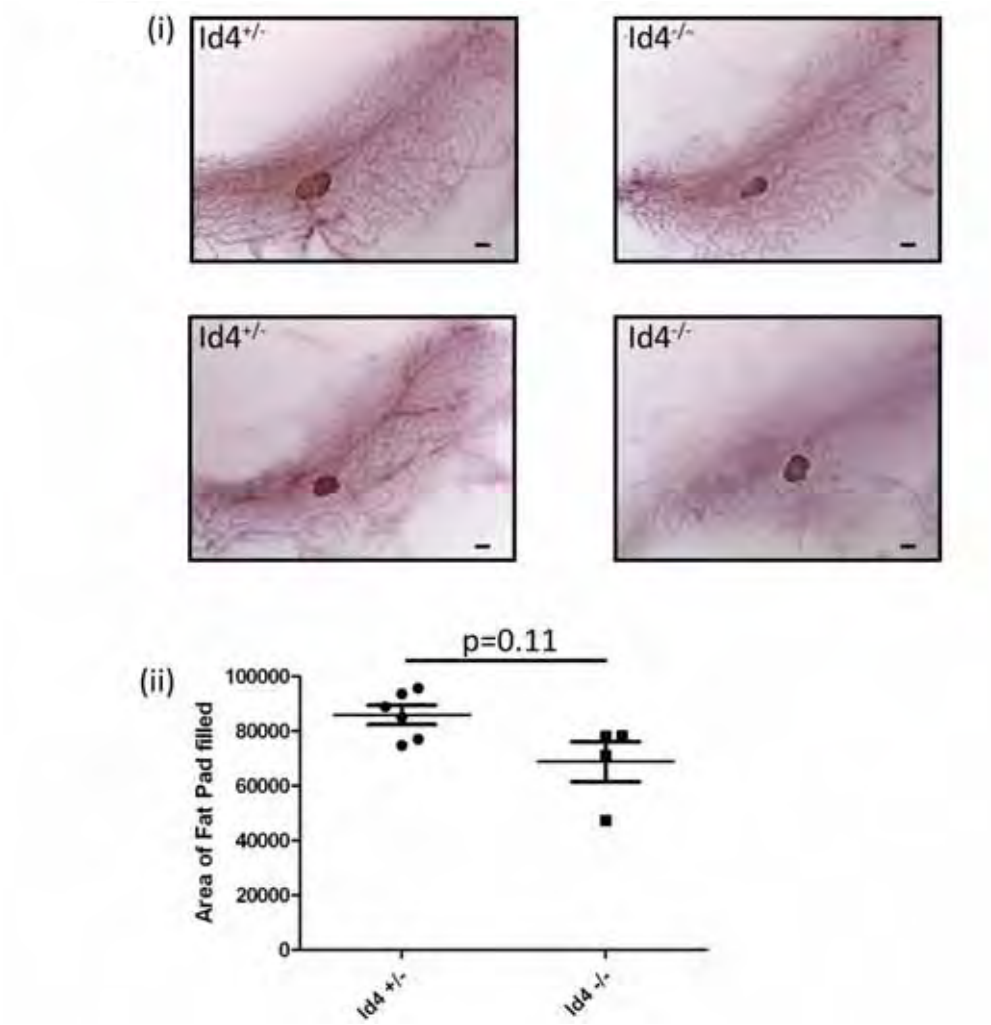


Figure 4-10. Mammary glands from Id4 null mice show a severe defect in pubertal development.
(i) Wholemounted mammary glands from mature virgin 12 week old FVB/N Id4 null ($Id4^{-/-}$), and Id4 heterozygous ($Id4^{+/-}$) mice and (Scale bar=1mm) (ii) area of fat pad filled (arbitrary units). Graph points indicate individual mice. Bars indicate average +/- SEM. *p<0.05 unpaired t-test.

4.2.2.2 Mammary phenotype of Id4 null mammary glands appears to be epithelial cell specific

To determine if the mammary phenotype of the Id4 null mice was due to an epithelial specific function of Id4 as opposed to a systemic effect of Id4 deficiency we used a transplantation approach. Small pieces (~$1mm^3$) of mammary gland were harvested from donor mice (12 weeks old) either Id4 null or heterozygous (C57BL/6 background), these were then transplanted into the 4^{th} fat pad of recipient mice cleared of endogenous epithelium. For each of the 5 donors per genotype mammary epithelial pieces were transplanted into 4-5 recipient mice, spread over two separate experiments. For each recipient mouse one side was transplanted with an Id4 null mammary piece and the other side with an Id4 heterozygous mammary piece, so that each Id4 null transplant experienced the same hormonal milieu throughout the experiment as the control transplant. To mimic the 15g/5 week old phenotype

we harvested the glands 5 weeks after transplantation and analysed these by wholemount histology (Figure 4-11). The area of fat pad filled and the number of ducts were quantified, averaged for each donor mouse and then plotted and analysed for statistical significance (Figure 4-11). A trend to reduced fat pad filling (31.5% reduction) and number of ducts (39% reduction) was seen in the Id4 null transplants, however this did not reach statistical significance. Further replicates would need to be performed to see if this would reach statistical significance, however in a later section we instead utilise a limiting dilution approach to look at mammary repopulating and growth characteristics (see section 4.2.4.3).

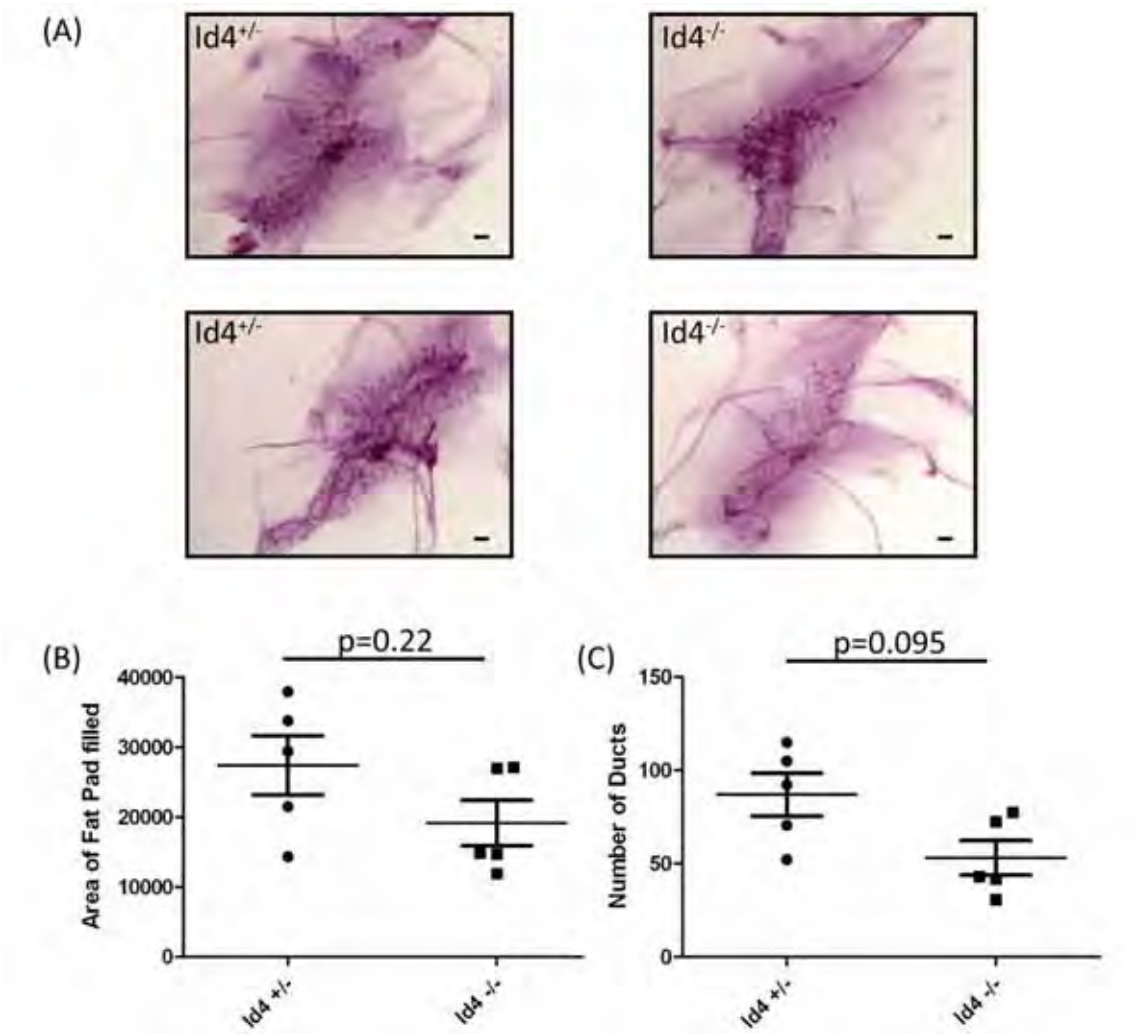


Figure 4-11. Defect in mammary gland development in the Id4 null mice appears to be epithelial cell intrinsic.

(A) Wholemount analysis of mammary epithelial transplant outgrowths from Id4 heterozygous ($Id4^{+/-}$) and null ($Id4^{-/-}$) C57BL/6 donor mice 5 weeks after mammary epithelial transplant. Scale bar=1mm. Analysis of (B) area of fat pad filled by mammary epithelium (arbitrary units), and (C) number of ducts from Id4 heterozygous and null. This analysis was performed by first averaging the counts per recipient mouse for each of the five donors then plotting the average results from each of the five donors individually. Bars indicate the average of the 5 donors ± SEM. p values were calculated using the unpaired t-test.

4.2.2.3 Loss of Id4 does not significantly affect the ability of mammary epithelial cells to differentiate into milk producing alveolar structures

There are two major stages of mammary gland development after birth, one is during puberty and the second during pregnancy. Since normal Id4 expression is crucial for mammary gland development during puberty we wanted to see if Id4 contributed to development during pregnancy. To do this we again adopted a transplantation approach. We transplanted small pieces of mammary gland from heterozygous or Id4 null mice into the cleared fat pad of 3 week old C57BL/6 wild type mice, allowed the transplant to engraft for 2 months and then the mice were time-mated, whereby mice were checked for vaginal plugs to confirm the date of pregnancy. Glands were then harvested at day 18 of pregnancy and wholemounted. There was no significant change in alveolar development when the wholemounts were analysed (Figure 4-12A). Further analysis of sections stained with haematoxylin and eosin (H&E) demonstrated no major morphological differences between the genotypes, and furthermore when tissue sections were stained for milk proteins the level of milk production appeared identical between the genotypes (Figure 4-12B-C). Again there was a trend to reduced fat pad filling in the Id4 null mammary transplants (Figure 4-12D). Even though there were only 6 transplants per group from 2 donors this data almost reached significance. This suggests that Id4 is not necessary for pregnancy induced mammary gland development and further demonstrates a role for Id4 in ductal elongation.

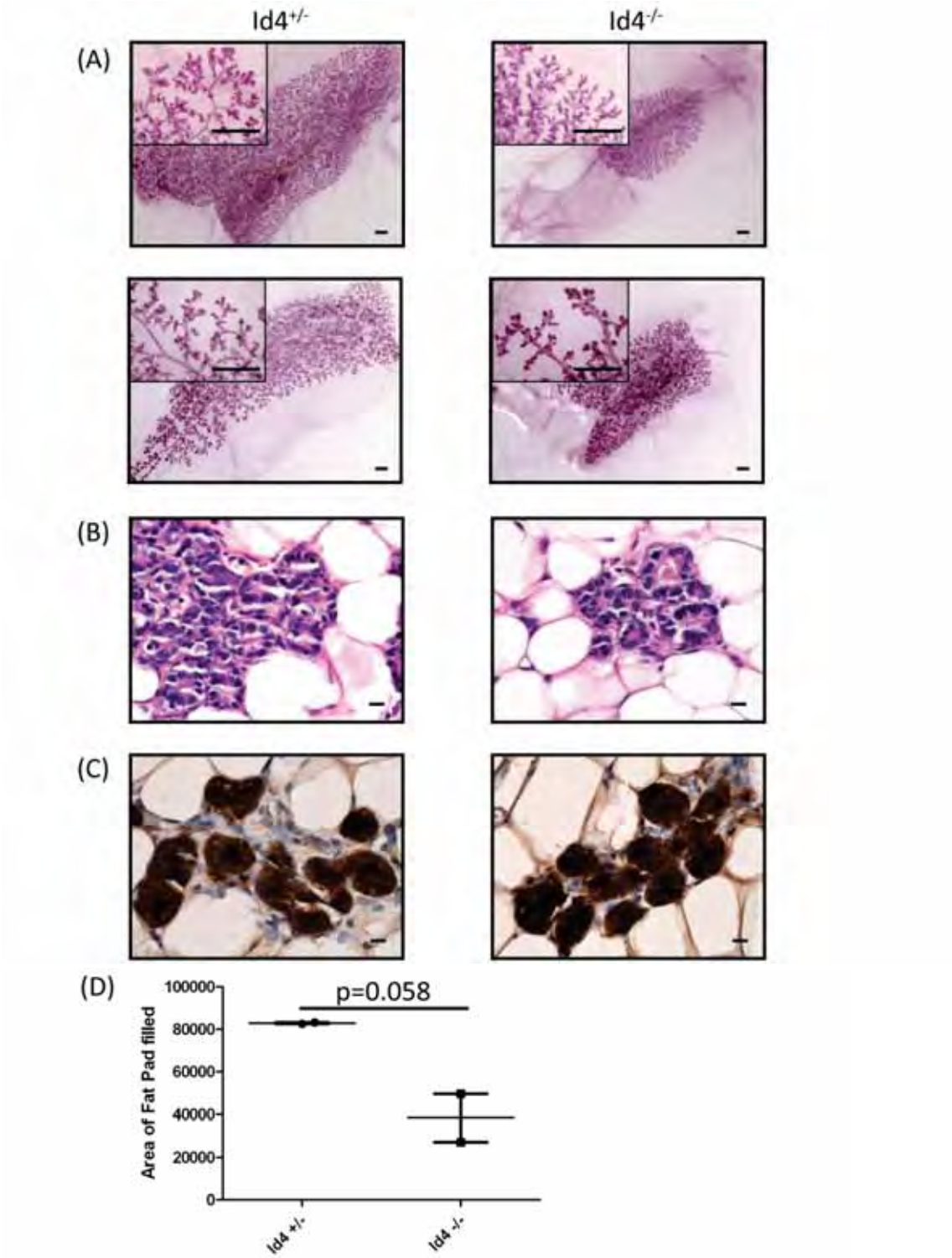


Figure 4-12. Loss of Id4 does not affect mammary epithelial cells ability to differentiate into alveolar structures.

(A) Wholemounted mammary transplants from Id4 heterozygous ($Id4^{+/-}$) and null ($Id4^{-/-}$) donors, into C56BL/6 recipient mice taken from mice after 8 weeks of engraftment and 18 days of pregnancy, with higher magnification inserts (Scale bar=1mm). (B) Representative images of H&E stained histology of alveolar structures and (C) IHC for milk proteins (brown staining) in alveolar structures (Scale bar=10μm). (D) Analysis of area of fat pad filled by mammary epithelium (arbitrary units), area of fat pad filled was first averaged per donor mouse and then averaged for each genotype this was then plotted. p=0.058 unpaired t-test.

4.2.3 Loss of Id4 does not affect levels of myoepithelial cell markers

As Id4 is primarily expressed in the myoepithelial cells of the mammary gland I investigated whether loss of Id4 impacted the expression of other myoepithelial cell markers. It is of note that the mammary ducts that do form in the Id4 null mice have a normal bilayered epithelial structure as is observed in sections stained with H&E (Figure 4-13). The levels of the myoepithelial cell markers cytokeratin 14 (CK14), cytokeratin 5 (CK5), and the functional myoepithelial marker αSMA were compared between the Id4 null and heterozygous control mice in mature virgin mammary glands. No differences in the spatial organisation or the expression levels of these markers could be observed (Figure 4-13), suggesting that Id4 is not required for normal myoepithelial cell development.

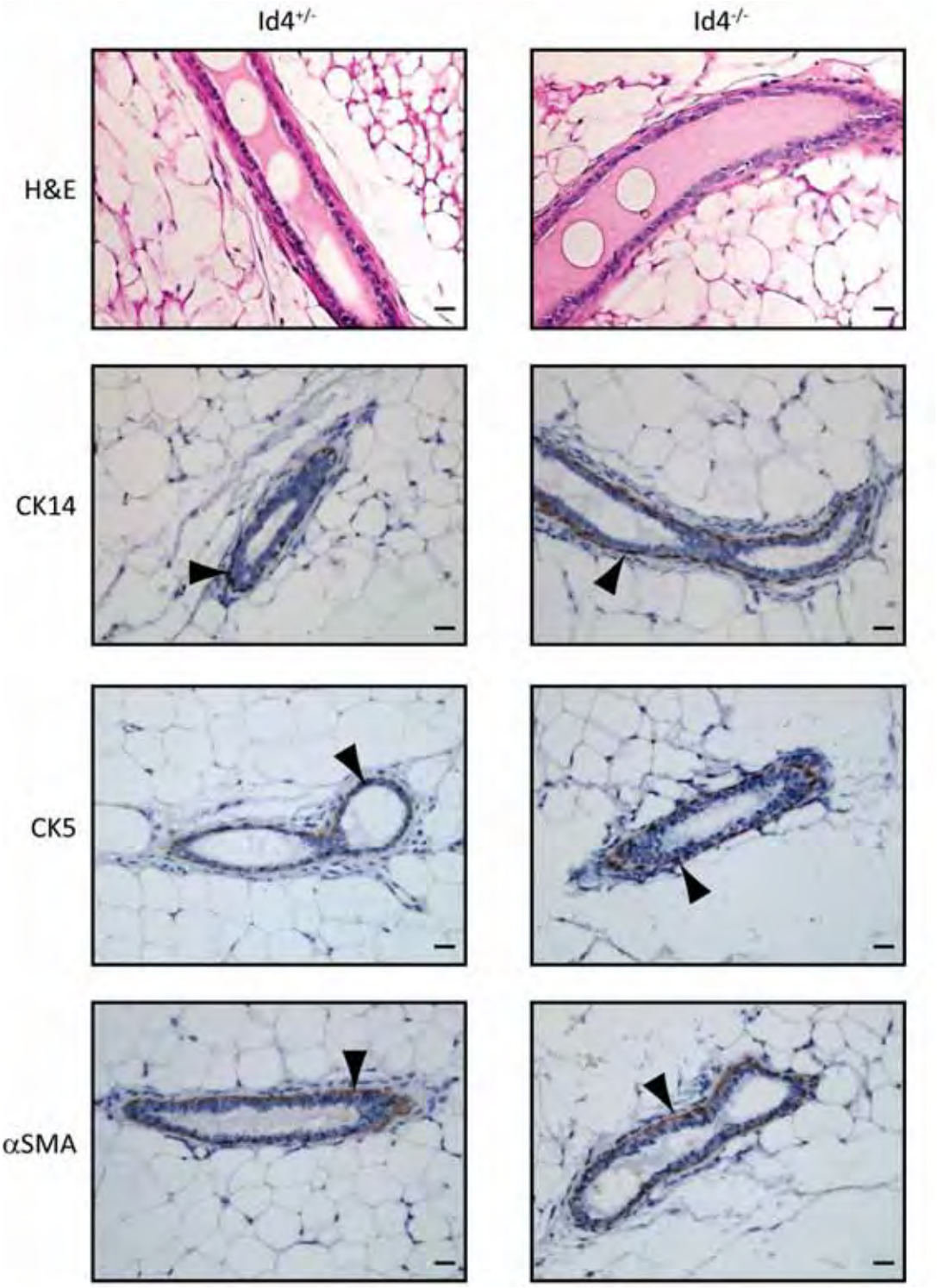


Figure 4-13. Loss of Id4 does not affect expression levels of myoepithelial cell markers.
Control Id4 heterozygous ($Id4^{+/-}$) tissue left panels, and Id4 null ($Id4^{-/-}$) tissue right panels from 12 week old mice. Normal ductal histology of the Id4 null mammary gland stained with haematoxylin and eosin. IHC for the myoepithelial markers (brown staining) CK14, CK5 and αSMA in Id4 null and heterozygous mammary glands from mature virgin mice are shown, as indicated. Arrowheads indicate examples of positive staining in the myoepithelial cell layer. Scale bars = 10μm.

4.2.4 Loss of Id4 leads to an increase in luminal progenitor cells and reduction in stem cells

To determine if loss of Id4 is leading to a defect in mammary stem cell and differentiation pathways we used *in vivo* and *in vitro* models to characterise these cell populations in the Id4 null mammary gland.

4.2.4.1 Loss of Id4 leads to an increase in CD24^{+} CD29^{low} CD61^{+} luminal progenitor cells

By using cell surface markers and flow cytometry it is possible to differentiate mature luminal cells from luminal progenitor cells and the stem/myoepithelial cells (Asselin-Labat et al., 2007). Unfortunately there are currently no cell surface markers that can delineate myoepithelial cells from myoepithelial progenitor cells or from mammary stem cells so these cannot be analysed by flow cytometry. Since the level of Id4 fluctuates through the oestrus cycle, and the percentages of the various cell populations also vary during the oestrus cycle, we restricted our analysis to mice at oestrus by staging the mice using vaginal swab cytometry (Joshi et al., 2010). Single cell suspensions were made from the mammary glands of 10-12 week old Id4 null mice and their heterozygous littermates by collagenase and trypsin digestion. These single cell suspensions were stained for the cell surface markers CD24, CD29 and CD61 along with lineage markers for the exclusion of stromal cells (CD45, CD31, Ter119 and BP-1). DAPI was added so that dead cells could be excluded. By gating on live, lineage negative, CD24 positive cells, it was possible to divide the epithelial cells into three populations, a mature luminal cell population (CD29 low and CD61 negative), a luminal progenitor cell population (CD29 low CD61 positive), and a stem/myoepithelial cell population (CD29 high CD61 positive) (Figure 4-14A). There was a 37% increase in the percentage of cells in the luminal progenitor population and a concomitant 20% reduction in the percentage of mature luminal cells in the Id4 null mammary gland (Figure 4-14B). These differences were significant if analysed using a paired t-test, however additional replicates will be performed to see if this represents a robust increase in luminal progenitors. There was no change in the percentage of stem/myoepithelial cells between the Id4 null and the heterozygous controls (Figure 4-14B).

Using this analysis it was also possible to determine the population of cells where the Id4 promoter is active as measured by GFP expression. Id4 promoter activity was restricted to the stem/myoepithelial fraction of mammary epithelia cells (Figure 4-14C). It was surprising however, that GFP expression was only observed in 15-20% of the stem/myoepithelial cells in heterozygous mice considering that Id4 protein, as measured by IHC, is detected in the vast

majority if not all myoepithelial cells. Taken together, Id4 is expressed in the stem/myoepithelial cell population of the mammary epithelial cells and when absent there is an increased percentage of luminal progenitor cells and a decreased percentage of mature luminal cells.

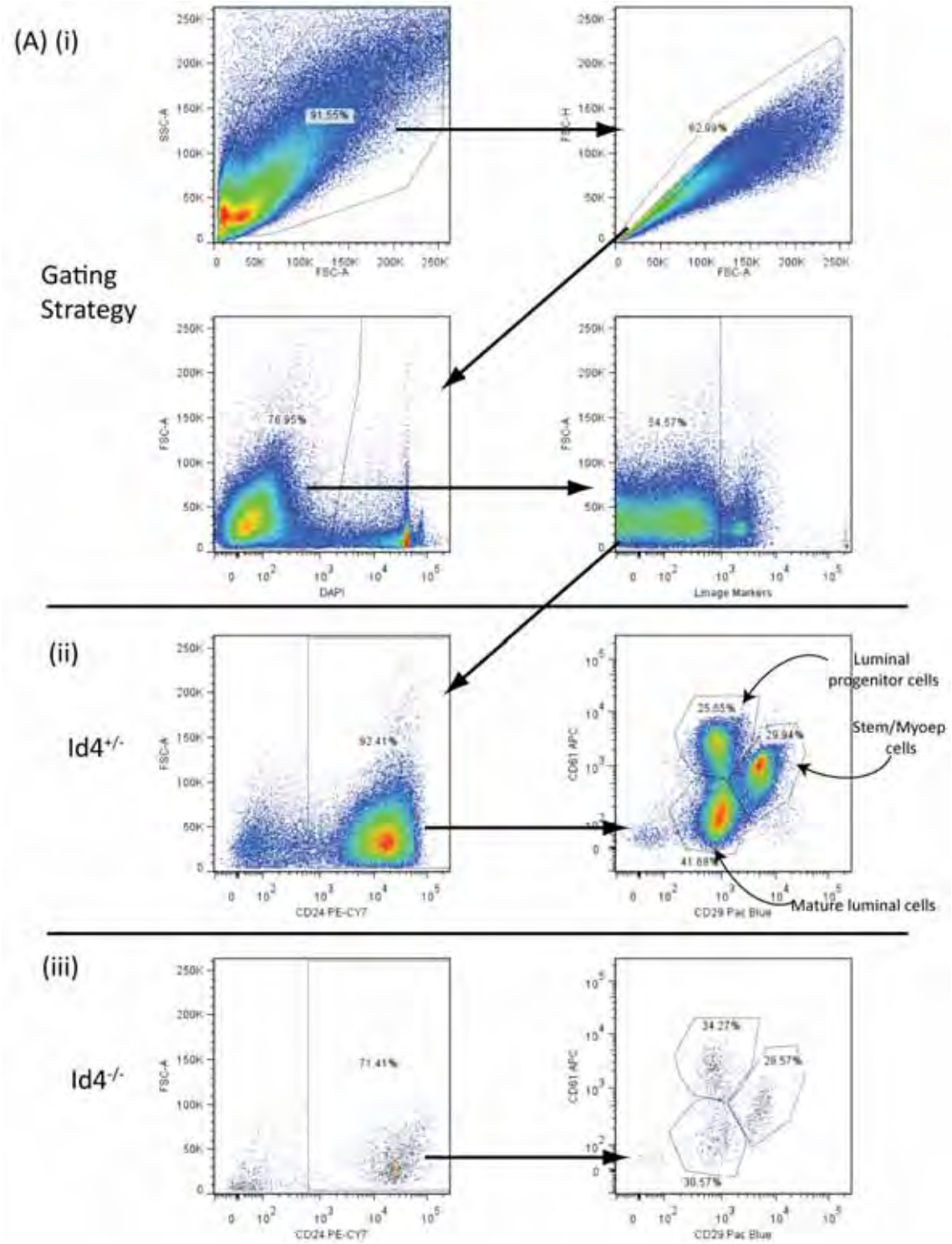


Figure 4-14. Loss of Id4 leads to an increase in luminal progenitor cells as measured by flow cytometry.

(A) (i) To examine mammary epithelial subsets we used the following gating strategy: Small particles were excluded based on forward scatter area (FSC-A) and side scatter area (SSC-A), single cells were selected based on FSC-A verses forward scatter height (FSC-H). Dead DAPI positive cells were excluded, as well as cells positive for the stromal linage markers. CD24 positive epithelial cells were then gated. (ii) Representative plots of the percentages of CD24 positive cells and CD29 and CD61 positive cells from Id4 heterozygous mice ($Id4^{+/-}$)and (iii) Id4 null mice ($Id4^{-/-}$). Mature luminal cell, luminal progenitor cell, and stem/myoepithelial cell populations are indicated with arrows. (Figure continues on following 2 pages)

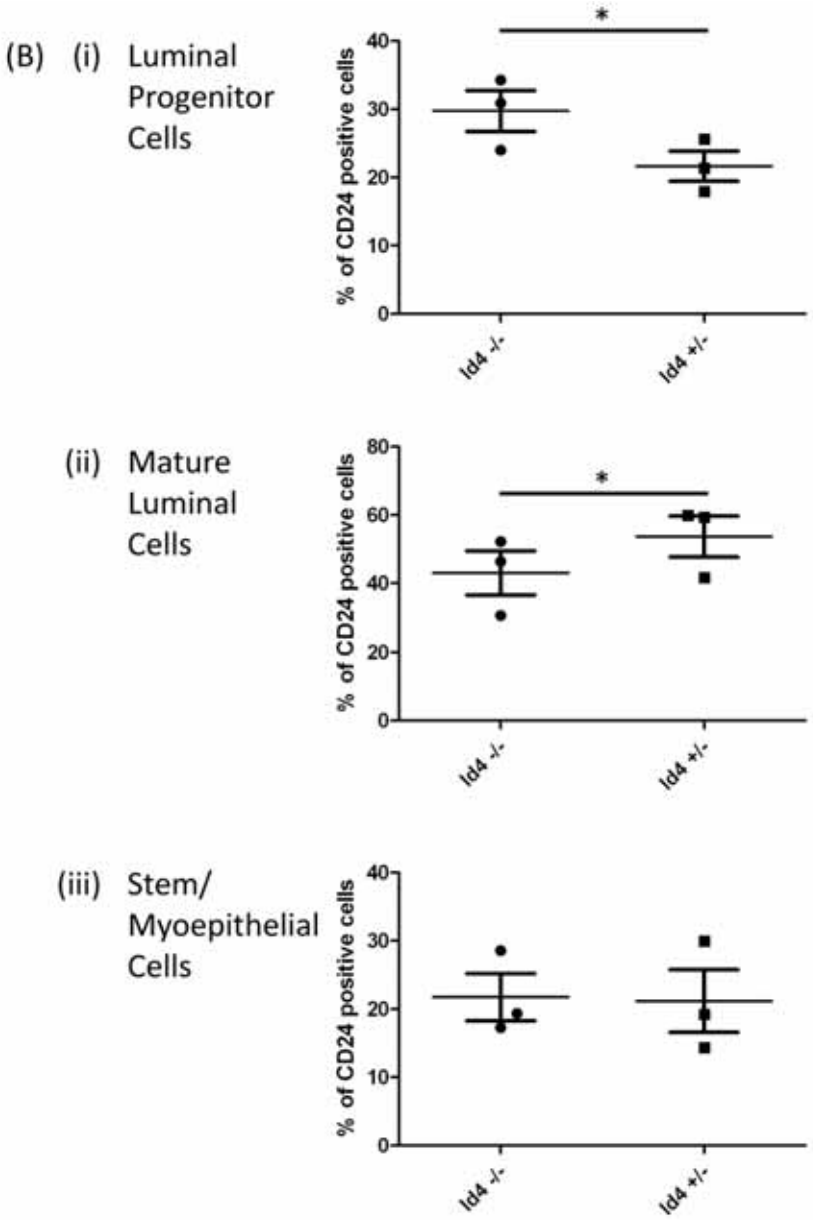


Figure 4–14 Continued. Loss of Id4 leads to an increase in luminal progenitor cells as measured by flow cytometry.

(B) Average percentage of epithelial cell populations from multiple experiments (i) luminal progenitors, (ii) mature luminal and (iii) stem/myoepithelial cells, from Id4 knockout mice ($Id4^{-/-}$) and Id4 heterozygous mice ($Id4^{+/-}$). Error bars indicate SEM. *p<0.05 paired t-test. (Figure continued on following page)

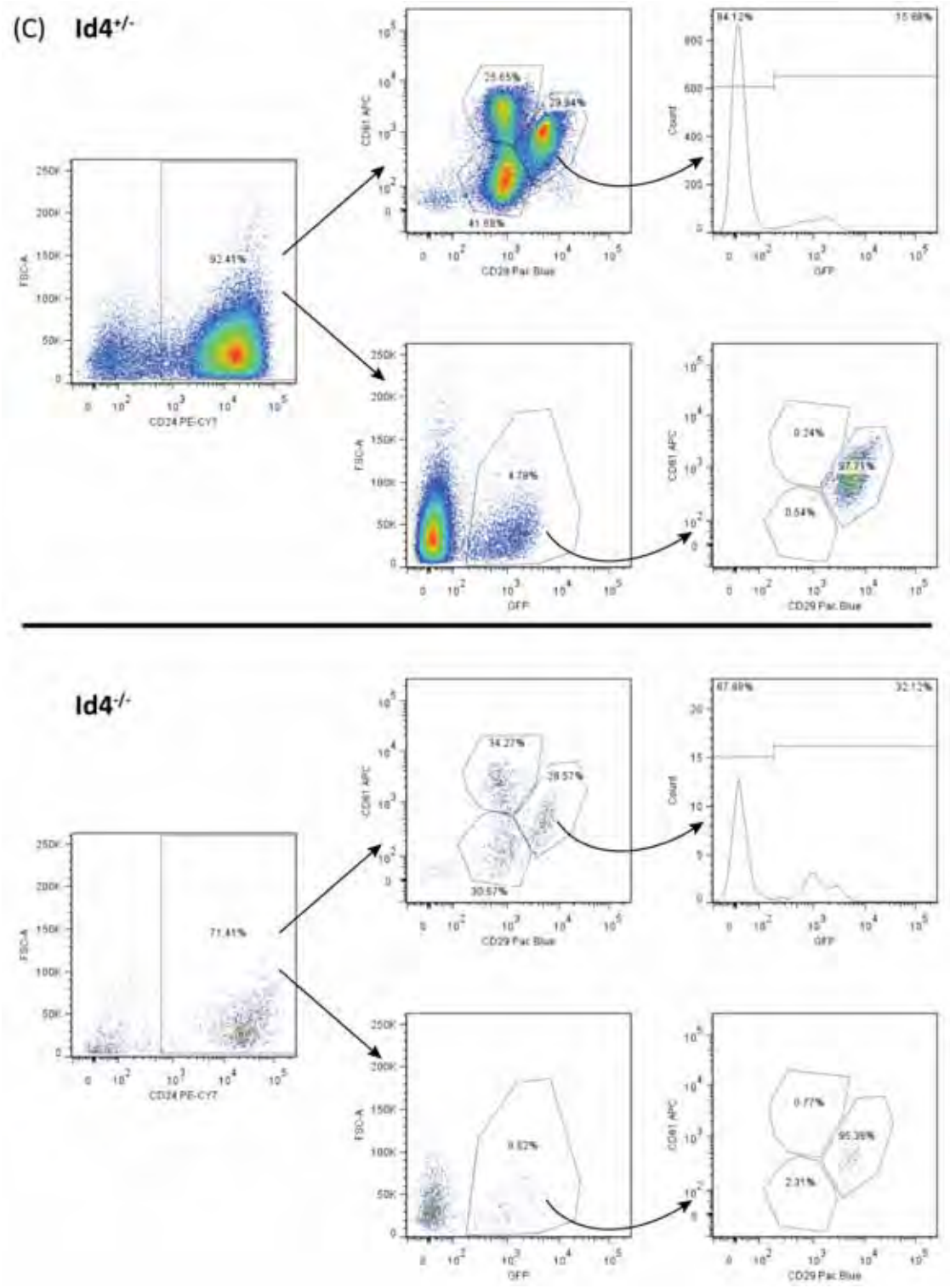


Figure 4–14 Continued. Loss of Id4 leads to an increase in luminal progenitor cells as measured by flow cytometry.

(C) Analysis of Id4 promoter activity in the cell populations, as measured by GFP expression from the Id4 promoter in the $Id4^{+/-}$ and $Id4^{-/-}$ mice.

4.2.4.2 Loss of Id4 leads to an increase in colonies as measured by CFC assay

To further understand the changes in mammary epithelial cell differentiation we used the colony forming capacity (CFC) assay to measure the frequency of luminal progenitor cells (Stingl, 2009). There are a number of variants on this assay, however, the one that we used in our experiments involved plating disassociated mammary epithelial cells at low density *in vitro* with a feeder layer of irradiated fibroblasts. The irradiated fibroblasts provide growth and other factors that help the epithelial cells grow, however, due to the non-lethal irradiation these cells have undergone cellular senescence and do not proliferate. This technique has previously been shown to primarily be a read-out of luminal progenitor cell activity (Sleeman et al., 2007). In duplicate, 2500 freshly isolated mammary epithelial cells from either Id4 null or heterozygous mice were plated onto 6 cm dishes along with 800 000 irradiated 3T3 fibroblasts and cultured for 6 days. Colonies were then stained using a modified Giemsa stain (DiffQuik) and counted using a dissecting microscope. In three independent experiments a reproducible significant increase in colonies was observed in the mMECs derived from the Id4 null mice (Figure 4-15). This corroborates the flow cytometry data showing that loss of Id4 leads to an increase in luminal progenitors, and that Id4 deficiency leads to precocious luminal differentiation.

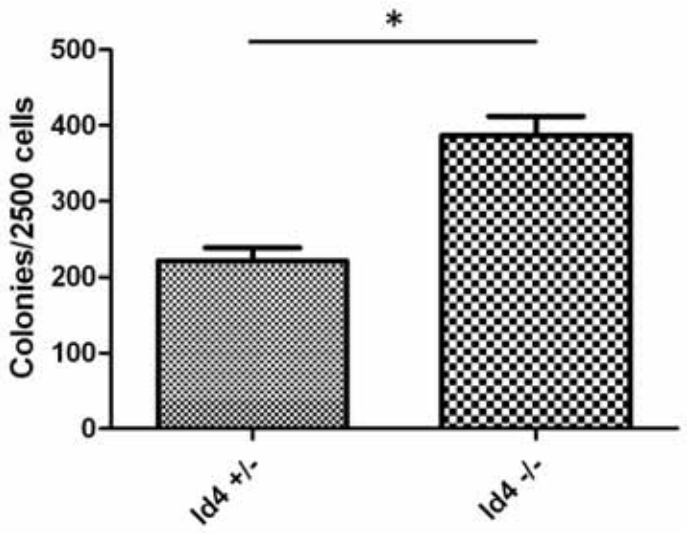


Figure 4-15. Loss of Id4 leads to an increase in colony formation.
Primary mammary epithelial cells (2500) from Id4 heterozygous ($Id4^{+/-}$) and Id4 null ($Id4^{-/-}$) mice were plated with a feeder layer of irradiated fibroblasts into 6cm dishes and allowed to grow for 7 days in culture. Mammary epithelial cell colonies were counted and the average plotted. This is a representative graph from 1 of 3 independent experiments. *$p<0.05$ unpaired t-test.

4.2.4.3 Loss of Id4 leads to a decrease in stem cell activity as measured by mammary repopulating unit assay

The defect in ductal elongation during mammary gland development in the Id4 null mice could also be due to reduced numbers of mammary epithelial stem cells. Furthermore the precocious luminal differentiation seen in the Id4 null mice could actually lead to a depletion of mammary stem cells due to increased differentiation and reduced self-renewal. There is currently no cell surface marker that can effectively differentiate mammary stem cells from myoepithelial cells. Additionally since the stem cells make up a very small percentage of the stem/myoepithelial fraction it is hard to determine changes in stem cell number using flow cytometry. Instead, limiting dilution transplantation is the gold standard technique to quantify mammary repopulating units (MRU). MRU activity is thought to represent true stem cell activity as it demonstrates the ability to recapitulate an entire mammary epithelial ductal tree with all the different lineages of mammary epithelial cells. Single cell suspensions were generated from the mammary glands of 10-12 week old Id4 heterozygous and null mice. I then performed limiting dilution experiments with 10 000, 5 000, 1000, 500, and 100 cells transplanted into the cleared fatpad of syngeneic wildtype FVB/N mice. Each recipient mouse was transplanted with Id4 null cells on one side and Id4 heterozygous cells into the contralateral gland. Eight weeks after transplant the mammary glands were harvested and analysed for any outgrowths by wholemount histology. The analysis revealed that the mMECs from the Id4 null mice had a reduced number of outgrowths and therefore a reduction in MRUs, however after two independent experiments this did not reach statistical significance ($p=0.09$) (Figure 4-16A). Further replicates are currently being performed and this may lead to a statistically significant difference. As well as having fewer successful transplants the outgrowths that formed from Id4 null mMECs were 43% smaller in size than the heterozygous controls (Figure 4-16B). This provides evidence that Id4 is important in maintaining mammary stem cells in addition to inhibiting precocious luminal differentiation and promoting ductal elongation.

(A)

No. of cells injected/fat pad	No. of positive outgrowths/ Total number of fat pads injected	
	Id4 +/-	Id4 -/-
10 000	6/6	5/6
5 000	6/9	5/7
1 000	7/11	2/10
500	0/4	0/4
100	1/13	0/13
MRU Frequency	**1/2529**	**1/4980**
(95% CI)	(1/1452-1/4406)	(1/2702-1/9181)

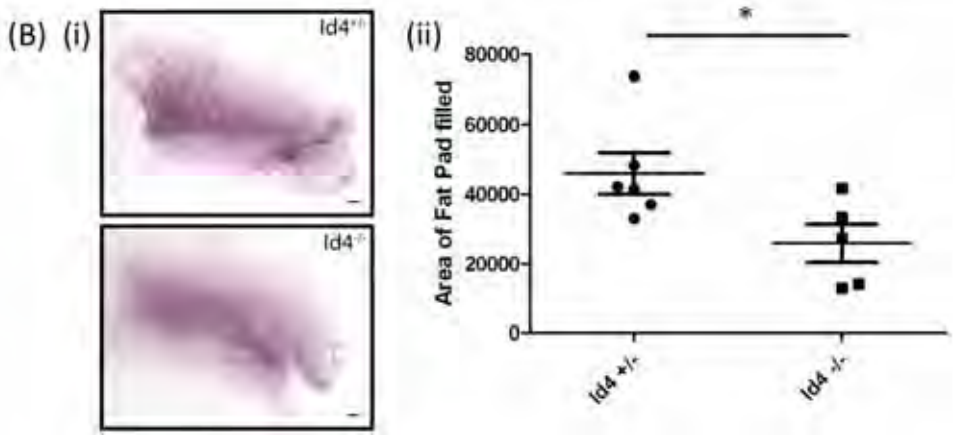


Figure 4-16. Loss of Id4 leads to a reduction in mammary repopulating units.
(A) Limiting dilution transplant analysis of mammary epithelial cells from Id4 null and heterozygous mice. p=0.09. (B) Wholemount histology on representative mammary glands from (A). (i) Id4 null ($Id4^{-/-}$) mammary epithelial cells form smaller outgrowths than heterozygous controls ($Id4^{+/-}$) following transplantation into wild type mice (Scale bar=1mm), and (ii) analysis of area of fat pad filled (arbitrary units) from the 10 000 cell transplants with Id4 heterozygous and Id4 knockout cells. Bars indicatethe average +/- SEM. *p<0.05 unpaired t-test.

4.3 Discussion:

Mammary gland development during puberty is a highly regulated process that is driven primarily by the hormone oestrogen. ER positive luminal cells receive the signalling from the hormone and relay it via paracrine signalling factors to the surrounding epithelium and stroma. Various pathways are activated in the different cell types in response to hormone signalling that result in the co-ordinated proliferation and differentiation of the mammary epithelium to form the ductal tree of the mature mammary gland. My results suggest that Id4 is a crucial regulator of oestrogen signalling in the stem/myoepithelial cell compartment. Id4 is clearly regulated through the oestrus cycle and has previously been shown to be up regulated in response to progesterone (Fernandez-Valdivia et al., 2008). We now present data showing that in the absence of Id4 normal pubertal mammary gland development is delayed. Our results go on to show that in the absence of Id4 there is a reduction of mammary epithelial stem cells and that this is likely due to precocious differentiation along the luminal lineage. The relationship between Id4 and mammary epithelial differentiation will be further

investigated in Chapter 5. Our results also suggest that Id4 is not necessary for myoepithelial cell differentiation since normal levels of myoepithelial cell markers are present in Id4 null mammary glands and the ducts have a normal bilayered structure. Previously it has been shown that the regulation of gene expression in the mammary epithelium by oestrogen is tightly regulated in a time dependant manner (Silberstein et al., 2006). This suggests that the role of Id4 in regulating normal mammary biology is partly in response to its regulation by the oestrus cycle, leading to a cyclical block and release of luminal differentiation (Figure 4-17).

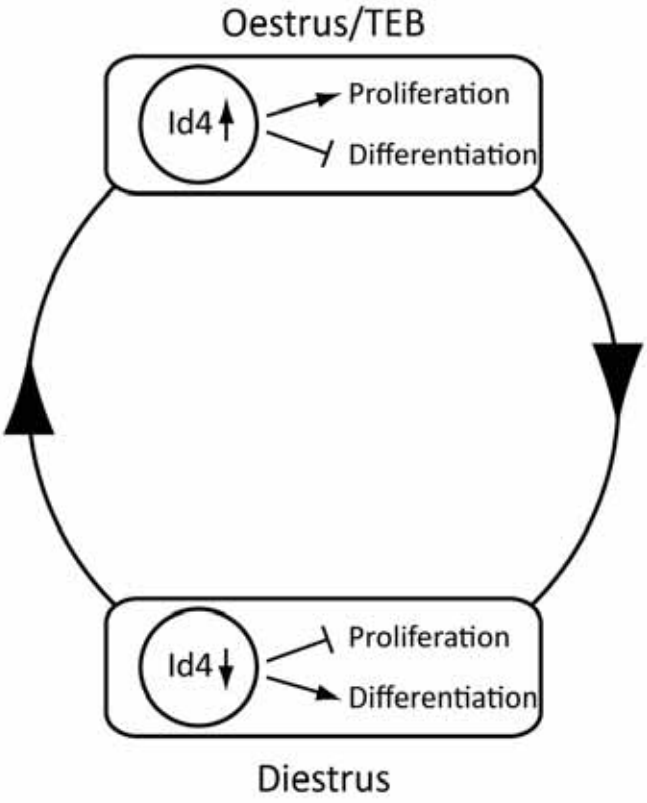


Figure 4-17. Model of regulation of proliferation and differentiation signals by Id4 in the myoepithelial cells during mammary development in the terminal end bud (TEB) and during the oestrus cycle.

A study published in 2011 also examined the role of Id4 in mammary gland development (Dong et al., 2011). This study confirmed my results that Id4 was primarily expressed by the myoepithelial cells of the mammary gland and that in its absence there was a significant defect in pubertal mammary gland development. They went on to show that loss of Id4 led to an increase in apoptosis and a decrease in proliferation in the epithelial cells of the terminal end bud (TEB) and that this was a consequence of increased p38MAPK activation (Dong et al., 2011). Although Dong et al demonstrate that Id4 is important for normal mammary gland development they fail to fully explain how the loss of Id4 causes this defect and how the loss of

Id4 as a transcriptional regulator is linked to increased p38MAPK phosphorylation. Although they do show that blocking p38MAPK activation can rescue the apoptosis/proliferation phenotype *in vivo* they do not show that it rescues the defect in mammary gland development. Furthermore they demonstrated that the activation of p38MAPK occurs in both the luminal (Id4 negative) and the myoepithelial (Id4 positive) cells of the mammary gland. In their discussion Dong et al categorically state that the function of Id4 during mammary gland development is not due to it regulating differentiation pathways, although they offer no evidence to support this claim. The activation of p38MAPK does however fit with our model: that Id4 regulates epithelial differentiation pathways by blocking luminal differentiation. In other tissues such as the lung p38MAPK activation promotes differentiation and in its absence there is a block in differentiation and a build up of progenitor cells (Hui et al., 2007; Ventura et al., 2007b). Furthermore p38MAPK is known to play a role in lumen clearing in the developing mammary gland which would also explain the increase in apoptosis seen by Dong et al and the decrease in mature luminal cells seen by us. Therefore we suspect that the increase in p38MAPK phosphorylation is a consequence of abnormal luminal differentiation rather than the cause of the mammary gland phenotype. To further address the relationship between Id4 and p38MAPK it would be interesting to FACS purify the different mammary epithelial cell populations based on CD24, CD29 and CD61 expression and determine in which cell population that loss of Id4 most increases p38MAPK activity.

Id4 has previously been shown to be a progesterone target gene and I have shown that it is regulated through the oestrus cycle, however, since Id4 is not expressed in oestrogen/progesterone receptor positive (luminal) cells this must be via a paracrine mechanism. Our studies do not address what paracrine signalling molecule regulates Id4 expression. RANKL, Wnt ligands, and amphiregulin are candidate paracrine factors that are known to be secreted by luminal sensor cells upon oestrogen/progesterone signalling. The signalling pathways that could lead to Id4 up regulation are summarised in Figure 4-18. Loss of RANKL has limited impact on pubertal mammary development and Id4 is not directly regulated by RANKL signalling (Mukherjee et al., 2010; Stingl), thus the Wnt ligands or amphiregulin appear to be the more likely candidates for paracrine signalling from ER/PR positive cells to up regulate Id4 expression in the myoepithelial cells. Amphiregulin was shown to primarily signal through stromal cells rather than directly to other epithelial cells which adds another layer of complexity to this process (Wiesen et al., 1999). The mammary gland defect in the Id4 null mouse to some extent phenocopies the amphiregulin null mouse suggesting amphiregulin could be involved in Id4 activation. Although that would have to be via a downstream

signalling molecule secreted from stromal cells in response to amphiregulin, an interesting candidate for this would be BMP2 or BMP4 which are known activators of Id4 expression (Cheng et al., 2007). The Wnt pathway could also be involved in regulating Id4 as mammary tumours from the MMTV-Wnt1 transgenic mouse have elevated levels of Id4 (Fernandez-Valdivia et al., 2008). Experiments targeting the Wnt and amphiregulin pathways either with inhibitors or gene knockdown strategies could be performed to better understand the regulation of Id4 expression by the oestrus cycle.

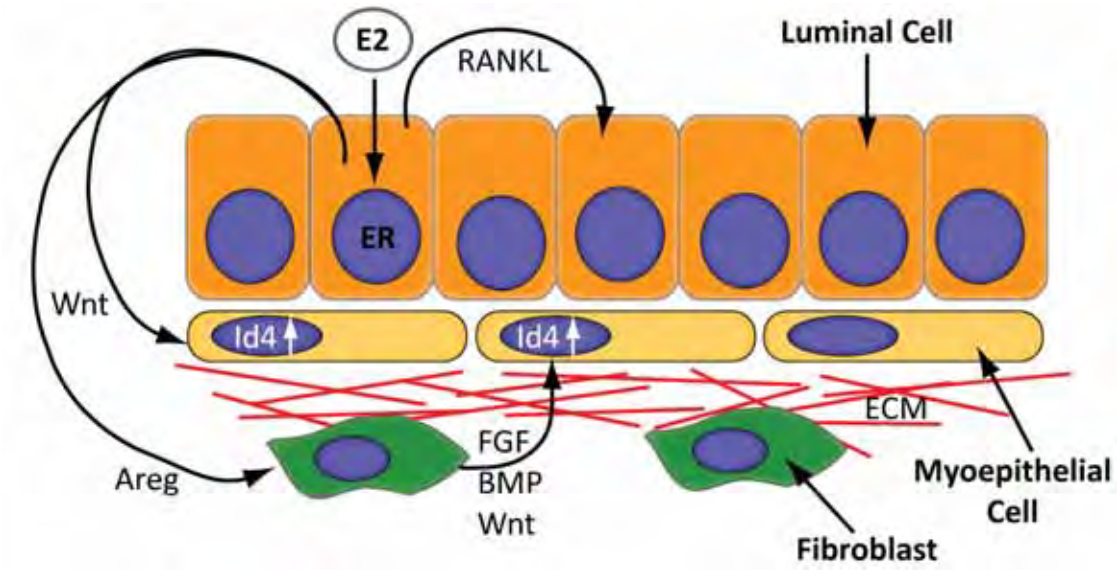


Figure 4-18. Model of paracrine regulation of Id4 expression in the mammary gland.
Oestrogen (E2) signalling to the oestrogen receptor (ER) positive luminal cells stimulates these cells to secrete RANKL, Wnt proteins, and amphiregulin (Areg). Either Wnt proteins from the ER positive cell or FGF, BMP or Wnt proteins from stromal fibroblasts then signal to the myoepithelial cell and lead to an up regulation of Id4.

A number of transcription factors have been identified that regulate the differentiation of stem cells down the luminal/alveolar lineage such as NOTCH, GATA3, BRCA1, and ELF-5, however, less is known about the transcription factors that determine the stem and myoepithelial cell lineages. One transcription factor that is required for the myoepithelium is p63 (Li et al., 2008). We saw that Id4 co-localised with p63 in the myoepithelial cells but it is unknown whether there is any direct link between the expression of the two. Recently it was shown that the epithelial to mesenchymal transition (EMT) factor Slug (SNAI2) plays a crucial role in maintaining stemness in mammary epithelial cells and Id4 was shown to be up regulated upon the ectopic expression of Slug, however, it was not shown whether Id4 was a direct or functional target of Slug (Guo et al., 2012). Understanding how Id4 regulates these other transcription factors in the mammary epithelial hierarchy will give us a better idea of

how it is regulating the differentiation pathways of the mammary gland and will be investigated in more detail in the following chapter: “Understanding the function of Id4 using the mouse normal mammary epithelial cell line Comma-Dβ”.

Chapter 5. Understanding the function of Id4 using the mouse normal mammary epithelial cell line Comma-Dβ

5.1 Introduction:

In my previous chapter I described data demonstrating that loss of Id4 in the mouse mammary epithelium leads to a significant defect in ductal elongation during mammary gland development. To better understand this phenomenon, we wanted to identify the mechanism of action of Id4 in mammary epithelial cells and to discover the transcriptional targets of Id4. The transcriptional targets of Id4 in epithelia are currently unknown. As Id4 does not directly bind to DNA to regulate transcription, its transcriptional targets are regulated through the dominant negative effect of Id4 binding to bHLH and ETS transcription factors. Thus, depending on the specific transcription factors expressed in a particular cell type the response to modulation of Id4 expression will be different. Some transcriptional targets for Id4 have been identified in the brain. One suggested target for Id4 in glioma stem cells is miR-9, and another in glioblastoma is the Notch pathway (Jeon et al., 2008; Jeon et al., 2011). In this chapter, I will examine the phenotype that results from Id4 overexpression and knockdown on the Comma-Dβ cells and then use these systems to discover the transcriptional targets of Id4 in mammary epithelial cells.

5.1.1 The Comma-Dβ cell line model

The mammary epithelial Comma-D cell line was derived from the mammary epithelium of a mid pregnant BALB/C mouse (Danielson et al., 1984). Comma-D cells can differentiate and produce milk proteins *in vitro* and, when transplanted back into the cleared fat pad of BALB/C mice, they can form normal mammary ductal outgrowths (Danielson et al., 1984). The Comma-Dβ cells are a sub-line of the Comma-D cell line generated by infection with a retrovirus containing a fusion of β-galactosidase and neomycin-resistance genes, and then resistant clones were selected for with G418. The advantage of the Comma-Dβ cell line over the parental cell line is that they have the ability to be passaged longer in culture while still retaining the ability to form normal ductal outgrowths when transplanted into syngeneic mice (Deugnier et al., 2006). Another sub-line of the Comma-D cells is the HC-11 cell line. The HC-11 cell line, however, appears to be more restricted to the luminal epithelial cell lineage. Although Comma-Dβ cells are considered normal, they do contain p53 mutations and can be passaged in culture without undergoing senescence. Both copies of p53 are mutated in the Comma-D cell line with one copy containing a substitution of tryptophan for cysteine at

position 138, and the other copy having a 7 amino acid deletion from amino acid 123 to 129 (Jerry et al., 1994). Despite initial reports to the contrary, Comma-D cells do have a tendency to form tumours spontaneously after transplantation into mice (Jerry et al., 1994).

The Comma-D cells have been used by several research groups to gain a better understanding of mammary epithelial biology. The stem cell characteristics and expression of various luminal and myoepithelial cell markers was investigated in a number of different *in vitro* and *in vivo* systems by one group, demonstrating a number of similarities to primary mammary epithelial cells (Deugnier et al., 2006). Two groups have investigated changes in microRNA expression as the cells differentiated in culture, revealing many candidates that were also known to play a role in breast cancer, as well as novel candidates (Greene et al., 2010; Ibarra et al., 2007). The Comma-Dβ cells have also been used to further understand the role of amphiregulin on mammary epithelial cell growth and self renewal (Booth et al., 2009).

We therefore used the Comma-Dβ cells to further understand the role of Id4 in mammary epithelial differentiation pathways and to identify its transcriptional targets. To identify the transcriptional targets of Id4 in mammary epithelial cells we decided to use this cell line model instead of primary cells as it was more amenable to overexpression and knockdown studies. It also offers the advantage of being a simplified system that contains no contaminating stromal cells. Transcriptional targets identified in the Comma-Dβ model were then further characterised in the mammary glands from mice.

5.1.2 Hypothesis:

Modulating Id4 levels in the Comma-Dβ cell line will modulate differentiation pathways and the transcription factors that regulate these processes.

5.1.3 Aims:

The specific aims of this chapter are to determine:

1. The effect of Id4 overexpression on the Comma-Dβ cell line.
2. The effect of Id4 knockdown on the Comma-Dβ cell line.
3. The transcriptional targets of Id4 by transcript profiling the overexpressing and knockdown cell lines.

5.2 Results:

5.2.1 Comma-Dβ cells endogenously express Id4

To first validate that the Comma-Dβ cell line was an appropriate model for studying Id4 in mammary epithelial biology we needed to confirm that this cell line expressed Id4 in a similar manner to that seen in primary mammary epithelial cells. Comma-Dβ cells were analysed for Id4 protein expression by IHC and by western blot. Id4 protein could easily be detected by both techniques (Figure 5-1A). To further determine if the Comma-Dβ cell line was a relevant model for studying the role of Id4 in the mammary gland, we transplanted 50 000 cells into the cleared fatpad of 3 week old BALB/C mice. After allowing the outgrowths to develop for 8 weeks, the transplants were harvested and analysed both by wholemount histology and Id4 IHC on FFPE sections. Wholemount histology showed that the Comma-Dβ cells could indeed form normal outgrowths when transplanted into the fat pad of BALB/C mice (Figure 5-1B). Furthermore normal ductal structures could be observed in H&E stained sections, and strong nuclear expression for Id4 could be observed in the myoepithelial but not in the luminal cells in these sections (Figure 5-1C). These results demonstrate that the Comma-Dβ cell line expresses Id4, and furthermore when these cells are transplanted into the mammary fat pad that they can differentiate into both Id4 positive myoepithelial cells and Id4 negative luminal epithelial cells.

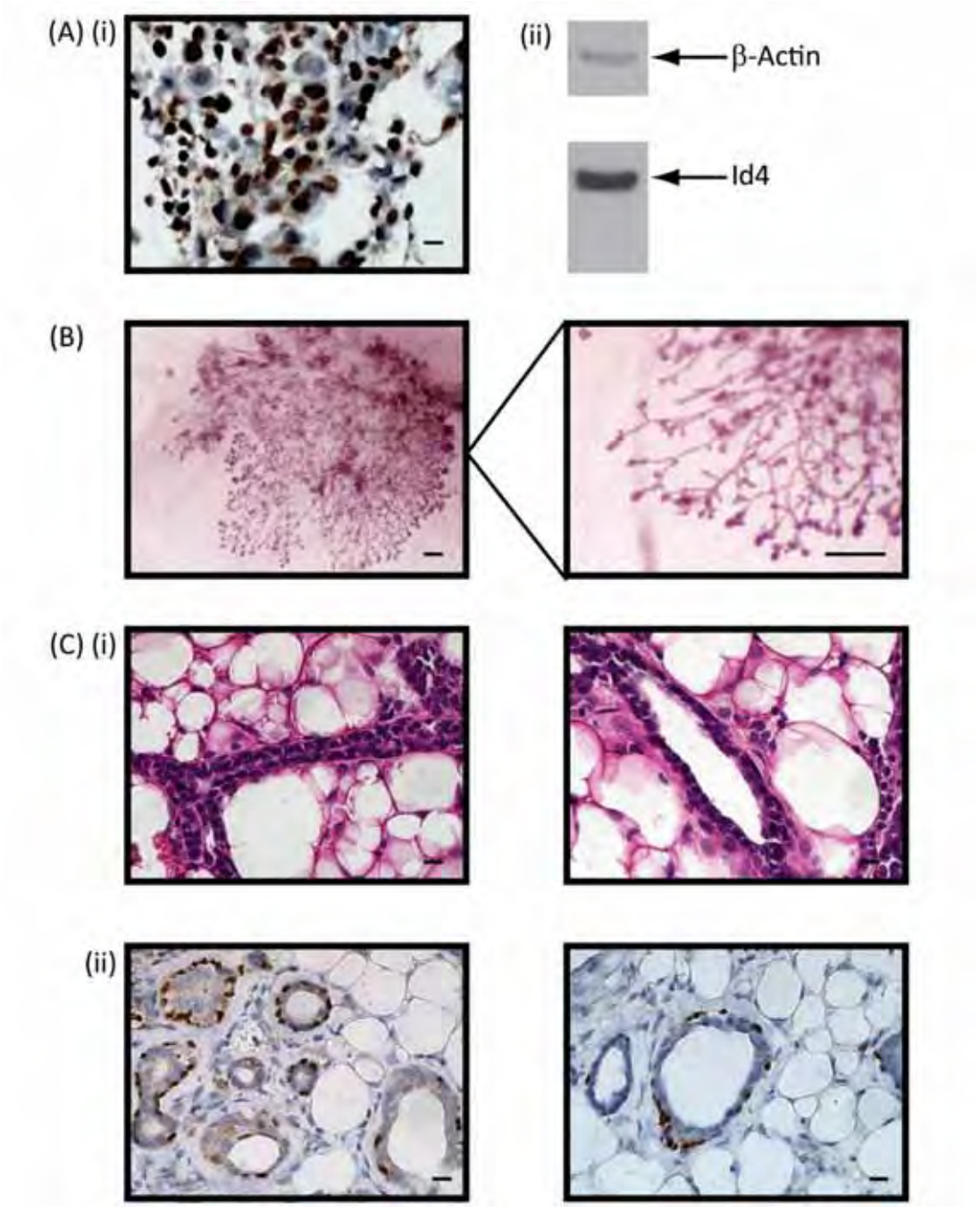


Figure 5-1 Endogenous levels of Id4 expression in Comma-Dβ cells.
(A) (i) FFPE Comma-Dβ cells stained for Id4 expression (brown) and counterstained with haematoxylin. Scale bar = 10μm. (ii) Western blot analysis of protein lysate from Comma-Dβ cells for Id4 expression. (B) Wholemount histology of Comma-Dβ outgrowths from cells transplanted into BALB/c mice. Scale bars = 500μm. (C) (i) H&E analysis of ducts formed by Comma-Dβ cells shows a normal bilayered structure. (ii) Id4 IHC analysis of ducts from Comma-Dβ outgrowths shows nuclear myoepithelial staining (brown). Scale bars = 10μm.

5.2.2 Overexpression of Id4 in a normal mouse mammary cell line

As shown in section 5.2.1, the Comma-Dβ cells already express Id4 protein but can differentiate into Id4 negative cells. We therefore wanted to determine the effect of enforced

overexpression of Id4 on the phenotype of these cells. Comma-Dβ cells stably overexpressing Id4 were generated by retroviral transduction with the pMSCV-Id4-DSRed construct or a negative control pMSCV-DSRed construct. Cells that had taken up the construct were then cell sorted based on high DS Red fluorescence. A flow cytometric purity check showed that high DS Red fluorescence was observed in 92% of the sorted cells (Figure 5-2A). Id4 overexpression in these sorted cells was then confirmed by western blot and these cells were used for the subsequent experiments (Figure 5-2B).

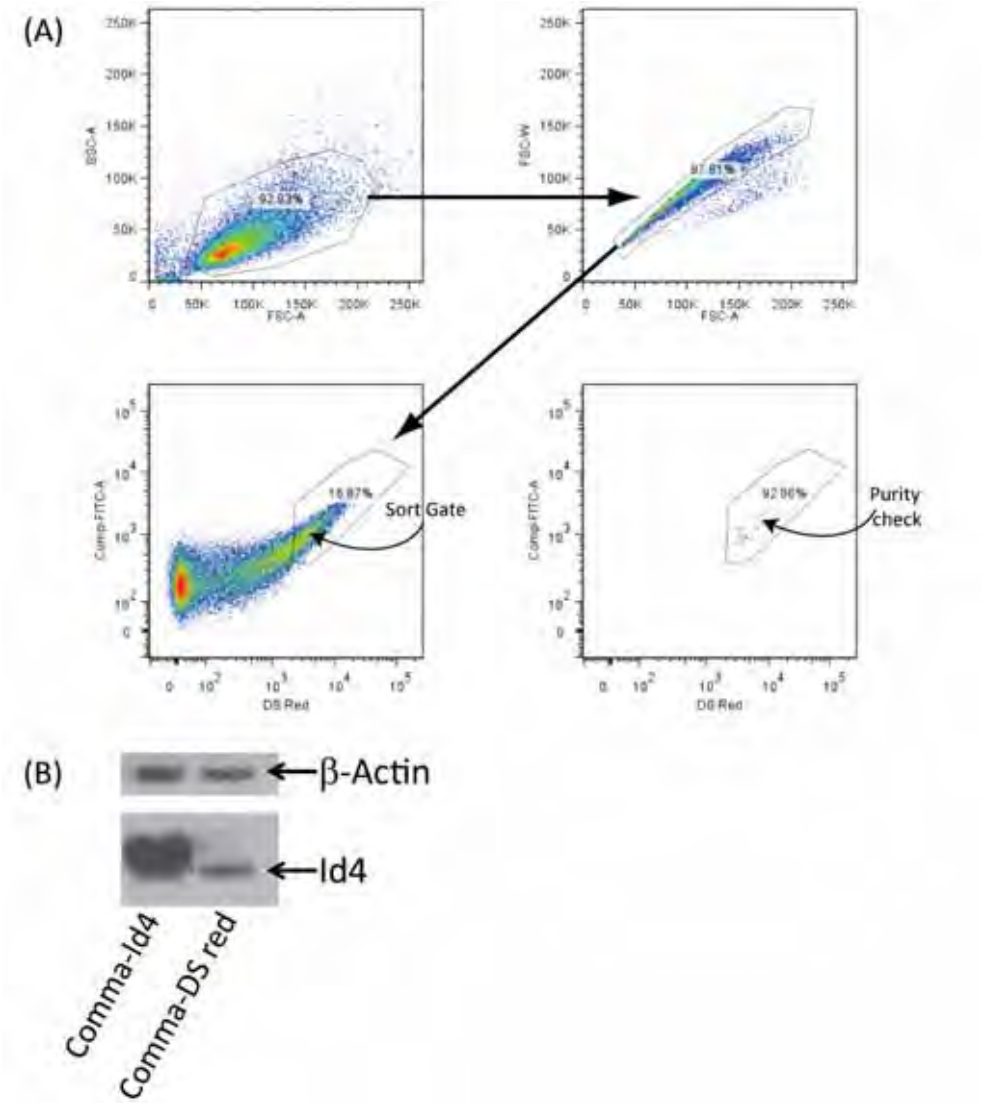


Figure 5-2. Enforced overexpression of Id4 in Comma-Dβ cells.
(A) (i) Gates used for sorting DS red positive cells: Small particles were excluded based on forward scatter area (FSC-A) and side scatter area (SSC-A), then single cells were selected based on FSC-A verses forward scatter width (FSC-W). The sort gate was selected based on high DS Red expression against the FITC-A channel, and the purity check following cell sorting showed 92% of cells within the sort gate. (B) Western blot analysis for Id4 protein levels in lysates from the Comma-Id4 cells compared to Comma-DSred control cells.

5.2.2.1 Overexpression of Id4 has minimal effect on Comma-Dβ in vitro growth rate

To first characterise the Comma-Id4 cells we tested whether Id4 overexpression had any effect on proliferation. Proliferation was compared between the Id4 overexpressing and DS Red control cells. Cell proliferation was measured over an 8 day time course by two independent methods firstly by cell counts and secondly using the MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay. Both methods showed a modest reduction in proliferation initially after the seeding of the cells but showed no difference at later time points (Figure 5-3).

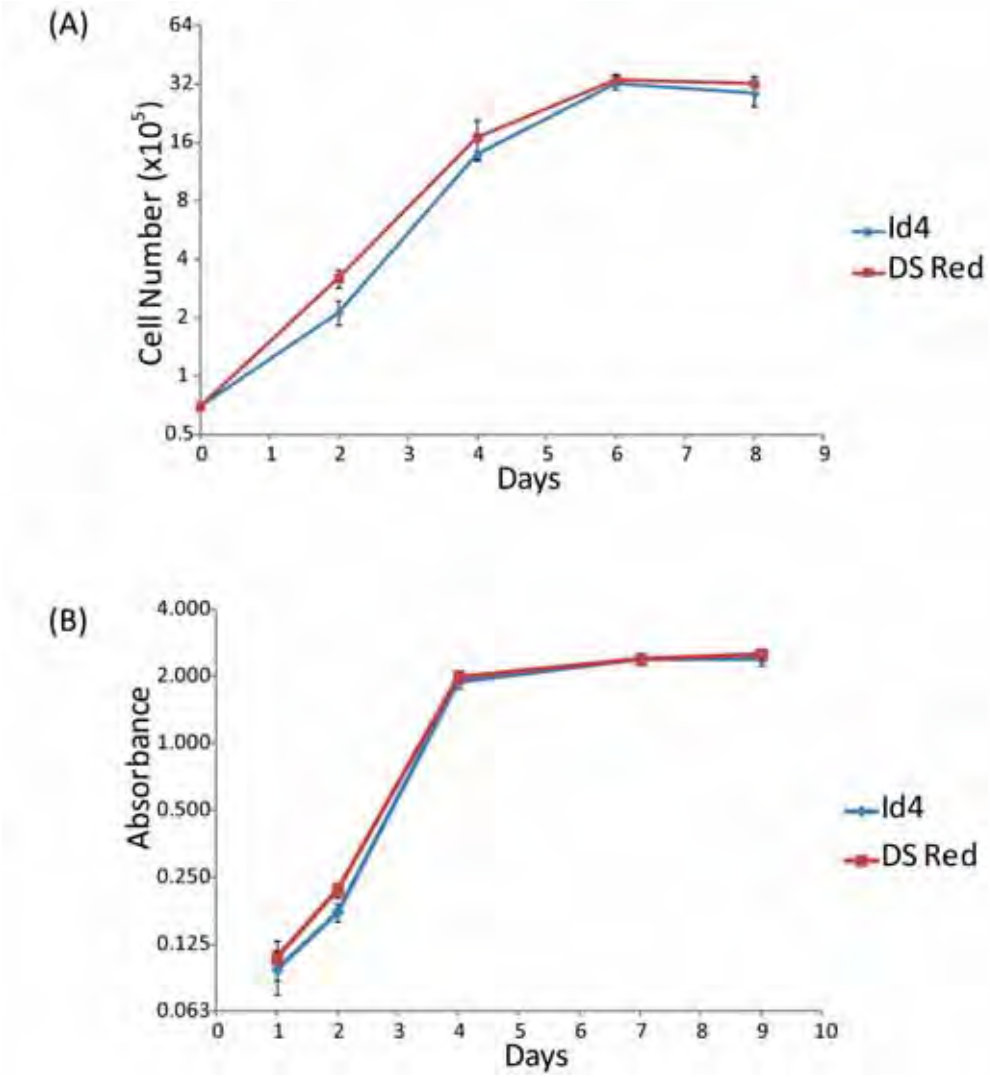


Figure 5-3. Enforced overexpression of Id4 has minimal effect on Comma-Dβ growth *in vitro*.
(A) Enumeration of Comma-Id4 cells compared to Comma-DSred cells over 8 days. Cell concentration per ml was measured by haemocytometer and plotted on a log2 scale. Graph depict the average of 2 replicates ±SEM. (B) Proliferation of Comma-Id4 cells compared to Comma-DSred control cells over 7 days as measured by MTS assay. Graph depict the average of 6 replicates ±SEM.

5.2.2.2 Overexpression of Id4 inhibits luminal differentiation and milk production

The studies described in Chapter 4 demonstrated that Id4 plays a role in blocking luminal differentiation in the mouse mammary gland. We therefore hypothesised that overexpression of Id4 would inhibit luminal differentiation and milk protein production in Comma-Dβ cells. Using the well established *in vitro* differentiation assay (Danielson et al., 1984; Lee et al., 2011a), Comma-Id4 and DS red control cells were seeded at a density that allowed them to reach near confluence within two days, they were then subjected to EGF withdrawal with or without the addition of the lactogenic hormone prolactin to promote luminal/alveolar differentiation. Protein and RNA lysates were collected just prior to addition of prolactin (Day 0) as well as 2 and 4 days later. During the differentiation process the Comma-Id4 cells

showed dramatically different morphology to controls. The Id4 overexpressing cells retained a monolayer cell morphology throughout the differentiation process, whereas the DS red control cells piled up and formed three dimensional structures (Figure 5-4A). Milk protein production was used as a readout of functional luminal differentiation, overexpression of Id4 led to a dramatic reduction in milk protein production as measured by western blot for three specific milk proteins: α-casein, β-casein and whey acidic protein (WAP), as indicated with arrows (Figure 5-4B). This was then confirmed at the mRNA level, where there was also a dramatic reduction of expression of *WAP* mRNA as measured by Taqman and also a more subtle reduction of *β-casein* mRNA (Figure 5-4C). Id4 overexpression also suppressed expression of *Elf-5*, a transcription factor that drives alveolar differentiation and *cytokeratin 8* (*CK8*), a marker of luminal mammary epithelial cells (Figure 5-4D).

In addition to examining how Id4 modulated *in vitro* differentiation at the protein and mRNA level we also set up chamber slides to determine CK8 and CK14 expression levels by immunofluorescence at the individual cell level. Day 4 prolactin treated Id4 overexpressing and control Comma-Dβ cells were imaged using confocal microscopy and eight high powered images were examined for CK8, CK14, and CK8 + CK14 double positive cells. Overexpression of Id4 led to a significant 4-fold reduction in the percentage of CK8 positive cells ($p=0.0022$, unpaired t test) and there was a trend to increased numbers of CK14 positive and CK8 + CK14 double positive cells (Figure 5-4E). These experiments corroborate the *in vivo* results described in Chapter 4, demonstrating that Id4 plays a role in inhibiting luminal cell differentiation.

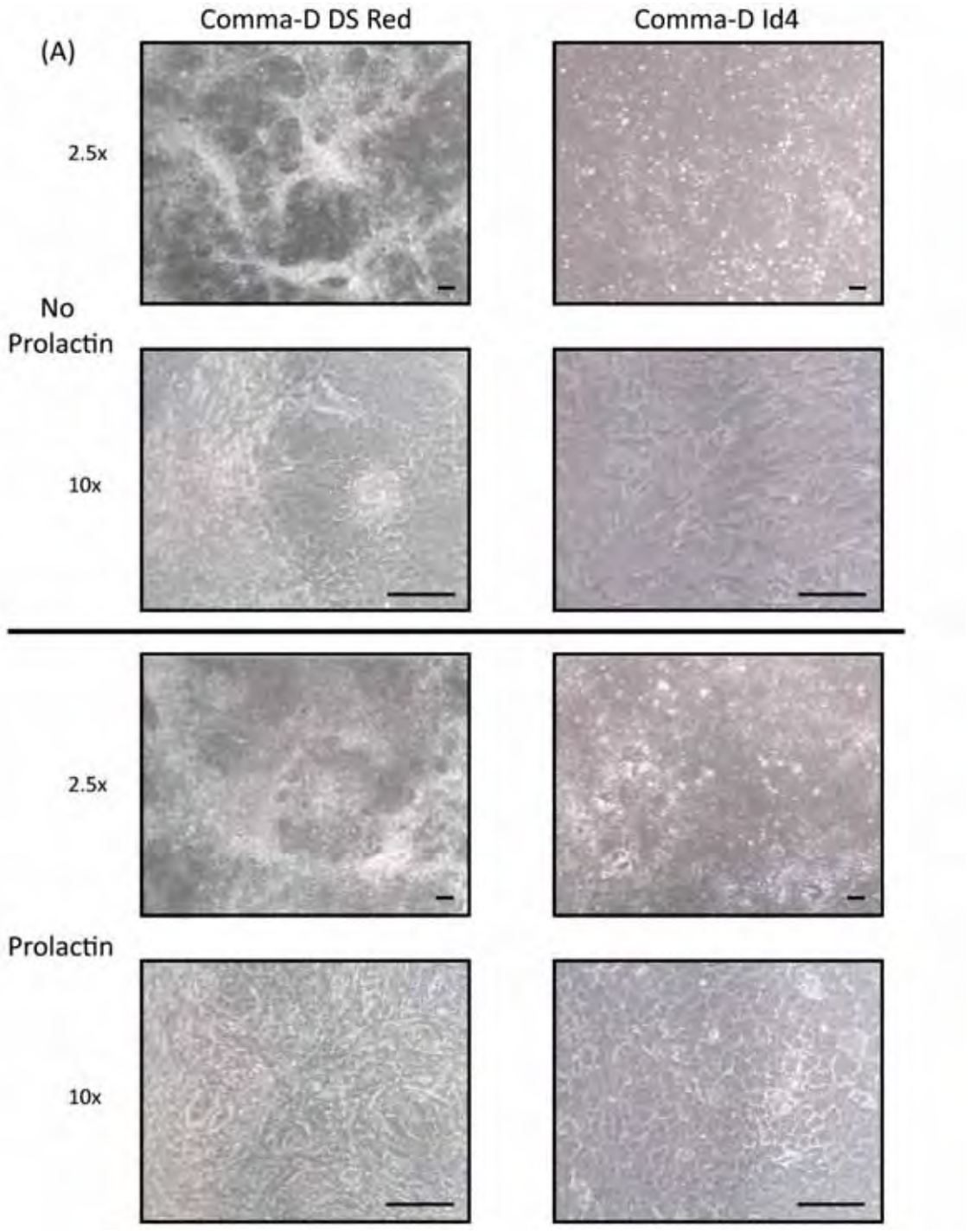


Figure 5-4. Overexpression of Id4 blocks luminal/alveolar differentiation in the Comma-Dβ cells.
Comma-Id4 and DS red control cells were seeded at a density that allowed them to reach near confluence within two days, they were then subjected to EGF withdrawal with or without the addition of prolactin (0.5μg/ml). (A). Changes in Comma-Dβ cell morphology during *in vitro* differentiation due to Id4 overexpression. Light microscopy images of cells grown in culture taken on Day 2 of the assay at low and high magnification. Scale bar=125μm. (Figure continued on following 2 pages)

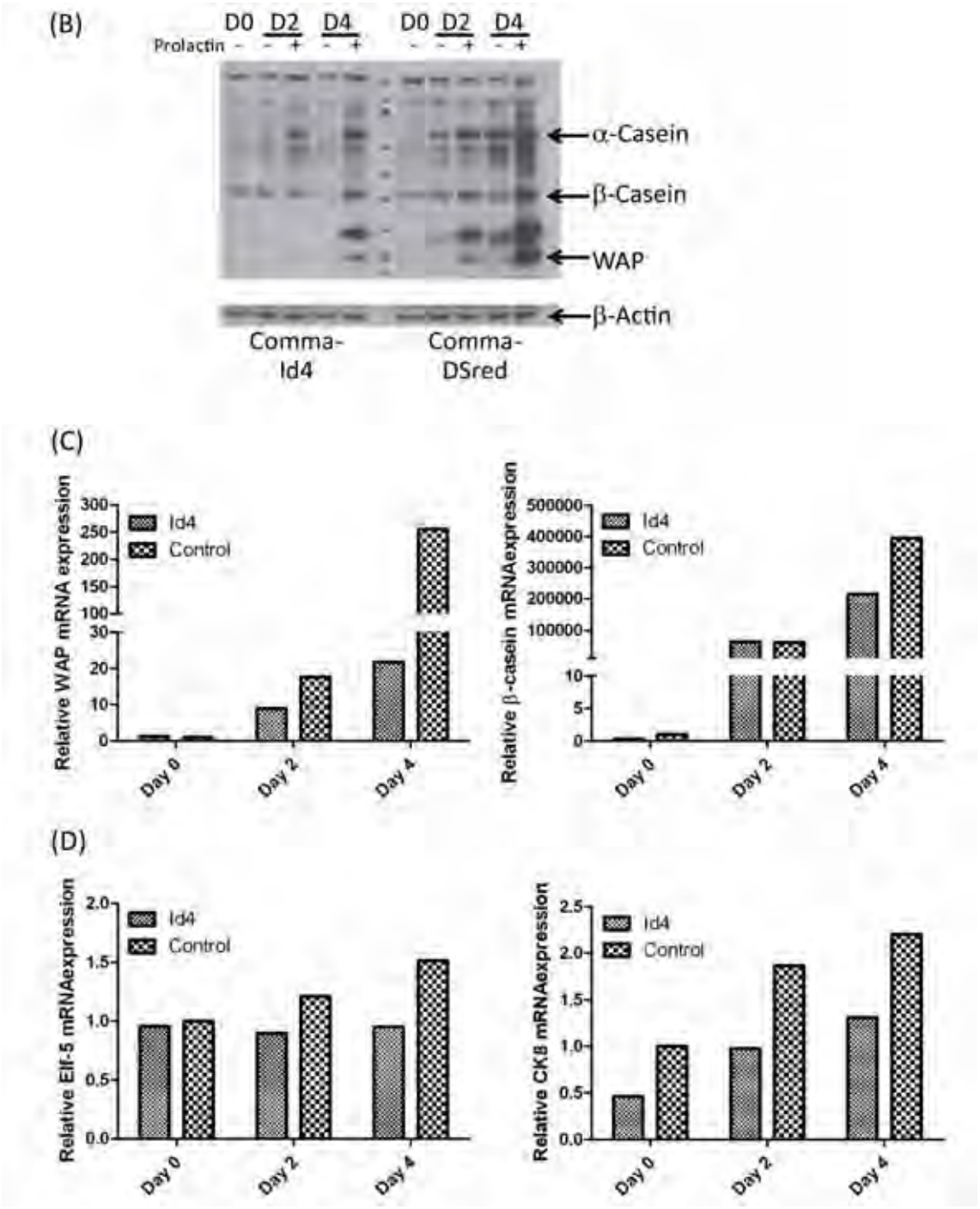


Figure 5-4 Continued. Overexpression of Id4 blocks luminal/alveolar differentiation in the Comma-Dβ cells.

(B) Western blot analysis for milk protein production following *in vitro* differentiation with or without the lactogenic hormone prolactin. The specific milk proteins α-casein, β-casein and whey acidic protein (WAP) are indicated with arrows. (C) Taqman analysis of WAP and β-casein mRNA expression levels and (D) Elf-5 and cytokeratin 8 (CK8) levels during *in vitro* differentiation in the presence of prolactin. (Figure continued on following page)

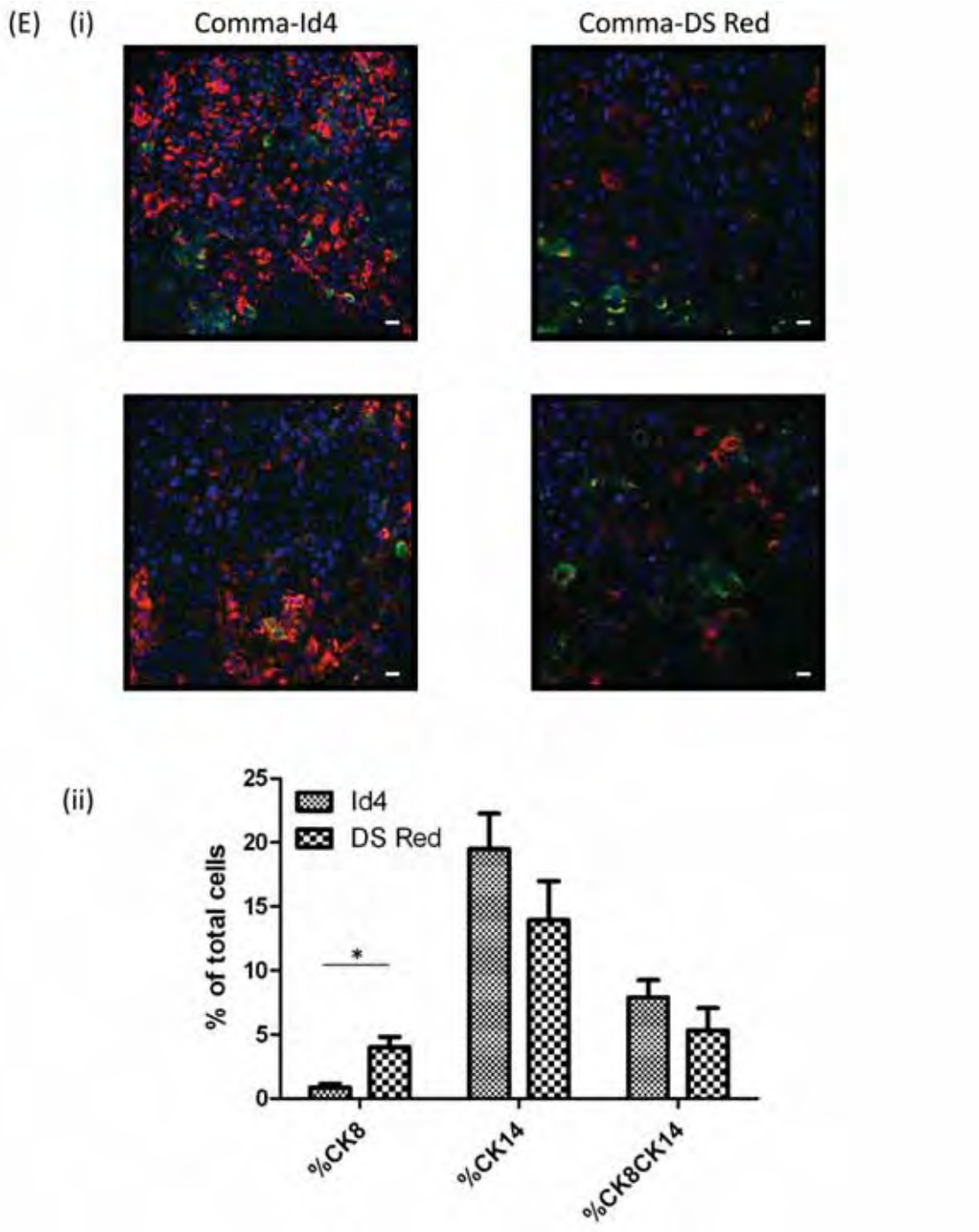


Figure 5-4. Continued. Overexpression of Id4 blocks luminal/alveolar differentiation in the Comma-Dβ cells.

(E) (i) Immunofluorescence analysis of cytokeratin 8 (CK8) in green and cytokeratin 14 (CK14) in red and nuclei stained with DAPI in blue in Comma-Id4 and DS red control cells after 4 days of prolactin treatment. Scale bar=15μm. (ii) Quantification of CK8 positive, CK14 positive, and CK8 + CK14 double positive cells averaged from 8 high powered fields of view ±SEM, *p=0.0022 unpaired t test.

5.2.3 Id4 overexpression promotes tumour formation in the Comma-Dβ cells

To determine the effect of enforced Id4 overexpression on the ability of the Comma-Dβ cells to form ductal outgrowths *in vivo* we transplanted 50 000 Comma-Id4 cells into the cleared 4th mammary fat pad of three week old Balb/C mice, control Comma-DSred cells were transplanted into the contralateral mammary gland. The transplants were allowed to engraft for 8 weeks with the exception of two Comma-DSred transplants that formed large tumours and were harvested at 6 weeks. The remaining glands were collected for analysis by wholemount histology. All of the Id4 overexpressing Comma-Dβ transplants had formed small tumour-like growths, while the remaining Comma-DSred controls had formed normal mammary outgrowths (Figure 5-5A). The wholemounted transplants were then paraffin embedded and sections were cut and stained with haematoxylin and eosin. While the control transplants contained normal ductal/alveolar epithelial structures the Id4 overexpressing constructs contained a large amount of extracellular matrix (ECM) deposition, immune infiltrate and dysplastic epithelial growth (Figure 5-5B). When analysed by a breast cancer pathologist, Dr Sandra O'Toole, the two tumours formed by Comma-DSred control cells were described as spindle cell metaplastic carcinomas, whereas the Comma-Id4 tumours were described as either an adenosis or a low grade carcinoma with areas of calcification (Figure 5-5C). When this experiment was repeated, again all of the Id4 overexpressing Comma-Dβ cell transplants generated small tumour-like growths. Thus, over the two experiments, all 15 Comma-Id4 transplants generated small tumours. In contrast 13 of the 15 Comma-DS transplants formed normal ductal outgrowths and the remaining 2 formed rapidly growing carcinomas. These results show that overexpression of Id4 prevents normal mammary epithelial differentiation and promotes the generation of low grade carcinomas.

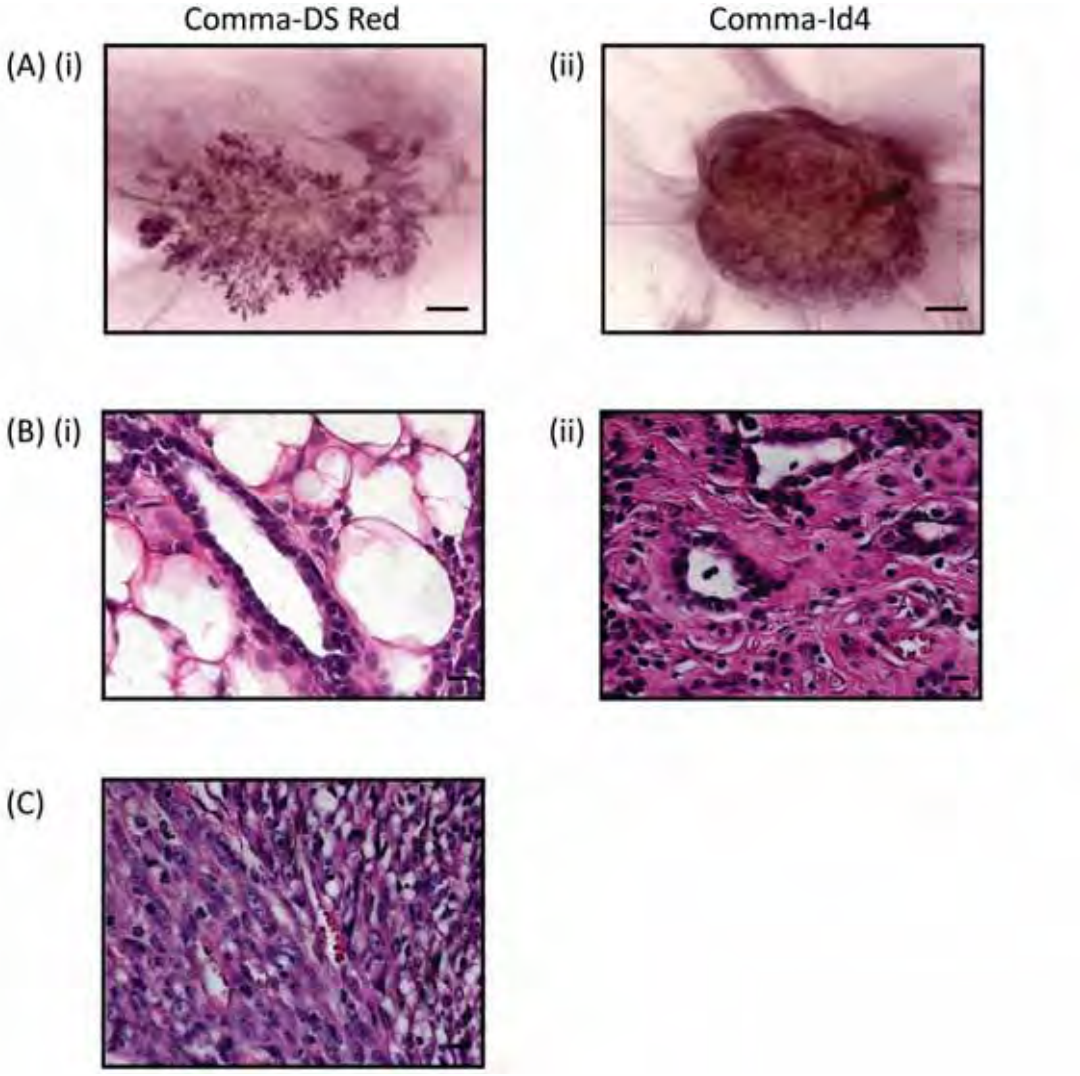


Figure 5-5. Overexpression of Id4 by Comma-Dβ cells *in vivo* leads to increased collagen deposition and tumour formation.

50 000 Comma-Id4 cells were transplanted into the cleared 4th mammary fat pad of three week old Balb/C mice, control Comma-DSred cells were transplanted into the contralateral mammary gland. The transplants were allowed to engraft for 8 weeks with the exception of two Comma-DSred transplants that formed large tumours and were harvested at 6 weeks. (A) Images of wholemounts of Comma-DSred (i) and Comma-Id4 (ii) transplants (Scale bar=1mm). (B) H&E stained sections of FFPE tissue from Comma-DSred (i) and Comma-Id4 (ii) transplants (Scale bar=10μm). (C) H&E stained sections of FFPE tumour from Comma-DSred transplant (Scale bar=10μm).

5.2.4 Knockdown of Id4 in a normal mouse mammary cell line

5.2.4.1 Determining the best knockdown strategy for Id4 in Comma-Dβ cells

To further characterise the mechanism of action of Id4 we also wanted to analyse the effect of Id4 gene knockdown on the Comma-Dβ cells. We first compared three strategies for knocking down Id4, these were: (i) a constitutive lentiviral shRNA system pMISSION (Figure. 5-6A), (ii) an inducible lentiviral shRNA system pSLIK (Figure. 5-6B), and (iii) using transfected siRNAs

(Figure. 5-6C). The pMISSION and the siRNAs were both purchased commercially and contained 5 and 4 different knockdown sequences, respectively. The pSLIK-shId4 construct had been previously generated by another member of our lab (Wee Teo) but had not been validated for knockdown efficiency. The pSLIK system is doxycycline inducible and has the shRNA cloned into a microRNA backbone so that the shRNA is processed in a similar manner to endogenous microRNAs (Shin et al., 2006). The pSLIK construct also constitutively expresses the Venus fluorescent protein, so following infection, the cells were sorted for Venus fluorescence. For the pMISSION system and siRNAs we firstly screened which hairpin (Figure. 5-6A i) or siRNA had the greatest knockdown efficiency (Figure. 5-6C i) and then we analysed Id4 knockdown with the two best constructs/siRNAs over several time points (Figure. 5-6A ii and Figure. 5-6C ii). Knockdown efficiency was measured by western blotting for Id4 and densitometry. It was determined that the lentiviral pMISSION shRNA system generated the greatest knockdown, however, it was observed that cells that continued to grow following puromycin selection had escaped knockdown of Id4 (Figure. 5-6A iv). It was therefore decided that acute knockdown without selection with the pMISSION system led to the greatest knockdown efficiency, and this technique was used for the subsequent experiments. During these experiments it was noted that during the time course the level of Id4 protein initially increased before decreasing by the final time point in the control cells (Figure. 5-6C iii). This suggested that Id4 levels are regulated in a dynamic way depending on the density/differentiation state of these cells in culture.

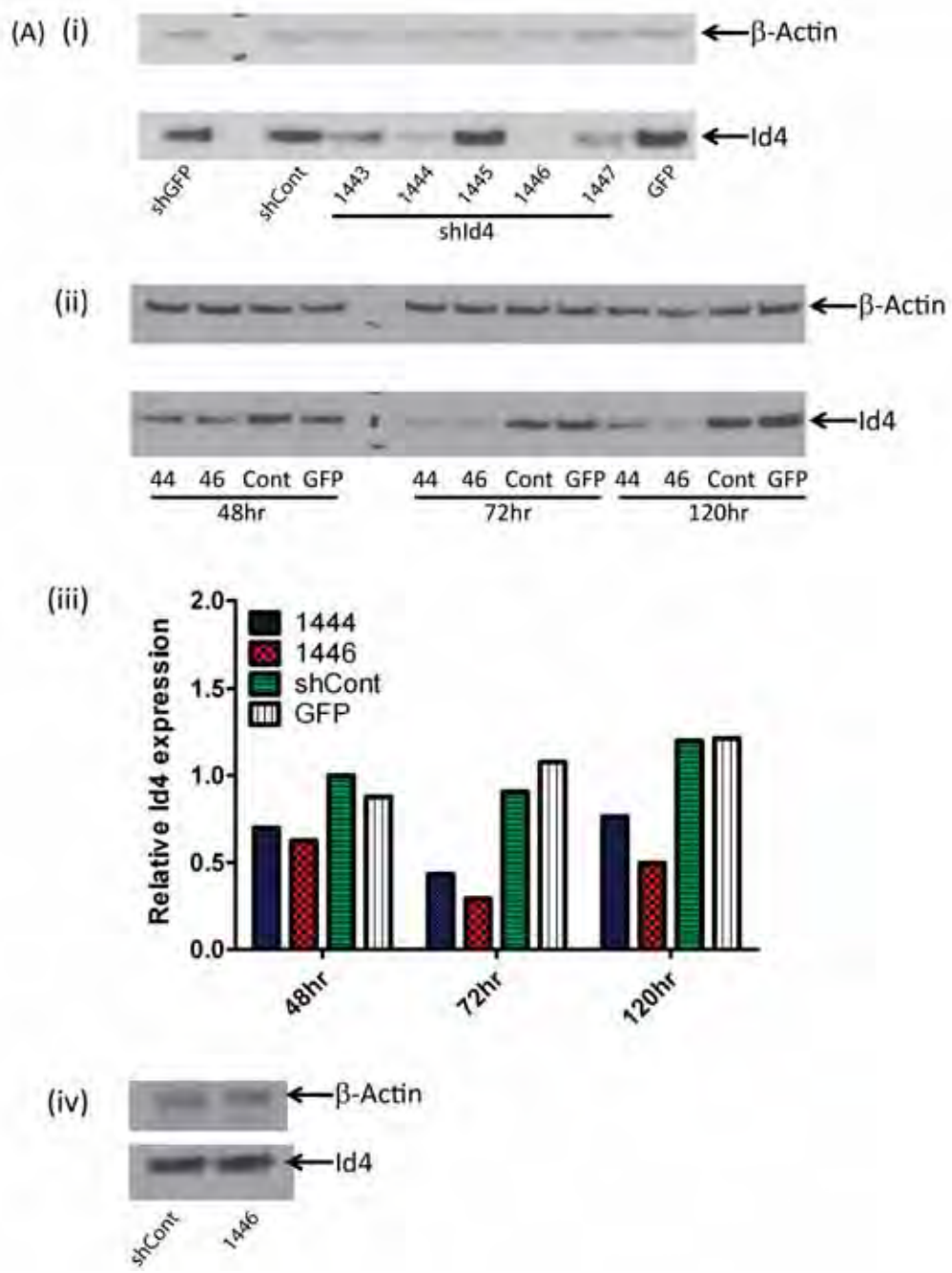


Figure. 5-6. Optimising gene knockdown strategies for Id4 in Comma-Dβ cells.
(A) (i) Analysis of Id4 knockdown using the Sigma pMISSION constitutive shRNA constructs shId4 1443, 1444, 1445, 1446, and 1447 by western blot 72 hours after infection. (ii) Time course of Id4 knockdown using shId4 1444, and 1446 and shCont analysed by Id4 western blot. (iii) Densitometry analysis of knockdown efficiency relative to shCont at the 48hr time point of the different pMISSION constructs over the time course. (iv) Western blot analysis for Id4 of Comma-Dβ shId4 1446 cells selected with puromycin for 7 days. (Figure continued on following 2 pages).

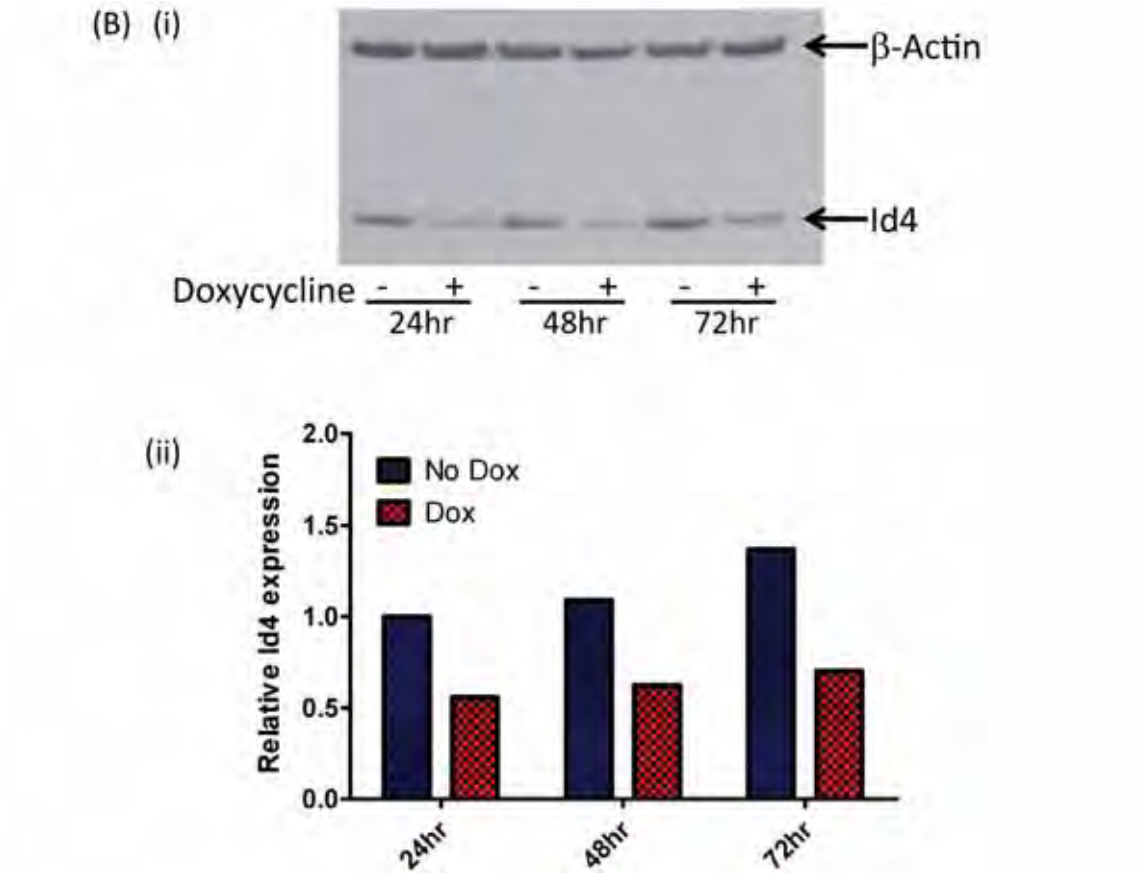


Figure.5-6. Continued. Optimising gene knockdown strategies for Id4 in Comma-Dβ cells.
(B) Analysis of Id4 knockdown using the inducible pSLIK-shId4 construct with or without doxycycline to induce expression of the shId4 construct over a 3 day time course. (i) Western blot for Id4 in protein lysates treated with doxycycline (Dox) or untreated (No Dox) at time points as indicated (ii) Densitometry analysis of knockdown efficiency relative to No Dox at 24 hours through the time course. (Figure continued on following page).

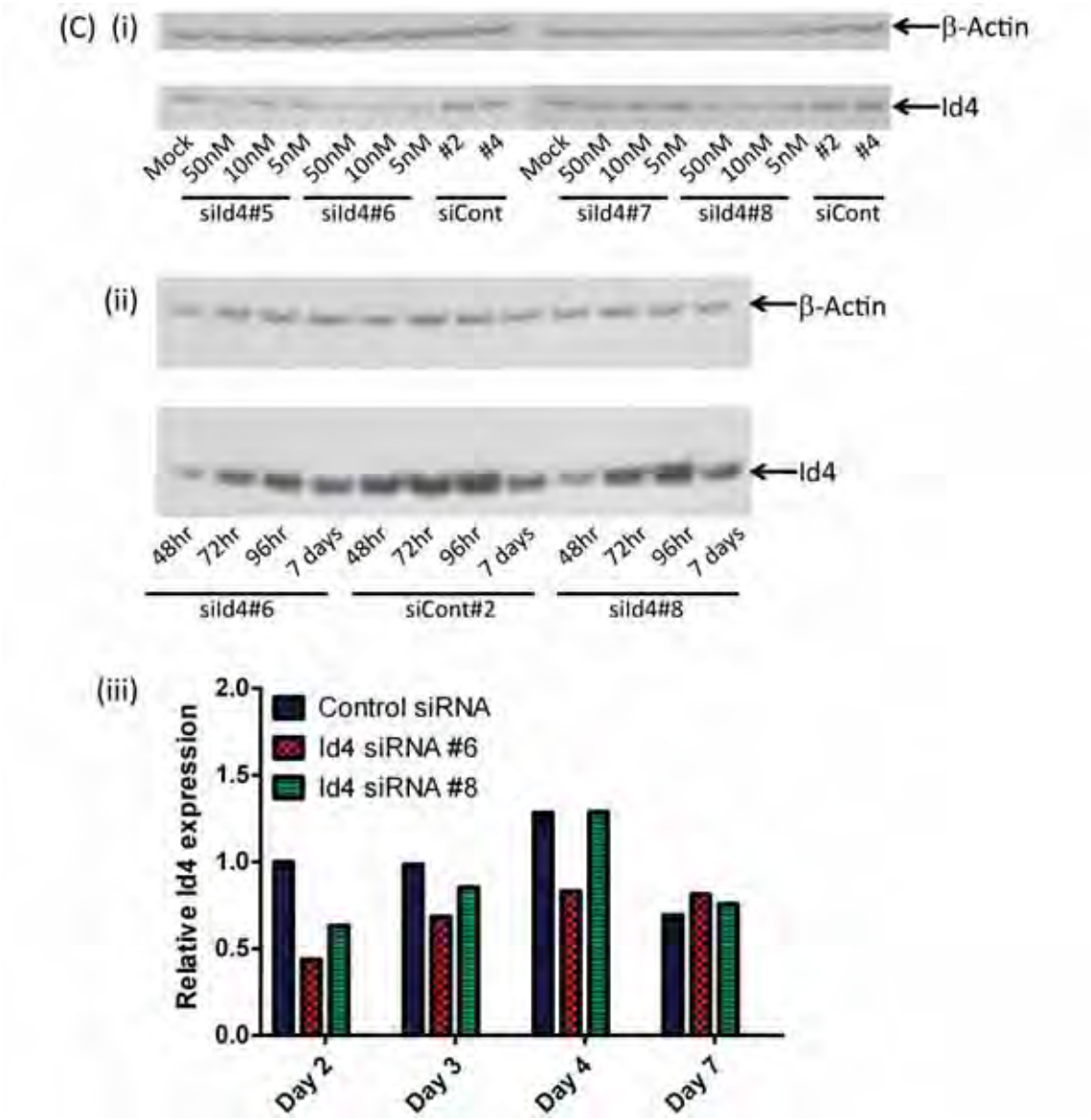


Figure.5-6. Continued. Optimising gene knockdown strategies for Id4 in Comma-Dβ cells.
(C) (i) Id4 knockdown using the Dharmacon siRNAs to Id4 #5, #6, #7 and #8, as well as control siRNAs #2 and #4 transfected at the indicated concentrations using Lipofectamine 2000 and, 48 hours later, lysates were collected and analysed by Id4 western blot. (ii) Time course of Id4 knockdown using siRNAs (20nM) to Id4 #6 and #8, and siRNA control #2 analysed by western blot. (iii) Densitometry analysis of Id4 levels in (ii) to determine knockdown efficiency.

5.2.4.2 Knockdown of Id4 has a dramatic effect on Comma-Dβ growth at subconfluent densities in vitro

Comma-Dβ cells were transduced with lentiviral vectors expressing one of two Id4 shRNA constructs (sh1444 or sh1446), or a control non-targeting shRNA (shCont). Seventy two hours later, at the time point of maximal Id4 knockdown (as shown in Figure. 5-6B), the cells were passaged into 2 wells each of a 6 well plate at a density of 7 x 10^4 cells/well. Four days later, the cell numbers were counted using a haemocytometer and using trypan blue to exclude

dead cells. Cell proliferation was dramatically reduced in the Id4 knockdown cell lines (Figure 5-7A). There did not, however, appear to be a change in the percentage of trypan blue positive cells suggesting the cells had undergone cell cycle arrest and were not undergoing apoptosis. In contrast, when the cells were not split but allowed to grow in the same wells there was no difference in proliferation in the five days following infection (Figure 5-7B). In this second experiment cells were plated into a 12 well plate at 4.5 x 10^4 cells/well and 6 wells were infected with sh1446 and 6 wells were infected with shCont, cells were then counted 3, 4, and 6 days following plating, equivalent to 2, 3, and 5 days following infection. In the second experiment Id4 levels would have started to drop 48 hours post infection (Day 3) and be maximally reduced at 72 hours post infection (Day 4) as was shown in the time course experiment in Figure. 5-6A.

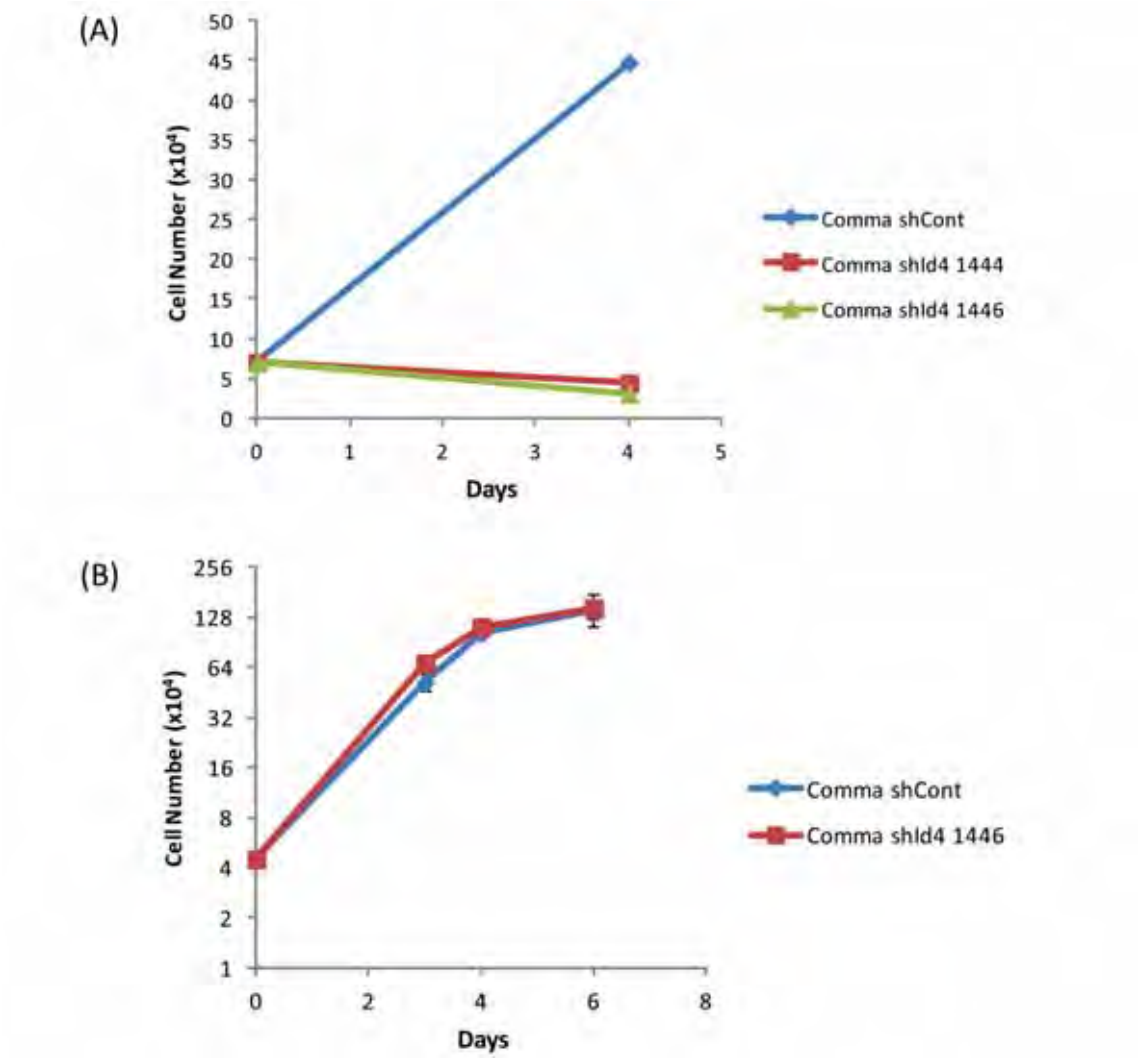


Figure 5-7. Knockdown of Id4 blocks proliferation in sub-confluent Comma-Dβ cells.
(A) Comma-Dβ cell were transduced with lentiviral vectors expressing either sh1444 or sh1446 Id4 shRNA, or a control non-targeting shRNA (shCont). Seventy two hours later the cells were passaged into 2 wells each of a 6 well plate at a density of 7 x 10^4 cells/well. Four days later, the cell numbers were counted using a haemocytometer and using trypan blue to exclude dead cells. (B) Comma-Dβ cell growth following infection with shId4 1446 or shCont lentivirus, without subsequent passage.

5.2.4.3 Knockdown of Id4 enhances luminal differentiation and milk protein production in Comma-Dβcells in vitro

As we had seen the suppression of milk protein production and luminal differentiation in the Comma-Dβ cell line when Id4 was overexpressed, we hypothesised that if Id4 protein was knocked down we would see an increase in milk protein production and luminal differentiation. The *in vitro* differentiation assay was repeated as described in section 5.2.2.2, except that in this experiment the cells were seeded at a slightly higher density and then infected with either shId4 1446 or shCont the day after seeding. The cells were seeded at a higher density because the infection process inhibited the proliferation of the Comma-Dβ cells

thus preventing them from reaching confluence at the correct time for the differentiation assay. It was observed that the Id4 knockdown cells did not form as many 3-dimensional structures as the control cells during the *in vitro* differentiation assay. Protein lysates were analysed for milk protein production. The Day 0 level of β-Casein was dramatically increased in the Id4 knockdown cells and to a lesser extent the levels of WAP were also enhanced (Figure 5-8). Despite the initial higher levels of some of the milk proteins in the Id4 knockdown cells milk protein production did not increase in response to prolactin to the same extent as the control cells during the assay (Figure 5-8). This suggests that while Id4 suppression enhances basal milk protein production by the Comma-Dβ cells it does not enhance their response to lactogenic hormones.

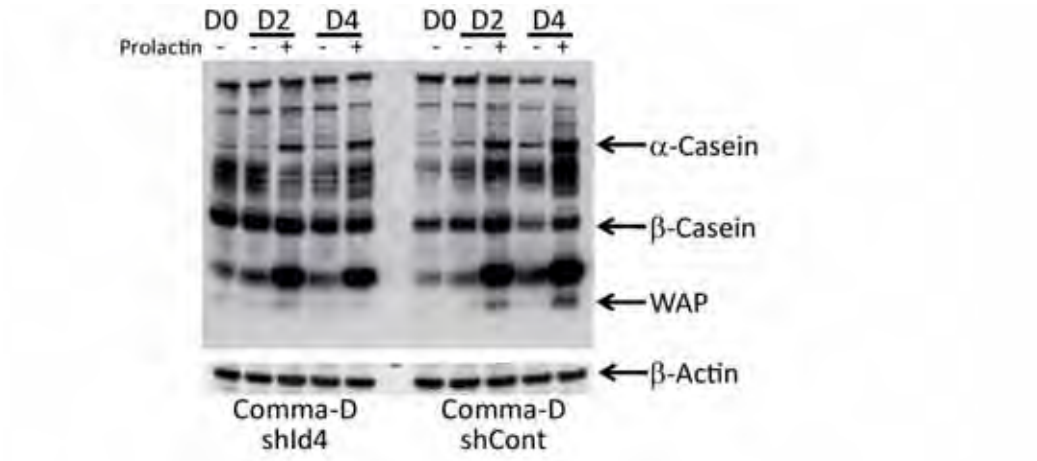


Figure 5-8. Knockdown of Id4 enhances β-casein production *in vitro*.
Comma-Dβ cells were infected with lentiviruses encoding Id4 shRNA (shId4 1446) or control shRNA (shCont). 72 hours after infection, day 0 (D0) samples were taken and the remaining cells either continued in culture untreated or prolactin (0.5μg/ml) treated (as indicated). Samples were also taken on day 2 (D2) and day 4 (D4) of prolactin treatment. Western blot analysis for milk protein production during the *in vitro* differentiation assay is shown. The specific proteins α-casein, β-casein, and WAP are indicated with arrows. β-Actin was used as a loading control.

5.2.5 Overexpression and knockdown of Id4 in a normal mouse mammary cell line does not regulate p38MAPK phosphorylation

As discussed in Chapter 4, it was proposed by Dong et al that the reason for the defect in mammary gland development in Id4 null mice was due to an increase in phosphorylation of p38MAPK and a subsequent increase in apoptosis and reduction in proliferation (Dong et al., 2011). To further investigate this, we looked at levels of phospho-p38MAPK by western blot in samples of Comma-Dβ cells where Id4 levels had either been overexpressed or knocked down using siRNA. The samples analysed for the overexpression experiment were harvested 3 days after plating into 6 well plates at $7x10^4$ cells/well. The samples used for the siRNA knockdown

experiment were the 48 hour time point samples from section 5.2.4.1. No difference in the phosphorylation levels of p38MAPK were observed when Id4 levels were modulated by overexpression. However, there was a slight reduction in phospo-p38MAPK when Id4 was knocked down (Figure 5-9). Similar results were found when cells were harvested at different time points and also using the lentiviral shId4 1446 knockdown construct (data not shown). This demonstrates that decreased p38MAPK phosphorylation is not a direct mode of Id4 action and that regulation of p38MAPK is not likely to be a central component of Id4 action in mammary epithelial cells.

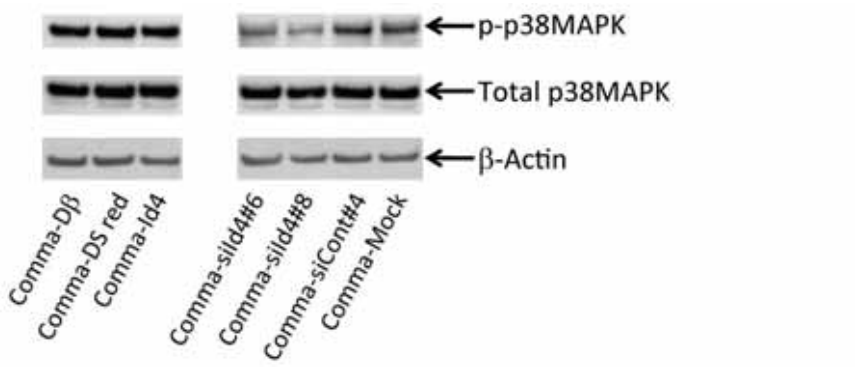


Figure 5-9. Modulating Id4 levels in the Comma-Dβ cell line has no effect on phosphorylation levels of p38MAPK.

Western blot analysis for phospo-p38MAPK, total-p38MAPK, and β-Actin in Comma-Dβ parental cells, Comma-DSred, Comma-Id4, Comma-siId4 #6, Comma-siId4 #8, Comma-siCont #2, and Comma-Mock transfected cells. Cells in the first panel (overexpression) were harvested 3 days after plating into 6 well plates at $7x10^4$ cells/well. The samples used in the second panel (siRNA knockdown) were the 48 hour time point samples from section 5.2.4.1.

5.2.6 Genome-wide determination of Id4 target genes and pathways

Having identified and characterised systems for overexpressing and knocking down Id4 in the Comma-Dβ cell line and having seen that each caused a distinct phenotype, we went on to identify the global transcriptional targets of Id4. We reasoned that combining transcriptional analysis of cells from overexpression and knockdown studies would narrow down the direct transcriptional targets of Id4 rather than from secondary downstream effects of Id4 modulation.

5.2.6.1 Overexpressing Id4 in Comma-Dβ cells for profiling

To determine the most appropriate time point and conditions for analysing Id4 target genes we examined the results from our overexpression studies. From the *in vitro* differentiation experiments discussed above (Section 5.2.2.2) it was noted that Comma-Dβ cell line undergoes limited luminal differentiation even without lactogenic hormones and Id4 overexpression

suppressed this. A distinct morphological phenotype and a number of gene expression changes were observed in the D2 without prolactin samples from the *in vitro* differentiation assay (see Figure). Thus, this time point was chosen for the transcript profiling experiments. Four independent overexpression experiments were set up with low passage number Comma-Id4 and Comma-DSred cells. They were subjected to EGF withdrawal 3 days after seeding and then the cells were harvested for RNA 3 days after EGF withdrawal. The RNA was extracted using a column based kit (QIAGEN) and RNA integrity was checked using the bioanalyser. This confirmed that the RNA was of high quality and an aliquot was sent to the Ramaciotti Centre, UNSW for hybridisation to Affymetrix Mouse 1.0 ST whole genome arrays. Overexpression was confirmed by western blotting protein from cells set up in parallel with the cells for RNA extraction (data not shown).

5.2.6.2 Knocking down Id4 in Comma-Dβ cells for profiling

Significant knockdown of Id4 was observed at 72 hours after lentiviral infection with the pMISSION-shId4-1446 construct; this was also the time point when the control cells had the highest level of endogenous Id4 protein expression (Figure 5-6). It was also seen that Id4-knockdown cells split at this time point failed to proliferate upon reseeding (Figure 5-7), thus this time point was selected for transcript profiling. Four independent knockdown experiments were set up with low passage number Comma-Dβ cells, these were infected either with the shId4 constructs 1444 and 1446 or the control constructs shCont and pLV4311, media was changed 18 hours later and the cells were harvested for RNA 3 days after lentiviral infection. The RNA was extracted using a column based kit and high RNA integrity was confirmed using a bioanalyser. Id4 knockdown was first checked using a Taqman probe, in the samples infected with shId4 1446 50% knockdown was confirmed in two of the samples and 30% in the other two samples, none of the samples infected with shId4 1444 showed any reduction in Id4 transcript levels. It was decided that the three best shId4 1446 samples and their respective controls would be used for the microarray analysis. An aliquot of these samples were sent to the Ramaciotti Centre, UNSW for hybridisation to Affymetrix Mouse 1.0 ST whole genome gene chips.

5.2.7 Analysing profiling results

5.2.7.1 Basic analysis

The profiling was analysed using Genepattern software (Broad Institute), the data from the Affymetrix files was normalised using NormalizeAffymetrixST, and then analysed for significantly changed gene expression using LimmaGP using the default settings. There were

3661 significantly up regulated and 3159 significantly down regulated genes in the Id4 overexpression analysis (Q<0.05), and there were 555 significantly up regulated and 416 significantly down regulated genes in the knockdown analysis (Q<0.05). The Q-value utilises a more stringent criteria for significance than the equivalent p-value, as it takes into account the false discovery rate in profiling studies. The top ten up and down regulated genes based on the Genepattern LimmaGP t-score are listed in Table 5-1 for the overexpression analysis and Table 5-2 for the knockdown analysis. The t-score is determined based on a combination of the Q-value and the fold change in expression. Interestingly, a large number of the top candidates are involved in extracellular pathways such as secretory pathways (eg. VAMP5, PTPRN), cell surface receptors (eg. IL1RL, VCAM1), and secreted factors (eg. EFNA3, ADAMTSL3). This is particularly the case for the overexpression study.

Gene	Direction	Fold Change	P value
VAMP5: vesicle-associated membrane protein 5	UP	2.5142	1.498E-13
NEBL: nebulette	UP	2.7438	2.404E-13
KCNA6: potassium voltage-gated channel, shaker-related, subfamily, member 6	UP	3.1741	6.069E-13
EFNA3: ephrin A3	UP	3.192	9.681E-13
PROM2: prominin 2	UP	2.9641	1.216E-12
BBOX1: butyrobetaine (gamma), 2-oxoglutarate dioxygenase 1 (gamma-butyrobetaine hydroxylase)	UP	2.8906	2.655E-12
ADAMTSL3: ADAMTS-like 3	UP	2.1736	3.725E-12
C1QTNF3: C1q and tumor necrosis factor related protein 3	UP	7.0468	5.473E-12
TMEM154: transmembrane protein 154	UP	2.5505	6.461E-12
TMEFF2: transmembrane protein with EGF-like and two follistatin-like domains 2	UP	1.9981	1.246E-11
IL1RL1: interleukin 1 receptor-like 1	DOWN	14.8991	< 2.2E-16
DCN: decorin	DOWN	8.5786	2.695E-16
--- (Affymetrix probeset: #10468877)	DOWN	25.3942	4.068E-16
CSN3: casein kappa	DOWN	8.7245	2.602E-14
LTF: lactotransferrin	DOWN	6.0399	2.744E-14
C13orf33: RIKEN cDNA 6330406I15 gene	DOWN	8.6853	6.732E-14
VCAM1: vascular cell adhesion molecule 1	DOWN	3.0561	1.902E-13
STC1: stanniocalcin 1	DOWN	6.6948	4.52E-13
CDH11: cadherin 11	DOWN	3.4072	5.934E-13
RUNX1T1: runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	DOWN	4.2303	8.711E-13

Table 5-1 Top 20 regulated genes in the Id4 overexpression experiment in Comma-Dβ cells.

Gene	Direction	Fold Change	P value
PTPRN: protein tyrosine phosphatase, receptor type, N	UP	1.9528	1.8E-10
C20orf194: RIKEN cDNA 4930402H24 gene	UP	1.689	1.6E-09
LMCD1: LIM and cysteine-rich domains 1	UP	1.8286	2.6E-09
TSPAN18: tetraspanin 18	UP	2.0104	4.8E-09
CNN1: calponin 1	UP	2.2703	5.6E-09
PTK2B: PTK2 protein tyrosine kinase 2 beta	UP	1.6688	1.5E-08
PLEKHO1: pleckstrin homology domain containing, family O member 1	UP	1.6228	5.8E-08
FAM26E: family with sequence similarity 26, member E	UP	1.4309	1E-07
TNRC6A: trinucleotide repeat containing 6a	UP	1.4868	1.2E-07
ORM1: orosomucoid 1	UP	1.9097	1.7E-07
YWHAB: tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	DOWN	1.8569	2.2E-11
MSRB3: methionine sulfoxide reductase B3	DOWN	1.9849	1.8E-10
EIF4G1: eukaryotic translation initiation factor 4, gamma 1	DOWN	1.6981	2.3E-10
FAM149A: family with sequence similarity 149, member A	DOWN	1.7284	1E-09
KLHDC5: kelch domain containing 5	DOWN	1.7995	2.6E-09
LARP4B: La ribonucleoprotein domain family, member 4B	DOWN	1.497	6E-09
BIRC5: baculoviral IAP repeat-containing 5	DOWN	1.4789	1.2E-08
VCP: valosin containing protein	DOWN	1.5323	2.4E-08
HIPK3: homeodomain interacting protein kinase 3	DOWN	1.6028	3.1E-08
STAMBP: STAM binding protein	DOWN	1.568	3.7E-08

Table 5-2 Top 20 regulated genes in the Id4 knockdown experiment in Comma-Dβ cells.

5.2.7.2 Gene set enrichment analysis

To determine if certain sets of genes related to relevant biological processes were being regulated in these experiments we performed gene set enrichment analysis (GSEA) using GenePattern on both the overexpression and knockdown experiments. GSEA is a computational method that determines whether predetermined set of genes show statistically significant, concordant differences between two samples. Curated gene sets are available from the molecular signatures database (MSigDB). I have summarised the results for the GSEA using the gene ontology (GO) term gene sets in Table 5-3 and Table 5-4. GSEA was also performed using the other available gene sets however no obvious conclusions could be made from these (data not shown). Genes up regulated in response to overexpression of Id4 were associated with differentiation and developmental pathways, while genes down regulated by overexpression of Id4 were related to cell cycle and proliferation pathways. This suggests that overexpression of Id4 promotes differentiation while reducing cell proliferation as a result. Whilst this may seem to contradict earlier results showing that Id4 is necessary for proliferation, this is a cell density dependent phenomenon where, at low density, Id4 is required for proliferation while at high density including the time point that was used in these

profiling experiments it appears to slow proliferation, as indicated by a lower density of Id4 overexpressing cells in Figure 5-4A. Genes that are up regulated in response to Id4 knockdown appear to be associated with responses to extracellular and intracellular signalling events as well as cytoskeletal processes, while genes down regulated by knockdown of Id4 were also related to cell cycle and proliferation pathways.

GO Gene Set	Direction	Normalised Enrichment Score	P value
POSITIVE_REGULATION_OF_CELL_DIFFERENTIATION	UP	1.861024	<0.0001
MYOBLAST_DIFFERENTIATION	UP	1.854269	<0.0001
DEVELOPMENTAL_MATURATION	UP	1.80778	0.017544
VOLTAGE_GATED_POTASSIUM_CHANNEL_ACTIVITY	UP	1.783918	<0.0001
SKELETAL_MUSCLE_DEVELOPMENT	UP	1.75993	<0.0001
CYTOSKELETON_DEPENDENT_INTRACELLULAR_TRANSPORT	UP	1.73647	<0.0001
ACTIN_FILAMENT_BINDING	UP	1.71234	<0.0001
STRIATED_MUSCLE_DEVELOPMENT	UP	1.70996	<0.0001
TRANSMEMBRANE_RECEPTOR_PROTEIN_SERINE_THREONINE_KINASE_ SIGNALING_PATHWAY	UP	1.694175	<0.0001
SH3_SH2_ADAPTOR_ACTIVITY	UP	1.691741	<0.0001
CELL_CYCLE_PROCESS	DOWN	-2.4642	<0.0001
MITOSIS	DOWN	-2.41522	<0.0001
SPINDLE	DOWN	-2.40156	<0.0001
DNA_REPLICATION	DOWN	-2.37723	<0.0001
M_PHASE_OF_MITOTIC_CELL_CYCLE	DOWN	-2.34504	<0.0001
CELL_CYCLE_PHASE	DOWN	-2.31869	<0.0001
M_PHASE	DOWN	-2.26752	<0.0001
CARBOHYDRATE_BINDING	DOWN	-2.25119	<0.0001
POLYSACCHARIDE_BINDING	DOWN	-2.24989	<0.0001
GLYCOSAMINOGLYCAN_BINDING	DOWN	-2.23859	<0.0001

Table 5-3. Gene sets enriched in the Id4 overexpression experiment.

GO Gene Set	Direction	Normalised Enrichment Score	P value
INTERLEUKIN_RECEPTOR_ACTIVITY	UP	2.043898	<0.0001
ACTIN_CYTOSKELETON	UP	2.025346	<0.0001
HEMATOPOIETIN_INTERFERON_CLASS__D200_DOMAIN__CYTOKINE_ RECEPTOR_ACTIVITY	UP	2.01985	<0.0001
COLLAGEN	UP	1.99253	<0.0001
TRANSMEMBRANE_RECEPTOR_ACTIVITY	UP	1.989691	<0.0001
STRESS_ACTIVATED_PROTEIN_KINASE_SIGNALING_PATHWAY	UP	1.979269	<0.0001
PHOSPHORIC_DIESTER_HYDROLASE_ACTIVITY	UP	1.97638	<0.0001
ACTIN_BINDING	UP	1.975147	<0.0001
REGULATION_OF_SMALL_GTPASE_MEDIATED_SIGNAL_TRANSDUCTION	UP	1.971086	<0.0001
RECEPTOR_SIGNALING_PROTEIN_ACTIVITY	UP	1.969638	<0.0001
CHROMOSOME	DOWN	-2.30401	<0.0001
M_PHASE	DOWN	-2.24394	<0.0001
CHROMOSOMAL_PART	DOWN	-2.20812	<0.0001
SPINDLE	DOWN	-2.18905	<0.0001
CONDENSED_CHROMOSOME	DOWN	-2.12572	<0.0001
REPLICATION_FORK	DOWN	-2.11426	<0.0001
ENDONUCLEASE_ACTIVITY	DOWN	-2.09205	<0.0001
DNA_REPLICATION	DOWN	-2.09	<0.0001
CELL_CYCLE_PROCESS	DOWN	-2.08896	<0.0001
MEIOTIC_CELL_CYCLE	DOWN	-2.08283	<0.0001

Table 5-4. Gene sets enriched in the Id4 knockdown experiment

5.2.7.3 Using Venn Diagrams to determine co-regulated genes

Genes that were reciprocally regulated were determined using Venn diagrams. To do this we relaxed our significance cut offs ($p<0.05$) then selected the genes up regulated by Id4 knockdown and down regulated by Id4 overexpression (Figure 5-10A), or genes down regulated by Id4 knockdown and up regulated by Id4 overexpression (Figure 5-10B). The Venn Diagrams show that there were 522 genes that were up regulated in the knockdown study and down regulated in the overexpression study, and there were 327 genes that were down regulated in knockdown study and up regulated in the overexpression study. The top 10 co-regulated genes from each Venn diagram in Figure 5-10 are listed in Table 5-5 and Table 5-6, these top 10 genes were determined by averaging the GenePattern LimmaGP t-scores from the two experiments. The full list can be found in appendices 1 and 2. Again, it was noted that the majority of the co-regulated genes in Table 5-5 have extracellular functions including ECM remodelling (eg. PLAT and MMP9), extracellular signalling molecules (eg. THBS2), and cell surface receptors (eg. VCAM1 and GPR124)

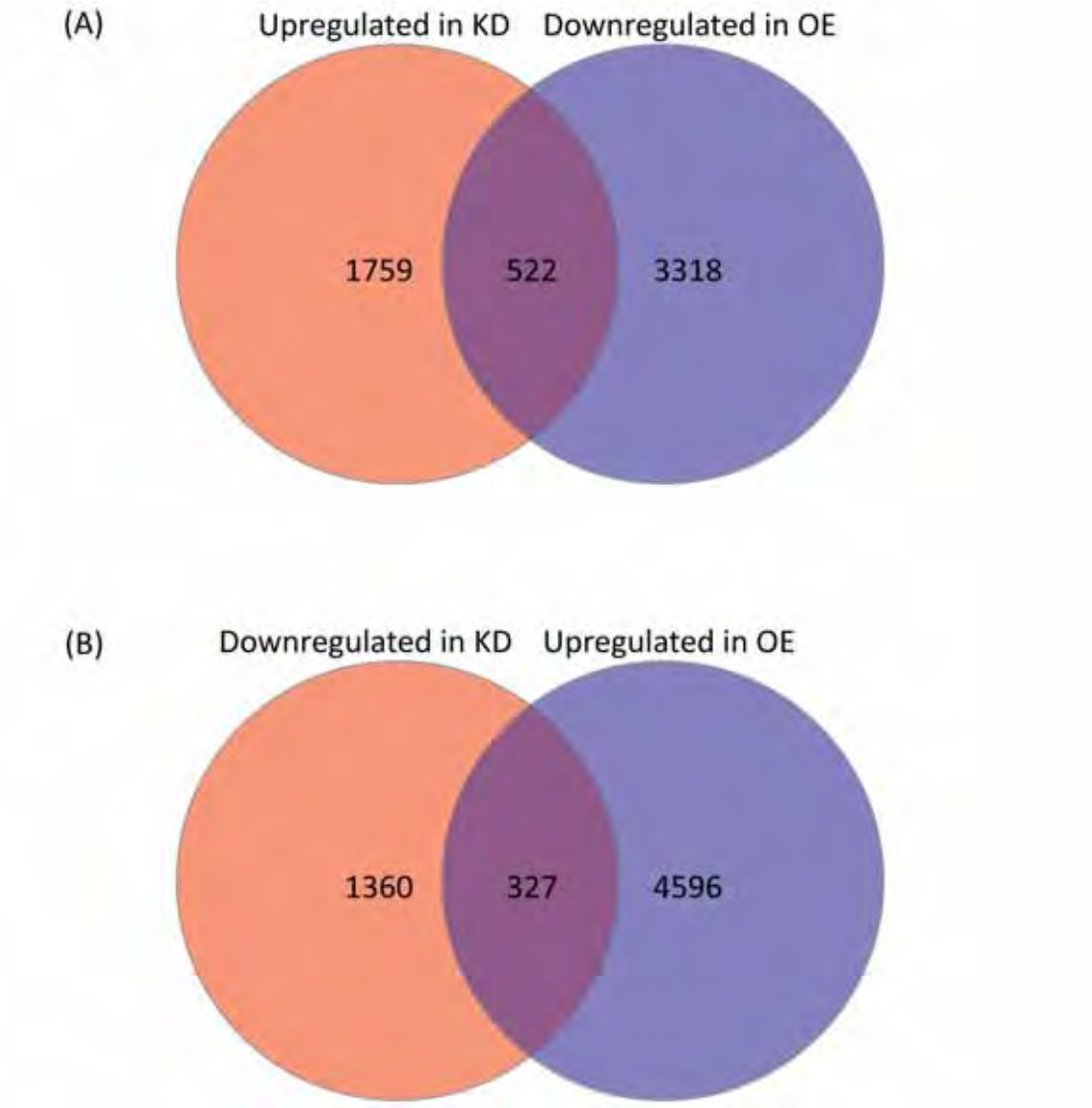


Figure 5-10. Genes that show evidence of reciprocal regulation in the 2 transcript profiling datasets. (A) Venn diagram showing the number of genes that are up regulated by Id4 knockdown (KD) and down regulated by Id4 overexpression (OE). (B) Venn diagram showing the number of genes that are down regulated by Id4 knockdown and up regulated by Id4 overexpression.

Gene	Overexpression p value	Knockdown p value
PTPRN: protein tyrosine phosphatase, receptor type, N	5.465E-09	1.771E-10
VCAM1: vascular cell adhesion molecule 1	1.902E-13	0.0049286
PLAT: plasminogen activator, tissue	9.854E-12	0.000007493
THBS2: thrombospondin 2	2.152E-12	0.0001423
SERPINF1: serine (or cysteine) peptidase inhibitor, clade F, member 1	4.862E-12	0.0000939
MMP9: matrix metallopeptidase 9	7.766E-10	3.668E-07
SEMA7A: sema domain, immunoglobulin domain (Ig), and GPI membrane anchor, (semaphorin) 7A	3.086E-10	0.000001099
GPR124: G protein-coupled receptor 124	2.196E-12	0.0010342
LGI2: leucine-rich repeat LGI family, member 2	9.556E-11	0.000008784
CDH11: cadherin 11	5.934E-13	0.0392414

Table 5-5. Top 10 co-regulated genes up regulated by Id4 knockdown and down regulated by Id4 overexpression.

Gene	Overexpression p value	Knockdown p value
MSRB3: methionine sulfoxide reductase B3	0.0004439	1.809E-10
KLHDC5: kelch domain containing 5	0.000001563	2.605E-09
NEBL: nebulette	2.404E-13	0.0479929
EFNA3: ephrin A3	9.681E-13	0.0283418
PRELP: proline arginine-rich end leucine-rich repeat	1.686E-09	0.000009423
ZDHHC23: zinc finger, DHHC domain containing 23	2.767E-10	0.00004472
FHL1: four and a half LIM domains 1	1.077E-10	0.0002875
LARP4B: La ribonucleoprotein domain family, member 4B	0.00002908	5.974E-09
GYG1: glycogenin	0.0010689	4.753E-08
GNAI1: guanine nucleotide binding protein (G protein), alpha inhibiting 1	3.452E-10	0.0023039

Table 5-6. Top 10 co-regulated genes up regulated by Id4 overexpression and down regulated by Id4 knockdown.

5.2.7.4 Using Ingenuity pathway analysis

To gain a better understanding of the pathways and processes that were being modulated by Id4 overexpression and knockdown we performed Ingenuity pathway analysis (IPA, Ingenuity Systems Inc). IPA performs a number of analyses on the data and can associate gene expression changes with certain disease processes, canonical pathways, and the activation or inhibition of specific transcription factors. The results from the overexpression and knockdown experiments are summarised in Table 5-7 and Table 5-8 respectively. Interestingly cancer was the top disease process associated with both experiments, and furthermore, gastrointestinal disease and developmental disorders were also associated with both experiments. When we specifically analysed the subset of co-regulated genes from Figure 5-10A (down regulated by overexpression/up regulated by knockdown) we saw that the gene targets of the transcription factor Myc were inhibited (p=0.04) and that the gene targets of the transcription factors Snail (p=0.01), Notch-1 (p=0.01), and EZH2 (p=0.1) were activated.

Disease processes	Canonical Pathways	Transcription Factors (Activated)	Transcription Factors (Inhibited)
Cancer	Cell Cycle Control of Chromosomal Replication	CDKN2A	TBX2
Reproductive System Disease	Glycolysis/Gluconeogenesis	TCF3	STAT4
Gastrointestinal Disease	Oestrogen-mediated S-phase Entry	Rb	E2F1
Dermatological Diseases	Aryl Hydrocarbon Receptor Signalling	TP53 (includes EG:22059)	FOXM1
Developmental Disorder	Hepatic Fibrosis/Hepatic Stellate Cell Activation	RB1	STAT1

Table 5-7. Summary of Ingenuity pathway analysis of the Id4 overexpression experiment.
Top 5 results for each of the column categories.

Disease processes	Canonical Pathways	Transcription Factors (Activated)	Transcription Factors (Inhibited)
Cancer	Molecular Mechanisms of Cancer	CDKN2A	MYC
Gastrointestinal Disease	Hepatic Fibrosis/Hepatic Stellate Cell Activation	TP53 (includes EG:22059)	TBX2
Inflammatory Response	Retanoic acid Mediated Apoptosis Signalling	RELA	FOXM1
Developmental Disorder	Signalling by Rho Family GTPases	NFkB (complex)	E2F2
Skeletal and Muscular Disorders	Antiproliferative Role of TOB in T Cell Signalling	TCF3	E2f

Table 5-8. Summary of Ingenuity pathway analysis of the Id4 knockdown experiment.
Top 5 results for each of the column categories.

5.2.8 Validation of transcript profiling results

To validate the transcription profiling results we chose to analyse expression of a number of the top hits and other interesting candidate genes. Since Id4 is considered a negative regulator of transcription factors we considered the genes up regulated by Id4 knockdown and down regulated by Id4 overexpression the most likely direct targets of Id4. We initially confirmed the profiling results by analysing mRNA expression levels of *ptprn*, *mmp9* and *notch1* using specific Taqman probes in the same RNA samples used for the profiling experiments, as well as in RNA extracted from the mammary glands of Id4 null (n=2) and heterozygous (n=4) mice. These experiments were performed with the assistance of Jessica Yang. Both *ptprn* and *mmp9* were significantly down regulated by Id4 overexpression and up regulated by Id4 knockdown in the transcript profiling experiment. This was confirmed using Taqman assays on the cell line samples (Figure. 5-11). *Ptprn* expression was barely detectable in the primary mouse mammary gland samples coming up between the 33rd and 37th cycle. There was a trend to an increase in *mmp9* expression in the mammary glands from the knockout mice. *Notch1* was not significantly changed in either of the profiling experiments.

However, Ingenuity Pathway Analysis had suggested that NOTCH1 target genes were being activated by Id4 knockdown so we wanted to determine if the more sensitive Taqman assays could detect a change in *notch1* expression. Indeed, *notch1* expression was up regulated in the knockout mouse, suggesting that it could be a downstream target of Id4. The Taqman results confirm the results that we saw in transcript profiling experiments, however the number of knockout mice need to be increased before coming to any definitive conclusions about the results from the whole mammary glands.

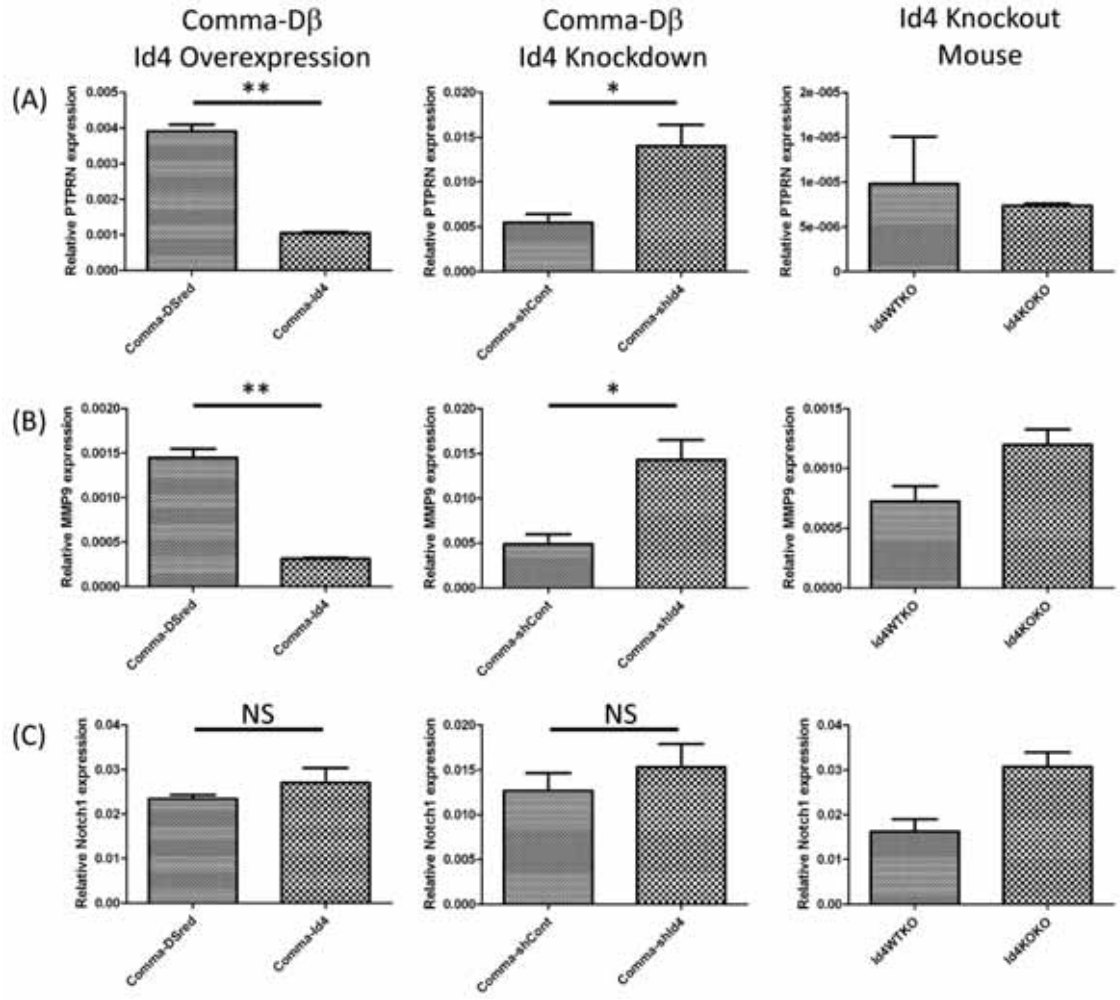


Figure. 5-11. Validation of array results by RT-PCR using Taqman probes.
Relative mRNA expression of PTPRN (A), MMP9 (B), and NOTCH1 (C) in RNA from the Id4 overexpression experiment, the Id4 knockdown experiment and in RNA extracted from Id4 heterozygous and knockout mammary glands, as indicated. **p<0.0001, *p<0.05, unpaired t-test.

5.3 Discussion:

I have established here that the Comma-Dβ cell line is a good model for examining the function of Id4 in mammary epithelial cells. This cell line expresses Id4 protein and these cells can differentiate into Id4 positive myoepithelial cells and Id4 negative luminal cells when

transplanted *in vivo*. Modulation of Id4 expression in this model demonstrated that overexpression of Id4 had little effect on Comma-Dβ proliferation yet had a dramatic effect on the ability of these cells to differentiate into milk producing cells in response to lactogenic hormones. Overexpression of Id4 also had a pro-tumourigenic effect when these cells were transplanted into the fat pad of wildtype BALB/C mice. In contrast, knocking down Id4 led to a dramatic attenuation in proliferation when cells were subconfluent but allowed for enhanced milk protein production in the absence of lactogenic hormones.

These results correlate well with a number of my findings for the role of Id4 using the Id4 knockout mouse described in Chapter 4. Id4 loss in the mouse led to precocious luminal differentiation and a reduction in stem cells. Using the Comma-Dβ cell line, I demonstrated that knocking down Id4 inhibited proliferation and promoted luminal differentiation, and furthermore that overexpression of Id4 inhibited luminal differentiation. We thus determined that the Comma-Dβ cell line was a good model for Id4 function in the mammary gland and could be used to examine the transcriptional targets of Id4. There were a number of advantages to transcript profiling the Comma-Dβ cell line as opposed to whole mammary glands or primary cells from the Id4 null mouse. Firstly, the cell line approach eliminates the problems associated with stromal cell contamination. It also offers a tractable system for the analysis of both knockdown and overexpression of Id4.

Working with primary tissue or cells brings a number of other considerations; firstly whether to analyse the whole mammary gland or dissociated (and possibly sorted) mammary epithelial cells. Analysing whole mammary glands has the advantage of allowing for the snap freezing of the tissue following dissection with minimal chance of any impacts on gene expression in the cells introduced by processing the tissue. Whole mammary gland profiling does however have the problem of combining the mRNAs from all the different stromal and mammary epithelial cell sub-populations. In contrast, dissociating the mammary epithelial cells and then sorting them for the different epithelial subpopulations may introduce artefacts due to the collagenase/trypsin digestion and sorting processes. Secondly, when working with primary tissue it would be necessary to decide whether to control for mouse age or developmental stage between the knockout and control mice, neither of which would be perfect in the study of the Id4 knockout mouse. Finally, when studying germline knockout mice it is also possible that some transcriptional targets would have been activated/inhibited by compensatory pathways or altered cell differentiation would contribute to apparent changes in expression.

Despite the Comma-Dβ cell line offering a cleaner analysis of Id4 transcriptional targets it has to be noted that cell lines in general have their own considerations, in particular how well growth in 2-dimensional culture can represent the 3-dimensional complex structure of the mammary epithelial ductal tree. Furthermore both copies of the p53 gene are mutated in the Comma-Dβ cell line. We decided that the advantages of using the Comma-Dβ cell line model outweighed the disadvantages and thus we used this model for our analyses.

When we examined the pathways regulated by Id4 using transcript profiling, the results suggested that as well as having cell intrinsic effects on proliferation and differentiation, Id4 also appeared to regulate the interaction with the extracellular matrix and stromal cells. The profiling results did not definitively point to one obvious gene target known to regulate luminal versus myoepithelial cell differentiation or a specific bHLH transcription factor that Id4 was regulating. However, although they were not necessarily the top targets, it was noted using Ingenuity Pathway analysis that the targets of NOTCH-1 were activated by Id4 knockdown. NOTCH-1 is a key regulator of luminal differentiation from the stem/common progenitor cells (Bouras et al., 2008; Chakrabarti et al., 2012; Raafat et al., 2011; Tao et al., 2010; Yalcin-Ozuysal et al., 2010). Interestingly *Notch-1* mRNA was also significantly up regulated in the mammary gland from the knockout mouse compared with wild type controls. Ingenuity pathway analysis also showed that the targets of MYC, a known stem cell factor, were down regulated by Id4 knockdown and up regulated by Id4 overexpression, suggesting a role for Id4 in regulating the stem cell phenotype of mammary epithelial cells. This correlates well with the reduction in mammary repopulating units seen in the Id4 null mouse and with the reduction in proliferation of Comma-Dβ cells following Id4 knockdown. The other two transcription factors activated by Id4 knockdown/inhibited by Id4 overexpression are SNAIL and EZH2. Both of these genes have been associated with a stem or EMT like phenotype which, at least initially, appears counter intuitive. However, it is known that Snail is a downstream target of hedgehog and Notch signalling, both of which are key regulators of luminal differentiation. Furthermore, it has been shown that transgenic overexpression of EZH2 under the MMTV promoter leads to an increase in GATA3, another key regulator of luminal differentiation (Katoh and Katoh, 2008; Li et al., 2009). One transcription factor that was down regulated by both Id4 overexpression and knockdown was *foxm1*. FOXM1 was recently shown to be strongly associated with the luminal progenitor lineage, to promote stem to luminal progenitor differentiation but to negatively regulate terminal luminal differentiation through inhibition of GATA3 (Carr et al., 2012). FOXM1 thus needs to be tightly regulated to promote luminal progenitor differentiation but then be suppressed to allow for terminal

luminal differentiation. This might indicate that the Comma-Dβ cell line is enriched in luminal progenitor cells and modulation of Id4 can induce the cells either to dedifferentiate (Id4 overexpression) or terminally differentiate (Id4 knockdown). These results suggest to us that Id4 regulates multiple pathways that in combination lead to an inhibition of luminal differentiation.

PTPRN was our top candidate gene that was significantly down regulated by Id4 overexpression and up regulated by Id4 knockdown. PTPRN, also known as PTP35, ICA512, and IA2, is a member of the protein tyrosine phosphatase receptor family, however, it lacks any phosphatase activity (Magistrelli et al., 1996). Interestingly, PTPRN is an auto-antigen that is associated with type-1 diabetes (Mziaut et al., 2008). PTPRN is known to be expressed by neuroendocrine cells such as pancreatic beta-cells and has been shown to locate to the secretory granules and to regulate insulin secretion (Torii, 2009). Part of its role in regulating insulin secretion is through the promotion of STAT-3 and -5 phosphorylation (Mziaut et al., 2008). The STAT proteins have been demonstrated to be key regulators of mammary epithelial differentiation down the luminal and alveolar lineages (Visvader, 2009). Thus it is possible that, through its activation of STAT-5, PTPRN could be promoting luminal cell differentiation in mammary epithelial cells (Yamaji et al., 2009). PTPRN possibly regulates STAT-3 and -5 phosphorylation through the binding to, and inhibition of, active receptor tyrosine phosphatases (Gross et al., 2002). Expression of PTPRN in primary mammary epithelial cells and its regulation of STAT-5 phosphorylation remains to be confirmed. There is also the interesting possibility that PTPRN could be regulating secretory granules containing the MMPs involved in ECM remodelling process.

MMP9 was the 4th most up regulated gene in the knockdown study and the 6th most significantly co-regulated gene in both experiments. MMP9 encodes a matrix metalloproteinase that is a type IV collagenase. As well as being differentially expressed in these transcript profiling experiments, MMP9 was also chosen for further validation as it is known to regulate tumour progression and metastasis (Coussens et al., 2000; Kessenbrock et al., 2010). The exact role MMP9 plays in mammary gland development also needs further elucidation (Coussens et al., 2000; Khokha and Werb, 2011). It was one of many genes encoding ECM remodelling proteins that were differentially expressed by Id4 overexpression and knockdown, these included a large number of MMPs, ADAMS, ADAMTS, TIMPs, and plasminogen activators along with a number of ECM components such as nidogen and collagen. Due to the highly dynamic nature of the mammary gland through the different

developmental stages, as well as through the oestrus cycle, there is almost constant tissue remodelling (Khokha and Werb, 2011) and our data suggest that Id4 could be a major regulator of this remodelling process.

The results presented above indicate multiple roles for Id4 through regulation of proliferation, differentiation pathways, and thirdly through its regulation of ECM remodelling. These three aspects of Id4 function need not be mutually exclusive, with ECM integrin interactions known to regulate mammary epithelial differentiation pathways, and the regulation of growth factor release such as latent TGFβ and amphiregulin being crucial for growth and differentiation responses (Khokha and Werb, 2011; Muschler and Streuli, 2010; Naylor et al., 2005). Thus Id4 could either be regulating differentiation and proliferation through its regulation of ECM remodelling or it could instead be modulating both pathways independently. The pathways observed to be regulated by Id4 in these experiments fit well with the cellular proliferation and differentiation phenotypes observed during our studies of the Comma-Dβ cell line and also with the phenotype of the knockout mouse.

Chapter 6. The role of Id4 in breast cancer

6.1 Introduction:

It is becoming increasingly clear that factors that regulate developmental processes are often perturbed during cancer. This is particularly well established in breast cancer, as exemplified by the central role oestrogen plays in mammary gland development as well as in the majority of breast cancers. More recently, a number of other factors regulating mammary development have been shown to play a role in breast cancer such as GATA3, BRCA1 and Notch1 (Bouras et al., 2008; Kouros-Mehr et al., 2008; Lim et al., 2009). The role of Id4 in breast cancer is currently not well understood. Results presented in Chapter 5 demonstrate that the overexpression of Id4 in the Comma-Dβ cell line promotes tumourigenesis *in vivo*, however there is conflicting evidence in the literature as to the role Id4 plays in breast cancer with some reports suggesting that Id4 is a tumour suppressor gene while others suggest that it promotes tumourigenesis. While it is known that Id4 mRNA expression negatively correlates with both ER and BRCA1 expression in sporadic breast cancers and that it also positively correlates with basal-like breast cancer, no one has characterised Id4 protein expression across the full range of breast cancer subtypes (de Candia et al., 2006; Roldan et al., 2006; Turner et al., 2007).

6.1.1 Id4 as a tumour suppressor in breast cancer

Two relatively small clinical cohorts (n=60 and n=170) have been analysed to determine whether there is an association between Id4 promoter methylation and patient outcome. Both studies demonstrated that Id4 promoter methylation was associated with shorter patient survival (Noetzel et al., 2008; Umetani et al., 2005). Furthermore it was seen that Id4 promoter methylation associated with an increased risk of lymph node metastasis (Umetani et al., 2005). Both studies conclude that Id4 acts as a tumour suppressor gene, however neither study directly showed any tumour suppressor activity of Id4 nor proved a direct causal relationship between Id4 promoter methylation and survival. Furthermore neither of these studies stratified patients based on breast cancer subtype and thus could not determine whether Id4 promoter methylation differentiated one more aggressive subtype of breast cancer from less aggressive subtypes of breast cancer rather than acting as a tumour suppressor.

6.1.2 Id4 as a tumour promoter in breast cancer

There have been a number of *in vitro* and *in vivo* studies that suggest that Id4 actually acts as a tumour promoter. Increased Id4 expression has also been associated with tumour progression in a carcinogen induced rat mammary cancer model, with Id4 nuclear protein expression correlating with increased proliferation, invasiveness and tumour weight (Shan et al., 2003). Additionally enforced Id4 expression in the mouse mammary cell line HC11 showed increased proliferation, increased ability to form colonies in soft agar, and an inhibition of differentiation in response to lactogenic hormones (Shan et al., 2003). High Id4 expression has also been associated with tumours that develop in the MMTV-Wnt1 mouse model of breast cancer (Fernandez-Valdivia et al., 2008). Wnt proteins are important for the maintenance of the stem cell compartment during mammary gland development and a recent study suggests that Id4 can regulate the stem cell phenotype. This study demonstrated that Id4 promotes cancer stem cell maintenance and expansion in the 4T1 mouse model of breast cancer, these researchers also found that higher Id4 mRNA expression was associated with tumour recurrence in breast cancer patients using publically available microarray data (Park et al., 2011).

A number of studies have also linked Id4 to BRCA1. BRCA1 is the most commonly mutated gene in familial breast and ovarian cancer and BRCA1 down regulation is a common characteristic of basal-like breast cancer (Turner et al., 2004). Firstly, in an unbiased ribozyme based screen Id4 was discovered to be a potent negative regulator of BRCA1 (Beger et al., 2001). This study however was performed in an ovarian cell line PA-1 and confirmed in the luminal Her2 positive breast cancer cell line SK-BR3, thus its relevance to basal-like breast cancer is uncertain. Secondly Berger and colleagues demonstrate that miR-335 negatively regulates Id4 as well as several other genes that regulate BRCA1 in the luminal breast cancer cell line MCF7, suggesting a coordinated regulation of Id4 and BRCA1 (Heyn et al., 2011). Id4 was linked to BRCA1 in a third study where overexpression of BRCA1 increased the expression level of Id4, suggesting a negative feed-back loop in BRCA1 signalling. These experiments were performed in a kidney cell line and thus its relevance to breast cancer needs to be confirmed (Welcsh et al., 2002).

More recent studies have proposed links between Id4 and the tumour microenvironment. Id4 can been induced by overexpression of mutant p53 in a lung cancer cell line, and that siRNA knockdown of mutant p53 in breast cancer cell lines leads to the down regulation of Id4. This induction of Id4 by mutant p53 was demonstrated to lead to an increase in angiogenesis

through stabilisation of mRNA transcripts for IL8 and GROalpha through direct binding by Id4 (Fontemaggi et al., 2009). The stabilisation of mRNAs by Id proteins is not a mechanism of action that has been reported by others and will require confirmation. Id4 expression has also been shown to be highly induced by TIMP-1 overexpression in MDA-MB-231 breast cancer cells when grown as xenografts but not when these cells were grown in 2D culture (Bigelow et al., 2009). This suggests that Id4 expression in the tumour cells can be modulated by crosstalk with the stromal compartment in response to TIMP-1 activity.

The evidence above suggests that Id4 does play a role in the aetiology of breast cancer, and further investigation is needed to determine its actual function. From my previous results looking at the role of Id4 in mammary gland development it seems likely that the dysregulation of Id4 in breast cancer affects the normal differentiation processes in mammary epithelial cells and that this could be influencing the development and phenotype of breast cancer. Some of the conflicting data regarding Id4 in breast cancer could be due to the different methodologies used and also due to differences in the subtypes of breast cancer studied. It is thus important to establish what methodologies are most appropriate and which subtypes of breast cancer are most relevant for the study of Id4.

6.1.3 Hypothesis:

That Id4 functionally associates with ER-negative breast cancer and that measurement of Id4 protein expression is more accurate than mRNA expression for associating Id4 with outcome in breast cancer patients.

6.1.4 Aims:

The specific aims of this chapter are to:

1. Characterise Id4 promoter methylation and Id4 expression level at the mRNA and protein level in a panel of breast cancer cell lines
2. Determine whether Id4 protein is differentially expressed by the different subtypes of breast cancer by IHC
3. Determine whether Id4 expression correlates with patient survival and/or other clinical parameters

6.2 Results:

6.2.1 Analysis of Id4 promoter methylation, mRNA expression and protein expression in a panel of breast cancer cell lines

Promoter hypermethylation is a mechanism of gene silencing that frequently occurs during tumorigenesis to inactivate tumour suppressor genes (Momparler and Bovenzi, 2000). Id4 promoter methylation has previously been associated with poor outcome in two independent cohorts of breast cancer patients (Noetzel et al., 2008; Umetani et al., 2005), however, this has not been correlated with the subtype of breast cancer or the level of Id4 protein using a well validated monoclonal anti-Id4 antibody. To better understand the correlation of Id4 mRNA and protein expression in different subtypes of breast cancer, we screened a large panel of breast cancer cell lines for Id4 expression at both the mRNA and protein levels. Furthermore, we analysed a subset of these cell lines for the level of Id4 promoter methylation. Unlike primary breast cancers, cell lines have a more limited number of molecular subtypes. The three breast cancer cell line subtyes are: Luminal, Basal-A, and Basal-B. Evidence suggests that Basal-B cell lines correspond to the Claudin-low subtype in humans (Taube et al., 2010).

6.2.1.1 Id4 promoter methylation is observed in a subset of breast cancer cell lines

DNA from a small panel of breast cancer cell lines (listed in Figure 6-5A) was obtained from the Garvan Cancer Program Cell Line DNA bank. The DNA was bisulfite converted and analysed for Id4 promoter methylation using the Sequenom system with the assistance of Shalima Nair from the Epigenetics group, Cancer Program, Garvan Institute. Bisulfite conversion changes unmethylated cytosines into uracils, which can then be analysed either by a PCR and sequencing approach or by a PCR and mass spectrometry approach. The later approach was chosen utilising the Sequenom MassArray system, which analyses the level of methylation across specifically amplified regions of bisufite converted DNA using MALDI TOF mass spectroscopy. Four regions of Id4 were analysed: the first 124 base pair region started at the transcription start site 385bp upstream of the ATG, the second 189 base pair region was located across the transcription start site, the third 147 base pair region started 227 base pairs down-stream of the ATG, and the fourth 130 base pair region started 1022 base pairs down-stream of the ATG, located within an intron (Figure 6-5Ai). Primer pairs 3 and 4 were designed by the EpiDesigner software (Sequenom), however as neither of these primer pairs were located in the promoter region two additional primer pairs (1 and 2) were designed manually. Results showed that the Id4 promoter was hypermethylated across 3 of the 4 regions analysed in both of the luminal cell lines (MCF7 and T47D) and one of the basal cell

lines (MDA-MB-231) (Figure 6-5Aii). The level of hypermethylation of region 3 was lower in the MDA-MB-231 and MCF7 cell lines when compared to the T47D cell line. The other two basal cell lines (HCC1954 and MDA-MB-468) were unmethylated across all four regions (Figure 6-5Aii). Control Sss1 methylated DNA was methylated across all four regions and control whole genome PCR amplified unmethylated DNA was unmethylated across all four regions.

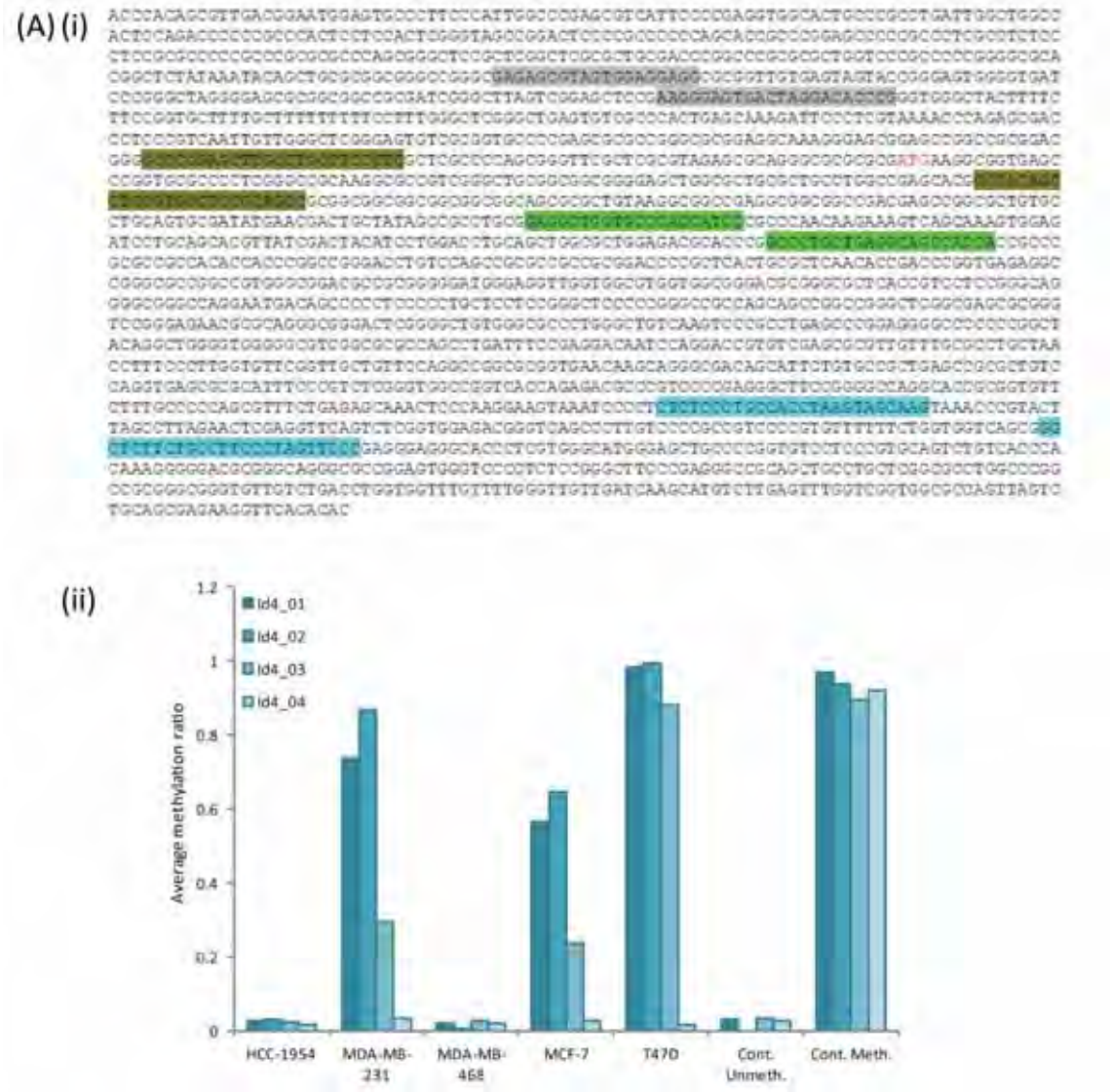


Figure 6-1. Id4 promoter methylation, mRNA level and protein level in a panel of breast cancer cell lines.

(A) (i) The four regions amplified by primers (highlighted in four different colours) for the analysis of methylated CpG residues, ATG translation start codon indicated in red text. (ii) ratio of methylated:non-methylated CpG residues for each of the four regions across Id4 in 5 breast cancer cell lines, a whole genome PCR amplified unmethylated negative control and an Sss1 methyl-transferase treated positive control (as indicated). (Continued on next 4 pages)

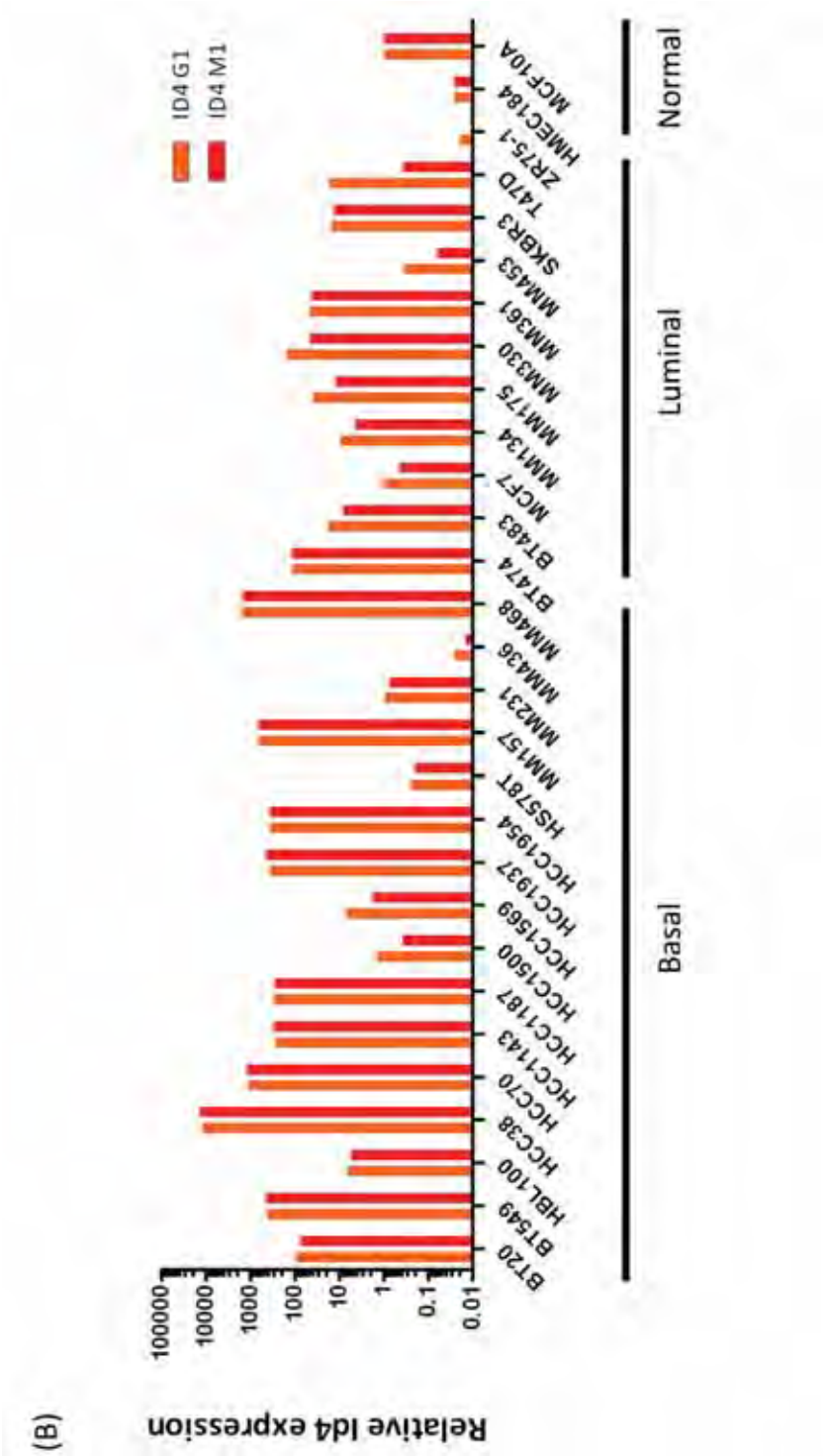


Figure 6-2.Continued. Id4 promoter methylation, mRNA level and protein level in a panel of breast cancer cell lines.

(B) Id4 mRNA levels in a large panel of breast cancer cell lines as measured by two independent Taqman probe sets (m1 and g1, as indicated), signal is normalised to GAPDH and plotted relative to Id4 mRNA expression in the MCF10a "normal" cell line.

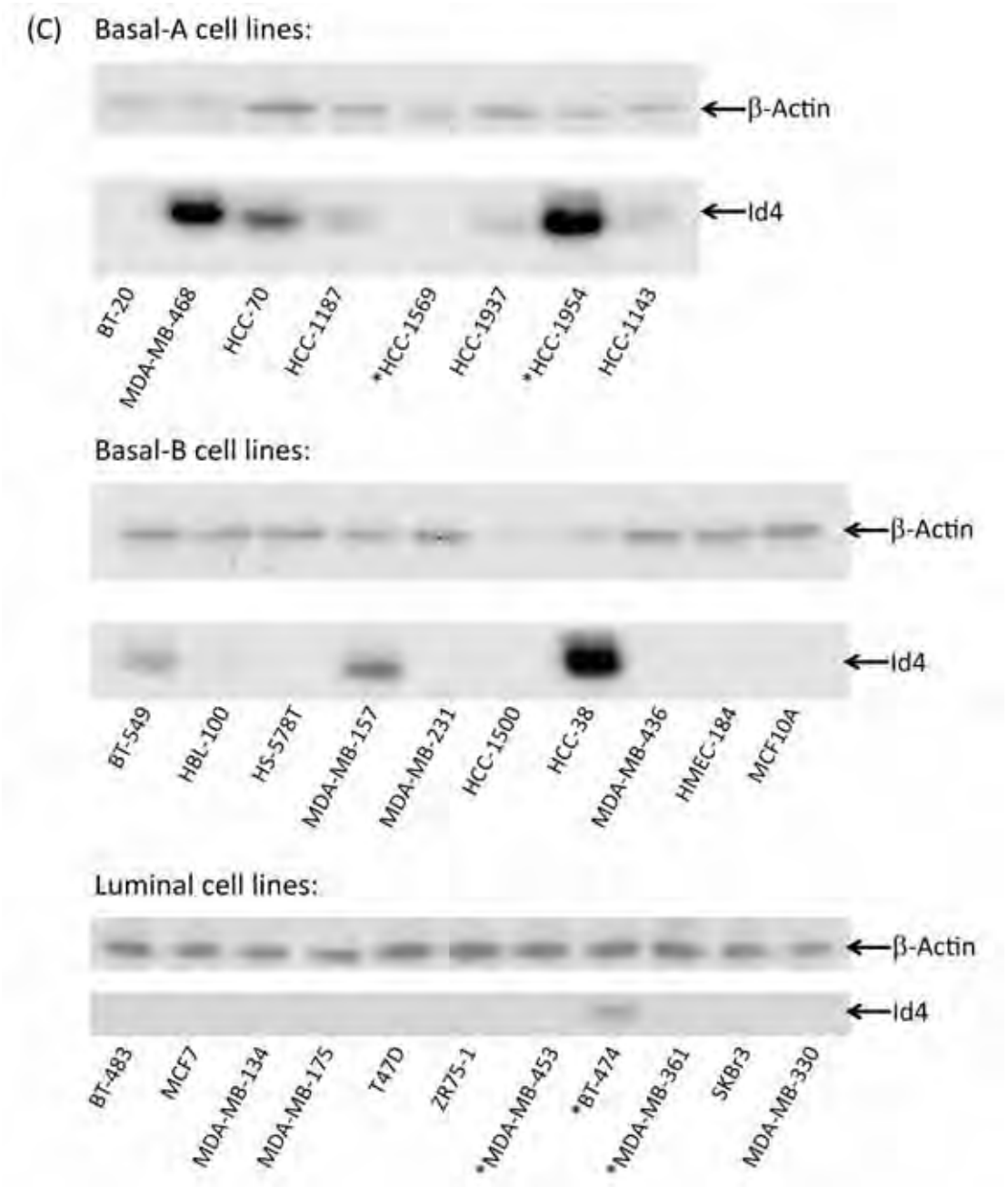


Figure 6-3. Continued. Id4 promoter methylation, mRNA level and protein level in a panel of breast cancer cell lines.

(C) Id4 protein levels in a large panel of breast cancer cell lines as measured by anti-Id4 western blot. Samples are divided into the different subtypes and Her2 amplified samples are marked with an asterisk*.

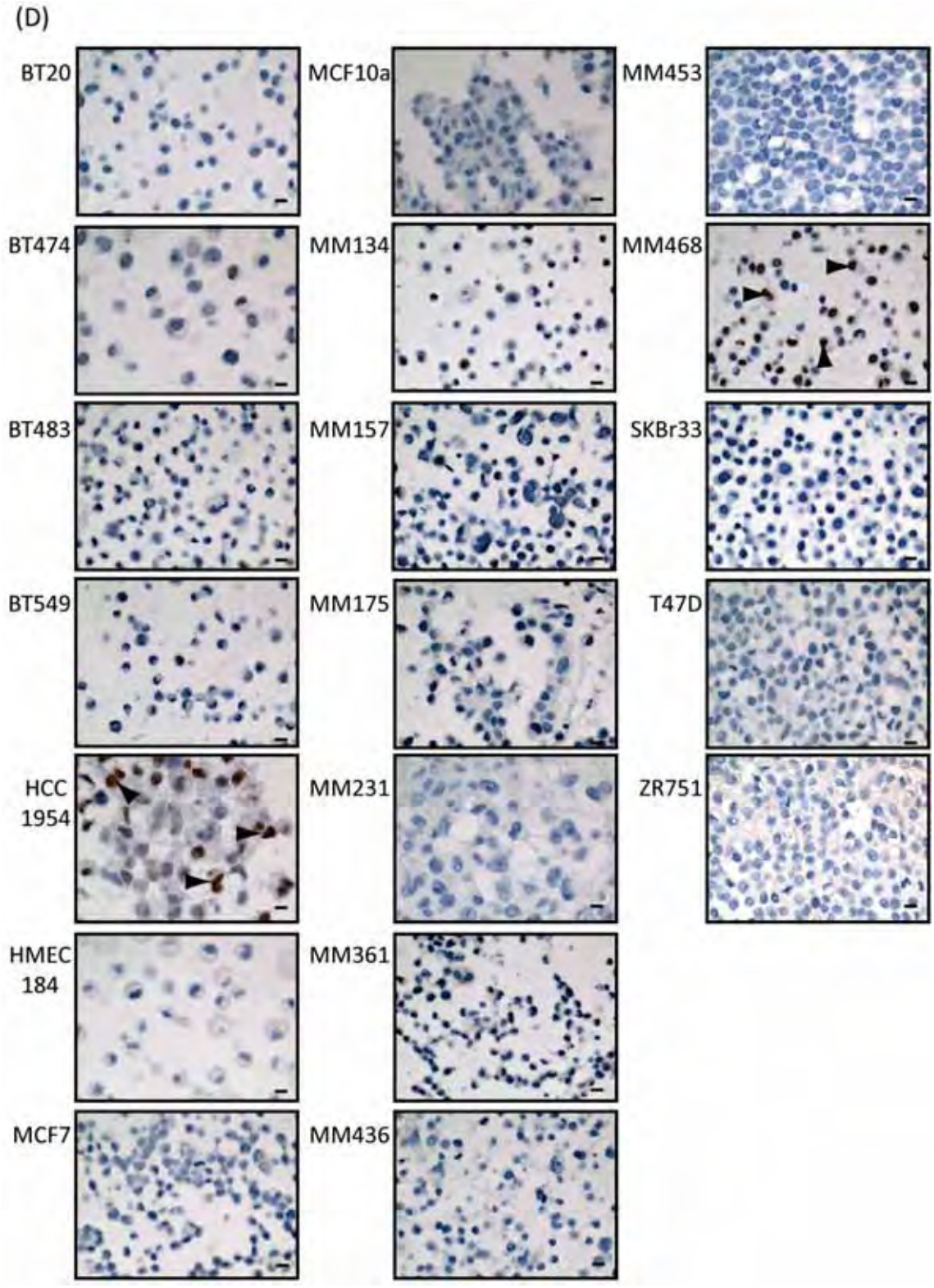


Figure 6-4. Continued. Id4 promoter methylation, mRNA level and protein level in a panel of breast cancer cell lines.

(D) Id4 IHC (brown staining) on a subset of the breast cancer cell lines. Arrow heads indicate brown cells. Scale bars=10μm.

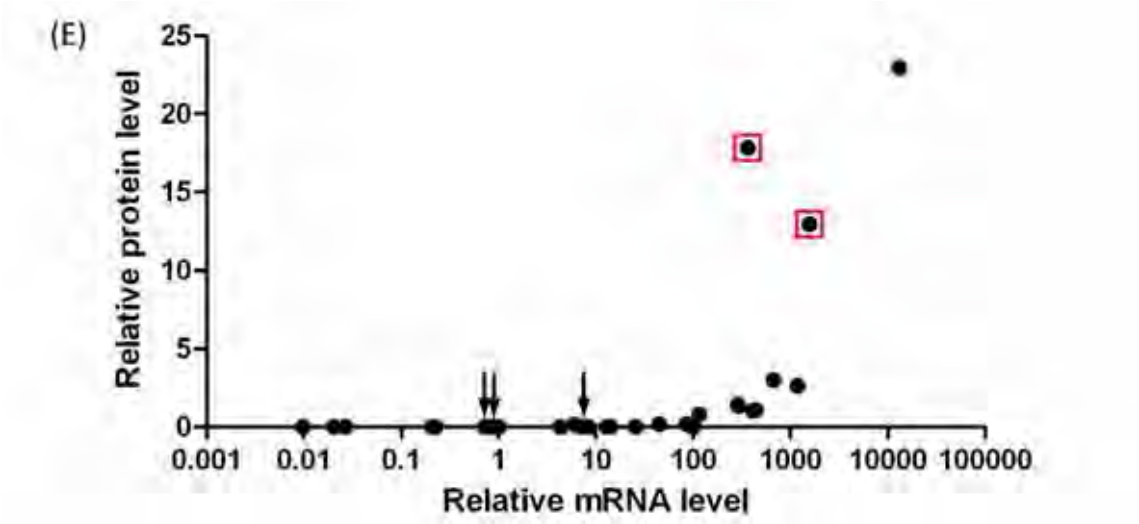


Figure 6-5. Continued. Id4 promoter methylation, mRNA level and protein level in a panel of breast cancer cell lines.

(E) Scatter plot examining the correlation between mRNA levels (by RT-PCR) and protein levels (by densitometry of western blots), with the methylated samples indicated with black arrows and unmethylated samples indicated with red squares.

6.2.1.2 Id4 mRNA levels in a panel of breast cancer cell lines

To determine the level of Id4 mRNA expression in breast cancer cell lines we obtained RNA from a large panel of breast cancer cell lines (Table 6-1) from the Garvan Cancer Program Cell Line RNA bank. The RNA was reverse transcribed using oligo dT primers and Id4 mRNA levels were measured using two different Taqman probe sets and normalised to GAPDH levels, with the help of Laura Baker. Id4 mRNA levels were plotted relative to the transformed normal mammary epithelial cell line MCF10a on a logarithmic scale (Figure 6-5B). Both of the Id4 Taqman probe sets generated similar results in all but a few of the cell lines. This discrepancy may arise from the fact that the G1 probe set detects a region in the 3'-UTR of Id4, which is expressed variably in different cell types (van Cruchten et al., 1998). The highest Id4 mRNA expression was observed in some of the basal cell lines however there was not a clear differentiation of expression levels between basal and luminal cell lines. There also did not appear to be a correlation between Id4 mRNA expression and either of the Basal A or Basal B subtypes of basal cell lines.

Cell Line	Molecular Subtype	Her2 amplification	Tumour type	Id4 mRNA	Id4 protein	Id4 promoter methylation
HMEC-184	Normal	-	N	-	-	ND
BT-474	Lu	+	IDC	++	+/-	ND
BT-483	Lu	-	IDC, pap	+	-	ND
MCF7	Lu	-	IDC	-	-	Methylated
MDA-MB-134	Lu	-	IDC	-	-	ND
MDA-MB-175	Lu	-	IDC	+	-	ND
MDA-MB-330	Lu	+	LC	++	-	ND
MDA-MB-361	Lu	+	AC	+	-	ND
MDA-MB-453	Lu	-	AC	-	-	ND
SKBR3	Lu	+	AC	+	-	ND
T47D	Lu	-	IDC	-	-	Methylated
ZR751	Lu	-	IDC	-	-	ND
BT-20	BaA	-	IDC	+	-	ND
HCC-70	BaA	-	Duc.Ca	+++	++	ND
HCC-1143	BaA	-	Duc.Ca	++	+	ND
HCC-1187	BaA	-	Duc.Ca	++	+	ND
HCC-1569	BaA	+	MC	-	-	ND
HCC-1937	BaA	-	Duc.Ca	++	+	ND
HCC-1954	BaA	+	Duc.Ca	++	+++	Unmethylated
MDA-MB-468	BaA		AC	+++	+++	Unmethylated
BT-549	BaB	-	IDC, pap	++	+	ND
HBL100	BaB	-	N	-	-	ND
HCC-38	BaB	-	Duc.Ca	+++	+++	ND
HCC-1500	BaB	-	Duc.Ca	-	-	ND
HS578T	BaB	-	IDC	-	-	ND
MCF10a	BaB	-	F	-	-	ND
MDA-MB-157	BaB	-	MC	++	++	ND
MDA-MB-231	BaB	-	AC	-	-	Methylated
MDA-MB-436	BaB	-	IDC	-	-	ND

Table 6-1. Id4 expression and promoter methylation levels in a panel of breast cancer cell lines.
mRNA expression relative to MCF10A "normal" cell line. -: <10 fold, +: 10-100 fold, ++: 100-1000 fold, +++: >1000 fold. Protein level was determined by western blot and densitometry: -: 0, +/-: 0-0.9, +: 0.9-2.5, ++: 2.5-5, +++: >5 arbitrary units. IDC: Invasive ductal carcinoma, Pap: Papillary, N: Normal, Duc.Ca: Ductal carcinoma, MC: Metaplastic carcinoma, AC: Adenocarcinoma, LC: lobular carcinoma, F: Fibrocystic disease. ND: Not Done.

6.2.1.3 Id4 protein levels in a panel of breast cancer cell lines

To determine how Id4 protein levels correlate with mRNA levels and promoter methylation we obtained protein lysates from the same panel of breast cancer cell lines (Table 6-1) from the Garvan Cancer Program Cell Line protein lysate bank. Protein lysates were analysed by SDS-PAGE and western blotted for Id4. The same rabbit monoclonal anti-Id4 antibody that was validated in section 3.2.1.1 as being specific to Id4 was used, this antibody also cross reacts with human Id4. Id4 protein was detected at varying levels in a number of the basal cell lines but only a low level of expression was detected in one of the luminal cell lines BT-474 (Figure 6-5C). To confirm the western blot results Id4 protein levels were also analysed in 18 of these breast cancer cell lines by IHC on FFPE sections from cell pellets. The IHC results confirmed the western blot results, where samples were matched. HCC-1954 and MDA-MB-468 cell lines, which had high Id4 expression by western blot, were the only cell lines tested that had extensive Id4 expression by IHC (Figure 6-5D). Rare lightly positive cells were also observed in the BT-474, BT-549 and MDA-MB-157 cell lines. Thus, of the cell lines that were analysed by IHC every cell line demonstrated to express Id4 by western blot also showed expression of Id4 by IHC. Interestingly the IHC results also demonstrated that Id4 expression was not homogeneous within the positive cell lines, with individual cells within a cell line expressing variable amounts of Id4 and with some cells having strong nuclear expression while other cells had no Id4 expression (Figure 6-5D).

6.2.1.4 Correlation between methylation, mRNA expression and protein levels of Id4

Protein levels from the western blots in Figure 6-5C were calculated by densitometry and compared to methylation level and mRNA expression levels. Id4 protein expression correlated very well with promoter methylation in the small subset of breast cancer cell lines analysed for promoter methylation, with high levels of methylation being associated with no protein expression (Table 6-1; Figure 6-5E). High Id4 mRNA expression also correlated with high protein expression ($p<0.001$ $R^2=0.5839$) (Figure 6-5E). However the R^2 value of 0.58 indicates that the correlation between Id4 mRNA expression and protein expression is not linear. This suggests that some post-transcriptional regulation of Id4 protein expression occurs in breast cancer cell lines.

6.2.2 Id4 is highly expressed by a subset of triple-negative and Her2 amplified breast cancer samples

Id4 mRNA expression had previously been associated with basal-like breast cancer (Turner et al., 2007), and the data described shows that Id4 protein is predominantly expressed by a

subset of basal breast cancer cell lines. However, it is unknown whether Id4 protein expression is associated with a particular subtype of breast cancer in patient samples or with patient outcome. Furthermore, mRNA studies rely on tissue lysates that are composed of numerous cell types, IHC on the other hand can localise protein expression to a particular cell type. In all of the following studies on primary human breast cancer samples, the subtypes have been determined by the following criteria:

- **Luminal A:** ER and/or PR positive by IHC;
- **Luminal B:** ER and/or PR positive by IHC as well as Her2 amplified as determined by fluorescence *in situ* hybridisation (FISH);
- **Her2 amplified:** ER and PR negative by IHC and Her2 amplified as determined by FISH;
- **Triple-negative:** ER, PR and Her2 negative;
- **Basal-like:** ER, PR, and Her2 negative but CK5/6 and/or EGFR positive by IHC.

To determine whether Id4 expression is associated with a particular subtype of breast cancer we obtained paraffin sections of 15 Luminal A, 10 Luminal B, 21 Her2 amplified, and 28 triple negative breast cancers from the Australian Breast Cancer Tissue Bank (ABCTB). These were analysed for Id4 expression by IHC. Strong nuclear Id4 expression by epithelial cells was observed in a number of the breast cancer samples (Figure 6-6A). Id4 expression was then scored by a breast cancer pathologist (Dr Sandra O'Toole), using the H-score which is determined by multiplying the percentage of positive epithelial cells (identified morphologically) by the staining intensity (graded on a scale of 0, 1, 2, or 3). Thus an H-score ranges from 0 to 300. This analysis showed that 13 of the 28 triple negative tumours had moderate to high Id4 protein expression (H-score>10), 7 of the 21 Her2 amplified tumours had moderate to high Id4 expression, but only 1 of 15 Luminal A tumours and 1 of 10 Luminal B tumours had moderate Id4 protein expression (Figure 6-6 B). This showed that high Id4 protein expression was associated with a subset of the triple-negative and Her2 amplified subtypes of breast cancer. The division between high Id4 expression and very low Id4 expression in the triple-negative and Her2 amplified subtypes of breast cancer led us to question whether the expression level of Id4 had any clinical significance. Unfortunately we did not have any clinical follow up data on this ABCTB cohort of patients thus we could not correlate Id4 expression with any clinical outcomes or parameters. Instead we went on to look at other cohorts.

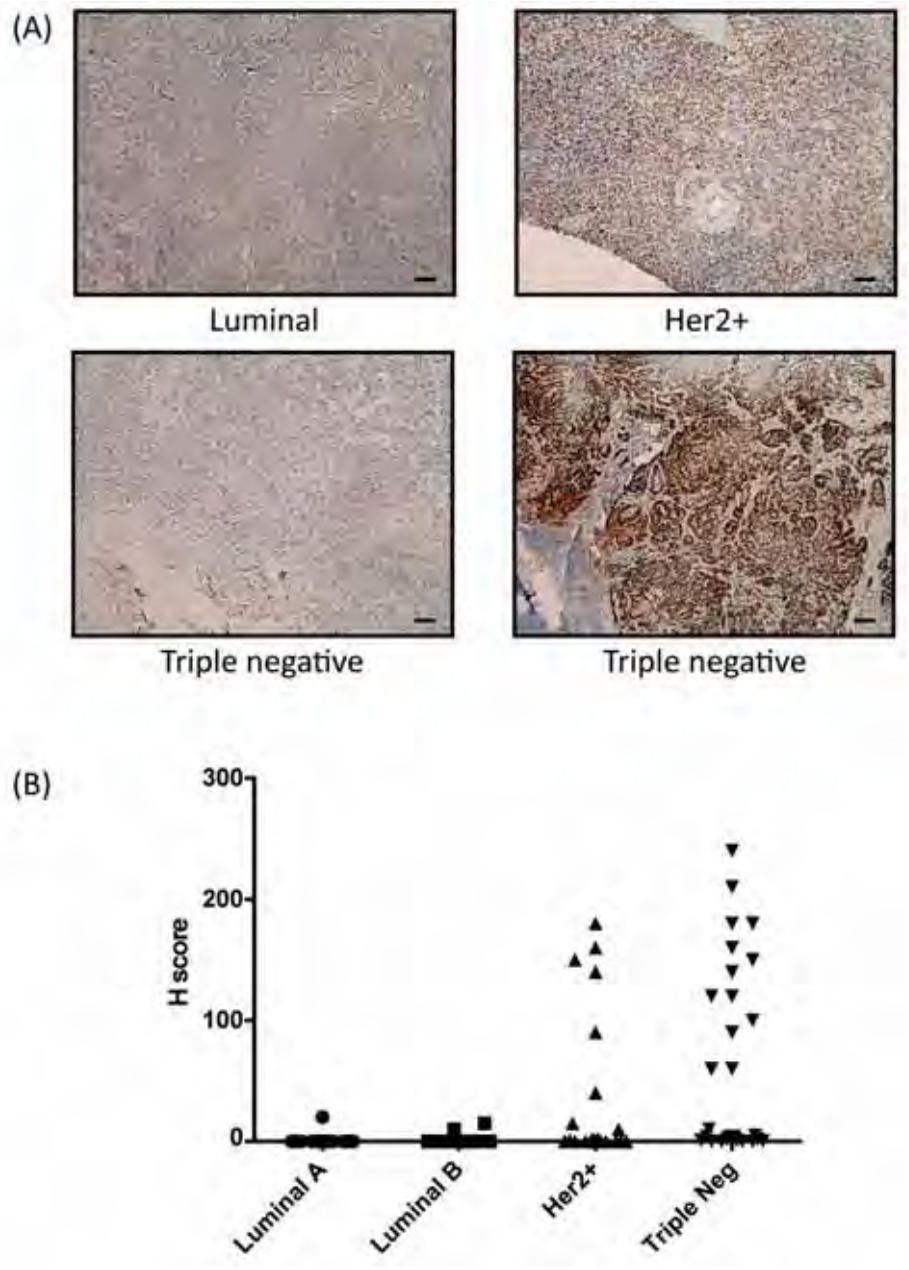


Figure 6-6 Id4 protein expression in the different subtypes of breast cancer.
(A) Example Id4 IHC staining (brown) in human breast cancer subtypes imaged at low magnification from the ABCTB cohort of patients. Scale bars=100μm. (B) Scoring of Id4 protein expression in the subtypes of breast cancer. H−score = % of positive epithelial cells multiplied by the staining intensity with a range of 0-3.

6.2.3 High Id4 mRNA level correlates with better survival within the basal-like and Her2 subtypes of breast cancer

As Id4 mRNA expression correlated with protein expression within the basal breast cancer cell lines, we analysed a publically available microarray dataset, the NKI 295 dataset (van de Vijver et al., 2002), to determine whether there was any association of Id4 mRNA expression with patient outcome. We specifically wanted to look at patients with basal-like and Her2 amplified

subtypes of breast cancer since this is where we had seen protein expression in the ABCTB cohort. This analysis was done with the bioinformatic assistance of Dr Tim Molloy. Similar to what we saw in the breast cancer cell lines, there was a range of expression of Id4 across all the subtypes of breast cancer. However unlike what we saw in the cell lines, or the previously published results (Turner et al., 2007), Id4 mRNA expression did not correlate with the basal-like subtype of breast cancer (Figure 6-7A). When we looked at survival in relation to Id4 expression across all the subtypes of breast cancer we saw no association between Id4 mRNA expression level and either recurrence free survival (RFS) or overall survival (OS) (Figure 6-7B). As Id4 protein expression is limited to the basal-like and Her2 subtypes of breast cancer cell lines we looked specifically at patient survival within these subtypes. The basal-like subtype (n=44) of patients were split half way into Id4 high and low and then this was correlated with outcome. We reasoned that since ~50% of the triple negative patients in the ABCTB cohort had high Id4 expression that this was a good cut point for basal-like breast cancer. When we determined survival we observed that high Id4 mRNA expression correlated strongly with improved patient outcome: Recurrence free survival (RFS) p=0.00017, Hazard ratio (HR)=0.251, Overall survival (OS) p=0.004, HR=0.275 within the basal-like subtype (Logrank test) (Figure 6-7C). The HR indicates the relative risk of death or recurrence when compared to control, thus in the patients with high Id4 expression have a quarter of the risk of recurrence when compared to the patients with low Id4 expression in their tumours. Furthermore, high Id4 expression also correlated with better OS but not RFS when we looked within the Her2 amplified subset (n=31). RFS p=0.0775 HR=0.405, OS p=0.0476 HR=0.328 within the Her2 amplified subtype (Logrank test) (Figure 6-7D). For this analysis we similarly stratified the data around the median expression in the Her2 amplified cohort to define Id4 high verses low, however, this may not be the most appropriate cut point for the Her2 subtype due to less than 50% of the patients showing high Id4 expression in the ABCTB cohort.

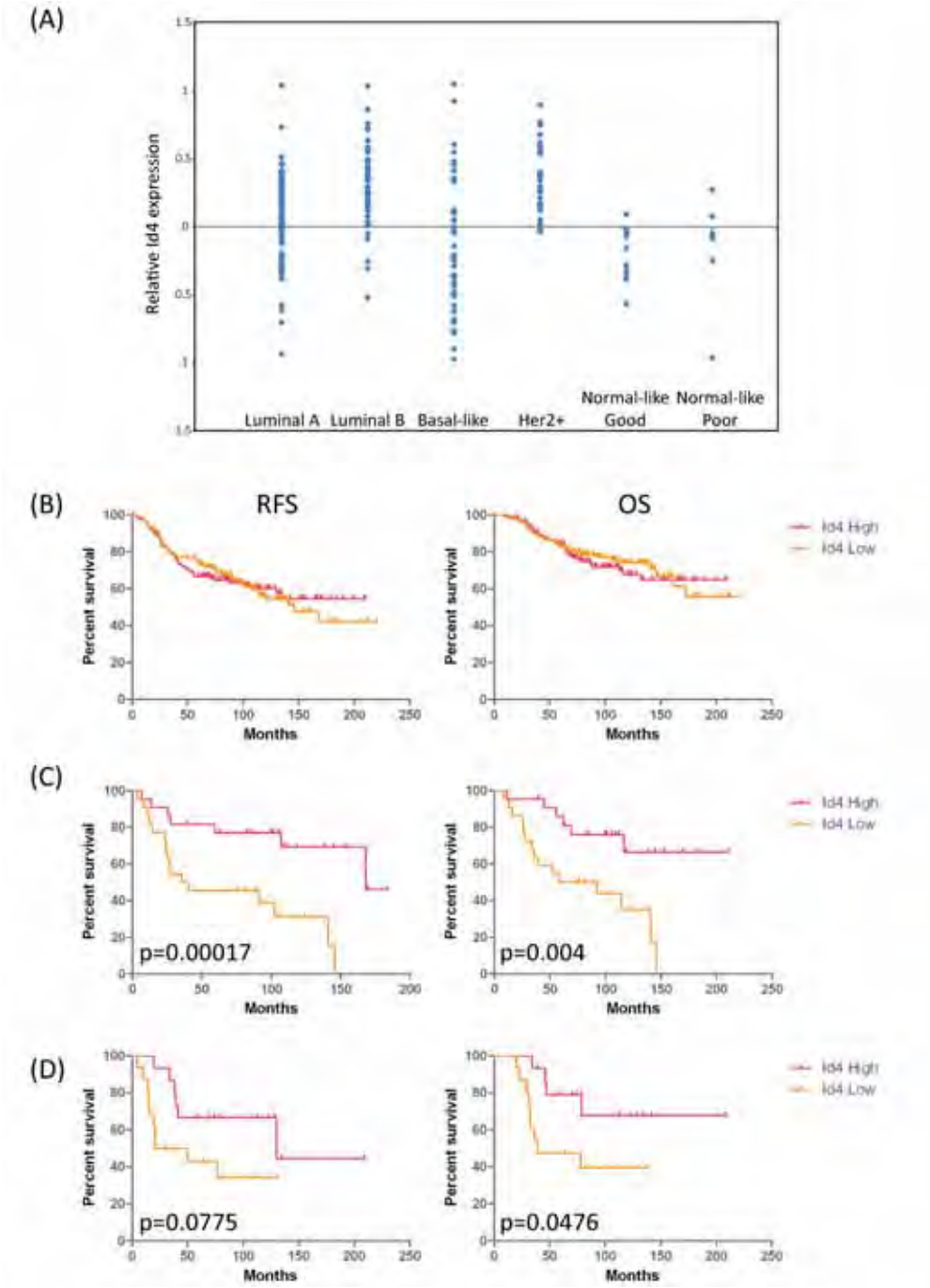


Figure 6-7. Id4 mRNA expression level correlates with outcome within the basal-like and Her2 amplified subtypes of breast cancer in the NKI295 cohort of breast cancer patients.

(A) Id4 mRNA level for each cancer sample is given relative to the median Id4 level over the entire cohort. (B) Id4 high and low populations were designated based on the Id4 expression relative to the median expression level in the whole cohort, as indicated by the grey line in (A). Data from all subtypes was combined and recurrence free survival (RFS) or overall survival (OS) is plotted for Id4 high and low populations. (C) Id4 high and low populations were designated relative to the median expression within the basal-like dataset and data plotted as in (B). High Id4 mRNA expression within the basal-like subtype of breast cancer correlates with overall survival as well as relapse free survival. (D) Id4 high and low populations were designated relative to the median expression within the Her-2 amplified dataset and data plotted as in (B).High Id4 mRNA expression within the Her2 amplified subtype of breast cancer correlates with better overall survival.

6.2.4 High Id4 protein levels correlate with better survival in basal-like breast cancer patients

To better understand the association of Id4 with patient outcome we looked at two cohorts of breast cancer patients to determine if high Id4 protein expression correlated with better patient outcome.

Tissue microarray sections from the Garvan Crea breast cancer cohort of 255 patients were cut and analysed for Id4 protein expression by IHC. The Crea cohort is a single surgeon cohort and these samples were collected prior to the use of molecular targeted therapies (Lopez-Knowles et al., 2010). Follow-up data on the patients in the Crea cohort has been collected over at least 10 years and all the subtypes of breast cancer are represented. Subtypes of breast cancer were again defined by IHC and FISH analysis as described in section 6.2.2. Furthermore, this cohort has been extensively studied and analysed for expression of many other breast cancer associated genes. Nuclear Id4 IHC staining was scored by breast cancer pathologist Dr Sandra O'Toole, and the cohort was divided into Id4 low (H-score≤10) or Id4 high (H-score>10). Id4 expression was again limited to the triple negative (made up of basal-like and five marker negative subtypes) and Her2 amplified subtypes of breast cancer (Figure 6-8A). Similar to the ABCTB cohort, the triple-negative patient samples could be divided into two groups: one with high Id4 expression and one with little to no Id4 expression. Among the 44 samples from triple negative patients and the subset of these that were basal-like (29 samples) there was, again, an association with better survival in the patients with high Id4 expression. When Kaplan-Mayan analysis was performed on these groups of patients it was determined that high Id4 protein expression correlated with better prognosis in the basal-like subtype (p=0.0062 Logrank test, HR=0.097) and within the triple-negative subtype as a whole (p=0.018 Logrank test, HR=0.251) (Figure 6-8B&C). There were too few Her2 amplified cases (n=22) within the Crea cohort to determine if Id4 protein expression correlated with patient outcome.

Within the basal-like subset of breast cancer patients, we also assessed the correlation of Id4 expression with proliferative index (Ki67 expression in tissues with an H score>10), lymph node metastasis and p53 status. It has previously been demonstrated using the whole Crea cohort that a higher proliferative index is associated with poorer outcome (Reyal et al., 2008). Surprisingly, within the basal-like subtype of breast cancer this was not the case. We observed that high Id4 expression correlated with a higher proliferative index (p=0.013 Fisher's exact test). Despite a report that mutant p53 can induce Id4 expression (Fontemaggi et al., 2009) we saw no association of Id4 expression and p53 status (d.n.s). We also looked to see if there was

an association between Id4 protein expression levels and the occurrence of lymph node metastasis. In line with the increased patient survival observed for Id4 high basal-like tumours, we saw that there was a very strong negative correlation between Id4 expression and the presence of lymph node metastasis at diagnosis, with 2 out of 13 Id4 high patients having lymph node metastasis compared to 12 out of 16 Id4 low patients ($p=0.0025$ Fisher's exact test). The negative association between Id4 expression and lymph node metastasis was so strong that it remained, even when analysis was performed across all the subtypes of breast cancer with 11 out of 37 Id4 high patients having lymph node metastasis and 106 out of 215 Id4 low patients ($p=0.032$ Fisher's exact test).

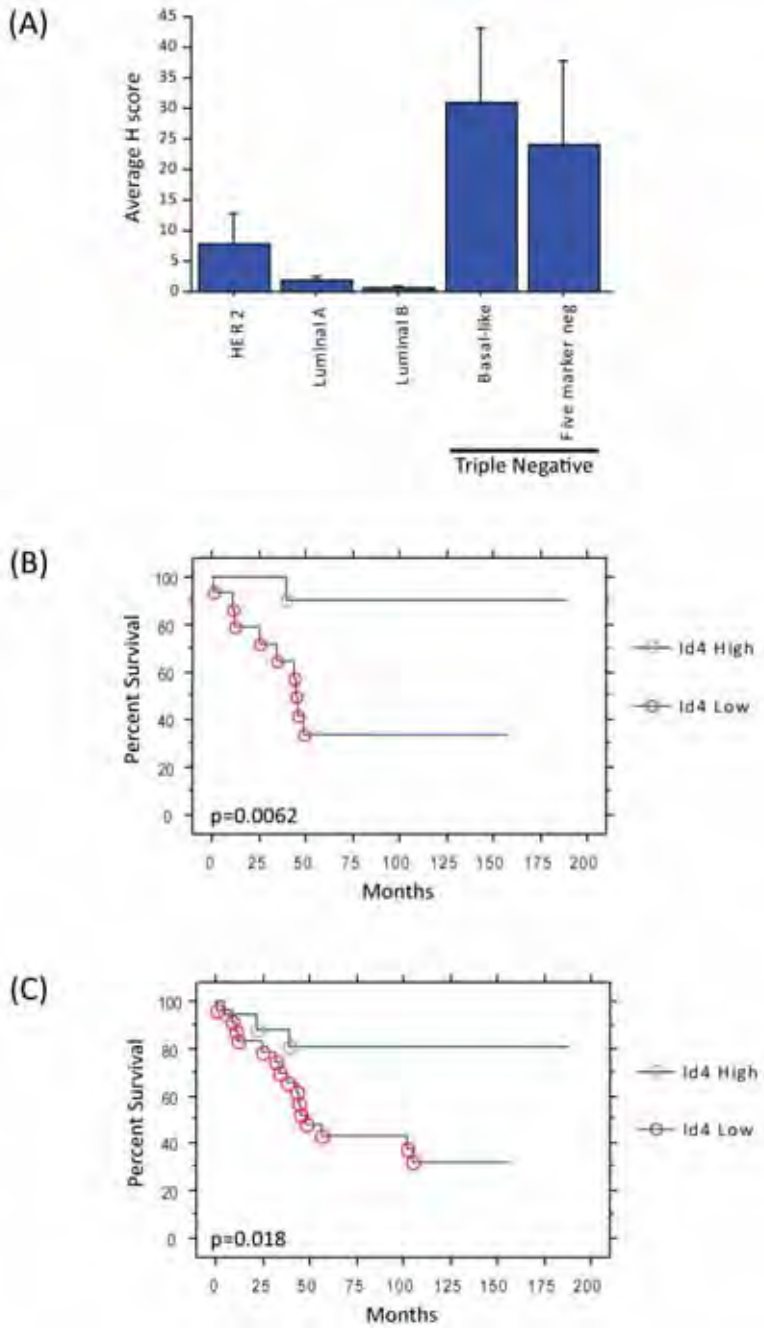


Figure 6-8. Id4 protein expression correlates with better survival in the basal-like breast cancer patients within the Crea cohort of breast cancer patients.

(A) Id4 protein level was assessed by IHC and each sample assigned and H score. Average H scores in each of the breast cancer subtypes is indicated, error bars indicate SEM. (B) Samples that were classified as basal-like (n=29) were divided into Id4 high (H>10) and Id4 low (H<10) groups. Survival curves are plotted for the each groups as indicated. (C) The complete triple-negative subtype (n=44) was assessed as for (B).

6.2.5 Association of better survival with chemotherapy response

One factor suggested to contribute to better survival outcomes in breast cancer patients is that their tumours are less proliferative. However, this was not the case for the basal-like breast cancer patients with high Id4 expression who had better prognosis but elevated proliferative index. We therefore sought other explanations for their improved survival. One alternate explanation was that they respond better to chemotherapy. Only a subset of the patients analysed had received chemotherapy as well as having at least 5 years of follow up data and these were again analysed for survival based on Id4 protein expression. Despite only 35 patients combined from the Crea cohort and a separate Royal Prince Alfred (RPA) hospital cohort fitting this criteria there was a very dramatic and significant difference in survival curves. The patients with high Id4 expression (n=20) had a much better outcome than patients with low Id4 expression (n=15; p=0.0373; HR=3.91) with 80% of the Id4 high patients surviving ten years after diagnosis and only 20% of the Id4 low patients surviving at this point (Figure 6-9). These striking results obviously need to be verified in a larger cohort of patients, but suggest that basal-like breast cancer patients with high Id4 expression respond better to chemotherapy.

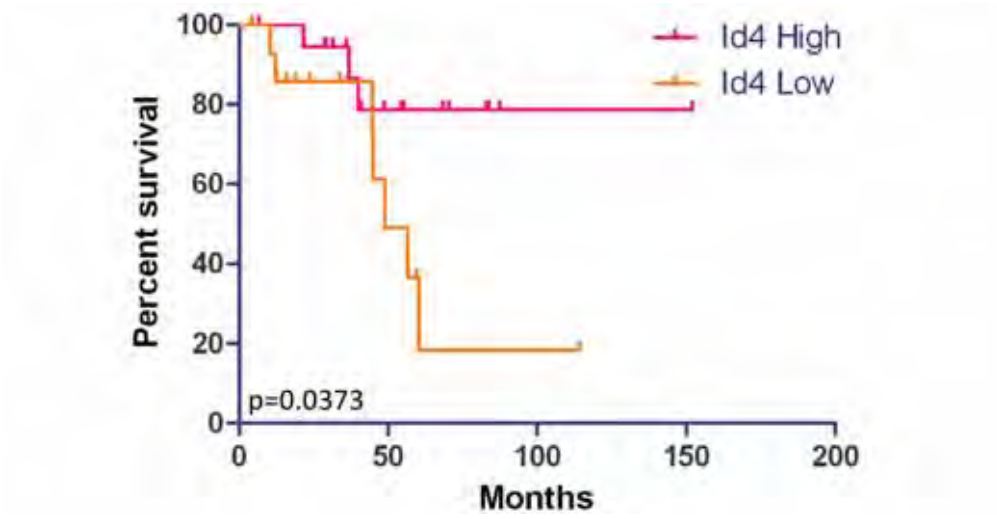


Figure 6-9. High Id4 protein expression predicts better response to chemotherapy in basal-like breast cancer patients.

Kaplan-Meyer analysis of survival was performed in a cohort of chemotherapy treated basal-like breast cancer patients. Data from the Crea and RPA cohorts were combined and patients were divided into Id4 high (H>10) and Id4 low (H<10) groups, as indicated (n=35).

6.2.6 Id4 protein expression levels in a number of mouse models of breast cancer

We intend to utilise mouse models to further understand the role of Id4 in breast cancer. To determine which mouse model would be most appropriate in the study of Id4 we first examined Id4 expression in each model by IHC. The following models were examined: PyMT,

Neu, Myc, Neu+Myc, C3TAg transgenic, and p53 null mice, as well as the transplantable tumour cell lines 4T1 and M6 (Aslakson and Miller, 1992; Donehower, 1996; Green et al., 2000; Holzer et al., 2003; Welm et al., 2005; Welm et al., 2007). The PyMT, Neu, Myc, and Neu+Myc models were generated by retroviral transduction of primary mammary epithelial cells with the relevant oncogene and all of these phenocopied their transgenic equivalents with the exception of the Neu+Myc model, which has no corresponding transgenic model (Welm et al., 2005; Welm et al., 2007). The mouse models have been associated with the different subtypes of human breast cancer by transcript profiling and hierarchical clustering (Herschkowitz et al., 2007). The PyMT, Neu, and Myc models clustered with luminal breast cancer, the Neu+Myc model was not examined in this paper but probably represents a luminal B subtype of breast cancer. The p53 null and C3TAg models mostly cluster with the basal-like subtype of breast cancer.

Multiple tumours from each model were stained for Id4 expression by IHC. Extensive cytoplasmic Id4 expression with rare strong nuclear Id4 expression was observed in the PyMT model. No Id4 expression was observed in the Neu model. A small percentage of cells with strong nuclear expression were observed in the Myc model and a larger percentage of cells with strong nuclear expression were observed in the Neu+Myc model compared with Myc alone. Little Id4 expression was seen in any of the p53 null tumours observed (FigureA). In the C3TAg model there was a high degree of variability between tumours, with some having a high percentage of tumour cells with strong nuclear Id4 expression and other tumours with little expression (FigureB). Neither of the cell line models showed abundant expression of Id4, although a rare subpopulation of Id4 positive cells was seen in the M6 model (FigureC). As the C3TAg model closely resembled what we saw in human basal-like breast cancer we chose this model to further analyse the role of Id4 in breast cancer. We are currently crossing the Id4 null mice with the C3TAg mice, and we plan to analyse any changes in tumour frequency or latency in these mice compared to heterozygous controls. However, this work is beyond the scope of this dissertation.

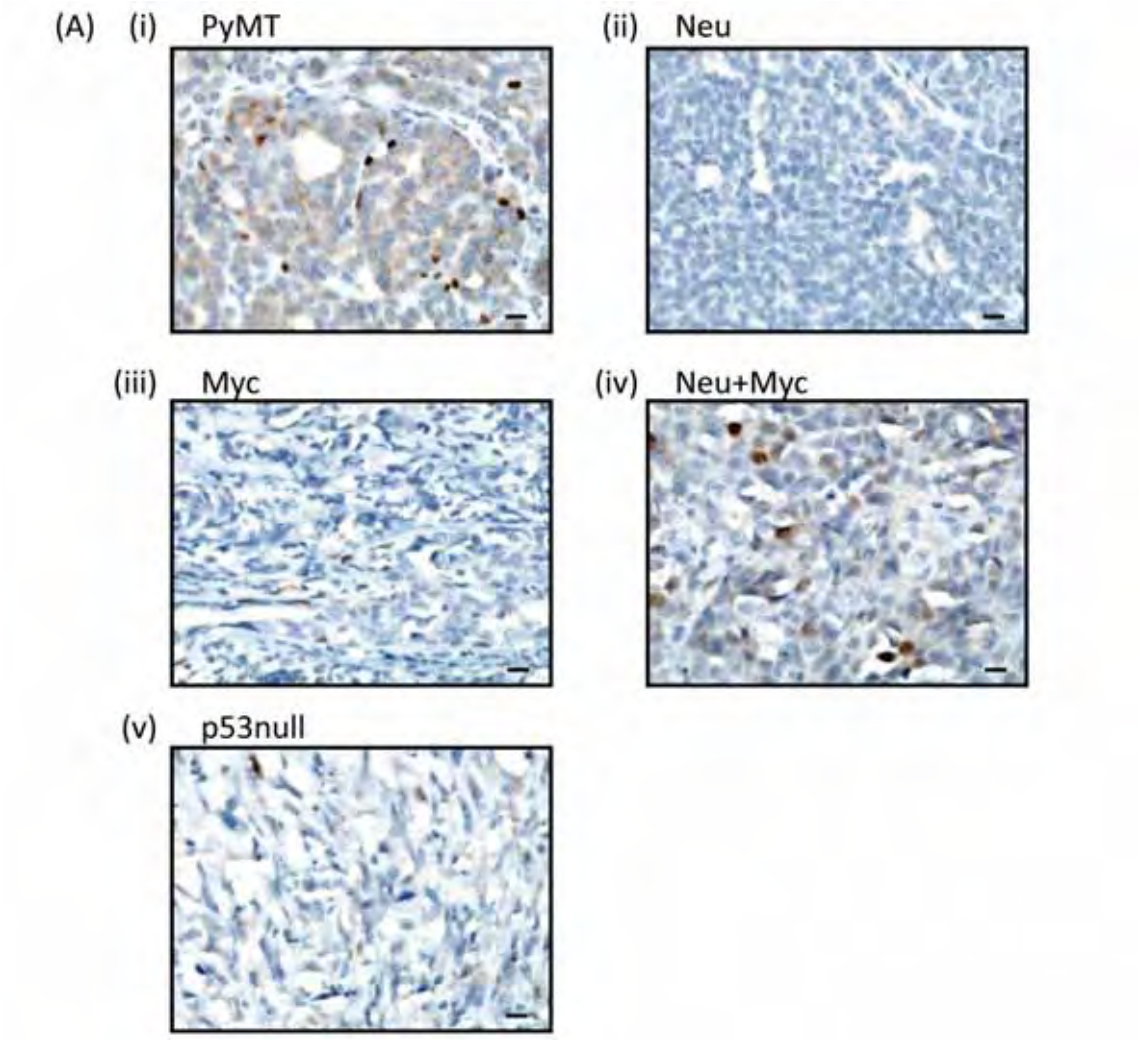


Figure 6-10. Id4 expression patterns in a variety of mouse models of breast cancer.
(A) Id4 expression was detected by IHC (brown staining) in the tumour sections from Pymt (i), Neu (ii), Myc (iii), Neu+Myc (iv), and p53 null (v) models of breast cancer. Scale bar=10μm. (Figure continues on next page)

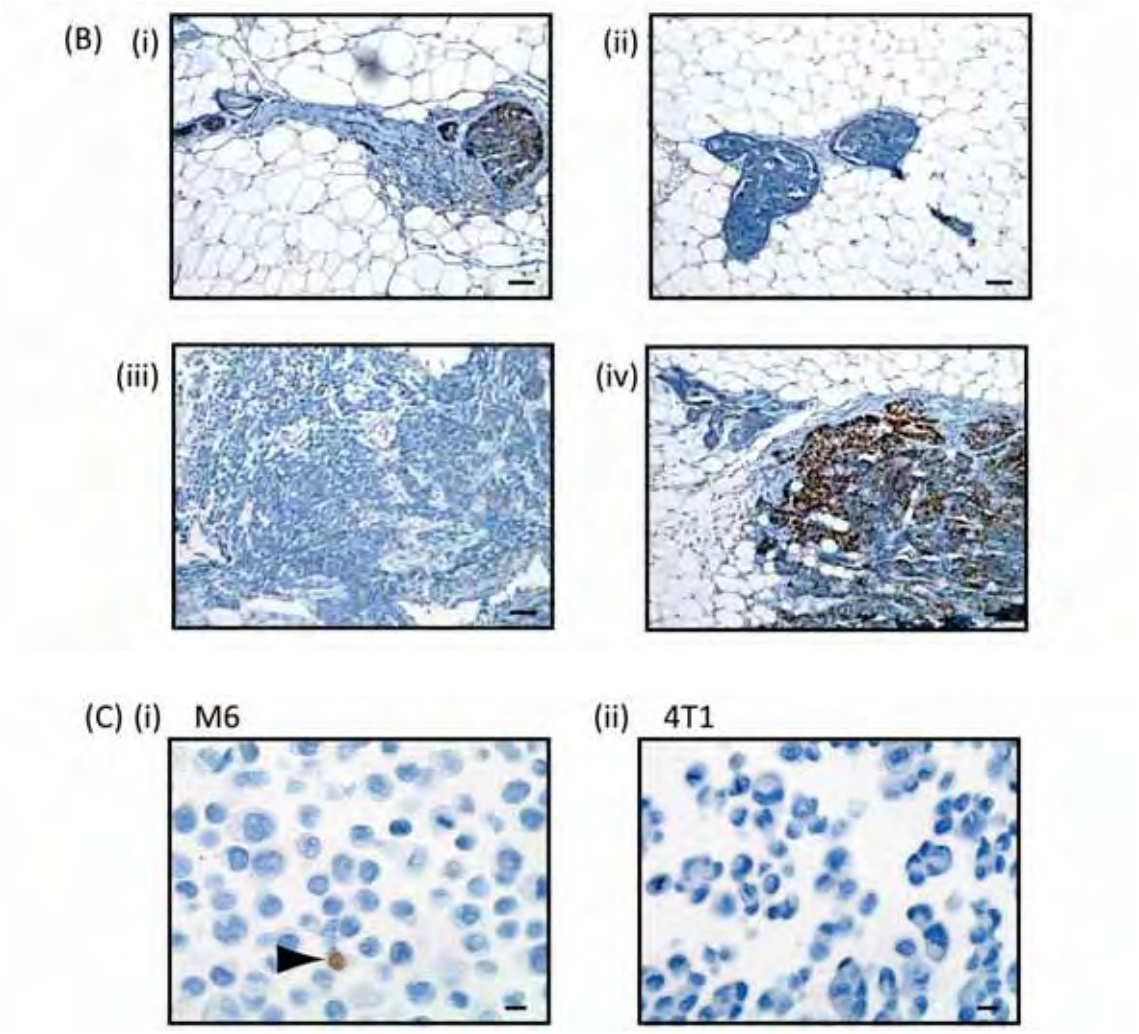


Figure 6–6. Continued. Id4 expression patterns in a variety of mouse models of breast cancer.
(B) Id4 IHC in the C3TAg model, variable Id4 expression in DCIS (i-ii) and carcinomas (iii-iv). Scale bar=50μm. (C) Id4 IHC on M6 (i) and 4T1 (ii) cell lines. Rare positive cell indicated with an arrowhead. Scale bar=10μm.

6.3 Discussion:

The results described in this chapter demonstrate, using a highly specific anti-Id4 antibody, that Id4 protein expression is limited to the triple-negative/basal-like and Her2-amplified subtypes of breast cancer. Id4 protein expression correlates with mRNA expression in a large number of breast cancer cell lines, albeit in a non-linear manner, and correlates strongly with the level of promoter methylation in the small number of cell lines analysed. However, Id4 protein detection appears to be the most robust method for inferring functional Id4 expression. These data suggest that Id4 expression is regulated at several levels including promoter methylation, and also at a post transcriptional level by a currently unknown mechanism. I have also shown that within the basal-like subtype of breast cancer higher Id4 protein and mRNA levels associate with better outcome, and that this may be due in part to an

improved response to chemotherapy. High Id4 expression was also associated with a high proliferative index and reduced lymph node metastasis within the basal-like subtype of breast cancer. The increased proliferative index is surprising since high proliferative index correlates with poor patient outcome in the cohort as a whole (Sandra O'Toole personal communication). Interestingly, within basal-like breast cancer patients those with Brca-1 mutations also have higher proliferation and better response to chemotherapy, in particular platinum based chemotherapy (Byrski et al., 2009; Gonzalez-Angulo et al., 2011). The low incidence of lymph node metastasis in Id4 high patients is in line with improved patient outcome.

A recent publication has confirmed our observations that Id4 protein expression is associated with the triple negative subtype of breast cancer (Wen et al., 2012). This group used the same monoclonal anti-Id4 antibody that we used and, while the staining pattern appeared similar to what we saw, they did not find any association between Id4 expression and patient outcome. This could be due to differences in the composition and treatment of their cohort or could be due to different analysis methods. In their study they divided the patients based on a percentage of Id4 positive cells greater than 5%, whereas we used an H-score of greater than 10. They also saw a strong association between Id4 expression and higher proliferation but saw no association with lymph node metastasis. Interestingly they did see an association with the basal cell marker CK14 suggesting that Id4 expression could be marking a basal/myoepithelial subtype of triple negative breast cancer.

The exact role of Id4 in breast cancer is not well understood and this is partly due to contradictory reports in the literature. My results raise further questions about some of the Id4 literature and in some cases directly contradict aspects of previously published results. As I have analysed Id4 expression at multiple levels and have a very robust antibody for detection of Id4 protein I am very confident that my results are correct. Furthermore, on another level I think my research can bring together some of the disparate observations about Id4 in breast cancer. I will now go on to discuss these points in more detail.

Firstly, a number of studies have analysed the role of Id4 in cell lines that I have shown not to express any detectable Id4 protein and that also have a hypermethylated Id4 promoter region. For example the studies that showed that Id4 could negatively regulate BRCA1 were performed in SK-BR3 and MCF-7 cells, and the study showing mutant p53 could activate Id4 expression was performed in MDA-MB-231 cells (Beger et al., 2001; Fontemaggi et al., 2009; Heyn et al., 2011). One possible explanation for this discrepancy is that modulating Id4 levels even below the sensitivity of protein detection by western blot can still have dramatic effects

on the phenotype of these cells. If this is the case, it raises the question as to why some cells would express Id4 to much higher levels if this does not affect Id4 function. Cell line differences can also be accounted for by differences in culturing conditions, and genetic drift. All of our cell lines were sourced from ATCC and have been validated using microsatellite markers. Another possible explanation is that Id4 promoter methylation and expression level is regulated in a dynamic manner and that in the specific experimental conditions used in these studies Id4 is expressed by these cell lines. A third explanation, which is not mutually exclusive from the first, is that Id4 is expressed by a rare cell (cancer stem cell?) population within the cell line and that the effects of Id4 are modulated through this cell population. Using the 4T1 transplantable cell line mouse model of breast cancer, Id4 has been suggested to play a role in the cancer stem cell phenotype (Park et al., 2011). Again, I did not see expression of Id4 in this cell line. I have, however, seen rare Id4 positive cells in some other cell lines such as the M6 cell line.

Some of the discrepancies between my research and others could be due to differences in the methodology of analysis. For example, one group has suggested an enrichment of Id4 mRNA expression in basal-like breast cancer patients. Our analysis of publically available datasets does not support this, our results actually showed that the Her2 subset had the highest average Id4 mRNA expression level (Figure 6-7A) (Turner et al., 2007). This could be explained by differences in methodologies (RT-PCR vs microarray) and could be due to differences in the regions of Id4 interrogated by the different primers or array probes. In cell lines we do see some enrichment of Id4 expression in the basal breast cancer cell lines using Taqman probes. We also show that there is an association between mRNA expression and protein level in breast cancer cell lines, however the observation that this is not a linear relationship suggests that Id4 expression is regulated in a post-transcriptional manner. The different Id4 transcripts with variable length 3'UTRs that allow for differential regulation by microRNAs could be one mechanism affecting the post transcriptional regulation of Id4 (van Cruchten et al., 1998). Analysis of patient samples with a more comprehensive set of primers or by northern blot to determine the relative abundance of the different Id4 transcripts could indicate whether 3'UTR length correlates with Id4 protein expression level.

There also appears to be a discrepancy between our results and others in regard to Id4 promoter methylation in cell lines. In our study, both MCF-7 and MDA-MB-231 cells were heavily methylated however MDA-MB-231 cells have previously been reported to have an unmethylated Id4 promoter and MCF7 cells were shown to have a partially methylated

promoter (Noetzel et al., 2008; Umetani et al., 2005). The discrepancy in methylation results again is likely due to differences in technique. With our Sequenom analysis covering a greater range, we can detect methylation at a range of sites not analysed by these other groups. This is supported by a recent study comparing methodologies of detecting Id4 methylation in acute myeloid leukaemia, which determined that the Sequenom MassArray system was far more accurate than methylation specific PCR (Claus et al., 2012). It is also curious that the same group that saw low levels of Id4 methylation in the MDA-MB-231 cells also saw relatively high mRNA expression in this cell line (Umetani et al., 2005), again this is the exact opposite of my results. If they are not detecting all cases that have a methylated Id4 promoter, this raises doubts about their larger patient studies associating Id4 methylation with outcome (Umetani et al., 2005). It would be interesting to further analyse the association between Id4 promoter methylation and outcome by dividing the patient cohorts based on breast cancer subtype. As Id4 protein expression is limited to ER negative subtypes of breast cancer, methylation studies should be limited to these subtypes. Furthermore, Id4 may mark a subtype of basal-like breast cancer that is derived from a different cell of origin to Id4 negative basal-like breast cancers.

A problem with using methylation of a single gene and associating it with patient outcomes is that cancer genomes are generally hypermethylated and therefore it can be misleading to infer a tumour suppressor nature to a gene just based on its methylation status. It is therefore possible that Id4 promoter methylation is a read-out of general hypermethylation of the cancer genome that frequently occurs across large regions of the genome (Coolen et al., 2010).

Despite the differences in cell line results mentioned above the main results from my studies on patient cohorts agree with the previous studies on patient cohorts. I have shown that Id4 protein expression correlates with improved patient survival and reduced lymph node metastasis. The previous studies demonstrated the same link between Id4 and survival except by utilising Id4 silencing via promoter methylation instead of protein expression to demonstrate the link between Id4 and patient outcome and lymph node metastasis (Noetzel et al., 2008; Umetani et al., 2005). Our results further suggest that there is a direct association between reduced Id4 protein expression and poor outcome in breast cancer patients, suggesting that Id4 associates with a less aggressive subset of basal-like breast cancer.

This chapter does not address the functional role of Id4 in breast cancer, rather, it correlates Id4 expression with various characteristics of breast cancer. The results from Chapter 5 however suggest that Id4 does have a functional role in tumourigenesis that promotes the transformation of the Comma-Dβ cell line. Furthermore using the knowledge that I have

obtained about its expression patterns in breast cancer cell lines and mouse models of breast cancer I now have the tools available to better understand the function of Id4 in breast cancer. Firstly we could investigate whether overexpression of Id4 is sufficient to transform normal mammary epithelial cells, similar to what I saw in the Comma-Dβ mouse system. Moreover, we have identified cell lines with high Id4 expression that can be used to analyse the effect of knocking down Id4 expression on cell phenotype. These knockdown experiments could be analysed for the expression of Id4 target candidate genes such as BRCA1 and the novel Id4 regulated genes I identified in Chapter 5 such as MMP9. Finally I have also identified a mouse model of basal-like breast cancer, the C3TAg model, in which a proportion of tumours show high Id4 expression and this model is currently being crossed to our Id4 null mice. We will use the C3TAg model to determine the effect of Id4 deletion on tumour incidence and latency. Whether Id4 plays a role in promoting breast cancer through its suppression of BRCA1 and its promotion of angiogenesis will need to be proven in clinical samples and more relevant cell lines (Beger et al., 2001; Fontemaggi et al., 2009).

One of the most exciting findings of this study is that high Id4 expression is clearly associated with better prognosis and that this could be due to improved response to chemotherapy. This is very clinically relevant since predicting response to therapy means that Id4 expression levels could be used to tailor therapy for individual patients. Of course, these findings need to be corroborated in a larger cohort before this could be translated into clinical practice. We are currently obtaining samples from a cohort of patients that have had biopsies taken pre- and post- neo adjuvant chemotherapy as well as during surgical resection (Schneider-Kolsky et al., 2010). This will allow us to better understand the association between Id4 expression and the response to chemotherapy and it will also allow us to determine whether Id4 expression level is modulated in frequency or intensity by chemotherapy. Of note, the Id4 IHC procedure is very robust and could easily be incorporated into current IHC and pathological examination of patient samples. Our results in breast cancer however stand in stark contrast to what was observed in glioma cells by Jeon et al (Jeon et al., 2011). In glioma cells high Id4 expression was associated with increased chemotherapy resistance through the repression of mIR-9, activation of SOX2 and the subsequent activation of ABCC3 and ABCC6 drug efflux transporters. These ABC drug efflux transporters were not up regulated in the Id4 overexpression study in Comma-Dβ cells suggesting that this mechanism might be specific to glioma cells. Other ABC transporters were modulated (both up and down) in the Comma-Dβ profiling studies, the relevance of these would need to be validated in human breast cancer

cell lines and ideally primary breast cancer samples to determine if they are playing a role in response to chemotherapy drugs.

Taken together these results demonstrate that Id4 protein expression is associated with ER negative breast cancer and that high Id4 expression is associated with better patient outcome possibly due to improved response to chemotherapy. Whether Id4 is directly modulating the phenotype of breast cancer or is a surrogate marker associated with outcome requires further confirmation. Our results from Chapter 5 however suggest that Id4 does influence the aetiology of ER negative breast cancer and can act as an oncogene *in vivo* and thus Id4 is not a tumour suppressor gene as has been postulated (Noetzel et al., 2008; Umetani et al., 2005). This does bring together the results of the majority of studies looking into Id4 function in breast cancer and demonstrates that examining its expression at the protein level gives the most meaningful insights into its function. So, while Id4 cannot be described as a tumour suppressor gene, it does appear to promote a less aggressive form of breast cancer. With further confirmation, these findings could lead to better patient care where Id4 expression could be used as a prognostic biomarker for patient outcome and possibly also as a predictive biomarker of response to chemotherapy.

Chapter 7. Discussion

Prior to the studies described in this thesis, the role for Id4 during mammary gland development had not been investigated and its role in breast cancer was controversial. The results presented here show that Id4 is a critical regulator of mammary gland development through its control of differentiation, proliferation and extracellular matrix (ECM) remodelling pathways. In addition, they also go some way to clarifying the role of Id4 in breast cancer by showing that Id4 protein expression clearly associates with the Her2 and basal-like subtypes of breast cancer and, within these subsets, high Id4 expression significantly correlates with improved patient survival. Despite high Id4 expression associating with improved patient survival, our results demonstrate that Id4 nonetheless has the capacity to promote tumour progression.

This work has shown that studying the role of transcription factors such as Id4 in the developmental context can lead to a better understanding of important aspects of breast cancer phenotype. Furthermore, by understanding the function of Id4 we can also gain a greater understanding of fundamental processes of mammary gland biology. I will now discuss in more detail some of the implications of my results in the context of mammary gland biology before going on to discuss how these pathways influence breast cancer aetiology.

The transcription factors that regulate myoepithelial cell differentiation in the mammary gland are not well understood. ΔNp63 is known to play a role, however, unlike luminal cell differentiation, the factors that regulate the differentiation from common progenitors to myoepithelial progenitors and terminally differentiated myoepithelial cells is still a mystery (Li et al., 2008; Visvader, 2009; Yalcin-Ozuysal et al., 2010). When we discovered that Id4 expression was restricted to the myoepithelial cells, we hypothesised that it could be a key regulator of myoepithelial cell differentiation. However, from my studies analysing the ducts of Id4 null mice we discovered a normal architecture of the myoepithelial cell layer and normal expression of the myoepithelial cell markers p63, αSMA, Keratin-5 and Keratin-14. Moreover, evidence from both *in vivo* and *in vitro* assays demonstrated that Id4 was in fact blocking luminal differentiation, with Id4 loss promoting luminal progenitor differentiation *in vivo* and Id4 overexpression blocking luminal differentiation *in vitro*. Loss of Id4 also led to a reduction in stem cell activity *in vivo*, suggesting that in its absence precocious luminal differentiation depletes the stem cell pool. It was demonstrated that ΔNp63 is required for stem cell maintenance in the skin (Senoo et al., 2007). Id4 and p63 co-localise to the myoepithelial cell

layer in the mammary gland, in which the epithelium is derived from the skin during embryogenesis. It is thus possible that Id4 and ΔNp63 co-operate to maintain the stem cell pool and block luminal differentiation. As both p63 and Id4 are expressed in all myoepithelial cells, this fits with a model proposed by John Stingl that all myoepithelial cells can be induced to have stem cell activity (personal communication). This would suggest that, unlike luminal epithelial cells, myoepithelial cells exist in a plastic state, and this allows for the transition from stem/progenitor to differentiated myoepithelial and back again rather than a linear hierarchy.

When examining the normal expression pattern of Id4 during development it was noted that high Id4 protein expression was detected both in the highly proliferative stem cell enriched TEBS and also in the quiescent ducts at oestrus when stem cell activity is at its lowest (Joshi et al., 2010; Kouros-Mehr et al., 2006). However when considered in more detail this potential dichotomy can be explained. While oestrus is when the least stem cells are present, it is also when the proliferative stage of the oestrus cycle is initiated. This proliferation is presumably initially activated in the stem/common progenitor cells followed by the more differentiated progenitor cells, although this needs to be proven experimentally. At diestrus, epithelial cells switch from a proliferative program to an apoptotic program (Khokha and Werb, 2011). Therefore, while stem cell number may be at its lowest at oestrus it is at this point that proliferation and expansion begins. It is also interesting to note a similarity between TEBs and the mammary gland at oestrus: when analysing wholemount images of mammary glands at the different stages of the oestrus cycle (Joshi et al., 2010; Khokha and Werb, 2011), I have observed that the terminal ducts appear to have taken on a mini-TEB like morphology at oestrus (unpublished data).

The high expression of Id4 in TEBs and during oestrus in the ducts suggests that it is a downstream mediator of oestrogenic effects in the myoepithelial cells. The expression of Id4 appears to be highest when levels of oestrogen are also at their highest and the Id4 null mouse has a similar, but less severe pubertal ductal phenotype to the ER null mouse and also the amphiregulin (AREG) null mouse (the main paracrine signalling molecule of oestrogen signalling). Despite Id4 being a progesterone target gene it does not appear to be critical to progesterone signalling (Fernandez-Valdivia et al., 2008). Progesterone is not required for pubertal development but is critical for pregnancy induced lobular alveolar development (Lydon et al., 1995). My results suggest that Id4 is not necessary for this process. In my studies, Id4 expression in the mammary gland was unaffected by loss of PR. Additionally, progesterone treatment of ovariectomised mice leads to proliferation in the Id4 negative

(luminal cells) but not the (Id4 positive) myoepithelial cells (Fernandez-Valdivia et al., 2008). These data combined with my results suggest that oestrogen is the key regulator of Id4 both during pubertal development and through the oestrus cycle. Intriguingly, since Id4 expression is limited to ER negative subtypes of breast cancer then it appears that Id4 expression must become independent of oestrogen signalling during tumourigenesis.

The mRNA profiling data from Id4 modulation in the Comma-Dβ cell line has resulted in a greater appreciation of the dynamic link between the epithelium and stroma of the mammary gland. The Id4 expressing myoepithelial cells are located between the luminal epithelial cells and the ECM, and play an important role in the regulation of signalling between the two (Fata et al., 2004). The dynamic state of the mammary gland requires the intimate cooperation between epithelial cell proliferation/apoptosis with ECM remodelling and changes in adipocity. My profiling results suggest that Id4 plays a role in regulating ECM stability. Id4 up regulation suppressed the expression of a number of key remodelling enzymes including MMP2, MMP3 and MMP9 (Fata et al., 1999; Hens et al., 2009; Laffin et al., 2008; Thompson et al., 1994; Wiseman et al., 2003). ECM remodelling and metalloproteinases are important both for the structural changes within the mammary gland as well as in regulating the release of sequestered growth factors such as AREG and TGF-β. In the bone marrow the MMPs have also been shown to regulate the transition of a quiescent stem cell niche to a proliferative one (Sneddon and Werb, 2007). It therefore appears that high Id4 expression would inhibit the ability of mammary epithelial cells to degrade the ECM and expand out into the surrounding tissue. As high Id4 expression is seen in the cap cells of the highly proliferative and motile TEBs during development this was somewhat surprising. However expression of these factors by TEBs may not be required, since immune cells such as eosinophils, macrophages and mast cells orchestrate this ECM remodelling in a tightly regulated manner (Coussens and Pollard, 2011; Gouon-Evans et al., 2000; Lilla et al., 2009). In particular, during pubertal development these immune cells control ECM remodelling in a manner that promotes an organised mammary ductal invasion of the fat pad with regular bifurcations to generate the mature ductal tree (Coussens and Pollard, 2011; Gouon-Evans et al., 2002). During development, the ECM must be remodelled in a way that allows for directional invasion at the TEB while constraining the mature ducts that form behind it and the specific localisation of the different immune cell subsets as well as the proteases, and growth factors that they secrete, regulate this process (Coussens and Pollard, 2011; Hinck and Silberstein, 2005).

During the oestrus cycle regular side branching occurs, with tertiary side branches appearing following oestrus and reaching their maximum size at diestrus. Following diestrus in the absence of pregnancy, organised apoptosis and tissue remodelling then returns the mammary gland to its oestrus state (Ferguson et al., 1992; Khokha and Werb, 2011; Lilla et al., 2009). MMP3 is a key regulator of this process (Wiseman et al., 2003), and my profiling results, combined with the observation that levels of Id4 change during the oestrus cycle, suggest that Id4 may be a master regulator of this process. Reduced branching was also observed in the Id4 null mouse during pubertal development, however, the regulation of tertiary branching during the oestrus cycle has not yet been examined in the Id4 null mouse. The above mentioned studies combined with my results suggest the following model for Id4 action during the oestrus cycle. It appears that the proliferative part of the oestrus cycle is initiated around oestrus when Id4 is highest, then as Id4 levels drop, MMP and ECM remodelling proteinases rise allowing for the expansion of the epithelial compartment. Following diestrus, the levels of Id4 again increase reducing the levels of proteinases and returning the mammary gland to its pre-oestrus structure. Id4 expression was also shown to be necessary for Comma-Dβ cells to proliferate *in vitro* and to block luminal differentiation. Therefore one would expect induction of Id4 to allow for stem/myoepithelial cell proliferation, and then as Id4 levels drop luminal differentiation would be enhanced followed by proliferation and expansion of the luminal/alveolar cell compartments.

The role of Id4 in regulating differentiation pathways in other tissues, including during neural development and glioblastoma, have been described in a number of elegant studies (Bedford et al., 2005; Kondo and Raff, 2000; Samanta and Kessler, 2004; Yun et al., 2004). I have shown here that, in a similar manner, Id4 regulates cell fate decisions during mammary epithelial cell differentiation, with Id4 blocking luminal differentiation *in vitro* and Id4 loss leading to precocious luminal differentiation *in vivo.* In addition, Id4 loss led to a reduction in stem cell activity *in vivo*. It is now clear that factors that regulate developmental and differentiation pathways are often deregulated in breast cancer. Dysregulation of ER was the first of these to be discovered and this has had a huge impact in the treatment of breast cancer through anti-oestrogen therapy. There is now a growing list of other developmental regulators including Hedgehog, FoxM1, Gata-3, and Brca-1 that are deregulated in cancer (Carr et al., 2012; Kouros-Mehr et al., 2008; Lim et al., 2009; O'Toole et al., 2011). These factors are mutated, epigenetically silenced or overexpressed in specific subtypes of breast cancer and, aside from ER, their dysregulation follows the common theme of inhibiting the terminal luminal differentiation of mammary epithelial cells. Through its inhibition of luminal differentiation,

Id4 can now be added to this list. High Id4 expression is significantly associated with increased proliferation both in breast cancer samples and in the highly proliferative cap cells located within the TEBs during mammary gland development. Promotion of breast cancer by Id4 could also be due in part to its suppression of Brca-1, as has been proposed by a number of groups (Beger et al., 2001; Gilbert et al., 2010; Wen et al., 2012). Other transcriptional regulators modulate tumourigenesis in this manner such as HMGA1 (Baldassarre et al., 2003). While Brca-1 was significantly down regulated by Id4 overexpression in the Comma-Dβ cell line, it was not one of the most highly regulated genes and its expression was unaffected by Id4 knockdown. Further validation of the link between Id4 expression and Brca-1 inhibition in relevant breast cancer cell lines such as the MDA-MB-468 cell line is required to determine whether Id4 regulates Brca-1 in cancer. These studies are currently being performed by our laboratory. While Brca-1 is an interesting potential target for Id4 its down regulation may be a consequence of an earlier blockade in luminal differentiation rather than the specific mechanism through which Id4 blocks luminal differentiation.

The results obtained by transcription profiling the Comma-Dβ cell line have expanded our understanding of the transcriptional networks controlled by Id4 in the developmental and cancer contexts. One of the striking examples is the elucidation of the role of Id4 in the regulation of numerous ECM related genes. Overexpression of Id4 in the Comma-Dβ cell line led to reduced expression of a number of ECM remodelling proteinases such as MMP2, MMP3 and MMP9. In cancer these enzymes through ECM remodelling regulate multiple pathways to promote invasion and metastasis, including the release of growth and angiogenic factors that can further promote tumour growth and metastasis (Chang and Werb, 2001; Egeblad and Werb, 2002). In addition several factors controlling the epithelial to mesenchymal transition (EMT) such as Snail, Twist1, Twist2, Zeb1, and Zeb2 were also significantly down regulated by Id4 overexpression. The down regulation of these factors could also be reducing the metastatic potential of the tumours that highly express Id4 (Foubert et al., 2010; Muschler and Streuli, 2010). Further validation of these results in relevant breast cancer cell lines and mouse models of breast cancer will give us additional insights into their role in breast cancer. These profiling results, however, do provide an explanation for the reduced incidence of lymph node metastasis and improved survival in patients with breast cancers that have high Id4 expression.

These experiments have demonstrated that Id4 is a regulator of multiple pathways within the mammary gland. The three main pathways that Id4 regulates are proliferation, differentiation, and ECM remodelling. These roles also help us develop models for how Id4 may contribute to

breast cancer phenotype. By promoting proliferation and inhibiting luminal differentiation Id4 can act as a tumour promoter, however, by inhibiting ECM remodelling Id4 may also inhibit metastatic tumour progression. Teasing out the impact of these different roles of Id4 on breast cancer aetiology will be challenging. However, overexpression and knockdown of its key transcriptional targets may give further insights into the relevance of the individual pathways controlled by Id4. These discoveries also help to bring together some of the seemingly contradictory functions ascribed to Id4 in the breast cancer literature, demonstrating that it may both promote tumourigenesis whilst still leading to improved patient outcome. While many of these discoveries may not be directly translational to the clinic, they will improve our understanding of a subset of basal-like and Her2 amplified breast cancers. Furthermore, as Id4 appears to be a predictive biomarker for chemotherapy within the basal-like subtype of breast cancer, confirming whether Id4 expression predicts susceptibility to chemotherapy in ER negative breast cancer is very clinically relevant. Thus, while the Id4 gene is not generally mutated or amplified in cancer, when it is aberrantly expressed it nonetheless may be a critical controller of cancer phenotype. Importantly by studying its expression pattern and function during mammary gland development I have shed light on its possible functions during breast carcinogenesis.

Chapter 8. References

Alani, R.M., Young, A.Z., and Shifflett, C.B. (2001). Id1 regulation of cellular senescence through transcriptional repression of p16/Ink4a. Proceedings of the National Academy of Sciences *98*, 7812-7816.

Allred, D.C., Mohsin, S.K., and Fuqua, S.A. (2001). Histological and biological evolution of human premalignant breast disease. Endocrine-Related Cancer *8*, 47-61.

Appolloni, I., Calzolari, F., Barilari, M., Terrile, M., Daga, A., and Malatesta, P. (2012). Antagonistic modulation of gliomagenesis by Pax6 and Olig2 in PDGF-induced oligodendroglioma. International Journal of Cancer, n/a-n/a.

Aslakson, C.J., and Miller, F.R. (1992). Selective Events in the Metastatic Process Defined by Analysis of the Sequential Dissemination of Subpopulations of a Mouse Mammary Tumor. Cancer Research *52*, 1399-1405.

Asselin-Labat, M.L., Sutherland, K.D., Barker, H., Thomas, R., Shackleton, M., Forrest, N.C., Hartley, L., Robb, L., Grosveld, F.G., van der Wees, J., Lindeman, G.J., and Visvader, J.E. (2007). Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat Cell Biol *9*, 201-209.

Asselin-Labat, M.L., Vaillant, F., Sheridan, J.M., Pal, B., Wu, D., Simpson, E.R., Yasuda, H., Smyth, G.K., Martin, T.J., Lindeman, G.J., and Visvader, J.E. (2010). Control of mammary stem cell function by steroid hormone signalling. Nature *465*, 798-802.

Badders, N.M., Goel, S., Clark, R.J., Klos, K.S., Kim, S., Bafico, A., Lindvall, C., Williams, B.O., and Alexander, C.M. (2009). The Wnt receptor, Lrp5, is expressed by mouse mammary stem cells and is required to maintain the basal lineage. PLoS One *4*, e6594.

Baldassarre, G., Battista, S., Belletti, B., Thakur, S., Pentimalli, F., Trapasso, F., Fedele, M., Pierantoni, G., Croce, C.M., and Fusco, A. (2003). Negative regulation of BRCA1 gene expression by HMGA1 proteins accounts for the reduced BRCA1 protein levels in sporadic breast carcinoma. Mol Cell Biol *23*, 2225-2238.

Barrett, Lindy E., Granot, Z., Coker, C., Iavarone, A., Hambardzumyan, D., Holland, Eric C., Nam, H.-s., and Benezra, R. (2011). Self-Renewal Does Not Predict Tumor Growth Potential in Mouse Models of High-Grade Glioma. Cancer Cell *21*, 11-24.

Bedford, L., Walker, R., Kondo, T., van CrÃ¼chten, I., King, E.R., and Sablitzky, F. (2005). Id4 is required for the correct timing of neural differentiation. Developmental Biology *280*, 386-395.

Beger, C., Pierce, L.N., Kruger, M., Marcusson, E.G., Robbins, J.M., Welcsh, P., Welch, P.J., Welte, K., King, M.C., Barber, J.R., and Wong-Staal, F. (2001). Identification of Id4 as a regulator of BRCA1 expression by using a ribozyme-library-based inverse genomics approach. Proc Natl Acad Sci U S A *98*, 130-135.

Beleut, M., Rajaram, R.D., Caikovski, M., Ayyanan, A., Germano, D., Choi, Y., Schneider, P., and Brisken, C. (2010). Two distinct mechanisms underlie progesterone-induced proliferation in the mammary gland. Proc Natl Acad Sci U S A *107*, 2989-2994.

Bellido, M., Aventin, A., Lasa, A., Estivill, C., Carnicer, M.J., Pons, C., Matias-Guiu, X., Bordes, R., Baiget, M., Sierra, J., and Nomdedeu, J.F. (2003). Id4 is deregulated by a t(6;14)(p22;q32)

chromosomal translocation in a B-cell lineage acute lymphoblastic leukemia. Haematologica *88*, 994-1001.

Bigelow, R.L., Williams, B.J., Carroll, J.L., Daves, L.K., and Cardelli, J.A. (2009). TIMP-1 overexpression promotes tumorigenesis of MDA-MB-231 breast cancer cells and alters expression of a subset of cancer promoting genes in vivo distinct from those observed in vitro. Breast Cancer Res Treat *117*, 31-44.

Blanks, R.G., Moss, S.M., McGahan, C.E., Quinn, M.J., and Babb, P.J. (2000). Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990-8: comparison of observed with predicted mortality. BMJ *321*, 665-669.

Bloom, H.J., and Richardson, W.W. (1957). Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer *11*, 359-377.

Booth, B.W., Boulanger, C.A., Anderson, L.H., Jimenez-Rojo, L., Brisken, C., and Smith, G.H. (2009). Amphiregulin mediates self-renewal in an immortal mammary epithelial cell line with stem cell characteristics. Exp Cell Res *316*, 422-432.

Borrello, M.G., Degl'Innocenti, D., and Pierotti, M.A. (2008). Inflammation and cancer: The oncogene-driven connection. Cancer Letters *267*, 262-270.

Bouras, T., Pal, B., Vaillant, F., Harburg, G., Asselin-Labat, M.L., Oakes, S.R., Lindeman, G.J., and Visvader, J.E. (2008). Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. Cell Stem Cell *3*, 429-441.

Braig, M., Lee, S., Loddenkemper, C., Rudolph, C., Peters, A.H.F.M., Schlegelberger, B., Stein, H., Dorken, B., Jenuwein, T., and Schmitt, C.A. (2005). Oncogene-induced senescence as an initial barrier in lymphoma development. Nature *436*, 660-665.

Brennan, M., Wilcken, N., French, J., Ung, O., and Boyages, J. (2005). Management of early breast cancer--the current approach. Aust Fam Physician *34*, 755-760.

Brisken, C., Heineman, A., Chavarria, T., Elenbaas, B., Tan, J., Dey, S.K., McMahon, J.A., McMahon, A.P., and Weinberg, R.A. (2000). Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. Genes Dev *14*, 650-654.

Brisken, C., Kaur, S., Chavarria, T.E., Binart, N., Sutherland, R.L., Weinberg, R.A., Kelly, P.A., and Ormandy, C.J. (1999). Prolactin controls mammary gland development via direct and indirect mechanisms. Dev Biol *210*, 96 - 106.

Brisken, C., and O'Malley, B. (2010). Hormone Action in the Mammary Gland. Cold Spring Harbor Perspectives in Biology *2*.

Britt, K.L., Kendrick, H., Regan, J.L., Molyneux, G., Magnay, F.A., Ashworth, A., and Smalley, M.J. (2009). Pregnancy in the mature adult mouse does not alter the proportion of mammary epithelial stem/progenitor cells. Breast Cancer Res *11*, R20.

Byrski, T., Gronwald, J., Huzarski, T., Grzybowska, E., Budryk, M., Stawicka, M., Mierzwa, T., Szwiec, M., Wisniowski, R., Siolek, M., Dent, R., Lubinski, J., and Narod, S. (2009). Pathologic complete response rates in young women with BRCA1-positive breast cancers after neoadjuvant chemotherapy. J Clin Oncol *28*, 375-379.

Cai, L., Morrow, E.M., and Cepko, C.L. (2000). Misexpression of basic helix-loop-helix genes in the murine cerebral cortex affects cell fate choices and neuronal survival. Development *127*, 3021-3030.

Carr, J.R., Kiefer, M.M., Park, H.J., Li, J., Wang, Z., Fontanarosa, J., DeWaal, D., Kopanja, D., Benevolenskaya, E.V., Guzman, G., and Raychaudhuri, P. (2012). FoxM1 Regulates Mammary Luminal Cell Fate. Cell Reports.

Castro, M., Grau, L., Puerta, P., Gimenez, L., Venditti, J., Quadrelli, S., and Sanchez-Carbayo, M. (2010). Multiplexed methylation profiles of tumor suppressor genes and clinical outcome in lung cancer. J Transl Med *8*, 86.

Chakrabarti, R., Wei, Y., Romano, R.-A., DeCoste, C., Kang, Y., and Sinha, S. (2012). Elf5 Regulates Mammary Gland Stem/Progenitor Cell Fate by Influencing Notch Signaling. Stem Cells, N/A-N/A.

Chan, A.S., Tsui, W.Y., Chen, X., Chu, K.M., Chan, T.L., Li, R., So, S., Yuen, S.T., and Leung, S.Y. (2003). Downregulation of ID4 by promoter hypermethylation in gastric adenocarcinoma. Oncogene *22*, 6946-6953.

Chang, C., and Werb, Z. (2001). The many faces of metalloproteases: cell growth, invasion, angiogenesis and metastasis. Trends in Cell Biology *11*, S37-S43.

Chaudhary, J., Johnson, J., Kim, G., and Skinner, M.K. (2001). Hormonal regulation and differential actions of the helix-loop-helix transcriptional inhibitors of differentiation (Id1, Id2, Id3, and Id4) in Sertoli cells. Endocrinology *142*, 1727-1736.

Chen, S.-S., Claus, R., Lucas, D.M., Yu, L., Qian, J., Ruppert, A.S., West, D.A., Williams, K.E., Johnson, A.J., Sablitzky, F., Plass, C., and Byrd, J.C. (2011). Silencing of the inhibitor of DNA binding protein 4 (ID4) contributes to the pathogenesis of mouse and human CLL. Blood *117*, 862-871.

Chen, Z., Trotman, L.C., Shaffer, D., Lin, H.-K., Dotan, Z.A., Niki, M., Koutcher, J.A., Scher, H.I., Ludwig, T., Gerald, W., Cordon-Cardo, C., and Paolo Pandolfi, P. (2005). Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nature *436*, 725-730.

Cheng, X., Wang, Y., He, Q., Qiu, M., Whittemore, S.R., and Cao, Q. (2007). Bone Morphogenetic Protein Signaling and Olig1/2 Interact to Regulate the Differentiation and Maturation of Adult Oligodendrocyte Precursor Cells. Stem Cells *25*, 3204-3214.

Chu, E.Y., Hens, J., Andl, T., Kairo, A., Yamaguchi, T.P., Brisken, C., Glick, A., Wysolmerski, J.J., and Millar, S.E. (2004). Canonical WNT signaling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis. Development *131*, 4819-4829.

Chua, A.C., Hodson, L.J., Moldenhauer, L.M., Robertson, S.A., and Ingman, W.V. (2010). Dual roles for macrophages in ovarian cycle-associated development and remodelling of the mammary gland epithelium. Development *137*, 4229-4238.

Clark, S.J., Harrison, J., Paul, C.L., and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. Nucleic Acids Research *22*, 2990-2997.

Ciarloni, L., Mallepell, S., and Brisken, C. (2007). Amphiregulin is an essential mediator of estrogen receptor alpha function in mammary gland development. Proc Natl Acad Sci U S A *104*, 5455-5460.

Claus, R., Wilop, S., Hielscher, T., Sonnet, M., Dahl, E., Galm, O., Jost, E., and Plass, C. (2012). A systematic comparison of quantitative high-resolution DNA methylation analysis and methylation-specific PCR. Epigenetics *7*, 0--1.

Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. Nat Med *17*, 313-319.

Cook, K.L., Shajahan, A.N., and Clarke, R. (2011). Autophagy and endocrine resistance in breast cancer. Expert Review of Anticancer Therapy *11*, 1283-1294.

Coolen, M.W., Statham, A.L., Gardiner-Garden, M., and Clark, S.J. (2007). Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. Nucleic Acids Research *35*, e119.

Coolen, M.W., Stirzaker, C., Song, J.Z., Statham, A.L., Kassir, Z., Moreno, C.S., Young, A.N., Varma, V., Speed, T.P., Cowley, M., Lacaze, P., Kaplan, W., Robinson, M.D., and Clark, S.J. (2010). Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. Nat Cell Biol *12*, 235-246.

Coussens, L.M., and Pollard, J.W. (2011). Leukocytes in mammary development and cancer. Cold Spring Harb Perspect Biol *3*.

Coussens, L.M., Tinkle, C.L., Hanahan, D., and Werb, Z. (2000). MMP-9 Supplied by Bone Marrow Derived Cells Contributes to Skin Carcinogenesis. Cell *103*, 481-490.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., Graf, S., Ha, G., Haffari, G., Bashashati, A., Russell, R., McKinney, S., Caldas, C., Aparicio, S., Brenton, J.D., Ellis, I., Huntsman, D., Pinder, S., Purushotham, A., Murphy, L., Bardwell, H., Ding, Z., Jones, L., Liu, B., Papatheodorou, I., Sammut, S.J., Wishart, G., Chia, S., Gelmon, K., Speers, C., Watson, P., Blamey, R., Green, A., Macmillan, D., Rakha, E., Gillett, C., Grigoriadis, A., di Rinaldis, E., Tutt, A., Parisien, M., Troup, S., Chan, D., Fielding, C., Maia, A.T., McGuire, S., Osborne, M., Sayalero, S.M., Spiteri, I., Hadfield, J., Bell, L., Chow, K., Gale, N., Kovalik, M., Ng, Y., Prentice, L., Tavare, S., Markowetz, F., Langerod, A., Provenzano, E., and Borresen-Dale, A.L. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature.

Danielson, K.G., Oborn, C.J., Durban, E.M., Butel, J.S., and Medina, D. (1984). Epithelial mouse mammary cell line exhibiting normal morphogenesis in vivo and functional differentiation in vitro. Proc Natl Acad Sci U S A *81*, 3756-3760.

Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987-1000.

de Candia, P., Akram, M., Benezra, R., and Brogi, E. (2006). Id4 messenger RNA and estrogen receptor expression: inverse correlation in human normal breast epithelium and carcinoma. Hum Pathol *37*, 1032-1041.

de Candia, P., Benezra, R., and Solit, D.B. (2004). A role for Id proteins in mammary gland physiology and tumorigenesis. Adv Cancer Res *92*, 81-94.

de Visser, K.E., Korets, L.V., and Coussens, L.M. (2005). De novo carcinogenesis promoted by chronic inflammation is B lymphocyte dependent. Cancer Cell *7*, 411-423.

DeNardo, D.G., Barreto, J.B., Andreu, P., Vasquez, L., Tawfik, D., Kolhatkar, N., and Coussens, L.M. (2009). CD4+ T Cells Regulate Pulmonary Metastasis of Mammary Carcinomas by Enhancing Protumor Properties of Macrophages. Cancer Cell *16*, 91-102.

Desprez, P.-Y., Sumida, T., and Coppé, J.-P. (2003). Helix-Loop-Helix Proteins in Mammary Gland Development and Breast Cancer. Journal of Mammary Gland Biology and Neoplasia *8*, 225-239.

Deugnier, M.A., Faraldo, M.M., Teuliere, J., Thiery, J.P., Medina, D., and Glukhova, M.A. (2006). Isolation of mouse mammary epithelial progenitor cells with basal characteristics from the Comma-Dbeta cell line. Dev Biol *293*, 414-425.

Donehower, L.A. (1996). The p53-deficient mouse: a model for basic and applied cancer studies. Seminars in Cancer Biology *7*, 269-278.

Dong, J., Huang, S., Caikovski, M., Ji, S., McGrath, A., Custorio, M.G., Creighton, C.J., Maliakkal, P., Bogoslovskaia, E., Du, Z., Zhang, X., Lewis, M.T., Sablitzky, F., Brisken, C., and Li, Y. (2011). ID4 regulates mammary gland development by suppressing p38MAPK activity. Development *138*, 5247-5256.

Egeblad, M., and Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. Nat Rev, Cancer *2*, 161 - 174.

Ewan, K.B., Shyamala, G., Ravani, S.A., Tang, Y., Akhurst, R., Wakefield, L., and Barcellos-Hoff, M.H. (2002). Latent Transforming Growth Factor-[beta] Activation in Mammary Gland: Regulation by Ovarian Hormones Affects Ductal and Alveolar Proliferation. The American Journal of Pathology *160*, 2081-2093.

Faraldo, M.M., Teuliere, J., Deugnier, M.A., Taddei-De La Hosseraye, I., Thiery, J.P., and Glukhova, M.A. (2005). Myoepithelial cells in the control of mammary development and tumorigenesis: data from genetically modified mice. J Mammary Gland Biol Neoplasia *10*, 211-219.

Fata, J., Werb, Z., and Bissell, M. (2004). Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes. Breast Cancer Res *6*, 1 - 11.

Fata, J.E., Kong, Y.-Y., Li, J., Sasaki, T., Irie-Sasaki, J., Moorehead, R.A., Elliott, R., Scully, S., Voura, E.B., Lacey, D.L., Boyle, W.J., Khokha, R., and Penninger, J.M. (2000). The Osteoclast Differentiation Factor Osteoprotegerin-Ligand Is Essential for Mammary Gland Development. Cell *103*, 41-50.

Fata, J.E., Leco, K.J., Moorehead, R.A., Martin, D.C., and Khokha, R. (1999). Timp-1 is important for epithelial proliferation and branching morphogenesis during mouse mammary development. Dev Biol *211*, 238 - 254.

Ferguson, J.E., Schor, A.M., Howell, A., and Ferguson, M.W. (1992). Changes in the extracellular matrix of the normal human breast during the menstrual cycle. Cell Tissue Res *268*, 167 - 177.

Ferlay, J., Forman, D., Mathers, C.D., and Bray, F. (2012). Breast and cervical cancer in 187 countries between 1980 and 2010. The Lancet *379*, 1390-1391.

Fernandez-Valdivia, R., Jeong, J., Mukherjee, A., Soyal, S.M., Li, J., Ying, Y., Demayo, F.J., and Lydon, J.P. (2009a). A mouse model to dissect progesterone signaling in the female reproductive tract and mammary gland. Genesis *48*, 106-113.

Fernandez-Valdivia, R., Mukherjee, A., Creighton, C.J., Buser, A.C., DeMayo, F.J., Edwards, D.P., and Lydon, J.P. (2008). Transcriptional response of the murine mammary gland to acute progesterone exposure. Endocrinology *149*, 6236-6250.

Fernandez-Valdivia, R., Mukherjee, A., Ying, Y., Li, J., Paquet, M., DeMayo, F.J., and Lydon, J.P. (2009b). The RANKL signaling axis is sufficient to elicit ductal side-branching and alveologenesis in the mammary gland of the virgin mouse. Dev Biol *328*, 127-139.

Fontemaggi, G., Dell'Orso, S., Trisciuoglio, D., Shay, T., Melucci, E., Fazi, F., Terrenato, I., Mottolese, M., Muti, P., Domany, E., Del Bufalo, D., Strano, S., and Blandino, G. (2009). The execution of the transcriptional axis mutant p53, E2F1 and ID4 promotes tumor neo-angiogenesis. Nat Struct Mol Biol *16*, 1086-1093.

Foubert, E., De Craene, B., and Berx, G. (2010). Key signalling nodes in mammary gland development and cancer. The Snail1-Twist1 conspiracy in malignant breast cancer progression. Breast Cancer Res *12*, 206.

Fujita, P.A., Rhead, B., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Cline, M.S., Goldman, M., Barber, G.P., Clawson, H., Coelho, A., Diekhans, M., Dreszer, T.R., Giardine, B.M., Harte, R.A., Hillman-Jackson, J., Hsu, F., Kirkup, V., Kuhn, R.M., Learned, K., Li, C.H., Meyer, L.R., Pohl, A., Raney, B.J., Rosenbloom, K.R., Smith, K.E., Haussler, D., and Kent, W.J. (2010). The UCSC Genome Browser database: update 2011. Nucleic Acids Research.

Gallego, M.I., Binart, N., Robinson, G.W., Okagaki, R., Coschigano, K.T., Perry, J., Kopchick, J.J., Oka, T., Kelly, P.A., and Hennighausen, L. (2001). Prolactin, Growth Hormone, and Epidermal Growth Factor Activate Stat5 in Different Compartments of Mammary Tissue and Exert Different and Overlapping Developmental Effects. Developmental Biology *229*, 163-175.

Gilbert, P.M., Mouw, J.K., Unger, M.A., Lakins, J.N., Gbegnon, M.K., Clemmer, V.B., Benezra, M., Licht, J.D., Boudreau, N.J., Tsai, K.K., Welm, A.L., Feldman, M.D., Weber, B.L., and Weaver, V.M. (2010). HOXA9 regulates BRCA1 expression to modulate human breast tumor phenotype. J Clin Invest *120*, 1535-1550.

Gomez Del Pulgar, T., Valdes-Mora, F., Bandres, E., Perez-Palacios, R., Espina, C., Cejas, P., Garcia-Cabezas, M.A., Nistal, M., Casado, E., Gonzalez-Baron, M., Garcia-Foncillas, J., and Lacal, J.C. (2008). Cdc42 is highly expressed in colorectal adenocarcinoma and downregulates ID4 through an epigenetic mechanism. Int J Oncol *33*, 185-193.

Gonzalez-Angulo, A.M., Timms, K.M., Liu, S., Chen, H., Litton, J.K., Potter, J., Lanchbury, J.S., Stemke-Hale, K., Hennessy, B.T., Arun, B.K., Hortobagyi, G.N., Do, K.-A., Mills, G.B., and Meric-Bernstam, F. (2011). Incidence and Outcome of BRCA Mutations in Unselected Patients with Triple Receptor-Negative Breast Cancer. Clinical Cancer Research *17*, 1082-1089.

Gouon-Evans, V., Lin, E.Y., and Pollard, J.W. (2002). Requirement of macrophages and eosinophils and their cytokines/chemokines for mammary gland development. Breast Cancer Res *4*, 155-164.

Gouon-Evans, V., Rothenberg, M.E., and Pollard, J.W. (2000). Postnatal mammary gland development requires macrophages and eosinophils. Development *127*, 2269 - 2282.

Green, J.E., Shibata, M.A., Yoshidome, K., Liu, M.L., Jorcyk, C., Anver, M.R., Wigginton, J., Wiltrout, R., Shibata, E., Kaczmarczyk, S., Wang, W., Liu, Z.Y., Calvo, A., and Couldrey, C. (2000). The C3(1)/SV40 T-antigen transgenic mouse model of mammary cancer: ductal epithelial cell targeting with multistage progression to carcinoma. Oncogene *19*, 1020-1027.

Greene, S.B., Gunaratne, P.H., Hammond, S.M., and Rosen, J.M. (2010). A putative role for microRNA-205 in mammary epithelial cell progenitors. J Cell Sci *123*, 606-618.

Grimm, S.L., Bu, W., Longley, M.A., Roop, D.R., Li, Y., and Rosen, J.M. (2006). Keratin 6 is not essential for mammary gland development. Breast Cancer Res *8*, R29.

Gross, S., Blanchetot, C., Schepens, J., Albet, S., Lammers, R., den Hertog, J., and Hendriks, W. (2002). Multimerization of the Protein-tyrosine Phosphatase (PTP)-like Insulin-dependent Diabetes Mellitus Autoantigens IA-2 and IA-2Î² with Receptor PTPs (RPTPs). Journal of Biological Chemistry *277*, 48139-48145.

Grumolato, L., Liu, G., Mong, P., Mudbhary, R., Biswas, R., Arroyave, R., Vijayakumar, S., Economides, A.N., and Aaronson, S.A. (2010). Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. Genes Dev *24*, 2517-2530.

Gu, B., Sun, P., Yuan, Y., Moraes, R.C., Li, A., Teng, A., Agrawal, A., RhÃ©aume, C., Bilanchone, V., Veltmaat, J.M., Takemaru, K.-I., Millar, S., Lee, E.Y.H.P., Lewis, M.T., Li, B., and Dai, X. (2009). Pygo2 expands mammary progenitor cells by facilitating histone H3 K4 methylation. The Journal of Cell Biology *185*, 811-826.

Guerra, C., Mijimolle, N., Dhawahir, A., Dubus, P., Barradas, M., Serrano, M., Campuzano, V., and Barbacid, M. (2003). Tumor induction by an endogenous K-ras oncogene is highly dependent on cellular context. Cancer Cell *4*, 111-120.

Guo, W., Keckesova, Z., Donaher, Joana L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zarrer-Hardi, U., Bell, G., Tam, Wai, L., Mani, Sendurai, A., van Oudenaarden, A., and Weinberg, Robert A. (2012). Slug and Sox9 Cooperatively Determine the Mammary Stem Cell State. Cell *148*, 1015-1028.

Gupta, P.B., Chaffer, C.L., and Weinberg, R.A. (2009). Cancer stem cells: mirage or reality? Nat Med *15*, 1010-1012.

Hanahan, D., and Weinberg, Robert A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646-674.

Harris, J., Stanford, P.M., Sutherland, K., Oakes, S.R., Naylor, M.J., Robertson, F.G., Blazek, K.D., Kazlauskas, M., Hilton, H.N., Wittlin, S., Alexander, W.S., Lindeman, G.J., Visvader, J.E., and Ormandy, C.J. (2006). Socs2 and elf5 mediate prolactin-induced mammary gland development. Mol Endocrinol *20*, 1177-1187.

Hatsell, S.J., and Cowin, P. (2006). Gli3-mediated repression of Hedgehog targets is required for normal mammary development. Development *133*, 3661-3670.

He, Y., Dupree, J., Wang, J., Sandoval, J., Li, J., Liu, H., Shi, Y., Nave, K.A., and Casaccia-Bonnefil, P. (2007). The Transcription Factor Yin Yang 1 Is Essential for Oligodendrocyte Progenitor Differentiation. Neuron *55*, 217-230.

Hens, J., Dann, P., Hiremath, M., Pan, T.-C., Chodosh, L., and Wysolmerski, J. (2009). Analysis of gene expression in PTHrP−/− mammary buds supports a role for BMP signaling and MMP2 in the initiation of ductal morphogenesis. Developmental Dynamics *238*, 2713-2724.

Hens, J.R., Dann, P., Zhang, J.-P., Harris, S., Robinson, G.W., and Wysolmerski, J. (2007). BMP4 and PTHrP interact to stimulate ductal outgrowth during embryonic mammary development and to inhibit hair follicle induction. Development *134*, 1221-1230.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S., Backlund, M.G., Yin, Y., Khramtsov, A.I., Bastein, R., Quackenbush, J., Glazer, R.I., Brown, P.H., Green, J.E., Kopelovich, L., Furth, P.A., Palazzo, J.P., Olopade, O.I., Bernard, P.S., Churchill, G.A., Van Dyke, T., and Perou, C.M. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol *8*, R76.

Heyn, H., Engelmann, M., Schreek, S., Ahrens, P., Lehmann, U., Kreipe, H., Schlegelberger, B., and Beger, C. (2011). MicroRNA miR-335 is crucial for the BRCA1 regulatory cascade in breast cancer development. Int J Cancer.

Higa, G.M., and Abraham, J. (2007). Lapatinib in the treatment of breast cancer. Expert Review of Anticancer Therapy *7*, 1183-1192.

Hinck, L., and Silberstein, G.B. (2005). Key stages in mammary gland development: the mammary end bud as a motile organ. Breast Cancer Res *7*, 245-251.

Hirai, S., Miwa, A., Ohtaka-Maruyama, C., Kasai, M., Okabe, S., Hata, Y., and Okado, H. (2012). RP58 controls neuron and astrocyte differentiation by downregulating the expression of Id1-4 genes in the developing cortex. EMBO J *31*, 1190-1202.

Hogg, K., Etherington, S.L., Young, J.M., McNeilly, A.S., and Duncan, W.C. (2010). Inhibitor of Differentiation (Id) Genes Are Expressed in the Steroidogenic Cells of the Ovine Ovary and Are Differentially Regulated by Members of the Transforming Growth Factor-β Family. Endocrinology *151*, 1247-1256.

Hollestelle, A., Elstrodt, F., Nagel, J.H.A., Kallemeijn, W.W., and Schutte, M. (2007). Phosphatidylinositol-3-OH Kinase or RAS Pathway Mutations in Human Breast Cancer Cell Lines. Molecular Cancer Research *5*, 195-201.

Hollestelle, A., Pelletier, C., Hooning, M., Crepin, E., Schutte, M., Look, M., Collee, J.M., Nieuwlaat, A., Dorssers, L.C.J., Seynaeve, C., Aulchenko, Y.S., Martens, J.W.M., Van Den Ouweland, A.M.W., and Weidhaas, J.B. (2011). Prevalence of the variant allele rs61764370 T>G in the 3'UTR of KRAS among Dutch BRCA1, BRCA2 and non-BRCA1/BRCA2 breast cancer families. Breast Cancer Research and Treatment *128*, 79-84.

Holzer, R.G., MacDougall, C., Cortright, G., Atwood, K., Green, J.E., and Jorcyk, C.L. (2003). Development and Characterization of a Progressive Series of Mammary Adenocarcinoma Cell Lines Derived from the C3(1)/SV40 Large T-antigen Transgenic Mouse Model. Breast Cancer Research and Treatment *77*, 65-76.

Hon, G.C., Hawkins, R.D., Caballero, O.L., Lo, C., Lister, R., Pelizzola, M., Valsesia, A., Ye, Z., Kuan, S., Edsall, L.E., Camargo, A.A., Stevenson, B.J., Ecker, J.R., Bafna, V., Strausberg, R.L., Simpson, A.J., and Ren, B. (2012). Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Research *22*, 246-258.

Htun, H., Holth, L.T., Walker, D., Davie, J.R., and Hager, G.L. (1999). Direct Visualization of the Human Estrogen Receptor α Reveals a Role for Ligand in the Nuclear Distribution of the Receptor. Molecular Biology of the Cell *10*, 471-486.

Huang, J., Vogel, G., Yu, Z., Almazan, G., and Richard, S.p. (2011). Type II Arginine Methyltransferase PRMT5 Regulates Gene Expression of Inhibitors of Differentiation/DNA

Binding Id2 and Id4 during Glial Cell Differentiation. Journal of Biological Chemistry *286*, 44424-44432.

Hudis, C.A. (2007). Trastuzumab - Mechanism of Action and Use in Clinical Practice. New England Journal of Medicine *357*, 39-51.

Hui, L., Bakiri, L., Stepniak, E., and Wagner, E.F. (2007). p38α: A Suppressor of Cell Proliferation and Tumorigenesis. Cell Cycle *6*, 2429-2433.

Ibarra, I., Erlich, Y., Muthuswamy, S.K., Sachidanandam, R., and Hannon, G.J. (2007). A role for microRNAs in maintenance of mouse mammary epithelial progenitor cells. Genes Dev *21*, 3238-3243.

Iyer, V., Klebba, I., McCready, J., Arendt, L.M., Betancur-Boissel, M., Wu, M.-F., Zhang, X., Lewis, M.T., and Kuperwasser, C. (2012). Estrogen promotes ER-negative tumor growth and angiogenesis through mobilization of bone marrow-derived monocytes. Cancer Research.

Jang, K.-S., Han, H.X., Paik, S.S., Brown, P.H., and Kong, G. (2006). Id-1 overexpression in invasive ductal carcinoma cells is significantly associated with intratumoral microvessel density in ER-negative/node-positive breast cancer. Cancer Letters *244*, 203-210.

Jen, Y., Manova, K., and Benezra, R. (1996). Expression patterns of Id1, Id2, and Id3 are highly related but distinct from that of Id4 during mouse embryogenesis. Dev Dyn *207*, 235-252.

Jeon, H.M., Jin, X., Lee, J.S., Oh, S.Y., Sohn, Y.W., Park, H.J., Joo, K.M., Park, W.Y., Nam, D.H., DePinho, R.A., Chin, L., and Kim, H. (2008). Inhibitor of differentiation 4 drives brain tumor-initiating cell genesis through cyclin E and notch signaling. Genes Dev *22*, 2028-2033.

Jeon, H.M., Sohn, Y.W., Oh, S.Y., Kim, S.H., Beck, S., Kim, S., and Kim, H. (2011). ID4 Imparts Chemoresistance and Cancer Stemness to Glioma Cells by Derepressing miR-9*-Mediated Suppression of SOX2. Cancer Res *71*, 3410-3421.

Jerry, D., Medina, D., and Butel, J. (1994). p53 Mutations in COMMA-D cells. In Vitro Cellular & Developmental Biology - Animal *30*, 87-89.

Johnson, A.L., Haugen, M.J., and Woods, D.C. (2008). Role for Inhibitor of Differentiation/Deoxyribonucleic Acid-Binding (Id) Proteins in Granulosa Cell Differentiation. Endocrinology *149*, 3187-3195.

Joshi, P.A., Jackson, H.W., Beristain, A.G., Di Grappa, M.A., Mote, P.A., Clarke, C.L., Stingl, J., Waterhouse, P.D., and Khokha, R. (2010). Progesterone induces adult mammary stem cell expansion. Nature *465*, 803-807.

Junankar, S.R., Eichten, A., Kramer, A., de Visser, K.E., and Coussens, L.M. (2006). Analysis of immune cell infiltrates during squamous carcinoma development. J Investig Dermatol Symp Proc *11*, 36-43.

Kang, T.W., Yevsa, T., Woller, N., Hoenicke, L., Wuestefeld, T., Dauch, D., Hohmeyer, A., Gereke, M., Rudalska, R., Potapova, A., Iken, M., Vucur, M., Weiss, S., Heikenwalder, M., Khan, S., Gil, J., Bruder, D., Manns, M., Schirmacher, P., Tacke, F., Ott, M., Luedde, T., Longerich, T., Kubicka, S., and Zender, L. (2011). Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. Nature *479*, 547-551.

Katoh, Y., and Katoh, M. (2008). Hedgehog signaling, epithelial-to-mesenchymal transition and miRNA (review). Int J Mol Med *22*, 271-275.

Kee, B.L. (2009). E and ID proteins branch out. Nat Rev Immunol *9*, 175-184.

Kelly, C.M., Bernard, P.S., Krishnamurthy, S., Wang, B., Ebbert, M.T.W., Bastien, R.R.L., Boucher, K.M., Young, E., Iwamoto, T., and Pusztai, L. (2012). Agreement in Risk Prediction Between the 21-Gene Recurrence Score Assay (Oncotype DX®) and the PAM50 Breast Cancer Intrinsic Classifier in Early-Stage Estrogen Receptor α Positive Breast Cancer. The Oncologist *17*, 492-498.

Kessenbrock, K., Plaks, V., and Werb, Z. (2010). Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. Cell *141*, 52-67.

Khokha, R., and Werb, Z. (2011). Mammary Gland Reprogramming: Metalloproteinases Couple Form with Function. Cold Spring Harbor Perspectives in Biology *3*.

Kim, N.-S., Kim, H.-T., Kwon, M.-C., Choi, S.-W., Kim, Y.-Y., Yoon, K.-J., Koo, B.-K., Kong, M.-P., Shin, J., Cho, Y., and Kong, Y.-Y. (2011). Survival and Differentiation of Mammary Epithelial Cells in Mammary Gland Development Require Nuclear Retention of Id2 Due to RANK Signaling. Molecular and Cellular Biology *31*, 4775-4788.

Klinowska, T.C., Soriano, J.V., Edwards, G.M., Oliver, J.M., Valentijn, A.J., Montesano, R., and Streuli, C.H. (1999). Laminin and beta1 integrins are crucial for normal mammary gland development in the mouse. Dev Biol *215*, 13 - 32.

Kondo, T., and Raff, M. (2000). The Id4 HLH protein and the timing of oligodendrocyte differentiation. EMBO J *19*, 1998-2007.

Korkaya, H., Liu, S., and Wicha, M.S. (2011). Breast cancer stem cells, cytokine networks, and the tumor microenvironment. The Journal of Clinical Investigation *121*, 3804-3809.

Kouros-Mehr, H., Bechis, S.K., Slorach, E.M., Littlepage, L.E., Egeblad, M., Ewald, A.J., Pai, S.-Y., Ho, I.C., and Werb, Z. (2008). GATA-3 Links Tumor Differentiation and Dissemination in a Luminal Breast Cancer Model. Cancer Cell *13*, 141-152.

Kouros-Mehr, H., Slorach, E.M., Sternlicht, M.D., and Werb, Z. (2006). GATA-3 Maintains the Differentiation of the Luminal Cell Fate in the Mammary Gland. Cell *127*, 1041-1055.

Kuzontkoski, P.M., Mulligan-Kehoe, M.J., Harris, B.T., and Israel, M.A. (2010). Inhibitor of DNA binding-4 promotes angiogenesis and growth of glioblastoma multiforme by elevating matrix GLA levels. Oncogene *29*, 3793-3802.

Laffin, B., Wellberg, E., Kwak, H.-I., Burghardt, R.C., Metz, R.P., Gustafson, T., Schedin, P., and Porter, W.W. (2008). Loss of Singleminded-2s in the Mouse Mammary Gland Induces an Epithelial-Mesenchymal Transition Associated with Up-Regulation of Slug and Matrix Metalloprotease 2. Molecular and Cellular Biology *28*, 1936-1946.

Langers, I., Renoux, V.M., Thiry, M., Delvenne, P., and Jacobs, N. (2012). Natural killer cells: role in local tumor growth and metastasis. Biologics *6*, 73-82.

Larsson, L.-G. (2011). Oncogene- and tumor suppressor gene-mediated suppression of cellular senescence. Seminars in Cancer Biology *21*, 367-376.

Lasorella, A., Uo, T., and Iavarone, A. (2001). Id proteins at the cross-road of development and cancer. Oncogene *20*, 8326-8333.

Lee, H.J., Hinshelwood, R.A., Bouras, T., Gallego-Ortega, D., Valdés-Mora, F., Blazek, K., Visvader, J.E., Clark, S.J., and Ormandy, C.J. (2011a). Lineage Specific Methylation of the Elf5 Promoter in Mammary Epithelial Cells. Stem Cells *29*, 1611-1619.

Lee, S., Schmitt, C.A., and Reimann, M. (2011b). The Myc/macrophage tango: Oncogene-induced senescence, Myc style. Seminars in Cancer Biology *21*, 377-384.

Leong, C.O., Vidnovic, N., DeYoung, M.P., Sgroi, D., and Ellisen, L.W. (2007). The p63/p73 network mediates chemosensitivity to cisplatin in a biologically defined subset of primary breast cancers. J Clin Invest *117*, 1370-1380.

Li, N., Singh, S., Cherukuri, P., Li, H., Yuan, Z., Ellisen, L.W., Wang, B., Robbins, D., and DiRenzo, J. (2008). Reciprocal intraepithelial interactions between TP63 and hedgehog signaling regulate quiescence and activation of progenitor elaboration by mammary stem cells. Stem Cells *26*, 1253-1264.

Li, X., Gonzalez, M.E., Toy, K., Filzen, T., Merajver, S.D., and Kleer, C.G. (2009). Targeted Overexpression of EZH2 in the Mammary Gland Disrupts Ductal Morphogenesis and Causes Epithelial Hyperplasia. The American Journal of Pathology *175*, 1246-1254.

Liedtke, C., Mazouni, C., Hess, K.R., Andre, F., Tordai, A., Mejia, J.A., Symmans, W.F., Gonzalez-Angulo, A.M., Hennessy, B., Green, M., Cristofanilli, M., Hortobagyi, G.N., and Pusztai, L. (2008). Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol *26*, 1275-1281.

Lilla, J.N., Joshi, R.V., Craik, C.S., and Werb, Z. (2009). Active Plasma Kallikrein Localizes to Mast Cells and Regulates Epithelial Cell Apoptosis, Adipocyte Differentiation, and Stromal Remodeling during Mammary Gland Involution. Journal of Biological Chemistry *284*, 13792-13803.

Lilla, J.N., and Werb, Z. (2010). Mast cells contribute to the stromal microenvironment in mammary gland branching morphogenesis. Dev Biol *337*, 124-133.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., Feleppa, F., Huschtscha, L.I., Thorne, H.J., Fox, S.B., Yan, M., French, J.D., Brown, M.A., Smyth, G.K., Visvader, J.E., and Lindeman, G.J. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat Med *15*, 907-913.

Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M.L., Vaillant, F., Yagita, H., Lindeman, G.J., Smyth, G.K., and Visvader, J.E. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res *12*, R21.

Lin, C.Q., Singh, J., Murata, K., Itahana, Y., Parrinello, S., Liang, S.H., Gillett, C.E., Campisi, J., and Desprez, P.-Y. (2000). A Role for Id-1 in the Aggressive Phenotype and Steroid Hormone Response of Human Breast Cancer Cells. Cancer Research *60*, 1332-1340.

Lindvall, C., Evans, N.C., Zylstra, C.R., Li, Y., Alexander, C.M., and Williams, B.O. (2006). The Wnt signaling receptor Lrp5 is required for mammary ductal stem cell activity and Wnt1-induced tumorigenesis. J Biol Chem *281*, 35081-35087.

Liu, K.J., and Harland, R.M. (2003). Cloning and characterization of Xenopus Id4 reveals differing roles for Id genes. Dev Biol *264*, 339-351.

Lopez-Knowles, E., Zardawi, S.J., McNeil, C.M., Millar, E.K.A., Crea, P., Musgrove, E.A., Sutherland, R.L., and O'Toole, S.A. (2010). Cytoplasmic Localization of β-Catenin is a Marker of Poor Outcome in Breast Cancer Patients. Cancer Epidemiology Biomarkers & Prevention *19*, 301-309.

Lund, L.R., Romer, J., Thomasset, N., Solberg, H., Pyke, C., Bissell, M.J., Dano, K., and Werb, Z. (1996). Two distinct phases of apoptosis in mammary gland involution: proteinase-independent and -dependent pathways. Development *122*, 181-193.

Lyden, D., Young, A.Z., Zagzag, D., Yan, W., Gerald, W., O'Reilly, R., Bader, B.L., Hynes, R.O., Zhuang, Y., Manova, K., and Benezra, R. (1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts. Nature *401*, 670-677.

Lydon, J.P., DeMayo, F.J., Funk, C.R., Mani, S.K., Hughes, A.R., Montgomery, C.A., Shyamala, G., Conneely, O.M., and O'Malley, B.W. (1995). Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities. Genes Dev *9*, 2266 - 2278.

Magistrelli, G., Toma, S., and Isacchi, A. (1996). Substitution of Two Variant Residues in the Protein Tyrosine Phosphatase-like PTP35/IA-2 Sequence Reconstitutes Catalytic Activity. Biochemical and Biophysical Research Communications *227*, 581-588.

Mailleux, A.A., Overholtzer, M., Schmelzle, T., Bouillet, P., Strasser, A., and Brugge, J.S. (2007). BIM Regulates Apoptosis during Mammary Ductal Morphogenesis, and Its Absence Reveals Alternative Cell Death Mechanisms. Developmental cell *12*, 221-234.

Mailleux, A.A., Spencer-Dene, B., Dillon, C., Ndiaye, D., Savona-Baron, C., Itoh, N., Kato, S., Dickson, C., Thiery, J.P., and Bellusci, S. (2002). Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo. Development *129*, 53-60.

Mallepell, S., Krust, A., Chambon, P., and Brisken, C. (2006). Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. Proc Natl Acad Sci U S A *103*, 2196-2201.

Maller, O., Martinson, H., and Schedin, P. (2010). Extracellular Matrix Composition Reveals Complex and Dynamic Stromal-Epithelial Interactions in the Mammary Gland. Journal of Mammary Gland Biology and Neoplasia *15*, 301-318.

Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R.M. (1995). The nuclear receptor superfamily: the second decade. Cell *83*, 835-839.

Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. Nature *454*, 436-444.

Marin-Husstege, M., He, Y., Li, J., Kondo, T., Sablitzky, F., and Casaccia-Bonnefil, P. (2006). Multiple roles of Id4 in developmental myelination: Predicted outcomes and unexpected findings. Glia *54*, 285-296.

McBryan, J., Howlin, J., Napoletano, S., and Martin, F. (2008). Amphiregulin: role in mammary gland development and breast cancer. J Mammary Gland Biol Neoplasia *13*, 159-169.

McCaffrey, L.M., and Macara, I.G. (2009). The Par3/aPKC interaction is essential for end bud remodeling and progenitor differentiation during mammary gland morphogenesis. Genes Dev *23*, 1450-1460.

McNally, S., and Martin, F. (2011). Molecular regulators of pubertal mammary gland development. Ann Med *43*, 212-234.

Michaloglou, C., Vredeveld, L.C.W., Soengas, M.S., Denoyelle, C., Kuilman, T., van der Horst, C.M.A.M., Majoor, D.M., Shay, J.W., Mooi, W.J., and Peeper, D.S. (2005). BRAFE600-associated senescence-like cell cycle arrest of human naevi. Nature *436*, 720-724.

Momparler, R.L., and Bovenzi, V. (2000). DNA methylation and cancer. Journal of Cellular Physiology *183*, 145-154.

Montesano, R., Sarkozi, R., and Schramek, H. (2008). Bone morphogenetic protein-4 strongly potentiates growth factor-induced proliferation of mammary epithelial cells. Biochem Biophys Res Commun *374*, 164-168.

Mueller, S.O., Clark, J.A., Myers, P.H., and Korach, K.S. (2002). Mammary Gland Development in Adult Mice Requires Epithelial and Stromal Estrogen Receptor α. Endocrinology *143*, 2357-2365.

Mukherjee, A., Soyal, S.M., Li, J., Ying, Y., He, B., DeMayo, F.J., and Lydon, J.P. (2010). Targeting RANKL to a specific subset of murine mammary epithelial cells induces ordered branching morphogenesis and alveologenesis in the absence of progesterone receptor expression. FASEB J *24*, 4408-4419.

Murad, J.M., Place, C.S., Ran, C., Hekmatyar, S.K., Watson, N.P., Kauppinen, R.A., and Israel, M.A. (2010). Inhibitor of DNA binding 4 (ID4) regulation of adipocyte differentiation and adipose tissue formation in mice. J Biol Chem *285*, 24164-24173.

Murre, C. (2005). Helix-loop-helix proteins and lymphocyte development. Nat Immunol *6*, 1079-1086.

Muschler, J., and Streuli, C.H. (2010). Cell-Matrix Interactions in Mammary Gland Development and Breast Cancer. Cold Spring Harbor Perspectives in Biology *2*.

Mziaut, H., Kersting, S., Knoch, K.-P., Fan, W.-H., Trajkovski, M., Erdmann, K., Bergert, H., Ehehalt, F., Saeger, H.-D., and Solimena, M. (2008). ICA512 signaling enhances pancreatic β-cell proliferation by regulating cyclins D through STATs. Proceedings of the National Academy of Sciences *105*, 674-679.

Nair, R., Junankar, S., O'Toole, S., Shah, J., Borowsky, A.D., Bishop, J.M., and Swarbrick, A. (2010). Redefining the expression and function of the inhibitor of differentiation 1 in mammary gland development. PLoS One *5*, e11947.

Naylor, M.J., Li, N., Cheung, J., Lowe, E.T., Lambert, E., Marlow, R., Wang, P., Schatzmann, F., Wintermantel, T., Schaetz, G.n., Clarke, A.R., Mueller, U., Hynes, N.E., and Streuli, C.H. (2005). Ablation of β1 integrin in mammary epithelium reveals a key role for integrin in glandular morphogenesis and differentiation. The Journal of Cell Biology *171*, 717-728.

Noetzel, E., Veeck, J., Niederacher, D., Galm, O., Horn, F., Hartmann, A., Knuchel, R., and Dahl, E. (2008). Promoter methylation-associated loss of ID4 expression is a marker of tumour recurrence in human breast cancer. BMC Cancer *8*, 154.

Norton, J.D. (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. J Cell Sci *113 (Pt 22)*, 3897-3905.

O'Toole, S.A., Machalek, D.A., Shearer, R.F., Millar, E.K.A., Nair, R., Schofield, P., McLeod, D., Cooper, C.L., McNeil, C.M., McFarland, A., Nguyen, A., Ormandy, C.J., Qiu, M.R., Rabinovich, B., Martelotto, L.G., Vu, D., Hannigan, G.E., Musgrove, E.A., Christ, D., Sutherland, R.L., Watkins, D.N., and Swarbrick, A. (2011). Hedgehog Overexpression Is Associated with Stromal Interactions and Predicts for Poor Outcome in Breast Cancer. Cancer Research *71*, 4002-4014.

Oakes, S.R., Naylor, M.J., Asselin-Labat, M.L., Blazek, K.D., Gardiner-Garden, M., Hilton, H.N., Kazlauskas, M., Pritchard, M.A., Chodosh, L.A., Pfeffer, P.L., Lindeman, G.J., Visvader, J.E., and Ormandy, C.J. (2008). The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev *22*, 581-586.

Oatley, M.J., Kaucher, A.V., Racicot, K.E., and Oatley, J.M. (2011). Inhibitor of DNA Binding 4 Is Expressed Selectively by Single Spermatogonia in the Male Germline and Regulates the Self-Renewal of Spermatogonial Stem Cells in Mice. Biology of Reproduction *85*, 347-356.

Ormandy, C.J., Camus, A., Barra, J., Damotte, D., Lucas, B., Buteau, H., Edery, M., Brousse, N., Babinet, C., Binart, N., and Kelly, P.A. (1997). Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. Genes & Development *11*, 167-178.

Pagliuca, A., Bartoli, P.C., Saccone, S., Della Valle, G., and Lania, L. (1995). Molecular cloning of ID4, a novel dominant negative helix-loop-helix human gene on chromosome 6p21.3-p22. Genomics *27*, 200-203.

Pagliuca, A., Cannada-Bartoli, P., and Lania, L. (1998). A role for Sp and helix-loop-helix transcription factors in the regulation of the human Id4 gene promoter activity. J Biol Chem *273*, 7668-7674.

Park, S.J., Kim, R.J., and Nam, J.S. (2011). Inhibitor of DNA-binding 4 contributes to the maintenance and expansion of cancer stem cells in 4T1 mouse mammary cancer cell line. Lab Anim Res *27*, 333-338.

Perk, J., Gil-Bazo, I., Chin, Y., de Candia, P., Chen, J.J.S., Zhao, Y., Chao, S., Cheong, W., Ke, Y., Al-Ahmadie, H., Gerald, W.L., Brogi, E., and Benezra, R. (2006). Reassessment of Id1 Protein Expression in Human Mammary, Prostate, and Bladder Cancers Using a Monospecific Rabbit Monoclonal Anti-Id1 Antibody. Cancer Research *66*, 10870-10877.

Perk, J., Iavarone, A., and Benezra, R. (2005). Id family of helix-loop-helix proteins in cancer. Nat Rev Cancer *5*, 603-614.

Pierce, D.F., Johnson, M.D., Matsui, Y., Robinson, S.D., Gold, L.I., Purchio, A.F., Daniel, C.W., Hogan, B.L., and Moses, H.L. (1993). Inhibition of mammary duct development but not alveolar outgrowth during pregnancy in transgenic mice expressing active TGF-beta 1. Genes & Development *7*, 2308-2317.

Pietersen, A.M., Evers, B., Prasad, A.A., Tanger, E., Cornelissen-Steijger, P., Jonkers, J., and van Lohuizen, M. (2008). Bmi1 Regulates Stem Cells and Proliferation and Differentiation of Committed Cells in Mammary Epithelium. Current Biology *18*, 1094-1099.

Polyak, K., and Kalluri, R. (2011). The Role of the Microenvironment in Mammary Gland Development and Cancer. Cold Spring Harbor Perspectives in Biology *2*.

Prat, A., and Perou, C.M. (2010). Deconstructing the molecular portraits of breast cancer. Molecular Oncology *5*, 5-23.

Qian, T., Lee, J.Y., Park, J.H., Kim, H.J., and Kong, G. (2010). Id1 enhances RING1b E3 ubiquitin ligase activity through the Mel-18/Bmi-1 polycomb group complex. Oncogene *29*, 5818-5827.

Raafat, A., Goldhar, A.S., Klauzinska, M., Xu, K., Amirjazil, I., McCurdy, D., Lashin, K., Salomon, D., Vonderhaar, B.K., Egan, S., and Callahan, R. (2011). Expression of Notch receptors, ligands, and target genes during development of the mouse mammary gland. Journal of Cellular Physiology *226*, 1940-1952.

Reddy, J., and Li, Y. (2011). Oncogene-Induced Senescence and its Role in Tumor Suppression. Journal of Mammary Gland Biology and Neoplasia *16*, 247-256.

Regan, J.L., Kendrick, H., Magnay, F.A., Vafaizadeh, V., Groner, B., and Smalley, M.J. (2011). c-Kit is required for growth and survival of the cells of origin of Brca1-mutation-associated breast cancer. Oncogene.

Reyal, F., van Vliet, M.H., Armstrong, N.J., Horlings, H.M., de Visser, K.E., Kok, M., Teschendorff, A.E., Mook, S., van 't Veer, L., Caldas, C., Salmon, R.J., van de Vijver, M.J., and Wessels, L.F. (2008). A comprehensive analysis of prognostic signatures reveals the high predictive capacity of the proliferation, immune response and RNA splicing modules in breast cancer. Breast Cancer Res *10*, R93.

Riechmann, V., van Cruchten, I., and Sablitzky, F. (1994). The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3. Nucleic Acids Res *22*, 749-755.

Rigolet, M., Rich, T., Gross-Morand, M.S., Molina-Gomes, D., Viegas-Pequignot, E., and Junien, C. (1998). cDNA cloning, tissue distribution and chromosomal localization of the human ID4 gene. DNA Res *5*, 309-313.

Roarty, K., and Rosen, J.M. (2010). Wnt and mammary stem cells: hormones cannot fly wingless. Curr Opin Pharmacol *10*, 643-649.

Roarty, K., and Serra, R. (2007). Wnt5a is required for proper mammary gland development and TGF-beta-mediated inhibition of ductal growth. Development *134*, 3929-3939.

Rodier, F., and Campisi, J. (2011). Four faces of cellular senescence. The Journal of Cell Biology *192*, 547-556.

Roldan, G., Delgado, L., and Muse, I.M. (2006). Tumoral expression of BRCA1, estrogen receptor alpha and ID4 protein in patients with sporadic breast cancer. Cancer Biol Ther *5*, 505-510.

Romero-Lanman, E.E., Pavlovic, S., Amlani, B., Chin, Y., and Benezra, R. Id1 maintains embryonic stem cell self-renewal by up-regulation of nanog and repression of brachyury expression. Stem Cells Dev *21*, 384-393.

Rottenberg, S., Nygren, A.O., Pajic, M., van Leeuwen, F.W., van der Heijden, I., van de Wetering, K., Liu, X., de Visser, K.E., Gilhuijs, K.G., van Tellingen, O., Schouten, J.P., Jonkers, J., and Borst, P. (2007). Selective induction of chemotherapy resistance of mammary tumors in a conditional mouse model for hereditary breast cancer. Proc Natl Acad Sci U S A *104*, 12117-12122.

Ruchusatsawat, K., Wongpiyabovorn, J., Protjaroen, P., Chaipipat, M., Shuangshoti, S., Thorner, P.S., and Mutirangura, A. (2011). Parakeratosis in skin is associated with loss of inhibitor of differentiation 4 via promoter methylation. Hum Pathol *42*, 1878-1887.

Russell, L.J., Akasaka, T., Majid, A., Sugimoto, K.-j., Loraine Karran, E., Nagel, I., Harder, L., Claviez, A., Gesk, S., Moorman, A.V., Ross, F., Mazzullo, H., Strefford, J.C., Siebert, R., Dyer, M.J.S., and Harrison, C.J. (2008). t(6;14)(p22;q32): a new recurrent IGH@ translocation involving ID4 in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Blood *111*, 387-391.

Ruzinova, M.B., and Benezra, R. (2003). Id proteins in development, cell cycle and cancer. Trends Cell Biol *13*, 410-418.

Sablitzky, F., Moore, A., Bromley, M., Deed, R.W., Newton, J.S., and Norton, J.D. (1998). Stage- and subcellular-specific expression of Id proteins in male germ and Sertoli cells implicates distinctive regulatory roles for Id proteins during meiosis, spermatogenesis, and Sertoli cell function. Cell Growth Differ *9*, 1015-1024.

Samanta, J., and Kessler, J.A. (2004). Interactions between ID and OLIG proteins mediate the inhibitory effects of BMP4 on oligodendroglial differentiation. Development *131*, 4131-4142.

Sarkisian, C.J., Keister, B.A., Stairs, D.B., Boxer, R.B., Moody, S.E., and Chodosh, L.A. (2007). Dose-dependent oncogene-induced senescence in vivo and its evasion during mammary tumorigenesis. Nat Cell Biol *9*, 493-505.

Schlee, M., Hölzel, M., Bernard, S., Mailhammer, R., Schuhmacher, M., Reschke, J., Eick, D., Marinkovic, D., Wirth, T., Rosenwald, A., Staudt, L.M., Eilers, M., Baran-Marszak, F., Fagard, R., Feuillard, J., Laux, G., and Bornkamm, G.W. (2007). c-MYC activation impairs the NF-κB and the interferon response: Implications for the pathogenesis of Burkitt's lymphoma. International Journal of Cancer *120*, 1387-1395.

Schneider-Kolsky, M.E., Hart, S., Fox, J., Midolo, P., Stuckey, J., Hofman, M., and Ganju, V. (2010). The role of chemotherapeutic drugs in the evaluation of breast tumour response to chemotherapy using serial FDG-PET. Breast Cancer Res *12*, R37.

Schoppmann, S.F., Schindl, M., Bayer, G., Aumayr, K., Dienes, J., Horvat, R., Rudas, M., Gnant, M., Jakesz, R., and Birner, P. (2003). Overexpression of Id-1 is associated with poor clinical outcome in node negative breast cancer. International Journal of Cancer *104*, 677-682.

Senoo, M., Pinto, F., Crum, C.P., and McKeon, F. (2007). p63 Is essential for the proliferative potential of stem cells in stratified epithelia. Cell *129*, 523-536.

Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. (1997). Oncogenic ras Provokes Premature Cell Senescence Associated with Accumulation of p53 and p16INK4a. Cell *88*, 593-602.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84-88.

Shan, L., Yu, M., Qiu, C., and Snyderwine, E.G. (2003). Id4 regulates mammary epithelial cell growth and differentiation and is overexpressed in rat mammary gland carcinomas. Am J Pathol *163*, 2495-2502.

Shin, K.J., Wall, E.A., Zavzavadjian, J.R., Santat, L.A., Liu, J., Hwang, J.I., Rebres, R., Roach, T., Seaman, W., Simon, M.I., and Fraser, I.D. (2006). A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. Proc Natl Acad Sci U S A *103*, 13759-13764.

Silberstein, G.B., Van Horn, K., Hrabeta-Robinson, E., and Compton, J. (2006). Estrogen-triggered delays in mammary gland gene expression during the estrous cycle: evidence for a novel timing system. J Endocrinol *190*, 225-239.

Simian, M., Hirai, Y., Navre, M., Werb, Z., Lochter, A., and Bissell, M.J. (2001). The interplay of matrix metalloproteinases, morphogens and growth factors is necessary for branching of mammary epithelial cells. Development *128*, 3117 - 3131.

Skinner, M.K., Rawls, A., Wilson-Rawls, J., and Roalson, E.H. (2010). Basic helix-loop-helix transcription factor gene family phylogenetics and nomenclature. Differentiation *80*, 1-8.

Sleeman, K.E., Kendrick, H., Robertson, D., Isacke, C.M., Ashworth, A., and Smalley, M.J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. J Cell Biol *176*, 19-26.

Smart, C.E., Wronski, A., French, J.D., Edwards, S.L., Asselin-Labat, M.L., Waddell, N., Peters, K., Brewster, B.L., Brooks, K., Simpson, K., Manning, N., Lakhani, S.R., Grimmond, S., Lindeman, G.J., Visvader, J.E., and Brown, M.A. (2010). Analysis of Brca1-deficient mouse mammary glands reveals reciprocal regulation of Brca1 and c-kit. Oncogene.

Smith, G.H., and Medina, D. (2008). Re-evaluation of mammary stem cell biology based on in vivo transplantation. Breast Cancer Res *10*, 203.

Sneddon, J.B., and Werb, Z. (2007). Location, Location, Location: The Cancer Stem Cell Niche. Cell Stem Cell *1*, 607-611.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Thorsen, T., Quist, H., Matese, J.C., Brown, P.O., Botstein, D., Lonning, P.E., and Borresen-Dale, A.-L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences *98*, 10869-10874.

Soyal, S., Ismail, P.M., Li, J., Mulac-Jericevic, B., Conneely, O.M., and Lydon, J.P. (2002). Progesterone's role in mammary gland development and tumorigenesis as disclosed by experimental mouse genetics. Breast Cancer Res *4*, 191-196.

Sternlicht, M.D. (2006). Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. Breast Cancer Res *8*, 201.

Sternlicht, M.D., Kouros-Mehr, H., Lu, P., and Werb, Z. (2006). Hormonal and local control of mammary branching morphogenesis. Differentiation *74*, 365-381.

Stingl, J. (2009). Detection and analysis of mammary gland stem cells. J Pathol *217*, 229-241.

Stingl, J. (2011). Estrogen and Progesterone in Normal Mammary Gland Development and in Cancer. Hormones and Cancer *2*, 85-90.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993-997.

Swarbrick, A., Akerfeldt, M.C., Lee, C.S.L., Sergio, C.M., Caldon, C.E., Hunter, L.-J.K., Sutherland, R.L., and Musgrove, E.A. (2004). Regulation of cyclin expression and cell cycle progression in breast epithelial cells by the helix-loop-helix protein Id1. Oncogene *24*, 381-389.

Swarbrick, A., Roy, E., Allen, T., and Bishop, J.M. (2008). Id1 cooperates with oncogenic Ras to induce metastatic mammary carcinoma by subversion of the cellular senescence response. Proc Natl Acad Sci U S A *105*, 5402-5407.

Tabar, L., Dean, P.B., Kaufman, C.S., Duffy, S.W., and Chen, H.H. (2000). A new era in the diagnosis of breast cancer. Surg Oncol Clin N Am *9*, 233-277.

Taddei, I., Deugnier, M.-A., Faraldo, M.M., Petit, V., Bouvard, D., Medina, D., Fassler, R., Thiery, J.P., and Glukhova, M.A. (2008). [beta]1 Integrin deletion from the basal compartment of the mammary epithelium affects stem cells. Nat Cell Biol *10*, 716-722.

Tao, L., Roberts, A.L., Dunphy, K.A., Bigelow, C., Yan, H., and Jerry, D.J. (2010). Repression of Mammary Stem/Progenitor Cells by P53 is Mediated by Notch and Separable from Apoptotic Activity. Stem Cells.

Taube, J.H., Herschkowitz, J.I., Komurov, K., Zhou, A.Y., Gupta, S., Yang, J., Hartwell, K., Onder, T.T., Gupta, P.B., Evans, K.W., Hollier, B.G., Ram, P.T., Lander, E.S., Rosen, J.M., Weinberg, R.A., and Mani, S.A. (2010). Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proceedings of the National Academy of Sciences *107*, 15449-15454.

Thompson, E.W., Yu, M., Bueno, J., Jin, L., Maiti, S.N., Palao-Marco, F.L., Pulyaeva, H., Tamborlane, J.W., Tirgari, R., and Wapnir, I. (1994). Collagen induced MMP-2 activation in human breast cancer. Breast Cancer Res Treat *31*, 357 - 370.

Tokuzawa, Y., Yagi, K., Yamashita, Y., Nakachi, Y., Nikaido, I., Bono, H., Ninomiya, Y., Kanesaki-Yatsuka, Y., Akita, M., Motegi, H., Wakana, S., Noda, T., Sablitzky, F., Arai, S., Kurokawa, R., Fukuda, T., Katagiri, T., Schonbach, C., Suda, T., Mizuno, Y., and Okazaki, Y. (2010). Id4, a new candidate gene for senile osteoporosis, acts as a molecular switch promoting osteoblast differentiation. PLoS Genet *6*, e1001019.

Torii, S. (2009). Expression and Function of IA-2 Family Proteins, Unique Neuroendocrine-specific Protein-tyrosine Phosphatases. Endocrine Journal *56*, 639-648.

Turner, N., Tutt, A., and Ashworth, A. (2004). Hallmarks of 'BRCAness' in sporadic cancers. Nat Rev Cancer *4*, 814-819.

Turner, N.C., Reis-Filho, J.S., Russell, A.M., Springall, R.J., Ryder, K., Steele, D., Savage, K., Gillett, C.E., Schmitt, F.C., Ashworth, A., and Tutt, A.N. (2007). BRCA1 dysfunction in sporadic basal-like breast cancer. Oncogene *26*, 2126-2132.

Umetani, N., Mori, T., Koyanagi, K., Shinozaki, M., Kim, J., Giuliano, A.E., and Hoon, D.S. (2005). Aberrant hypermethylation of ID4 gene promoter region increases risk of lymph node metastasis in T1 breast cancer. Oncogene *24*, 4721-4727.

Umetani, N., Takeuchi, H., Fujimoto, A., Shinozaki, M., Bilchik, A.J., and Hoon, D.S.B. (2004). Epigenetic Inactivation of ID4 in Colorectal Carcinomas Correlates with Poor Differentiation and Unfavorable Prognosis. Clinical Cancer Research *10*, 7475-7483.

van Cruchten, I., Cinato, E., Fox, M., King, E.R., Newton, J.S., Riechmann, V., and Sablitzky, F. (1998). Structure, chromosomal localisation and expression of the murine dominant negative helix-loop-helix Id4 gene. Biochim Biophys Acta *1443*, 55-64.

van de Vijver, M.J., He, Y.D., van't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W., Schreiber, G.J., Peterse, J.L., Roberts, C., Marton, M.J., Parrish, M., Atsma, D., Witteveen, A., Glas, A.,

Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E.T., Friend, S.H., and Bernards, R. (2002). A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med *347*, 1999 - 2009.

Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. Nature *479*, 189-193.

Veltmaat, J.M., Relaix, F.d.r., Le, L.T., Kratochwil, K., Sala, F.d.r.G., van Veelen, W., Rice, R., Spencer-Dene, B., Mailleux, A.A., Rice, D.P., Thiery, J.P., and Bellusci, S. (2006). Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes. Development *133*, 2325-2335.

Veltmaat, J.M., Van Veelen, W., Thiery, J.P., and Bellusci, S. (2004). Identification of the mammary line in mouse by Wnt10b expression. Developmental Dynamics *229*, 349-356.

Ventura, A., Kirsch, D.G., McLaughlin, M.E., Tuveson, D.A., Grimm, J., Lintault, L., Newman, J., Reczek, E.E., Weissleder, R., and Jacks, T. (2007a). Restoration of p53 function leads to tumour regression in vivo. Nature *445*, 661-665.

Ventura, J.J., Tenbaum, S., Perdiguero, E., Huth, M., Guerra, C., Barbacid, M., Pasparakis, M., and Nebreda, A.R. (2007b). p38[alpha] MAP kinase is essential in lung stem and progenitor cell proliferation and differentiation. Nat Genet *39*, 750-758.

Visvader, J.E. (2009). Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev *23*, 2563-2577.

Visvader, J.E. (2011). Cells of origin in cancer. Nature *469*, 314-322.

Wang, H., Wang, X.-Q., Xu, X.-P., and Lin, G.-W. (2010). ID4 methylation predicts high risk of leukemic transformation in patients with myelodysplastic syndrome. Leukemia Research *34*, 598-604.

Wang, S., Sdrulla, A., Johnson, J.E., Yokota, Y., and Barres, B.A. (2001). A Role for the Helix-Loop-Helix Protein Id2 in the Control of Oligodendrocyte Development. Neuron *29*, 603-614.

Watson, C.J., and Khaled, W.T. (2008). Mammary development in the embryo and adult: a journey of morphogenesis and commitment. Development *135*, 995-1003.

Watson, C.J., and Kreuzaler, P.A. (2011). Remodeling mechanisms of the mammary gland during involution. Int J Dev Biol *55*, 757-762.

Welcsh, P.L., Lee, M.K., Gonzalez-Hernandez, R.M., Black, D.J., Mahadevappa, M., Swisher, E.M., Warrington, J.A., and King, M.C. (2002). BRCA1 transcriptionally regulates genes involved in breast tumorigenesis. Proc Natl Acad Sci U S A *99*, 7560-7565.

Welm, A.L., Kim, S., Welm, B.E., and Bishop, J.M. (2005). MET and MYC cooperate in mammary tumorigenesis. Proceedings of the National Academy of Sciences of the United States of America *102*, 4324-4329.

Welm, A.L., Sneddon, J.B., Taylor, C., Nuyten, D.S.A., van de Vijver, M.J., Hasegawa, B.H., and Bishop, J.M. (2007). The macrophage-stimulating protein pathway promotes metastasis in a mouse model for breast cancer and predicts poor prognosis in humans. Proceedings of the National Academy of Sciences *104*, 7570-7575.

Wen, Y., Ho, A., Patil, S., Akram, M., Catalano, J., Eaton, A., Norton, L., Benezra, R., and Brogi, E. (2012). Id4 protein is highly expressed in triple-negative breast carcinomas: possible implications for BRCA1 downregulation. Breast Cancer Research and Treatment, 1-10.

Wiesen, J.F., Young, P., Werb, Z., and Cunha, G.R. (1999). Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development. Development *126*, 335-344.

Williams, Samuel A., Maecker, Heather L., French, Dorothy M., Liu, J., Gregg, A., Silverstein, Leah B., Cao, Tim C., Carano, Richard A.D., and Dixit, Vishva M. (2011). USP1 Deubiquitinates ID Proteins to Preserve a Mesenchymal Stem Cell Program in Osteosarcoma. Cell *146*, 918-930.

Wiseman, B.S., Sternlicht, M.D., Lund, L.R., Alexander, C.M., Mott, J., Bissell, M.J., Soloway, P., Itohara, S., and Werb, Z. (2003). Site-specific inductive and inhibitory activities of MMP-2 and MMP-3 orchestrate mammary gland branching morphogenesis. The Journal of Cell Biology *162*, 1123-1133.

Wiseman, B.S., and Werb, Z. (2002). Stromal effects on mammary gland development and breast cancer. Science *296*, 1046 - 1049.

Woodward, W.A., Chen, M.S., Behbod, F., and Rosen, J.M. (2005). On mammary stem cells. Journal of Cell Science *118*, 3585-3594.

Wooster, R., and Weber, B.L. (2003). Breast and Ovarian Cancer. New England Journal of Medicine *348*, 2339-2347.

Wu, Y., Richard, J.-P., Wang, S.D., Rath, P., Laterra, J., and Xia, S. (2012). Regulation of glioblastoma multiforme stem-like cells by inhibitor of DNA binding proteins and oligodendroglial lineage-associated transcription factors. Cancer Science, 103: 1028–1037.

Xue, W., Zender, L., Miething, C., Dickins, R.A., Hernando, E., Krizhanovsky, V., Cordon-Cardo, C., and Lowe, S.W. (2007). Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. Nature *445*, 656-660.

Yalcin-Ozuysal, O., Fiche, M., Guitierrez, M., Wagner, K.U., Raffoul, W., and Brisken, C. (2010). Antagonistic roles of Notch and p63 in controlling mammary epithelial cell fates. Cell Death Differ *17*, 1600-1612.

Yamaji, D., Na, R., Feuermann, Y., Pechhold, S., Chen, W., Robinson, G.W., and Hennighausen, L. (2009). Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. Genes & Development *23*, 2382-2387.

Yokota, Y., and Mori, S. (2002). Role of Id family proteins in growth control. Journal of Cellular Physiology *190*, 21-28.

Yu, L., Liu, C., Vandeusen, J., Becknell, B., Dai, Z., Wu, Y.Z., Raval, A., Liu, T.H., Ding, W., Mao, C., Liu, S., Smith, L.T., Lee, S., Rassenti, L., Marcucci, G., Byrd, J., Caligiuri, M.A., and Plass, C. (2005). Global assessment of promoter methylation in a mouse model of cancer identifies ID4 as a putative tumor-suppressor gene in human leukemia. Nat Genet *37*, 265-274.

Yuen, H.F., Chua, C.W., Chan, Y.P., Wong, Y.C., Wang, X., and Chan, K.W. (2006). Id proteins expression in prostate cancer: high-level expression of Id-4 in primary prostate cancer is associated with development of metastases. Mod Pathol *19*, 931-941.

Yun, K., Mantani, A., Garel, S., Rubenstein, J., and Israel, M.A. (2004). Id4 regulates neural progenitor proliferation and differentiation in vivo. Development *131*, 5441-5448.

Zeng, Y.A., and Nusse, R. (2010). Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. Cell Stem Cell *6*, 568-577.

Zeps, N., Bentel, J.M., Papadimitriou, J.M., D'Antuono, M.F., and Dawkins, H.J.S. (1998). Estrogen receptor-negative epithelial cells in mouse mammary gland development and growth. Differentiation *62*, 221-226.

Appendix 1

Up regulated by Id4 knockdown and down regulated by Id4 overexpression.

Gene.Symbol	Adjusted Average t
PTPRN: protein tyrosine phosphatase, receptor type, N	18.07527
VCAM1: vascular cell adhesion molecule 1	14.95808
PLAT: plasminogen activator, tissue	14.68257
THBS2: thrombospondin 2	14.3399
SERPINF1: serine (or cysteine) peptidase inhibitor, clade F, member 1	14.17895
MMP9: matrix metallopeptidase 9	14.11304
SEMA7A: sema domain, immunoglobulin domain (Ig), and GPI membrane anchor, (semaphorin) 7A	13.86503
GPR124: G protein-coupled receptor 124	13.85752
LGI2: leucine-rich repeat LGI family, member 2	13.34722
CDH11: cadherin 11	13.30668
AQP1: aquaporin 1	13.0584
PAPLN: papilin, proteoglycan-like sulfated glycoprotein	12.76218
GDPD2: glycerophosphodiester phosphodiesterase domain containing 2	12.70516
FSTL1: follistatin-like 1	12.52027
GLIPR2: GLI pathogenesis-related 2	11.7751
H2-GS10: MHC class I like protein GS10	11.75913
SGIP1: SH3-domain GRB2-like (endophilin) interacting protein 1	11.70429
NID1: nidogen 1	11.68817
HHIP: Hedgehog-interacting protein	11.20038
E330013P04RIK: RIKEN cDNA E330013P04 gene	11.14428
ADAMTS14: a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 14	11.1192
STEAP1: six transmembrane epithelial antigen of the prostate 1	11.09477
PRRX1: paired related homeobox 1	10.94361
GSTT1: glutathione S-transferase, theta 1	10.91673

S1PR1: sphingosine-1-phosphate receptor 1	10.903
STAMBPL1: STAM binding protein like 1	10.79118
ABCB1B: ATP-binding cassette, sub-family B (MDR/TAP), member 1B	10.73276
TMEM119: transmembrane protein 119	10.72418
PDGFRA: platelet derived growth factor receptor, alpha polypeptide	10.71078
ST3GAL1: ST3 beta-galactoside alpha-2,3-sialyltransferase 1	10.64472
PCDH7: protocadherin 7	10.48115
FLNC: filamin C, gamma	10.43773
HAS2: hyaluronan synthase 2	10.36514
MMP16: matrix metallopeptidase 16	10.2545
PSMB8: proteasome (prosome, macropain) subunit, beta type 8 (large multifunctional peptidase 7)	9.922156
GPR176: G protein-coupled receptor 176	9.86456
IL6R: interleukin 6 receptor, alpha	9.849632
EDNRA: endothelin receptor type A	9.832244
ODZ3: odd Oz/ten-m homolog 3 (Drosophila)	9.778826
LRCH2: leucine-rich repeats and calponin homology (CH) domain containing 2	9.775172
ADAMTS4: a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 4	9.537532
KIF21B: kinesin family member 21B	9.44952
PDPN: podoplanin	9.428428
GRIK2: glutamate receptor, ionotropic, kainate 2 (beta 2)	9.401988
SLC41A2: solute carrier family 41, member 2	9.286534
CACNA1C: calcium channel, voltage-dependent, L type, alpha 1C subunit	9.267682
STAC2: SH3 and cysteine rich domain 2	9.245252
PDGFRB: platelet derived growth factor receptor, beta polypeptide	9.239442
TF: transferrin	9.208244
C1R: complement component 1, r subcomponent B	9.157276
FAM131B: family with sequence similarity 131, member B	9.127492

USP11: ubiquitin specific peptidase 11	9.111134
LOXL2: lysyl oxidase-like 2	9.058566
TNS1: tensin 1	9.004248
IGLON5: IgLON family member 5	8.98523
CDYL2: chromodomain protein, Y chromosome-like 2	8.976184
C15orf48: expressed sequence AA467197	8.88399
MMP19: matrix metallopeptidase 19	8.828552
CTSO: cathepsin O	8.786136
ANKRD1: ankyrin repeat domain 1 (cardiac muscle)	8.78514
ENO2: enolase 2, gamma neuronal	8.73448
ACSS3: acyl-CoA synthetase short-chain family member 3	8.669444
ABCB1: ATP-binding cassette, sub-family B (MDR/TAP), member 1A	8.557034
TNFRSF1B: tumor necrosis factor receptor superfamily, member 1b	8.552112
CRABP1: cellular retinoic acid binding protein I	8.545576
PHACTR1: phosphatase and actin regulator 1	8.528602
OLFML3: olfactomedin-like 3	8.4947
RBP1: retinol binding protein 1, cellular	8.477214
PSMB9: proteasome (prosome, macropain) subunit, beta type 9 (large multifunctional peptidase 2)	8.452906
D4S234E: neuron specific gene family member 1	8.447762
STEAP2: six transmembrane epithelial antigen of prostate 2	8.424778
GBE1: glucan (1,4-alpha-), branching enzyme 1	8.386318
HK2: hexokinase 2	8.326054
C3: complement component 3	8.262998
CREB3L1: cAMP responsive element binding protein 3-like 1	8.220828
SCN5A: sodium channel, voltage-gated, type V, alpha	8.145456
PVRL2: poliovirus receptor-related 2	8.049806
H2-Q6: histocompatibility 2, Q region locus 6	8.04384
CDC42EP2: CDC42 effector protein (Rho GTPase binding) 2	8.041382
SLCO3A1: solute carrier organic anion transporter family, member 3a1	8.030666
TMEM229B: transmembrane protein 229B	7.903308

KIAA1199: RIKEN cDNA 9930013L23 gene	7.886352
IGSF10: immunoglobulin superfamily, member 10	7.863256
OLFML2B: olfactomedin-like 2B	7.84347
GGTA1: glycoprotein galactosyltransferase alpha 1, 3	7.806592
GEM: GTP binding protein (gene overexpressed in skeletal muscle)	7.801976
CD248: CD248 antigen, endosialin	7.78094
CLCA2: chloride channel calcium activated 2	7.773868
CXCL6: chemokine (C-X-C motif) ligand 5	7.726318
MAN1A1: mannosidase 1, alpha	7.613358
SIRPA: signal-regulatory protein alpha	7.609572
PAPPA: pregnancy-associated plasma protein A	7.58164
SP110: Sp110 nuclear body protein	7.570246
LAMB1: laminin B1	7.543952
CCL13: chemokine (C-C motif) ligand 2	7.502674
ABHD2: abhydrolase domain containing 2	7.50102
WISP1: WNT1 inducible signaling pathway protein 1	7.498314
IL13RA1: interleukin 13 receptor, alpha 1	7.489474
IRF1: interferon regulatory factor 1	7.475458
C19orf66: RIKEN cDNA A230050P20 gene	7.46119
CRISPLD2: cysteine-rich secretory protein LCCL domain containing 2	7.454726
TMEM173: transmembrane protein 173	7.454182
GPM6B: glycoprotein m6b	7.444992
ZEB1: zinc finger E-box binding homeobox 1	7.428174
ALDH1L1: aldehyde dehydrogenase 1 family, member L1	7.372892
SFXN5: sideroflexin 5	7.35222
CASP4: caspase 4, apoptosis-related cysteine peptidase	7.34923
GNG2: guanine nucleotide binding protein (G protein), gamma 2	7.338824
EVI2A: ecotropic viral integration site 2a	7.278604
COL5A1: collagen, type V, alpha 1	7.268014
LAMC1: laminin, gamma 1	7.251682
SOCS3: suppressor of cytokine signaling 3	7.1934

ADAMTS7: a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 7	7.165642
CTHRC1: collagen triple helix repeat containing 1	7.148332
ADAMTS9: a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 9	7.059632
COL5A2: collagen, type V, alpha 2	7.046382
IL24: interleukin 24	7.027138
ACPP: acid phosphatase, prostate	7.022562
PLA1A: phospholipase A1 member A	7.019714
CAPN6: calpain 6	6.951942
MPP2: membrane protein, palmitoylated 2 (MAGUK p55 subfamily member 2)	6.944702
DOCK5: dedicator of cytokinesis 5	6.93553
XDH: xanthine dehydrogenase	6.871186
CEP170: centrosomal protein 170	6.866698
LRRC15: leucine rich repeat containing 15	6.861134
E430024C06RIK: RIKEN cDNA E430024C06 gene	6.855516
PRRG1: proline rich Gla (G-carboxyglutamic acid) 1	6.843482
C1RA: complement component 1, r subcomponent A	6.832696
CYP4V2: cytochrome P450, family 4, subfamily v, polypeptide 3	6.827202
NMNAT2: nicotinamide nucleotide adenylyltransferase 2	6.794494
TNFRSF14: tumor necrosis factor receptor superfamily, member 14 (herpesvirus entry mediator)	6.76943
1110012D08RIK: RIKEN cDNA 1110012D08 gene	6.769356
1700001L05RIK: RIKEN cDNA 1700001L05 gene	6.75913
ERRFI1: ERBB receptor feedback inhibitor 1	6.756948
C10orf11: RIKEN cDNA 1700112E06 gene	6.72905
MAP1A: microtubule-associated protein 1 A	6.701708
ADAMTS12: a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 12	6.6954
ZCCHC11: zinc finger, CCHC domain containing 11	6.666628
CD274: CD274 antigen	6.64963
TAP1: transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	6.633806

IL1RL2: interleukin 1 receptor-like 2	6.616754
TWIST2: twist homolog 2 (Drosophila)	6.614856
TGM2: transglutaminase 2, C polypeptide	6.614796
ZWINT: ZW10 interactor	6.589234
ADAM8: a disintegrin and metallopeptidase domain 8	6.515504
PLA2G15: phospholipase A2, group XV	6.500676
MAP3K14: mitogen-activated protein kinase kinase kinase 14	6.490254
ZEB2: zinc finger E-box binding homeobox 2	6.488012
EPB41L3: erythrocyte protein band 4.1-like 3	6.48196
IL2RB: interleukin 2 receptor, beta chain	6.478894
OGDHL: oxoglutarate dehydrogenase-like	6.466006
CNN2: calponin 2	6.44535
GBP1: guanylate binding protein 1	6.423322
RGS17: regulator of G-protein signaling 17	6.395304
IL10RB: interleukin 10 receptor, beta	6.385582
PFKL: phosphofructokinase, liver, B-type	6.37849
SPRED3: sprouty-related, EVH1 domain containing 3	6.365388
PSMB10: proteasome (prosome, macropain) subunit, beta type 10	6.326606
IL17RC: interleukin 17 receptor C	6.301282
ABCG2: ATP-binding cassette, sub-family G (WHITE), member 2	6.27304
HLA-G: histocompatibility 2, M region locus 3	6.270868
PTPRJ: protein tyrosine phosphatase, receptor type, J	6.25899
WISP2: WNT1 inducible signaling pathway protein 2	6.218288
ACSL4: acyl-CoA synthetase long-chain family member 4	6.180854
BIN1: bridging integrator 1	6.167242
LPGAT1: lysophosphatidylglycerol acyltransferase 1	6.150548
SYT13: synaptotagmin XIII	6.110474
MOGAT1: monoacylglycerol O-acyltransferase 1	6.107488
RGS4: regulator of G-protein signaling 4	6.104788
NBL1: neuroblastoma, suppression of tumorigenicity 1	6.09802
C1S: complement component 1, s subcomponent	6.096338
MAGED1: melanoma antigen, family D, 1	6.04951
MAPK10: mitogen-activated protein kinase 10	6.045484
MX1: myxovirus (influenza virus) resistance 1	6.032856

GBP2: guanylate binding protein 2	5.99567
TTPA: tocopherol (alpha) transfer protein	5.979902
IGF1: insulin-like growth factor 1	5.937076
GVIN1: GTPase, very large interferon inducible 1	5.913264
CSPRS: component of Sp100-rs	5.910162
PARVB: parvin, beta	5.909594
EFNB2: ephrin B2	5.903746
ROR1: receptor tyrosine kinase-like orphan receptor 1	5.88385
MAP4K3: mitogen-activated protein kinase kinase kinase kinase 3	5.872332
MAGED2: melanoma antigen, family D, 2	5.86719
DPY19L3: dpy-19-like 3 (C. elegans)	5.848032
RARB: retinoic acid receptor, beta	5.839536
WNT10A: wingless related MMTV integration site 10a	5.813974
SEC23A: SEC23A (S. cerevisiae)	5.772238
TCEAL3: transcription elongation factor A (SII)-like 3	5.770892
DNM1: dynamin 1	5.769356
ROBO2: roundabout homolog 2 (Drosophila)	5.767156
DTX4: deltex 4 homolog (Drosophila)	5.76639
GM7609: predicted pseudogene 7609	5.762914
LRRK2: leucine-rich repeat kinase 2	5.76126
ANO1: anoctamin 1, calcium activated chloride channel	5.760528
CMBL: carboxymethylenebutenolidase-like (Pseudomonas)	5.753436
NLRC5: NLR family, CARD domain containing 5	5.697468
FAM46A: family with sequence similarity 46, member A	5.697402
PRDX5: peroxiredoxin 5	5.68189
MOV10: Moloney leukemia virus 10	5.673906
GLIPR1: GLI pathogenesis-related 1 (glioma)	5.669162
QSOX1: quiescin Q6 sulfhydryl oxidase 1	5.665106
LRP1: low density lipoprotein receptor-related protein 1	5.662242
RGL1: ral guanine nucleotide dissociation stimulator,-like 1	5.6615
IL1A: interleukin 1 alpha	5.652912
SMARCA1: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	5.628166
AOX1: aldehyde oxidase 1	5.626658

DGAT2: diacylglycerol O-acyltransferase 2	5.610228
LAMA4: laminin, alpha 4	5.605086
ENSMUST00000083804: ncrna:snoRNA chromosome: NCBIM37:3:153574574:153574654:-1 gene:ENSMUSG00000065738	5.600088
IFI47: interferon gamma inducible protein 47	5.580488
CCL5: chemokine (C-C motif) ligand 5	5.56656
WNT5A: wingless-related MMTV integration site 5A	5.562396
SMOX: spermine oxidase	5.560542
5730471H19RIK: RIKEN cDNA 5730471H19 gene	5.554104
CLCA1: chloride channel calcium activated 1	5.550486
NDUFV3: NADH dehydrogenase (ubiquinone) flavoprotein 3	5.550022
UGT1A7: UDP glucuronosyltransferase 1 family, polypeptide A9	5.543746
MICALL2: MICAL-like 2	5.543596
DHX40: DEAH (Asp-Glu-Ala-His) box polypeptide 40	5.528458
ALG11: asparagine-linked glycosylation 11 homolog (yeast, alpha-1,2-mannosyltransferase)	5.518064
FHOD1: formin homology 2 domain containing 1	5.514536
SPSB1: splA/ryanodine receptor domain and SOCS box containing 1	5.514256
SOX12: SRY-box containing gene 12	5.509004
GDA: guanine deaminase	5.503634
PSME1: proteasome (prosome, macropain) 28 subunit, alpha	5.503404
TSKU: tsukushin	5.501144
GRAMD1B: GRAM domain containing 1B	5.497222
SFMBT2: Scm-like with four mbt domains 2	5.434594
EHD1: EH-domain containing 1	5.432856
H2-Q8: histocompatibility 2, Q region locus 8	5.390638
KCNK5: potassium channel, subfamily K, member 5	5.3873
PLEKHB1: pleckstrin homology domain containing, family B (evectins) member 1	5.381404
PLOD1: procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1	5.378574
MTMR11: myotubularin related protein 11	5.378006
GM12250: predicted gene 12250	5.354878

IL11: interleukin 11	5.350254
HDAC7: histone deacetylase 7	5.32821
SECTM1: secreted and transmembrane 1A	5.323478
BRSK1: BR serine/threonine kinase 1	5.318198
ADHFE1: alcohol dehydrogenase, iron containing, 1	5.30729
NTRK3: neurotrophic tyrosine kinase, receptor, type 3	5.307188
PDK3: pyruvate dehydrogenase kinase, isoenzyme 3	5.303276
CFB: complement factor B	5.301412
MASP1: mannan-binding lectin serine peptidase 1	5.29542
PGM1: phosphoglucomutase 2	5.285684
APOL10B: apolipoprotein L 10b	5.281456
PTPLAD2: protein tyrosine phosphatase-like A domain containing 2	5.278678
C130026I21RIK: RIKEN cDNA C130026I21 gene	5.27069
KCND1: potassium voltage-gated channel, Shal-related family, member 1	5.252898
DOCK6: dedicator of cytokinesis 6	5.240372
C1S: predicted gene 5077	5.227192
KHNYN: KH and NYN domain containing	5.214064
TNFRSF19: tumor necrosis factor receptor superfamily, member 19	5.210566
NDST3: N-deacetylase/N-sulfotransferase (heparan glucosaminyl) 3	5.20305
LYN: Yamaguchi sarcoma viral (v-yes-1) oncogene homolog	5.194046
DOCK11: dedicator of cytokinesis 11	5.176526
AKR1B15: aldo-keto reductase family 1, member B8	5.170028
TWIST1: twist homolog 1 (Drosophila)	5.153024
CH25H: cholesterol 25-hydroxylase	5.14672
SESN2: sestrin 2	5.133962
PCBD1: pterin 4 alpha carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1) 1	5.130876
ZNF217: zinc finger protein 217	5.108424
FAP: fibroblast activation protein	5.09971
EMP3: epithelial membrane protein 3	5.09465

NAV3: neuron navigator 3	5.056046
KCNJ2: potassium inwardly-rectifying channel, subfamily J, member 2	5.049972
ITPR1: inositol 1,4,5-triphosphate receptor 1	5.046436
PDLIM4: PDZ and LIM domain 4	5.037228
GSDMD: gasdermin D	5.002584
PTPLA: protein tyrosine phosphatase-like (proline instead of catalytic arginine), member a	4.997184
BCL9: B-cell CLL/lymphoma 9	4.993898
ADCY6: adenylate cyclase 6	4.98849
LOC100507705: histocompatibility 2, D region locus 1	4.985236
USP47: ubiquitin specific peptidase 47	4.980352
CAPN1: calpain 1	4.976448
MXRA8: matrix-remodelling associated 8	4.967828
SEC24D: Sec24 related gene family, member D (S. cerevisiae)	4.94578
EHBP1L1: EH domain binding protein 1-like 1	4.929738
MYO1B: myosin IB	4.929192
CACNA2D1: calcium channel, voltage-dependent, alpha2/delta subunit 1	4.925598
SARDH: sarcosine dehydrogenase	4.916744
KSR1: kinase suppressor of ras 1	4.916438
DHRS7: dehydrogenase/reductase (SDR family) member 7	4.913564
H2-Q7: histocompatibility 2, Q region locus 7	4.895526
SYTL5: synaptotagmin-like 5	4.871142
SHISA4: shisa homolog 4 (Xenopus laevis)	4.85625
MAF: avian musculoaponeurotic fibrosarcoma (v-maf) AS42 oncogene homolog	4.851246
CD33: CD33 antigen	4.823432
ENSMUST00000082947: ncrna:snRNA chromosome:NCBIM37:1:194571765:194571880:1 gene:ENSMUSG00000064881	4.820484
PODNL1: podocan-like 1	4.819872
SERPINB6B: serine (or cysteine) peptidase inhibitor, clade B, member 6b	4.810342

NOL3: nucleolar protein 3 (apoptosis repressor with CARD domain)	4.791564
5430435G22RIK: RIKEN cDNA 5430435G22 gene	4.780464
GSN: gelsolin	4.779
SIPA1: signal-induced proliferation associated gene 1	4.757904
TRAF5: TNF receptor-associated factor 5	4.749348
MAP3K6: mitogen-activated protein kinase kinase kinase 6	4.745088
HDAC1: histone deacetylase 1	4.742912
TNN: tenascin N	4.740232
UACA: uveal autoantigen with coiled-coil domains and ankyrin repeats	4.736762
DAPK1: death associated protein kinase 1	4.734688
UGT8: UDP galactosyltransferase 8A	4.733262
MIF: macrophage migration inhibitory factor	4.727292
TMEM26: transmembrane protein 26	4.726652
TFAP2B: transcription factor AP-2 beta	4.723448
MOGAT2: monoacylglycerol O-acyltransferase 2	4.721468
H2-T9: MHC class Ib T9	4.692952
VAV3: vav 3 oncogene	4.681828
NFIL3: nuclear factor, interleukin 3, regulated	4.681006
C14orf159: RIKEN cDNA 9030617O03 gene	4.662966
LIF: leukemia inhibitory factor	4.657554
LHFP: lipoma HMGIC fusion partner	4.657268
MGST1: microsomal glutathione S-transferase 1	4.643716
FAM38B: family with sequence similarity 38, member B	4.627878
COL16A1: collagen, type XVI, alpha 1	4.62164
TRIP10: thyroid hormone receptor interactor 10	4.615604
ITGA8: integrin alpha 8	4.586038
PVR: poliovirus receptor	4.582454
NCAM1: neural cell adhesion molecule 1	4.563038
9330175E14RIK: RIKEN cDNA 9330175E14 gene	4.556136
GLRX: glutaredoxin	4.554842
EIF4E3: eukaryotic translation initiation factor 4E member 3	4.551438

MICAL2: microtubule associated monoxygenase, calponin and LIM domain containing 2	4.52316
HSPA12A: heat shock protein 12A	4.520906
ZBP1: Z-DNA binding protein 1	4.470346
ZFPM2: zinc finger protein, multitype 2	4.465252
MAP1B: microtubule-associated protein 1B	4.460526
ADAM5: a disintegrin and metallopeptidase domain 5	4.457218
RNF152: ring finger protein 152	4.442844
CCNJL: cyclin J-like	4.43516
ZNF25: zinc finger protein 9	4.420936
LRRC51: leucine rich repeat containing 51	4.413628
SAMD12: sterile alpha motif domain containing 12	4.412232
CLIP2: CAP-GLY domain containing linker protein 2	4.41084
EXT1: exostoses (multiple) 1	4.400332
CREBZF: CREB/ATF bZIP transcription factor	4.393872
FOXF2: forkhead box F2	4.388796
SLCO2A1: solute carrier organic anion transporter family, member 2a1	4.371776
GM12185: predicted gene 12185	4.363836
IL2RG: interleukin 2 receptor, gamma chain	4.359164
ING5: inhibitor of growth family, member 5	4.355152
CHN1: chimerin (chimaerin) 1	4.333484
GDNF: glial cell line derived neurotrophic factor	4.33084
GALK1: galactokinase 1	4.31503
OLFR920: olfactory receptor 920	4.300274
ENSMUST00000131638: cdna:pseudogene chromosome:NCBIM37:18:60380497:60381712:1 gene:ENSMUSG00000085977	4.271202
SRGAP3: SLIT-ROBO Rho GTPase activating protein 3	4.269712
DZIP3: DAZ interacting protein 3, zinc finger	4.267668
AGPAT4: 1-acylglycerol-3-phosphate O-acyltransferase 4 (lysophosphatidic acid acyltransferase, delta)	4.266268
ABCC1: ATP-binding cassette, sub-family C (CFTR/MRP), member 1	4.258624

MUM1L1: melanoma associated antigen (mutated) 1-like 1	4.23835
TMSB10: thymosin, beta 10	4.210836
TTC7B: tetratricopeptide repeat domain 7B	4.19573
MSL3L2: male-specific lethal 3-like 2 (Drosophila)	4.190086
HSPA13: heat shock protein 70 family, member 13	4.189054
WSB1: WD repeat and SOCS box-containing 1	4.186552
FOXO3: forkhead box O3	4.184868
RASA3: RAS p21 protein activator 3	4.177786
APAF1: apoptotic peptidase activating factor 1	4.17759
GENSCAN00000019204: cdna:Genscan chromosome:NCBIM37:12:78724538:78725888:-1	4.176874
SAMD14: sterile alpha motif domain containing 14	4.161346
TAP2: transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	4.145106
PRUNE2: prune homolog 2 (Drosophila)	4.14501
RTN2: reticulon 2 (Z-band associated protein)	4.124482
SERPINH1: serine (or cysteine) peptidase inhibitor, clade H, member 1	4.10895
RNF128: ring finger protein 128	4.105938
H2-BL: histocompatibility 2, blastocyst	4.092138
MATN2: matrilin 2	4.088498
ENSMUST00000120523: cdna:pseudogene chromosome:NCBIM37:4:26428590:26428806:-1 gene:ENSMUSG00000081985	4.088426
MTHFR: 5,10-methylenetetrahydrofolate reductase	4.087428
CPT1C: carnitine palmitoyltransferase 1c	4.071418
ADAM19: a disintegrin and metallopeptidase domain 19 (meltrin beta)	4.069138
SYNGAP1: synaptic Ras GTPase activating protein 1 homolog (rat)	4.06779
GNL1: guanine nucleotide binding protein-like 1	4.058942
MMS19: MMS19 (MET18 S. cerevisiae)	4.033364
PTK7: PTK7 protein tyrosine kinase 7	4.02937
ZNF580: zinc finger protein 580	4.02728
MAP3K13: mitogen-activated protein kinase kinase kinase 13	4.02382

ZNF462: zinc finger protein 462	3.996878
FRMD6: FERM domain containing 6	3.985256
ILF3: interleukin enhancer binding factor 3	3.98373
GFOD1: glucose-fructose oxidoreductase domain containing 1	3.975788
RPS6KA3: ribosomal protein S6 kinase polypeptide 3	3.970404
NES: nestin	3.967854
SLC39A4: solute carrier family 39 (zinc transporter), member 4	3.96744
MYC: myelocytomatosis oncogene	3.962424
RAP1A: RAS-related protein-1a	3.9326
TMEM88: transmembrane protein 88	3.92855
UBFD1: ubiquitin family domain containing 1	3.928042
TNFAIP3: tumor necrosis factor, alpha-induced protein 3	3.91999
IL7: interleukin 7	3.90937
1600029D21RIK: RIKEN cDNA 1600029D21 gene	3.90887
NT5DC2: 5'-nucleotidase domain containing 2	3.903428
PRSS22: protease, serine, 22	3.89565
DLG4: discs, large homolog 4 (Drosophila)	3.8843
S1PR2: sphingosine-1-phosphate receptor 2	3.884098
PLAUR: plasminogen activator, urokinase receptor	3.881686
STAT2: signal transducer and activator of transcription 2	3.870038
PSME2: proteasome (prosome, macropain) 28 subunit, beta	3.863824
THY1: thymus cell antigen 1, theta	3.851126
TMEM200A: transmembrane protein 200A	3.849522
GM8773: predicted gene 8773	3.84766
NTN1: netrin 1	3.829342
KLRB1A: killer cell lectin-like receptor subfamily B member 1A	3.827652
---	3.826734
MAP3K1: mitogen-activated protein kinase kinase kinase 1	3.825828
PTK2: PTK2 protein tyrosine kinase 2	3.82391
MIR542: microRNA 542	3.816764
ARNTL2: aryl hydrocarbon receptor nuclear translocator-like 2	3.809436
UBE2Q2: ubiquitin-conjugating enzyme E2Q (putative) 2	3.7575
SLC30A4: solute carrier family 30 (zinc transporter), member 4	3.753386
ABCB6: ATP-binding cassette, sub-family B (MDR/TAP), member 6	3.733434

C4orf31: RIKEN cDNA A930038C07 gene	3.722774
CD63: CD63 antigen	3.70621
GSTT3: glutathione S-transferase, theta 3	3.700208
PION: pigeon homolog (Drosophila)	3.69354
ROBO1: roundabout homolog 1 (Drosophila)	3.678134
EPN2: epsin 2	3.676534
TIMP1: tissue inhibitor of metalloproteinase 1	3.671724
IPO5: importin 5	3.669546
ENSMUST00000158294: ncrna:miRNA chromosome:NCBIM37:4:48775748:48775833:-1 gene:ENSMUSG00000088919	3.65477
PPP1R12B: protein phosphatase 1, regulatory (inhibitor) subunit 12B	3.63964
NCOA7: nuclear receptor coactivator 7	3.632974
PLSCR4: phospholipid scramblase 4	3.619856
NUDT7: nudix (nucleoside diphosphate linked moiety X)-type motif 7	3.618948
MAP4K1: mitogen-activated protein kinase kinase kinase kinase 1	3.616794
FNBP1: formin binding protein 1	3.615896
MUSTN1: musculoskeletal, embryonic nuclear protein 1	3.613478
KIAA1949: RIKEN cDNA 2310014H01 gene	3.612302
RUSC2: RUN and SH3 domain containing 2	3.610172
SEPT11: septin 11	3.60922
TCIRG1: T-cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 protein A3	3.596978
ADARB1: adenosine deaminase, RNA-specific, B1	3.59415
ZFAND1: zinc finger, AN1-type domain 1	3.591156
SIPA1L2: signal-induced proliferation-associated 1 like 2	3.581286
COL5A3: collagen, type V, alpha 3	3.558438
CLIC1: chloride intracellular channel 1	3.55021
CYR61: cysteine rich protein 61	3.547526
ELOVL5: ELOVL family member 5, elongation of long chain fatty acids (yeast)	3.541982
RNF114: ring finger protein 114	3.532772

STYK1: serine/threonine/tyrosine kinase 1	3.531864
ZSWIM4: zinc finger, SWIM domain containing 4	3.523504
CDK2AP2: CDK2-associated protein 2	3.519688
DYRK1B: dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1b	3.518778
FMNL3: formin-like 3	3.515658
DOCK10: dedicator of cytokinesis 10	3.472358
OSMR: oncostatin M receptor	3.46149
IPO7: importin 7	3.456632
GPD1: glycerol-3-phosphate dehydrogenase 1 (soluble)	3.45458
FCHSD2: FCH and double SH3 domains 2	3.438284
INTU: inturned planar cell polarity effector homolog (Drosophila)	3.437812
KARS: lysyl-tRNA synthetase	3.415458
FBXL19: F-box and leucine-rich repeat protein 19	3.403526
HS3ST1: heparan sulfate (glucosamine) 3-O-sulfotransferase 1	3.398818
RRBP1: ribosome binding protein 1	3.395398
AGK: acylglycerol kinase	3.394408
RDH5: retinol dehydrogenase 5	3.36793
CD47: CD47 antigen (Rh-related antigen, integrin-associated signal transducer)	3.364326
ENSMUST00000119692: cdna:pseudogene chromosome:NCBIM37:X:96276752:96277741:-1 gene:ENSMUSG00000083310	3.346672
KCNAB2: potassium voltage-gated channel, shaker-related subfamily, beta member 2	3.306238
BLVRB: biliverdin reductase B (flavin reductase (NADPH))	3.305088
PLCB1: phospholipase C, beta 1	3.295772
ENSMUST00000082933: ncrna:snRNA chromosome:NCBIM37:3:86188169:86188275:1 gene:ENSMUSG00000064867	3.29292
ARNT: aryl hydrocarbon receptor nuclear translocator	3.29249
PLA2R1: phospholipase A2 receptor 1	3.28831
TSPYL2: TSPY-like 2	3.280206
DDX60: DEAD (Asp-Glu-Ala-Asp) box polypeptide 60	3.248216

PBX2: pre B-cell leukemia transcription factor 2	3.215128
ITM2C: integral membrane protein 2C	3.203024
CDKN1B: cyclin-dependent kinase inhibitor 1B	3.199272
ERCC1: excision repair cross-complementing rodent repair deficiency, complementation group 1	3.196868
MORF4L1: mortality factor 4 like 1	3.183468
MPG: N-methylpurine-DNA glycosylase	3.177062
ACSS2: acyl-CoA synthetase short-chain family member 2	3.170412
FSTL3: follistatin-like 3	3.139956
ENSMUST00000083377: ncrna:rRNA chromosome:NCBIM37:7:106646228:106646332:1 gene:ENSMUSG00000065311	3.108382
FAM13C: family with sequence similarity 13, member C	3.101654
CFLAR: CASP8 and FADD-like apoptosis regulator	3.077924
MAP3K12: mitogen-activated protein kinase kinase kinase 12	3.064196
PROSC: proline synthetase co-transcribed	3.037702
SYT1: synaptotagmin I	3.034176
FYN: Fyn proto-oncogene	3.026374
WLS: wntless homolog (Drosophila)	3.001854
PICALM: phosphatidylinositol binding clathrin assembly protein	2.971864
CHST1: carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	2.968118
ATF6: activating transcription factor 6	2.95465
GAMT: guanidinoacetate methyltransferase	2.948754
RLIM: ring finger protein, LIM domain interacting	2.85346

Appendix 2

Down regulated by Id4 knockdown and up regulated by Id4 overexpression.

Gene.Symbol.x	Adjusted average t
MSRB3: methionine sulfoxide reductase B3	14.82609
KLHDC5: kelch domain containing 5	14.33727
NEBL: nebulette	14.26102
EFNA3: ephrin A3	13.96849
PRELP: proline arginine-rich end leucine-rich repeat	12.89036
ZDHHC23: zinc finger, DHHC domain containing 23	12.80888
FHL1: four and a half LIM domains 1	12.11966
LARP4B: La ribonucleoprotein domain family, member 4B	11.61751
GYG1: glycogenin	11.23247
GNAI1: guanine nucleotide binding protein (G protein), alpha inhibiting 1	10.53526
GSTA1: glutathione S-transferase, alpha 1 (Ya)	10.23805
PPIF: peptidylprolyl isomerase F (cyclophilin F)	10.14284
HSD17B4: hydroxysteroid (17-beta) dehydrogenase 4	9.6323
GRAMD2: GRAM domain containing 2	9.601078
SLC46A3: solute carrier family 46, member 3	9.597712
STAMBP: STAM binding protein	9.53788
GM1006: predicted gene 1006	9.507778
ASB13: ankyrin repeat and SOCS box-containing 13	9.454784
EIF5A2: eukaryotic translation initiation factor 5A2	9.274902
SLC44A1: solute carrier family 44, member 1	8.883482
ALDH3B2: aldehyde dehydrogenase 3 family, member B2	8.855544
TMEM20: transmembrane protein 20	8.792032
FYB: FYN binding protein	8.684138
SPRYD4: SPRY domain containing 4	8.665534
NUTF2: nuclear transport factor 2	8.644532
MATN4: matrilin 4	8.562804

DSTYK: dual serine/threonine and tyrosine protein kinase	8.515326
FN3KRP: fructosamine 3 kinase related protein	8.28647
PERP: PERP, TP53 apoptosis effector	8.148354
RAB2B: RAB2B, member RAS oncogene family	8.034464
PPARG: peroxisome proliferator activated receptor gamma	7.899976
POSTN: periostin, osteoblast specific factor	7.861616
DSG1: desmoglein 1 alpha	7.682772
SPINK5: serine peptidase inhibitor, Kazal type 5	7.638164
UGP2: UDP-glucose pyrophosphorylase 2	7.590292
NQO1: NAD(P)H dehydrogenase, quinone 1	7.563116
MILL1: MHC I like leukocyte 1	7.559496
PHTF2: putative homeodomain transcription factor 2	7.466502
PDGFA: platelet derived growth factor, alpha	7.388368
MECOM: MDS1 and EVI1 complex locus	7.385816
UBXN2A: UBX domain protein 2A	7.375708
MED30: mediator complex subunit 30	7.194626
SUMF1: sulfatase modifying factor 1	7.032166
FBP2: fructose bisphosphatase 2	7.019036
MARCH8: membrane-associated ring finger (C3HC4) 8	6.980202
CENPV: centromere protein V	6.969924
MOBKL1A: MOB1, Mps One Binder kinase activator-like 1A (yeast)	6.929948
TIMP2: tissue inhibitor of metalloproteinase 2	6.874424
RAB7L1: RAB7, member RAS oncogene family-like 1	6.779438
MST4: RIKEN cDNA 2610018G03 gene	6.771242
ZFP2: zinc finger protein 2	6.733252
ATP6V0A4: ATPase, H+ transporting, lysosomal V0 subunit A4	6.728264
C16orf89: expressed sequence AU021092	6.703818
SLC45A3: solute carrier family 45, member 3	6.651644
MAVS: mitochondrial antiviral signaling protein	6.561406
MTMR12: myotubularin related protein 12	6.488698
ZNF239: zinc finger protein 239	6.486854
OBFC2B: oligonucleotide/oligosaccharide-binding fold containing 2B	6.475738

MGP: matrix Gla protein	6.469208
LPIN1: lipin 1	6.465144
GLTP: glycolipid transfer protein	6.449718
HRASLS: HRAS-like suppressor	6.396292
BARX2: BarH-like homeobox 2	6.395054
RUFY3: RUN and FYVE domain containing 3	6.373824
NKD2: naked cuticle 2 homolog (Drosophila)	6.31073
CXCL16: chemokine (C-X-C motif) ligand 16	6.304684
ATP6V1C2: ATPase, H+ transporting, lysosomal V1 subunit C2	6.254556
MBLAC2: metallo-beta-lactamase domain containing 2	6.230718
HEATR6: HEAT repeat containing 6	6.18235
ANKRD52: ankyrin repeat domain 52	6.164772
RNF39: ring finger protein 39	6.12008
SLC9A3R2: solute carrier family 9 (sodium/hydrogen exchanger), member 3 regulator 2	6.04874
PRPS2: phosphoribosyl pyrophosphate synthetase 2	6.044924
GALC: galactosylceramidase	6.014402
VPS54: vacuolar protein sorting 54 (yeast)	6.005556
MYO19: myosin XIX	5.981204
CERK: ceramide kinase	5.96388
DLX6: distal-less homeobox 6	5.954028
PIP4K2C: phosphatidylinositol-5-phosphate 4-kinase, type II, gamma	5.874204
SH3YL1: Sh3 domain YSC-like 1	5.854632
SLC25A42: solute carrier family 25, member 42	5.827054
ZC3H12C: zinc finger CCCH type containing 12C	5.791696
C18orf32: cDNA sequence BC031181	5.766444
CLOCK: circadian locomotor output cycles kaput	5.718766
TNFRSF10A: tumor necrosis factor receptor superfamily, member 10b	5.69573
ST3GAL6: ST3 beta-galactoside alpha-2,3-sialyltransferase 6	5.686692
EIF2C4: eukaryotic translation initiation factor 2C, 4	5.627992
SERTAD4: SERTA domain containing 4	5.626894
GRPEL2: GrpE-like 2, mitochondrial	5.617178

KIAA1543: RIKEN cDNA 2310057J16 gene	5.5895
PACS2: phosphofurin acidic cluster sorting protein 2	5.565166
VSNL1: visinin-like 1	5.456132
PLEKHA1: pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 1	5.450674
GPD1L: glycerol-3-phosphate dehydrogenase 1-like	5.433508
WRB: tryptophan rich basic protein	5.408784
RFNG: RFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase	5.390726
SOCS4: suppressor of cytokine signaling 4	5.382486
PENK: preproenkephalin	5.345086
LNPEP: leucyl/cystinyl aminopeptidase	5.340102
PRKAR2A: protein kinase, cAMP dependent regulatory, type II alpha	5.332088
B4GALT7: xylosylprotein beta1,4-galactosyltransferase, polypeptide 7 (galactosyltransferase I)	5.321962
ANAPC11: anaphase promoting complex subunit 11	5.30827
IMPAD1: inositol monophosphatase domain containing 1	5.292542
---	5.252052
---	5.252052
GPR63: G protein-coupled receptor 63	5.2471
GM10524: predicted gene 10524	5.202054
WBSCR27: Williams Beuren syndrome chromosome region 27 (human)	5.199004
RHBDD1: rhomboid domain containing 1	5.188768
KBTBD8: kelch repeat and BTB (POZ) domain containing 8	5.156434
TMEM144: transmembrane protein 144	5.149952
UFM1: ubiquitin-fold modifier 1	5.133878
MARCH3: membrane-associated ring finger (C3HC4) 3	5.122206
SMAD2: MAD homolog 2 (Drosophila)	5.10897
A930004D18RIK: RIKEN cDNA A930004D18 gene	5.103106
L2HGDH: L-2-hydroxyglutarate dehydrogenase	5.10104
ATG14: VATG14 autophagy related 14 homolog (S. cerevisiae)	5.07677
TMOD2: tropomodulin 2	5.056854

SNX13: sorting nexin 13	5.055612
FOXJ2: forkhead box J2	5.050878
---	5.04943
---	5.04943
---	5.04943
---	5.04943
---	5.04943
PEX1: peroxisomal biogenesis factor 1	5.033288
RNF144B: ring finger protein 144B	5.030834
CDK19: cyclin-dependent kinase 19	5.02904
GNA15: guanine nucleotide binding protein, alpha 15	4.982206
---	4.97163
---	4.97163
MOBKL2B: MOB1, Mps One Binder kinase activator-like 2B (yeast)	4.958082
NRBF2: nuclear receptor binding factor 2	4.95286
LY6D: lymphocyte antigen 6 complex, locus D	4.942166
LMLN: leishmanolysin-like (metallopeptidase M8 family)	4.915868
2310001H12RIK: RIKEN cDNA 2310001H12 gene	4.88471
ZMAT3: zinc finger matrin type 3	4.881564
TBCE: tubulin-specific chaperone E	4.870868
DHX35: DEAH (Asp-Glu-Ala-His) box polypeptide 35	4.856396
STK38: serine/threonine kinase 38	4.831016
CISD1: CDGSH iron sulfur domain 1	4.824342
ADAMTSL5: ADAMTS-like 5	4.819638
GM10516: predicted gene 10516	4.81863
C4orf52: RIKEN cDNA 1810013D10 gene	4.812322
GATC: glutamyl-tRNA(Gln) amidotransferase, subunit C homolog (bacterial)	4.782896
NPR2: natriuretic peptide receptor 2	4.761634
TLCD1: TLC domain containing 1	4.76062
KRT6B: keratin 6B	4.75765
RECQL5: RecQ protein-like 5	4.746678

ENSMUST00000101817: ncrna:snRNA chromosome:NCBIM37:9:65044379:65044494:-1 gene:ENSMUSG00000075763	4.737582
CDS2: CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 2	4.731908
RTN4IP1: reticulon 4 interacting protein 1	4.731888
RNF38: ring finger protein 38	4.688422
MAP1LC3B: microtubule-associated protein 1 light chain 3 beta	4.653978
MTUS1: mitochondrial tumor suppressor 1	4.646828
MMGT2: membrane magnesium transporter 2	4.634848
EPB41L4B: erythrocyte protein band 4.1-like 4b	4.630294
PRPSAP1: phosphoribosyl pyrophosphate synthetase-associated protein 1	4.629298
GSTA2: glutathione S-transferase, alpha 2 (Yc2)	4.628322
SMAGP: small cell adhesion glycoprotein	4.614052
ACOT6: acyl-CoA thioesterase 6	4.592936
SLC26A7: solute carrier family 26, member 7	4.583098
YKT6: YKT6 homolog (S. Cerevisiae)	4.576852
TUBA4A: tubulin, alpha 4A	4.568474
ID4: inhibitor of DNA binding 4	4.558324
ZNF300: zinc finger protein 300	4.552476
ATP11A: ATPase, class VI, type 11A	4.537288
SGCB: sarcoglycan, beta (dystrophin-associated glycoprotein)	4.534894
APLP1: amyloid beta (A4) precursor-like protein 1	4.521844
GENSCAN00000019148: cdna:Genscan supercontig:NCBIM37:NT_161895:89616:108906:1	4.51252
FAM73A: family with sequence similarity 73, member A	4.463606
TP63: transformation related protein 63	4.4604
ZNF398: zinc finger protein 398	4.447646
SPNS2: spinster homolog 2 (Drosophila)	4.43786
TMEM50B: transmembrane protein 50B	4.430924
KLC1: kinesin light chain 1	4.421802
MBNL2: muscleblind-like 2	4.421568
HOXC5: homeobox C5	4.417606

COMMD8: COMM domain containing 8	4.37827
C5orf41: RIKEN cDNA A930001N09 gene	4.364942
GM2A: GM2 ganglioside activator protein	4.363854
---	4.361422
NCKIPSD: NCK interacting protein with SH3 domain	4.336262
RAB4A: RAB4A, member RAS oncogene family	4.326588
1700012B15RIK: RIKEN cDNA 1700012B15 gene	4.32466
DBP: D site albumin promoter binding protein	4.324048
B4GALT4: UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 4	4.307852
C6orf192: RIKEN cDNA 1110021L09 gene	4.29423
TPP1: tripeptidyl peptidase I	4.293852
ZNF32: zinc finger protein 637	4.289418
ZFYVE9: zinc finger, FYVE domain containing 9	4.286978
ARHGEF18: rho/rac guanine nucleotide exchange factor (GEF) 18	4.282958
DNLZ: DNL-type zinc finger	4.266216
DYNLT3: dynein light chain Tctex-type 3	4.263848
TTL: tubulin tyrosine ligase	4.189294
MOXD1: monooxygenase, DBH-like 1	4.184032
ZNF22: zinc finger protein 422	4.176046
AMIGO1: adhesion molecule with Ig like domain 1	4.17209
N4BP2: NEDD4 binding protein 2	4.163562
HUS1: Hus1 homolog (S. pombe)	4.142514
TTC32: tetratricopeptide repeat domain 32	4.141218
UFSP1: UFM1-specific peptidase 1	4.137902
GDE1: glycerophosphodiester phosphodiesterase 1	4.131998
CNOT6: CCR4-NOT transcription complex, subunit 6	4.117182
COMMD6: COMM domain containing 6	4.107798
ZNF124: zinc finger protein 825	4.100834
SIRT1: sirtuin 1 (silent mating type information regulation 2, homolog) 1 (S. cerevisiae)	4.09656
TDRD3: tudor domain containing 3	4.088292
ENSMUST00000102439: ncrna:miRNA chromosome:NCBIM37:11:64881211:64881309:-1	4.080094

gene:ENSMUSG00000076394	
ZNF664: zinc finger protein 664	4.075942
ZNF76: zinc finger protein 523	4.07577
ABHD13: abhydrolase domain containing 13	4.073086
PYGO1: pygopus 1	4.06553
ARL2BP: ADP-ribosylation factor-like 2 binding protein	4.051354
C3orf39: expressed sequence C85492	4.046972
AK138466: Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched library, clone:A330022I22 product:unclassifiable, full insert sequence.	4.044548
MCTP2: multiple C2 domains, transmembrane 2	4.018874
SFT2D3: SFT2 domain containing 3	4.017644
C3orf23: DNA segment, Chr 9, ERATO Doi 402, expressed	3.999242
SP2: Sp2 transcription factor	3.997732
ADI1: acireductone dioxygenase 1	3.994146
RANBP9: RAN binding protein 9	3.99307
SLC35A5: solute carrier family 35, member A5	3.986542
C11orf65: RIKEN cDNA 4930550C14 gene	3.98296
TSPAN15: tetraspanin 15	3.976478
KLHL23: kelch-like 23 (Drosophila)	3.950026
TESK1: testis specific protein kinase 1	3.949002
ENSMUST00000078052: cdna:novel chromosome:NCBIM37:5:148220550:148220998:1 gene:ENSMUSG00000057157	3.926356
EAF1: ELL associated factor 1	3.905982
SLC26A1: solute carrier family 26 (sulfate transporter), member 1	3.873012
MTR: 5-methyltetrahydrofolate-homocysteine methyltransferase	3.870542
DNAJB11: DnaJ (Hsp40) homolog, subfamily B, member 11	3.848822
GM10265: predicted gene 10265	3.843608
RNF44: ring finger protein 44	3.840194
SLC5A1: solute carrier family 5 (sodium/glucose cotransporter), member 1	3.829612
DNAJC12: DnaJ (Hsp40) homolog, subfamily C, member 12	3.822092

SLC39A3: solute carrier family 39 (zinc transporter), member 3	3.818032
1300003B13RIK: RIKEN cDNA 1300003B13 gene	3.811648
IFNG: interferon gamma	3.773334
LYRM7: LYR motif containing 7	3.77138
ZBTB39: zinc finger and BTB domain containing 39	3.758498
FAM185A: family with sequence similarity 185, member A	3.751532
HELQ: helicase, POLQ-like	3.74913
USP16: ubiquitin specific peptidase 16	3.730524
AB010352: cDNA sequence AB010352	3.729252
MIR132: microRNA 132	3.71711
RRAGC: Ras-related GTP binding C	3.709698
MRAS: muscle and microspikes RAS	3.70625
RPL31: ribosomal protein L31	3.70616
UBL4A: ubiquitin-like 4	3.70373
TRAV14D-1: T-cell receptor alpha variable region 14D-1	3.68406
ENSMUST00000082828: ncrna:snRNA chromosome:NCBIM37:4:67843566:67843736:1 gene:ENSMUSG00000064762	3.683928
RNF181: ring finger protein 181	3.680696
GM6086: predicted gene 6086	3.652612
GM5434: ubiquitin-conjugating enzyme E2F (putative) pseudogene	3.61962
3110052M02RIK: RIKEN cDNA 3110052M02 gene	3.612254
PCYT1A: phosphate cytidylyltransferase 1, choline, alpha isoform	3.607606
KLHL8: kelch-like 8 (Drosophila)	3.604236
KDM4C: lysine (K)-specific demethylase 4C	3.603752
NPN2: neoplastic progression 2	3.585034
TMEM8A: transmembrane protein 8 (five membrane-spanning domains)	3.583206
AI987944: expressed sequence AI987944	3.560498
ULK2: Unc-51 like kinase 2 (C. elegans)	3.553376
C20orf27: RIKEN cDNA 1700037H04 gene	3.552148
C2orf42: expressed sequence C87436	3.533984
NAT6: N-acetyltransferase 6	3.489972

MOCS2: molybdenum cofactor synthesis 2	3.489244
PECR: peroxisomal trans-2-enoyl-CoA reductase	3.478756
OCEL1: occludin/ELL domain containing 1	3.46504
GENSCAN00000034109: cdna:Genscan chromosome:NCBIM37:13:79093443:79094053:-1	3.464086
AGGF1: angiogenic factor with G patch and FHA domains 1	3.452252
GALT: galactose-1-phosphate uridyl transferase	3.44694
SLC16A13: solute carrier family 16 (monocarboxylic acid transporters), member 13	3.442482
GPR174: G protein-coupled receptor 174	3.4259
OSBPL8: oxysterol binding protein-like 8	3.42577
GLO1: glyoxalase 1	3.412582
SPAG11A: sperm associated antigen 11A	3.412372
GNAQ: guanine nucleotide binding protein, alpha q polypeptide	3.40335
TPRA1: transmembrane protein, adipocyte asscociated 1	3.40182
GM5341: predicted pseudogene 5341	3.384628
DICER1: Dicer1, Dcr-1 homolog (Drosophila)	3.380064
CREBL2: cAMP responsive element binding protein-like 2	3.380012
C19orf25: RIKEN cDNA 2310011J03 gene	3.368326
MFF: mitochondrial fission factor	3.361098
NFE2L3: nuclear factor, erythroid derived 2, like 3	3.348414
LYZ1: lysozyme 1	3.342006
PAIP2B: poly(A) binding protein interacting protein 2B	3.33876
GUK1: guanylate kinase 1	3.337732
GM5959: predicted gene 5959	3.32669
KRTAP5-2: keratin associated protein 5-2	3.315538
TMX2: thioredoxin-related transmembrane protein 2	3.3153
SCAMP2: secretory carrier membrane protein 2	3.306884
GZMG: granzyme G	3.305128
ENSMUST00000103701: cdna:novel chromosome:NCBIM37:14:54795946:54796000:1 gene:ENSMUSG00000076889	3.295644
GENSCAN00000024777: cdna:Genscan chromosome:NCBIM37:18:64448012:64448467:1	3.286576

C8orf84: predicted gene 106	3.28359
KRT10: keratin 10	3.273668
E330021D16RIK: ubiquitin-conjugating enzyme E2Q (putative) 2 pseudogene	3.272496
BEND3: BEN domain containing 3	3.266642
6330416L07RIK: RIKEN cDNA 6330416L07 gene	3.236748
HPCAL4: hippocalcin-like 4	3.232494
GENSCAN00000028428: cdna:Genscan chromosome:NCBIM37:14:121214348:121214833:-1	3.216832
A730017L22RIK: RIKEN cDNA A730017L22 gene	3.188972
SLC25A45: solute carrier family 25, member 45	3.18697
4921517L17RIK: RIKEN cDNA 4921517L17 gene	3.186094
HSDL1: hydroxysteroid dehydrogenase like 1	3.171156
GSDMCL-PS: gasdermin C-like, pseudogene	3.158374
TULP4: tubby like protein 4	3.158156
NPFFR2: neuropeptide FF receptor 2	3.136456
ENSMUST00000094539: cdna:known chromosome:NCBIM37:7:48590751:48591593:1 gene:ENSMUSG00000082080	3.100622
TMEM90B: transmembrane protein 90B	3.083788
TMEM43: transmembrane protein 43	3.082214
PSMD11: proteasome (prosome, macropain) 26S subunit, non-ATPase, 11	3.035036
U90926: cDNA sequence U90926	2.946268
UROS: uroporphyrinogen III synthase	2.932754
MT4: metallothionein 4	2.896826
PPAPDC2: phosphatidic acid phosphatase type 2 domain containing 2	2.871094
NCOA4: nuclear receptor coactivator 4	2.65185