Investigating RUNX transcription factors in mammary gland development and breast cancer

Nicola Ferrari

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Beatson Institute for Cancer Research
College of Medical, Veterinary and Life Sciences
University of Glasgow
Glasgow
UK

October 2013

Summary

Breast cancer is the third most common cause of cancer death in the UK, accountable for more than 11000 deaths in 2010 alone (www.cancerresearchuk.org). Developmental pathways commonly required for normal development are often hijacked during tumour progression, so a better understanding of mammary gland development is necessary to fully understand the roots of breast cancer. The *Runx* gene family are known to be important regulators of development in different lineages. In particular RUNX1 and RUNX2 have been widely studied in the context of haematopoiesis and osteogenesis respectively, but their role in epithelial tissue is much less well understood. In this thesis a role for RUNX1 and RUNX2 in mammary development and breast cancer has been identified.

The first part of this study is focused on characterizing the expression and function of the *Runx* genes in the mammary epithelium. RUNX1 and RUNX2 protein levels fluctuate during embryonic and adult mammary development, and an *in vivo* conditional knockout strategy shows that both genes are important for maintenance of mammary epithelium homeostasis. Moreover, combined loss of RUNX1 and RUNX2 significantly perturbs the normal mammary architecture with an expansion of the basal population *in vivo* and the appearance of preneoplastic lesions in aged mammary glands. An exciting new role for RUNX2 in mammary stem cells has also been revealed showing that RUNX2 is important for the regenerative potential of mammary epithelial cells *in vitro*. Evidence is also presented to indicate that RUNX2 could be linked to regulation of quiescence and *Wnt* signalling in the stem cell compartment and during transformation.

Finally, the role of these genes in breast cancer is discussed demonstrating involvement of RUNX1 and RUNX2 specifically in the triple negative (ER-PR-HER2-) subtype. In particular, for the first time, RUNX1 is revealed as an independent prognostic indicator correlating with poor prognosis in triple negative tumours. Meanwhile, evidence from various mouse models demonstrates that RUNX2 may be specifically involved in the squamous metaplastic form of this disease.

Table of Contents

Summary 2
Table of Contents 3
List of Figures 6
Acknowledgements 9
Author's Declaration 10
Abbreviations 11

1 Introduction 15
1.1 Mammary gland development 15
1.1.1 Function and evolution 15
1.1.2 Mammary gland development: an overview 16
1.1.3 Embryonic development 16
1.1.4 Virgin mammary development 20
1.1.5 Pregnancy and lactation 24
1.1.6 Involution 26
1.2 Mammary stem cells 29
1.2.1 Birth of the mammary stem cell concept 30
1.2.2 The troubled pathway towards isolation of MaSCs 31
1.2.3 MaSCs and pregnancy 35
1.2.4 In vitro assays for the study of MaSCs 36
1.2.5 MaSCs regulators: signalling pathways and transcription factors. 37
1.3 Breast cancer 40
1.3.1 Aetiology, progression of disease and treatment 40
1.3.2 Molecular classification of breast cancer 42
1.3.3 Understanding the molecular portrait of breast cancer 45
1.3.4 The origin of breast cancer subtypes 46
1.4 The *RUNX* Genes 47
1.4.1 The Runx genes: evolution, structure and regulation 47
1.4.2 RUNX Genes in Development 56
1.4.3 RUNX Genes and Cancer: an Overview. 62
1.5 *RUNX* genes in mammary development 70
1.6 *RUNX* genes in breast cancer 71

Aims of the thesis 75

2 Materials & Methods 76
2.1 Animals 76
2.1.1 Characterization of the Runx2$^{flx/flx}$ mouse model 76
2.1.2 Other mouse strains 77
2.2 Wholemount/histological analysis of mammary glands 80
2.3 *In vivo* imaging 80
2.4 Fat pad transplantation 81
2.5 Cell lines 81
2.5.1 Cell line transfections 81
2.5.2 Cell line assays 82

2.6 Primary mouse mammary cells 83
2.6.1 Flow cytometry/cell sorting 83
2.6.2 MMECs 2D cultures. 87
2.6.3 Mammospheres 87
2.6.4 WNT3a treatments on mammospheres and 2D-MMECs 88
2.6.5 Matrigel Colony forming assay 88
2.7 Immunohistochemistry 89
2.8 Immunofluorescence 89
2.9 Non quantitative PCR 90
2.9.1 PCR determination of Runx2 status. 90
2.9.2 PCR detection of Runx2 mRNA. 91
2.10 Quantitative RT-PCR 91
2.11 Western blot 92
2.12 Human breast samples/tissue microarray 93
2.13 Statistical analysis 93

3 Characterization of the *RUNX* genes in the mammary gland. 94
3.1 INTRODUCTION 94
3.1.1 RUNX proteins in mammary gland development. 94
3.1.2 Experimental Aims. 95
3.2 RESULTS 96
3.2.1 Characterization of RUNX expression in the normal mammary gland. 96
3.2.2 Mammary specific deletion of RUNX1 and RUNX2 102
3.2.3 Effects of combined deletion of Runx1 and Runx2 on virgin mammary development. 112
3.3 DISCUSSION 120

4 A new role for RUNX2 in mammary stem cells 127
4.1 INTRODUCTION 127
4.1.1 Runx genes and stem cells. 127
4.1.2 RUNX, mammary stem cells and Wnt signalling. 129
4.1.3 Experimental Aims. 129
4.2 RESULTS 130
4.2.1 Expression of Runx1 and Runx2 in mammospheres. 130
4.2.2 In vivo deletion of Runx2 in the mammary basal compartment does not affect normal mammary development. 135
4.2.3 In vivo deletion of Runx2 in the K14+ population impairs in vitro MMECs regenerative potential. 139
4.2.4 Runx2 is a downstream target of Wnt signalling in mammary stem cells. 148
4.3 DISCUSSION 151

5 The role of RUNX2 in metaplastic squamous breast cancer. 156
5.1 INTRODUCTION 156
5.1.1 Runx2 in breast cancer 156
5.1.2 Experimental Aims 157
5.2 RESULTS 158
5.2.1 Runx2 expression in mouse models of breast cancer 158
5.2.2 Runx2 in a mouse model of mammary squamous metaplasia. 161
5.2.3 RUNX2 deletion in an in vitro model of triple negative metastatic breast cancer. 167
5.3 DISCUSSION 171

6 A role for RUNX1 in breast cancer 175
6.1 INTRODUCTION 175
6.1.1 RUNX1 in epithelial cancer. 175
6.1.2 RUNX1 in breast cancer 176
6.1.3 Experimental Aims. 177
6.2 RESULTS 178
6.2.1 Expression of RUNX1 in human breast cancer. 178
6.2.2 Effects of RUNX1 overexpression in vitro. 184
6.3 DISCUSSION 190

7 Conclusions and Future Directions 193

Bibliography 201

List of Figures

Figure 1-1 The mammary cell hierarchy 39

Figure 1-2 Post-transcriptional regulation of RUNX proteins. 55

Figure 2-1 Characterization of the *Runx2*$^{flx/flx}$ mouse model. 78

Figure 2-2 *Runx2* deletion after *in vitro* Cre-recombination. 79

Figure 2-3 FACS gating strategy for mammary population profiling. 85

Figure 2-4 Validation of FACS sorting strategy. 86

Figure 3-3-1 Expression of *Runx* genes in mammary luminal and basal populations. 98

Figure 3-3-2 Expression of RUNX1 and RUNX2 in human and murine mammary gland. 99

Figure 3-3 RUNX1 and RUNX2 expression in murine embryonic development. 100

Figure 3-4 RUNX1 and RUNX2 expression in adult mammary development. 101

Figure 3-5 Characterization of BLG-Cre expression through the use of a GFP reporter. 103

Figure 3-6 Characterization of *Runx1* and *Runx2* conditional knock-out mice. 105

Figure 3-7 Effects of loss of RUNX1 and RUNX2 in the virgin mammary gland. 107

Figure 3-8 Effects of RUNX1 and RUNX2 loss in mammary epithelium *in vivo*. 108

Figure 3-9 Effects of RUNX1 and RUNX2 loss on luminal and basal populations *in vivo*. 109

Figure 3-10 Effects of *in vivo Runx2* deletion during the lactation cycle. 111

Figure 3-11 Effects of combined loss of RUNX1 and RUNX2 in the virgin mammary gland.....113

Figure 3-12 Effects of combined loss of RUNX1 and RUNX2 in mammary epithelium *in vivo*. 114

Figure 3-13 Combined loss of RUNX1 and RUNX2 on luminal and basal populations *in vivo*. 115

Figure 3-14 Changes in GFP reporter expression in RUNX1-RUNX2 double knock-out glands.....116

Figure 3-15 Double deletion of RUNX1 and RUNX2 causes alveolar hyperplasia in virgin mice.....118

Figure 3-16 Double deletion of *Runx1* and *Runx2* causes pre-neoplastic lesions in virgin mice.....119

Figure 4-1 Characterization of *Runx* expression in mammospheres. 131

Figure 4-2 RUNX2 protein is enriched in mammospheres. 132

Figure 4-3 Effects of *Runx2* loss on mammospheres formation. 134

Figure 4-4 Effects of loss of *Runx2* in the basal layer of the virgin mammary gland. 136

Figure 4-5 Analysis of *Runx2* loss in the basal lineage using a GFP reporter....137

Figure 4-6 K14-Cre is expressed in the luminal population of the virgin mammary gland.....138

Figure 4-7 *In vivo* loss of *Runx2* in the basal lineage impairs mammosphere formation. 140

Figure 4-8 *Runx1*, *Runx2* and *p21* expression on mammospheres derived from K14-Cre+/*Runx2*$^{flx/flx}$ mice. 141

Figure 4-9 *Runx* expression in matrigel colony-forming assays.....143

Figure 4-10 *In vivo* loss of *Runx2* in the K14+ cells impairs matrigel colony formation. 144

Figure 4-11 *Runx1* and *Runx2* expression on K14-Cre+/*Runx2*$^{flx/flx}$ matrigel colonies. 145

Figure 4-12 *In vivo* regeneration activity of *Runx2* deleted mammary cells. ..147

Figure 4-13 Effects of *Wnt* signalling on mammosphere cultures.....149

Figure 4-14 Effects of Wnt signalling activation on *Runx2* expression in mammospheres. 150

Figure 5-1 RUNX2 expression in mouse models of breast cancer. 159

Figure 5-2 Examples of RUNX2 basal pattern of expression in squamous lesions.....160

Figure 5-3 Effects of ß-catenin stabilization at different stages of mammary development. 163

Figure 5-4 Histological analysis of ß-catenin-induced squamous lesions. 164

Figure 5-5 Effects of ß-catenin stabilization on RUNX2 expression. 165

Figure 5-6 Effects of RUNX2 loss on squamous metaplasia. 166

Figure 5-7 Creation of a stable RUNX2 knockdown in MDA-MB-231 cells. 169

Figure 5-8 Effects of RUNX2 knockdown in MDA-MB-231 cells. 170

Figure 6-1 *RUNX1* expression on multi-cancer gene expression profiling studies. 179

Figure 6-2 *RUNX1* expression in normal versus invasive breast carcinoma. 180

Figure 6-3 Examples of RUNX1 expression in a breast cancer cohort. 181

Figure 6-4 Correlation of RUNX1 expression and survival in a breast cancer cohort. 182

Figure 6-5 RUNX1 expression in different subtypes of breast cancer. 185

Figure 6-6 Relationship between RUNX1 status and standard clinical, pathological, and biological features of triple-negative breast cancer. 186

Figure 6-7 Relationship between RUNX1 status and inflammatory infiltrate/blood vessel invasion in triple-negative breast cancer. 187

Figure 6-8 Expression of RUNX1 in a panel of human breast cancer cell lines. 188

Figure 6-9 *In vitro* effects of RUNX1 overexpression in hMEC-TERT. 189

Acknowledgements

I am particularly indebted to my supervisor Dr. Karen Blyth for excellent supervision and support during the course of my studies, and for teaching me the rigorous application of the scientific method. I would also like to thank my advisor Professor Owen Sansom and Professor Ewan Cameron for useful discussions and input of new ideas.

I am grateful to all the members of my laboratory for the invaluable help offered in developing my technical skills and for their expert advice. A big thanks to Dimitris for help with everything concerning RNA and PCR; Laura for immunohistochemistry and animal work and a special thank you to Susan for teaching me everything else. I would also like to thank all the colleagues at the Institute who have helped me during the course of my studies (too many to be listed).

A big thank you to Tamara, family and friends.

The studies described in this thesis were carried out at the Beatson Institute for Cancer Research with the support of Cancer Research UK.

Author's Declaration

I hereby declare that all the work reported in this thesis is my own unless otherwise stated. None of the work has been previously submitted for any other degree at any other institution. All sources of information used in the preparation of this thesis are indicated by reference.

Nicola Ferrari

October 2013

Abbreviations

AKT	v-akt murine thymoma viral oncogene homolog
ALL	acute lymphoblastic leukaemia
AML	acute myeloid leukaemia
APC	adenomatous polyposis coli (gene)
APC	allophycocyanin (fluorochrome)
AREG	amphiregulin
BC	breast cancer
BLG	beta-lactoglobulin
BMP	bone morphogenetic protein
BRCA1	breast cancer 1, early onset protein
BSA	bovine serum albumin
BSP	bone sialoprotein
CBF-ß	core-binding factor, beta subunit
CCD	cleidocranial dysplasia
cdk(s)	cyclin-dependent kinase(s)
cDNA	complementary DNA
CK	cytokeratin
Cre	cre recombinase protein
DAB	3,3-diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle's Medium
DNA	deoxyribonucleic acid
EGF	epidermal growth factor
EGTA	ethylene glycol tetra-acetic acid
ELF5	E74-like factor 5 (ets domain transcription factor)

EMT	epithelial-mesenchymal transition
EpCAM	epithelial cell adhesion molecule
ER	oestrogen receptor
EtOH	ethanol
F12	Ham's F12 nutrient mixture
FACS	fluorescence-activated cell sorting
FCS	fetal calf serum
FGF	fibroblast growth factor
Flx	floxed
FSC	forward scatter
GAPDH	glyceraldehyde phosphate dehydrogenase
GATA3	GATA binding protein 3
GFP	green fluorescent protein
GH	growth hormone
HER2	v-erb-b2 avian erythroblastic leukaemia viral oncogene homolog 2
HFSC	hair follicle stem cell
HRP	horseradish peroxidase
Ig	immunoglobulin
IGF-1	insulin-like growth factor 1
IHC	immunohistochemistry
IKKß	IKB kinase beta
JAK2	janus kinase 2
kDa	kilodalton
µg	microgram
µl	microliter
µm	micrometer
µM	micromolar

LacZ	beta-D-galactosidase
LEF	lymphoid enhancer-binding factor
LGR5	leucine rich repeat containing G protein coupled receptor 5
LIF	leukaemia inhibitory factor
Lin-	lineage negative
M	molar
MaSC	mammary stem cell
MBC	metaplastic breast cancer
MDS	myelodysplastic syndrome
MEF	mouse embryonic fibroblast
mg	milligram
ml	millilitre
mm	millimetre
mM	millimolar
MMECs	mouse mammary epithelial cells
MMP	matrix metalloproteinase
MMTV	mouse mammary tumour virus
mRNA	messenger ribonucleic acid
MUC1	mucin 1
Myc	v-myc avian myelocytomatosis viral oncogene homolog
NF-kB	nuclear factor of kappa light polypeptide gene enhancer in B-cells
ng	nanogram
nm	nanometre
NMTS	nuclear matrix targeting sequence
P21	cyclin-dependent kinase inhibitor 1A
PBS	phosphate buffered saline
PCR	polymerase chain reaction

Pen/strep	penicillin/streptomycin
PI-MEC	parity-identified mammary cells
PR	progesterone receptor
PRL	prolactin
PTHrP	parathyroid hormone-related protein
Puro	puromycin
RANKL	receptor activator of nuclear factor κB (NF-κB)-ligand
RNA	ribonucleic acid
RT	room temperature
RT-PCR	reverse transcriptase-polymerase chain reaction
RUNX	runt-related transcription factor
sh-RNA	short hairpin RNA
SMBC	squamous metaplastic breast cancer
SRC	v-src avian sarcoma viral oncogene homolog
SSC	side scatter
STAT	signal-transducer and activator of transcription protein
SV40	simian virus 40
TEBs	terminal end buds
TGFß	transforming growth factor beta
TIMP	tissue inhibitors of metalloproteinases
TMA	tissue microarray
TN	triple negative
WAP	whey acidic protein
WNT	wingless-type MMTV integration site
WT	wild type

1 Introduction

1.1 Mammary gland development

To introduce mammary gland development nothing could be more appropriate than the inspiring and elegant words of Daniel Medina, a pioneer of mammary gland research (Medina 1996):

> *"The problems of mammary gland development and function have attracted the attention of scientists for over a century. Since the complete anatomical description of the gland in 1845, it has been studied by physiologists to understand milk and milk secretion, by cell and electron microscopists to understand structure and secretory function, by developmental biologists to understand organ development in the postnatal organism, by endocrinologists to understand hormone action, by molecular biologists to understand regulation of gene expression, by biotechnologists to develop in vivo bioreactors, by cancer biologists and virologists to study the causes and processes of neoplastic transformation, and by clinical scientists to examine, treat, and cure breast cancer. The organ interests and excites the experimental scientist because of its unique development and function, its experimental applicability, its complex biological and cellular interactions, and, despite all this, its inherent simplicity and beauty."*

1.1.1 Function and evolution

The mammary gland is the organ designated to the production, secretion and delivery of milk to the newborn offspring. In 1758, Linnaeus was the first to recognize the uniqueness of mammary glands and on this basis he created the class *Mammalia*, or creatures with 'mammae' or 'breasts'. Several theories have been proposed to explain the origin and the evolution of this glandular organ. On one side it is thought that the mammary gland evolved through the combination of different skin gland populations into a new functional organ (Blackburn 1993) while others support the view that the mammary gland most likely derives from an ancestral apocrine-like gland (Oftedal 2002). However, the exact evolutionary origin of the mammary gland is still unknown due to the absence of any direct fossil evidence of mammary glands.

1.1.2 Mammary gland development: an overview.

The mammary gland is a very dynamic tissue which undergoes dramatic changes during different developmental phases (Richert, 2000; Hennighausen & Robinson, 2005; Watson & Khaled, 2008). From a small ectodermal invagination during the first phases of embryonic development, the mammary gland develops into a highly plastic and expanding duct tree under the influence of pubertal hormones (Brisken & O'Malley 2010). The ductular appearance of the virgin epithelium is then replaced during pregnancy and lactation through a massive event of cell proliferation and differentiation which lead to the formation of mature alveoli, the milk-producing unit of the mammary gland. The end of lactation sees the beginning of involution, a strictly regulated process involving apoptosis and tissue remodelling, which will bring the mammary gland back to a virgin-like state (Watson & Kreuzaler 2011). Several cell types take part in this process; comprising epithelial cells forming the ductal structures which themselves are surrounded by a fatty stroma made up by adipocytes, vascular endothelial cells, blood vessels, fibroblasts and a variety of immune cells (Wiseman & Werb 2002). Given the high complexity and plasticity of this organ, the creation of a mammary gland requires the coordination of many biological processes such as proliferation, invasion and cell death among others. Interestingly all of these developmental pathways are hijacked during the development of breast cancer. Hence a better understanding of the basic biological mechanism regulating normal mammary gland development will help to shed more light on the neoplastic transformation. Given its advantages as an experimental model (Smith, 2012), much of the characterization of mammary gland development has been done in mice. Despite some unavoidable differences, such as tissue architecture and hormonal control, the mouse mammary gland shares similar developmental features with the human one.

1.1.3 Embryonic development

During early embryogenesis, the ectoderm represents the external germ layer, which is composed of the surface ectoderm, neural crest, and neural tube. The surface ectoderm gives rise to the epidermis and other stratified epithelia such as the oral epithelium (Jiménez-Rojo et al. 2012). Then, in those new-formed tissues, a continuous cross-talk between the epithelium and the mesenchyme

gives rise to diverse specialized structures, called ectodermal appendages, such as hair follicles, mammary glands, salivary glands and teeth (Jiménez-Rojo et al. 2012). Interestingly all ectodermal appendages share common morphological features and similar developmental processes during early organogenesis, while later stages of development are characterized by structure-specific morphogenetic programs (Mikkola & Millar 2006).

Embryonic mammary development starts on embryonic day (E) 10.5, with the formation of the milk or mammary lines, ectodermal ridges running between the fore and hind limbs in both male and females embryos (for excellent reviews on embryonic mammary development see Robinson 2007; Watson & Khaled 2008; Cowin & Wysolmerski 2010). Around E11.5, the mammary line resolves into five pairs of placodes, lens-shaped thickenings of the surface ectoderm, symmetrically positioned along the milk line in characteristic locations (three thoracic and two inguinal). Each pair of placodes develops symmetrically following a characteristic temporal sequence. Interestingly different placodes seem to be specified by different molecular determinants as shown by genetic studies on mice; for example *Lef1*-deficient animals do not form placodes 2 and 3 (van Genderen et al. 1994).

Wnt signalling is necessary for the specification of the mammary line and placode development (Chu et al. 2004). The mammary line can be visualized *in situ* by *Wnt10b* staining (Veltmaat et al. 2004). Moreover the use of transgenic mice expressing the *Wnt* reporter gene TOPGAL, confirmed *Wnt* activation in mesenchymal and epithelial components of the mammary line followed by an increased expression in the placodes until E15.5 (Chu et al. 2004). Activation of *Wnt* signalling through lithium chloride treatment on embryonic skin caused an increase in placode number and size while disruption of *Wnt* signalling through transgenic expression of the secreted *Wnt* inhibitor DKK1 inhibited placode formation (Chu et al. 2004). The placode-specific functions of *Wnt* signalling seem to be mediated by LEF1 as in *Lef1* knock-out embryos, *Wnt* signalling is not activated and placodes do not form or degenerate, resulting in lack of mammary epithelium (Boras-Granic et al. 2006). *Wnt* signalling is also necessary for hair follicle and tooth development, as shown by suppression of all placodes development after *in vivo* blockade of *Wnt* signalling through expression of *Wnt* inhibitor DKK (Andl et al. 2002). This result suggests that *Wnt* signalling is a key

lineage specifier during ectodermal appendage development where it acts to induce multipotent ectodermal cells to adopt a generic placodal cell fate.

Fgf signalling seems to act in parallel to *Wnt* signalling in specifying the mammary line as *Fgf10*-/- and *Fgfr2b*-/- mice fail to develop all but mammary bud #4 (Mailleux et al. 2002). While *Fgfr2b* is expressed in the mammary buds as early as E11.5, *Fgf10* is not detectable in the embryonic mammary epithelium but is first seen in the somites underlying the milk line (Mailleux et al. 2002). *Fgf10* is thought to exert its key role on milk line specification and positioning through paracrine signalling originating from somites and acting on a subset of ectodermal cells expressing its receptor *Fgfr2b* (Veltmaat et al. 2006). The role of *Fgf10* as the critical somitic factor required for mammary line specification is further supported by the phenotype of *Gli3* deleted mice (Hatsell & Cowin 2006). *Gli3* loss causes a reduction of *Fgf10* expression, abrogation of TOPGAL expression in the mammary line in the region of placode 3 and loss or misplacement of buds 3 and 5. In the current view, somitic *Gli3* represses *Hedgehog* signalling and regulates expression of *Fgf10*, which in turn signals to ectodermal *Fgfr2b* and then to *Wnt10b* to specify mammary line formation (Veltmaat et al. 2006).

Another key regulator of mammary placode specification and positioning is *Tbx3*, a member of the T-box family of transcription factors (Rowley et al. 2004). Interestingly embryos deleted for *Tbx3* do not develop mammary buds and fail to express *Wnt10b* and *Lef1* (Davenport et al. 2003). Interestingly in humans, TBX3 haploinsufficiency is associated with the ulnar-mammary syndrome, a dominant developmental disorder characterized by abnormal forelimb and apocrine gland development which lead sometimes to complete loss of mammary glands (Bamshad et al. 1997). Transgenic overexpression of *Ectodysplasin* (*Eda*), a member of the tumour necrosis factor family, in the ectoderm causes the development of supernumerary glands (Mustonen et al. 2003) while deletion of GATA3, a transcription factor involved in T-cell development, caused a variable loss of placodes and a failure to develop the nipple sheath (Asselin-Labat et al. 2007). Overall these studies have uncovered a web of molecular mechanisms underlying mammary line and placode specification.

By E13.5, morphologically distinct epithelial placodes are present along the mammary line in mouse embryos, and by E14.5 these buds have invaded into the underlying dermis. At this point signals generated from the mammary epithelium push the surrounding mesenchymal cells to differentiate into a mammary-specific mesenchyme (Robinson 2007). The mesenchymal cells become arranged in concentric layers around the epithelial bud and start to express receptors for oestrogen and androgens. The primary mammary mesenchyme plays a key role in mammary development; maintaining mammary epithelial cell identity, supporting ductal morphogenesis and in male mice causing the destruction of the mammary epithelial bud in the presence of testosterone (Robinson 2007). Parathyroid hormone-related protein (PTHrP), produced by the epithelial bud has been identified as a key inducer of mammary mesenchyme identity (Wysolmerski et al. 1998; Foley et al. 2001). PTHrP is expressed by the mammary epithelial bud as it starts to invaginate into the mesenchyme while its receptor, PTH1R, is expressed in the mesenchyme surrounding the placode. Deletion of PTHrP or PTH1R, leads to a failure of ductal growth into the mesenchyme resulting in arrest of mammary development at the late bud stage (Foley et al. 2001). In addition the epithelial cells lose their mammary-identity and revert to an epidermal fate as shown by expression of keratinocyte-specific keratins (Foley et al. 2001). Moreover, mammary placodes fail to regress in male embryos indicating that PTHrP production in the epithelium is required to induce androgen receptor expression within the mammary mesenchyme (Dunbar et al. 1999). PTHrP signalling is also necessary for the mammary mesenchyme to induce the overlying epidermis to form the nipple (Foley et al. 2001). These data underline the importance of epithelial/mesenchymal interactions during mammary morphogenesis and the pivotal role of PTHrP as a determinant of mammary mesenchyme.

The next phase of development, which starts around E16, is characterized by the beginning of ductal branching morphogenesis. In a tightly coordinated process of invasion, a solid cord of epithelial cells emerges from the mammary bud and extend down through the primary mammary mesenchyme into the final stromal compartment, the mammary fat pad (Cowin & Wysolmerski 2010). At this time two other morphological processes take place: ductal lumen formation and nipple morphogenesis. By E18.5 all the mammary placodes have developed into

small immature glands that bear between 10 and 15 small branches. At the conclusion of embryonic development, the mammary gland consists of a short primary duct ending in a small ductal tree that is embedded in one end of larger mammary fat pad. The branching morphogenesis of the embryonic mammary gland is hormone independent as shown by normal embryonic mammary development of mice that are deficient in either oestrogen, prolactin, growth hormone or progesterone receptor (Cowin & Wysolmerski 2010). Instead embryonic ductal morphogenesis requires soluble factors that are supplied by the mammary fat pad precursor, such as PTHrP (Hens et al. 2007).

1.1.4 Virgin mammary development

The rudimentary epithelial tree in the mammary fat pad remains quiescent until puberty, growing at a very slow rate to keep up with normal body growth (allosteric growth). Allosteric growth stops with the onset of puberty when ovarian steroid hormone production induces the complex process of pubertal ductal morphogenesis (Watson & Khaled 2008). At the tips of the ducts, highly proliferative structures, called terminal end buds (TEBs) develop and start to invade into the fat pad driving ductal elongation (Sternlicht 2006). TEBs bifurcate, forming secondary branches which will further invade into the fatty stroma until the entire fat pad is filled with a network of branched ducts. The fully grown mammary gland will then start cycling under the influence of the oestrous-induced hormones with waves of side-branching followed by apoptosis occurring within each oestrous cycle (Schedin et al. 2000). Thus the adult virgin mammary gland is an active remodelling tissue constantly involved in cycles of proliferation and apoptosis. In the pubertal mammary gland, the initial drive for ductal morphogenesis comes from three main circulating hormones: ovarian hormone oestrogen and progesterone and the pituitary growth hormone (GH) (Brisken & O'Malley 2010).

Oestrogen

Pubertal ductal outgrowth of the mammary gland requires oestrogen, a steroid hormone synthesized by the ovaries. Oestrogen binds to two distinct receptors, oestrogen receptor (ER) α and ERß, which become activated and function as transcription factors when bound to the steroid hormone (Heldring et al. 2007).

Ovariectomized mice, which cannot produce oestrogen, fail to develop a mammary ductal network and this phenotype is rescued upon implantation of slow-release oestrogen pellets into the mammary gland (Daniel et al. 1987). Moreover the ductal network in pubertal mice deficient in ERα fails to grow and invade the fat pad (Korach et al. 1996) while no effect is seen after ERβ deletion (Krege et al. 1998). Although ERα is expressed both in the mammary epithelium and stroma, transplantation experiments showed that only epithelial ER is required for pubertal outgrowth (Mallepell et al. 2006).

Progesterone

At puberty, the rise in gonadotrophin levels leads to progesterone secretion from the ovaries. Progesterone receptor (PR) is a member of the nuclear hormone receptor family of ligand-dependent transcription factors which regulates specific target genes through binding to *cis*-acting progesterone response elements (Li & O'Malley 2003). PR is expressed in both epithelial and stromal compartments in the mouse mammary gland (Haslam & Shyamala 1981). Several studies showed that epithelial PR is not required for primary ductal growth but is a key regulator of ductal branching and proliferation in the virgin mammary gland (Obr & Edwards 2012). The mammary glands of young virgin PR-deleted females have no obvious defect of ductal development (Lydon et al. 1995). However, since PR deleted females are unable to ovulate, the effects of PR loss on oestrous-induced branching could not be investigated. To overcome this defect, oestrous cycles where artificially simulated by treating WT and PR-deleted virgin females with estradiol and progesterone: interestingly the WT breast tissue responded with increased side-branching and lobuloalveolar development, whereas the mammary glands of PR deleted females remained essentially unchanged (Brisken et al. 1998). This suggested that PR is not required for initial ductal growth but it is essential for oestrous-dependent side-branching and alveologenesis. The essential role of progesterone during the expansion of the alveolar compartment during pregnancy is discussed below.

Growth Hormone

Although pubertal mammary development is started by the surge of oestrogen levels, the action of oestrogen depends on the presence of growth hormone

(GH). Experiments in the 1940s demonstrated that the pituitary gland was required for mammary gland development and that oestrogen-driven ductal development fails in hypophysectomised rodents (Reece et al. 1936; Kleinberg & Ruan 2008). Moreover transgenic mice lacking GH receptor displayed severely impaired mammary ductal outgrowth at puberty (Gallego et al. 2001). Interestingly GH signalling is required in the stroma since ductal epithelium arising from GH receptor null glands transplanted into the cleared fat pads of WT mice develop as normal (Gallego et al. 2001).

Other factors required in pubertal growth

Several other growth factors are required for pubertal ductal morphogenesis (reviewed in McNally & Martin 2011). During pubertal mammary development, the action of oestrogen depends on the presence of GH to drive ductal outgrowth. Further experiments showed that the IGF-1 signalling pathway is necessary downstream of GH to induce pubertal mammary development (Wood et al. 2000). IGF-1 deleted mice have a defective pubertal mammary development characterized by impaired ductal outgrowth (Ruan & Kleinberg 1999). Moreover IGF1-R null embryonic mammary buds transplanted into WT hosts possess limited outgrowth potential and show defects in TEB proliferation (Bonnette & Hadsell 2001). Rescue experiments also showed that IGF-1 is downstream of GH and oestrogen signalling where IGF-1 treatment rescued ductal development in oestrogen deficient (ovariectomized) and GH-deficient (hypophysectomized) animals (Ruan et al. 1992), while GH or oestrogen treatment did not rescue IGF-1-null glands (Ruan & Kleinberg 1999).

Fibroblast growth factor (FGF) signalling has been implicated in the control of postnatal mammary development (Hynes & Watson 2010). During ductal outgrowth multiple FGFs as well as FGFR1 and FGFR2 are expressed. TEBs have high levels of FGFR2, which is essential at this stage because FGFR2 null glands have a severe delay in adolescent ductal development, penetrating the fat pad more slowly and showing fewer branch points compared with WT controls (Lu, 2008). Moreover an analysis of genetic mosaicism reveals that epithelia without FGFR2 are eliminated from the ducts that do develop (Lu et al. 2008). Other FGF-deficient models (e.g., FGF-7) showed no mammary phenotype indicating

the presence of possible compensatory mechanisms between FGF family members (Sternlicht 2006).

Epidermal growth factor (EGF) ligands are produced as transmembrane precursors that after cleavage bind and activate EGF-R receptors. Several EGF-family ligands and all of the EGF-R receptors are expressed in the mammary gland during mammary development (Schroeder & Lee 1998). Interestingly, slow-release pellets delivering EGF restored ductal development in ovariectomized (Coleman et al. 1988) and ERα-deficient mice (Kenney et al. 2003) while exogenous oestrogen drives EGF-R activation in ovariectomized mice (Sebastian et al. 1998). These results suggest that EGF-R promotes mammary ductal outgrowth downstream of ERα. Moreover reciprocal transplant experiments between WT epithelium and EGF-R mutant stroma, revealed that only stromal EGFR is required for ductal morphogenesis (Wiesen et al. 1999). Amphiregulin (AREG) has been found to be the key EGF-family ligand to bind EGF-R and to mediate pubertal morphogenesis downstream of ER signalling. Ductal morphogenesis is not supported in *Areg* deleted mice (Luetteke et al. 1999) while ductal outgrowth can be induced in ovariectomized mice by treatment with recombinant AREG (Kenney et al. 1996).

Together with hormones and growth factors, immune cells are also important for pubertal mammary gland development (Coussens & Pollard 2011). Work from Pollard's laboratory showed that eosinophils and macrophages are recruited around the growing terminal end buds during postnatal development (Gouon-Evans et al. 2000). Furthermore, *in vivo* ablation of eosinophils and macrophages resulted in impaired TEBs development and reduced mammary branching demonstrating the requirement of immune cells during pubertal development (Gouon-Evans et al. 2000).

1.1.5 Pregnancy and lactation

Unlike the majority of other organs which mainly develop to a mature state during the embryonic phase, the mammary gland only reaches a mature fully functional state during the pregnancy-lactation cycle in the adult female. Undoubtedly lactation is the most important developmental stage of the breast as shown by the very high metabolic demand, requiring about 25% of daily maternal energy intake during lactation to produce milk (Hassiotou & Geddes 2013). Although a mini-remodelling of the breast occurs at each menstrual cycle it is only during the processes of pregnancy and lactation that the ductular appearance of the virgin epithelium is replaced through a massive event of cell proliferation and differentiation which lead to the formation of mature alveoli, the milk-producing unit of the mammary gland. Alveolar morphogenesis is initiated by coitus which provides nervous stimulation inducing prolactin secretion from the pituitary gland which in turn, sustains ovarian progesterone secretion (Exton et al. 2001). These hormones induce rapid proliferation of mammary epithelial cells followed by alveoli differentiation to form the secretory alveolar epithelium, capable of milk production and secretion during lactation.

Prolactin

Prolactin (PRL), produced by the lactotrophic cells of the anterior pituitary, is the major driver of development during pregnancy (Oakes et al. 2008). Grafting experiments with prolactin receptor (PRL-R) deleted epithelium showed no defects in ductal outgrowth and side branching (Brisken et al. 1999). However *Prl* signalling pathway is required for alveologenesis and differentiation of MECs into milk producing cells as showed by the total absence of alveolar differentiation and milk production in transplanted *Prl*-null epithelium (Ormandy et al. 1997). Confirming an epithelial role for prolactin signalling, WT epithelium grafted into *Prl-R* deleted stroma developed normally (Ormandy et al. 2003). *Jak2* and *Stat5* knockout mice showed failed lobuloalveolar development like *Prl-R* deleted mammary glands, indicating their critical role in Prl-R-mediated alveolar morphogenesis (Liu et al. 1997; Wagner et al. 2004). PRL binds the prolactin receptor, leading to receptor dimerization, phosphorylation and activation of the JAK2 kinase which in turn recruit and phosphorylates STAT5.

Phosphorylated STAT5 then dimerizes and translocates to the nucleus where it can activate the transcription of genes involved in alveolar morphogenesis (Wakao et al. 1994). The transcription factor ELF5 acts downstream of the PRLR to modulate alveolar morphogenesis and milk production, as shown by the impaired alveolar morphogenesis of transgenic mice heterozygous for *Elf5* (Zhou et al. 2005). Confirming their key role in alveologenesis, STAT5 and ELF5 have also been shown to regulate the specification of luminal progenitors, the cell population which will generate the alveolar lineage during pregnancy (Yamaji et al. 2009; Chakrabarti et al. 2012).

Progesterone

The ovarian hormone progesterone is required together with prolactin for the process of alveologenesis. Since PR knock-out mice are infertile (Lydon et al. 1995), transplantation of mammary epithelial cells derived from PR deleted mice into the cleared mammary fat pad of wild type female recipients has been used to study this. Transplanted mouse mammary epithelial cells (MMECs) from PR deleted mice into WT hosts, failed to undergo alveolar morphogenesis during pregnancy proving the essential role of PR signalling in the epithelial compartment during alveologenesis (Brisken et al. 1998). Progesterone is thought to act on PR positive epithelial cells stimulating the production of a paracrine signal which, in turn, induces the proliferation of neighbouring cells (Ismail et al. 2002). One candidate molecule likely responsible for this paracrine signalling is receptor activator of nuclear factor κB (NF-κB)-ligand (RANKL). In fact RANKL mRNA is induced by treatment with oestrogen and/or progesterone, in the mammary gland of ovariectomized adult mice, but not in PR-null mice (Mulac-Jericevic et al. 2003). Moreover ectopic expression of RANKL using a MMTV transgene results in side branching stimulation in the absence of pregnancy (Fernandez-Valdivia et al. 2009). These data suggested that RANKL is a paracrine mediator of progesterone-induced proliferation. Another molecule which has been suggested as a possible paracrine effector of progesterone signalling is WNT4. MMECs extracted from *Wnt4*-deleted mice showed lack of branching in transplanted outgrowth during pregnancy (Brisken et al. 2000). Further experiments showed that WNT4 expression is PR dependent; WNT4 is induced in the mammary gland of adult ovariectomized mice by progesterone

treatment while PR deleted cells showed loss of WNT4 expression (Mulac-Jericevic et al. 2003; Fernandez-Valdivia et al. 2008).

1.1.6 Involution

The end of lactation sees the beginning of involution, a strictly regulated process involving apoptosis and tissue remodelling, which will bring the mammary gland back to a virgin-like state (Watson 2006; Stein et al. 2007; Watson & Kreuzaler 2011). In the last decades conditional knock-out animal models have been used extensively to unveil the molecular regulators of the involution process. Since natural weaning is often a long and gradual process not easily controllable, investigators have used forced weaning as a model for the study of involution. In this protocol pups are removed from the lactating dam after 7-10 days of lactation initiating the involution process which gives a controllable model in which to study the morphological and molecular features regulating this process. Teat sealing is another technique that has been used for the study of involution (Watson & Kreuzaler 2011). In particular this protocol allows the study of local regulators of involution (such as effects of milk accumulation) since the lactating hormonal milieu is maintained by the continued presence of suckling pups (Watson 2006).

Involution in the mouse mammary gland can be divided into two discrete phases called first phase and second phase (Watson, 2006; Watson & Kreuzaler, 2011). The first phase lasts approximately 48 h and can be reversed through re-suckling. The second phase instead is characterized by an irreversible remodelling programme which will bring the gland to a pre-pregnant state (Lund et al., 1996). Morphologically the first phase of involution is marked by accumulation of milk in the alveoli and the appearance of shed, dying cells within the alveolar lumen (Richert et al. 2000). The first phase of involution depends on local factors and probably it is triggered by the accumulation of milk into the alveoli (Li et al. 1997). Milk retention could lead to accumulation of secreted factors in the milk or to mechanical stretch of the alveolar epithelium which could signal the start of the involution process (Quaglino et al. 2009). Likely these processes are not mutually exclusive and in fact some evidence indicates that they could be mutually linked. For example, the plasma membrane calcium-ATPase 2 (PMCA2), which transports 60–70% of milk calcium

outside the alveolar cells, is regulated by changes in the shape of mammary epithelial cells (VanHouten et al. 2010) and it is dramatically downregulated during involution (Reinhardt & Lippolis 2009). Furthermore, mice deleted for PMCA2 exhibited precocious cell death at day 18 of pregnancy (VanHouten et al. 2010). In this scenario milk stasis could induce mechanical stress in the alveoli leading to a sharp drop in PMCA2 expression. This in turn could cause accumulation of calcium into the alveolar cells and initiation of calcium mediated cell death. Together with PMCA2 other factors have been found to play a role in the induction of involution. Not surprisingly involution is affected by interference with well-known regulators of the apoptotic program: for example deletion of the anti-apoptotic *Bcl2l1* gene accelerates involution while a delay is seen after loss of the pro-apoptotic BAX protein (Schorr et al. 1999; Walton et al. 2001). Moreover expression of constitutively active AKT in the mammary gland under the control of the MMTV promoter caused a delay in involution and a decrease in apoptosis (Schwertfeger et al. 2001). Work on genetically modified mice has found others factors which are involved in involution such as LIF, serotonin and TGFß-3 (Nguyen & Pollard 2000; Matsuda et al. 2004; Kritikou et al. 2003). Most of these studies exhibit a subtle involution phenotype, indicating that complex compensatory mechanisms are taking place in this phase of development. However two pathways seem to be crucial for involution: the JAK/STAT and the NF-kB pathways. Conditional deletion of STAT3 in the mouse mammary epithelium causes a block in cell death and tissue remodelling (Chapman et al. 1999), extending the reversible phase of involution until at least 6 days after forced involution (Humphreys et al. 2002). Similarly, conditional deletion of IKKß/2, the upstream regulator of the NF-kB pathway, caused a delayed involution and a reduction of cell death (Baxter et al. 2006). Since both pathways are necessary but not sufficient to trigger involution, these results indicate that cooperation between JAK/STAT and NF-kB pathways is required for a correct involution. The mechanism responsible for cell death during the first phase of involution has always presumed to be apoptosis. However the morphology of the dying cells that shed into the lumen is not classically apoptotic: shed cells are swollen, they lack any membrane blebbing and they have two hypercondensed nuclei (Watson & Kreuzaler 2011). Recently, a study from the Watson laboratory clarified this issue. A caspase-independent lysosomal pathway of cell death was found to be necessary for post-lactational

regression of the mammary gland (Kreuzaler et al. 2011). Interestingly the lysosomal pathway of cell death is independent of caspases 3, 6 and 7, but it requires STAT3 which upregulates the expression of intra-lysosomal specific proteases (cathepsin B and cathepsin L). Then leaky lysosomes leads to release of cathepsins into the cytoplasm and the initiation of cell death (Chwieralski et al. 2006). As a result cells are shed into the lumen where caspases are cleaved, probably in response to anoikis.

The overall structure of the mammary gland is not affected during the first phase of involution (Richert et al. 2000). However after 48 hours, the reversible phase ends and an irreversible process of widespread cell death and tissue remodelling signals the transition to the second phase. In this phase the alveoli start to collapse while adipocytes re-differentiate and begin to fill the fat pad (Richert et al. 2000). At the molecular level this phase is characterized by an increased protease activity which is responsible for the extensive tissue remodelling observed. Stromal cells increase their expression of MMP-2, MMP-3, MMP-9 and MMP-11 which will act on the tissue extracellular matrix causing alveolar collapse, cell detachment and cell death for anoikis (Pullan et al. 1996). Since the removal of extracellular matrix by MMPs is the main trigger of cell death in the second phase of mammary involution, fail-safe mechanism needs to be in place during the first phase to avoid early activation of MMPs. Tissue inhibitors of metalloproteinases (TIMPs) are direct inhibitors of zinc-dependent proteinases such as MMPs and key regulators of involution (Fata et al. 2001). *In vivo* loss of function of TIMP3, an inhibitor of MMP2, causes accelerated involution with loss of the reversible first phase (Fata et al. 2001). The wave of apoptosis generated by MMPs action creates a vast number of tissue and cell debris which needs to be removed in order to create space for the regenerating adipocytes; this process involves autophagy and phagocytosis carried out by activated mammary cells and recruited macrophages (Monks et al. 2005) and inflammatory cells. In fact different types of immune cells are recruited in the involuting mammary gland, starting with neutrophils at day 1 involution followed at day 4 by plasma cells, eosinophils and B-lymphocytes (Stein et al. 2004). The involvement of inflammation is also supported by microarray data that identified an immune cascade response during involution (Stein et al. 2004; Clarkson et al. 2004).

1.2 Mammary stem cells

The normal adult mammary gland is composed of a bi-layered structure consisting of an inner layer of luminal cells, and an outer layer of basal cells on a laminin-containing basement membrane. The luminal compartment consists of different subpopulations of hormone positive and negative progenitors which during pregnancy can differentiate into alveolar cells, the milk producing unit of the mammary gland (Richert et al. 2000). Besides its localization, the luminal population can be identified by the expression of cytokeratins (CK8, 18 and 19), MUC1 and EpCAM (Richert et al. 2000). The basal/myoepithelial population is responsible for the secretion of basement membrane components during all developmental stages and for the movement of milk through the alveoli and the ducts during lactation. These cells are characterized by the expression of CK14, CK5, smooth muscle actin and P63 (Moumen et al. 2011). However, under this apparently simple and dualistic structure of the mammary epithelium derived by histological studies, lies a complex hierarchical cellular system which has emerged from studies focused on mammary stem cells (Visvader & Smith 2010). Stem cells are emerging as fundamental players in both embryonic development and adult tissue homeostasis, especially in those organs characterized by high regenerative potential and cell turnover (Biteau, 2011). Stem cells are multipotent cells characterized by two unique features: self-renewability and wide differentiation capability. In adult organs the majority of the stem cell population is in a state of quiescence, which is necessary for stem cell maintenance preserving them from oxidative stress, accumulation of mutations and telomerase shortening (Blanpain et al. 2012). However, under appropriate stimuli, stem cells are able to enter a proliferation phase characterized by self-replication and lineage differentiation, (Biteau et al. 2011). The mammary gland is a very dynamic tissue which undergoes extensive cellular remodelling and changes in morphology throughout different developmental phases (Hennighausen & Robinson 2005). In such a dynamic context, different mammary epithelial cell populations with different specialized functions are continuously generated and replaced. This high tissue turnover implies the existence of a population of mammary stem/progenitor cells characterized by a high degree of self-renewability which are able to differentiate into the various cell types constituting the mammary epithelium. The existence of a mammary

stem cell (MaSC) population has been postulated since DeOme's transplantation studies in the late 50's (DeOme, 1959) but only in the last 10 years has solid evidence of their existence been found. These seminal experiments and the current status of the field are discussed.

1.2.1 Birth of the mammary stem cell concept

DeOme in the 1950s (DeOme, 1959) pioneered a technique which has become the standard assay in mammary stem cell biology: the cleared fat-pad transplantation. This procedure allows the identification of mammary stem cells by transplanting cells in limited dilution in a virgin fat pad cleared of the endogenous epithelium. It is based on the assumption that only cells characterized by high degree of self-renewal and regenerative potential will be able to proliferate and give rise to a fully functional new mammary epithelium. At 3 weeks of age, the mouse mammary fat pad is largely devoid of mammary epithelium: the rudimental epithelial tree consists of only a few branches confined to the vicinity of the nipple. Therefore, if the region between the nipple and the lymph node is surgically removed, no endogenous epithelium will be left to repopulate the mammary fat pad which is then described as "cleared". At this stage mammary cells can be injected in the cleared tissue where they can grow and repopulate the entire fat pad. DeOme and colleagues published the first successful mammary gland reconstitution experiment showing that small pieces of mammary epithelium could recreate a fully functional reconstituted gland (DeOme et al., 1959). Follow up experiments showed that the regenerative potential of normal mammary cells is limited. Daniel and colleagues performed up to seven serial transplantations over the course of two years showing that the ability of normal mammary epithelium to reconstitute a gland declines with time (Daniel, 1968). In the same paper, 5 independent serial transplants showed considerable variation in overall lifespan between individual outgrowths suggesting that mammary epithelial cells could differ in their regenerative potential (Daniel 1968). Interestingly the regenerating capability of mammary epithelial cells is independent of donor's age indicating the existence of precursor cells throughout the entire life span of the mammary gland (Hoshino & Gardner, 1967; Young, Medina, DeOme, & Daniel, 1971). Moreover fragments taken from different regions of the mammary gland, such as primary and tertiary branches, alveoli and terminal end buds were able to reconstitute fully

functional outgrowths, indicating the presence of repopulating cells throughout the ductal epithelial tree (Smith & Medina 1988). The development of tissue dissociation techniques allowed the transplant of dissociated mammary epithelial cells rather than tissue fragments (DeOme et al. 1978) and was followed by the introduction of the limiting dilution transplantation, the "gold standard" for defining stem cell activity. Using the limiting dilution transplantation technique Smith and co-workers, showed the existence of distinct mammary progenitor populations characterized by specific limited differentiation potential during pregnancy (Smith 1996). This experiment was the first suggestion of the existence of a cell hierarchy in the mammary epithelium with various degree of regenerative potential. Limiting dilution experiments suggested that a limited number of mammary cells were able to generate a duct tree. The first evidence that this regenerative activity could be attributed to a single cell came from experiments where fragments of MMTV-infected mammary tissue were transplanted into cleared mammary fat pads of uninfected mice. The ability of the MMTV virus to integrate randomly in the genome was used to investigate clonal expansion during the regenerative process. In particular the MMTV integration pattern was analysed through Southern blot in serially transplanted outgrowths, indicating that the majority of epithelial tissue in the reconstituted glands was likely to be derived from a single mammary cell (Kordon & Smith 1998). Further studies based on X-chromosome inactivation patterns in human samples showed that entire lobules and large ducts of normal breast tissue have the same X chromosome inactivated suggesting that clonal expansion was also responsible for human mammary gland development (Tsai et al. 1996). Taken together these early experiments provided a solid proof for the existence of a population characterized by high regenerative capability. These studies represent the foundations on which future mammary stem cell research was built and which resulted in the exciting discoveries of the last decades.

1.2.2 The troubled pathway towards isolation of MaSCs

Various experimental approaches have been carried out to characterize and isolate MaSCs (Smith et al. 2012). A first approach tried to identify cells with stem cell characteristics through cytological examination (Smith & Medina 1988). Smith and Medina identified a population of pale coloured cells with large nuclei

and clear cytoplasm. These cells were distributed sporadically among the mammary duct tree and generated cells capable of differentiating in the presence of lactogenic stimuli. Another morphological study identified potential mammary stem cells using electron microscopy techniques: a mammary epithelial subpopulation constituted by small light cells was proposed to be the putative stem cell compartment based on size, mitotic activity, absence of organelles and ability to give rise to darker cells, which were thought to be the differentiated population of cells (Chepko & Smith 1997). Furthermore, several independent studies identified long-term label retaining subpopulations in the mammary epithelium thought to represent stem cells (Zeps et al. 1996; Smith 2005). However none of these methods showed that a single label-retaining cell or pale cell could reconstitute a fully functional gland *in vivo*. In fact both the morphological and the label retention methods shared a common limitation: they do not allow the isolation of large numbers of relatively pure MaSC populations which can be further tested in *in vitro* or *in vivo* assays. To address this issue, lessons learned from haematopoietic stem cells were applied to the mammary stem cell field. In particular the use of fluorescence-activated cell sorting (FACS) allowed the isolation of different cell populations from dissociated mammary gland preparations, based on the expression of surface marker proteins. The subpopulations obtained were then assayed for repopulating ability *in vivo* by cleared fat pad transplantation of limiting numbers of cells. In 2006, a study from Visvader's laboratory (Shackleton et al. 2006) identified a subpopulation of cells that are enriched for mammary stem cells on the basis of high expression of CD29 (β1-integrin) and moderate levels of CD24 (heat stable antigen). When CD24+ $CD29^{hi}$ cells were transplanted into cleared fat pads of virgin mice, they were able to generate a functional gland with high efficiency. Indeed transplantation of single CD24+ $CD29^{hi}$ cells was sufficient to regenerate an entire mammary gland (6/102 transplants) which could undergo full alveolar differentiation during pregnancy. Interestingly the transplanted epithelial tree contained daughter stem cells with the same *in vivo* repopulating activity as the original CD24+ $CD29^{hi}$ population (Shackleton et al. 2006). At the same time, Stingl and co-workers, using a different set of markers, high expression of CD49f (α6-integrin) and moderate levels of CD24, identified a subpopulation of the mammary gland enriched for stem cells (Stingl et al. 2006). Again, the purified CD24+ $CD49f^{hi}$ population was able to generate the complete

functional mammary gland from a single cell (Stingl et al. 2006). Meanwhile CD24, used as a single marker, allowed the enrichment for fat pad repopulation capacity in the $CD24^{Low}$ population (Sleeman et al. 2006). However, based on CD24, CD29 or CD49f markers, the maximum enrichment for MaSCs achievable is far from purity (less than 1 MaSC in 100 cells). Placing GFP under the control of the s-SHIP promoter, which is expressed in embryonic and hematopoietic stem cells but not differentiated cells, allowed the isolation of a MaSC-enriched $GFP+CD49f^{hi}$ population (Bai & Rohrschneider 2010). These cells are characterized by an increased frequency of MaSC (1/48 cells). Recently another study reported an improved protocol for MaSC purification with the combination of an additional marker, CD1d, to further enrich the $CD24+CD29^{hi}$ population. The $CD24+CD29^{hi}CD1d+$ population showed an enrichment frequency to nearly single-cell level of 1 MaSC per 8 cells (dos Santos et al. 2013).

Besides being used to characterize the mammary stem cell population, FACS profiling of the mammary epithelium has also unveiled the different lineages which make up the mature mammary gland. The luminal cell compartment is very heterogeneous, being composed of a mixture of progenitor and differentiated cells which fluctuate in response to hormones (Schedin, 2000). FACS sorting has been used to isolate luminal progenitors based on the expression of CD61 (Asselin-Labat et al. 2007), c-KIT (Regan et al. 2012) and CD14 (Asselin-Labat et al. 2011). Luminal progenitor populations are characterized by high *in vitro* colony forming and proliferative potential, but limited ability to repopulate a cleared fat pad. Moreover Asselin-Labat and co-workers showed a dramatic decline in the proportion of CD61+ progenitor cells during pregnancy (Asselin-Labat et al. 2007). Overall these studies have shown that luminal progenitors are a highly dynamic population of the virgin mammary epithelium which is lost during pregnancy, through their differentiation into alveolar cells (Asselin-Labat et al. 2007). Like the luminal compartment, the basal lineage is also likely to be constituted by different subpopulations with various degrees of regenerative potential and differentiation; however the characterization of this compartment is still lagging behind. Only recently two different studies identified a subpopulation of LGR5+ cells with increased regenerative potential when compared to other basal cells and which were

necessary for postnatal mammary organogenesis (Plaks et al. 2013; Van Keymeulen et al. 2011).

Cell surface analyses have led to the concept of an existing differentiation hierarchy among mammary epithelial cells where a mammary stem cell population, which resides in the basal layer of the mammary gland, gives rise to progressively restricted progenitors which will finally differentiate into the mature luminal and basal mammary lineage (Visvader & Smith 2010). However this model has been challenged by lineage tracing experiments. Recently an elegant study from Blanpain's laboratory showed the existence of two lineage restricted mammary stem cell populations in the adult mammary gland: a basal-restricted and a luminal-restricted unipotent stem cell (Van Keymeulen et al. 2011). Furthermore, the only mammary multipotent stem cell capable of generating both basal and luminal population was detected uniquely during the embryonic mammary development (Van Keymeulen et al. 2011). Another lineage tracing study which followed the *Wnt* responsive population during mammary development reached the same conclusion showing that the basal cell layer in the mature virgin gland does not contribute to the luminal cell layer (van Amerongen et al. 2012). These evidences support the view that both luminal and basal compartments, under normal physiological conditions, are maintained by separate stem/progenitor compartments. Moreover these studies demonstrate a critical limitation of fat pad transplantation experiments: the regenerative potential showed by transplanted cells is not reflecting their real physiological behaviour in the intact tissue. This raises the question of whether previous studies using FACS sorting and fat pad transplantation to demonstrate the existence of multipotent mammary stem cells are indeed mere artefact. However it should be realised that lineage tracing has its own limitations. For example the reporter may not be expressed in stem cells (epigenetic silencing) or could alter stem cell function *in vivo* (Baens et al. 2006). Moreover MaSCs may not be targeted by the promoters used to drive the reporter expression. Along these lines it is interesting that none of the previous studies have reported labelled cap cells in terminal end buds.

1.2.3 MaSCs and pregnancy

MaSCs have also been studied during pregnancy. The first evidence for the involvement of a stem cell population during pregnancy came from experiments using the WAP-Cre model in combination with the Rosa26LacZ reporter mice. In this model a LacZ-marked lobular-limited progenitor, which was able to persist after involution and take part in the remodelling of the involuting mammary epithelium, was identified in the pregnant mouse mammary gland (Wagner et al. 2002). These LacZ-positive cells, called parity-identified mammary cells (PI-MEC), showed self-renewal and multipotent potential, producing outgrowth with both luminal and basal cells, following serial transplantation in epithelium-free mammary fat pads (Boulanger et al. 2005). Progesterone is one of the key hormonal regulators of mammary epithelium alveolar development and interestingly MaSC pool increases 14-fold during maximal progesterone levels at the dioestrus phase of the mouse oestrous cycle (Joshi et al. 2010). Moreover the stem-cell-enriched CD49f^{hi} population expands after treatment with exogenous progesterone, demonstrating a key role for progesterone in driving MaSCs expansion (Joshi et al. 2010). Confirming these data, work from the Visvader/Lindeman laboratory showed that the MaSC pool increases under the hormonal milieu of pregnancy, with a significant 10- fold increase in mid-pregnancy (Asselin-Labat et al. 2010). By the end of pregnancy, the number of stem cells return to levels observed in virgin mice, suggesting that the MaSC expansion is necessary to sustain the high proliferative first phase of pregnancy but it is not required during the terminal differentiation phase (Asselin-Labat et al. 2010). These results indicate that MaSCs undergo strict hormonal regulation during oestrous cycle and pregnancy. However, using FACS sorting techniques it was shown that MaSCs do not express oestrogen receptor or progesterone receptor (Asselin-Labat et al. 2006). Hence MaSCs are likely under control of paracrine signalling molecules which are secreted from other mammary compartments in response to hormones. A study from Joshi and co-workers suggested a mechanism for this paracrine signalling where pregnancy-induced progesterone activates a subset of luminal cells, which in turn secrete WNT4 and RANKL signals. These molecules act on the basal and MaSC population which respond by upregulating transcriptional targets and cell cycle markers (Joshi et al. 2010). Intriguingly, a recent study provides evidence confirming the link between pregnancy, *Wnt* signalling and MaSC expansion (van Amerongen et al.

2012). Van Amerongen and co-workers showed the existence of multipotent stem cells residing in the basal layer which become *Wnt*-activated during pregnancy and contribute to both basal and luminal lineages. Taken together these data are suggesting the existence of MaSC with multipotent potential which, in normal physiological conditions are restricted to unipotency but can switch from unipotent to multipotent potential under particular circumstances such as pregnancy.

1.2.4 *In vitro* assays for the study of MaSCs

In vivo mammary transplantation and lineage tracing experiments, besides the limitations listed before, are relatively costly and time consuming. These restrictions have driven the development of *in vitro* 2D or 3D assays as simplified models for investigating mammary stem cell properties (Smith et al. 2012). One of the first assays to be developed was based on plating mammary cells in 2D at low density, using media containing epidermal growth factor and adding a layer of irradiated fibroblasts to achieve maximal cloning efficiencies (Stingl et al. 2006). In this culture model, robust growth of colonies with luminal, basal or bipotent characteristics, based on cytokeratin expression, can be achieved (Stingl et al. 2006; Shackleton et al. 2006). In 2D assays colony-forming cells are at least 100-fold more numerous than cells detected in fat pad transplantation showing that the more permissive conditions of this assay probably allow the expansion of more differentiated progenitor cells (Makarem et al. 2013). Mammosphere culture is another *in vitro* culture technique which allows the enrichment for MaSC. It is based on the use of non-adherent conditions which induce anoikis on differentiated cells and serum free media supplemented with EGF and bFGF (Dontu et al. 2003). This assay was first developed in the neural stem cell research field, allowing the growth of a multipotent population of neural cells in suspension which were called neurospheres. Neurospheres were shown to consist of 4%–20% stem cells, the rest of the population representing progenitor cells in various stages of differentiation (Reynolds & Weiss 1996). Applying the same technique to the mammary field, Dontu and colleagues showed that human mammary cells cultivated in low adherent and serum free conditions formed floating colonies which were called mammospheres (Dontu et al. 2003). Interestingly whereas freshly isolated mammary cells contain only 8% multilineage progenitors, mammosphere cultures are enriched to ~68% in

primary mammospheres, and virtually 100% in secondary and later-passage mammospheres. A key result confirming the enrichment for cells with high regenerative capability achievable with this assay was obtained by the Pelicci's laboratory demonstrating that cleared fat pad injection of a single sphere was able to generate mammary outgrowth in 11 out of 18 transplantations (Cicalese et al. 2009). However several limitations are linked to this assay. For example the floating culturing conditions and the absence of any matrix means mammary stem cells are grown in extremely artificial conditions which could influence their behaviour. Moreover given the very strong tendency of mammary epithelial cells to adhere to one another, it can be difficult to distinguish between aggregation versus clonal growth (Pastrana et al. 2011). To overcome the clonogenicity issue mammary cells have been immobilized in a semi-solid matrix such as Matrigel were they can grow and develop ducts-like structures. These colony-forming assays in Matrigel allow assessment of the clonal growth of mammary cells in a 3D environment which more closely resembles the physiological conditions (Zeng & Nusse 2010; Guo et al. 2012).

1.2.5 MaSCs regulators: signalling pathways and transcription factors.

Over the past few years the role of various signalling pathways has been evaluated in the context of mammary stem and progenitor cells. Several studies indicated *Wnt* signalling to be involved in MaSC regulation. Overexpression of *Wnt1* using the mammary-specific MMTV promoter resulted in a 6-fold increase in the number of MaSCs (Shackleton et al. 2006). Moreover treatment with recombinant WNT protein leads to an expansion of mammary stem cells *in vitro* as shown by their ability to reconstitute functional glands in transplantation assays (Zeng & Nusse 2010). In addition LGR5+ cells have recently been identified as a basal subpopulation characterized by high regenerative potential, underlying the key regulatory role of *Wnt* pathway in MaSCs (Plaks et al. 2013). On the other hand *Notch* pathway normally plays a role in restricting expansion of MaSCs. In fact deletion of *Cbf-1*, a canonical effector in the *Notch* pathway, caused higher stem cell repopulating activity *in vivo*, accompanied by aberrant terminal end buds and ductal branching development (Bouras et al. 2008). In addition the *Notch* pathway is required for luminal cell fate determination by promoting commitment of MaSCs to the luminal cell lineage at the expense of

the basal lineage (Bouras et al. 2008). Hedgehog signalling also appears to play a negative regulatory role in MaSCs. Expression of constitutive activated human SMO (*SmoM2*) under the MMTV promoter in transgenic mice caused a decrease in the frequency of regenerative stem cells in MMTV-*SmoM2* epithelium relative to wild type (Moraes et al. 2007).

The concept of a mammary epithelial hierarchy implies that specific transcription factors are needed to determine the correct specification of the different mammary subpopulations. Indeed several transcription factors have been found to play a role in mammary stem cells and lineage differentiation (Siegel & Muller 2010). For example the transcription factor p53 regulates the self-renewal of MaSCs, with loss of p53 leading to increased symmetric cell division and expansion of the stem cell pool (Cicalese et al. 2009). *Slug* and *Sox9* have also been found to be key determinants of the mammary stem cell fate with transient co-expression sufficient to convert differentiated luminal cells into MaSCs with long-term mammary gland-reconstituting ability (Guo et al. 2012). The list of transcription factors found to be involved in the regulation of mammary stem cells is constantly expanding and now includes MYC, C/EBPβ, and STAT3 among others (Staniszewska et al. 2012; Moumen et al. 2012; LaMarca et al. 2010). Nevertheless, given the emerging complexity of the mammary epithelial hierarchy, it is tempting to predict the existence of several other transcription factors involved in the mammary stem cell differentiation process. The identification of new transcription factors with a role in mammary lineage specification could translate to important advances in breast cancer research, allowing the identification of new potent regulators of mammary epithelium cell-fate which could be hijacked during the neoplastic process. In fact transcription factors, due to their biological role as regulators of wide arrays of target genes and as convergence nodes of different pathways, represent very promising therapeutic targets (Koehler 2010).

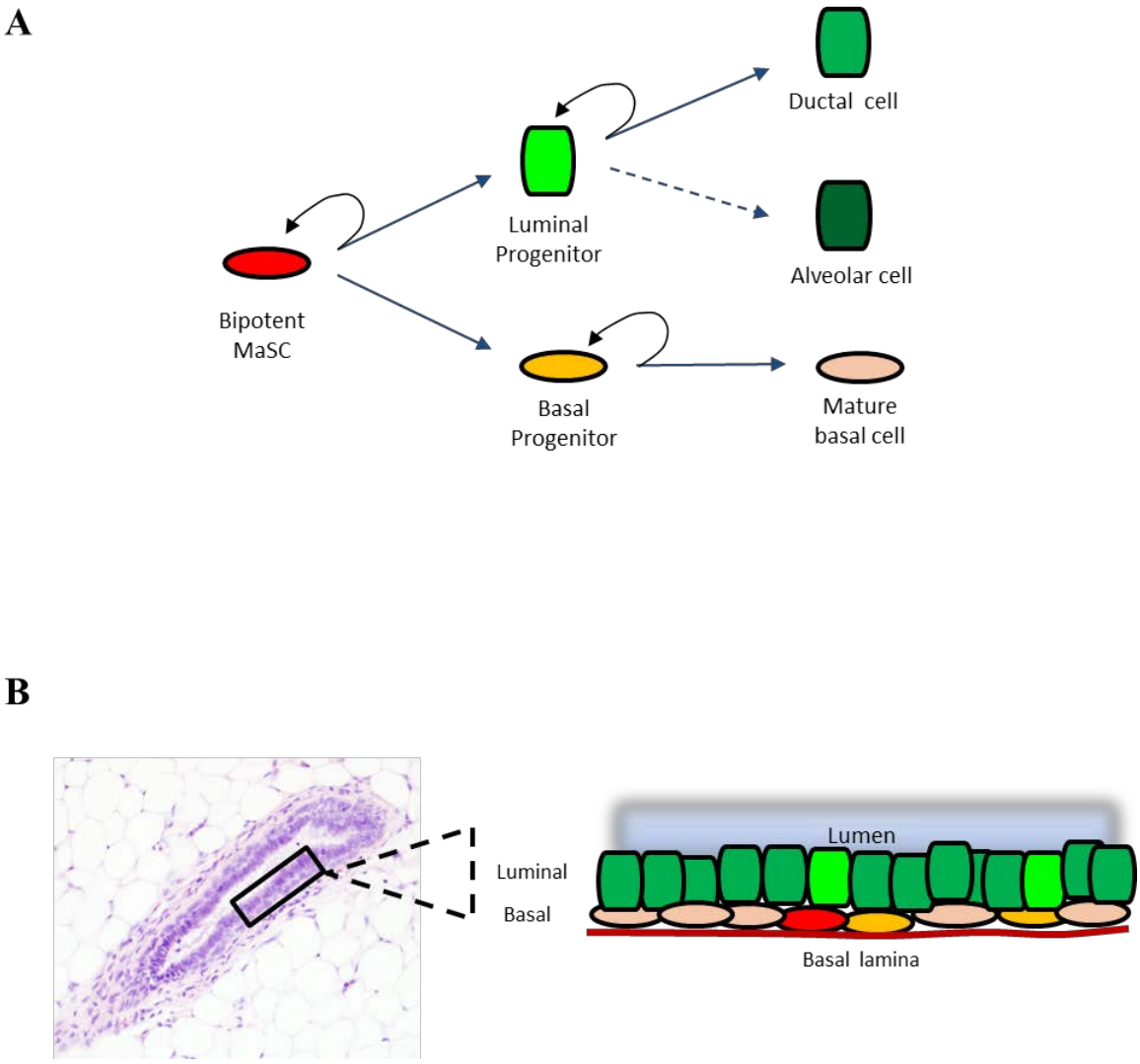


Figure 1-1 The mammary cell hierarchy

Diagram showing an overview of the mammary cell hierarchy. A multipotent stem cell can give rise to two lineage-restricted progenitors (luminal and basal). The luminal progenitor can further differentiate into mature ductal cells and, during pregnancy, into alveolar cells. The basal progenitor can differentiate into mature basal cells. Only bipotent stem cells and progenitors have self-renewal capacity as opposed to terminally differentiated cells (A). Cartoon showing how luminal and basal mammary lineages are heterogeneous populations composed of different cells with variable proliferative and self-renewal potential (B).

1.3 Breast cancer

Breast cancer is the leading cause of cancer death in women, accounting for 22% of all female cancers (Jemal et al. 2010). Every year more than 1.3 million women worldwide are diagnosed for breast cancer, making it the second most common type of cancer behind lung cancer (Kamangar et al. 2006). In the last decades much investment in funding has led to a research boost into our understanding of breast cancer origin, pathology and treatment options. These efforts have resulted in a marked improvement in the survival rates for breast cancer where in the sixties, only 35% of women diagnosed with breast cancer in the United States would have been alive ten years later whereas that percentage increased to 77% by the mid-nineties. However half-a-million women still die from this disease each year indicating that further improvements in breast cancer treatment are still needed.

1.3.1 Aetiology, progression of disease and treatment

Although the aetiology of human breast cancer remains largely unknown, risk factors associated with this disease can be grouped into three broad categories: hereditary, hormonal and reproductive factors, and environmental (Hankinson et al. 2004). Family history is one of the strongest determinants of risk. Hereditary factors are associated primarily with early-onset premenopausal breast cancer. In particular the *BRCA1* gene has been implicated in the pathogenesis of hereditary breast cancer whereby women with mutations in either *BRCA1* or *BRCA2* have a 60% to 85% lifetime risk for developing breast or ovarian cancer (Fackenthal et al. 2007). Germline mutations in p53 account for 1% of early onset breast cancer (Gasco et al. 2002). Interestingly women with Li-Fraumeni syndrome (carrying a mutated p53 gene) who survive childhood cancers often develop breast cancer, underlying the genetic connection between loss of p53 and this pathology (Gasco et al. 2002). A link between hormonal and reproductive status of patients and breast cancer is strongly suggested by the association between the incidence of this tumour and the age of menarche, menopause and first pregnancy (Guinee et al. 1994). In particular nulliparous women, or those who become pregnant after the age of 35, have a twofold to threefold higher risk of breast cancer than those women whose pregnancy occurred before age 25. Several environmental factors have been found to

influence breast cancer risk such as smoking, dietary factors and chemical compounds (DeBruin & Josephy 2002).

The natural history of breast cancer involves progression through defined pathological and clinical stages, beginning with ductal hyperproliferation, which can then evolve into *in situ* and invasive carcinomas (Allred et al. 2001). Invasive breast cancer spreads regionally by direct extension through the chest wall or via lymphatic channels to regional lymph nodes (Lee, 1983). The axillary lymph nodes are the most common site of regional metastasis and breast cancer patients are usually stratified into node negative or positive based on excision of axillary lymph nodes (Devitt, 1965). Metastatic breast cancer can then spread from regional sites to colonize distant organs; the most common sites of distant metastasis are lung and pleura, liver, bone, skin, adrenal and brain (Lee 1983). Interestingly, unlike many tumours, breast cancers can recur and metastasize decades after primary tumour excision (Gerber et al. 2010). Not surprisingly the presence of distant metastasis correlates with poor prognosis since no treatments currently exist for metastatic breast cancer.

Several factors determine the choice of treatment for a newly diagnosed breast cancer such as the age of the patient, involvement of axillar lymph nodes, the size of the tumour, histological grade, expression of hormonal receptors and HER2 status. The first type of treatment for breast cancer is usually surgery followed by adjuvant chemotherapy, radiotherapy or in some cases, hormone or targeted treatments. The choice of treatment is strictly dependent on the type of breast cancer identified. Focusing on the receptor status human breast cancer can be subdivided into three main groups: oestrogen receptor positive (ER+), epidermal growth factor receptor 2 positive (HER2+) and triple negative (ER-/PR-/HER2-). ER+ and HER2+ patients benefit from targeted treatments such as Tamoxifen and Trastuzumab which have consistently improved disease outcome (Howard et al. 2012). On the other hand, the triple negative subtype lacks any specific targeted therapy and it has been associated with worse overall prognosis in comparison with the other subtypes (Foulkes et al. 2010).

1.3.2 Molecular classification of breast cancer

Human breast cancers have been categorized by pathologists into various histological subtypes based on a set of features of the primary tumour at the time of diagnosis such as lesion size, necrosis, nuclear grade and mitotic index. However clinical responses of patients to therapy showed considerable heterogeneity among the same subtype, indicating the need for a new, better classification method (Stingl & Caldas 2007). In fact the histological classification of human breast tumours is confounded by a number of factors including scoring subjectivity and intra-inter tumour heterogeneity. In recent times a new way to classify human breast tumours, based on gene expression profiling by microarray analysis, has provided a new perspective on breast cancer. Breast cancer is no longer a single disease but a collection of distinct malignancies which vary from the clinical, morphological and molecular point of view (Stingl & Caldas 2007). A pivotal study from Perou and co-workers was the first to provide a molecular classification for breast cancer (Perou et al. 2000). Gene expression profiling of 38 breast cancer cases revealed four distinct molecular subtypes: luminal, HER2, basal-like and normal breast (Perou et al. 2000). A follow-up study using a larger cohort of patients refined this classification showing that the luminal subgroup could be divided into at least two groups (luminal A and B), and that different molecular subtypes were associated with different prognosis (Sorlie et al. 2003). The main molecular subtypes are described below.

Luminal A

The luminal A group comprises 50–60% of all diagnosed breast tumours, representing the most common subtype. It is characterized by the expression of genes typically expressed in the luminal epithelium and which are linked to ER transcriptional regulation (Prat & Perou 2011). In particular the immunohistochemistry profile of the luminal A subtype is characterized by expression of ER, PR, Bcl-2, GATA3 and cytokeratin CK8/18, and absence of HER2 expression. Moreover these tumours show a low rate of proliferation and a low histological grade. Luminal A patients have a generally good prognosis with a relapse rate significantly lower than that of other subtypes (27.8%) (Colleoni et al. 2012). However this subtype of breast cancer has the higher incidence of

bone metastases (18.7%) compared to other sites such as liver and lungs (Kennecke et al. 2010). The treatment of luminal A tumours is mainly based on hormonal aromatase inhibitors, which work by inhibiting the action of the enzyme aromatase, responsible for the synthesis of oestrogen, and selective oestrogen receptor modulators like Tamoxifen, a competitive inhibitor of the oestrogen-oestrogen receptor binding (Guarneri & Conte 2009).

Luminal B

The luminal B group makes up 10–20% of all breast tumours. Compared to the luminal A group, they have a more aggressive phenotype, higher histological grade and worse prognosis (Colleoni et al. 2012). Both luminal A and B express ER and the main biological difference between the two subtypes is an increased expression of proliferation genes, such as Ki67 and cyclin-B1, and growth factor receptors EGFR and HER2 in the luminal B group. Bone is still the most common site of recurrence (30%), together with a high recurrence rate in other organs such as the liver (Kennecke et al. 2010). Luminal B tumours are treated with Tamoxifen and aromatase inhibitor. However the worse prognosis compared to luminal A tumours, underlines the need of new therapeutic options for this subgroup.

HER2 positive

HER2 positive tumours represent 15-20% of breast cancers. They are characterized by a high expression of the HER2 gene and other genes associated with the HER2 pathway. Morphologically, these tumours are highly proliferative, with a high histological grade and frequent p53 mutations (Montemurro et al. 2013). HER2 amplified tumours are a heterogeneous group and have been further subdivided into three separate subtypes; one characterized by worse prognosis with 12% of patients alive after 10 years, compared to the 50–55% survival in the other two groups (Staaf et al. 2010). From the clinical point of view, the HER2 subtype is characterized by a poor prognosis, although the introduction of anti-HER2 treatment has substantially improved survival in both primary and metastatic diseases (Slamon et al. 2001).

Basal-like and triple negative

The basal-like subtype accounts for 15% to 20% of newly diagnosed breast cancer cases. One of the most relevant features of this subtype is the lack of the three key receptors in breast cancer: oestrogen, progesterone and HER2. For this reason the basal-like group is also defined as triple-negative breast cancer (TNBC). However some controversies exist on the definition of TNBC and basal-like breast cancers (Gusterson 2009) and although triple-negative is not synonymous with basal-like tumours both groups show a wide range of similarities (Rakha et al. 2007). At the morphologic level, TNBC and basal-like tumours share similar characteristics such as larger tumour size, higher grade, presence of necrosis, pushing borders of invasion, and lymphocytic infiltrate (Livasy et al. 2006). The majority of TNBCs are invasive ductal carcinoma, but less common histologic subtypes (for example medullary and metaplastic) are also represented in this group (Bertucci et al. 2006; Weigelt et al. 2009). Basal-like and TNBC tumours usually express genes characteristic of mammary myoepithelial cells, including cytokeratins CK5, CK14 and CK17, P-cadherin, Caveolin 1 and 2, Vimentin and EGF-R (Nielsen et al. 2004). However low levels of genes characteristic of luminal epithelium such as CK8/18 and cKIT can be found in these subtypes (Livasy et al. 2006). Moreover TNBC and basal-like tumours frequently show high expression of proliferation markers (Ki67) and activation of the beta-catenin pathway (Geyer et al. 2011; Arnedos et al. 2012). Interestingly, tumours with germ-line mutations in the *BRCA1* gene belong to these groups (Sorlie et al. 2003).

From a clinical point of view they are characterized by appearance at an early age (<40 years old), and predominantly in women of African origin (Metzger-Filho et al. 2012). The pattern of metastatic relapse is aggressive, mainly localized to lungs, central nervous system and lymph nodes (Kennecke et al. 2010). Basal-like tumours have a worse prognosis compared to the luminal subgroups, with a higher relapse rate in the first 3 years despite showing a high response to chemotherapy (Rouzier et al. 2005). Hence the identification of new therapeutic targets and treatment strategies for TNBC is urgently needed. Given the extreme heterogeneity of TNBC and basal-like tumours, effort has been put into trying to identify biologically distinct TNBC subgroups. In 2007, a new triple negative molecular subtype named claudin-low, was identified (Herschkowitz et

al. 2007). These distinct tumours were characterized by the low gene expression of tight junction proteins claudin 3, 4 and 7 and E-cadherin. A follow up study showed that claudin-low tumours are enriched in epithelial-to-mesenchymal transition (EMT) features, immune system responses, and stem cell-associated genes (Prat et al. 2010). In addition four independent gene clusters were identified in the TNBC subgroup using the transcriptome data from 21 independent breast cancer studies (Lehmann et al. 2011). These clusters were defined by mesenchymal features, immune system–related genes, DNA damage response genes and activated androgen receptor signalling.

1.3.3 Understanding the molecular portrait of breast cancer

The implementation of high throughput technologies has opened our eyes on the intimate molecular heterogeneity of breast cancer. However the complete molecular portrait of breast cancer is still far from being completed. The continuous development and improvement of high-throughput techniques means that more powerful, fast and reliable methods are available every year, helping to dig more deeply into the molecular architecture of breast cancer (Gray & Druker 2012). In 2012, five independent whole-genome analyses, through massively parallel DNA and RNA sequencing of various breast cancers, refined our current tumour-classification system with the finding of new subtypes among the categories established by early transcriptome profiling studies (Ellis et al. 2012; Stephens et al. 2012; Shah et al. 2012; Curtis et al. 2012; Banerji et al. 2012). For example Curtis and co-workers identified new breast-cancer subtypes that are associated with different patient outcomes, based on the combination of copy number and gene expression analysis (Curtis et al. 2012). Banejerji and co-workers instead, identified three new molecular aberrations implicated in breast cancer (mutations in the genes CBF-ß, RUNX1 and the MAGI3–AKT3 gene fusion) (Banerji et al. 2012). In a recent study, a diverse set of breast tumours were assayed using DNA methylation, microRNA (miRNA) expression and protein expression, combined with mRNA expression profiling, DNA copy number analysis and massively parallel sequencing (Cancer Genome Atlas Network 2012). The information generated was then integrated across platforms demonstrating the existence of four main breast cancer classes, each of which shows significant molecular heterogeneity. Moreover somatic mutations in only three genes (*TP53*, *PIK3CA* and *GATA3*) occurred in more than 10% of all breast cancers while

several subtype-specific and novel gene mutations were identified (Cancer Genome Atlas Network 2012). Overall these studies show that individual breast cancers typically carry a few consistent characterized abnormalities, along with tens to thousands of other changes that are rare or unique to the individual tumour and about which little is known. The challenge in the future will be to understand which of these genes are responsible for tumourigenesis and how this ensemble of aberrations collaborates to drive tumour growth and response to therapy.

1.3.4 The origin of breast cancer subtypes

The previous studies showed that breast cancer can be viewed as a collection of multiple different diseases, each one characterized by specific molecular, histological and clinical features. Two competing models have been suggested to explain this extensive heterogeneity: the clonal evolution and the cancer stem cell hypotheses (Shackleton et al. 2009). According to the clonal evolution hypothesis breast cancer subtypes do not mirror the features of normal cell populations and the unique subtype-specific gene expression patterns are rather generated by evolution of transformed cells during decades of clonal selection influenced by a particular tumour genetic background. According to the cancer stem cell hypothesis instead, accumulation of specific mutations in a particular cell type of the normal mammary epithelium generates transformed multipotent cells (cancer stem cells) which will then give rise to a specific breast cancer subtype (Smalley et al. 2003). In this view the molecular features of each subtype are mirroring the characteristic of the cell type which was originally transformed. For example MaSC are thought to be the cell of origin of triple negative breast cancer based on their shared features such as lack of expression of hormone receptors and expression of basal cytokeratins (Perou 2010). On the other hand luminal A tumours, are thought to derive from relatively well-differentiated cells of the ER+ lineage, whereas luminal B tumours are believed to develop from less differentiated luminal progenitors (Polyak 2007). The theory of the cell of origin of breast cancer is still controversial and further studies are needed to corroborate this hypothesis. In particular a better characterization of the mammary epithelial cell hierarchy, through the identification of new markers to identify stem, progenitor and differentiated cells of the mammary epithelium, is essential to allow a better understanding of

the cellular targets of breast cancer mutations. By determining the identity of the original cell population that was transformed to form the cancer stem cell population, investigators will be able to identify early key steps in the formation of breast cancer, which could potentially translate into novel therapeutic targets.

1.4 The *RUNX* Genes

The *Runx* genes represent a family of different transcription factors characterized by a common feature: a highly conserved 128 amino acid sequence known as the Runt domain. The Runt domain, located in the amino-terminal region of these proteins, is required for their DNA binding activity and for protein-protein interactions (Levanon & Groner 2004). *Runx* genes regulate the transcription process binding the DNA as a multicomponent complex together with the DNA binding partner CBF-β cofactor and a wide variety of tissue specific co-repressors and co-activators (Chuang et al. 2012). RUNX transcription factors are involved in several fundamental processes during animal development, playing a key role in the regulation of cell differentiation and proliferation (Coffman 2009). In addition and strictly connected to this decision, *Runx* genes have been found to take part in cancer development, although their ambiguous roles of oncogenes or tumour suppressors have yet to be clarified (Blyth et al., 2005; Blyth et al., 2010).

1.4.1 The *Runx* genes: evolution, structure and regulation

1.4.1.1 The Runx genes: an evolutionary look.

The RUNX proteins are well represented among metazoan organisms, from mouse to sea urchin and sponges, underlining the importance of this transcriptional regulatory system (Braun & Woollard 2009; Sullivan et al. 2008). While lower organisms such as *C. elegans* usually have a single *Runx* gene, more evolved organisms possess additional *Runx* genes probably deriving from ancient duplication events. These genes have then acquired different lineage and tissue specificity while becoming intertwined to distinct cell regulatory systems. As a result RUNX family members have developed exclusive gene-specific functions.

The first *Runx* gene to be discovered, *runt*, was identified as a major player in *Drosophila* early segmentation process (Gergen & Butler 1988) and later in neuronal development and sex determination (Duffy et al. 1991). Another *Drosophila Runx* homolog, *lozenge*, was found to regulate cell fate specification during haematopoiesis and to control cell patterning during eye development (Daga et al. 1996; Lebestky et al. 2000). Moreover the *Drosophila* genome project identified 2 additional Runx homologs, CG34145 and CG42267, whose function is still unknown (Rennert et al. 2003).

The sea urchin *S. purpuratus* has two *Runx* genes but only one, called *SpRunt-1* has been characterised so far (Fernandez-Guerra et al. 2006). *SpRunt-1* is involved in the regulation of both cell proliferation and terminal differentiation in all major tissues during different stages of embryogenesis (Coffman et al. 2004). In particular *SpRunt-1* is generally required in early embryonic development to support cell division, and it is required later on to activate expression of structural genes involved in terminal differentiation in a specific subset of tissues (Coffman et al. 2004). Further studies in sea urchin have unveiled an anti-apoptotic role for RUNX during gastrulation (Dickey-Sims et al. 2005) and the existence of a crosstalk between *Runx* and *Wnt* signalling during early development (Robertson et al. 2008).

Since its genome contains only one *Runx* gene, *rnt-1*, *C. elegans* is an invertebrate model organism commonly used for investigating the role of *Runx* in development avoiding redundancy problems. Genetic analysis have shown a key role of *rnt-1* as a regulator of cell division patterns in seam cells, a population of well characterized neuroectodermal stem cells located in the worm epidermis (Nimmo & Woollard 2008). In particular in *rnt-1* mutants, symmetrical cell division of seam stem cells is drastically impaired, leading to a reduction in their number in adult animals (Nimmo et al. 2005). Moreover overexpression of either *rnt-1* or *bro-1*, a homologue of mammalian *CBF-β*, leads to increase of both asymmetrical and symmetrical divisions among seam cells, resulting in massive hyperplasia and a tumour-like phenotype (Kagoshima et al. 2007).

The key role of *Runx* genes as important regulators of stem cells is also emerging from studies on Planarians (Rink 2013). Planarian flatworms are capable of regenerating any missing body part and represent an attractive model for the

investigation of the molecular pathway underlying tissue regeneration. Planarian regeneration uses a population of regenerative cells (neoblasts) which includes pluripotent stem cells. Interestingly, after wounding *runt-1* is activated directly within neoblasts where it is necessary for specifying different cell types during regeneration (Wenemoser et al. 2012).

The recurrent theme of *Runx* genes as specific lineage determinants is further confirmed in more complex vertebrate animal models. In particular in Zebrafish and *Xenopus Runx1* is expressed in hematopoietic progenitors where it controls stem cell specification (Tracey et al. 1998; Kalev-Zylinska et al. 2002; Burns et al. 2005). Moreover in *Xenopus*, *Runx1* is also required for the development of Rohon/Beard neurons, a particular class of sensory neurons which mediate response to touch in the larval stage (Park et al. 2012). In both fish and frogs, *Runx2* is detected in developing skeletal elements where it regulates chondrogenesis (Flores et al. 2004; Flores et al. 2006; Kerney et al. 2007) while in Zebrafish, *Runx3* has been studied in the context of haematopoiesis and chondrocyte differentiation (Kalev-Zylinska et al. 2003).

The mammalian genome contains three different *Runx* genes: *Runx1*, *Runx2* and *Runx3*. These genes maintain strong structural similarities but they have acquired lineage-specific roles. RUNX1 is required for the development of hematopoietic lineages; RUNX2 is a master regulator of bone formation while RUNX3 has a role in neuronal development (Cohen 2009). Taken together, the evidence gathered from different animal models, separated by millions of years of evolution, helps to build a general picture of *Runx* genes as key transcription factors involved in the regulation of lineage determination and self/renewal through a finely tuned action on cell proliferation and differentiation. Such a key role in diverse developmental processes implies that *Runx* genes need to be tightly controlled at the transcriptional and post-transcriptional level.

1.4.1.2 The Runx Genes: genomic and protein structure.

Runx genes are located on different chromosomes which are species-specific. *RUNX*1 is localised on human chromosome 21, *RUNX*2 on chromosome 6 and *RUNX*3 on chromosome 1 while in mouse they are located on chromosomes 16, 17 and 4 respectively (Stock & Otto 2005). The three mammalian *RUNX* genes

share a similar genomic structure which can also be found in two neighbouring paralog genes outside the *RUNX* locus: *CLIC* and *DSCR* (Levanon & Groner 2004). *RUNX*1 is the biggest gene with 9 exons and 12 possible isoforms, *RUNX*2 has 8 exons and 12 possible isoforms, while *RUNX*3 is the smallest one with 6 exons and few isoforms (Bangsow et al. 2001).

RUNX proteins have an average size of 50 KDa, ranging from the 44 KDa of RUNX3 to the 57KDa of RUNX2. Despite the difference in size, RUNX proteins contain several conserved regions. At the N-terminus is the highly conserved Runt domain, which is shared by all members of the RUNX family with a degree of homology close to 90%. RUNX proteins bind through their Runt domains, to a conserved nucleotide sequence (R/TACCRCA). Structural studies have determined the three-dimensional conformation of the Runt domain in its DNA-bound state showing that it forms an immunoglobulin (Ig) fold (Nagata et al. 1999). The Ig fold is a common domain, mainly involved in molecular recognition and binding, that can be found in the DNA-binding domain of other transcription factors such as P53, NF-kB, NFAT, and STAT (Williams & Barclay 1988). However, unlike other transcription factors, the Runt domain uses loop regions located at the end of the Ig fold for DNA binding (Berardi et al. 1999). Since the Runt domain has a low DNA-binding affinity, during the transcriptional process DNA binding usually occurs together with the heterodimeric binding partner CBF-β.

In addition to the Runt domain, necessary for protein-DNA binding capability, RUNX proteins share other common features. All these proteins have a large C-terminal transactivation domain and a C-terminal inhibitory domain (ID) which downregulates protein expression. Other domains, localized in the C-terminal region, further contribute to RUNX regulation: a NMTS sequence, required for nuclear targeting of RUNX activated proteins (Zaidi et al. 2001), a VWRPY sequence, which binds to the co-repressor Groucho/TLE (Imai et al. 1998) and a PY sequence, a proline-rich motif required for protein interactions (Terry et al. 2004). Studies conducted on RUNX1 also show transactivation properties in the N-terminal and the α-helix of Runt domain (Liu et al. 2006). Interestingly RUNX family members show less homology in the areas outside the Runt domain: the presence of isoform-specific regulatory elements and protein binding sequences may account for the functional differences observed between RUNX proteins

(Levanon & Groner 2004). For example, RUNX2 is the only RUNX protein with a specific glutamine-alanine repeat domain (QA), localized in N-terminal, which is required for osteoblast lineage specification (Thirunavukkarasu et al. 1998).

1.4.1.3 Transcriptional regulation of RUNX genes.

Given their essential role in development, *RUNX* genes are integrated into a complex regulatory network which acts both at the transcriptional and post-transcriptional level. *RUNX* genes have two different promoters, P1 (distal) and P2 (proximal), resulting in two main isoforms with distinct 5'-UTRs sequences (Levanon & Groner 2004). Interestingly, both P1 and P2 genomic regions contain several *RUNX* binding sites suggesting the possibility of cross-regulation between the *RUNX* genes (Otto et al. 2003). Indeed cross-regulation has been reported whereby *RUNX*3 overexpression downregulates *RUNX*1 in human B cell lines (Spender et al. 2005) while *RUNX*2 represses *RUNX*3 in tooth development (X.-P. Wang et al. 2005). Moreover autoregulation of *RUNX* genes has been shown *in vitro* with overexpression of *RUNX*2 in NIH3T3 cells downregulating *RUNX*2 transcription in a negative feedback loop (Drissi et al. 2000).

1.4.1.4 Translational and post-translational regulation of RUNX genes.

RUNX gene expression is tightly regulated at the translational level through the two distinct 5'UTRs sequences (called P1-5'UTR and P2-5'UTR) of the P1 and P2 isoforms. In *RUNX*1 P1-5'UTR directs cap-mediated translation while P2-5'UTR contains an internal ribosomal entry sequence (IRES) which regulates its translation (Pozner et al. 2000). In *RUNX*2 both P1 and P2-5'UTR contain IRES sequences (Xiao et al. 2003) and *RUNX*3 P2-5'UTR exhibits internal translation initiation capability (Bone et al. 2010). IRES elements are involved in translational regulation when cap-dependent system is impaired, such as during mitosis, suggesting a particular role for these isoforms in proliferating cells (Vagner et al. 2001).

Transcription factors can be further regulated through a wide variety of post-translational modifications such as phosphorylation, acetylation, ubiquitination, SUMOylation, and methylation, which are necessary for fine-tuning of transcriptional regulation. These modifications control various aspects of transcriptional factors activity such as auto inhibition, dimerization, proteolytic

cleavage, cellular localization and ubiquitin-mediated degradation. Not surprisingly, RUNX activity is dynamically regulated by a growing list of post-translational modifications (reviewed in Wang et al. 2009). RUNX proteins can be regulated through phosphorylation mediated by various kinases such as ERK, SRC and Cyclin/Cdks (Chuang et al. 2012). For example ERK1/2 dependent phosphorylation of RUNX2 is critical for osteoblast-specific gene expression (Xiao et al. 2000) and RUNX2 activity is stimulated after phosphorylation by ERK1/2 and p38 MAP kinase (Ge et al. 2012). Moreover, the serine/threonine kinase PIM-1, which has been implicated in cytokine-dependent signalling in hematopoietic cells, interacts and phosphorylates RUNX1 and RUNX3 enhancing their transactivation activity (Aho et al. 2006), while SRC-mediated phosphorylation causes a re-localization of RUNX3 to the cytoplasm in breast cancer cells (Goh et al. 2010). Acetylation of transcription factors, mediated by histone acetyltransferase (HAT) proteins, leads to changes in protein-protein and protein-DNA interaction, which subsequently influence gene expression (Sterner & Berger 2000). All mammalian RUNX proteins can be acetylated by p300 resulting in increased DNA binding ability, protein stability and stimulated transactivation properties (Yamaguchi et al. 2004; Jin et al. 2004; Jeon et al. 2006). The necessity of a fine control on RUNX activity in response to different stimuli means that RUNX proteins levels are controlled post-translationally by ubiquitin-mediated proteasome degradation. Various E3 ubiquitin ligases such as MDM2, promote degradation of RUNX proteins (Zhao et al. 2003; Chi et al. 2009), while proteins such as CBF-β and SIN3A interact with RUNX preventing their proteasomal degradation (Huang et al. 2001).

1.4.1.5 Co-repressor and co-activators.

RUNX and CBFβ heterodimers are relatively weak transcriptional factors but they can increase their efficiency through recruitment of a wide array of co-activators and co-repressors. Moreover the vast array of these co-regulators is pivotal to guarantee specificity and accuracy in RUNX-mediated transcriptional regulation. Several transcription factors have been found to interact with *RUNX* genes to drive transcriptional activation or repression depending on the context. The list of transcription factors interacting with RUNX proteins contains several which play key roles in development such as AP1 (c-Fos and c-Jun), BMP-responsive SMADs (SMAD1 and SMAD5), Hes1, GATA-1, STATs and Oct-1 among

others (Chuang et al. 2012). Interestingly many of these transcription factors such as STAT5 and SOX9, interact with RUNX via the Runt domain, suggesting that they may compete with each other for binding to RUNX (Zhou et al. 2006; Ogawa et al. 2008). It is generally believed that transcription factors cooperate with RUNX proteins to facilitate the recruitment of additional co-regulators and the assembly of higher-order transactivation complexes. In fact, other RUNX interactors do not bind to the DNA directly, but are recruited to the template by DNA-bound RUNX proteins which act as a scaffold. TLE proteins, the human homologs of *Drosophila* Groucho and Grg proteins in mice, are broadly expressed co-repressors which are recruited to promoters/enhancers by numerous transcription factors (Cinnamon & Paroush 2008). The TLE/Groucho family is one of the first characterized groups of co-repressors, found to interact with RUNX through the highly conserved VWRPY sequence, present at the C-terminus of all Runt domain proteins (Aronson et al. 1997). Interestingly the interaction between TLE/Groucho and RUNX proteins results in transcriptional repression of osteoblast (Javed et al. 2000) and T-cell specific genes (Seo et al. 2012). In addition RUNX proteins can also recruit different co-repressors such as mSIN3A (Lutterbach et al. 2000) and CoAA (Li et al. 2009). The mechanism whereby those co-repressors inhibit RUNX-mediated transcription is still not clear.

A growing list of chromatin-modifying proteins has been shown to interact with RUNX proteins acting as both activator and repressor. RUNX proteins interact with the acetyltransferase p300 through their C-terminal region, leading to stimulation of RUNX-mediated transcription (Kitabayashi et al. 1998; Jin et al. 2004; Boumah et al. 2009). Others acetyltransferases, such as PCAF (Wang et al. 2013), MOZ and MORF (Pelletier et al. 2002; Kitabayashi et al. 2001) physically interact with RUNX proteins, strongly stimulating their transactivation activity. RUNX proteins can also associate with histone-modifying enzymes, such as histone acetyltransferase HDAC, to induce transcription repression (Ali et al. 2012). RUNX proteins have also been shown to collaborate with the SWI/SNF chromatin modelling complex for transcription activation. In particular RUNX1 and RUNX2 recruit the SWI/SNF complex to specific target genes to regulate respectively hematopoietic and osteoblast specific genes (Villagra et al. 2006; Bakshi et al. 2010). The repertoire of proteins that cooperate with RUNX in

transcriptional regulation continues to increase year after year, reflecting the necessity of fine tuning of RUNX activity according to specific cell contexts.

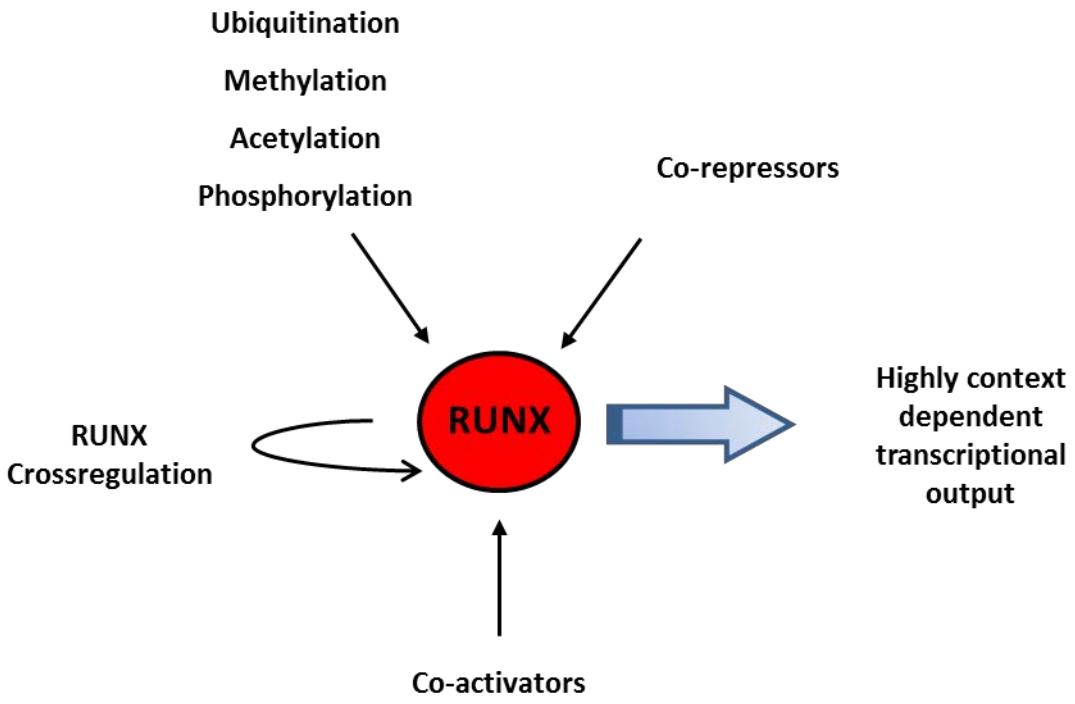


Figure 1-2 Post-transcriptional regulation of RUNX proteins.

RUNX proteins can be regulated through various post-transcriptional modifications and by interactions with tissue-specific co-repressors and co-activators. In addition RUNX proteins can interact and cross-regulate with each other. Thus RUNX-mediated transcription is highly context dependent.

1.4.2 *RUNX* Genes in Development

RUNX genes are fundamental in mammalian development, as underlined by the severity of individual gene knockout phenotypes in transgenic mice. In particular *Runx1* is required for haematopoiesis (Okuda et al. 1996), *Runx2* for osteogenesis (Otto et al. 1997) and *Runx3* for neurogenesis (Inoue et al. 2002). Several studies have also uncovered various developmental roles for *Runx* genes in a wide variety of tissues and organs.

1.4.2.1 RUNX1

RUNX1 has a primary role in haematopoiesis (reviewed in Swiers et al. 2010) since *Runx1* null mice fail to generate haematopoietic cells of all lineages (Okuda et al. 1996). *Runx1* deficient embryos die at E12.5 with diffuse haemorrhages within the ventricle of the central nervous system and vertebral canal. Interestingly blood cells in the sites of haemorrhaging consist only of primitive erythrocytes, since all definitive blood cells are absent from *Runx1* deficient embryos. This failure is due to a blockage in haematopoietic stem cell differentiation during embryo development (North et al. 1999). Further studies revealed the precise step of definitive haematopoiesis where *Runx1* is required: *Runx1* is expressed in a specific subset of endothelial cells where it is essential for the formation of intra-aortic haematopoietic clusters and their differentiation into haematopoietic progenitor and stem cells (Chen et al. 2009). This key finding helped to confirm the highly debated theory of "hemogenic endothelium" (Jordan 1916): during embryogenesis, haematopoietic stem and progenitor cells are generated, in a *Runx1*-dependent way, from a unique population of vascular endothelium termed hemogenic endothelial cells (Antas et al. 2013). Others studies using conditional knockout mouse models have confirmed the master role of *Runx1* in haematopoiesis, not only at an embryonic stage but also during differentiation of various adult lineages. Haematopoietic precursors differentiate into mature T-cells through discrete stages which can be defined based on expression of T-cell surface antigens CD4 and CD8. Immature thymocytes lacking CD4 and CD8 co-receptors (CD4-CD8-), progress through a double positive stage (CD4+CD8+), and are then selected to become either T-helper cells (CD4+CD8-) or cytotoxic T-cells (CD4-CD8+). *Runx1* is necessary for T-cell specification with a role in repressing CD4 expression in immature

thymocytes (Taniuchi et al. 2002). Moreover inducible *Runx1* deletion in the bone marrow showed that *Runx1* is necessary for megakaryocytic maturation and differentiation of T- and B-cells (Ichikawa et al. 2004; Niebuhr et al. 2013).

In addition to its role in the haematopoietic system *Runx1* takes part in developmental processes regulating lineage specification and differentiation in other organs. During embryonic bone development *Runx1* is expressed in pre-chondrocytic tissue, after birth it is expressed in resting zone chondrocytes and suture lines of the calvarium, and in the adult, in periosteal and perichondral membranes of all bones but is absent in mature cartilage or mineralized bone (Yamashiro et al. 2002; Lian et al. 2003). The expression of *Runx1* at sites of cartilage growth, suggests that *Runx1* expression may be related to chondroprogenitor cell differentiation (Y. Wang et al. 2005; Soung et al. 2012). However, *in vivo, Runx1* conditional knockout mice showed that *Runx1* is not essential for major skeletal growth, but plays a role in the development of the sternum and some skull elements (Kimura et al. 2010; Liakhovitskaia et al. 2010).

Genetic studies on Drosophila demonstrated that *Runx* genes are involved in neuronal development (Duffy et al. 1991) with expression in neural cells of the developing nervous system of Zebrafish and mouse (Zagami et al. 2009). Dorsal root ganglion neurons (DRG) convey peripheral somatosensory stimuli to the spinal cord. They can be subdivided in three major subpopulations – nociceptive, mechanoreceptive, and proprioceptive. *Runx1* is specifically expressed in DRG nociceptive neurons where it controls lineage specification and regulates axonal outgrowth and guidance in the developing embryo (Kramer et al. 2006; Marmigère et al. 2006). Outside DRG neurons, *Runx1* plays a role in regulating survival of post-mitotic neurons of the embryonic central and peripheral nervous system (Theriault et al. 2004) and regulates neural progenitor cell proliferation in olfactory receptor neurons (Theriault et al. 2005).

In recent years, studies on the hair follicle linked RUNX1 to lineage differentiation of epithelial tissues (Scheitz & Tumbar 2012). *Runx1* is expressed in a subset of cells during embryonic and adult hair follicle formation (Osorio et al. 2011). Interestingly, lineage tracing experiments showed that embryonic *Runx1* positive epithelial cells contribute to all lineages of hair follicle during,

morphogenesis and adult homeostasis, proving that *Runx1* expressing cells are precursors of adult hair follicle stem cells (HFSCs) (Osorio et al. 2011). Moreover, while the majority of adult epithelial *Runx1* expressing cells are short-lived progenitors (Osorio et al. 2011), Scheitz et al. 2012 showed that the adult bulge cells expressing *Runx1* are HFSCs since they contribute long-term to HF homeostasis as well as to physical injury repair. Induced epithelial knockout of *Runx1* has been used to characterize the role of *Runx1* in skin development. However *Runx*1 loss was only found to affect hair structure (Raveh et al. 2006) and to cause a temporary delay in morphogenesis and hair cycle which is overcome with age and injury (Osorio et al. 2008). Remarkably, a much more profound effect on hair follicle integrity was caused by *Runx1* loss in the embryonic skin mesenchyme, an important HFSCs niche component. After *Runx1* deletion in the mesenchyme, the hair follicle shows no abnormalities during morphogenesis, but in the first hair cycle, when adult stem cells generate the differentiated hair lineages, hair follicles are converted to enormous sebaceous cysts (Osorio et al. 2011). Overall these studies identified RUNX1 as a key transcriptional factor regulating hair follicle differentiation. Focusing on different epithelia, RUNX1 expression has been observed in other tissues specifically in the basal layer of the oral epithelium, and in LGR5+ cells of the intestinal crypt and in the more differentiated villus cells (Scheitz et al. 2012). RUNX1 expression has been detected also in mammary epithelium as will be discussed in detail in Section 1.5.

1.4.2.2 RUNX2

RUNX2 has a primary developmental role as a key lineage determinant for osteoblast differentiation. In fact *Runx2* knockout mice exhibit complete lack of bone formation and die soon after birth because of asphyxia caused by lack of ossification in the ribs (Komori et al. 1997; Otto et al. 1997). Moreover, deletion of a nuclear targeting signal located in the C-terminal domain of *Runx2* phenocopied *Runx2*-null mice, highlighting the importance of proper nuclear localization of this transcription factor (Choi et al. 2001). Mutations of *RUNX*2 resulting in a hypomorphic allele, are linked to a congenital human disease, cleidocranial dysplasia (CCD), characterized by abnormal clavicles, supernumerary teeth, short stature, and a variety of other skeletal changes (Mundlos et al. 1997). These mutations affect only one allele of *RUNX*2 leading

to a decrease in wild-type *RUNX*2 levels and activity. Interestingly *Runx*2 haploinsufficiency in mice causes similar phenotypes as observed in CCD families (Komori et al. 1997). Moreover *RUNX*2 is one of few genes undergoing specific modifications in the modern human lineage compared to Neanderthal genomes (Green et al. 2010) suggesting that an evolutionary change in *RUNX*2 was involved in the determination of morphological traits which characterize modern human skeletal development. Those evolutionary observations are confirming the key role for *RUNX*2 as a main regulator of cranial and skeletal features. Studies have clarified the role of *Runx2* in osteoblast differentiation (Long 2011). During embryonic development *Runx2* is expressed in osteochondroprogenitors; bipotent progenitors which have the capacity to differentiate into osteoblasts or chondrocytes (Marie 2008). Then *Runx2* expression decreases in cells which will differentiate into chondrocytes while remains expressed at high levels in cells of the osteoblast lineage and perichondrium. Interestingly, the perichondrium, which normally contains bipotent osteochondroprogenitors, becomes hypoplastic in *Runx2*-null mice demonstrating its requirement for the production and/or maintenance of the bone progenitors (Komori et al. 1997). Thus during skeletal development *Runx2* activates a differentiation pathway in bone marrow-derived mesenchymal stem cells which is necessary for osteoblast differentiation. *Runx2* expression needs to be downregulated to guarantee correct osteoblast terminal differentiation exemplified by mouse models where bone formation is impaired by *Runx2* overexpression (Liu et al. 2001). In particular overexpression of *Runx2* results in increased osteoblast number but inhibits their terminal maturation, resulting in accumulation of less mature osteoblasts and consequent osteopenia (Liu et al. 2001). Moreover osteoclastogenesis is stimulated, possibly by the increased production of RANKL and MMP-13 by the immature osteoblasts (Geoffroy et al. 2002). In addition *Runx2* is also necessary for the proper function of terminally differentiated osteoblasts: mature mice in which active *Runx2* levels have been reduced, exhibit decreased expression of the genes encoding several bone matrix proteins such as BSP, Osteocalcin and Osteopontin (Ducy et al. 1999). These results indicate that RUNX2 proteins levels need to be finely tuned throughout the entire process of osteoblast differentiation to guarantee a correct skeletal development.

Runx2 is also expressed in extra-skeletal tissues where its function is less well understood. Expression of *Runx2* is high in the haematopoietic stem cell population, where it is expressed at higher levels than *Runx1*, and sharply decreases during myeloid differentiation (Kuo et al. 2009). Moreover forced *Runx2* expression in *in vitro* differentiation assays, blocks myeloid progenitor differentiation capacity (Kuo et al. 2009). Besides myeloid differentiation, *Runx2* is also involved in the regulation of the lymphoid lineage. *Runx2* is expressed at the earliest stage of thymocyte development (Satake et al. 1995; Blyth et al. 2010) and enforced expression of *Runx2* in transgenic mice under the CD2 promoter affects T cell development, resulting in an expansion of double-negative and CD8 immature single-positive cells (Vaillant et al. 2002). Moreover additional evidence suggests a role for *Runx2* in B-cell differentiation since *Runx2* transcripts are enriched in a subpopulation of memory B cells (Ehrhardt et al. 2008).

The development of ectoderm-derived appendages results in several highly specialized organs such as hair follicles, mammary glands and teeth (Jiménez-Rojo et al. 2012). Interestingly *Runx2* is emerging as a common regulator of ectodermal-derived epidermal appendages development. It is found expressed during tooth development (Jiang et al. 1999; Bronckers et al. 2001) where it is necessary for tooth morphogenesis and odontoblast differentiation (Camilleri & McDonald 2006). *Runx2*-deficient mice have an impaired tooth development, which is blocked at late bud stage (D'Souza et al. 1999). Moreover, patients with *RUNX*2 mutations show a wide variety of dental disorders, with supernumerary teeth, abnormal tooth eruption, and tooth hypoplasia (Mundlos et al. 1997). *Runx2* is not involved in the initiation of tooth formation, but it is necessary for the regulation of the epithelial –mesenchymal crosstalk required to control tooth morphogenesis (Aberg et al. 2004). Only one study so far has analysed the role of *Runx2* in skin and hair follicle development. Glotzer et al., 2008 found that *Runx2* is expressed in embryonic and adult hair follicles, and it cycles during hair follicle development. Moreover hair follicle maturation is slightly delayed in the absence of *Runx2* and overall skin and epidermal thickness of *Runx2* null embryos is reduced (Glotzer et al. 2008). The role of *Runx2* in mammary development will be discussed in a separate chapter.

In addition *RUNX2* expression has been found in vascular endothelial cells, testis, sperm and other reproductive organs such as placenta, ovary, Mullerian duct and prostate (Sun et al. 2001; Jeong et al. 2008; Blyth et al. 2010). A functional role for *Runx2* in those tissues however has yet to be clarified.

1.4.2.3 RUNX3

In the vertebrate somatosensory system, stimuli are transmitted from the periphery to the spinal cord by sensory neurons located in dorsal root ganglia (DRG) that flank the spinal cord. Two independent studies demonstrated that *Runx3* knockout mice manifest severe limb ataxia due to defective development of proprioceptive neurons in the dorsal root ganglia (DRG) (Inoue et al. 2002; Levanon et al. 2002). During DRG neurogenesis, proprioceptive tyrosine kinase receptor C + (TrkC+) and mechanoreceptive (TrkB+) neurons are derived from a common bipotent precursor (TrkB+, TrkC+). During segregation of the two complementary sensory populations, *Runx3* is specifically expressed in TrkC+ neurons where it represses TrkB expression, acting as a lineage specifier for proprioceptive neurons (Inoue et al. 2002; Levanon et al. 2002). *Runx3* also regulates axonal outgrowth and/or axonal guidance of proprioceptive DRG neurons (Levanon et al. 2002; Chen et al. 2006). As with *Runx1* and *Runx2*, *Runx3* has been found to play a role in T-cell development (Durst & Hiebert 2004). *Runx3* knockout mice have reduced numbers of CD8+ T-cells in the thymus and in the circulating T-cell population together with increased expression of CD4 in the peripheral CD8+ cells (Woolf et al. 2003). In addition, when T-cells were analysed in immunodeficient mice reconstituted with haematopoietic stem cells from the foetal livers of embryos lacking *Runx3* mice, CD4 was de-repressed in the peripheral cytotoxic T cells (Taniuchi et al. 2002). Those studies show how *Runx3* is necessary for CD4 silencing and CD8 T-cell maturation during T-cell development.

In addition to neuronal defects, *Runx3* deleted mice from Ito's group developed hyperplasia of the gastric mucosa and died shortly after birth apparently due to starvation (Li et al. 2002). The increased proliferation rate in the gastric epithelia of *Runx3* knockout mice was attributed to suppression of apoptosis and reduced sensitivity to the growth inhibitory effects of TGF-ß1 (Li et al. 2002). These results suggest a possible role for *Runx3* in regulation of gastric epithelium

homeostasis. However, this phenotype was not found in the *Runx3* knockout mice produced by Levanon et al (2002). The reason for the discrepancy between the different *Runx*3 knockout phenotypes is still not clear, but one factor could be the different genetic backgrounds used in these studies (Bae & Ito 2003).

This overview of *Runx* genes in developmental processes shows how the field has progressed in recent years, moving from a confined view of RUNX proteins as master gene regulators of a few specific lineages to a wider role for RUNX transcription factors in multiple tissues and cell types. Consequently future studies using conditional knockout strategies, will likely find new developmental roles for this widely conserved family of transcription factors. We are just beginning to discover and understand a more global role for the RUNX proteins as regulators of tissue function and maintenance.

1.4.3 *RUNX* Genes and Cancer: an Overview.

As shown previously, *RUNX* genes act as important regulators during mammalian development in a lot of different tissues, through regulation of cell proliferation and differentiation. In addition *RUNX* genes are deeply involved in tuning stem cell fate and lineage determination (Appleford & Woollard 2009; Wang et al. 2010). Therefore it is not surprising that *RUNX* genes have been implicated in cancer. One of the interesting features of *RUNX* genes in cancer is their contrasting behaviour as both oncogenes and oncosuppressors in different types of cancers (Blyth et al. 2005; Pratap et al. 2006; Chuang et al. 2012). An overview is presented here of the evidence linking *RUNX* genes to different types of tumours while the role of *RUNX* genes in breast cancer will be discussed later.

1.4.3.1 RUNX1

The first evidence that linked *RUNX* genes and cancer was the discovery of *RUNX*1 as one of the genes most frequently targeted by chromosomal translocation in acute myeloid leukaemia (AML) (Miyoshi et al. 1991) with three of the most common chromosomal translocations in acute leukaemia involving *RUNX*1 (Lam & Zhang 2012). The first translocation to be discovered, present in 10–20% of adult AML, is the t(8;21) chromosomal translocation, which results in fusion of the N-terminal portion of *RUNX*1, including the Runt domain, to a

heterologous partner protein, ETO (Miyoshi et al. 1991). Another common translocation, *TEL-RUNX*1, is a result of t(12;21) and it is present in about 25% of patients with childhood pre-B cell acute lymphoblastic leukaemia (ALL). *TEL-RUNX*1 produces a fusion with the N-terminal domain of the TEL protein and almost the entire RUNX1 protein, including its DNA binding and transactivation domains (Golub et al. 1995). The third most common translocation involving *RUNX*1 is t(3;21), which causes the fusion of the N-terminal portion of *RUNX*1 including its Runt domain with one of three genes on chromosome three including *EVI*, *MDS1*, or *EAP*. This translocation was first discovered in patients affected by chronic myelogenous leukaemia and in approximately 3% of therapy-related myelodysplastic syndrome (MDS) and AML (Rubin et al. 1987). There is strong evidence that the products arising from translocations involving *RUNX*1 have dominant negative activity with respect to the endogenous *RUNX*1 product. For example fusion of *RUNX*1 to ETO leads to recruitment of co-repressors and active repression of *RUNX*1-mediated gene transcription while *TEL–RUNX*1 has also been shown to function as a constitutive repressor of *RUNX* target genes (Meyers et al. 1995; Hiebert et al. 1996). Some *in vivo* genetic studies are also supporting the dominant negative function of *RUNX*1 translocation as shown by *RUNX*1– *ETO* knock-in mice which produced a very similar phenotype to loss of *RUNX*1 (Yergeau et al. 1997). In the current view, *RUNX*1 translocations cause the block of haematopoietic stem/progenitor cells at an immature developmental stage, where the accumulation of additional collaborating mutations is required to drive the malignancy. Supporting this hypothesis, two studies showed that expression of *RUNX*1–*ETO* in human haematopoietic stem cells or primary erythroid cells greatly increases their survival and self-renewal *in vitro* (Mulloy et al. 2002; Tonks et al. 2003). In addition to translocations, *RUNX*1 is also targeted by somatic point mutations that have been identified in *de novo* and therapy-related AML and MDS (Osato et al. 1999; Harada et al. 2003). Again the majority of these mutations were found to be clustered within the Runt domain probably resulting in dominant-negative forms of *RUNX*1, while other mutations have been found in the C-terminal portion. Additional hints pointing for a tumour suppressor role for *RUNX*1 in the haematopoietic system come from the hereditary disease familial platelet disorder (FPD), an autosomal dominant condition characterized by platelet defects, and propensity to develop AML (Song et al. 1999). FPD patients frequently show monoallelic *RUNX*1

mutations, and biallelic mutations resulting from a second hit are often associated with progression to AML (Preudhomme et al. 2009). Besides a clear tumour suppressive role, several evidences show how *RUNX*1 can function as an oncogene in haematopoietic malignancies (Blyth et al. 2005). Retroviral mutagenesis studies identified *Runx1* as a common insertion site for murine leukaemia virus (MLV), mostly in T- or B-cell lymphomas (Li et al. 1999). Those insertions result in overexpression of a full-length, gene product (Wotton et al. 2002). Supporting the data coming from murine models, human cancers also manifest evidence of a possible oncogenic role for *RUNX*1. The strongest indication so far comes from a small subset (3–5%) of childhood B-ALL with poor prognosis, in which *RUNX*1 is affected by amplification of a large segment of chromosome 21q (Niini et al. 2000; Robinson et al. 2003). *RUNX*1 amplification has also been reported, although much more rarely, in myeloid leukaemias (Roumier et al. 2003). Interestingly *in vitro* studies showed that *Runx1* overexpression causes senescence in wild-type mouse embryonic fibroblasts but induces a transformed phenotype in the absence of functional p53 (Wotton et al. 2004). These observations suggest a scenario where the oncogenic consequences of *RUNX*1 overexpression are manifested only in the presence of a specific genetic background. In addition to haematopoietic malignancies, RUNX1 has also been linked to various epithelial cancers, including breast cancer (discussed in Chapter 3).

1.4.3.2 RUNX2

The oncogenic potential of RUNX2 has been demonstrated in lymphoma models, with the discovery of *Runx2* as a frequent target for viral insertions in T-cell lymphomas of CD2-MYC mice (Stewart et al. 1997). Follow up studies confirmed that *Runx2* overexpression interferes with murine immature T-cell differentiation but other collaborating mutations such as *Myc* activation or *p53* deletion are required for cancer development (Vaillant et al. 1999; Blyth et al. 2001). In particular, the mechanism underlying the potent synergy between *Myc* and *Runx2* in lymphoma development has been studied in depth (Blyth et al. 2006). Double transgenic animals co-expressing *Myc* and *Runx2* uniformly develop tumours by 36 days on average (Blyth et al. 2001). In preneoplastic cells co-expression of MYC counteracts the negative effects of RUNX2 on cell growth and proliferation, releasing the T-cell differentiation block conferred by *Runx2*

overexpression. At the same time, RUNX2 collaborates with MYC by inhibiting its ability to induce apoptosis in T-cell tumours (Blyth et al. 2006). Additional evidence of an oncogenic role for RUNX2 comes from studies on a model for CBFβ-SMMHC (also known as inv16), a common translocation in human AML (Castilla et al. 2004). Retroviral infection of this model identified *Runx2* as a target for insertional mutagenesis (Kuo et al. 2009). Moreover full-length *Runx2* cooperated with CBFβ-SMMHC in leukaemia development in transplantation assays and conversely, *Runx2* haplo-insufficiency delayed the onset and reduced the incidence of acute myeloid leukaemia (Kuo et al. 2009). Focusing on the human disease, so far RUNX2 expression has been detected in plasmacytoid dendritic cell malignancy (Dijkman et al. 2007) and in multiple myeloma (Colla et al. 2005) so although mouse studies indicate a possible oncogenic role for RUNX2 in haematopoietic malignancies, little clinical evidence supporting this hypothesis has been found in human disease.

Focusing on other malignancies, *RUNX*2 genomic locus 6p21 is amplified in osteosarcoma (Lau et al. 2004) and increased expression of *RUNX*2 in osteosarcoma biopsies has been associated with increased tumourigenicity, metastases, lower survival, and poor prognosis (Sadikovic et al. 2010; Kurek et al. 2010). *RUNX*2 is consistently upregulated in papillary carcinomas and thyroid carcinoma cell lines compared with normal thyroid tissue (Endo et al. 2008; Dalle Carbonare et al. 2012). In addition *RUNX*2 silencing on thyroid carcinoma cells caused a decrease in EMT-related molecules and angiogenic factors (Niu et al. 2012). In recent years several studies have also linked RUNX2 to prostate cancer (Pratap et al. 2006; Blyth et al. 2010). In the conditional *Pten*-knockout mouse model, a widely used mouse model of prostate cancer, *Runx2* levels increased with growth of prostate tumour suggesting an oncogenic role for this transcription factor (M. Lim et al. 2010). Furthermore increased RUNX2 expression positively correlated with Gleason scores and metastatic potential of human prostate cancer (Chua et al. 2009; Akech et al. 2010). Several *in vitro* studies have focused on understanding the possible oncogenic mechanism of *Runx2* in prostate cancer. In the PC-3 cell line, RUNX2 can mediate N to E-cadherin switch, activation of AKT and enhanced invasive capacity (Chua et al. 2009). Conversely, knocking down of RUNX2 expression in the PC3 cell line led to increased adhesion to fibronectin and reduced invasion through matrigel (Akech

et al. 2010). In a similar way transcriptomic analysis of gene expression in response to RUNX2 overexpression in the C4-2B cell line, showed up-regulation of genes implicated in cancer progression and cellular migration (Baniwal et al. 2010). Those results showed how RUNX2 can regulate invasive features of prostate cancer cell lines. In addition, RUNX2 positively regulated Survivin, an important inhibitor of apoptosis, thus facilitating cancer cell survival *in vitro* (Akech et al. 2010, Lim et al. 2010).

1.4.3.3 RUNX3

A tumour suppressor role for RUNX3 has been proposed following studies using transgenic knockout mouse models. In particular *Runx3* deficiency in the gastric epithelia is associated with a preneoplastic state characterized by loss of chief cells (Li et al. 2002). *Runx3* deleted mice, unlike wild-type mice, consistently develop gastric cancer after chemical-induced (N-methyl-N-nitrosourea) carcinogenesis (Ito et al. 2011). These results, published by Ito's group, have caused some controversy on the actual role of *Runx3* in stomach cancer. In fact a *Runx3* knock-out model created by Yoram Groner's laboratory showed no evidence of gastric abnormalities (Levanon et al. 2002). Moreover RUNX3 protein was not detectable at the protein level in the gastric epithelium by Groner's group (Levanon et al. 2011), questioning the tumour suppressor role for this protein in this organ. Heterozygous inactivation of *Runx3* has also been shown to induce colon adenoma at a frequency similar to $Apc^{Min/+}$ mice (Ito et al. 2008), data again not reproducible by other laboratories. *Runx3* has also been shown to be required in the lung during bronchiolar epithelial cell differentiation and *Runx*3 deletion leads to lung adenoma development in aging mice (Lee et al. 2010). One recent study has analysed the role of *Runx3* in neuroblastoma: high levels of *RUNX*3 expression contribute to the favourable outcome in patients with neuroblastoma. Furthermore *Runx3* tumour suppressor role in this malignancy could depend on its direct repression of MYCN oncogenic activity (Yu et al. 2013). All the evidence presented so far suggests a causal link between loss of *Runx3* and cancer development. However, compared to *RUNX*1 and *RUNX*2, point mutations in the *RUNX*3 gene are rarely identified. This phenomenon could be explained by the finding that *RUNX*3 inactivation in cancer seems to occur mainly through aberrations in DNA methylation or histone modification (Lee 2011). Importantly, multiple studies have shown that

hypermethylation and subsequent silencing of *RUNX*3 gene expression is prevalent in solid tumours of breast, colon, lung, bladder and gastric origins (Chuang & Ito 2010). These results suggest a strong tumour suppressive role for *RUNX*3. However evidence indicates that *RUNX*3 can also act as an oncogene (Blyth et al. 2005). In lymphomas *Runx3* is targeted by viral insertions, which drive protein overexpression (Wotton et al. 2002). Moreover *Runx3* is expressed in a considerable proportion of pancreatic tumours suggesting that *Runx3* might play a role in the pathogenesis of pancreatic ductal adenocarcinoma (Li et al. 2004). *RUNX*3 is also overexpressed in skin (Lee et al. 2011), head and neck (Kudo et al. 2011) and ovarian cancers (Nevadunsky et al. 2009) again supporting an oncogenic role for RUNX3 in those tumours. The evidences linking RUNX3 to breast cancer will be discussed in detail later in the thesis.

1.4.3.4 RUNX genes in tumours: mechanism and pathways.

Runx genes can regulate cell growth and proliferation through a tight control on the cell cycle: not surprisingly the effects of *Runx* genes on the cell cycle are highly context dependent and dose sensitive (Coffman 2009). RUNX protein levels oscillate throughout the cell cycle indicating that variations in RUNX activity influence cell cycle progression. Indeed RUNX1 has been found to stimulate G1 to S progression in haematopoietic cells (Strom et al. 2000) while overexpression of RUNX2 in MC3T3-E1 osteoblastic cells delays progression from G1 to S phase (Galindo et al. 2005). Another way RUNX proteins regulate the cell cycle is through the control of cell cycle regulators: for example RUNX1 represses *p21* transcription, inducing cell cycle progression, in hair follicle stem cells *in vivo* (Lee et al. 2013). On the other hand RUNX activity is modulated by cell cycle regulators such as cyclins and cyclin-dependent kinases (Cdks): for example RUNX protein is ubiquitinated and degraded following phosphorylation by Cdk–cyclin complexes such as Cdk4/cyclin D1 and Cdk1/cyclin B (Biggs et al. 2006). Furthermore Cdk/cyclin complexes can affect RUNX function through other mechanisms such as disruption of RUNX-DNA interaction, displacing of co-receptors and inhibition of RUNX transactivation ability (Chuang et al. 2012). In addition to their role on cell cycle, RUNX proteins can control cell growth through expression of genes required for protein synthesis, such as rRNA genes (Young et al. 2007). Beyond regulation of cell growth and proliferation, RUNX can influence tumour growth through control on apoptosis and senescence. In

particular RUNX proteins can either activate apoptosis through induction of pro-apoptotic proteins such as BAX and BIM (Eliseev et al. 2008) or inhibit MYC-induced or p53-dependent cell death (Blyth et al. 2006; Ozaki et al. 2013). In addition *Runx* genes play a key role in oncogene-induced senescence, as shown by *in vitro* experiments on primary mouse embryonic fibroblasts (MEFs) where, after aberrant *Ras* expression, all RUNX proteins are necessary to induce a senescence-like growth arrest (Kilbey et al. 2008). *RUNX* genes are also playing a role in the regulation of genes that are intimately associated with tumour progression, invasion and metastasis including osteopontin, bone sialoprotein and matrix metalloproteinase. Again the role of *Runx* is highly isoform- and context- dependent: RUNX2 has been shown to promote invasiveness of breast and prostate cancer cells (Pratap et al. 2005), whereas RUNX3 expression inhibits metastasis of colon cancer cells (Peng et al. 2008) and decreases cell migration and invasion abilities of renal cell carcinoma cells (Chen et al. 2013).

Runx genes are intertwined with major intercellular signalling pathways associated with animal development, each of which is also known to be hijacked during neoplastic transformation.

TGFb signalling

Transforming growth factor-ß (TGFß) signalling is a key pathway involved in various developmental programmes where it regulates multiple processes such as differentiation, proliferation, apoptosis and adhesion (Massagué 2012). Interestingly like *Runx*, the TGF-ß pathway can play opposite roles as it serves to both inhibit the growth of many normal cells and induce highly malignant tumour processes such as EMT. Multiple lines of evidence support cooperation between *Runx* and the two major branches of TGF-ß superfamily: TGF-ß and bone morphogenetic proteins (BMPs) (Ito & Miyazono 2003). All RUNX members interact with the SMAD transcription factors which are key effectors of the TGF-ß pathway. RUNX proteins physically interact with Smad2/3 acting on the TGF-ß signalling pathway and cooperatively stimulating the synthesis of IgA (Pardali et al. 2000). In addition, the ability of RUNX3 to augment the TGF-ß pathway is essential for its tumour suppressor role in gastric cancer since gastric hyperplasia in *Runx3* knockout mice occurs, in part, through defective TGF-ß -mediated apoptosis (Li et al. 2002). In another example, RUNX2 interacts with

SMAD5 to perform critical roles in BMP-induced bone formation (Lee et al. 2000). RUNX2 can also regulate the expression TGF-ß type I receptor (Ji et al. 1998).

Wnt Signalling

Wnt signalling is a key regulator of stem cell differentiation and maintenance in different systems (Cadigan & Peifer 2009). Interestingly several evidences, coming from different experimental models, are linking RUNX and *Wnt* signalling. In particular, in sea urchin embryos, *Runt-1* expression promotes the expression of several *Wnt* genes (Robertson et al. 2008). In *C.Elegans*, *Rnt-1* collaborates with *Wnt* signalling to regulate asymmetric divisions in the T-blast stem-cell lineage (Kagoshima et al. 2007). In osteoblasts and chondrocytes *Wnt* signalling induces differentiation and activates RUNX2 where *Wnt* induces chondrocyte hypertrophy through RUNX2 upregulation (Dong et al. 2006) while during osteogenesis, *RUNX*2 is a direct target of ß-catenin/TCF1 for the stimulation of bone formation (Gaur et al. 2005). Finally in the skin epithelium, RUNX1 is an activator of *Wnt* signalling since *in vivo* loss of RUNX1 results in a generalized decrease in LEF1 protein and in canonical *Wnt* signalling (Osorio et al. 2011).

Notch signalling

The *Notch* signalling pathway plays a key role in modulating cell fate decisions throughout the development of invertebrate and vertebrate species (Fortini 2009). Studies in Zebrafish showed that *Notch* signalling activates RUNX1 expression, which in turn drives haematopoietic stem cell and progenitor expansion (Burns et al. 2005). In *Drosophila* haemocytes (blood cells), where *Notch* promotes crystal cell differentiation, the RUNX protein Lozenge (Lz) directs *Notch* to activate a combination of target genes which drives cells into the differentiation programme (Terriente-Felix et al. 2013). In bone development, *Notch* signalling in bone marrow acts to maintain a pool of mesenchymal progenitors by suppressing osteoblast differentiation through HES or HEY proteins, which diminish RUNX2 transcriptional activity via physical interaction (Hilton et al. 2008).

The Hippo/MST2 pathway

The Hippo/MST2 signalling pathway is a potent regulator of organ growth, and its deregulation leads to tumourigenesis (Zhao et al. 2010). Several studies have uncovered a crosstalk between RUNX and components of the Hippo pathway. YAP1 (Yes-associated protein), the main downstream target of the mammalian Hippo pathway, interacts with the PY motif of RUNX proteins via its WW domain (Yagi et al. 1999). Following this interaction, RUNX proteins recruit YAP1 to RUNX target promoters such as Osteocalcin, to stimulate transcription (Cui et al. 2003). In addition, RUNX3 forms a complex with two other components of the Hippo/MST2 pathway, scaffold protein SAV1 and tumour suppressor LATS2 kinase and it is necessary for Hippo-mediated cell death (Min et al. 2012).

1.5 *RUNX* genes in mammary development

The first evidence of *Runx* expression in mammary tissue came from a *Runx2* knock-out mouse where expression of LacZ from the targeted allele was detected in the embryonic mammary epithelium (Otto et al. 1997). Later work showed that RUNX2 is expressed in normal mammary epithelial, and some breast cancer, cell lines (Selvamurugan & Partridge 2000; Inman & Shore 2003). Moreover RUNX2 can directly regulate several mammary specific genes such as Osteopontin (Inman & Shore 2003) and Beta-casein (Inman et al. 2005). All three *Runx* genes were shown to be expressed in the adult mouse mammary epithelium with expression fluctuating during different stages of mammary development (Blyth et al. 2010). At the transcript level, *Runx1* is the most abundant gene and all three *Runx* transcripts follow a similar pattern of expression in whole gland extracts; decreasing during late pregnancy and lactation and rising again during involution. These stage-specific fluctuations in transcript levels suggest that RUNX proteins undergo a specific spatial and temporal regulation in mammary epithelium, underlining a possible new role for these transcription factors in the context of mammary development. The expression of *RUNX* genes has also been investigated in MCF10A cells, a human mammary epithelial cell line derived from a basal-like fibrocystic disease; *RUNX*1 mRNA is 15 fold higher than *RUNX*2 while *RUNX*3 is undetectable (Wang et al. 2011). Interestingly, RUNX2 overexpression in the MCF10A cell line cultured in a 3D culture model disrupts normal acini development resulting in increased cell proliferation and hyperplasia (Pratap et

al. 2009). On the other hand, confirming the exquisite cell context dependence of *RUNX* genes, a different study showed that loss of *RUNX*1 in MCF10A resulted in hyperplastic acini (Wang et al. 2011). These results suggest key independent roles for both RUNX1 and RUNX2 in regulating normal mammary epithelium growth and homeostasis, however direct *in vivo* genetic evidence, through conditional knock-out or knock-in transgenic models is still lacking. A study from our lab is the first to address this need (McDonald et al, submitted). Using a transgenic mouse model in which *Runx2* is placed under the transcriptional control of an MMTV promoter (MMTV-*Runx2* mouse) we have shown that ectopic *Runx*2 expression leads to a defect in alveolar development and lack of milk production in post-parturient transgenic mice. Downstream analysis showed that *Runx*2 overexpression could block lactation through its negative effects on the prolactin signalling pathway as shown by downregulation of prolactin receptor and decreased pSTAT5 activation in MMTV-*Runx*2 lactating glands. These results suggest that RUNX2 is involved in the regulation of mammary alveolar progenitor differentiation *in vivo*. Moreover MMTV-*Runx*2 transgenic glands are characterized by delayed ductal elongation during virgin development and a reduction in tertiary side-branching in the mature gland. These results represent the first evidence of an *in vivo* role for a member of the RUNX transcription family in mammary tissue. A natural extension to these studies using conditional knock-out models has been addressed in this thesis.

1.6 *RUNX* genes in breast cancer

A growing body of evidence has linked *RUNX* genes with breast cancer (Ferrari et al. 2013; Chimge & Frenkel 2012). The most studied *RUNX* gene in the context of breast cancer is *RUNX*2 while recent studies are also indicating a possible role for *RUNX*1 and *RUNX*3. The most compelling evidence linking *RUNX*1 to breast cancer comes from two recent sequencing studies which identified *RUNX1* somatic mutations in a small number of human breast cancer cases (Ellis et al. 2012; Banerji et al. 2012). In addition *RUNX*1 was identified as a downregulated gene in a 17-gene signature associated with metastasis of adenocarcinoma nodules of diverse origin, including breast cancer (Ramaswamy et al. 2003). Decreased expression of *RUNX*1 was also detected in BC compared to normal breast epithelial cells, with a more pronounced decrease with increasing tumour invasiveness (Kadota et al. 2010). Finally, *RUNX*1 silencing in MCF10A cells led

to the formation of hyperplastic acinar structures in 3D cultures (Wang et al. 2011). There is however also studies suggesting a possible oncogenic role for *RUNX*1 in breast cancer (Janes 2011) which is discussed in greater detail in Chapter 6.

In human breast cancers expression of *RUNX*3 decreases during human BC progression and reduced *RUNX*3 expression predicts worse survival, suggesting a putative tumour suppressive role for *RUNX*3 in BC (Jiang et al. 2008; Bai et al. 2013). Moreover *in vivo* studies showed that 20% of *Runx*3 heterozygous knockout mice develop spontaneous ductal carcinoma at an average age of 14.5 months (Huang et al. 2012). However, as work from our laboratory is unable to detect *Runx*3 gene expression in purified mammary epithelial cells from normal mice (see Chapter 3) we question the role of *Runx*3 as a tumour suppressor in this cell lineage. As the study by Huang and colleagues use a constitutive knockout model it is possible that reduced gene dosage in the non-autonomous tissue (e.g. stroma) may promote tumourigenesis. This discordance is somewhat mirroring the unsolved controversy about the possible tumour suppressive role of *Runx*3 in gastric and colorectal cancers (Ito et al. 2008; Li et al. 2002; Levanon et al. 2011). Interestingly all these *in vivo* cancer studies are based on the same *Runx*3 knock-out mouse (Li et al. 2002) and these results have not been confirmed by other laboratories (Levanon et al. 2003). Further studies are needed to clarify this discrepancy.

While RUNX1 and RUNX3 have a supposedly tumour suppressor role in the mammary epithelium, evidence indicates that RUNX2 could have oncogenic properties in breast cancer. RUNX2 is upregulated in some human breast cancer cell lines compared to normal mammary epithelial cells (Lau et al. 2006; Inman & Shore 2003). In addition *RUNX*2 was identified as one of the most upregulated genes in a comparative transcriptomic analysis of invasive versus non-invasive breast cancer cell lines (Nagaraja et al. 2006). Since the metastatic cell line MDA-MB-231 expresses high levels of RUNX2, much of the research published to date to understand the role of RUNX2 in breast cancer, focus on this *in vitro* model. RUNX2 inhibition in these cells results in a less invasive phenotype *in vitro* together with reduced osteolytic disease *in vivo* (Javed et al. 2005). Conversely RUNX2-negative MCF-7 cells become more invasive after RUNX2 overexpression (Pratap et al. 2006; Leong et al. 2010). Inhibition of RUNX2 in a

3D culture model reverted the disorganized structures formed by MDA-MB-231 cells to a more normal phenotype resulting in formation of acinar-like structures (Pratap et al. 2009). RUNX2 deletion was also shown to inhibit tumour growth in mammary fat pad transplantations of MDA-MB-231 cells (Pratap et al. 2009). In line with these observations, RUNX2 regulates genes that are closely associated with tumour invasiveness, metastasis and angiogenesis, such as BSP, OPN, MMPs and VEGF (Pratap et al. 2005; Inman & Shore 2003; Javed et al. 2005; Barnes et al. 2003). Interestingly many of these RUNX2-dependent genes, seen to be upregulated in breast cancer cell lines, are also the same genes which are required for bone development and turnover (MMPs, RANKL, OPN, BSP) (Pratap et al. 2006). This observation has led to the formulation of the 'osteomimicry' theory in which RUNX2-mediated expression of bone-specific genes confers osteoblast-like features to breast cancer cells allowing them to home and thrive in the bone microenvironment, leading to bone metastasis formation (Barnes et al. 2003). However the *in vivo* proof of this theory is still lacking and only a few *in vitro* studies support this hypothesis. During bone metastasis, breast cancer cells must induce osteoclast activity and block osteoblast differentiation in order to create a vital space in which to grow (Suva et al. 2011). Interestingly the ability of two breast cancer cell lines to inhibit osteoblast differentiation and enhance osteoclast differentiation is abrogated upon deletion of RUNX2 in *in vitro* co-culture assays (Barnes et al. 2004). Moreover, in the MDA-MB-231 model, the RUNX2/CBF-ß complex was shown to inhibit osteoblast differentiation through the induction of Sclerostin, a known inhibitor of bone formation (Mendoza-Villanueva et al. 2011). Taken together these data are indicating RUNX2 as a key regulator of the invasive and metastatic potential of a few breast cancer cell lines *in vitro.* Recent publications have also linked RUNX2 in primary breast cancer. In particular, our work and that of others has found expression of RUNX2 restricted to specific subsets of breast cancer. In a study from Das and co-workers, nuclear RUNX2 expression was detected in 28% of breast tumours and this expression was significantly associated with ER/PR positive, Grade 2 tumours (Das et al. 2009). However, this study is in contrast with other publications (McDonald et al. submitted; Khalid et al. 2008) which identified a positive association between RUNX2 and ER-negative disease. Meta-analysis of gene expression in 779 breast cancer biopsies showed a negative correlation between the expression of ER and RUNX2 target genes (Khalid et al.

2008). In a small TMA study, high RUNX2 expression was associated with worse clinical outcome, and this correlation was more marked in ER negative patients (Onodera et al. 2010), while unpublished data from our laboratory confirm the correlation between RUNX2 expression and ER negative patients in human breast cancer. Assessment of RUNX2 expression in a large tissue microarray constituted by 459 primary human breast cancers showed that high RUNX2 expression correlates with poor overall patient survival and ER negative tumours (McDonald et al. submitted). Supporting evidence of a role for RUNX2 in ER negative disease comes from *in vitro* analysis where high RUNX2 expression is found specifically in "basal-like" (ER-negative) cell lines, and not in those of a 'luminal-like' subtype (Lau et al. 2006). Confirming the negative correlation between RUNX2 and oestrogen, recent *in vitro* experiments showed that RUNX2 expression can decrease oestrogen-driven colony formation of breast cancer cells in soft agar colony assays (Chimge et al. 2012) while estradiol treatment antagonizes RUNX2-induced EMT and invasiveness *in vitro*, in part through decreased expression of the EMT-inducer SNAI2 (Chimge et al. 2011).

To conclude, the picture of *RUNX* genes in breast cancer is still far from complete and their dualistic role as oncogenes and tumour suppressors (Blyth et al, 2005), highlight their extreme context dependency. It is likely that each of the Runx proteins will play different, and even opposing, roles in different subtypes of breast cancers. Better characterization of *RUNX* function in small subgroups of breast cancer will hopefully translate into more targeted therapies which could greatly benefit subsets of patients.

Aims of the thesis

The mammary gland is a complex secretory organ that undergoes dramatic changes during its entire development. In order to maintain such complexity, mammary gland development is a tightly controlled process that involves a wide variety of molecular regulators. A better understanding of the regulatory system involved in mammary development is needed to improve our knowledge of breast cancer initiation and progression. This could translate in the identification of new therapeutic targets for the treatment of this deadly disease. Accumulating evidence suggests a possible role for RUNX transcription factors in mammary gland development and in breast cancer. However this evidence is mainly coming from *in vitro* studies and no *in vivo* study has been carried out so far.

The aim of the first part of this thesis was to determine the physiological role of *Runx* genes in normal mammary gland development through understanding the expression of *Runx* genes in the mammary epithelium and utilising an *in vivo* conditional deletion strategy to create *Runx*-specific knock-out in the mouse gland.

Besides being key developmental regulators, *Runx* genes have recently emerged as major players in stem cell biology, coordinating cell-fate decisions and lineage differentiation. It was intriguing then that results achieved in the first part of this thesis suggested *Runx* expression may correlate with mammary stem cells. Therefore the involvement of RUNX proteins in mammary stem cell physiology was investigated in Chapter 4 through a combination of *in vitro* (mammospheres, matrigel colony-forming assay) and in vivo assays.

Finally, considering that genes involved in normal tissue development often take part in cancer initiation and progression, the role of RUNX1 and RUNX2 in breast cancer was characterized through a combination of *in vitro* and *in vivo* modelling as discussed in Chapter 5 and Chapter 6.

2 Materials & Methods

2.1 Animals

All animal work was carried out under UK Home Office guidelines in line with Animal (Scientific Procedures) Act 1986 and the EU Directive 2010. Experimental cohorts and breeding stocks were maintained for defined periods of time and health of animals was checked at least two times weekly. Animals were euthanized by cervical dislocation or by carbon dioxide (CO_2) asphyxiation. Mouse ear notching and general maintenance (food, water and housing) was carried out by the Biological Services Unit at the Beatson Institute. Genotyping was carried out by Transnetyx, Inc. (Cordova, TN, US).

2.1.1 Characterization of the *Runx2*$^{flx/flx}$ mouse model

The *Runx2*$^{flx/flx}$ mouse was made in the lab of Professor Mike Owen by Theresa Higgins and Iain Rosewell, (ICRF labs, London) and characterised here. To generate *Runx2*$^{flx/flx}$ mice one *loxP* site was inserted in the intron 5' of exon 3 and a *ploxPneo* cassette in the intron 3' of exon 3 of the *Runx2* gene. Exon 3 was chosen because that was the one encoding for the RUNX2 DNA binding domain (FIGURE 2-1-A). A gPCR on DNA extracted from the tail of *Runx2*$^{flx/flx}$ mice was carried out to amplify each *loxP* site using the following primers: *loxP1* (F-GACCTCCTCCCTACAGCTTCG) and R-CCCTCGCGTTTCAAGGTGCCG), *loxP2* (F-CAGCTGACAGCAGGTTGAAA; R-TGTCTGTTGTGCCCAGTCAT) and *loxP3* (F-TCGCTAACTTGTGGCTGTTG; R-TGGATGCTCTCAAAAAGGAAA). The PCR products were purified from contaminants (primers, nucleotides, enzymes, salts and other impurities from DNA samples) with QIAquick PCR Purification Kit (Qiagen). Then the location and the integrity of the *loxP* sequences were confirmed through DNA sequencing of the PCR products (FIGURE 2-1-B). To confirm the recombination capability of the *Runx2*$^{flx/flx}$ mouse model, mouse embryonic fibroblasts (MEFs) from *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ embryos, were retrovirally transfected with a Cre-expressing plasmid (pBABE-*Puro-Cre*) or a control empty-

vector (pBABE-*Puro*). After 5 days of selection in Puromycin (2µg/ml) transfected MEFs were harvested and *Runx2* expression was assessed both at the RNA and at the protein level (FIGURE 2-2).

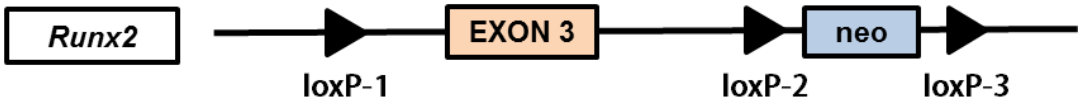
2.1.2 Other mouse strains

Runx1$^{flx/flx}$ mice (Growney et al. 2005), BLG-Cre mice (Selbert et al. 1998),and Z/EG reporter mice (Novak et al, 2000) have all been described previously. K14-Cre mice (Dassule et al. 2000) were obtained from The Jackson Laboratory.

MMTV-*PyMT* (Guy et al. 1992) and MMTV-*Her2* (Muller et al. 1988) tumour samples were obtained from mouse cohorts bred in our facilities. BLG-Cre/*Brca1*$^{flx/flx}$/*p53*$^{+/-}$ mouse tumour samples were kindly provided by Prof. Matt Smalley (European Cancer Stem Cell Research Institute, Cardiff). *Apc*^{1572T} mouse tumour samples were kindly provided by Prof. Riccardo Fodde (Erasmus Medical Centre, Rotterdam). BLG-Cre/*Pten*$^{flx/flx}$ *Apc*$^{flx/flx}$ mouse tumour samples and the *Catnb*$^{+/lox(ex3)}$ mouse model (Harada et al. 1999) were kindly provided by Prof. Owen Sansom (Beatson Institute for Cancer Research, Glasgow).

FVB wild type, SCID and CD1-nude mice were obtained from Charles River Laboratories (UK).

A



B

LoxP Sequence : ATAACTTCGTATAATGTATGCTATACGAAGTTATC (34bp)

LoxP-1

CNAGCGGCCGGCGGGCGGCGGCAGCGGCTGCGATCCGCAGACTCCAGATNNNNNNNTAA
CTTCGTATAATGTATGCTATACGAAGTTATCCCGATNNNTNNNCCCGAGATCCGCTCTCCC
CCCGCCCCCACTTACCTCCCGGCACCTTGAAACGCN

LoxP-2

CTGCTCGACATTGGGTGGAAACATTCCAGGCCTGGGTGGAGAGGCTTTTTGCTTCCTCTTG
AAAACCACACTGCTCGACCTGCAGCCAAGCTAGCTTGGCTGGACGTAAACTCCTCTTCANA
CCTAATAACTTCGTATAGCATACATTATACGAAGTTATATTAAGGGTTATTGAATATGATCG
GAATTCCTCGAGTCTAGAGTCGAGCC

LoxP-3

GGCAATAAAAAGACAGAATAAAACGCACGGGTGTTGGGTCGTTTGTTCGGATCTCGAGGAA
TTCCGATCATATTCAATAACCCTTAATATAACTTCGTATAATGTATGCTATACGAAGTTATTA
GGTCTGAAGAGGAGTTTACGTCCAGCCAAGCTAGCTTGGCTGCAGGTCGAGGTACCGAGC
TCGAATTCATCGATGG

Figure 2-1 Characterization of the *Runx2*$^{flx/flx}$ mouse model.

Schematic representation of the Runx2$^{flx/flx}$ mouse (exon3 genomic locus), with the relative position of *loxP* sequences and Neomycin cassette (neo) (A). Sequences of the three *loxP* sites as obtained from DNA sequencing (highlighted in yellow). The expected *loxP* sequence of 34bp is shown above (B).

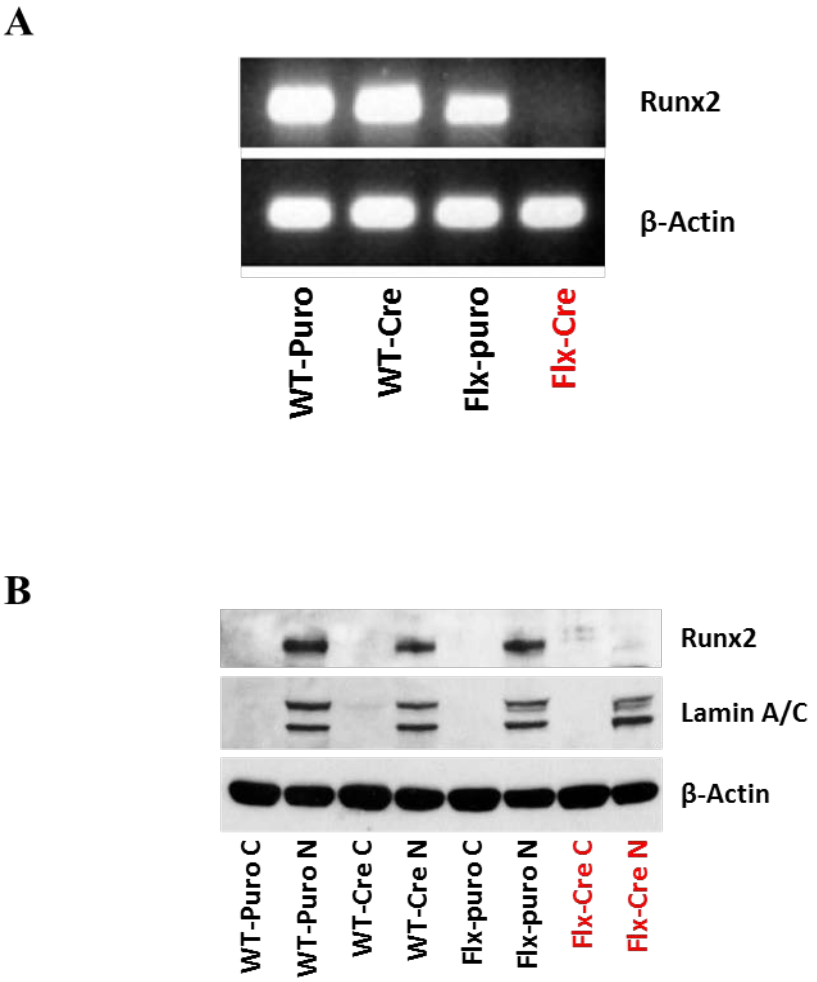


Figure 2-2 *Runx2* deletion after *in vitro* Cre-recombination.

RT-PCR on MEFs extracted from *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ embryos and retrovirally transfected with a Cre-expressing plasmid (WT-Cre, Flx-Cre) or a control empty vector (WT-Puro, Flx-Puro). ß-Actin was used as loading control (A). Western blot on cytoplasmic (C) and nuclear (N) extracts extracted from MEFs derived from *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ embryos; MEFs were retrovirally transfected with a Cre-expressing plasmid (WT-Cre, Flx-Cre) or a control empty vector (WT-Puro, Flx-Puro). Lamin A/C was used as a nuclear control. ß-Actin was used as a total loading control (B).

2.2 Wholemount/histological analysis of mammary glands

For wholemount analysis, inguinal mammary glands were dissected, air dried onto a glass slide and fixed in Carnoy's (300ml 70% EtOH, 150ml chloroform, 50 ml acetic acid). Glands were rehydrated through ethanol and stained with carmine alum overnight. Glands were dehydrated in increasing concentration of EtOH, cleared in xylene, mounted in Permount (Thermo Fisher) and captured with a Zeiss stereomicroscope. For pregnancy samples the state of pregnancy was determined by checking vaginal plugs (as carried out by BSU staff). For involution studies a 7 day forced involution strategy was used where pups were standardised to 6, culled 7 d after parturition to initiate involution and the dam was taken at the selected time points. For histological analysis, adult mammary glands were dissected into 10% neutral buffered formalin and processed for haematoxylin and eosin (H&E) staining. For histological analysis of embryonic mammary glands, the age of embryos were determined by checking vaginal plugs. Whole embryos were taken at the different time points and fixed into neutral buffered formalin for at least 1 week. Embryos were then cut transversally at the level of the 4^{th} and 3^{rd} mammary gland, embedded in paraffin blocks and serial sections were taken to identify the embryonic mammary duct tree. All histopathological processing was carried out by Mr. Colin Nixon and staff.

2.3 *In vivo* imaging

In vivo GFP imaging was carried out on freshly dissected mammary tissue. Briefly, gland #4 was dissected and spread onto a glass microscope slide. The gland was then imaged using OV-100 Imaging System (Olympus).

CD-1 nude mice were injected with 10^6 cells/mouse in the tail vein (procedure carried out by Derek Miller). Metastatic growth was monitored weekly by *in vivo* luciferase imaging. Mice were anaesthetized with isofluorane and subcutaneously injected with 150 mg/kg of D-luciferin (PerkinElmer, US) in PBS. Bioluminescence images were acquired 5 min after injections, to allow the distribution of D-luciferin in the entire body of the animal, by using the IVIS Spectrum Imaging System (PerkinElmer, US).

2.4 Fat pad transplantation

Single cell suspensions of freshly extracted mammary epithelial cells (see below) were re-suspended into a solution of PBS/25% Matrigel™ Matrix Phenol Red-Free (BD Biosciences, CAT 356237) at the desired concentration to allow a final injection volume of 10µl. Cells were injected using a 10µl Hamilton syringe (Hamilton, CH) into the inguinal fat pads of 3-week-old SCID females cleared of endogenous epithelium by surgical intervention. At the point of clearing, fat pads were mounted on a slide and stained for wholemount analysis to confirm the clearing of endogenous epithelium. The outgrowths were analysed 7 weeks after transplantation by GFP imaging (Leica M205 FA) and wholemount analysis (Zeiss Semi 2000-C).

2.5 Cell lines

Mouse embryonic fibroblasts from *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ embryos (previously extracted by Dr. Karen Blyth) were expanded in culture using DMEM (Gibco), 10% Fetal Calf Serum (FCS), 1% Pen/strep (Gibco) and 1% L-Glutamine (Gibco). 3SS (mouse leukaemia cell line) nuclear extracts (kindly provided by Prof. Ewan Cameron) were used as a negative control in RUNX1 western blots. MDA-MB-231-luc-D3H2LN cells (kind gift of Dr Dan Croft) were grown in Hyclone MEM/EBSS media (Thermoscientific), 10% FCS; 1% MEM/NEAA (non-essential amino acid); 1% of Sodium Pyruvate; 1% Pen/strep and 1% L-Glutamine. MDA-MB-468, HCC-70, BT-549, T47D, MDA-MB-361 and BT-474 nuclear extracts were generated by Susan Mason. hMEC-TERT cell line (a kind gift of Barbara Chaneton) was grown in HuMEC complete media (Gibco). All cell lines were grown in a Galaxy+ incubator (RS Biotech) at 37^{o}C with 5% CO_2. Cells were dissociated enzymatically and passaged using a solution of 0.05% trypsin (Gibco).

2.5.1 Cell line transfections

To generate RUNX2 stable knock-out, MDA-MB-231-luc-D3H2LN cells were transfected with 4 different *RUNX2* Sh-RNAs and 1 scrambled control (HuSH^{TM}, Origene) through electroporation using Nucelofector Kit V, program X-013 (Amaxa, Lonza). After electroporation, cells were allowed to recover for 24h

and then selected in puromycin selection media (1µg/ml) for 2 weeks. Selected cells were then FACS-sorted for GFP expression for at least three times to achieve maximum population purity.

To generate RUNX1 overexpressing cells, hMEC-TERT were transfected with p-BABE-*Puro-Runx1* or p-BABE-*Puro* (kindly provided by Anna Kilbey) through electroporation using Nucelofector Kit V, program T-013 (Amaxa, Lonza). After electroporation, cells were allowed to recover for 24h and then selected in puromycin selection media (10µg/ml) for 2 weeks.

2.5.2 Cell line assays

2.5.2.1 2D Growth curve

Cells were seeded in triplicate in 12 well plates ($5x10^4$ cells/well) and harvested at 24, 48, 72 and 96 hours. Cells were trypsinised and counted using the Trypan Blue (Life Technologies) exclusion method on haemocytometer.

2.5.2.2 Tumourspheres

Cells were dissociated enzymatically (Trypsin) and mechanically by pipetting to single-cell suspension and plated on nonadherent plates (Corning) at a concentration of 1000 cells/ml. Cells were grown in a serum-free Hyclone MEM/EBSS media (Minimal Essential Medium with Earle's, Thermoscientific) with 1% Pen/strep and 1% L-Glut (Gibco) supplemented with B27® (Gibco), 20 ng/ml EGF (Sigma), 20 ng/ml b-FGF (Sigma), BSA 0.4% and 4 µg/ml Heparin (Sigma). Tumourspheres were grown for 7 days and colonies were counted under a bright field microscope.

2.6 Primary mouse mammary cells

Mouse mammary epithelial cells (MMECs) were extracted from mammary glands of adult virgin mice (more than 12 weeks of age) unless otherwise specified. At least 2 mice x group were used for virgin analysis, 1 per group for other developmental stages. Glands #4 were dissected and the lymph node excised to avoid contamination with non-mammary cells; glands #2-3 were dissected avoiding the surrounding muscular tissue. Dissected glands were placed into 50ml falcon tubes with cold 15ml DMEM (0% FCS, + Pen/Strep) and kept on ice. Then glands were finely minced using a McIlwain tissue chopper (Mickle Laboratories, UK) at maximum speed and force until a liquid slurry with white fat on the top was formed. The minced tissue was then transferred into a 50ml falcon tube and digested for 90 mins in 15ml of Collagenase (300 U/ml, Sigma) plus Hyaluronidase (100 U/ml, Sigma) solution in DMEM 0% FCS at 37˚C under shaking. Tubes were then spun down at 300g for 5 min, pellets were resuspended in 5ml DMEM 0% FCS and transferred to a 15ml falcon tube. After an additional centrifugation (300g for 5 min), pellets were incubated for 5min at RT in 2ml NH_4Cl solution (0.8% in H_2O) to eliminate red blood cells. The reaction was stopped with 3ml of DMEM 0%FCS and cells were spun down (300g for 5 min). After eliminating the supernatant, 2ml of TEG (0.25% trypsin and 1mM EGTA in PBS) plus 10% DNAse (1mg/ml in PBS, Worthington) were added to the pellet followed by serial pipetting to allow dissociation. The digestion solution was then incubated for 10 min at 37˚C in a water bath. After incubation all remaining DNA clumps were dissolved by pipetting or by additional DNAse treatment. The digestion was stopped by addition of DMEM 10% FCS followed by filtering with a 70µm filter to eliminate remaining clumps of tissue. After an additional centrifugation step (250g for 5 min) pellets were resuspended in the required final media. The single cell suspension was checked and counted using the Trypan Blue exclusion method on haemocytometer.

2.6.1 Flow cytometry/cell sorting

For mammary population profiling virgin mammary glands were dissected (at least 3 groups of n≥3 for *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ cohorts); for pregnancy, lactation and involution stages (at last 3 groups of n=1 for *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ cohorts). Tissues were processed to obtain single cell suspensions of

mammary epithelial cells (see above). Cells were labelled with CD24-Phycoerythrin (PE) (1:200, BD Bioscience), CD29-PerCP-eFluor® 710 (1:100, eBioscience), CD31- Allophycocyanin (APC) (1:100, BD Bioscience) and CD45-APC (1:200, BD Bioscience) in preparation for flow cytometry (1h incubation on ice in the dark). DAPI (1:250, Sigma) was added to the samples at the end of antibody incubation, 5 minutes before analysis. For each day of FACS analysis, single labelled controls were prepared to confirm the gating strategy. Live Lin- cells (DAPI/CD31/CD45 negative) were gated on a BD FACSAria, doublet exclusion was sequentially performed for FSC (FSC-A vs FSC-H plot) and single cells were analysed to assess GFP+ populations in $CD24^{hi}CD29^{lo}$ (luminal) and $CD24^{hi}CD29^{hi}$ populations (basal/myoepithelial) using FlowJo software (FIGURE 2-3). For cell sorting, $CD24^{hi}CD29^{lo}$ and $CD24^{hi}CD29^{hi}$ populations were collected and processed for RNA extraction and quantitative RT-PCR. The efficacy of the sorting strategy was confirmed by RT-PCR for a basal (CK5) and a luminal (CK18) marker (FIGURE 2-4).

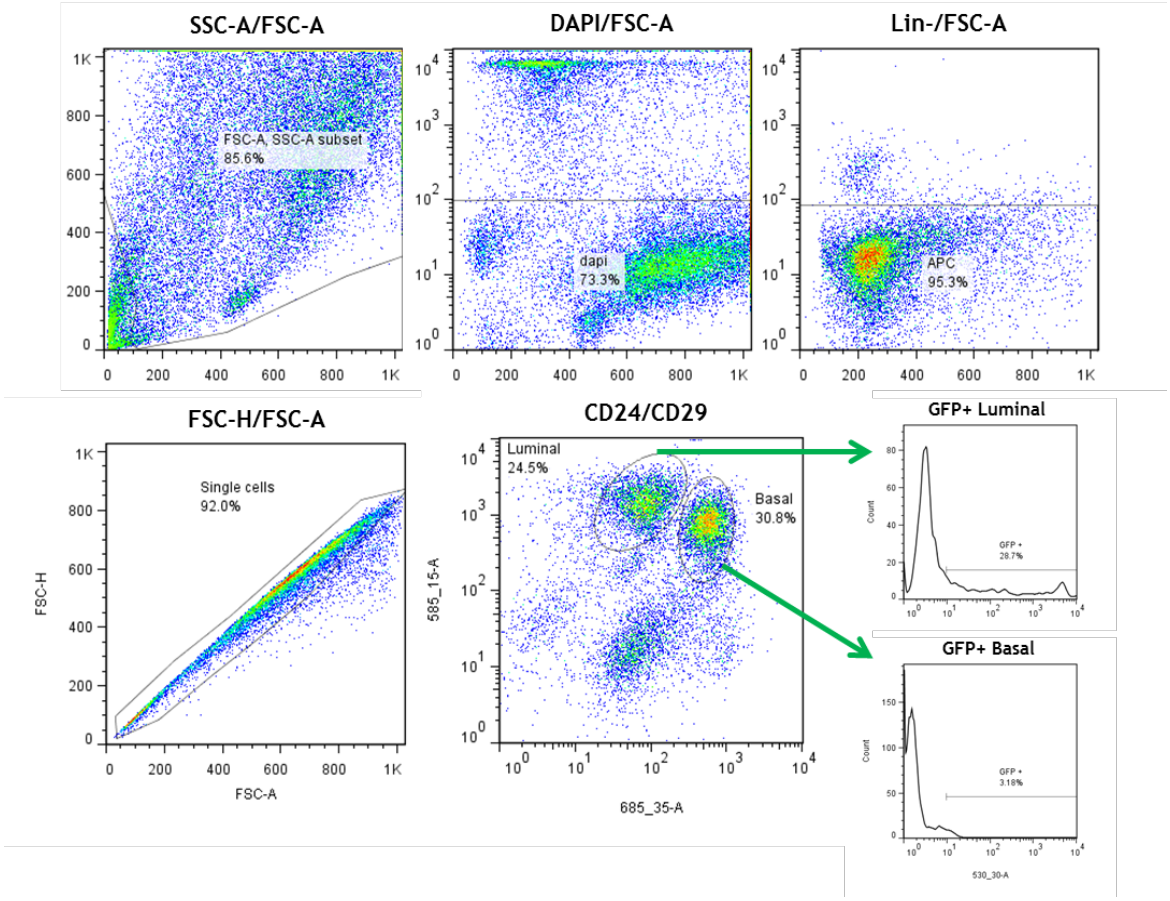


Figure 2-3 FACS gating strategy for mammary population profiling.

Cells were first gated based on forward and size scatter to exclude debris (SSC-A/FSC-A). Alive cells were selected based on DAPI staining (DAPI/FSC-A). Blood and endothelial cells were excluded based on CD31 and CD45 staining (Lin-/FSC-A). Doublets and triplets were excluded based on forward scatter height vs area (FSC-H/FSC-A). Luminal and basal population were resolved based on CD24 and CD29 markers (CD24/CD29). % of GFP+ cells was assessed in the luminal and in the basal population (GFP+ Luminal and GFP+ Basal).

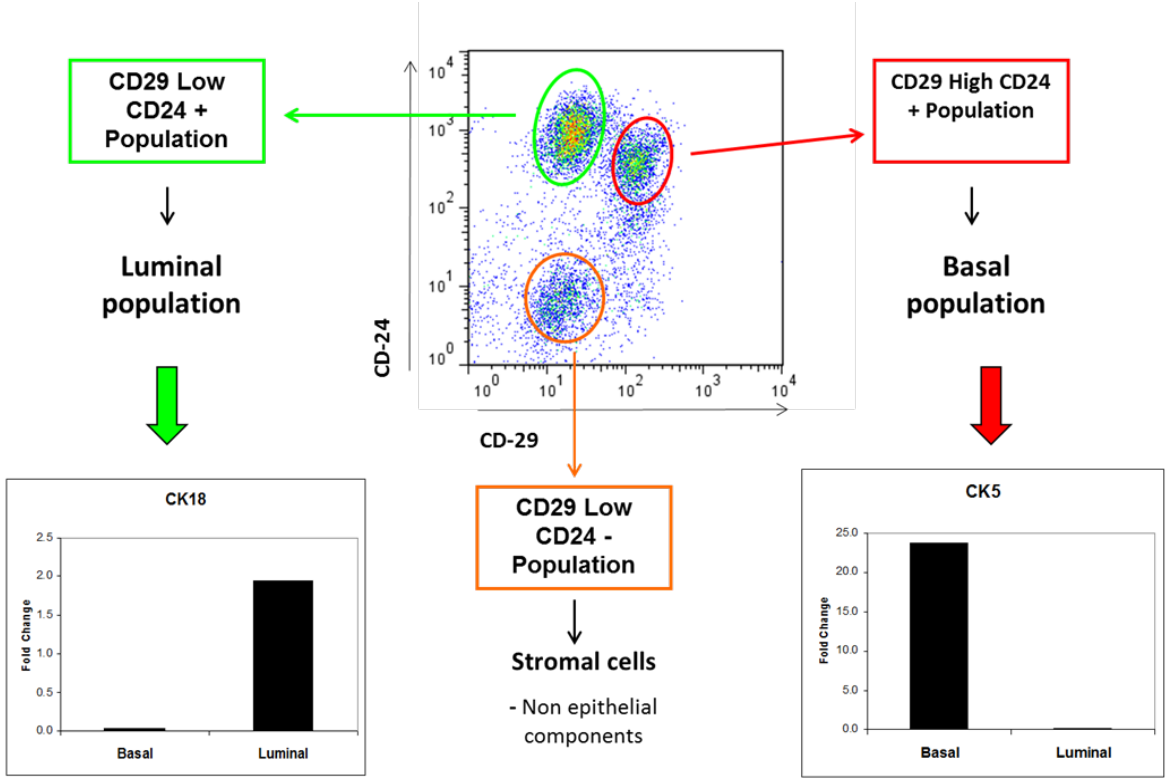


Figure 2-4 Validation of FACS sorting strategy.

Luminal and basal populations from 6 weeks virgin FVB mice were sorted based on CD24/CD29 expression. RNA was extracted from those populations and qRT-PCR was performed for a luminal (CK18) and a basal (CK5) marker. As expected CK18 was enriched in the luminal population, while CK5 was enriched in the basal one.

2.6.2 MMECs 2D cultures.

Single cells extracted from 12 week old virgin mice were plated onto 6-well plates ($3x10^5$ cells seeded per well) and cultured in DMEM/F12 10% FCS with Pen/strep and L-Glut, supplemented with 10 ng/ml EGF (Sigma), 5 µg/ml Insulin (Roche) and 10 ng/ml Cholera Toxin (Sigma). For quantitative RT-PCR analysis, 2D MMECs were dissociated enzymatically after 7 days and processed for RNA extraction.

2.6.3 Mammospheres

Single cells extracted from 12 week old virgin mice, were plated in ultra-low adherent 24-wells plates (Corning) at a density of 20,000 viable cells/ml for primary passage and 1000 cells/ml for second passage. Cells were grown in a serum-free medium DMEM/F12 with Pen/Strep and L-Glut (Gibco), supplemented with B27® (Gibco), 20 ng/ml EGF (Sigma), 20 ng/ml bFGF (Sigma), 0.4% BSA and 4 µg/ml Heparin (Sigma). Growth factors were re-added fresh to the culture every 3-4 days. Plates were left untouched in the incubator to avoid cell clumping due to excessive manipulation. After 7 days, mammospheres were counted under a bright field microscope, collected by gentle centrifugation (200 g for 5 minutes) and dissociated enzymatically (10 min in TEG at 37˚C in water bath) and mechanically, by pipetting. The cells obtained from dissociation were checked for single-cellularity, counted using the Trypan Blue exclusion method on haemocytometer and seeded again to generate secondary mammospheres. For quantitative RT-PCR analysis, mammospheres were processed for RNA extraction after 7 days. For histologic analysis, mammospheres were fixed on day 7 with 500 µl of 2% paraformaldehyde for 15 minutes. After centrifugation pellets were resuspendend in 150 µl of 3% UltrPure™ low-melting agarose (Invitrogen) and left for 20 minutes at room temperature to solidify. The agarose plug was then put into 70% EtOH and embedded in paraffin blocks. RUNX2 staining (1:500 45 min at RT, Sigma) was carried out on mammosphere blank sections using citrate buffer antigen retrieval and standard IHC protocols (see below). For size measurement, primary and secondary mammospheres colonies were photographed with a bright field microscope (Olympus CKX41) after 7 days in culture and colony size was assessed using Axiovision software (Zeiss).

For adenoviral infection of mammospheres, MMECs single cell suspensions from *Runx2*$^{flx/flx}$ and *Runx2*$^{WT/WT}$ 12 week old virgin mice, were spin-infected with Ad5CMVCre-eGFP (Iowa University) (3.30 hours at 300g, RT; MOI=100) in 200 µl DMEM/F12 with 2% FCS, 20 ng/ml EGF (Sigma) and 20 ng/ml bFGF (Sigma). After infection, cells were seeded at 20,000 cells/ml in ultra-low adherent 24-well plates (Corning). After 7 days mammospheres were counted under a bright field microscope and processed for RNA extraction and quantitative RT-PCR.

2.6.4 WNT3a treatments on mammospheres and 2D-MMECs

For long-term WNT3a treatment, primary mammospheres (from 12 week old FVB virgin mice) were grown for 1 week in mammosphere media supplemented with 50 ng/µl of recombinant WNT3a (Sigma, SRP3259) or vehicle (H_2O). After 7 days mammospheres were counted under a bright field microscope and processed for RNA extraction and quantitative RT-PCR. For size measurement, mammospheres colonies were photographed with a bright field microscope (Olympus CKX41) and colony size was assessed using Axiovision software (Zeiss) after 7 days in culture. For short-term WNT3a treatment, mammospheres and 2D-MMECs were grown for 3 days in normal conditions. At day 4, 100ng/µl of recombinant WNT3a or vehicle (H_2O) was added to the media and cells were processed for RNA extraction 24h after.

2.6.5 Matrigel Colony forming assay

Single cells extracted from 12 week old virgin mice were re-suspended in ice-cold Growth-factor-reduced Matrigel (BD Biosciences, CAT 356237). 8000 cells/well were seeded for primary Matrigel colony formation. 5000 cells/well were seeded for secondary Matrigel colony formation. A drop of 20 µl of Matrigel/cells was added in the centre of each well of a 24 well plate and the plate was returned to the incubator for 5 minutes to allow the Matrigel to solidify. 1 ml of DMEM/F12 with Pen/Strep and L-Glut supplemented with 20 ng/ml EGF (Sigma) was added to each well. Fresh EGF was re-added to the culture every 3-4 days. After 7 days, Matrigel colonies were counted under a bright field microscope and dissociated enzymatically (10 min in TEG at 37˚C in water bath) and mechanically, by pipetting. The cells obtained from dissociation were checked for single-cellularity, counted using the Trypan Blue exclusion

method on haemocytometer and seeded again to generate secondary Matrigel colonies. For quantitative RT-PCR analysis, Matrigel colonies were processed for RNA extraction after 7 days. For histologic analysis, Matrigel colonies were fixed with 500 µl of 2% paraformaldehyde for 15 minutes on day 7. After centrifugation pellets were resuspendend in 150 µl of 3% UltrPure™ low-melting agarose (Invitrogen) and left for 20 minutes at room temperature to solidify. The agarose plug was then put into 70% EtOH and embedded in paraffin blocks. For size measurement, primary and secondary Matrigel colonies were photographed after 7 days in culture with a bright field microscope (Olympus CKX41) and colony size was assessed using Axiovision software (Zeiss).

2.7 Immunohistochemistry

Paraffin-embedded tissue sections were rehydrated before antigen retrieval using pH6 sodium citrate buffer (or 1mM EDTA pH8 for RUNX2). After washing with Tris buffered saline and blocking endogenous peroxidase (Peroxidase blocking solution, DAKO), sections were incubated with the following HRP-conjugated primary antibodies: RUNX1 (Sigma HPA004176; 1/200 1h RT), RUNX2 (Sigma HPA022040; 1/100 overnight at 4°C), ß-Catenin (BD Bioscience, 610154; 1/1000 1h RT), GFP (Abcam ab6556; 1/2000 overnight at 4°C). Anti-rabbit secondary (Dako EnVision) was used for all antibodies except for ß-Catenin (anti-mouse; Dako EnVision). Sections were incubated with secondary antibodies for 1 hour at RT, treated with DAB (Dako) and counterstained with haematoxylin. Images were captured using a Zeiss AX10 or an Olympus BX51 microscope.

2.8 Immunofluorescence

Paraffin-embedded tissue sections were rehydrated before antigen retrieval using pH6 sodium citrate buffer and permeabilized with incubation in 1% Triton® X-100 (T8787, Sigma) in PBS for 10 minutes. Slides were then blocked with a solution of 0.5% Triton® X-100 and 5% Goat Serum (DAKO, X0907) in PBS for 1h at RT. Primary antibodies were diluted in Dako REAL™ Antibody Dilutent and

incubated on sections overnight at 4°C. Primary antibodies used were anti-cytokeratin 8/18 (Fitzgerald 20R-CP004; 1/400), anti-GFP (Abcam ab6556; 1/250), anti-CK14 (Abcam ab7800; 1/250). After washing with Tris buffered saline sections were incubated with Alexa 488 (goat-anti-rabbit-IgG), Alexa 594 (goat-anti-mouse-IgG) and Alexa 647 (goat-anti-guinea pig-IgG) secondary antibodies (1:250 in PBS, Invitrogen) for 1 hr at room temperature. Slides were mounted using Vectashield mounting media with DAPI (Vector, H-1200). Confocal images were captured using a Zeiss 710 confocal microscope.

2.9 Non quantitative PCR

2.9.1 PCR determination of *Runx2* status.

To test the presence of the recombined *Runx2* allele, genomic DNA was extracted using isopropanol precipitation. Samples were resuspended in 500 µl of chilled isopropanol to precipitate DNA. After centrifugation (16 rcf for 10 min at 4˚C) the supernatant was poured off and 500 µl of 70% EtOH added to wash the DNA pellet. After an additional centrifugation (16 rcf for 5 min at 4˚C), the supernatant was poured off and samples were left to dry for 20 min. The final DNA pellet was resuspended in 100 µl of dH_2O and dissolved at 37˚C for 15 min. A set of primers was then used to check for the recombination of the *Runx2* genomic locus; Flx probes: F-TCAGCTTTAGCGTCGTCAGA, R-CAAGCTAACGGG ACTTGGAA. WT probe: R-TGTCCGCCACCGCCAAG. PCR was performed in a 25 µl reaction mixture containing 12.5 µl of 2 x Reddymix ThermoPrime Taq DNA Polymerase master mix (Thermo), 0.4 µM of each of the primers and 1 µl of DNA. The reaction mixture with added water instead of the DNA template was run as a negative control. The cycling conditions were as follows: 95°C for 3 min, followed by 34 cycles of three steps consisting of denaturation at 95°C for 30 s, primer annealing at 60°C for 30 s and extension at 72°C for 60s plus a final step of extension at 72°C for 10 min. PCR reactions were loaded onto a 2% agarose gel and bands visualized under a GelDoc-IT TS Imaging System (UVP, US).

2.9.2 PCR detection of *Runx2* mRNA.

RNA was isolated using a Qiagen RNeasy Mini Kit (Qiagen, UK) according to the manufacturer's instructions and quantified using a Nanovue spectrophotometer. RNA was cleared from any DNA contamination with DNAse treatment and reverse transcribed to cDNA using Quantitect® Reverse Transcription Kit (Qiagen, 205311). A PCR for genomic DNA was performed to confirm the purity of the cDNA obtained. RT-PCR was performed in a 25 µl reaction mixture containing 12.5 µl of 2 x Reddymix ThermoPrime Taq DNA Polymerase master mix (Thermo), 1 µM of each of the primers and 1 µl of cDNA. Primers used were *Runx2* (F- AGAGGGCACAAGTTCTATCTG; R-GCTGTTGCTGTTGTTGCT) and ß-Actin (F- AGAGGGAAATCGTGCGTGAC; R-CAATAGTGATGACCTGGCCGT). The reaction mixture with added water instead of the cDNA template was run as a negative control. The cycling conditions were as follows: 95°C for 3 min, followed by 34 cycles of three steps consisting of denaturation at 95°C for 30 s, primer annealing at 60°C for 30 s and extension at 72°C for 60s plus a final step of extension at 72°C for 10 min. PCR reactions were loaded onto a 2% agarose gel and bands visualized under a GelDoc-IT TS Imaging System (UVP, US).

2.10 Quantitative RT-PCR

RNA was isolated using a Qiagen RNeasy Mini Kit (Qiagen, UK) according to the manufacturer's instructions and quantified using a Nanovue spectrophotometer. RNA was cleared from any DNA contamination with DNAse treatment and reverse transcribed to cDNA using Quantitect® Reverse Transcription Kit (Qiagen, 205311). SYBR Green based quantitative PCR was performed in triplicate in a 20 µl reaction mixture containing 10 µl of 2 x SYBR®Green Jumpstart™ Taq (Sigma, S4438) master mix, 0.5 µM of each of the primers and 5 µl of 10x diluted cDNA. The reaction mixture with water instead of the template was run as a control. The cycling conditions were as follows: 94°C for 2 min, followed by 40 cycles of two steps consisting of denaturation at 94°C for 15 s and primer annealing/extension at 60°C for 60s plus a final step of extension at 72°C for 5 min. A melting curve analysis was performed from 70°C to 95°C in 0.3°C intervals to demonstrate the specificity of each amplicon and to identify the

formation of primer dimers. Data were acquired using an MJ Chromo4 (BioRad) and analysed using MJ Opticon Monitor 3 software (BioRad). The following primers were used: *Runx1* (Qiagen Quantitect Assay QT0010000380), *Runx2* (Qiagen Quantitect assay QT00102193), *Runx3* (F-GCA CCG GCA GAA GAT AGA AGAC; R-GGTTTAAGAAGCCTTGGATTGG), *Axin2* (F-GCTCCAGAAGATCACAAAGAG; R-AGCTTTGAGCCTTCAGCATC), *Hes1* (F-CAGGAGGGAAAGGTTATTTTGACG; R-TAG TTGTTGAGATGGGAGACCAGGCG), *Slug* (F-CTCACCTCGGGAGCATACA; R-GACTTAC ACGCCCCAAGGATG), *p21* (F-CAAGAGGCCCAGTACTTCC; R-TGGAGTGATAGAA ATCTGTCAGG) and GAPDH (PrimerDesign kit). All reactions were performed in triplicate, and expression was normalised to GAPDH.

2.11 Western blot

Nuclear extracts were prepared from mammary cell lines using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Scientific, Cat No 78833) as per kit instructions. Briefly, addition of the first two reagents of the kit (CER1, CER2) to a cell pellet causes cell membrane disruption and release of cytoplasmic contents. After recovering the intact nuclei from the cytoplasmic extract by centrifugation, the nuclei are lysed with a third reagent (NER) to yield the nuclear extract. At least 15 micrograms of protein extract were resolved on 10% NuPAGE Novex Bis-Tris gels (Life Technologies) and transferred to Hybond-ECL nitrocellulose membranes (Amersham). Membranes were probed with antibodies to RUNX1 (HPA004176, Sigma), RUNX2 (HPA022040, Sigma and R&D Systems, AF2006), GAPDH (Cell Signalling), ß-Actin (Cell signalling), Lamin A/C (Cell Signalling), E-Cadherin (Cell Signalling), Vimentin (Sigma).

2.12 Human breast samples/tissue microarray

All expression studies in human tissues were ethically approved (REC Ref: Project Number 02/SG007(10), R and D project: RN07PA001).

Tissue microarrays used in the current study were kindly provided by Dr. Joanne Edwards who has given a brief outline of the procedure. 0.6 mm^2 cores of breast cancer tissue, identified by the pathologist, were removed from representative areas of the tumour taken from breast cancer patients at the time of surgical resection. All tissue microarray blocks were constructed in triplicate. All patients were diagnosed with operable invasive breast carcinoma between 1980 and 1999 in the Greater Glasgow and Clyde area. These patients received standard adjuvant treatment according to protocols at the time of diagnosis. Patient follow-up details included information on clinical attendances, recurrence and metastasis, date and cause of death as well as adjuvant therapy details. ER, PR (Mohammed et al. 2012), HER2 status (Mohammed et al. 2012), immune and inflammatory infiltrate (Mohammed et al. 2012) were already available for this cohort. The tissue microarrays were stained for RUNX1 protein by immunohistochemistry and quantified using the weighted histoscore method [(0 x % negative staining) + (1 x % weak staining) + (2 x % moderate staining) + (3 x % strong staining)] to give a value of 0 – 300 (Kirkegaard et al. 2006).

2.13 Statistical analysis

Statistical significance ($p<0.05$) of differential findings between experimental groups was determined by a Student's t test (unless otherwise specified) using Minitab or GraphPad software. For the human studies SPPS-19 software was employed and disease-specific survival rates were generated using the Kaplan-Meier method. The log-rank test was used to compare significant differences between subgroups using univariate analysis. Interrelationships between RUNX1 expression and clinical parameters, hormonal status and inflammatory features were calculated using Cox-regression and chi square test.

3 Characterization of the *RUNX* genes in the mammary gland.

3.1 INTRODUCTION

3.1.1 RUNX proteins in mammary gland development.

Given its first discovery as the "bone-master regulator" gene, most of the research on RUNX2 has been focused on its role in osteoblast lineage specification and regulation of bone development (Long 2011). Similarly most research on RUNX1 has been focused on its role in the haematopoietic system while RUNX3 has been mostly studied in the context of its role as a putative tumour suppressor in epithelial lineages and in neuronal development (Blyth et al. 2005). However *RUNX* genes are expressed in a wide variety of tissues and cell types indicating possible new roles for these transcription factors in different systems (Blyth et al. 2010; Scheitz & Tumbar 2012). In particular several evidences coming from *in vitro* and *in vivo* studies are suggesting a new role for *RUNX* genes in mammary gland development (see Chapter 1). Briefly, all three *RUNX* transcripts have been shown to fluctuate during the different phases of mammary development, decreasing during pregnancy and lactation to rise again in involution (Blyth et al. 2010). In addition, RUNX2 regulates genes characteristic of mammary epithelial cells such as ß-Casein and Osteopontin in the HC11 mammary cell line (Inman et al. 2005; Inman & Shore 2003). Moreover RUNX2 is likely to be expressed in normal murine mammary cells in the embryonic mammary gland (Otto et al. 1997) and it is enriched in the terminal end buds (TEB) of young pubertal mammary glands (Kouros-Mehr & Werb 2006).

3.1.2 Experimental Aims.

So far the early lethality of RUNX knock-out animals has hindered the study of the role of those transcription factors in the mammary gland, an organ which mainly develops after birth. The aim of this study was to characterize the function of *RUNX* genes in mammary gland development through a combination of *in vitro* and *in vivo* loss of function approaches based on the use of RUNX-specific conditional knock-out models.

3.2 RESULTS

3.2.1 Characterization of RUNX expression in the normal mammary gland.

To better understand the role of RUNX proteins in mammary gland development it is important to improve our knowledge of their normal physiological expression pattern in mammary epithelial tissue. The mammary epithelium is composed of two main lineages: luminal (ductal and alveolar) and basal cells (myoepithelial) (Richert et al. 2000). Recent advances in cell surface marker analysis through FACS allow the isolation of specific cell lineages from mouse mammary epithelium. In particular specific enrichment for the basal/myoepithelial or luminal compartment can be achieved by using a combination of cell surface markers (CD31, TER119 and CD45 to exclude haematopoietic and endothelial cells together with the epithelial markers CD24 and CD29) (Shackleton et al. 2006). To investigate RUNX expression in the two main mammary lineages FACS sorting for luminal and basal populations was performed (see Material and Methods) on fresh-extracted mouse mammary epithelial cells (MMECs) from 6 week old mice. qRT-PCR on the sorted populations showed that *Runx*1 represents the more expressed isoform while *Runx*3 transcripts are expressed at very low levels or undetectable in both compartments (FIGURE 3-1-A). Moreover both *Runx1* and *Runx2* transcripts are enriched in the basal population which interestingly has been showed to be enriched in mammary stem cells (Shackleton et al. 2006). Since the pubertal virgin mammary gland at 6 weeks is an actively growing and expanding organ characterized by high numbers of TEBs, expression in the mature virgin mammary gland was also assessed. qRT-PCR on FACS-sorted MMECs extracted from 12 week old mature virgin mice confirmed the expression pattern for all *Runx* isoforms as observed in the younger mice (FIGURE 3-1-B). From these results it was decided to concentrate future analysis on RUNX1 and RUNX2 proteins, since they are the major RUNX proteins expressed in mammary tissue. To confirm RUNX1 and RUNX2 expression at the protein level, immunohistochemistry (IHC) was performed showing that both RUNX1 and RUNX2 are indeed expressed in discrete subpopulations of mouse and human mammary

epithelium (FIGURE 3-2). RUNX1 and RUNX2 expression was then characterized during different phases of embryonic and adult mammary development. Interestingly RUNX1 and RUNX2 have the opposite pattern of expression during embryonic development, with RUNX2 being expressed specifically at embryonic day E12 while RUNX1 was apparent from E16 onwards (FIGURE 3-3). However, in adult mouse mammary development both RUNX1 and RUNX2 show a similar expression pattern, decreasing during pregnancy and lactation to rise again during involution (FIGURE 3-4) which confirms with what has been shown at the RNA level (Blyth et al. 2010).

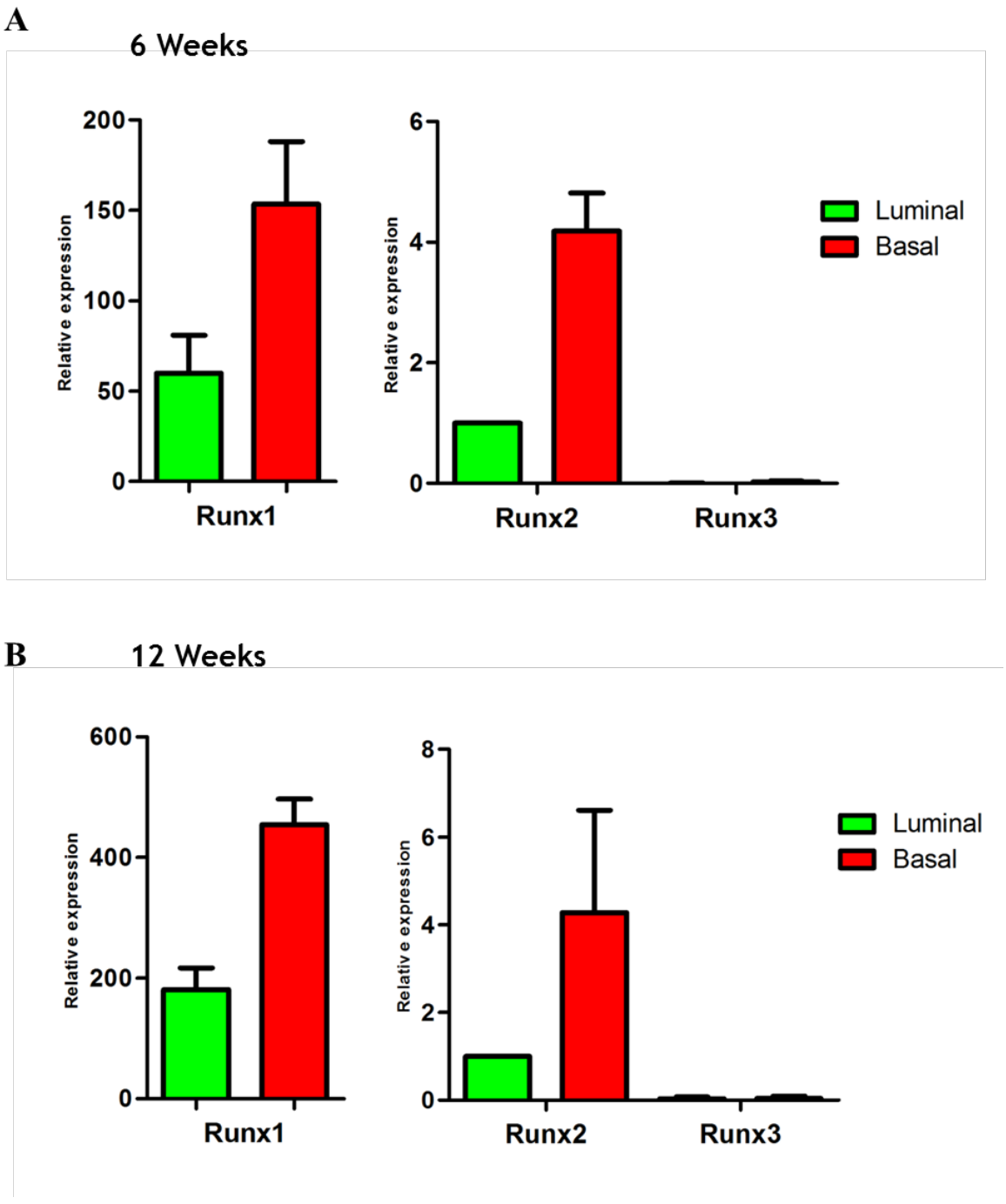


Figure 3-3-1 Expression of *Runx* genes in mammary luminal and basal populations.

The relative levels of *Runx1*, *Runx2* and *Runx3* mRNA measured by real-time RT-PCR on mouse mammary luminal and basal populations as sorted by FACS, based on expression of CD29 and CD24 surface markers. MMECs were extracted from 6 weeks (A) and 12 weeks (B) virgin FVB female mice. Data are expressed as mean relative expression (±SD). *Runx*1 is plotted on a different Y-axis due to the higher levels of expression. Expression normalised to *Gapdh* is relative to luminal *Runx2*; each bar represents at least 3 separate experiments (3 independent groups of 3 mice each).

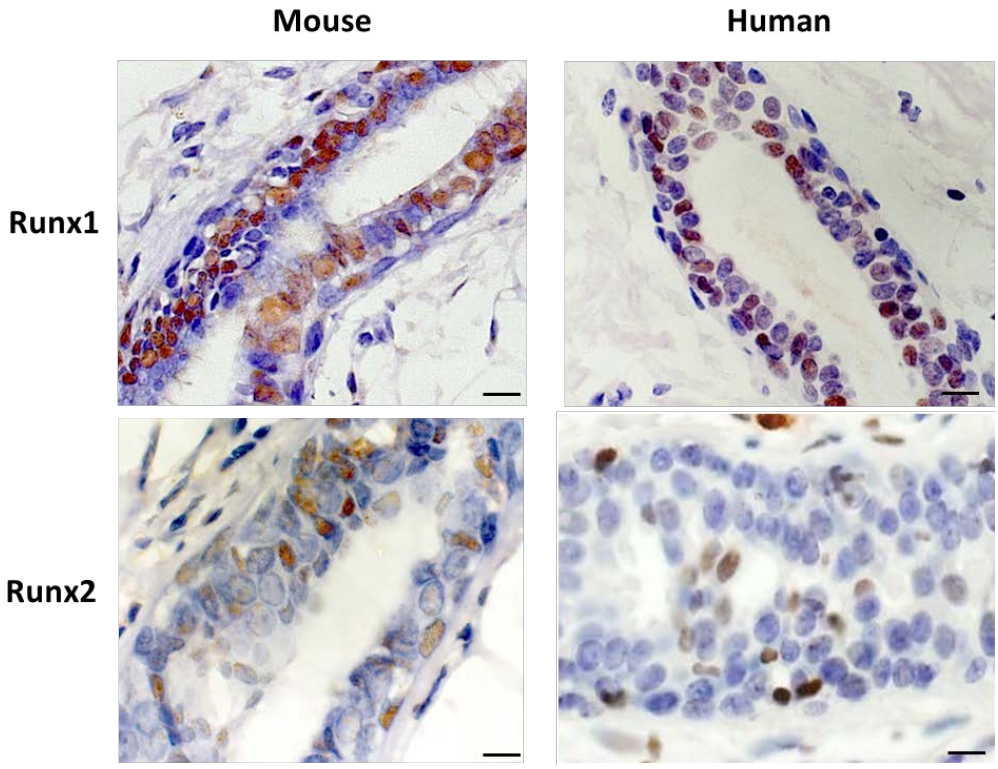


Figure 3-3-2 Expression of RUNX1 and RUNX2 in human and murine mammary gland.

Immunohistochemistry for RUNX1 and RUNX2 on mouse and human virgin mammary epithelium. Sections were counterstained with haematoxylin. Scale bars represent 10 μm.

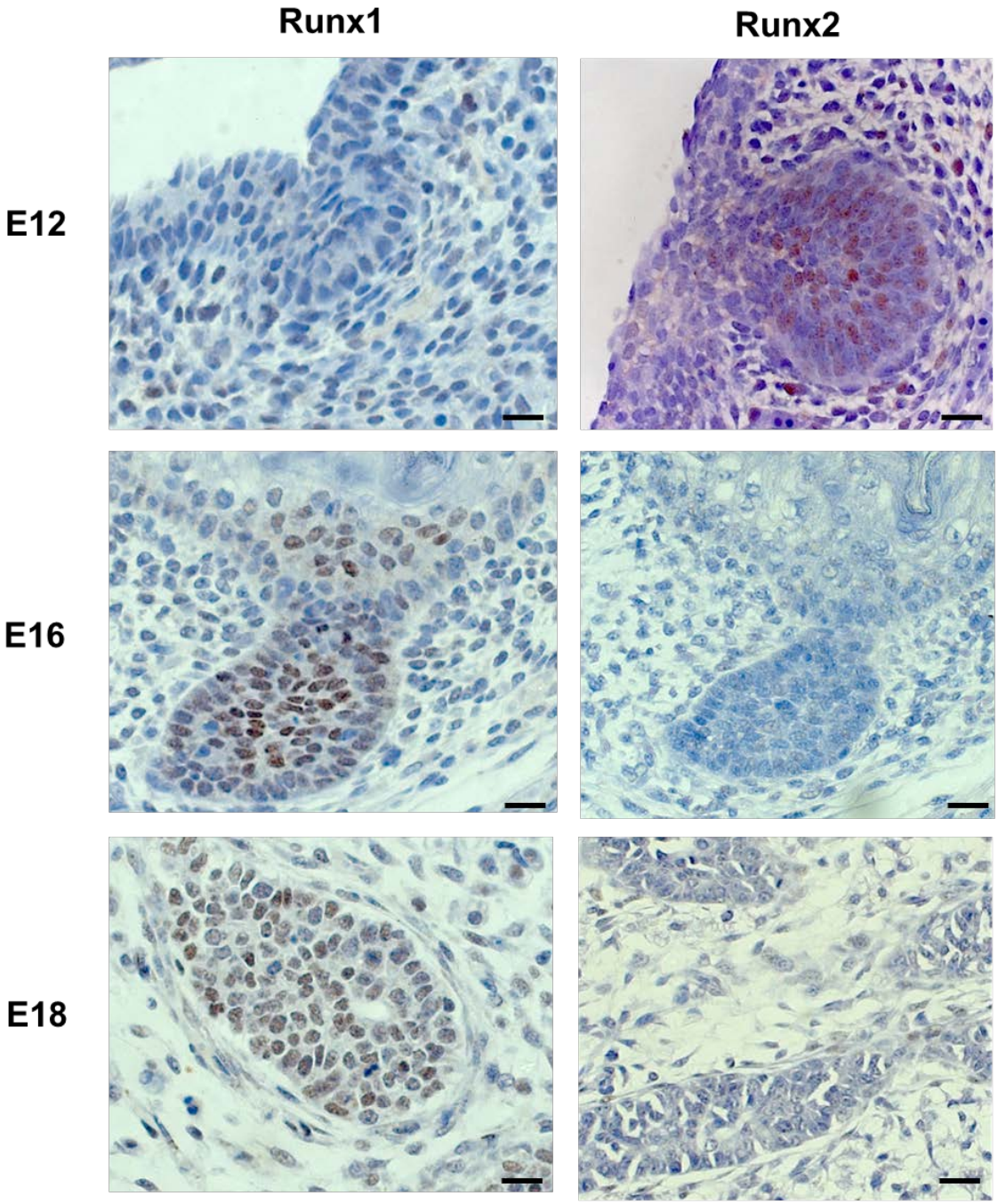


Figure 3-3 RUNX1 and RUNX2 expression in murine embryonic development.

Immunohistochemistry for RUNX1 and RUNX2 on mouse embryonic mammary epithelium, taken at the selected timepoints (e.g. E12 is 12 days post-coitus). Sections were counterstained with haematoxylin. Scale bars represent 20 μm.

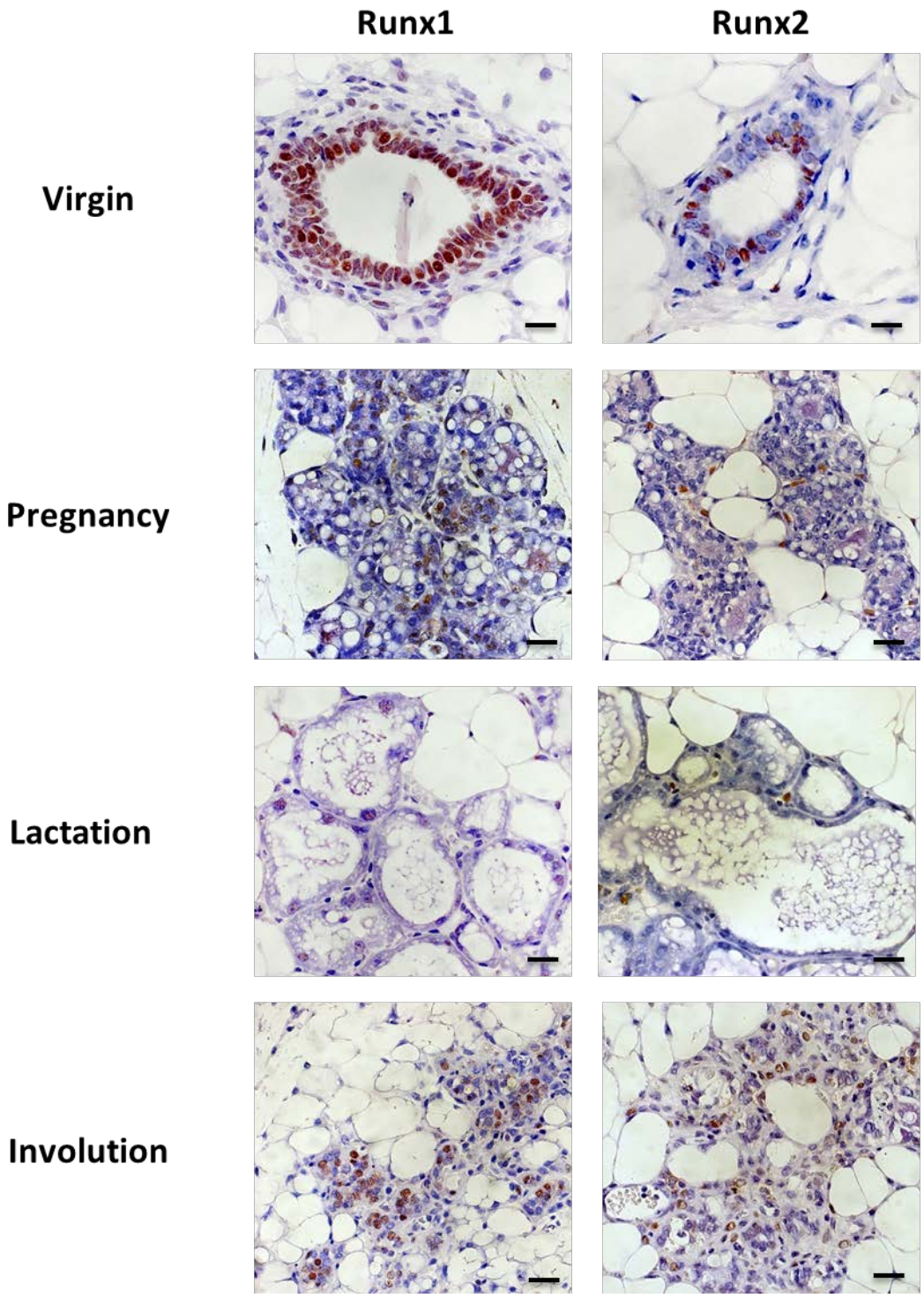


Figure 3-4 RUNX1 and RUNX2 expression in adult mammary development.

Immunohistochemistry for RUNX1 and RUNX2 on adult mouse mammary epithelium, taken at the selected timepoints (12 week old virgin, d17 pregnancy, d1 lactation, d7 involution). Sections were counterstained with haematoxylin. Scale bars represent 30 μm.

3.2.2 Mammary specific deletion of RUNX1 and RUNX2

To understand the functional role of *Runx* genes in mouse mammary gland biology *in vivo* and given the early lethality of RUNX knock-out animals, conditional knock-out mouse models for *Runx1*$^{flx/flx}$ (Growney et al. 2005) and *Runx2*$^{flx/flx}$ (unpublished, see Material and Methods for characterization) in combination with a mammary specific Cre recombinase (BLG-Cre) were used. In those mice Cre DNA-recombinase is under the control of the mammary gland specific promoter of the ovine beta-lactoglobulin (BLG) gene. BLG-Cre is mainly activated during pregnancy and lactation and is commonly used as a specific Cre for targeting the mammary luminal lineage (Selbert et al. 1998). First BLG-Cre expression in the mammary gland was characterized using the Z/EG reporter mouse which utilises a lox-stop-lox GFP reporter cassette (Novak et al., 2000). *In vivo* GFP imaging on BLG-Cre:Z/EG+ mice showed consistent GFP expression in the mammary epithelium indicating that the transgenic models were working (FIGURE 3-5-A). To identify the mammary lineages which are targeted by BLG-Cre expression, FACS mammary lineage profiling was performed using GFP expression as a tracker for BLG-Cre activation. Confirming published data (Molyneux et al. 2010), we found that BLG-Cre is mainly expressed in the virgin luminal population. However, a small but detectable percentage of the basal population in which BLG-Cre is active was also detected (FIGURE 3-5-B). Immunofluorescence was used to confirm this data showing double positive GFP+/CK14+ cells in BLG-Cre:Z/EG+ mice (FIGURE 3-5-C). This discrepancy with published data could be due to differences in mouse strains or the use of a more sensitive reporter gene compared to ß-gal as used previously (Molyneux et al. 2010). BLG-Cre expression was then characterized at different stages of mammary development confirming that Cre expression is activated during pregnancy and lactation and decreases during involution (FIGURE 3-5-D).

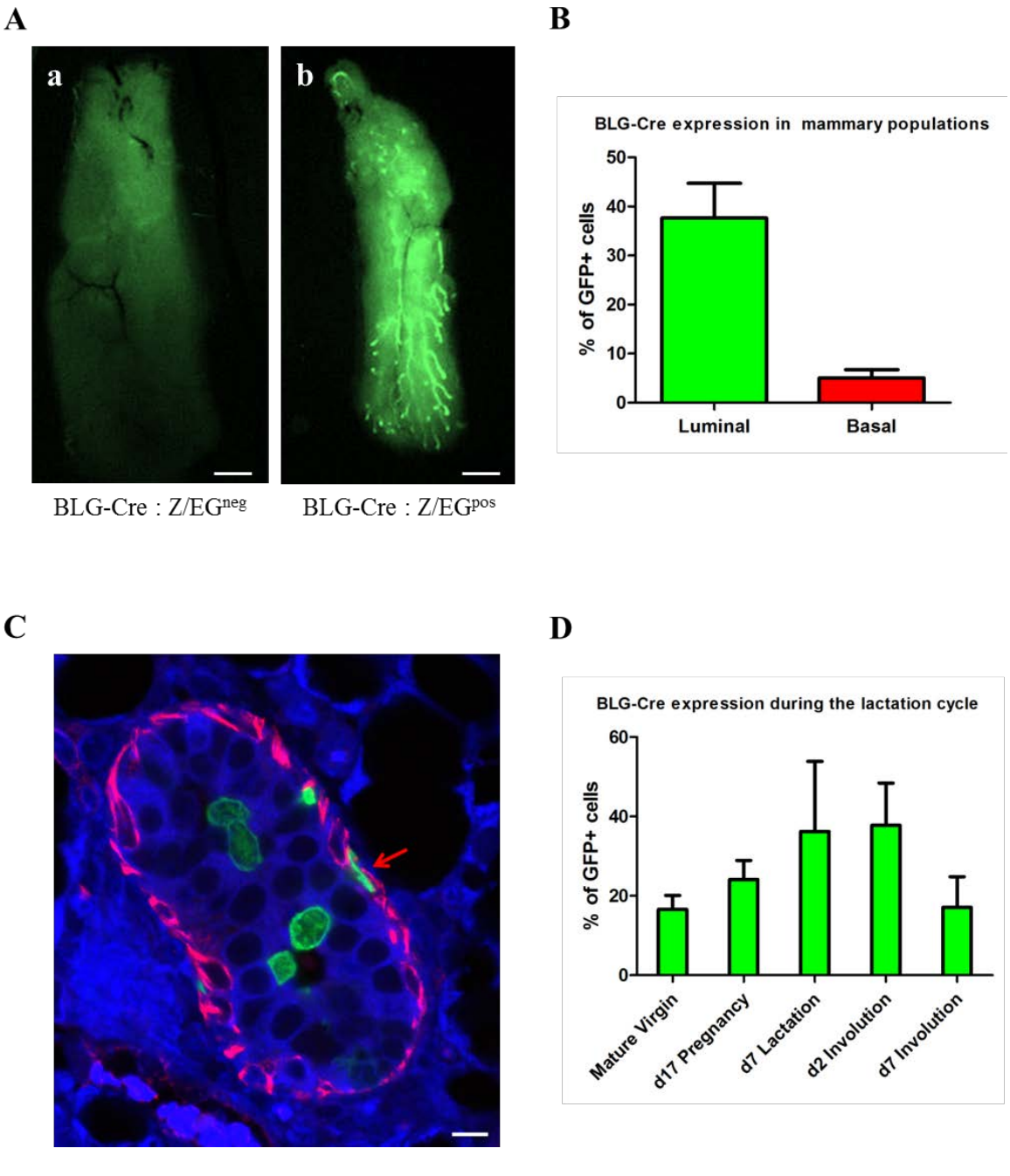


Figure 3-5 Characterization of BLG-Cre expression through the use of a GFP reporter.

In vivo GFP expression as a surrogate for Cre expression in 12 week virgin BLG-Cre:Z/EG negative (a) and BLG-Cre:Z/EG+ (b) mice. Scale bars represent 5 mm (A). Reporter GFP expression as analysed by FACS, gated on CD24+/CD29low luminal population (green) and CD24+/CD29high basal population (red) from 12 week BLG-Cre+:Z/EG+ mice. Data are expressed as % of GFP+ cells over total (±SD). 8 groups of 2 mice each analysed (B). CK14 (red) and GFP (green) co-immunofluorescence on mammary gland tissue from BLG-Cre+:Z/EG+ mice. Red arrow indicates a GFP+ (green) CK14+ (red) basal cell. Glands were counterstained with DAPI (blue). Scale bar represents 10 µm (C). Reporter GFP expression, as analysed by FACS, on total cells extracted from mammary glands of BLG-Cre+:Z/EG+ mice at different developmental stages (n=3 for each stage). Data are expressed as % of GFP+ cells over total (±SD) (D).

Using the RUNX conditional knock-out mice two mouse cohorts were generated: BLG-Cre:Z/EG:*Runx1*$^{flx/flx}$ and BLG-Cre:Z/EG:*Runx2*$^{flx/flx}$ (FIGURE 3-6-A). IHC on BLG-Cre *Runx1*$^{flx/flx}$ mice showed a reduction of RUNX1 protein expression in the adult virgin epithelium (FIGURE 3-6-B). Moreover PCR for *Runx2* genomic locus on MMECs extracted from BLG-Cre *Runx2*$^{flx/flx}$ mice showed the presence of a recombinant band, indicating that an *in vivo* Cre-mediated deletion of the *Runx2* gene was occurring (FIGURE 3-6-C). These results confirm that the conditional knock-out strategy is working *in vivo*.

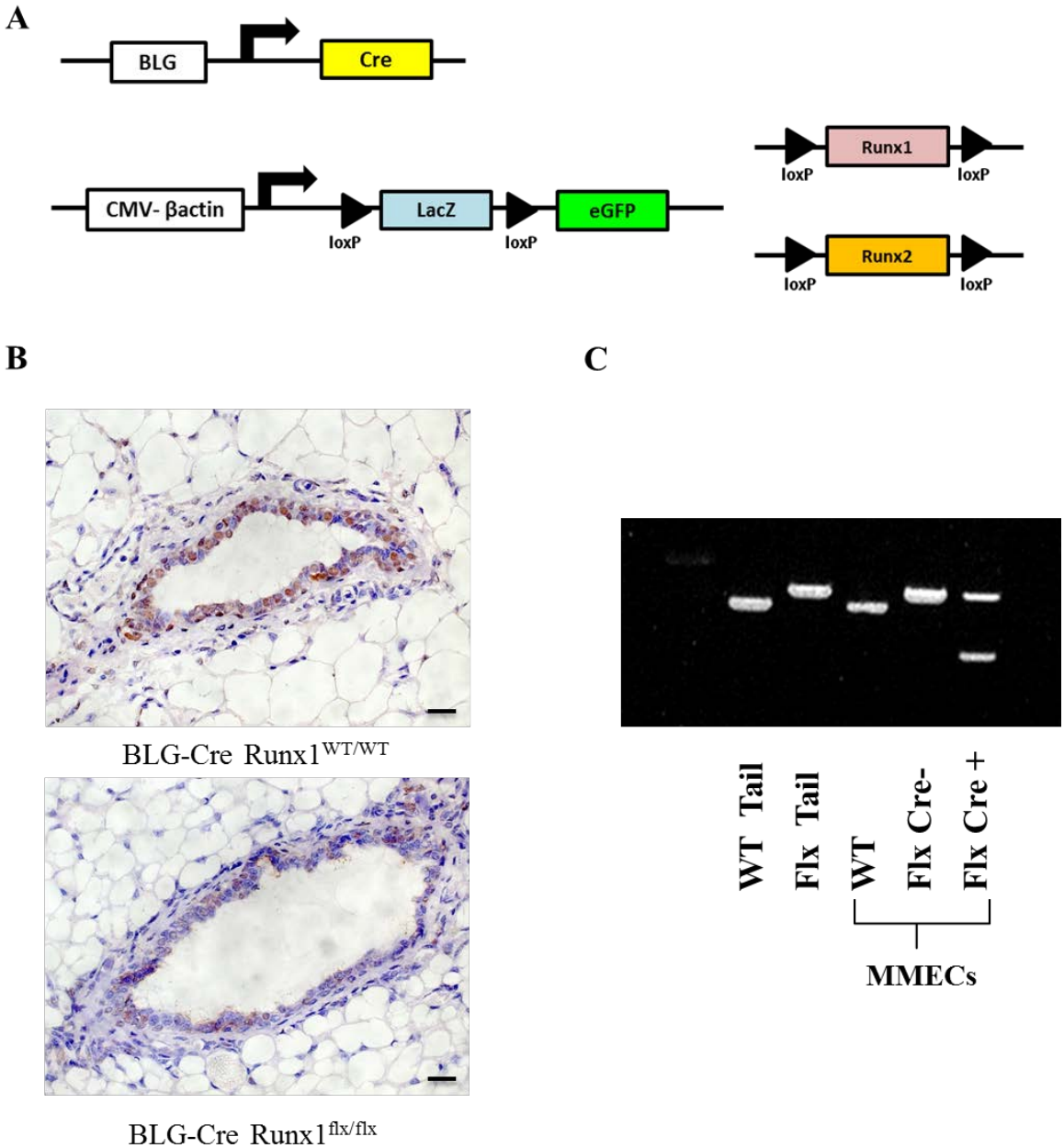


Figure 3-6 Characterization of *Runx1* and *Runx2* conditional knock-out mice.

Schematic representation of the conditional knock-out strategy used to create a mammary-specific deletion of *Runx1* and *Runx2* combined with a GFP reporter. The BLG promoter will drive Cre expression specifically in the mammary gland. In the target tissue Cre recombinase will drive recombination at the LoxP sites resulting in *Runx1* or *Runx2* loss. In addition Cre recombinase activity will cause loss of the LacZ cassette and activation of eGFP expression (A). Immunohistochemistry for RUNX1 on 12 weeks old virgin mammary epithelium from BLG-Cre *Runx1*$^{WT/WT}$ and BLG-Cre *Runx1*$^{flx/flx}$ mice. Scale bar represents 30 µm (B). PCR on gDNA extracted from BLG-Cre *Runx2*$^{flx/flx}$ MMECs and tail. In the BLG-Cre *Runx2*$^{flx/flx}$ MMECs (Flx Cre+) the 2 bands correspond to the recombined *Runx2* locus (lower band) and the unrecombined cells (upper band). The size of the band in *Runx2*$^{flx/flx}$ cells is larger than the one from *Runx2*$^{WT/WT}$ because of the presence of LoxP sites (C).

No obvious phenotype was detectable in mature virgin mice (>12 weeks of age) from the two cohorts as assessed by whole gland analysis and histological analysis (FIGURE 3-7). However, the lack of phenotype in *Runx* deleted mice could be due to loss of recombined cells and re-population of the gland by wild type (unrecombined) cells. Furthermore, since BLG-Cre is only expressed in a small percentage of the virgin mammary epithelium (around 15% of mammary epithelial cells, Figure 1-5-D) the phenotype may be too mild to be observed at the histological level. To better analyse the recombinant glands, the fate of *Runx*-deleted cells was tracked in the mammary epithelium using the lox-stop-lox GFP reporter cassette. FACS analysis of GFP expression levels in MMECs extracted from mature virgins of BLG-Cre *Runx1*$^{flx/flx}$ and BLG-Cre *Runx2*$^{flx/flx}$ females was performed. Interestingly both RUNX1 and RUNX2 loss caused a significant decrease in overall GFP levels compared to BLG-Cre+ *Runx1*$^{WT/WT}$ *Runx2*$^{WT/WT}$ (FIGURE 3-8). Thus, RUNX1 and RUNX2 loss is impairing mammary epithelial cell maintenance *in vivo*. To determine if RUNX1 and RUNX2 are playing diverse roles in different mammary cell types, mammary population profiling through FACS using CD24 and CD29 cell surface markers, was performed to separate basal and luminal lineages. The GFP+ in each of the gated basal and luminal population is then used to specifically ascertain levels of RUNX-deleted cells. Interestingly loss of RUNX1 caused a reduction in GFP+ cells in the luminal compartment but did not affect the basal population (FIGURE 3-9-A, 3-9-B). In contrast BLG-Cre *Runx2*$^{flx/flx}$ mice showed a reduction in both basal and luminal GFP+ cells compared to BLG-Cre *Runx2*$^{WT/WT}$ mice (FIGURE 3-9-C, 3-9-D). These results suggest that RUNX1 and RUNX2 expression is important for maintenance of specific subpopulations of the virgin mammary epithelium *in vivo*.

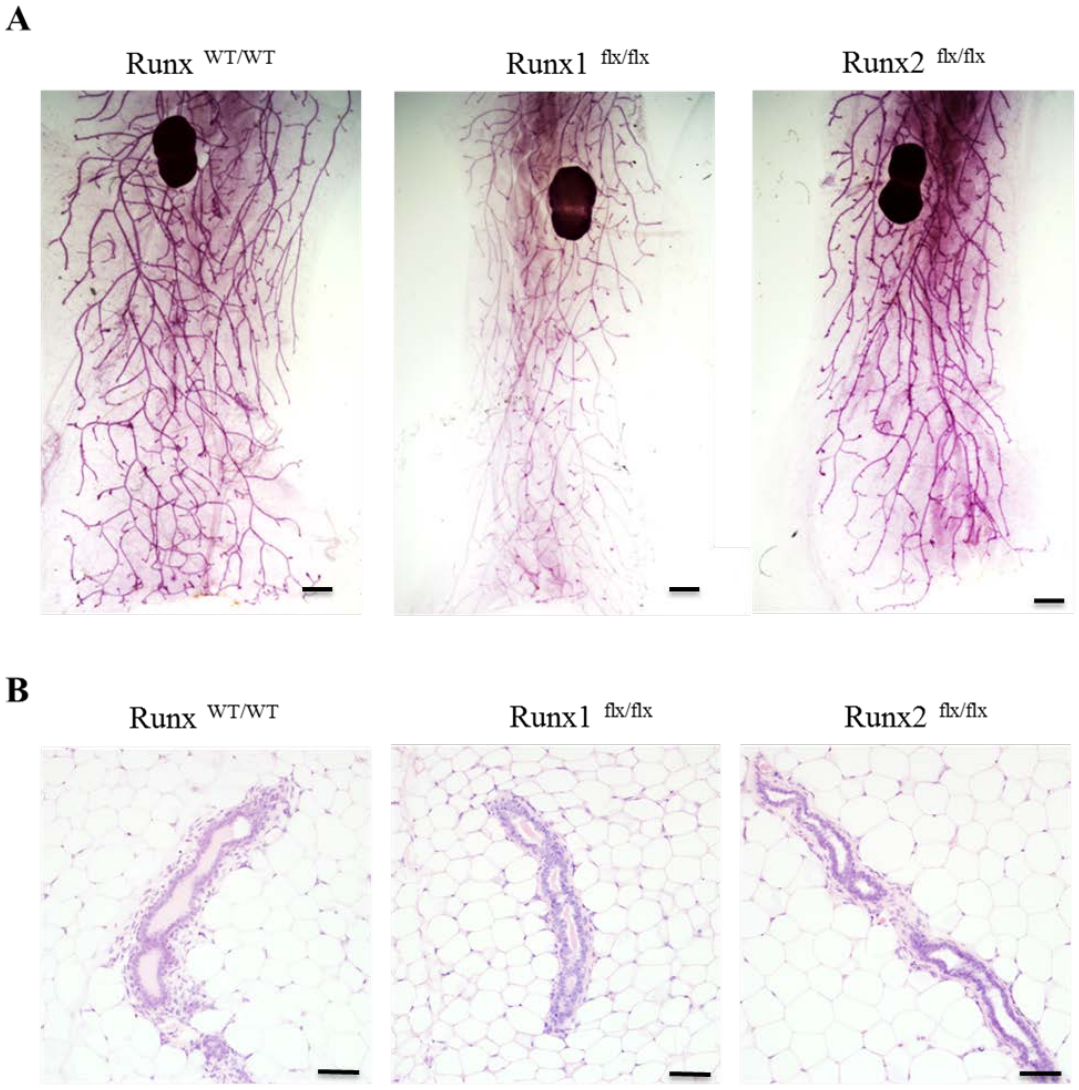


Figure 3-7 Effects of loss of RUNX1 and RUNX2 in the virgin mammary gland.

Wholemounts (A) and histological sections (B) of mammary gland from BLG-Cre *Runx*$^{WT/WT}$, *Runx1*$^{Flx/Flx}$ and *Runx2*$^{Flx/Flx}$ mice extracted from 12 week virgin mice (n>7 for each genotype). No differences were detected between the three genotypes. Scale bars in (A) represent 1 mm. Scale bars in (B) represent 50 µm.

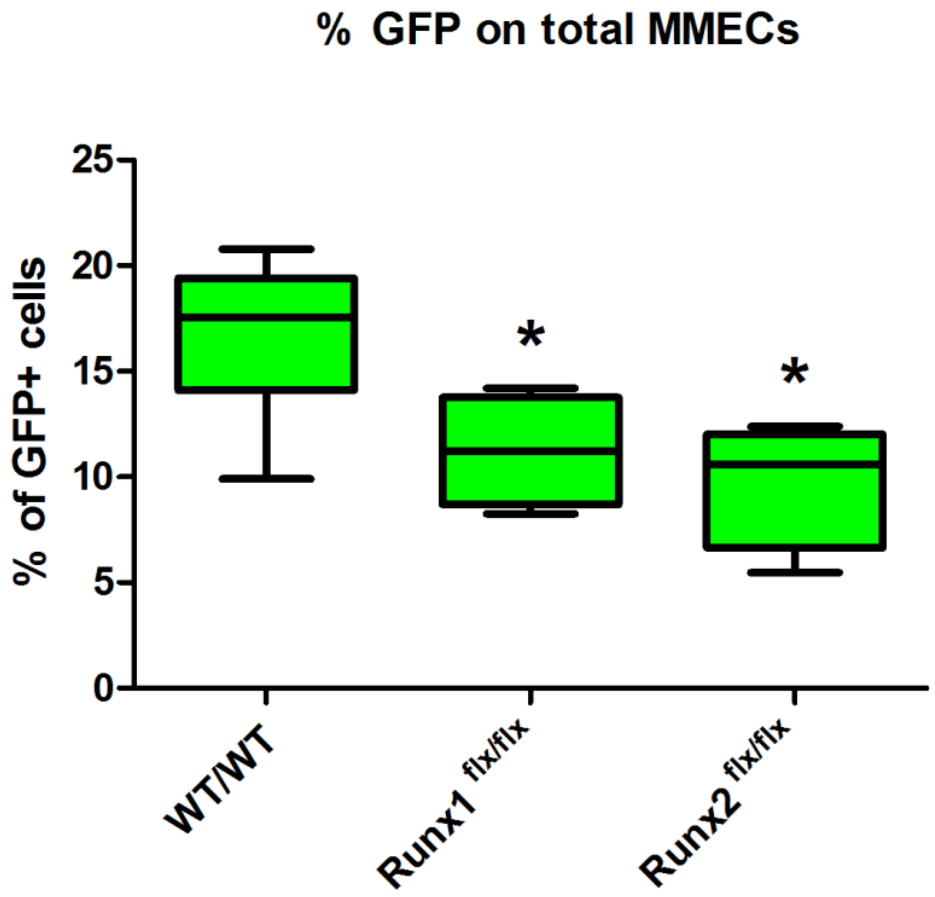


Figure 3-8 Effects of RUNX1 and RUNX2 loss in mammary epithelium *in vivo*.

Reporter GFP expression as analysed by FACS, on cells extracted from whole mammary glands of 12 week BLG-Cre *Runx*$^{WT/WT}$, BLG-Cre *Runx1*$^{Flx/Flx}$ and BLG-Cre *Runx2*$^{Flx/Flx}$ mice. Data are expressed as % of GFP+ cells over total (±SD). n≥4 for each genotype. *p<0.05, Mann Whitney test.

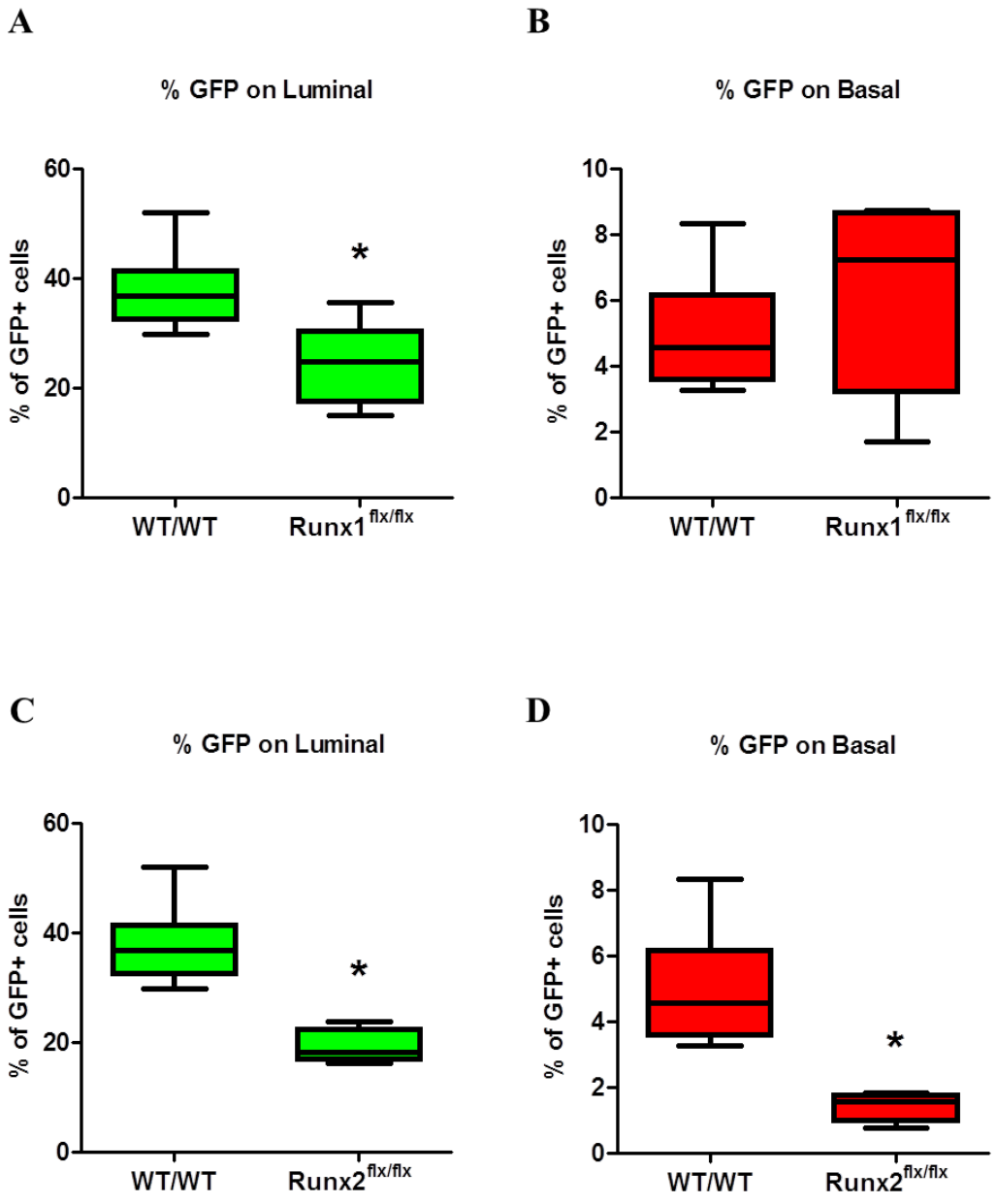


Figure 3-9 Effects of RUNX1 and RUNX2 loss on luminal and basal populations *in vivo*.

Reporter GFP expression as analysed by FACS, gated on CD24+/CD29^{low} luminal population (green) and CD24+/CD29^{high} basal population (red). MMECs were extracted from 12 week BLG-Cre *Runx1*$^{Flx/Flx}$ (A, B) and BLG-Cre *Runx2*$^{Flx/Flx}$ (C,D) mice and compared to the BLG-Cre *Runx*$^{WT/WT}$ cohort . Data are expressed as % of GFP+ cells over total in the selected population (±SD). n≥4 for each stage. *p<0.01, Mann Whitney test.

To investigate if *Runx*2 is involved at different stages of mammary development the BLG-Cre *Runx2*$^{flx/flx}$ cohort was analysed during the lactation cycle. Conditional *Runx2* knock-out mice do not have any obvious defect during pregnancy and lactation and they can lactate and nurse their pups as normal. FACS analysis of GFP expression levels in MMECs extracted from different phases of the lactation cycle of the BLG-Cre *Runx2*$^{flx/flx}$ mice showed no difference in GFP levels compared to BLG-Cre *Runx2*$^{WT/WT}$ mice (FIGURE 3-10). This result shows that RUNX2 is not necessary for the process of alveolar differentiation and milk production *in vivo*. Further work is currently focusing on the characterization of the effects of RUNX2 loss on the involution process and on the analysis of RUNX1 during the different phases of the lactation cycle.

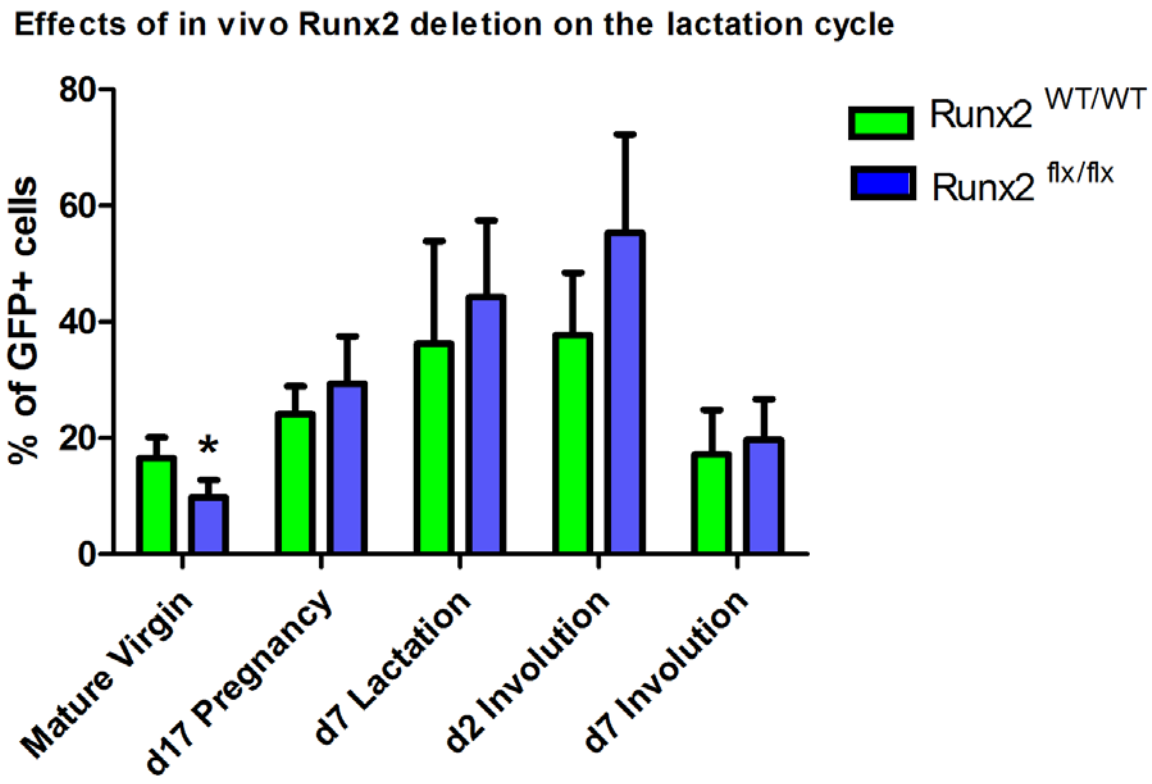


Figure 3-10 Effects of *in vivo Runx2* deletion during the lactation cycle.

Reporter GFP expression, as analysed by FACS, on total mammary gland cells extracted from BLG-Cre *Runx2*$^{WT/WT}$ (green) and BLG-Cre *Runx2*$^{Flx/Flx}$ (blue) mice at different developmental stages. Data are expressed as % of GFP+ cells over total (±SD). n≥3 for each stage. *p<0.05, Mann Whitney test.

3.2.3 Effects of combined deletion of *Runx1* and *Runx2* on virgin mammary development.

All three RUNX transcription factors bind to the same DNA sequence and some cross-regulation between different RUNX family members has been found in the haematopoietic system (Spender et al. 2005). Hence some redundancy could also occur in the mammary gland where RUNX1 and RUNX2 follow a similar pattern of expression. To avoid any compensatory effect of single RUNX deficiency *in vivo*, a double knockout mouse BLG-Cre:*Runx1*$^{flx/flx}$:*Runx2*$^{flx/flx}$ was generated. No obvious phenotype was detectable in young virgin females from this cohort (FIGURE 3-11). FACS analysis of GFP expression levels was then performed in MMECs extracted from mature virgins from BLG-Cre:*Runx1*$^{flx/flx}$:*Runx2*$^{flx/flx}$ mice. Interestingly combined loss of both *Runx1* and *Runx2* rescued the significant decrease in overall GFP levels caused by the single knock-outs (FIGURE 3-12). To understand the cellular dynamics underlying this rescue mammary population profiling through FACS was performed using GFP+ cells as a marker for deleted cells. The percentage of GFP+ cells in the luminal compartment was reduced in the double knock-out gland compared to BLG-Cre *Runx*$^{WT/WT}$ mice, a similar phenotype to the single *Runx1* and *Runx2* knock-outs (FIGURE 3-13-A). However, the basal population was significantly affected in the double knock-out gland with a 5-fold increase in the GFP+ basal population (FIGURE 3-13-B). GFP and CK14 co-immunofluorescence confirmed the increase in basal GFP expression in the BLG-Cre:*Runx1*$^{flx/flx}$:*Runx2*$^{flx/flx}$ glands (FIGURE 3-14-A). Interestingly the overall pattern of GFP expression in the double knock-out mice showed a drastic change with appearance of many ducts with high basal GFP positivity and ducts showing GFP expression only in the basal layer (FIGURE 3-14-B). This result shows that combined loss of RUNX1 and RUNX2 is causing a deregulation in normal homeostasis of mammary lineages driving an expansion of the basal population.

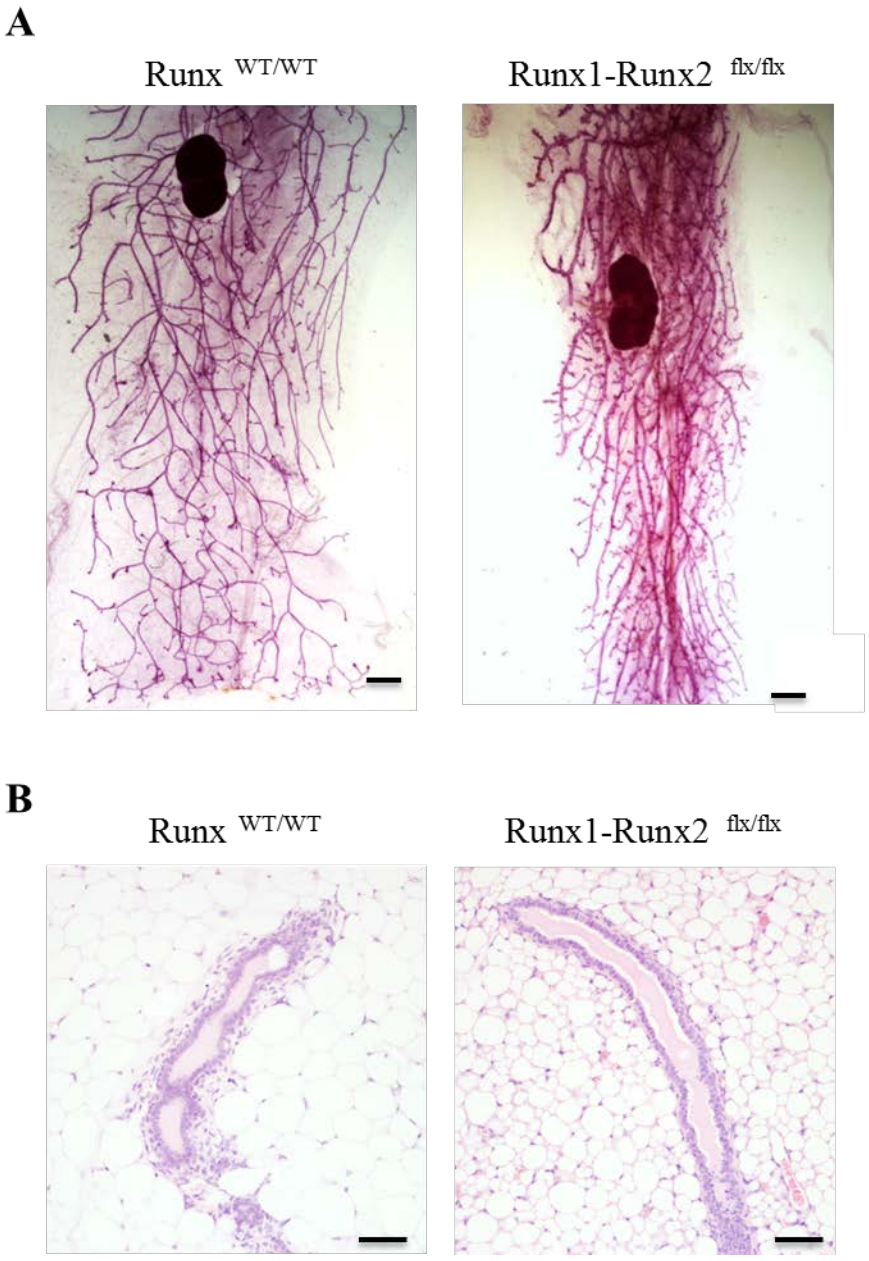


Figure 3-11 Effects of combined loss of RUNX1 and RUNX2 in the virgin mammary gland.

Representative wholemounts (A) and histological sections (B) of mammary gland from BLG-Cre *Runx*$^{WT/WT}$ and BLG-Cre:*Runx1*$^{Flx/Flx}$:*Runx2*$^{Flx/Flx}$ mice extracted from 12 week virgin mice (n>7 for each stage). No differences were detected between the two genotypes. Scale bars in (A) represent 1 mm. Scale bars in (B) represent 50 μm.

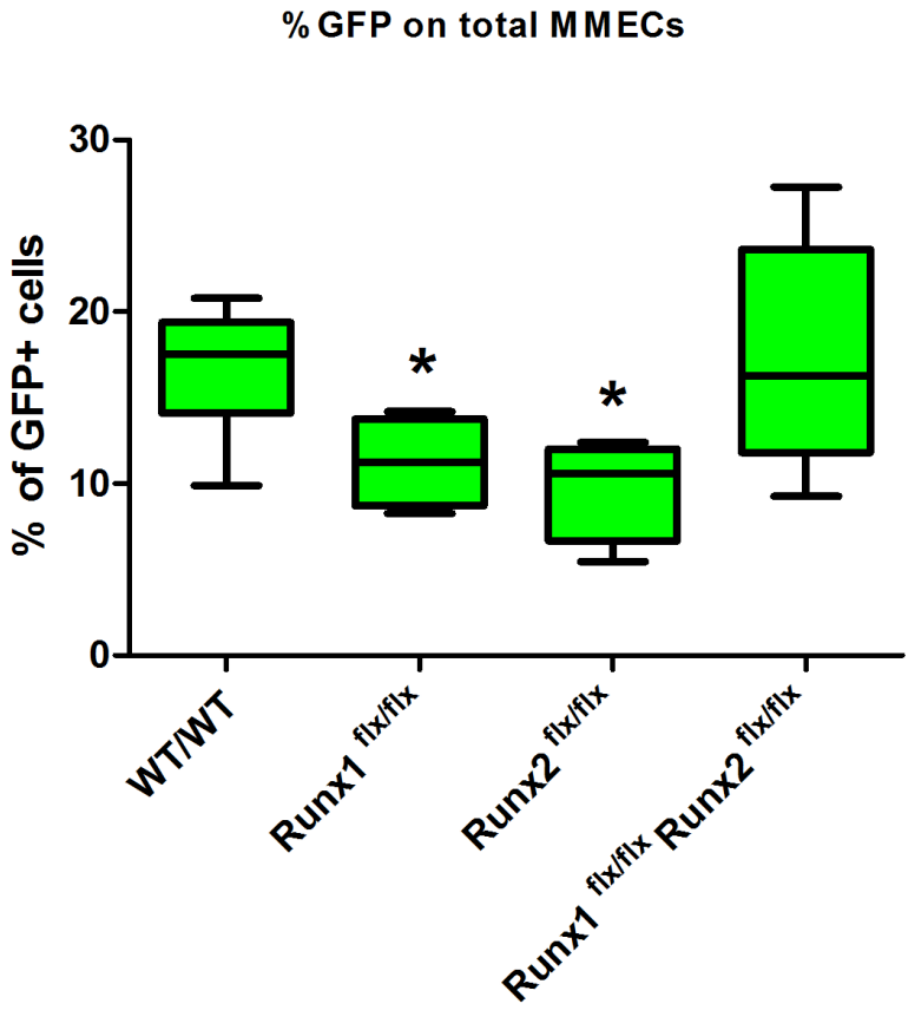


Figure 3-12 Effects of combined loss of RUNX1 and RUNX2 in mammary epithelium *in vivo*.

Reporter GFP expression as analysed by FACS, on total mammary gland extracted from 12 week BLG-Cre *Runx1*$^{Flx/Flx}$, BLG-Cre *Runx2*$^{Flx/Flx}$, BLG-Cre:*Runx1*$^{Flx/Flx}$:*Runx2*$^{Flx/Flx}$ mice, compared to the BLG-Cre *Runx*$^{WT/WT}$ cohort. Data are expressed as % of GFP+ cells over total (±SD). n≥4 for each stage. *p<0.05, Mann Whitney test. No significance was found comparing the double knock-out with the single knock-out cohorts.

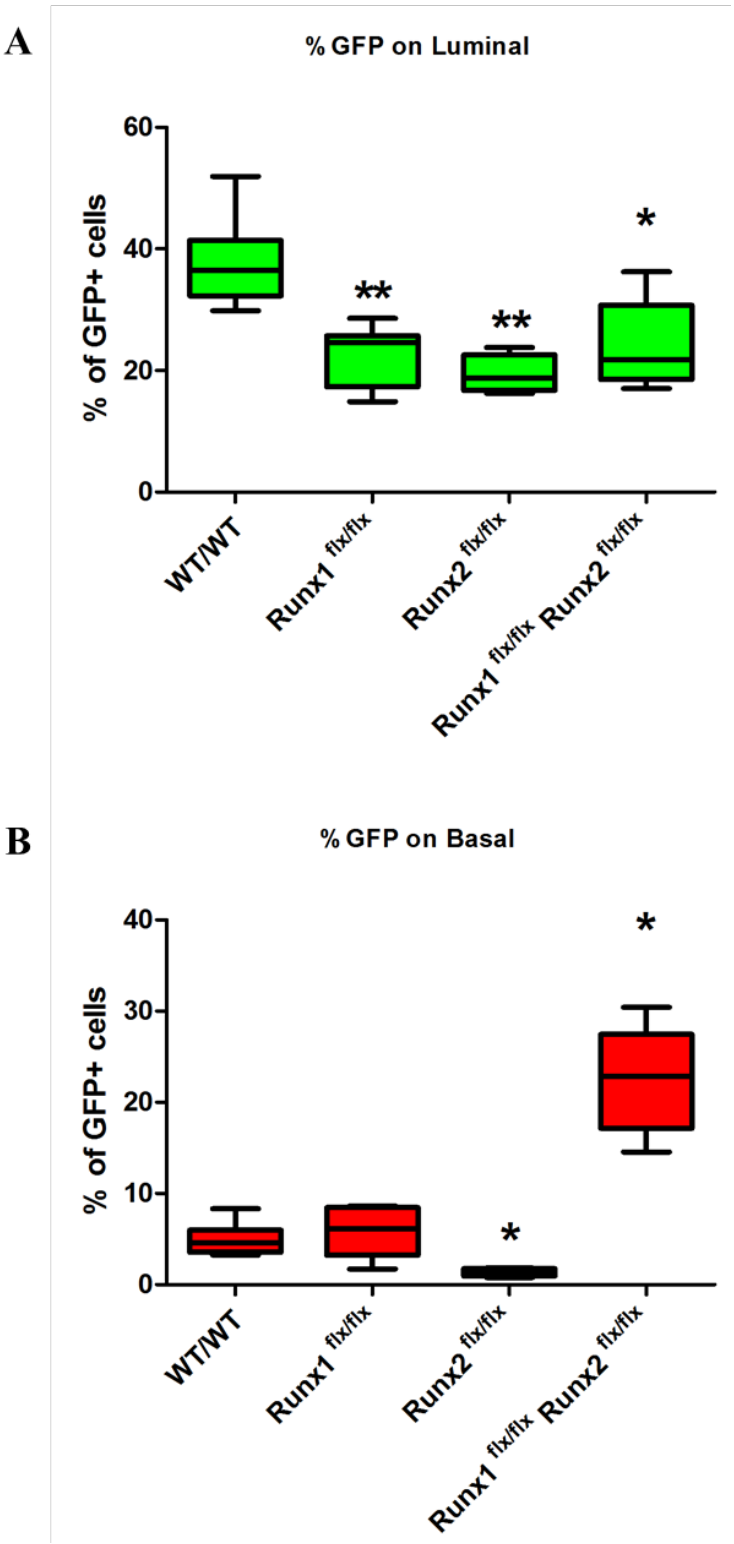


Figure 3-13 Combined loss of RUNX1 and RUNX2 on luminal and basal populations *in vivo*.

Reporter GFP expression as analysed by FACS, on CD24+/CD29low luminal population (A) and CD24+/CD29high basal population (B). MMECs were extracted from 12 week BLG-Cre *Runx1*$^{Flx/Flx}$, BLG-Cre *Runx2*$^{Flx/Flx}$, BLG-Cre:*Runx1*$^{Flx/Flx}$:*Runx2*$^{Flx/Flx}$ mice and compared to the BLG-Cre *Runx*$^{WT/WT}$ cohort. Data are expressed as % of GFP+ cells over total in the selected population (±SD). n≥4 for each stage. In graph A, *=p<0.05, **=p<0.01 when compared to the WT/WT. In graph B, *=p<0.005 when compared to the WT/WT, Mann Whitney test. No significance was found comparing the double knock-out with the single knock-out cohorts in the luminal compartment. The double knock-out was statistically different from the single knock-out cohorts in the basal compartment. p<0.02, Mann Whitney test.

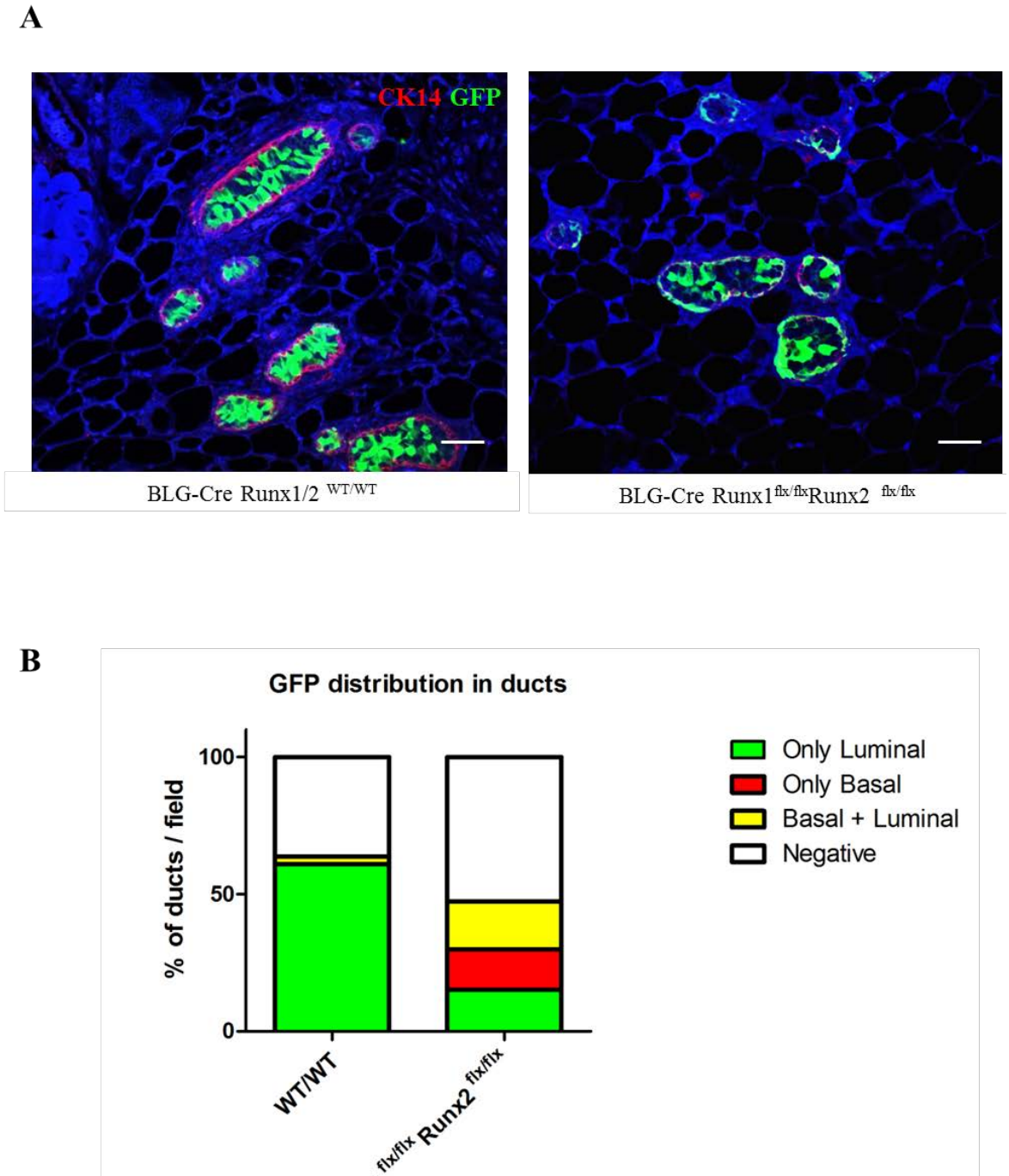


Figure 3-14 Changes in GFP reporter expression in RUNX1-RUNX2 double knock-out glands.

CK14 (red) and GFP (green) co-immunofluorescence on mammary glands extracted from BLG-Cre+ *Runx*$^{WT/WT}$ and BLG-Cre+*Runx1*$^{Flx/Flx}$*Runx2*$^{Flx/Flx}$ mice. Glands were counterstained with DAPI (blue). Representative images shown. Scale bars represent 50 µm (A). Bar chart showing the distribution of GFP+ cells in BLG-Cre+ *Runx*$^{WT/WT}$ and BLG-Cre+ *Runx1*$^{Flx/Flx}$ *Runx2*$^{Flx/Flx}$ mice as determined by GFP IHC (n=4 per each genotype) (B).

To determine the long term effects of the population imbalance caused by *Runx1* and *Runx2* single and double deletion, cohorts of BLG-Cre *Runx1*$^{flx/flx}$, BLG-Cre *Runx2*$^{flx/flx}$ and BLG-Cre:*Runx1*$^{flx/flx}$:*Runx2*$^{flx/flx}$ mice were aged to 6 months and compared to a BLG-Cre wild-type cohort. Histological analysis showed an increase in alveolar structures in the double knock-out glands compared to wild-type and single knock-outs (FIGURE 3-15). Moreover 33% of double knock-out mice develop rare lesions in the mammary gland compared to none of the controls or the single *Runx*-deleted models (Chi square test p=0.0089) (Figure 3-16-A). Histological analysis showed that all lesions are hyperplastic intraductal preneoplastic lesions, characterized by nuclear atypia, squamous metaplasia or necrosis in the middle combined with an abundant immune reaction (Figure 1-16-B).

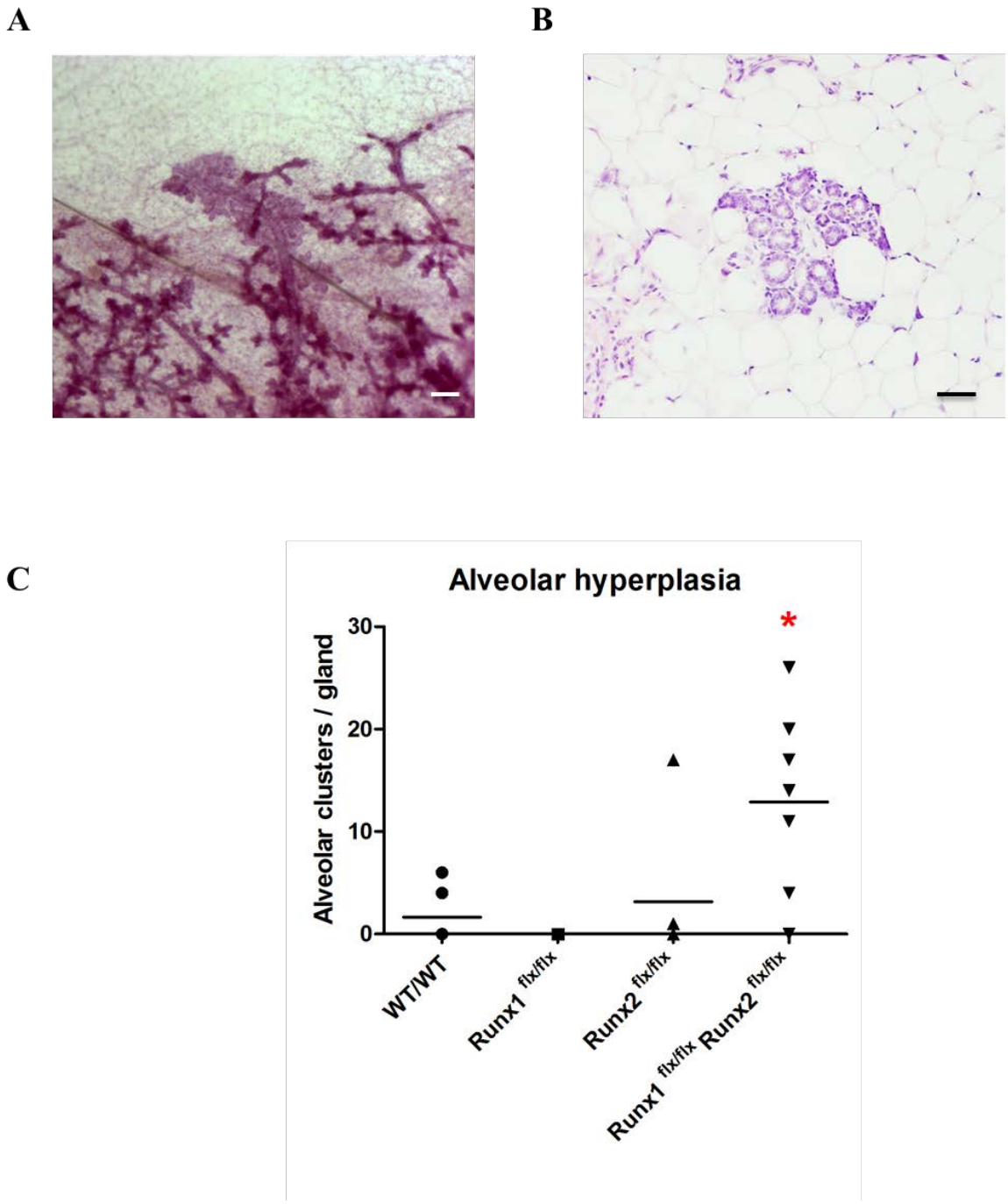


Figure 3-15 Double deletion of RUNX1 and RUNX2 causes alveolar hyperplasia in virgin mice.

Representative examples of alveolar structures detected in wholemounts (A) and histological sections (B) of mammary glands extracted from 6 month virgin mice from the BLG-Cre:*Runx1* $^{Flx/Flx}$:*Runx2*$^{Flx/Flx}$ cohort. Scale bar in (A) represents 200 µm. Scale bar in (B) represents 25 µm. Scatter plot showing the number of alveolar clusters per gland detected in 6 month old BLG-Cre *Runx1*$^{Flx/Flx}$, BLG-Cre *Runx2*$^{Flx/Flx}$, BLG-Cre:*Runx1*$^{Flx/Flx}$:*Runx2*$^{Flx/Flx}$ mice, compared to the BLG-Cre *Runx*$^{WT/WT}$ controls. Mice (n≥6 per group) were oestrous matched (metestrous-diestrous) through vaginal smear (C). Alveolar clusters per gland were counted under a bright-field microscope on an H&E section. * = $p<0.05$. t-test with Welch's correction.

A

Mouse cohort	Preneoplastic lesions
BLG-Cre Runx $^{WT/WT}$	0 (n=9)
BLG-Cre Runx1 $^{flx/flx}$	0 (n= 13)
BLG-Cre Runx2 $^{flx/flx}$	0 (n=9)
BLG-Cre Runx1 $^{flx/flx}$ Runx2 $^{flx/flx}$	5 (n=15)

B

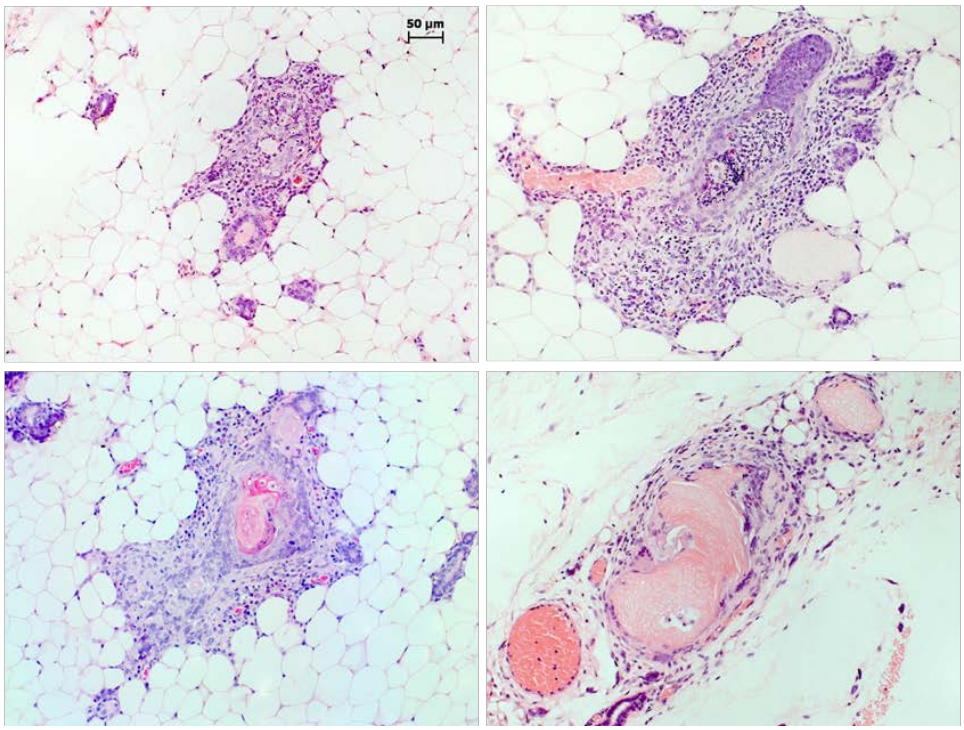


Figure 3-16 Double deletion of *Runx1* and *Runx2* causes pre-neoplastic lesions in virgin mice.

Table showing the incidence of preneoplastic lesions found in the 6 month aging cohort (number of mice with detectable preneoplastic lesions, total number assessed in brackets). Chi square test p=0.0089 (A). H&E stained examples of lesions developed by BLG-Cre:*Runx1*$^{Flx/Flx}$:*Runx2*$^{Flx/Flx}$ mice (B).

3.3 DISCUSSION

The results presented in this chapter represent an answer to the unmet need of better characterization of RUNX expression in normal mammary development. Increasing the basic understanding of the role of *Runx* genes in the regulation of normal mammary epithelial homeostasis will expand the knowledge of the molecular controls acting in this tissue and would hopefully translate into new possible biomarkers or targeted therapies for breast cancer treatment.

Studies on FACS purified mammary epithelial cells investigate for the first time expression of *Runx* genes in enriched populations of mammary epithelial cells. These findings demonstrate that Runx1 is the main expressed *Runx* transcript while *Runx2* levels are around 50 fold lower. Moreover *Runx3* transcript is not detectable in the purified mammary populations. Supporting this pattern of expression, IHC on mouse and human mammary epithelium showed high widespread levels of RUNX1 staining, while RUNX2 was expressed at lower levels in discrete subpopulations of luminal and basal cells. RUNX3 expression could not be assessed due to the lack of good specific antibodies for IHC. However given the absence of detectable *Runx3* RNA in the mammary epithelium it seems unlikely that any protein can be translated. Interestingly, confirmation of this pattern of expression comes from another study of qRT-PCR on MCF10A cells, a human mammary epithelial cell line derived from a basal-like fibrocystic disease, where a similar pattern of *Runx* expression with *RUNX*1 mRNA being 15 fold higher than *RUNX*2 and undetectable levels of *RUNX*3 was found (Wang et al. 2011). However, the lack of detectable *Runx3* expression in normal mammary epithelial cells is conflicting with some published observations. Firstly this differs with the expression pattern published by Blyth and colleagues (Blyth et al. 2010) which indicates that all three *Runx* genes are expressed in the adult mouse mammary gland. However this study was based on qRT-PCR analysis on whole mammary extracts where contamination with non-epithelial cells, like stromal, immune and endothelial cells would be expected. Secondly, some reports have indicated Runx3 as a tumour suppressor in breast cancer. In particular one study shows that about 20% of female *Runx3*(+/-) heterozygous mice spontaneously developed ductal carcinoma at an average age of 14.5 months (Huang et al. 2012). However the authors failed to provide convincing evidence of *Runx3* expression in normal mammary epithelium. Thus the

phenotype observed could be caused by the effects of *Runx3* haploinsufficiency on non-epithelial stromal components of the mammary gland.

Another interesting data coming from our analysis of FACS purified mammary populations is the finding that both *Runx1* and *Runx2* are enriched in the mammary basal population: a lineage which has been shown to be enriched in mammary stem cells (Shackleton et al. 2006). Moreover several proteins which have been found enriched in the basal population (such as SLUG, ß1-Integrin, LGR5) have also been found to play a role in the regulation of mammary stem cells (Plaks et al. 2013; Guo et al. 2012; Taddei et al. 2008). So it is tempting to speculate an involvement for *Runx1* and *Runx2* in the mammary stem cell population. The characterization of *Runx1* and *Runx2* in mammary stem cells will be discussed in the next chapter.

Our studies characterize for the first time RUNX1 and RUNX2 protein expression during mouse mammary gland development. Focusing on embryonic mammary development, expression of RUNX1 and RUNX2 seems to follow an opposite pattern of expression. Mouse embryonic mammary development starts around embryonic day 10 (E10) with the formation of the milk line, a slight thickening of the ectoderm which at E11.5 develops into individual placodes which sink deeper into the dermis (Robinson 2007). This phase of mammary embryonic development, between E10 and E13, is when the growing epithelium acquires its identity as mammary tissue (Wansbury et al. 2011). Interestingly RUNX2 epithelial expression is detectable specifically at E12 and is then turned off. This result indicates that RUNX2 could play a role in the early specification of the mammary lineage. Between E13 to E15, the bud undergo a period of relative quiescence, to then start proliferating again at E15.5 when the primary sprout starts invading the mammary mesenchyme towards the fat pad. At E18.5 the elongating duct has grown into a small rudimental ductal system (Cowin & Wysolmerski 2010). These last stages of mammary embryonic development, going from E16 to E18.5, are the ones characterized by RUNX1 expression. This pattern would suggest that RUNX1 is not involved in early mammary epithelium specification. Hence its expression could mark an embryonic population of mammary epithelial cells which have already acquired their lineage specifications. Further investigations are needed to investigate the functional role of RUNX1 and RUNX2 in mammary embryonic development.

Focusing on RUNX1 and RUNX2 protein expression in the adult mammary gland, RUNX1 IHC showed that this protein is widely expressed at high levels in the adult mammary epithelium. On the other hand, RUNX2 expression is confined to specific subpopulations of luminal and basal cells which are scattered in mammary epithelium. Moreover the RUNX2 positive population greatly varies between different areas of the same mammary gland and between mammary glands extracted from different mice, suggesting waves of expansion and contraction of RUNX2 positive cells. This dynamic pattern of expression is reminiscent of the cycling changes which characterize the mammary gland tissue during the oestrous cycle. Thus it is tempting to speculate that the spatial and temporal regulation of RUNX2 in adult mammary epithelium could be linked to the oestrous cycle: interestingly preliminary results from our laboratory seem to indicate an increase in RUNX2 expression during metestrous and diestrous (data not shown). This phase of the oestrous cycle is characterized by an increase in progesterone levels together with WNT-4 which are driving the branching expansion characteristic of that stage (Brisken et al. 2000). So RUNX2 expression could mark a highly dynamic progenitor subpopulation of the adult mammary epithelium which could undergo expansion in the second phase of the oestrous cycle and could participate in the formation of secondary and tertiary branching. Further experiments using ovariectomized mice treated with single injections of oestrogen and progesterone to induce a controlled oestrous cycle are needed to investigate the possible link between RUNX2 expression and oestrous cycle. Focusing on RUNX1 and RUNX2 expression during the lactational cycle, both genes were found to follow a similar pattern of expression, dropping during pregnancy and lactation to rise again in involution. Interestingly RUNX1 and RUNX2 expression are at the lowest during the highest differentiated phases of mammary gland development (pregnancy and lactation) where the majority of the mammary gland is comprised of alveolar cells. This data suggests that RUNX genes need to be turned off to allow a complete alveolar differentiation program. The *in vivo* proof of this concept can be found in the phenotype observed in transgenic mice overexpressing RUNX2 under a luminal promoter (MMTV-*Runx2*). These mice showed impaired lactation capability due to a block in alveolar differentiation as showed by lack of functional alveolar structures and terminal differentiation markers such as WAP and ß-Casein (McDonald et al, submitted).

To study the role of RUNX1 and RUNX2 in the mammary epithelium we adopted a combination of conditional RUNX knock-out models targeted to the mammary epithelium. No overall effects after deletion of both RUNX1 and RUNX2 were observed in the mammary epithelium. The lack of a gross detectable phenotype is not surprising considering that BLG-Cre is only acting in a small percentage of mammary epithelial cells at the virgin stage. Moreover Cre expression in this model peaks during pregnancy and lactation, stages where the maximum amount of recombination is achieved, but which are both characterized by lack of expression of RUNX proteins. To achieve a better recombination in the virgin mammary gland, our lab is currently generating conditional knock-out models using a different mammary specific Cre (MMTV-Cre) (Wagner et al. 2001). This model will help identify any possible gross defects after RUNX loss in the virgin stage. The use of a genetic marker (GFP expression) to trace the fate of RUNX-deleted cells *in vivo* allowed us to show for the first time the requirement for both RUNX1 and RUNX2 in normal mammary epithelial maintenance. In fact both loss of RUNX1 and RUNX2 caused a reduction of overall GFP levels in the virgin mammary gland. Moreover focusing on mammary populations, individual loss of either RUNX1 or RUNX2 caused a decrease in GFP positivity in the luminal compartment. However, only RUNX2 deletion caused a reduction in GFP+ cells in the basal layer, suggesting a possible specific role for RUNX2 in this population. The mechanism underlying the loss of RUNX-deleted cells is still unknown. However the most likely explanation could be induction of quiescence and block of the cell cycle after loss of *Runx* genes. RUNX proteins are well known regulators of the cell cycle in different systems: for example, *in vitro* studies showed that RUNX1 stimulates the G1 to S phase transition in haematopoietic cells (Strom et al. 2000) and RUNX2 acts as transcriptional repressor of the *p21Cip1/Waf1* cyclin-dependent kinase inhibitor (Westendorf et al. 2002). I*n vivo* studies also showed evidence of a link between RUNX and quiescence where in the adult skin RUNX1 promotes hair follicle stem cell proliferation through repression of negative cell cycle regulators such as *p21*, *p27*, *p57*, and *p15* (Lee et al. 2013). In our model, a block in the cell cycle would be sufficient to cause a decrease of GFP populations in long term maintenance. In particular while the RUNX wild-type cells will keep dividing and growing in number, the RUNX-deleted quiescent population will be slowly diluted down and the mammary gland will be repopulated by non-recombined cells. Transcriptional profiling of

sorted GFP+ populations from the *Runx* knock-out mice would help to clarify the molecular mechanisms underlying loss of RUNX-deleted cells in the virgin mammary gland. It should be acknowledged that the use of reporter mice to track the fate of conditional-deleted cells may bring some intrinsic flaws as shown by work from Anton Berns' lab. Here the existence of marked differences in the recombination frequencies of different loci within the same cell was demonstrated (Vooijs et al. 2001). A different frequency of recombination at the GFP locus compared to the *Runx* locus cannot be ruled out in the models presented in this chapter. To prove the efficiency of GFP reporter expression as direct readout for *in vivo* RUNX deletion, freshly sorted GFP+ populations could be tested for recombination at the RUNX locus by genomic PCR.

To avoid any compensatory effect of single *Runx* gene deficiency *in vivo*, a double *Runx*1-*Runx*2 knockout mouse was generated. This model is characterized by the appearance of gross phenotypes detectable both in mature virgins and aging cohorts indicating that compensation between RUNX isoforms is happening in the single knock-out models. The first striking phenotype of the double *Runx*1-*Runx*2 knockout mouse is the rescue of overall GFP positivity in *Runx*1-*Runx*2 knock-out glands: this is not caused by a rescue in the luminal population but it is caused by a 5 fold specific expansion of the basal population. Immunofluorescence showed that GFP+ basal cells maintain their normal localization as a single layer surrounding the duct suggesting that those cells are not transformed. This result indicates *Runx* genes as new key regulators of the basal population. However, the overall picture of the role of *Runx1* and *Runx2* in the mammary basal population is not clear: in fact both *Runx1* and *Runx2* are enriched in the basal lineage but only loss of *Runx2* leads to a decrease in the basal population. Moreover combined loss of *Runx1* and *Runx2* causes basal cell expansion. Those results could be explained taking in account the vast array of isoform specific cofactors which are required for RUNX-mediated transcriptional control. In speculation RUNX2, but not RUNX1, may be a key regulator of proliferation of a mammary basal subpopulation and it does so cooperating with an unknown transcriptional cofactor which we can call cofactor-X. In a normal gland RUNX2 and cofactor-X cooperate and control the growth of a basal subpopulation. After loss of RUNX1, as a compensatory effect, RUNX2, which is usually expressed at lower levels than RUNX1, is upregulated together with

cofactor-X. Then the complex RUNX2-cofactor-X compensates for RUNX1 loss keeping basal cell growth under control. In the scenario of loss of RUNX2 instead, cofactor-X will not be transcribed in basal cells. RUNX1 will then try to compensate RUNX2 loss but, since cofactor-X is not translated, RUNX1 will not be able to maintain normal basal cell homeostasis and it will drive a different transcriptional program resulting in a block of proliferation or induction of quiescence. In the case of combined loss of RUNX1 and RUNX2 instead, basal cells lose all key controllers of cell proliferation (RUNX2, cofactor-X and RUNX1) leading to an expansion and uncontrolled growth of that subpopulation. In another scenario, the luminal lineage could be the one responsible for the expanded basal subpopulation: combined loss of RUNX1 and RUNX2 could start a dedifferentiation program pushing luminal cells into the basal lineage. It will be interesting to investigate the actual mechanism for this phenotype further.

At 12 weeks of age double *Runx1-Runx2* knockout mice showed no gross abnormality. However more severe phenotypes started to appear when mice were aged for 6 months. The first detectable phenotype was an increased alveolar hyperplasia. Alveolar hyperplasia is the most common preneoplastic lesion in mouse mammary glands (Medina 2002), characterized by appearance of foci of hyperplastic lobuloalveolar development in a virgin mammary gland. Activation of several oncogenes such as *CycID1*, *Wnt1* and *Tnf,* have been found to cause alveolar hyperplasia in the mouse mammary gland. Moreover *Runx1-Runx2* knockout mice also develop rare preneoplastic lesions characterized by elements of necrosis, squamous metaplasia and extensive inflammation and stromal cell recruitment. The rarity of these lesions is not surprising considering the low amount of recombination achieved by BLG-Cre in the virgin mammary gland. The downstream molecular causes underlying alveolar hyperplasia and preneoplastic lesions development are currently under investigation. However the squamous metaplastic appearance of some of the lesions would suggest an involvement of *Wnt* signalling, whose activation has already been shown to cause extensive squamous metaplasia in the mammary gland (Miyoshi, Shillingford, et al. 2002). Further experiments will help to clarify the underlying mechanism of preneoplastic lesion formation after *Runx1* & *Runx2* loss. In particular fat pad transplantation of FACS sorted basal GFP+ cells from the double *Runx1-Runx2* knockout mice will test the ability of *Runx*-deleted cells to regenerate an entire

mammary epithelium and will clarify the oncogenic potential of this population. Moreover double *Runx1-Runx2* knockout mice will be crossed with transgenic lines carrying oncogenes (*Myc*, *Wnt*) targeted specifically to the mammary gland. These models will test the tumour suppressor/oncogenic role of *Runx* isoforms in the context of breast cancer and how different oncogenic signalling pathways can affect this behaviour.

4 A new role for RUNX2 in mammary stem cells

4.1 INTRODUCTION

4.1.1 *Runx* genes and stem cells.

An exciting role for RUNX proteins in stem cell biology has emerged recently from different studies using diverse animal models, from sea urchins to mammals (Braun, 2009). In invertebrate systems, which represent a good simplified model for the study of basic developmental processes, *Runx* genes are involved in several processes linked to stem cell regulation and homeostasis; in particular RUNX proteins have been found to play key roles in stem cell maintenance, lineage differentiation and organ regeneration in Planarians, *C.elegans* and *Drosophila* (Kagoshima et al. 2007; Braun & Woollard 2009; Wenemoser et al. 2012). Additional studies are unveiling a role for RUNX transcription factors in stem cell biology in different mammalian systems such as haematopoietic stem cells, hair follicle stem cells and mesenchymal stem cells (Wang et al. 2010). Several features make *Runx* genes an ideal hub for the control of stem cell homeostasis; *Runx* genes regulate transcription through interaction with a wide variety of co-repressors and co-activators which bring to this transcriptional network high flexibility and context dependency, features necessary for the regulation of dynamic entities such as stem cells (Chuang et al. 2012). Moreover RUNX transcription factors can control different signalling pathways such as *Wnt*, *Notch* and *Hedgehog*, which are all fundamental in stem cell biology (Coffman 2009). RUNX proteins can also interact with several chromatin-modifying proteins, such as acetyltransferases, helping to shape the stem cell epigenetic landscape (Lunyak & Rosenfeld 2008). Finally RUNX proteins can function as potent cell cycle regulators, activating and/or repressing cell proliferation and quiescence (Coffman 2003).

Mammary gland stem cells represent a poorly characterized population of the adult mammary gland characterized by the ability to differentiate into the multiple cell lineages which make up the mammary epithelium, and the capacity to self-renew in order to maintain a stable pool of tissue stem cells (Rosen 2012). Improving our knowledge of mammary stem cell biology is of pivotal importance for a deeper understanding of mammary gland and breast cancer development. Recently some evidence has emerged in the literature, suggesting a possible role for RUNX transcription factors in mammary stem cell biology. *Runx2* transcript is enriched in the mammary basal population, a compartment characterized by high stem cell content. In particular, microarray analysis on FACS sorted murine mammary cells showed that *Runx2* is enriched in the CD24+/Low *Sca1*- basal population (Kendrick et al. 2008). Interestingly, *Runx2* expression is also upregulated in the luminal ER negative population ($CD24^{+/high}$ $SCA1^{-}$) where the luminal-restricted progenitors are thought to reside (Kendrick et al. 2008). Moreover, qRT-PCR on the same populations confirmed that *Runx2* transcript is upregulated in the basal/myoepithelial and in the stem-cell enriched mammary population (Molyneux et al. 2010). Embryonic mammary development is characterized by actively proliferating stem cells which will generate all the different cell types and structures of the mature ductal epithelium. Intriguingly, work from the Howard lab identified *Runx1* as a key node of the transcriptional network which guides mammary embryonic development (Wansbury et al. 2011). Moreover, gene expression analysis on microdissected adult mammary tissue revealed that both *Runx1* and *Runx2* RNA are enriched in the terminal end buds, virgin-specific mammary structures which are thought to contain a high number of activated mammary stem cells (Kouros-Mehr & Werb 2006). *In vitro* experiments using HC11 cells, a murine mammary epithelial cell line, added some evidence to a possible involvement of *Runx* genes in mammary stem cell biology. HC11 cells have been widely used as a model to study changes in gene expression during the mammary stem cell differentiation process (Ball et al. 1988). *Runx2* transcript decreases during *in vitro* mammary differentiation in HC11 cells, suggestive of a possible role in stem cell maintenance (Williams, 2009). Several observations from our previous experiments also point to a role for RUNX proteins in mammary stem cells (previous chapter). First, we confirmed that not only *Runx*2 but also *Runx*1 is enriched in the mammary basal population. Second, both RUNX1 and RUNX2 are

expressed during embryonic mammary development, a stage characterized by a high degree of stemness and epithelial plasticity. Finally our conditional knock-out mouse model shows that loss of RUNX2 dramatically affects the basal cell population *in vivo*, indicating a possible key role for this transcription factor in this lineage.

4.1.2 RUNX, mammary stem cells and Wnt signalling.

Wnt signalling is a key regulator of stem cell differentiation and maintenance in different systems (Cadigan & Peifer 2009). Moreover the *Wnt* pathway has been shown to play a role in mammary gland development and in mammary stem cell regulation and maintenance (Zeng & Nusse 2010). Interestingly evidence coming from different experimental models links RUNX2 and *Wnt* signalling (see Introduction). Studies show that *Wnt* and RUNX2 follow similar expression patterns in the mammary gland while *Wnt* signalling pathway is enriched in the mammary basal population (E. Lim et al. 2010; van Amerongen et al. 2012), the same population which we have shown to be enriched in *Runx2*. Also, similar to RUNX2 expression, *Wnt* signalling is active until E15.5 in mammary embryonic epithelial cells and then is switched off (Chu et al. 2004).

4.1.3 Experimental Aims.

The aim of this study was to characterize RUNX1 and RUNX2 in mammary stem cell biology. In addition the interaction between *Wnt* signalling and RUNX2 in mammary stem cells has been investigated.

4.2 RESULTS

4.2.1 Expression of *Runx1* and *Runx2* in mammospheres.

To investigate a possible role for *Runx* genes in mammary stem cells, mammospheres were used as a tool for the generation of mammary stem cell-enriched cultures (see Introduction and Material and Methods for details). As a control, primary mammary cells were grown in differentiating conditions (from now on called 2D MMECs) as adherent 2D-culture on plastic with foetal calf serum (FCS) and a cocktail of growth factors (Insulin, EGF and Cholera toxin). Fresh-extracted MMECs, after one week in mammosphere-culturing conditions form floating spherical colonies as expected (FIGURE 4-1-A & B). First the pattern of expression of all three *Runx* genes was analysed by qRT-PCR to check if *in vitro* culture was affecting their normal pattern of expression. Again *Runx1* transcript represented the main isoform while *Runx3* levels were very low, close to the limit of detectability (FIGURE 4-1-C). This data confirmed our previous results on fresh sorted MMECs and showed that *in vitro* culturing conditions are not causing drastic changes in the pattern of expression of *Runx* genes. *Runx* expression was then assessed in primary cells grown in differentiating conditions and those grown as primary and secondary mammospheres. Since mammosphere passaging has been shown to enrich for mammary cells with stem cell features (Dontu et al. 2003), this experiment allowed to test a possible involvement of *Runx* genes in the MaSC population. Interestingly *Runx2* was enriched in primary and secondary mammospheres when compared to 2D MMECs (FIGURE 4-1-D). On the other hand *Runx1* showed no enrichment in primary mammospheres and a decrease in secondary mammospheres (FIGURE 4-1-E). Western blot analysis confirmed RUNX2 enrichment in primary mammospheres at the protein level (FIGURE 4-2-A). IHC analysis was used to characterize the pattern of RUNX2 expression in mammospheres. High RUNX2 expression in primary and secondary mammospheres was detectable in a subgroup of cells, usually located at the centre of the spheres (FIGURE 4-2-B & C). Together these results indicate a correlation between elevated RUNX2 expression and stem cell-enriched cultures.

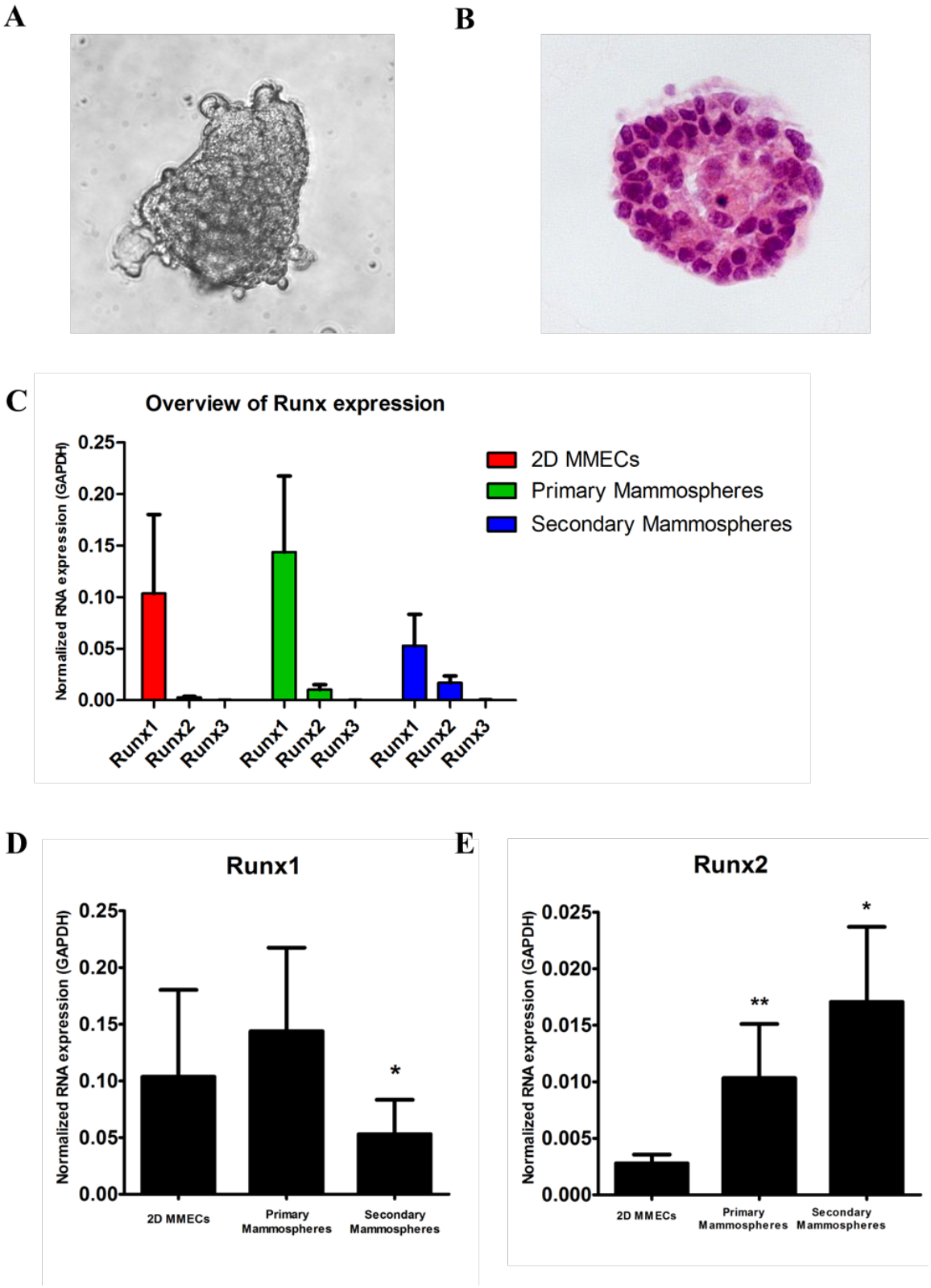


Figure 4-1 Characterization of *Runx* expression in mammospheres.

Bright field (A) and H&E (B) of primary mammospheres after 1 week of culture. qRT-PCR for *Runx1*, *Runx2* and *Runx3* on MMECs grown in 2D (2D MMECs), primary and secondary mammospheres (C-E). RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). n≥5 for each group. Expression in primary and secondary mammospheres was compared to 2D MMECs; in (E), *=p<0.001. **=p<0.0005. Expression in primary mammospheres was also compared to secondary mammospheres; in (D), p<0.02. In (E), p<0.05.

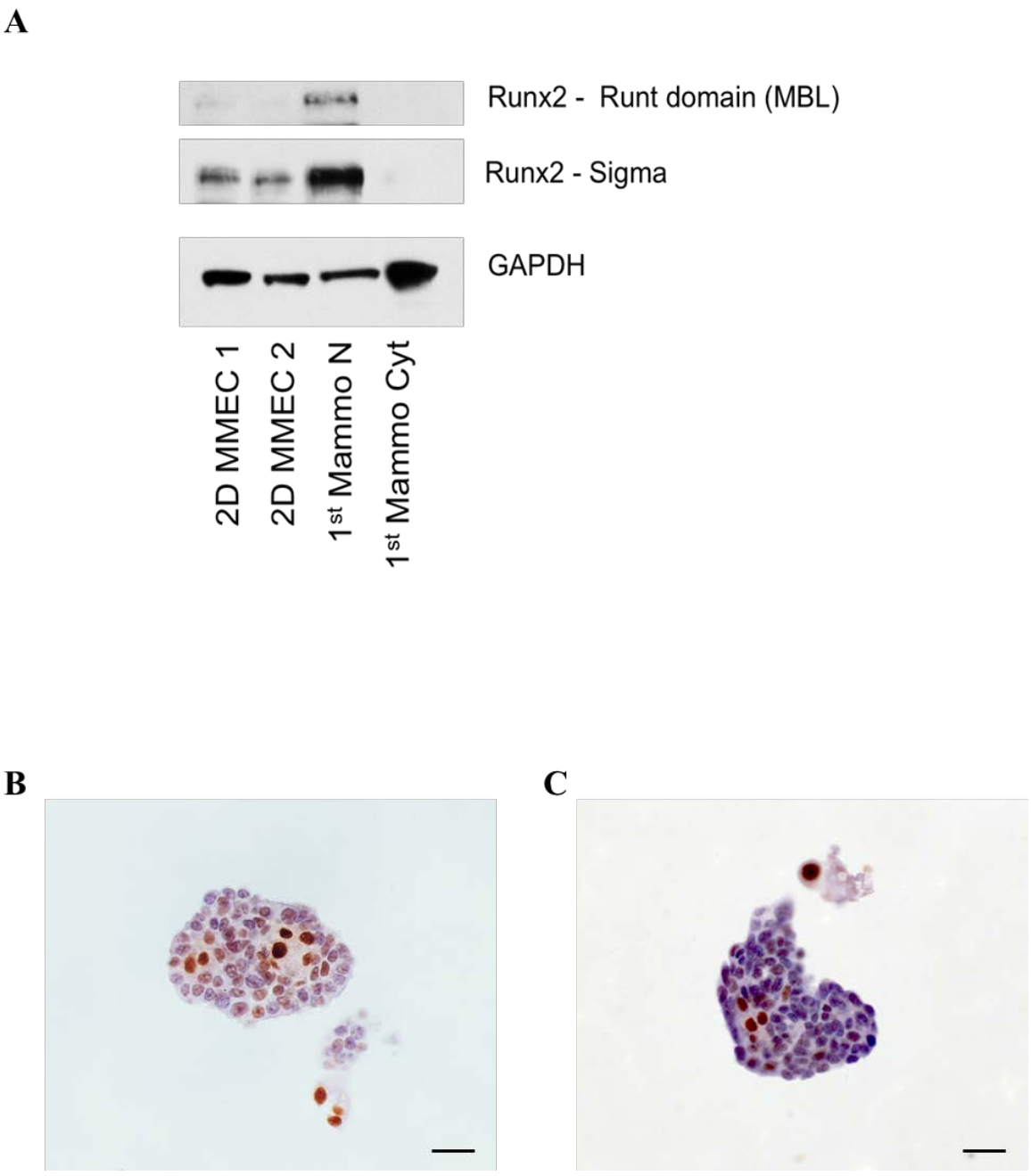


Figure 4-2 RUNX2 protein is enriched in mammospheres.

Western blot on MMECs grown in 2D from two separate experiments (2D MMEC1, 2D MMEC2) and primary mammospheres nuclear and cytoplasmic extract (1stMammo N, 1stMammo Cyt). Two different antibodies against RUNX2 were used as specified in the figure (A). RUNX2 IHC on primary (B) and secondary mammospheres (C). Scale bars represent 30 μm.

To test if RUNX2 is necessary for mammospheres formation a loss of function experiment was performed whereby MMECs extracted from Cre-negative *Runx2*$^{WT/WT}$ and *Runx2*$^{Flx/Flx}$ mice were infected with adenoviral Cre-GFP and grown as mammospheres. There was no difference in total number of primary and secondary mammospheres generated from *Runx2*$^{WT/WT}$ and *Runx2*$^{Flx/Flx}$ cells (FIGURE 4-3-A & B). However RT-PCR on RNA extracted from mammospheres derived from *Runx2*$^{Flx/Flx}$ cells showed that the partial *Runx*2 knock-down induced by Cre expression *in vitro* is rescued in secondary mammospheres (FIGURE 4-3-C & D). This result indicates that *Runx2*-deleted cells are selected against in this assay suggesting a functional role for *Runx2* in mammosphere formation.

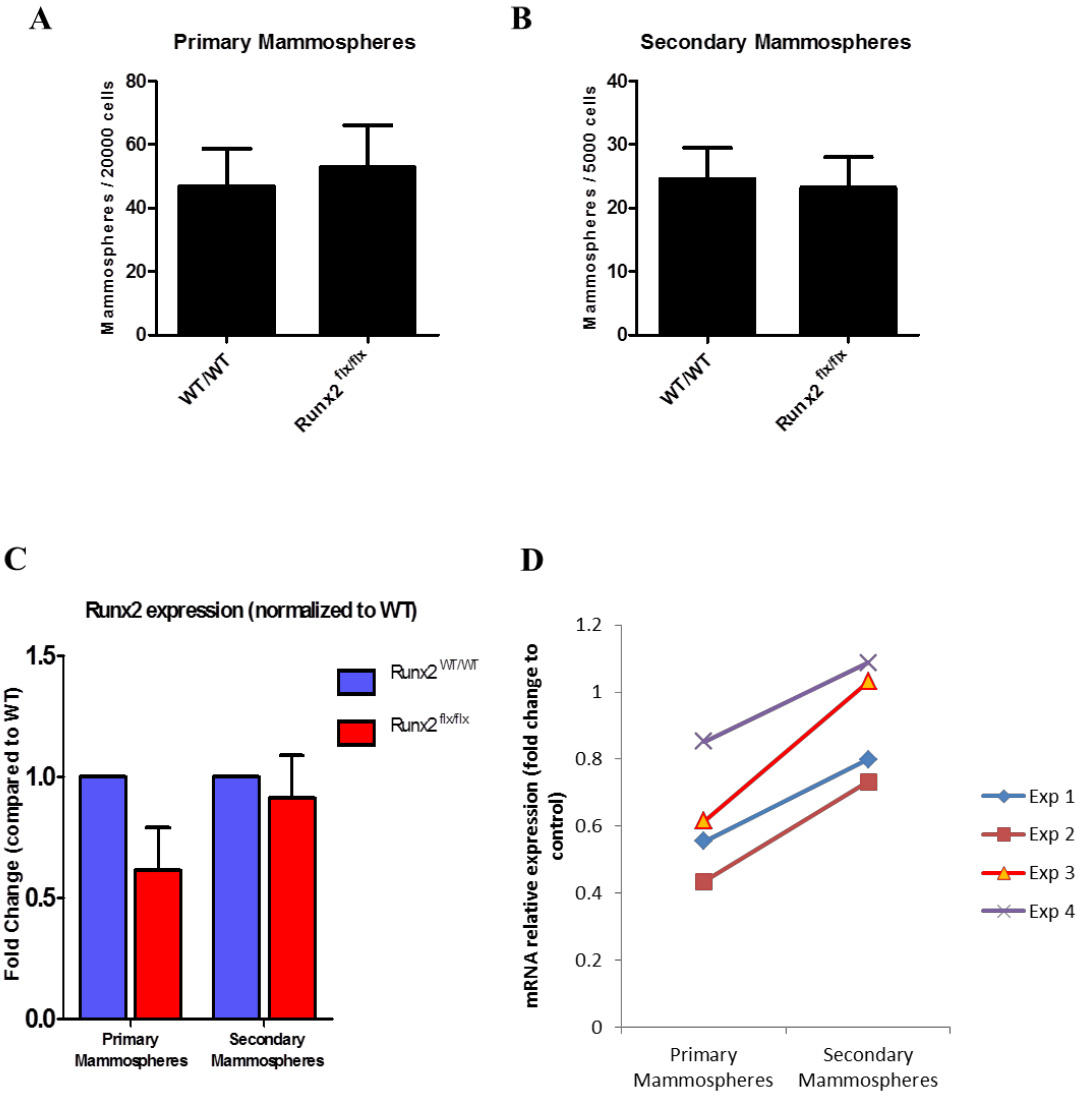


Figure 4-3 Effects of *Runx2* loss on mammospheres formation.

Quantification of primary (A) and secondary (B) mammospheres on MMECs extracted from *Runx2*$^{WT/WT}$ and *Runx2*$^{Flx/Flx}$ mice and infected with Adenoviral Cre. MMECs were spin-infected with adenoviral Cre-GFP and grown in non-adherent conditions as mammospheres. Primary and secondary mammospheres were counted after 7 days in culture under a bright field microscope. *Runx2* mRNA expression from primary and secondary mammospheres (C) *Runx2* expression is shown as fold change compared to the WT control. Data are expressed as mean fold expression (±SD), average of 4 independent experiments. Chart showing rescue of *Runx2* expression in 4 independent experiments (D).

4.2.2 *In vivo* deletion of *Runx2* in the mammary basal compartment does not affect normal mammary development.

To investigate the role of *Runx2* in mammary stem cells *in vivo,* a loss of function mouse model targeted to the mammary basal population was created. Here K14-Cre mice, which express the Cre recombinase under the control of the human keratin 14 promoter, were used. Under this promoter, specific deletion of *loxP* flanked target sequences can be achieved in the basal layer of different epithelial tissues such as skin, tongue and mammary gland (Dassule et al. 2000). To achieve a targeted deletion of *Runx2* in the basal population of the mammary epithelium the K14-Cre mouse was combined together with the *Runx2*$^{flx/flx}$ mouse. No obvious phenotype was detectable in K14-Cre+/*Runx2*$^{flx/flx}$ glands at the histological level in mature virgin, lactating and involuting stages when compared to K14-Cre+/*Runx2*$^{WT/WT}$ controls (FIGURE 4-4-A). The lactating capability of K14-Cre+/*Runx2*$^{flx/flx}$ mice was also tested by weighing litters at 7 days of age. No difference in weight was found between pups nursed by K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{flx/flx}$ females indicating that loss of *Runx2* in the basal layer of the mammary gland is not affecting milk production (FIGURE 4-4-B). The fate of K14-Cre/*Runx2* deleted cells in the mammary epithelium was then tracked using the GFP reporter as previously described (Chapter 3). FACS analysis on MMECs extracted from mature virgins showed no difference in total GFP expression levels between K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{flx/flx}$ mice (FIGURE 4-5-A). Interestingly the levels of GFP expression driven by K14-Cre (~40% of total MMECs) are much higher compared to BLG-Cre (~16% of total MMECs, see figure 3-8) indicating that K14-Cre is highly active in the mature virgin state. To identify possible aberrations in mammary lineages induced by loss of *Runx2* in the basal population, mammary population profiling through FACS looking at GFP expression specifically in basal and luminal lineages was carried out. Loss of *Runx2* causes no effects on GFP levels in either the basal or luminal population (FIGURE 4-5-B & C). FACS analysis confirmed that K14-Cre is expressed at high levels in the basal layer of the mammary gland: however a considerable amount of GFP+ cells were also detected in the luminal compartment (FIGURE 4-5-B). GFP expression in luminal cells was confirmed through immunofluorescence showing double positive GFP+/CK18+ cells in K14-Cre Z/EG+ mice (FIGURE 4-6).

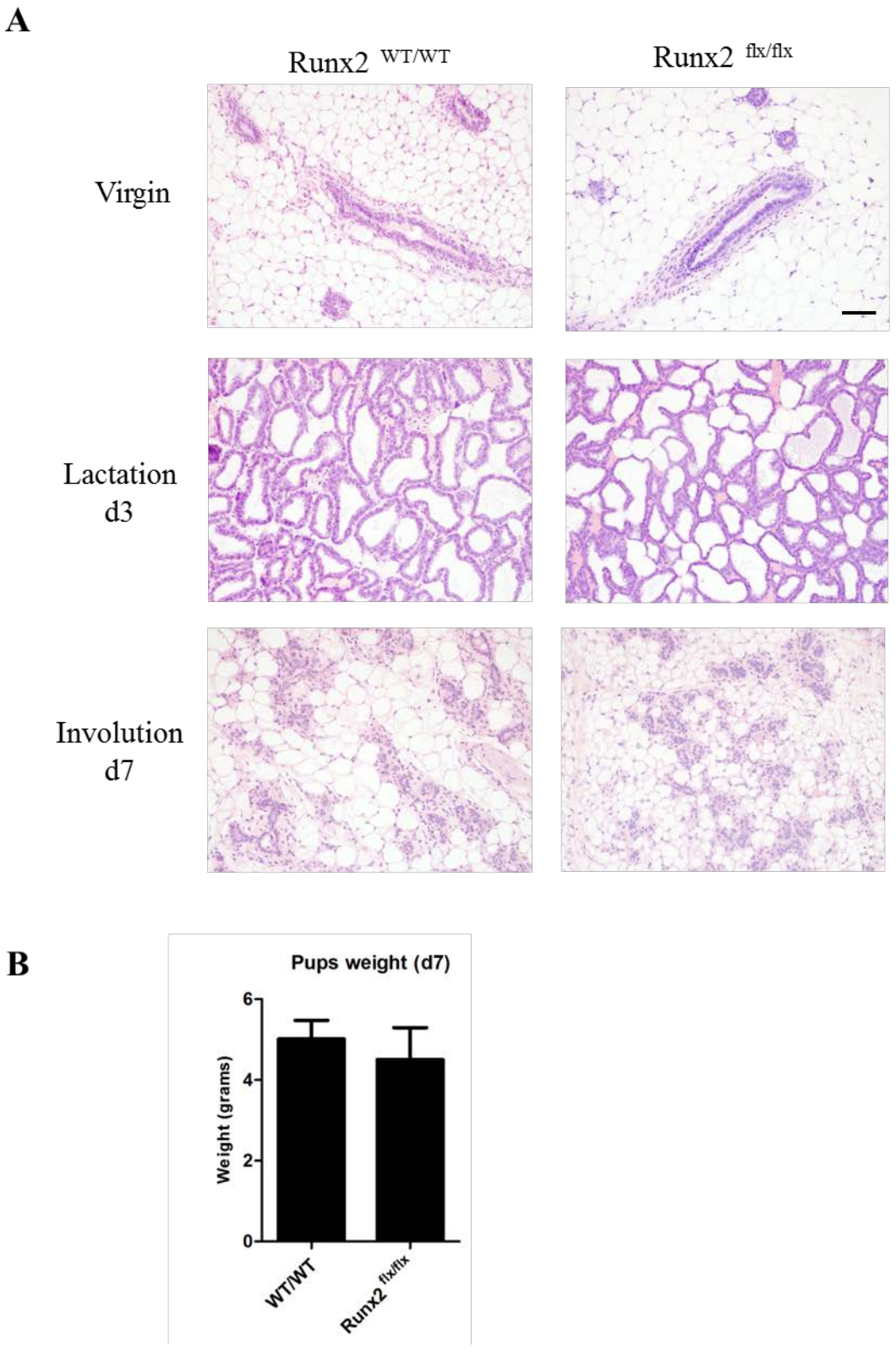


Figure 4-4 Effects of loss of *Runx2* in the basal layer of the virgin mammary gland.

Representative histological sections of mammary glands from K14-Cre+/*Runx*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice extracted from mice at 12 week old virgin, lactation day 3 and involution day7 (n>4 for each stage). Scale bar represents 50 μm (A). Bar chart showing the weight of pups from K14-Cre+/*Runx*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ litters expressed in grams. Litters (n≥5) were normalized to 6 and pups were weighed after 7 days from the start of lactation (B).

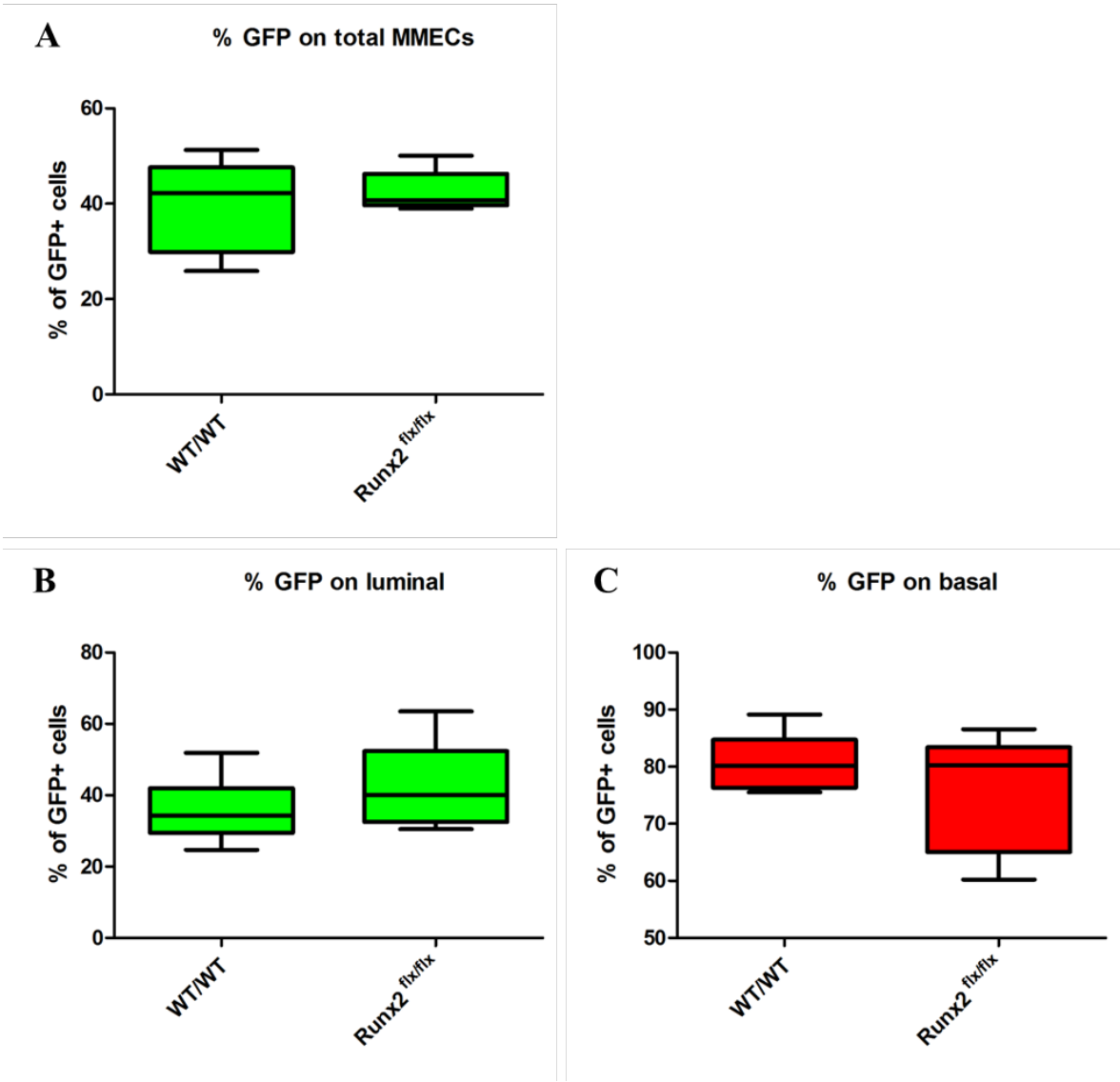


Figure 4-5 Analysis of *Runx2* loss in the basal lineage using a GFP reporter.

Reporter GFP expression as analysed by FACS, on whole mammary gland extraction (A), and gated on CD24+/CD29^{low} luminal population (B) and CD24+/CD29^{high} basal population (C). MMECs were isolated from 12 week K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice. Data are expressed as % of GFP+ cells over total in the selected population (±SD). n≥5 for each stage.

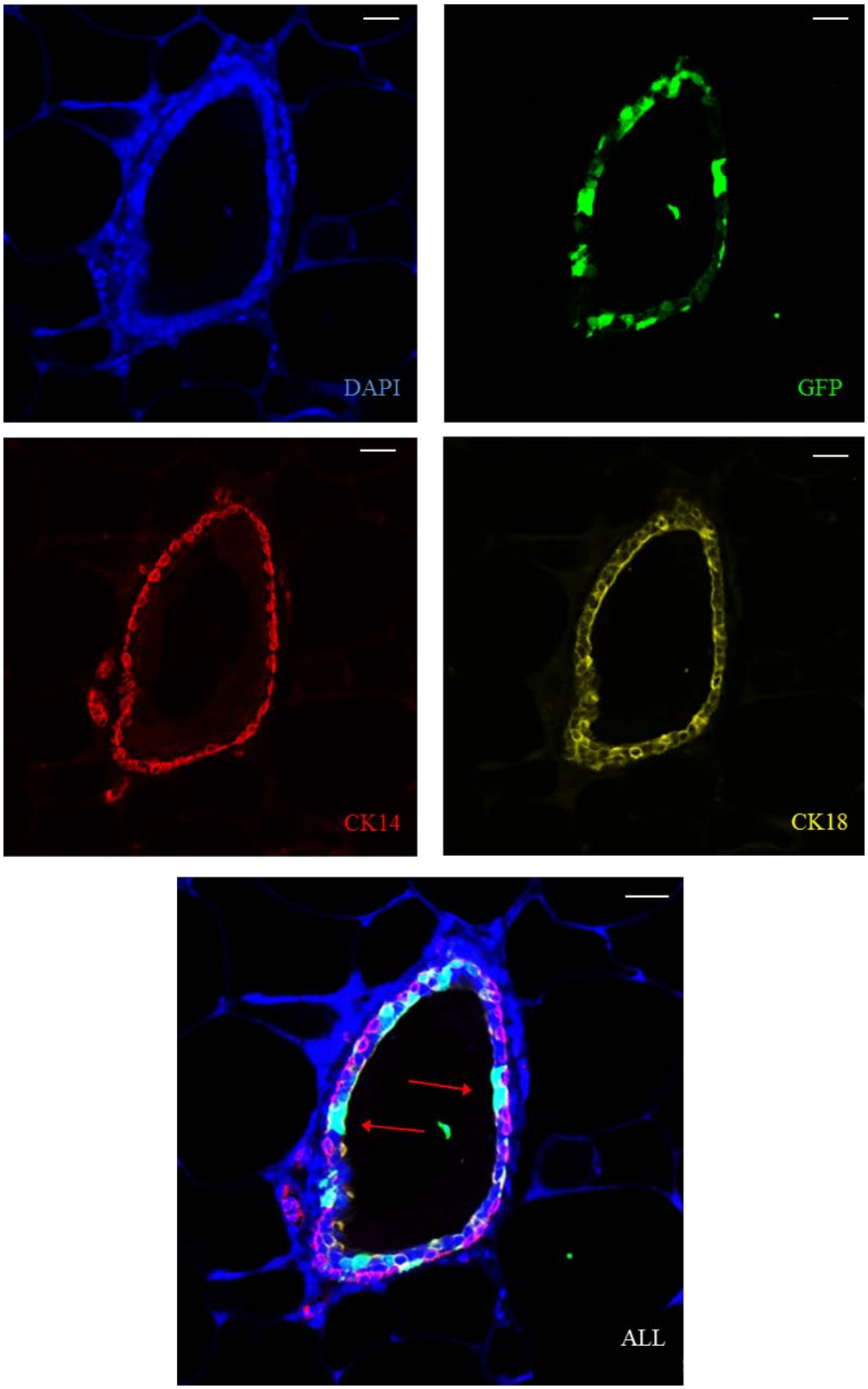


Figure 4-6 K14-Cre is expressed in the luminal population of the virgin mammary gland.

Immunofluorescence of CK14 (red), CK18 (yellow) and GFP (green) on mammary glands extracted from K14-Cre+/*Runx2*$^{WT/WT}$ mice. Red arrows indicate GFP+ luminal cells on the merge image. Glands were counterstained with DAPI (blue). Scale bar represents 20 μm.

4.2.3 *In vivo* deletion of *Runx2* in the K14+ population impairs *in vitro* MMECs regenerative potential.

Overall *Runx2* loss with a K14-Cre driven promoter, and high recombination in the basal lineage, causes no detectable phenotypes in the mammary gland. However loss of genes involved in stem cells and regeneration can often result in no overall phenotype during normal development, with the defect becoming apparent only after challenging the recombinant tissue through regenerative assays (Taddei et al. 2008). To test the regenerative potential of *Runx2* deleted mammary epithelial cells *in vitro* stem cell assays were carried out. In fresh extracted MMECs from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice, loss of *Runx2* caused a significant reduction in number and size of primary mammospheres (FIGURE 4-7-A-C). In addition secondary mammosphere growth was also affected (FIGURE 4-7-D & E). qRT-PCR on RNA extracted from primary and secondary mammospheres confirmed *Runx2* deletion in K14-Cre+/*Runx2*$^{flx/flx}$ derived spheres while *Runx1* levels were not affected (FIGURE 4-8-A & B). Considering that *Runx* genes are key regulators of cell cycle progression (Coffman 2009), RNA levels of *p21*, a cyclin-dependent kinase inhibitor, were checked to test if loss of *Runx2* was affecting mammospheres proliferation status. qRT-PCR for *p21* on RNA extracted from primary mammospheres from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice showed that *p21* levels are reduced after *Runx2* deletion (FIGURE 4-8-C).

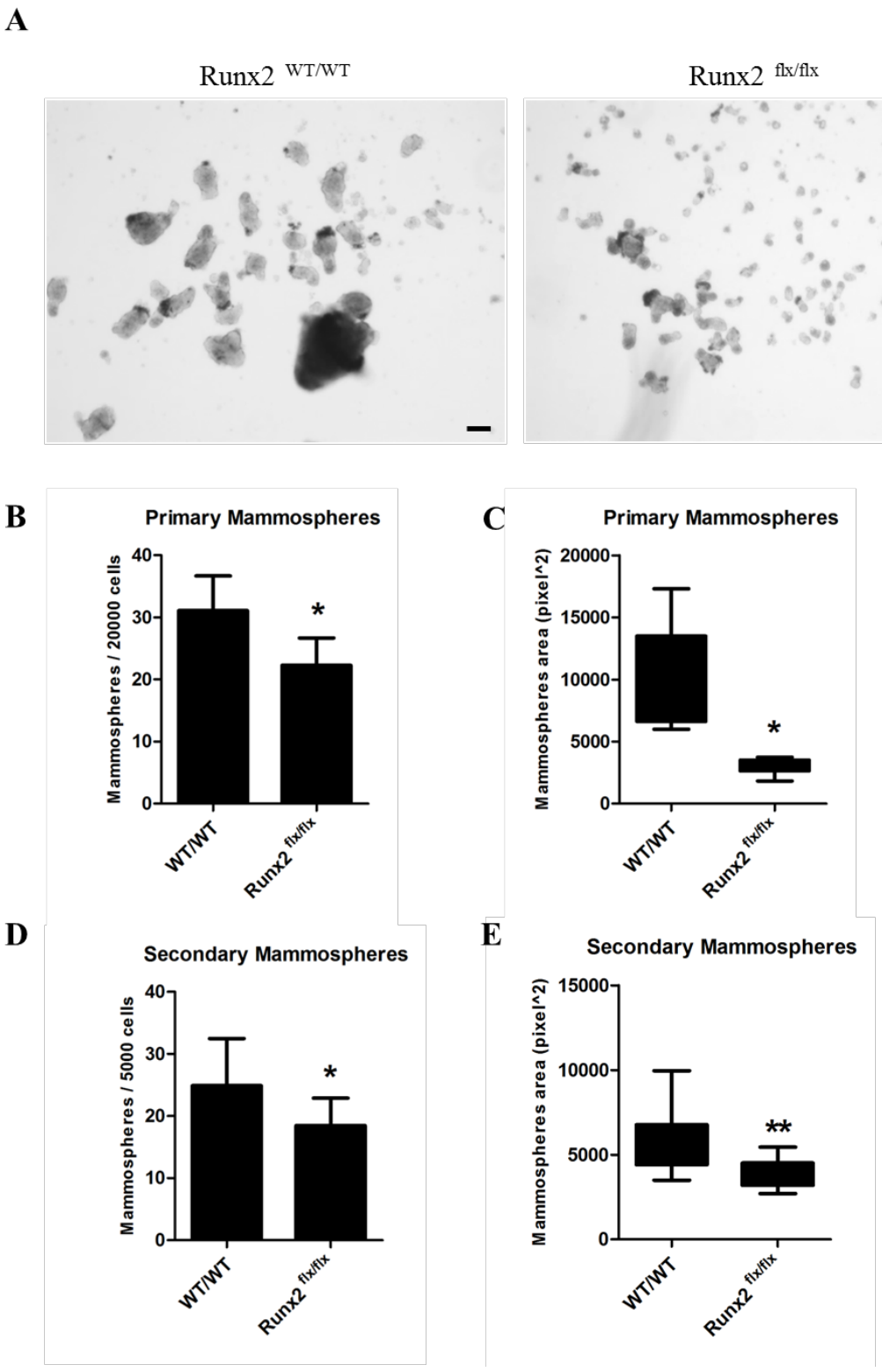


Figure 4-7 *In vivo* loss of *Runx2* in the basal lineage impairs mammosphere formation.

Bright field images of mammosphere cultures derived from K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice. Scale bars represent 100 µm (A). Quantification and size of primary (B-C) and secondary (D-E) mammospheres on MMECs extracted from K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice. Primary and secondary mammospheres were counted and measured after 7 days in culture. Data are expressed as mean (±SD). Four independent experiments for each group. *= $p<0.001$; **=$p<0.01$.

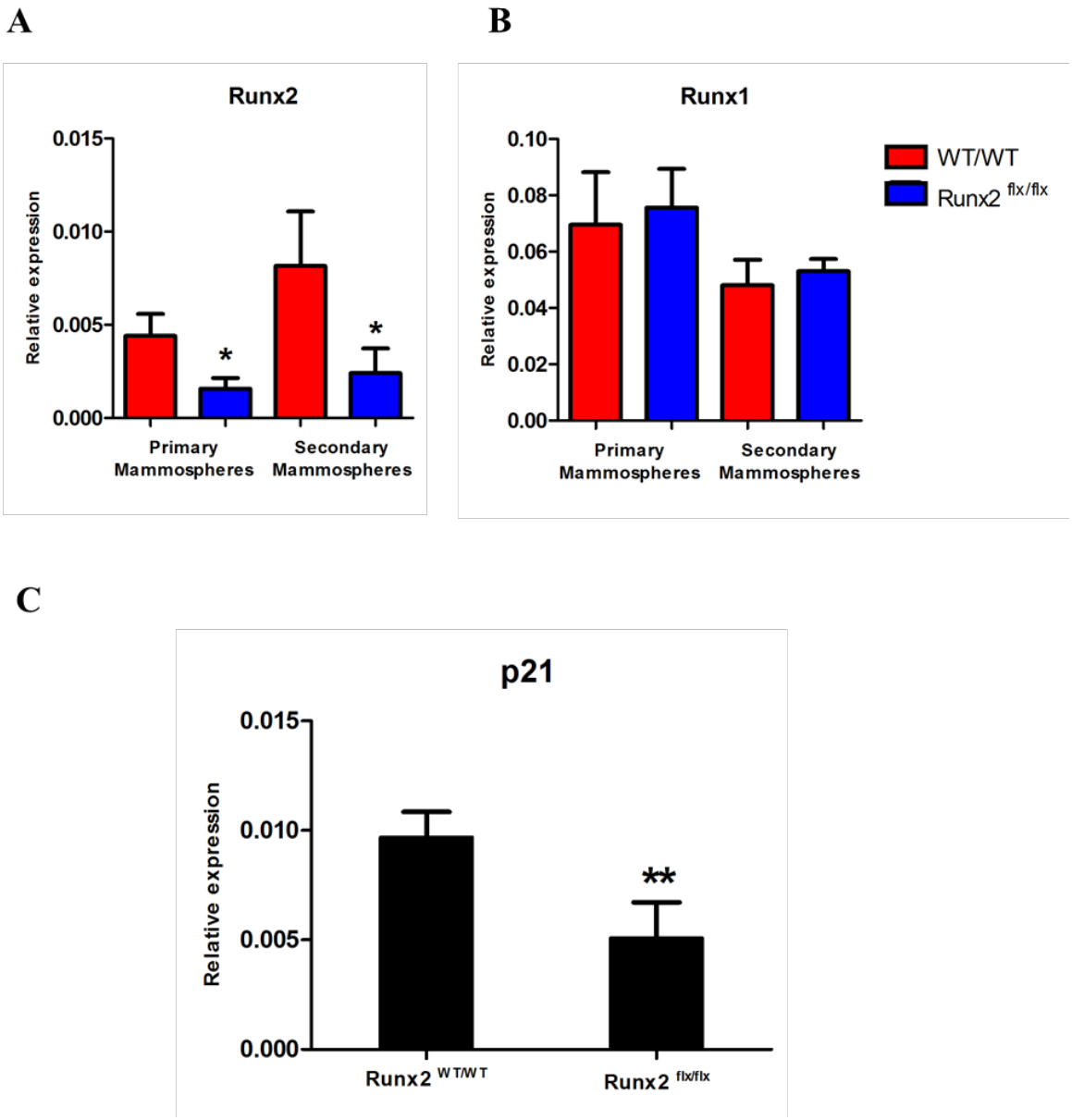


Figure 4-8 *Runx1*, *Runx2* and *p21* expression on mammospheres derived from K14-Cre+/*Runx2*$^{flx/flx}$ mice.

qRT-PCR for *Runx2* (A) and *Runx1* (B) mRNA expression in primary and secondary mammospheres derived from K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice. *p21* mRNA expression on primary mammospheres derived from K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice (C). RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). n=4 for each group. *=p<0.05. **=p<0.01.

In vitro clonogenic assays are widely used in the stem cell field as a surrogate to identify putative stem/progenitor cells (Grompe 2012). To further characterize the role of *Runx2* in mammary stem cells, the self-renewal potential of *Runx2* deleted cells *in vitro* using colony-forming assays in Matrigel (Guo et al. 2012) was performed. Specifically, fresh MMECs were extracted and seeded as single cells in Matrigel with serum-free medium supplemented with epidermal growth factor (EGF) (Zeng & Nusse 2010). In this assay, after one week in culture, MMECs form 3D colonies which develop into a ring of cells surrounding a hollow lumen, a structure resembling a mammary duct tree (FIGURE 4-9-A & B). These colonies can be further dissociated into single cells and plated again to form secondary colonies. *Runx1* and *Runx2* expression was assessed in this system. RT-PCR on primary and secondary colonies grown from wild type mice confirmed that *Runx1* is the most expressed of the *Runx* genes while *Runx3* transcripts are undetectable (FIGURE 4-9-C). Moreover, in a nice parallel with findings in the mammospheres system, *Runx2* is enriched in secondary colonies while *Runx1* expression is not (FIGURE 4-9-D & E).

Fresh extracted MMECs from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice were then assessed in matrigel colony-forming assays. Interestingly MMECs extracted from K14-Cre+/*Runx2*$^{flx/flx}$ mice formed less primary and secondary colonies when compared to controls (FIGURE 4-10-A-C). No difference was detectable in the size of matrigel colonies between the two groups (FIGURE 4-10-D,E). *Runx2* deletion was confirmed by RT-PCR on RNA extracted from primary matrigel cultures (FIGURE 4-11-A-B). Again *Runx1* levels were not affected indicating the specificity of the targeted deletion approach. These results indicate that *Runx2* expression in the K14+ cells of the mammary epithelium is required for sustaining MMECs growth in matrigel colony-forming assays.

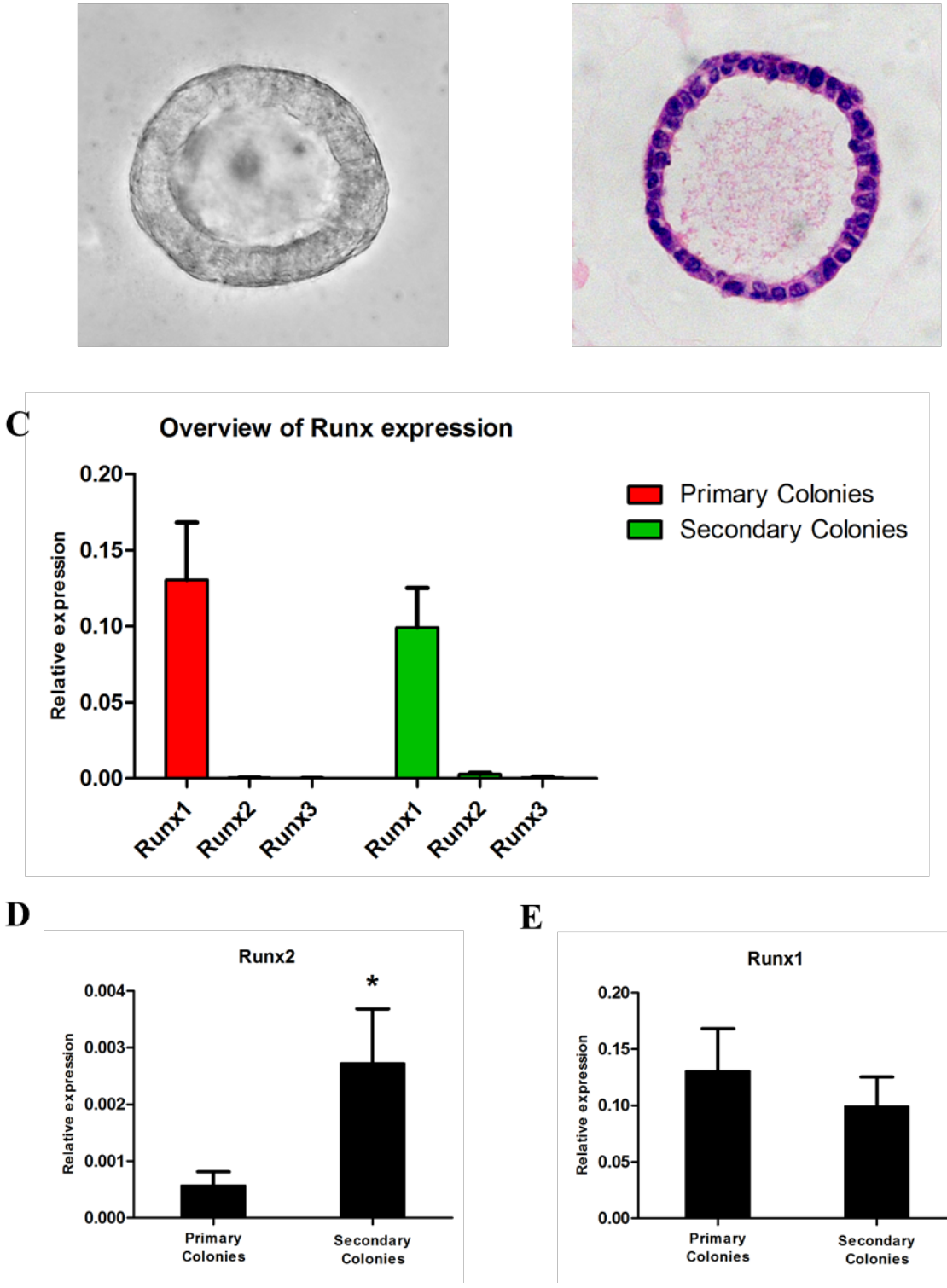


Figure 4-9 *Runx* expression in matrigel colony-forming assays.

Bright field (A) and H&E (B) of primary colonies after 1 week in culture. qRT-PCR for *Runx1*, *Runx2* and *Runx3* on primary and secondary matrigel colonies (C). qRT-PCR for *Runx2* (D) and *Runx1* (E) on primary and secondary matrigel colonies. RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). n=4 for each group. *=p<0.001.

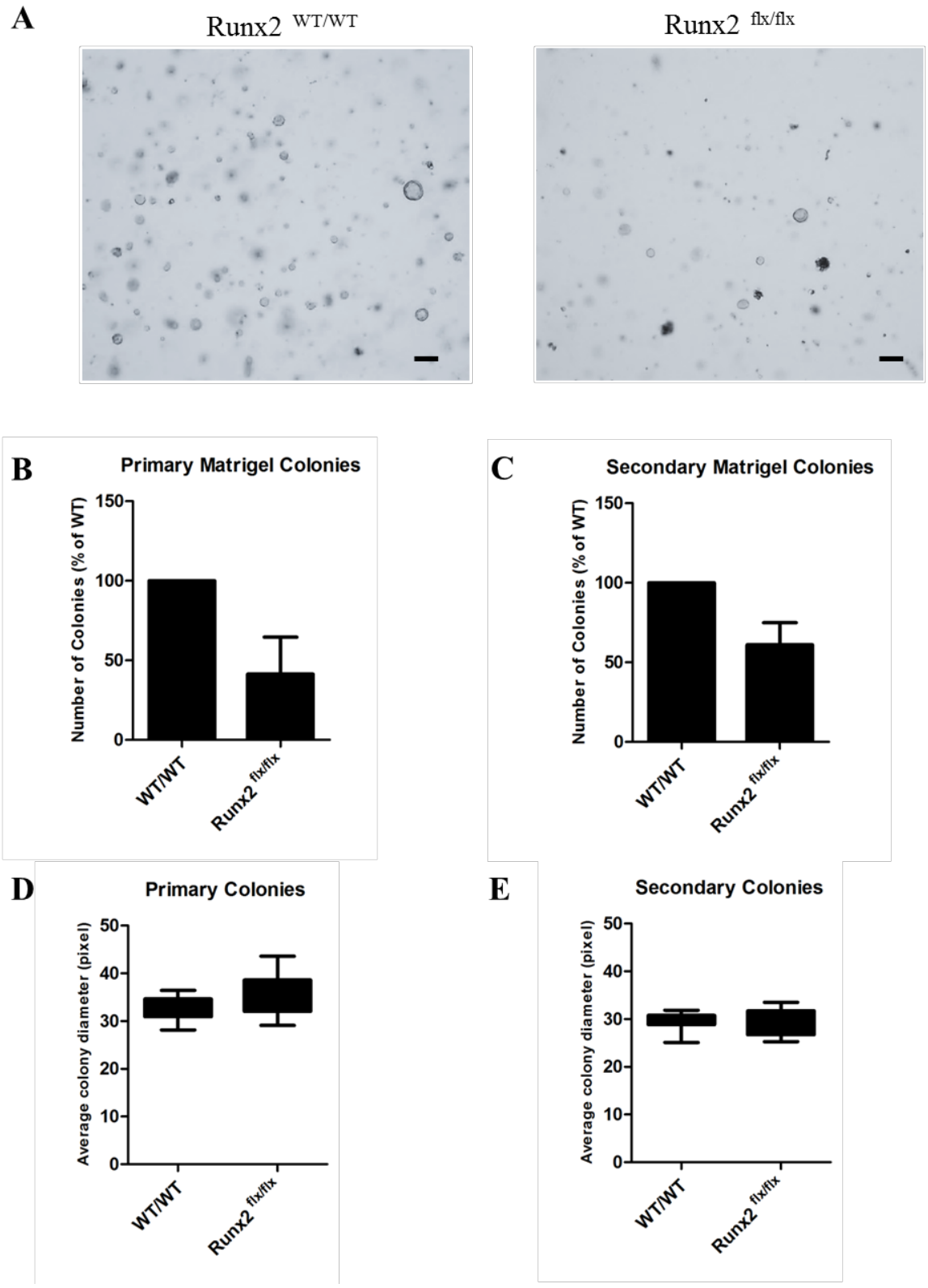


Figure 4-10 ***In vivo*** **loss of** ***Runx2*** **in the K14+ cells impairs matrigel colony formation.**

Bright field images of primary matrigel colonies derived from K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice. Scale bars represent 100 µm (A). Quantification and size of primary (B&D) and secondary (C&E) matrigel colonies from K14-Cre+/*Runx2* $^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ MMECs. Primary and secondary matrigel colonies were counted and measured after 7 days in culture. Data are expressed as mean number of colonies (% over the WT) and mean colony diameter (±SD). Four independent experiments for each group.

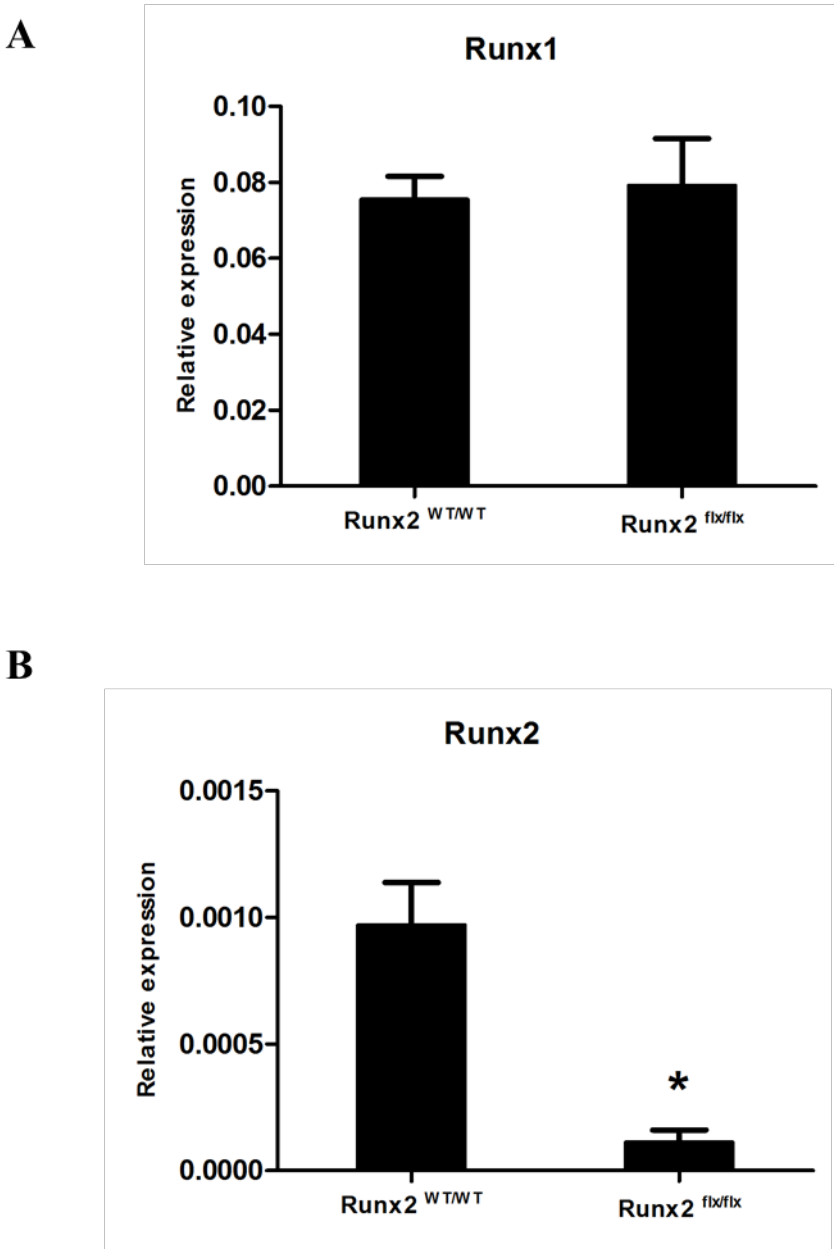


Figure 4-11 *Runx1* and *Runx2* expression on K14-Cre+/*Runx2*$^{flx/flx}$ matrigel colonies.

qRT-PCR of *Runx1* (A) and *Runx2* (B) expression on primary matrigel colonies derived from K14-Cre+/*Runx2*$^{WT/WT}$ and K14-Cre+/*Runx2*$^{Flx/Flx}$ mice after 7 days in culture. RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). n=4 for each group. *=p<0.005.

To determine if these *in vitro* findings were transferable to an *in vivo* setting and to test if *Runx2* is indeed involved in the regulation of mammary stem cell regenerative potential, the cleared fat pad transplantation assay was performed. MMECs were extracted from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice bearing the Z/EG transgene. Two different concentrations of cells ($1x10^4$ and $1x10^5$ MMECs) were injected into the cleared fat pad of 3 week old SCID female mice. After 7 weeks epithelial duct trees were detectable at the same frequency in both control and *Runx2* knock-out groups (FIGURE 4-12-A). Moreover *in vivo* GFP analysis indicated that *Runx2* deleted cells contributed to the ductal outgrowth in a similar manner to the WT (FIGURE 4-12-B). This data indicates that *Runx2* deleted cells can contribute to mammary epithelial regeneration when injected together with WT cells.

A

Number of transplanted cells	K14Cre/Runx2$^{WT/WT}$	K14Cre/Runx2$^{flx/flx}$
1 x 10^5	3/4	4/5
1 x 10^4	4/5	4/5

B

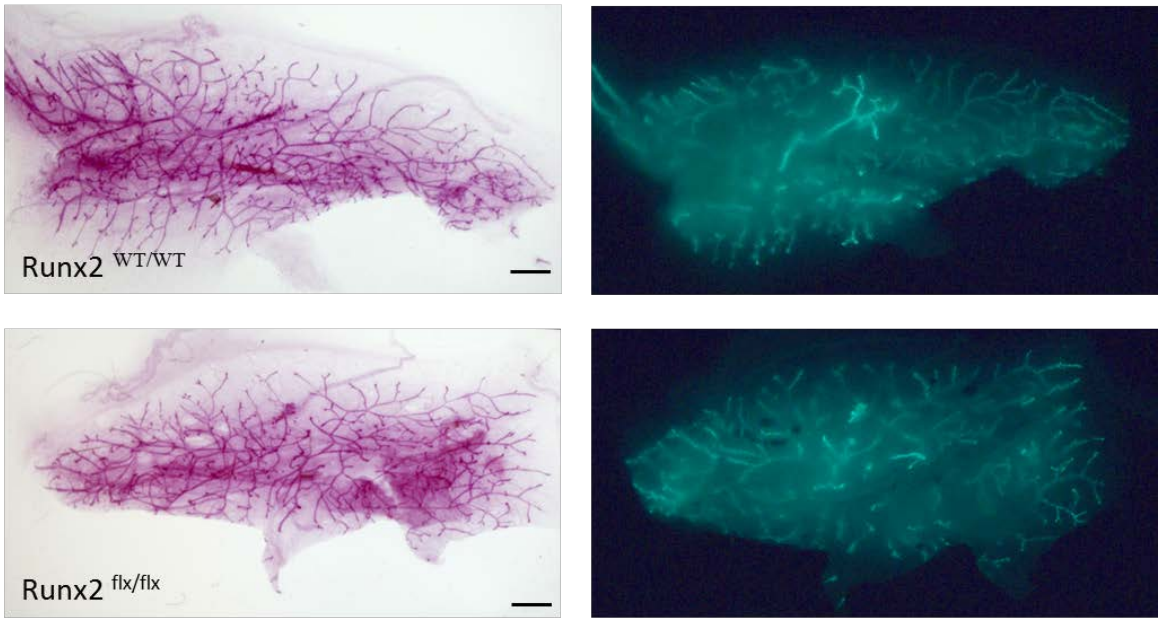


Figure 4-12 ***In vivo*** **regeneration activity of** ***Runx2*** **deleted mammary cells.**

Table showing the results obtained from cleared fat pad transplantation of total MMECs extracted from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice bearing the Z/EG transgene. Freshly isolated total MMECs were transplanted at the two concentrations specified above (A). Examples of reconstituted glands from the two cohorts: whole mounts (left) and *in vivo* GFP imaging of glands extracted from K14-Cre+/*Runx2*$^{WT/WT}$ (top) and K14-Cre+/*Runx2*$^{flx/flx}$ (bottom) mice. Scale bars represent 1mm (B).

4.2.4 *Runx2* is a downstream target of Wnt signalling in mammary stem cells.

The evidence showed so far suggests that RUNX2 is important for the regenerative potential of a mammary stem/progenitor population. Since *Wnt* signalling has been shown to regulate mammary stem cells and evidence links *Runx* and *Wnt* signalling pathways (see Introduction) the role of *Runx2* as a possible mediator or regulator of *Wnt* signalling in mammary stem cells was investigated. *Wnt* signalling status was checked in mammosphere culture using *Axin2* expression, one of the main downstream targets of *Wnt* signalling (Jho et al. 2002), as a readout for *Wnt* pathway activation. RT-PCR showed increased levels of *Axin2* in mammospheres culture compared to MMECs grown in differentiating conditions (FIGURE 4-13-A). This result indicates that *Wnt* signalling is more active in a population enriched in mammary stem cells. The effects of *Wnt* signalling activation on the mammosphere-forming capacity of MMECs was then tested. Treatment of fresh extracted MMECs grown as mammospheres with recombinant WNT3A caused an increase in mammosphere number and size (FIGURE 4-13-B-D). These data suggest a growth promoting effect of *Wnt* pathway on mammary stem cells. *Runx2* as a downstream target of *Wnt* signalling in MMECs was tested by performing a 24h treatment with recombinant WNT3A on MMECs grown in 2D differentiating conditions and on MMECs grown as mammospheres. After incubation with the recombinant protein, qRT-PCR for *Axin2* was used as a readout to confirm activation of *Wnt* pathway in both systems (FIGURE 4-14-A,B). Interestingly WNT3A treatment did not affect *Runx2* transcript levels in MMECs grown in differentiating conditions (FIGURE 4-14-C). However WNT3A treatment caused a significant up-regulation of *Runx2* expression in mammospheres (FIGURE 4-14-D). To confirm the specificity of that up-regulation levels of *Slug*, another transcription factor involved in mammary stem cells, were unchanged (FIGURE 4-14-E). Moreover RT-PCR for *Hes1*, one of the main transcriptional targets of Notch pathway, showed no difference, confirming the specific activation of *Wnt* signalling (FIGURE 4-14-F). These results are indicating *Runx2* as a potential downstream target of *Wnt* signalling specifically in mammary stem cell-enriched cultures.

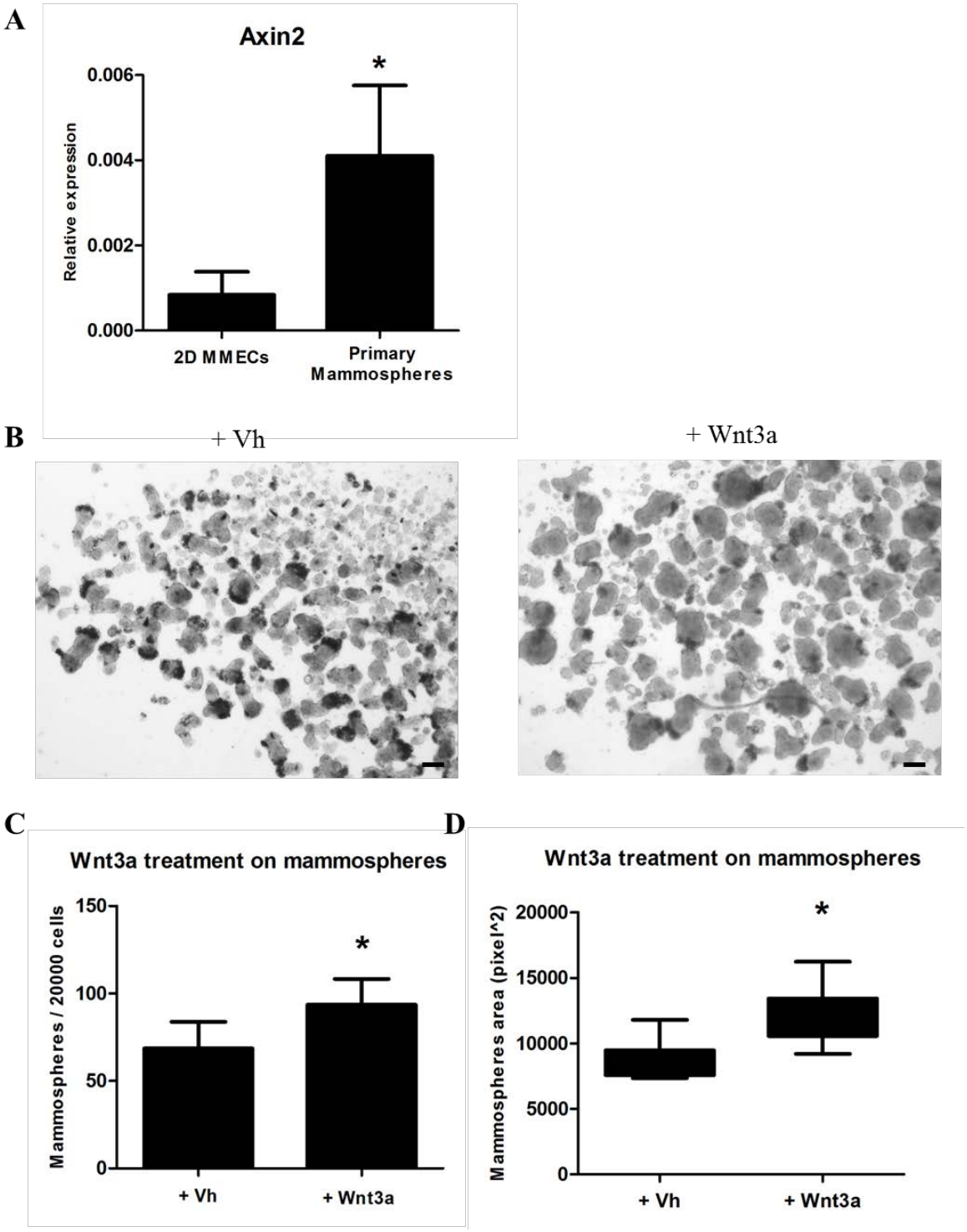


Figure 4-13 Effects of *Wnt* signalling on mammosphere cultures.

qRT-PCR for *Axin2* of MMECs grown in 2D (2D MMECs) and primary mammospheres. RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). n≥5 for each group (A). Bright field images of mammosphere cultures treated for 1 week, with either vehicle (Vh) or WNT3A (B). Quantification (C) and size (D) of primary mammospheres extracted from 12 week old virgin FVB mice and treated for 1 week, with either vehicle (Vh) or WNT3A. Data are expressed as mean (±SD). 3 independent MMECs extractions for each group. *=p<0.005.

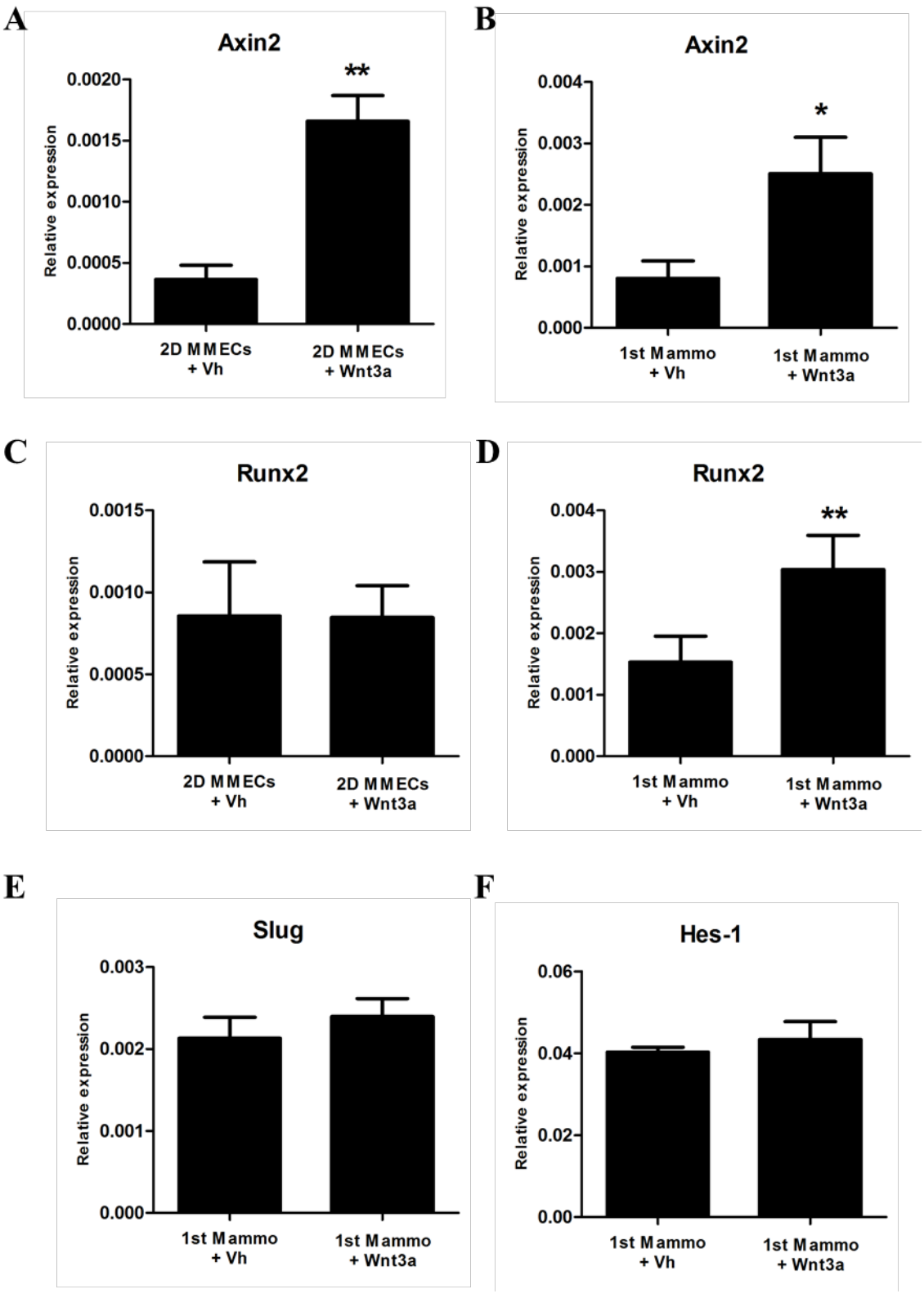


Figure 4-14 Effects of Wnt signalling activation on *Runx2* expression in mammospheres.

qRT-PCR for *Axin2* (A,B) and *Runx2* (C,D) on MMECs grown in 2D (2D MMECs) and primary mammospheres treated for 24h, with either vehicle (Vh) or WNT3A. qRT-PCR for *Slug* (E) and *Hes1* (F) on primary mammospheres treated for 24h, with either vehicle (Vh) or WNT3A. RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). 3 independent MMECs extractions for each group.*=p<0.05. **=p<0.0005.

4.3 DISCUSSION

In this study mammosphere culture was used as a tool to study the role of *Runx* genes in the mammary stem cell population. Several studies have shown a link between mammospheres and MaSC: in particular cleared fat pad injection of a single sphere is able to generate mammary outgrowth in 11 out of 18 transplantations, indicating that enrichment for cells with high regenerative capability is achievable in this assay (Cicalese et al. 2009). Moreover, the progenitor/stem cell content of mammospheres has been shown to increase with passage. Indeed whereas freshly isolated mammary cells contain less than 10% multi-lineage progenitors when cultured under differentiating conditions, these are enriched to 70% in primary mammospheres, and virtually 100% in secondary mammospheres (Dontu et al. 2003). Results presented here showed that *Runx2* expression is enriched in primary and further enriched in secondary mammosphere cultures. This data indicates that *Runx2* expression increases together with the stem cell content of mammospheres. Characterization of RUNX2 expression through immunohistochemistry showed that RUNX2 is expressed at high levels in a few cells usually located at the centre of the sphere. Intriguingly work from Wicha's lab showed that cytokeratin 5 (CK5) and CD49f, two known markers for the basal stem cell population, followed a similar pattern of expression (Dontu et al. 2003). Confirming these observations in a different *in vitro* system, *Runx2* is also upregulated in secondary matrigel colonies. These results unveiled a correlation between cells characterized by a certain degree of self-renewal *in vitro* and upregulation of *Runx2* expression. Interestingly this upregulation is specific for *Runx2* since *Runx1* showed no sign of enrichment in either mammosphere or matrigel cultures. Rather, *Runx1* was significantly downregulated in secondary mammospheres suggesting a possible role for this transcription factor in a more differentiated population of the mammary epithelium. The specific enrichment for *Runx2* but not *Runx1* in mammosphere cultures is also fitting with the data coming from the BLG-Cre/*Runx2*$^{flx/flx}$ cohorts indicating that only loss of *Runx2* (and not *Runx1*) perturbs the basal compartment. Taken together, the experiments above show a correlation between high *Runx2* expression and stem cell enriched cultures. To investigate the functional role of *Runx2* in mammospheres *in vitro, Runx2* knock-

out experiments, using adenoviral infection were performed. These experiments showed that *Runx2* is required for mammospheres formation since *Runx2*-deleted cells undergo a negative selection in the mammosphere assay.

Given the correlative and functional link discovered between *Runx2* and stem cells and since stem cells are enriched in the basal population we decided to create an *in vivo* deletion of *Runx2* specific for the basal lineage using the K14-Cre system. K14-Cre/*Runx2*$^{flx/flx}$ mice didn't show any phenotype at the histological level in all the tested developmental stages. This lack of phenotype could be due to a sufficient number of wild-type cells still present in the *Runx2*-deleted gland which can compensate for *Runx2* loss of function. Moreover lineage profiling with GFP marker showed no alterations in the mammary populations after *Runx*2 deletion indicating that loss of *Runx2 per se* in the basal compartment is not deleterious. This is in apparent contrast with the reduction of GFP+ basal cells seen after *Runx2* loss in the BLG-Cre model. However this could be explained by the fact that *Runx2* is expressed only in a small subpopulation of basal cells while K14-Cre is expressed in the majority of them. Hence the lack of a detectable reduction in the number of basal cells in the K14-Cre/*Runx2*$^{flx/flx}$ mouse could be due to the inherent lack of sensitivity of this system; in other words the reduction in GFP+ cells which may be caused by loss of *Runx2* in the basal layer would be smaller than the biological variation of K14-Cre expression. In the BLG-Cre model instead, the small percentage of basal cells that are targeted and which could represent the RUNX2 positive basal population, allows us to detect their loss after RUNX2 deletion.

Although no gross phenotype was detectable in normal mammary development of K14-Cre/*Runx2*$^{flx/flx}$ mice, interesting results started to appear when those cells were challenged in *in vitro* stem cell assays. MMECs extracted from K14-Cre/*Runx2*$^{flx/flx}$ mice generated fewer colonies in the mammosphere assay. This is showing that *Runx2* expression in the mammary basal compartment is required for the mammosphere generating potential of the virgin mammary gland. Moreover a reduction in colony size suggests that *Runx2* deleted colonies are somewhat arrested in development. Two possible scenarios could explain the reduced size and number of colonies observed; firstly, *Runx2* loss forces MaSC into quiescence hindering mammosphere formation. Or in fact *Runx2* deleted stem cells may be unable to enter quiescence; this would cause stem cell

exhaustion *in vivo* in the K14-Cre+/*Runx2*$^{flx/flx}$ mammary epithelium. Hence MMECs extracted from the *Runx2* deleted glands will be depleted of stem cells and this will cause the defective mammospheres forming capability. Preliminary results showing reduced levels of *p21*, a key regulator of quiescence (Cheng et al. 2000), in *Runx2* deleted mammospheres cultures support the second scenario. Furthermore, evidence in the literature suggests a role for *Runx2* in the induction of quiescence. For example *Runx2* is upregulated during quiescence in osteoblasts and *Runx2* can induce quiescence in prostate cancer cells (Galindo et al. 2005; Baniwal et al. 2010). Further experiments such as cell cycle profiling through FACS of primary mammospheres from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice will help to confirm if loss of *Runx2* is affecting MaSC cell cycle.

To further test if *Runx2* could affect the regenerative potential of mammary cells in a different *in vitro* culture model, 3D matrigel colony forming assay was performed. This assay is commonly used in mammary stem cell research by several labs for the study of MaSC regenerative potential *in vitro* (Guo et al. 2012; Zeng & Nusse 2010; Bai & Rohrschneider 2010). Interestingly results from these experiments confirmed that loss of *Runx2* impairs *in vitro* mammary regenerative potential causing a reduction in the number of colonies formed after one week in culture. How can loss of *Runx2* cause a reduction of regenerative potential in *in vitro* experiments without affecting normal mammary development? Looking into the literature this situation appears to be a common theme of different studies focusing on the role of stem cell genes in mammary gland and skin: no obvious gross defects are detectable during normal development with a phenotype becoming apparent only when cells are tested in regenerative assays *in vitro* and *in vivo* (Taddei et al. 2008; Jensen et al. 2009). This behaviour would suggest that, at least in some tissues, normal epithelial adult homeostasis does not require activation of the stem cell population, whose regenerative potential is instead necessary in more extreme conditions which can be physiological, such as tissue repair after wounding, or experimentally induced, such as stem cell assays. Besides forcing stem cell regenerative potential, i*n vitro* assays also serve as a good tool to eliminate confounding and redundant effects of the tissue microenvironment on the targeted population which could result in the rescue of the phenotype. Overall *in vitro* assays

represent a way to simplify *in vivo* systems, unmasking phenotypes which would be otherwise hidden by the complexity of tissue compensatory systems.

However *in vitro* assays also come with flaws; cells are grown in artificial conditions, no microenvironment is present and the structures formed are only reminiscent of the real complex architecture of a mature mammary epithelium. To prove if *Runx2* is really needed for the generation of a functional mammary duct tree *in vivo*, a cleared fat pad transplantation experiment was required. Epithelial duct trees were detectable at the same frequency in both control and *Runx2* knock-out groups showing that *Runx2*-deleted cells can take part in the regeneration process and be incorporated in the newly formed duct tree. However this experiment does not definitively prove that *Runx2* is not essential for mammary regenerative potential as the cells injected were not FACS purified so they represented a mixed population of *Runx2* wild-type and *Runx2* deleted cells. Thus the duct tree reconstitution capability of the *Runx2* deleted group could be due to the regenerative potential of the non-recombined population with *Runx2*-deleted cells being passively incorporated in the duct tree. Serial fat pad transplantation experiments using a pure population of FACS sorted GFP+ basal cells extracted from K14-Cre+/*Runx2*$^{flx/flx}$ and K14-Cre+/*Runx2*$^{WT/WT}$ mice will clarify if *Runx2* is absolutely required for the regenerative potential of mammary stem cells.

The fat pad transplantation technique also comes with limitations: in this assay MMECs undergo several stresses (tissue dissociation, enzymatic digestion, FACS sorting) which could affect the MaSC population. Moreover recent lineage tracing studies showed that fat pad transplantation can force cells to acquire stem cell behaviours which they do not normally have in physiological conditions (Van Keymeulen et al. 2011; van Amerongen et al. 2012). The final experiment which would show if *Runx2* is really involved in a mammary stem population would be lineage tracing using a tetracycline-inducible GFP, driven by the endogenous *Runx2* promoter. This would allow a precise localization and characterization of RUNX2 positive cells in the mammary epithelium. Using pulse and chase experiments at different stages of development (embryonic, virgin, pregnancy) would allow the clarification of the actual contribution of RUNX2 positive cells to mammary development and homeostasis (van Amerongen et al. 2012; Van Keymeulen et al. 2011).

The signalling mechanism by which *Runx2* can exert its effects on stem cell regulation has also been investigated. With evidence linking *Runx2* with *Wnt* signalling (Dong et al. 2006; Gaur et al. 2005) this seemed to be a promising area to investigate. First it was shown that *Wnt* signalling is activated in mammospheres compared to 2D-MMECs and that *Wnt* treatment stimulates the growth of mammospheres. These results are in line with the published evidence indicating *Wnt* signalling as a key regulator of MaSC (Zeng & Nusse 2010). Moreover, treating mammospheres for 24h with recombinant WNT3A induces *Runx2* transcription indicating that *Runx2* is a downstream target of *Wnt* signalling. The specificity of *Runx2* activation was confirmed by looking at *Slug*, another gene involved in MaSC which did not show any change after *Wnt* treatment. Furthermore no effects on *Runx2* expression were seen on differentiating 2D MMECs suggesting that *Runx2* is downstream to *Wnt* signalling only in MaSC enriched cultures. Probably MaSC cells are characterized by a different epigenetic conformation or a specific set of coactivators which allows the increase of *Runx2* expression in response to *Wnt* signalling.

5 The role of RUNX2 in metaplastic squamous breast cancer.

5.1 INTRODUCTION

5.1.1 Runx2 in breast cancer

An oncogenic role for RUNX2 was first reported in the haematopoietic lineage, where RUNX2 overexpression was shown to promote lymphoma development in mice (Stewart et al. 1997; Vaillant et al. 1999). In addition several studies have also suggested a possible role for RUNX2 as an oncogene in metastatic breast and prostate cancer (Pratap et al. 2006; Blyth et al. 2010). In particular some human breast cancer cell lines overexpress RUNX2 and its inhibition leads to less invasive potential *in vitro* and to decreased osteolytic metastatic activity *in vivo* (Barnes et al. 2003). Furthermore RUNX2 regulates several genes that are involved in breast cancer metastasis such as collagenase3, bone sialoprotein, osteopontin and MMP-9 (Pratap et al. 2005; Barnes et al. 2003). More hints regarding a role for RUNX transcription factors in the mammary gland are emerging from studies using three-dimensional epithelial culture systems. Interestingly RUNX2 overexpression in a 3D culture model using the MCF10A cell line disrupts normal acini development resulting in increased cell proliferation and hyperplasia (Pratap et al. 2009). However all the published studies to date have analysed the effects of RUNX2 in an *in vitro* context, using a limited number of cell lines. Hence *in vivo* evidence of the pro-tumourigenic role of RUNX2 in breast cancer is still lacking. A study from our laboratory is the first to provide *in vivo* confirmation of an oncogenic role for RUNX2 in breast cancer. McDonald *et al* showed that *in vivo* overexpression of *Runx2* caused hyperplasia and induces DCIS in old virgin mice (McDonald, Ferrari *et al*, submitted). Only a few studies have looked at RUNX2 in human breast cancer with contrasting results (Onodera et al. 2010; Das et al. 2009; Khalid et al. 2008). Unpublished

data from our laboratory show that RUNX2 is highly expressed in a small percentage of human breast cancers and interestingly its expression correlates with triple negative disease. Triple negative breast cancers, which account for around 15% of all BCs, are a mixed group of tumours defined by lack of expression of oestrogen, progesterone, and ERBB2 receptors. This subgroup is associated with poor outcome and lack of specific targeted therapy (Foulkes et al. 2010).

5.1.2 Experimental Aims

A few previous reports (our lab included) have indicated RUNX2 as a new oncogene in a specific subgroup of human breast cancer, the triple negative subtype. The aim of this study was to better characterize the role of RUNX2 in triple negative breast cancer through a combination of *in vitro* and *in vivo* modelling.

5.2 RESULTS

5.2.1 Runx2 expression in mouse models of breast cancer

RUNX2 is highly expressed in a small percentage of human breast cancers and interestingly its expression correlates with triple negative tumours (McDonald et al, unpublished). These results indicate that the role of RUNX2 in breast cancer could be restricted to a particular subgroup of the triple negative subtype. To better identify the breast cancer subgroup in which RUNX2 may play a role, six mouse models representing different breast cancer subtypes (Herschkowitz et al. 2007) were stained for RUNX2. Luminal-like models (MMTV-*PyMT* and MMTV-*Her2*) as well as a basal-like BRCA1 tumour model (BLG-*Cre*/*Brca1*$^{flx/flx}$/*p53*$^{+/-}$) showed low or negative RUNX2 staining (FIGURE 5-1). Conversely, *Apc*^{1572T} and BLG-*Cre*/*Pten*$^{flx/flx}$*Apc*$^{flx/flx}$ mice, two mouse models of metaplastic breast cancer, a rare subtype of triple negative BC (Gaspar et al. 2009), showed high RUNX2 positivity, especially in correlation with squamous metaplastic lesions (FIGURE 5-1). In addition, in the *Apc*^{1572T} model, RUNX2 expression is very high in the basal layer of the squamous metaplastic lesions which has been shown to contain a high percentage of the Ki67+ cells (Kuraguchi, 2009) (FIGURE 5-2). Intriguingly, both *Apc*^{1572T} and BLG-*Cre*/*Pten*$^{flx/flx}$*Apc*$^{flx/flx}$ models, are characterized by truncating mutations or deletions in the *Apc* tumour suppressor gene, the main negative regulator of the *Wnt*/*β-catenin* pathway (Cadigan & Peifer 2009). Thus both models are characterized by the constitutive activation of canonical *Wnt* signalling, suggesting a link between *Runx2* and this pathway.

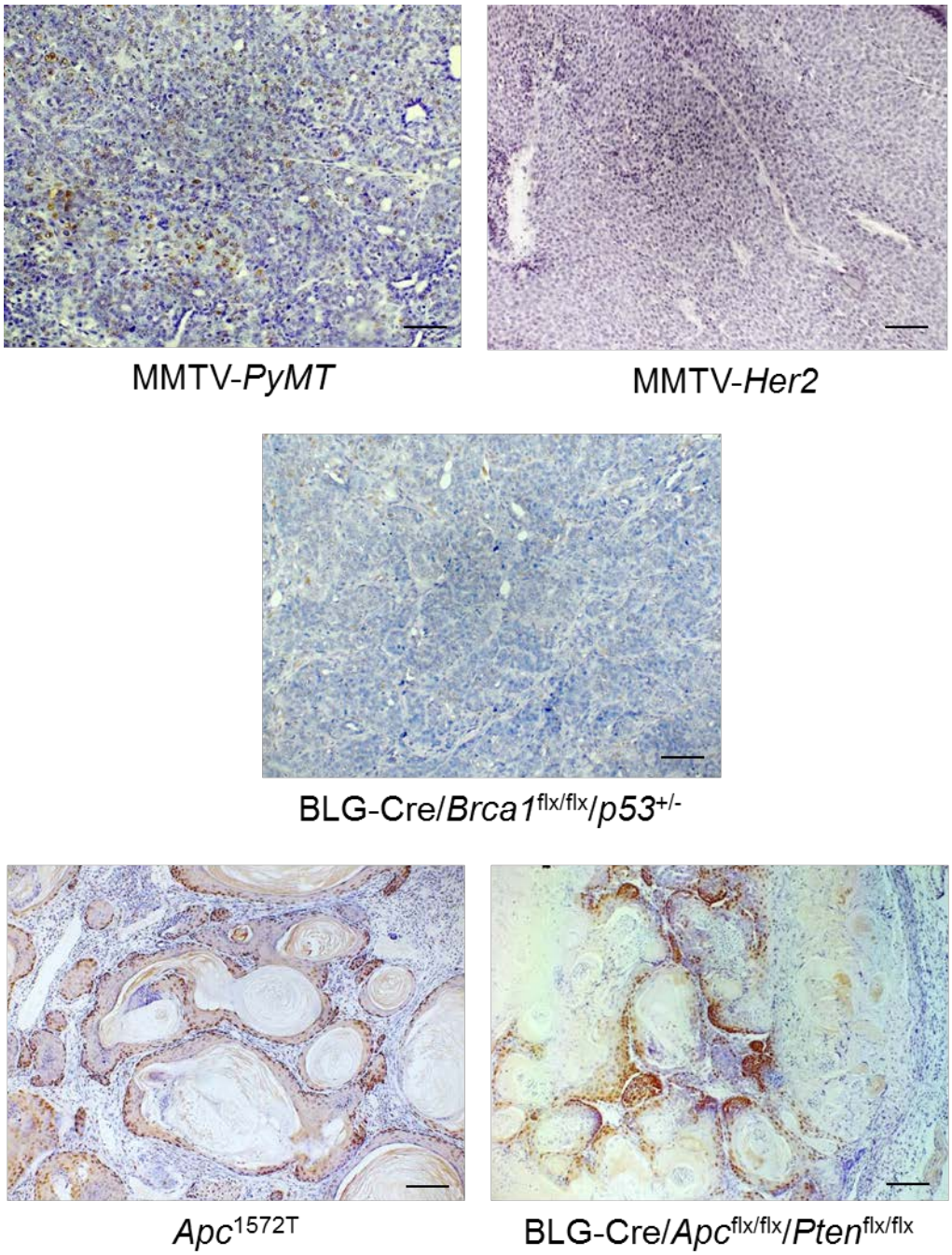


Figure 5-1 RUNX2 expression in mouse models of breast cancer.

RUNX2 IHC on five mouse models of breast cancer. MMTV-*PyMT* and MMTV-*Her2* are luminal-like BC while BLG-Cre/*Brca*1$^{flx/flx}$/*p53*+/-, *Apc*^{1572T} and BLG-Cre/*Apc*$^{flx/flx}$*Pten*$^{flx/flx}$ are basal-like models. Three to five independent tumours have been stained from each genotype. Scale bars represent 50μM.

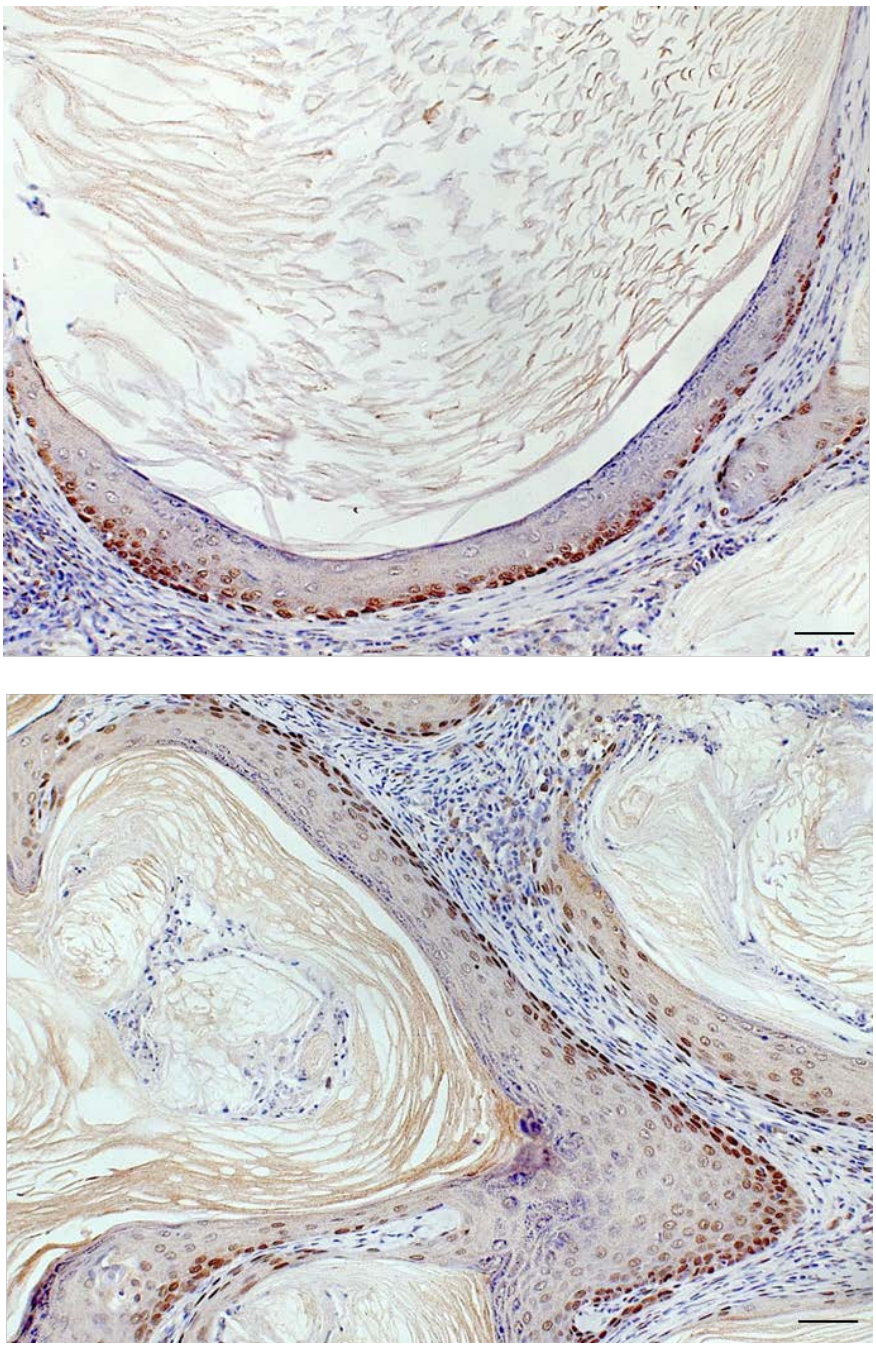


Figure 5-2 Examples of RUNX2 basal pattern of expression in squamous lesions.

RUNX2 expression on *Apc*^{1572T} tumours as assessed by IHC. Scale bars represent 50μM.

5.2.2 Runx2 in a mouse model of mammary squamous metaplasia.

These data suggested a link between RUNX2 expression, squamous metaplastic breast cancer (SMBC) and activated *Wnt* signalling in mouse models of breast cancer. To further investigate a possible involvement of RUNX2 in metaplastic breast cancer, a mouse model of *Wnt*-induced squamous metaplasia was chosen (Miyoshi, Shillingford, et al. 2002), where endogenous ß-catenin is stabilized through the deletion of exon 3 (amino acids 5-80) of the *ß-catenin* gene (*Catnb*$^{+/lox(ex3)}$ mice). This model has been reported to give rise to extensive squamous metaplasia, but not adenocarcinomas, upon *ß-catenin* activation during pregnancy (Miyoshi, Shillingford, et al. 2002). As the Cre recombinase used here was different from the ones published previously, it was important to first characterize the model, choosing to look at virgin, lactating and involuting stages. Activation of ß-catenin driven by BLG-Cre was confirmed to induce squamous metaplastic lesions at all stages tested (FIGURE 5-3). These lesions are characterized by the presence of asymmetrical keratinized structures usually with a thicker layer of squamous cells at one edge (FIGURE 5-4-A). Like in normal epidermis (Blanpain & Fuchs 2009), the squamous cells located in the basal region undergo a constant process of maturation. Actively cycling basal cells continuously produce progenitors which are pushed upwards towards the centre of the cysts, while undergoing a terminal differentiation program. As they mature, cells acquire a flatten morphology and start to produce large amounts of epidermal keratins such as CK1 (FIGURE 5-4-B). At the end of the process these cells will become embedded in a very dense keratin structure, similar to hair matrix and will die, as shown by the presence of ghost cells in the middle of the lesions (FIGURE 5-4-C). To investigate if *Wnt* signalling activation in the virgin epithelium induces *Runx2* expression, qRT-PCR was performed on RNA extracted from virgin glands of BLG-Cre/*Catnb*$^{+/lox(ex3)}$ and BLG-Cre/*Catnb*$^{+/+}$ mice. Interestingly a significant upregulation of *Runx2* transcript was found in the mammary epithelium expressing the activated form of ß-catenin, corroborating the idea of *Runx2* as a downstream target of *Wnt* signalling in the mammary gland (FIGURE 5-5-A). RUNX2 activation at the protein level was then confirmed through immunohistochemistry showing that ß-catenin stabilization induces RUNX2 expression during the lactation phase, a stage of development which is normally devoid of RUNX2 (FIGURE 5-5-B). RUNX2 expression was usually located in the basal layer of the squamous lesions (FIGURE 5-5-C). Interestingly

ß-catenin expression was restricted to this compartment and not expressed in the terminal differentiating cells (FIGURE 5-5-D).

To investigate how RUNX2 loss could affect development of mammary squamous metaplasia three cohorts, BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{WT/WT}$, BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{WT/flx}$ and BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{flx/flx}$, were generated. Since BLG-Cre is mainly activated during pregnancy, and to achieve high levels of *Wnt* activation in the majority of the mammary epithelium, mice were taken through one round of pregnancy and then sacrificed at involution day 7. Loss of RUNX2 caused a significant increase in the number of squamous lesions compared to the wild-type and heterozygous cohort (FIGURE 5-6). This data could suggest that RUNX2 is activated by *Wnt* signalling and acts to repress the formation of squamous lesions. In this situation, RUNX2 would restrict the mammary to epidermal trans-differentiation signal mediated by aberrant *Wnt* signalling in the mammary epithelium.

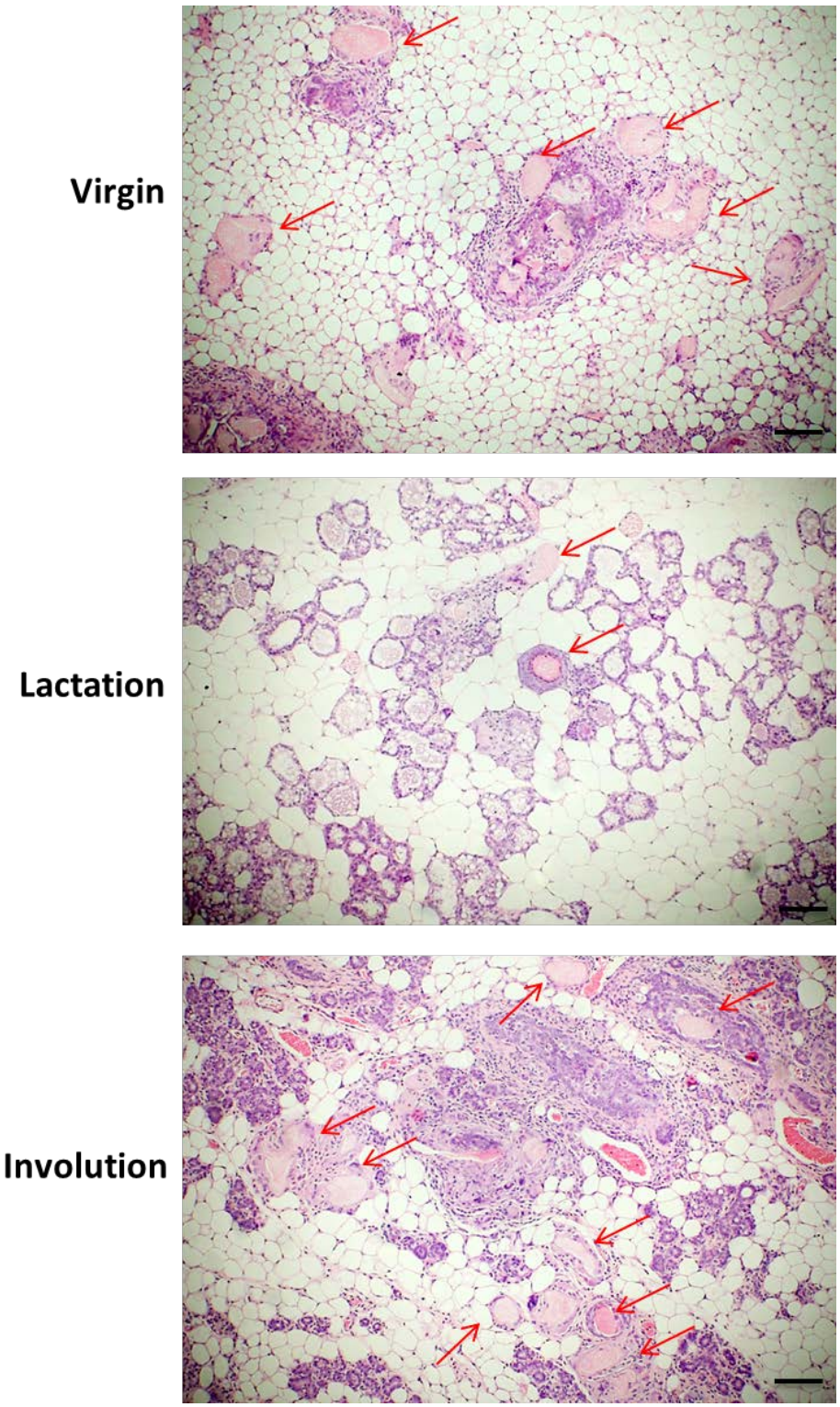


Figure 5-3 Effects of ß-catenin stabilization at different stages of mammary development.

H&E examples of 12 weeks virgin, lactating d1 and involuting day 7, BLG-Cre/*Catnb*$^{+/lox(ex3)}$ mice. Red arrows indicate keratinized squamous lesions. Scale bars represent 100µM. n≥2 per group.

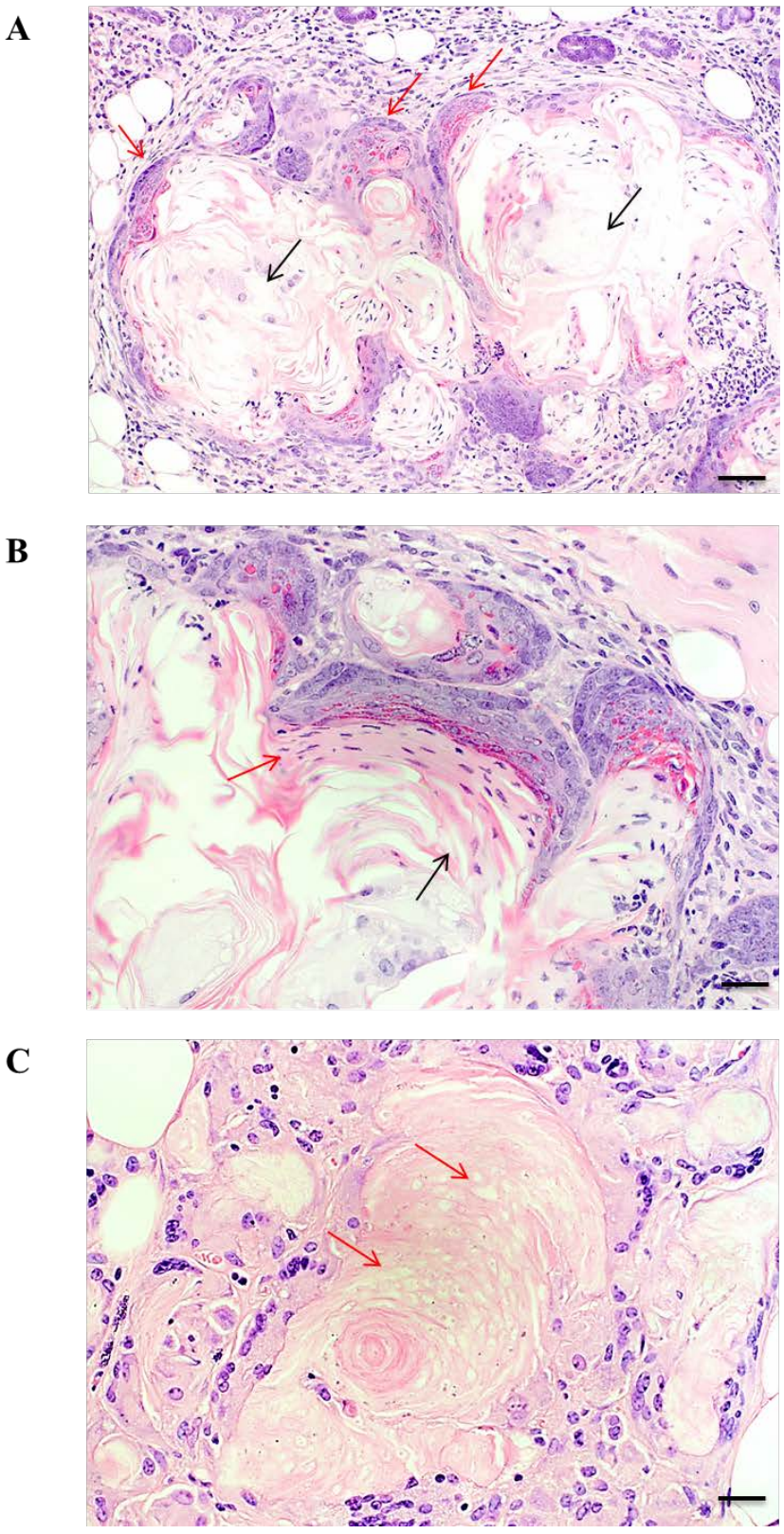


Figure 5-4 Histological analysis of ß-catenin-induced squamous lesions.

H&E examples of squamous lesions developed by BLG-Cre/*Catnb*$^{+/lox(ex3)}$ mice. Red arrows indicate layers of differentiating squamous cells. Black arrows indicate the keratin deposit in the middle of the lesion (A). H&E of a squamous lesion. Red arrow indicates the layer of differentiating cells which have acquired a flatten morphology. Black arrow indicates completely keratinized, terminally differentiated layers, which are released into the lesion (B). H&E of a mature squamous lesion. Red arrows indicate ghost cells as white traces left on the keratin matrix by dead cells (C). Scale bar in (A) represents 50µM. Scale bars in (B-C) represent 25µM.

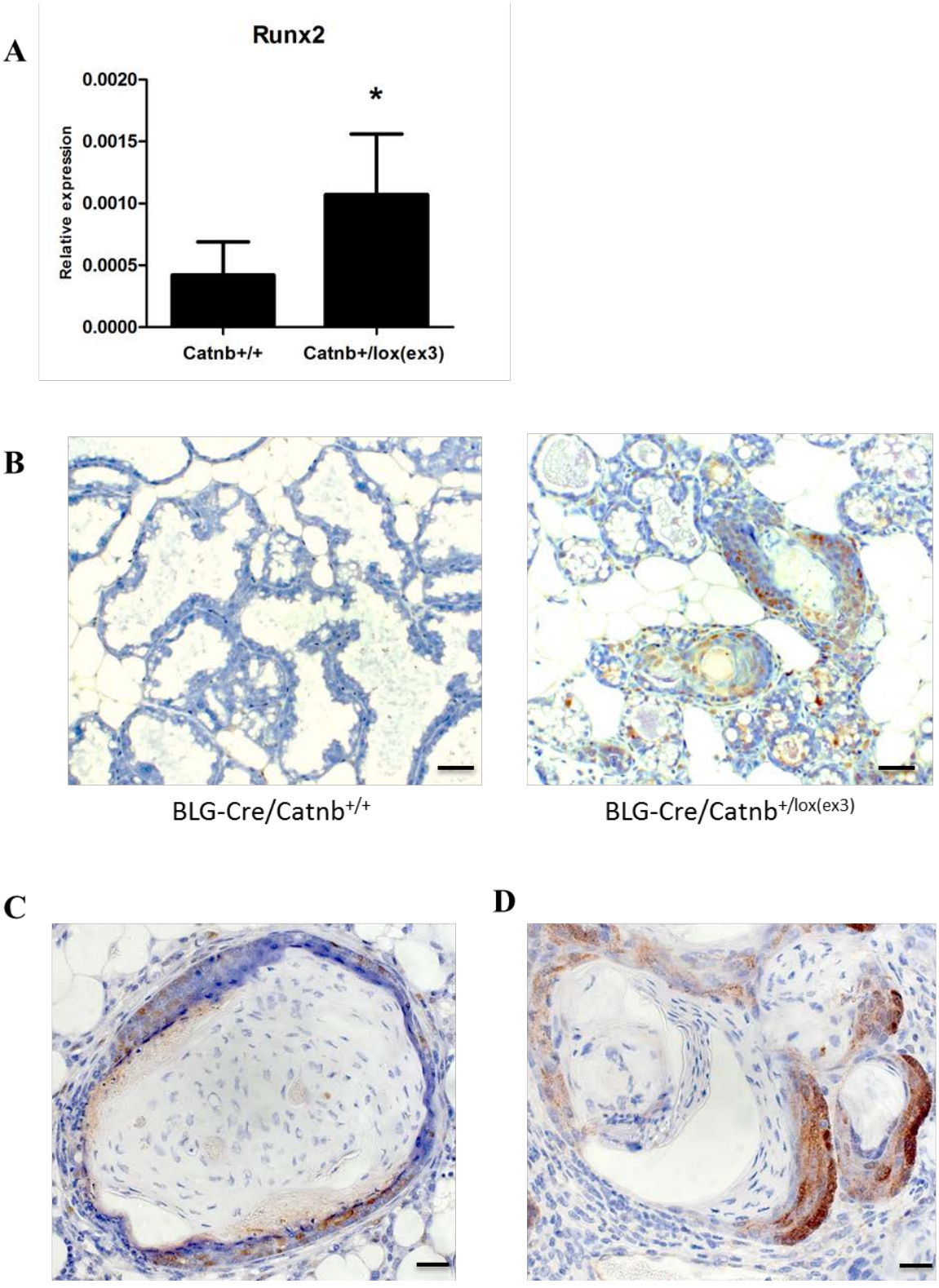


Figure 5-5 Effects of ß-catenin stabilization on RUNX2 expression.

qRT-PCR for *Runx2* on RNA extracted from BLG-Cre/*Catnb*$^{+/+}$ and BLG-Cre/*Catnb*$^{+/lox(ex3)}$ virgin glands. RNA levels were normalized to GAPDH. Data are expressed as mean relative expression (±SD). n≥3 for each group. $p<0.05$, Mann-Whitney test (A). RUNX2 immunohistochemistry on lactating day 1 BLG-Cre/*Catnb*$^{+/+}$ and BLG-Cre/*Catnb*$^{+/lox(ex3)}$ glands (B). RUNX2 (C) and ß-catenin (D) IHC on BLG-Cre/*Catnb*$^{+/lox(ex3)}$ squamous lesions. Scale bar in (B) represents 50µM. Scale bars in (C-D) represent 20µM.

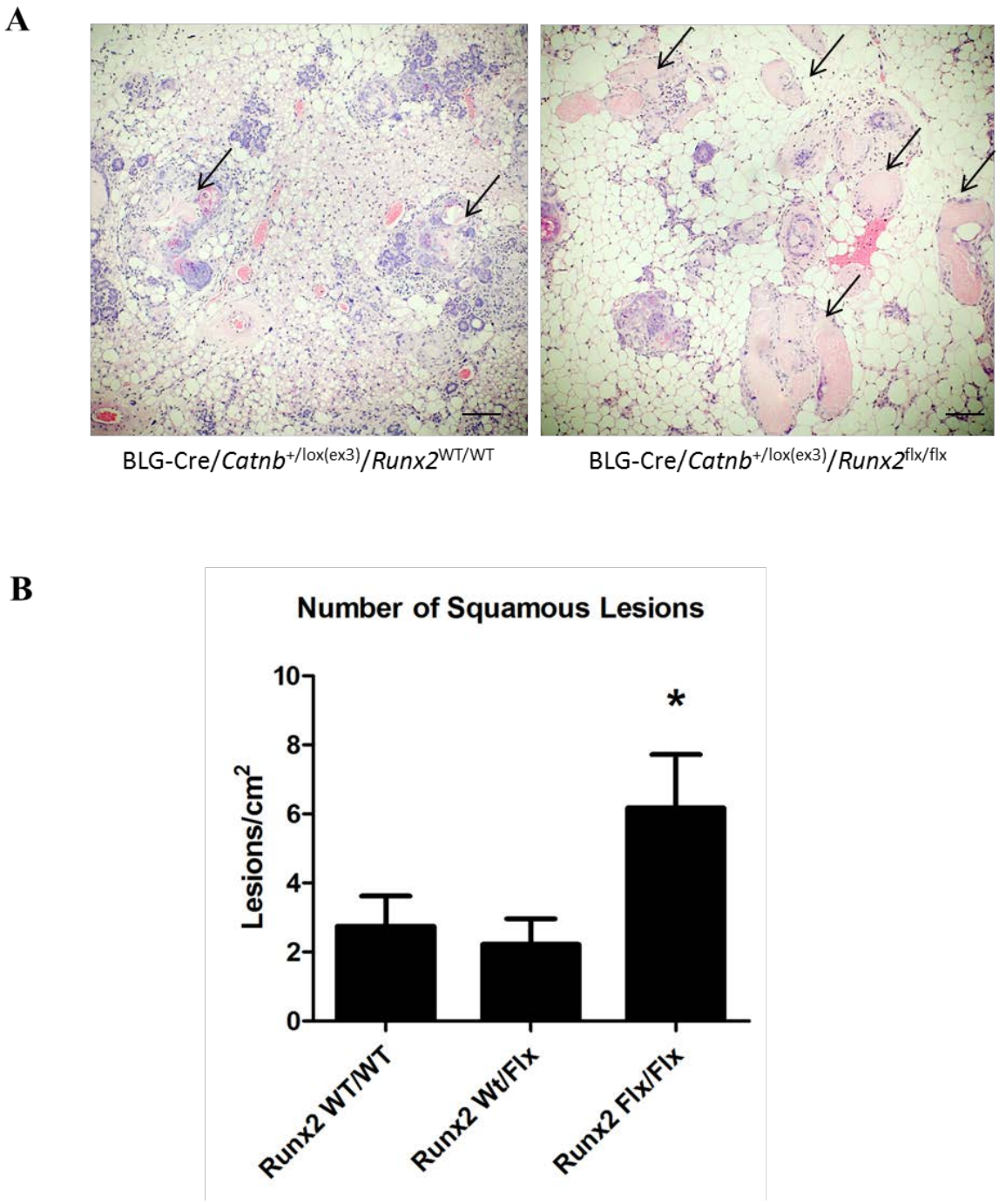


Figure 5-6 Effects of RUNX2 loss on squamous metaplasia.

Representative H&E images of BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{WT/WT}$ and BLG-Cre/*Catnb*$^{+/lox(ex3)}$/ Runx2$^{flx/flx}$ glands taken at involution day 7. Black arrows indicate squamous lesions. Scale bars represent 100μM (A). Bar-chart showing the number of squamous lesions/cm^2, developed by BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{WT/WT}$ (Runx2 WT/WT), BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{WT/flx}$ (Runx2 WT/Flx) and BLG-Cre/*Catnb*$^{+/lox(ex3)}$/ Runx2$^{flx/flx}$ (Runx2 Flx/Flx) cohorts. n≥4 for each group. $p<0.05$, Mann-Whitney test (A).

5.2.3 RUNX2 deletion in an *in vitro* model of triple negative metastatic breast cancer

The human breast cancer cell line MDA-MB-231 is a commonly used *in vitro* model representative of the triple-negative subtype as it lacks expression of the ER, PR and HER2 markers (Holliday & Speirs 2011). This cell line expresses high levels of RUNX2 (Nagaraja et al. 2006) and was used to investigate the role of RUNX2 in triple-negative cells. The subline MDA-MB-231-luc-D3H2LN, a highly metastatic variant of MDA-MB-231 which has been derived from a spontaneous lymph node metastasis of a MDA-MB-231 mammary fat pad tumour (Jenkins, 2005) was utilised due to its high metastatic potential. Furthermore, the presence of a luciferase reporter in this cell line is advantageous for *in vivo* imaging in experimental metastasis models (intravenous and intracardiac) and orthotopic mammary fat pad models (Jenkins et al. 2005). RUNX2 has been shown to be an important regulator of invasive and tumourigenic features of MDA-MB-231 (Pratap et al. 2009), however no study so far has looked at its role in the metastatic seeding process. Using short-hairpin RNA technology, stable MDA-MB-231 RUNX2 knock-out cell lines were generated, one carrying a scramble sh-RNA as a control (shSCR) and two with different RUNX2-shRNAs (shRUNX2-1, shRUNX2-2). The RUNX2 knock-down was confirmed through western blot (FIGURE 5-7-A). No difference in 2D growth was detected between the control and the two knock-down (FIGURE 5-7-B).

Previous data (Chapter 4) indicated that RUNX2 is a regulator of stemness in the mammary epithelium. To test if loss of RUNX2 was affecting the cancer stem cell properties of MDA-MB-231-luc-D3H2LN cells, a tumoursphere assay was used as an marker of stemness (Ponti et al. 2005). In this assay cells are grown in non-adherent conditions and in absence of serum; conditions which allow enrichment in cancer stem cells. Interestingly loss of RUNX2 caused a drastic reduction in tumoursphere-forming capacity of MDA-MB-231 cells (FIGURE 5-8). This data is suggesting that RUNX2 loss could hinder cancer stem cell properties of MDA-MB-231-luc-D3H2LN cells. Several studies have indicated a strong link between cancer stem cells and metastasis formation (Gao et al. 2012; Malanchi et al. 2011). To asses if loss of RUNX2 was affecting the metastatic behaviour of MDA-MB-231 cells, *in vivo* experimental metastasis experiments (tail vein injections) were performed taking advantage of MDA-MB-231 luciferase expression to follow

the metastatic disease through *in vivo* imaging. Interestingly, in a pilot experiment (3 mice per cohort), tail-vein injection of shSCR and shRUNX2-1 cells showed a potential decrease of lung metastasis formation after Runx2 knock-down (FIGURE 5-8).

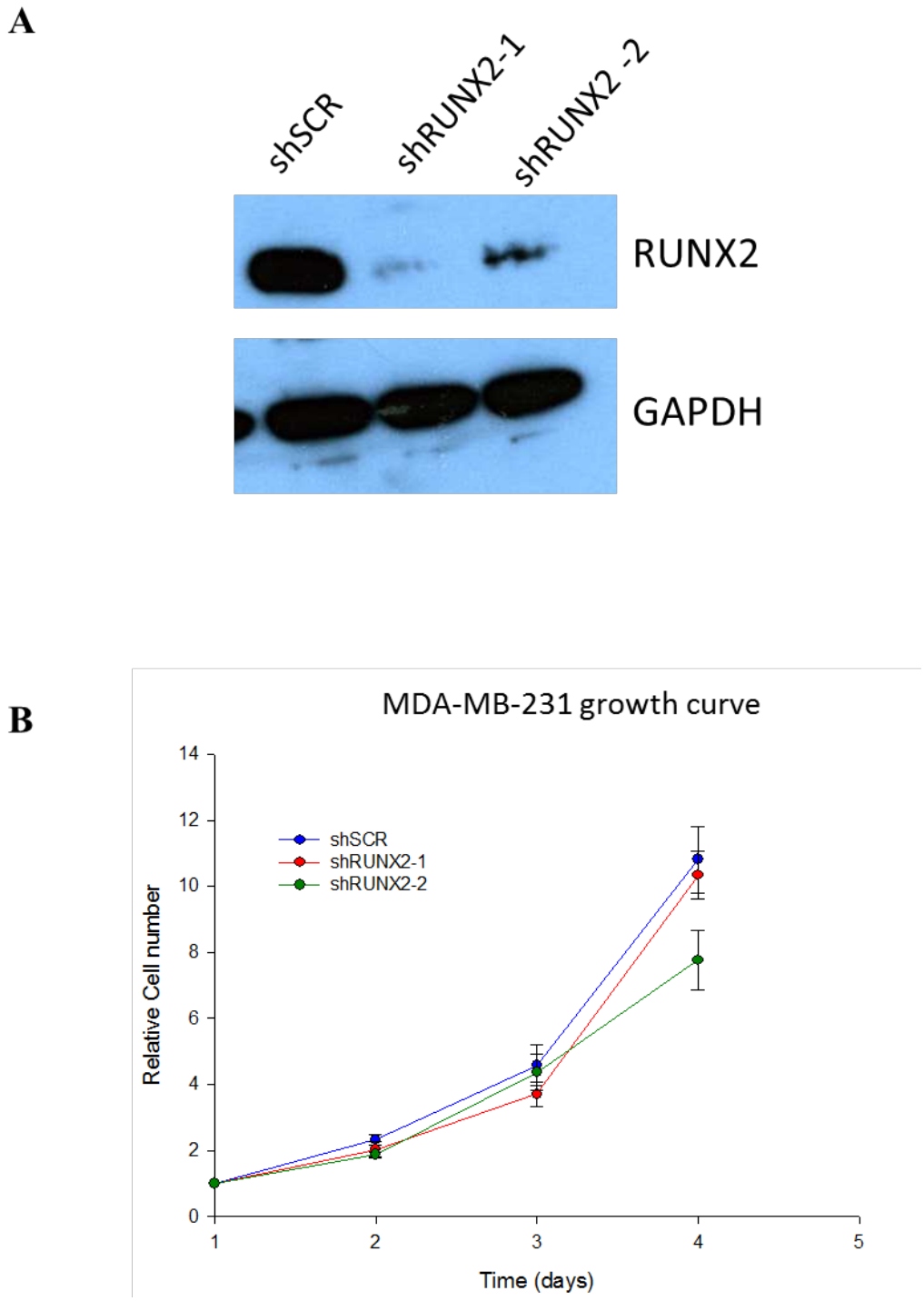


Figure 5-7 Creation of a stable RUNX2 knockdown in MDA-MB-231 cells.

RUNX2 western blot on MDA-MB-231 cells, transfected with sh-scrambled (shSCR) or 2 independent sh-RNAs targeting RUNX2 (shRUNX2-1, sh-RUNX2-2) (A). 2D Growth curve of MDA-MB-231 cells transfected with shSCR, or sh-RUNX2 1 and sh-RUNX2-2 (B). Cell numbers were counted daily in quadruplicate for each time point, for each cell line. Graph representative of two independent experiments.

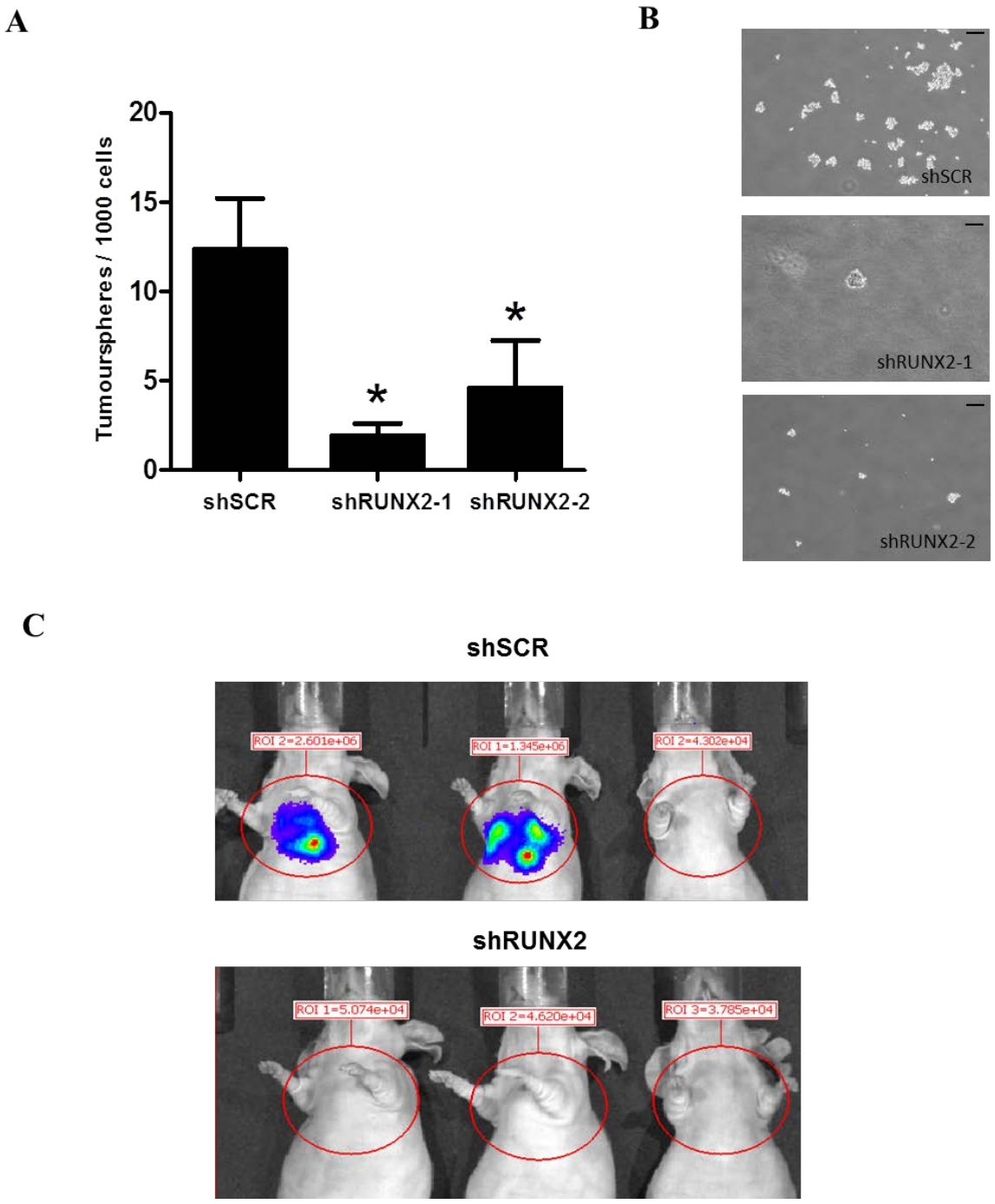


Figure 5-8 Effects of RUNX2 knockdown in MDA-MB-231 cells.

Tumoursphere quantification from shSCR, and sh-RUNX2 cells grown in non-adherent conditions for 7 days. Tumourspheres were counted under a bright field microscope. Data are expressed as mean +/- SD. Graph representative of four independent experiments. $p<0.05$ (A). Bright field images of tumourspheres. Scale bars represent 150µM (B). *In vivo* luciferase imaging of nude mice (n=3 per group), 12 weeks after being tail vein injected with MDA-MB-231 cells transfected with shSCR or shRUNX2-1 (C).

5.3 DISCUSSION

Unpublished data from our laboratory indicate that RUNX2 expression correlates with triple negative samples in human breast cancer. Metaplastic breast cancer (MBC) is a heterogeneous subtype of triple negative breast carcinoma characterized by poor outcome, and few therapeutic options (Lee et al. 2012). One feature of MBC is the co-existence of both carcinoma and non-epithelial components which can include fibrosarcomatous, cartilaginous, osseus, muscular, or squamous differentiation (Cooper et al. 2013). MBC is rare relative to invasive ductal carcinoma, representing less than 1% of all breast cancers. However, since the total number of BC diagnosed in 2011, in the US alone (http://www.cancer.org) was 288130, more than 2000 women were expected to be affected by metaplastic breast cancer. Moreover the 5-year survival rate for MBC is approximately 65% compared to 89% for invasive ductal cancer, indicating the need of new treatment options (Verma 2012). Therefore, a better understanding of the molecular alterations underlying the distinctive morphological and clinical characteristics of this aggressive subtype is needed in order to identify potential new targets for treatment.

RUNX2 immunohistochemistry on a panel of mouse breast cancer models showed that BRCA1-deleted triple negative breast cancers lack RUNX2 expression while a strong staining was detected in two different models of *Wnt*-induced squamous metaplastic breast cancer (*Apc*^{1572T} and BLG-Cre/*Pten*$^{flx/flx}$ *Apc*$^{flx/flx}$). Moreover in a nice confirmation of our unpublished human data, RUNX2 expression was negligible in two luminal-like breast cancer models (Herschkowitz et al. 2007). The development of squamous metaplasia in the mouse mammary gland is induced by aberrant *Wnt* signalling (Miyoshi, Rosner, et al. 2002; Miyoshi, Shillingford, et al. 2002; Gaspar et al. 2009), which activates a process of trans-differentiation of the mammary epithelium resulting in epidermal-like structures with extensive keratinization and basosquamous/pilar histological structures (Kuraguchi et al. 2009). Interestingly RUNX2 expression in both squamous metaplastic breast cancer (SMBC) models was limited to the tumour cells positioned in the basal layer of the lesions. This compartment is characterized by cells in a high proliferative state when compared with the other squamous

part of the tumour which are mainly quiescent (Kuraguchi et al. 2009). Moreover in the normal epidermis this is the compartment were the stem cell population resides (Blanpain & Fuchs 2009). These results suggest a link between RUNX2 expression and a highly proliferative population of SMBC, which could be endowed with cancer stem cell features.

To understand the possible role of RUNX2 expression in squamous metaplasia a published mouse model was chosen, where expression of an active form of ß-catenin (*Catnb*$^{+/lox(ex3)}$) in the mammary gland causes constitutive *Wnt* signalling activation and widespread squamous metaplasia (Miyoshi, et al. 2002). Intriguingly, aberrant ß-catenin activation driven by BLG-Cre expression induced RUNX2 expression *in vivo* in the mammary gland. This result is supporting previous *in vitro* data obtained from WNT3A treatment on mammospheres (Chapter 4), suggesting that *Runx2* is a direct downstream target of *Wnt* signalling in the mammary epithelium. Moreover these data are also in line with various findings from other systems which link RUNX proteins to *Wnt* signalling (Gaur et al. 2005; Osorio et al. 2011). The *Wnt*-induced trans-differentiation process of the mammary epithelium is fascinating: activation of a single molecule, such as ß-catenin, in a mature gland is able to reprogram mammary epithelial cells into a different lineage characterized by a completely different structure and function (i.e. epidermis). This could be explained by the fact that ß-catenin is a key specifier of hair follicle fate (Gat, 1998; Huelsken, 2001); hence its over-activation is sufficient to induce aberrant epidermal/follicular trans-differentiation in the mammary epithelium. However the underlying molecular mechanism for this mammary to epidermal trans-differentiation has not been investigated. A better understanding of this process could help identify new regulators of mammary epithelial identity and expand our basic knowledge of mammary epithelium plasticity. Furthermore, since squamous metaplasia is a key feature of a subgroup of TN breast cancers (Weigelt et al. 2009), these studies could identify new potential targets for this type of tumour.

The experiments described in this chapter unveil a potential role for RUNX2 in *Wnt*-induced squamous metaplasia development. Conditional deletion of RUNX2 combined with ß-catenin activation led to an increase of squamous lesions compared to controls. This result could suggest that RUNX2 is counteracting the ß-catenin-induced epidermal trans-differentiation process. When RUNX2 is lost

however, its repressor function is missing and the trans-differentiation process can proceed uncontrolled, leading to increased squamous metaplasia. A possible mechanism for RUNX2's repressor role in this system could be through a block of terminal differentiation of ß-catenin-activated cells, via induction of quiescence. Interestingly previous experiments, where loss of *Runx2 in vivo* caused a decrease in mammosphere size and number followed by a decrease in p21 (Chapter 4) are suggesting RUNX2 as a possible positive regulator of quiescence in mammary cells. A way to test if RUNX2 deletion could cause exit from quiescence after ß-catenin activation, would be the analysis of cell cycle regulators such as p21 and Cyclin D1 via qRT-PCR and western blot on BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{WT/WT}$ and BLG-Cre/*Catnb*$^{+/lox(ex3)}$/*Runx2*$^{flx/flx}$ glands.

Overall these experiments suggest a new role for RUNX2 in the mammary gland as a downstream target of *Wnt* signalling and as an essential repressor of squamous metaplasia development. Further studies to characterize the exact molecular mechanism of increased trans-differentiation after RUNX2 loss are currently underway. Preliminary data on multiparous experiments, where females from control and RUNX2-deleted cohorts have been taken through multiple rounds of pregnancies to achieve maximum levels of recombination confirm the involution data, showing a dramatic increase in squamous metaplasia in RUNX2-deleted glands. Moreover the almost complete absence of proliferation (as assessed by ki67 staining) in RUNX2-deleted glands support the view that loss of RUNX2 is pushing cells towards terminal differentiation. Molecular analysis of the involuting and multiparous glands through qRT-PCR and western blot will help to clarify the mechanism of the observed phenotype. To better investigate the role of RUNX2 in squamous metaplastic breast cancer, future experiments could use the *Apc*^{1572T} mouse model, which develops squamous metaplastic breast cancer, in combination with the K14-Cre/*Runx2*$^{flx/flx}$ mouse. If the same mechanism observed in the *Catnb*$^{+/lox(ex3)}$ model, where loss of RUNX2 induces differentiation, is conserved, we could speculate that less aggressive and more differentiated tumours will develop after RUNX2 deletion. The correlation between high RUNX2 expression, *Wnt* pathway activation and metaplastic breast cancer seems to be conserved also in human disease; transcriptomic analysis on human metaplastic breast cancer listed RUNX2 among the upregulated gene signature (Hennessy et al. 2009) while

Wnt signalling was found to be activated in 95% of metaplastic breast cancer (Hayes et al. 2008). As a future priority, RUNX2 expression needs to be confirmed at the protein level on human SMBCs samples through immunohistochemistry.

In an attempt to understand the functional role of RUNX2 in human triple negative disease we performed *in vitro* experiments using a well-established model of the MDA-MB-231 cell line (Holliday & Speirs 2011). Interestingly, in a nice parallel with previous data on mammospheres, RUNX2 inhibition caused decreased tumoursphere capacity without affecting MDA-MB-231 growth in 2D. This result indicates that RUNX2 could be a widespread regulator of regenerative potential in both normal and breast cancer cells. Cancer stemness and metastatic potential are strongly linked (Malanchi et al. 2011; Gao et al. 2012). In fact only cancer cells characterized by a certain degree of regenerative potential will be able to seed and form new tumour colonies in a foreign tissue. RUNX2's role in metastatic cancer cell seeding has not been investigated since the only "metastatic" assays used were orthotopic injections which do not involve seeding (Javed et al. 2005). RUNX2 has been shown to be important for the osteolytic potential of MDA-MB-231 cells when injected directly into the intramedullary space of the tibia (Barnes et al. 2004; Pratap et al. 2008). In addition MDA-MB-231 cells stably expressing shRNA-*Runx2*, when injected in the mammary fat pad, showed a significant reduction in tumour growth when compared with a control group (Pratap et al. 2009). Our pilot experiment showed that RUNX2 inhibition impairs lung colonization in an experimental metastasis experiment suggesting that RUNX2 is also required for the ability of MDA-MB-231 cells to seed and grow in the lung. Further experiments transplanting MDA-MB-231 cells into the fat pad with stable RUNX2 knockdown, and thereafter removal of primary tumour will allow the assessment of metastatic growth and clarify the role of RUNX2 in breast cancer metastatic disease. Furthermore, since the majority of the research conducted so far has always been limited to the use of a single cell line (MDA-MB-231) which cannot recapitulate the disease complexity, there is a need for new and better *in vitro* and *in vivo* models for the study of RUNX2 role in breast cancer.

6 A role for RUNX1 in breast cancer

6.1 INTRODUCTION

6.1.1 RUNX1 in epithelial cancer.

RUNX1 has been extensively studied in the haematopoietic system where it is essential for the establishment of definitive haematopoiesis and the generation of haematopoietic stem cells during embryonic development (Swiers et al. 2010). Moreover *RUNX1* is the most frequently mutated gene in human leukaemia, usually through chromosomal translocations which interfere with RUNX1 transcriptional activity (Lam & Zhang 2012). Hence, the mainstream of RUNX1 research has been focused on its tumour suppressive function in haematopoietic malignancies (Lam & Zhang 2012). However, in recent years, a new role for RUNX1 outside the haematopoietic system has started to emerge with several studies indicating how this transcription factor could be more broadly implicated in cancer than previously thought (Scheitz & Tumbar 2012; Taniuchi et al. 2012). In particular RUNX1 has been identified as a key regulator of tumourigenesis in various epithelial cancers. RUNX1 is expressed at high levels in mouse skin papilloma and squamous cell carcinoma, and *Runx1* deficiency impairs mouse skin tumourigenesis (Hoi et al. 2010). Moreover deletion of *Runx1* in a mouse model of head and neck cancer caused a significant delay in tumour formation (Scheitz et al. 2012). RUNX1 is also highly expressed in primary and metastatic epithelial ovarian cancer compared to normal tissue (Keita et al. 2013) while in prostate cancer, RUNX1 expression increases with pathological stage (Yeh et al. 2009). These data are indicating an oncogenic role for RUNX1 in different epithelial cancers. In contrast, RUNX1 seems to have a tumour suppressor role in the intestine since both wild-type and Apc^{Min} mice develop significantly more colon and small intestine tumours after *Runx1* deletion (Fijneman et al. 2011). These studies serve to highlight the context dependence

of *Runx* genes (Blyth et al. 2005), being able to orchestrate both an oncogenic and a tumour suppressive program probably through collaboration with different combinations of tissue-specific transcriptional co-regulators.

6.1.2 RUNX1 in breast cancer

A role for RUNX1 in breast cancer is also starting to emerge. However the evidences published so far are discordant. Wang *et al* using 3D culture models showed that *RUNX1* deletion in MCF10A resulted in increased cell proliferation and abnormal morphogenesis (Wang et al. 2011). Furthermore two independent whole-exome sequencing studies on human breast cancers discovered recurrent RUNX1 mutations and deletions in a small subgroup of human tumours (Ellis et al. 2012; Banerji et al. 2012) while Kadota *et al* showed that RUNX1 deletion is associated with high-grade primary breast tumours (Kadota et al. 2010). These studies indicate a possible tumour suppressor role for RUNX1 in breast cancer. On the other hand, work from Tumbar's laboratory, found RUNX1 among the top 1% highly expressed genes in tumour versus normal tissue in a meta-analysis of microarray studies of various cancers, including breast (Scheitz et al. 2012). *RUNX1* was also identified among the upregulated transcripts in breast tumours and short term tumour cultures when compared with established breast cancer cell lines (Dairkee et al. 2004). This study suggests that RUNX1 could be lost in established cell lines and poses a word of caution on the use of *in vitro* models for the study of RUNX1 involvement in BC. Other transcriptomic studies on human breast cancer support an oncogenic role for RUNX1 in this malignancy. In human breast cancer oestrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor 2 (HER2) are well-established prognostic and predictive markers and testing for them is now considered standard of care (Taneja et al. 2010). Based on the receptor status, human breast cancer can be subdivided into three main groups: oestrogen receptor positive (ER+), epidermal growth factor receptor 2 positive (HER2+) and triple negative (ER-/PR-/HER2-). ER+ and HER2+ patients benefit of targeted treatments such as Tamoxifen and/or Trastuzumab which have consistently improved disease outcome (Higgins & Baselga 2011). On the other hand, the triple negative subtype lacks any specific targeted therapy and is associated with worse overall prognosis in comparison with the other subtypes (Foulkes et al. 2010). Interestingly RUNX1 is among a 264 gene signature which correlates with bad prognosis in triple

negative breast cancer (Karn et al. 2011) and indeed was seen to specifically correlate with bad prognosis in the claudin-low subtype of triple negative breast cancers (Rody et al. 2011). Finally, RUNX1 is among the top 20% differentially expressed genes in two triple negative subtypes: the mesenchymal stem-like (MSL), and a luminal androgen receptor (LAR) (Lehmann et al. 2011). Interestingly the MSL subtype also displays low expression of claudins 3, 4, and 7, confirming a possible link between RUNX1 expression and the claudin-low subtype. Overall these transcriptomic studies suggest a possible oncogenic role for RUNX1.

6.1.3 Experimental Aims.

The role of RUNX1 in breast cancer is still unresolved with different studies suggesting a possible oncogenic role while others point towards a tumour suppressive function. Since breast cancer is a very heterogeneous disease, constituted by different subtypes, each characterized by specific molecular alterations (Sorlie et al. 2003; Prat & Perou 2011), it is tempting to speculate that the discordant data on RUNX1 in this malignancy could reflect different subtype-specific roles for this transcription factor. The aim of this study was to better characterize RUNX1 in human breast cancer with a particular focus on the different tumour subtypes.

6.2 RESULTS

6.2.1 Expression of RUNX1 in human breast cancer.

To get a preliminary overview of *RUNX1* expression in human breast cancer, data-mining in an online cancer gene microarray database (https://www.oncomine.org/) was performed. Two multi-cancer gene expression profiling studies showed an upregulation of *RUNX1* in breast cancer tissue (Yu et al. 2008; Su et al. 2001). *RUNX1* is amongst the top 2% upregulated genes in breast cancer compared to other tumours (FIGURE 6-1-A) and in the top 3% upregulated genes in breast cancer when tumours were compared to adjacent normal mammary tissue (FIGURE 6-1-B). *RUNX1* is also in the top 4% upregulated genes in invasive breast carcinoma when compared to normal tissue (FIGURE 6-2) (Cancer Genome Atlas Network 2012). These analyses are suggesting that RUNX1 could be upregulated in breast cancer. However the association of *RUNX1* expression with clinical outcome and its significance as a prognostic factor in human breast cancer is still unclear. To investigate if RUNX1 expression influenced clinical outcome in primary breast tumours, a tissue microarray (TMA) containing 449 patients with operable invasive ductal breast cancer (Mohammed et al., 2012) was stained for RUNX1. RUNX1 antibody was first validated by western blot and immunohistochemistry, confirming its specificity (data not shown). The invasive cancers showed different degrees of RUNX1 expression predominantly localised to the nucleus (FIGURE 6-3). RUNX1 expression was determined by histoscore (as described in material and methods) and patients were divided into two groups: RUNX1 negative (histoscore = 0, n=109) and RUNX1 positive (histoscore > 0, n=340). The relationship between RUNX1 expression and clinical outcome was then assessed by looking at overall survival (OS) in the cohort. Survival analyses showed no difference between RUNX1 negative (154.9 months - 95% confidence interval, 142–164 months) and RUNX1 positive tumours (148.7 months - 95% confidence interval, 142.6–153.8 months) in the full cohort (FIGURE 6-4-A). The distribution of RUNX1 positive and negative samples in relation to hormonal status (ER/PR/HER2) of the full cohort is showed in FIGURE 6-4-B.

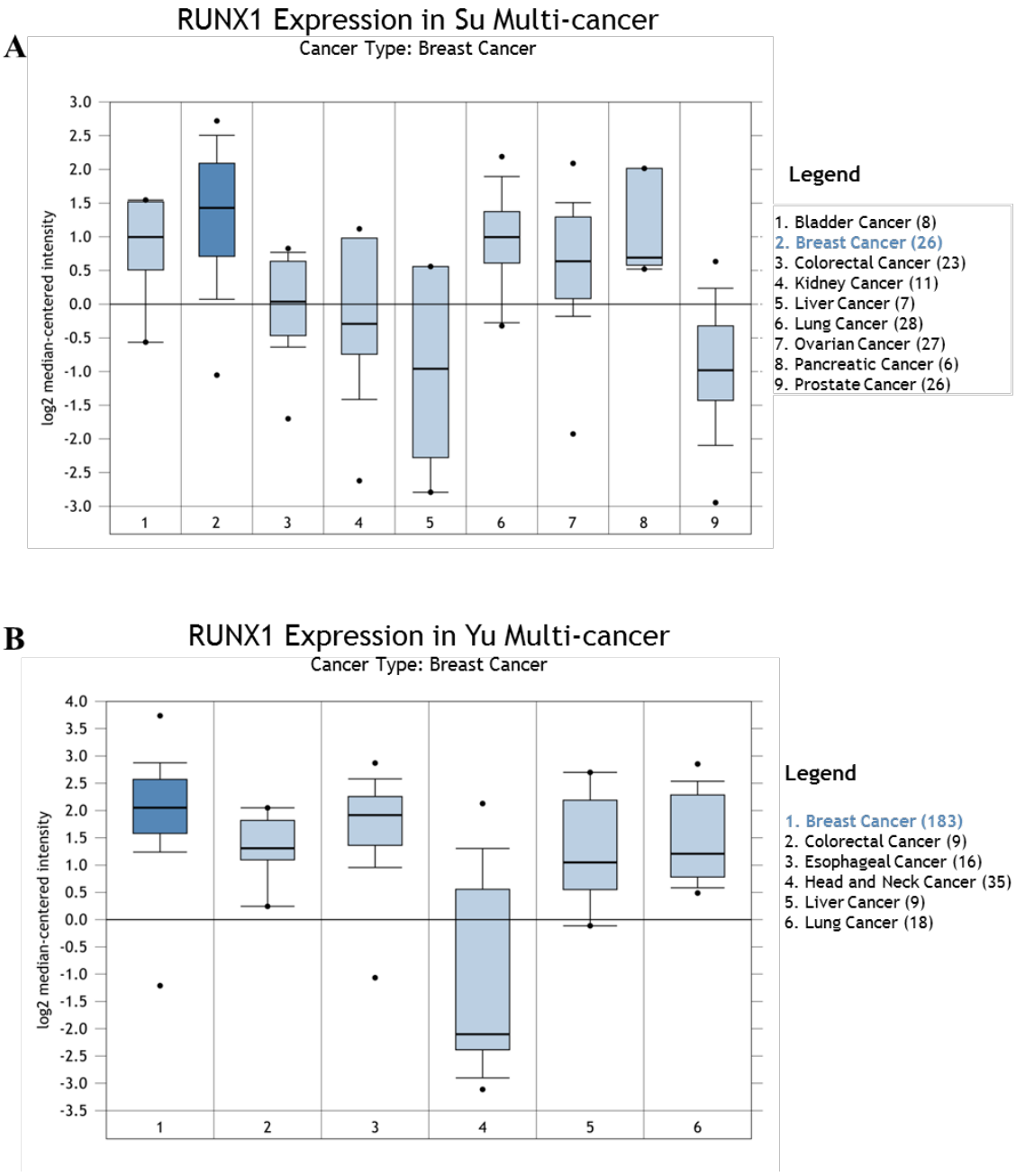


Figure 6-1 *RUNX1* expression on multi-cancer gene expression profiling studies.

RUNX1 gene expression on a panel of 162 tumours. *RUNX1* is overexpressed (2.24 fold-change) in breast cancer compared to *RUNX1* median expression across the database. p = 5.94E-7 (A). *RUNX1* gene expression on a panel of 270 tumour samples as shown in legend, compared to correspondent adjacent normal tissues from breast (n=13), colon (n=9), oesophagus (n=13), liver (n=8), lung (n=12), and thyroid gland (n=16). *RUNX1* is overexpressed (3.01 fold-change) in breast cancer compared to *RUNX1* median expression (cancer vs non-malignant tissue) across the database. p = 4.58E-13 (B). Information taken from www.oncomine.org.

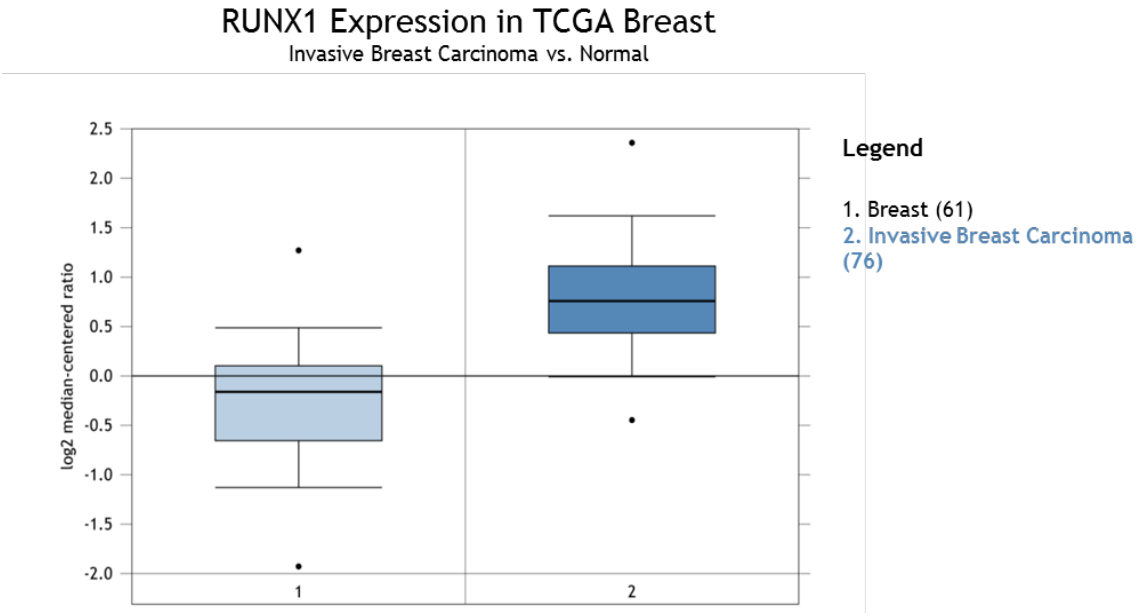


Figure 6-2 *RUNX1* expression in normal versus invasive breast carcinoma.

RUNX1 gene expression on a panel of normal breast (61) and invasive breast carcinoma (76) samples. *RUNX1* is overexpressed (2.1 fold-change) in invasive breast carcinoma compared to normal tissue. p = 2.03E-17. Information taken from www.oncomine.org.

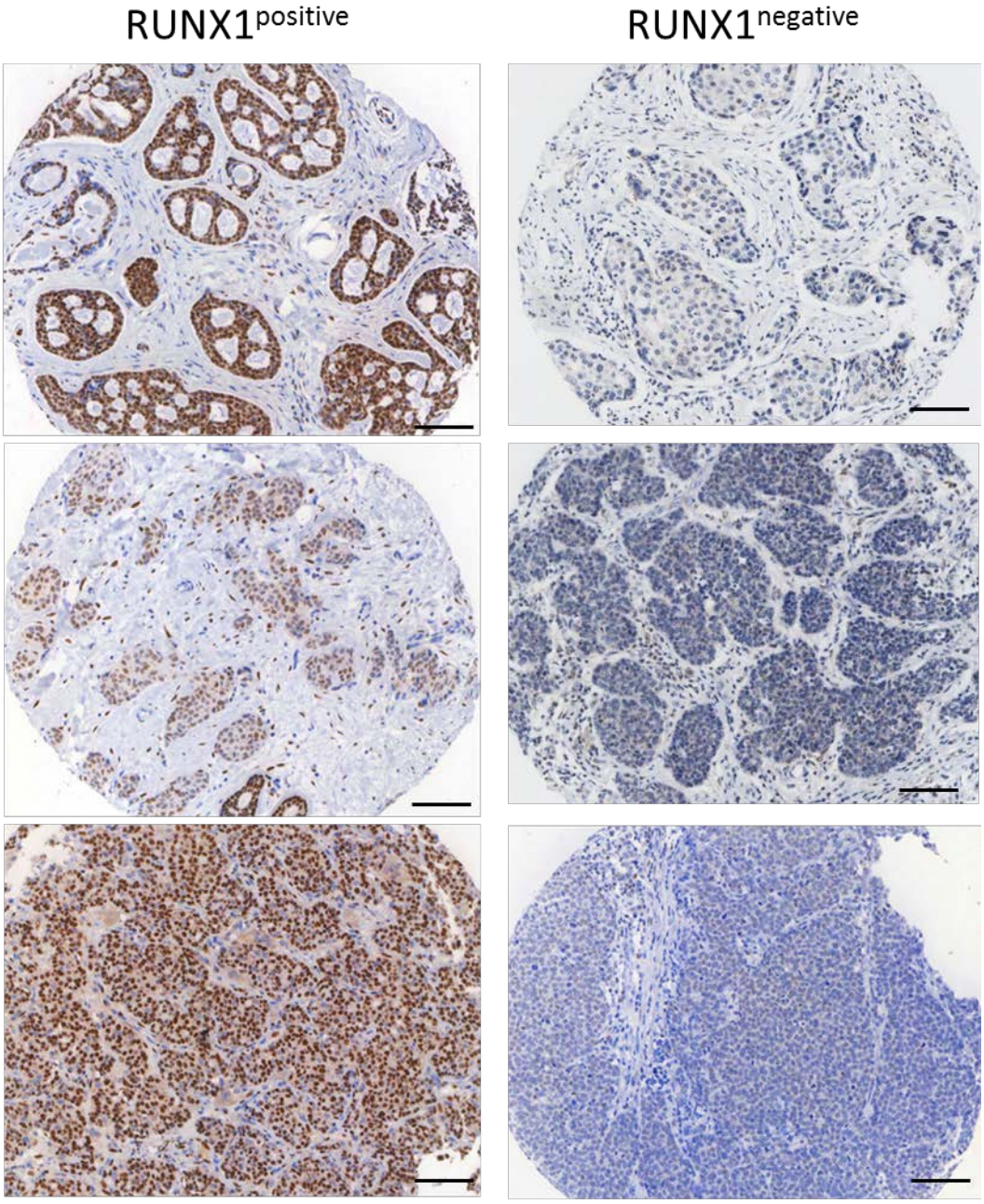
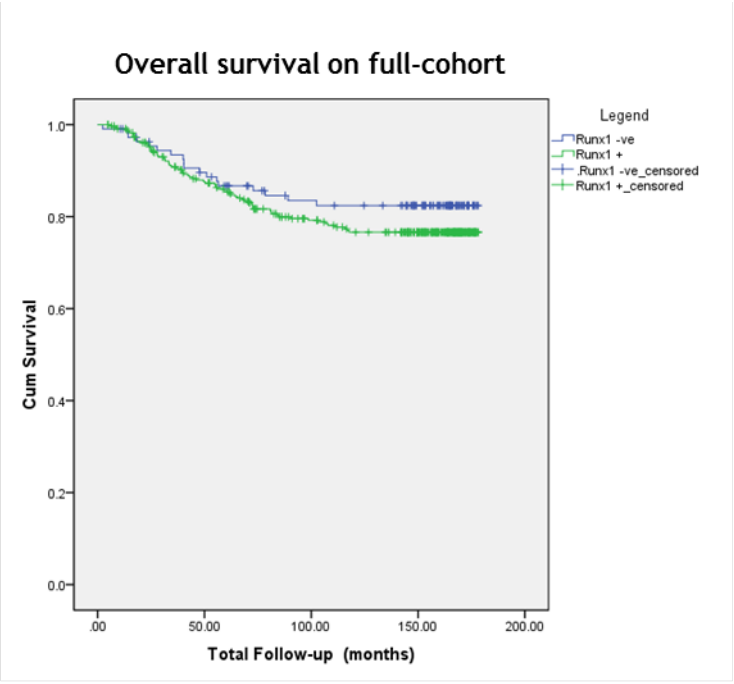


Figure 6-3 Examples of RUNX1 expression in a breast cancer cohort.

Invasive breast carcinomas from a tumour tissue microarray (TMA) were stained for expression of RUNX1. Examples of tumours with positive RUNX1 staining are shown on the left column. Examples of tumours with negative RUNX1 staining are shown on the right column. Scale bar represents 100μM.

A



B

Number of patients (percentage)

	Total population number (percentage)	Runx1 +	Runx1 -
Total	449 (100.0)	340 (75.7)	109 (24)
ER			
Negative	161 (35.8)	115 (71.4)	46 (28.6)
Positive	273 (60.8)	216 (79.1)	57 (20.9)
PR			
Negative	240 (53.4)	171 (71.2)	69 (28.8)
Positive	194 (43.2)	158 (81.4)	36 (18.5)
HER2			
Negative	366 (81.5)	277 (75.7)	89 (24.3)
Positive	68 (15.1)	52 (76.5)	16 (23.5)
Triple negative status			
ER/PR/HER2 neg	110 (24.5)	80 (72.8)	30 (27.2)

Figure 6-4 Correlation of RUNX1 expression and survival in a breast cancer cohort.

Kaplan Meier of patient survival in a cohort of 449 breast cancers. Survival is plotted for patients with cancers scored positive for RUNX1 (green), or negative for RUNX1 expression (blue). p>0.1. p-value calculated using Log Rank (Mantel-Cox) test (A). Table showing the distribution of RUNX1 positive and negative patients in relation to their hormonal status (B).

To define the prognostic impact of RUNX1 expression in different breast cancer subtypes, the patient cohort was divided into 4 subgroups accordingly to their receptor status (ER+, PR+, HER2+ and ER-/PR-/HER2-). The relationship between RUNX1 expression and clinical outcome was then assessed by looking at overall survival in each breast cancer subtype. Survival analyses showed no difference between the RUNX1 positive and negative groups in the ER+, PR+ and HER2+ patients (FIGURE 6-5). However RUNX1 positivity significantly correlated with poorer prognosis in the receptor negative patients (RUNX1+, 127.3 months - 95% CI, 111.7–143 months vs RUNX1-, 162 months - 95% CI, 145.3–178.8 months - p=0.017) (FIGURE 6-5-D).

To test if RUNX1 was an independent prognostic factor in the triple negative cohort we performed a correlation analysis using the Cox regression model. Interestingly RUNX1 was an independent prognostic factor (p=0.009) together with size (p=0.005) and lymph node status (p=0.003) in the triple negative (TN) cohort. A bivariate analysis (Pearson's chi-square) was then performed to determine the presence of an association between clinicopathological history and RUNX1 expression in TN samples. RUNX1 expression was not significantly associated with age (p=0.564), tumour type (p=0.442), tumour size (p=0.456), grade (p=0.352), nodal status (p=0.348), necrosis (p=0.736), Ki67 (p=0.225) and TUNEL (p=0.257) (FIGURE 6-6). This data indicates that RUNX1 is an independent prognostic factor specific for the triple negative subgroup of breast cancers.

Besides being influenced by well recognised host and tumour related factors such as patient age, histological type and grade, tumour size, lymph node and hormonal status, the prognosis of breast cancer is also dependent on other factors such as lymphocytic infiltrate and blood vascular invasion (Z M A Mohammed, Going, et al. 2012). To determine any association between inflammatory reaction/blood vascular invasion and RUNX1 expression in the triple negative subgroup a new bivariate analysis (Pearson's chi-square) was performed on a set of markers previously published (Z M A Mohammed, Going, et al. 2012; Klintrup et al. 2005; Mohammed et al. 2013). The inflammatory reaction in this cohort had been previously assessed in Joanne Edwards laboratory by Klintrup-Makinen scoring (Z M A Mohammed, Going, et al. 2012), a measure of inflammation at the invasive margin of the tumours, together with scoring for lymphocytes (CD4+ and CD8+), plasma cells (CD138+) and

macrophages (CD68+). Vascular invasion was also determined by endothelial cell scoring (CD34+). RUNX1 expression was not significantly associated with Klintrup-Maninen scoring (p=0.081), CD8 (p=0.675), CD138 (p=0.251), CD68 (p=0.710), or CD34 (p=0.928). Interestingly RUNX1 expression was correlated with lymphocytic CD4 staining (p=0.016) (FIGURE 6-7).

6.2.2 Effects of RUNX1 overexpression *in vitro*.

RUNX1 expression was tested on a panel of breast cancer cell lines by western blot analysis. The chosen cell lines included normal human mammary epithelial cells derived from primary tissue and immortalized with TERT expression (hMEC-TERT), 4 basal-like (MDA-MB-468, HCC-70, MDA-MB-231, BT-549) and 3 luminal-like (T47D, MDA-MB-361, BT-474) breast cancer cell lines. Significantly, RUNX1 expression was not detectable in normal hMEC-TERT but was overexpressed in all basal-like cancer cell lines and in 1 out of 3 of the luminal cell lines (FIGURE 6-8). These results suggest an oncogenic role for RUNX1 in human breast cancer, with a particular association with the basal like subgroup of cancers.

To further investigate the putative oncogenic role for RUNX1 in breast cancer, the effects of overexpressing RUNX1 in immortalized mammary epithelial cells (hMEC-TERT) was assessed. hMEC-TERT were transfected with a plasmid encoding the mouse RUNX1 protein (pBABE-puro-Runx1) and vector alone as control (pBABE-puro). Initially both hMEC-TERT control (hMEC-Puro) and hMEC-TERT overexpressing *Runx1* (hMEC-*Runx1*) showed the same morphology. However with passaging, cells with an elongated shape and a fibroblastic appearance started to appear in hMEC-*Runx1* cultures. The control hMEC-Puro instead, maintained the same flat and rounded epithelial morphology (FIGURE 6-9-A). RUNX1 overexpression was confirmed by western blot (FIGURE 6-9-B). This result suggests that RUNX1 overexpression is activating an EMT program in a subpopulation of hMEC-TERT. To confirm that RUNX1 overexpression elicits EMT changes in hMEC-TERT cells, two classical EMT markers were analysed by western blot. Vimentin was strongly upregulated while E-Cadherin levels were decreased in hMEC-*Runx1* compared to hMEC-Puro (FIGURE 6-9-C).

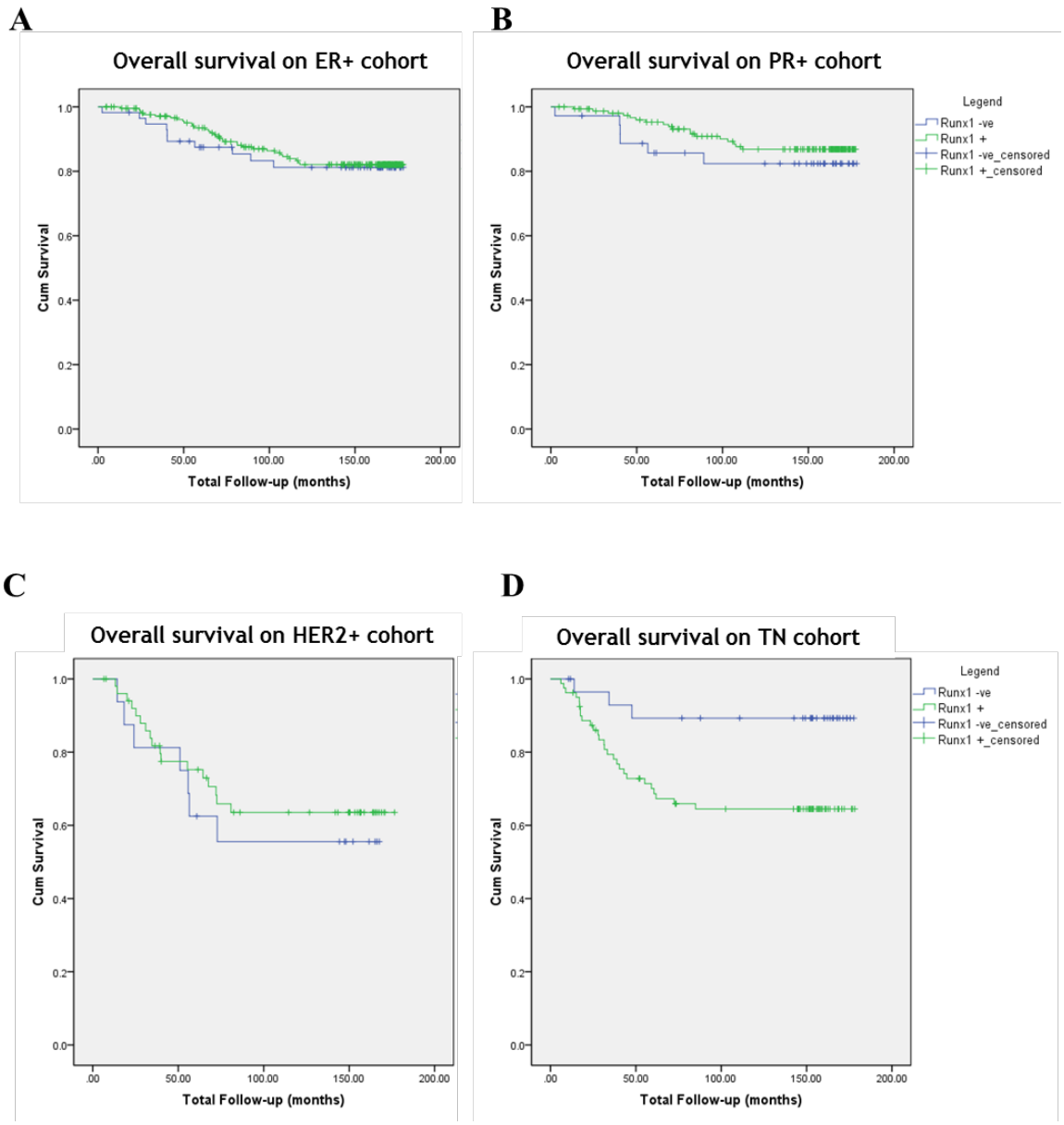


Figure 6-5 RUNX1 expression in different subtypes of breast cancer.

Kaplan Meier of patient survival in the ER+ (A), PR+ (B), HER2+ (C) and triple negative (TN) cohort (D). Survival is plotted for patients with cancers positive for RUNX1 (green), or those with no RUNX1 expression (blue). ER+ cohort, p=0.721. PR+ cohort, p=0.378. HER2+ cohort, p=0.551. TN cohort, p=0.017. p-value calculated using Log Rank (Mantel-Cox) test.

Number of patients (percentage)				
	Total population number (percentage)	**Runx1 +**	**Runx1 -**	**P value[a]**
Total	115 (100.0)	84 (73)	31 (27)	
Age, years				
≤ 50	42 (36.5)	32 (76.2)	10 (23.8)	NS
> 50	73 (63.5)	52 (71.2)	21 (28.8)	
Invasive grade				
I	3 (2.7)	3 (100)	0 (0)	NS
II	22 (19.5)	14 (63.6)	8 (36.4)	
III	88 (77.9)	65 (73.9)	23 (26.1)	
Histology				
Special type	10 (85.8)	6 (86.0)	4 (85.5)	NS
Lobular	2 (6.4)	2 (6.3)	0 (6.6)	
Ductal	103 (7.8)	76 (7.7)	27 (7.9)	
Lymph node status				
Negative	66 (57.4)	46 (69.7)	20 (30.3)	NS
Positive	49 (42.6)	38 (77.6)	11 (22.4)	
Tumor size				
≤ 20 mm	58 (50.9)	42 (72.4)	16 (27.6)	NS
21-50 mm	52 (45.6)	37 (71.2)	15 (28.8)	
> 50 mm	4 (3.5)	4 (100)	0 (0)	
Necrosis				
Negative	20 (17.4)	14 (70)	6 (30)	NS
Positive	95 (82.6)	70 (73.7)	25 (26.3)	
Ki67				
Low	88 (77.9)	67 (76.1)	21 (23.9)	NS
High	25 (22.1)	16 (64)	9 (36)	
TUNEL				
Low	62 (71.3)	48 (77.4)	14 (22.6)	NS
High	25 (28.7)	17 (68)	8 (32)	

Figure 6-6 Relationship between RUNX1 status and standard clinical, pathological, and biological features of triple-negative breast cancer.

p-value calculated using Chi-squared test (linear by linear association). NS, not significant. Where information is not available on the full cohort (n=115), the number of patients with information is specified in brackets: Invasive grade (113), Tumour size (114), Ki67 (113), TUNEL (87).

Number of patients (percentage)				
	Total population number (percentage)	**Runx1 +**	**Runx1 -**	**P value[a]**
Total	115 (100.0)	84 (73)	31 (27)	
Klintrup				
Low grade	51 (43.2)	41 (80.4)	10 (19.6)	NS
High grade	67 (56.8)	45 (67.2)	22 (32.8)	
CD4				
Low	43 (37.4)	25 (58.1)	18 (41.9)	0.016
Medium	20 (17.4)	17 (85)	3 (15)	
High	52 (45.2)	42 (80.8)	10 (19.2)	
CD8				
Low	39 (33.9)	28 (71.8)	11 (28.2)	NS
Medium	23 (20)	16 (69.6)	7 (30.4)	
High	53 (46.1)	40 (75.5)	13 (24.5)	
CD138				
Low	61 (53)	42 (68.9)	19 (31.1)	NS
Medium	11 (9.6)	8 (72.7)	3 (27.3)	
High	43 (37.4)	34 (79.1)	9 (20.9)	
CD68				
Low	50 (43.5)	36 (72)	14 (28)	NS
Medium	24 (20.9)	17 (70.8)	7 (29.2)	
High	41 (35.7)	31 (75.6)	10 (24.4)	
CD34				
Low	38 (33.3)	30 (76.1)	8 (23.9)	NS
Medium	28 (24.6)	17 (64)	11 (36)	
High	48 (42.1)	37 (77.1)	11 (22.9)	

Figure 6-7 Relationship between RUNX1 status and inflammatory infiltrate/blood vessel invasion in triple-negative breast cancer.

p-value calculated using Chi-squared test (linear by linear association). NS, not significant. Where information is not available on the full cohort (n=115), the number of patients with information is specified in brackets: CD34 (114).

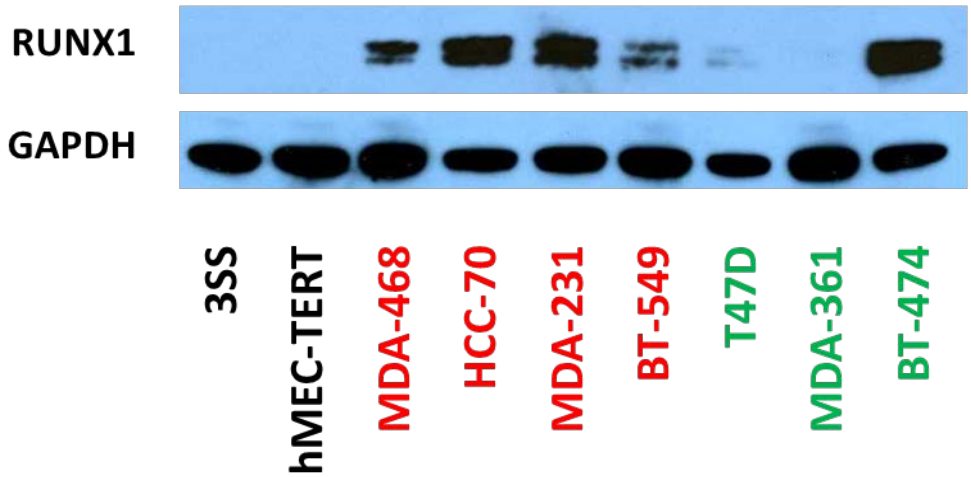


Figure 6-8 Expression of RUNX1 in a panel of human breast cancer cell lines.

RUNX1 western blot on a panel of human breast cell lines: hMEC-TERT (immortalized human mammary epithelial cells), basal-like (red) and luminal-like (green) breast cancer cell lines. 3SS (a leukaemia cell line deleted for RUNX1) was used as a negative control. Human GAPDH was used as a loading control.

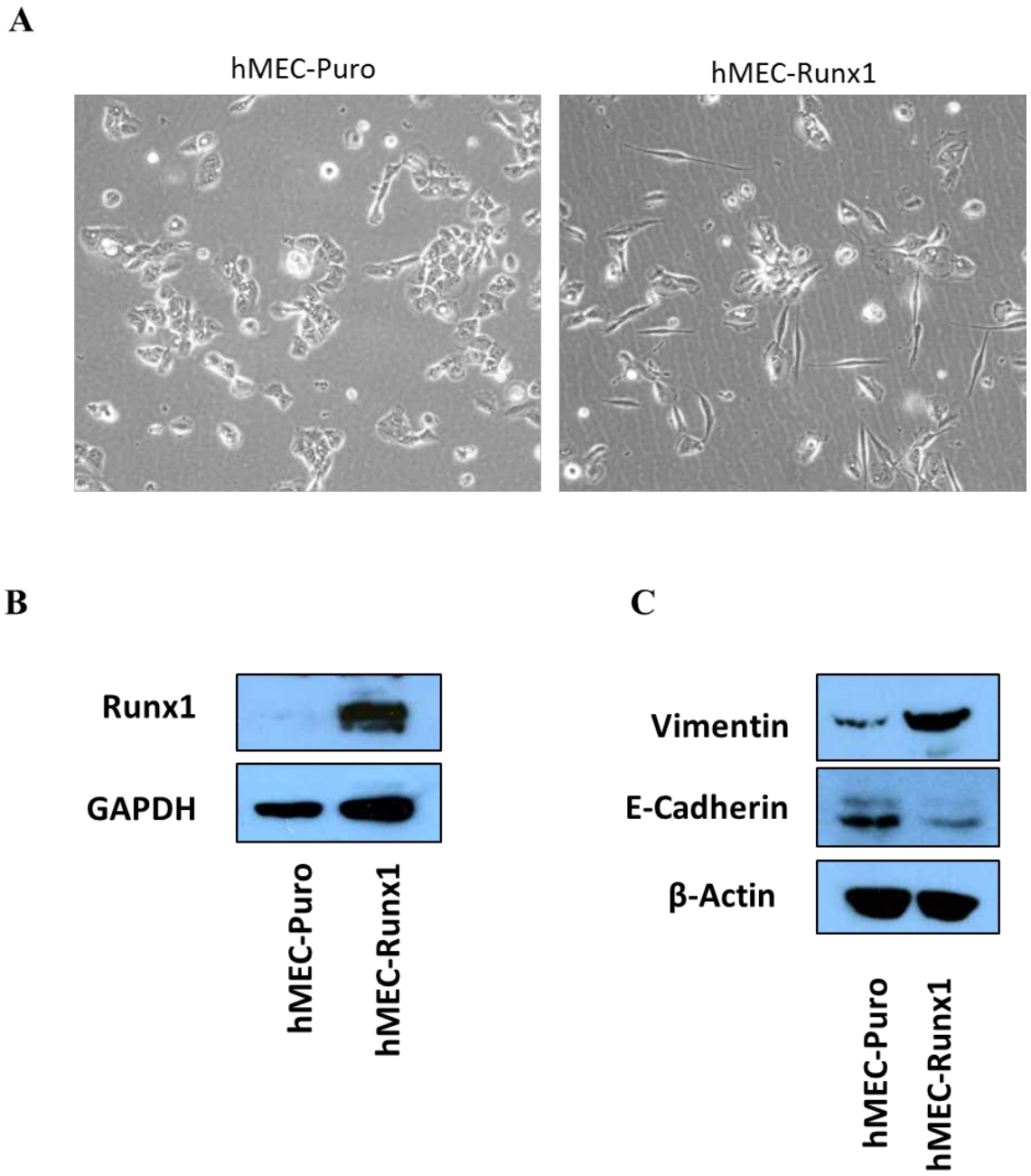


Figure 6-9 *In vitro* effects of RUNX1 overexpression in hMEC-TERT.

Bright field images of hMEC-TERT transfected with pBABE-puro (hMEC-puro) or pBABE-puro-*Runx1* (hMEC-*Runx1*) vector, after 4 passages in culture (A). RUNX1 western blot on hMEC-puro and hMEC-puro-*Runx1*. GAPDH was used as a loading control (B). Western blot for Vimentin and E-Cadherin on hMEC-puro and hMEC-puro-Runx1 cells. Actin was used as loading control (C).

6.3 DISCUSSION

Triple-negative breast cancer (TNBC), which accounts for 15% to 20% of breast cancers, is an aggressive disease, associated with a significantly higher probability of relapse and poorer overall survival when compared with other breast cancer subtypes (Arnedos et al. 2012). Moreover, the lack of identified molecular targets in the majority of TNBCs implies that chemotherapy remains the treatment of choice for patients with TNBC. However early relapse after chemotherapy is common in patients with TNBC (Gluz et al. 2009). Hence there is an urgent need for identification of new prognostic markers which could translate into novel druggable targets specific for this lethal subgroup (Foulkes et al. 2010).

Recently a number of studies indicated RUNX1 as a possible tumour suppressor gene in the mammary epithelium while others suggested a putative oncogenic role (see Chapter 1). To clarify the role for RUNX1 in human breast cancer, invasive breast carcinomas from a tumour tissue microarray were stained for expression of RUNX1. No difference in overall survival was detectable in the full cohort or in ER+, PR+ and HER2+ subgroups. However RUNX1 expression was significantly associated with poorer overall survival in the triple negative (ER-/PR-/HER2-) group of patients. Interestingly this data is supported by several transcriptomic studies which have identified RUNX1 as a possible oncogene in TNBC (Karn et al. 2011; Rody et al. 2011; Lehmann et al. 2011). Moreover *in vitro* analysis showed that all basal-like cell lines tested overexpressed RUNX1 compared to only 1 out 3 luminal-like. Correlation analysis using the Cox regression model, confirmed that RUNX1 is an independent prognostic factor together with size and lymph node status in the TNBC subgroup. Moreover bivariate analysis (Pearson's chi-square) did not find any association between clinicopathological history and RUNX1 expression in TNBC samples confirming the independent prognostic value of RUNX1 expression. Inflammation has been shown to represent a critical component of tumour progression (Grivennikov et al. 2010). Interestingly, RUNX1 expression correlates with the presence of lymphocytic CD4+ infiltrate in TNBC. RUNX1 is one of the key factors that drives various aspects of T-cell differentiation including regulation of cytokine

production (Wong et al. 2012). We could speculate that in TNBC highly positive for RUNX1 that RUNX1 would drive a transcriptional program in breast cancer cells resulting in production and secretion of high levels of cytokines which would then lead to recruitment of lymphocytic cells at the tumour site. Further studies will clarify the significance of the correlation between RUNX1 and CD4 lymphocytes in TNBC. Overall the results indicate that RUNX1 could be a new prognostic biomarker in triple negative breast cancers. The usefulness of RUNX1 as a new prognostic marker is reinforced by the fact that RUNX1 positive samples represent the majority (73%) of the triple negative cohort. Interestingly, no reliable markers have been identified as having a predictive role for the prognosis of TNBC patients so far (Arnedos et al. 2012). So RUNX1 could represent one of the first prognostic factors specific for triple negative breast cancers. Follow-up studies looking at RUNX1 expression in different cohorts of breast cancer patients from different institutions, will help to confirm the prognostic value of RUNX1. RUNX1 overexpression induced EMT in immortalized normal mammary cells suggesting that the worse prognosis in TN patients could be due to increased invasiveness of RUNX1 expressing tumours. These data are in line with reports correlating RUNX1 expression to the claudin-low tumours, a TN subtype characterized by mesenchymal features (Lehmann et al. 2011; Rody et al. 2011). Moreover RUNX2 expression has been shown to induce EMT in a breast cancer cell line (Chimge et al. 2011) suggesting a possible broad role of Runx transcription factors as EMT inducers.

The widespread expression of RUNX1 in TNBC suggests new therapeutic avenues for the treatment of TNBCs. In particular, work from Tumbar's laboratory showed that Runx1 overexpression leads to Stat3 activation and is necessary for skin and oral cancer growth (Scheitz et al. 2012). Interestingly STAT3 is involved In human breast cancer with high STAT3 levels correlating with poorer survival (Diaz et al. 2006). It would be interesting to ascertain if RUNX1 overexpression and STAT3 activation was also conserved in human breast cancer and in TNBC in particular. Further studies are needed to investigate this correlation between STAT3 and RUNX1 expression in TNBC which could pave the way for new treatment options, based on the use of STAT3 inhibitors (Fagard et al. 2013). Moreover the development of small-molecule inhibitors which bind to CBFß and inhibit RUNX1 activity opens the possibility of a RUNX1-specific targeted therapy

for the treatment of TNBC (Gorczynski et al. 2007; Cunningham et al. 2012). Taken together these results are indicating *RUNX1* as a new prognostic biomarker in TNBC and are opening exciting possibilities for the development of new targeted therapies for this lethal subgroup.

7 Conclusions and Future Directions

The mammary gland is a very dynamic tissue composed of different mammary epithelial cell populations with different specialized functions (Visvader & Smith 2010). In order to maintain lineage heterogeneity, various transcription factors are used to determine the correct specification of different mammary subpopulations (Siegel & Muller, 2010). Our current understanding of mammary lineage determination is far from complete; thus the identification of new transcription factors with a role in mammary lineage specification is required to improve our knowledge of normal mammary epithelial development and homeostasis.

This study identifies for the first time RUNX1 and RUNX2, two transcription factors previously linked to haematopoiesis and bone development (Blyth et al. 2005), as new regulators of the mammary epithelium. RUNX1 and RUNX2 follow a similar pattern of expression during mammary development, decreasing during pregnancy and lactation to rise again during involution. In addition, mammary population profiling showed that both RUNX1 and RUNX2 are enriched in the basal compartment. However, only loss of RUNX2 *in vivo* is detrimental for the basal compartment while both genes affect the luminal population. These results suggest a possible independent role for RUNX1 and RUNX2 in the maintenance of specific mammary lineages. The mechanisms driving the loss of RUNX-deleted cells are still not clear and, given the extreme context dependency of *Runx* genes (Blyth et al. 2005), they are likely to be different in divergent compartments of the mammary gland. For example RUNX1 loss in a subpopulation of luminal cells could cause apoptosis while RUNX2 loss in basal cells could cause induction of differentiation. Both *in vivo* and *in vitro* approaches to unravel the role of RUNX1 and RUNX2 in the mammary gland will be informative. Here combined infection of FACS sorted basal and luminal MMECs with lentivirus carrying sh-RNA targeting *Runx1* or *Runx2* could be used to create luminal and basal RUNX knock-out cells. These cells could then be grown in different culture conditions (2D, Matrigel) and analysed for any growth defect when compared to sh-scrambled controls. In addition, transcriptional profiling of

early time points cultures will help to show the main transcriptional changes occurring after RUNX loss in luminal and basal mammary cells.

However the ultimate experiment which will clarify the role of RUNX1 and RUNX2 in the mammary epithelium will be the generation of lineage tracing models. As an example of this technique, Van Amerongen et al., performed lineage tracing to characterize the contribution of *Wnt*-activated cells to different stages of mammary development, using an *Axin2*^{CreERT2} mouse combined with fluorescent reporters (van Amerongen et al. 2012). In the same way the construction of transgenic mice carrying the *Runx1*^{CreERT2} and *Runx2*^{CreERT2} constructs (which will drive an inducible Cre from the endogenous *Runx1* and *Runx2* promoters) combined with GFP and RFP fluorescent reporters will allow us to mark any RUNX1 or RUNX2 positive cell in the entire mammary gland. This will allow a precise localization of RUNX positive cells at different stages of mammary development through a combination of *in vivo* imaging, IHC and FACS population profiling. In addition transcriptional profiling of pure RUNX positive populations, FACS sorted based on fluorescent marker expression, will help to characterize those populations at the molecular level. Overall this *in vivo* approach will help to characterize the role and the dynamics of RUNX positive lineages in the mammary epithelium, with the possibility to further extend lineage tracing analysis to other organs in which RUNX proteins could be implicated, such as prostate (Blyth et al. 2010).

One of the most interesting phenotypes observed in this thesis is the dramatic expansion of the basal population when both *Runx1* and *Runx2* are deleted in the BLG-Cre driven model, especially given that *Runx2* loss alone diminishes this pool. These results suggest the existence of compensatory mechanisms between RUNX proteins in the mammary gland and they are supporting a role for *Runx* genes as lineage regulators in mammary epithelium. Since BLG-Cre is expressed in both luminal and basal compartments, this effect could be due to an expansion of the basal population or a dedifferentiation of luminal cells into the basal lineage. The creation of conditional *Runx1-Runx2* mouse models specific for the basal and the luminal population would address this point. A luminal specific K8-Cre (Van Keymeulen et al. 2011) combined with the *Runx1*$^{flx/flx}$*Runx2*$^{flx/flx}$ mouse and a fluorescent reporter would prove if loss of RUNX1 and RUNX2 is sufficient to drive luminal to basal dedifferentiation *in vivo*.

Conversely, the creation of a double *Runx1* and *Runx2* knock-out in the basal layer of the mammary gland using a K14-Cre would show if combined loss of RUNX1 and RUNX2 can drive an expansion of the basal population *in vivo*. However, the creation of K14-Cre/*Runx1*$^{flx/flx}$ mice was found to be problematic since all animals displayed severe growth defects (data not shown), an interesting phenotype which we have not characterized further. This problem could be avoided by using a Tamoxifen or doxycycline-inducible K14-Cre in combination with the *Runx1*$^{flx/flx}$*Runx2*$^{flx/flx}$ mouse.

Another interesting phenotype was the appearance of preneoplastic lesions in double *Runx1* and *Runx2* knock-out mice. The frequency of lesions (33%) is quite high, taking into consideration the low amount of recombination driven by BLG-Cre in the virgin gland (~15% of recombined MMECs over total). This result is pointing to a tumour suppressive role for RUNX1 and RUNX2 in mammary epithelium which, however, only becomes apparent when both genes are deleted. Importantly, this study is the first to adopt a double conditional knock-out strategy for the combined deletion of two *Runx* genes *in vivo*. The interesting phenotypes observed in this model suggest that the extent to which different *Runx* genes compensate for each other during development of different organs in single knock-out mice remains largely unknown; in the future we can envisage that the use of combined double or triple *Runx* conditional knock-out models will likely identify new compensatory mechanisms between RUNX proteins in the regulation of multiple tissues. Further experiments are needed to characterize the mechanism underlying the formation of neoplastic lesions in aged mice after combined loss of RUNX1 and RUNX2. Fat pad transplants of FACS sorted luminal or basal RUNX1-RUNX2 deleted cells will show if the oncogenic potential lies in the basal or in the luminal compartment. Molecular characterisation of these lesions will also help to identify the signalling pathways which are perturbed (for example *Wnt)*. For this a candidate pathway approach could be done using IHC which has the advantage that spatial information can be made apparent. Alternatively by laser microdissecting the lesions a microarray could be carried out to get a broader overview of the genetic features of the hyperplastic areas.

In Chapter 4 the possible role for RUNX1 and RUNX2 in the mammary stem cell population was examined. Widespread expression of RUNX1 in the mammary

epithelium, and the fact that its expression decreases in mammosphere culture, suggests that this transcription factor could be expressed in a more differentiated population of the virgin mammary gland. Future experiments will need to investigate if RUNX1 expression is necessary for the differentiation of mammary cells. The COMMA-1D cell line could be used to clarify this point. COMMA-1D cell line exhibits several characteristics distinctive of normal mammary epithelial cells, including induction of casein synthesis *in vitro* and normal duct morphogenesis in the cleared mammary fat pads of syngeneic mice (Danielson et al. 1984). The creation of stable RUNX1 deleted and RUNX1 overexpressing COMMA-1D cells, followed by differentiation assays will address if RUNX1 is involved in mammary differentiation *in vitro*. These studies will then be expanded *in vivo* through the use of RUNX1 mammary conditional knock-out models.

In vitro assays indicate that RUNX2 has a role in the regenerative potential of mammary epithelial cells. This is in line with several reports from different systems which indicate RUNX proteins as key regulators of stem cell biology (Appleford & Woollard 2009). However the final proof for a role of RUNX2 in the regulation of mammary regenerative potential *in vivo* is still lacking. Fat pad experiments using pure FACS sorted populations of RUNX2-deleted cells will clarify this point. In particular GFP+ basal cells will be extracted from K14-Cre *Runx2*$^{flx/flx}$ mice and implanted at limiting dilutions in cleared fat pads (Shackleton et al. 2006). This will prove if RUNX2 has a role in mammary regenerative potential *in vivo*. In addition lineage tracing experiments (see above) will help to unveil the real behaviour of the RUNX2 positive mammary population *in vivo*. Single induction of mice carrying the *Runx2*^{CreERT2} allele in combination with a fluorescent reporter at different developmental time points will result in the permanent labelling of RUNX2 positive population in the mammary epithelium. The behaviour of labelled RUNX2 positive cells will then be followed *in vivo* to investigate if they are maintained over time and after several rounds of involution, two classic features of stem cell populations (Van Keymeulen et al. 2011).

The functional role for RUNX2 in a MaSC population is still unknown, but some pieces of evidence indicate that RUNX2 could be linked to regulation of quiescence and *Wnt* signalling in those cells. HC11, a murine mammary

epithelial stem-like cell line, could be used as an *in vitro* model to test RUNX2's molecular regulation of stem cells. In fact *Runx2* transcript has been shown to decrease during the *in vitro* differentiation process of HC11 cells (Williams et al. 2009). The creation of stable HC11 RUNX2 knock-out cells using sh-RNAs, followed by growth analysis, differentiation assays and transcriptional profiling will address possible molecular mechanisms of RUNX2-mediated regulation of mammary stem cells. The actual relevance of RUNX2-dependent pathways identified in HC11 cells could then be confirmed in primary cells, using RUNX2 deleted mammosphere cultures.

Focusing on the link to *Wnt*, RUNX2 could contribute to the regulation of the MaSC population acting as a downstream mediator of *Wnt* signalling, as shown by RUNX2 induction after WNT3A treatments on mammospheres. Future work is needed to understand the role of RUNX2 on the *Wnt* pathway in this cell type. For example, is RUNX2 acting as a repressor or as a co-activator of the *Wnt* signal and what is the outcome of *Wnt*-induced RUNX2 expression in MaSC? Possible answers to those questions come from the *in vivo* studies of aberrant *Wnt* signalling activation in the mammary gland performed in Chapter 5. *In vivo* ß-Catenin stabilization in the mammary gland induces RUNX2 expression and loss of RUNX2 correlates with increased trans-differentiation of the mammary epithelium into squamous metaplastic lesions. This result could suggest that RUNX2 acts as a repressor of *Wnt*-induced squamous differentiation.

In the context of normal mammary development *Wnt* signalling is finely tuned to guarantee the correct balance between MaSC maintenance, luminal progenitor expansion and alveolar differentiation (Brisken et al. 2000; Joshi et al. 2010). During pregnancy, progesterone induces WNT4 production from PR+ cells which then signal in a paracrine fashion to the MaSC compartment stimulating their expansion (Joshi et al., 2010). MaSC expansion during the first phases of pregnancy is probably linked to their differentiation and creation of a large pool of luminal progenitors which will differentiate into alveolar cells (Asselin-Labat et al. 2010). One hypothesis from our results could be that RUNX2 is expressed in a subpopulation of mammary stem cells/progenitors where it is required to maintain their regenerative potential. In this scenario RUNX2 is activated during the waves of progesterone-induced WNT4 which characterize oestrous cycle and act as a gatekeeper of stemness restricting MaSC expansion and differentiation.

During pregnancy instead RUNX2 expression is downregulated and progesterone-induced WNT4 can stimulate MaSC expansion and differentiation in luminal progenitors which will drive the subsequent phases of alveolar differentiation. This scenario is in line with the lactational defects observed after ectopic RUNX2 expression in transgenic MMTV-*Runx2* animals (McDonald, Ferrari *et al*. submitted) where RUNX2 overexpression in the pregnant mammary gland blocks alveolar development. Intriguingly, such a hypothetical role for RUNX2 in mammary lineage specification has a nice parallel in the osteoblast maturation process where during skeletal development RUNX2 activates a differentiation pathway in bone marrow-derived mesenchymal stem cells which is necessary for osteoblast differentiation (Long 2011). Interestingly, RUNX2 expression needs to be downregulated to guarantee a correct osteoblast terminal differentiation (Liu et al. 2001). If this is indeed paralleled in the mammary lineage then the role of RUNX2 as a key lineage specifier, regulating the fine balance of proliferation, quiescence and differentiation in different organs and tissues is highly relevant.

The role of RUNX1 and RUNX2 in breast cancer was discussed in Chapter 5 and 6. Interestingly both RUNX1 and RUNX2 have been linked to the TN subtype. RUNX1 is expressed in the majority of TNBC (~80%) while RUNX2 is detectable in a smaller group of the same cohort (15%, McDonald *et al*, submitted). In addition, no overlap or exclusiveness between the two *RUNX* genes has been detected in the patient cohort analysed (data not shown).

Triple negative breast cancer is an highly heterogeneous group (Metzger-Filho et al. 2012). Metaplastic breast cancer (MBC) is a subtype of TNBC characterized by poor outcome, and few therapeutic options (Lee et al. 2012). Staining of various mouse models of BCs showed squamous metaplastic breast cancer (SMBC) to be highly positive for RUNX2. This result suggests that the 15% of TN cases with high RUNX2 expression identified by McDonald *et al*., could belong to the metaplastic subgroup. To confirm this correlation, it will be interesting to examine human SMBC for RUNX2 to test if this expression pattern found in mouse models translates to the human disease. In addition the use of an *in vivo* model of SMBC (APC^{1572T}) (Gaspar et al. 2009) combined with the $Runx2^{flx/flx}$ model will help to clarify the role of RUNX2 in cancer development and progression. This study could lead to important clinical outcomes such as the identification of RUNX2 as a new biomarker and/or therapeutic target for SMBC.

Bone metastasis are a frequent outcome of aggressive breast cancer, leading to intense pain and severe complications (Suva et al. 2011). For many years RUNX2 has been described as a possible mediator of breast bone metastasis formation based on work from the Stein group (Javed et al. 2005; Barnes et al. 2004). However these studies have several limitations; first they are based on intra-tibia injections, an orthotopic model which cannot recapitulate the seeding and homing of BC cancer cells in a bone microenvironment. Secondly, all these studies have been conducted using a single breast cancer cell line (MDA-MB-231). To address the first point, intracardiac injections, an *in vivo* model of bone metastasis needs to be used (Kretschmann & Welm 2012). In this model MDA-MB-231 deleted for RUNX2 will be injected intracardically in nude mice; this will allow the spreading of BC cells into the entire body of the animal and the development of bone metastasis. In such a system the contribution of RUNX2 to the homing and the growth of MDA-MB-231 cells in the bone microenvironment can be assessed by *in vivo* luciferase imaging and histological analysis on bone lesions. The second point is not currently addressable with current available models (Kretschmann & Welm 2012); in fact excluding MDA-MB-231 cells, no alternative system for the study of breast cancer bone metastasis exist. Since a large percentage of all patients dying of breast cancer(~70%) have evidence of metastatic bone disease (Coleman 2006) more research is needed to address this need.

With recent deep sequencing efforts showing that RUNX1 is a putative tumour suppressor in breast cancer (Ellis et al. 2012; Banerji et al. 2012) it is timely to investigate the role of RUNX1 in human breast cancer as carried out in Chapter 6. For the first time an assessment of RUNX1 protein expression on a large cohort of human breast cancers has been carried out. Interestingly RUNX1 was identified as an independent prognostic factor in TN patients, whose expression correlates with poorer survival. This result suggests a possible oncogenic role for RUNX1 in TNBC and it is in apparent contrast with the recent sequencing studies reporting inactivating *RUNX1* mutations in human BC (Ellis et al. 2012; Banerji et al. 2012). However, Ellis *et al.* was focused on ER+ breast cancers while Banerji *et al.* on BC samples from a mixture of different subtypes. So these studies are not focused on TN breast cancer and that could be a reason why RUNX1 oncogenic role has not been identified. Taken together these studies suggest

that RUNX1 could play different roles in different cancer subtypes: tumour suppressor in luminal-like BC and oncogene in TNBC. The result from the human TMA suggests a possible oncogenic role for RUNX1 in TNBC but cannot exclude if RUNX1 is merely a marker of TNBC characterized by worse prognosis or it has an actual biological role in this subtype. Preliminary *in vitro* experiments tried to address this question; RUNX1 overexpression induced EMT in non-tumourigenic normal human mammary cells suggesting that this could be one of the mechanisms by which RUNX1 expression increases cancer aggressiveness. Taken together our results are indicating Runx1 as a new prognostic biomarker in TNBC. Moreover the availability of small-molecule inhibitors targeting RUNX1 (Gorczynski et al. 2007; Cunningham et al. 2012) are opening exciting possibilities for the development of new targeted therapies for this lethal subgroup.

Bibliography

Aberg, T. et al., 2004. Runx2 mediates FGF signaling from epithelium to mesenchyme during tooth morphogenesis. *Developmental biology*, 270(1), pp.76–93.

Aho, T.L.T. et al., 2006. Pim-1 kinase phosphorylates RUNX family transcription factors and enhances their activity. *BMC cell biology*, 7, p.21.

Akech, J. et al., 2010. Runx2 association with progression of prostate cancer in patients: mechanisms mediating bone osteolysis and osteoblastic metastatic lesions. *Oncogene*, 29(6), pp.811–21.

Ali, S.A. et al., 2012. A RUNX2-HDAC1 co-repressor complex regulates rRNA gene expression by modulating UBF acetylation. *Journal of cell science*, 125(Pt 11), pp.2732–9.

Allred, D.C., Mohsin, S.K. & Fuqua, S.A., 2001. Histological and biological evolution of human premalignant breast disease. *Endocrine-related cancer*, 8(1), pp.47–61.

Andl, T. et al., 2002. WNT signals are required for the initiation of hair follicle development. *Developmental cell*, 2(5), pp.643–53.

Antas, V.I. et al., 2013. Hemogenic endothelium: a vessel for blood production. *The international journal of biochemistry & cell biology*, 45(3), pp.692–5.

Appleford, P.J. & Woollard, A., 2009. RUNX genes find a niche in stem cell biology. *Journal of cellular biochemistry*, 108(1), pp.14–21.

Arnedos, M. et al., 2012. Triple-negative breast cancer: are we making headway at least? *Therapeutic advances in medical oncology*, 4(4), pp.195–210.

Aronson, B.D. et al., 1997. Groucho-dependent and -independent repression activities of Runt domain proteins. *Molecular and cellular biology*, 17(9), pp.5581–7.

Asselin-Labat, M.-L. et al., 2010. Control of mammary stem cell function by steroid hormone signalling. *Nature*, 465(7299), pp.798–802.

Asselin-Labat, M.-L. et al., 2007. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nature cell biology*, 9(2), pp.201–9.

Asselin-Labat, M.-L. et al., 2011. Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. *Molecular and cellular biology*, 31(22), pp.4609–22.

Asselin-Labat, M.-L. et al., 2006. Steroid hormone receptor status of mouse mammary stem cells. *Journal of the National Cancer Institute*, 98(14), pp.1011–4.

Bae, S.C. & Ito, Y., 2003. Comment on Levanon et al., "Runx3 knockouts and stomach cancer", in EMBO reports (June 2003). *EMBO reports*, 4(6), pp.538–9.

Baens, M. et al., 2006. The dark side of EGFP: defective polyubiquitination. D. Davis, ed. *PloS one*, 1(1), p.e54.

Bai, J. et al., 2013. RUNX3 is a prognostic marker and potential therapeutic target in human breast cancer. *Journal of cancer research and clinical oncology*.

Bai, L. & Rohrschneider, L.R., 2010. s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes & development*, 24(17), pp.1882–92.

Bakshi, R. et al., 2010. The human SWI/SNF complex associates with RUNX1 to control transcription of hematopoietic target genes. *Journal of cellular physiology*, 225(2), pp.569–76.

Ball, R.K. et al., 1988. Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line. *The EMBO journal*, 7(7), pp.2089–95.

Bamshad, M. et al., 1997. Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome. *Nature genetics*, 16(3), pp.311–5.

Banerji, S. et al., 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*, 486(7403), pp.405–9.

Bangsow, C. et al., 2001. The RUNX3 gene--sequence, structure and regulated expression. *Gene*, 279(2), pp.221–32.

Baniwal, S.K. et al., 2010. Runx2 transcriptome of prostate cancer cells: insights into invasiveness and bone metastasis. *Molecular cancer*, 9, p.258.

Barnes, G.L. et al., 2004. Fidelity of Runx2 activity in breast cancer cells is required for the generation of metastases-associated osteolytic disease. *Cancer research*, 64(13), pp.4506–13.

Barnes, G.L. et al., 2003. Osteoblast-related Transcription Factors Runx2 (Cbfa1/AML3) and MSX2 Mediate the Expression of Bone Sialoprotein in Human Metastatic Breast Cancer Cells. *Cancer Res.*, 63(10), pp.2631–2637.

Baxter, F.O. et al., 2006. IKKbeta/2 induces TWEAK and apoptosis in mammary epithelial cells. *Development (Cambridge, England)*, 133(17), pp.3485–94.

Berardi, M.J. et al., 1999. The Ig fold of the core binding factor alpha Runt domain is a member of a family of structurally and functionally related Ig-

fold DNA-binding domains. *Structure (London, England : 1993)*, 7(10), pp.1247–56.

Bertucci, F. et al., 2006. Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers. *Cancer research*, 66(9), pp.4636–44.

Biggs, J.R. et al., 2006. AML1/RUNX1 phosphorylation by cyclin-dependent kinases regulates the degradation of AML1/RUNX1 by the anaphase-promoting complex. *Molecular and cellular biology*, 26(20), pp.7420–9.

Biteau, B., Hochmuth, C.E. & Jasper, H., 2011. Maintaining Tissue Homeostasis: Dynamic Control of Somatic Stem Cell Activity. *Cell Stem Cell*, 9(5), pp.402–411.

Blackburn, D.G., 1993. Lactation: historical patterns and potential for manipulation. *Journal of dairy science*, 76(10), pp.3195–212.

Blanpain, C. et al., 2012. Stem cells assessed. *Nature reviews. Molecular cell biology*, 13(7), pp.471–6.

Blanpain, C. & Fuchs, E., 2009. Epidermal homeostasis: a balancing act of stem cells in the skin. *Nature reviews. Molecular cell biology*, 10(3), pp.207–17.

Blyth, K. et al., 2006. Runx2 and MYC collaborate in lymphoma development by suppressing apoptotic and growth arrest pathways in vivo. *Cancer research*, 66(4), pp.2195–201.

Blyth, K. et al., 2010. Runx2 in normal tissues and cancer cells: A developing story. *Blood cells, molecules & diseases*, 45(2), pp.117–23.

Blyth, K. et al., 2001. Runx2: a novel oncogenic effector revealed by in vivo complementation and retroviral tagging. *Oncogene*, 20(3), pp.295–302.

Blyth, K., Cameron, E.R. & Neil, J.C., 2005. The RUNX genes: gain or loss of function in cancer. *Nature reviews. Cancer*, 5(5), pp.376–87.

Bone, K.R. et al., 2010. Translation regulation of Runx3. *Blood cells, molecules & diseases*, 45(2), pp.112–6.

Bonnette, S.G. & Hadsell, D.L., 2001. Targeted disruption of the IGF-I receptor gene decreases cellular proliferation in mammary terminal end buds. *Endocrinology*, 142(11), pp.4937–45.

Boras-Granic, K. et al., 2006. Lef1 is required for the transition of Wnt signaling from mesenchymal to epithelial cells in the mouse embryonic mammary gland. *Developmental biology*, 295(1), pp.219–31.

Boulanger, C.A., Wagner, K.-U. & Smith, G.H., 2005. Parity-induced mouse mammary epithelial cells are pluripotent, self-renewing and sensitive to TGF-beta1 expression. *Oncogene*, 24(4), pp.552–60.

Boumah, C.E. et al., 2009. Runx2 recruits p300 to mediate parathyroid hormone's effects on histone acetylation and transcriptional activation of the matrix metalloproteinase-13 gene. *Molecular endocrinology (Baltimore, Md.)*, 23(8), pp.1255–63.

Bouras, T. et al., 2008. Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. *Cell stem cell*, 3(4), pp.429–41.

Braun, T. & Woollard, A., 2009. RUNX factors in development: lessons from invertebrate model systems. *Blood cells, molecules & diseases*, 43(1), pp.43–8.

Brisken, C. et al., 1998. A paracrine role for the epithelial progesterone receptor in mammary gland development. *Proceedings of the National Academy of Sciences of the United States of America*, 95(9), pp.5076–81.

Brisken, C. et al., 2000. Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. *Genes & development*, 14(6), pp.650–4.

Brisken, C. et al., 1999. Prolactin controls mammary gland development via direct and indirect mechanisms. *Developmental biology*, 210(1), pp.96–106.

Brisken, C. & O'Malley, B., 2010. Hormone action in the mammary gland. *Cold Spring Harbor perspectives in biology*, 2(12), p.a003178.

Bronckers, A.L. et al., 2001. Cell-specific patterns of Cbfa1 mRNA and protein expression in postnatal murine dental tissues. *Mechanisms of development*, 101(1-2), pp.255–8.

Burns, C.E. et al., 2005. Hematopoietic stem cell fate is established by the Notch-Runx pathway. *Genes & development*, 19(19), pp.2331–42.

Cadigan, K.M. & Peifer, M., 2009. Wnt signaling from development to disease: insights from model systems. *Cold Spring Harbor perspectives in biology*, 1(2), p.a002881.

Camilleri, S. & McDonald, F., 2006. Runx2 and dental development. *European journal of oral sciences*, 114(5), pp.361–73.

Cancer Genome Atlas Network, 2012. Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418), pp.61–70.

Castilla, L.H. et al., 2004. Identification of genes that synergize with Cbfb-MYH11 in the pathogenesis of acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 101(14), pp.4924–9.

Chakrabarti, R. et al., 2012. Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem cells (Dayton, Ohio)*, 30(7), pp.1496–508.

Chapman, R.S. et al., 1999. Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. *Genes & development*, 13(19), pp.2604–16.

Chen, A.I., de Nooij, J.C. & Jessell, T.M., 2006. Graded activity of transcription factor Runx3 specifies the laminar termination pattern of sensory axons in the developing spinal cord. *Neuron*, 49(3), pp.395–408.

Chen, F. et al., 2013. RUNX3 suppresses migration, invasion and angiogenesis of human renal cell carcinoma. R. Samant, ed. *PloS one*, 8(2), p.e56241.

Chen, M.J. et al., 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature*, 457(7231), pp.887–91.

Cheng, T. et al., 2000. Hematopoietic stem cell quiescence maintained by p21cip1/waf1. *Science (New York, N.Y.)*, 287(5459), pp.1804–8.

Chepko, G. & Smith, G.H., 1997. Three division-competent, structurally-distinct cell populations contribute to murine mammary epithelial renewal. *Tissue & cell*, 29(2), pp.239–53.

Chi, X.-Z. et al., 2009. Runt-related transcription factor RUNX3 is a target of MDM2-mediated ubiquitination. *Cancer research*, 69(20), pp.8111–9.

Chimge, N.-O. et al., 2012. Opposing effects of Runx2 and estradiol on breast cancer cell proliferation: in vitro identification of reciprocally regulated gene signature related to clinical letrozole responsiveness. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 18(3), pp.901–11.

Chimge, N.-O. et al., 2011. Regulation of breast cancer metastasis by Runx2 and estrogen signaling: the role of SNAI2. *Breast cancer research : BCR*, 13(6), p.R127.

Chimge, N.-O. & Frenkel, B., 2012. The RUNX family in breast cancer: relationships with estrogen signaling. *Oncogene*.

Choi, J.Y. et al., 2001. Subnuclear targeting of Runx/Cbfa/AML factors is essential for tissue-specific differentiation during embryonic development. *Proceedings of the National Academy of Sciences of the United States of America*, 98(15), pp.8650–5.

Chu, E.Y. et al., 2004. Canonical WNT signaling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis. *Development (Cambridge, England)*, 131(19), pp.4819–29.

Chua, C.-W. et al., 2009. Suppression of androgen-independent prostate cancer cell aggressiveness by FTY720: validating Runx2 as a potential antimetastatic drug screening platform. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 15(13), pp.4322–35.

Chuang, L.S.H., Ito, K. & Ito, Y., 2012. RUNX family: Regulation and diversification of roles through interacting proteins. *International journal of cancer. Journal international du cancer*.

Chuang, L.S.H. & Ito, Y., 2010. RUNX3 is multifunctional in carcinogenesis of multiple solid tumors. *Oncogene*, 29(18), pp.2605–15.

Chwieralski, C.E., Welte, T. & Bühling, F., 2006. Cathepsin-regulated apoptosis. *Apoptosis : an international journal on programmed cell death*, 11(2), pp.143–9.

Cicalese, A. et al., 2009. The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell*, 138(6), pp.1083–95.

Cinnamon, E. & Paroush, Z., 2008. Context-dependent regulation of Groucho/TLE-mediated repression. *Current opinion in genetics & development*, 18(5), pp.435–40.

Clarkson, R.W.E. et al., 2004. Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. *Breast cancer research : BCR*, 6(2), pp.R92–109.

Coffman, J.A. et al., 2004. Evaluation of developmental phenotypes produced by morpholino antisense targeting of a sea urchin Runx gene. *BMC biology*, 2, p.6.

Coffman, J.A., 2009. Is Runx a linchpin for developmental signaling in metazoans? *Journal of cellular biochemistry*, 107(2), pp.194–202.

Cohen, M.M., 2009. Perspectives on RUNX genes: an update. *American journal of medical genetics. Part A*, 149A(12), pp.2629–46.

Coleman, R.E., 2006. Clinical features of metastatic bone disease and risk of skeletal morbidity. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 12(20 Pt 2), p.6243s–6249s.

Coleman, S., Silberstein, G.B. & Daniel, C.W., 1988. Ductal morphogenesis in the mouse mammary gland: evidence supporting a role for epidermal growth factor. *Developmental biology*, 127(2), pp.304–15.

Colla, S. et al., 2005. Human myeloma cells express the bone regulating gene Runx2/Cbfa1 and produce osteopontin that is involved in angiogenesis in multiple myeloma patients. *Leukemia*, 19(12), pp.2166–76.

Colleoni, M. et al., 2012. Outcome of special types of luminal breast cancer. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*, 23(6), pp.1428–36.

Cooper, C.L. et al., 2013. Molecular alterations in metaplastic breast carcinoma. *Journal of clinical pathology*, 66(6), pp.522–8.

Coussens, L.M. & Pollard, J.W., 2011. Leukocytes in mammary development and cancer. *Cold Spring Harbor perspectives in biology*, 3(3), p.a003285–.

Cowin, P. & Wysolmerski, J., 2010. Molecular mechanisms guiding embryonic mammary gland development. *Cold Spring Harbor perspectives in biology*, 2(6), p.a003251.

Cui, C. Bin et al., 2003. Transcriptional coactivation of bone-specific transcription factor Cbfa1 by TAZ. *Molecular and cellular biology*, 23(3), pp.1004–13.

Cunningham, L. et al., 2012. Identification of benzodiazepine Ro5-3335 as an inhibitor of CBF leukemia through quantitative high throughput screen against RUNX1-CBFβ interaction. *Proceedings of the National Academy of Sciences of the United States of America*, 109(36), pp.14592–7.

Curtis, C. et al., 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403), pp.346–52.

D'Souza, R.N. et al., 1999. Cbfa1 is required for epithelial-mesenchymal interactions regulating tooth development in mice. *Development (Cambridge, England)*, 126(13), pp.2911–20.

Daga, A. et al., 1996. Patterning of cells in the Drosophila eye by Lozenge, which shares homologous domains with AML1. *Genes & development*, 10(10), pp.1194–205.

Dairkee, S.H. et al., 2004. A molecular "signature" of primary breast cancer cultures; patterns resembling tumor tissue. *BMC genomics*, 5(1), p.47.

Dalle Carbonare, L. et al., 2012. Runx2 mRNA expression in the tissue, serum, and circulating non-hematopoietic cells of patients with thyroid cancer. *The Journal of clinical endocrinology and metabolism*, 97(7), pp.E1249–56.

Daniel, C.W. et al., 1968. The in vivo life span of normal and preneoplastic mouse mammary glands: a serial transplantation study. *Proceedings of the National Academy of Sciences of the United States of America*, 61(1), pp.53–60.

Daniel, C.W., Silberstein, G.B. & Strickland, P., 1987. Direct action of 17 beta-estradiol on mouse mammary ducts analyzed by sustained release implants and steroid autoradiography. *Cancer research*, 47(22), pp.6052–7.

Das, K. et al., 2009. Positive association between nuclear Runx2 and oestrogen-progesterone receptor gene expression characterises a biological subtype of breast cancer. *European journal of cancer (Oxford, England : 1990)*, 45(13), pp.2239–48.

Dassule, H.R. et al., 2000. Sonic hedgehog regulates growth and morphogenesis of the tooth. *Development (Cambridge, England)*, 127(22), pp.4775–85.

Davenport, T.G., Jerome-Majewska, L.A. & Papaioannou, V.E., 2003. Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in

human ulnar mammary syndrome. *Development (Cambridge, England)*, 130(10), pp.2263–73.

DeBruin, L.S. & Josephy, P.D., 2002. Perspectives on the chemical etiology of breast cancer. *Environmental health perspectives*, 110 Suppl , pp.119–28.

DeOme, K.B. et al., 1978. Detection of inapparent nodule-transformed cells in the mammary gland tissues of virgin female BALB/cfC3H mice. *Cancer research*, 38(7), pp.2103–11.

DEOME, K.B. et al., 1959. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer research*, 19(5), pp.515–20.

DEVITT, J.E., 1965. THE SIGNIFICANCE OF REGIONAL LUMPH NODE METASTASES IN BREAST CARCINOMA. *Canadian Medical Association journal*, 93, pp.289–93.

Diaz, N. et al., 2006. Activation of stat3 in primary tumors from high-risk breast cancer patients is associated with elevated levels of activated SRC and survivin expression. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 12(1), pp.20–8.

Dickey-Sims, C. et al., 2005. Runx-dependent expression of PKC is critical for cell survival in the sea urchin embryo. *BMC biology*, 3(1), p.18.

Dijkman, R. et al., 2007. Gene-expression profiling and array-based CGH classify CD4+CD56+ hematodermic neoplasm and cutaneous myelomonocytic leukemia as distinct disease entities. *Blood*, 109(4), pp.1720–7.

Dong, Y.-F. et al., 2006. Wnt induction of chondrocyte hypertrophy through the Runx2 transcription factor. *Journal of cellular physiology*, 208(1), pp.77–86.

Dontu, G. et al., 2003. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes & development*, 17(10), pp.1253–70.

Drissi, H. et al., 2000. Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. *Journal of cellular physiology*, 184(3), pp.341–50.

Ducy, P. et al., 1999. A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development. *Genes & development*, 13(8), pp.1025–36.

Duffy, J.B., Kania, M.A. & Gergen, J.P., 1991. Expression and function of the Drosophila gene runt in early stages of neural development. *Development (Cambridge, England)*, 113(4), pp.1223–30.

Dunbar, M.E. et al., 1999. Parathyroid hormone-related protein signaling is necessary for sexual dimorphism during embryonic mammary development. *Development (Cambridge, England)*, 126(16), pp.3485–93.

Durst, K.L. & Hiebert, S.W., 2004. Role of RUNX family members in transcriptional repression and gene silencing. *Oncogene*, 23(24), pp.4220–4.

Ehrhardt, G.R.A. et al., 2008. Discriminating gene expression profiles of memory B cell subpopulations. *The Journal of experimental medicine*, 205(8), pp.1807–17.

Eliseev, R.A. et al., 2008. Runx2-mediated activation of the Bax gene increases osteosarcoma cell sensitivity to apoptosis. *Oncogene*, 27(25), pp.3605–14.

Ellis, M.J. et al., 2012. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature*, 486(7403), pp.353–60.

Endo, T., Ohta, K. & Kobayashi, T., 2008. Expression and function of Cbfa-1/Runx2 in thyroid papillary carcinoma cells. *The Journal of clinical endocrinology and metabolism*, 93(6), pp.2409–12.

Exton, M.S. et al., 2001. Coitus-induced orgasm stimulates prolactin secretion in healthy subjects. *Psychoneuroendocrinology*, 26(3), pp.287–294.

Fackenthal, J.D. & Olopade, O.I., 2007. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nature reviews. Cancer*, 7(12), pp.937–48.

Fagard, R. et al., 2013. STAT3 inhibitors for cancer therapy: Have all roads been explored? *JAK-STAT*, 2(1), p.e22882.

Fata, J.E. et al., 2001. Accelerated apoptosis in the Timp-3-deficient mammary gland. *The Journal of clinical investigation*, 108(6), pp.831–41.

Fernandez-Guerra, A. et al., 2006. The genomic repertoire for cell cycle control and DNA metabolism in S. purpuratus. *Developmental biology*, 300(1), pp.238–51.

Fernandez-Valdivia, R. et al., 2009. The RANKL signaling axis is sufficient to elicit ductal side-branching and alveologenesis in the mammary gland of the virgin mouse. *Developmental biology*, 328(1), pp.127–39.

Fernandez-Valdivia, R. et al., 2008. Transcriptional response of the murine mammary gland to acute progesterone exposure. *Endocrinology*, 149(12), pp.6236–50.

Ferrari, N. et al., 2013. RUNX2 in mammary gland development and breast cancer. *Journal of cellular physiology*, 228(6), pp.1137–42.

Fijneman, R.J.A. et al., 2011. Runx1 is a tumor suppressor gene in the mouse gastrointestinal tract. *Cancer science*.

Flores, M.V. et al., 2006. A hierarchy of Runx transcription factors modulate the onset of chondrogenesis in craniofacial endochondral bones in zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists*, 235(11), pp.3166–76.

Flores, M.V. et al., 2004. Duplicate zebrafish runx2 orthologues are expressed in developing skeletal elements. *Gene expression patterns : GEP*, 4(5), pp.573–81.

Foley, J. et al., 2001. Parathyroid hormone-related protein maintains mammary epithelial fate and triggers nipple skin differentiation during embryonic breast development. *Development (Cambridge, England)*, 128(4), pp.513–25.

Fortini, M.E., 2009. Notch signaling: the core pathway and its posttranslational regulation. *Developmental cell*, 16(5), pp.633–47.

Foulkes, W.D., Smith, I.E. & Reis-Filho, J.S., 2010. Triple-negative breast cancer. *The New England journal of medicine*, 363(20), pp.1938–48.

Galindo, M. et al., 2005. The bone-specific expression of Runx2 oscillates during the cell cycle to support a G1-related antiproliferative function in osteoblasts. *The Journal of biological chemistry*, 280(21), pp.20274–85.

Gallego, M.I. et al., 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(1), pp.163–75.

Gao, H. et al., 2012. The BMP Inhibitor Coco Reactivates Breast Cancer Cells at Lung Metastatic Sites. *Cell*, 150(4), pp.764–779.

Gasco, M., Shami, S. & Crook, T., 2002. The p53 pathway in breast cancer. *Breast Cancer Res*, 4(2), pp.70–76.

Gaspar, C. et al., 2009. A targeted constitutive mutation in the APC tumor suppressor gene underlies mammary but not intestinal tumorigenesis. *PLoS genetics*, 5(7), p.e1000547.

Gat, U. et al., 1998. De Novo Hair Follicle Morphogenesis and Hair Tumors in Mice Expressing a Truncated ß-Catenin in Skin. *Cell*, 95(5), pp.605–614.

Gaur, T. et al., 2005. Canonical WNT signaling promotes osteogenesis by directly stimulating Runx2 gene expression. *The Journal of Biological Chemistry*, 280(39), pp.33132–33140.

Ge, C. et al., 2012. Interactions between extracellular signal-regulated kinase 1/2 and p38 MAP kinase pathways in the control of RUNX2 phosphorylation and transcriptional activity. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*, 27(3), pp.538–51.

Van Genderen, C. et al., 1994. Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes & development*, 8(22), pp.2691–703.

Geoffroy, V. et al., 2002. High bone resorption in adult aging transgenic mice overexpressing cbfa1/runx2 in cells of the osteoblastic lineage. *Molecular and cellular biology*, 22(17), pp.6222–33.

Gerber, B., Freund, M. & Reimer, T., 2010. Recurrent breast cancer: treatment strategies for maintaining and prolonging good quality of life. *Deutsches Ärzteblatt international*, 107(6), pp.85–91.

Gergen, J.P. & Butler, B.A., 1988. Isolation of the Drosophila segmentation gene runt and analysis of its expression during embryogenesis. *Genes & development*, 2(9), pp.1179–93.

Geyer, F.C. et al., 2011. ß-Catenin pathway activation in breast cancer is associated with triple-negative phenotype but not with CTNNB1 mutation. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*, 24(2), pp.209–31.

Glotzer, D.J., Zelzer, E. & Olsen, B.R., 2008. Impaired skin and hair follicle development in Runx2 deficient mice. *Developmental biology*, 315(2), pp.459–73.

Gluz, O. et al., 2009. Triple-negative breast cancer--current status and future directions. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*, 20(12), pp.1913–27.

Goh, Y.-M. et al., 2010. Src kinase phosphorylates RUNX3 at tyrosine residues and localizes the protein in the cytoplasm. *The Journal of biological chemistry*, 285(13), pp.10122–9.

Golub, T.R. et al., 1995. Fusion of the TEL gene on 12p13 to the AML1 gene on 21q22 in acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 92(11), pp.4917–21.

Gorczynski, M.J. et al., 2007. Allosteric Inhibition of the Protein-Protein Interaction between the Leukemia-Associated Proteins Runx1 and CBFß. *Chemistry & Biology*, 14(10), pp.1186–1197.

Gouon-Evans, V., Rothenberg, M. & Pollard, J., 2000. Postnatal mammary gland development requires macrophages and eosinophils. *Development*, 127(11), pp.2269–2282.

Gray, J. & Druker, B., 2012. Genomics: the breast cancer landscape. *Nature*, 486(7403), pp.328–9.

Green, R.E. et al., 2010. A draft sequence of the Neandertal genome. *Science (New York, N.Y.)*, 328(5979), pp.710–22.

Grivennikov, S.I., Greten, F.R. & Karin, M., 2010. Immunity, inflammation, and cancer. *Cell*, 140(6), pp.883–99.

Grompe, M., 2012. Tissue stem cells: new tools and functional diversity. *Cell stem cell*, 10(6), pp.685–9.

Growney, J.D. et al., 2005. Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood*, 106(2), pp.494–504.

Guarneri, V. & Conte, P., 2009. Metastatic breast cancer: therapeutic options according to molecular subtypes and prior adjuvant therapy. *The oncologist*, 14(7), pp.645–56.

Guinee, V.F. et al., 1994. Effect of pregnancy on prognosis for young women with breast cancer. *Lancet*, 343(8913), pp.1587–9.

Guo, W. et al., 2012. Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell*, 148(5), pp.1015–28.

Gusterson, B., 2009. Do "basal-like" breast cancers really exist? *Nature reviews. Cancer*, 9(2), pp.128–34.

Guy, C.T., Cardiff, R.D. & Muller, W.J., 1992. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Molecular and cellular biology*, 12(3), pp.954–61.

Hankinson, S.E., Colditz, G. a & Willett, W.C., 2004. Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones. *Breast cancer research : BCR*, 6(5), pp.213–8.

Harada, H. et al., 2003. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood*, 101(2), pp.673–80.

Harada, N. et al., 1999. Intestinal polyposis in mice with a dominant stable mutation of the beta-catenin gene. *The EMBO journal*, 18(21), pp.5931–42.

Haslam, S.Z. & Shyamala, G., 1981. Relative distribution of estrogen and progesterone receptors among the epithelial, adipose, and connective tissue components of the normal mammary gland. *Endocrinology*, 108(3), pp.825–30.

Hassiotou, F. & Geddes, D., 2013. Anatomy of the human mammary gland: Current status of knowledge. *Clinical anatomy (New York, N.Y.)*, 26(1), pp.29–48.

Hatsell, S.J. & Cowin, P., 2006. Gli3-mediated repression of Hedgehog targets is required for normal mammary development. *Development (Cambridge, England)*, 133(18), pp.3661–70.

Hayes, M.J. et al., 2008. Genetic changes of Wnt pathway genes are common events in metaplastic carcinomas of the breast. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 14(13), pp.4038–44.

Heldring, N. et al., 2007. Estrogen receptors: how do they signal and what are their targets. *Physiological reviews*, 87(3), pp.905–31.

Hennessy, B.T. et al., 2009. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer research*, 69(10), pp.4116–24.

Hennighausen, L. & Robinson, G.W., 2005. Information networks in the mammary gland. *Nature reviews. Molecular cell biology*, 6(9), pp.715–25.

Hens, J.R. et al., 2007. BMP4 and PTHrP interact to stimulate ductal outgrowth during embryonic mammary development and to inhibit hair follicle induction. *Development (Cambridge, England)*, 134(6), pp.1221–30.

Herschkowitz, J.I. et al., 2007. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome biology*, 8(5), p.R76.

Hiebert, S.W. et al., 1996. The t(12;21) translocation converts AML-1B from an activator to a repressor of transcription. *Molecular and cellular biology*, 16(4), pp.1349–55.

Higgins, M.J. & Baselga, J., 2011. Targeted therapies for breast cancer. *The Journal of clinical investigation*, 121(10), pp.3797–803.

Hilton, M.J. et al., 2008. Notch signaling maintains bone marrow mesenchymal progenitors by suppressing osteoblast differentiation. *Nature medicine*, 14(3), pp.306–14.

Hoi, C.S.L. et al., 2010. Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin. *Molecular and cellular biology*, 30(10), pp.2518–36.

Holliday, D.L. & Speirs, V., 2011. Choosing the right cell line for breast cancer research. *Breast cancer research : BCR*, 13(4), p.215.

HOSHINO, K. & GARDNER, W.U., 1967. Transplantability and Life Span of Mammary Gland during Serial Transplantation in Mice. *Nature*, 213(5072), pp.193–194.

Howard, J.H. & Bland, K.I., 2012. Current management and treatment strategies for breast cancer. *Current opinion in obstetrics & gynecology*, 24(1), pp.44–8.

Huang, B. et al., 2012. RUNX3 acts as a tumor suppressor in breast cancer by targeting estrogen receptor α. *Oncogene*, 31(4), pp.527–34.

Huang, G. et al., 2001. Dimerization with PEBP2beta protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation. *The EMBO journal*, 20(4), pp.723–33.

Huelsken, J. et al., 2001. beta-Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin. *Cell*, 105(4), pp.533–45.

Humphreys, R.C. et al., 2002. Deletion of Stat3 blocks mammary gland involution and extends functional competence of the secretory epithelium in the absence of lactogenic stimuli. *Endocrinology*, 143(9), pp.3641–50.

Hynes, N.E. & Watson, C.J., 2010. Mammary gland growth factors: roles in normal development and in cancer. *Cold Spring Harbor perspectives in biology*, 2(8), p.a003186.

Ichikawa, M. et al., 2004. AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nature medicine*, 10(3), pp.299–304.

Imai, Y. et al., 1998. TLE, the human homolog of groucho, interacts with AML1 and acts as a repressor of AML1-induced transactivation. *Biochemical and biophysical research communications*, 252(3), pp.582–9.

Inman, C.K., Li, N. & Shore, P., 2005. Oct-1 counteracts autoinhibition of Runx2 DNA binding to form a novel Runx2/Oct-1 complex on the promoter of the mammary gland-specific gene beta-casein. *Molecular and cellular biology*, 25(8), pp.3182–93.

Inman, C.K. & Shore, P., 2003. The osteoblast transcription factor Runx2 is expressed in mammary epithelial cells and mediates osteopontin expression. *The Journal of biological chemistry*, 278(49), pp.48684–9.

Inoue, K. et al., 2002. Runx3 controls the axonal projection of proprioceptive dorsal root ganglion neurons. *Nature Neuroscience*, 5(10), pp.946–954.

Ismail, P.M. et al., 2002. A novel LacZ reporter mouse reveals complex regulation of the progesterone receptor promoter during mammary gland development. *Molecular endocrinology (Baltimore, Md.)*, 16(11), pp.2475–89.

Ito, K. et al., 2011. Loss of Runx3 is a key event in inducing precancerous state of the stomach. *Gastroenterology*, 140(5), pp.1536–46.e8.

Ito, K. et al., 2008. RUNX3 attenuates beta-catenin/T cell factors in intestinal tumorigenesis. *Cancer cell*, 14(3), pp.226–37.

Ito, Y. & Miyazono, K., 2003. RUNX transcription factors as key targets of TGF-beta superfamily signaling. *Current opinion in genetics & development*, 13(1), pp.43–7.

Janes, K.A., 2011. RUNX1 and its understudied role in breast cancer. *Cell cycle (Georgetown, Tex.)*, 10(20).

Javed, A. et al., 2000. Groucho/TLE/R-esp proteins associate with the nuclear matrix and repress RUNX (CBF(alpha)/AML/PEBP2(alpha)) dependent activation of tissue-specific gene transcription. *Journal of cell science*, 113 (Pt 1, pp.2221–31.

Javed, A. et al., 2005. Impaired intranuclear trafficking of Runx2 (AML3/CBFA1) transcription factors in breast cancer cells inhibits osteolysis in vivo.

Proceedings of the National Academy of Sciences of the United States of America, 102(5), pp.1454–9.

Jemal, A. et al., Cancer statistics, 2010. *CA: a cancer journal for clinicians*, 60(5), pp.277–300.

Jenkins, D.E. et al., 2005. Bioluminescent human breast cancer cell lines that permit rapid and sensitive in vivo detection of mammary tumors and multiple metastases in immune deficient mice. *Breast cancer research : BCR*, 7(4), pp.R444–54.

Jensen, K.B. et al., 2009. Lrig1 expression defines a distinct multipotent stem cell population in mammalian epidermis. *Cell stem cell*, 4(5), pp.427–39.

Jeon, E.-J. et al., 2006. Bone morphogenetic protein-2 stimulates Runx2 acetylation. *The Journal of biological chemistry*, 281(24), pp.16502–11.

Jeong, J.-H. et al., 2008. Expression of Runx2 transcription factor in non-skeletal tissues, sperm and brain. *Journal of cellular physiology*, 217(2), pp.511–7.

Jho, E. et al., 2002. Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Molecular and cellular biology*, 22(4), pp.1172–83.

Ji, C. et al., 1998. CBFa(AML/PEBP2)-related elements in the TGF-beta type I receptor promoter and expression with osteoblast differentiation. *Journal of cellular biochemistry*, 69(3), pp.353–63.

Jiang, H. et al., 1999. Expression of core binding factor Osf2/Cbfa-1 and bone sialoprotein in tooth development. *Mechanisms of development*, 81(1-2), pp.169–73.

Jiang, Y. et al., 2008. Expression of RUNX3 gene, methylation status and clinicopathological significance in breast cancer and breast cancer cell lines. *Pathobiology : journal of immunopathology, molecular and cellular biology*, 75(4), pp.244–51.

Jiménez-Rojo, L. et al., 2012. Stem Cell Fate Determination during Development and Regeneration of Ectodermal Organs. *Frontiers in physiology*, 3, p.107.

Jin, Y.-H. et al., 2004. Transforming growth factor-beta stimulates p300-dependent RUNX3 acetylation, which inhibits ubiquitination-mediated degradation. *The Journal of biological chemistry*, 279(28), pp.29409–17.

Jordan, H.E., 1916. Evidence of hemogenic capacity of endothelium. *The Anatomical Record*, 10(5), pp.417–420.

Joshi, P.A. et al., 2010. Progesterone induces adult mammary stem cell expansion. *Nature*, 465(7299), pp.803–7.

Kadota, M. et al., 2010. Delineating genetic alterations for tumor progression in the MCF10A series of breast cancer cell lines. *PloS one*, 5(2), p.e9201.

Kagoshima, H., Shigesada, K. & Kohara, Y., 2007. RUNX regulates stem cell proliferation and differentiation: insights from studies of C. elegans. *Journal of cellular biochemistry*, 100(5), pp.1119–30.

Kalev-Zylinska, M.L. et al., 2002. Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. *Development (Cambridge, England)*, 129(8), pp.2015–30.

Kalev-Zylinska, M.L. et al., 2003. Runx3 is required for hematopoietic development in zebrafish. *Developmental dynamics : an official publication of the American Association of Anatomists*, 228(3), pp.323–36.

Kamangar, F., Dores, G.M. & Anderson, W.F., 2006. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 24(14), pp.2137–50.

Karn, T. et al., 2011. Homogeneous Datasets of Triple Negative Breast Cancers Enable the Identification of Novel Prognostic and Predictive Signatures S. Ranganathan, ed. *PLoS ONE*, 6(12), p.e28403.

Keita, M. et al., 2013. The RUNX1 transcription factor is expressed in serous epithelial ovarian carcinoma and contributes to cell proliferation, migration and invasion. *Cell cycle (Georgetown, Tex.)*, 12(6).

Kendrick, H. et al., 2008. Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. *BMC genomics*, 9, p.591.

Kennecke, H. et al., 2010. Metastatic behavior of breast cancer subtypes. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 28(20), pp.3271–7.

Kenney, N.J. et al., 2003. Effect of exogenous epidermal-like growth factors on mammary gland development and differentiation in the estrogen receptor-alpha knockout (ERKO) mouse. *Breast cancer research and treatment*, 79(2), pp.161–73.

Kenney, N.J. et al., 1996. Induction of ductal morphogenesis and lobular hyperplasia by amphiregulin in the mouse mammary gland. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*, 7(12), pp.1769–81.

Kerney, R., Gross, J.B. & Hanken, J., 2007. Runx2 is essential for larval hyobranchial cartilage formation in Xenopus laevis. *Developmental dynamics : an official publication of the American Association of Anatomists*, 236(6), pp.1650–62.

Van Keymeulen, A. et al., 2011. Distinct stem cells contribute to mammary gland development and maintenance. *Nature*.

Khalid, O. et al., 2008. Modulation of Runx2 activity by estrogen receptor-alpha: implications for osteoporosis and breast cancer. *Endocrinology*, 149(12), pp.5984–95.

Kilbey, A. et al., 2008. Oncogene-induced senescence: an essential role for Runx. *Cell cycle (Georgetown, Tex.)*, 7(15), pp.2333–40.

Kimura, A. et al., 2010. Runx1 and Runx2 cooperate during sternal morphogenesis. *Development (Cambridge, England)*, 137(7), pp.1159–67.

Kirkegaard, T. et al., 2006. Observer variation in immunohistochemical analysis of protein expression, time for a change? *Histopathology*, 48(7), pp.787–94.

Kitabayashi, I. et al., 2001. Activation of AML1-mediated transcription by MOZ and inhibition by the MOZ-CBP fusion protein. *The EMBO journal*, 20(24), pp.7184–96.

Kitabayashi, I. et al., 1998. Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. *The EMBO journal*, 17(11), pp.2994–3004.

Kleinberg, D.L. & Ruan, W., 2008. IGF-I, GH, and sex steroid effects in normal mammary gland development. *Journal of mammary gland biology and neoplasia*, 13(4), pp.353–60.

Klintrup, K. et al., 2005. Inflammation and prognosis in colorectal cancer. *European journal of cancer (Oxford, England : 1990)*, 41(17), pp.2645–54.

Koehler, A.N., 2010. A complex task? Direct modulation of transcription factors with small molecules. *Current opinion in chemical biology*, 14(3), pp.331–40.

Komori, T. et al., 1997. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell*, 89(5), pp.755–64.

Korach, K.S. et al., 1996. Estrogen receptor gene disruption: molecular characterization and experimental and clinical phenotypes. *Recent progress in hormone research*, 51, pp.159–86; discussion 186–8.

Kordon, E.C. & Smith, G.H., 1998. An entire functional mammary gland may comprise the progeny from a single cell. *Development (Cambridge, England)*, 125(10), pp.1921–30.

Kouros-Mehr, H. & Werb, Z., 2006. Candidate regulators of mammary branching morphogenesis identified by genome-wide transcript analysis. *Developmental dynamics : an official publication of the American Association of Anatomists*, 235(12), pp.3404–12.

Kramer, I. et al., 2006. A role for Runx transcription factor signaling in dorsal root ganglion sensory neuron diversification. *Neuron*, 49(3), pp.379–93.

Krege, J.H. et al., 1998. Generation and reproductive phenotypes of mice lacking estrogen receptor . *Proceedings of the National Academy of Sciences*, 95(26), pp.15677–15682.

Kretschmann, K.L. & Welm, A.L., 2012. Mouse models of breast cancer metastasis to bone. *Cancer metastasis reviews*, 31(3-4), pp.579–83.

Kreuzaler, P.A. et al., 2011. Stat3 controls lysosomal-mediated cell death in vivo. *Nature cell biology*, 13(3), pp.303–9.

Kritikou, E.A. et al., 2003. A dual, non-redundant, role for LIF as a regulator of development and STAT3-mediated cell death in mammary gland. *Development (Cambridge, England)*, 130(15), pp.3459–68.

Kudo, Y., Tsunematsu, T. & Takata, T., 2011. Oncogenic role of RUNX3 in head and neck cancer. *Journal of cellular biochemistry*, 112(2), pp.387–93.

Kuo, Y.-H. et al., 2009. Runx2 induces acute myeloid leukemia in cooperation with Cbfbeta-SMMHC in mice. *Blood*, 113(14), pp.3323–32.

Kuraguchi, M. et al., 2009. Genetic mechanisms in Apc-mediated mammary tumorigenesis. *PLoS genetics*, 5(2), p.e1000367.

Kurek, K.C. et al., 2010. Frequent attenuation of the WWOX tumor suppressor in osteosarcoma is associated with increased tumorigenicity and aberrant RUNX2 expression. *Cancer research*, 70(13), pp.5577–86.

Lam, K. & Zhang, D.-E., 2012. RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis. *Frontiers in bioscience : a journal and virtual library*, 17, pp.1120–39.

LaMarca, H.L. et al., 2010. CCAAT/enhancer binding protein beta regulates stem cell activity and specifies luminal cell fate in the mammary gland. *Stem cells (Dayton, Ohio)*, 28(3), pp.535–44.

Lau, C.C. et al., 2004. Frequent amplification and rearrangement of chromosomal bands 6p12-p21 and 17p11.2 in osteosarcoma. *Genes, chromosomes & cancer*, 39(1), pp.11–21.

Lau, Q.C. et al., 2006. RUNX3 is frequently inactivated by dual mechanisms of protein mislocalization and promoter hypermethylation in breast cancer. *Cancer research*, 66(13), pp.6512–20.

Lebestky, T. et al., 2000. Specification of Drosophila hematopoietic lineage by conserved transcription factors. *Science (New York, N.Y.)*, 288(5463), pp.146–9.

Lee, H. et al., 2012. Metaplastic breast cancer: clinicopathological features and its prognosis. *Journal of clinical pathology*, 65(5), pp.441–6.

Lee, J. et al., 2013. Runx1 and p21 synergistically limit the extent of hair follicle stem cell quiescence in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 110(12), pp.4634–9.

Lee, J.H. et al., 2011. Expression of RUNX3 in skin cancers. *Clinical and experimental dermatology*, 36(7), pp.769–74.

Lee, K.S. et al., 2000. Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. *Molecular and cellular biology*, 20(23), pp.8783–92.

Lee, K.-S. et al., 2010. Runx3 is required for the differentiation of lung epithelial cells and suppression of lung cancer. *Oncogene*, 29(23), pp.3349–61.

Lee, Y.M., 2011. Control of RUNX3 by histone methyltransferases. *Journal of cellular biochemistry*, 112(2), pp.394–400.

Lee, Y.T., 1983. Breast carcinoma: pattern of metastasis at autopsy. *Journal of surgical oncology*, 23(3), pp.175–80.

Lehmann, B.D. et al., 2011. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of clinical investigation*, 121(7), pp.2750–67.

Leong, D.T. et al., 2010. Cancer-related ectopic expression of the bone-related transcription factor RUNX2 in non-osseous metastatic tumor cells is linked to cell proliferation and motility. *Breast cancer research : BCR*, 12(5), p.R89.

Levanon, D. et al., 2011. Absence of Runx3 expression in normal gastrointestinal epithelium calls into question its tumour suppressor function. *EMBO molecular medicine*, 3(10), pp.593–604.

Levanon, D. et al., 2003. Runx3 knockouts and stomach cancer. *EMBO reports*, 4(6), pp.560–4.

Levanon, D. et al., 2002. The Runx3 transcription factor regulates development and survival of TrkC dorsal root ganglia neurons. *The EMBO journal*, 21(13), pp.3454–63.

Levanon, D. & Groner, Y., 2004. Structure and regulated expression of mammalian RUNX genes. *Oncogene*, 23(24), pp.4211–9.

Li, J. et al., 1999. Leukaemia disease genes: large-scale cloning and pathway predictions. *Nature genetics*, 23(3), pp.348–53.

Li, J. et al., 2004. RUNX3 expression in primary and metastatic pancreatic cancer. *Journal of clinical pathology*, 57(3), pp.294–9.

Li, M. et al., 1997. Mammary-derived signals activate programmed cell death during the first stage of mammary gland involution. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7), pp.3425–30.

Li, Q.L. et al., 2002. Causal relationship between the loss of RUNX3 expression and gastric cancer. *Cell*, 109(1), pp.113–24.

Li, X. et al., 2009. Co-activator activator (CoAA) prevents the transcriptional activity of Runt domain transcription factors. *Journal of cellular biochemistry*, 108(2), pp.378–87.

Li, X. & O'Malley, B.W., 2003. Unfolding the action of progesterone receptors. *The Journal of biological chemistry*, 278(41), pp.39261–4.

Liakhovitskaia, A. et al., 2010. The essential requirement for Runx1 in the development of the sternum. *Developmental biology*, 340(2), pp.539–46.

Lian, J.B. et al., 2003. Runx1/AML1 hematopoietic transcription factor contributes to skeletal development in vivo. *Journal of cellular physiology*, 196(2), pp.301–11.

Lim, E. et al., 2010. Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast cancer research : BCR*, 12(2), p.R21.

Lim, M. et al., 2010. Runx2 regulates survivin expression in prostate cancer cells. *Laboratory investigation; a journal of technical methods and pathology*, 90(2), pp.222–33.

Liu, H., Carlsson, L. & Grundström, T., 2006. Identification of an N-terminal transactivation domain of Runx1 that separates molecular function from global differentiation function. *The Journal of biological chemistry*, 281(35), pp.25659–69.

Liu, W. et al., 2001. Overexpression of Cbfa1 in osteoblasts inhibits osteoblast maturation and causes osteopenia with multiple fractures. *The Journal of cell biology*, 155(1), pp.157–66.

Liu, X. et al., 1997. Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes & development*, 11(2), pp.179–86.

Livasy, C.A. et al., 2006. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*, 19(2), pp.264–71.

Long, F., 2011. Building strong bones: molecular regulation of the osteoblast lineage. *Nature Reviews Molecular Cell Biology*, 13(1), pp.27–38.

Lu, P. et al., 2008. Genetic mosaic analysis reveals FGF receptor 2 function in terminal end buds during mammary gland branching morphogenesis. *Developmental biology*, 321(1), pp.77–87.

Luetteke, N.C. et al., 1999. Targeted inactivation of the EGF and amphiregulin genes reveals distinct roles for EGF receptor ligands in mouse mammary gland development. *Development (Cambridge, England)*, 126(12), pp.2739–50.

Lunyak, V. V & Rosenfeld, M.G., 2008. Epigenetic regulation of stem cell fate. *Human molecular genetics*, 17(R1), pp.R28–36.

Lutterbach, B. et al., 2000. A mechanism of repression by acute myeloid leukemia-1, the target of multiple chromosomal translocations in acute leukemia. *The Journal of biological chemistry*, 275(1), pp.651–6.

Lydon, J.P. et al., 1995. Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities. *Genes & Development*, 9(18), pp.2266–2278.

Mailleux, A.A. et al., 2002. Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo. *Development (Cambridge, England)*, 129(1), pp.53–60.

Makarem, M. et al., 2013. Developmental Changes in the in Vitro Activated Regenerative Activity of Primitive Mammary Epithelial Cells T. Schroeder, ed. *PLoS Biology*, 11(8), p.e1001630.

Malanchi, I. et al., 2011. Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature*, advance on.

Mallepell, S. et al., 2006. Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), pp.2196–201.

Marie, P.J., 2008. Transcription factors controlling osteoblastogenesis. *Archives of biochemistry and biophysics*, 473(2), pp.98–105.

Marmigère, F. et al., 2006. The Runx1/AML1 transcription factor selectively regulates development and survival of TrkA nociceptive sensory neurons. *Nature neuroscience*, 9(2), pp.180–7.

Massagué, J., 2012. TGFβ signalling in context. *Nature reviews. Molecular cell biology*, 13(10), pp.616–30.

Matsuda, M. et al., 2004. Serotonin regulates mammary gland development via an autocrine-paracrine loop. *Developmental cell*, 6(2), pp.193–203.

McNally, S. & Martin, F., 2011. Molecular regulators of pubertal mammary gland development. *Annals of medicine*, 43(3), pp.212–34.

Medina, D., 2002. Biological and molecular characteristics of the premalignant mouse mammary gland. *Biochimica et biophysica acta*, 1603(1), pp.1–9.

Medina, D., 1996. The mammary gland: A unique organ for the study of development and tumorigenesis. *Journal of Mammary Gland Biology and Neoplasia*, 1(1), pp.5–19.

Mendoza-Villanueva, D., Zeef, L. & Shore, P., 2011. Metastatic breast cancer cells inhibit osteoblast differentiation through the Runx2/CBFβ-dependent expression of the Wnt antagonist, sclerostin. *Breast cancer research : BCR*, 13(5), p.R106.

Metzger-Filho, O. et al., 2012. Dissecting the heterogeneity of triple-negative breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 30(15), pp.1879–87.

Meyers, S., Lenny, N. & Hiebert, S.W., 1995. The t(8;21) fusion protein interferes with AML-1B-dependent transcriptional activation. *Molecular and cellular biology*, 15(4), pp.1974–82.

Mikkola, M.L. & Millar, S.E., 2006. The mammary bud as a skin appendage: unique and shared aspects of development. *Journal of mammary gland biology and neoplasia*, 11(3-4), pp.187–203.

Min, B. et al., 2012. Identification of RUNX3 as a component of the MST/Hpo signaling pathway. *Journal of cellular physiology*, 227(2), pp.839–49.

Miyoshi, H. et al., 1991. t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proceedings of the National Academy of Sciences of the United States of America*, 88(23), pp.10431–4.

Miyoshi, K., Shillingford, J.M., et al., 2002. Activation of beta -catenin signaling in differentiated mammary secretory cells induces transdifferentiation into epidermis and squamous metaplasias. *Proceedings of the National Academy of Sciences of the United States of America*, 99(1), pp.219–24.

Miyoshi, K., Rosner, A., et al., 2002. Activation of different Wnt/beta-catenin signaling components in mammary epithelium induces transdifferentiation and the formation of pilar tumors. *Oncogene*, 21(36), pp.5548–56.

Mohammed, Z.M.A., Edwards, J., et al., 2012. Breast cancer outcomes by steroid hormone receptor status assessed visually and by computer image analysis. *Histopathology*, 61(2), pp.283–92.

Mohammed, Zahra M A, Going, J.J., et al., 2012. Comparison of visual and automated assessment of HER2 status and their impact on outcome in primary operable invasive ductal breast cancer. *Histopathology*, 61(4), pp.675–684.

Mohammed, Z.M.A. et al., 2013. Comparison of visual and automated assessment of microvessel density and their impact on outcome in primary operable invasive ductal breast cancer. *Human pathology*, 44(8), pp.1688–95.

Mohammed, Z M A, Going, J.J., et al., 2012. The relationship between components of tumour inflammatory cell infiltrate and clinicopathological factors and survival in patients with primary operable invasive ductal breast cancer. *British journal of cancer*, 107(5), pp.864–73.

Molyneux, G. et al., 2010. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell stem cell*, 7(3), pp.403–17.

Monks, J. et al., 2005. Epithelial cells as phagocytes: apoptotic epithelial cells are engulfed by mammary alveolar epithelial cells and repress inflammatory mediator release. *Cell death and differentiation*, 12(2), pp.107–14.

Montemurro, F., Di Cosimo, S. & Arpino, G., 2013. Human epidermal growth factor receptor 2 (HER2)-positive and hormone receptor-positive breast cancer: new insights into molecular interactions and clinical implications. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*.

Moraes, R.C. et al., 2007. Constitutive activation of smoothened (SMO) in mammary glands of transgenic mice leads to increased proliferation, altered differentiation and ductal dysplasia. *Development (Cambridge, England)*, 134(6), pp.1231–42.

Moumen, M. et al., 2011. The mammary myoepithelial cell. *The International journal of developmental biology*, 55(7-9), pp.763–71.

Moumen, M. et al., 2012. The Proto-Oncogene Myc is Essential for Mammary Stem Cell Function. *Stem cells (Dayton, Ohio)*.

Mulac-Jericevic, B. et al., 2003. Defective mammary gland morphogenesis in mice lacking the progesterone receptor B isoform. *Proceedings of the National Academy of Sciences of the United States of America*, 100(17), pp.9744–9.

Muller, W.J. et al., 1988. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. *Cell*, 54(1), pp.105–15.

Mulloy, J.C. et al., 2002. The AML1-ETO fusion protein promotes the expansion of human hematopoietic stem cells. *Blood*, 99(1), pp.15–23.

Mundlos, S. et al., 1997. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell*, 89(5), pp.773–9.

Mustonen, T. et al., 2003. Stimulation of ectodermal organ development by Ectodysplasin-A1. *Developmental biology*, 259(1), pp.123–36.

Nagaraja, G.M. et al., 2006. Gene expression signatures and biomarkers of noninvasive and invasive breast cancer cells: comprehensive profiles by representational difference analysis, microarrays and proteomics. *Oncogene*, 25(16), pp.2328–38.

Nagata, T. et al., 1999. Immunoglobulin motif DNA recognition and heterodimerization of the PEBP2/CBF Runt domain. *Nature structural biology*, 6(7), pp.615–9.

Nevadunsky, N.S. et al., 2009. RUNX3 protein is overexpressed in human epithelial ovarian cancer. *Gynecologic oncology*, 112(2), pp.325–30.

Nguyen, A. V & Pollard, J.W., 2000. Transforming growth factor beta3 induces cell death during the first stage of mammary gland involution. *Development (Cambridge, England)*, 127(14), pp.3107–18.

Niebuhr, B. et al., 2013. Runx1 is essential at two stages of early murine B-cell development. *Blood*, 122(3), pp.413–23.

Nielsen, T.O. et al., 2004. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 10(16), pp.5367–74.

Niini, T. et al., 2000. AML1 gene amplification: a novel finding in childhood acute lymphoblastic leukemia. *Haematologica*, 85(4), pp.362–6.

Nimmo, R., Antebi, A. & Woollard, A., 2005. mab-2 encodes RNT-1, a C. elegans Runx homologue essential for controlling cell proliferation in a stem cell-like developmental lineage. *Development (Cambridge, England)*, 132(22), pp.5043–54.

Nimmo, R. & Woollard, A., 2008. Worming out the biology of Runx. *Developmental biology*, 313(2), pp.492–500.

Niu, D.-F. et al., 2012. Transcription factor Runx2 is a regulator of epithelial-mesenchymal transition and invasion in thyroid carcinomas. *Laboratory investigation; a journal of technical methods and pathology*, 92(8), pp.1181–90.

North, T. et al., 1999. Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development (Cambridge, England)*, 126(11), pp.2563–75.

Novak, A. et al., 2000. Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. *Genesis (New York, N.Y. : 2000)*, 28(3-4), pp.147–55.

Oakes, S.R. et al., 2008. Prolactin regulation of mammary gland development. *Journal of mammary gland biology and neoplasia*, 13(1), pp.13–28.

Obr, A.E. & Edwards, D.P., 2012. The biology of progesterone receptor in the normal mammary gland and in breast cancer. *Molecular and cellular endocrinology*, 357(1-2), pp.4–17.

Oftedal, O.T., 2002. The mammary gland and its origin during synapsid evolution. *Journal of mammary gland biology and neoplasia*, 7(3), pp.225–52.

Ogawa, S., Satake, M. & Ikuta, K., 2008. Physical and functional interactions between STAT5 and Runx transcription factors. *Journal of biochemistry*, 143(5), pp.695–709.

Okuda, T. et al., 1996. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell*, 84(2), pp.321–30.

Onodera, Y. et al., 2010. Runx2 in human breast carcinoma: its potential roles in cancer progression. *Cancer science*, 101(12), pp.2670–5.

Ormandy, C.J. et al., 2003. Investigation of the transcriptional changes underlying functional defects in the mammary glands of prolactin receptor knockout mice. *Recent progress in hormone research*, 58, pp.297–323.

Ormandy, C.J. et al., 1997. Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. *Genes & development*, 11(2), pp.167–78.

Osato, M. et al., 1999. Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2alphaB gene associated with myeloblastic leukemias. *Blood*, 93(6), pp.1817–24.

Osorio, K.M. et al., 2008. Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation. *Development (Cambridge, England)*, 135(6), pp.1059–68.

Osorio, K.M., Lilja, K.C. & Tumbar, T., 2011. Runx1 modulates adult hair follicle stem cell emergence and maintenance from distinct embryonic skin compartments. *The Journal of cell biology*, 193(1), pp.235–50.

Otto, F. et al., 1997. Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell*, 89(5), pp.765–71.

Otto, F., Lübbert, M. & Stock, M., 2003. Upstream and downstream targets of RUNX proteins. *Journal of cellular biochemistry*, 89(1), pp.9–18.

Ozaki, T. et al., 2013. Runt-related transcription factor 2 (RUNX2) inhibits p53-dependent apoptosis through the collaboration with HDAC6 in response to DNA damage. *Cell death & disease*, 4, p.e610.

Pardali, E. et al., 2000. Smad and AML proteins synergistically confer transforming growth factor beta1 responsiveness to human germ-line IgA genes. *The Journal of biological chemistry*, 275(5), pp.3552–60.

Park, B.-Y. et al., 2012. Xaml1/Runx1 is required for the specification of Rohon-Beard sensory neurons in Xenopus. *Developmental biology*, 362(1), pp.65–75.

Pastrana, E., Silva-Vargas, V. & Doetsch, F., 2011. Eyes wide open: a critical review of sphere-formation as an assay for stem cells. *Cell stem cell*, 8(5), pp.486–98.

Pelletier, N. et al., 2002. MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. *Oncogene*, 21(17), pp.2729–40.

Peng, Z. et al., 2008. Inhibition of the growth and metastasis of human colon cancer by restoration of RUNX3 expression in cancer cells. *International journal of oncology*, 33(5), pp.979–84.

Perou, C.M. et al., 2000. Molecular portraits of human breast tumours. *Nature*, 406(6797), pp.747–52.

Perou, C.M., 2010. Molecular stratification of triple-negative breast cancers. *The oncologist*, 15 Suppl 5, pp.39–48.

Plaks, V. et al., 2013. Lgr5-Expressing Cells Are Sufficient and Necessary for Postnatal Mammary Gland Organogenesis. *Cell reports*.

Polyak, K., 2007. Breast cancer: origins and evolution. *The Journal of clinical investigation*, 117(11), pp.3155–63.

Ponti, D. et al., 2005. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer research*, 65(13), pp.5506–11.

Pozner, A. et al., 2000. Transcription-coupled translation control of AML1/RUNX1 is mediated by cap- and internal ribosome entry site-dependent mechanisms. *Molecular and cellular biology*, 20(7), pp.2297–307.

Prat, A. et al., 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast cancer research : BCR*, 12(5), p.R68.

Prat, A. & Perou, C.M., 2011. Deconstructing the molecular portraits of breast cancer. *Molecular oncology*, 5(1), pp.5–23.

Pratap, J. et al., 2009. Ectopic runx2 expression in mammary epithelial cells disrupts formation of normal acini structure: implications for breast cancer progression. *Cancer research*, 69(17), pp.6807–14.

Pratap, J. et al., 2006. Regulatory roles of Runx2 in metastatic tumor and cancer cell interactions with bone. *Cancer metastasis reviews*, 25(4), pp.589–600.

Pratap, J. et al., 2008. Runx2 transcriptional activation of Indian Hedgehog and a downstream bone metastatic pathway in breast cancer cells. *Cancer research*, 68(19), pp.7795–802.

Pratap, J. et al., 2005. The Runx2 osteogenic transcription factor regulates matrix metalloproteinase 9 in bone metastatic cancer cells and controls cell invasion. *Molecular and cellular biology*, 25(19), pp.8581–91.

Preudhomme, C. et al., 2009. High frequency of RUNX1 biallelic alteration in acute myeloid leukemia secondary to familial platelet disorder. *Blood*, 113(22), pp.5583–7.

Pullan, S. et al., 1996. Requirement of basement membrane for the suppression of programmed cell death in mammary epithelium. *Journal of cell science*, 109 (Pt 3, pp.631–42.

Quaglino, A. et al., 2009. Mechanical strain induces involution-associated events in mammary epithelial cells. *BMC cell biology*, 10, p.55.

Rakha, E.A. et al., 2007. Are triple-negative tumours and basal-like breast cancer synonymous? *Breast cancer research : BCR*, 9(6), p.404; author reply 405.

Ramaswamy, S. et al., 2003. A molecular signature of metastasis in primary solid tumors. *Nature genetics*, 33(1), pp.49–54.

Raveh, E. et al., 2006. Dynamic expression of Runx1 in skin affects hair structure. *Mechanisms of development*, 123(11), pp.842–50.

Reece, R.P., Turner, C.W. & Hill, R.T., 1936. Mammary Gland Development in the Hypophysectomized Albino Rat. *Experimental Biology and Medicine*, 34(2), pp.204–207.

Regan, J.L. et al., 2012. c-Kit is required for growth and survival of the cells of origin of Brca1-mutation-associated breast cancer. *Oncogene*, 31(7), pp.869–83.

Reinhardt, T.A. & Lippolis, J.D., 2009. Mammary gland involution is associated with rapid down regulation of major mammary Ca2+-ATPases. *Biochemical and biophysical research communications*, 378(1), pp.99–102.

Rennert, J. et al., 2003. The evolution of Runx genes I. A comparative study of sequences from phylogenetically diverse model organisms. *BMC evolutionary biology*, 3, p.4.

Reynolds, B.A. & Weiss, S., 1996. Clonal and population analyses demonstrate that an EGF-responsive mammalian embryonic CNS precursor is a stem cell. *Developmental biology*, 175(1), pp.1–13.

Richert, M.M. et al., 2000. An atlas of mouse mammary gland development. *Journal of mammary gland biology and neoplasia*, 5(2), pp.227–41.

Rink, J.C., 2013. Stem cell systems and regeneration in planaria. *Development genes and evolution*, 223(1-2), pp.67–84.

Robertson, A.J. et al., 2008. Runx expression is mitogenic and mutually linked to Wnt activity in blastula-stage sea urchin embryos. *PloS one*, 3(11), p.e3770.

Robinson, G.W., 2007. Cooperation of signalling pathways in embryonic mammary gland development. *Nature reviews. Genetics*, 8(12), pp.963–72.

Robinson, H.M. et al., 2003. Amplification of AML1 in acute lymphoblastic leukemia is associated with a poor outcome. *Leukemia*, 17(11), pp.2249–50.

Rody, A. et al., 2011. A clinically relevant gene signature in triple negative and basal-like breast cancer. *Breast cancer research : BCR*, 13(5), p.R97.

Rosen, J.M., 2012. On murine mammary epithelial stem cells: discovery, function, and current status. *Cold Spring Harbor perspectives in biology*, 4(4).

Roumier, C. et al., 2003. New mechanisms of AML1 gene alteration in hematological malignancies. *Leukemia*, 17(1), pp.9–16.

Rouzier, R. et al., 2005. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 11(16), pp.5678–85.

Rowley, M., Grothey, E. & Couch, F.J., 2004. The role of Tbx2 and Tbx3 in mammary development and tumorigenesis. *Journal of mammary gland biology and neoplasia*, 9(2), pp.109–18.

Ruan, W. & Kleinberg, D.L., 1999. Insulin-like growth factor I is essential for terminal end bud formation and ductal morphogenesis during mammary development. *Endocrinology*, 140(11), pp.5075–81.

Ruan, W., Newman, C.B. & Kleinberg, D.L., 1992. Intact and amino-terminally shortened forms of insulin-like growth factor I induce mammary gland differentiation and development. *Proceedings of the National Academy of Sciences of the United States of America*, 89(22), pp.10872–6.

Rubin, C.M. et al., 1987. Association of a chromosomal 3;21 translocation with the blast phase of chronic myelogenous leukemia. *Blood*, 70(5), pp.1338–42.

Sadikovic, B. et al., 2010. Expression analysis of genes associated with human osteosarcoma tumors shows correlation of RUNX2 overexpression with poor response to chemotherapy. *BMC cancer*, 10(1), p.202.

Dos Santos, C.O. et al., 2013. Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 110(18), pp.7123–30.

Satake, M. et al., 1995. Expression of the Runt domain-encoding PEBP2 alpha genes in T cells during thymic development. *Molecular and cellular biology*, 15(3), pp.1662–70.

Schedin, P., Mitrenga, T. & Kaeck, M., 2000. Estrous cycle regulation of mammary epithelial cell proliferation, differentiation, and death in the Sprague-Dawley rat: a model for investigating the role of estrous cycling in mammary carcinogenesis. *Journal of mammary gland biology and neoplasia*, 5(2), pp.211–25.

Scheitz, C.J.F. et al., 2012. Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer. *The EMBO journal*, 31(21), pp.4124–39.

Scheitz, C.J.F. & Tumbar, T., 2012. New insights into the role of Runx1 in epithelial stem cell biology and pathology. *Journal of cellular biochemistry*.

Schorr, K. et al., 1999. Gain of Bcl-2 is more potent than bax loss in regulating mammary epithelial cell survival in vivo. *Cancer research*, 59(11), pp.2541–5.

Schroeder, J.A. & Lee, D.C., 1998. Dynamic expression and activation of ERBB receptors in the developing mouse mammary gland. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*, 9(6), pp.451–64.

Schwertfeger, K.L., Richert, M.M. & Anderson, S.M., 2001. Mammary gland involution is delayed by activated Akt in transgenic mice. *Molecular endocrinology (Baltimore, Md.)*, 15(6), pp.867–81.

Sebastian, J. et al., 1998. Activation and function of the epidermal growth factor receptor and erbB-2 during mammary gland morphogenesis. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*, 9(9), pp.777–85.

Selbert, S. et al., 1998. Efficient BLG-Cre mediated gene deletion in the mammary gland. *Transgenic research*, 7(5), pp.387–96.

Selvamurugan, N. & Partridge, N.C., 2000. Constitutive expression and regulation of collagenase-3 in human breast cancer cells. *Molecular cell biology research communications : MCBRC*, 3(4), pp.218–23.

Seo, W. et al., 2012. Roles of VWRPY motif-mediated gene repression by Runx proteins during T-cell development. *Immunology and cell biology*, 90(8), pp.827–30.

Shackleton, M. et al., 2006. Generation of a functional mammary gland from a single stem cell. *Nature*, 439(7072), pp.84–8.

Shackleton, M. et al., 2009. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell*, 138(5), pp.822–9.

Shah, S.P. et al., 2012. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*, 486(7403), pp.395–9.

Siegel, P.M. & Muller, W.J., 2010. Transcription factor regulatory networks in mammary epithelial development and tumorigenesis. *Oncogene*, 29(19), pp.2753–9.

Slamon, D.J. et al., 2001. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *The New England journal of medicine*, 344(11), pp.783–92.

Sleeman, K.E. et al., 2006. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast cancer research : BCR*, 8(1), p.R7.

Smalley, M. & Ashworth, A., 2003. Stem cells and breast cancer: A field in transit. *Nature reviews. Cancer*, 3(11), pp.832–44.

Smith, B.A., Welm, A.L. & Welm, B.E., 2012. On the shoulders of giants: a historical perspective of unique experimental methods in mammary gland research. *Seminars in cell & developmental biology*, 23(5), pp.583–90.

Smith, G. & Medina, D., 1988. A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. *J. Cell Sci.*, 90(1), pp.173–183.

Smith, G.H., 1996. Experimental mammary epithelial morphogenesis in an in vivo model: evidence for distinct cellular progenitors of the ductal and lobular phenotype. *Breast cancer research and treatment*, 39(1), pp.21–31.

Smith, G.H., 2005. Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands. *Development (Cambridge, England)*, 132(4), pp.681–7.

Song, W.J. et al., 1999. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nature genetics*, 23(2), pp.166–75.

Sorlie, T. et al., 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences of the United States of America*, 100(14), pp.8418–23.

Soung, D.Y. et al., 2012. Runx1 dose-dependently regulates endochondral ossification during skeletal development and fracture healing. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*, 27(7), pp.1585–97.

Spender, L.C. et al., 2005. Transcriptional cross-regulation of RUNX1 by RUNX3 in human B cells. *Oncogene*, 24(11), pp.1873–81.

Staaf, J. et al., 2010. Identification of subtypes in human epidermal growth factor receptor 2--positive breast cancer reveals a gene signature prognostic of outcome. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 28(11), pp.1813–20.

Staniszewska, A.D. et al., 2012. Stat3 is required to maintain the full differentiation potential of mammary stem cells and the proliferative potential of mammary luminal progenitors. *PloS one*, 7(12), p.e52608.

Stein, T. et al., 2004. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. *Breast cancer research : BCR*, 6(2), pp.R75–91.

Stein, T., Salomonis, N. & Gusterson, B.A., 2007. Mammary gland involution as a multi-step process. *Journal of mammary gland biology and neoplasia*, 12(1), pp.25–35.

Stephens, P.J. et al., 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature*, 486(7403), pp.400–4.

Sterner, D.E. & Berger, S.L., 2000. Acetylation of Histones and Transcription-Related Factors. *Microbiology and Molecular Biology Reviews*, 64(2), pp.435–459.

Sternlicht, M.D., 2006. Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. *Breast cancer research : BCR*, 8(1), p.201.

Stewart, M. et al., 1997. Proviral insertions induce the expression of bone-specific isoforms of PEBP2alphaA (CBFA1): evidence for a new myc collaborating oncogene. *Proceedings of the National Academy of Sciences of the United States of America*, 94(16), pp.8646–51.

Stingl, J. et al., 2006. Purification and unique properties of mammary epithelial stem cells. *Nature*, 439(7079), pp.993–7.

Stingl, J. & Caldas, C., 2007. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nature reviews. Cancer*, 7(10), pp.791–9.

Stock, M. & Otto, F., 2005. Control of RUNX2 isoform expression: the role of promoters and enhancers. *Journal of cellular biochemistry*, 95(3), pp.506–17.

Strom, D.K. et al., 2000. Expression of the AML-1 oncogene shortens the G(1) phase of the cell cycle. *The Journal of biological chemistry*, 275(5), pp.3438–45.

Su, A.I. et al., 2001. Molecular classification of human carcinomas by use of gene expression signatures. *Cancer research*, 61(20), pp.7388–93.

Sullivan, J.C. et al., 2008. The evolutionary origin of the Runx/CBFbeta transcription factors--studies of the most basal metazoans. *BMC evolutionary biology*, 8, p.228.

Sun, L., Vitolo, M. & Passaniti, A., 2001. Runt-related gene 2 in endothelial cells: inducible expression and specific regulation of cell migration and invasion. *Cancer research*, 61(13), pp.4994–5001.

Suva, L.J. et al., 2011. Bone metastasis: mechanisms and therapeutic opportunities. *Nature reviews. Endocrinology*, 7(4), pp.208–18.

Swiers, G., de Bruijn, M. & Speck, N.A., 2010. Hematopoietic stem cell emergence in the conceptus and the role of Runx1. *The International journal of developmental biology*, 54(6-7), pp.1151–63.

Taddei, I. et al., 2008. Beta1 integrin deletion from the basal compartment of the mammary epithelium affects stem cells. *Nature cell biology*, 10(6), pp.716–22.

Taneja, P. et al., 2010. Classical and Novel Prognostic Markers for Breast Cancer and their Clinical Significance. *Clinical Medicine Insights. Oncology*, 4, pp.15–34.

Taniuchi, I. et al., 2002. Differential requirements for Runx proteins in CD4 repression and epigenetic silencing during T lymphocyte development. *Cell*, 111(5), pp.621–33.

Taniuchi, I., Osato, M. & Ito, Y., 2012. Runx1: no longer just for leukemia. *The EMBO journal*, 31(21), pp.4098–9.

Terriente-Felix, A. et al., 2013. Notch cooperates with Lozenge/Runx to lock haemocytes into a differentiation programme. *Development (Cambridge, England)*, 140(4), pp.926–37.

Terry, A. et al., 2004. Conservation and expression of an alternative 3' exon of Runx2 encoding a novel proline-rich C-terminal domain. *Gene*, 336(1), pp.115–25.

Theriault, F.M. et al., 2005. Role for Runx1 in the proliferation and neuronal differentiation of selected progenitor cells in the mammalian nervous system. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(8), pp.2050–61.

Theriault, F.M., Roy, P. & Stifani, S., 2004. AML1/Runx1 is important for the development of hindbrain cholinergic branchiovisceral motor neurons and selected cranial sensory neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 101(28), pp.10343–8.

Thirunavukkarasu, K. et al., 1998. Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfbeta. *Molecular and cellular biology*, 18(7), pp.4197–208.

Tonks, A. et al., 2003. The AML1-ETO fusion gene promotes extensive self-renewal of human primary erythroid cells. *Blood*, 101(2), pp.624–32.

Tracey, W.D. et al., 1998. A Xenopus homologue of aml-1 reveals unexpected patterning mechanisms leading to the formation of embryonic blood. *Development (Cambridge, England)*, 125(8), pp.1371–80.

Tsai, Y.C. et al., 1996. Contiguous patches of normal human mammary epithelium derived from a single stem cell: implications for breast carcinogenesis. *Cancer research*, 56(2), pp.402–4.

Vagner, S., Galy, B. & Pyronnet, S., 2001. Irresistible IRES. Attracting the translation machinery to internal ribosome entry sites. *EMBO reports*, 2(10), pp.893–8.

Vaillant, F. et al., 1999. A full-length Cbfa1 gene product perturbs T-cell development and promotes lymphomagenesis in synergy with myc. *Oncogene*, 18(50), pp.7124–34.

Vaillant, F. et al., 2002. Enforced expression of Runx2 perturbs T cell development at a stage coincident with beta-selection. *Journal of immunology (Baltimore, Md. : 1950)*, 169(6), pp.2866–74.

van Amerongen, R., Bowman, A.N. & Nusse, R., 2012. Developmental Stage and Time Dictate the Fate of Wnt/ß-Catenin-Responsive Stem Cells in the Mammary Gland. *Cell Stem Cell*.

VanHouten, J. et al., 2010. PMCA2 regulates apoptosis during mammary gland involution and predicts outcome in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 107(25), pp.11405–10.

Veltmaat, J.M. et al., 2006. Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes. *Development (Cambridge, England)*, 133(12), pp.2325–35.

Veltmaat, J.M. et al., 2004. Identification of the mammary line in mouse by Wnt10b expression. *Developmental dynamics : an official publication of the American Association of Anatomists*, 229(2), pp.349–56.

Verma, I.M., 2012. A clinicopathologic analysis of 45 patients with metaplastic breast cancer. *Cancer Research*, pp.72 (1024): P6–07–33. Available at: http://cancerres.aacrjournals.org/cgi/content/meeting_abstract/72/24_MeetingAbstracts/P6-07-33 [Accessed July 17, 2013].

Villagra, A. et al., 2006. Chromatin remodeling and transcriptional activity of the bone-specific osteocalcin gene require CCAAT/enhancer-binding protein beta-dependent recruitment of SWI/SNF activity. *The Journal of biological chemistry*, 281(32), pp.22695–706.

Visvader, J.E. & Smith, G.H., 2010. Murine Mammary Epithelial Stem Cells: Discovery, Function, and Current Status. *Cold Spring Harbor perspectives in biology*.

Vooijs, M., Jonkers, J. & Berns, A., 2001. A highly efficient ligand-regulated Cre recombinase mouse line shows that LoxP recombination is position dependent. *EMBO reports*, 2(4), pp.292–7.

Wagner, K.U. et al., 2001. Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice. *Transgenic research*, 10(6), pp.545–53.

Wagner, K.-U. et al., 2002. An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. *Development (Cambridge, England)*, 129(6), pp.1377–86.

Wagner, K.-U. et al., 2004. Impaired alveologenesis and maintenance of secretory mammary epithelial cells in Jak2 conditional knockout mice. *Molecular and cellular biology*, 24(12), pp.5510–20.

Wakao, H., Gouilleux, F. & Groner, B., 1994. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *The EMBO journal*, 13(9), pp.2182–91.

Walton, K.D. et al., 2001. Conditional deletion of the bcl-x gene from mouse mammary epithelium results in accelerated apoptosis during involution but does not compromise cell function during lactation. *Mechanisms of development*, 109(2), pp.281–93.

Wang, C.Q. et al., 2010. Runx family genes, niche, and stem cell quiescence. *Blood cells, molecules & diseases*, 44(4), pp.275–86.

Wang, C.-Y. et al., 2013. PCAF acetylates Runx2 and promotes osteoblast differentiation. *Journal of bone and mineral metabolism*, 31(4), pp.381–9.

Wang, L. et al., 2009. Post-translational modifications of Runx1 regulate its activity in the cell. *Blood cells, molecules & diseases*, 43(1), pp.30–4.

Wang, L., Brugge, J.S. & Janes, K.A., 2011. Intersection of FOXO- and RUNX1-mediated gene expression programs in single breast epithelial cells during morphogenesis and tumor progression. *Proceedings of the National Academy of Sciences of the United States of America*.

Wang, X.-P. et al., 2005. Runx2 (Cbfa1) inhibits Shh signaling in the lower but not upper molars of mouse embryos and prevents the budding of putative successional teeth. *Journal of dental research*, 84(2), pp.138–43.

Wang, Y. et al., 2005. Runx1/AML1/Cbfa2 mediates onset of mesenchymal cell differentiation toward chondrogenesis. *Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research*, 20(9), pp.1624–36.

Wansbury, O. et al., 2011. Transcriptome analysis of embryonic mammary cells reveals insights into mammary lineage establishment. *Breast cancer research : BCR*, 13(4), p.R79.

Watson, C.J., 2006. Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. *Breast cancer research : BCR*, 8(2), p.203.

Watson, C.J. & Khaled, W.T., 2008. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development (Cambridge, England)*, 135(6), pp.995–1003.

Watson, C.J. & Kreuzaler, P.A., 2011. Remodeling mechanisms of the mammary gland during involution. *The International journal of developmental biology*, 55(7-9), pp.757–62.

Weigelt, B., Kreike, B. & Reis-Filho, J.S., 2009. Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis. *Breast cancer research and treatment*, 117(2), pp.273–80.

Wenemoser, D. et al., 2012. A molecular wound response program associated with regeneration initiation in planarians. *Genes & development*, 26(9), pp.988–1002.

Westendorf, J.J. et al., 2002. Runx2 (Cbfa1, AML-3) interacts with histone deacetylase 6 and represses the p21(CIP1/WAF1) promoter. *Molecular and cellular biology*, 22(22), pp.7982–92.

Wiesen, J.F. et al., 1999. Signaling through the stromal epidermal growth factor receptor is necessary for mammary ductal development. *Development (Cambridge, England)*, 126(2), pp.335–44.

Williams, A.F. & Barclay, A.N., 1988. The immunoglobulin superfamily--domains for cell surface recognition. *Annual review of immunology*, 6, pp.381–405.

Williams, C. et al., 2009. Gene expression in murine mammary epithelial stem cell-like cells shows similarities to human breast cancer gene expression. *Breast cancer research : BCR*, 11(3), p.R26.

Wiseman, B.S. & Werb, Z., 2002. Stromal effects on mammary gland development and breast cancer. *Science (New York, N.Y.)*, 296(5570), pp.1046–9.

Wong, W.F. et al., 2012. Runx1 Deficiency in CD4+ T Cells Causes Fatal Autoimmune Inflammatory Lung Disease Due to Spontaneous Hyperactivation of Cells. *The Journal of Immunology*, 188(11), pp.5408–5420.

Wood, T.L. et al., 2000. The insulin-like growth factors (IGFs) and IGF binding proteins in postnatal development of murine mammary glands. *Journal of mammary gland biology and neoplasia*, 5(1), pp.31–42.

Woolf, E. et al., 2003. Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. *Proceedings of the National Academy of Sciences of the United States of America*, 100(13), pp.7731–6.

Wotton, S. et al., 2002. Proviral insertion indicates a dominant oncogenic role for Runx1/AML-1 in T-cell lymphoma. *Cancer research*, 62(24), pp.7181–5.

Wotton, S.F. et al., 2004. RUNX1 transformation of primary embryonic fibroblasts is revealed in the absence of p53. *Oncogene*, 23(32), pp.5476–86.

Wysolmerski, J.J. et al., 1998. Rescue of the parathyroid hormone-related protein knockout mouse demonstrates that parathyroid hormone-related protein is essential for mammary gland development. *Development (Cambridge, England)*, 125(7), pp.1285–94.

Xiao, G. et al., 2000. MAPK pathways activate and phosphorylate the osteoblast-specific transcription factor, Cbfa1. *The Journal of biological chemistry*, 275(6), pp.4453–9.

Xiao, Z.-S., Simpson, L.G. & Quarles, L.D., 2003. IRES-dependent translational control of Cbfa1/Runx2 expression. *Journal of cellular biochemistry*, 88(3), pp.493–505.

Yagi, R. et al., 1999. A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *The EMBO journal*, 18(9), pp.2551–62.

Yamaguchi, Y. et al., 2004. AML1 is functionally regulated through p300-mediated acetylation on specific lysine residues. *The Journal of biological chemistry*, 279(15), pp.15630–8.

Yamaji, D. et al., 2009. Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. *Genes & development*, 23(20), pp.2382–7.

Yamashiro, T. et al., 2002. Expression of Runx1, -2 and -3 during tooth, palate and craniofacial bone development. *Mechanisms of development*, 119 Suppl , pp.S107–10.

Yeh, H.-Y. et al., 2009. Identifying significant genetic regulatory networks in the prostate cancer from microarray data based on transcription factor analysis and conditional independency. *BMC medical genomics*, 2(1), p.70.

Yergeau, D.A. et al., 1997. Embryonic lethality and impairment of haematopoiesis in mice heterozygous for an AML1-ETO fusion gene. *Nature genetics*, 15(3), pp.303–6.

Young, D.W. et al., 2007. Mitotic occupancy and lineage-specific transcriptional control of rRNA genes by Runx2. *Nature*, 445(7126), pp.442–6.

Young, L.J.T. et al., 1971. The influence of host and tissue age on life span and growth rate of serially transplanted mouse mammary gland. *Experimental Gerontology*, 6(1), pp.49–56.

Yu, F. et al., 2013. RUNX3 interacts with MYCN and facilitates protein degradation in neuroblastoma. *Oncogene*.

Yu, K. et al., 2008. A precisely regulated gene expression cassette potently modulates metastasis and survival in multiple solid cancers. *PLoS genetics*, 4(7), p.e1000129.

Zagami, C.J., Zusso, M. & Stifani, S., 2009. Runx transcription factors: lineage-specific regulators of neuronal precursor cell proliferation and post-mitotic neuron subtype development. *Journal of cellular biochemistry*, 107(6), pp.1063–72.

Zaidi, S.K. et al., 2001. A specific targeting signal directs Runx2/Cbfa1 to subnuclear domains and contributes to transactivation of the osteocalcin gene. *Journal of cell science*, 114(Pt 17), pp.3093–102.

Zeng, Y.A. & 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(6), pp.568–77.

Zeps, N. et al., 1996. Detection of a population of long-lived cells in mammary epithelium of the mouse. *Cell and tissue research*, 286(3), pp.525–36.

Zhao, B., Li, L. & Guan, K.-L., 2010. Hippo signaling at a glance. *Journal of cell science*, 123(Pt 23), pp.4001–6.

Zhao, M. et al., 2003. E3 ubiquitin ligase Smurf1 mediates core-binding factor alpha1/Runx2 degradation and plays a specific role in osteoblast differentiation. *The Journal of biological chemistry*, 278(30), pp.27939–44.

Zhou, G. et al., 2006. Dominance of SOX9 function over RUNX2 during skeletogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 103(50), pp.19004–9.

Zhou, J. et al., 2005. Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. *The EMBO journal*, 24(3), pp.635–44.