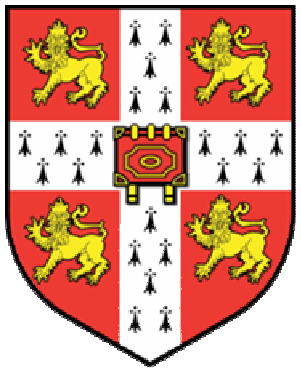
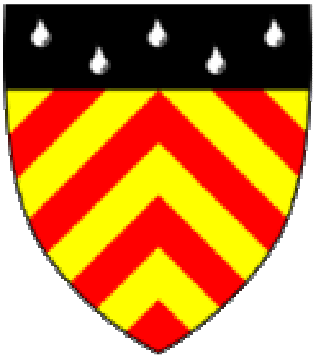
CHARACTERISATION OF BCL11 FUNCTIONS IN DEVELOPMENT USING GENETICALLY MODIFIED MICE

A Dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

By Song Choon Lee





University of Cambridge, Clare Hall College

The Wellcome Trust Sanger Institute

September 2008

DECLARATION

I hereby declare that this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specially indicated in the text. None of the material presented herein has been submitted previously for the purpose of obtaining another degree. I confirm that this thesis does not exceed 300 single sided pages of double spaced text, or 80,000 words.

Song Choon Lee

For always being there,

Dad, Mum, Angel and Shia

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest appreciation to my PhD supervisor, Dr Pentao Liu, for his patience and guidance in the projects that I have been involved in. It was great having such an enthusiastic mentor who is always full of creative ideas! I thank him for the challenging projects and the stimulating discussions we had in the lab. I would also like to thank Pentao's lab members, past and present, for their help and advice and most importantly for their encouragement when things were not going smoothly. Special mention goes to Dr Polly Chan and Dr Jacqui White, who were my first supervisors in the lab, and had taught me the basis of molecular biology and recombineering. I thank Dr Wei Wang and Dr Mariaestela Ortiz for the advice they have given me. Next, I would like to thank members of the RSF for their excellent technical help in taking care of my mice. In particular, I thank Dr Qin Si and Tina Hamilton for the blastocyst injection; Nick Harman, Michael Robinson, Paul Abbey and Peter Owers for day to day care of my mice. I also thank Dr Jeanne Estabel for teaching me the X-gal staining technique. I am grateful to Yvette Hooks and Kay Clarke for making the beautiful paraffin sections in this study. Thanks to Dr Huw Williams, Dr Sukit Chew, Dr Catherine Wilson and Dr Rebecca McIntyre for reading my thesis.

I am extremely fortunate to be able to collaborate with Dr Christine Watson, Dr John Stingl and Dr Walid Khaled from the Pathology Department, University of Cambridge, on the mammary studies. Their invaluable advice on my project and professional help with my experiments made everything went smoothly. I will miss the time spent in the 'boys' tissue culture room with John and Walid.

I would also like to show my appreciation to my thesis committee, Prof Allan Bradley, Dr Derek Stemple and Dr Bertie Gottgens for taking time off to attend the meetings and for their helpful suggestions and directions in my projects.

Last but not least, I would like to thank my parents, Kok Cheow Lee and Lar Keng Ng, and sister, Angel Lee for their unconditional love and support throughout these 4 years of my studies. I would also like to thank my significant other half, Li Shia Ng, who will be part of my family soon, for her words of encouragement and for always being there to support me.

ABSTRACT

Bcl11a and Bcl11b are two transcription factors that are essential for lymphocyte development and cell-fate decisions. To study the spatial expression patterns of these genes, I generated the *Bcl11-lacZ* tagged reporter mice. Using X-gal and FDG staining, the expression patterns of *Bcl11* genes in hematopoietic and mammary lineages were fully characterized. I found that *Bcl11a* and *Bcl11b* exhibited dynamic and contrasting expression patterns throughout mammary gland development. Both genes were among the earliest genes that were expressed specifically in the embryonic mammary placodes in the mouse. In the adult gland, *Bcl11a* was expressed in luminal progenitors and their differentiated derivatives while *Bcl11b* expression was predominantly restricted to basal cells and a small number of luminal progenitors.

Absence of *Bcl11a* and *Bcl11b* caused embryonic mammary placode defects. Deletion of *Bcl11a* in the virgin gland disrupted the mammary epithelial architecture and led to a decrease in Gata-3$^+$ cells and an increase in the number of ERα^+ cells. Loss of *Bcl11a* in the lactation gland led to the loss of secretory cells in the lobulo-alveoli, lactation failure, and premature onset of involution. This suggests that *Bcl11a* is essential for maintenance of terminally differentiated luminal secretory cell fate. In contrast, deletion of *Bcl11b* in the virgin gland led to precocious alveologenesis and a basal to luminal lineage switch in the basal cells. Transient over-expression of *Bcl11b* was sufficient to induce expression of basal cell specific genes. These results demonstrate that *Bcl11b* promotes and maintains basal identity, and also suppresses the luminal lineage.

At the molecular level, deletion of *Bcl11a* in the lactation glands resulted in dysregulation of JAK-Stat and Notch signalling pathways. In the *Bcl11a*-deficient lactation gland, there was an absence of phosphorylated Stat5-positive cells and a dramatic increase in phosphorylated Stat3-positive cells. Interestingly, loss of *Bcl11a* in lactation gland resulted in over-expression of *Notch1* but down-regulation of *Notch3* expression. These results demonstrate that different Notch receptors/ligands play different roles in maintaining mammary cell fate and that *Bcl11a* might potentially regulate JAK-Stat signalling via the Notch signalling pathway. This study thus identified *Bcl11a* and *Bcl11b* as critical regulators of the mammary epithelium.

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX

LIST OF TABLES

LIST OF ABBREVIATIONS

4-OHT	4-hydroxytamoxifen
AGM	Aorta-gonand-mesonephros
Amp	Ampicillin
BAC	Bacterial artificial chromosome
Bcl	B-cell lymphoma/leukaemia
BLG	β-lactoglobulin
BM	Bone marrow
Bmp	Bone morphogenic protein
BSA	Bovine serum albumin
Bsd	Blasticidin
C/EBP	CCAAT/enhancer-binding protein
cDNA	Complementary deoxyribose nucleic acid
ChIP	Chromatin immunoprecipitation
Cko	Conditional knockout
CLP	Common lymphoid progenitor
CMP	Common myeloid progenitor
CNS	Central nervous system
COUP-TF	Chicken ovalbumin upstream promoter transcription factor
Cre-ERT	Cre-estrogen receptor (Tamoxifen-inducible Cre)
CTIP	COUP-TF interacting protein
ddH_2O	Double-distilled H_2O
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DN	Double negative ($CD4^-CD8^-$)
dNTP	Deoxyribonucleotide triphosphate
Dox	Doxycycline
DP	Double positive ($CD4^+CD8^+$)
Dpc	Days post-coitum
DTT	Dithiothreitol

EDTA	Ethylene-diamine-tetra-acetic acid
ER	Estrogen receptor
ERK	Extracellular signal-regulated kinase
Evi	Ecotopic virus integration site
FACS	Fluorescent-activated cell sorting
FAM	6-carboxyfluorescein
FCS	Fetal calf serum
FDG	Fluorescein di-β-D-galactopyranoside
Fgf	Fibroblast growth factor
FIAU	1-2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-iodouracil
GAS	γ-interferon activation sites
GH	Growth hormone
Hes	Hairy enhancer of split
Hh	Hedgehog
HSC	Hematopoietic stem cell
HSVtk	Herpes Simplex Virus thymidine kinase
Hygro	Hygromycin
ICN	Notch intracellular domain
IGF	Insulin-like growth factor
IGFBP	Insulin-like growth factor binding protein
IL	Interleukin
IRES	Internal ribosome entry site
JAK	Janus kinase
Kan	Kanamycin
KLS	Lineage-negative c-kit$^+$ Sca1$^+$ cells
Lef	Lymphoid enhancing factor
LIF	Leukemia inhibitory factor
LMPP	Lymphoid-primed multi-potent progenitor
Ma-CFC	Mammary colony-forming-cell
MAML	Mastermind-like protein
MAP	Mitogen-activated protein

MaSC	Mammary stem cell
MEF	Mammary epithelial cell
MEP	Megakaryocyte/erythroid progenitor
MMTV	Mouse mammary tumour virus
MPP	Multi-potent progenitor
MSCV	Murine stem cell virus
MTA	Metastasis-associated protein
NBT/BCIP	Nitro blue tetrazolium chloride/5-Bromo-4-chloro-3-indolyl phosphate, toluidine salt
NCoR	Nuclear receptor co-repressor
Neo	Neomycin
NP-40	Nonidet P-40
Nrg	Neuregulin
NuRD	Nucleosome remodelling and deacetylase
PAC	P1 artificial chromosome
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline with 0.1% Tween-20
PCR	Polymerase chain reaction
PDK1	Phosphoinositide-dependent kinase 1
PFA	Paraformaldehyde
PH	Pleckstrin homology
PI3K	Phosphatidylinositol-3-OH-kinase
PIAS	Protein inhibitor of activated Stat
PKB	Protein kinase B
PR	Progesterone receptor
Prl	Prolactin
PtdIns-3,4,5-P_3	Phosphatidylinositol-3,4,5-triphosphate
PtdIns-4,5-P_2	Phosphatidylinositol-4,5-biphosphate
PTHrP	Parathyroid hormone related peptide
Puro	Puromycin

QPCR	Quantitative real-time PCR
qRT-PCR	Quantitative real-time Reverse Transcription PCR
QTL	Quantitative trait locus
RBP-Jκ	Recombination-binding protein-Jκ
Rit	Radiation induced tumour suppressor
RT-PCR	Semi-quantitative Reverse Transcription PCR
rtTA	Reverse tetracycline-controlled transactivator
SA	Splice acceptor
Sca1	Stem cell antigen 1
SDS	Sodium dodecyl sulphate
SH2	Src-homology 2
SHP	SH2 containing phosphatase
Ska	Scaramanga
SMA	Smooth muscle actin
SMRT	Silencing mediator for retinoid and thyroid hormone receptor
SOCS	Suppressor of cytokine signalling
Stat	Signal transducer and activator of transcription
TAM	Tamoxifen
TAMRA	Tetramethyl-6-carboxyrhodamine
Tbx	T-box
TE	Tris-EDTA
TEB	Terminal end bud
Tet	Tetracycline
TetR	Tetracycline repressor protein
TGF	Transforming growth factor
TNF	Tumour necrosis factor
TRE	Tetracycline response element
tTA	Tetracycline-controlled transactivator
WAP	Whey acidic protein
Wnt	Wingless and Int
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

TABLE OF CONTENT

CHAPTER 1:
INTRODUCTION

1.1 Mouse as a genetic tool

1.1.1 A brief history

The inbred laboratory mouse has become an indispensable tool in modern biological and medical research. Among mammals, the mouse is ideally suited for genetic analysis because of its small size (average adult weight of 25-40 g) which allows mice to be housed in high density, its short generation time (about 10 weeks from being born to giving birth) and prolificacy in breeding (5-10 pups per litter and an immediate postpartum estrus). In addition, the docile nature of the mouse makes handling easy and the deposition of a vaginal plug upon mating with a male allows pregnancies to be timed accurately. William Ernest Castle was the pioneer mouse geneticist who carried out the first systematic analysis of Mendelian inheritance and genetic variation in mice at the Bussey Institute, Harvard University (Snell and Reed, 1993). One important milestone in the history of mouse genetics was the establishment of inbred mouse strains. An inbred strain is defined as one that has been maintained for more than 20 generations by brother-to-sister mating and is essentially homozygous at all genetic loci, except for mutations arising spontaneously. After 20 generations of inbreeding, approximately 98.7% of the genomic loci in each animal will be homozygous (Silver, 1985). Subsequently, further inbreeding will result in decreasing heterozygosity at the rate of 19.1% with each generation, and by the 40^{th} generation of inbreeding, essentially 99.98% of the genome will be homozygous. To date, all the commonly used inbred strains have been inbred for at least 60 generations; hence all siblings are essentially 100% identical. The inbred mouse strains have revolutionized studies in cancer research, tissue transplantation, and immunology by eliminating genetic variability and making it possible to compare and analyze data from different laboratories worldwide.

1.1.2 Using the mouse to model human disease

Similarities in both the anatomy and physiology between the mouse and human make the mouse the most popular model organism used to study basic biological

processes, development, immunology and behaviour. Sequencing and analysis of the mouse and human genomes have led to the identification of ~25,000 genes as well as hundreds of conserved non-coding regions. It was found that about 99% of mouse genes have at least one orthologue in the human genome (Lander et al., 2001; Venter et al., 2001; Waterston et al., 2002). Mice and humans diverged from a common ancestor about 65 million years ago. The sequenced mouse genome is about 14% smaller than the human genome (2.5×10^9 bases compared to 2.9×10^9 bases). Even though evolutionary divergence between the two genomes has occurred over time, comparative genomics analysis has identified syntenic regions between the human and mouse genomes (Waterston et al., 2002). As such, more than 90% of the human and mouse genomes can be clustered into segments of conserved synteny, indicating the high degrees of conservation of genomic organisation (Waterston et al., 2002). Thus the mouse represents a good model to study human gene function *in vivo*.

1.2 Methodologies to genetically manipulate the mouse genome

1.2.1 Pronuclear injections and transgenic mice

The development of many genetic and genomic tools has played a critical role in the widespread use of the mouse for biomedical research. In 1966, Teh Ping Lin first demonstrated that following the direct injection of macromolecules into the pronuclei of mouse zygotes, embryos developed and were viable, thus establishing the technical basis for generation of transgenic mice (Lin, 1966). Following this seminal discovery, several groups showed that pronuclear injection of DNA into mouse zygotes resulted in expression of the introduced 'transgenes' in the mouse; a process called transgenic mouse production (Brinster et al., 1981; Wagner et al., 1981a; Wagner et al., 1981b). In general, transgenic DNA injected into mouse pronuclei will concatamerize and integrate into the genome. Consequently, the integration sites of the transgene are random and this can be a problem because the insertion sites may drastically influence the expression of the transgene. Additionally, copy number variation can also influence the expression of the transgene. In some cases, transgenes are used for expression studies, where it is essential to include all the regulatory elements of the gene to faithfully recapitulate expression of the endogenous locus. It is estimated that in 5-10% of transgenic lines, homozygosity for

particular transgenes may cause developmental anomalies, including lethality. These recessive phenotypes are likely to be the result of disruption of vital genes at the insertion sites of the transgenes. In addition, pronuclear injection of DNA can lead to genomic DNA rearrangements which could complicate the identification of the mutant loci. Despite these drawbacks, pronuclear injection of DNA was one of the most significant breakthroughs in mouse genetics as it demonstrated that experimental modification of the animal germ line was possible and has provided fundamental insights into many biological processes.

1.2.2 Development of mouse embryonic stem (ES) cells

The most commonly used platform in genetic engineering came with the isolation of embryonic stem cells and the development of gene targeting. One of the landmarks in mouse genetics was the isolation of pluripotent mouse embryonic stem cells (ES) from mouse blastocysts (Evans and Kaufman, 1981; Martin, 1981). The concept of pluripotent ES cells was established when teratocarcinomas were found to contain cells (embryonic carcinoma, EC cells) that gave rise to new tumours (Solter et al., 1970). Teratomas contain a mixture of differentiated cell types derived from all three germ layers and these are present in the tumour mass in a disorganised manner. Malignant teratomas (teratocarcinomas) were shown to contain undifferentiated stem cells that formed tumours upon transplantation into secondary recipients (Solter et al., 1970). These teratocarcinoma stem cells formed secondary teratocarcinomas upon the transfer of individual undifferentiated cells, demonstrating their clonal potency (Kleinsmith and Pierce, 1964). When EC cells were injected into a blastocyst, these cells became incorporated into the embryo and contributed to the tissues of the developing fetus (Brinster, 1974). This indicates that the proliferation of undifferentiated EC cells can be controlled in response to appropriate *in vivo* cues. Importantly, EC cells can be cultured and maintained *in vitro* on a feeder layer of fibroblasts (Martin and Evans, 1975). Based on these initial observations, pluripotent mouse ES cells were isolated successfully from the inner cell mass of 3.5 days post-coitum (dpc) wild-type embryos (Evans and Kaufman, 1981; Martin, 1981). These ES cells were initially characterized as teratocarcinoma stem cell lines due to their similarities in morphology, culture conditions

and pluripotency. However, it became apparent that these ES cells were distinct from EC cells because they were more stable and could be controlled. The most important difference between ES cells and EC cells is that ES cells are euploid and are able to repopulate somatic tissues and also transmit through the mouse germ-line during embryogenesis (Bradley et al., 1984). A major advance in ES cell technology is that these cultured ES cells maintain their pluripotency even after modification of their genomes by the use of retroviral vectors (Robertson et al., 1986). Therefore, the desired modifications can be introduced into the ES cells before they are injected into mouse blastocysts. Subsequently, the development of homologous recombination technology permitted targeted mutagenesis in ES cells, thus facilitating the potential disruption of any gene in the mouse (Capecchi, 1989; Koller and Smithies, 1989; Smithies et al., 1985; Thomas et al., 1986). These recent technological advances have allowed us to study the consequences of individual gene loss to deduce the gene's function in the mouse and likely, its role in humans. These approaches, together with the transgenic technique of zygote injection, are classified as reverse genetics. With the development of ES cell technology and the discovery of homologous recombination, reverse genetics became the main approach to study gene function in the mouse and in subsequent years, numerous gene functions were elucidated and mouse models of human diseases were created, advancing the understanding of gene function and human diseases. In recognition of their pioneering contributions to the field of mouse genetics, Mario R. Capecchi, Sir Martin J. Evans and Oliver Smithies were awarded the Nobel Prize in Physiology or Medicine in 2007 'for their discoveries of principles for introducing specific gene modifications in mice by the use of embryonic stem cells'.

1.2.3 Homologous recombination and gene targeting

Homologous recombination is the ability of the cellular machinery to combine DNA fragments that share stretches of similar or identical sequences. Early studies, which focused on extrachromosomal recombination between transfected DNA molecules, showed that mammalian cells possess effective machinery for mediating the homologous recombination of introduced DNA fragments. Although the mechanism of homologous recombination is not fully understood, recombination between introduced DNA

molecules and homologous sequences in the chromosome can be exploited to precisely alter genes of interest in mouse ES cells. Using homologous recombination, Smithies *et al*. introduced an exogenous DNA sequence into the human chromosomal *β-globin* locus (Smithies et al., 1985) and shortly afterwards, Thomas *et al*. corrected a mutant neomycin resistance gene (*neo*) in the host genome with an injected *neo* gene fragment (Thomas et al., 1986). These key experiments paved the way for the current gene targeting technology, thus enabling the analysis of the effects of loss- or gain-of-function of a specific gene.

A myriad of sophisticated gene targeting strategies in ES cells have since been developed to manipulate the mouse genome (Bradley and Liu, 1996). In addition to the conventional use of ES cells to generate knockout mice, it can also be used for transgene expression analysis. Using ES cells as a genetic vehicle facilitated experiments that were not possible by pronuclear injections; for example where a dominant lethal phenotype is caused by transgene expression (Warren et al., 1994). Furthermore using gene knock-in strategies, Hanks *et al*. showed that the heterologous genes *engrailed 1/2* were capable of functional compensation by targeting one gene to the locus of its homologue and observing rescue of the knockout phenotype (Hanks et al., 1995). In the last two decades, gene targeting has been widely utilized in ES cells to make a variety of genetic mutations in many loci, allowing the phenotypic analyses of the genetic modifications to be determined. The combination of gene targeting and ES cell technology provides two vital components of the current mouse gene knockout technology.

1.2.4 Conventional gene knockout technology

The procedure for generating mice that have been genetically modified using gene targeting strategies is essentially the same regardless of the specific targeting strategy used (Figure 1.1). Firstly, the targeting construct which contains DNA fragments homologous to the targeted gene and a selectable marker is generated. The construct is then linearized and electroporated into wild-type ES cells which are cultured in the presence of a selection agent to select for transfectants which have stably integrated the construct into their genome. The surviving ES cell colonies are then isolated and examined for presence of the targeted allele using long range PCR or Southern blot

analysis to confirm that the desired recombined event has occurred. ES cell clones containing the correctly targeted allele are then expanded and injected into 3.5 dpc blastocysts and transplanted into the uteri of pseudopregnant surrogate females. The resulting pups are then examined for their degree of chimerism (percentage of genetic makeup of the mouse that was contributed by the ES cell). Male chimera showing high levels of chimerism are then mated with wild-type females to check for germ-line transmission of the targeted allele in the F1 offspring. The F1 heterozygotes can then be inter-crossed to breed to homozygosity.

Introducing a loss-of-function mutation is the most common method for determining the function of a gene. In order to ablate the function of a gene of interest (to generate a null allele), the most common experimental strategy is to replace all or part of the coding sequence of the gene with a selectable marker. There are two main types of targeting vectors for mutating a gene through homologous recombination in ES cells: (1) replacement vectors and (2) insertion vectors (Thomas and Capecchi, 1986). A replacement vector consists of a positive selection marker such as Neomycin (*Neo*), Puromycin (*Puro*) or Blasticidin (*Bsd*) resistance markers, which is flanked by isogenic DNA homologous arms (5-8 kb), and a linearization site outside of the homologous arms of the vector (Figure 1.2A). Usually, a negative selectable marker Herpes Simplex Virus thymidine kinase gene (*HSVtk*) is also used to enrich for correctly targeted events from random integrations. The *HSVtk* is inserted next to the homology arms. Random insertion events are likely to retain the functional *HSVtk* gene which would encode for thymidine kinase (Thomas and Capecchi, 1986). Negative selection is then performed by adding chemical compounds, such as gancyclovir or FIAU [1-2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-iodouracil]. These drugs are transformed into cytotoxic substances by the HSVtk. Thus, the cells that have the random integration events and thus harbour the *HSVtk* gene are eliminated. Those cells in which homologous recombination have occurred do not harbour the *HSVtk* gene and thus survive the selection. The final post-recombination product is the replacement of the targeted region with all the components of the vector. The basic elements of insertion vectors are a genomic DNA fragment and a selection marker. Unlike the replacement vector where the linearization site is outside the homology arms, the site of linearization of the insertion vector is within the homologous

DNA sequence. Homologous recombination between the insertion vector and its target locus results in insertion of the entire vector into the locus, leading to a duplication of the genomic sequences (Figure 1.2B).

Several parameters are known to affect the efficiency of the targeting events in ES cells. Firstly, the length of homologous sequence on the targeting vector; generally, the longer the homologous sequence, the higher the efficiency and typically, 7-8 kb in total (or 3-4 kb on each side of the replacement vector) is sufficient for successful targeting (Deng and Capecchi, 1992). Secondly, the use of isogenic DNA from the strain of mouse from which the ES cells are derived (such as 129/Sv for AB2.2 ES cells) has been shown to enhance the recombination efficiencies during targeting (van Deursen and Wieringa, 1992). Finally, the employment of promoter or poly-adenylation (polyA) trapping has also been shown to enhance targeting efficiencies (Donehower et al., 1992). In promoter trapping, the positive selection marker is promoter-less which allows for the selection of insertions that occur next to an endogenous promoter. However, this limits the targeting events to genes that are expressed in ES cells. To overcome this problem, polyA trapping which uses a positive selection marker that does not have a polyA signal and allows for the selection of integrations into genes, regardless of whether or not they are expressed in ES cells, was developed. However in our lab, neither promoter nor polyA trapping strategies are required for successful gene targeting in mouse ES cells.

The genetic changes created by conventional gene targeting vectors as discussed above are usually null mutations. This is usually accomplished by the deletion of exons and/or the introduction of a selection marker into the coding sequences which disrupts the reading frame of the gene, resulting in premature termination of the transcript or frame-shifts. Analyses of these null mutant mice have provided fundamental insights into mammalian gene function. Modification of the mouse genome can also produce other types of mutations; for example, hypomorphic alleles of a gene are also desirable to complement the null allele. In addition, subtle mutations such as point mutations in a gene are necessary in order to study functional domains of proteins and to accurately mirror some human diseases. Importantly, for genes that are critical for embryonic development, null mutants usually die *in uteri*, thereby precluding the study of their functions in late development or in specific tissues in the adult mice. To circumvent this

problem and to enable the investigation of gene function in a temporal and spatial manner, conditional knockout (cko) approaches have been developed (Glaser et al., 2005; Rajewsky et al., 1996). The cko approach allows a more accurate mouse model of human disease and sporadic cancer initiation and progression to be created.

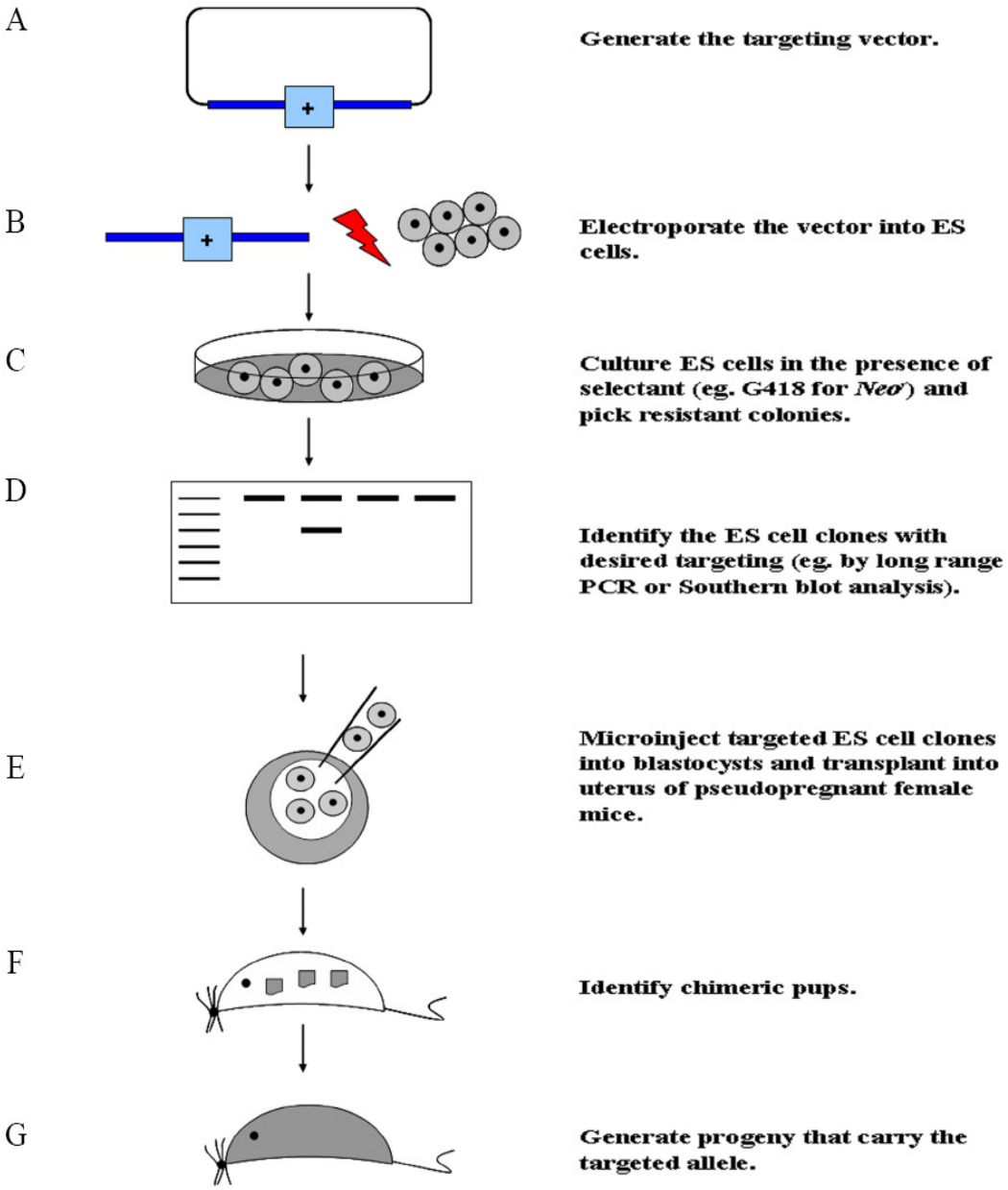


Figure 1.1. General procedure for generation of genetically modified mice by gene targeting strategies. (A) Generation of a targeting vector containing a positive (+) selection cassette and sequences of homology with the target locus (blue line). **(B)** The vector is linearized and electroporated into ES cells. **(C)** Correct transformants are selected for in the presence of a selectant (eg. G418 if a neomycin resistance cassette is present in the targeting vector). **(D)** Correctly targeted ES cell clones are then identified and genetically characterized using long range PCR or Southern blot analysis. **(E)** The selected ES cell clones are then microinjected into 3.5 days post-coitum blastocysts and transplanted into the uteri of pseudopregnant females. **(F)** Chimera obtained from the microinjections are mated with wild-type mice to establish germ-line transmission of the modified allele. **(G)** Progeny derived from the chimeras are characterized using long range PCR or Southern blot analysis, and a mutant mouse line that carries the desired targeted allele is established.

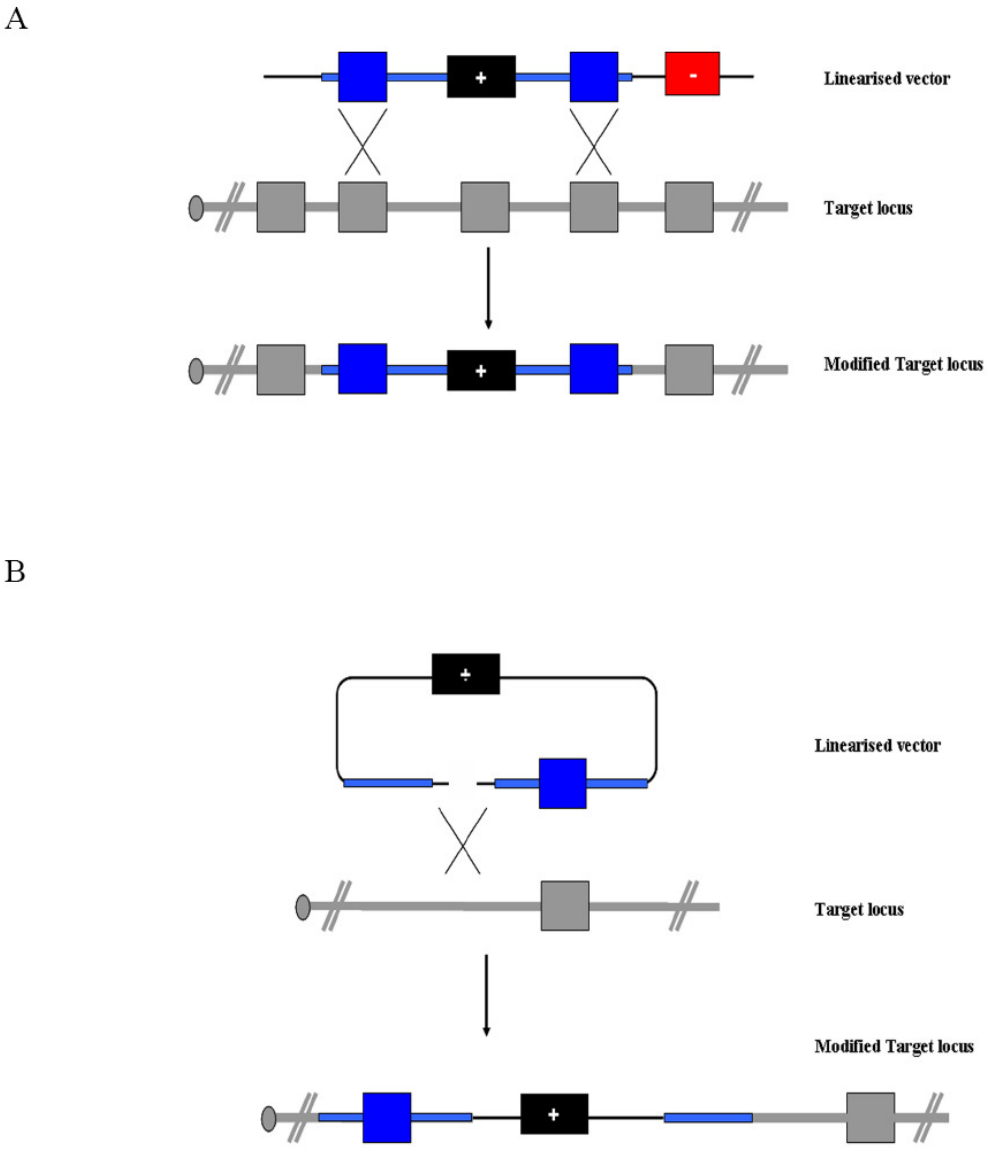


Figure 1.2. Replacement and insertion type vectors. (A) Replacement vectors target the locus by double reciprocal recombination, and they insert only the genetic sequences contained within the homologous sequences (blue). The positive selection markers (+) can be used to replace an exon or disrupt an exon. A negative selection marker (-) can be used to select against the random insertion of vectors. **(B)** Insertion vectors target the locus by single reciprocal recombination and insert the entire vector sequence. The vector contains regions of DNA (blue lines) homologous to the target locus (grey line) and a positive selection marker (+). X represents recombination between vector and the genomic loci.

1.2.5 Conditional knockout and Cre-*lox*P recombination system

Site-specific recombinases catalyze the recombination between two recombinant sites. By placing these sites strategically within the gene loci, deletion, insertion, inversion or translocation of specific regions of DNA can be produced after expression of the specific recombinase. There are two commonly used site-specific recombinase system in the mice: (1) the Cre/*lox*P system from the bacteriophage P1 (Figure 1.3A) and (2) the Flp/*FRT* system from the budding yeast *S. cerevisiae* (Figure 1.3B). Cre recombinase appears more effective in recombining its substrates *in vitro* and *in vivo* than Flp and it is the more widely used site-specific DNA system in both ES cells and mice (Nagy, 2000). This prokaryotic protein was first shown to work in mammalian cells by Sauer *et al.* (Sauer and Henderson, 1988) and subsequently in the mouse by Lakso *et al* (Lakso et al., 1992). Cre recombinase cleaves DNA between two 34-bp-long specific consensus sites known as *lox*P sites (Figure 1.3A). The *lox*P site has a unique structure that contains two palindromic 13-bp repeats flanking an 8-bp core sequence. The asymmetry of the core confers the orientation of the *lox*P site which is important for gene targeting as the directionality of the *lox*P sites determines the final configuration of the DNA sequence after expression of Cre recombinase. For example, if the two *lox*P sites are in the same orientation, expression of Cre recombinase would result in deletion of the intervening sequences (Figure 1.3C). In contrast, if the two *lox*P sites are in opposite orientations, Cre recombinase expression would result in an inversion (Figure 1.3D). Additionally, if the *lox*P sites are in *trans* (on different chromosomes), expression of Cre recombinase will lead to reciprocal exchange of the two chromosomes resulting in a translocation (Ramirez-Solis et al., 1995) (Figure 1.3E).

The Cre/*lox*P system has several advantages that make it suitable as a DNA recombination system in the mouse. Firstly, the *lox*P site is short enough to be targeted to intronic regions without disrupting endogenous transcription of the gene, but long enough to avoid the random occurrence of intrinsic *lox*P sites in the mouse genome. Analysis of the genomes of the mouse, rat and zebrafish revealed no perfectly matched intrinsic *lox*P sites. However, it has been noted that cryptic *lox*P sites are present in the mouse genome. Recombination between the cryptic *lox*P sites and *lox*P sites is expected to occur at a much lower efficiency, but the efficiency has not been comprehensively investigated. Cre

recombinase has also been shown to be functional in a wide range of cell types and can act on various forms of DNA substrates (supercoiled, relaxed or linear). *In vitro* Cre-mediated excision has been shown to facilitate genomic deletion of up to 400 kb, and recombinants can be identified without selection (Nagy, 2000). *In vivo* Cre-mediated excision is also extremely efficient and many Cre transgenic mouse lines have been generated to facilitate Cre-mediated excision in a wide range of tissues and over numerous developmental time points (Nagy, 2000).

One of the main uses of the Cre/*lox*P system in the mouse is to generate cko alleles. The first step to generating a cko allele is to identify the critical exon(s) of the gene which encodes for the essential functional domains of the protein. The critical exon(s) is then flanked by *lox*P sites which are targeted to the intronic regions flanking the critical exon(s). Mice containing the cko alleles should be phenotypically normal unless they are bred to a Cre-expressing transgenic line. Depending on the nature of the promoter driving Cre recombinase expression (ubiquitous or tissue-specific), deletion of the floxed region can occur in all cells or in specific cells/tissues or at certain developmental stages (Nagy, 2000).

To design a successful cko allele, there are several considerations that need to be addressed. Firstly, the critical exon(s) which encodes the essential functional domains has to be selected and flanked by *lox*P sites such that Cre-mediated excision would result in a null allele. The ideal scenario would be to flank the entire gene with *lox*P sites but this is usually not technically feasible as the efficiency of Cre-mediated excision decreases dramatically with increasing distance (Nagy, 2000). Another method would be to target the first exon of the gene. However there are pitfalls regarding this strategy such as placing a *lox*P site upstream of the ATG start site might disrupt endogeneous promoter sequences and there might also be alternative start sites downstream of the first exon to generate alternative transcripts of the gene. In addition, one also has to consider the strategy of targeting *lox*P sites to the gene locus and keeping the modified allele functional. The selection cassette used in the targeting step is usually removed because the presence of the selection cassette can result in a hypomorphic allele or affect the expression of neighbouring genes (Fiering et al., 1995). Such interference can generate a

phenotype that is normally not associated with loss-of-function of the targeted allele and can cause misinterpretation of the results.

The more sophisticated techniques of generating cko alleles use both Cre/*lox*P and Flp/*FRT* systems. In this method, the selection marker is flanked by two *FRT* sites and expression of Flp can be carried out either *in vitro* (ES cells using a Flp-expression plasmid) or *in vivo* (Flp-expressing mice) to remove the selection cassette (Figure 1.4). Following expression of Flp recombinase, the selection cassette is removed, leaving behind two *lox*P sites flanking the critical exon, generating a cko allele. The disadvantage of this system is that Flp recombinase is not as efficient as Cre recombinase; however with the generation of the enhanced Flp recombinase (Flpe) and mouse codon-optimised Flp recombinase (Flpo), the recombination efficiency of Flp has improved and is closer to that of Cre recombinase (Buchholz et al., 1998; Raymond and Soriano, 2007).

The cko alleles would not be possible unless there were well-characterized Cre transgenic lines which permitted the Cre-mediated excision of cko alleles. Thus, a major challenge is to identify promoters that are characteristic for different cell types, lineages and developmental events (Nagy, 2000). The Cre transgenic lines generated from these promoters will have to be validated to ensure their efficacy as well as specificity. This is because Cre-mediated excision does not usually reach 100% efficiency and partial excision creates mosaicism that can complicate interpretation of the observed phenotype. Nevertheless, these tissue- and cell-lineage-specific Cre transgenic lines are invaluable resources for the cko technology. A further improvement to Cre transgenic technology was achieved with the creation of inducible Cre systems. The inducible Cre system allows for the regulated control of the induction of Cre expression, thus permitting temporal deletion of the conditional alleles. One such system is the tamoxifen (TAM) or 4-hydroxytamoxifen (4-OHT) inducible Cre-ERT (Metzger and Chambon, 2001). The Cre recombinase has been fused to a mutated ligand binding domain of the human estrogen receptor (ER), resulting in a tamoxifen-dependent Cre recombinase, Cre-ERT, that is activated by TAM or 4-OHT, but not by endogenous estrogen or progesterone. In the absence of TAM or 4-OHT, the Cre-ERT protein resides in the cytoplasm, whereas upon treatment, binding of the ligand to the Cre-ERT protein results in a transformational change, causing the Cre-ERT protein to translocate into the nucleus where it executes its

function (Metzger and Chambon, 2001). Another widely used inducible Cre technology is the tetracycline (Tet)-dependent regulatory system (Tet-on or Tet-off) (Gossen and Bujard, 1992). The Tet systems use a chimeric transactivator to control transcription of the gene of interest from a silent promoter and are based on two regulatory elements derived from the *E. coli* Tet resistance operon: the Tet repressor protein (TetR) and the Tet operator sequence (*tetO*). The Tet-off system uses a plasmid that expresses a fusion protein known as the tetracycline-controlled transactivator (tTA), which is composed of TetR and the VP16 activation domain of the herpes simplex virus. The tTA binds to the tetracycline response element (TRE) and activates transcription of the target gene in the absence of the inducer, doxycyclin (Dox) (Gossen and Bujard, 1992). In contrast, the Tet-on system uses a form of TetR containing four amino acid changes that results in altered binding characteristics and creates the reverse TetR (rTetR). The rTetR binds the TRE and activates transcription of the target gene in the presence of Dox (Gossen and Bujard, 1992). The temporal control of Cre activity makes it an attractive system to combine inducibility with cell- or lineage-specific expression of Cre recombinase, thus allowing the temporal and spatial control of the deletion of the cko allele.

While expression of Cre recombinase is not normally associated with lethality, it is important to note that there are reported Cre toxicity effects. Cre expression in cultured mammalian cells results in a markedly reduced rate of proliferation and this effect is dependent on the endonuclease activity of Cre (Loonstra et al., 2001). *In vivo* titration experiments revealed that toxicity is due to the level of Cre activity (Loonstra et al., 2001). Hence proper controls are required to ensure that the observed phenotypes in cko mice are due to deletion of the targeted allele and not due to side effects of Cre toxicity.

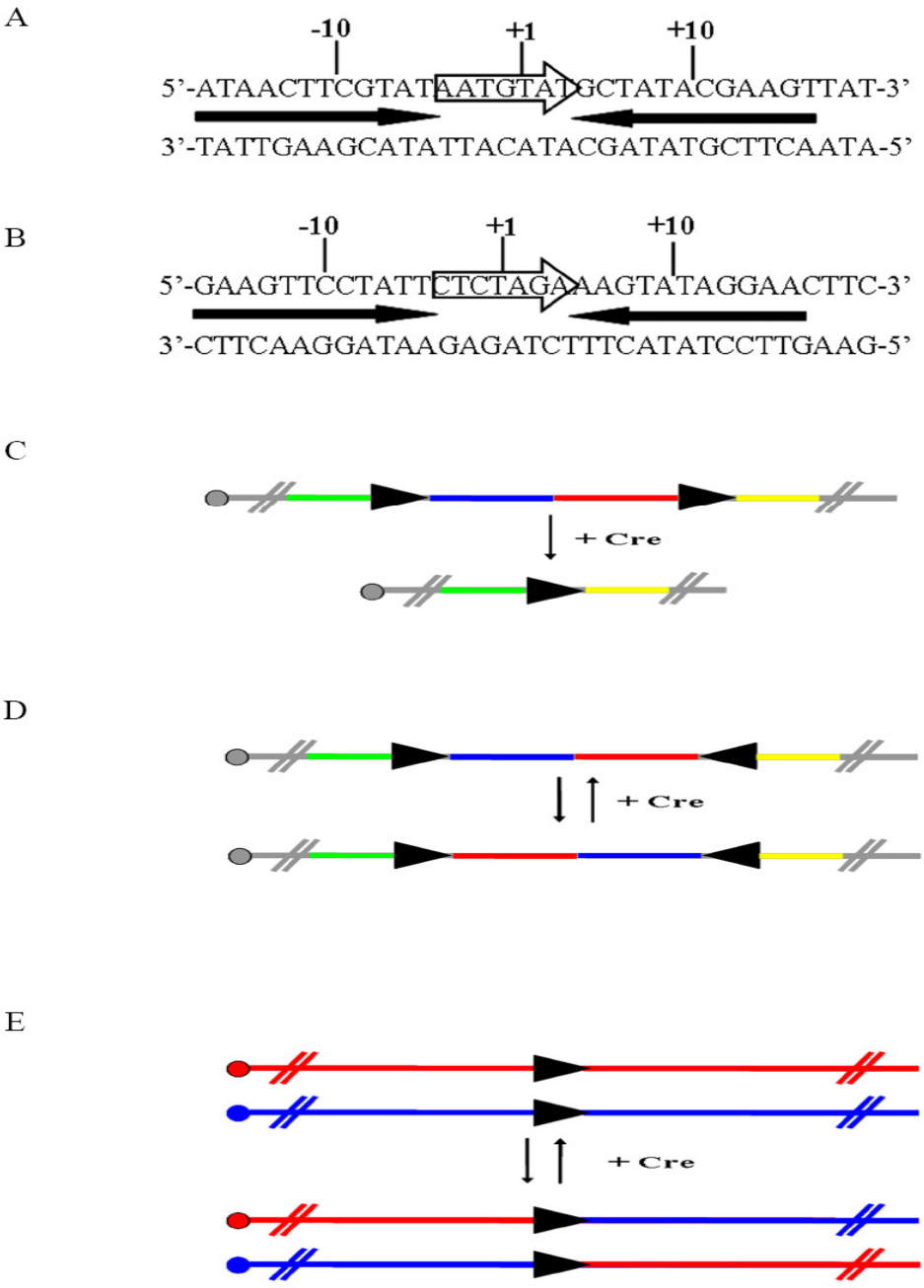


Figure 1.3. Properties of Cre recombinase. The consensus DNA sequence is recognized by **(A)** Cre recombinase and **(B)** Flp recombinase. Target sites (*lox*P of Cre; *FRT* of Flp) contain inverted 13-bp symmetry elements (indicated by the bold arrows) flanking an 8-bp A:T rich non-palindromic core (indicated by the open arrow). One recombinase monomer binds to each symmetry element, while the core sequence provides the site of strand cleavage, exchange, and ligation. The asymmetry of the core region (open arrow) imparts directionality on the reaction. Depending on the directionality of the *lox*P sites, expression of Cre recombinase can result in **(C)** deletion of the intervening sequences if *lox*P sites are in *cis*, **(D)** an inversion if the *lox*P sites are oriented in opposite directions or **(E)** reciprocal exchange of the regions that flank the *lox*P sites if the two sites are in *trans*. The arrows between the recombinase substrates and the products indicate the reversible nature of each reaction.

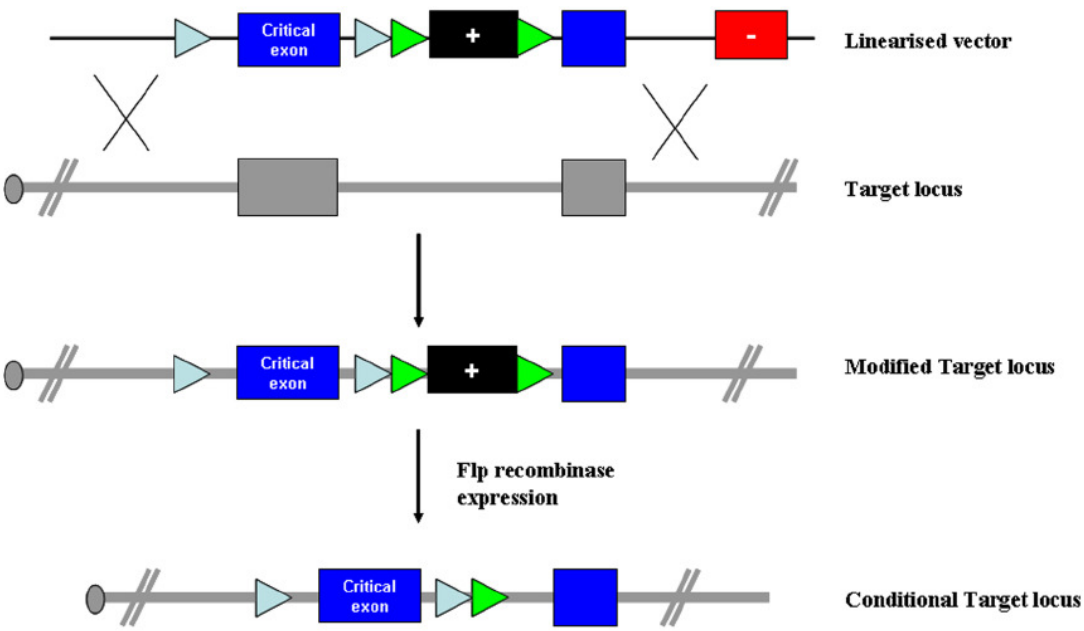


Figure 1.4. Strategy for generation of a conditional knockout allele. This strategy employs the use of both Cre/*lox*P and Flp/*FRT* systems. The selection marker (+) is flanked by two *FRT* sites (green arrow heads) and expression of Flp recombinase (*in vitro* or *in vivo*) results in the excision of the selection marker, leaving behind a single *FRT* site and two *lox*P sites (blue arrow heads) flanking the critical exon(s). (-) indicates negative selection marker. X represents recombination between vector and the genomic loci.

1.2.6 Recombineering technology

The development of recombinogenic engineering or recombineering has greatly reduced the time required to generate DNA constructs for making knockout mice. Recombineering is based on the *E. coli* phage homologous recombination systems which enable modification of genomic DNA such as those carried on bacterial artificial chromosomes (BACs) or P1 artificial chromosomes (PACs) without any restriction enzymes or ligases. Traditionally, *E. coli* has been the organism of choice for carrying out the genetic modifications required for making targeting vectors. Numerous restriction digestion and ligation steps are required to piece different DNA fragments together in order to complete the targeting vectors. This is usually tedious and requires the use of specific compatible restriction sites. Using homologous recombination for DNA modification was first widely used in *S. cerevisiae* to combine transformed, linear, double-stranded DNA (dsDNA) with homologous regions in the yeast genome (Lafontaine and Tollervey, 1996). However, unlike in the yeast, linear dsDNA are highly unstable in *E. coli* due to the presence of the ATP-dependent, linear-dsDNA exonuclease RecBCD (Copeland et al., 2001). RecBCD unwinds and degrades DNA to generate 3' single-stranded DNA (ssDNA) tails, which are used by RecA to initiate recombination. Therefore transformation of linear dsDNA into wild-type *E. coli* would result in degradation of the exogenous DNA, making homologous recombination unfeasible. To overcome this, *E. coli* strains that contain the *RecBC* (hence lacking RecBCD), *sbcB* and *sbcC* mutations (which restore recombination capabilities) were initially used for genetic manipulation (Copeland et al., 2001). However, loss of the RecBCD complex generally impairs cellular integrity and as the recombination pathway is constitutively active in these strains, rearrangements and deletions between the repeat sequences found within BACs and PACs usually occur. To overcome this deficiency, a temperature-sensitive shuttle-vector-based system based on the RecA pathway was developed (O'Connor et al., 1989). The *E. coli RecA* gene was cloned into a temperature-sensitive shuttle vector that has the temperature-sensitive origin of replication which functions at 30^{o}C but not at the higher restrictive temperature of 42-44^{o}C. Hence, transformation of this plasmid, which also contains cloned sequences homologous to the target genomic DNA, into the *RecA*$^{-}$ *E. coli* containing the BAC confers recombination-competence to the *E. coli*, allowing

homologous recombination to occur (O'Connor et al., 1989). As loss of *RecA* in *E coli* generally impairs general cellular integrity, an improved recombineering approach by transient expression of *RecA* has been developed to generate targeting constructs (Wang et al., 2006).

In 1998, an important advancement to the field of homologous recombination was made by Francis Steward and colleagues who developed the Rac-encoded RecET system (also known as Red recombination and lambda-mediated cloning). They showed that PCR-amplified fragment of linear dsDNA, flanked by short homology arms (42 bp homologous sequence to target plasmid or BAC) can be efficiently targeted to a plasmid or BAC by electroporating the dsDNA into *recBC sbcA E. coli* strains (Zhang et al., 1998). Instead of using *RecBC* mutants as discussed above, RecBC can be inactivated either by the *sbcA* mutation, which removes a repressor for the endogenous Rac prophage to induce expression of *recE* and *recT* (Lloyd, 1974) or the *gam* protein of λ bacteriophage (Murphy, 1998). Recombination functions encoded by *recE* and *recT* genes enabled genomic DNA to be modified using PCR products with short homology arms (Zhang et al., 1998). The overview of the RecET system is depicted in Figure 1.5. Firstly, PCR amplification is used to generate the homology arms (to the target plasmid or BAC DNA) flanking the insert DNA. Secondly, phage-derived recombineering competence was conferred to the BAC-containing *E. coli* host strain. Finally, the PCR-generated cassette is introduced into the host by electroporation to initiate homologous recombination and generate the recombinant cassette (Figure 1.5). Bacteriophage λ contains an efficient recombination system known as the Red system. The key genes in this system are the *redα/exo* and *redβ/bet*. Exo is a 5'-3' exonuclease that acts on linear dsDNA while Bet binds to the 3' ssDNA (single-stranded DNA) overhangs generated by Exo and promotes annealing to a complementary strand. The Red system is assisted by the bacteriophage-encoded Gam protein that inhibits host RecBCD activity (Murphy, 1998).

After the initial publication by Steward, several labs developed other methods based on expression of the Exo, Bet and Gam proteins to maximise efficient homologous recombination efficiency (Datsenko and Wanner, 2000; Yu et al., 2000). In the defective prophage-based system, *exo*, *bet* and *gam* are expressed from a defective prophage that

has integrated into the *E. coli* genome and their expression is regulated under the tight control of the temperature-sensitive λ-CI857 repressor (this strain is known as DY380) (Yu et al., 2000). At 32^{o}C, the repressor is active and there is no expression of the three genes, thus the *E. coli* cells do not have recombineering activity. In contrast, upon shifting the culture to 42^{o}C for 15 min, the CI repressor is inactivated and the *red* genes are co-ordinately expressed from a λ-*pL* promoter at high levels, conferring recombineering-competence to the cells. The λ prophage-based recombineering system proved to be extremely efficient because all three recombineering genes (*exo*, *bet* and *gam*) are present as a single copy and each is expressed from their natural operon and from a strong promoter *pL*, which means that they are present in molar ratio suitable for forming a complex *in vivo*. This tightly regulated expression ensures that the Exo and Bet proteins are only present after induction, thus minimizing the deleterious effects of rearrangement or deletion of the BAC. To facilitate *in vitro* expression of Cre and Flpe recombinase, arabinose-inducible Cre and Flpe recombinases have also been introduced into DY380 strain to generate EL350 and EL250 strains respectively (Lee et al., 2001). Transient expression of Cre or Flpe recombinases after induction with arabinose would mediate recombination between *lox*P and *FRT* sites respectively, allowing the removal of undesired selection markers from targeting cassettes. Using these bacterial strains, a highly efficient strategy for generation of conditional targeting vectors was developed (Figure 1.6) (Liu et al., 2003a). In this method, longer homology arms (200-500 bp for retrieving and 100-300 bp for targeting) were generated by PCR amplification using BAC DNA as a template. These homology arms were used to subclone BAC DNA into high-copy plasmid backbone (pBluescript) by gap repair using recombineering in the bacterial strain EL350. Similarly, long homology arms were also generated by PCR and used to generate the mini-targeting vector (containing selection markers and *lox*P sites) and used for the first targeting to the retrieved BAC DNA in pBluescript backbone. Subsequently, the selection marker was excised by the expression of Cre recombinase in EL350 and a final targeting step was carried out to complete the cko targeting construct. By using pBluescript as a backbone for the vector construct, the problems caused by *lox*P sites present in the BAC backbone can be eliminated. In addition, the long homology arms increase the recombineering efficiency when compared to that obtained with

conventional 45 bp arms, and typically >95% of the colonies obtained after recombineering contain the correct DNA constructs (Liu et al., 2003a). A drawback of this method is that high-copy plasmids are used for retrieving BAC DNA and they can be unstable when harbouring large pieces of DNA. Since cloning steps are still involved, this method is not suitable for high throughput vector construction.

Subsequently, our lab developed a new set of recombineering reagents and protocols that would facilitate the high-throughput generation of cko targeting constructs in a rapid and highly efficient manner (Chan et al., 2007). Two new mobile recombineering reagents were generated, the first of which is a complete λ phage that is replication defective in BAC-harbouring DH10B *E. coli* but still retains its full temperature-inducible homologous recombination functions. In addition, a set of low-copy plasmids (pSim), which contains the three recombineering genes in their native operon, *pL*, under the control of the λ CI repressor. BAC-harbouring cells can be made recombineering-competent by a simple infection (with λ phage) or transformation (with pSim), allowing genetic manipulation to be carried out to facilitate the generation of targeting constructs in a 96-well plate. The overall strategy for using mobile recombineering reagents to generate cko targeting constructs is depicted in Figure 1.7. Briefly, the critical exons to be deleted are identified and the BAC clone containing the region of interest is then made recombineering-competent either by λ phage infection or by transformation with pSim. Next, 70-80 bp of homology sequence (homologous to the target BAC DNA) is included in the primers used to amplify the 5' and 3' targeting cassettes (5' – *Bsd* cassette; 3' – *Neo* cassette). Two sequential targeting steps are then carried out to target the 5' and 3' cassettes to the BAC DNA (Figure 1.7A) and the targeted BAC region is then retrieved into a pBR322-based backbone (PL611) (Figure 1.7B). Finally, the 5' *Bsd* cassette is replaced by a reporter cassette (*lacZ*) to generate the final targeting vector (Figure 1.7C). This type of targeting vector would generate a multi-purpose allele that can serve as a reporter, null and conditional allele (Figure 1.7D). Our new recombineering system enables numerous cko targeting constructs to be generated simultaneously in a 96-well format and has solved a key technical bottleneck in genome-wide targeted mutagenesis programmes of the mouse (Collins et al., 2007).

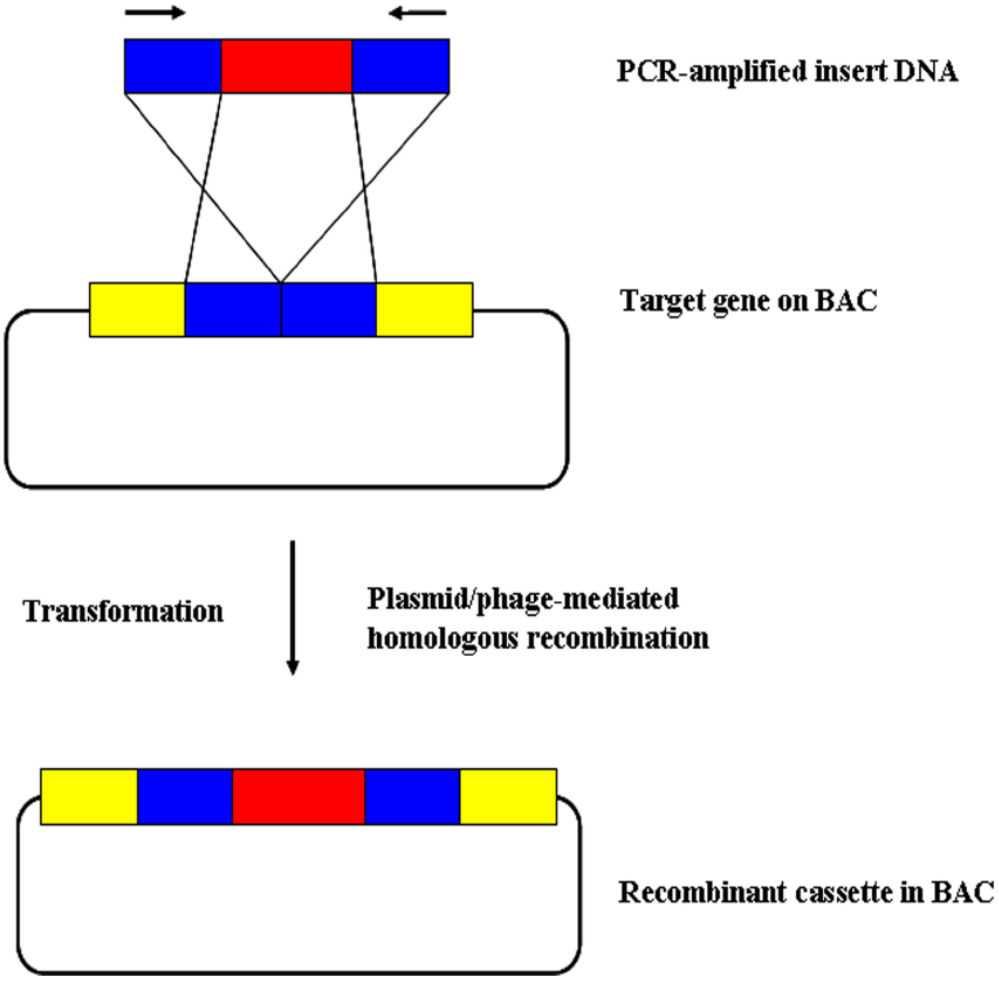


Figure 1.5. Overview of recombineering using Rac-encoded RecET system. PCR amplification is first carried out to generate homology arms (to the target plasmid or BAC DNA) flanking the cassette of interest (such as a selection marker). Next, phage-derived recombination functions are introduced into the target host bacterial strain or the BAC can also be transformed into a bacterial strain that is recombination competent. Finally, the PCR products are electroporated into the target host bacterial strain and recombinant products can be detected by selection or counter-selection.

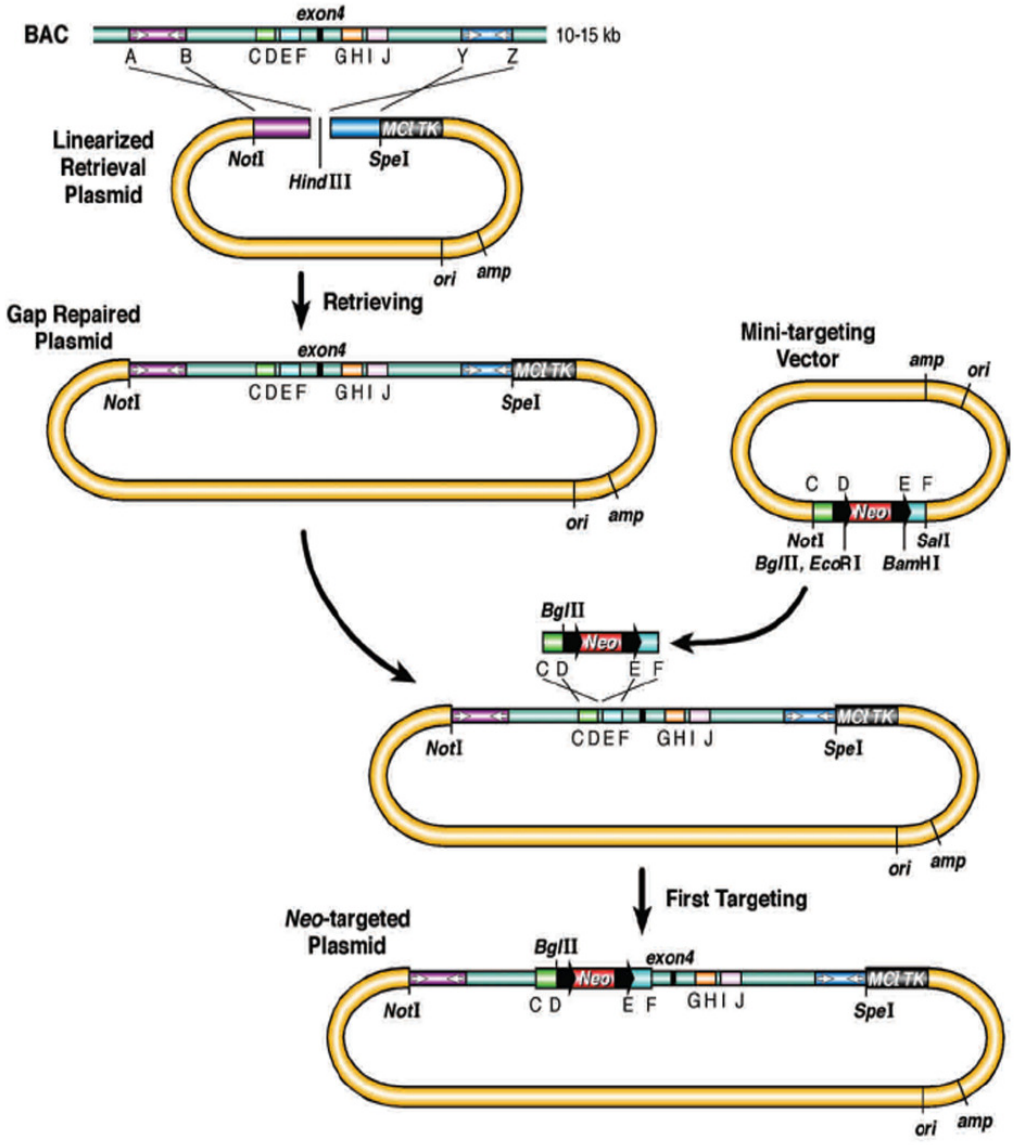
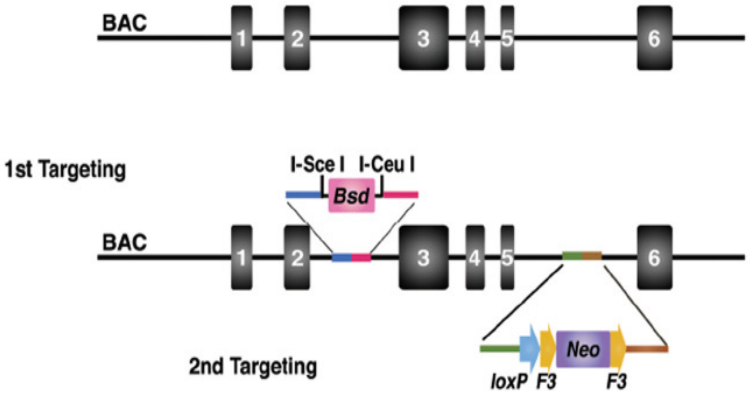
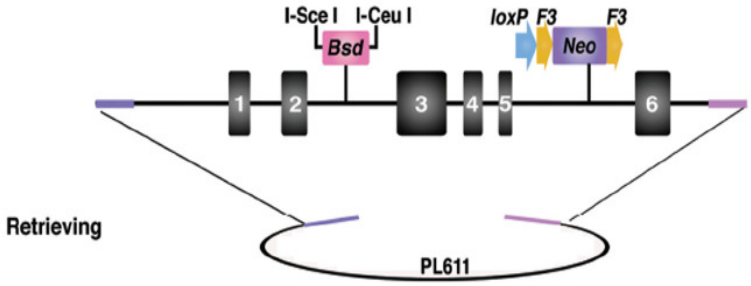


Figure 1.6. Generation of conditional targeting vectors using the Red system. DNA from a BAC containing the region of interest is sub-cloned from the BAC and is used for the construction of conditional targeting vectors. The homology arms used for subcloning and for targeting are generated by PCR using BAC DNA as a template. The two homology arms (purple or dark blue), which are amplified using primers A and B or primers Y and Z, are cloned into an *MC1-TK*-containing plasmid. The gap repair plasmid is then linearized with *Hind*III to create a double-strand break for gap repair. A mini-targeting vector is constructed by ligation of the two PCR products generated by amplification of BAC DNA with primers C and D (light green) or primers E and F (blue), with a floxed selection marker and pBluescript. A *Bgl*II restriction site is included in the mini targeting vector for diagnosing gene targeting in ES cells. The black arrows denote *lox*P sites. The targeting cassette is excised by *Not*I and *Sal*I digestion, or by PCR amplification, using primers C and F. The gap-repaired plasmid and the excised targeting cassette are cotransformed into recombineering-competent DY380 or EL350 cells. The recombinants have a floxed *Neo* cassette inserted between primers D and E and can be selected on kanamycin plates. The *Neo* cassette is excised with Cre recombinase, leaving a single *lox*P site at the targeted locus. Figure obtained from (Liu et al., 2003a).

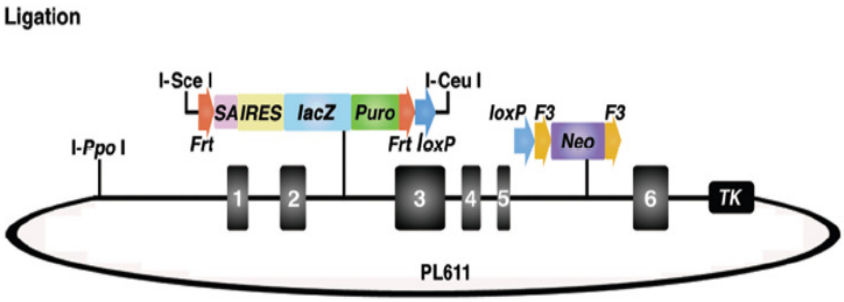
A


BAC
1
2
3
4
5
6
1st Targeting
I-Sce I
I-Ceu I
Bsd
BAC
2nd Targeting
Neo
loxP
F3
F3

B


I-Sce I
I-Ceu I
Bsd
loxP
F3
F3
Neo
1
2
3
4
5
6
Retrieving
PL611

C


Ligation
I-Sce I
I-Ceu I
SA
IRES
lacZ
Puro
loxP
F3
F3
Neo
Frt
Frt
loxP
I-Ppo I
1
2
3
4
5
6
TK
PL611

D

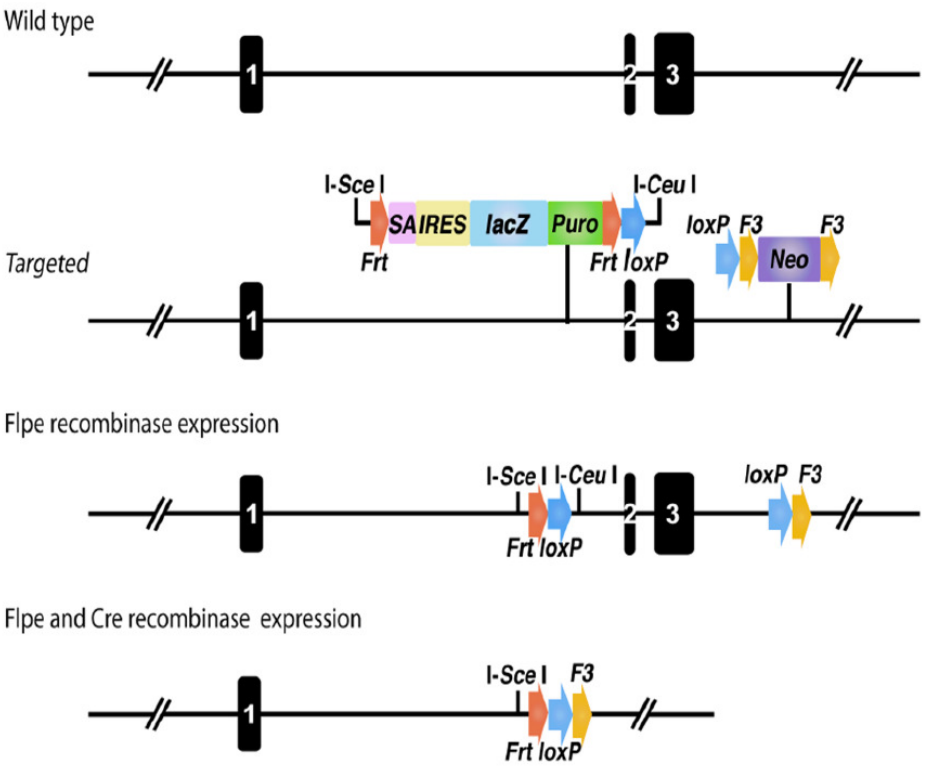


Figure 1.7. Generation of conditional targeting vectors using mobile reagents. (A) The critical exon(s) to be deleted in the cko allele is identified and in this example, exons 3-5 are to be deleted. The *Bsd* cassette, which is flanked by the two rare cutter sites, *I-Sce*I and *I-Ceu*I, is targeted to the 5' side of the intended deletion region. Subsequently, the *lox*P-*F3-PGK-EM7-Neo-F3* (*Neo*) cassette is targeted to the 3' side of the deletion region. Coloured lines represent the short homology arms used for recombineering. **(B)** The genomic DNA fragment is then retrieved from the BAC to PL611, which has the Amp^R gene. **(C)** The *Bsd* cassette can be conveniently replaced by a reporter, i.e. *lacZ*, in a simple ligation reaction. The final targeting vector has the reporter flanked by two *FRT* sites followed by a *lox*P site at the 5' side of the intended deletion region, and a *F3* flanked *Neo* cassette that provides for positive selection in ES cells. The negative selection marker *TK* is added to the vector backbone by recombineering. The vector is then linearized with the rare-cutter *I-Ppo*I before targeting to ES cells. **(D)** The targeted allele is a reporter allele in which endogenous regulatory elements control the expression of *lacZ* and staining for X-gal reveals the spatial expression patterns of the gene. The conditional knockout allele can be obtained by expression of Flpe recombinase either *in vitro* or *in vivo* to delete the intervening sequences between *FRT*/*F3* sites, leaving behind two *lox*P sites flanking critical exons of the gene. The conditional allele can then be crossed to a cell or lineage-specific Cre recombinase-expressing mice. Figure obtained from Chan *et al.* (Chan et al., 2007).

1.3 B-cell lymphoma/leukaemia 11(*Bcl11*) gene family

1.3.1 Discovery of the *Bcl11* genes

Loss-of-function analyses using knockout mice generated using the techniques discussed above have provided great insights into the specific roles of genes at different stages of development and in different tissues. One such example is the *Bcl11* family. The *Bcl11* family consists of two members in the mouse and human, *Bcl11a* and *Bcl11b*, which encode for Krüppel-like transcription factors that contain both single and double C_2H_2 zinc finger motifs. C_2H_2 zinc fingers are modular protein domains that generally confer sequence-specific DNA binding activity, and transcriptional regulatory domains. Structurally, the C_2H_2 zinc finger motif is comprised of a β-hairpin followed by an α-helix that folds around a single zinc ion (Suzuki et al., 1994). Sequence-specific recognition of DNA is mediated primarily by interactions between the variable amino acids within and around the α-helix and nucleotides within the major groove of DNA (Suzuki et al., 1994). Outside of these zinc finger domains, the rest of the coding sequences of Bcl11a and Bcl11b showed no homology to known protein domains in the database, suggesting that the *Bcl11* genes encode for two novel proteins.

Bcl11a and Bcl11b proteins were shown to interact with all members of the chicken ovalbumin upstream promoter transcription factor (COUP-TF) subfamily of orphan nuclear receptors and were termed COUP-TF-interacting proteins 1 and 2, respectively (CTIP1 and CTIP2) (Avram et al., 2000). These studies identified *Bcl11a* (*Evi9*/*Ctip1*) and *Bcl11b* (*Rit1*/*Ctip2*) as C_2H_2 zinc finger transcription factors and preliminary analyses suggest that they may function as transcription repressors.

Bcl11a was first identified as a common site of retroviral integration in BXH2 murine myeloid leukemias and was initially called Ectropic viral integration site 9 (*Evi9*) (Nakamura et al., 2000). Retroviral integration into the *Evi9* locus leads to the development of murine leukaemia either by enhancing the expression of proto-oncogenes or by disrupting the expression of tumour suppressor genes. Because proviral integration into the genomic loci leaves behind a genetic tag, the integration site can be readily determined and the affected gene(s) identified. Retroviral mutagenesis screens in the mouse have been instrumental in identifying numerous leukaemia disease genes (Jonkers and Berns, 1996). Sequence analysis showed that proviral integrations were located

within the first intron of *Evi9* in the reverse transcriptional orientation. This location and orientation is frequently found when retroviral integration activates gene transcription through an enhancer mechanism (Jonkers and Berns, 1996). Northern blots of cell lines with proviral integration at *Evi9* showed elevated levels of *Evi9* expression, consistent with the hypothesis that proviral integration results in the up-regulation of *Evi9* expression, suggesting that *Evi9* may be a putative proto-oncogene (Nakamura et al., 2000). Subsequently, *Evi9* was shown to possess oncogenic properties as transformation of NIH 3T3 cells with *Evi9* resulted in the growth of anchorage-independent colonies in soft agar (Nakamura et al., 2000).

In addition to retroviral screens, radiation-induced malignancies in the mouse, such as thymic lymphomas, can be studied to define the regions of allelic loss and delineate the putative tumour suppressor(s) within the deletion region. In a γ-ray induced mouse thymic lymphoma, *Bcl11b* (also known as Radiation-induced tumour suppressor gene 1, *Rit1*) was identified as a novel tumour suppressor gene located on mouse chromosome 12 (Shinbo et al., 1999; Wakabayashi et al., 2003a). Homozygous deletions of exons 2 and 3 of *Rit1* and point mutations of *Rit1* were detected in these thymic lymphomas and resulted in reduced or no expression of Rit1 protein (Wakabayashi et al., 2003a). In the same study, Rit1 proteins were found to suppress tumour cell growth and there was also a preferential inactivation of *Rit1* in *p53* wild-type lymphomas, suggesting that *Rit1* may be a potential tumour suppressor.

1.3.2 Properties of *Bcl11* genes

The human *BCL11A* gene is located on chromosome 2 while its mouse orthologue is found on chromosome 11. There is a high level of conservation between the human and mouse Bcl11a proteins (99% identity), suggesting that there is conservation of function in both species. There are at least three known isoforms of Bcl11a in both the human and mouse (XL – 835 aa; L - 773 aa and S - 243 aa) (Nakamura et al., 2000; Satterwhite et al., 2001; Weniger et al., 2006), though bioinformatical prediction and Northern blot analysis suggest there could be three more isoforms (Figure 1.8A). The three common Bcl11a isoforms are derived from five exons and all three isoforms contain the first three exons. The longest isoform (XL) contains the sequences from exons 1 to 4 only and is

predicted to have six zinc finger domains. Alternative splicing within exon 4 to exon 5 is thought to result in formation of the other two common isoforms. *BCL11B*, the paralogue of *BCL11A*, is found on chromosome 14 while its mouse orthologue is located on chromosome 12. Similarly, a high level of conservation between the human and mouse Bcl11b proteins is observed (93% identity), again suggesting conservation of function in both species. There are at least two isoforms of Bcl11b in both the human and mouse (α – 894/884 aa; β – 823/812 aa in human/mouse) (Figure 1.8A), which are derived from four exons; the longest isoform (α) contains exons 1 to 4 while the shorter isoform (β) only contains exons 1, 2 and 4.

Bcl11a and Bcl11b share 67% similarity at the nucleotide and 63% similarity at the amino acid levels, but the level of similarity increases to 95% between the zinc finger domains (Figure 1.8B and 1.8C). Exon 4 of both genes, which encodes 75% of the protein sequence, constitutes the main functional domain of these proteins. Bcl11a (XL) contains six zinc fingers, a proline-rich domain between zinc fingers 1 and 2 and an acidic domain between zinc fingers 3 and 4 that contains a run of 21 consecutive residues (Figure 1.8B-D). Zinc fingers 1 and 6 are different from zinc finger 2, 3, 4 and 5 in that they have 4 amino acids separating the two zinc-binding histidines, whereas the others have 3 amino acids (Satterwhite et al., 2001). The internal zinc fingers (2, 3, 4 and 5) are arranged in pairs, with each pair being separated by a canonical 'linker' sequence. The Bcl11a (L) isoform contains only four zinc fingers as it lacks the two C_2H_2 zinc fingers present at the C-terminus of Bcl11a (XL), and the shortest isoform (S), has only one zinc finger domain. The two Bcl11b isoforms contain six zinc fingers domains, a proline-rich domain between zinc fingers 1 and 2, and an acidic domain between zinc fingers 3 and 4 that contains a run of 13 consecutive residues (Figure 1.8B-D). Thus, the Bcl11 transcription factors have similar sequences, structures and physical properties.

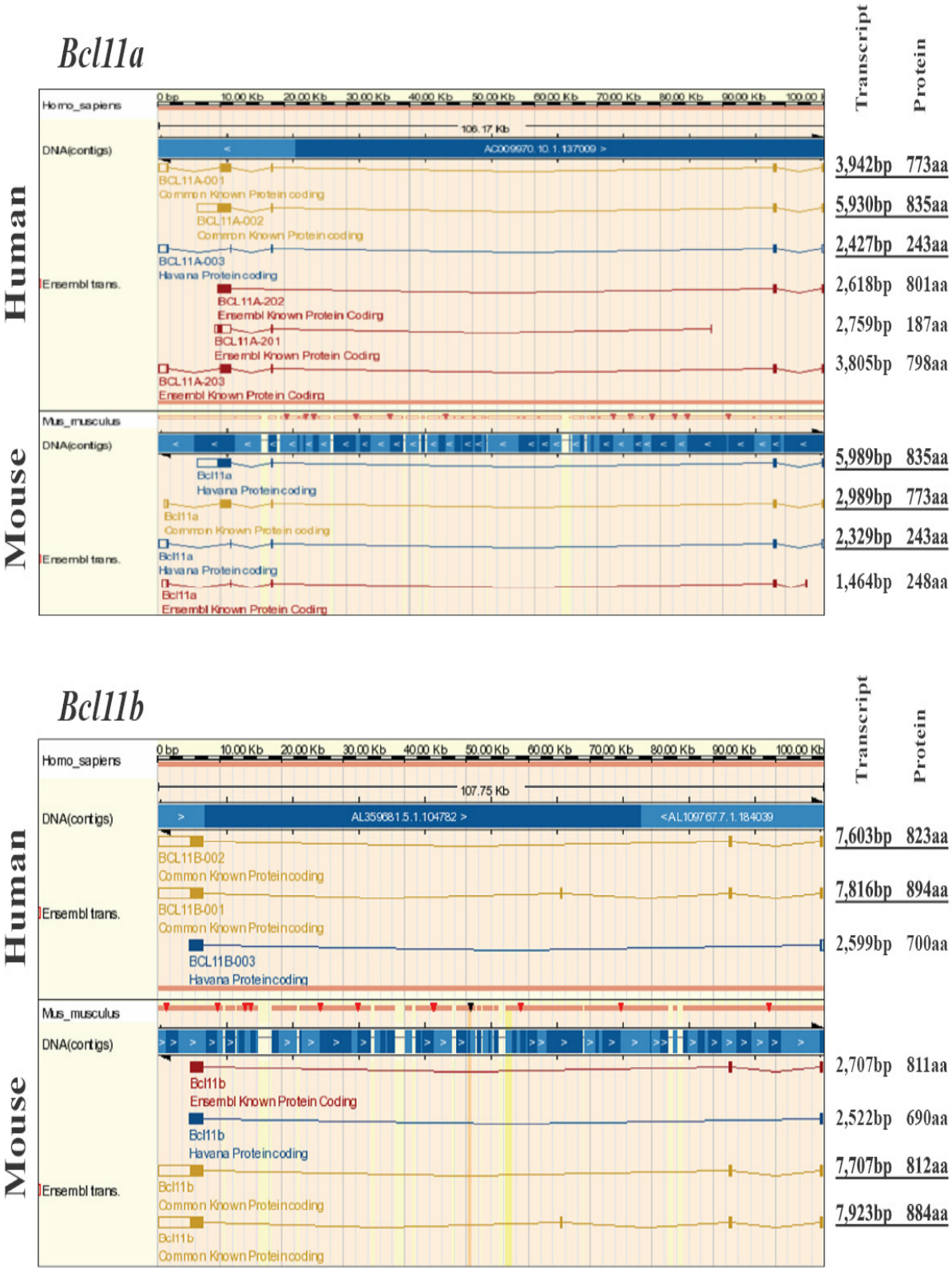
Different murine Bcl11a isoforms were shown to have different subcellular localization patterns. For example, one isoform (L) was shown to be located exclusively in the nucleus and was distributed within spherical nuclear structures, 50 to 500 nm in diameter (Nakamura et al., 2000). In contrast, the shortest isoform (S) was detected predominantly in the cytoplasm. In addition, Bcl11a (L) colocalized with BCL6 but not with PML or Sp100 (Nakamura et al., 2000). Further analysis by coimmunoprecipitation

and GST pull-down assay showed that Bcl11a and BCL6 interact directly, suggesting that Bcl11a may function in part through its interaction with BCL6 (Nakamura et al., 2000). Similarly, different human isoforms of BCL11A also show varying subcellular localization: the human BCL11A (XL) was shown to localize exclusively within nuclear dots (paraspeckles), and colocalized with the major fraction of BCL6 found inside the nucleus, while BCL11A (L) showed fewer nuclear dots and more diffuse nuclear staining, but nevertheless, BCL11A (L) colocalized with BCL6 (Liu et al., 2006; Pulford et al., 2006). In contrast, the shortest BCL11A (S) isoform is only relocalized from the cytoplasm to the nucleus by interaction with the other isoforms.

Putative target DNA binding sites of Bcl11a proteins have also been determined by reiterative oligonucleotide selection. The human BCL11A (XL) isoform binds to the core consensus 5'C-C-C/T-A/G-C-3' (Liu et al., 2006), which differs from the consensus binding sequence (5'-GGCCGGAGG-3') determined for the mouse Bcl11a (L) isoform (Avram et al., 2002). However, neither of these binding sites has been validated *in vivo*. As discussed above, both Bcl11 proteins were identified as COUP-TF interaction partners (Avram et al., 2000). COUP-TF family members generally mediate transcriptional repression by recruiting nuclear receptor co-repressor (NCoR) and/or silencing mediator for retinoid and thyroid hormone receptor (SMRT) to the template (Avram et al., 2000). NCoR and SMRT are components of a larger repressor complex that also includes mSin3A/B and a trichostatin-sensitive histone deacetylase (Heinzel et al., 1997). Avram *et al*. investigated the nature of the interaction between Bcl11a and ARP1, a COUP-TF member in yeast, and found that Bcl11a contained two independent ARP1 interaction domains, ID1 and ID2 that bound to the putative AF-2 domain of ARP1 *in vitro* (Avram et al., 2000). Bcl11a exhibited a punctuate distribution in the nucleus and recruited co-transfected ARP1 to these foci. Interestingly, Bcl11a potentiated the transcriptional repression activity of ARP1 in HEK293 cells independently of trichostatin-sensitive histone deacetylation, suggesting that Bcl11a does not mediate transcriptional repression by acting through recruitment of trichostatin-sensitive class I or II histone deacetylase(s). These results demonstrate that Bcl11a and Bcl11b transcription factors may either bind directly to the target sites in a sequence-specific manner or be recruited to the target sites by a COUP-TF family member. Recent studies demonstrated that Bcl11a and Bcl11b

recruited sirtuin 1 (SIRT1), a trichostatin-insensitive, nicotinamide-sensitive class III histone deacetylase, to the promoter region of a reporter gene template in HEK293 cells (Senawong et al., 2003; Senawong et al., 2005). Additionally, SIRT1 was shown to catalyze the deacetylation of histones H3 and/or H4 on the reporter gene template, suggesting that SIRT1 contributes at least partially to the transcriptional repression activities of both Bcl11 proteins. Interestingly, Bcl11b was also shown to be associated with the nucleosome remodelling and deacetylase (NuRD) complex in T lymphocytes (Cismasiu et al., 2005). Further analysis showed that the endogenous Bcl11b complexes from $CD4^+$ T lymphocytes harbour histone deacetylase activity sensitive to trichostatin A and that both metastasis-associated proteins MTA1 and MTA2 interact directly with BCL11B. In summary, the current data suggests that the transcriptional repression activities of Bcl11 proteins could be mediated by recruitment of SIRT1 (trichostatin-insensitive) and/or NuRD complex (trichostatin-sensitive) depending on the cellular context.

A


Bcl11a
Human
Mouse
Transcript
Protein
Homo_sapiens
DNA(contigs)
Ensembl trans.
106.17 Kb
AC009970.10.1.137009 >
BCL11A-001
Common Known Protein coding
3,942bp 773aa
BCL11A-002
Common Known Protein coding
5,930bp 835aa
BCL11A-003
Havana Protein coding
2,427bp 243aa
BCL11A-202
Ensembl Known Protein Coding
2,618bp 801aa
BCL11A-201
Ensembl Known Protein Coding
2,759bp 187aa
BCL11A-203
Ensembl Known Protein Coding
3,805bp 798aa
Mus_musculus
Bcl11a
Havana Protein coding
5,989bp 835aa
Bcl11a
Common Known Protein coding
2,989bp 773aa
Bcl11a
Havana Protein coding
2,329bp 243aa
Bcl11a
Ensembl Known Protein Coding
1,464bp 248aa
Bcl11b
Human
Mouse
Homo_sapiens
DNA(contigs)
Ensembl trans.
107.75 Kb
AL359681.5.1.104782 >
<AL109767.7.1.184039
BCL11B-002
Common Known Protein coding
7,603bp 823aa
BCL11B-001
Common Known Protein coding
7,816bp 894aa
BCL11B-003
Havana Protein coding
2,599bp 700aa
Mus_musculus
Bcl11b
Ensembl Known Protein Coding
2,707bp 811aa
Bcl11b
Havana Protein coding
2,522bp 690aa
Bcl11b
Common Known Protein coding
7,707bp 812aa
Bcl11b
Common Known Protein coding
7,923bp 884aa

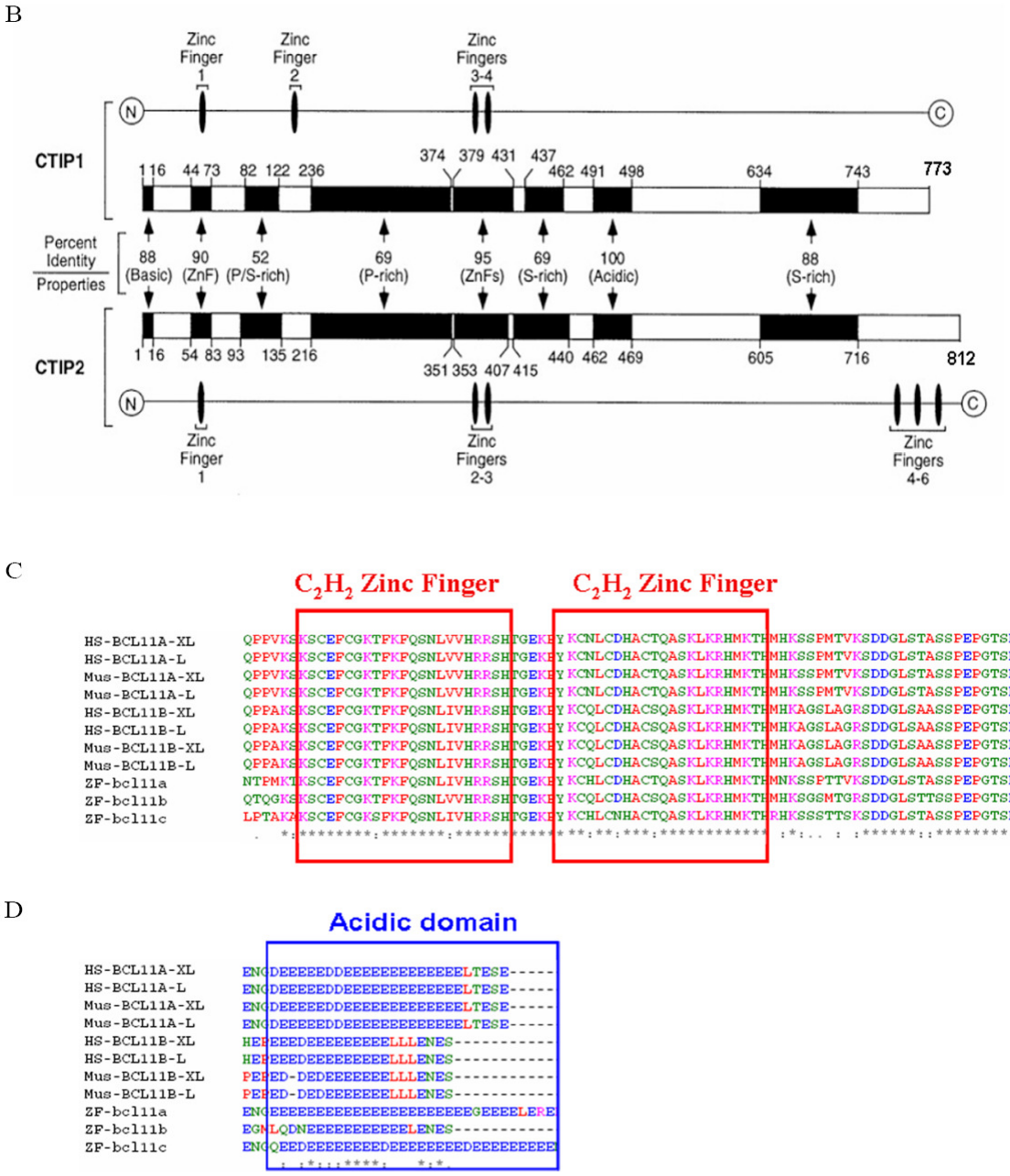


Figure 1.8. Properties of *Bcl11* genes. (A) The human *BCL11A* gene is located on chromosome 2 while the mouse *Bcl11a* gene is found on chromosome 11. The human *BCL11B* gene is located on chromosome 14 while the mouse *Bcl11b* gene is found on chromosome 12. Each gene is predicted to have multiple transcripts and proteins; the length and size of the transcripts and proteins are shown beside the diagrammatic representation of each gene. (Underlined numbers – transcripts and proteins that have been detected experimentally). **(B)** Schematic diagram of Bcl11a and Bcl11b amino acid alignment. The homologous regions are represented by black boxes and the percentage of identity between each region of Bcl11a and Bcl11b is indicated. Bcl11a is predicted to have three C_2H_2 zinc finger domains while Bcl11b is predicted to have an extra three zinc finger domains at its C-terminus. (Figure adapted from Avram *et al.*) (Avram et al., 2000). Clustal w aligment of **(C)** C_2H_2 zinc finger domains and **(D)** acidic domains of human, mouse and zebrafish Bcl11 proteins. These alignments show that the C_2H_2 zinc finger and acidic domains are highly conserved between the three species.

1.3.3 Bcl11 and tumorigenesis

The discovery of the *Bcl11* genes has brought about great interest because their expression has been linked to the etiology of disease in both mice and humans. Over-expression of Bcl11a following proviral integration resulted in the development of myeloid leukaemia in mice (Nakamura et al., 2000). In addition, over-expression of Bcl11a transformed NIH 3T3 cells and this transformation event is thought to be partly facilitated by the physical interaction of Bcl11a with BCL6. Importantly, deregulated expression of *BCL11A* has been consistently found in rare but clinically aggressive cases of B-cell malignancy with chromosomal translocation t(2; 14) (p13; q32.3) (Satterwhite et al., 2001). These data suggest that *Bcl11a* is a proto-oncogene in both the mouse and human. Moreover, recent studies have implicated *BCL11A* in other human diseases, for example, two mutations in *BCL11A* were identified and validated in human breast cancers (Wood et al., 2007) and a quantitative trait locus (QTL) that influences F-cell production (that measures the presence of fetal haemoglobin) was mapped to *BCL11A* in human thalassemia patients (Menzel et al., 2007).

The majority of studies have suggested that, in contrast to *Bcl11a*, *Bcl11b* may be a tumour suppressor. p53 is an important transcription factor that functions as a tumour suppressor and is critical in preventing the genomic instability that leads to cancer. An association between p53 and Bcl11b in mice has been implicated in the development of thymic lymphomas. It was reported that inactivation of *Bcl11b* in normal thymocytes triggered cell proliferation and this was accompanied by a profound p53-dependent apoptosis, whereas *Bcl11b* deficiency appeared to activate cell growth in transformed lymphocytes (Wakabayashi et al., 2003b). In a separate study, preferential inactivation of *Bcl11b* was found in *p53* wild-type lymphomas, suggesting that loss of *Bcl11b* may contribute to oncogenesis only in *p53*-proficient lymphocytes or that it may not be required in p53-null lymphomas (Wakabayashi et al., 2003a). Another functional link between p53 and Bcl11b was demonstrated by Okazuka *et al.* who showed that inactivation of p53 is sufficient for $CD4^-CD8^-$ double-negative ($CD44^-CD25^+$) DN3 stage thymocytes to differentiate into the $CD4^-CD8^+$ immature single-positive, but not to the $CD4^+CD8^+$ double-positive stage of thymocyte development in $Bcl11b^{-/-}$ mice (Okazuka et al., 2005). In addition, $Bcl11b^{+/-}p53^{+/-}$ mice exhibited greater susceptibility to

lymphomas than *Bcl11b*$^{+/+}$*p53*$^{+/-}$ mice, suggesting that functional loss of one *Bcl11b* allele confers a selective advantage for tumour growth (Kamimura et al., 2007). *Bcl11b* has also been shown to be involved in other human disorders, for example BCL11B was shown to repress Tat-mediated transcriptional activation and to inhibit HIV-1 replication in human microglial cells (Rohr et al., 2003). *BCL11B* was also disrupted in a novel chromosomal aberration, inv(14)(q11.2q32.31), in T-cell acute lymphoblastic leukaemia (T-ALL), which resulted in the absence of wild-type BCL11B (Przybylski et al., 2005). Taken together, these studies underline the importance of *Bcl11* genes in human diseases, and underscore the need for further studies on their functional and expression patterns during development.

1.3.4 Hematopoiesis and the roles of *Bcl11* genes in hematopoietic development

As discussed above, hematopoietic malignancies account for most of the abnormalities observed in humans and mice when *Bcl11* genes are dysregulated. As such, the majority of the research to-date has focused on dissecting the roles of *Bcl11* genes in hematopoietic lineages. All the hematopoietic cells, including the red blood cells, platelets and white blood cells, are derived from the same progenitor or precursor cells – the hematopoietic stem cell (HSC) (Figure 1.9A). The self-renewal HSCs are mostly quiescent but can be rapidly mobilized from their bone marrow (BM) niche to proliferate and differentiate into lineages of the innate and adaptive immune systems, as well as into red blood cells and platelets. In the mouse, the contribution of the yolk-sac to the formation of adult HSCs remains controversial and it is thought that most HSCs arise from the aorta-gonand-mesonephros (AGM) region of the embryo (Godin et al., 1995). These HSCs then seed the fetal liver where embryonic hematopoiesis occurs and the stem cells proliferate and differentiate into myeloid and lymphoid lineages (Godin and Cumano, 2002). Subsequently, the HSCs migrate out of the fetal liver and populate the spleen and eventually the BM, where a specialized microenvironment forms – the endosteal niche - where HSCs become quiescent (Godin and Cumano, 2002).

The continual generation and replenishment of the hematopoietic system involves the sequential commitment of the HSCs to the more restricted multi-potent progenitor

cells (lineage-restricted progenitor cells such as common lymphoid progenitors and myeloid progenitors) and eventually to the functionally distinct mature blood cells (Cantor and Orkin, 2001). The myeloid progenitor is the precursor of the granulocytes, macrophages and some dendritic cells of the immune system (Figure 1.9A). Granulocytes (neutrophils, eosinophils and basophils) are involved in the clearing of bacterial and parasitic infections and as such, their production is increased during the immune response. Macrophages are the mature form of monocytes which play a critical role in innate immunity. Dendritic cells are specialized antigen-presenting cells that take up antigens and display them to lymphocytes.

The common lymphoid progenitor gives rise to the lymphocytes: T lymphocytes (T cells), B lymphocytes (B cells) and natural killer cells. T cells can be classified into two main classes: Cytotoxic T cells, which kill cells infected with viruses, and Helper T cells, which differentiate into cells that activate other cells such as B cells and macrophages. T cell progenitors are produced in the BM but migrate to the thymus where they undergo maturation (Rothenberg, 2007a). T cell development proceeds through a series of stages that are well defined by the expression of cell surface makers (Rothenberg, 2007a) (Figure 1.9B). There are two types of thymocytes that are normally present within the thymus: those expressing αβ T cell receptor (αβ T cells) and those expressing γδ (γδ T cells). The αβ T cells later develop into two distinct functional subsets, CD4 and CD8 single positive mature T cells. Development of the prospective αβ T cells proceeds through several stages, one of which is the double-negative (DN) stage, when both CD4 and CD8 are not expressed. This DN stage can be further subdivided into at least four stages (DN1-4) depending on the expression of the adhesion molecules CD44 and CD25 (the α chain of the IL-2 receptor) (Figure 1.9B): the DN1 stage is characterized by expression of CD44 but not CD25; the DN2 stage by expression of both CD44 and CD25; the DN3 stage by expression of CD25 and down-regulation of CD44 and finally, the DN4 stage by down-regulation of both CD44 and CD25 (Rothenberg, 2007a). An important checkpoint occurs at the transition between DN3 to DN4 where thymocytes that failed to produce a functional TCRβ protein undergo apoptosis. Subsequently, thymocytes begin to express both CD4 and CD8 (double-positive, DP) and

eventually, these DP cells mature and express either CD4 or CD8; becoming single-positive thymocytes.

B cell development occurs in the BM, and after immature B cells migrate to the spleen, they further develop into antibody-secreting memory B cells and plasma cells (Hardy and Hayakawa, 2001). Similar to T cell development, B cell maturation occurs through several distinct stages that can be identified by the expression of cell surface markers (Figure 1.9C). The earliest B-lineage cells are known as pre/pro-B cells and can be characterized by the expression of B220 and CD43 cell surface markers. These cells develop into pro-B cells and up-regulate the expression of CD19 (Figure 1.9C). The pro-B cells can be further divided into 'early' or 'late' pro-B cells depending on their expression of BP-1: early Pro-B cells do not express BP-1 but late Pro-B cells do. Next, the pro-B cell develops into a pre-B cell which no longer expresses CD43. Pre-B cells will have successfully undergone immunoglobulin (Ig) loci rearrangement to express IgM. These immature B cells begin to express IgD in addition to IgM and are now known as mature (naïve) B cells. Naïve B cells circulate in the blood and peripheral lymphoid tissues such as the spleen, where upon encountering antigens, these cells differentiate into antibody-producing plasma cells or memory B cells.

Since both *Bcl11a* and *Bcl11b* are involved in hematopoietic disorders, each of these genes was knocked out in the mouse to characterize their roles in hematopoietic development (Liu et al., 2003b; Wakabayashi et al., 2003b). Interestingly, even though Bcl11a and Bcl11b are highly similar in terms of their sequence, structure and physical properties, the knockout mice have different phenotypes. As shown by *Liu* et al., loss of *Bcl11a* in the mouse resulted in a complete block of B cell development, which occurred at the earliest stage of B cell development (Liu et al., 2003b). This phenotype, together with the observation in humans that over-expression of *BCL11A* causes B cell tumourigenesis, demonstrates that *Bcl11a* is essential for B cell development (Satterwhite et al., 2001; Weniger et al., 2006). Although proviral insertion in the *Bcl11a* locus resulted in over-expression of *Bcl11a*, leading to myeloid leukaemia (Nakamura et al., 2000); there was no obvious defect in myeloid development in the knockout mice. In addition to the B cell defects, T cell development was also affected in the absence of *Bcl11a*. The proportion of αβ T cells to γδ T cells was perturbed and there was a three- to

four-fold increase in the number of γδ T cells concomitant with a loss of αβ T cells (Liu et al., 2003b). Loss of *Bcl11a* also affected subsets of αβ T cells as there was an increase in the population of $CD4^-CD8^-$ and $CD4^{lo}CD8^+$ cells and an absence of $CD4^{hi}CD8^-$ thymocytes. Intriguingly, a 'non-cell autonomous tumour suppressor function' for *Bcl11a* was identified in the same study. When mutant *Bcl11a* fetal liver cells were transplanted into lethally-irradiated host mice, the transplanted mice developed $CD4^+CD8^+$ T cell leukaemia three months after transplantation. Further analysis of these tumours showed that most of the cells originated from the wild-type host and not from *Bcl11a* mutant cells (Liu et al., 2003b). This suggests that mutant *Bcl11a* behaves in a cell non-autonomous manner in the development of T cell leukaemia in these mice. In addition, in these T cell leukemic cells, there was a dramatic up-regulation of *Notch1* transcripts. As Notch1 signalling has been shown to inhibit B cell development and promote T cell development (Radtke et al., 2004), a possible explanation for the *Bcl11a* knockout phenotype would be that *Bcl11a* normally inhibits Notch1 to permit B cell development; loss of *Bcl11a* would release the brake on *Notch1*, which would result in a block in B cell development. However, formal proof of this hypothesis is still lacking.

On the other hand, knockout of *Bcl11b* in the mouse resulted in a block in thymocyte development (Wakabayashi et al., 2003b). Only αβ T cell development was affected while γδ T cells and B cells were unaffected in the *Bcl11b* mutant mice. Further analysis showed that rearrangement of the TCRβ locus in *Bcl11b* mutant double negative (DN) thymocytes is incomplete – while they successfully rearranged from D to J, they did not rearrange from V to D. T cell development was predominantly arrested at the DN3 stage ($CD44^-CD25^+$) and some at the immature single positive stage. Host mice transplanted with mutant *Bcl11b* fetal liver cells showed reconstitution of B cells and γδ T cells as well as immature (DN1 and DN2), but not mature T cells, suggesting that the block in T cell development occurred after the γδ T cell lineage commitment. In addition, extensive apoptosis was observed in *Bcl11b* mutant thymocytes. These results suggest that *Bcl11b* is a critical regulator of both development and survival of thymocytes.

Collectively, the knockouts of *Bcl11a* and *Bcl11b* clearly establish their essential roles at different stages of lymphopoietic development and illustrate the importance of transcription factors in hematopoiesis.

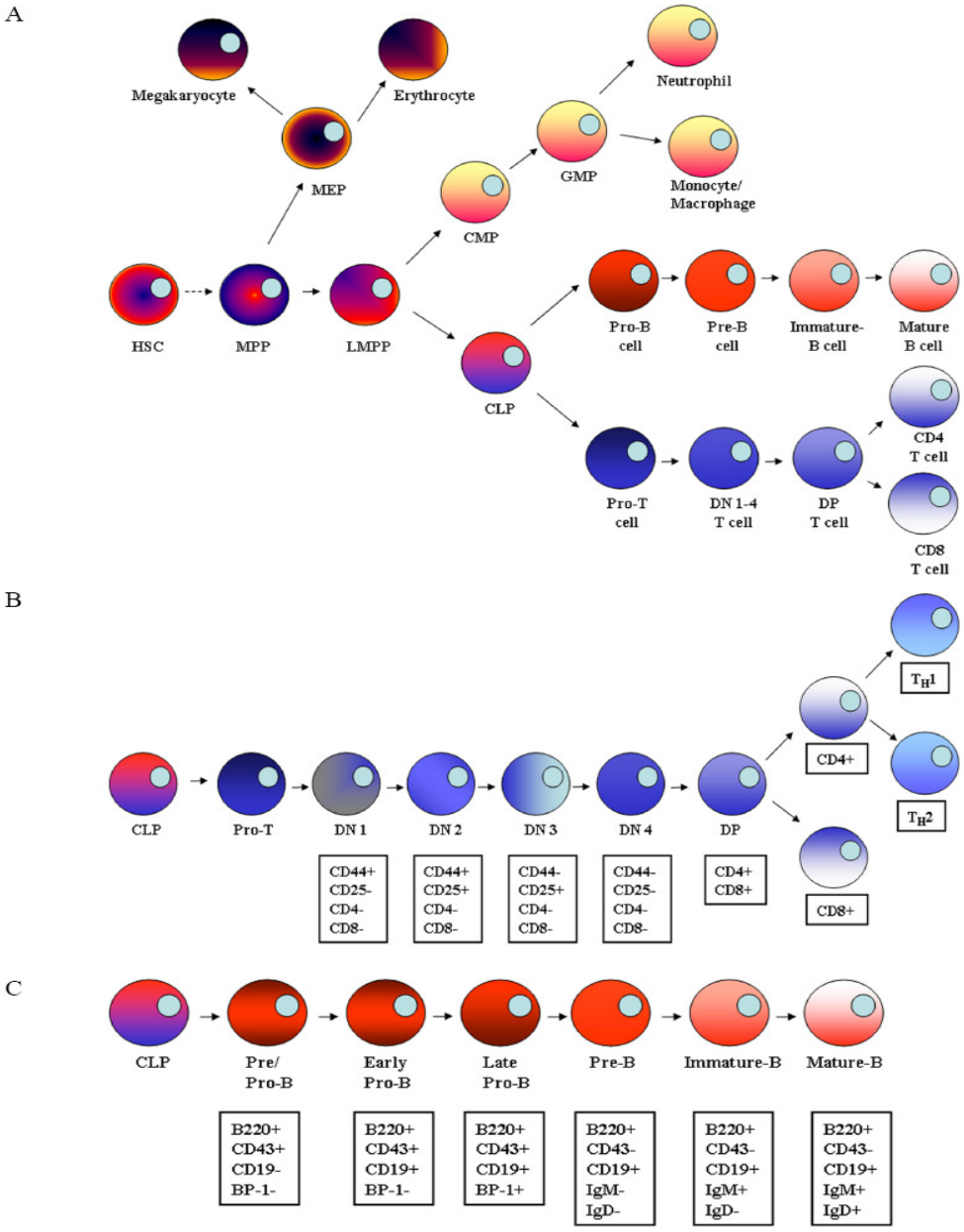


Figure 1.9. Schematic diagram showing hematopoietic development. (A) All lineages of the blood cells including red blood cells and platelets are believed to derive from the hematopoietic stem cells (HSCs). Proliferation and differentiation of the HSCs generate the more lineage-restricted multi-potent progenitors such as megakaryo-erythrocyte progenitor (MEP); common myeloid progenitor (CMP); common lymphoid progenitor (CLP). These lineage-restricted progenitors are responsible for setting up the entire repertoire of the hematopoietic system. **(B)** Schematic diagram depicting various stages of T lymphocyte development. From the CLP, thymocyte development can be defined into different stages characterized by the expression of cell surface markers CD4, CD8, CD44 and CD25. **(C)** Schematic diagram showing B lymphocyte development from the CLP. The various B cell development stages can be defined by expression of cell surface markers B220, CD43, BP-1, IgM and IgD. Abbreviation used: HSC: Hematopoietic stem cell; MPP: Multi-potent progenitor; LMPP: Lymphoid-primed MPP; CLP: Common lymphoid progenitor; DN: CD4/CD8 double-negative T cell; DP: CD4/CD8 double-positive T cell; CMP: Common myeloid progenitor; GMP: Granulocyte/macrophage progenitor; MEP: Megakaryocyte/erythroid progenitor.

1.3.5 Transcription factors in hematopoiesis

Understanding the molecular mechanisms that determine lineage specification of the entire repertoire of blood cells from HSCs is critical to the field of hematopoiesis. Extensive studies have shown that lineage-specific transcription factors play integral roles in hematopoietic lineage decisions (Cantor and Orkin, 2001; Orkin and Zon, 2008; Rothenberg, 2007a). A simplistic view of cell lineage commitment would be that certain lineage-specific transcription factors ('master regulators') act principally in a positive manner to specify lineage choice by activating unique programs of gene expression. Recent studies have shown that this interpretation is overly simplistic; instead, transcription factors act in a combinatorial fashion to direct cell lineage commitment and the establishment of certain cell fate (Cantor and Orkin, 2001; Orkin and Zon, 2008). Transcription factors are key components of complex gene regulatory networks. They orchestrate the entire developmental process through a series of co-ordinated hierarchical stages, allowing the determination of different cell fates from multi-potent stem or progenitor cells. Gene regulatory networks consist of many different yet interconnected hierarchical states and each of these different states is determined by a unique combinatorial code of transcription factors (Singh et al., 2005). Higher up the regulatory hierarchy, expression of certain transcription factors at a crucial time is critical to activate target genes that are the executioner of the lineage specification process. For example, Notch1 which is important for the cell-lineage specification of T cells from the common lymphoid progenitor (CLP), is not required to maintain γδ T cells past the DN2 stage nor for αβ T cell development after β-selection, except for a much later role in promoting T helper 2 (T_H2) subset differentiation in response to antigen (Artavanis-Tsakonas et al., 1999; Radtke et al., 2004). In both the initial T cell specification and T_H2 cell differentiation, Notch1 works in part by inducing expression of *Gata-3* (Amsen et al., 2007; Fang et al., 2007). However, in the CD4 versus CD8 T cell lineage decision, Notch1 is not required for the Gata-3 dependent determination of CD4 cell fate (Radtke et al., 2004; Rothenberg, 2007a). Taken together, these studies illustrate two pertinent points: (1) homeostasis of the level of transcription factors at each cell state is critical to initiate and define the specific cell fate; and (2) temporal and spatial regulation of the expression of these transcription factors is also vital to the cell-fate decisions.

In addition, it is becoming apparent that lineage-specific transcription factors exhibit cross-antagonism. Whilst specifying a certain lineage by activating lineage-specific genes, transcription factors simultaneously exert inhibitory effects on the alternative lineage gene programs by directly antagonizing the actions of opposing transcription factors. Thus, there appears to be a dynamic balance of forces that ultimately determine the phenotype of a cell. Such cross-antagonism of transcription factors in determining the cell fate has been reported in the hematopoietic system. For example, *Blimp1* and *Pax5* act in a mutually repressive opposing fashion to control the mature B cell versus plasma cell developmental switch (Davis, 2007); in T-helper cells, *Gata-3* and *T-bet* specify the T_H2 and T_H1 cell fates, respectively (Ho and Pai, 2007; Zheng and Flavell, 1997).

In summary, the gene regulatory networks play critical roles in establishing the hematopoietic system; precise, co-ordinated regulation of the levels of transcription factors facilitate the cell-fate decisions of progenitors and thus the generation of the entire repertoire of differentiated hematopoietic progenies from the HSC. However, the 'information content' of a transcription factor's activity depends on the cellular context and is not permanent or fixed for a particular lineage. This means a transcription factor can participate in the control of completely different sets of gene targets (Orkin and Zon, 2008; Rothenberg, 2007a). The functional flexibility of transcription factors depends on 'local factors' such as the chromatin status of a cell, which controls the accessibility of different target genes in different cell types, thereby allowing the same transcription factor to perform different functions at respective anatomical locations. Frequently, key transcription factors have been found to play essential roles in different tissue systems. For example, in addition to its role in T cell development, *Gata-3* is also important for skin, nervous system, adrenal and jaw development (Kaufman et al., 2003; Lim et al., 2000; Zheng and Flavell, 1997). Recently, *Gata-3* has also been shown to be a critical factor in epithelial cell-lineage determination in the mammary gland (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). These studies also demonstrate that the genetic control between lymphocyte development and mammary development is likely to be conserved.

1.4 Mammary gland biology

The mammary gland serves to provide nourishment and passive immunity via milk to newborns until they are capable of independent feeding. In order to fulfil its function, the mammary gland has to develop an extensive network of milk-producing cells and interconnecting ducts to produce and deliver the milk to the young. Development of the mammary gland occurs in distinct phases that can be broadly classified into embryonic, postnatal (pre-pubertal and pubertal), pregnancy, lactation and involution stages. The entire developmental cycle involves specification of progenitors, proliferation and differentiation of these progenitors to generate the entire mammary gland and also apoptosis, de-differentiation and tissue remodelling during post-lactational regression. Hence the mammary gland is an excellent model system to study the gene regulatory networks that control specification, proliferation, differentiation, survival and death of cells. Understanding the molecular mechanisms that control these processes is important not only from a developmental biologist's point of view but also allows us to elucidate how dysregulation of these processes can lead to breast cancer. Eventually, this may facilitate the design of better therapeutics to treat breast cancer, which is the most commonly occuring cancer in women in the UK (Cancer Research UK). The mammary developmental process is tightly regulated by temporal and spatial stimuli, such as steroid and peptide hormones, through various signalling networks (Hennighausen and Robinson, 1998, 2005). Interestingly, recent studies have shown that transcription factors that are important in hematopoiesis also function in various aspects of mammary lineage commitment (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006; Oakes et al., 2008). The relative contributions of hormonal regulation and transcription factors to mammary lineage commitment, differentiation and development of the mammary gland will be discussed in greater detail below.

1.4.1 Embryonic development

1.4.1.1 Overview of development

The initial stages of mammary development are independent of systemic cues and instead depend on reciprocal signalling between the mammary epithelium and the underlying mesenchyme. Embryonic mammary gland development in the mouse is

initiated around mid-gestation, at 10.5 days post-coitum (dpc), with the formation of mammary (milk) lines that run in an anteroposterior fashion along the ventral side of both the male and female embryos (Figure 1.10A). By 11.5 dpc, the mammary lines have given way to five pairs of mammary placodes that form at specific locations along the mammary line which eventually invaginate into the underlying mesenchymal layer to form mammary buds at 12.5 dpc (Figure 1.10B). Further development of the mammary bud and the surrounding mesenchymal cells continues between 13.5-15.5 dpc, when the mesenchymal cells around the mammary bud condense and differentiate to form the mammary mesenchyme and when the mammary bud begins to elongate and form a sprout (Figure 1.10C). Importantly, while the female embryonic mammary gland remains quiescent until 15.5 dpc, the dense mammary mesenchyme in the male embryos expresses androgen receptors that respond to testosterone and result in the disruption of the mammary anlage (Kratochwil and Schwartz, 1976). Between 15.5-18.5 dpc, the mammary sprout continues to penetrate into the underlying fat pad and undergoes arborization to eventually form a rudimentary ductal tree consisting of a primary duct and 15-20 ductules at birth (Hens and Wysolmerski, 2005).

1.4.1.2 Specification of the milk line

The milk line consists of an ectodermal thickening that arises in the skin and can be visualised using *in situ* staining for the wingless gene, *Wnt10b*, which is one of the earliest markers of mammary lineage (Veltmaat et al., 2004). Using the TOP-GAL Wnt reporter line (Chu et al., 2004), the dynamics of early mammary gland development was monitored and it was found that specification and formation of the milk line depends on canonical Wnt signalling. Forced activation of Wnt signalling promotes development of mammary placodes while inhibition of Wnt signalling blocks initiation of placode formation (Chu et al., 2004). Following the specification of the milk line, a cascade of Wnt expression was shown to be induced (includes *Wnt10b*, *Wnt10a* and *Wnt6*) within the milk line and eventually becomes localized to the mammary placodes (Chu et al., 2004). In addition to Wnt signalling, the fibroblast growth factor (Fgf) signalling pathway has also been suggested to be important for specification of the milk lines. Between 10.5-11.5 dpc, *Fgf10* is expressed in the most ventral-lateral reaches of the dermatomyotome

of the somites adjacent to the developing mammary line while *Fgfr2b* is expressed within the mammary epithelial placodes (Mailleux et al., 2002). Knocking out either *Fgf10* or its receptor *Fgfr2b* in the mouse disrupted four out of the five pairs of the mammary placodes (Numbers 1, 2, 3 and 5) (Mailleux et al., 2002). In addition, inhibition of Wnt signalling does not alter the expression of *Fgf10* or *Fgfr1* signalling (Mailleux et al., 2002), suggesting that Fgf signalling is important for the earliest stage of mammary development and probably acts in concert with Wnt signalling, rather than downstream of it. The T-box transcription factor, *Tbx3,* is also expressed in the mammary line from 10.25-10.5 dpc and formation of mammary placodes is abolished in *Tbx3* knockout mice (Eblaghie et al., 2004). In addition, *Tbx3* expression was shown to be induced by both Fgf and Wnt signalling within the mammary line of cultured mouse embryos (Eblaghie et al., 2004), thus establishing a link between *Tbx3* and the Fgf and Wnt pathways in specification of the milk line.

1.4.1.3 Formation of placodes

By 11.5 dpc, the mammary lines have given way to five pairs of mammary placodes that appear at specific positions along the ventral side of the embryos (Figure 1.10B). Placodes are visible as lens-shaped structures and consist of an ectodermal thickening. Even though the placodes appear at highly reproducible positions in the embryos, they are formed asynchronously in a specific order: the third pair is the first to appear, followed by the fourth pair, and then the first and fifth pairs, which appear simultaneously, and eventually the second pair appears. Using knockout mice, several genes were shown to be implicated in the formation of the mammary placodes as deletion of these genes resulted in mice missing all or some of the mammary placodes. As discussed above, loss of *Fgf10*/*Fgfr2* and *Tbx3* resulted in absence of certain mammary placodes (Eblaghie et al., 2004; Mailleux et al., 2002). Lymphoid enhancing factor 1 (*Lef1*) is a nuclear target of the canonical Wnt signalling pathway where it interacts with β-catenin (Behrens et al., 1996). It was found that *Lef1*-null mice form primitive mammary placodes that eventually degenerate, although the fourth pair of placodes is sometimes retained (van Genderen et al., 1994). This highlights the importance of *Lef1* as a survival factor in the mammary anlage. In the scaramanga (*ska*) mutant mouse, aberrant

bud formation was observed. Interestingly, in addition to the absence of the third pair of mammary buds, supernumerary buds were observed at the location normally occupied by the fourth pair of mammary buds (Howard et al., 2005). The '*ska'* gene was later identified as the neuregulin 3 (*Nrg3*) gene, which encodes the ligand for the *Erbb4* (*Her4*) tyrosine kinase receptor. It has been proposed that Nrg3 could function to transmit signals from Fgf10 and/or Tbx3 to the precursor mammary epithelial cells which appear to express Wnt10b (Veltmaat et al., 2004).

1.4.1.4 Bud formation

From 11.5-12.5 dpc, several morphogenic processes in the embryonic mammary gland begin to occur. Firstly, the mammary placodes began to invaginate into the underlying mesenchymal layer and develop into mammary buds. Secondly, the mesenchymal cells surrounding the mammary bud condense and differentiate into the dense mammary mesenchyme that surrounds the mammary bud. In addition, the mammary mesenchyme up-regulates the expression of androgen receptors which are important for the regression of male mammary buds in response to androgens (Kratochwil and Schwartz, 1976). As mentioned above, *Lef1* mutant embryos have primitive placodes but they eventually regress by 12.5 dpc, suggesting that Wnt signalling continues to play an important role in mammary bud development. Studies of TOP-GAL reporter mice showed that Wnt signalling is detectable in the mammary bud right up to 15.5 dpc: *Lef1* is expressed in the mammary placodes and bud until 15.5 dpc and in the mammary mesenchyme at 14.5 dpc. The homeodomain-containing transcription factors, *Msx1* and *Msx2* are both expressed in the mammary buds with the latter also expressed in the mammary mesenchyme (Satokata et al., 2000). *Msx* genes appear to have necessary but redundant functions in mammary development: knockout of a single *Msx* gene did not affect mammary bud formation (Satokata et al., 2000) but knocking-out both genes simultaneously disrupted mammary bud formation (Satokata et al., 2000). Recently, the hedgehog (Hh) signalling pathway has also been implicated in mammary bud formation (Hatsell and Cowin, 2006). *Gli3*, which encodes a microtubule-bound transcription factor, is an effector of the Hh signalling pathway (Hooper and Scott, 2005). Depending on its phosphorylation status, it can either become a transcription

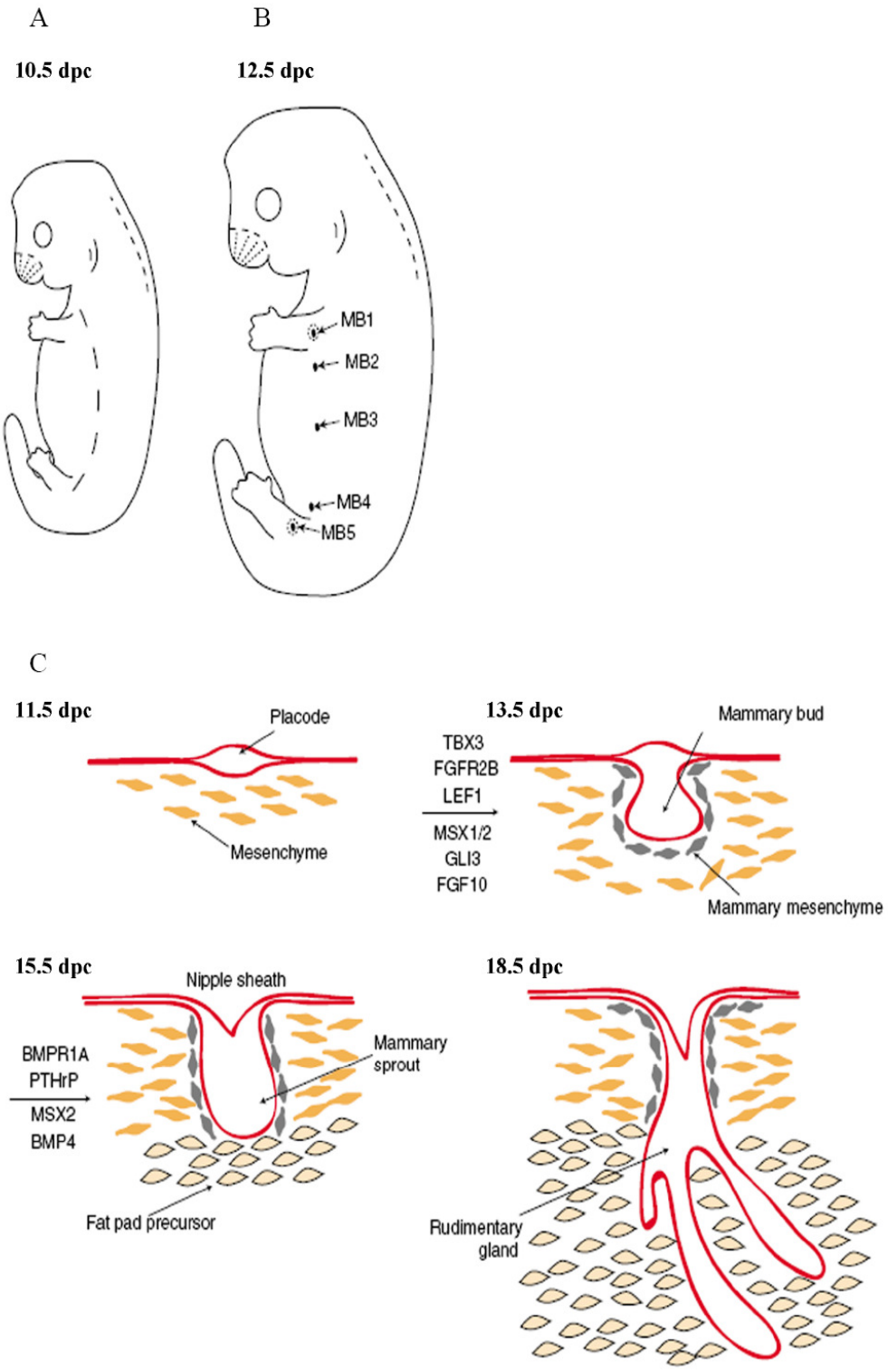
activator (Gli^A; phosphorylated form) or a transcription repressor (Gli^R; cleaved form) to mediate the signals from interactions between the secreted ligands (sonic hedgehog, Indian hedgehog or desert hedgehog) and their receptors (Patched1 and Patched2) (Hooper and Scott, 2005). The Gli^A/Gli^R ratio of Gli3 was found to be required to repress Hh target genes that are involved in patterning and bud formation for the third and fifth pairs of mammary buds (Hatsell and Cowin, 2006), suggesting that different inductive signals are required for different mammary buds. Another transcription factor, *Gata-3*, has also been shown to be essential for mammary bud formation. Expression of *Gata-3* was detected in the mammary buds from 12.5 dpc and conditional deletion of *Gata-3* using Cre recombinase driven by the keratin 14 (K14) promoter resulted in variable loss of mammary placodes and a failure to form the nipple sheath (Asselin-Labat et al., 2007).

1.4.1.5 Sprout formation and early ductal morphogenesis

Between 13.5-15.5 dpc, the mammary bud begins to elongate and penetrate into the underlying fat pad to form a sprout-like structure and nipple formation is initiated (Figure 1.10C). Epithelial-mesenchymal signalling through parathyroid hormone related peptide (PTHrP) and its receptor (PTHrPR) plays a critical role in mammary sprout and nipple sheath formation (Foley et al., 2001). PTHrP is expressed and secreted from the mammary epithelial bud which then binds to PTHrP receptors found in the mammary mesenchyme (Figure 1.10D) (Foley et al., 2001). This up-regulates the expression of bone morphogenic protein receptor 1A (*Bmpr1A*) in the mammary mesenchyme which makes the mesenchyme responsive to bone morphogenic protein 4 (Bmp4) signalling (Figure 1.10D). Binding of Bmp4 to Bmpr1A stimulates epithelial bud elongation and also up-regulates the expression of *Msx2*. This in turn inhibits the formation of hair follicles within the nipple sheath and permits the development of the nipple sheath (Foley et al., 2001). Thus, PTHrP is the first mammary epithelium-specific secreted signalling molecule that has been found to influence cell fate decisions in the surrounding mesenchyme to-date. From 15.5 dpc, the elongation and arborisation of the sprout continues to be stimulated by the interactions between PTHrP and Bmp4 signalling pathways. At birth, the mammary sprout has developed into a small rudimentary

mammary network with a primary duct and about 10-15 ductules. Mammary gland development is then arrested and does not recommence until puberty.

In summary, from the appearance of the milk line at 10.5 dpc to the formation of the rudimentary mammary tree structure at birth, embryonic mammary gland development is a highly co-ordinated developmental process that requires the interplay between several key signalling pathways such as the Wnt, Fgf, Bmp and PTHrP pathways. Studies have also shown that the interaction between the developing mammary bud and the surrounding mammary mesenchyme is also critical in ensuring proper mammary development. A more detailed functional analysis of genes that are expressed in the mammary buds and/or mammary mesenchyme is being studied in mouse knockouts and this would undoubtedly add new players to the mammary developmental process. Therefore, one major challenge would be to integrate these new genes into the existing signalling pathways in order to obtain a better understanding of the initiation, specification and differentiation process of formation of the embryonic mammary gland.


A
B
10.5 dpc
12.5 dpc
MB1
MB2
MB3
MB4
MB5
C
11.5 dpc
Placode
Mesenchyme
13.5 dpc
TBX3
FGFR2B
LEF1
MSX1/2
GLI3
FGF10
Mammary bud
Mammary mesenchyme
15.5 dpc
Nipple sheath
BMPR1A
PTHrP
MSX2
BMP4
Mammary sprout
Fat pad precursor
18.5 dpc
Rudimentary gland

D

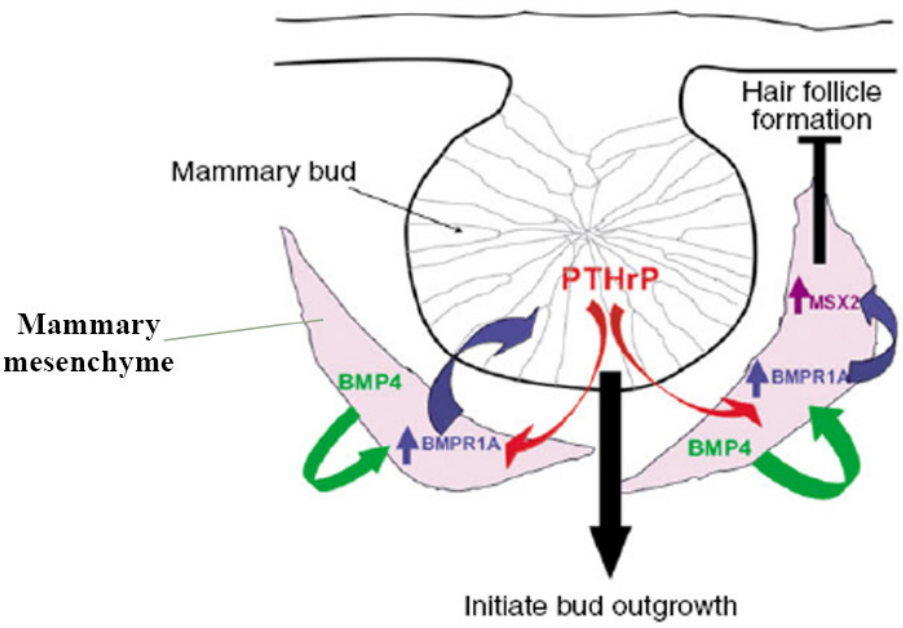


Figure 1.10. Overview of murine embryonic mammary gland development. (A) Diagrammatic view of the presumptive locations of the mammary milk lines which appear in an anteroposterior axis at around 10.5 dpc. **(B)** At 12.5 dpc, five pairs of mammary placodes which have become mammary buds (MB1-5) are visible at specific positions along the anteroposterior ventral side of the embryo. MB1 and MB5 are hidden behind the fore- and hind-limb respectively. **(C)** Overview of the murine embryonic development process from the milk line. At 11.5 dpc, the milk line has developed into an ectodermal thickening known as the placode. By 13.5 dpc, the placodes have invaginated into the underlying mesenchymal layer (orange) to form mammary buds. The mesenchymal cells (orange) that enclose the mammary buds begin to condense to form the mammary mesenchyme (grey). By 15.5 dpc, these buds have developed into sprouts which form a lumen with an opening to the skin, characterized by the formation of the nipple sheath. By 18.5 dpc, the mammary gland is present as a rudimentary arborized structure which has several small ducts that have invaded the mammary fat pad (buff). **(D)** Schematic cross-section of a mammary bud between 13.5-15.5 dpc. Within these time points, mammary buds begin to elongate and invaginate into the underlying fat pad stimulated by the reciprocal epithelial-mesenchymal signalling between PTHrP and its receptor. Binding of the PTHrP (secreted from the mammary bud) to its receptor (expressed on mammary mesenchyme) up-regulates expression of BMPR1A in the mammary mesenchyme which is now responsive to BMP4 signalling. This stimulates epithelial outgrowth, elevates *MSX2* expression, and inhibits hair follicle formation within the nipple sheath. Images obtained from (Watson and Khaled, 2008).

1.4.2 Postnatal morphogenesis

Unlike other organs, development of the mammary gland is completed during adolescent rather than during embryonic development. Mammary ductal elongation and branching develops predominantly after puberty, whereas alveolar proliferation occurs during pregnancy and functional differentiation of these alveolar cells is completed with parturition and lactation. Prior to the onset of puberty (which occurs at about four weeks after birth), the mammary gland enters an allometric growth phase where the mammary ducts elongate into the fat pad at a similar rate to the overall growth of the animal (Figure 1.11A). Accelerated ductal elongation and side branching occurs at the start of puberty, about four weeks after birth. Terminal end buds (TEBs), which are club-shaped-like structures found at the leading edges of the mammary ducts, are prominent throughout this stage of development; they lead to the elongation and penetration of the ducts through the fat pad. The TEB is a specialized structure that consists of two histologically distinct cell types: (1) cap cells, which are precursors of myoepithelial cells and are thought to contain mammary stem/progenitor cells, and (2) body cells which give rise to mammary epithelial cells (Figure 1.11A) (Humphreys et al., 1996). The TEB plays an important role in the generation of the mature mammary duct. A highly regulated process of proliferation and apoptosis in the TEB generates a hollow lumen surrounded by a single layer of luminal epithelial cells that serves as a channel for the transport of milk during lactation and the surrounding myoepithelial layers, which serve to provide contractile forces to facilitate transport of milk to the nipple (Humphreys et al., 1996). Additionally, ductal arborization is initiated from the highly proliferative TEBs which form secondary and tertiary branches (Daniel et al., 1987). At the end of puberty, the TEBs have disappeared and the entire fat pad is now invaginated by a network of mammary ducts (Figure 1.11B).

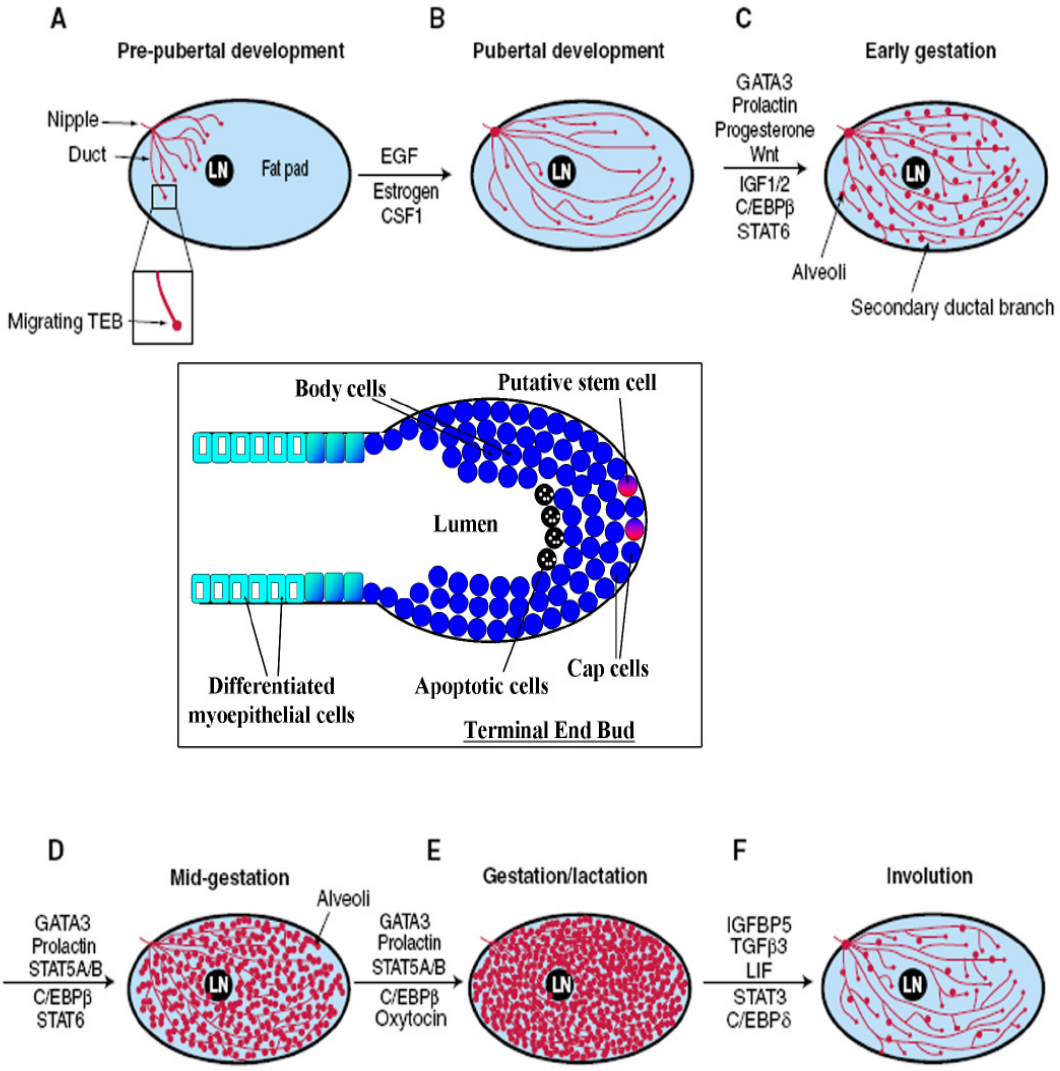


Figure 1.11. Overview of murine adult mammary development. Schematic diagrams depicting the major phases of adult mammary gland development: **(A)** Pre-pubertal, **(B)** Pubertal, **(C)** Early gestation, **(D)** Mid-gestation, **(E)** Gestation/Lactation and **(F)** Involution. Some of the important factors that are involved in the various phases are highlighted. Insert shows a diagrammatic representation of the cap and body cell layers of TEB. LN: lymph node; TEB: terminal end bud.

1.4.2.1 Hormonal regulation

From the onset of puberty, functional development of the mammary gland is essentially defined by the hormonal status of the animal. The combinatorial effects of systemic hormones such as growth hormone (GH), estrogen and progesterone, together with the stroma are major influences in the development of the postnatal mammary gland (Sakakura et al., 1987). Pituitary GH, which is already present before the pubertal up-regulation of estrogen, acts via its receptor on mammary stromal cells to elicit expression of insulin-like growth factor-1 (IGF-1). IGF-1 then stimulates TEB formation and epithelial branching in a paracrine manner (Sternlicht, 2006).

Estrogen plays an important role in branching morphogenesis in the virgin gland. There are two estrogen receptors in the mouse, ERα and ERβ, with ERα playing a more important role in ductal morphogenesis. Development of the murine mammary ductal network is completed by four months, whereas the mammary glands of the ERα knockout females displayed a rudimentary epithelium devoid of any TEBs, demonstrating that estrogen is required for ductal outgrowth (Bocchinfuso and Korach, 1997; Korach, 2000). Interestingly, embryonic tissue recombination studies suggest that only stromal ERα is required, whereas adult tissue transplants indicate that both epithelial and stromal ERα are necessary for ductal morphogenesis (Cunha and Hom, 1996; Mueller et al., 2002). However, this original ERα knockout was not a true null allele as alternative splicing generated a trans-activation protein and a failure of estrogen signalling through the hypothalamic/pituitary axis resulted in reduced levels of prolactin, which complicates analysis (Bocchinfuso and Korach, 1997). Subsequently, a complete ERα knockout mouse was generated and TEBs were completely absent and ducts failed to penetrate the mammary fat pad (Mallepell et al., 2006). Recently, conditional deletion of ERα at various stages of mammary development demonstrated that estrogen is required for both pubertal development and development of alveolar structures during late gestation and lactation (Feng et al., 2007).

Progesterone mediates its effects through two receptors, PR-A and PR-B, which are encoded by a single gene containing two distinct promoters. Deletion of both PR isoforms in the mouse resulted in severe defects in the reproductive tissues (Lydon et al., 1995). Mammary gland development was also affected in PR knockout mice. While

ductal growth was normal, side branching was nearly completely absent. The main effects of progesterone appear to be during pregnancy (See Chapter 1.4.3.1). Collectively, the combined effects of GH, estrogen and progesterone ensure the proper development of the ductal system - setting up a mature mammary network that is primed for further proliferation and differentiation during pregnancy.

1.4.3 Lobulo-alveolar development

The most dramatic changes in the mammary gland occur during pregnancy in preparation for the birth of the pups. From the onset of pregnancy, the mammary epithelial cells, stimulated by prolactin and placental lactogens, proliferate and undergo differentiation to form lobulo-alveolar structures that eventually fill up the entire fat pad (Figure 1.11C-F). The maximal rate of proliferation of the epithelial cells is observed between days six and ten of gestation. Extensive branching morphogenesis occurs from day six of gestation, producing secondary and tertiary side branches from the primary duct; this is followed by proliferation and differentiation of alveolar buds (Figure 1.11C). By mid-gestation, clusters of alveoli are evident throughout the entire fat pad and further differentiation of the alveoli results in the production of milk proteins such as β-casein (Figure 1.11D) (Hennighausen and Robinson, 2005; Watson and Khaled, 2008). By late gestation, the mammary epithelial cells account for about 90% of the cells in the mammary gland, most of which have began to produce other milk proteins such as whey acidic protein (WAP) (Watson and Khaled, 2008).

1.4.3.1 Progesterone signalling

In the PR knockout mouse, ductal growth was normal, however, pregnancy induced side branching and alveolar development were completely absent, suggesting that progesterone is necessary for side branching and alveolar differentiation (Lydon et al., 1995). Reciprocal transplantation experiments showed that the effects of progesterone are mediated via the mammary epithelium. Notably, Wnt4 was shown to be required for tertiary side-branching and is also regulated by progesterone, hence Wnt4 might function downstream of the progesterone receptor (Brisken et al., 2000).

1.4.3.2 Prolactin signalling

Prolactin signalling is essential for the proliferation and functional differentiation of lobulo-alveolar structures during pregnancy (Topper and Freeman, 1980). Prolactin is produced and secreted from the anterior pituitary glands of the mouse and binding of prolactin to its receptor results in receptor dimerization and activation of the Janus kinase 2 (JAK2), Fyn, and the mitogen-activated protein (MAP) kinase *in vitro* (Hennighausen and Robinson, 1998). The *in vivo* effects of prolactin signalling are mediated predominantly through the JAK-Stat (signal transducer and activator of transcription) pathway, which will be discussed in detail below (Chapter 1.4.3.3) (Liu et al., 1997). Deletion of *Prl* (prolactin) in the mouse curtailed ductal arborisation in adult virgins, suggesting that ductal morphogenesis is affected; this could however be due to indirect effects because *Prl*-deficient females do not ovulate (Horseman et al., 1997). However, the effects of *Prl* deletion during pregnancy and lactation cannot be examined because *Prl*-deficient mice are infertile. Using *PrlR* (prolactin receptor) knockout mice, Ormandy *et al.* showed that prolactin signalling is important for mammary gland development (Ormandy et al., 1997). Females with only one copy of *PrlR* allele failed to lactate after their first pregnancy, indicating that levels of prolactin are important for epithelial cell proliferation and differentiation (Ormandy et al., 1997). However, mammary gland development was normal after the second pregnancy or in older females, and dams were able to lactate and nurture their pups. This demonstrates that continued hormonal stimuli were sufficient to ensure complete functional development of the mammary gland. Collectively, these studies demonstrate that prolactin signalling is essential for reproductive functions and mammary gland development in the mouse. This was confirmed by studying various mouse knockouts of different components of the prolactin signalling pathway, in particular, components of the JAK-Stat pathway.

1.4.3.3 JAK/Stat signalling

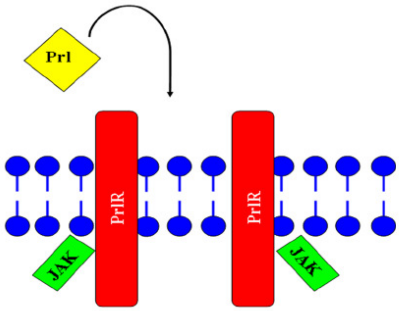
Prolactin signalling in the mammary gland is mediated largely through the JAK-Stat pathway (Liu et al., 1997). Binding of prolactin to its receptor results in receptor dimerization and activation of the receptor-associated kinase (JAK) which then phosphorylates specific tyrosine residues of the receptor (Figure 1.12A and 1.12B).

Subsequently, Stats are recruited to the receptor through their Src-homology-2 (SH2) domains and become phosphorylated on tyrosine 694 by JAK (Figure 1.12C). Phosphorylated Stats form hetero- and/or homo-dimers and translocate into the nucleus where they activate transcription of target genes (Figure 1.12D) (Hennighausen and Robinson, 1998, 2005). There are seven different members of the Stat family (Stat1, Stat2, Stat3, Stat4, Stat5a, Stat5b and Stat6) and four known JAK proteins (JAK1, JAK2, JAK3 and JAK4) (Darnell, 1997). Most Stats bind to and induce transcription of genes containing γ-interferon activation sites (GAS), $TTCn_3GAA$ with only Stat6 binding to a unique consensus sequence $TTCn_4GAA$ (Darnell, 1997). Stat proteins have six conserved domains with the DNA binding domain in the centre and the SH2 domain at the C-terminus of the protein. Despite the common consensus binding sequence, Stats proteins achieve specificity in their activation of target genes by the different combinations of co-activators and co-repressors that are associated with Stats.

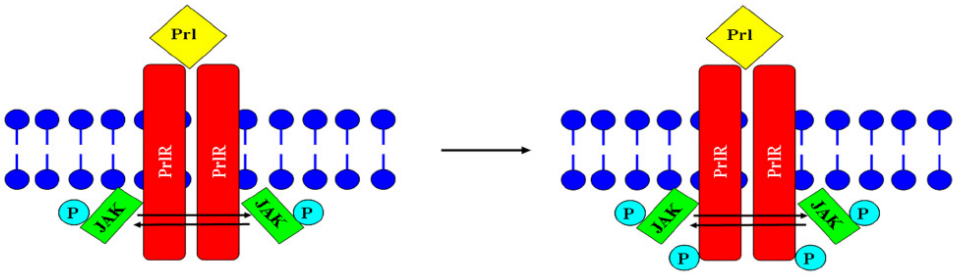
Two major mechanisms function to prevent precocious or sustained JAK-Stat signalling during pregnancy: (1) suppressor of cytokine signalling (SOCS) family of proteins and (2) caveolin-1 (Cav1). There are eight members of the SOCS family [SOCS1-7 and cytokine-inducible SH2 protein (CIS)] that interact with JAKs (Figure 1.12E1) and cytokine receptors (Figure 1.12E2) to control the activation of Stat proteins (Kubo et al., 2003). *In vitro* studies showed that SOCS proteins bind to tyrosine residues in cytokine receptors and block the binding and hence the activation of Stat proteins. SOCS1 has been shown to be a critical negative regulator of alveologenesis (Lindeman et al., 2001). Loss of *Socs1* in the mouse resulted in excessive alveologenesis during gestation and elevated Stat5 activity during lactation, suggesting that SOCS1 normally inhibits prolactin signalling. To demonstrate the direct modulation of prolactin signalling by SOCS1, Lindeman *et al.* showed that an equivalent loss of one *Socs1* allele in *PrlR* hemizygous mice rescued alveologenesis and Stat5 activity (Lindeman et al., 2001). Similarly, homozygous null mutants of *Socs2* rescued the alveologenesis defects in *PrlR* hemizygous mice, suggesting that SOCS2 is also a negative regulator of prolactin signalling (Harris et al., 2006). SOCS3 proteins function in a negative-feedback-loop that involves Stat5 and the gp130 receptor that is the shared subunit of cytokine receptors for interleukin 6 (IL6) and leukaemia inhibitory factor (Lif). Activation of gp130 signalling

occurs during involution which leads to activation of Stat3 (Humphreys et al., 2002). Expression of *Socs3* is up-regulated by Stat5 during pregnancy, resulting in the binding of SOCS3 to the tyrosine residues of gp130 to prevent the activation of Stat3 (Humphreys et al., 2002). *Socs5* and *Cav1* null mutants showed accelerated alveolar development and dramatic activation of Stat5 leading to precocious milk production in the pregnant gland (Khaled et al., 2007). Hence regulation of the activity of SOCS and Cav-1 proteins is also critical to the maintenance and duration of the JAK-Stat signalling pathway. Other negative regulators of JAK-Stat pathways in the mammary gland include SH2 containing phosphatases (SHP) and protein inhibitor of activated Stat (PIAS) family (Wormald and Hilton, 2004).

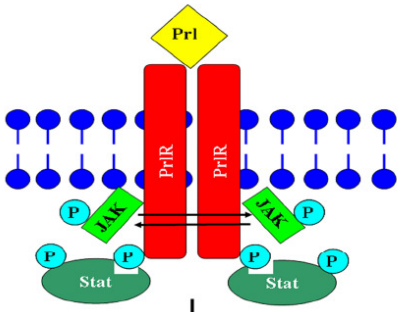
A


Prl
PrlR
PrlR
JAK
JAK

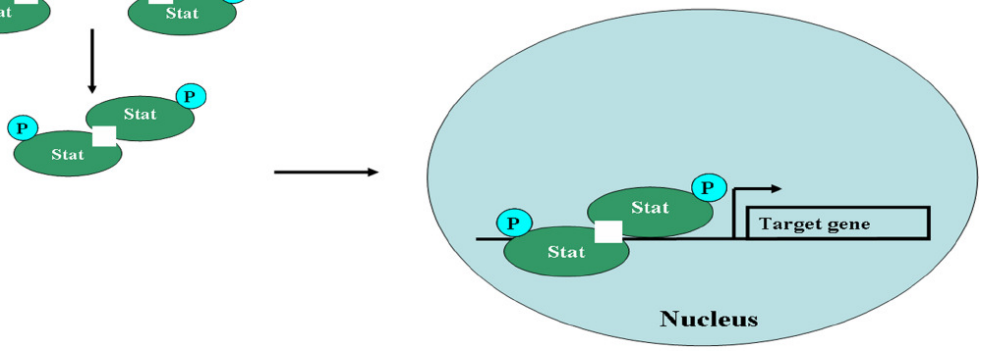
B


Prl
PrlR
PrlR
JAK
JAK
P
Prl
PrlR
PrlR
JAK
JAK
P

C


Prl
PrlR
PrlR
JAK
JAK
P
Stat
Stat

D


Stat
Stat
Stat
Stat
Target gene
Nucleus

E

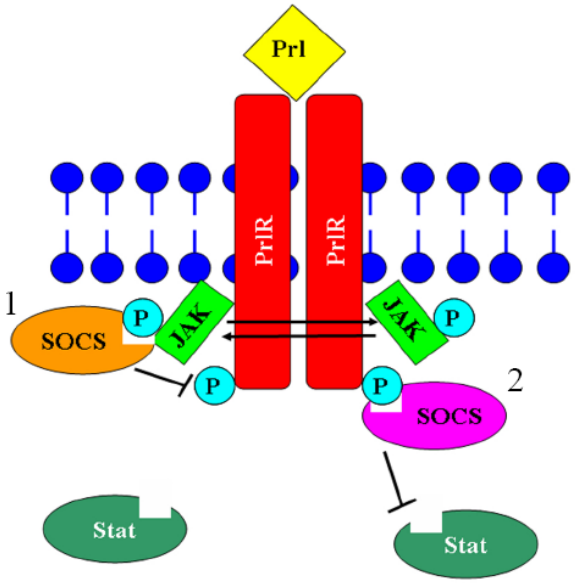


Figure 1.12. Stat signalling pathway. (A) PrlR with the associated tyrosine kinase JAK remains inactive in the absence of ligand (Prl). **(B)** Binding of Prl results in dimerization of the receptor and activation of JAK which in turn phosphorylates specific tyrosine residues on the receptor, creating docking sites for Stats. **(C)** Stat is recruited to the phosphorylated PrlR through its SH2 domains, where it becomes phosphorylated. **(D)** Phosphorylated Stat dimerizes and translocates into the nucleus where it activates the transcription of target genes. **(E)** SOCS proteins are induced in response to cytokine signalling. Different SOCS proteins negatively regulate Stat signalling via different methods: (1) SOCS can bind to phosphorylated JAKs to prevent activation of the receptor and subsequently phosphorylation of Stats; (2) SOCS can also bind directly to Stat binding sites on the activated receptor, preventing docking and hence activation of Stat proteins. Prl: Prolactin; PrlR: Prolactin receptor; JAK: Janus kinase; Stat: Signal transducer and activator of transcription; SOCS: Suppressor of cytokine signalling; P: phosphorylated residue.

1.4.3.4 Stat5 as a key mediator of lobulo-alveolar development

The effects of prolactin signalling during pregnancy are predominantly mediated via Stat5 (Hennighausen and Robinson, 2008). Stat5a and Stat5b are encoded by two juxtaposed genes and are highly similar; they show a 96% similarity at the amino acid level (Hennighausen and Robinson, 2008). Differential expression patterns of both genes were observed in the mouse: *Stat5a* is the predominantly expressed in mammary tissue while *Stat5b* is more abundant in the muscle and liver (Hennighausen and Robinson, 2008). Consistent with their expression patterns, knockout of *Stat5a* in mice resulted in the attenuation of alveolar development and failure to lactate (Liu et al., 1997) while knockout of *Stat5b* in mice did not affect mammary development but led to stunted body growth (Udy et al., 1997). These tissue-specific defects reflect either the tissue-specific expression of *Stat5* genes or the tissue-specific function of these genes. After multiple pregnancies, some functional mammary development was observed in *Stat5a* knockout mice, coinciding with an increase in the levels of Stat5b (Liu et al., 1998), which suggests that there might be partial compensation from *Stat5b* in the mutant gland. Deletion of *Stat5a* and *Stat5b* simultaneously in the mouse confirmed functional redundancy of both isoforms in the mammary gland (Miyoshi et al., 2001; Teglund et al., 1998). As the original *Stat5a/b* double knockout females were infertile due to non-functional corpora lutea, *Stat5a/b* conditional knockout mice were used to analyse the mammary phenotypes during pregnancy (Cui et al., 2004). Transplantation of *Stat5a/b* null epithelial cells into a cleared mammary fat pad showed that while ductal growth was unaffected, cell proliferation and alveologenesis during gestation were severely affected (Miyoshi et al., 2001). In addition, conditional deletion of *Stat5a/b* during late pregnancy led to premature cell death of the alveoli, suggesting a role for Stat5a/b in cell survival (Cui et al., 2004). Hence, *Stat5a/b* is a crucial mediator of proliferation and differentiation of mammary luminal cells and also a regulator of cell survival and function during lactation.

1.4.4 Lactation

At the onset of lactation, the mammary gland reaches maturity and produces and secretes milk to feed the newborn pups. Proliferation and expansion of the alveolar structures occurs at parturition and the entire mammy gland becomes populated by

alveoli with large lumens (Figure 1.11E). The mature alveoli acquire a secretory function and the luminal secretory cells are surrounded by basal myoepithelial cells, which contract and force the milk into the ducts that transport the milk to the nipple. Activation of Stat5 by prolactin signalling is required for β-casein production (Kazansky et al., 1995). Direct binding of activated Stat5 and recruitment of CCAAT/enhancer-binding protein (C/EBPβ) (Raught et al., 1995) to the β-casein promoter up-regulates β-casein transcript levels during pregnancy and lactation. In addition to milk protein production, the ability to expel the secreted milk into the mammary ducts is also critical to successful lactation. The systemic hormone, oxytocin (OT) plays an important role in this process by inducing the contraction of the myoepithelial layer surrounding the alveoli and ducts to facilitate delivery of the milk to the young (Hennighausen and Robinson, 1998). Mice lacking OT failed to lactate and undergo precocious involution most likely due to an accumulation of milk in the alveoli (Wagner et al., 1997). Therefore, successful lactation is dependent on the terminal differentiation of luminal epithelial cells, which produce and secrete milk, and on the contractile functionality of the myoepithelial layers that enables the milk to be transported to the nipple to nurse the young.

1.4.5 Involution

At the end of the lactation phase following the weaning of the pups, the milk-producing epithelial cells are no longer required and are removed (Figure 1.11F). This post-lactational mammary gland regression is known as involution and is mediated by apoptosis (Watson, 2006b). The main features of involution include removal of most of the secretory epithelial cells by apoptosis, clearing of the dead cells and milk, and re-differentiation of the adipocytes such that the mammary gland is remodelled to a pre-pregnant state in preparation for subsequent pregnancies. Failure to remove redundant lobulo-alveoli during involution could result in inflammation and tissue damage, therefore regulation of involution is an extremely important process (Watson, 2006a). Detailed analysis of the involution process has allowed the identification of morphological features and molecular events that occur in the first six days of involution. Studies from forced weaning (removal of pups from lactating females at day 10 lactation) and glucocorticoid administration in mice revealed two distinct phases of involution

(Watson, 2006a). In the first phase (first 48 hours of involution), no remodelling of the mammary gland is observed and only apoptosis is initiated, resulting in the detection of apoptotic cells in the alveolar lumens. The second phase of involution (72 hours after initiation of involution) is characterized by apoptosis together with remodelling of the mammary gland where the alveoli collapse and adipocytes re-differentiate and begin to refill the gland (Watson, 2006b). If pups are returned to the involuting female within 48 hours of initiation of involution, apoptosis is curtailed and lactation is re-initiated (Watson, 2006a). The first phase of involution is regulated by local factors such as milk accumulation within each gland and is not dependent on circulating hormones (Li et al., 1997; Marti et al., 1997). In contrast, the second phase of involution is dependent on circulating hormones and can be blocked by the administration of progesterone (Feng et al., 1995; Lund et al., 1996). In addition, the second phase is also dependent on the activity of matrix metalloproteinases whose function is inhibited by the expression of tissue inhibitors of metalloproteinases (Watson, 2006a). Immuno-infiltration of the mammary gland is also characteristic of the second phase where phagocytes clear the gland of cellular debris and milk components.

1.4.5.1 Lif/Stat3-mediated apoptosis during involution

Apoptosis mediated by the activation of the LIF/Stat3 pathway plays a critical role in the involution process (Watson, 2006a; Watson, 2006b). Apoptosis is a form of programmed cell death in multicellular organisms that was first coined by Kerr, Wyllie and Currie in 1972 (Kerr et al., 1972). Apoptosis has been extensively studied and the morphological changes that occur within an apoptotic cell are well-defined. The cells undergo cellular shrinking, membrane blebbing, chromatin condensation, nuclear fragmentation and chromosomal DNA fragmentation (Kerr et al., 1972). Caspases, which are cysteine proteases, are responsible for most of the morphological changes described in the apoptotic cells (Kerr et al., 1972). Several signalling pathways and apoptotic regulators have been implicated in the first phase of involution in the mammary gland such as members of the Bcl-2 family, Bcl-x and Bax. Conditional deletion of the anti-apoptotic *Bcl-x* in the mouse resulted in accelerated apoptosis (Walton et al., 2001), whereas inactivation of the pro-anti-apoptotic *Bax* delayed involution (Schorr et al.,

1999). However, the main signalling pathway that is involved in the first phase of involution is the JAK-Stat pathway, in particular, Stat3, which is the main player in the involution process (Chapman et al., 2000).

Stat3 is a member of the Stat family and activation of Stat3 occurs at the start of the first phase of involution (Liu et al., 1996). Conditional deletion of *Stat3* in the mammary glands of mice led to a dramatic reduction in epithelial apoptosis, resulting in delayed mammary gland involution (Chapman et al., 1999). This suggests that activation of Stat3 is a critical event in the initiation of apoptosis and involution. Subsequently, both the upstream and downstream targets of the Stat3 signalling pathway were identified (Kritikou et al., 2003; Thangaraju et al., 2005; Tonner et al., 2002). Leukemia inhibitory factor (Lif), a 40 kDa glycoprotein belonging to the IL6 family of cytokines, binds to the Lif receptor β (LifRβ) which forms a complex with gp130 to activate Stat3 *in vivo* (Figure 1.13A) (Kritikou et al., 2003). *Lif* is up-regulated dramatically within 12 hours of initiation of involution and deletion of *Lif* in the mouse resulted in the absence of Stat3 activation, decreased apoptosis and delayed involution (Kritikou et al., 2003). Interestingly in *Lif* knockout mice, there was precocious lobulo-alveolar development in the pregnant gland (when activation of Stat3 is normally undetected) (Kritikou et al., 2003). Activation of extracellular signal-regulated kinase 1/2 (ERK1/2) was dramatically reduced in the pregnant glands of *Lif* knockout mice, suggesting that Lif could act via the ERK pathway during lobulo-alveolar development (Kritikou et al., 2003). In addition to Lif, transforming growth factor (TGF) β3 has also been shown to be capable of activating Stat3 in the first phase of involution (Nguyen and Pollard, 2000).

Downstream targets of Stat3 have also been identified and they include insulin-like-growth-factor-binding protein 5 (IGFBP-5), C/EBPδ and regulatory subunits of the phosphatidylinositol-3-OH kinase (PI3K) pathway. Over-expression of *Igfbp-5* led to impaired lobulo-alveolar development in pregnant mice, an increase in expression of cleaved Caspase 3 (pro-apoptotic) and a decrease in expression of *Bcl-2* and *Bcl-xL* (Tonner et al., 2002). In contrast, deletion of *C/EBPδ* delayed involution due to the inactivation of pro-apoptotic genes and continued expression of anti-apoptotic genes (Thangaraju et al., 2005). Activation of Stat3 also up-regulates the expression of PI3K regulatory subunits *p55α* and *p50α* (Figure 1.13B) (Abell et al., 2005). PI3K is a family

of enzymes that catalyze the production of 3' phosphoinositide lipids in the cell membrane upon their recruitment in response to activation of growth factor receptors (Figure 1.13C) (Katso et al., 2001). Class I PI3Ks catalyze the phosphorylation of phosphatidylinositol-4,5-biphosphate (PtdIns-4,5-P_2) to phosphatidylinositol-3,4,5-triphosphate (PtdIns-3,4,5-P_3), which in turn acts as a docking site for adaptor proteins that contain the pleckstrin homology (PH) domain such as the serine-threonine kinase, protein kinase B (PKB)/Akt (Figure 1.13D) and phosphoinositide-dependent kinase 1 (PDK1) (Katso et al., 2001). After recruitment to the membrane, Akt is phosphorylated and consequently activated by PDK1 and is now able to mediate phosphorylation of target proteins on serine and threonine residues. Through phosphorylation of these target proteins, Akt functions as a key regulator of cellular proliferation and survival (Katso et al., 2001). PI3K consists of a p110 (α/β/δ isoforms) catalytic subunit associated with either of the p85 (α/β isoforms), p55 (α/γ isoforms) and p50α regulatory subunits (Funaki et al., 2000). Both p55α and p50α are alternative splice variants of the p85α subunit. Association of p110 with p85 isoforms render the enzyme active while p55α and p50α are negative regulators of p110 catalytic activity (Funaki et al., 2000). The levels of *p85α* remained constant throughout mammary gland development while expression of *p55α* and *p50α* were up-regulated at the onset of involution (Figure 1.13B) (Abell et al., 2005). Activation of Stat3 directly up-regulates the transcription of *p55α* and *p50α* via an internal Stat3 binding site in *p85α* (Abell et al., 2005). The increased levels of p55α and p50α correlate with a decrease in Akt phosphorylation during involution (Abell et al., 2005). Consistent with this, in the *Stat3* mutant mammary glands, levels of p55α and p50α were reduced, thus demonstrating that the Stat3-mediated apoptosis is regulated by a subunit switch in the PI3K enzyme (Abell et al., 2005).

In summary, mammary gland development consists of a tightly co-ordinated series of events. Embryonic development sets up the basis of a rudimentary mammary structure and at puberty, ductal morphogenesis stimulated by steroid hormones, generates the entire mature mammary network. Terminal differentiation of the mammary epithelium occurs during pregnancy and lactation. During involution, the redundant lobulo-alveolar structures regress by apoptosis to return the mammary epithelium to a pre-pregnant state, in preparation for the next pregnancy.

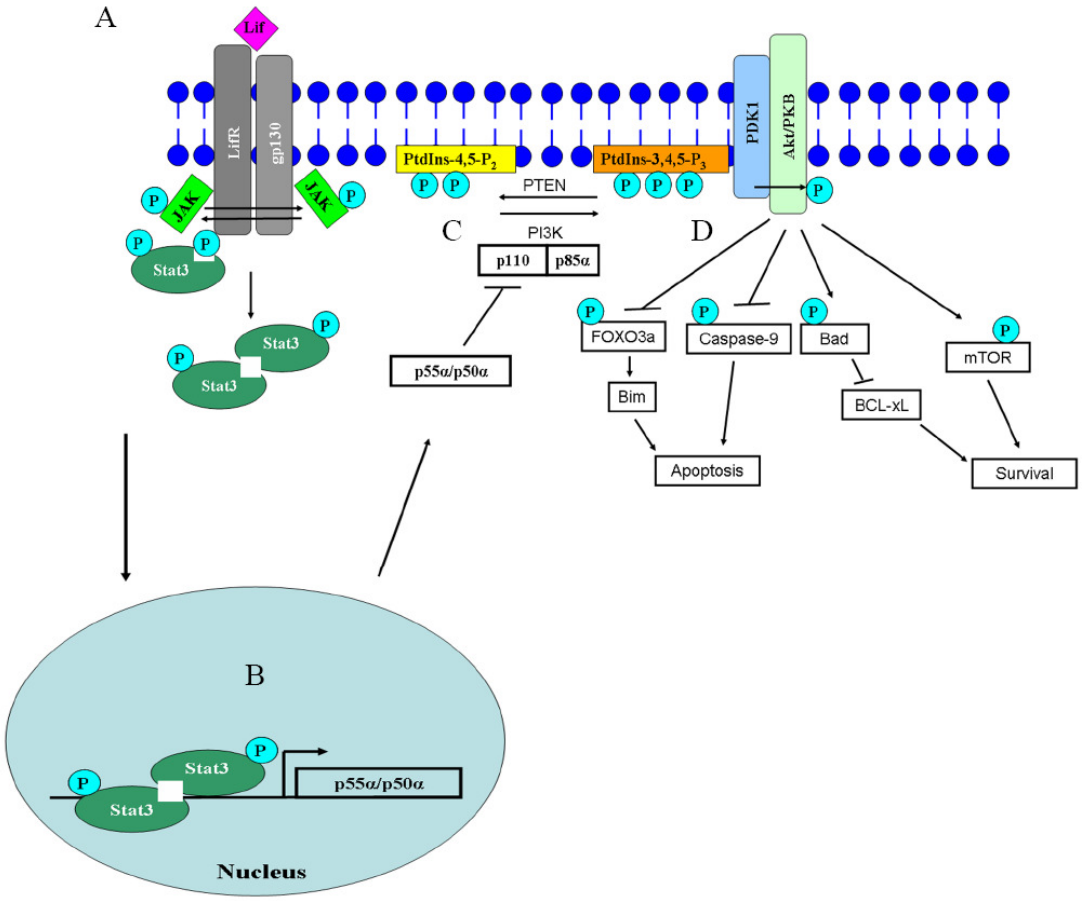


Figure 1.13. Overview of Lif/Stat3-mediated apoptosis. (A) Binding of Lif to its receptor and gp130 results in the activation of JAK which in turn phosphorylates specific tyrosine residues on the LifR, creating docking sites for Stat3. Recruitment of Stat3 to the receptor results in the phosphorylation and activation of Stat3. (B) Phosphorylated Stat3 dimerizes and translocates into the nucleus where it activates the transcription of target genes such as *p55α* and *p50α*. The p55α and p50α proteins are negative regulators of PI3K. (C) PI3K catalyses the phosphorylation of PtdIns-4,5-P_2 to PtdIns-3,4,5-P_3. The reverse reaction is catalysed by PTEN. (D) PtdIns-3,4,5-P_3 creates docking sites for the recruitment of PKB/Akt and PDK1. Akt becomes phosphorylated by PDK1 and is now able to activate/inhibit a plethora of targets such as FOXO3a transcription factor, Caspase-9, pro-apoptotic protein Bad and mTOR, mediating effects of survival. Lif: Leukaemia inhibitory factor; LifR: Lif receptor; JAK: Janus kinase; Stat3: Signal transducer and activator of transcription 3; PtdIns-4,5-P_2: phosphatidylinositol-3,4-biphosphate; PtdIns-3,4,5-P_3: phosphatidylinositol-3,4,5-triphosphate; PDK1: phosphoinositide-dependent kinase 1; P: phosphorylated residue.

1.5 Genetic control of mammary cell fate

1.5.1 Historical perspectives

The mammary epithelium primarily consists of two main cell types: (1) luminal epithelial cells which line the lumen and (2) myoepithelial/basal epithelial cells which form a surrounding sleeve encompassing the ductal epithelium. Similar to the hematopoietic system, where the various different blood lineages are generated from the HSC, the entire mammary epithelium is also derived from the mammary stem cell (MaSC). Likewise, different signals are required for the specification and cell fate determination of mammary stem and progenitor cells during lineage commitment throughout development. In addition to steroid hormones, transcription factors and cytokines also play important roles in mammary cell fate and lineage commitment (Watson and Khaled, 2008).

The first evidence of the possible existence of adult mammary stem cells was revealed in 1959 when DeOme *et al*. showed that virtually any part of the adult mammary gland could reconstitute the cleared mammary fat pad to generate the entire mammary ductal tree (Deome et al., 1959). Importantly, the cleared fat pad transplantation technique developed in this study formed the basis of an *in vivo* functional test for MaSCs, similar to the transplantation of hematopoietic cells into sub-lethally irradiated host mice. Subsequently, Smith and Medina demonstrated that mammary cells capable of repopulating the mammary gland exist throughout the life span of the adult mice (Smith and Medina, 1988). In addition, using retroviral-tagged mammary epithelial cells, Kordon and Smith revealed the existence of MaSCs and also the presence of three multi-potent but distinct mammary epithelial progenitors (Kordon and Smith, 1998), similar to that described in an earlier study (Smith, 1996). In general, the field of MaSC biology has lagged behind that of HSC research because of a lack of definitive and exclusive stem/progenitor cell markers. However, a breakthrough was achieved in 2006 when two groups demonstrated that a single cell isolated by a combination of cell surface marker was able to reconstitute the entire mammary gland (Shackleton et al., 2006; Stingl et al., 2006). This is the first direct evidence of the existence of the MaSC.

1.5.2 Mammary epithelial hierarchy

1.5.2.1 Characteristics of the mammary stem cell

The isolation of MaSCs was based on expression of the cell surface marker CD24 (heat-stable antigen) and either CD49f (α_6-integrin) (Stingl et al., 2006) or CD29 (β_1-integrin) (Shackleton et al., 2006). Interestingly, CD49f and CD29 are likely to form a functional α_6-β_1-integrin hetero-dimeric complex that mediates the interaction between the epithelial cells and stroma. The subset of lineage-depleted mammary epithelial cells (depletion of hematopoietic and endothelial cells) which expresses medium levels of CD24 and high levels of CD49f/CD29 ($CD24^{med}CD49f^{hi}/CD29^{hi}$) (Figure 1.14A) were found to be enriched in MaSCs (termed mammary repopulation units - MRUs) as determined by transplantation of these cells into cleared fat pads. Stingl *et al.* estimated the frequency of these MaSCs to be 1 per 20 cells of the $CD24^{med}CD49f^{hi}$ subset (Stingl et al., 2006) while Shackleton *et al.* estimated a frequency of 1 per 64 cells of the $CD24^{med}CD29^{hi}$ subset (Shackleton et al., 2006). Consequently, a single cell derived from the MaSC-subset ($CD24^{+}CD49f^{hi}/CD29^{hi}$) was shown to reconstitute the entire mammary gland without the use of supporting cells (Stingl et al., 2006) (Shackleton et al., 2006). A functional role for CD29 in the maintenance of MaSC in the epithelium has been demonstrated recently (Taddei et al., 2008). Deletion of CD29 from the basal compartment of the epithelium abolished the regenerative potential of MaSC, suggesting a critical role of CD29 in the maintenance of MaSC population by mediating basal cell interactions with the extracellular matrix.

The MaSC-enriched subset also expresses low levels of stem cell antigen (Sca1) and appears to occupy a basal position within the mammary epithelium (Stingl et al., 2006) (Shackleton et al., 2006). In addition, the MaSC-enriched subset was also found not to exclude the $Hoechst_{33342}$ dye (Shackleton et al., 2006; Stingl et al., 2006). This is somewhat surprising because the ability to exclude dyes is a characteristic of stem cells. Previous studies have also identified a side population phenotype that contains undifferentiated cells which can give rise to luminal and basal cells (Alvi et al., 2003; Welm et al., 2002). However, this side-population is most likely to represent bi-potent progenitor cells because it is depleted of expression of $CD24^{med}CD49f^{hi}/CD29^{hi}$ and transplantation of these cells at limiting dilutions showed a low repopulating frequency

(Shackleton et al., 2006; Stingl et al., 2006). Interestingly, the MaSC subset appears to be actively cycling as determined by $Hoechst_{33342}$ and pyronin Y staining (Stingl et al., 2006) and this is consistent with another study by Smith *et al.* (Smith, 2005), which showed that *in vivo* long-term label-retaining mammary cells appeared to retain their template DNA strands through asymmetrical cell division. However, the precise relationship between MaSCs and long-term label-retaining mammary cells is unknown. It has been postulated that a minor population of MaSCs that resides in a quiescent state (G0) *in vivo* may be present (Visvader and Lindeman, 2006).

Estrogen plays an important role in ductal morphogenesis and also in alveologenesis as discussed above (Chapter 1.4.2.1). ERα is an important prognostic marker for breast cancer and ERα-negative breast cancers are those that are resistant to therapy and usually have a bad prognosis. Expression of ERα appears to be restricted to 10-30% of luminal cells in the ducts and alveoli (Asselin-Labat et al., 2006; Mallepell et al., 2006) and ERα-positive cells appear to be non-dividing cells that instruct adjacent epithelial cells to proliferate via paracrine regulatory effects (Clarke et al., 1997; Mallepell et al., 2006; Russo et al., 1999; Seagroves et al., 1998). The basal MaSC-enriched subset was found to be ERα-negative and PR-negative (Asselin-Labat et al., 2006). It has been hypothesized that these ERα-negative and PR-negative MaSCs undergo asymmetric division to generate ERα-positive cells, which subsequently proliferate in response to estrogen and secrete paracrine factors that regulate adjacent ERα-negative cells (Dontu et al., 2004).

Using the same cell surface markers (CD24/CD49f), the luminal and basal cells of the mammary epithelium can be clustered into two distinct populations: $CD24^{hi}CD49f^{+}$ (luminal) and $CD24^{+}CD49f^{hi}$ (basal) (Figure 1.14B). The ability to separate these two populations is critical as it allows the systematic characterization of the progenitor and differentiated cell populations within the mammary epithelium.

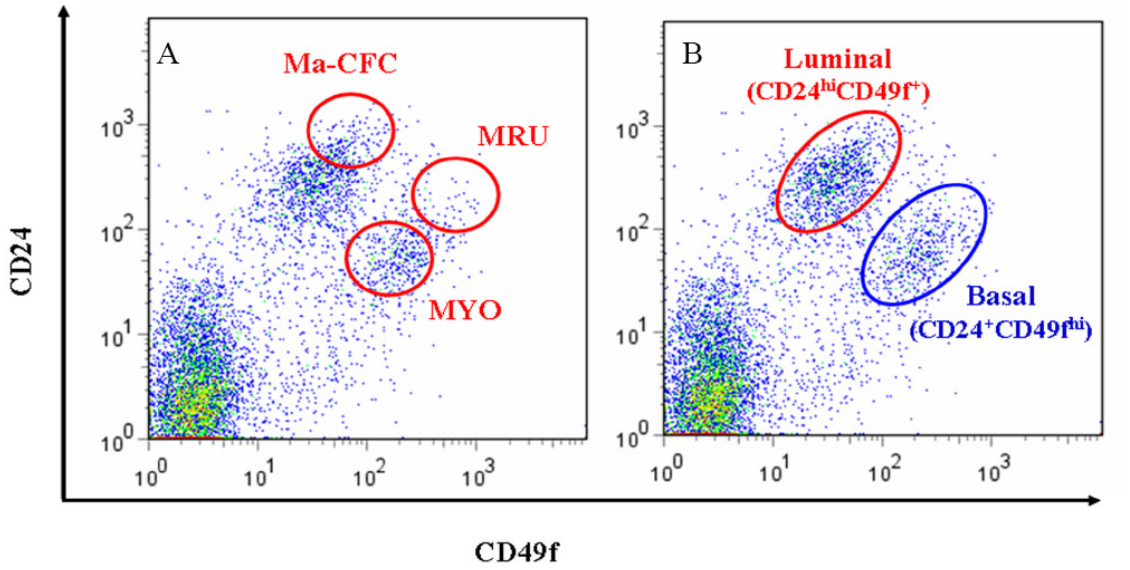


Figure 1.14. Characterization of murine epithelial cells using cell surface markers. (A) Distribution of mammary repopulating units (MRU)-CD24^{med}CD49f^{hi}; mammary-colony forming cells (Ma-CFC) C24^{hi}CD49f^{lo} and myoepithelial cells (MYO) CD24^{lo}CD49f^{lo}. **(B)** Distribution of CD45^{-}Ter119^{-}CD31^{-} epithelial cells according to their CD24 and CD49f expression. Luminal cells are characterized by CD24^{hi}CD49f^{+} whereas basal cells are characterized by C24^{+}CD49f^{hi}.

1.5.2.2 Mammary luminal progenitors

Different types of mammary progenitor cells can be detected under various *in vitro* culture conditions (Dontu et al., 2003; Smalley et al., 1998; Stingl et al., 2001). Mammary colony forming cells (Ma-CFCs) refer to progenitors that can form discrete mammary colonies *in vitro* (Stingl et al., 2001). In the mouse mammary epithelium, most Ma-CFCs are localized within the $CD24^{hi}CD49f^{lo}$ subset (Figure 1.14A) (Stingl et al., 2006). This subset lacks any repopulating capability and is found within the luminal profile and has been shown to express high levels of cytokeratin 6, a putative progenitor cell marker (Hu et al., 2006; Stingl et al., 2006). Thus the $CD24^{hi}CD49f^{lo}$ subset most likely represents luminal progenitors which are responsible for generating the luminal lineages.

Transplantation of mammary epithelial cells into a cleared mammary fat pad generates outgrowths which consist of (1) lobular (secretory) cells only, (2) ducts only and (3) both lobular cells and ducts (Smith, 1996). This demonstrates that three distinct multi-potent mammary cell types exist in the murine mammary gland (Smith, 1996). All three cell types are most likely derived from a common ancestor, but the precise progenitor cells that generate the lobule-restricted or the ductal-restricted outgrowths have not been identified. Recently, CD61 has been identified as a luminal epithelial progenitor marker (Asselin-Labat et al., 2007). Lineage-negative $CD29^{lo}CD24^{+}CD61^{+}$ subset is highly enriched for Ma-CFCs (Asselin-Labat et al., 2007). Another luminal progenitor population which is enriched for Ma-CFCs has also been identified (Sleeman et al., 2007). This subset of progenitor cells exhibits low levels of CD133/prominin expression, suggesting that luminal differentiation is accompanied by reduced CD61 but increased CD133 expression, resulting in a mature luminal ($CD61^{-}CD133^{+}$) luminal cell (Vaillant et al., 2007). Interestingly, a subset of luminal progenitors was found to express ERα (Asselin-Labat et al., 2007). Consistent with this study, an ERα-positive luminal progenitor subset has been identified and characterized by John Stingl (Stingl and Watson; manuscript in preparation). Using cell surface markers CD49b (α2-integrin) and Sca1 (Ly-6A/E), the luminal subset can be classified into progenitor ($CD49b^{+}$) or differentiated ($CD49b^{-}$) cells based on the presence/absence of CD49b (Figure 1.15) (Stingl and Watson; manuscript in preparation). In addition, the luminal progenitor

population can be separated into Sca1$^-$ and Sca1$^+$ luminal progenitors. The CD49b$^+$Sca1$^+$ subset shows expression of ERα while almost all the CD49b$^+$Sca1$^-$ cells are ERα-negative (Stingl and Watson; manuscript in preparation). Further analysis of the Sca1$^-$ and Sca1$^+$ luminal progenitors revealed that the Sca1$^-$ fraction showed relatively high levels of expression of milk proteins. Hence, these results revealed that two functionally distinct luminal progenitor populations exist in the mammary epithelium and it has been postulated that the CD49b$^+$Sca1$^-$ (ERα-negative) subset represents alveolar progenitors while the CD49b$^+$Sca1$^+$ (ERα-positive) subset represents ductal progenitors (Stingl and Watson; manuscript in preparation). Definitive proof of this hypothesis is needed by transplanting these two different luminal progenitor populations into mammary fat pad.

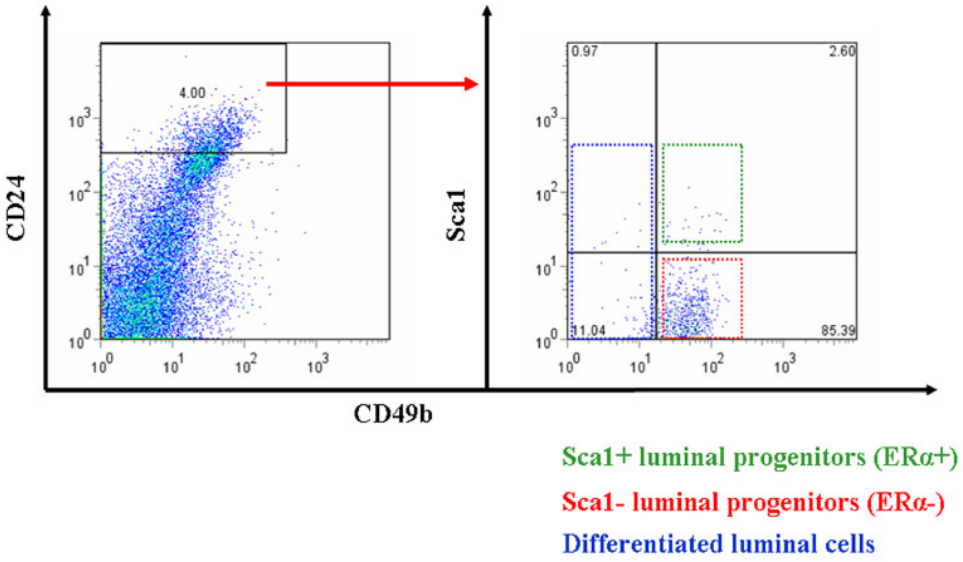


Figure 1.15. FACS profile of murine luminal epithelial cells. Luminal mammary epithelial cells (CD24^{hi}) can be separated into luminal progenitors (CD49b-positive) and differentiated luminal (CD49b-negative) populations. The luminal progenitors can be further classified by their expression of Sca1. The Sca1$^+$ and Sca1$^-$ mammary epithelial cells form colonies *in vitro*, demonstrating their progenitor property.

1.5.2.3 Myoepithelial/basal cells

The myoepithelial/basal cells occupy a $CD24^{+}CD49f^{hi}/CD29f^{hi}$ position in the mammary flow cytometric profile which partially overlaps with the MaSC-enriched subset (Figure 1.14B). Hence, the MaSC-enriched subset also contains myoepithelial/basal cells (Shackleton et al., 2006; Stingl et al., 2006). The Ma-SC-enriched and myoepithelial/basal subsets showed elevated expression in cytokeratin 5 and 14, smooth muscle actin, smooth muscle myosin, vimentin and laminin transcripts (Stingl et al., 2006). Thus the myoepithelial/basal subset contains a mixture of MaSC and committed myoepithelial/basal cells but may also contain a putative common luminal-myoepithelial progenitor and/or myoepithelial progenitor.

1.5.2.4 Current model of mammary epithelial hierarchy

Based on the recent data, a model of the hierarchy of mammary epithelial cells has been proposed (Figure 1.16). The basal ERα/PR-negative MaSC may divide asymmetrically to generate a putative common bi-potent progenitor. This common progenitor would then give rise to the luminal and myoepithelial progenitors. The luminal progenitor consists of both ERα-positive and ERα-negative subsets which may represent ductal and alveolar progenitors respectively. Differentiation of the ERα-positive ductal progenitor produces ERα-positive ductal cells. In addition, the ERα-positive ductal progenitor may secrete paracrine factors that may induce the proliferation and/or differentiation of adjacent ERα-positive and ERα-negative luminal progenitors. The ERα-negative alveolar progenitors proliferate during pregnancy and differentiate to form secretory alveolar cells. The myoepithelial progenitors would give rise to myoepithelial cells.

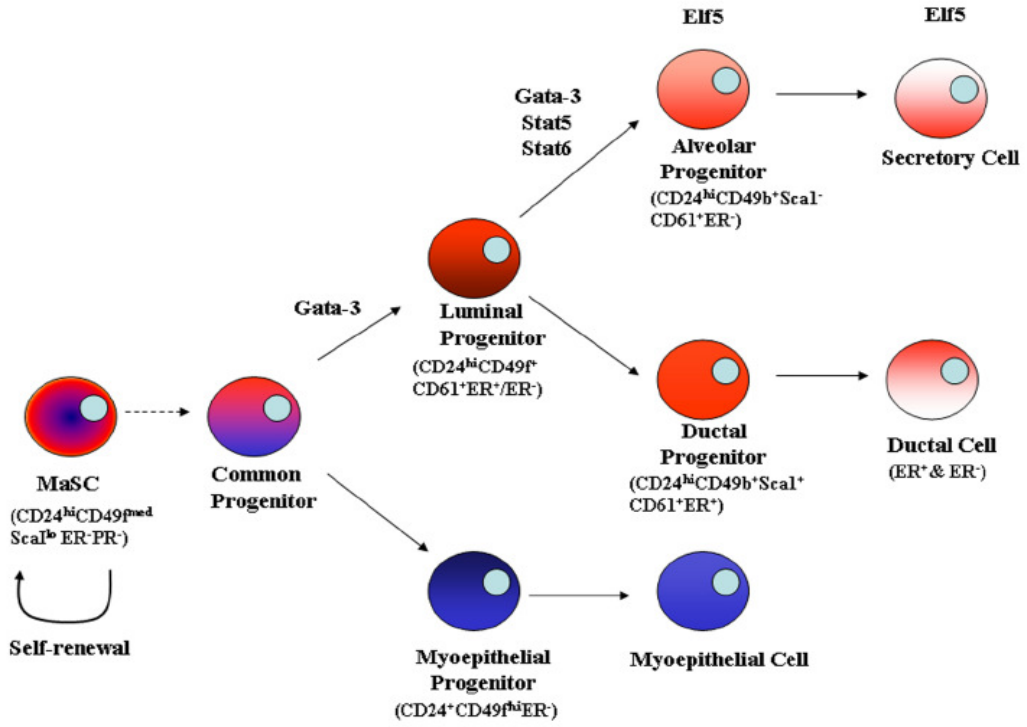


Figure 1.16. Hierarchy of murine epithelial cells. Model of epithelial hierarchy and transcription factors implicated in lineage commitment in the murine mammary gland. The mammary stem cell (MaSC) exhibits high levels of integrins and is negative for the estrogen receptor and the progesterone receptor. Proliferation and differentiation of the MaSC gives rise to a common progenitor which would generate the lineage-restricted luminal and myoepithelial progenitors. Two phenotypically distinct populations of luminal progenitors exist in the epithelium: $CD49b^{+}Sca1^{-}(ER\alpha^{-})$-putative alveolar progenitors and $CD49b^{+}Sca1^{+}$ $(ER\alpha^{+})$-putative ductal progenitors. Proliferation and differentiation of these progenitors generates the differentiated luminal cells. MaSC: Mammary stem cells; ER: estrogen receptor α; PR: progesterone receptor; Stat: Signal transducer and activator of transcription.

1.5.3 Lineage commitment during puberty and pregnancy

1.5.3.1 Transcription factors

Although the mammary epithelial hierarchy has been proposed, the precise molecular mechanisms that regulate mammary stem/progenitor cell proliferation and differentiation into the various epithelial lineages remains poorly understood. However, recent publications have identified several transcription factors that are important for lineage commitment during mammary gland development (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006; Oakes et al., 2008), adding a different perspective to the existing paradigm that steroid hormones and prolactin are the principal temporal regulators of mammary gland lineage commitment and differentiation.

The Gata family of transcription factors plays an important role in cell fate specification in many different tissues such as the skin and the immune system. In the immune system, *Gata-3* has been shown to be critical for T cell specification and in T_H2 cell differentiation (Rothenberg, 2007a). Recently, *Gata-3* has been shown to be essential for mammary gland development and the maintenance of differentiated luminal cell fate (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). In the first study, *Gata-3* was identified to be the most highly enriched transcription factor in the developing and mature mammary luminal cells (Kouros-Mehr et al., 2006). Conditional deletion of *Gata-3* using MMTV-Cre resulted in the failure of TEB formation and profound structural defects (Kouros-Mehr et al., 2006). In addition, there was a selection pressure to retain a functional *Gata-3* allele in the surviving mammary outgrowths, suggesting that *Gata-3* is essential for mammary development. Subsequently, conditional deletion of *Gata-3* using WAP-rtTA-Cre (where Cre expression is induced by doxycycline administration) revealed that *Gata-3* is necessary for maintenance of the luminal epithelium in the adult mammary gland (Kouros-Mehr et al., 2006). Acute loss of *Gata-3* (5 days after doxycycline treatment) led to the expansion of undifferentiated luminal cells as demonstrated by a dramatic reduction of the levels of differentiated luminal markers (Kouros-Mehr et al., 2006). In contrast, chronic loss of *Gata-3* (14 days after doxycycline treatment) led to luminal cell death, leading to severe lactational defects. Importantly, these defects occurred because the lobulo-alveoli did not develop properly, rather than because of any apoptosis or abnormalities in cell proliferation. Similar results were

obtained in a separate study where *Gata-3* was deleted using the WAP-Cre mice where Cre recombinase was expressed in the epithelium during pregnancy (Asselin-Labat et al., 2007). Deletion of *Gata-3* resulted in failed alveoli development and lactogenesis. Additionally, consistent with the first study which demonstrated that loss of *Gata-3* resulted in an increase in undifferentiated luminal cells, Asselin-Labat *et al.* revealed that loss of *Gata-3* caused a dramatic increase in the pool of $CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitors, suggesting a block in luminal differentiation (Asselin-Labat et al., 2007). Over-expression of *Gata-3* in the $CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitors was sufficient to induce terminal differentiation of these cells, eventually cumulating in the expression of milk protein transcripts such as *β-casein* and *Wap* (Asselin-Labat et al., 2007). Collectively, these results indicated that *Gata-3* is required for the maintenance and differentiation of the luminal lineage (Asselin-Labat et al., 2007). The effects of loss of *Gata-3* in the mammary gland resembled those of *ERα* deficiency, suggesting that these two genes may function in the same pathway (Mallepell et al., 2006). Subsequently, Gata-3 was shown to bind directly to the promoter of forkhead transcription factor (*FOXA1*), which has been suggested to be essential for estrogen signalling in the mammary gland and required for binding of ERα to chromatin (Carroll et al., 2005). It has also been shown that *Gata-3* is required for estrogen-mediated breast cancer cell growth and ERα was shown to directly stimulate *Gata-3* expression (Eeckhoute et al., 2007); however whether this feedback loop applies to the normal mammary gland remains to be established.

The prolactin-regulated Ets transcription factor *Elf5* has recently been shown to specify alveolar cell fate (Oakes et al., 2008). Loss of one allele of *Elf5* in the mouse resulted in females that displayed defective lobulo-alveoli development and reduced milk secretion (Zhou et al., 2005) and retroviral over-expression of *Elf5* in *PrlR*-deficient mammary cells rescued alveologenesis (Harris et al., 2006). Loss of both alleles of *Elf5* caused embryonic lethality; therefore Oakes *et al.* transplanted $Elf5^{-/-}$ mammary cells into cleared fat pads of $Rag1^{-/-}$ host mice and analyzed the effects of *Elf5* deletion on mammary development (Oakes et al., 2008). While ductal morphogenesis was unaffected, alveologenesis during pregnancy was severely affected (Oakes et al., 2008). Analysis of these mammary glands by flow cytometry revealed an accumulation of

$CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitors, similar to that observed in the *Gata-3*-deficient mammary glands, suggesting that loss of *Elf5* resulted in the blockade of luminal differentiation. Quantitative PCR analysis of the $CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitors in both the virgin and 12.5 dpc gestation mammary glands demonstrated expression of both *Gata-3* and *Elf5*. However, immuno-staining showed that they were expressed in different cell populations (Oakes et al., 2008). The differential expression patterns of both genes were also observed in the $CD29^{lo}CD24^{+}CD61^{-}$ mature luminal cells. Whereas expression of *Gata-3* was high in the $CD29^{lo}CD24^{+}CD61^{-}$ mature luminal cells in the virgin but decreased at late gestation, *Elf5* exhibited the exact opposite expression patterns – low expression in the virgin glands but high expression in the late gestation mammary glands. In addition, over-expression of *Elf5* in the virgin glands disrupted ductal morphogenesis and caused precocious alveologenesis and milk production (Oakes et al., 2008). This prompted the authors to suggest that different populations of Gata-3-responsive and Elf5-responsive cells exist within the $CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitor population and conclude that *Elf5* is important for specifying alveolar cell fate. Whether Gata-3 and Elf-5 co-operate in luminal progenitor cells to regulate alveoli differentiation is still unknown.

Several other transcription factors such as the helix-loop-helix inhibitor (*Id2*), *C/EBPβ*, *Stat5* and *Stat6* have also been implicated in alveolar differentiation. Khaled *et al.* demonstrated that Stat6 and its upstream cytokines IL4 and IL13 are required for the expansion of luminal lineage during early gestation (Khaled et al., 2007). Loss of *Stat6* or double knockout of *Il4* and *Il13* resulted in a dramatic reduction in the number of alveoli during gestation, with up to 70% reduction at day five of gestation compared to the wild-type mice (Khaled et al., 2007). Conversely, loss of *Socs5*, a negative regulator of Stat6 caused precocious alveolar development. Hence, Stat6 signalling is important for the regulation of mammary luminal progenitors during lactation.

An interesting perspective to mammary lineage commitment has been added with the recent studies of Gata-3 and Stat6 in mammary development (Watson and Khaled, 2008). As discussed above, both Gata-3 and Stat6 are required for development of the luminal lineage in the mammary gland. In the hematopoietic system, naïve T cells can be polarized into T_H1 or T_H2 cells by cytokines. Both Gata-3 and Stat6 have been shown to

be required for T_H2 cell polarization (Ansel et al., 2006). Interestingly, mammary epithelial cells (MEC) secrete type-1 cytokines IL12a, interferon-γ (IFNγ) and tumour necrosis factor (TNF) in culture. However upon lactogenic stimuli, the MEC switched to secreting type-2 cytokines such as IL4, IL13 and IL5 (Khaled et al., 2007). This demonstrates a role for these immune cell cytokines in epithelial cell fate decision and suggests that the genetic control of mammary epithelial cell fate and hematopoietic cell fate might be conserved. Collectively, these studies demonstrate that temporal regulation of transcription factors in addition to steroid hormones and prolactin signalling, are also required for mammary cell specification and lineage commitment (Figure 1.16).

1.5.3.2 Notch signalling

Notch signalling plays important roles in cell fate decision and lineage specification in several tissues such as the hematopoietic system where temporal regulation of Notch signalling is critical for lymphocyte development (Radtke et al., 2004). Not surprisingly, Notch signalling is also required for T_H1/T_H2 cell fate determination (Ansel et al., 2006; Radtke et al., 2004). In mammals, there are four Notch receptors (Notch1-4) and five ligands (Jagged1 and 2, Delta-like-ligand1, 3 and 4). Notch receptors are synthesized as precursor proteins that are cleaved upon transport to the cell membrane and are expressed as hetero-dimers. Ligand binding to the extracellular portion of the Notch receptor triggers a cascade of proteolytic reaction that cumulates in the liberation of the cytoplasmic domain of Notch receptor (ICN). The ICN translocates to the nucleus and binds to the transcription factor recombination-binding protein-Jκ (RBP-Jκ), converting it from a transcription repressor to a transcription activator. ICN also recruits several co-activators such as mastermind-like protein (MAML) and p300, eventually resulting in the activation of Notch target genes, including the hairy enhancer of split (*Hes*) genes. Mammary fat pads transplanted with *Rbpj*κ-deficient mammary cells showed normal ductal morphogenesis, suggesting that RBP-Jκ is not required for the establishment of luminal and basal cells (Buono et al., 2006). Instead, it was revealed that lobulo-alveolar development during pregnancy was affected. Importantly, there was an accumulation of epithelial cells with expression of both luminal (cytokeratin 18) and basal (cytokeratin 14) markers and there was also an excess of luminal cells expressing

the p63 basal marker. In addition, there was a transient amplification of cytokeratin 6-positive luminal cells that could reflect either a block in differentiation or a slower progression through the luminal differentiation pathway (Buono et al., 2006). This study suggests that loss of *Rbpj*κ during pregnancy results in luminal cells that acquire basal-like features and that the Notch signalling pathway through RBP-Jκ is important for the maintenance, but not the establishment of luminal cells. Other studies have also implicated the importance of Notch signalling during lobulo-alveolar development during pregnancy. Over-expression of Notch1 and Notch3 ICN using the mouse mammary tumour virus (MMTV) provirus resulted in impaired ductal and lobulo-alveoli development during pregnancy and the transgenic mice developed mammary tumours, confirming the oncogeneic potential of Notch1 ICN *in vivo* (Hu et al., 2006). In addition, Notch3 has been shown to be critical for the commitment of bi-potent progenitors to the luminal lineage *in vitro* (Raouf et al., 2008). Collectively, these studies indicate the importance of Notch signalling in mammary development and suggest that temporal regulation of the levels of canonical Notch signalling pathway is crucial to proper lobulo-alveolar development.

1.6 Thesis project

The work of this thesis is based on a simple question: "Do Bcl11 proteins play important roles in development and cell fate decisions in the mammary gland?" The purpose of this project was to identify unique patterns of *Bcl11* expression in the mouse and to characterize their roles in development and lineage commitment.

Current expression data is derived primarily from reverse transcriptase PCR and *in situ* hybridization; whilst being informative, the resolution and sensitivity of these techniques at a single cell level is limited. Hence, the first aim of this thesis was to generate *lacZ*-tagged reporter alleles of both *Bcl11* genes using a new recombineering based approach (Chan et al., 2007). These mice were then used to characterize the spatial and cellular expression of *Bcl11* genes in the embryo and the adult tissues.

The second aim of this study was to characterize the roles of *Bcl11a* and *Bcl11b* in development and lineage commitment in specific tissues. The mammary gland provides a great model system to study various aspects of cellular proliferation,

differentiation and apoptosis. With the recent purification of mammary stem cells and the availabilities of different cell surface markers to isolate different subsets of mammary epithelial cells, the mammary gland has become an ideal system to study lineage commitment and cell fate decisions. These, together with the identification and validation of two *BCL11A* mutations in human breast cancer samples (Wood et al., 2007) prompted me to investigate the roles of *Bcl11* genes in mammary development.

In this thesis, I first characterized the expression of the *Bcl11* genes in various mammary lineages using *lacZ*-tagged reporter mice in combination with flow cytometric analysis. Subsequently, loss-of-function and gain-of-function approaches were used to dissect the roles of the *Bcl11* genes in embryonic and postnatal mammary development and in lineage commitment. Finally, I investigated the underlying molecular mechanisms implicated in the phenotypes.

CHAPTER 2:

MATERIALS AND METHODS

2.1 Vectors

2.1.1 Gene targeting vectors

The following vectors were obtained from Pentao Liu and used in the construction of targeting and retrieval constructs.

PL611 (Retrieval vector) (Appendix A.1):

PL611 is a retrieval vector derived from pBR322 (Covarrubias et al., 1981) which can support replication of relatively large pieces of mammalian genomic DNA in *E. coli*. This plasmid serves as the PCR template for generating the retrieval vector.

Bsd cassette (*I-Sce*I-*EM7-Bsd-I-Ceu*I selection plasmid) (Appendix A.2):

The *Bsd* (blasticidin) cassette serves as the PCR template for generating the 5' targeting vector. It contains two rare enzyme restriction sites (*I-Sce*I and *I-Ceu*I) flanking the Bsd resistance coding sequence (Bsd^R).

Neo cassette (*lox*P-*F3-PGK-EM7-Neo-F3* selection plasmid) (Appendix A.3):

The *Neo* (neomycin) cassette serves as the PCR template for generating the 3' targeting vector. It contains a single *lox*P site and two *F3* sites flanking the Neo resistance coding sequence (Neo^R).

PL613 (*I-Sce*I-*FRT-SA-IRES-lacZ-PGK-EM7-Puro-FRT-lox*P-*I-Ceu*I plasmid) (Appendix A.4):

PL613 is the *lacZ* reporter cassette plasmid which contains the bacterial *lacZ* gene with a nuclear localisation signal and the puromycin resistance coding sequence ($Puro^R$). The *lacZ* and *Puro* selection marker are flanked by two *FRT* sites. Restriction digestion of this plasmid with two rare enzyme restriction sites (*I-Sce*I and *I-Ceu*I) would release the *lacZ* reporter module.

Cm-TK cassette (*MC1-TK-Cm* plasmid) (Appendix A.5):

The *Cm-TK* cassette contains the chloramphenicol resistance coding sequence (Cm^R) and the negative selection marker, *MC1-TK*. *Cm-TK* cassette is flanked by two 600 bp homology regions to PL611, therefore *MC1-TK* can be added to the targeting vector backbone by recombineering.

pSim6/pSim18 (Appendix A.6):

pSim6 was constructed by recombination of the λ-defective prophage into pSC101 plasmid backbone (Datta et al., 2006). The ampicillin resistance coding sequence (Amp^R) is present as a substitute for the *rexAB* genes. pSim18 was constructed by replacing the Amp^R in pSim6 with either the Bsd^R or hygromycin-resistant coding sequences (Hyg^R) using the recombineering functions expressed from the plasmid.

2.1.2 Over-expression vectors

PL417 (*MSCV-IRES-eGFP* plasmid) and PL419 (*MSCV-Bcl11a* cDNA-*IRES-eGFP* plasmid) were obtained from Pentao Liu. *MSCV-Bcl11b* cDNA-*IRES-eGFP* plasmid was constructed as follows: Firstly, two pairs of primers were designed for amplification of *Bcl11b* cDNA (Appendix A.7A) from total mouse brain oligo-dT cDNA library (constructed as detailed in Chapter 2.6.3). PCR amplification was carried out using Pwo master (Roche) and performed using PTC-225 PCR machine (Peltier Thermal Cycler) with the following settings: 94°C for 2 min, this was followed by 35 cycles of 94°C for 30 sec, 52-62°C for 30 sec and 72°C for 45 sec (*Bcl11b* exons 1-2-3) or 2.5 min (*Bcl11b* exons 3-4). This was then followed by 72°C for 5 min. PCR products, containing *Bcl11b* exons 1-2-3 (529 bp; Appendix A.7B) and exons 3-4 (2240 bp; Appendix A.7C) respectively, were run through a 1% agarose gel and appropriate bands were purified using Qiagen gel extraction kit. 4 µl of purified PCR products were used for cloning into pCR-BluntII-TOPO vector and transformed into TOP10 cells using manufacturer's protocol (Invitrogen). 1.5 µg of each plasmid were digested with *Bgl*II and *Cla*I (*Bcl11b* exons 1-2-3; Appendix A.7D) or *Cla*I and *Eco*RI (*Bcl11b* exons 3-4; Appendix A.7E) and ligated with 1.5 µg of PL417 that has been cut with *Bgl*II and *Eco*RI. 2 µl of ligation products were transformed into DH10B cells and transformants were selected on LB-

Amp plates. Plasmids obtained from minipreparations of Amp^R colonies were checked by restriction digestion (Appendix A.7F) and sequencing to ensure that there were no mutations in the *Bcl11b* cDNA.

2.1.3 *Bcl11-lacZ* reporter conditional null vectors

2.1.3.1 Construction of targeting and retrieval vectors

A novel PCR method using 100 mers PAGE-purified primers to construct the targeting and retrieval vectors was employed. These primers consist of 80 bases of homologous sequence to designated *Bcl11* targeted/retrieval regions (Ensembl *Mus musculus,* Sanger Institute) and 20 bases of complementary sequence to targeting/retrieval vector plasmids (Appendix A.1-A.3). Primers used are listed in Table 2.1. The templates for PCR amplification were prepared as follows: The linear retrieval vector (PL611, *Eco*RI and *Bam*HI digested; Appendix A.1), *loxP-F3-Neo-F3* (*Neo* cassette, *Not*I and *Sal*I digested; Appendix A.2) and *I-Sce*I-*Bsd-I-Ceu*I (*Bsd* cassette, *Eco*RI and *Bam*HI digested; Appendix A.3) selection cassettes were run through a 1% agarose gel and bands corresponding to 560 bp (*Bsd),* 2 kb (*Neo*) and 3 kb (PL611) were excised and purified using Qiagen purification kit. To avoid the background from uncut cassettes of the PCR templates, 1 ng of the restriction digested and gel-purified retrieval vector and the two selection cassettes were transformed into DH10B cells. If there were any background drug resistant colonies, the cassettes were re-purified and tested again until they were clean.

One nanogram of each of the above linear purified products was used as a template for PCR reaction to generate PCR products for targeting or retrieving. PCR amplification was carried out using Extensor Hi-Fidelity PCR Master Mix 2 (ABgene). 25 µl of the master mix was added to 1 µl of template (1 ng), 2 µl of each primer (10 µM) and 20 µl of PCR-grade water. PCR was performed using PTC-225 PCR machine (Peltier Thermal Cycler) with the following settings: 94^{o}C for 4 min, followed by 35 cycles of 94^{o}C for 30 sec, 60^{o}C for 30 sec and 68^{o}C for 1 min (*Bsd*) or 2-3 min (*Neo* and retrieval backbone). This was then followed by 68^{o}C for 5 min. After PCR reactions, 0.5 µl of exonuclease I (10U, New England Biolabs, NEB) was added per 50 µl of PCR products and incubated at 37^{o}C for 1 hour followed by heat inactivation at 80^{o}C for 20 min. The

PCR products were then purified using Qiagen mini-preparation columns and eluted in 50 µl of PCR-grade water to obtain the targeting and retrieval vectors.

Table 2.1. List of targeting primers used.

Targeting Primers	Primer Sequence (5'– 3')
Bcl11a-lacZ-Ret-1	GAGACTTGGTTCAAGAAACAAATATGTGTCCCTTTTGTTGTTTGTGCTAAA TTGGGAGTGAGGTTTAAAAAAAAAAATCAGATACGACTCACTATAGGGAG
Bcl11a-lacZ-Ret-2	TAATGCCTTTTATCCAAAGCCAGGAGACTTTTATCTTTTTAAGCATCGGCA AAGTAAGGTGTTTGGCTCTTACTTTTATTTTAGTGAGGGTTAATTATCG
Bcl11a-lacZ-5'-F	TCTAGCCTCACATAGGGGAGAAAGTGTATTTCTCAGTTATACTTTAAGCCC TGGCATTTTTTTAAAGTGTCTGGGACATTCTAGGGATAACAGGGTAATG
Bcl11a-lacZ-5'-R	TAACCAACAGTTTACCAGCCAACTGCAACATTTAAGGATGTAGAAGAGAC AATGGCCTAGGGACAAGGATGAGTCTAGCTTCGCTACCTTAGGACCGTTA
Bcl11a-lacZ-3'-F	AAAGTTTTATTTTATTTTATTTTTAAATGGTTATCAAATTGAATGTGAAATG TGCAAAGGCCCTGGAATGTGATGAAATATGTAAAACGACGGCCAGTGA
Bcl11a-lacZ-3'-R	TCTACTTCTTTTGGCGCCAGAATTTCATTAAATGCATCATTTTAAACAAGTA TTGTCACAAGATGAACTTCTTGCTAATGAGGAAACAGCTATGACCATG
Bcl11b-lacZ-Ret-1	CCTGCAGTTCAATCTAGTCTGGGGCGTTTCCTTAGCTTTTGCAAGCAGACA CTTCTGTTCCCTTCAGACTTGACACTGCAATACGACTCACTATAGGGAG
Bcl11b-lacZ-Ret-2	CGTGATTTTGGTGCTTTCTGTGTCTGCCCTGCATAGCCAAGGCACATTGATC TCTGCGGTTTGCTTTTGTACTGGCAAGGTTAGTGAGGGTTAATTATCG
Bcl11b-lacZ-5'-F	GCATTCCTGGGGACTGGGCTTGAGTCCTCAGACACGCTAAGCAAGCATTC CACGACGGAGCATCATAGCCATTTTTGTTACTAGGGATAACAGGGTAATG
Bcl11b-lacZ-5'-R	TGAAAATCACTCCACTGCATCCTCCCCTTGGCAACTACTGACTCTGTTGCTT TGCAAACTTAAAAAAAAAGAAAAAAAATTCGCTACCTTAGGACCGTTA
Bcl11b-lacZ-3'-F	CAGAGGCGTCTCTGTGTGGAATTCAACTTTCATGCTATGGAAGGAAAATGT TGAGGGGGGCGTGGCAAACGAGCCTCTCCTGTAAAACGACGGCCAGTGA
Bcl11b-lacZ-3'-R	CTCCTGGTAACACACAATTGCAGGATGTGGGGGCTGTTCCTAGAGTCACCC TGGAAAAGATTCTCGGGGTCCCAAGTGGGAGGAAACAGCTATGACCATG
Bcl11a_Bsd_PCR F	TCTAGCCTCACATAGGGGAGAAAGT
Bcl11a_Bsd_PCR R	CAGTTCAGCTAGTCAACTCTTTGGC
Bcl11a_Neo_PCR F	GGCTGTTTCGCCTAAAGTTTTATTT
Bcl11a_Neo_PCR R	CACAGACATAGATACTGCTGCTCTT
Bcl11b_Bsd_PCR F	GCATTCCTGGGGACTGGGCTTGAGT
Bcl11b_Bsd_PCR R	CGTCCCCTGTAGGTGTTCTTATTTT
Bcl11b_Neo_PCR F	CCTTTGCCCTACAGTGCCCATCAGC
Bcl11b_Neo_PCR R	CCTTTGCCCTACAGTGCCCATCAGC

Underlined sequences - complementary to retrieval/targeting plasmid templates

2.1.3.2 Construction of final targeting vectors

The final step in the construction is the replacement of the *Bsd* cassette on the retrieved plasmids with the *lacZ* reporter. Restriction digestion of the retrieved plasmids (1.5 µg) and PL613 (1.5 µg; Appendix A.4) with *I-Sce*I and *I-Ceu*I (NEB) was first carried out. The restriction digestion reaction of PL613 was run through a 1% agarose gel and the *lacZ* reporter cassette (7 kb band) was purified using QIAquick Gel Extraction kit (Qiagen). Next, ligation reaction of purified digestion products using 12 µl of *lacZ* reporter cassette (600 ng), 10 µl of purified retrieved plasmid (50 ng), 2.5 µl of T4 DNA

ligase buffer with 1 µl of T4 DNA ligase (NEB) was set up and incubated at room temperature for 2 hours. 5 µl of the ligation products was added to chemical competent TOP10 cells (Invitrogen) and incubated on ice for 30 min. The cells were then heat shocked at 42^{o}C for 30 sec and 250 µl of SOC was added. The transformation mixture was incubated at 37^{o}C for 1 hour before plating out the cells on a LB-Puro-Kan plate. The plate was incubated at 37^{o}C overnight and colonies are typically observed after 16-24 hours.

2.2 Recombineering

2.2.1 Using λ-phage

The replication-defective λ prophage (strain LE392) λ CI_{857} *ind1* $Cro_{TYR26amber}$ $P_{GLN59amber}$ *rex<>tetRA*, was created by combining several different mutations (Chan et al., 2007). The prophage backbone contains the three Red genes, *exo*, *bet* and *gam* and the tetracycline resistance coding sequence (Tet^R). Cells which have been successfully infected with the prophage are Tet^R and recombineering competent. BAC clones containing *Bcl11a* (bMQ-202K13 and bMQ-44A03) and *Bcl11b* (bMQ-188F4) regions were identified and obtained from Wellcome Trust Sanger Institute 129/AB2.2 BAC library. Each BAC clone was inoculated into 1 ml of LB with Cm (in a 15 ml polypropylene tube) for overnight growth at 37^{o}C with shaking at 244 G. Cells were then spun at 16,000 G for 25 sec and washed with 1 ml of SM buffer (100 mM $CaCl_2$; 8 mM $MgSO_4.7H_2O$ and 50 mM Tris-Cl, pH 7.5) before being re-suspended in 100 µl of SM buffer. 2 µl of the prophage lysate was added and the suspension was incubated at 32^{o}C for 20 min. 500 µl of LB was then added and the suspension was incubated at 32^{o}C for another 20 min. The cells were then plated out onto LB-Tet plates and incubated overnight at 32^{o}C.

Tet^R BAC clones were inoculated into 1 ml of LB with Tet and grown overnight at 32^{o}C with shaking at 243 G. The next day, 30 µl of the overnight culture was inoculated into 1 ml of LB and incubated at 32^{o}C with shaking at 243 G till OD_{600} = 0.2 (~2 hours). The cells were then heat shocked in a 42^{o}C water bath for 15 min and cooled immediately on ice for 2 min. Cells were spun down at 16,000 G for 25 sec and washed sequentially with 1 ml of ice-cold 0.1 M $CaCl_2$ and $MgCl_2$ and 1 ml of ice-cold 0.1 M

$CaCl_2$. Finally, the cell pellet was re-suspended in 50 µl of targeting/retrieval long primer PCR products (~4 µg) and 5 µl of 1 M $CaCl_2$. This mixture was heat shocked at 42°C for 2 min and cooled immediately on ice. 1 ml of LB was added and cells were incubated at 32°C with shaking at 243 G for 2 hours before plating them out on agar plates containing the appropriate antibiotics. Plates were incubated overnight at 32°C.

2.2.2 Using pSim plasmids

pSim18 (Appendix A.6) has the pSC101 replication origin which is low copy and temperature sensitive. This plasmid contains the three Red genes, *exo*, *bet* and *gam*, together with *CI857* repressor and *pL* promoter. Additionally, pSim18 has the Hyg^R cassette inserted between *CI857* and *pL* promoter. Consequently, transformants of pSim18 are selected with Hyg and these cells are now recombineering-competent. The same BAC clones as described in Chapter 2.2.1 were inoculated in 1 ml of LB with Cm (in a 15 ml polypropylene tube) for overnight growth at 37°C with shaking at 243 G. Cells were then spun at 16,000 G for 25 sec and washed thrice with ice-cold water. Cells were re-suspended in 50 µl of ice-cold water with 1 ng of pSim18 and electroporation was carried out using the following conditions: 1.75 kV, 25 F with the pulse controller set at 200 Ω. After electroporation, 1 ml of LB was added to the cuvette which was subsequently incubated at 32°C for one hour. Cells were then plated on a LB-Hyg plate and incubated overnight at 32°C.

Hyg^R BAC clones were inoculated into 1 ml of LB with Hyg and grown overnight at 32°C with shaking at 243 G. The next day, 30 µl of the overnight culture was inoculated into 1 ml of LB and incubated at 32°C with shaking at 243 G till $OD_{600} = 0.2$ (~2 hours). The cells were then heat shocked in a 42°C water bath for 15 min and cooled immediately on ice for 2 min. Following which, cells were spun at 16,000 G for 25 sec and washed thrice with ice-cold water and finally re-suspended in 50 µl of ice-cold water with 100 ng of targeting/retrieval long primer PCR products and electroporation was performed. After electroporation, 1 ml of LB was added to the cuvette which was subsequently incubated at 32°C for one hour. Cells were then plated onto agar plates containing the appropriate antibiotics and incubated overnight at 32°C.

2.2.3 Expression of Cre/Flpe recombinase

EL250 and EL350 are two DH10B derived *E. coli* strains that contain a defective prophage with *pL* operon encoding *gam* and the red recombination genes, *exo* and *bet*, under tight control of the temperature sensitive λ repressor (*CI857*). EL250 and EL350 have the Flpe recombinase or the Cre recombinase under the tight control of *AraC* and P_{BAD} respectively. Upon addition of arabinose, expression of Flpe or Cre recombinase will be induced. One colony of EL250 or EL350 was inoculated into 1 ml of LB and incubated overnight at 32°C with shaking at 243 G. The next day, 400 µl of overnight culture was inoculated into 10 ml of fresh LB and incubated at 32°C with shaking at 243 G for 2 hours. 100 µl of 10% L(+) arabinose (Sigma) was added to the culture (final concentration, 0.1%) and incubated for 1 hour at 32°C with shaking at 243 G. Following which, cells were spun at 16,000 G for 25 sec and washed thrice with ice-cold water and finally re-suspended in 50 µl of ice-cold water with 100 ng of the final targeting constructs and electroporation was performed. After electroporation, 1 ml of LB was added to the cuvette which was subsequently incubated at 32°C for one hour. Cells were then plated onto agar plates containing the appropriate antibiotics and incubated overnight at 32°C.

2.2.4 Antibiotics

Antibodies were used at the following concentration: Ampicillin (50 µg/ml); Kanamycin (30 µg/ml); Tetracycline (12.5 µg/ml); Blasticidin (50 µg/ml); Chloramphenicol (25 µg/ml) and Hygromycin (50 µg/ml).

2.3 ES cell culture

2.3.1 Culture condition

ES cell culture was performed as described before (Ramirez-Solis et al., 1993). Briefly AB2.2 (129 S7/SvEv$^{Brd\text{-}Hprtb\text{-}m2}$) wild-type ES cells and their derivatives were always maintained on SNL76/7 feeder cell layers that had been mitotically inactivated by mitomycin C treatment. ES cells were grown in M15 media (Table 2.2) and maintained at 37°C/5% CO_2. ES cell media was changed daily unless otherwise specified.

ES cells were passaged once they reached 80-85% confluence. ES cell media was changed 2 hours prior to passaging the ES cells. ES cell media was then aspirated off and the cells were washed once with 10 ml PBS. 2 ml of trypsin was added to each 9 cm plate and the plate was incubated at 37°C for 15 min. 8 ml of fresh M15 media was added to each plate to inactivate the trypsin and the cells were dispersed by repeated pipetting. The ES cell suspension was then centrifuged at 1,585 G for 5 min and the supernatant was aspirated. The cell pellet was re-suspended in 1 ml of M15 media before redistributing the ES cell suspension to four 9 cm feeder plates. The plates were maintained at 37°C/5% CO_2.

Table 2.2. Composition of cell culture medium.

Medium Name	Recipe	Purpose
M15	Knockout Dulbecco's Modified Eagle's Medium (DMEM, Gibco/Invitrogen), supplemented with 15% foetal bovine serum (FBS, Gibco/Invitrogen), 2 mM L-Glutamine, 50 U/ml penicillin, 40 µg/ml streptomycin and 100 µM β-mercaptoethanol	Culture of undifferentiated ES cells
M10	Knockout Dulbecco's Modified Eagle's Medium (DMEM, Gibco/Invitrogen), supplemented with 10% foetal bovine serum (FBS, Gibco/Invitrogen), 2 mM L-Glutamine, 50 U/ml penicillin, 40 µg/ml streptomycin and 100 µM β-mercaptoethanol	Culture of feeder cells

2.3.2 Chemicals used for selection of ES cells

G418: Geneticin (Invitrogen) was obtained as a sterile stock solution containing 50 mg/ml of active ingredients.

Blasticidin: Blasticidin S HCl (Invitrogen), 1000X stock (5 mg/ml) was made in PBS. The stock solution was filter sterilized using a 0.2 µm syringe filter.

Puromycin: $C_{22}H_{29}N_7O_5.2HCl$ (Sigma), 100X stock (3 mg/ml) was made in MiliQ water. The stock solution was filter sterilized using a 0.2 µm syringe filter.

FIAU: 1-(2'-deoxy-2'-fluoro-b-d-arabinofuranosyl)-5-iodouracil (Oclassen Pharmaceuticals, Inc.), 1000X stock was made in PBS and NaOH was added dropwise

till the powder dissolved to give a final stock concentration of 200 mM. The stock solution was filter sterilized using a 0.2 μm syringe filter.

Trypsin: To make 5 L of trypsin, 35 g of NaCl, 5 g of D-glucose, 0.9 g of $Na_2HPO_4.7H_2O$, 1.85 g of KCl, 1.2 g of H_2PO_4, 2 g of EDTA, 12.5 g of Trypsin (1:250) and 15 g of Tris base were mixed. The pH was adjusted from 8.71 to 7.6 with HCl and phenol was added to obtain a pink colouration. Trypsin was then filter sterilized and aliquoted into 50 ml falcon tubes and stored at -20 °C.

2.3.3 Transfection of DNA into ES cells by electroporation

Prior to transfection of DNA into ES cells, ~20 μg of each targeting construct was linearized with the appropriate enzymes in a 100 μl total volume at 37°C for 2 hours. The DNA was then purified by ethanol precipitation and air-dried at room temperature. The air-dried DNA was re-suspended in sterile 1X TE (pH 8) to a final concentration of 1 μg/μl. 20 μg of DNA was used for each electroporation. ES cell electroporation was performed according to standard protocols (Ramirez-Solis et al., 1993). Briefly, ES cells (80-85% confluent) were fed 2-3 hours before harvesting. Immediately before electroporation, ES cell media was then aspirated off and the cells were washed once with 10 ml PBS. 2 ml of trypsin was added to each 9 cm plate and the plate was incubated at 37°C for 15 min. 8 ml of fresh M15 media was added to each plate to inactivate the trypsin and the cells were dispersed by repeated pipetting. The ES cell suspension was then centrifuged at 1,585 G for 5 min and the supernatant was aspirated. The cell pellet was re-suspended in PBS to a final concentration of 1 x 10^7 cells/ml. 1 x 10^7 cells were transferred into a 0.4 cm gap cuvette (Biorad) together with 20 μg of DNA. Electroporation was carried out using the Biorad gene pulser at 230 V, 500 μF. After electroporation, the cells were plated onto a 9 cm feeder plate and cultured for 10 days to allow for formation of single ES cell colonies. Drug selection was applied to the culture only 24-36 hours after electroporation.

2.3.4 Picking ES cell colonies

ES cell colonies were picked from the 9 cm feeder plates into a 96-well round bottom plate. 50 µl of trypsin was added to each well of the 96-well round bottom plate. Next, media was aspirated off the 9 cm plate and washed once with 10 ml of PBS. 8 ml of PBS was added to cover the plate and ES cell colonies were picked using a P20 Pipetman set at 10 µl and transferred into the wells of the 96-well round bottom plate (with trypsin). After completing a 96-well plate, the plate was incubated at 37°C for 15 min before adding 150 µl of fresh M15 media to each well. Next, the colonies were dispersed by vigorous pipetting using a multi-channel pipette. The ES cell suspension was then transferred to a fresh 96-well feeder plate and incubated at 37°C with 5% CO_2.

2.3.5 Passaging ES cells

When ES cells in most wells of the 96-well feeder plate reached 80-85% confluence, the plate was judged to be ready for passaging. Fresh media was added to each well 2-3 hours prior to passaging the cells. The media was then aspirated off, and the wells were washed once with PBS. Next, 50 µl of trypsin was added to each well of the 96-well plate and the plate was incubated at 37°C for 15 min. 150 µl of fresh M15 media was added to each well and the cells were dispersed by vigorous pipetting using a multi-channel pipette. The ES cell suspension was then evenly distributed to three 96-well gelatinised/feeder plates before incubating the plates at 37°C/5% CO_2.

2.3.6 Freezing ES cells

When the ES cells reached 80-85% confluence, they were ready for freezing. The media was changed 2 hours before adding trypsin to the plate. The media was then aspirated off and the plate was washed once with PBS. 50 µl of trypsin was added to each well of the 96-well plate and the plate was incubated at 37°C for 15 min. Next, 50 µl of 2X Freezing Media (60% DMEM, 20% DMSO, 20% FCS) was added to each well and the cells were dispersed by repeated pipetting using a multi-channel pipette. The plate was then sealed with tape and put into a polystyrene box with lid and frozen at -80°C.

2.3.7 Thawing ES cells

To thaw frozen ES cell clones, the 96-well plate was taken out of the -80^{o}C freezer and placed immediately into the 37^{o}C incubator. After ensuring that all the wells had thawed completely, the required clones were transferred to individual wells of a 24-well feeder plate containing 2 ml of M15 media. For maximum recovery of cells, another 200 µl of M15 media was added to rinse each well of the 96-well plate and the cell suspension was transferred into corresponding wells of the 24-well feeder plate. The plates were then incubated at 37^{o}C/5% CO_2.

2.3.8 Generation of targeted ES cell lines

Bcl11a-lacZ (AB2.2 targeted with *Bcl11a-lacZ* construct):

Twenty micrograms of *Bcl11a-lacZ* construct was linearized with *I-Ppo*I and electroporated into AB2.2 cells. The transfectants were selected with G418 and FIAU simultaneously for 8 days. 96 G418-resistant clones were picked and expanded on a 96-well feeder plate. Genomic DNA was extracted for long range PCR analysis. The expected sizes of the PCR products were 6.435 kb for 5' PCR and 3.446 kb for 3' PCR. The correctly targeted clones were expanded.

Bcl11b-lacZ (AB2.2 targeted with *Bcl11b-lacZ* construct):

Twenty micrograms of *Bcl11b-lacZ* construct was linearized with *I-Ppo*I and electroporated into AB2.2 cells. The transfectants were selected with G418 and FIAU simultaneously for 8 days. 96 G418-resistant clones were picked and expanded on a 96-well feeder plate. Genomic DNA was extracted for long range PCR analysis. The expected sizes of the PCR products were 7.141 kb for 5' PCR and 3.969 kb for 3' PCR. The correctly targeted clones were expanded.

2.4 Mouse techniques

2.4.1 Animal husbandry

Mice were treated in accordance with local ethical committee and the UK Animals (Scientific Procedures) Act, 1986 and all procedures were approved by the British Home Office Inspectorate. *Bcl11-lacZ* reporter lines were generated as detailed

above (Chapter 2.2). *Bcl11* conditional knockout mice were generated by Pentao Liu. Briefly, exon 1 and exon 4 of *Bcl11a* and *Bcl11b* respectively were flanked by *lox*P sites; upon expression of Cre recombinase, the intervening sequences between the *lox*P sites would be deleted, resulting in a null allele. The Rosa26-CreERT2 line, where Cre expression is driven by *Rosa26* promoter, was a kind gift from David Adams. In the Rosa26-CreERT2 line, Cre has been fused to the ligand-binding domain of a mutated human estrogen receptor (ERT) that recognizes tamoxifen or its derivative 4-hydroxytamoxifen. Therefore, Cre-estrogen receptor (CreERT) fusion proteins are retained in the cytoplasm but translocated to the nucleus on addition of the synthetic ligand. The BLG-Cre line, where the Cre transgene is driven by the promoter of the milk gene *β-lactoglobulin* (BLG), was obtained from Christine Watson (Selbert et al., 1998). Virgin female mice, 8 to 14 weeks old were mated and plug checked to confirm mating and pregnancy was confirmed post-mortem to avoid pseudo-pregnancies. At least three mice of each genotype and each time point were analysed.

2.4.2 Tamoxifen preparation and injection

One gram of tamoxifen-free base (Sigma) was re-suspended in 5 ml of ethanol and then dissolved in 50 ml of sunflower oil to obtain a final stock concentration of 20 mg/ml. This stock solution was then sonicated on ice for 2 min (at 15 sec interval) to ensure that the tamoxifen had completely dissolved. 1 mg of tamoxifen-free base in sunflower oil was administrated to each mouse by intraperitoneal injection for 3 consecutive days. Food mash was given to injected mice at the end of third injection and their weights were closely monitored.

2.5 DNA methods

2.5.1 Extraction of DNA from BAC clones

BAC clones were inoculated in 5 ml of LB with Cm at 37^{o}C with shaking at 244 G. The overnight culture was spun down at 16,000 G for 5 min and the pellet was washed with 1 ml of STE (100 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 8). The pellet was then re-suspended in 250 µl of P1 buffer (Qiagen) and clones were lysed with an equal volume of P2 buffer (Qiagen). Next, 350 µl of N3 (Qiagen) was added and the cell lysates were

spun at 16,000 G for 5 min. The supernatant was transferred to a fresh eppendoff tube and DNA was precipitated with 0.7 volumes of isopropanol. The DNA was pelleted by spinning at 16,000 G for 10 min and washed once with 1 ml of 70% ethanol. Finally, the DNA was air dried before re-suspending in an appropriate volume of TE buffer.

2.5.2 Extraction of DNA from ES cells

ES cells were incubated in 50 µl of lysis buffer [10 mM Tris (pH 7.5), 10 mM EDTA (pH 8), 10 mM NaCl, 0.5% Sarcosyl, and 1 mg/mL Proteinase K (added fresh)] at 65^{o}C overnight in a humidified chamber. Next, 100 µl of freshly prepared solution of 75 mM NaCl in ethanol was added and the suspension was incubated at room temperature for 30 min. The cell suspension was then spun at 3,000 G for 5 min and washed twice with 70% ethanol. The DNA was air dried before re-suspended in an appropriate volume of TE buffer.

2.5.3 Extraction of DNA from tissues

Tissues (ear or tail biopsies) were incubated in 600 µl of lysis buffer [50 mM Tris (pH 8), 100 mM NaCl, 25 mM EDTA (pH 8), 0.5% SDS, and 0.5 mg/ml Proteinase K (added fresh)] at 65^{o}C for 2 hours. 500 µl of buffered phenol/chloroform/isoamyl alcohol (25:24:1 parts by volume) was then added and the suspension was mixed well before centrifuging at 16,000 G for 5 min. The aqueous phase was then carefully removed and added to 420 µl of isopropanol to precipitate the DNA. The DNA pellet was washed once with 1 ml of 70% ethanol and air dried before being re-suspended in an appropriate volume of TE buffer.

2.5.4 Polymerase chain reaction (PCR)

2.5.4.1 Long range PCR

PCR amplification was carried out using Extensor Hi-Fidelity PCR Master Mix 2 (2X, ABgene). 25 µl of the master mix was added to 1 µl of template (1-5 ng), 2 µl of each primer (10 µM) and 20 µl of PCR-grade water. PCR was performed using PTC-225 PCR machine (Peltier Thermal Cycler) with the following settings: 94^{o}C for 4 min, this was followed by 35 cycles of 94^{o}C for 30 sec, 60^{o}C for 30 sec and 68^{o}C for 5 min (5’

PCR) or 3 min (3' PCR). This was then followed by 68^{o}C for 10 min. The PCR reaction was mixed with 6X loading dye and visualised on a 0.8% agarose gel. Primers used are shown in Table 2.3.

2.5.4.2 Genotyping PCR

PCR amplification was carried out using Extensor Hi-Fidelity PCR Master Mix 2 (2X, ABgene). 5 µl of the master mix was added to 1 µl of template (1-5 ng), 1 µl of each primer (10 µM) and 2 µl of PCR-grade water. PCR was performed using PTC-225 PCR machine (Peltier Thermal Cycler) with the following settings: 94^{o}C for 4 min, this was followed by 35 cycles of 94^{o}C for 30 sec, 60^{o}C for 30 sec and 68^{o}C for 45 sec. This was then followed by 68^{o}C for 10 min. The PCR reaction was mixed with 6X loading dye and visualised on a 1% agarose gel. Primers used are shown in Table 2.3.

Table 2.3. List of genotyping primers used.

Long range PCR primers	Primer sequence (5' – 3')	Size (bp)
Bcl11a-5'-Fwd	GAGACAAGAACAGGTGCAAGAGTGGATT	6,435
Bcl11a-5'-Rev	CAAGGAAACCCTGGACTACTGCGCCCTA	
Bcl11a-3'-Fwd	GAAAGAACCAGCTGGGGCTCGACTAGAG	3,446
Bcl11a-3'-Rev	CAGCGAGGTCCCCTTTCTCACTAAAAAT	
Bcl11b-5'-Fwd	GCTTAACTGGACACCATGTGCATACCCC	7,141
Bcl11b-5'-Rev	CAAGGAAACCCTGGACTACTGCGCCCTA	
Bcl11b-3'-Fwd	GAAAGAACCAGCTGGGGCTCGACTAGAG	3,969
Bcl11b-3'-Rev	CTGTAAGAGACAAACCCTCTAGGCAAAC	
Genotyping PCR primers	**Primer sequence (5' – 3')**	**Size (bp)**
Bcl11a-lacZ-Fwd	GCTTGCTTTGGAATATGAATGTTTG	363 (Wt); 437 (Mut)
Bcl11a-lacZ-Rev1	CATATATGGGGTTTATGGAGTAACC	
Bcl11a-lacZ-Rev2	CAAGGAAACCCTGGACTACTGCGCCCTA	
Bcl11a-Wt-Fwd	CCATGACGGCTCTCCCACAAT	248
Bcl11a-Wt-Rev	GCGAGAATTCCCGTTTGCTT	
Bcl11a-Mut-Fwd	ACGAGTTCTTCTGAGGGGATC	420
Bcl11a-Mut-Rev	ATCCCGACTCCAGACTGGGAC	
Bcl11a-cko-Fwd	TAGCTCCTGCTAGCCAGGTTTCTT	377 (Wt);
Bcl11a-cko-Rev	CGAGGCTTGCAGAAACAGAAAGAT	470 (Cko);
Bcl11a-cko-Del	CTCGAAGGGAGGTTTCGGTATTGTG	700 (Del)
Bcl11b-lacZ-Fwd	GACACCAAACCGCTCTTTTAGACTA	644 (Wt); 527 (Mut)
Bcl11b-lacZ-Rev1	CATTGTGTAAGAGACAGGGTATAAG	
Bcl11b-lacZ-Rev2	CAAGGAAACCCTGGACTACTGCGCCCTA	
Bcl11b-Wt-Fwd	GCGTTTTCATCATTTACTTTGTCC	500
Bcl11b-Wt-Rev	TGGAAAAGATTCTCGGGGTCC	
Bcl11b-Mut-Fwd	GCCATTGTCACTGTCTCCCATC	500
Bcl11b-Mut-Rev	CATTGTTAGTCTCTTCTTTCCCGC	
Bcl11b-cko-Fwd	TGAGTCAATAAACCTGGGCGAC	243 (Wt);
Bcl11b-cko-Rev	GGAATCCTTGGAGTCACTTGTGC	345 (Cko);
Bcl11b-cko-Del	TCCTGGTAACACACAATTGC	450 (Del)
Rosa26-ERT-Cre-Fwd	TGTGGACAGAGGAGCCATAAC	500 (Wt); 300 (Mut)
Rosa26-ERT-Cre-Rev1	CATCACTCGTTGCATCGACC	
Rosa26-ERT-Cre-Rev2	AAGACCCAACCAACAGCAG	
BLG-Cre-Fwd	TCGTGCTTCTGAGCTCTGCAG	280
BLG-Cre-Rev	GCTTCTGGGGTCTACCAGGAA	
BAC Validation primers	**Primer sequence (5' – 3')**	**Size (bp)**
SpeI-Evi9-ext4-Ret-3'-2	TCTACTAGTCTCACCACCTGTACAGTAAGT	618
H3-Evi9-ext4-Ret-3'-1	GTCAAGCTTTGAGCCAGATCGAGCTAGGTC	
NotI-Ctip2-Ret-5'-1	ATATGCGGCCGCTTCCAGTAAGATGTGTCT	371
H3-Ctip2-Ret-5'-2	ATGTAAGCTTCTGGCTCACTTCGCGTGGTTA	
Bsd/Neo Validation primers	**Primer sequence (5' – 3')**	**Size (bp)**
Bcl11a-Bsd-PCR-F	TCTAGCCTCACATAGGGGAGAAAGT	420 (Wt)
Bcl11a-Bsd-PCR-R	CAGTTCAGCTAGTCAACTCTTTGGC	947 (Bsd); 7,251 (lacZ)
Bcl11a-Neo-PCR-F	GGCTGTTTCGCCTAAAGTTTTATTT	454 (Wt)
Bcl11a-Neo-PCR-R	CACAGACATAGATACTGCTGCTCTT	2,406 (Mut)
Bcl11b-Bsd-PCR-F	GCATTCCTGGGGACTGGGCTTGAGT	437 (Wt)
Bcl11b-Bsd-PCR-R	CGTCCCCTGTAGGTGTTCTTATTTT	964 (Bsd); 7,268 (lacZ)
Bcl11b-Neo-PCR-F	CAGAGGCGTCTCTGTGTGGAATTCA	498 (Wt)
Bcl11b-Neo-PCR-R	CCTTTGCCCTACAGTGCCCATCAGC	2,450 (Mut)

2.5.5 Transfection

Transfection was carried out using Lipofectamine LTX reagent (Invitrogen) according to manufacturer's protocol. Briefly, one day prior to transfection, KIM-2 cells were plated in 500 µl of growth medium so that cells would be 50-80% confluent on the day of transfection. 500 ng of plasmid DNA (MSCV-egfp; MSCV-Bcl11a-egfp and MSCV-Bcl11b-egfp) was diluted in 100 µl of Opti-MEM I reduced serum medium without serum and 0.5 µl of PLUS reagent (Invitrogen) was added and incubated at room temperature for 5 min. 1.25 µl of Lipofectamine LTX reagent was added directly to the DNA and the mixture was incubated at room temperature for 30 min. 100 µl of the DNA-Lipofectamine LTX complexes was added to each well of KIM-2 cells and cells were incubated at 37^{o}C/5% CO2 for 48 hours. GFP positive cells were then sorted out and RNA was extracted using PicoPure kit (as detailed below in Chapter 2.6.1).

2.6 RNA methods

2.6.1 Extraction of total RNA from cells

RNA from sorted cells was extracted using the PicoPure RNA extraction kit according to manufacturer instructions (Arcturus). Cells were pelleted by centrifugation at 5,000 G for 5 min and the supernatant was removed carefully. 100 µl of Extraction Buffer was added and cell mixture was incubated at 42^{o}C for 30 min. After centrifugation at 3,000 G for 2 min, the supernatant was removed and transferred to RNAse-free tubes. RNA purification columns were pre-conditioned by adding 250 µl of Conditioning Buffer to the membrane and incubating the columns at room temperature for 5 min before centrifugation at 16,000 G for 1 min. 100 µl of 70% ethanol was added into the cell extract and mixed well by pipetting up and down. The cell extracts were then pipetted onto the pre-conditioned RNA purification columns and centrifuged at 100 G to bind the RNA to the membrane. This was immediately followed by centrifugation at 16,000 G for 30 sec to remove the flow through. 50 µl of Wash Buffer 1 (W1) was pipetted into the purification columns and columns were spun at 8,000 G for 1 min. Next, 2.5 µl of DNase I stock solution in 10 µl of RDD buffer (Qiagen) was added to the purification column membrane and the columns were incubated at room temperature for 15 min. 50 µl of W1 was then added to the columns and columns were spun at 8,000 G for 1 min. The

columns were washed with 100 µl of Wash Buffer 2 before elution with 22 µl of Elution Buffer.

2.6.2 Extraction of total RNA from tissues

Frozen tissues were ground in a mortar using a pestle under liquid nitrogen and the powder was transferred into RNAse-free tubes. 1 ml of Trizol (Invitrogen) was added to the homogenized tissues and the cell suspension was incubated at room temperature for 5 min. 200 µl of chloroform was added to separate out the phenol phase (includes the genomic DNA). After centrifugation at 12,000 G for 15 min at 4^{o}C, the aqueous phase was removed and transferred to RNAse-free tubes. RNA was then precipitated by adding 250 µl of isopropyl alcohol and centrifugation at 12,000 G for 10 min at 4^{o}C. The RNA pellet was washed once with 1 ml of 75% ethanol with centrifugation at 7,400 G and air-dried before being re-suspended in an appropriate volume of RNAse-free water.

2.6.3 First strand cDNA synthesis

RNA quality and quantity was verified using a Nanodrop ND-100 Spectrophotometer (Thermo Scientific). Typically, 1-2 µg of RNA was made up to 11 µl with dd$H_2$0 and incubated with 1 µl of oligo-dT at 70^{o}C for 10 min and then placed on ice. 4 µl of first strand reverse transcription buffer, 2 µl of 0.1 M DTT and 1 µl of dNTPs were added and the mixture was incubated at 42^{o}C for 2 min. 1 µl of Superscript II Reverse Transcriptase was then added and the mixture was incubated at 42^{o}C for 1 hour. 1 µl of RNase H (Roche) was added and the mixture was incubated at 37^{o}C for 30 min to complete the first strand synthesis reaction. The samples were diluted using dd$H_2$0 and a PCR using the β-actin/Cyclophilin A primers was always performed to check for cDNA successful synthesis and concentration. All reagents were purchased from Invitrogen unless otherwise stated.

2.6.4 Semi-quantitative Real Time PCR

To investigate changes in expression of certain genes, semi-quantitative Reverse Transcription PCR (RT-PCR) was performed on cDNA using the primers listed in Table 2.4. PCR amplification was carried out using Extensor Hi-Fidelity PCR Master Mix 2

(2X, ABgene). 5 µl of the master mix was added to 1 µl of template (1-5 ng), 1 µl of each primer (10 µM) and 2 µl of PCR-grade water. PCR was performed using PTC-225 PCR machine (Peltier Thermal Cycler) with the following settings: 94^{o}C for 4 min, this was followed by 35 cycles of 94^{o}C for 30 sec, annealing temperature (See table 2.4) for 30 sec and 68^{o}C for 45 sec. This was then followed by 68^{o}C for 10 min. The PCR reaction was mixed with 6X loading dye and visualised on a 1.5% agarose gel.

Table 2.4. List of RT-PCR primers used.

RT-PCR primers	Primer Sequence (5'-3')	Annealing Temperature (^{o}C)	Cycles
α-casein-Fwd	CATCATCCAAGACTGAGCCAG	60	25
α-casein-Rev	CCTGTGGAAAGTAAGCCCAAAG		
β-casein-Fwd	GGCACAGGTTGTTCAGGCTT	60	25
β-casein-Rev	AAGGAAGGGTGCTACTTGCTG		
Whey acidic protein-Fwd	GCAGCATTTCATGTTGCCA	60	25
Whey acidic protein-Rev	TCGTTCTTGGCCTGCTGG		
α–lactalbumin-Fwd	TCGTTCCTTTGTTCCTGGTG	60	25
α-lactalbumin-Rev	GGGCTTGTAGGCTTTCCAGT		
Prolactin-Fwd	GTGGCCCCAATTCCTGTTTCTTTA	58	30
Prolactin-Rev	ATTTCTCCTGGCCCCATCTACTCC		
Prolactin-Receptor-Fwd	CACACGCGCAGATCTCCTTACC	58	30
Prolactin-Receptor-Rev	CGCTGTGTTCTGGGCTCGTG		
Gata3-Fwd	TGGGTGGGGCCTCATCCTCAG	60	30
Gata3-Rev	ACCGGGTCCCCATTAGCGTTCCT		
Stat6-Fwd	CTCTGTGGGGCCTAATTTCCA	60	30
Stat6-Rev	CATCTGAACCGACCAGGAACT		
Stat5-Fwd	CGCCAGATGCAAGTGTTGTAT	58	30
Stat5-Rev	TCCTGGGGATTATCCAAGTCAAT		
Stat3-Fwd	CAATACCATTGACCTGCCGAT	58	30
Stat3-Rev	GAGCGACTCAAACTGCCCT		
Akt1-Fwd	ATGAACGACGTAGCCATTGTG	61	30
Akt1-Rev	TTGTAGCCAATAAAGGTGCCAT		
Akt2-Fwd	ACGTGGTGAATACATCAAGACC	61	30
Akt2-Rev	GCTACAGAGAAATTGTTCAGGGG		
Notch1-Fwd	CAGCATGGCCAGCTCTGGTT	58	30
Notch1-Rev	AGCAGCATCCACATTGTTCA		
Notch2-Fwd	CATTGACGAGTGCACTGAGAGCTCCTGC	61	30
Notch2-Rev	AGTGCTGGCACAAGTGTTCAACAGGT		
Notch3-Fwd	ACACTGGGAGTTCTCTGTGAGATCA	60	30
Notch3-Rev	GTCTGCTGGCATGGGATACCCACTG		
Notch4-Fwd	CAGAACGTGGATCCCCTCAAGTTGCCTG	61	30
Notch4-Rev	GGCAGAGAGAGGGCAAGGAGTCATCAGC		
Jagged1-Fwd	CTGAGTCTTCTGCTCGCCCT	60	30
Jagged1-Rev	CGGCTAGGGTTTATCATGCC		
Jagged2-Fwd	CTGTGCAGCGTGTTCAGTG	61	35
Jagged2-Rev	GTGTCCACCATACGCAGATAAC		
Delta-like-ligand1-For	ACCGCAACTGCTGCCGCGGG	61	30
Delta-like-ligand1-Rev	GGCCGCTACTGTGAAGGTCC		
Delta-like-ligand3-For	CTGGTGTCTTCGAGCTACAAAT	61	35
Delta-like-ligand3-Rev	TGCTCCGTATAGACCGGGAC		
Delta-like-ligand4-For	TTCCAGGCAACCTTCTCCGA	62	35
Delta-like-ligand4-Rev	ACTGCCGCTATTCTTGTCCC		

Hes1-Fwd	ATAGCTCCCGGCATTCCAAG	61	30
Hes1-Rev	GCGCGGTATTTCCCCAACA		
Hes2-Fwd	CTGAAGGGTCTCGTATTGCCG	62	35
Hes2-Rev	CGCAGGTGCTCTAGTAGGC		
Hes3-Fwd	GCACGCATCAACGTGTCAC	61	35
Hes3-Rev	TGAGTTCTGGAGGCTTCTCAT		
Hes5-Fwd	AGTCCCAAGGAGAAAAACCGA	61	35
Hes5-Rev	GCTGTGTTTCAGGTAGCTGAC		
Hes6-Fwd	ACCACCTGCTAGAATCCATGC	61	30
Hes6-Rev	GCACCCGGTTTAGTTCAGC		
Hes7-Fwd	CGGGAGCGAGCTGAGAATAG	61	35
Hes7-Rev	CACGGCGAACTCCAGTATCT		
Elf5-For	ATGTTGGACTCCGTAACCCAT	61	20
Elf5-Rev	GCAGGGTAGTAGTCTTCATTGCT		
p63-Fwd	AGATCCCTGAACAGTTCCGAC	60	30
p63-Rev	CGACGAGAATCCATGTCAAAGTT		
NKCC1-Fwd	GGTTCTCCAAACTCACGG	60	25
NKCC1-Rev	GTCTTGCCATCCTCTTCCTC		
Muc1-Fwd	GGCATTCGGGCTCCTTTCTT	60	25
Muc1-Rev	TGGAGTGGTAGTCGATGCTAAG		
CK18-Fwd	CAGCCAGCGTCTATGCAGG	58	25
CK18-Rev	CTTTCTCGGTCTGGATTCCAC		
CK14-Fwd	TGAGAGCCTCAAGGAGGAGC	58	30
CK14-Rev	TCTCCACATTGACGTCTCCAC		
Smooth muscle actin-α-Fwd	AGATTGTCCGTGACATCAAGG	60	25
Smooth muscle actin-α-Rev	TTGTGTGCTAGAGGCAGAGC		
LIF-Fwd	GCTATGTGCGCCTAACATGAC	63	35
LIF-Rev	CGCTCAGGTATGCGACCAT		
ERα-Fwd	CCTCCCGCCTTCTACAGGT	63	30
ERα-Rev	CACACGGCACAGTAGCGAG		
Bcl11a-Fwd	TGGTATCCCTTCAGGACTAGGT	61	35
Bcl11a-Rev	TCCAAGTGATGTCTCGGTGGT		
Bcl11b-Fwd	CCCGACCCTGATCTACTCAC	61	35
Bcl11b-Rev	CTCCTGCTTGGGACAGATGCC		
Cyclophilin A-Fwd	CCTTGGGCCGCGTCTCCTT	55	25
Cyclophilin A-Rev	CACCCTGGCACATGAATGGTG		
β-actin-Fwd	GTGGGCCGCTCTAGGCACCAA	58	25
β-actin-Rev	CTCTTTGATGTCACGCACGA		
p85α-Fwd	GCCCCGTGCTTTTCAGATTTC	60	30
p85α-Rev	TCCTGCTGGTATTTGGACACTGGGTAG		
p55α-Fwd	GTTACAGTGCGGGCCGTATAGGTTTTA	60	30
p55α-Rev	TCCTGCTGGTATTTGGACACTGGGTAG		
p50α-Fwd	CTGGCAGTTCAAAGCGAAACCGT	60	30
p55α-Rev	TCCTGCTGGTATTTGGACACTGGGTAG		

2.6.5 Quantitative Real Time PCR

2.6.5.1 TaqMan

Quantitative real-time TaqMan PCR was used for detection of deletion efficiency in genomic DNA following Cre recombinase expression. The PCR was carried out with the addition of 12.5 µl of ABsolute QPCR mix (Thermo) with 1 µl of each primer (10 µM) and 0.5 µl of probes (0.25 µM; MWG, Ebersberg, Germany). The real-time PCR reactions were run in an ABI PRISM 7900HT (Applied Biosystems) in triplicate. Probes

were labelled with reported dye (FAM) at the 5' end and the quencher dye TAMRA at the 3' end. Primers and probes used are listed in Table 2.5.

2.6.5.2 SYBR Green

Real time PCR was performed to quantify the relative amounts of mRNA expression of gene of interest against *Cyclophilin* A as an internal control. Bio-Rad iCycler PCR (Bio-Rad) machine with a Halogen Tungsten lamp were used and data was analysed using iCycler iQ software (Bio-Rad). The PCR was performed using iQ Supermix (Bio-Rad) which contains the Taq enzyme, buffer and dNTPs. SYBR Green I (Bio-Rad) DNA fluorescent dye was used with high affinity to dsDNA and an excitation wavelength of 488 nm. Primers used are listed in Table 2.5

Table 2.5. List of qRT-PCR primers used.

qRT-PCR (TaqMan)	Primer Sequence (5'-3')
Bcl11a-Fwd	AAAGGCACTGATGAAGATATTTTCTCT
Bcl11a-Rev	CGACGGCTCGGTTCACAT
Bcl11a-Probe	TCTCCTTCTTTCTAACCCGGCTCTCCC
Bcl11b-Fwd	AGTGCCTTCGACCGAGTCAT
Bcl11b-Rev	GAAGTCCATGGCAGGAGAGTCT
Bcl11b-Probe	CGCCTGAACCCCATGGCCA
β-actin-Fwd	TTCAACACCCCAGCCATGTA
β-actin-Rev	TGTGGTACGACCAGAGGCATAC
β-actin-Probe	TAGCCATCCAGGCTGTGCTGTCCC
qRT-PCR (SYBR-Green)	**Primer Sequence (5'-3')**
Bcl11a-Fwd	TGGTATCCCTTCAGGACTAGGT
Bcl11a-Rev	TCCAAGTGATGTCTCGGTGGT
Bcl11b-Fwd	CCCGACCCTGATCTACTCAC
Bcl11b-Rev	CTCCTGCTTGGGACAGATGCC

2.6.6 *In situ* hybridization

2.6.6.1 Cloning and synthesis of probes

Primers were used to PCR amplify probes from a mouse embryo 10.5 days post-coitum total cDNA library constructed as detailed in Chapter 2.6.3. PCR products were then cloned into either pCR-II-TOPO or pCR-Blunt-II-TOPO vectors (Invitrogen) and transformed into TOP10 cells (Invitrogen). Transformants were selected on LB agar containing either Amp (pCR-II-TOPO) or Kan (pCR-Blunt-II-TOPO). 8 clones were picked for each probe and verified using restriction digestion and sequencing to

determine the orientation of the insert in the vector backbone. Appropriate enzymes were then chosen to linearize the plasmids to enable synthesis of the antisense strand by *in vitro* transcription. 2 µg of plasmid was digested with 1 µl of appropriate enzyme with 2 µl of buffer (NEB) in a total volume of 20 µl. Linearized plasmids were then purified using a mini-preparation spin column (Qiagen) and eluted in 20 µl of ddH_2O. 1 µg of linearized probe was then incubated at 37°C for 2 hours with 2 µl of transcription buffer (10X), 2 µl of NTP labelling mixture, 1 µl of RNase inhibitor, and 2 µl of SP6 or T7 RNA polymerase. After *in vitro* synthesis, 2 µl of DNase I was added and the mixture was incubated at 37°C for 15 min before purifying the probes using ProbeQuant G-50 Micro columns (GE Healthcare). Reagents were obtained from Roche unless otherwise stated.

Table 2.6. List of primers used for cloning *in situ* probes.

In situ Probe	Primer sequence (5' – 3')	Size (bp)	Vector	Linearization Enzyme	Polymerase
Fgf10-probe-Fwd	AATGTGGAAATGGATACTGACACAT	631	pCRII-TOPO	*Not* I	SP6
Fgf10-probe-Rev	CTATGTTTGGATCGTCATGGGGAGG				
Fgfr2TK-probe-Fwd	ACAGTTCTGCCAGCGCCTGTGAGAG	949	pCRII-TOPO	*Not* I	SP6
Fgfr2TK-probe-Rev	CTTCAGGAGCCATCCACTTGACTGG				
Lef1-probe-Fwd	TCATCTTTGGTTAACGAGTCCGAAA	359	pCR-Blunt II-TOPO	*Hind* III	T7
Lef1-probe-Rev	ATGTGTGACGGGTGGGATCCCGGAG				
Notch1-probe-Fwd	CAGCATGGCCAGCTCTGGTT	846	pCR-Blunt II-TOPO	*Xho* I	SP6
Notch1-probe-Rev	AGCAGCATCCACATTGTTCA				
Wnt10b-probe-Fwd	GGAGGGCAGCGCCAGAGTTCC	257	pCR-Blunt II-TOPO	*Bam* HI	T7
Wnt10b-probe-Rev	AGGCTGCCACAGCCATCCAAGAGG				
Tbx3-probe-Fwd	GAGATGGTCATCACGAAGTC	567	pCRII-TOPO	*Not* I	SP6
Tbx3-probe-Rev	CTGCAATGCCCAATGTCTCG				

2.6.6.2 Whole mount hybridization

Embryos were dissected from time mated females and staged before fixing them in 4% paraformaldehyde (PFA) at 4°C for 48 hours. The embryos were then washed sequentially in PBST (PBS with 0.1% Tween-20), 25%, 50%, 75% Methanol (MeOH)/PBST and 100% MeOH for 5 min each at room temperature (with shaking). Embryos were rehydrated by sequential washes in 75%, 50%, 25% MeOH/PBST and PBST for 5 min at room temperature (with shaking). Next, embryos were bleached using 6% hydrogen peroxide (in PBST) for 1 hour at room temperature, washed thrice with PBST before treatment with 10 µg/ml of Proteinase K (in PBST) for 15 min at room temperature. Subsequently, embryos were washed thrice with PBST, post-fixed with 4%

PFA/0.2% Glutaldehyde/PBST for 20 min at room temperature and washed twice with PBST prior to incubating them in hybridization solution (50% formamide, 5X SSC, 1% SDS, 50 µg/ml tRNA, 50 µg/ml Heparin) for 1 hour at 70°C. Finally, embryos were incubated overnight at 70°C in preheated hybridization solution containing 1 µg/ml of antisense riboprobe. On the second day, the embryos were washed twice with preheated solution 1 (50% formamide, 5X SSC, 1% SDS) at 70°C for 30 min each, thrice with preheated solution 2 (50% formamide, 2X SSC) at 65°C for 30 min each and thrice with TBST (1X; 8 g of NaCl, 25 ml of Tris-HCl, pH 7.5, 0.1% of Tween-20). The embryos were then pre-blocked with 10% sheep serum (in TBST) for 90 min before incubating with anti-DIG antibody (Roche) overnight at 4°C. On the third day, the embryos were washed five times with TBST (with 50 mg/ml of Levamisole, Sigma) for 1 hour each (with shaking) at room temperature. Embryos were left overnight in TBST at room temperature to improve background. To develop the staining, embryos were incubated in NBT/BCIP solution (Roche) at room temperature in the dark until desired reaction solution was achieved. Reaction was stopped by washing the embryos in TE (pH 8) and fixing in 4% PFA for 1 hour at 4°C. Embryos were then imaged using a LEICA MZ75 light microscope.

2.7 Protein methods

2.7.1 Protein extraction

Frozen tissues were ground in a mortar using a pestle under liquid nitrogen and the powder was transferred into a 1.5 ml tubes. Tissue powder was re-suspended in 100-300 µl of RIPA buffer [50 mM Tris (pH 7.4), 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM EDTA, 1X Cocktail protease inhibitors (Roche), 1 mM Na_3VO_4, and 1 mM NaF] and transferred to a 1.5 ml tube. The lysate was mechanically ground using a micropestle (Eppendorf) and then passed through a syringe 5-8 times and left on ice for 40 min. Lysate was spun for 15 min at 16,000 G using a bench-top centrifuge. The supernatant was collected and frozen. Also, 10 µl of the supernatant was collected in a separate tube for BCA protein concentration assay.

2.7.2 BioCinchominic Acid (BCA) protein concentration assay (Pierce)

Bovine Serum Albumin (BSA) (Sigma) was made at required concentrations of 0.1 mg, 0.2 mg, 0.5 mg, 1.0 mg and 2.0 mg which were used as standards for the assay. The BCA reagents were mixed according to manufacturer's protocol in a ratio of 50:1 (A:B) respectively. 200 ul of the reagent mix was added to 5 ul of standard or sample in a 96 well plate. Each standard or sample was measured in triplicate. The sample were diluted in water prior to usage in the assay, the dilution factor depended on how concentrated the sample was. The range of dilution was between 1:4 and 1:16. This step is crucial for accuracy of the results. The reaction was incubated at room temperature for 20 min before the light absorbance was measured at a wavelength of 570 nm using a DYNATECH-MR5000 plate reader. The read out from the standards less the background, was used to plot a graph and the slope of which was used to calculate the concentration of the samples.

2.7.3 SDS-PAGE

The protein samples were made up to a final concentration of 30 µg using RIPA buffer and Loading Buffer [125 mM Tris-HCl (pH 6.8), 2.5% SDS, 20% glycerol, 0.002% Bromophenol blue and 5% β-mercaptoethanol]. The samples were boiled at 95^{o}C for 5 min and then were kept on ice. SDS-PAGE Criterion precast Tris-HCl resolving gels (Biorad) were used. Samples were loaded and gels were run at 100 V for 90-120 min at room temperature using Biorad powepac 300. 1X Running buffer (3.03 g Tris, 14.4 g Glycine, 1 g SDS and ddH_2O to 1 L).

2.7.4 Immunoblotting

Proteins were immunoblotted onto PVDF membranes (Millipore-immobilon FL) using the Biorad wet transfer tanks. The gels were placed in a cassette facing the PVDF membrane and electric current was applied at 400 mA for 1 hour. The transfer buffer used varied depending on the size of protein of interest. For high molecular weight proteins, low percentage methanol was used while for low molecular weight proteins

higher percentage of methanol was used. 1X Transfer buffer (192 mM Glycine, 25 mM Tris, 0.1% SDS, 5-20% v/v 100% Methanol, pH 8.3).

2.7.5 Primary antibody incubation

The PVDF membranes were incubated in blocking solution for 1 hour at room temperature to reduce background signal in the consequent steps. The blocking buffer and dilution differs between antibodies used are summarised in Table 2.7. All primary antibodies were incubated overnight at 4^{o}C whilst shaking.

Table 2.7. List of antibodies used for Western blot analysis.

Antibody	Clone	Source	Dilution	Blocking Buffer	Species
Bcl11a	BL1796	Bethyl Laboratories	1:1000	5% Milk	Rabbit
Stat5	9352	Cell Signalling	1:1000	5% Milk	Mouse
pStat5A/B (phospho Tyr)	9351	Cell Signalling	1:1000	5% Milk	Rabbit
Stat3	9139	Cell Signalling	1:1000	5% Milk	Mouse
pStat3 (phospho Tyr)	9131	Cell Signalling	1:1000	5% Milk	Rabbit
Stat6	M-20	Santa Cruz	1:1000	5% Milk	Rabbit
pStat6 (phospho Tyr)	177C322	Abcam	1:8000	5% BSA	Mouse
Gata-3	H-48	Santa Cruz	1:1000	5% Milk	Rabbit
Notch1	C-20	Santa Cruz	1:1000	5% Milk	Rabbit
Notch3	M-134	Santa Cruz	1:1000	5% Milk	Rabbit
Jagged1	H114	Santa Cruz	1:1000	5% Milk	Rabbit
Pan-p85	N-SH2	Millipore	1:1000	5% Milk	Rabbit
Akt	9272	Cell Signalling	1:1000	5% Milk	Rabbit
pAkt (phosphor Ser 473)	4058	Cell Signalling	1:1000	5% Milk	Rabbit
WAP	M-16	Santa Cruz	1:1000	5% BSA	Goat
β-casein	-	Gift from Bertr Binas	1:1000	5% BSA	Goat
Cleaved caspase 3	9664	Cell Signalling	1:1000	5% Milk	Rabbit
β-actin	N-21	Santa Cruz	1:1000	5% Milk	Rabbit

2.7.6 Secondary antibody incubation and detection

The primary antibodies were removed and the membranes were rinsed once with 1X PBS-0.1%Tween (Sigma) (PBST). The appropriate secondary antibody was diluted to a final concentration of 1:2000 in the blocking buffer and incubated with the membrane for 45 min. The secondary antibodies used were all conjugated with horseradish peroxidase (HRP) (DAKO). The membrane was then washed thrice with 1X PBST for 5 min each (or TBST in case of pAKT and pStat6) to get rid of non-specific binding. ECL or ECL-plus (Amersham) was used according to the manufacture's instructions as a substrate for the HRP enzyme. The chemi-luminescence was detected using Hyper-film

(Amersham). All western blots were scanned and edited using Jasc Paint Shop Pro software.

2.7.7 Stripping PVDF membranes

To allow the probing of a single membrane with multiple antibodies, the previous primary antibody has to be striped off. The membrane was incubated with 25 ml of strip buffer and 175 µl of 2-Mercaptoethanol for 1 hour at 50°C. The membrane was then washed thoroughly using ddH_2O, appropriate blocking buffer applied and incubated with primary antibody overnight. Strip buffer (40 g SDS, 15 g Tris, ddH_2O to 2 L, pH 6.8).

2.8 Whole mount X-gal staining

2.8.1 Embryo

Embryos were removed and dissected with the yolk sac and placenta intact before fixation with 4% PFA overnight at 4°C with shaking. Fixation time (7.5-11.5 dpc: 15 min; 12.5-14.5 dpc: 30 min; 15.5-17.5 dpc: 60 min). Embryos were then washed thrice with ice-cold PBS (pH 8) for 20 min each at 4°C with shaking before staining in X-gal staining solution [2mM $MgCl_2.6H_2O$, 0.01% deoxycholic acid, 0.02% IGEPAL CA-630, 5 mM potassium ferrocyanide ($K_4Fe(CN)_6.3H_2O$), 5 mM potassium ferricyanide ($K_3Fe(CN)_6$), 1 mg/ml X-gal in dimethylformamide, PBS pH 8] at 4°C with shaking for up to 48 hours. Stained embryos were then washed briefly in PBS (pH 8) before post-fixing in 4% PFA overnight at 4°C with shaking. Embryos were then cleared with 50% glycerol (in PBS, pH 7.4) overnight at 4°C with shaking and stored in 70% glycerol (in PBS, pH 7.4, 0.01% sodium azide) at room temperature for a week prior to imaging.

2.8.2 Adult tissues

Adult mice were anesthetized with isofluorane and a midline incision was made to visualize the heart and the great vessels. The abdominal aorta was cannulated with a 30-gauge needle coupled to PE-10 tubing to allow aortic retrograde perfusion of 4% PFA. Following perfusion, various tissues such as the brain, spinal cord and the internal organs were dissected out and fixed in 4% PFA for 1 hour. Tissues were then washed thrice with ice-cold PBS (pH 8) at 4°C with shaking before staining in X-gal staining solution [2mM

$MgCl_2.6H_2O$, 0.01% deoxycholic acid, 0.02% IGEPAL CA-630, 5 mM potassium ferrocyanide ($K_4Fe(CN)_6.3H_2O$), 5 mM potassium ferricyanide ($K_3Fe(CN)_6$), 1 mg/ml X-gal in dimethylformamide, PBS pH 8) at 4°C with shaking for up to 48 hours. Stained tissues were then washed briefly in PBS (pH 8) before post-fixing in 4% PFA overnight at 4°C with shaking. Tissues were then cleared with 50% glycerol (in PBS, pH 7.4) overnight at 4°C with shaking and stored in 70% glycerol (in PBS pH 7.4, 0.01% sodium azide) at room temperature for a week prior to imaging.

2.8.3 Mammary gland

Mammary glands were dissected out from female mice and spread onto a clean glass slide before fixing in 4% PFA for 1 hour at 4°C. Glands were then rinsed briefly in PBS (pH 8) and washed thrice with wash buffer (PBS with 2 mM $MgCl_2$) for 5 min each at room temperature with shaking. Glands were then incubated in permeabilization solution (PBS with 0.2% NP-40, 0.01% deoxycholic acid, 2 mM $MgCl_2$) for 1 hour at room temperature with shaking. Next, glands were incubated in X-gal mixer [25 mM potassium ferrocyanide ($K_4Fe(CN)_6.3H_2O$), 25 mM potassium ferricyanide ($K_3Fe(CN)_6$), 0.2% NP-40, 0.01% deoxycholic acid, 2 mM $MgCl_2$, PBS] for 1.5 hours at 37°C. 1 mg/ml of X-gal (Invitrogen) was then added to the X-gal mixer and glands were stained for 1-3 days at 37°C. After staining, glands were post-fixed in 4% PFA at 4°C overnight and then processed through sequential 1 hour washes of acetone, 70% ethanol, 95% ethanol, 100% ethanol followed by overnight clearing in xylene. Glands were then imaged prior to being embedded in paraffin for sectioning.

2.9 Histology and immunohistochemistry

2.9.1 Tissue carnoys fixative and whole mount carmine alum staining

Mammary glands were dissected out and spread out onto a glass slide and incubated in carnoys fixative (6 parts 100% ethanol, 3 parts chloroform, 1 part glacial acetic acid) overnight. The slide was then placed in carmine alum stain (1 g carmine, 2.5 g aluminium potassium sulphate and ddH_2O) overnight. The slide was then washed with

ethanol and cleared in xylene for 1 day. After photographic documentation the slide was stored in xylene.

2.9.2 Paraffin sections and Hematoxylin and Eosin (H&E) staining

Mammary glands were collected as described earlier. Glands were fixed in 4% PFA for 24 hours at 4^{o}C. The glands were then transferred into 70% ethanol and stored at -20^{o}C till sectioning. All tissues were embedded in wax and sectioned at 5 µm thickness on a glass slide. The sections were then de-waxed using three 5 min xylene washes and re-hydrated in a decreasing alcohol series and rinsed in ddH_2O before staining in Hematoxylin solution for 5 min and placed under running tap water till it turned blue. The slides were then placed in 1% aqueous Eosin for 1 min and then washed in running tap water.

2.9.3 Fluorescence immunohistochemistry

Paraffin embedded sections were de-waxed in three 5 min xylene washes and re-hydrated using 5 min washes in serial dilutions of ethanol/H_2O/PBS in the following order 100%, 100%, 95%, 95%, 50%, H_20, H_20, PBS, PBS (0.1% Triton X100) and PBS. The sections were then boiled in 10 mM Tri-Sodium Citrate Buffer (pH 6) using a steam cooker under pressure for 12 min. The sections were left in the cooling buffer for 20 min, washed once in PBS for 5 min and then dried before drawing wax barriers around the sections using a pap pen (Vector laboratories). Blocking solution (10% goat serum in PBS) was then applied on the tissue sections and incubated in a humidified chamber for 1 hour at room temperature. Following one rinse in PBS, the sections were then either incubated with primary antibody or blocking buffer overnight at 4^{o}C in a humidified chamber. Primary antibodies used are listed in Table 2.8. Following the overnight incubation with the primary antibody, the sections were washed thrice in PBS for 5 min each and incubated with either Alexa Fluor 488 (Invitrogen) or Cy3 (Sigma) conjugated secondary antibody diluted in blocking buffer and incubated in the humidified chamber in the dark at room temperature for 1 hour. The sections were then washed thrice with PBS for 5 min each and incubated with Bisbenzimide-Hoechest 33342 (Sigma) to stain the DNA for 5 min in the dark at room temperature. After a wash with PBS the sections were

mounted with 50% PBS-50% Glycerol solution and cover-slip was placed on top and sealed using clear nail polish (Worldwide glass resources Ltd). Pictures were taken immediately afterwards or within two days of mounting.

Table 2.8. List of antibodies used for immunohistochemistry.

Antibody	Clone	Source	Dilution	Species
pStat5A/B (phospho Tyr)	9351	Cell Signalling	1:100	Rabbit
pStat3 (phospho Tyr)	9131	Cell Signalling	1:100	Rabbit
Gata-3	HG3-31	Santa Cruz	1:100	Mouse
Notch1	C-20	Santa Cruz	1:100	Rabbit
Jagged1	H-114	Santa Cruz	1:100	Rabbit
β-casein	-	Gift from Bertr Binas	1:1000	Goat
Cleaved Caspase 3	9664	Cell Signalling	1:100	Rabbit
Aquaporin 5	178615	Calbiochem	1:50	Rabbit
Cytokeratin 14	LL002.	Abcam	1:100	Mouse
Cytokeratin 18	CY90	AbD Serotec	Single drop	Mouse
Smooth muscle actin	Ab-1	NeoMarkers	1:1000	Mouse
p63	4A4	Abcam	1:50	Mouse
Estrogen receptor-α	MC-20	Santa Cruz	1:50	Rabbit
Ki67	KLH	Abcam	1:50	Rabbit

2.10 FACS staining and analysis

2.10.1 Hematopoietic cells

8-12 weeks old adult mice were used. Bone marrow cells were flushed from the long bones (tibias and femurs) using a 26G syringe with PBS (without calcium or magnesium), supplemented with 1% heat-inactivated serum (GIBCO). Spleen or thymus were placed in PBS and gently homogenized with the end of a syringe. Cells were triturated and filtered through a 30 μm mesh to obtain single-cell suspension. Lysis of the red blood cells in ammonium chloride (NH_4Cl) was carried out before using cells for staining.

Prior to staining for hematopoietic stem cells in the bone marrow, depletion of mature hematopoietic cells such as T cells, B cells, monocytes/macrophages, granulocytes, and erythrocytes as well as their committed precursors from whole bone marrow cells using the Lineage cell depletion kit (Miltenyi Biotec). Cells were first stained with a cocktail of biotinylated antibodies against a panel of lineage antigens [CD5, CD45R (B220), CD11b, Gr-1 (Ly-6G/C), 7-4, and TER-119] and incubated on ice in the dark for 10 min. Anti-Biotin MicroBeads were then added and cells were incubated in the fridge at 4^{o}C for a further 15 min. 1 ml of MACS buffer (0.5% BSA, 5 mM EDTA

in PBS) was added and cells were spun at 375 G for 5 min. The supernatant was decanted and cells were re-suspended in 500 µl of MACS buffer and added to a pre-washed MS column (Miltenyi Biotec). The flow-through containing the lineage-negative cells was collected and used for staining with Sca1 and c-kit primary antibodies. Prior to staining for T cell progenitors in the thymus, CD4-positive and CD8-positive thymocytes were removed using CD4 (L3T4) and CD8 (Ly2) depletion kit (Miltenyi Biotec). Cells were first stained with biotinylated anti-CD4 and anti-CD8 antibodies and incubated on ice in the dark for 10 min. Anti-Biotin MicroBeads were then added and cells were incubated in the fridge at 4^{o}C for a further 15 min. 1 ml of MACS buffer was added and cells were spun at 375 G for 5 min. The supernatant was decanted and cells were re-suspended in 500 µl of MACS buffer and added to a pre-washed MS column (Miltenyi Biotec). The flow-through containing the CD4 and CD8-negative thymocytes was collected and used for staining with CD44 and CD25 primary antibodies.

Primary antibodies used are listed in Table 2.9. Cells were stained on ice in the dark for 20 min. 1 ml of PBS (with 1% FCS) was then added and cells were centrifuged at 375 G for 5 min. Supernatant was then decanted and cells were re-suspended in 500 µl of PBS (with 1% FCS) before analysis. FACS analysis was done using CyAN ADP (DakoCytomation) or MoFlo (DakoCytomation) and gates were set to exclude >99.9% of cells labelled with isoform-matched control antibodies conjugated with the corresponding fluorochromes. Three mice of each genotype were analyzed.

2.10.2 Mammary cells

Mammary glands from 4-14-week-old virgin female mice were dissected and mammary epithelial cell suspensions were prepared as previously described (Stingl et al., 2006). Briefly, dissected mammary glands were digested for 8 hours at 37^{o}C in DMEM:F12 (Gibco) with collagenase (300 U/ml) and hyaluronidase (100 U/ml). After vortexing and lysis of the red blood cells in NH_4Cl, epithelial cell clumps were first treated by gentle pipetting for 2 min in 0.25% trypsin, and then for 2 min in dispase II (5 mg/ml) plus DNase I (0.1 mg/ml; Sigma) followed by filtration through a 30 µm mesh. Primary antibodies used are listed in Table 2.9. Secondary antibody used: Strepavidin-PE-Texas Red (PE-TR, Molecular Probes). Apoptotic cells were excluded by elimination

of propidium iodide (PI) positive cells. All antibody staining was carried out on ice for 10 min in dark. All reagents were from StemCell Technologies Inc. unless otherwise specified. FACS analysis was done using CyAN ADP (DakoCytomation) and all sorts were performed using MoFlo (DakoCytomation) and gates were set to exclude >99.9% of cells labelled with isoform-matched control antibodies conjugated with the corresponding fluorochromes.

2.10.3 FDG staining

Fluorescein di-β-D-galactopyranoside (FDG; Sigma), a fluorescent substrate for β-galactosidase, was used for flow cytometric analysis to determine expression of *Bcl11* in hematopoietic and mammary epithelial cells. Hematopoietic or mammary cells were first stained with primary and/or secondary antibodies as detailed above (Chapter 2.10.1 and 2.10.2) before being used for FDG staining. Cell samples and FDG stock (2 mM in DMSO) were pre-warmed at 37°C for 5 min. Next, an equal volume of FDG stock was added and the cell mixture was incubated at 37°C for 1 min before adding 2 ml of HBSS and incubating on ice for 1 hour in the dark. For staining of *Bcl11b*$^{lacZ/+}$ thymocytes, phenylethyl-β-d-thiogalactopyranoside (PETG, 200 mM; Invitrogen), an inhibitor to β-galactosidase, was added to the cell mixture after 10 min.

Table 2.9. List of antibodies used for FACS analysis.

Antibody	Fluorescent dye	Clone	Source	Dilution
B220	PE	RA3-6B2	BD Biosciences	1 μg/ml
CD19	PE	1D3	BD Biosciences	1 μg/ml
Gr1	PE	RB6-8C5	BD Biosciences	1 μg/ml
Ter119	PE	TER-119	BD Biosciences	1 μg/ml
CD4	PE	GK1.5	BD Biosciences	1 μg/ml
CD8	PE-Cy5	53-6.7	BD Biosciences	1 μg/ml
CD44	PE	IM7	BD Biosciences	1 μg/ml
CD25	APC	PC61	BD Biosciences	1 μg/ml
Sca1	PE	D7	BD Biosciences	1 μg/ml
c-kit	APC	2B8	BD Biosciences	1 μg/ml
CD45	Biotin	30-F11	eBioscience	1 μg/ml
Ter119	Biotin	TER-119	eBioscience	1 μg/ml
CD31	Biotin	390	eBioscience	1 μg/ml
CD24	PE	M1/69	eBioscience	1 μg/ml
CD49f	AF647	GoH3	eBioscience	1 μg/ml
CD49b	FITC	HMa2	eBioscience	1 μg/ml
Sca1	AF647	D7	eBioscience	1 μg/ml

2.11 Mammary colony-forming-cell (Ma-CFC) assay

For Ma-CFC assays, mammary epithelial cells were prepared and stained with CD24, CD49f and FDG as described above (Chapter 2.10.2). Next, lineage negative $CD24^{+}CD49f^{+}$ epithelial cells were sorted into FDG^{-} (3,000 cells) and FDG^{+} (500 cells) fractions using MoFlo (DakoCytomation). The freshly-sorted cells were then plated onto irradiated feeders (10,000 feeders per ml of media) in NSA media and maintained for a week at $37^{o}C/5\%$ CO_2 as described previously (Stingl et al., 2006). To terminate the assays, the plates were washed once with PBS and fixed in ice-cold acetone:methanol (1:1) for 5 min. Plates were then washed once with water and stained with Giemsa (Merck) for 5 min before a final rinse with water. The number of colonies was enumerated under a dissecting microscope. Composition of NSA media: 9 parts NeuroCult NSA base medium (Human) (StemCell Tech); 1 part NeuroCult NSA supplements (StemCell Tech); 1:100 dilution of N2 supplements (Gibco); 10 ng/ml basic fibroblast growth factor (Sigma); 10 ng/ml epidermal growth factor (Sigma); 10 ng/ml insulin (Sigma) and 5% FBS (StemCell Tech).

CHAPTER 3:
GENERATION OF THE *BCL11-LACZ* CONDITIONAL NULL REPORTER MICE

3.1 Introduction

3.1.1 Current *Bcl11* knockout mice

The original *Bcl11a* and *Bcl11b* knockout mice were reported by Pentao Liu (Liu et al., 2003b) and Yuichi Wakabayashi (Wakabayashi et al., 2003b) respectively. In each of these knockout mice, exon 1 of *Bcl11a* and *Bcl11b* was targeted and deleted using conventional gene targeting strategies. *Bcl11a* and *Bcl11b* heterozygous mutant mice were fertile and viable; however, homozygous mutant mice died within a few hours of birth from unknown reasons. Southern blot analysis of genomic DNA from homozygous mutant mice indicated the presence of the targeted alleles and Western blots showed expression of Bcl11a and Bcl11b only in the wild-type and heterozygous mutant mice, confirming that *Bcl11a* and *Bcl11b*-deficient mice were obtained. *Bcl11a* homozygous mutant embryos lacked B cells and had alterations in several T cells. Further studies suggested that Bcl11a functions upstream of Ebf1 and Pax5 in the B cell development pathway (Liu et al., 2003b). In contrast, *Bcl11b* homozygous mutant embryos showed a block at the $CD4^-CD8^-$ double-negative stage of thymocyte development without any impairment to B cell and γδ T cell lineages (Wakabayashi et al., 2003b). These results suggest that *Bcl11* genes are essential for lymphocyte development.

Exon 1 encodes the translation start site and the first 18 amino acids of both Bcl11a and Bcl11b. However, bioinformatical analysis of the *Bcl11* genomic regions showed that alternative transcriptional start sites exist within the first intron, downstream of exon 1. Semi-quantitative RT-PCR analysis of cDNA obtained from *Bcl11* homozygous mutant embryos detected the presence of transcripts that were derived from alternative transcriptional start sites (Personal communication with Pentao, Liu). These results suggest that the original *Bcl11* knockout mice are not complete null alleles but are severe hypomorphic alleles. Therefore, I decided to generate new null alleles for both *Bcl11a* and *Bcl11b* as part of the high-throughput recombineering vector construction

program (Chan et al., 2007). In contrast to exon 1, exon 4 of *Bcl11a* and *Bcl11b* encodes for at least 75% of the total protein-coding sequences of either Bcl11 proteins which contains the two main functional C_2H_2 zinc fingers. Therefore, to generate the *Bcl11-lacZ* conditional null reporter mice, I decided to target exon 4 of both *Bcl11* genes as it encodes the main functional domains of the proteins.

3.1.2 New recombineering reagents

Analysing knockout mice produced through gene targeting in mouse ES cell is still the most widely used approach to understand mammalian gene function. Knockout mutant lines have been generated for hundreds of mouse genes and analyses of these mutant lines have provided invaluable insights into mammalian gene functions (Austin et al., 2004). The first step in generating a knockout allele (null or conditional) is to construct a targeting vector in *E. coli* that is subsequently electroporated into mouse ES cells for homologous recombination. Our lab and others have previously described methods for constructing targeting vectors using recombineering, which is based on highly efficient homologous recombination systems from bacteriophages as detailed in Chapter 1.2.6 (Angrand et al., 1999; Lee et al., 2001; Liu et al., 2003a). A key bottleneck for genome-wide targeted mutagenesis programmes in the mouse is the generation of targeting constructs (Austin et al., 2004; Auwerx et al., 2004). We have recently developed a high-throughput recombineering system which can generate conditional knockout targeting vectors in a rapid and efficient manner (Chan et al., 2007). A key reagent for this new system includes a complete λ phage that is replication-defective in BAC-harbouring DH10B *E. coli* but still retains its heat-inducible homologous recombination functions. Another reagent includes a set of low-copy plasmids (pSim) that contains the genes encoding for recombineering functions in their native operon, *pL* under the control of λ CI repressor (Datta et al., 2006). By using cells containing either of these reagents, thousands of BAC clones can be made recombineering-competent either by transfection (λ phage) or transformation (pSim). Further improvement of recombineering in these systems allowed the steps for making targeting vectors to be performed in 96-well plates, making the whole process suitable for high-throughput operations. I decided to generate *Bcl11a-lacZ* and *Bcl11b-lacZ* reporter conditional null

mice using the new recombineering reagents in order to obtain (1) reporter alleles to study the expression patterns of these genes in other tissues and (2) exon 4 null alleles of *Bcl11* genes.

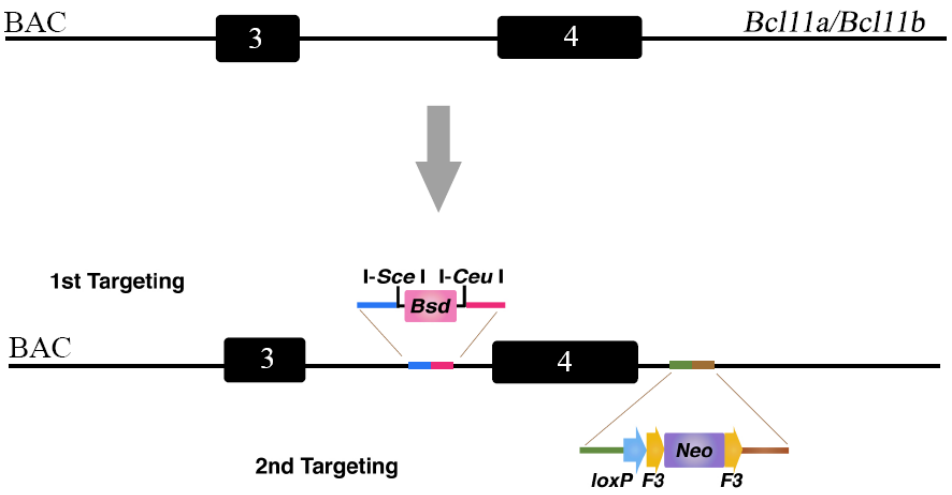
3.1.3 Targeting strategy for generating *Bcl11-lacZ* reporter alleles

The overall strategy for the construction of targeting and retrieval vectors for *Bcl11a* and *Bcl11b* is shown in Figure 3.1. Firstly, the BAC clones containing the *Bcl11a* and *Bcl11b* genomic regions were identified and verified. These BAC clones were then made recombineering-competent either by λ phage transfection or by transformation of pSim plasmids. Next, in order to generate targeting vectors with homologous sequences to the targeted region, PCR reactions using long primers were carried out using *Bsd*/*Neo*/retrieval cassettes as templates. Each primer used in the PCR reactions was 100 mers long, consisting of 80 bases of homologous sequence to the targeted region and 20 bases of homologous sequence to the *Bsd*/*Neo*/retrieval cassette templates. For both *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting vectors, ~5-kb and ~3 kb of genomic DNA were chosen as the left and right homology arms respectively. The genomic DNA region to be deleted is 4,565 bp and 3,615 bp for *Bcl11a* and *Bcl11b* respectively and deletion would result in a frameshift. The PCR products were then purified and used for targeting to the BAC clones using recombineering. The *Bsd* cassette was first targeted to the BAC clones and successful transformants were selected using LB agar containing blasticidin (Bsd; Figure 3.1A). The Bsd-resistant (Bsd^R) BAC clones were then verified by checking for their resistance to tetracycline (Tet^R) to ensure that they had retained recombineering-competence before the second round of targeting. The *Neo* cassette containing *lox*P-*F3-PGK-EM7-Neo-F3* was then targeted to the BAC clones using recombineering and successful transformants were selected on LB agar containing kanamycin (Kan; Figure 3.1A). The Bsd^R and Kan-resistant (Kan^R) BAC clones were again verified to ensure that they still retained recombineering-competence. Next, the correctly *Bsd*- and *Neo*-targeted BAC DNA were retrieved into the retrieval plasmid PL611 using recombineering and the successfully retrieved clones were selected on LB agar containing ampicillin (Amp; Figure 3.1B). The Bsd^R–Kan^R–Amp^R (Amp-resistant) retrieval plasmids were then cut with two rare restriction homing endonucleases *I-Sce*I and *I-Ceu*I and ligated with a

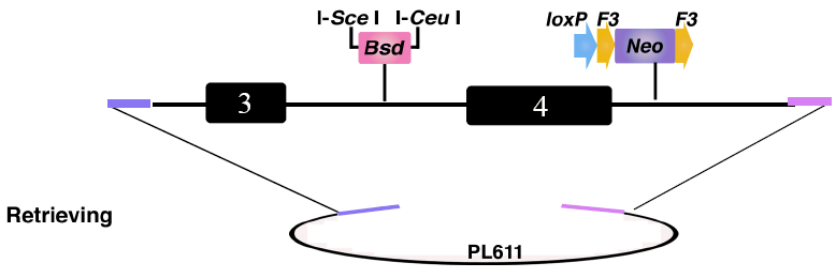
similarly cut *lacZ-Puro* cassette to generate the *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting constructs (Figure 3.1C). Finally, the negative selection marker Thymidine kinase (*TK*) with the chloramphenicol acetyltransferase gene (*Cm*) was added to the backbone of the retrieval vector by recombineering; replacing the Amp^R coding sequence. The final *Bcl11a-lacZ* and *Bcl11b-lacZ* reporter targeting constructs were then verified by restriction digestion and sequencing before targeting to ES cells.

The flexibility of these targeting vectors enables the generation of a multi-purpose allele that can serve as a reporter, knockout and a conditional knockout allele. The expression of *lacZ* is driven by the endogenous *Bcl11* regulatory elements, functioning as a reporter allele (Figure 3.1D). By staining $Bcl11^{lacZ/+}$ heterozygous embryos or tissues with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), *Bcl11* expression can be detected by the presence of blue staining. To generate a conditional knockout allele, the *lacZ* reporter mouse line can be crossed to a ubiquitous Flpe recombinase-expressing mouse line to excise the *lacZ-Puro* and the *Neo* cassettes. Upon expression of Flpe recombinase, the intervening sequences between the respective *FRT* and *F3* sites are deleted, leaving behind two *lox*P sites flanking exon 4 of *Bcl11a* and *Bcl11b*, generating a conditional knockout allele (Figure 3.1E). Subsequently, the mouse line with the conditional knockout allele can be crossed to a tissue-specific Cre recombinase-expressing mouse line. Upon expression of Cre recombinase, the intervening sequences between the two *lox*P sites would be deleted, producing a null allele (Figure 3.1F).

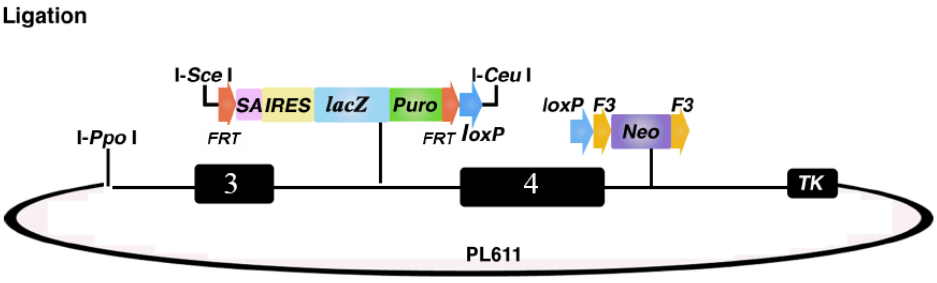
A


BAC
3
4
Bcl11a/Bcl11b
1st Targeting
I-Sce I
I-Ceu I
Bsd
BAC
3
4
2nd Targeting
Neo
loxP
F3
F3

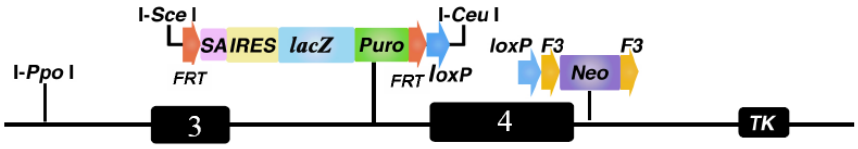
B

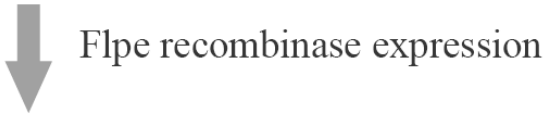

I-Sce I
I-Ceu I
Bsd
loxP
F3
F3
Neo
3
4
Retrieving
PL611

C

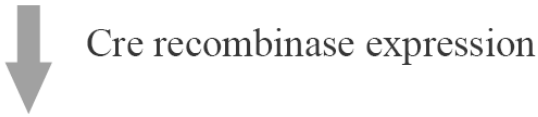

Ligation
I-Sce I
I-Ceu I
SA
IRES
lacZ
Puro
loxP
F3
F3
Neo
I-Ppo I
FRT
FRT
loxP
3
4
TK
PL611

D





E



F

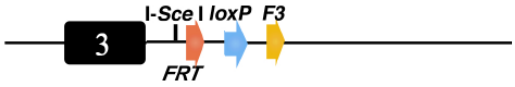


Figure 3.1. Construction of conditional knockout targeting vectors using the new recombineering reagents. (A) The genomic structure of the *Bcl11* locus with exon 4 to be deleted in the conditional knockout (cko) allele. The *Bsd* cassette flanked by two rare cutter sites, *I-Sce*I and *I-Ceu*I, is targeted to the 5' side of the intended deletion region. Subsequently, the *lox*P-*F3-PGK-EM7-Neo-F3* (*Neo*) cassette is targeted to the 3' side of the deletion region. Coloured lines represent the short homology arms used for recombineering. **(B)** The genomic DNA fragment is then retrieved from the BAC to PL611, which has the Amp^R gene. **(C)** The *Bsd* cassette can be conveniently replaced by a reporter, i.e. *lacZ*, in a simple ligation reaction. The final targeting vector has the reporter flanked by two *FRT* sites followed by a *lox*P site at the 5' side of the intended deletion region, and a *F3* flanked *Neo* cassette providing positive selection in ES cells. The negative selection marker *TK* is added to the vector backbone by recombineering. **(D)** The vector is then linearized with the rare-cutter *I-Ppo*I before targeting to ES cells. **(E)** Expression of Flpe recombinase would lead to recombination between the specific *FRT/F3* sites, resulting in removal of the *lacZ-Puro* reporter (*FRT*) and the *Neo* cassette (*F3*) sequences, leaving behind a single *FRT* and *F3* site at the 5' and 3' end of exon 4. This creates the cko alleles for *Bcl11a* and *Bcl11b* as exon 4 of either gene is now flanked by *lox*P sites and deletion of this exon can be mediated by tissue-specific expression of Cre recombinase. **(F)** Expression of Cre recombinase would lead to recombination between the two *lox*P sites and result in deletion of the intervening sequences (including exon 4 of each *Bcl11* gene), leaving behind a single *lox*P site. (Modified from Chan *et al.;* 2007).

3.2 Results

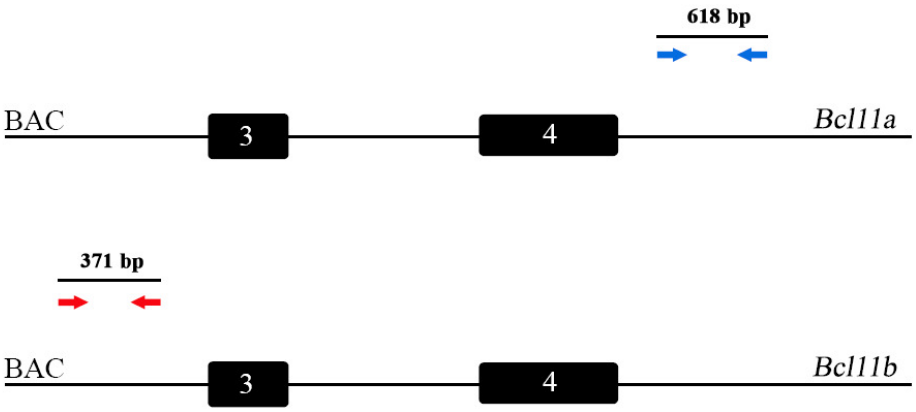
3.2.1 Construction of targeting and retrieval vectors

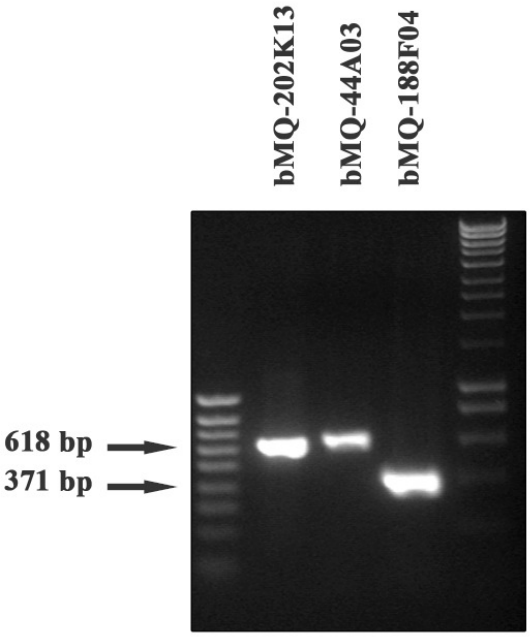
BAC clones containing the *Bcl11a* (bMQ-202K13 and bMQ-44A03) and *Bcl11b* (bMQ-188F04) genomic regions were identified from Wellcome Trust Sanger Institute 129/AB2.2 BAC library. BAC DNA were extracted and verified with sequence specific primers to ensure that they indeed contained the genomic regions of interest (Figure 3.2A). Next, templates for the PCR reactions were generated by restriction digestion of the *Bsd*/*Neo*/retrieval plasmids with appropriate enzymes and purification of the fragments of interest (Figure 3.2B-D). PCR reactions were then carried out using 1 to 5 ng of template (purified *Bsd*/*Neo*/retrieval cassettes) and PAGE-purified primers. The PCR products were detected in a 1% agarose gel and the appropriate bands purified to obtain the recombineering substrates (Figure 3.2E-G). The verified BAC clones were either transfected with λ phage (strain LE392) or transformed with pSim18 to deliver recombineering-competence. Transformants were selected on LB agar containing Tet (λ phage) or Hygromycin (Hyg; pSim18). Typically, there were between 500 to 1000 Tet^R (Tet-resistance) or Hyg^R (Hyg-resistance) colonies obtained for each BAC clone. One Tet^R or Hyg^R BAC clone was then used for the first targeting with the *Bsd* PCR product and successful transformants were selected on LB agar containing Bsd. There were between 20 to 100 Bsd^R colonies obtained for each BAC clone. Two Bsd^R colonies were picked for each clone and streaked onto LB agar containing Amp to determine if Bsd^R was due to contamination of the uncut *Bsd* cassette. None of the clones were Amp^R, suggesting that these Bsd^R clones were derived from successful *Bsd* cassette targeting. Next, the Bsd^R clones were checked for their susceptibility to Tet or Hyg to determine if the BAC clones had retained their recombineering competence after the first round of targeting. None of the Bsd^R clones were Tet^R or Hyg^R, suggesting that they had lost their recombineering competence. These Bsd^R clones were again either transfected with λ phage or transformed with pSim18 to restore recombineering-competence. Successful transformants were used for the next round of targeting.

After targeting of the *Bsd* cassette to the BAC DNA, the *Neo* PCR product, which is flanked by two *F3* sites and one *lox*P site, was targeted to the 3' side of the deletion region. *F3* is a mutant *FRT* site and in the presence of Flpe recombinase, recombination

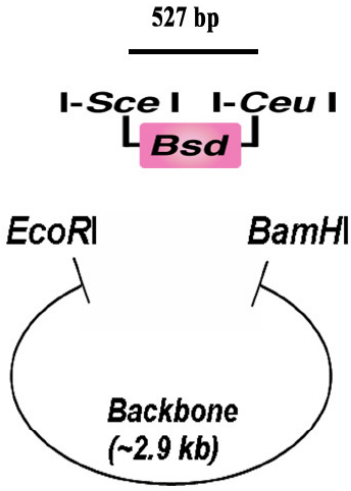
occurs only between two *F3* sites or wild-type *FRT* sites but not between *F3* and *FRT* sites (Schlake and Bode, 1994). The *Neo* cassette is driven by a *PGK* promoter and an *EM7* promoter and is therefore functional in both mammalian cells and *E. coli*. There were between 50 to 100 Kan^R colonies obtained for each BAC clone. Two Kan^R colonies were picked for each clone and tested for their susceptibility to Amp and Bsd. All the clones were Bsd^R but susceptible to Amp, suggesting that these clones were derived from successful *Bsd* and *Neo* cassette targeting. Following completion of the two targeting steps, the modified BAC was subsequently retrieved to a modified pBR322 plasmid (PL611). pBR322 replication origin was used instead of pUC19 to reduce the instability problems associated with cloning mouse genomic DNA into high-copy pUC-typed plasmid. There were between 10 to 50 Amp^R colonies obtained from this round of retrieving.

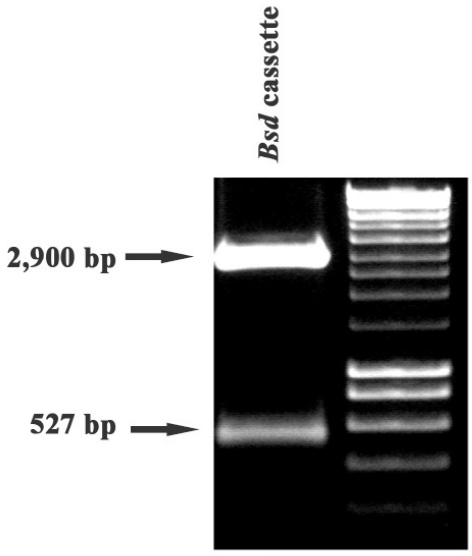
A


618 bp
BAC
3
4
Bcl11a
371 bp
BAC
3
4
Bcl11b

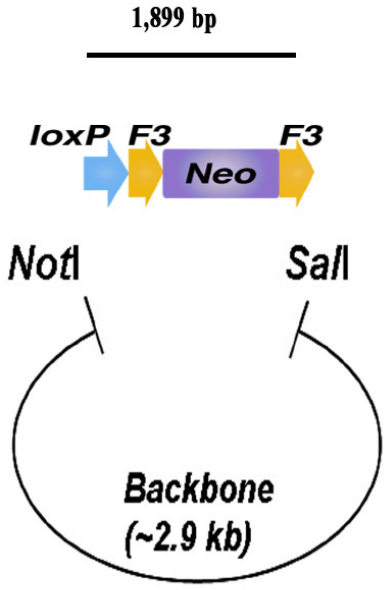

bMQ-202K13
bMQ-44A03
bMQ-188F04
618 bp
371 bp

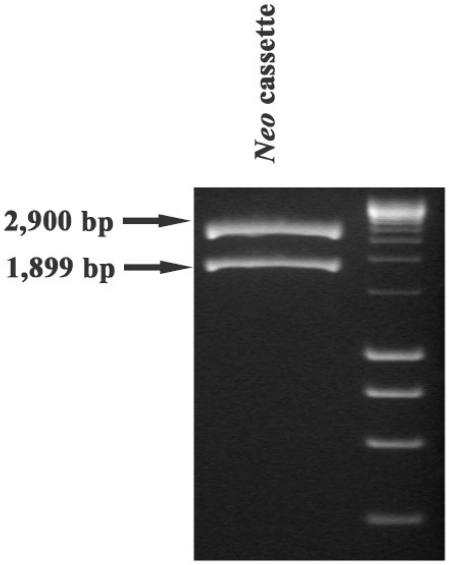
B


527 bp
I-Sce I
I-Ceu I
Bsd
EcoRI
BamHI
Backbone
(~2.9 kb)

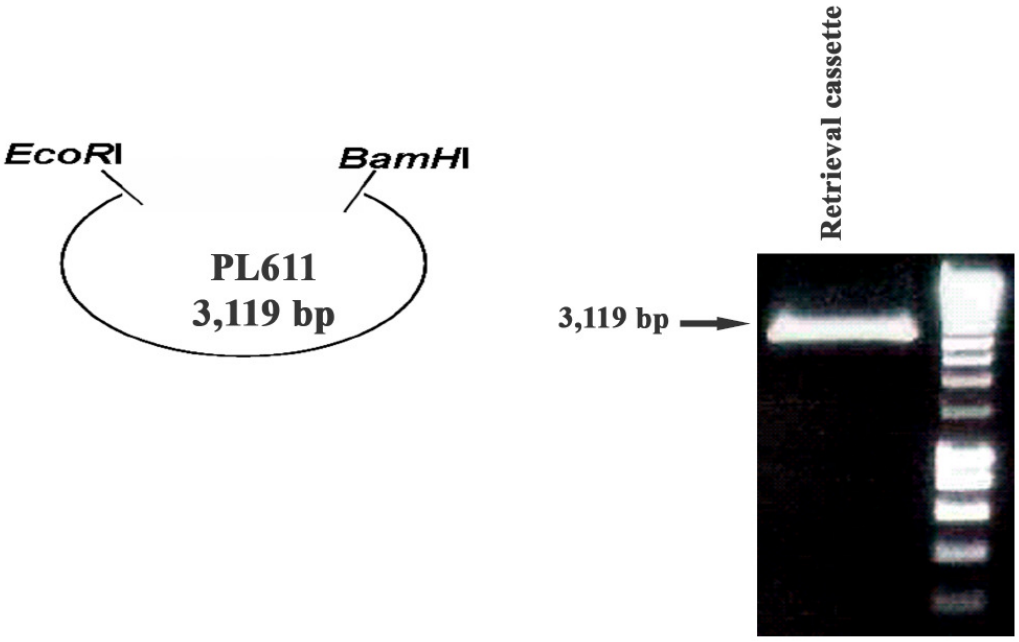

Bsd cassette
2,900 bp
527 bp

C

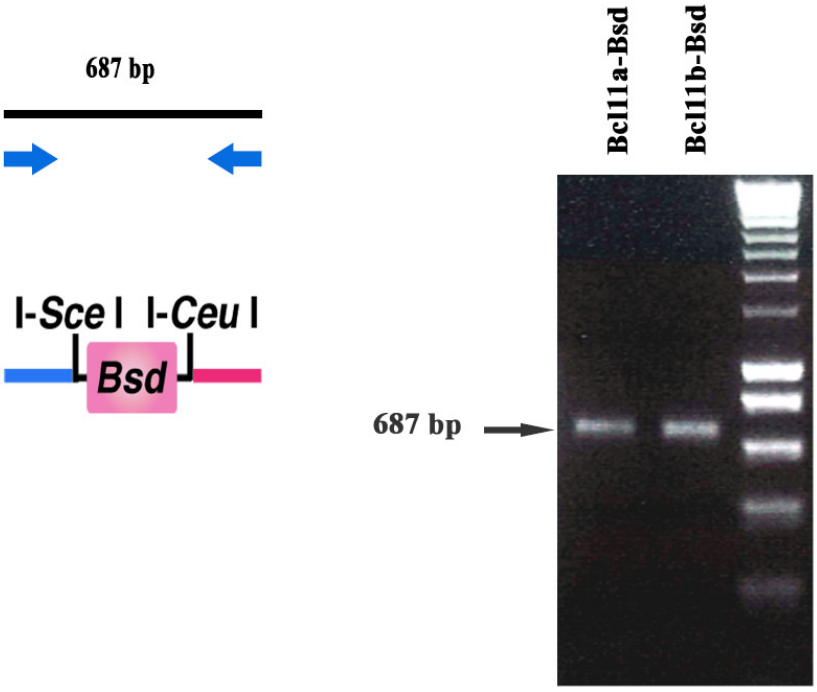

1,899 bp
loxP
F3
Neo
F3
NotI
SalI
Backbone
(~2.9 kb)


Neo cassette
2,900 bp
1,899 bp

D


EcoRI
BamHI
PL611
3,119 bp
Retrieval cassette
3,119 bp

E


687 bp
I-SceI
I-CeuI
Bsd
Bcl11a-Bsd
Bcl11b-Bsd
687 bp

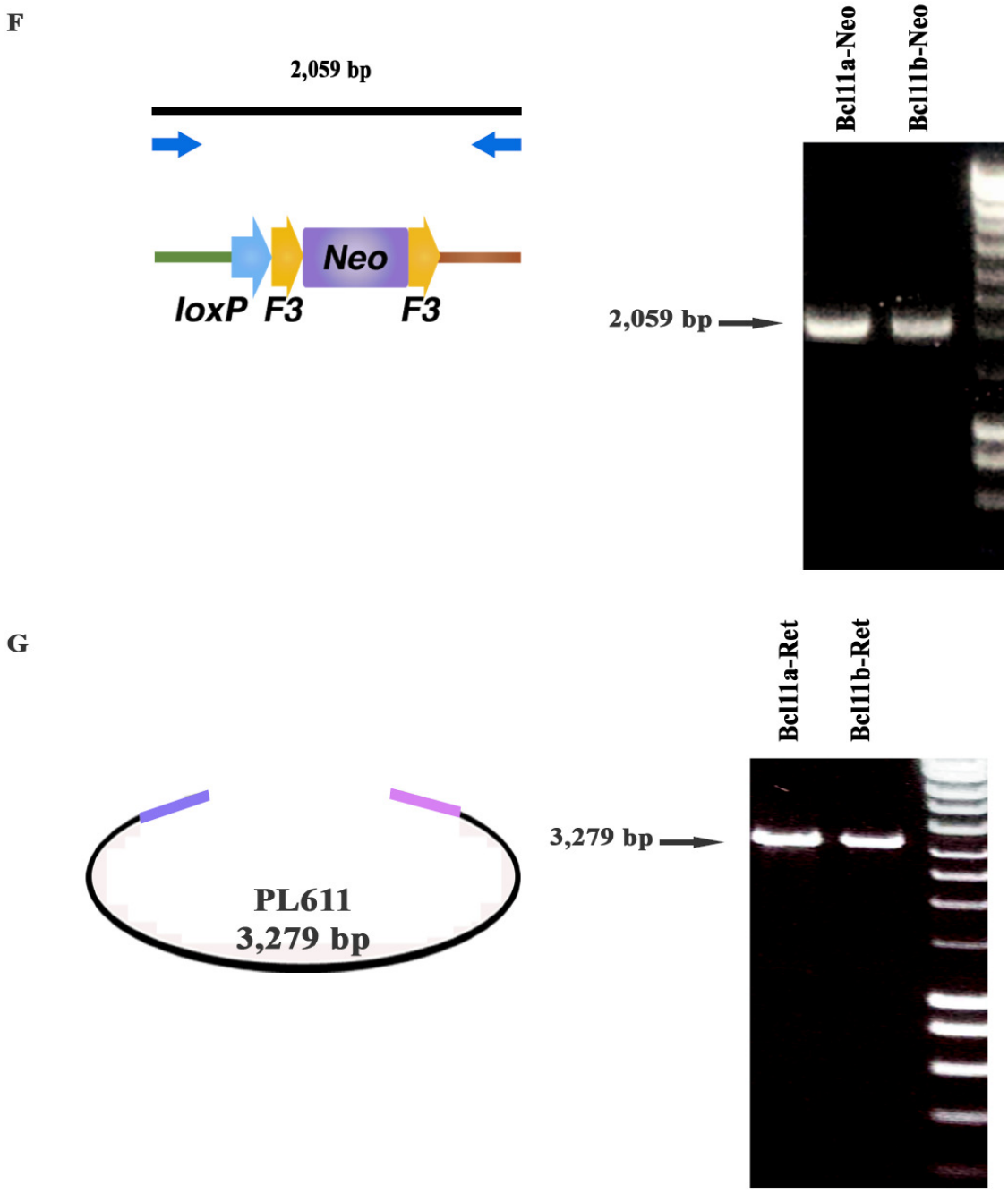


Figure 3.2. Construction of recombineering substrates by PCR. (A) Schematic diagrams depicting positions of primers used to validate BAC clones. Gel image showing PCR validation of the *Bcl11a* (bMQ-202K13 and bMQ-44A03) and *Bcl11b* (bMQ-188F04) BAC clones by sequence-specific primers. Gel images showing digestion products of **(B)** *Bsd* **(C)** *Neo* **(D)** retrieval plasmids after restriction digestion with appropriate enzymes as indicated in schematic diagrams. Fragments corresponding to **(B)** 527 bp **(C)** 1,899 bp and **(D)** 3,119 bp of *Bsd*, *Neo* and retrieval (Ret) cassette are purified and used as templates for PCR reactions. Gel images showing PCR products of **(E)** *Bsd* (687 bp), **(F)** *Neo* (2,059 bp) and **(G)** Ret (3,279 bp) cassettes. Homology arms of each product are shown as coloured blocks. PCR products are purified and used for recombineering.

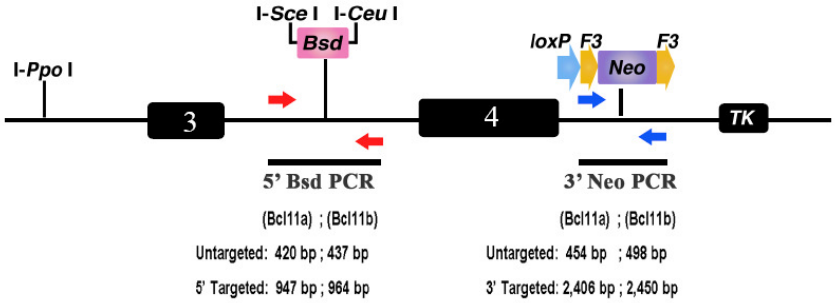
The retrieved doubly-targeted plasmids were then verified by pairs of gene-specific primers as shown in Figure 3.3A. Each of these primer pair were designed to span 200 to 250 bp either side of the 5' *Bsd* and 3' *Neo* targeted sites. Therefore, PCR amplification of plasmids containing non-modified genomic regions would produce a 420/454 bp (*Bcl11a/Bcl11b*) and/or a 437/498 bp (*Bcl11a/Bcl11b*) band respectively. In contrast, the BACs with correctly targeted 5' *Bsd* and 3' *Neo* cassettes would produce a 947/964 bp (*Bcl11a/Bcl11b*) and a 2,406/2,450 bp (*Bcl11a/Bcl11b*) band respectively (Figure 3.3A). As shown in Figure 3.3B, each of the two retrieved doubly-targeted plasmids contained the *Bsd* and *Neo* cassettes correctly targeted to the 5' and 3' regions.

Next, the 5' *Bsd* cassette was replaced by the *lacZ-Puro* reporter cassette. The *lacZ-Puro* cassette, with its splice acceptor (*SA*) and its polyadenylation site, would disrupt full transcription of the endogenous allele, creating a loss-of-function allele. Both the *Bsd* cassette and the *lacZ-Puro* cassette are flanked by two rare cutter sites, *I-Sce*I and *I-Ceu*I. Restriction digestion of the retrieved BAC clones with *I-Sce*I and *I-Ceu*I would result in the release of the 5' *Bsd* cassette (Figure 3.3C). The *lacZ-Puro* cassette was also restriction digested with *I-Sce*I and *I-Ceu*I (Figure 3.3D) and the purified fragment was ligated into the digested retrieved plasmids to generate the *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting constructs. Selection of *E. coli* cells containing the *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting constructs were done on LB agar containing Kan and Puro. Double antibiotic selection was used to eliminate any background. There were between 10 to 25 Kan^R and $Puro^R$ colonies obtained, usually after a 16 to 24 hour incubation period. Two Kan^R and $Puro^R$ colonies of each construct were picked and the plasmids were restriction digested with either *I-Ppo*I or *I-Sce*I in combination with *I-Ceu*I to confirm that they contained the *lacZ-Puro* insert. Figure 3.3E showed that all 4 clones contained the *lacZ-Puro* insert (6,814 bp).

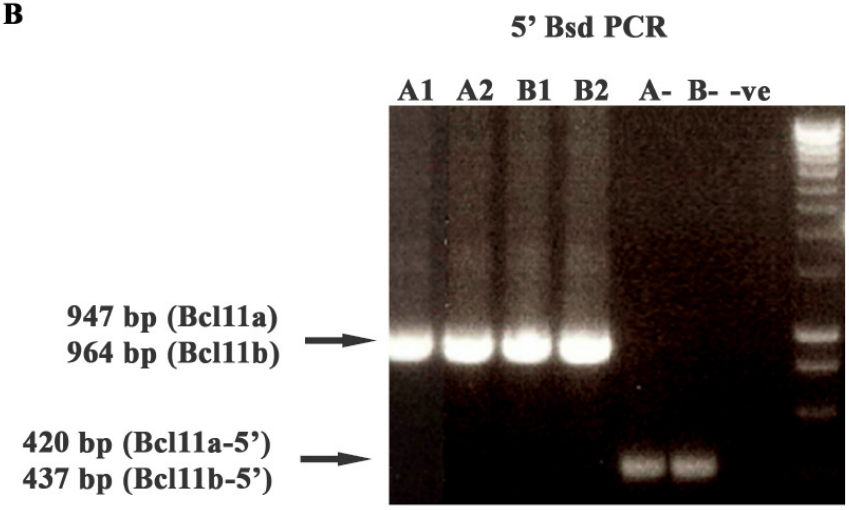
Finally, the negative selection marker *MC1-TK* was targeted to the backbone of the *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting constructs, replacing the Amp^R coding sequence. The *MC1-TK* serves as a negative selection marker for targeting in ES cells. This strategy relies on the principal that those cells in which random integration has occurred is likely to retain the *TK* gene and will thus be specifically eliminated by the antiviral agent gancyclovir (Thomas et al., 1986). By this means, enrichment in properly

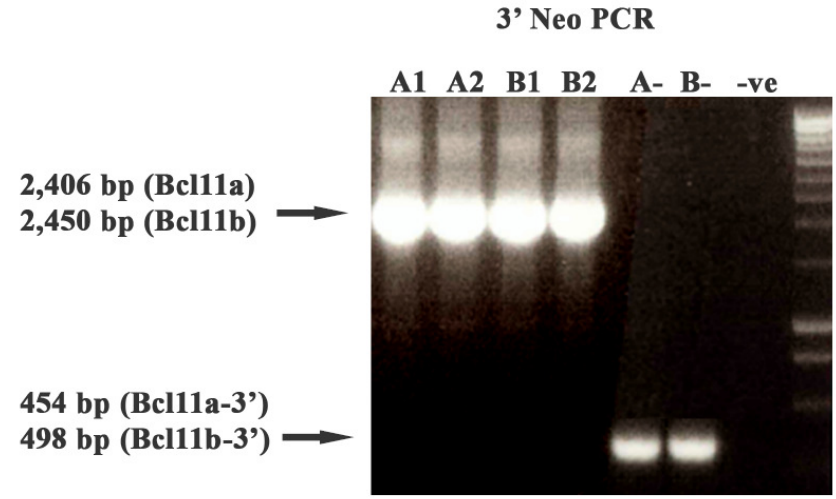
targeted clones could be obtained. The *Cm-TK* cassette was restriction digested (Figure 3.3F) and the required fragment was purified and used to target *Cm-TK* to the backbone of the retrieved plasmids by recombineering. Successful transformants were selected on LB agar containing Kan and Cm. There were about 50 to 100 Kan^R and Cm^R colonies. Four Kan^R and Cm^R colonies from each construct were selected and tested for their resistance to Puro, Kan, Cm and Amp. Clones containing the final targeting constructs with *Cm-TK* were $Puro^R$, Kan^R, Cm^R but susceptible to Amp. All the selected clones were found to contain the correct final targeting constructs. The final targeting constructs were then sequenced to ensure that there were no mutations within the *lacZ* reporter, *lox*P, *FRT* and *F3* sequences.

A

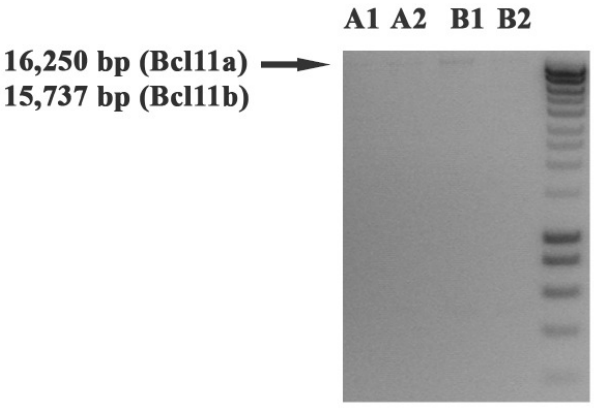

I-Sce I
I-Ceu I
Bsd
loxP
F3
F3
Neo
I-Ppo I
3
4
TK
5' Bsd PCR
(Bcl11a) ; (Bcl11b)
Untargeted: 420 bp ; 437 bp
5' Targeted: 947 bp ; 964 bp
3' Neo PCR
(Bcl11a) ; (Bcl11b)
Untargeted: 454 bp ; 498 bp
3' Targeted: 2,406 bp ; 2,450 bp

B

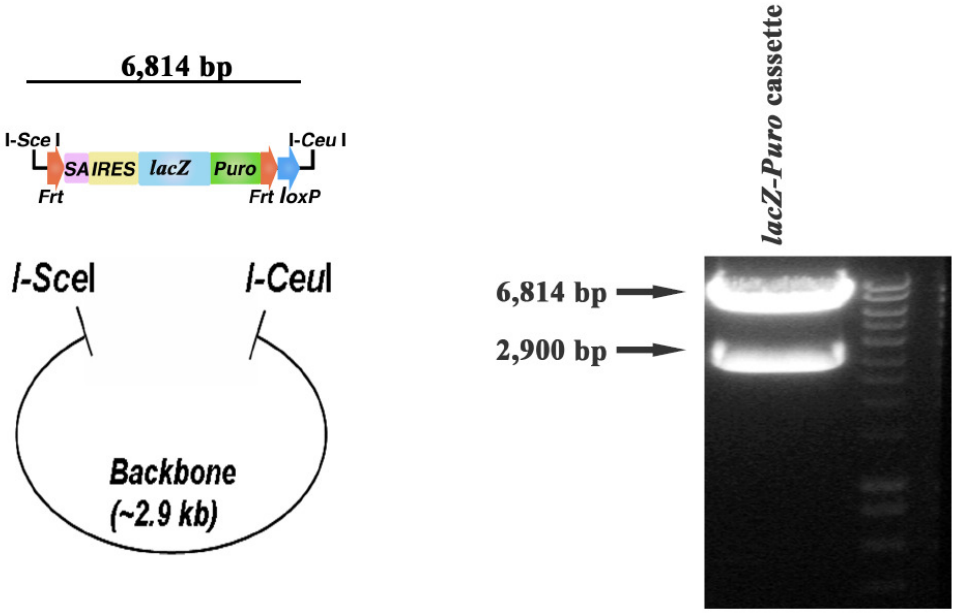

5' Bsd PCR
A1 A2 B1 B2 A- B- -ve
947 bp (Bcl11a)
964 bp (Bcl11b)
420 bp (Bcl11a-5')
437 bp (Bcl11b-5')


3' Neo PCR
A1 A2 B1 B2 A- B- -ve
2,406 bp (Bcl11a)
2,450 bp (Bcl11b)
454 bp (Bcl11a-3')
498 bp (Bcl11b-3')

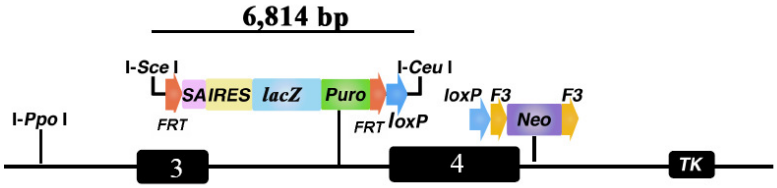
C

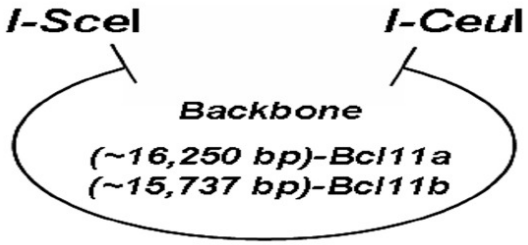

A1 A2 B1 B2
16,250 bp (Bcl11a)
15,737 bp (Bcl11b)

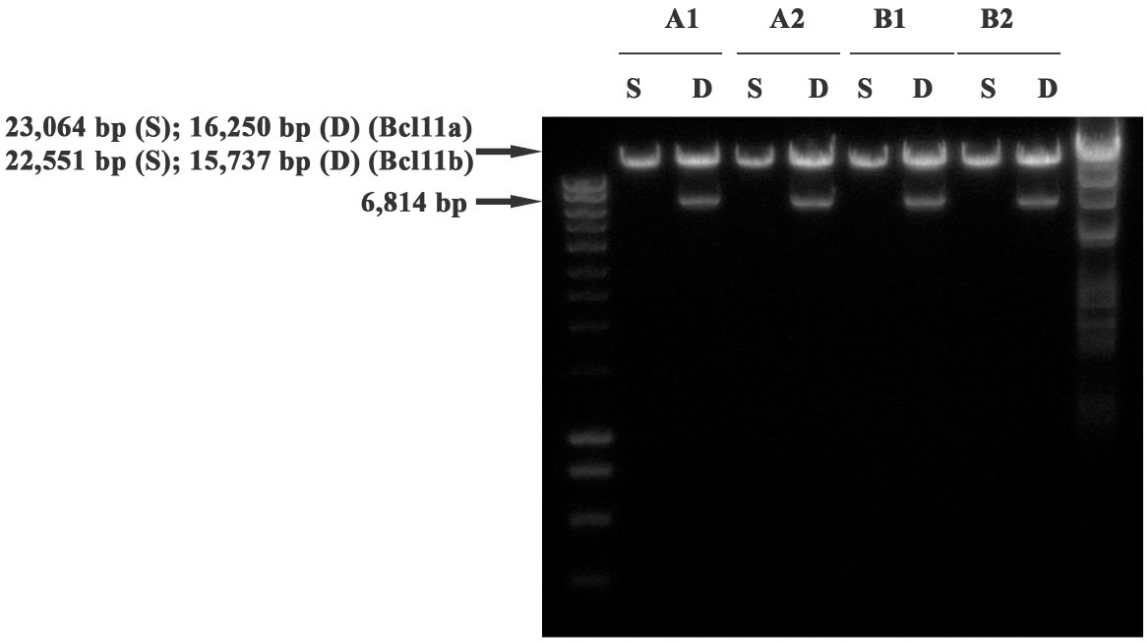
D


6,814 bp
I-Sce I
I-Ceu I
SA
IRES
lacZ
Puro
Frt
Frt
loxP
I-SceI
I-CeuI
Backbone
(~2.9 kb)
lacZ-Puro cassette
6,814 bp
2,900 bp

E







S: *I-Ppo*I digest

D: *I-Sce*I +*I-Ceu*I digest

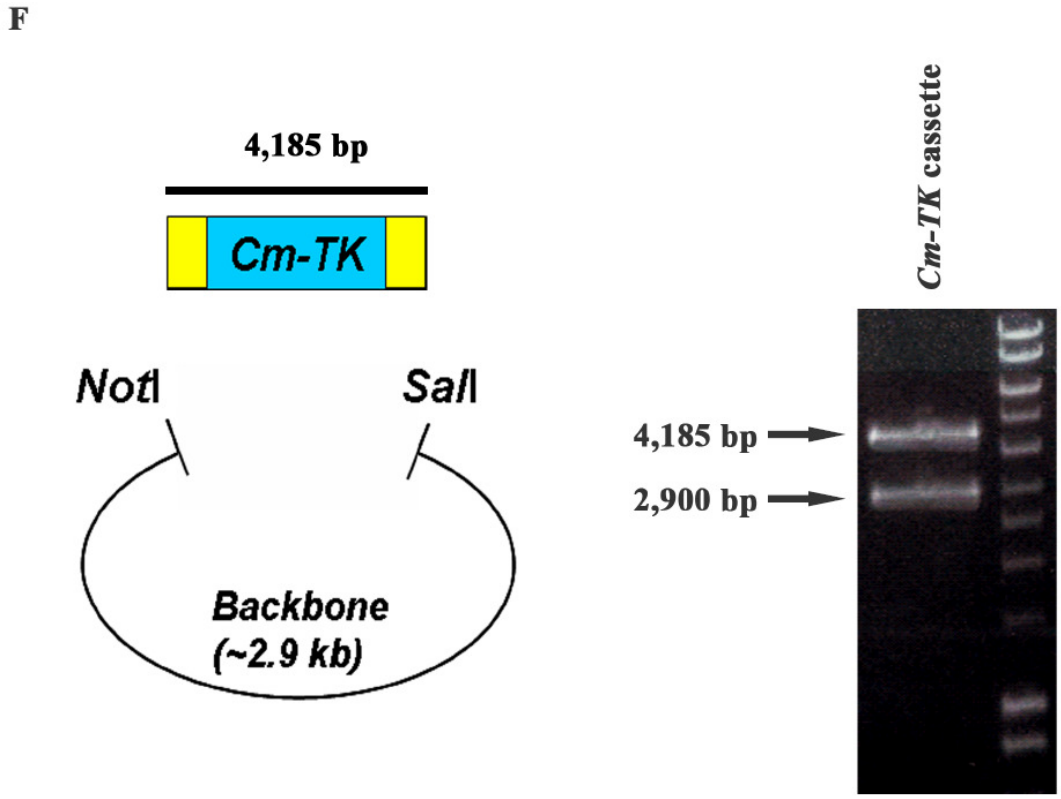


Figure 3.3. Validation of targeting constructs. (A) Red arrows indicate primer pairs for validating 5' *Bsd* targeting while blue arrows indicate primer pairs for validating 3' *Neo* targeting. Expected PCR products from untargeted and targeted regions are as shown in figure. **(B)** Gel image showing validation PCR products from *Bcl11a* (A1/A2) and *Bcl11b* (B1/B2) targeting vectors. *Bcl11a* constructs with successful *Bsd* and *Neo* targeting show 947 bp and 2,406 bp PCR products with *Bsd* and *Neo* validation primers respectively. Similarly, *Bcl11b* constructs with successful *Bsd* and *Neo* targeting show 964 bp and 2,450 bp PCR products with *Bsd* and *Neo* validation primers respectively. A-/B-/-ve indicate no template controls. **(C)** Restriction digestion confirmation of plasmids with successful retrieval of *Bsd* and *Neo* targeted *Bcl11a* and *Bcl11b* BAC regions. **(D)** Restriction digestion of *lacZ-Puro* reporter plasmids with *I-Sce*I and *I-Ceu*I produces a 6,814 bp fragment that contains the *lacZ-Puro* reporter. This fragment is purified and used for ligation with *Bcl11a* and *Bcl11b* targeting vectors that have been cut with *I-Sce*I and *I-Ceu*I. **(E)** Confirmation of *Bcl11a* and *Bcl11b* targeting vectors that have been successfully ligated with *lacZ-Puro* reporter cassette. Linearization of the targeting constructs with *I-Ppo*I produces a 23,064 bp band for *Bcl11a* and a 22,551 bp band for *Bcl11b*. In contrast, restriction digestion of the targeting constructs with *I-Sce*I and *I-Ceu*I results in the release of the *lacZ-Puro* insert (6, 814 bp) and a backbone of 16,250 bp and 15,737 bp for *Bcl11a* and *Bcl11b* constructs respectively. **(F)** Restriction digestion of *Cm-TK* plasmid with *Not*I and *Sal*I releases the *Cm-TK* insert (4,185 bp) from the pBluescript backbone. This fragment is purified and used for recombineering. Yellow areas denote homologous sequences to Amp^R sequence on PL611 backbone.

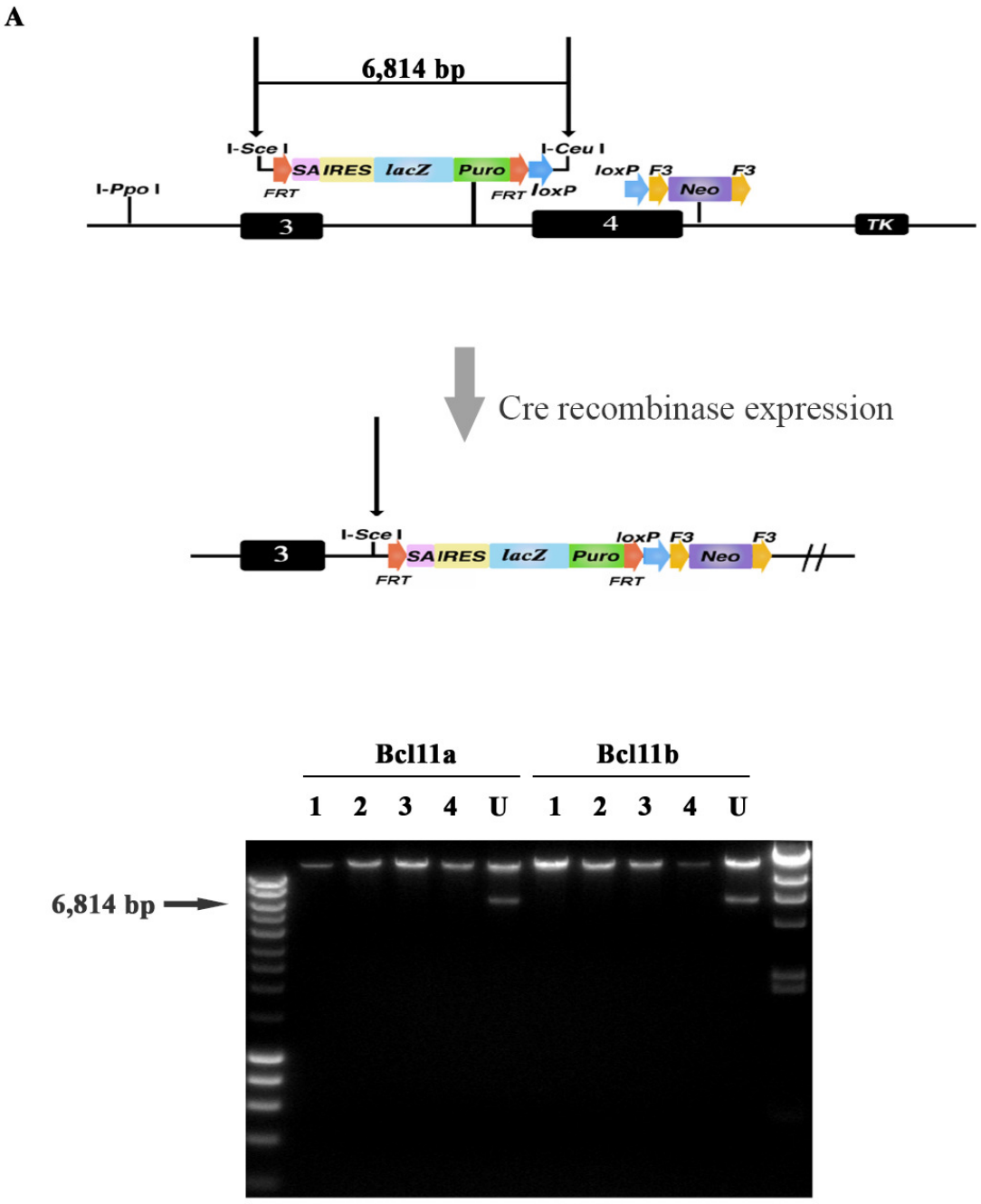
3.2.2 Verification of functionality of targeting constructs

The final targeting constructs described above would allow me to generate a flexible multi-purpose allele in the mouse. The *lacZ*-tagged alleles also serve as null alleles since the targeted *lacZ-Puro* cassette contains a splice acceptor (*SA*) which would disrupt endogeneous *Bcl11* transcription. In addition, by expression of Flpe recombinase, a conditional knockout allele can be obtained. The sequence and fidelity of *lox*P/*FRT*/*F3* sites is critical for recognition by Cre and Flpe recombinase. The efficacy of *lox*P sites of the targeting constructs was first tested using EL350 which was derived from DY380 (Lee et al., 2001). Besides the λ phage recombination genes found in DY380, EL350 also contained a tightly controlled arabinose-inducible Cre recombinase. Expression of Cre recombinase was induced by addition of arabinose and this facilitated recombination between the two *lox*P sites, resulting in deletion of the intervening sequences. The *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting constructs contained two *lox*P sites. Upon expression of Cre recombinase, the intervening sequence between the *lox*P sites, which contained exon 4 and the *I-Ceu*I restriction digestion site would be deleted (Figure 3.4A). Hence, after Cre recombinase expression, constructs which contained functional *lox*P sites would show a single band (21,529 bp – Bcl11a; 21,807 bp– Bcl11b) following double digestion with *I-Sce*I and *I-Ceu*I. The final targeting constructs were first electroporated into EL350 cells that were either induced or uninduced with arabinose for Cre expression. Cells were then plated into LB agar plates containing Cm. Four Cm^R colonies were picked for each construct and analysed by restriction digestion. As shown in Figure 3.4A, all four clones of each targeting constructs showed a single band after *I-Sce*I and *I-Ceu*I double digestion following Cre recombinase expression. In contrast, constructs that were electroporated into EL350 without Cre recombinase expression retained the *I-Ceu*I site. Therefore restriction digestion with *I-Sce*I and *I-Ceu*I resulted in the presence of two bands, the vector backbone (~21 kb) and the *lacZ-Puro* insert (6,814 bp) (Figure 3.4A). Hence, the *lox*P sites in both *Bcl11a-lacZ* and *Bcl11b-lacZ* constructs were functional.

Next, I verified the fidelity and efficacy of the *FRT*/*F3* sites using EL250. EL250 was also derived from DY380; however, in contrast to EL350, EL250 contained a tightly controlled arabinose-inducible Flpe recombinase gene (Lee et al., 2001). Expression of

Flpe recombinase was induced by addition of arabinose and this would facilitate recombination between the two identical *FRT*/*F3* sites, resulting in removal of the *lacZ-Puro* and *Neo* cassettes from the *Bcl11a-lacZ* and *Bcl11b-lacZ* constructs. The *Bcl11a-lacZ* and *Bcl11b-lacZ* targeting constructs were first electroporated into EL250 cells that were either induced or uninduced with arabinose for Flpe expression. Next, the transformants were plated out onto LB agar plates containing either Kan or Cm. Clones that contained constructs in which the *lacZ-Puro* and *Neo* cassettes were successfully excised would be sensitive to Puro and Kan but resistant to Cm because the Cm^R coding sequence was found on the PL611 plasmid backbone. In contrast, clones which contained constructs with partial/no excision of the *lacZ-Puro* and *Neo* cassettes would remain $Puro^R$ and/or Kan^R. The number of Kan^R colonies observed would reflect the efficiency of Flpe recombination reactions. Flpe recombinase-mediated recombination was fairly efficient as there were fewer than 10 Kan^R colonies. In contrast, there were about 1000 Cm^R colonies. Four Cm^R colonies from each construct were selected and checked for their Kan^R. All four clones from the *Bcl11a-lacZ* construct were susceptible to Kan. In contrast, three of the Cm^R colonies from the *Bcl11b-lacZ* construct showed resistance to Kan (plasmid 2, 3 and 4). Next, these clones were tested using the 5' and 3' primer pairs (Figure 3.4B) to confirm that the *lacZ-Puro* and *Neo* cassettes were properly excised. Constructs with *lacZ-Puro* (5' PCR) and *Neo* (3' PCR) cassettes successfully excised would show a 5' PCR band of 454/471 bp (*Bcl11a*/*Bcl11b*) and a 3' PCR band of 488/532 bp (*Bcl11a*/*Bcl11b*) respectively (Figure 3.4B). All four Cm^R clones of the *Bcl11a-lacZ* construct showed complete excision of *lacZ-Puro* and *Neo* cassettes following Flpe recombinase expression (Figure 3.4B). In contrast, for the *Bcl11b-lacZ* construct, all four clones showed complete *Puro-lacZ* cassette excision but only one showed complete *Neo* cassette excision (Figure 3.4B). This result confirmed the previous observation that only one of the four Cm^R clones was susceptible to Kan (plasmid 1). As the other three clones (plasmids 2, 3 and 4) had incomplete/no excision of the *Neo* cassette, resistance to Kan was retained. This result also suggests that recombination between *F3* sites is not as efficient as *FRT* sites. However, this inefficiency would not pose a problem when using germline Flpe mice as Flpe recombinase would be expressed constitutively. In addition, in agreement with the study by Schlake *et al.* (Schlake and

Bode, 1994), no recombination between *FRT* and *F3* sites were detected, suggesting that recombination occurred only between two *F3* sites or wild-type *FRT* sites.



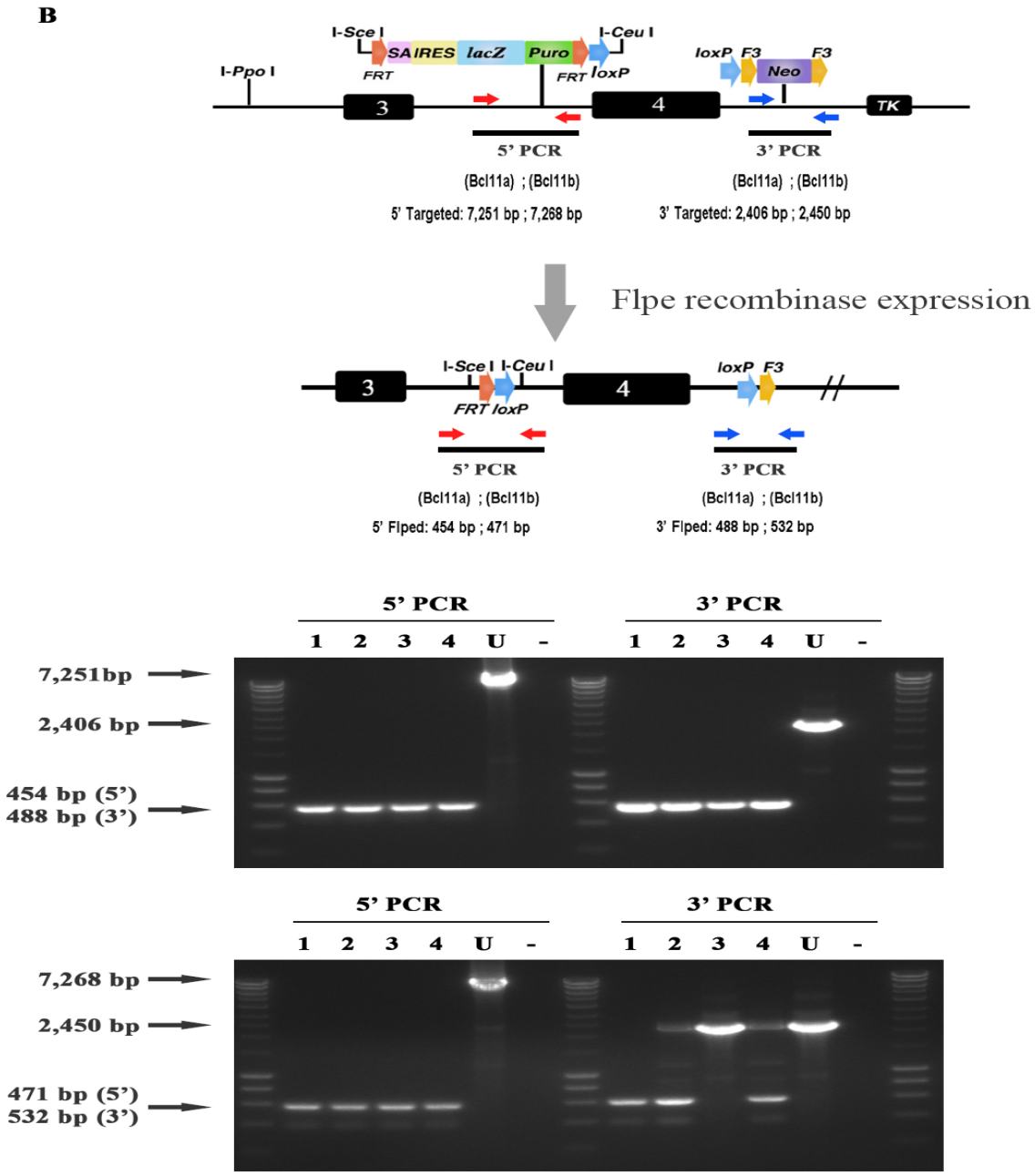
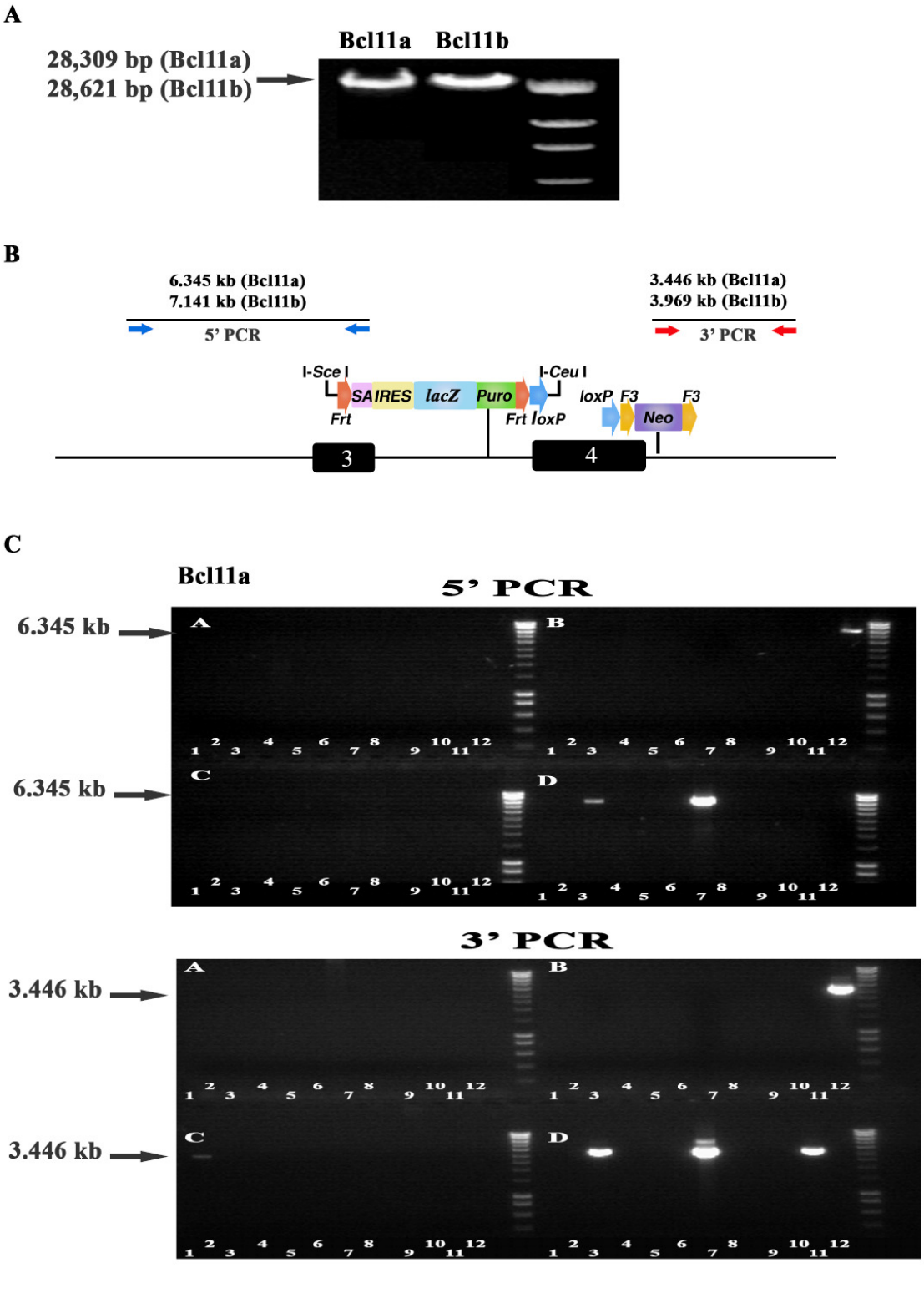


Figure 3.4. Verification of *lox*P/*FRT*/*F3* functionality. (A) Restriction digestion patterns of *Bcl11a* and *Bcl11b* targeting constructs with *I-Sce*I and *I-Ceu*I following Cre recombinase expression. Successful recombination between *lox*P sites results in elimination of *I-Ceu*I enzyme site. Therefore restriction digestion with *I-Sce*I and *I-Ceu*I produces a single band (Lanes 1-4) without the presence of the 6,814 bp *lacZ-Puro* insert as seen in the original constructs (Lane U). **(B)** Gel images showing PCR verification of *Bcl11a* and *Bcl11b* constructs following Flpe recombinase expression. Primer pairs used are as shown. Blue and red arrows indicate location of primers used for 5' and 3' PCR respectively. Successful recombination between *FRT* sites or *F3* sites results in elimination of the *lacZ-Puro* and *Neo* cassettes respectively. Hence PCR amplification of these plasmids (Lanes 1-4) produces a 454 bp product (5' PCR) and a 488 bp product (3' PCR) for *Bcl11a* and a 471 bp product (5' PCR) and a 532 bp product (3' PCR) for *Bcl11b*. Lane U shows the PCR product obtained from original targeting vectors without Flpe recombinase expression.

3.2.3 Targeting to ES cells and verification of targeted clones

The *Bcl11a-lacZ* and *Bcl11b-lacZ* final targeting constructs were subsequently linearized by *I-Ppo*I (Figure 3.5A) and electroporated into AB2.2 wild-type ES cells. Gene targeting events were identified by selecting for cells resistant to G418. DNA from G418-resistant ($G418^R$) colonies were analysed for the desired homologous recombination events with two independent long-range PCR reactions (5' and 3' PCR). Each of these primer pairs consists of a common primer that resides within the *SA* (5' PCR) or *Neo* cassette (3' PCR) and a gene-specific primer that resides outside the homology arms (Figure 3.5B). Correctly targeted ES cell clones would produce a 6,345/7,141 bp (*Bcl11a*/*Bcl11b*) and a 3,446/3,696 bp (*Bcl11a*/*Bcl11b*) band for the 5' and 3' long range PCR reactions respectively. As shown in Figure 3.5C, for *Bcl11a-lacZ* targeting, 3 of the 48 $G418^R$ colonies analysed showed positive long range PCR results for both 5' and 3' reactions, giving an overall targeting efficiency of at least 6.25%. In contrast, for *Bcl11b-lacZ* targeting, 8 of the 48 $G418^R$ colonies analysed showed positive long range PCR results for both 5' and 3' reactions, giving an overall targeting efficiency of at least 16.67% (Figure 3.5D).


A
Bcl11a Bcl11b
28,309 bp (Bcl11a)
28,621 bp (Bcl11b)
B
6.345 kb (Bcl11a)
7.141 kb (Bcl11b)
5' PCR
3.446 kb (Bcl11a)
3.969 kb (Bcl11b)
3' PCR
I-Sce I
SA
IRES
lacZ
Puro
I-Ceu I
loxP
F3
F3
Neo
Frt
Frt
loxP
3
4
C
Bcl11a
5' PCR
6.345 kb
6.345 kb
A
B
C
D
3' PCR
3.446 kb
3.446 kb

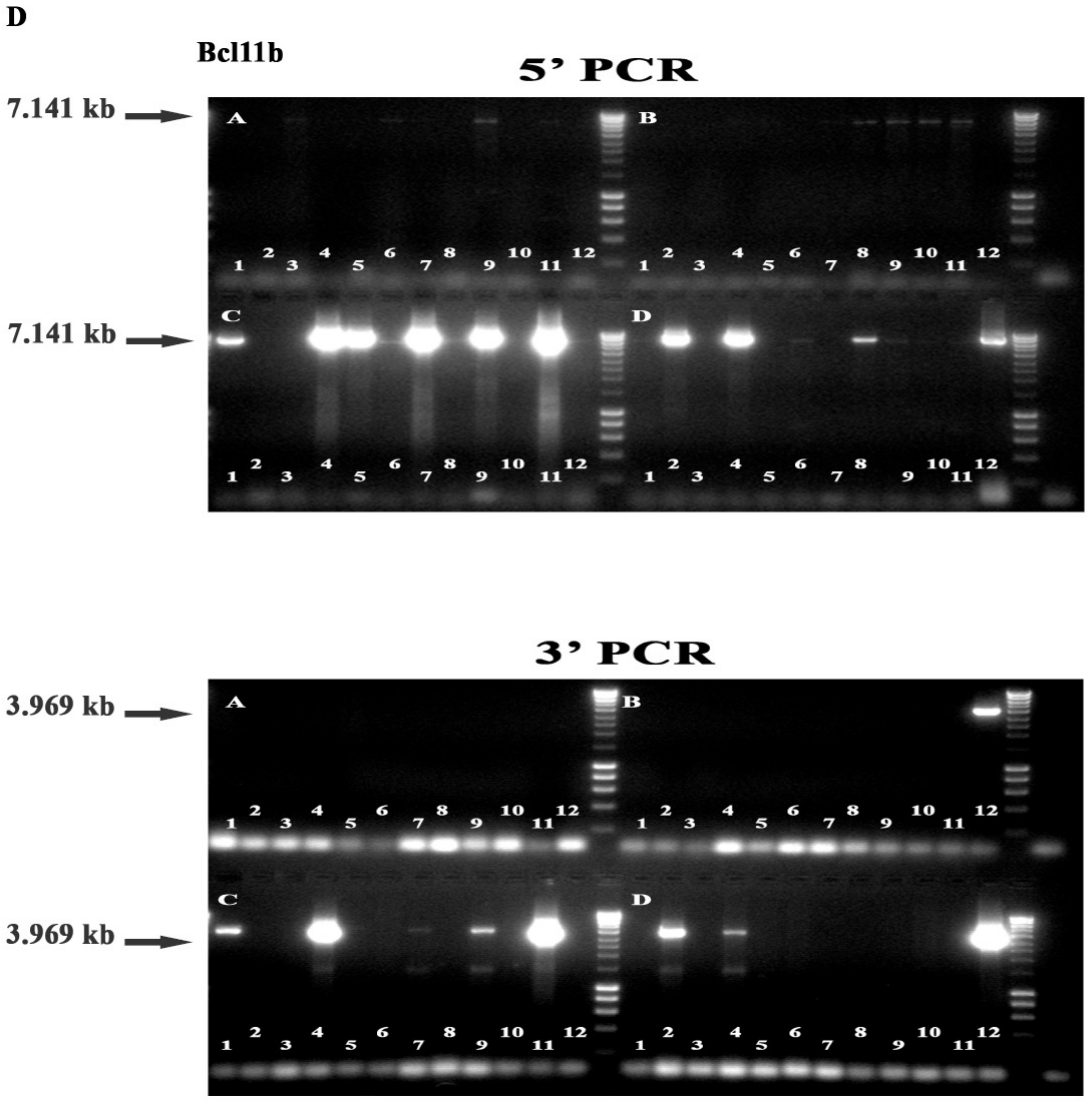


Figure 3.5. ES cell targeting. (A) Gel image showing linearized *Bcl11a* and *Bcl11b* final targeting vectors prior to electroporation into ES cells. **(B)** Schematic diagram showing the *Bcl11-lacZ* reporter alleles. Blue and red arrows indicate location of primers used for long range PCR confirmation of 5' and 3' targeting respectively. Long range PCR verification of **(C)** *Bcl11a-lacZ* and **(D)** *Bcl11b-lacZ* targeted ES clones using 5' and 3' primers. Numbers represent individual colony picked and analysed.

3.2.4 Confirmation of germline transmission of targeted alleles

Two independent targeted clones for each construct were injected into 3.5 dpc C57BL6/J mouse blastocyts which were subsequently implanted into uteri of pseudopregnant foster females (Microinjections were performed by Dr Si Qin and Tina Hamilton). Chimeras were obtained from each microinjection (Table 3.1). F0 chimeric mice are not typically phenotyped directly because they are comprised of a mixture of wild-type and mutant cells. Male chimeras were therefore bred to wild-type C57BL6/J females to generate F1 generation offspring. Genomic DNA was extracted from F1 offspring and genotyped using long-range PCR with primers as described in Figure 3.5B to assay for germline transmission of targeted alleles. As shown in Figure 3.6A, germline transmission of *Bcl11a-lacZ* and *Bcl11b-lacZ* targeted alleles were obtained in F1 mice. These heterozygotes were then intercrossed to produce F2 generation heterozygotes and/or homozygotes which were characterized to assess gene function. Having confirmed germline transmission of the targeted alleles by long-range PCR, mice were subsequently genotyped using two short PCR reactions. The position of primers used were as illustrated in Figure 3.6B; with a pair of gene-specific forward (Fwd) and reverse (Rev2) primers amplifying the wild-type band (*Bcl11a* – 363 bp; *Bcl11b* – 644 bp) and the gene-specific forward (Fwd) primer with a common reverse (Rev1) primer located within *SA* of the *lacZ-Puro* cassette amplifying the mutant band (*Bcl11a* – 437 bp; *Bcl11b* – 527 bp). PCR of genomic DNA from heterozygotes would produce two bands; a wild-type and a mutant band while genomic DNA from wild-type would produce only a wild-type band (Figure 3.6C).

Table 3.1. Chimeras obtained from microinjections

Allele	Clone/Passage number	% chimera
Bcl11a-lacZ	B12/P6	20
Bcl11a-lacZ	B12/P6	40
Bcl11a-lacZ	B12/P6	40
Bcl11a-lacZ	B12/P6	50
Bcl11a-lacZ	D7/P5	50
Bcl11a-lacZ	D7/P5	30
Bcl11b-lacZ	C11/P6	30
Bcl11b-lacZ	C11/P6	10
Bcl11b-lacZ	D2/P5	60
Bcl11b-lacZ	D2/P5	30

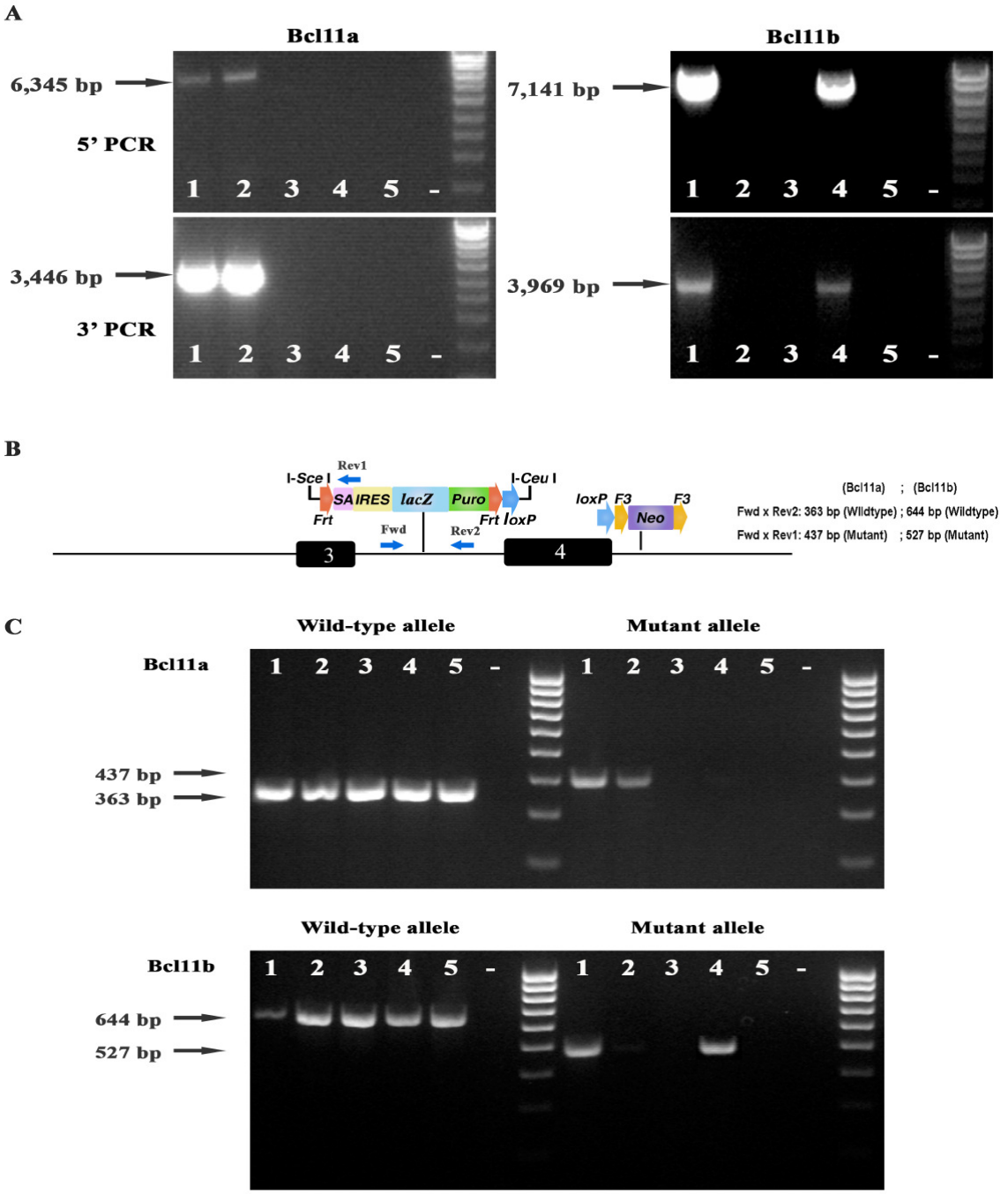


Figure 3.6. Confirmation of germline transmission of targeted alleles in F1 mice. Five F1 mice were characterized for targeted *Bcl11a-lacZ* and *Bcl11b-lacZ* alleles. **(A)** Gel images showing 5' and 3' long range PCR confirmation of germline transmission of targeted alleles in *Bcl11a-lacZ* and *Bcl11b-lacZ* F1 mice using long-range PCR. **(B)** Genotyping using short-range PCR. Blue arrows indicate positions of new genotyping primers. PCR amplification of tail genomic DNA with Fwd and Rev2 primer pairs produces a 363 bp (*Bcl11a*) and a 644 bp (*Bcl11b*) wild-type band. PCR amplification of tail genomic DNA with Fwd and Rev1 primer pair produces a 437 bp (*Bcl11a*) and a 527 bp (*Bcl11b*) mutant band. **(C)** The 363 bp wild-type band is present in all five *Bcl11a-lacZ* mice while the 437 bp mutant band is present only in mice 1 and 2 indicating these two mice are heterozygotes. The 644 bp wild-type band is present in all five *Bcl11b-lacZ* mice while the 527 bp mutant band is present only in mice 1 and 4 indicating these two mice are heterozygotes.

3.3 Discussion

I have used the new mobile recombineering reagents (λ phage and pSim) to generate targeting constructs for *Bcl11a-lacZ* and *Bcl11b-lacZ* reporter conditional null alleles. These two constructs were verified by expression of Cre and Flpe recombinase using EL350 and EL250 and the functionality and fidelity of *lox*P, *FRT* and *F3* sites were confirmed. I then electroporated these targeting constructs to wild-type ES cells and successfully isolated ES cells which contained the targeted alleles. These ES cell clones were confirmed by long range PCR genotyping. Two independent clones for each construct were then microinjected into 3.5 dpc mouse blastocyts and transplanted into uteri of pseudopregnant females. The chimeras obtained were then bred to wild-type females and the resultant F1 mice born were genotyped by long range PCR to confirm germline transmission of the targeted alleles. Heterozygous *Bcl11a-lacZ* and *Bcl11b-lacZ* mice were obtained.

During verification of the functionality and fidelity of the constructs, I observed that the recombination efficiency between *F3* sites was not as efficient as that of *FRT* sites. Three out of the four *Bcl11b-lacZ* clones picked showed partial recombination between *F3* sites (Figure 3.4B). However, it should be noted that expression of Flpe recombinase was only induced for one hour (in this experiment) and this inefficiency could be negated by increasing the duration of expression of Flpe recombinase. Moreover, this inefficiency should not pose a problem *in vivo* because Flpe recombinase is driven by the *Rosa26* promoter; hence Flpe recombinase is expressed constitutively and complete recombination between *F3* sites is expected and indeed achieved.

In Chapter 4, I will detail the expression patterns of *Bcl11* genes using the *Bcl11-lacZ* reporter mice. I will characterize the spatial expression patterns using whole mount X-gal staining and also delineate the expression patterns at a single cell level using flow cytometric analysis.

CHAPTER 4:
EXPRESSION PATTERNS OF *BCL11* GENES IN MICE

4.1 Introduction

4.1.1 Current knowledge of *Bcl11* genes expression patterns

Bcl11a and Bcl11b are transcription factors and dysregulation of either protein has been associated with etiology of disease in both human and mouse. Over-expression of Bcl11a following proviral integration resulted in the development of myeloid leukaemia in mice (Nakamura et al., 2000). This transformation event may be partially mediated by the physical interaction of Bcl11a with BCL6 (Nakamura et al., 2000). In contrast, homozygous deletions and point mutations of murine *Bcl11b* resulted in thymic lymphomas (Wakabayashi et al., 2003a). Chromosomal translocation of *BCL11A* was also shown to be involved in lymphoid malignancies in humans (Satterwhite et al., 2001). Additionally, both Bcl11a and Bcl11b have essential roles in murine lymphocyte development (Liu et al., ; Wakabayashi et al., 2003b). Recently, studies have implicated *BCL11A* in other human diseases. For example, *BCL11A* mutations have been identified in human breast cancers (Wood et al., 2007) and a quantitative trait locus (QTL) influencing F cell production maps to the *BCL11A* locus in human thalassemia patients (Menzel et al., 2007). These reports underline the potential importance of *Bcl11* genes in several tissues in human and mouse, and clearly emphasize the need for further characterization of their function and expression.

To date, several studies have reported the expression patterns of Bcl11a and Bcl11b proteins (Avram et al., 2000; Gunnersen et al., 2002; Leid et al., 2004; Nakamura et al., 2000). These studies provide useful information regarding the expression patterns of *Bcl11* genes. For example, expression of both *Bcl11* genes in mice is detected from 10.5 days post-coitum (dpc) and this expression persists till adulthood. During embryogenesis, *Bcl11a* is expressed in both mouse and rat cortex and may be required for neuronal development and differentiation (Gunnersen et al., 2002). In the adult mouse, *Bcl11a* mRNA is detected in the brain and spleen, and found at lower levels in the heart,

liver, testis and lung (Avram et al., 2000; Nakamura et al., 2000). Expression of *BCL11A* is also detected in human hematopoietic cells such as myeloid precursors, B cells, monocytes and megakaryocytes (Saiki et al., 2000). Expression of Bcl11b is detected in the mouse skin during embryogenesis and in adulthood by immmunohistochemistry, suggesting that *Bcl11b* may play a role in development and/or homeostasis of the skin (Golonzhka et al., 2007).

However, there are several limitations in the previous studies. Firstly, with the exception of the report by Golonzhka *et al.*, the rest of published data were primarily based on RT-PCR studies, northern analyses and RNA anti-sense *in situ* hybridization. Hence these methods are not sensitive enough to detect expression of *Bcl11* genes at a single cell level. Next, in order to obtain spatial expression patterns of genes using adult tissues such as mammary tissues, penetration of *in situ* probes and antibodies becomes extremely difficult. Therefore it is not feasible to use *in situ* hybridization or antibody staining to detect whole mount spatial expression of genes. To overcome these technical limitations and/or difficulties, I chose the bacterial *lacZ* gene as the reporter gene and generated the *Bcl11-lacZ* reporter mice.

4.1.2 Using *E. coli lacZ* as a reporter in mice

The bacterial *lacZ* gene, encoding the enzyme β-galactosidase (β-gal), is a commonly used reporter gene in mouse genetics because β-gal activity can be readily assessed *in vivo*. By targeting *lacZ* to the *Bcl11* loci, the endogenous *Bcl11* regulatory elements would control expression of *lacZ*. Therefore spatial expression patterns of *Bcl11* genes can be easily detected by staining of the embryos and tissues with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) to produce a blue product, 5-bromo-4-chloro-indigo. Using this approach, fine-resolution visualization of cell lineage in vertebrate nervous system has been studied (Trainor et al., 1999). In addition, with the *lacZ* reporter mice, expression of the genes can also be detected at a cellular level using Fluorescein di-β-D-galactopyranoside (FDG). FDG is a fluorescent substrate of β-gal and is used in fluorescence activated cell sorting (FACS) analysis to detect β-gal activity in live cells. Therefore, using FDG in combination with other cell surface markers, one can determine the expression of *Bcl11* genes in specific cell types at a single cell level. Furthermore, β-

gal was found to be more effective in providing signal in the context of weak enhancers and to be extremely useful in high-resolution histochemical analysis (Timmons et al., 1997). Hence regions of low levels of *Bcl11* expression can be detected using the *Bcl11-lacZ* reporter mice.

In the first part of this Chapter, I will describe the use of *Bcl11-lacZ* reporter mice to characterize the spatial expression patterns of *Bcl11* genes in both embryonic and adult developmental stages in whole mount X-gal staining. Subsequently, I will describe the dynamic expression patterns of *Bcl11* genes in both the mammary epithelial and hematopoietic cells using FDG staining and flow cytometry.

4.2 Results

4.2.1 X-gal staining patterns faithfully recapitulate endogenous *Bcl11* expression

The X-gal staining patterns of *Bcl11a*$^{lacZ/+}$ and *Bcl11b*$^{lacZ/+}$ 10.5-11 dpc heterozygous embryos were first compared to whole mount *in situ* hybridization patterns (using antisense *Bcl11a* and *Bcl11b* RNA probes) obtained from VisiGene (http://genome.ucsc.edu/cgi-bin/hgVisiGene) (Figure 4.1). X-gal staining of *Bcl11a*$^{lacZ/+}$ and *Bcl11b*$^{lacZ/+}$ embryos revealed the expression of *Bcl11a* and *Bcl11b* in the forebrain, derivatives of the pharyngeal arches and limbs, which were similar to the RNA *in situ* hybridization patterns. Hence the X-gal staining pattern is a faithful recapitulation of endogenous *Bcl11* expression.

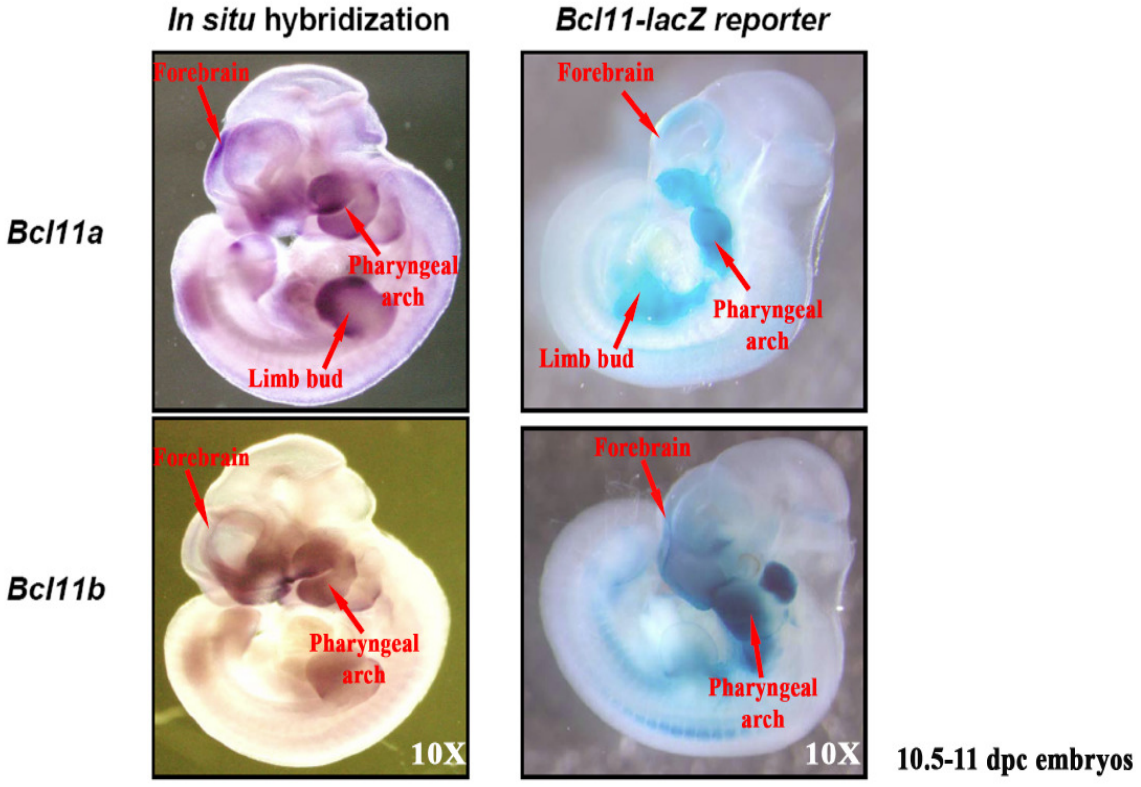


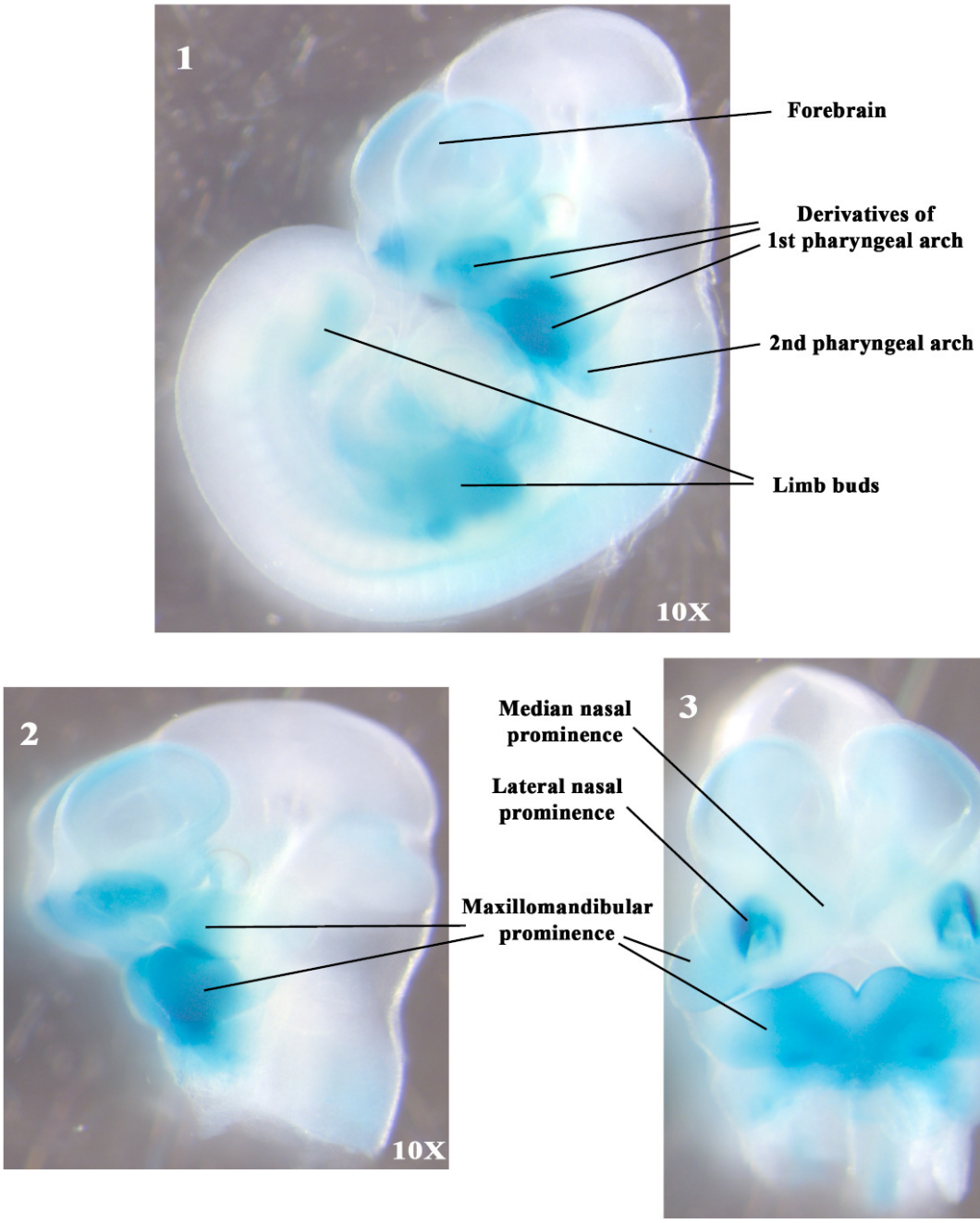
Figure 4.1. Validation of X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 10.5-11 dpc embryos. Images showing faithful recapitulation of X-gal staining patterns of both *Bcl11a*$^{lacZ/+}$ and *Bcl11b*$^{lacZ/+}$ embryos (right panels) as compared to *in situ* hybridization patterns (left panels; images obtained from Visigene).

4.2.2 *Bcl11* genes are expressed in early embryonic development

Expression of *Bcl11a* and *Bcl11b* was first observed at 10.5 dpc, consistent with a previous study (Nakamura et al., 2000). The expression patterns of the two genes partially overlapped at this stage (Figure 4.2A and 4.2B). Both genes were expressed in the forebrain and derivatives of the first and second pharyngeal arches (Figure 4.2A1 and Figure 4.2B1). In addition, high levels of expression of both genes were also detected in the lateral nasal and maxillomandibular prominence and at a lower level in the median nasal prominence (Figure 4.2A2-3 and Figure 4.2B2-3). Several regions of differential expression were also observed at this stage. For example, *Bcl11a* but not *Bcl11b*, was expressed in the limb buds (Figure 4.2A1 and Figure 4.2B1). In contrast, *Bcl11b* but not *Bcl11a* was observed in the trigeminal ganglion and nascent palate shelf (Figure 4B2 and Figure 4.2B4).

A

Bcl11a - 10.5 dpc



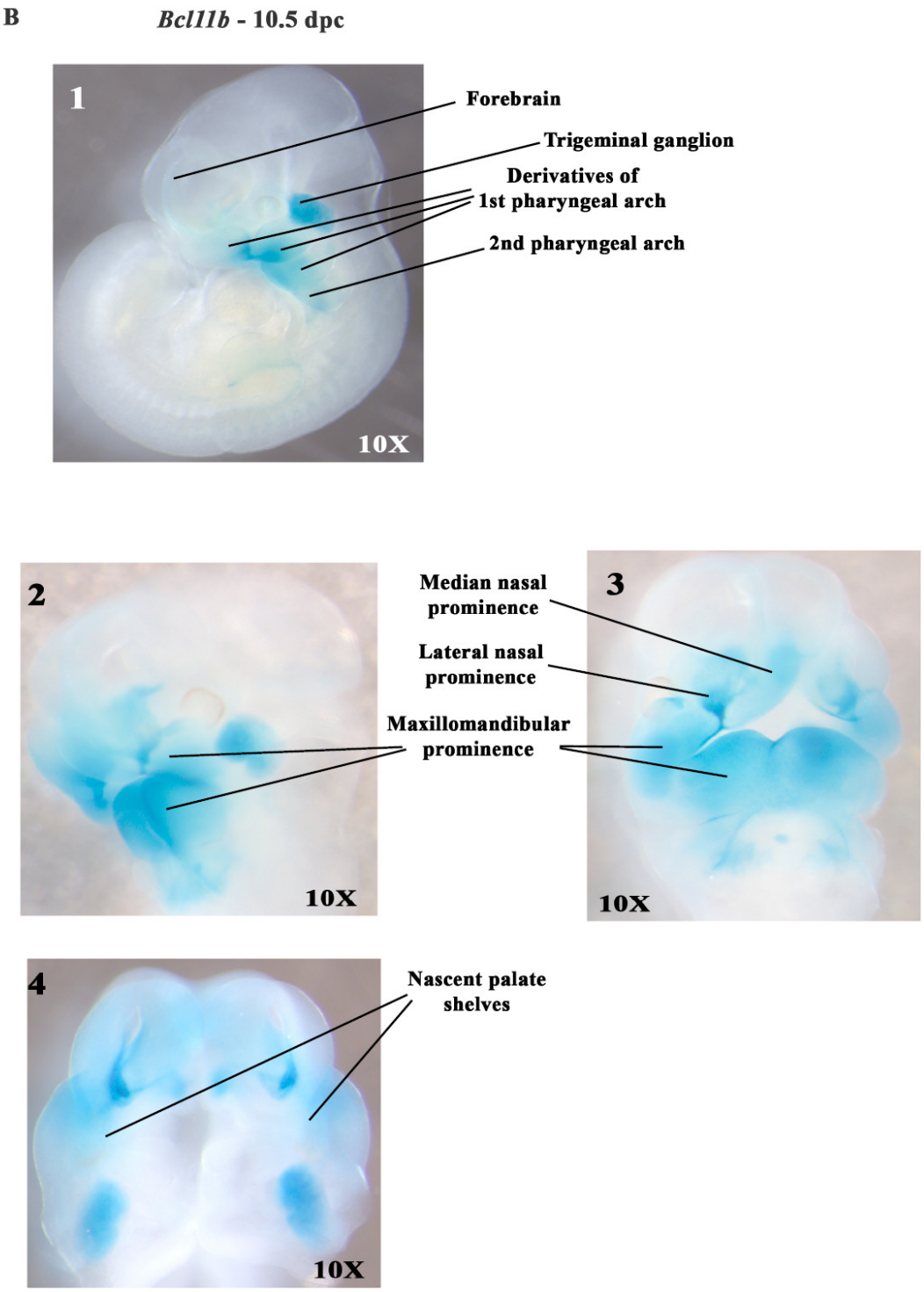
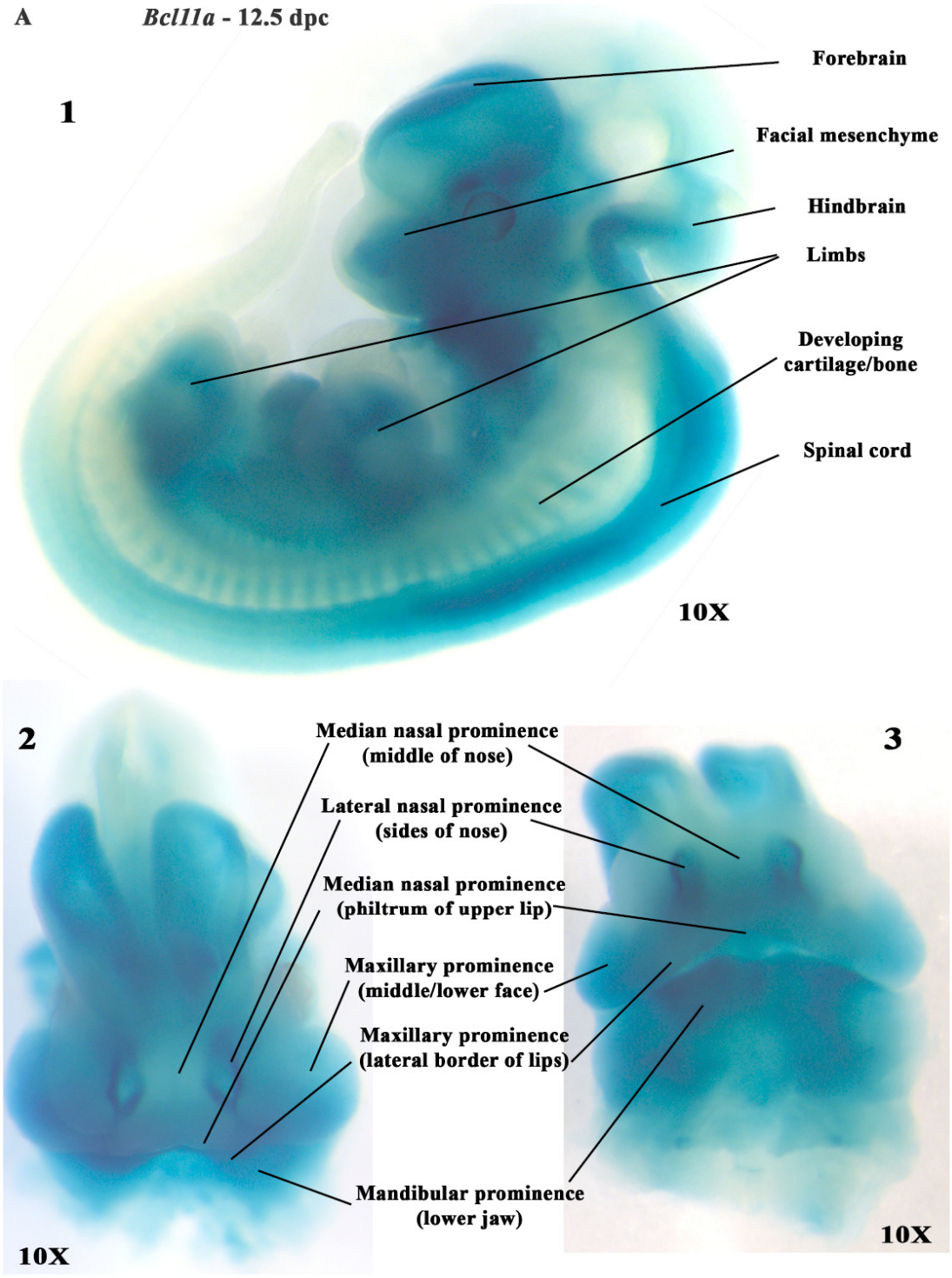


Figure 4.2. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 10.5-11 dpc embryos. **(A1)** *Bcl11a* is expressed in pharyngeal arches, limb buds and forebrain from 10.5 dpc. **(A2-3)** Expression of *Bcl11a* is detected in the maxillomandibular, medial and lateral nasal prominence. **(B1)** *Bcl11b* is expressed in pharyngeal arches, forebrain and trigeminal ganglion from 10.5 dpc. **(B2-3)** In addition, expression of *Bcl11b* is also detected in the maxillomandibular, medial and lateral nasal prominence and in **(B4)** the nascent palate shelves.

4.2.3 *Bcl11* genes are highly expressed in the brain and craniofacial regions

The expression patterns of both *Bcl11a* and *Bcl11b* were maintained at 12.5 dpc (Figure 4.3). Highest levels of *Bcl11* expression were detected in the central nervous system and the craniofacial mesenchyme. Within the brain, expression of *Bcl11a* was detected in the developing fore- and hind-brain (Figure 4.3A1). *Bcl11a* was found to be highly expressed in the lateral and median nasal prominence, the maxillary prominence and the mandibular prominence within the facial mesenchyme (Figure 4.3A2 and 4.3A3). Expression of *Bcl11b* in the facial mesenchyme overlapped with that of *Bcl11a*, suggesting that these genes might have complementary role(s) in development of the facial mesenchyme. In the brain, expression of *Bcl11b* was detected in the developing fore-brain (Figure 4.3B1-3). Intriguing, a differential spatial expression pattern of these two genes had begun to emerge at this stage. Within the CNS, while *Bcl11a* was highly expressed in the spinal cord (Figure 4.3A1); *Bcl11b* was found to be highly expressed in trigeminal ganglion and the dorsal root ganglion (Figure 4.3B1). This indicates that *Bcl11* genes may play different roles in development of specific regions of the CNS. Expression of *Bcl11* genes became more restricted from 14.5 dpc and high levels of expression were maintained in the CNS and craniofacial regions (Figure 4.4). Within the CNS, overlapping *Bcl11a* and *Bcl11b* expression was observed in forebrain (cerebral cortex), midbrain, hindbrain, the nasal epithelium and spinal cord (Figure 4.4A and 4.4B). Overlapping expression of *Bcl11* genes in the brain was maintained at 18.5 dpc (Figure 4.5) and in the adult (Figure 4.6). At 18.5 dpc, expression of both genes was observed in the cerebral hemisphere and in the medulla (Figure 4.5A and 4.5B). In the adult brain, expression of both genes was detected in cerebral cortex, olfactory bulb and in the cerebellum (Figure 4.6A1-2 and 4.6B1-2). Purkinje cells, which are large GABAergic neurons, are located within the purkinje cell layer of the cerebellum. These neurons extend dendritic projections upwards into the molecular layer and axonal projections downward to deep cerebellar nuclei. Purkinje cells are considered to be the principal output of the cerebellum controlling several components of descending motor pathways (Kreitzer and Regehr, 2001). Both *Bcl11* genes were expressed in the Purkinje cell layers (Figure 4.6A3 and 4.6B3). In summary, the unique expression of *Bcl11* genes is

maintained in the brain and craniofacial regions throughout embryonic and postnatal development, suggesting that *Bcl11* genes may play important, yet complementary roles in the brain.



B

Bcl11b - 12.5 dpc

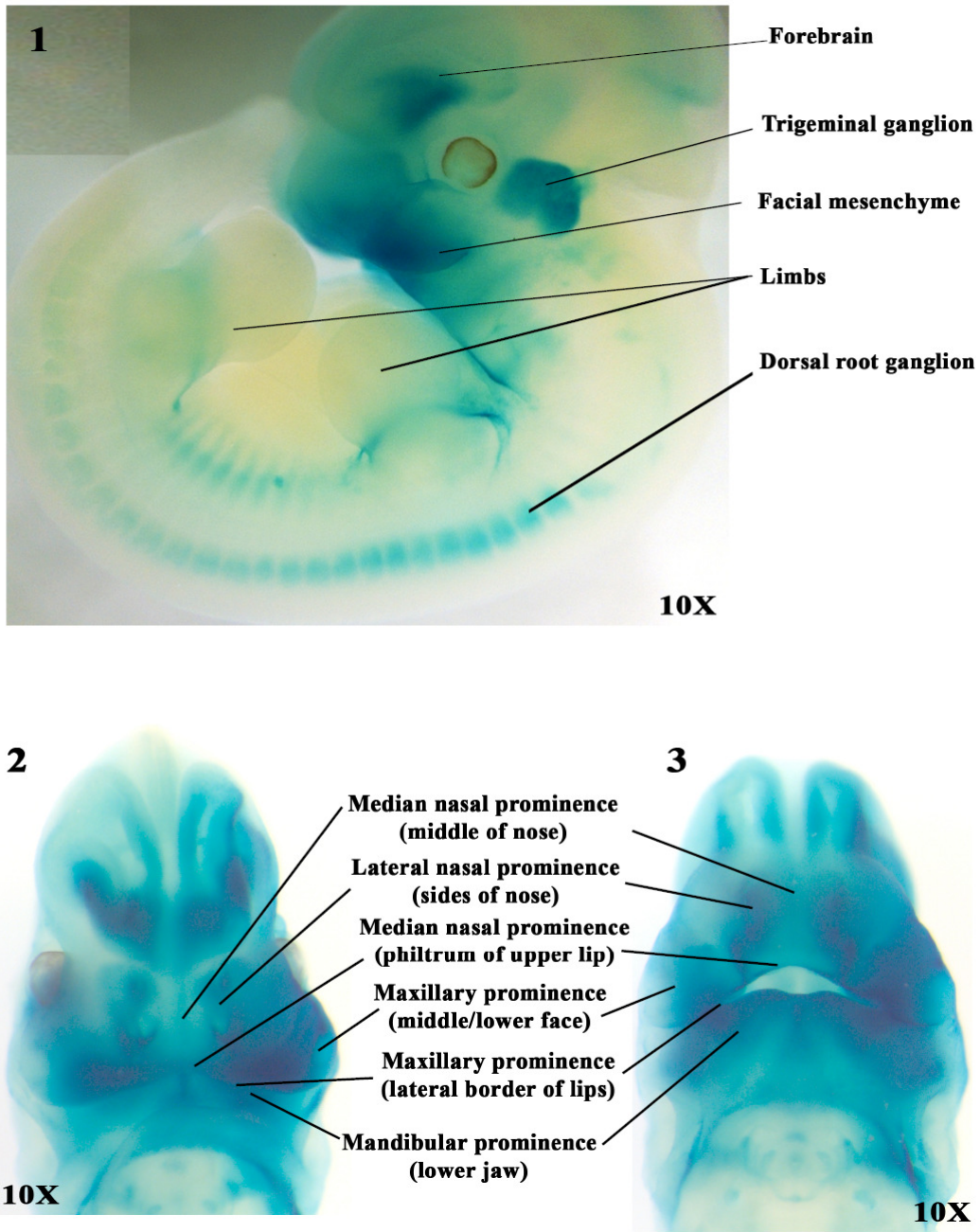
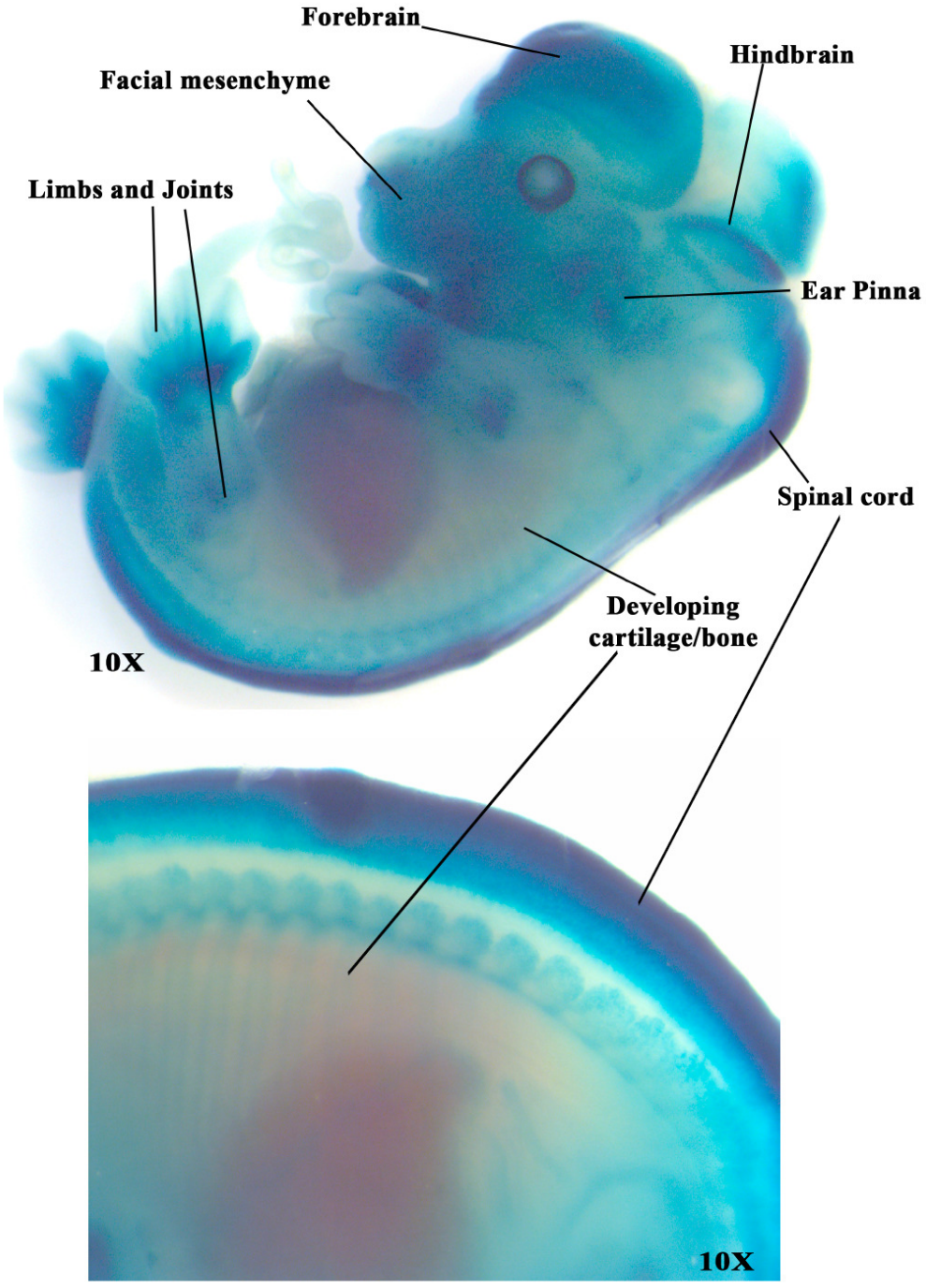


Figure 4.3. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 12.5-13 dpc embryos. **(A1)** *Bcl11a* is expressed in facial mesenchyme, limb buds, forebrain, hindbrain and spinal cord at 12.5 dpc. **(A2-3)** Expression of *Bcl11a* is maintained in all regions of the facial mesenchyme. **(B1)** *Bcl11b* is expressed in facial mesenchyme, forebrain, dorsal root ganglion and trigeminal ganglion from 12.5 dpc. **(B2-3)** In addition, expression of *Bcl11b* is also maintained in the facial mesenchyme, similar to *Bcl11a*.

A *Bcl11a* - 14.5 dpc



B ***Bcl11b*** **- 14.5 dpc**

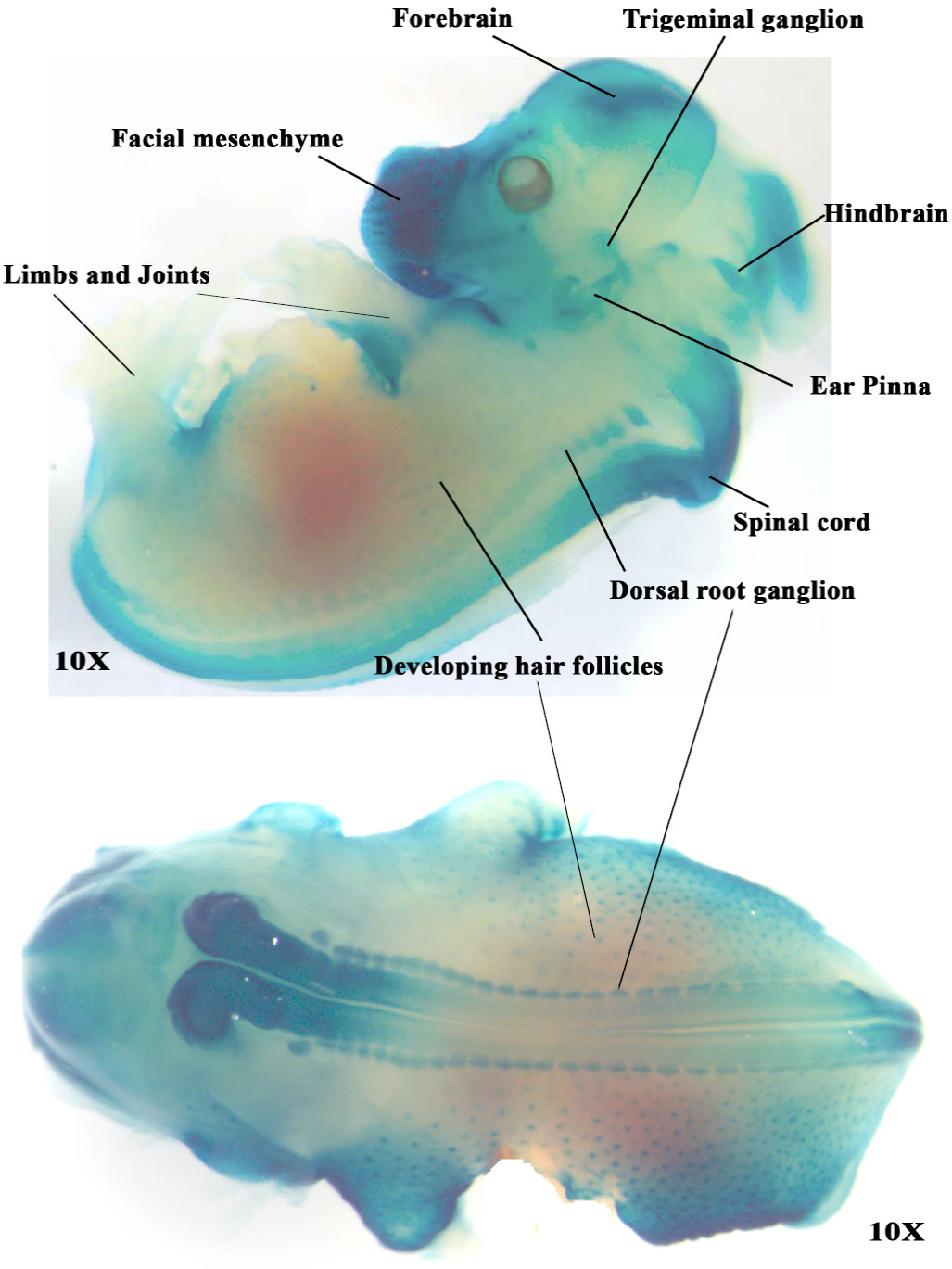


Figure 4.4. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 13.5-14.5 dpc embryos. (A) Expression of *Bcl11a* is maintained in facial mesenchyme, limbs and central nervous system. Expression of *Bcl11a* in facial mesenchyme is maintained. **(B)** Expression of *Bcl11b* is maintained in facial mesenchyme, central nervous system and also observed in hair follicles at 14.5 dpc.

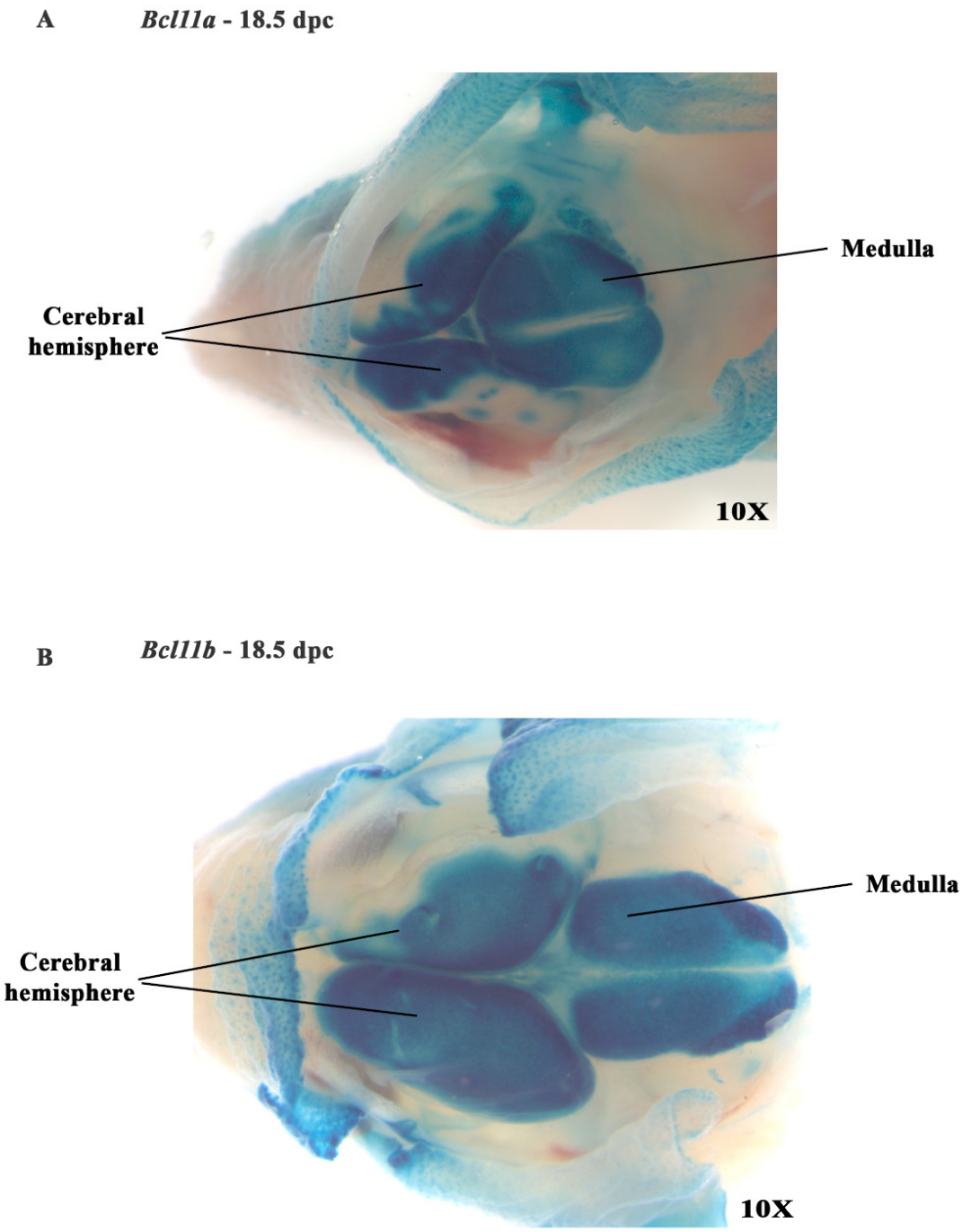


Figure 4.5. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 18.5 dpc embryos. **(A)** *Bcl11a* and **(B)** *Bcl11b* are highly expressed in cerebral hemispheres and medulla of the brain.

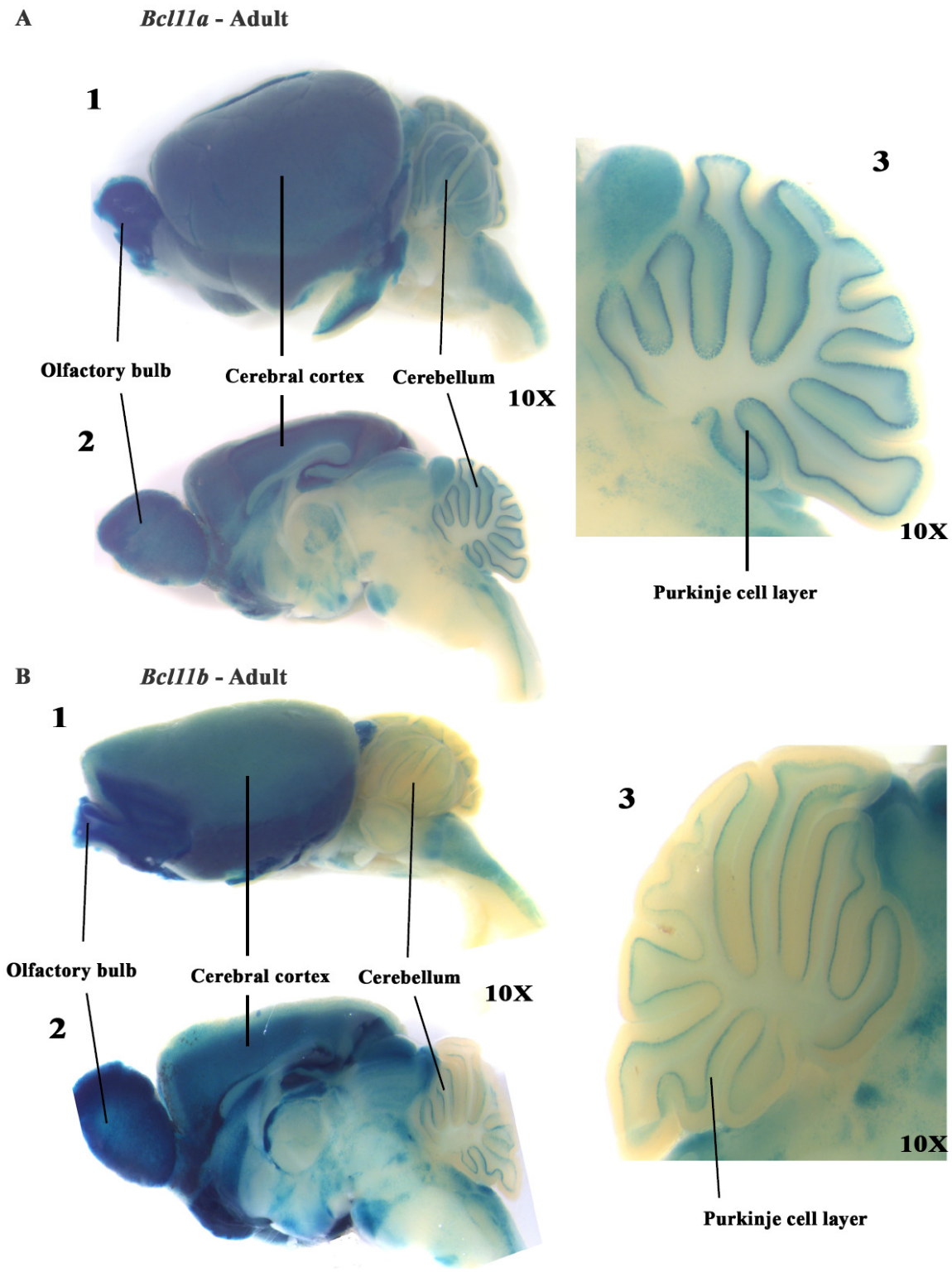


Figure 4.6. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ adult brain. Both **(A)** *Bcl11a* and **(B)** *Bcl11b* are highly expressed in the adult brain. Expression can be detected in the **(A1-2; B1-2)** olfactory bulb, cerebral cortex, cerebellum and also in **(A3 and B3)** purkinje cell layer of the cerebellum. **A1** and **B1** are lateral images of the brain while **A2** and **B2** are midline images of the brain. **A3** and **B3** are close-up images of the cerebellum.

4.2.4 *Bcl11* genes exhibit differential expression patterns in other tissues

Besides the craniofacial regions and the brain, expression of *Bcl11* genes was also detected in other tissues. At 12.5 dpc, *Bcl11a* but not *Bcl11b* was highly expressed in developing limb buds and cartilage/bone (Figure 4.3A1) which suggests that *Bcl11a* may play specific roles in limb and cartilage/bone formation. Similar to 12.5 dpc embryos, differential expression of *Bcl11* genes was maintained in the following regions: expression of *Bcl11a* was detected in the limbs and joints, developing cartilage/bones while expression of *Bcl11b* was detected in trigeminal ganglion at 14.5 dpc. Outside of the CNS, expression of both *Bcl11* genes was also detected in the ear pinna and the inner ear (Figure 4.4A and 4.4B), suggesting that *Bcl11* genes have roles in ear development. Additionally, *Bcl11b* was also found to be expressed in developing hair follicles (Figure 4.4B). This observation is consistent with the reported *Bcl11b* expression pattern using antibodies which showed that *Bcl11b* expression was detected in the rapidly dividing basal cell layer at 14.5 dpc (Golonzhka et al., 2007), suggesting that the possible involvement of *Bcl11b* in skin development.

Analyses of X-gal staining patterns of internal organs of the $Bcl11a^{lacZ/+}$ 14.5 dpc embryos revealed that *Bcl11a* was expressed in the heart, fetal liver and weakly in the thymus but not in the developing lungs (Figure 4.7A1-2). Hematopoiesis begins in the aorta-gonad-mesonephros (AGM) region from 8.5 dpc and continues in the fetal liver where hematopoietic stem cells differentiate into myeloid and lymphoid lineages. Expression of *Bcl11a* in the fetal liver and in hematopoietic cells (See below, Chapter 4.2.7.2) is consistent with its essential roles in hematopoiesis. In contrast, analyses of $Bcl11b^{lacZ/+}$ 14.5 dpc embryos revealed that *Bcl11b* was expressed in the oesophagus and in the developing lung but not in the heart (Figure 4.7B1-2). Intrathymic T cell development begins when fetal liver derived progenitor cells seed the thymus where they expand and differentiate into mature T cells (Rothenberg, 2007a, b). Expression of *Bcl11b* was detected in the developing thymus at 14.5 dpc at a time when most thymocytes are CD4 and CD8 double negative (T cell precursors), and its expression was maintained at 18.5 dpc (Figure 4.8B1) and in the adult thymus (Figure 4.9B). In addition, expression of *Bcl11b* was detected in thymocytes from $CD44^{+}CD25^{-}$ DN1 stage during T

cell development (See below, Chapter 4.2.7.2). These expression data are consistent with the *Bcl11b* knockout phenotype where *Bcl11b* is essential for αβ T cell development (Wakabayashi et al., 2003b). In contrast, expression of *Bcl11a* was only detected at low levels in the thymus at 18.5 dpc (Figure 4.8A1) and the $Bcl11a^{lacZ/+}$ adult thymus showed a punctate staining pattern (Figure 4.9A1). In the adult thymus, expression of *Bcl11a* was detected in only a small percentage of $CD44^{+}CD25^{-}$ thymocytes (DN1 stage) during T cell development (See below, Chapter 4.2.7.2). These observations, together with T cell defects in the *Bcl11a* knockout mice (Liu et al., 2003b) provide additional evidence that *Bcl11a* also plays an important role in T cell development.

Interestingly, dynamic reciprocal expression of *Bcl11* genes was detected in the developing lungs. Expression of *Bcl11a*, which was not detected in the lungs at 14.5 dpc, was detected at 18.5 dpc (Figure 4.7A2 and 4.8A2). In contrast, expression of *Bcl11b* was detected at 14.5 dpc and not in the lungs at 18.5 dpc (Figure 4.7B2 and 4.8B2). These dynamic expression patterns of *Bcl11* genes in the lungs highlight their potential roles during lung morphogenesis.

Taken together, *Bcl11a* and *Bcl11b* exhibit unique spatial and temporal expression patterns during embryonic development. Expression of both genes was detected from 10.5 dpc, with high levels of expression in the brain and derivatives of the first and second pharyngeal arches. The expression patterns in the brain and craniofacial mesenchyme were maintained right up to adulthood. Regions of differential *Bcl11* expression were observed in certain tissues, such as limbs, developing cartilage/bones, lungs, heart and thymus. Additionally, the differential expression patterns of *Bcl11* genes in the thymus were also maintained. *Bcl11a* was only expressed weakly in certain cells in the adult thymus while high levels of *Bcl11b* expression were maintained from the embryonic thymus to the adult thymus. Considered together, these findings suggest that these two genes may play complementary roles in the adult brain but not in the thymus.

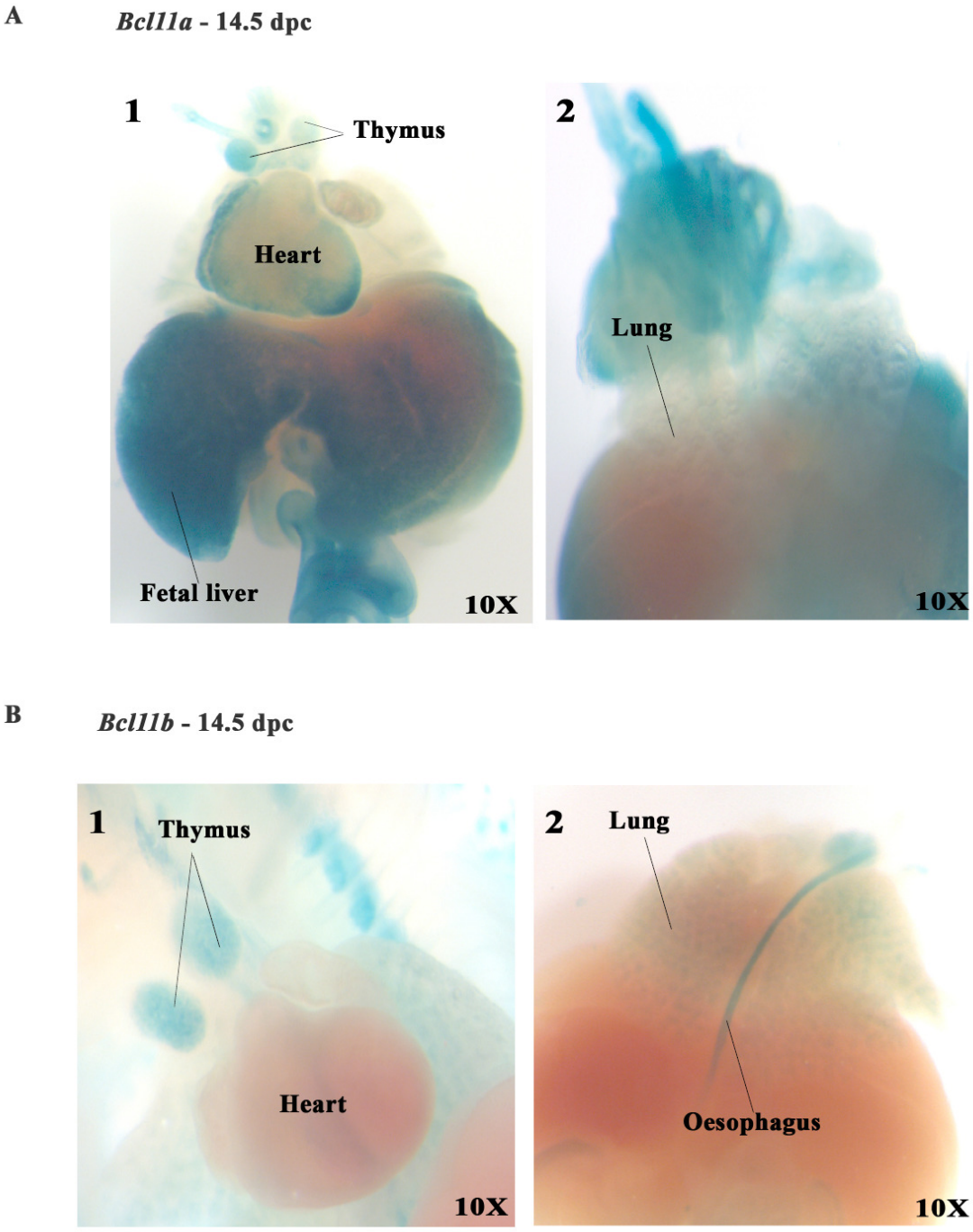


Figure 4.7. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 13.5-14.5 dpc tissues. Expression of *Bcl11a* is detected in the **(A1)** heart and thymus but not in the **(A2)** lungs. In contrast, expression of *Bcl11b* is not detected in the **(B1)** heart but only in the **(B1)** thymus and **(B2)** lungs.

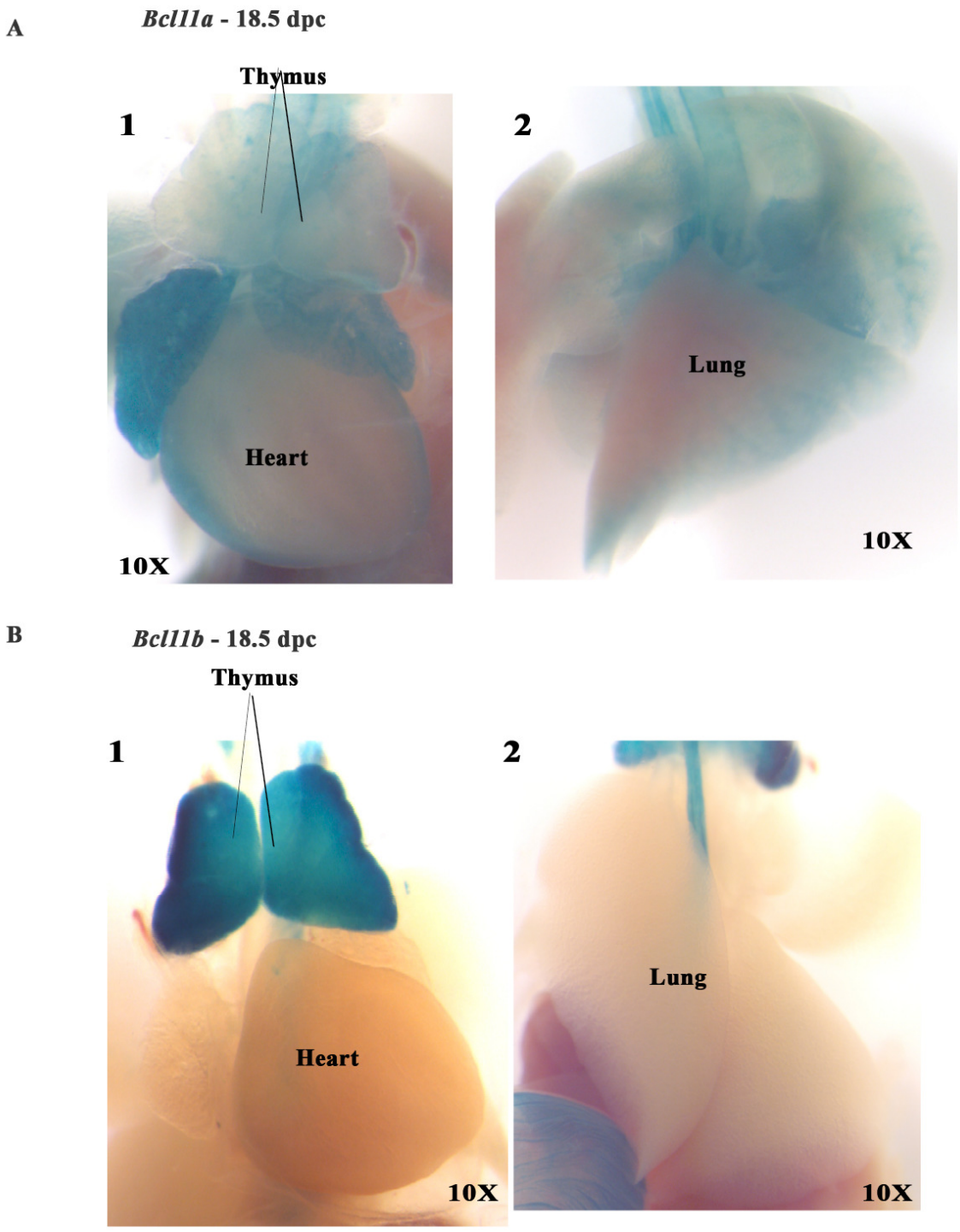


Figure 4.8. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ 18.5 dpc tissues. Expression of *Bcl11a* is maintained in the **(A1)** heart and certain regions of the thymus and also observed in **(A2)** certain regions of the lungs. Expression of *Bcl11b* is highly expressed in the **(B1)** thymus but not detected in the **(B1)** heart and **(B2)** lungs.

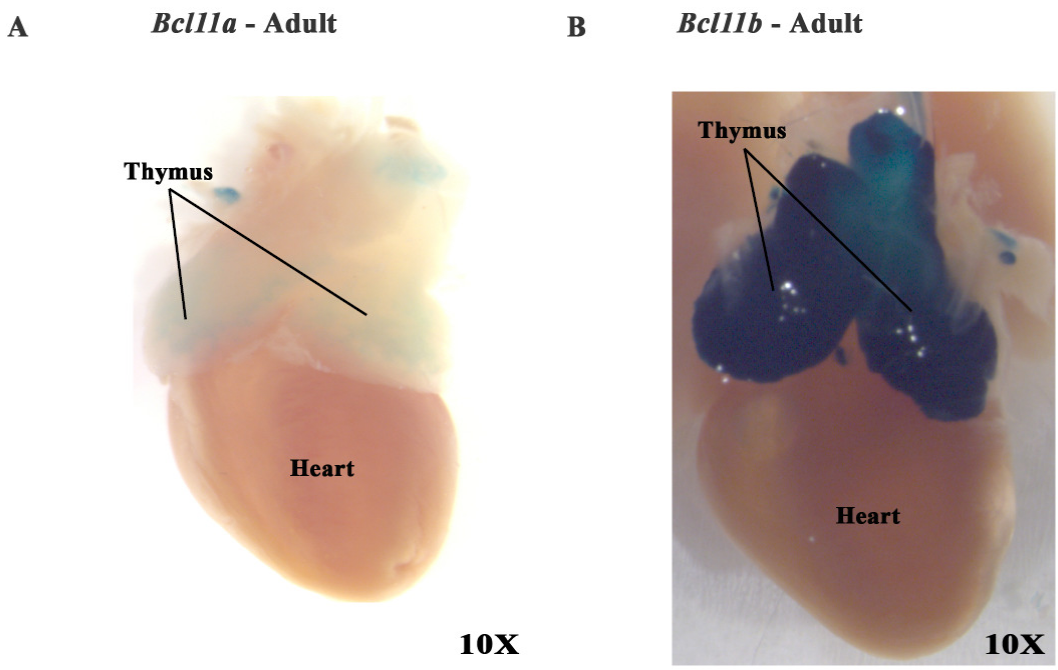


Figure 4.9. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ adult tissues. (A) Expression of *Bcl11a* is detected only in certain regions of the thymus but not in the heart. **(B)** *Bcl11b* is highly expressed in the thymus but not in the heart.

4.2.5 *Bcl11* genes are expressed specifically in embryonic mammary gland

Mammary gland development in the mouse is first observed at 10.5 dpc with the appearance of milk lines in both male and female embryos. The milk lines are two ridges of ectodermal thickenings that run in an anteroposterior (AP) direction between the fore- to the hind-limbs (Watson and Khaled, 2008). The development of the milk lines can be visualized using *in situ* hybridization with the *Wnt10b* and the *Lef1* probes (Foley et al., 2001; Veltmaat et al., 2004). These two genes are the earliest known markers of embryonic mammary development. Interestingly, expression of *Bcl11b* was detected in the presumptive regions of milk lines between the thoracic and inguinal regions at 10.5 dpc (Figure 4.10). Faint blue X-gal staining was observed to arise in an AP direction between the fore- and hind-limbs. No expression of *Bcl11a* was detected in the milk lines at 10.5 dpc. These observations indicated that *Bcl11b* is also one of the earliest genes to be specifically expressed in the milk lines at 10.5 dpc.

Mammary development becomes more distinctive from 11.5 dpc with the formation of five pairs of mammary placodes, each appearing in a specific order. By 12.5 dpc, the placodes have invaginated to form mammary buds (Watson and Khaled, 2008). Expression of *Bcl11a* remained undetectable in the mammary buds until 13.5-14.5 dpc (Figure 4.11A and 4.12A). Interestingly, expression of *Bcl11a* appeared to be in the surrounding mesenchyme of the mammary buds at 13.5-14.5 dpc (Figure 4.12A). After the initial expression in the milk line, *Bcl11b* became specifically expressed in developing mammary buds from 12.5 dpc (Figure 4.11B). This expression was maintained at 13.5-14.5 dpc where all five pairs of mammary buds expressed high levels of *Bcl11b* (Figure 4.12B). In summary, expression of *Bcl11b* started in the milk line at 10.5 dpc and from 12.5 dpc, its expression was detected in all the mammary buds. In contrast, expression of *Bcl11a* was detected only from 13.5 dpc onwards and its expression appeared to be in the mesenchyme of the mammary buds. These results clearly demonstrate that *Bcl11* genes are expressed in mammary lineages during early embryonic development.

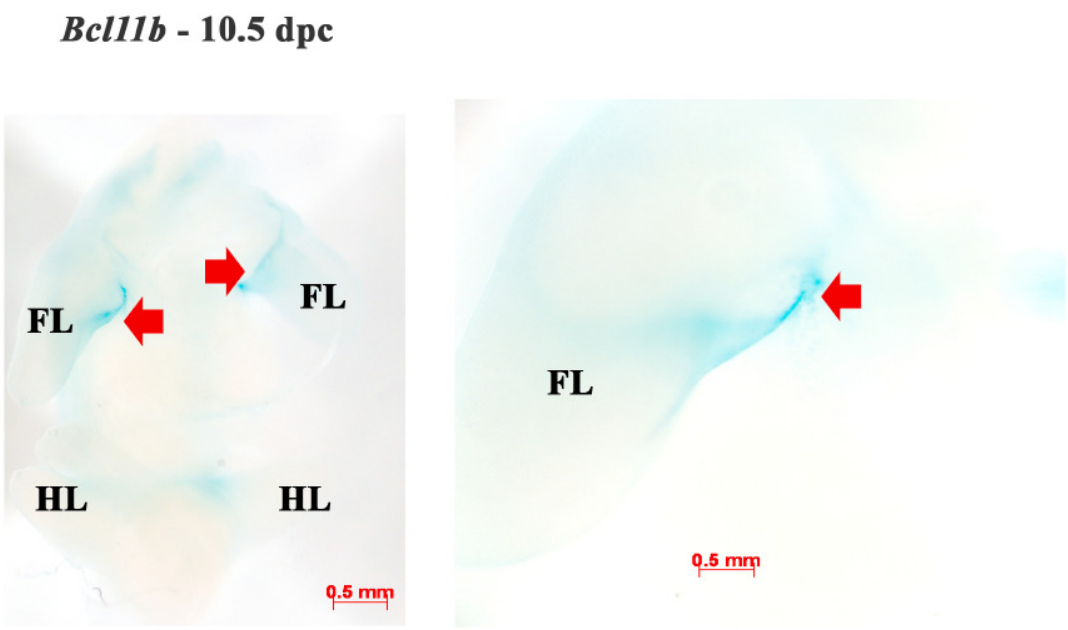


Figure 4.10. Expression of *Bcl11b* in milk line at 10.5 dpc. Expression of *Bcl11b* is detected in both the thoracic and inguinal milk lines from 10.5 dpc. Arrows indicate faint blue staining in the thoracic milk lines.

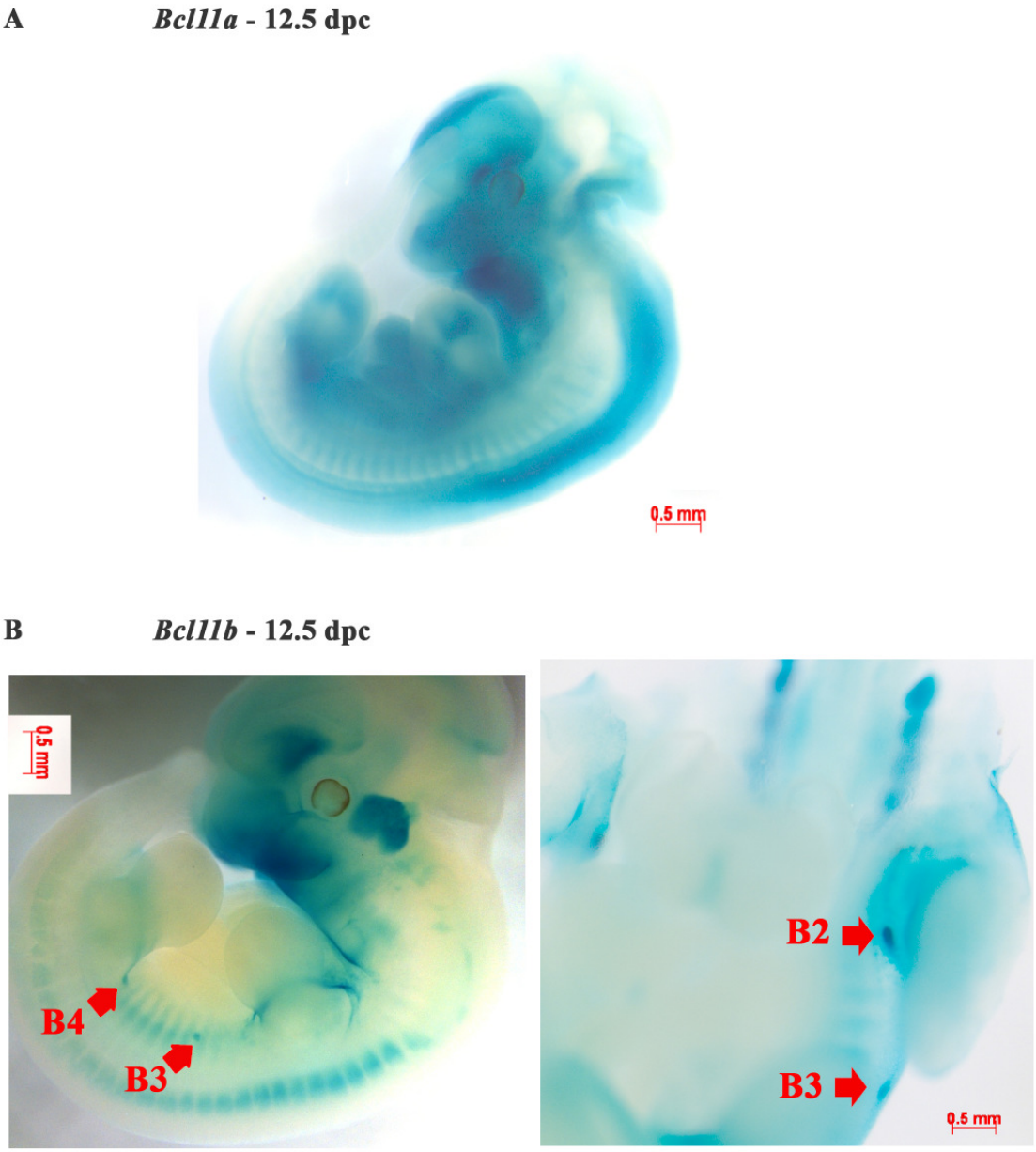


Figure 4.11. Expression of *Bcl11* genes in mammary lineages at 12.5 dpc. (A) Expression of *Bcl11a* is not detected at 12.5 dpc. **(B)** In contrast, expression of *Bcl11b* can be observed specifically in the mammary buds from 12.5 dpc. Arrows indicate positions of mammary buds. B2: 2nd pair of mammary buds; B3: 3rd pair of mammary buds; B4: 4th pair of mammary buds.

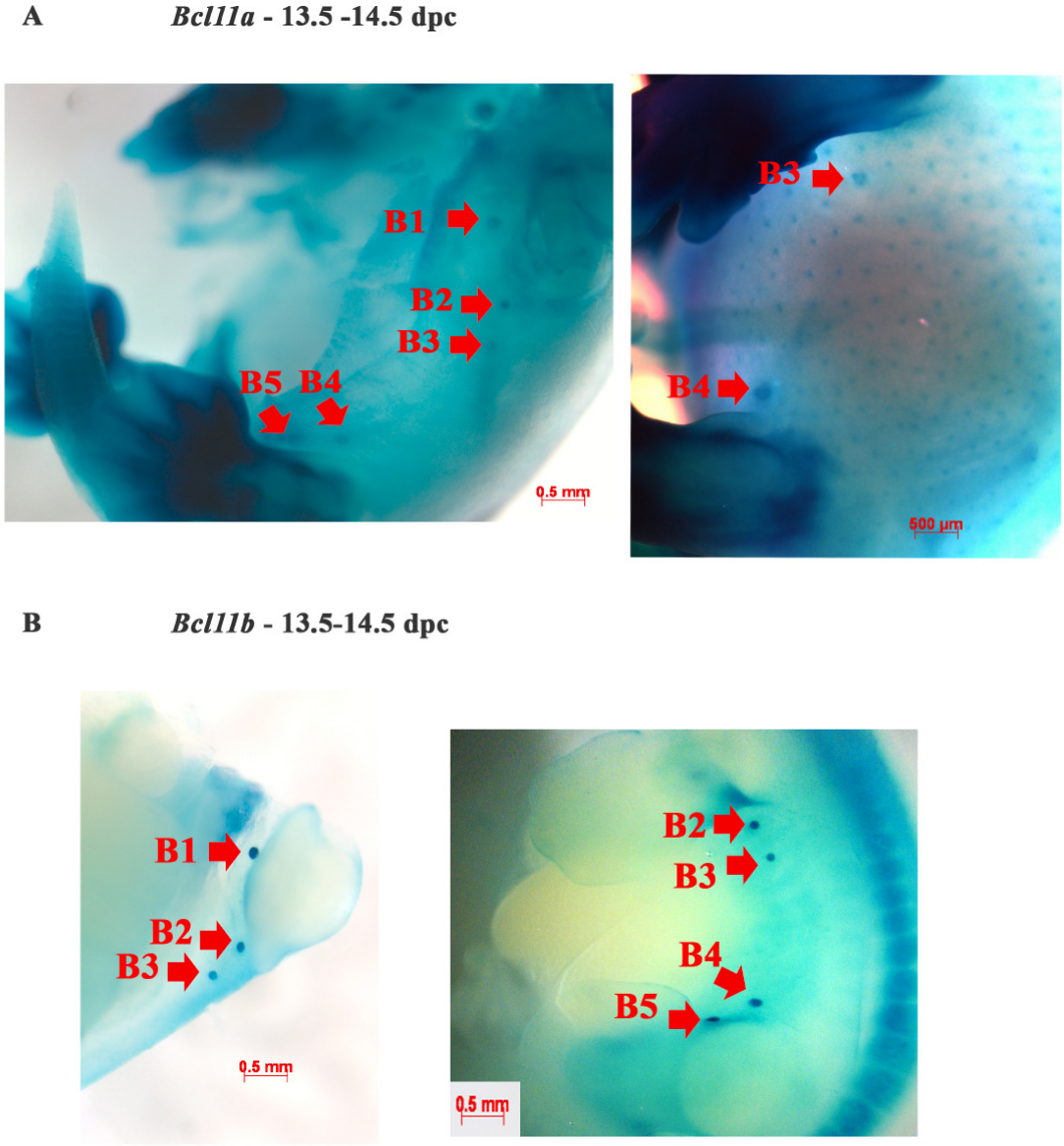


Figure 4.12. Expression of *Bcl11* genes in mammary lineages at 13.5-14.5 dpc. (A) Expression of *Bcl11a* is only detected in mammary buds from 13.5 dpc. Expression can be detected in all five pairs of mammary buds. Expression of *Bcl11a* appears to be predominantly in the surrounding mesenchyme.of mammary buds. **(B)** In contrast, expression of *Bcl11b* is detected in all five pairs of mammary buds. Arrows indicate positions of mammary buds. B1: 1st pair of mammary buds; B2: 2nd pair of mammary buds; B3: 3rd pair of mammary buds; B4: 4th pair of mammary buds; B5: 5th pair of mammary buds.

4.2.6 *Bcl11* genes exhibit unique and dynamic expression patterns in the mammary gland

4.2.6.1 *Bcl11a* is expressed in terminal end buds of mammary glands

Functional development of the mammary gland occurs in distinct stages; during embryonic development, the mammary anlage is established. By 16 dpc, the rudimentary mammary gland has become arborized and the ductules have begun to invade the underlying fat pad. After birth, ductal elongation occurs at a rate proportional to the overall growth rate of the animal. Following the onset of puberty at around 4 weeks, accelerated ductal elongation and branching morphogenesis occur, stimulated by estrogen hormone secretion. At this stage, terminal end buds (TEBs) which are large club-shaped structures appear at the end of growing ducts. The TEBs bifurcate to give rise to side branches and also lead the invasion of the fad pad by the growing ducts. As shown in Figure 4.13A1, *Bcl11a* was expressed in TEBs while expression of *Bcl11b* was detected in the neck region of TEBs (Figure 4.13A2). TEBs are specialized structures consisting of an outer layer of cap cells and the inner multi-layered body cells (Humphreys et al., 1996). The body cells give rise to mammary luminal epithelial cells and the cap cells are believed to contain mammary progenitor cells and precursors of myoepithelial cells. Sections of X-gal stained *Bcl11a*$^{lacZ/+}$ 4-5 weeks old mammary glands revealed that expression of *Bcl11a* was detected in both the cap and body cells of TEBs (Figure 4.13A3). In contrast, expression of *Bcl11b* was only detected in some cells found in the cap cell layer which are destined to become the basal/myoepithelial layer (Figure 4.13A4).

A highly regulated process of cell proliferation and apoptosis occurs within the highly proliferative TEBs (Humphreys et al., 1996). This ultimately generates the mature mammary epithelium which consists of two main cell types: the luminal cells which line the innermost layer of the lumen and forms the ducts and secretory alveoli, and the basal/myoepithelial cells that surrounds the luminal cells and provide contractile forces to facilitate transport of the milk to the nipple. In the mature mammary gland (8-12 weeks), *Bcl11a* was found to be expressed in both luminal and basal cells as well as in the alveolar buds during estrus cycle (Figure 4.13B1 and B3). Expression of *Bcl11b*, on the

other hand was detected primarily in the basal/myoepithelial layers in the mature epithelium (Figure 4.13B2 and B4). The differential expression patterns of *Bcl11* genes in the virgin mammary epithelium suggest that *Bcl11a* may be important in the establishment of the luminal and basal lineages while *Bcl11b* may be important for the basal lineage.

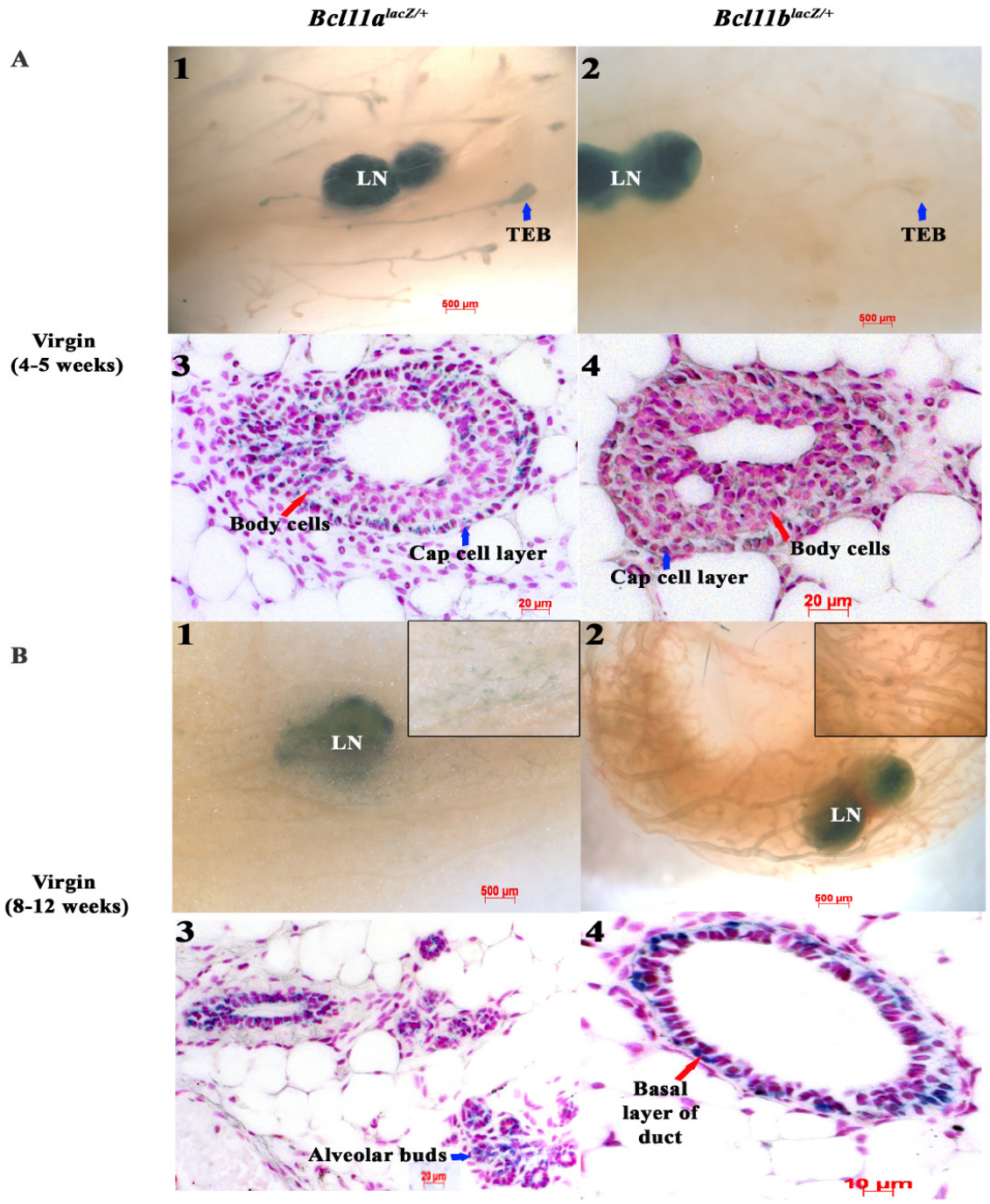


Figure 4.13. X-gal staining patterns of mammary tissues from *Bcl11*$^{lacZ/+}$ virgin glands. **(A1)** *Bcl11a* is highly expressed in terminal end buds (TEBs) of 4-5 weeks old virgin mammary glands. **(A3)** Sections show that *Bcl11a* is expressed in both cap and body cells of TEBs. **(A2)** In contrast, *Bcl11b* is not expressed in TEBs but at the neck regions of TEBs. **(A4)** Sections show the expression of *Bcl11b* is restricted to few cap cells and the developing myoepithelial/basal layers. In mature virgin females (8-12 weeks), **(B1)** low levels of *Bcl11a* expression is detected in the differentiated alveolar structures observed during estrus cycles and **(B3)** sections indicate that *Bcl11a* is detected in both luminal and basal layers of the mammary glands. However, expression of *Bcl11b* is restricted predominantly to **(B2 and B4)** basal layers of the mammary glands. TEB indicates terminal end bud; LN indicates lymph node.

4.2.6.2 Differential expression of *Bcl11* genes during pregnancy, lactation and involution

During pregnancy, the mammary gland undergoes extensive proliferation, differentiation and remodelling in preparation for lactation in order to nurse the pups. An important morphological change in the mammary gland is the production of lobulo-alveolar structures which eventually form secretory epithelial cells that produce milk to feed the newborn pups. Progesterone and prolactin signalling are critical in preparing the gland for gestation and lactation. Progesterone induces extensive side-branching and alveologenesis, while prolactin promotes differentiation of the alveolar structures. The alveolar structures, which consist of alveolar luminal cells and the surrounding myoepithelial cells, might arise from bi-potent ductal progenitors, though there is evidence that a distinct alveolar progenitor population exists within the mammary lineage hierarchy (Smith and Boulanger, 2003). Early in gestation (day 4-5), extensive proliferation and side-branching occur within the mammary gland. At this stage, up-regulation of both *Bcl11* genes was observed in the mammary epithelium (Figure 4.14A). Expression of *Bcl11a* was observed in the ducts and in the differentiating lobulo-alveolar structures (Figure 4.14A1) while expression of *Bcl11b* was observed only in the mammary ducts and not in the lobulo-alveolar structures (Figure 4.14A2). Sections of X-gal stained $Bcl11^{lacZ/+}$ gestation glands showed that *Bcl11a* was expressed in both the ductal luminal cells and in the differentiating alveoli (Figure 4.14A3) while expression of *Bcl11b* was restricted to the basal/myoepithelial layer of mammary ducts and not in alveolar cells (Figure 4.14A4). The differential expression patterns were maintained throughout gestation (Figure 4.14B1 and B2) where expression of *Bcl11a* was detected in luminal ductal and alveolar cells (Figure 4.14B3) while expression of *Bcl11b* remained restricted to the basal/myoepithelial layer of ducts (Figure 4.14B4). These results indicated that *Bcl11a* was highly expressed in differentiating luminal cells of the ducts and alveolar while expression of *Bcl11b* was found primarily in the basal/myoepithelial layer of mammary ducts, suggesting that *Bcl11a* may be important for luminal differentiation.

A lactogenic switch occurs during late pregnancy that is accompanied by the expression of milk proteins, whey acidic protein (WAP) and α-lactalbumin and by the

formation of lipid droplets (Watson and Khaled, 2008). During lactation, the mammary gland undergoes a morphological change. The alveolar structures become expanded that eventually filled up the entire mammary gland. The lumens of the alveoli become distended and filled with milk which is expelled from the secretory alveoli by the contraction of the myoepithelial layers surrounding the ducts and alveoli. Expression of *Bcl11a* was detected during lactation and sections of the X-gal stained *Bcl11a*$^{lacZ/+}$ day 3 lactation glands showed that *Bcl11a* was expressed in secretory luminal cells (Figure 4.15-1 and -3). Interestingly, not all the secretory luminal cells stained positive for *Bcl11a*. Expression of *Bcl11b* was undetectable during lactation (Figure 4.15-2 and -4).

Following lactation and weaning of the pups, the mammary gland undergoes a dramatic process involving apoptosis, dedifferentiation, tissue remodelling and immune infiltration called involution in order to remove the now redundant secretory cells (Watson, 2006a). At 72 hours post removal of the pups, expression of *Bcl11a* was present in the mammary gland undergoing involution (Figure 4.16-1). Surprisingly, *Bcl11b*, which was not detected during lactation, was up-regulated in the mammary gland at 72 hours post initiation of involution (Figure 4.16-2). Sections of X-gal stained *Bcl11*$^{lacZ/+}$ involution glands showed that *Bcl11* expression was detected in both epithelial cells and immune cells (Figure 4.16-3 and -4).

To confirm the X-gal staining, the expression changes of *Bcl11a* and *Bcl11b* during the adult mammary gland development cycle were determined using quantitative Real-Time PCR (qRT-PCR) (qRT-PCR performed by Dr Walid Khaled) (Figure 4.17). In total epithelial cells, there was dramatic up-regulation of *Bcl11* mRNA levels during early gestation. This is consistent with the X-gal staining patterns in the *Bcl11*$^{lacZ/+}$ adult mammary glands (Figure 4.14). *Bcl11a* mRNA levels remained high throughout gestation and lactation, suggesting that *Bcl11a* plays important roles in luminal epithelial cell differentiation. In contrast, *Bcl11b* expression levels decreased steadily over gestation and by lactation, its expression was virtually undetectable. During involution, there was a dramatic up-regulation of *Bcl11b* levels at 24 hours after initiation of involution, followed by a sharp decline before peaking again at 96 hours involution time point. These results suggest that *Bcl11b* plays important roles during involution, especially during the first phase of involution. On the other hand, levels of *Bcl11a* increased gradually during

involution and peaked at 72 hours involution, indicating that *Bcl11a* may also have a functional role during the second phase of involution. Taken together, qRT-PCR confirmed the dynamic expression of *Bcl11a* and *Bcl11b* during mammary gland development. These expression results laid the groundwork for the functional analysis of the roles of *Bcl11* genes in epithelial cell proliferation, differentiation and remodelling.

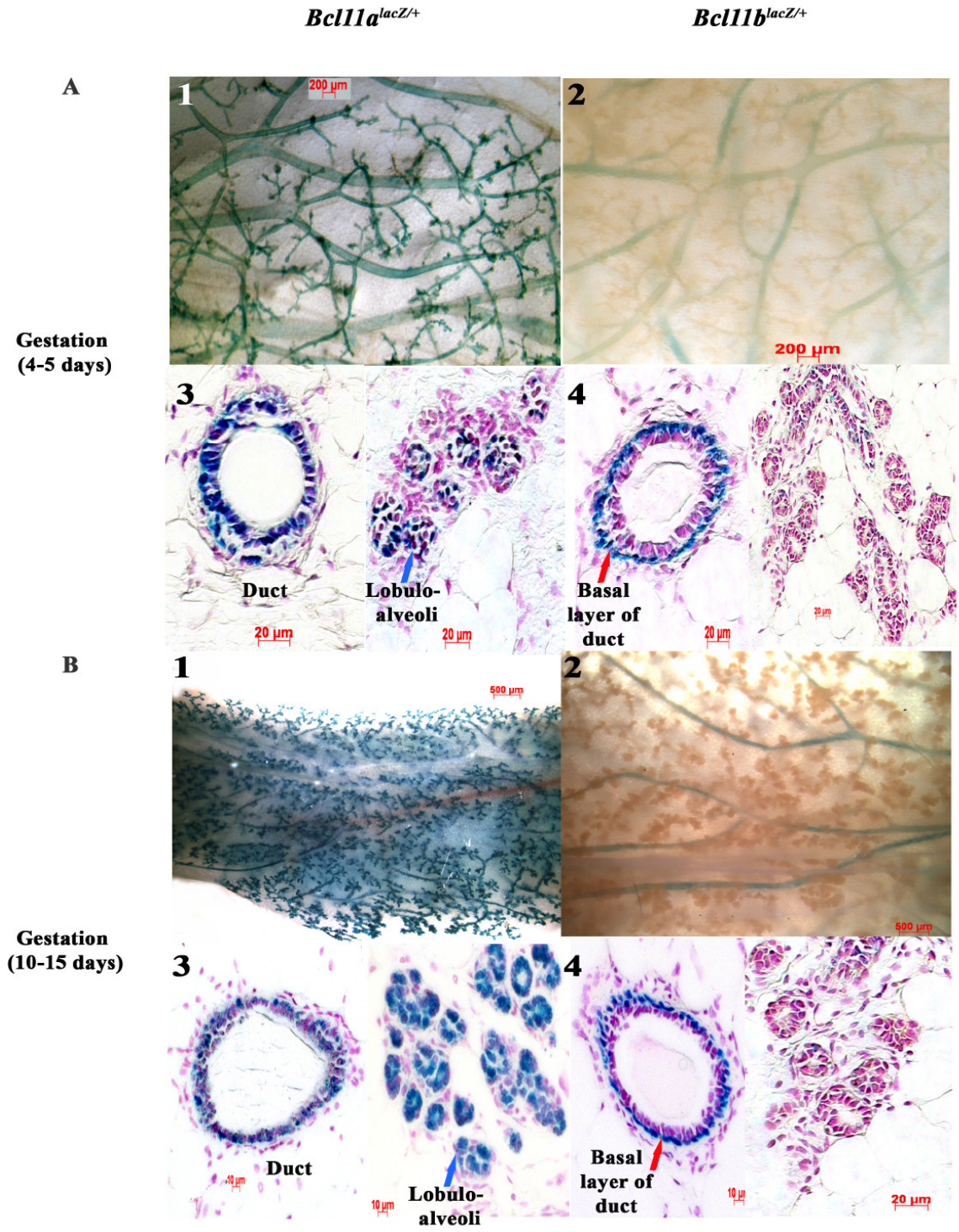


Figure 4.14. X-gal staining patterns of mammary tissues from *Bcl11*$^{lacZ/+}$ gestation glands. **(A1-2)** Both *Bcl11* genes are up-regulated during early gestation. **(A3)** Expression of *Bcl11a* is detected in both ductal luminal cells and in differentiating lobulo-alveolar structures. **(A4)** In contrast, expression of *Bcl11b* is detected only in ducts of the gestation glands and sections show that its expression is restricted primarily to basal layers of the ducts. **(B1-2)** The differential expression patterns are maintained throughout gestation. During late gestation, **(B3)** *Bcl11a* is detected in ductal cells and in differentiated lobulo-alveolar cells while **(B4)** *Bcl11b* is only detected in the basal layer of ducts but not in differentiated lobulo-alveolar cells.

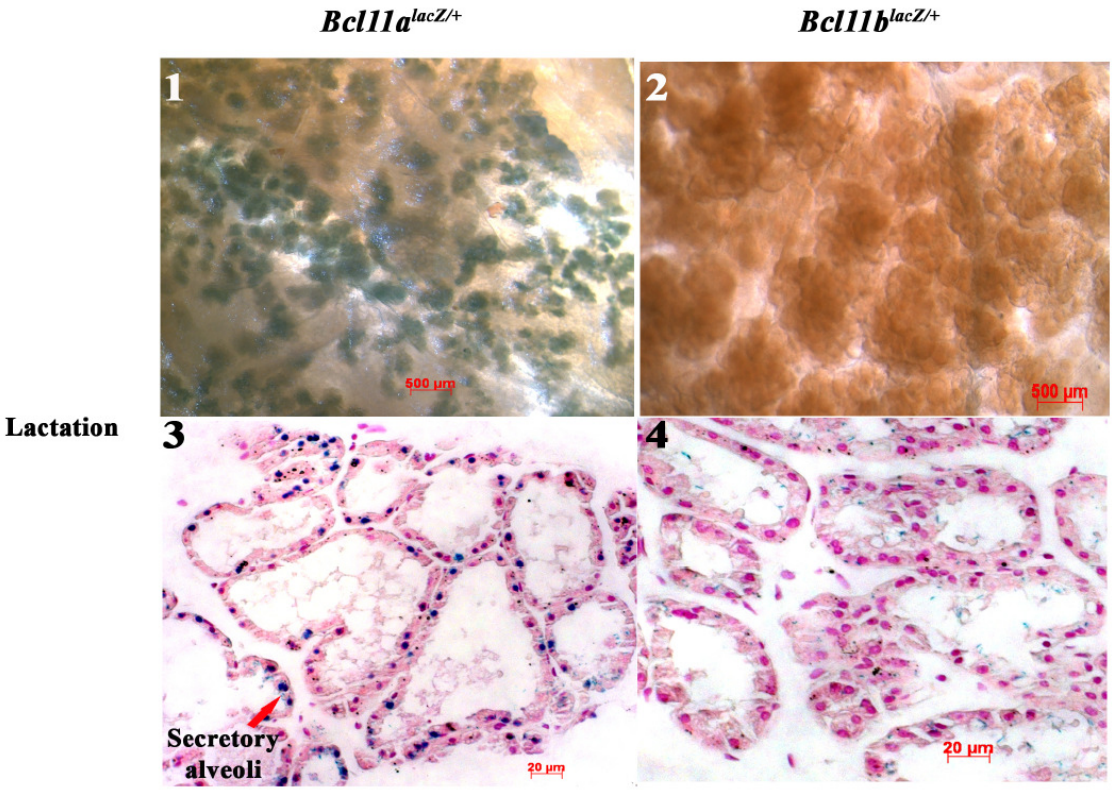


Figure 4.15. X-gal staining patterns of mammary tissues from *Bcl11*$^{lacZ/+}$ lactation glands. (1-2) Only expression of *Bcl11a* is detected during lactation. *Bcl11a* expression is detected in lobulo-alveolar structures and sections show that **(3)** *Bcl11a* is expressed in secretory luminal cells. **(4)** Expression of *Bcl11b* is undetectable in secretory luminal cells.

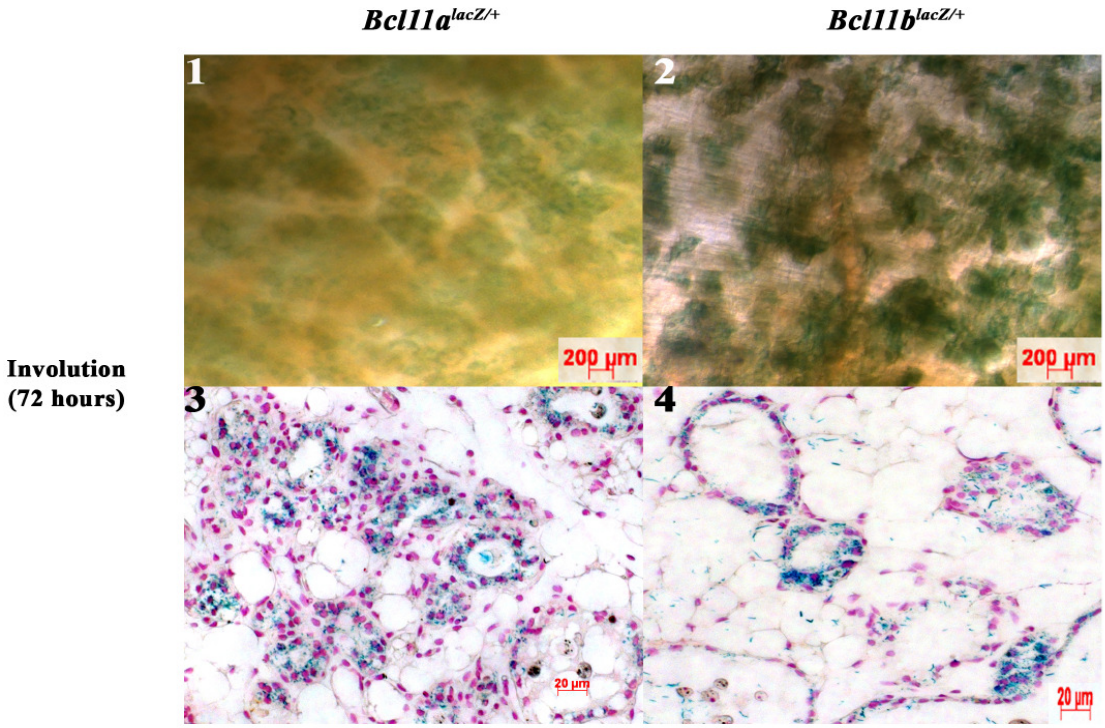


Figure 4.16. X-gal staining patterns of mammary tissues from *Bcl11*$^{lacZ/+}$ involution glands. (1-2) Expression of both *Bcl11* genes is detected during involution (72 hours). **(3-4)** Sections show that both genes are detected in some epithelial cells and also in the infiltrating immune cells.

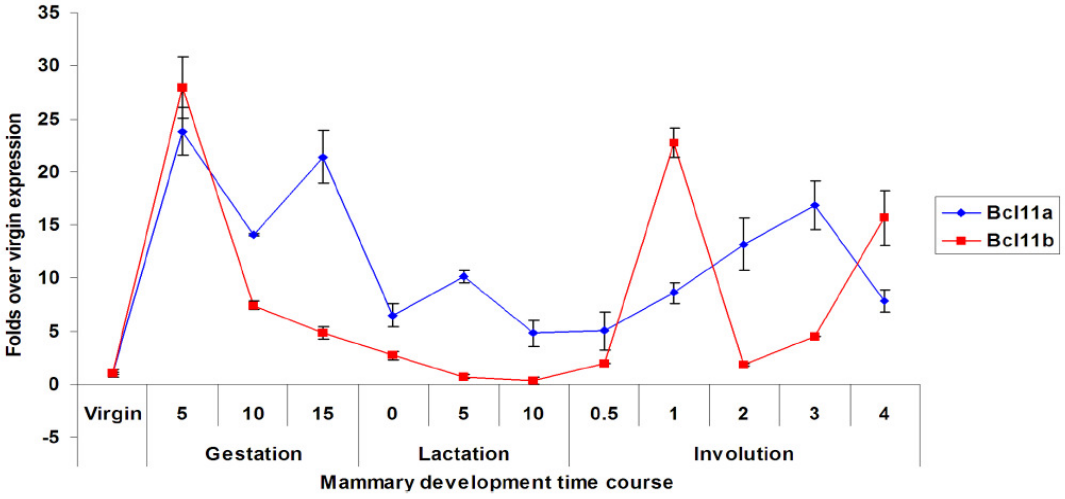


Figure 4.17. Expression patterns of *Bcl11* genes over Mammary Gland Development time course. Dynamic expression of *Bcl11* genes over mammary gland development time course as detected by quantitative real time PCR (qRT-PCR). Error bars denote standard deviation obtained from 3 independent samples. qRT-PCR performed by Dr Walid Khaled.

4.2.7 Expression of *Bcl11* genes in specific cell types

4.2.7.1 Characterization of *Bcl11*$^{lacZ/+}$ mammary epithelial cells

The mammary gland is a ductal epithelial organ that consists of two epithelial cell types: luminal epithelial cells, which line the ductal lumen and secrete milk proteins, and the myoepithelial or basal cells, which line the basal surface of the luminal cells and interact with the stroma. Both types of cells are thought to arise from a multi-potent stem or progenitor population that has been recently characterized (Kordon and Smith, 1998; Shackleton et al., 2006; Stingl et al., 2006). The dichotomy in mammary epithelium bears many similarities to hematopoiesis whereby both B cells and T cells are believed to derive from the common lymphoid progenitors (CLPs). Indeed several studies have already demonstrated that genes involved in lymphoid lineage specification play important roles in the mammary gland (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006). Similar to the field of hematopoiesis, the use of specific cell surface markers to delineate different populations of mammary epithelial cells has greatly facilitated characterization of the hierarchy of mammary epithelial population. Using antibodies to the heat stable antigen (CD24) in combination with either α6-integrin (CD49f) or β1-integrin (CD29), the mammary epithelial cells can be separated into the luminal ($CD24^{hi}CD29f^{+}/CD49f^{+}$) and basal/myoepithelial ($CD24^{+}CD29f^{hi}/CD49f^{hi}$) fractions using flow cytometry (Shackleton et al., 2006; Stingl et al., 2006) (Figure 4.18A).

To determine the expression of *Bcl11* genes in specific luminal and basal cell types, RT-PCR and qRT-PCR were performed on FACS-sorted mammary luminal/basal epithelial cells based on cell surface markers CD24 and CD49f (Stingl et al., 2006). The purity of each epithelial population was verified with lineage specific genes, *CK14* (basal) and *Muc1* (luminal). Consistent with the X-gal staining patterns as described above, *Bcl11a* transcripts were amplified from both the luminal and basal compartments of the virgin and gestation glands (Figure 4.18B). Up-regulation of *Bcl11a* expression was detected at day 5 gestation and this increase was predominantly in the luminal compartment (Figure 4.18B). Quantification using qRT-PCR showed that there was an approximately 10 fold increase in *Bcl11a* levels in the luminal compartment of the day 5 gestation gland compared to the virgin gland (Figure 4.18C). *Bcl11b* on the other hand

was amplified almost exclusively from the basal compartment in both the virgin and gestation glands (Figure 4.18B). Interestingly, there was a slight increase in the levels of *Bcl11b* in the luminal population during day 5 gestation as quantified by qRT-PCR (Figure 4.18C). As shown in Figure 4.19C, expression of *Bcl11b* in luminal lineages was detected in a small number of $Sca1^-$ ($ER\alpha^-$) putative alveolar progenitors. Therefore the increase of *Bcl11b* in the luminal fraction was most likely attributed to the expansion of putative alveolar progenitor population during early gestation.

To further reveal the expression of *Bcl11a* and *Bcl11b* in the mammary epithelium, I stained the epithelial cells with antibodies to several cell surface markers and incubated them with FDG before analysis with flow cytometry. FDG is a fluorescent substrate of β-galactosidase for detecting its activity in live cells and can be used to determine the expression of *Bcl11* genes in specific mammary epithelial cells. Consistent with the whole mount X-gal staining and RT-PCR, $Bcl11a^{lacZ/+}$ FDG positive epithelial cells in mice were located within both luminal ($CD24^{hi}CD49f^+$) and basal ($CD24^+CD49f^{hi}$) compartments while the $Bcl11b^{lacZ/+}$ FDG positive epithelial cells were distributed primarily in the basal compartment where the mammary stem cells (MaSC) and progenitor cells are thought to be localized (Stingl et al., 2006) (Figure 4.19A). Expression of *Bcl11a* was detected in about 14.3 $\pm$ 1.8 % of luminal cells and 6.8 $\pm$ 0.9% of basal cells. In contrast, expression of *Bcl11b* was detected in only 2.1 $\pm$ 1.2% of luminal cells and 6.8 $\pm$ 0.3% of basal cells.

I then analyzed the $Bcl11^{lacZ/+}$ FDG positive luminal epithelial cells ($CD24^{hi}$) in greater detail based on cell surface markers CD49b (α2-integrin) and Sca1 (Ly-6A/E) (Figure 4.19B). Luminal mammary epithelial cells can be classified into progenitor ($CD49b^+$) or differentiated ($CD49b^-$) cells based on the presence or absence of CD49b (Stingl and Watson; manuscript in preparation). The luminal progenitor population can be separated into $Sca1^-$ and $Sca1^+$ luminal progenitors. The $CD49b^+Sca1^+$ luminal progenitor subset has high levels of ERα expression while almost all the $CD49b^+Sca1^-$ cells are ERα-negative (Stingl and Watson; manuscript in preparation). Further analysis of the $Sca1^-$ and $Sca1^+$ luminal progenitors revealed that the $Sca1^-$ fraction has high levels of expression of milk proteins. These results conclude that there are two functionally distinct luminal progenitor populations in the mammary epithelium and it has been

postulated that the $CD49b^{+}Sca1^{-}$ (ERα-negative) subset represents putative alveolar progenitors while the $CD49b^{+}Sca1^{+}$ (ERα-positive) subset represents putative ductal progenitors (Stingl and Watson; manuscript in preparation).

Flow cytometric analysis revealed that that *Bcl11a*$^{lacZ/+}$ FDG positive luminal epithelial cells ($Lin^{-}CD24^{hi}FDG^{+}$) were distributed in both progenitor ($CD49b^{+}$) and differentiated ($CD49b^{-}$) regions of the FACS profile (Figure 4.19C), indicating that *Bcl11a* was expressed in both luminal progenitors and differentiated luminal epithelial cells. In contrast, there were very few *Bcl11b*$^{lacZ/+}$ FDG positive luminal epithelial cells ($Lin^{-}CD24^{hi}FDG^{+}$, 2.1% of luminal epithelial cells) and these $Lin^{-}CD24^{hi}FDG^{+}$ cells were distributed almost exclusively within the luminal progenitor ($CD49b^{+}$) regions of the FACS profile (Figure 4.19C). Expression of *Bcl11a* was detected in about 23.3 $\pm$ 6.3% of the total $Sca1^{-}$ ($ER\alpha^{-}$) luminal progenitors while expression of *Bcl11b* was detected in only 2.9 $\pm$ 0.4% of the same population. In the luminal progenitor population, expression of *Bcl11a* was detected in 14.1 $\pm$ 2.9% of total $Sca1^{+}$ ($ER\alpha^{+}$) luminal progenitors while expression of *Bcl11b* was detected in 2.3 $\pm$ 2.3% of the same population. Taken together, these results show that *Bcl11a* is expressed in luminal progenitors (both $Sca1^{+}$ and $Sca1^{-}$ progenitors) and in differentiated luminal cells. On the other hand, the few *Bcl11b* positive luminal cells are found within the $Sca1^{-}$ putative alveolar progenitor population.

Mammary glands contain different types of progenitor cells that can be detected under various conditions *in vitro* (Dontu et al., 2003; Smalley et al., 1998; Stingl et al., 2001). Mammary colony forming cells (Ma-CFCs) refer to progenitors that can form discrete mammary colonies *in vitro* (Stingl et al., 2001). In the mouse mammary epithelium, most Ma-CFCs are localized within the $CD24^{hi}CD49f^{lo}$ subpopulation of luminal cells (Stingl et al., 2006). Accordingly, I sorted out mammary epithelial cells based on *Bcl11a* and *Bcl11b* expression, and assayed for their Ma-CFC capabilities. I found that *Bcl11a*-expressing epithelial cells had a 6-fold enrichment in Ma-CFC cloning efficiencies over non-*Bcl11a*-expressing cells ($p<0.0011$), further confirming that many luminal progenitors express *Bcl11a* (Figure 4.19D). No enrichment in Ma-CFC cloning efficiencies was observed for *Bcl11b*-expressing luminal cells (Figure 4.19D).

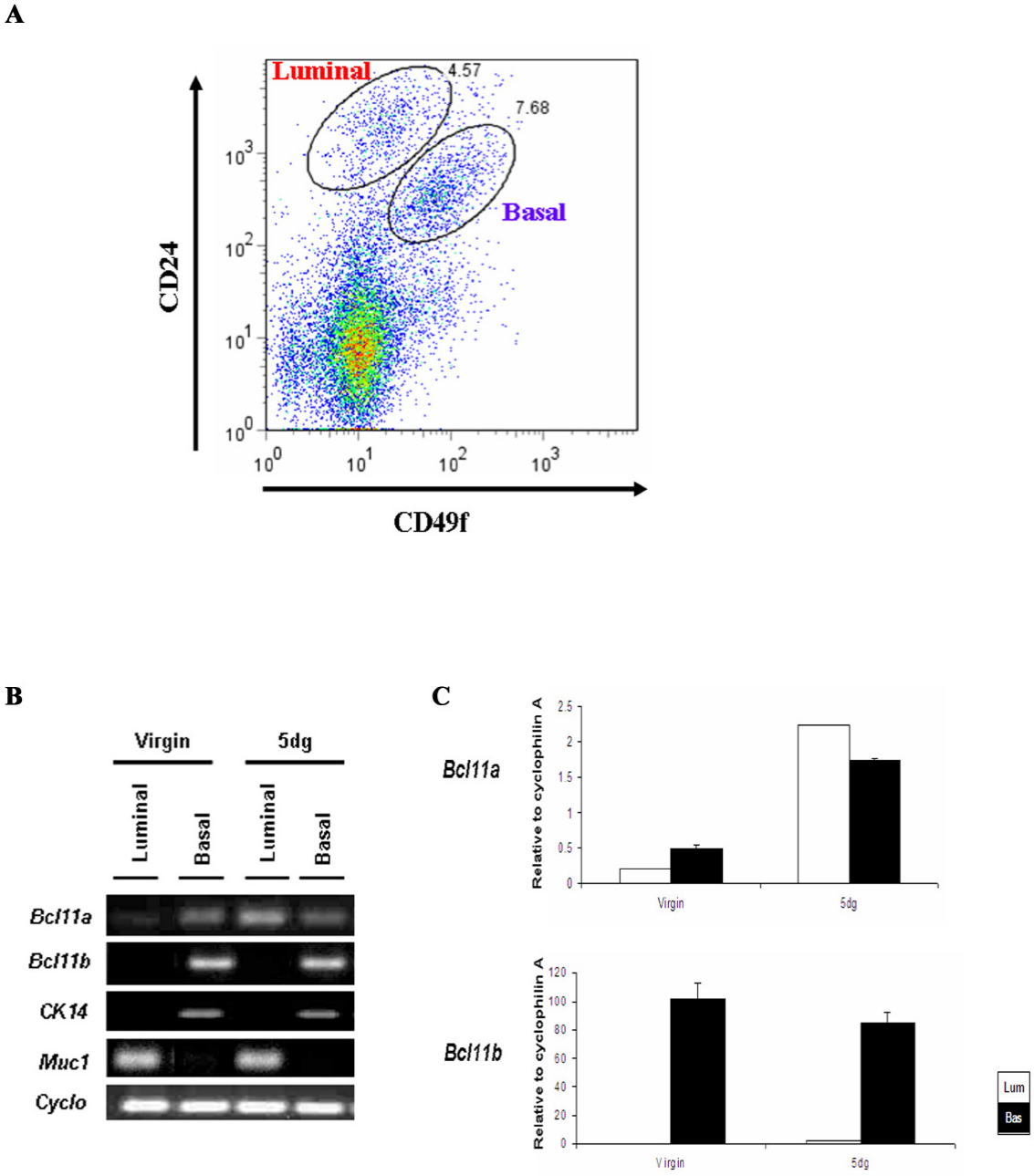
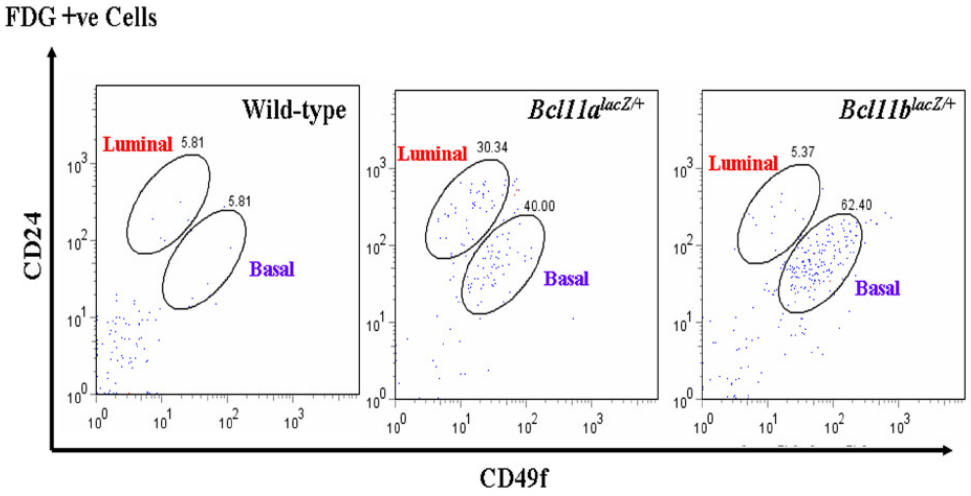
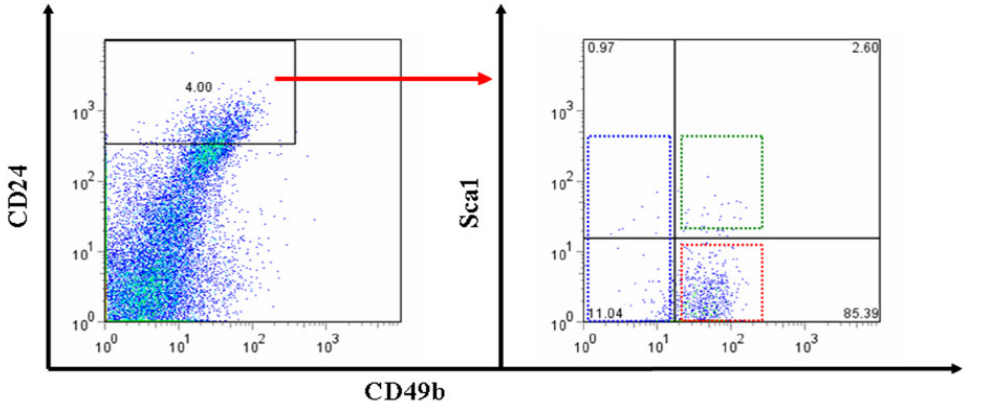


Figure 4.18. Expression of *Bcl11* genes in mammary epithelial cells. (A) Typical CD24 and CD49f mammary FACS profile of wild-type lineage-negative mammary epithelial cells. **(B)** Expression of *Bcl11* transcripts in sorted $CD24^{hi}CD49f^{+}$ (luminal) and $CD24^{+}CD49f^{hi}$ (basal) fractions from virgin and day 5 gestation (5dg) mammary gland. Purity of sorted cells determined by *Muc1* (luminal) and *CK14* (basal) markers. **(C)** qRT-PCR confirmation of *Bcl11* expression in sorted $CD24^{hi}CD49f^{+}$ (luminal) and $CD24^{+}CD49f^{hi}$ (basal) fractions from virgin and day 5 gestation mammary gland.

A



B

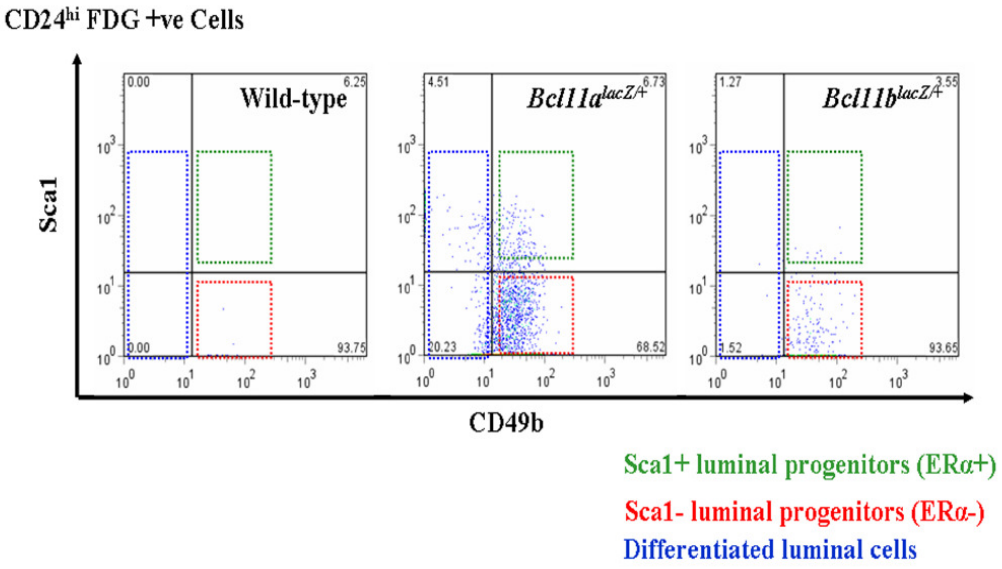


Sca1+ luminal progenitors (ERα+)

Sca1- luminal progenitors (ERα-)

Differentiated luminal cells

C



D

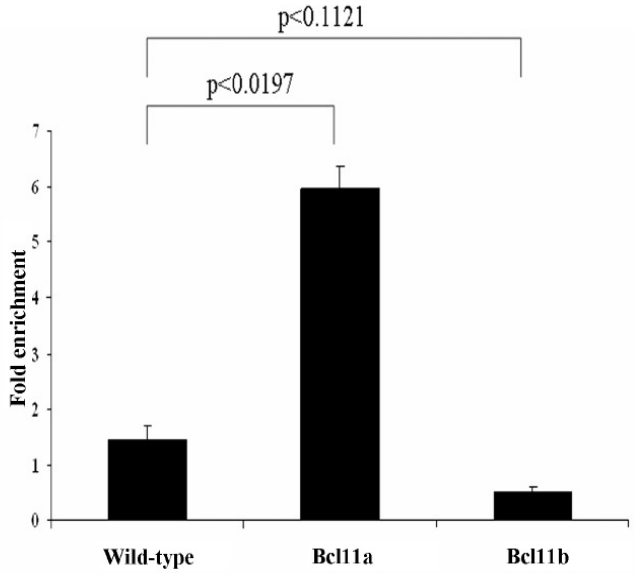


Figure 4.19. FACS-gal analysis of *Bcl11*$^{lacZ/+}$ epithelial cells. **(A)** Distribution of *Bcl11-lacZ* positive (FDG^{+}) mammary epithelial cells in a CD24/CD49f FACS profile. **(B)** Luminal mammary epithelial cells (CD24^{hi}) can be separated into luminal progenitors (CD49b-positive) and differentiated luminal (CD49b-negative) populations. **(C)** Distribution of *Bcl11-lacZ* positive (FDG^{+}) luminal (CD24^{hi}) mammary epithelial cells in a CD49b/Sca1 FACS profile. Analysis shows that *Bcl11a* is expressed in luminal progenitors and in differentiated luminal cells while *Bcl11b* is expressed only in luminal progenitors. **(D)** Graph showing relative enrichment in cloning efficiencies between *Bcl11*$^{lacZ/+}$ FDG-positive and FDG-negative epithelial cells. Error bars denote standard deviation. P values represent student T-test between *Bcl11a/Bcl11b* and wild-type samples (n=2).

4.2.7.2 Characterization of *Bcl11*$^{lacZ/+}$ hematopoietic cells

Both *Bcl11* genes have been shown to be essential for lymphocyte development (Liu et al., ; Wakabayashi et al., 2003b); however the expression of these two genes in other hematopoietic cells was not characterized in detail. Therefore, I performed flow cytometric analysis in combination with FDG staining to determine the expression patterns of *Bcl11* genes in different hematopoietic cells.

The cells of the immune system originate in the bone marrow (BM) where most of them also mature. They then migrate to guard the peripheral tissues, circulate in the blood and in specialized system of vessels called the lymphatic system. All the cellular components of blood and the immune system are derived from the hematopoietic stem cells (HSCs). The pluripotent HSCs are capable of self-renewal and can also differentiate into all the lineages of the blood cells to replenish cells that are damaged or are lost by attrition. HSCs are usually quiescent within the stem cell niche in the BM until they are mobilized from the niche to undergo proliferation and differentiation to generate cells of the immune system. HSCs are identified by their lack of lineage markers (Lin^-) and expression of Sca1 (Ly6A/E) and c-kit (CD117) markers ($Lin^-Sca1^+ckit^+$, KLS) (Spangrude et al., 1988). BM cells were harvested from *Bcl11a*$^{lacZ/+}$ and *Bcl11b*$^{lacZ/+}$ adult mice and lineage depleted as described in Chapter 2.10.1. After staining with antibodies to Sca1 and c-kit, cells were incubated with FDG before analysis with flow cytometry. As shown in Figure 4.20A, the percentage of KLS population in the BM of *Bcl11a*$^{lacZ/+}$ heterozygous mice was ~50% of that compared to the wild-type littermate control (n=3). This indicated that loss of one copy of *Bcl11a* allele resulted in a decrease in the percentage of KLS HSCs, suggesting that *Bcl11a* is a dosage-sensitive gene. About 85% of the KLS cells expressed *Bcl11a*, suggesting the *Bcl11a* was expressed in HSCs and/or MPPs (Figure 4.20B). In contrast, expression of *Bcl11b* in KLS cells was not detected. Most, if not all long term multi-lineage HSC activities reside within this minor KLS fraction of the murine BM. It has been noted that though this fraction is enriched for HSC activity when compared to whole BM, only ~1/30 of KLS cells are actually long-term reconstituting HSCs while the vast majority are multi-potent progenitors (MPPs) (Bryder et al., 2006). Taken together, *Bcl11a* is expressed in HSCs and/or MPPs, and loss of one

Bcl11a allele results in decreased numbers of KLS cells, indicating the importance of *Bcl11a* in maintenance of KLS cells.

HSCs can proliferate and divide to generate MPPs and progenitors with more limited potential such as the megakaryocyte/erythroid progenitors (MEPs), lymphoid-primed multi-potent progenitors (LMPPs), common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs). These progenitors then proliferate and differentiate to generate the entire hematopoietic lineages. The CMP is the precursor of granulocytes, monocytes/macrophages, eosinophils, neutrophils, dendritic cells and mast cells of the immune system.

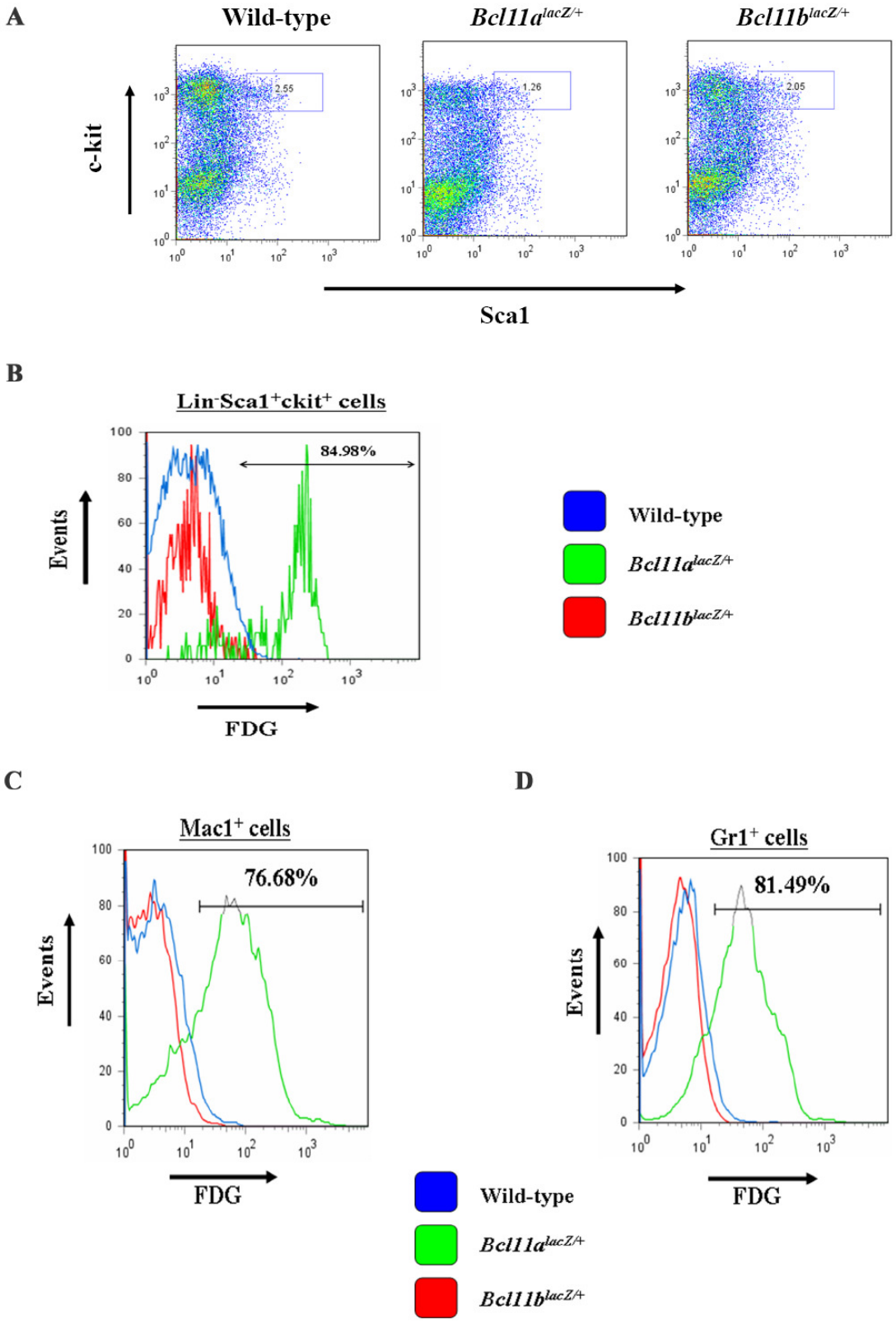
Expression of *Bcl11* genes was determined in the myeloid lineages using antibodies to Mac-1 (CD11b) and Gr-1 (Ly-6G and Ly-6C). Mac-1 is expressed on the surface of activated lymphocytes, monocytes, granulocytes and a subset of NK cells while Gr-1 is a myeloid differentiation marker and expression correlates with granulocyte differentiation and maturation. BM cells were harvested from $Bcl11a^{lacZ/+}$ and $Bcl11b^{lacZ/+}$ adult mice and stained with antibodies to Mac-1 (CD11b) and Gr-1 (Ly-6G and Ly-6C) as described in Chapter 2.10.1. After antibody staining, cells were incubated with FDG before analysis with flow cytometry. As shown in Figure 4.20C, about 77% of Gr-1 positive BM cells expressed *Bcl11a* but not *Bcl11b*. In addition, about 82% of Mac-1 positive BM cells expressed *Bcl11a* but not *Bcl11b* (Figure 4.20D). Taken together, these data suggest that *Bcl11a* but not *Bcl11b* is expressed in myeloid cells.

Next, expression of *Bcl11* genes in erythroid cells was determined using anti-TER119 antibody. TER-119 is expressed on erythroid cells from pro-erythroblast through mature erythrocyte stages. BM cells were harvested from $Bcl11a^{lacZ/+}$ and $Bcl11b^{lacZ/+}$ adult mice and stained with antibody to TER-119 as described in Chapter 2.10.1. After antibody staining, cells were incubated with FDG before analysis with flow cytometry. As shown in Figure 4.20E, about 61% of TER-119 positive BM cells expressed *Bcl11a*. Expression of *Bcl11b* in TER-119 positive BM cells was not detected. Interestingly, a quantitative trait locus (QTL) influencing F cell production maps to *BCL11A* in human thalassemia patients (Menzel et al., 2007). These results suggest that *Bcl11a* may play a role in erythroid development.

The CLPs give rise to the B lymphocytes (B cells) and the T lymphocytes (T cells) which are responsible for adaptive immunity. Activated B cells differentiate into plasma cells that secrete antibodies. Expression of *Bcl11* genes in B cells was detected using the B220 marker. The B220 antigen is expressed on all B cells from the Pro-B stage through to the mature B and activated B cell stages, but is decreased on plasma cells and a subset of memory B cells. Again, BM cells were harvested from $Bcl11a^{lacZ/+}$ and $Bcl11b^{lacZ/+}$ adult mice and stained with anti-B220 antibody as described in Chapter 2.10.1. After antibody staining, cells were incubated with FDG before analysis with flow cytometry. There was a 50% reduction in B220 positive cells in the BM of the $Bcl11a^{lacZ/+}$ heterozygous mice (Figure 4.20F). This indicated that B cells are sensitive to the levels of *Bcl11a* within the cells and that *Bcl11a* is a dosage sensitive gene. Expression of *Bcl11a* was detected in about 86% of B220 positive BM cells while expression of *Bcl11b* was not detected (Figure 4.20G).

In summary, expression of *Bcl11a* was detected in HSCs, myeloid, erythroid and B lymphocytes in the BM. In contrast, expression of *Bcl11b* was negligible in these lineages. Next, expression patterns of *Bcl11* genes in T lymphocytes were determined. Thymocytes were harvested from the thymus of $Bcl11a^{lacZ/+}$ and $Bcl11b^{lacZ/+}$ adult mice and stained with different T cell markers as described in Chapter 2.10.1. After antibody staining, cells were incubated with FDG before analysis with flow cytometry. Expression of *Bcl11* genes in T cell lineages was first determined using CD4 and CD8 cell surface markers. Expression of *Bcl11a* was only detected in about 3.45% of double-negative thymocytes (DN; $CD4^-CD8^-$) and virtually undetected in other T cell lineages (Figure 4.21A). Further analysis of the DN thymocytes using CD44 and CD25 showed that expression of *Bcl11a* was detected only at the DN1 stage (Figure 4.21B). In contrast, expression of *Bcl11b* was detected in all mature single-positive $CD4^+$ or $CD8^+$ T cells, double-positive $CD4^+CD8^+$ T cells and in 44.32% of DN T cells (Figure 4.21C). Analysis of the DN thymocytes revealed that *Bcl11b* was expressed in about 24.3% of $CD44^+CD25^-$ thymocytes at DN1 stage (Figure 4.21D). As these thymocytes mature, expression of *Bcl11b* was drastically increased and detected in more than 80% of all DN thymocytes from DN2 to DN4 stages (Figure 4.21D). Taken together, these results show that *Bcl11a* is expressed in early DN immature thymocytes at the DN1 stage. As levels of

Bcl11b increase from DN1 stage onwards, expression of *Bcl11a* is down-regulated and only *Bcl11b* is expressed in the T cell lineages (Figure 4.21E). This reciprocal expression of *Bcl11a* and *Bcl11b* during T cell development reflects their unique and potentially antagonistic function: *Bcl11a* might in general suppress T cell development while *Bcl11b* might promote T cell development. Taken together, these results suggest that *Bcl11b* may be critical in regulating T cell fate decision. Indeed, it has been suggested that *Bcl11b* is the key transcription factor that determine T cell fate (Rothenberg, 2007b).


A
Wild-type
Bcl11a lacZ/+
Bcl11b lacZ/+
c-kit
Sca1
2.55
1.26
2.05
B
Lin-Sca1+ckit+ cells
84.98%
Events
FDG
Wild-type
Bcl11a lacZ/+
Bcl11b lacZ/+
C
Mac1+ cells
76.68%
Events
FDG
D
Gr1+ cells
81.49%
Events
FDG
Wild-type
Bcl11a lacZ/+
Bcl11b lacZ/+

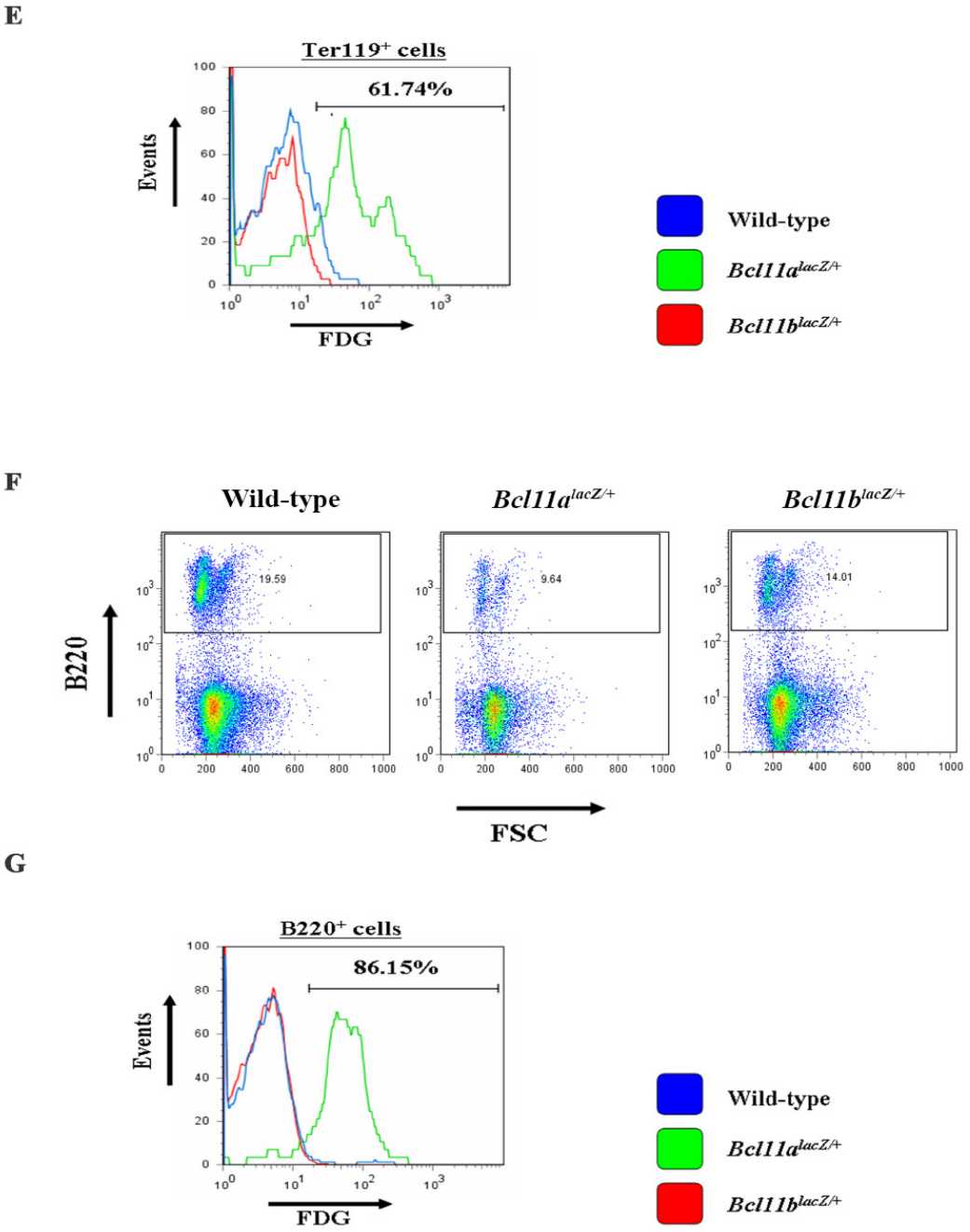
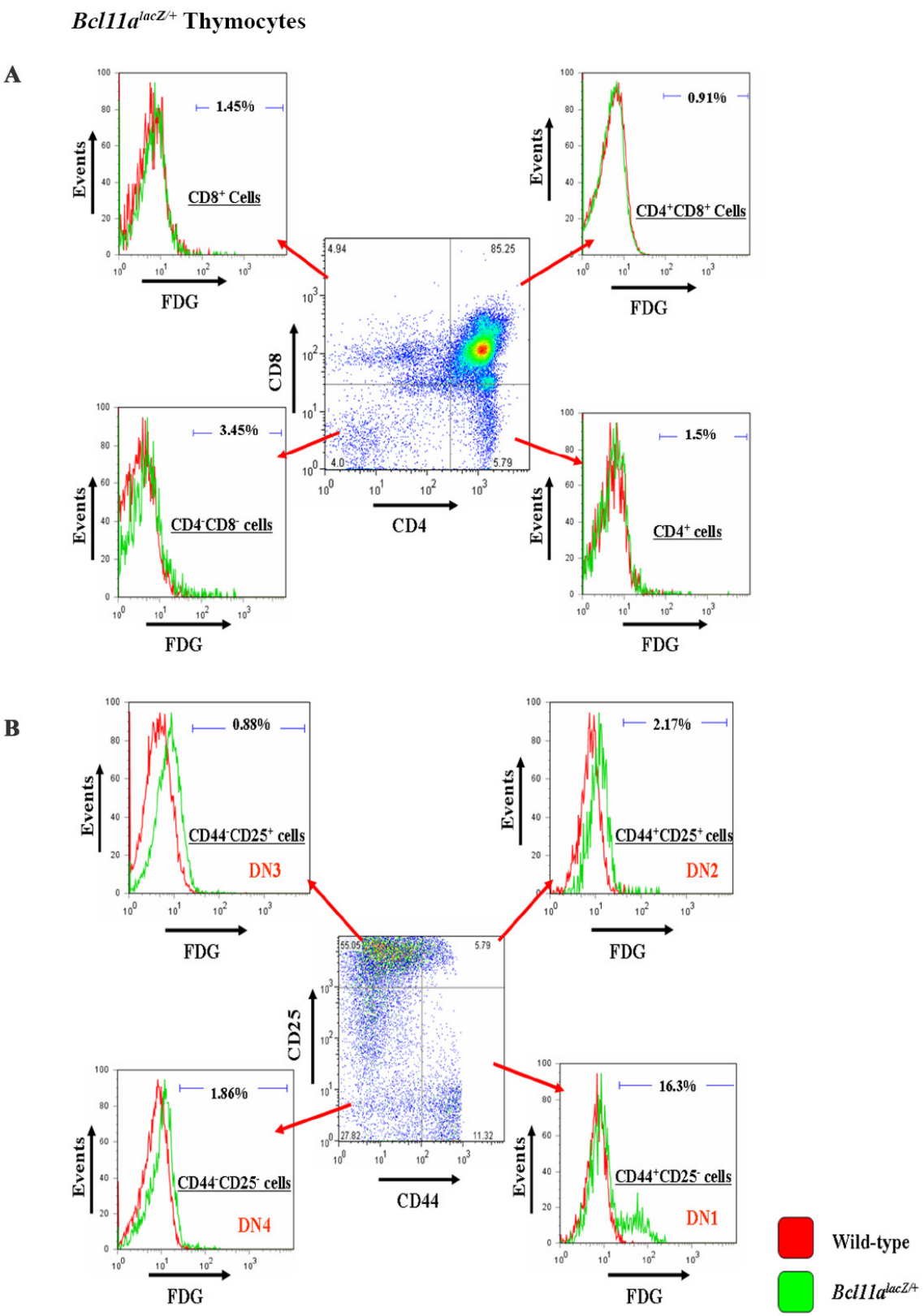
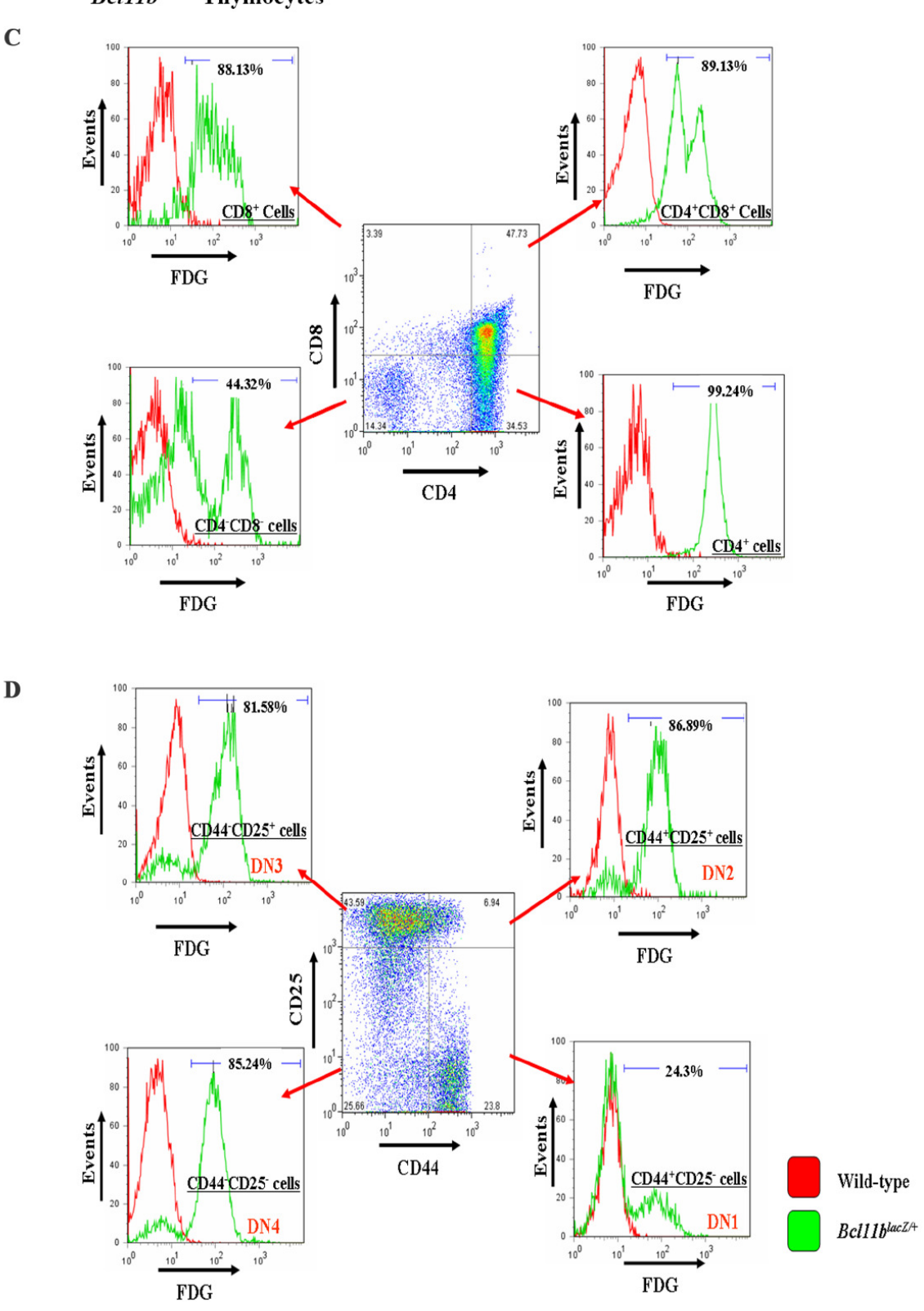


Figure 4.20. FACS-gal analysis of *Bcl11*$^{lacZ/+}$ bone marrow cells. **(A)** Flow cytometric analysis of hematopoietic cells shows that there is a decrease in the percentage of lineage-negative Sca1 and c-kit double-positive (KLS) hematopoietic stem cell population in *Bcl11a*$^{lacZ/+}$ bone marrow. **(B)** FACS-gal analysis shows that 84.98% of KLS cells express *Bcl11a*, suggesting that *Bcl11a* but not *Bcl11b* is expressed in KLS hematopoietic stem cells. FACS-gal analysis shows that *Bcl11a* but not *Bcl11b* is expressed in **(C)** Mac1-positive macrophages (76.68%), **(D)** Gr1-positive granulocytes (81.49%) and **(E)** TER119-positive red blood cells (61.74%). (%) indicates the % of Gr1/Mac1/TER119 positive cells that express *Bcl11a*. **(F)** Flow cytometric analysis of bone marrow cells also shows that there is about 50% reduction in B220-positive population (B cells) in *Bcl11a*$^{lacZ/+}$ bone marrow. **(G)** FACS-gal analysis shows that 86.15% of B220-positive cells express *Bcl11a*, indicating that *Bcl11a* but not *Bcl11b* is expressed in B cells.


Bcl11a$^{lacZ/+}$ Thymocytes
A
1.45%
CD8$^+$ Cells
0.91%
CD4$^+$CD8$^+$ Cells
4.94
85.25
4.0
5.79
CD8
CD4
3.45%
CD4$^-$CD8$^-$ cells
1.5%
CD4$^+$ cells
Events
FDG
B
0.88%
CD44$^-$CD25$^+$ cells
DN3
2.17%
CD44$^+$CD25$^+$ cells
DN2
55.05
5.79
27.82
11.32
CD25
CD44
1.86%
CD44$^-$CD25$^-$ cells
DN4
16.3%
CD44$^+$CD25$^-$ cells
DN1
Wild-type
Bcl11a$^{lacZ/+}$


Bcl11b^lacZ/+ Thymocytes
C
88.13%
CD8+ Cells
89.13%
CD4+CD8+ Cells
44.32%
CD4-CD8- cells
99.24%
CD4+ cells
3.39
47.73
14.34
34.53
CD8
CD4
Events
FDG
D
81.58%
CD44-CD25+ cells
DN3
86.89%
CD44+CD25+ cells
DN2
85.24%
CD44-CD25- cells
DN4
24.3%
CD44+CD25- cells
DN1
43.59
6.94
25.66
23.8
CD25
CD44
Wild-type
Bcl11b^lacZ/+

E

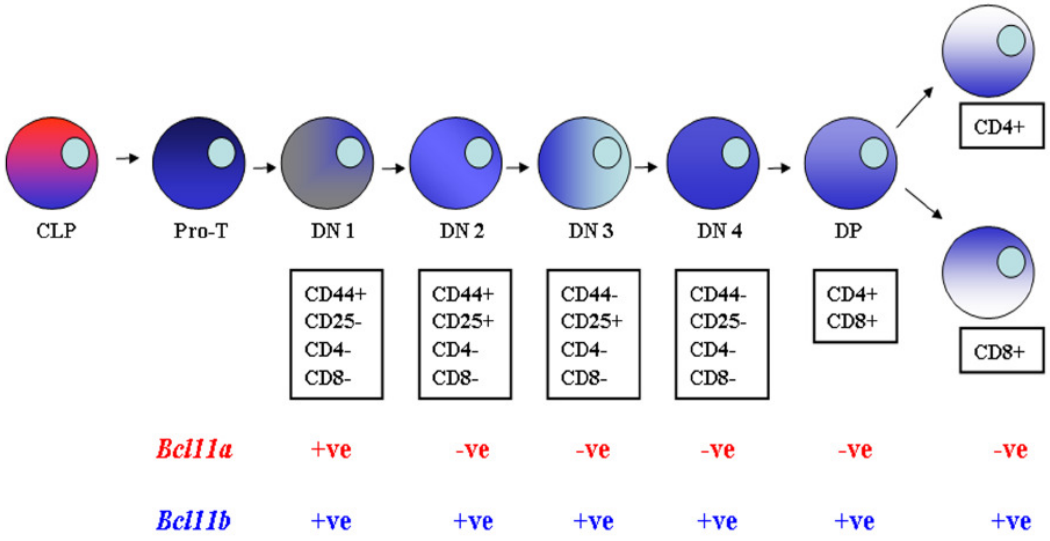


Figure 4.21. FACS-gal analysis of *Bcl11*$^{lacZ/+}$ thymocytes. **(A)** FACS-gal analysis of *Bcl11a*$^{lacZ/+}$ thymocytes shows that *Bcl11a* is expressed only in a small percentage of CD4 and CD8 double negative thymocytes (3.45%). **(B)** Further analysis of these CD4 and CD8 double negative thymocytes shows that expression of *Bcl11a* is only detected in the CD44^{+}CD25^{-} (DN1) thymocytes (16.3% of CD44^{+}CD25^{-} thymocytes express *Bcl11a*). **(C)** In contrast, expression of *Bcl11b* is detected in all thymocytes. About 45% of CD4 and CD8 double negative thymocytes express *Bcl11b*. **(D)** Analysis of the CD4 and CD8 double negative thymocytes shows that *Bcl11b* is expressed from DN1 stage (24.3% of CD44^{+}CD25^{-} thymocytes express *Bcl11b*) and the expression is maintained throughout T cell development stages. **(E)** Summary of expression of *Bcl11* genes in T cell developmental stages. +ve: expression; -ve: no expression.

4.3 Discussion

In this Chapter, I showed that the new gene targeting system and strategy that our lab had developed as described in Chapter 1 worked efficiently in generating *lacZ*-tagged reporter mice. X-gal staining patterns of *Bcl11*$^{lacZ/+}$ embryos faithfully recapitulated the endogenous *Bcl11* expression. Using whole mount X-gal staining, the spatial and temporal expression patterns of both *Bcl11a* and *Bcl11b* were determined. In addition, the expression of both genes in hematopoietic and mammary epithelial cells was also characterized at a cellular level using FDG staining and flow cytometry. *Bcl11* genes exhibited unique and dynamic expression patterns throughout the developmental time points studied.

4.3.1 Summary of embryonic expression patterns

Analysis of X-gal staining patterns of *Bcl11*$^{lacZ/+}$ embryos showed that both genes were expressed from 10.5 dpc. Expression of neither gene was detected at earlier stages. Some anatomical regions showed overlapping expression of both genes which was maintained throughout embryonic development; while other regions displayed reciprocal expression of the two genes. The primary regions of overlapping expression were in the brain and the craniofacial mesenchyme. Expression of *Bcl11* genes were first detected in the forebrain and in the derivatives of the pharyngeal arches from 10.5 dpc. At 12.5 dpc, expression of both genes became more restricted to regions of the developing brain and to the facial mesenchyme. This expression pattern was maintained at 14.5 dpc. Overlapping expression patterns within these regions suggest that functions of *Bcl11a* and *Bcl11b* might be needed in the development of brain and facial mesenchyme.

Differential expression patterns of *Bcl11* genes were detected outside of the brain and facial mesenchyme. Limb development occurs early in embryonic stages where the reciprocal interaction between the mesenchymal and the ectodermal cells is critical to create the limb buds (Johnson and Tabin, 1997). Expression of *Bcl11a* in the limb buds was detected from 10.5 dpc and its expression was maintained throughout later stages of limb development. In addition, expression of *Bcl11a* was also detected in the developing cartilage and bone. Nevertheless, no apparent abnormalities were observed in *Bcl11a* homozygous mutant embryos at 18.5 dpc. These observations suggest that *Bcl11a* might

not have a major role in the specification of limb fate or a compensatory mechanism exists for the loss of *Bcl11a*. Therefore, further studies are required to address whether *Bcl11a* plays a minor role in limb and/or skeletal development.

Expression of *Bcl11* genes was differentially regulated in the lungs during embryonic development (14.5–18.5 dpc). Transient expression of *Bcl11b* was detected in the lungs. It was expressed at 14.5 dpc and not at 18.5 dpc. In contrast, expression of *Bcl11a* in the lungs started from 18.5 dpc. Mouse lung bud formation occurs early at 9.5 dpc from the laryngotracheal groove and involves mesenchymal-epithelial cell interactions (Costa et al., 2001). During mouse lung development, the pseudoglandular stage (9.5-16.6 dpc) is characterized by formation of the bronchial and respiratory bronchiole tree, which is lined with undifferentiated epithelial cells juxtaposed to the splanchnic mesoderm. There is extensive branching of the distal epithelium and mesenchyme during the canalicular stage (16-17 dpc), resulting in formation of terminal sacs lined with epithelial cells integrating with the mesoderm-derived vasculature. The terminal sac stage [17.5 dpc to postnatal day 5] of lung development is characterized by a coordinated increase in terminal sac formation and vasculogenesis in conjunction with the differentiation of alveolar epithelial type I and II cells. The dynamic expression patterns of *Bcl11a* and *Bcl11b* during embryonic lung development suggest that each of these genes may be involved in different phases of lung development. *Bcl11b* may be involved in the pseudoglandular stage (initial generation of lung structures), while *Bcl11a* may be involved in the terminal sac stage (terminal differentiation of lung alveolar). Future studies could be carried out to determine the precise roles of these transcription factors in lung development.

In addition to the lung, differential expression patterns were also observed in the fetal liver, thymus and mammary gland. The murine fetal liver is organ of intense, but transient site of hematopoietic activity during mid-gestation. It is the initial site of hematopoiesis where hematopoietic cells are generated. T cell progenitors then exit the fetal liver and home to the thymus where the immature thymocytes undergo maturation. Expression of *Bcl11a* in the fetal liver and *Bcl11b* in the thymus was detected in early embryonic development. *Bcl11b* was detected in the developing thymus at 14.5 dpc, a time at which most thymocytes are T cell precursors, and its expression was maintained

at 18.5 dpc. These results together with the knockout phenotypes (Liu et al., ; Wakabayashi et al., 2003b) demonstrate that *Bcl11* genes are essential for the hematopoietic lineages.

4.3.2 Differential *Bcl11* expression patterns in hematopoietic lineages

Expression of *Bcl11a* and *Bcl11b* was studied in bone marrow hematopoietic cells using FDG staining. In addition to its known expression and function in B cells, expression of *Bcl11a* was also detected in KLS HSCs and/or multi-potent progenitors, myeloid cells and red blood cells; suggesting that *Bcl11a* may also play role(s) in these hematopoietic lineages. Recently, a quantitative trait locus (QTL) influencing F cell production mapping to *BCL11A* in human thalassemia patients has been identified (Menzel et al., 2007). In addition, *Bcl11a* is a dosage-sensitive gene where loss of one allele resulted in reduction in numbers of B cells and KLS HSCs (Figure 4.18A and 4.18F). These, together with some of my work which showed that deletion of *Bcl11a* in adult hematopoietic cells resulted in the initial expansion, followed by depletion of the KLS population (personal observations), suggest that *Bcl11a* plays an important role in maintenance of the hematopoietic lineages.

In the adult thymus, expression of *Bcl11a* was only detected in a small percentage of thymocytes. Using different molecular markers, I found that *Bcl11a* expression was only detected in 16.3% of DN1 immature thymocytes and not in mature T cells. In contrast, expression of *Bcl11b* was detected only in T cell lineages and its expression was detected at all stages of T cell development. *Bcl11b* expression was detected from DN1 immature thymocyte stage and was drastically up-regulated from the DN2 stage, a critical stage in T cell linage commitment. This expression pattern thus supports the proposal that *Bcl11b* is the main T cell lineage determinant in thymocyte development (Rothenberg, 2007a). The dynamic reciprocal expression patterns of *Bcl11* genes in immature thymocytes suggest that the interplay between the levels of these genes in hematopoietic progenitors may be critical for lymphocyte development.

4.3.3 Dynamic differential expression patterns in mammary lineages

Recent studies have shown that transcription factors that normally play critical roles in T helper cell lineage determination are also essential for mammary lineage determination (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006). This complements the existing paradigm that mammary development and cell fate decisions are determined by steroids and prolactin (Hennighausen and Robinson, 2005). Because *Bcl11* genes are expressed in a dynamic and contrasting pattern in hematopoietic cells and they are a pair of transcription factors that have been shown to be essential for B and T cell determination from a common lymphoid progenitor, I hypothesized that *Bcl11* genes also play important roles in mammary cell fate determination. Using the *Bcl11-lacZ* reporter mice, I found that *Bcl11* genes were among the earliest known genes to be expressed specifically in the mammary gland. Expression of *Bcl11b* was detected from 10.5 dpc in the milk line and from 12.5 dpc, its expression was detected within the mammary buds. In contrast, expression of *Bcl11a* within mammary buds was detected only from 13.5 dpc.

These differential expression patterns were maintained during postnatal mammary development where *Bcl11a* was expressed in both cap and body cells of the TEBs and *Bcl11b* was detected only in the neck region of TEBs in the cap cell layer. Consistent with these observations, *Bcl11a* was found to be expressed in both the luminal and basal layers of the mature virgin duct while expression of *Bcl11b* was detected predominantly in the basal layer. This was confirmed by both semi-quantitative and quantitative PCR using FACS sorted luminal and basal fractions. Interestingly, up-regulation of both genes was detected at early gestation. *Bcl11a* was detected in luminal layers of ducts and in differentiating alveoli. In contrast, expression of *Bcl11b* remained primarily in the basal layer of ducts and was not detected in the differentiating alveoli. This reciprocal expression pattern was maintained throughout gestation. During lactation, only expression of *Bcl11a* was detected in the differentiated luminal secretory cells. Further analysis using cell surface markers and flow cytometry showed that *Bcl11a* was expressed in luminal progenitors and their differentiated derivatives while only a small percentage of luminal progenitors expressed *Bcl11b*. In addition, a 6-fold enrichment in cloning efficiencies of *Bcl11a*-expressing epithelial cells which further confirmed that

many luminal progenitors expressed *Bcl11a*. These results suggest that *Bcl11a* may be important for differentiation of the luminal lineages while *Bcl11b* may be important for maintaining basal cell identity and/or suppressing luminal cell fate. During involution, there was a dramatic up-regulation of *Bcl11b* levels 24 hours after initiation of involution, followed by a sharp decline before peaking again at 96 hours involution time point. On the other hand, levels of *Bcl11a* increased gradually during involution and peaked at 72 hours involution. Thus *Bcl11* genes may play different roles during the involution phase. While *Bcl11b* may be important for the first phase of involution, expression pattern of *Bcl11a* suggests that it may play an important role during the second phase of involution. Taken together, the unique and dynamic expression patterns of *Bcl11* genes during mammary gland development suggest that these two genes play important roles in the mammary gland. The expression data also laid a solid ground for the characterization of the mammary phenotypes as described in Chapters 5 and 6.

In summary, I have detailed the spatial expression patterns of *Bcl11* genes during embryonic and adult development using the $Bcl11^{lacZ/+}$ mice that I have generated. X-gal staining patterns reveal that the brain and craniofacial mesenchyme are the two main regions of overlapping expression. Differential expression of the two genes was detected in the developing limbs, cartilage/bone, lung, hematopoietic lineages and mammary gland. Dynamic expression patterns of *Bcl11* genes during mammary gland development were observed and their expression in different epithelial populations suggests that they have roles in mammary lineage commitment. These results, together with a recent study that identified mutations in *BCL11A* in human breast cancer (Wood et al., 2007) prompted me to investigate the roles of *Bcl11* genes in mammary gland development. Based on these observations, I hypothesized that *Bcl11a* may be essential for luminal cell differentiation while *Bcl11b* may be important for maintaining basal identity and/or suppressing luminal cell fate. In Chapter 5, I will describe characterization of the functions of *Bcl11* genes in mammary development in the embryo and postnatal virgin glands.

CHAPTER 5:

BCL11 GENES ARE ESSENTIAL FOR NORMAL MAMMARY DEVELOPMENT

5.1 Introduction

5.1.1 Genetic control of lymphocyte and mammary development

Gene regulatory networks are essential to the developmental process, orchestrating the process through a series of co-ordinated hierarchical stages, enabling the specification, commitment and differentiation of different cell fates from multi-potent stem or progenitor cells. Transcription factors play intricate roles in the gene regulatory networks which are made up of complementary sets of transcription activators and repressors that play antagonistic roles by activating the transcription of effector genes of a given cell fate while repressing that of alternate cell fates (Singh et al., 2005). Therefore homeostasis of the levels and modifications of transcription factors at various checkpoints during the developmental process is critical for the specification and maintenance of a certain cell fate. The initial specification of cell fate arises in part through juxtacrine and paracrine signalling pathways, such as Wnt and Notch that activate the initial node in the regulatory network (Singh et al., 2005). These transcription factors then auto-regulate in positive- or negative-feedback loops and activate or repress other members in a hierarchical fashion in order to establish the differentiated state. Each cell fate is therefore specified by a unique combination of transcription factors; the level and activity of each contribute to the specification and differentiation of the cell.

With the well-defined molecular markers for various stem/progenitor and differentiated cells, together with a well-established transplantation protocol in the mouse, the hematopoietic system is currently the most comprehensively understood stem cell developmental system. A combination of reverse genetics by loss-of-function analyses with knockout mouse models and forward genetic screens with retroviral integrations that resulted in development of leukaemia have uncovered transcription factors that are critical for the control of different hematopoietic lineages (Cantor and Orkin, 2001; Orkin and Zon, 2008). The Gata family of transcription factors are crucial

players in the gene regulatory networks that control the specification of cell fates (Ho and Pai, 2007; Kaufman et al., 2003; Kouros-Mehr et al., 2008). In particular, Gata-1 and Gata-3 are involved in the specification of hematopoietic cell fates. While Gata-1 is involved in erythropoeisis, Gata-3 is required for T-cell specification and T_H2 differentiation. In these hematopoietic lineages, Gata transcription factors on one hand promote the fate of a specific cell lineage, and on the other hand repress alternate lineages. For example, in T-helper cells, Gata-3 and T-bet exhibit transcriptional cross-antagonism which help the bi-potent progenitors to commit to one fate and also reinforce the specification and maintenance of T_H2 and T_H1 cell fates, respectively (Ho and Pai, 2007; Zheng and Flavell, 1997).

Research on mammary gland development in the mouse has been focused on steroid and peptide hormones and the contribution of their signalling pathways on mammary gland differentiation (Chapter 1.4.2-3) (Hennighausen and Robinson, 2005). Recent studies have identified several transcription factors that normally functions in lineage commitment in lymphocyte development, also play essential roles in mammary lineage commitment (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006). Using mammary-specific transgenic Cre lines, *Gata-3* was conditionally deleted at various stages of mammary development and shown to be essential for the differentiation and maintenance of luminal cell fate (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). Prolactin signalling is essential for the proliferation and functional differentiation of lobulo-alveolar structures during pregnancy (Topper and Freeman, 1980). The effects of prolactin signalling during pregnancy are predominantly mediated via Stat5 as discussed in Chapter 1.4.3.3-4. Recently, Stat6 and its upstream cytokines, IL4 and IL13, have also been shown to be required for luminal epithelial cell development (Khaled et al., 2007). In the *Stat6*$^{-/-}$ and *Il4*$^{-/-}$/*Il13*$^{-/-}$ double mutant mice, a 70% reduction in the number of alveoli was observed in the day 5 gestation gland, whereas deletion of *Socs5*, a negative regulator of Stat6, resulted in precocious alveolar development (Khaled et al., 2007). Taken together, these results suggest that genetic control of cell fate determination in mammary and lymphocyte lineages is highly parallel.

Bcl11a and *Bcl11b* are essential in lymphoid lineage specification as discussed in Chapter 1.3.4. These observations, together with the fact that two *BCL11A* mutations

were identified in human breast cancer (Wood et al., 2007), prompted me to investigate the functions of *Bcl11a* and *Bcl11b* in the mammary gland.

5.1.2 Clues from expression patterns

As discussed in Chapter 4, *Bcl11a* and *Bcl11b* exhibited unique and differential expression patterns in the mammary gland. Expression of both genes in the mammary lineages was detected in early embryonic development. In the virgin gland, *Bcl11a* was expressed in TEBs and in both luminal and basal cells while *Bcl11b* was expressed predominantly in the basal layer. Further analysis showed that *Bcl11a* was expressed in differentiated luminal and luminal progenitor cells. *Bcl11b* on the other hand, was expressed in only a small number of $Sca1^-$ ($ER\alpha^-$) luminal progenitors. Up-regulation of both genes was detected early in gestation; while expression of *Bcl11a* was maintained throughout gestation and lactation, levels of *Bcl11b* declined steadily from gestation to lactation. In the gestation gland, expression of *Bcl11a* was detected in both ductal and differentiating lobulo-alveolar structures while expression of *Bcl11b* was restricted predominantly to the basal cells. During lactation, only *Bcl11a* was detected in differentiated luminal secretory lobulo-alveoli. Taken together, these expression data implicate *Bcl11* genes may play important but different roles in the mammary epithelium – while *Bcl11a* may be important for luminal cell differentiation and maintenance, *Bcl11b* may have a critical role in the basal cells.

In this Chapter, I will first describe the mammary phenotypes in the *Bcl11* homozygous mutant embryos. I will then discuss the mammary defects in the virgin gland of the *Bcl11* conditional knockout mice after deleting *Bcl11* genes.

5.2 Results

5.2.1 Loss of *Bcl11a* leads to abnormal mammary bud formation

To study the physiological role of *Bcl11a* and *Bcl11b* in embryonic development, heterozygous *Bcl11*$^{lacZ/+}$ breeding pairs were set up and female mice were subjected to timed pregnancies. Embryos were collected, staged and genotyped to identify the *Bcl11* homozygous mutant embryos for analysis.

Murine mammary development is evident at 10.5 day post-coitum (dpc) with the appearance of milk lines in the embryos (Veltmaat et al., 2004). At 11.5 dpc, the milk lines have developed into five pairs of mammary placodes that occupy specific positions along the ventral side of the embryo (Watson and Khaled, 2008). The placodes are formed asynchronously in a specific order. Several genes such as *Tbx3*, *Lef1*, *Fgf10* and *Fgfr2b* have been implicated in the formation of different placodes (Eblaghie et al., 2004; Mailleux et al., 2002; van Genderen et al., 1994).

I first analyzed the *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant embryos. Expression of *Bcl11a* was detected in the mammary buds from 13.5 dpc. Anatomical examinations of *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant embryos failed to identify the presence of mammary buds at 14.5 dpc. To determine the mammary phenotypes, *in situ* hybridization with *Fgfr2TK* antisense probe was performed. *Fgfr2TK* is expressed in the mammary primordial from 11.5 dpc (Mailleux et al., 2002) and *in situ* hybridization with this antisense probe showed that all five pairs of mammary buds were formed in the *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant 13.5-14 dpc embryos (Figure 5.1A). However, all the mammary buds in the *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant embryos had a flattened morphology whereas those in the wild-type embryos had a discrete lens-shaped structure (Figure 5.1A). *Notch1* is expressed specifically in the mammary buds during embryonic development (Personal communication with Christine Watson). Examination of wild-type and *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant 13.5-14 dpc embryos using *Notch1* antisense probe showed that expression of *Notch1* was detected only in the fourth and fifth pair of mammary buds. Consistent with the results obtained using *Fgfr2TK* probe, *in situ* hybridization with *Notch1* antisense probe confirmed that the mammary buds in *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant embryos had a less developed morphology as

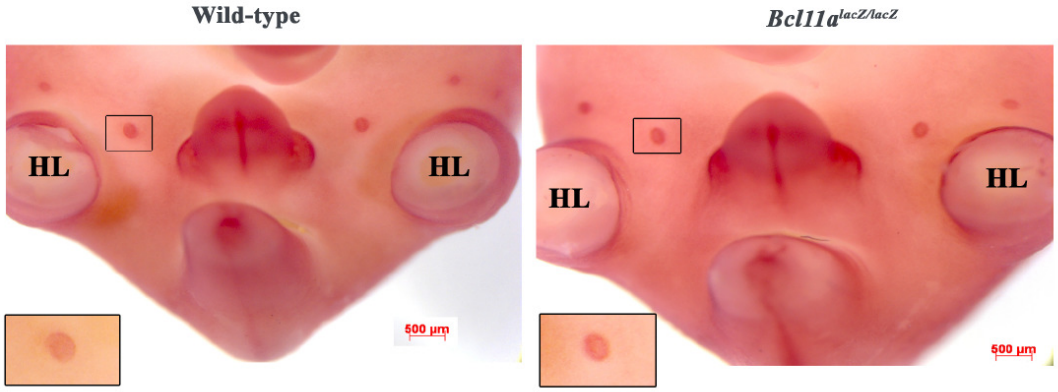
compared to the wild-type embryos which formed large, round knob-shaped structures (Figure 5.1B).

Lef1 exhibits a dynamic expression pattern during embryonic mammary development (Foley et al., 2001). Between 11-12 dpc, *Lef1* is expressed in basal cells of developing epidermis and in epithelial cells of the mammary buds and from 13.5 dpc, expression of *Lef1* is induced within mesenchymal cells of the developing mammary buds as it is fading away in the mammary epithelium. By 14.5-15.5 dpc, in the epidermis destined to become nipple sheath, *Lef1* expression becomes undetectable. At 14.5 dpc, expression of *Lef1* was detected only in the epidermis of the fifth pair of mammary buds in the wild-type embryos. In contrast, *Lef1* expression was detected in all the epidermis of the mammary buds in $Bcl11a^{lacZ/lacZ}$ homozygous mutant embryos (Figure 5.1C).

In male embryos at 15.5 dpc, the epithelial buds became separated from the epidermis under the influence of testosterone resulting in active induction of apoptosis by the mammary mesenchyme (Wysolmerski et al., 1998). This resulted in the regression of mammary buds in male embryos and *in situ* hybridization with *Fgfr2TK* antisense probe revealed no mammary epithelial remnants in wild-type male embryos at 16.5 dpc (n=1) (Figure 5.1D1-3). In contrast, presence of rudimentary mammary buds in $Bcl11a^{lacZ/lacZ}$ homozygous mutant male embryos at 16.5 dpc was detected using *in situ* hybridization with antisense *Fgfr2TK* probe (n=1) (Figure 5.1D4-6), suggesting that loss of *Bcl11a* resulted in failure of regression of mammary buds. *Lef1* has been shown to be an important mesenchymal survival factor (Boras-Granic et al., 2006) as $Lef1^{-/-}$ embryos showed arrested development and subsequent regression in mammary placodes. Therefore, the continual expression of *Lef1* in the mammary buds (possibly in both epidermis and mesenchyme) of $Bcl11a^{lacZ/lacZ}$ homozygous mutant 14.5 dpc embryos might explain the presence of rudimentary mammary buds in the male embryo at 16.5 dpc. Taken together, loss of *Bcl11a* results in defective embryonic mammary formation and failure in regression of mammary epithelial in the male embryo.

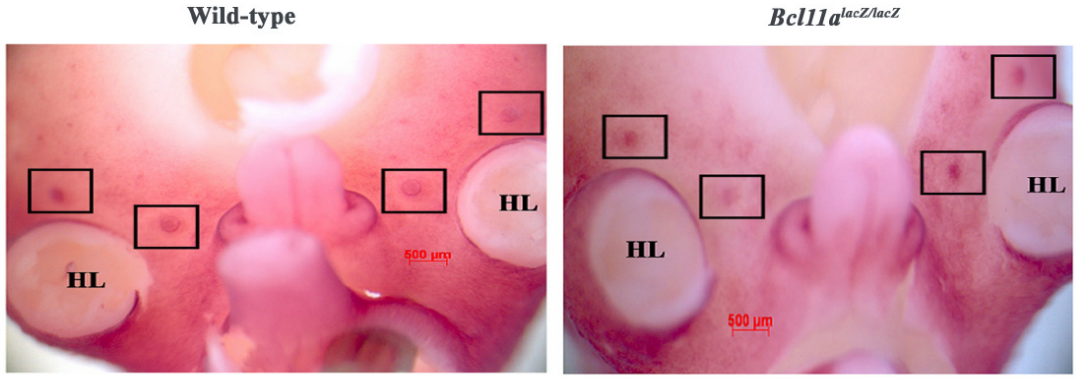
A

13.5-14 dpc (*Fgfr2TK*)



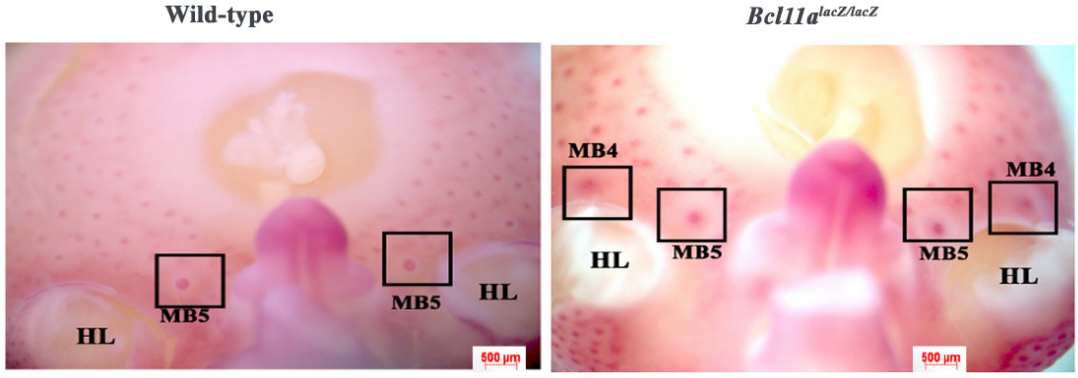
B

13.5-14 dpc (*Notch1*)



C

14.5 dpc (*Lef1*)



D

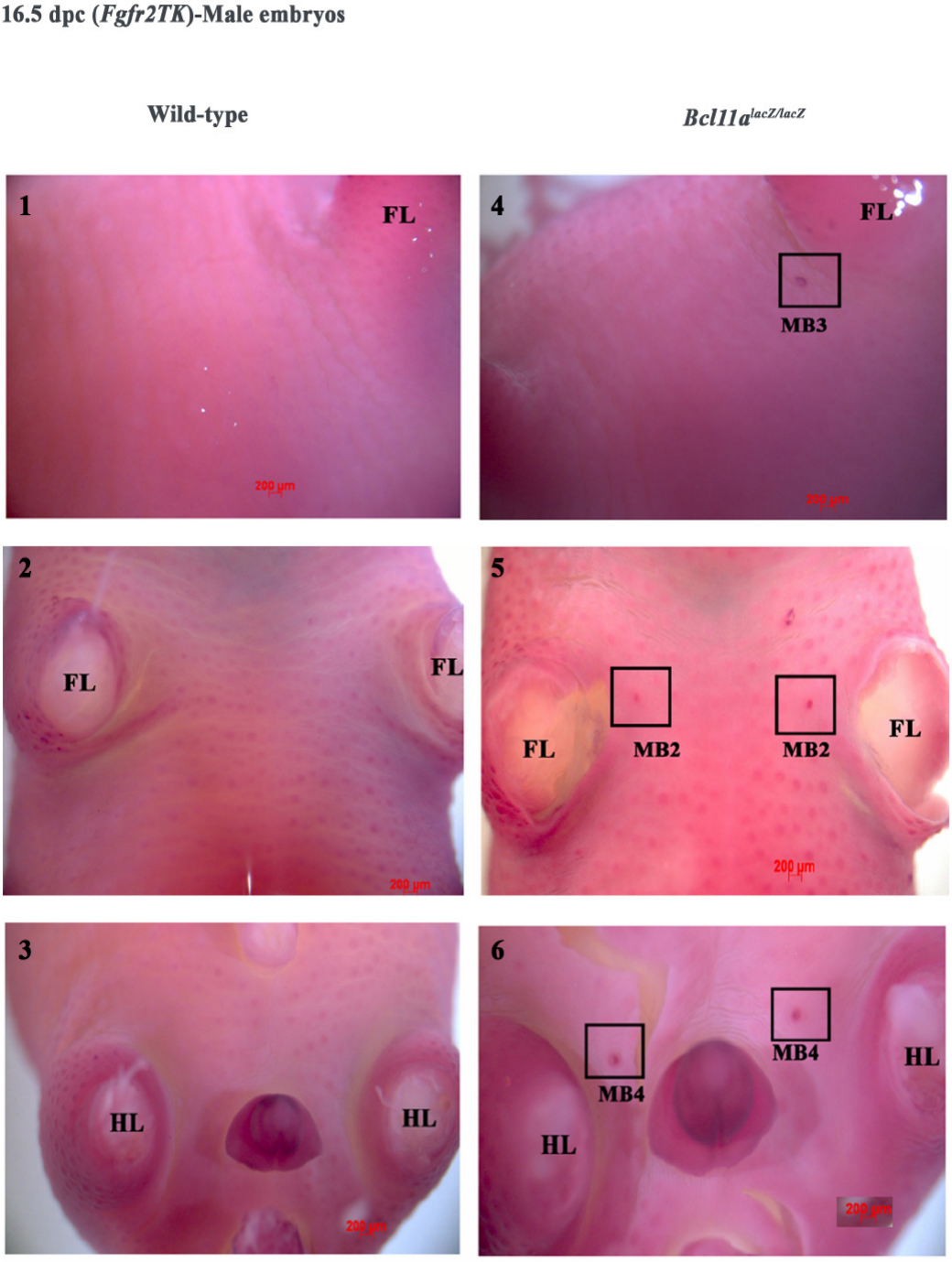


Figure 5.1. *In situ* hybridization of wild-type and *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant embryos. Whole mount images of *in situ* hybridization of wild-type and *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant 13.5-14 dpc embryos with **(A)** *Fgfr2TK*, **(B)** *Notch1* and **(C)** *Lef1* antisense probes. Inserts are higher magnification of mammary buds. Whole mount images of *in situ* hybridization of wild-type **(D1-3)** and *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant **(D4-5)** 16.5 dpc male embryo with *Fgfr2TK* antisense probe. Boxed regions indicate presence of buds in *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant male embryo. FL – Forelimb; HL – Hindlimb; MB – Mammary Bud.

5.2.2 *Bcl11b* is essential for formation of third pair of mammary buds

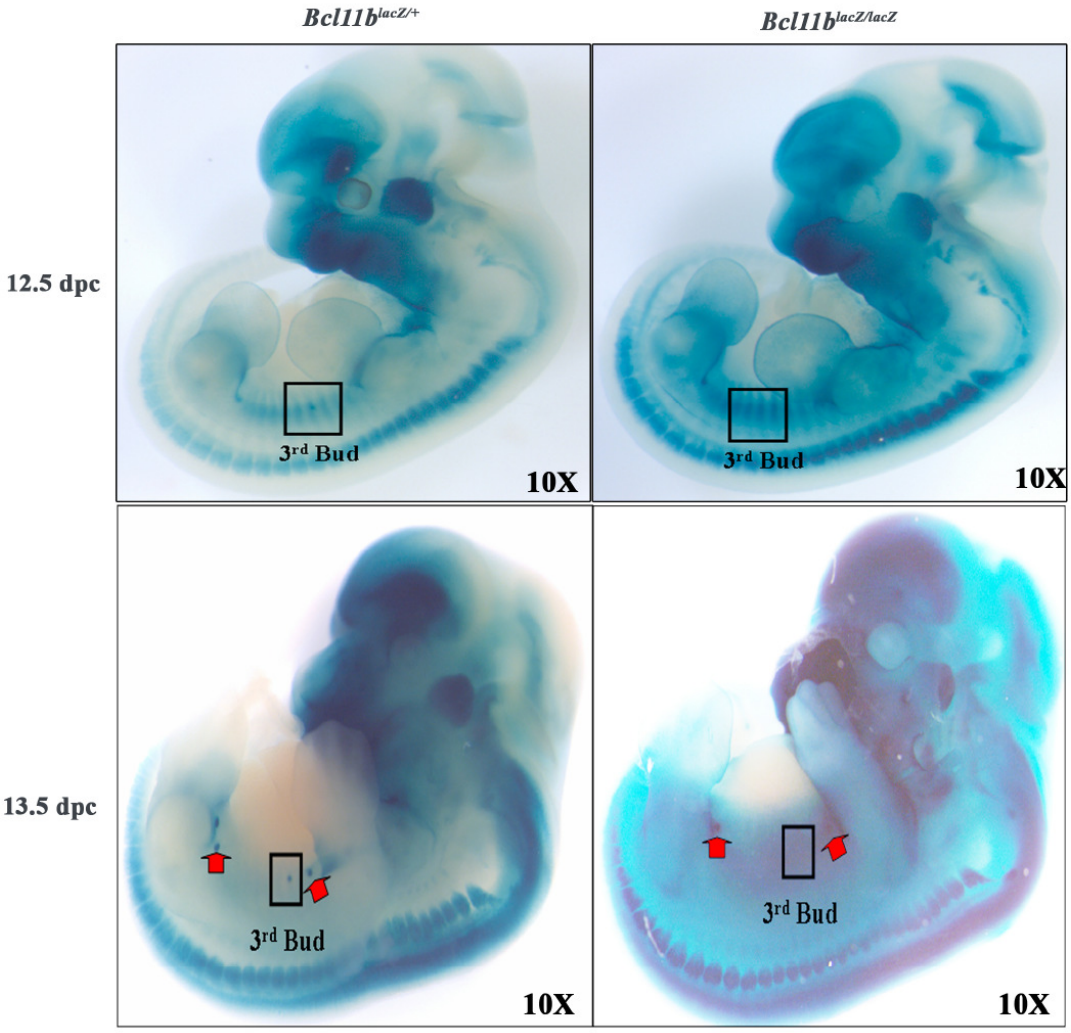
Bcl11b is expressed in all mammary placodes from 12.5 dpc. In the *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 12.5-13.5 dpc embryos, expression of *Bcl11b* was not detected in the third pair of mammary buds (Figure 5.2A). This suggests that the third pair of mammary buds may be missing in the *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant embryos. To confirm the effects of loss of *Bcl11b* on mammary buds, whole mount *in situ* hybridization using *Wnt10b* was performed. *Wnt10b* is one of the earliest markers of the mammary epithelium and is expressed from 10.5 dpc, beginning as a diffuse line (milk line) that eventually breaks up into individual spots (mammary placodes) by 12.5 dpc (Veltmaat et al., 2004). As shown in Figure 5.2B, expression of *Wnt10b* was abolished in the third pair of mammary placodes and also greatly reduced in the second and fourth pairs of mammary placodes in the *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 12.5 dpc embryos. This established that *Bcl11b* is essential for the formation of the third pair of mammary buds. This mammary defect was further confirmed by *in situ* hybridization with the antisense *Lef1* and *Fgfr2TK* probes. Similar to *Wnt10b*, expression of *Lef1* was absent in the third pair and reduced in the other pairs of mammary buds of the *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 12.5 dpc embryos (Figure 5.2C). In later stage embryos (14.5 dpc), expression of *Fgfr2TK* was not detected in the third pair and was greatly reduced in other pairs of mammary buds of the *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant embryos (Figure 5.2D). These results confirmed that in the absence of *Bcl11b*, formation of the third pair of mammary buds was completely abolished while formation of other mammary buds was also affected.

To further dissect the mammary development defects in *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant embryos, *in situ* hybridization on younger embryos (10.5 dpc) using other markers was performed. *Fgf10* expression is first detected in the presumptive mammary region at 10.5 dpc and its expression extends as a line from the forelimb bud to the level of somite 18 (where the third pair of mammary buds will subsequently form) in the dermamyotome (Mailleux et al., 2002). *Tbx3* expression is also observed as a thin line marking what appears to be the mammary line from 10.25 dpc (Eblaghie et al., 2004). Graded *Tbx3* expression is also observed with highest levels of transcripts present

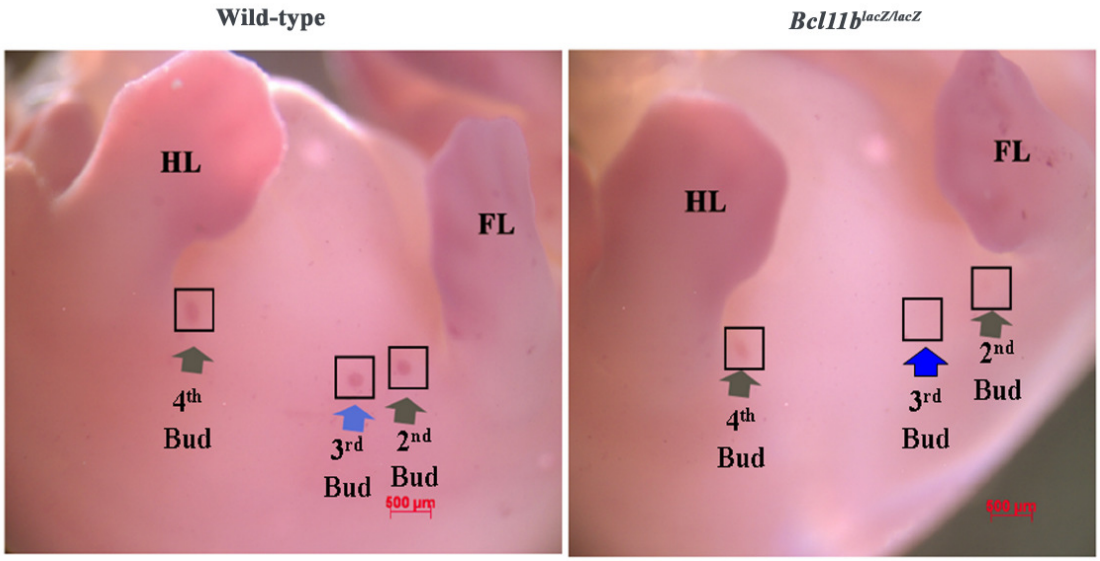
adjacent to the posterior margin of the forelimb bud and gradually diminishing down the flank to the anterior limit of the hindlimb bud. *Tbx3* expression at 10.5 dpc is stronger than at 10.25 dpc and from 11.5 dpc, expression of *Tbx3* is observed in the mammary placodes. *Fgf10* and *Tbx3* expression patterns were similar in *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant and wild-type embryos at 10.5 dpc, prior to mammary bud formation (Figure 5.2E-F). Expression of *Fgf10* in the lateral plate mesoderm was detected in both wild-type and *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 10.5 dpc embryos as a line extending from the forelimb bud to the level of the future site of the third pair of mammary placodes (Figure 5.2E). Similarly, expression of *Tbx3* was observed in the presumptive milk line in both wild-type and *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 10.5 dpc embryos (Figure 5.2F). These results indicated that expression of mesodermal mammary line markers, *Fgf10* and *Tbx3*, were unaffected in the *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant embryos, suggesting that specification of the mammary line was unaltered. A similar phenotype was observed in the *ska* homozygous mice where the gene affected in *ska* mice was mapped to Neuregulin3 (*Nrg3*) (Howard et al., 2005). In these mice, the third pair of mammary bud is hypoplastic from 11.75 dpc, leading to a reduced sized or absent bud. It has been suggested that *Nrg3* could act to transmit signals from *Fgf10* and/or *Tbx3* to the precursor mammary epithelial cells which appear to express *Wnt10b* (Veltmaat et al., 2004). Taken together, my results suggest that *Bcl11b* may function downstream of *Fgf10* and *Tbx3* but upstream of *Wnt10b* and *Lef1* in the specification and/or formation of number 3 pair of mammary buds and may function in concert with *Nrg3* to transmit signals from *Fgf10* and/or *Tbx3* to the precursor mammary epithelial cells. Alternatively, *Bcl11b* may also function in an alternative pathway for the formation of the third pair of mammary buds. Collectively, these results demonstrate the importance of both *Bcl11a* and *Bcl11b* in embryonic mammary development.

A

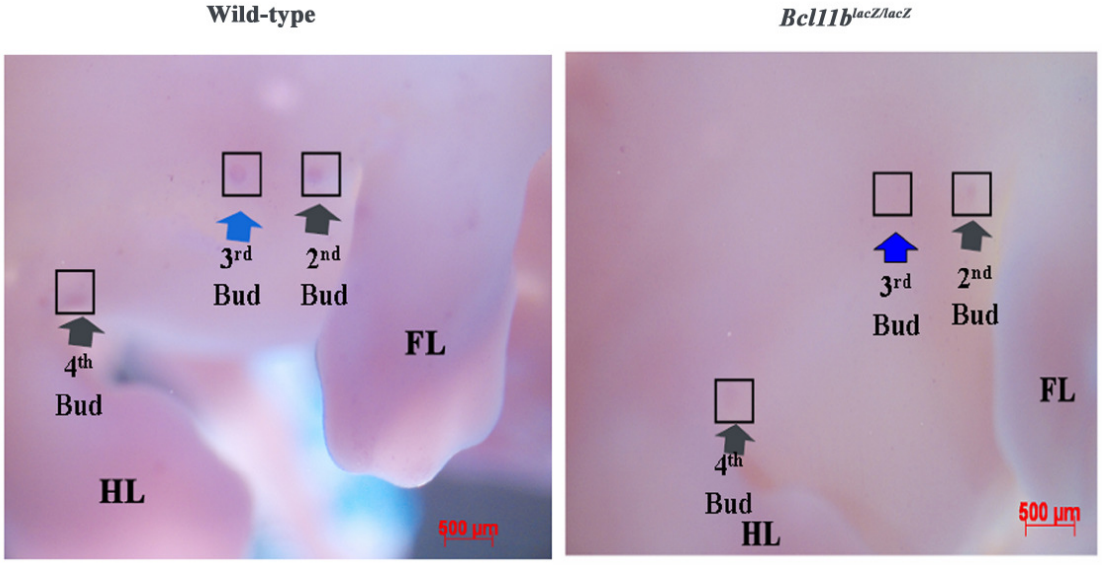
(Whole mount X-gal)



B 12.5 dpc (*Wnt10b*)

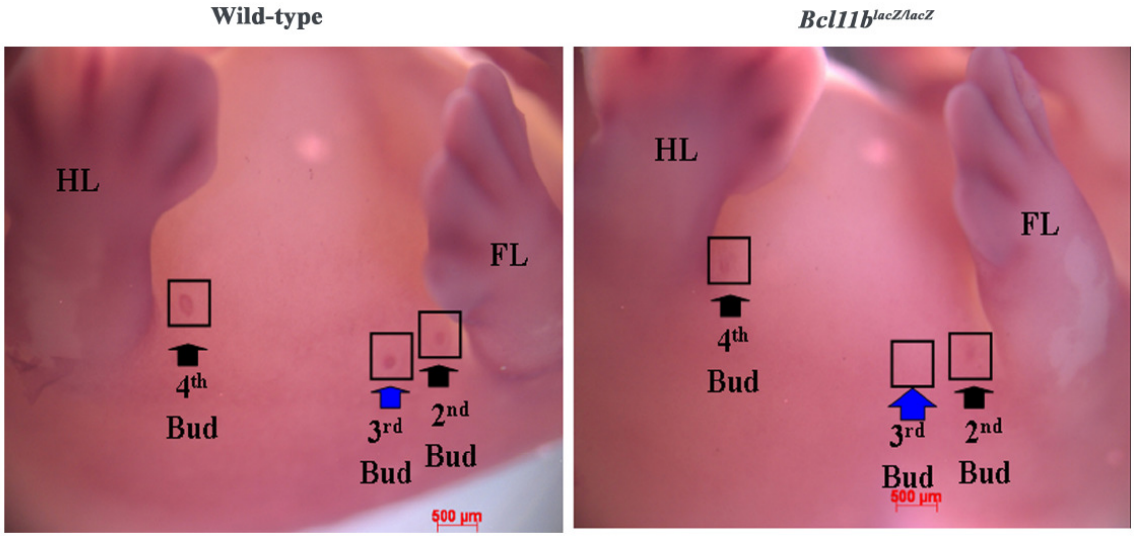


C 12.5 dpc (*Lef1*)



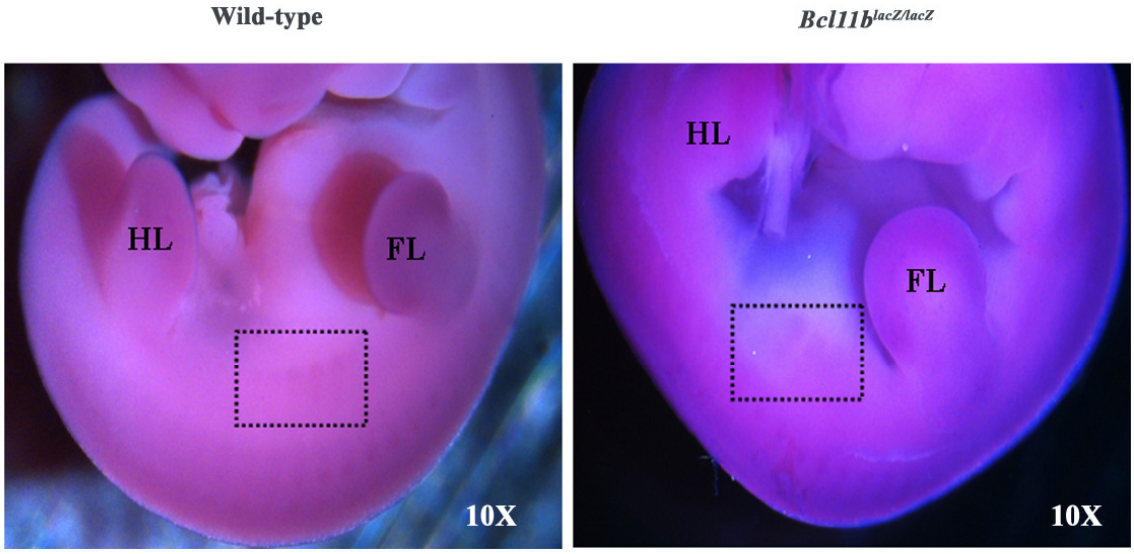
D

14.5 dpc (*Fgfr2TK*)



E

10.5 dpc (*Fgf10*)



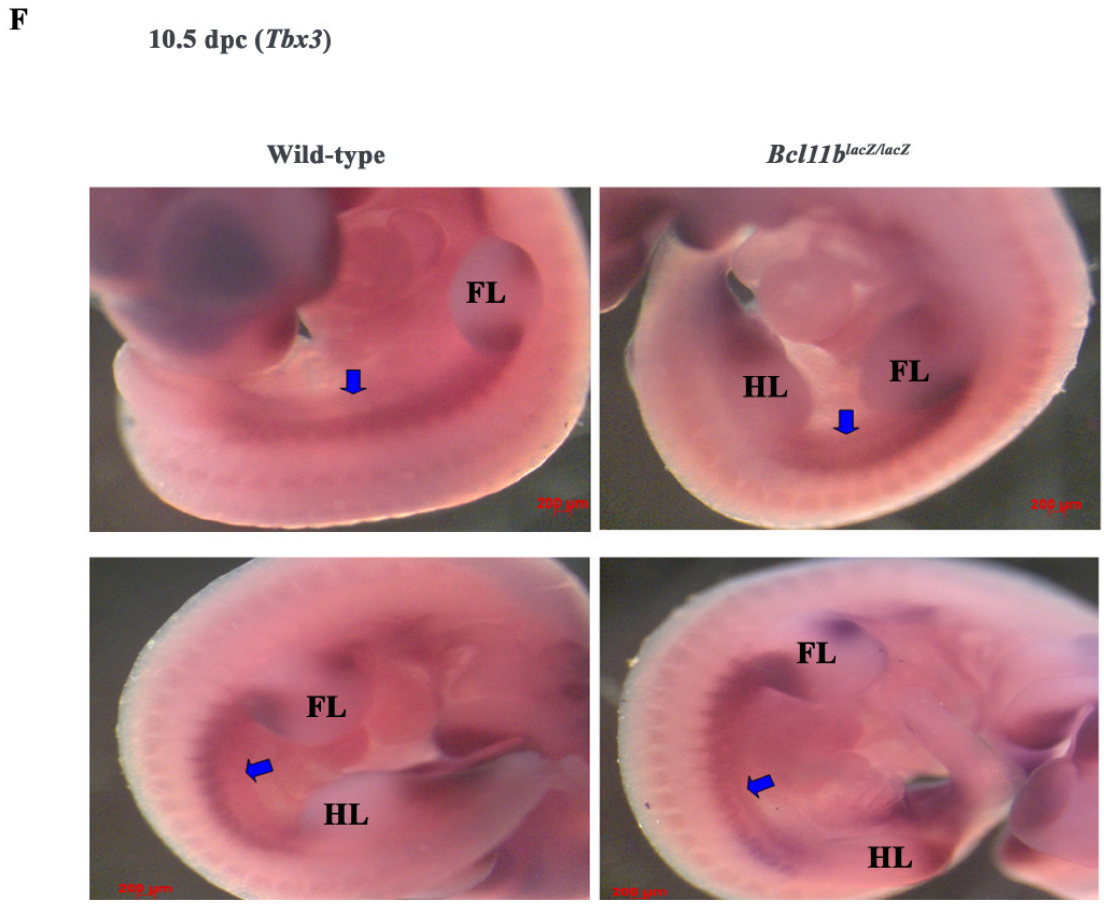


Figure 5.2. Analysis of *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant embryos using X-gal staining and *in situ* hybridization. **(A)** Whole mount images of X-gal staining of *Bcl11b*$^{lacZ/+}$ heterozygous and *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 12.5 dpc and 13.5 dpc embryos. Boxed regions show the approximate position of the third mammary bud. Red arrows indicate position of the second and fourth mammary buds. **(B-D)** Whole mount images of *in situ* hybridization of wild-type and *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 12.5 dpc embryos with **(B)** *Wnt10b* and **(C)** *Lef1* antisense probes; wild-type and *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 14.5 dpc embryos with **(D)** *Fgfr2TK* antisense probe. Boxed regions with black arrows show position of the second and fourth pairs of mammary buds while boxed regions with blue arrows show position of the third pair of mammary buds. Whole mount images of *in situ* hybridization of wild-type and *Bcl11b*$^{lacZ/lacZ}$ homozygous mutant 10.5 dpc embryos with **(E)** *Fgf10* and **(F)** *Tbx3* antisense probes. Boxed region shows the region of positive Fgf10 staining. Blue arrows show the location of the presumed milk line as indicated by Tbx3 staining in the embryos. FL – Forelimb; HL – Hindlimb.

5.2.3 Loss of *Bcl11a* in virgin glands causes disruption of mammary architecture

Both *Bcl11* null mice died several hours after birth from unknown reasons (Liu et al., 2003b; Wakabayashi et al., 2003b), so I was unable to use these mice to study postnatal mammary development. Therefore, our lab utilized recombineering to generate conditional knockout (cko) alleles of both *Bcl11* genes ($Bcl11^{flox/flox}$) (Mice were generated by Pentao Liu). In the cko mice, exon 1 and exon 4 of *Bcl11a* and *Bcl11b* were flanked by *lox*P sites respectively (Figure 5.3A-D); upon expression of Cre recombinase, the intervening sequences between the *lox*P sites will be deleted, resulting in inactivation of the genes. These cko mice were subsequently crossed to several Cre-lines to delete *Bcl11* genes temporally and spatially in the mammary gland.

To examine whether *Bcl11a* and *Bcl11b* are required for mammary development in the virgin gland, I first crossed the $Bcl11^{flox/flox}$ mice to the Rosa26-CreERT2 mice where *Cre-ERT2* fusion gene expression is controlled by the constitutive active *Rosa26* promoter (Hameyer et al., 2007). In the CreERT2 protein, Cre is fused to the ligand-binding domain of a mutated human estrogen receptor (ERT) that recognizes tamoxifen (TAM) or its derivative 4-hydroxytamoxifen (4-OHT). Homozygous floxed mice carrying the *Rosa26-CreERT2* allele (*CreERT2;* $Bcl11a^{flox/flox}$) were born in a normal Mendelian ratio. The females of the *CreERT2;* $Bcl11^{flox/flox}$ mice were able to nurse and nurture their pups. Whole mount and histological analyses of mammary glands from *CreERT2;* $Bcl11^{flox/flox}$ females appeared normal with no noticeable developmental defects. To gain insight into the function of *Bcl11a* in the virgin mammary gland, TAM was administrated to 4-8-week-old *CreERT2;* $Bcl11a^{flox/flox}$ and control (*CreERT2;* $Bcl11a^{flox/+}$ heterozygous and wild-type) females to induce *Bcl11a* deletion. The injected mice were analyzed three weeks post injection. Genomic DNA was extracted from TAM-injected mammary glands and Cre-mediated excision was detected using PCR primers as shown in Figure 5.4A. Cre-mediated excision of *Bcl11a* was confirmed by the presence of deletion bands in the mammary glands (Figure 5.4B) and quantitative real time PCR (qRT-PCR) amplification of genomic DNA derived from these mammary glands showed that the deletion efficiency was ~60%.

Whole-mount analysis of the mammary glands of control females did not reveal any obvious defects following TAM administration (Figure 5.5A). In contrast, the mammary glands of *CreERT2; Bcl11a*$^{flox/flox}$ homozygous females had enlarged primary ducts and reduced secondary and tertiary branches as compared to control glands (Figure 5.5A). Histological examination demonstrated that lumens of the mammary ducts were distended and lined by a thin epithelium with no clear distinction between the luminal and basal/myoepithelial layers as compared to the control glands (Figure 5.5B). The ducts appeared to have a thin cell layer in many areas instead of the distinct luminal and basal layers in the control ducts. To determine if basal and luminal cells were still present in this thin layer of epithelium cells, immunohistochemistry was conducted using antibodies to luminal marker Aquaporin 5 (Shillingford et al., 2003) and basal marker, cytokeratin 14 (CK14). The dilated ducts were stained positively for both markers, indicating that some luminal and basal epithelial cells were still present (Figure 5.6A). The disruption in mammary architecture was exemplified by the Aquaporin 5 positive luminal layer and CK14 positive basal layer forming a thin epithelial layer in the *Bcl11a*-deficient mammary glands (Figure 5.6A). This was confirmed by using additional basal markers such as p63 and smooth muscle actin (SMA), and luminal marker, cytokeratin 18 (CK18) (Figure 5.6B-D). Interestingly, some regions of the epithelium in *Bcl11a*-deficient glands appeared to be monolayered with basal and luminal cells found within the same layer (Figure 5.6A-D). Next, to determine whether deletion of *Bcl11a* led to abnormal proliferation or cell death, the mammary sections were stained with antibodies to cleaved Caspase 3 and Ki67. No obvious differences between the mutant and the control glands were observed, though a slight decrease in the number of Ki67 positive cells was observed in the mutant (Figure 5.7A and 5.7B). This suggests that loss of *Bcl11a* did not result in abnormal proliferation and apoptosis of mammary epithelial cells in the virgin glands. Alternatively, this may also suggest that the proliferative and apoptotic abnormalities occurred shortly after TAM injection and were thus not detected at the time of analysis (3 weeks after injection).

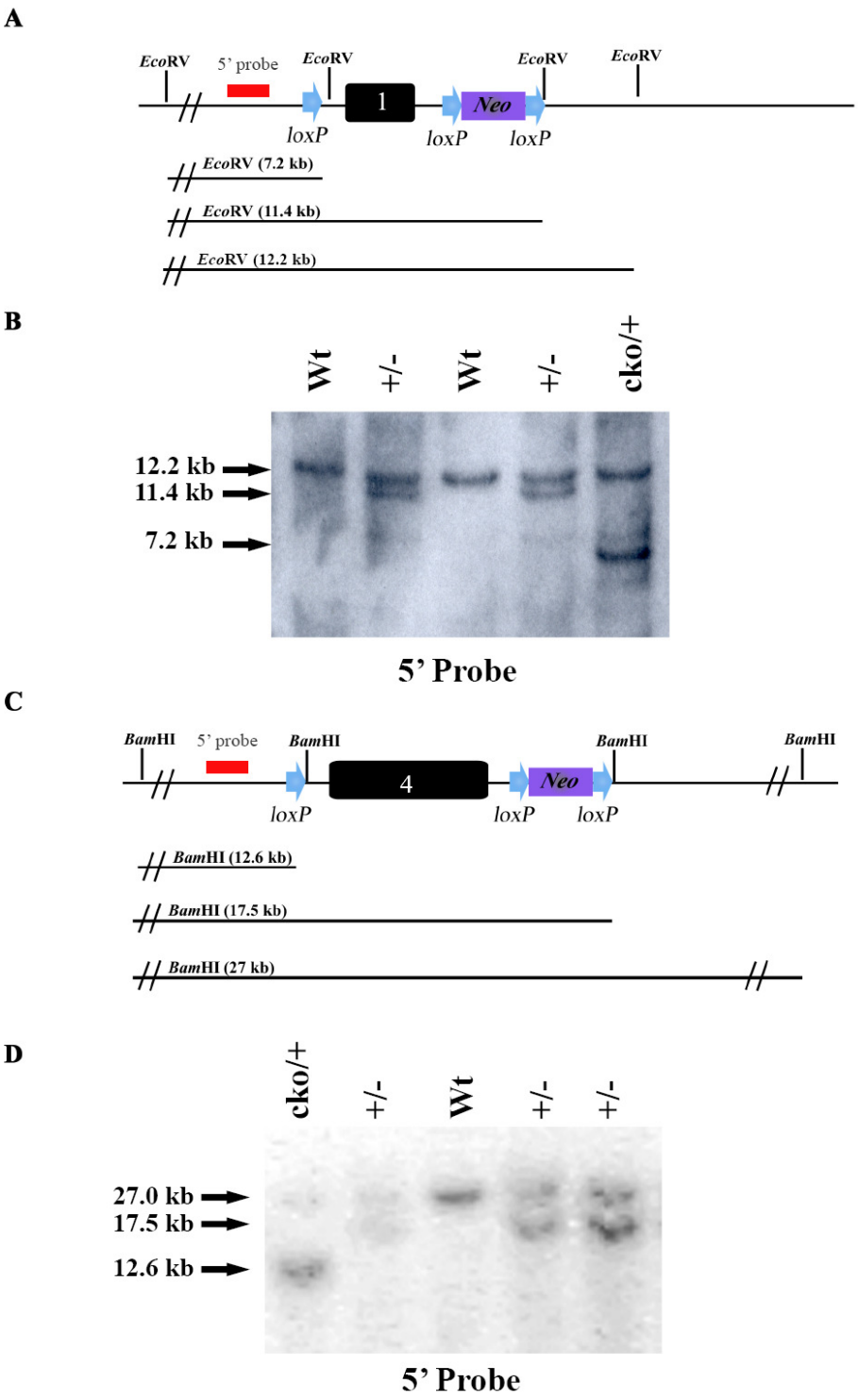
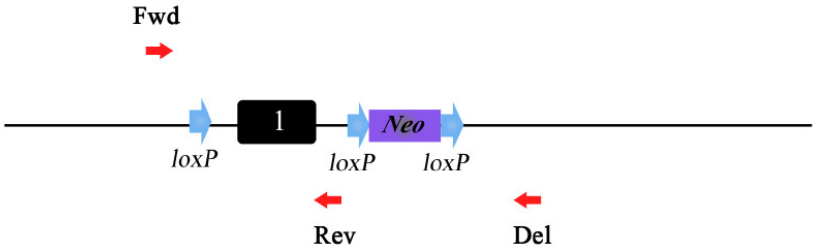


Figure 5.3. Generation of the *Bcl11* conditional knockout mice. **(A)** Schematic diagram showing targeting strategy of the *Bcl11a* conditional knockout (cko) allele. **(B)** Southern blot confirmation of the *Bcl11a* cko allele using 5' probe. Presence of cko allele is detected by a 7.2 kb band in the Southern analysis. **(C)** Schematic diagram showing targeting strategy of the *Bcl11b* cko allele. **(D)** Southern blot confirmation of the *Bcl11b* cko allele using 5' probe. Presence of cko allele is detected by a 12.6 kb band in the Southern analysis. (Southern blot analysis data provided by Pentao Liu).

A



Fwd x Rev = 470 bp (Cko band)
Fwd x Del = 700 bp (Deletion band)

B

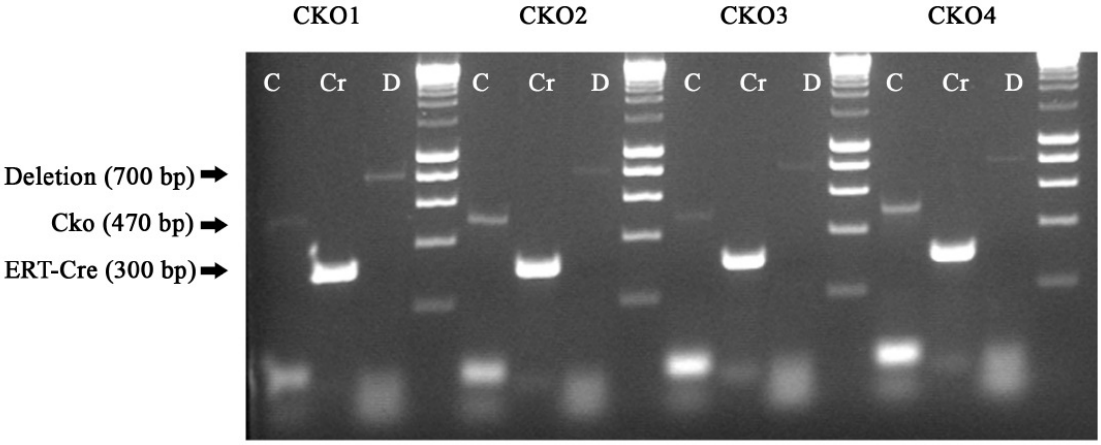


Figure 5.4. Detection of deletion of *Bcl11a* after Cre expression. **(A)** Schematic diagram showing relative position of primers in the *Bcl11a* locus used for detection of deletion band after Cre expression. **(B)** Gel image showing PCR products obtained with primers using genomic DNA extracted from mammary glands of tamoxifen injected *Cre-ERT2; Bcl11a*$^{flox/flox}$ mice. CKO1-4 represents mammary glands from four independent mice. C: cko PCR; Cr: Cre PCR; D: Deletion PCR.

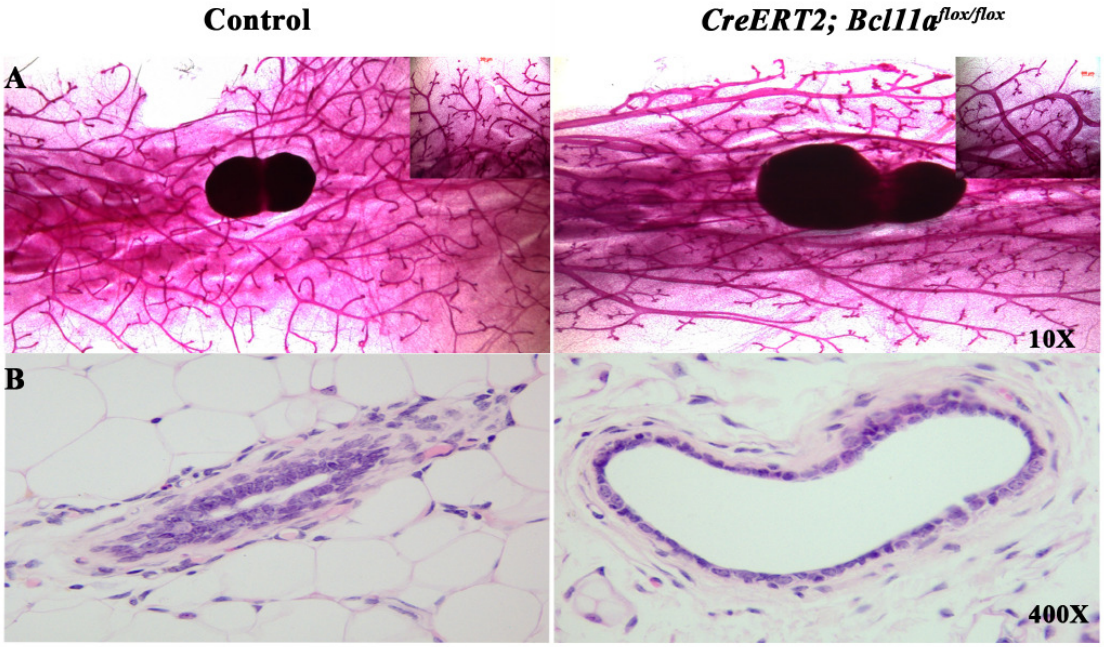


Figure 5.5. Morphological analysis of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands. (A) Whole mount carmine alum-stained and **(B)** H&E stained sections of TAM injected control (Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$) and *CreERT2*; *Bcl11a*$^{flox/flox}$ mammary glands.

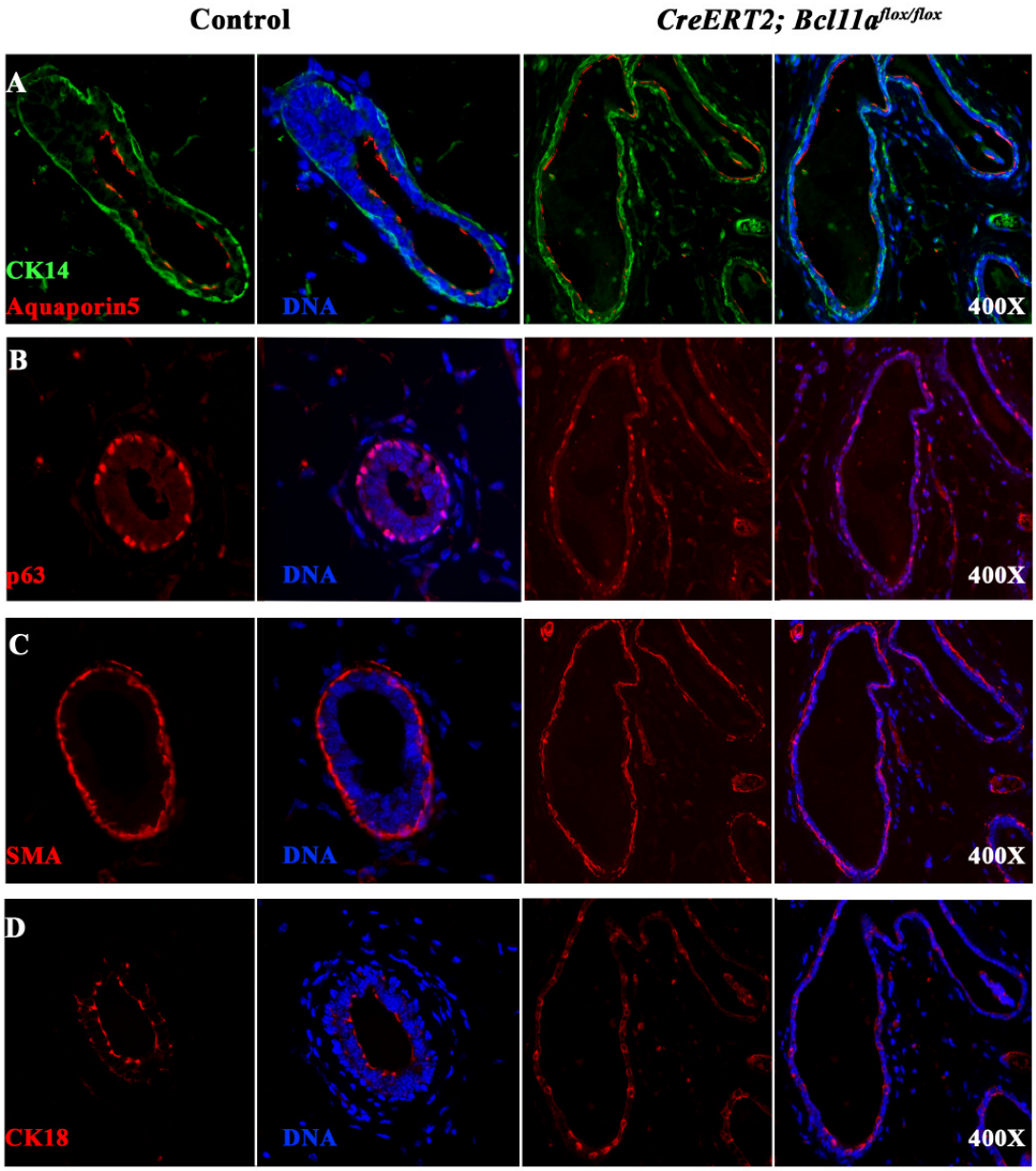


Figure 5.6. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands sections using luminal/basal markers. Immunostaining of sections using antibodies against **(A)** Aquaporin 5, CK14, **(B)** p63, **(C)** SMA and **(D)** CK18. Control: Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$.

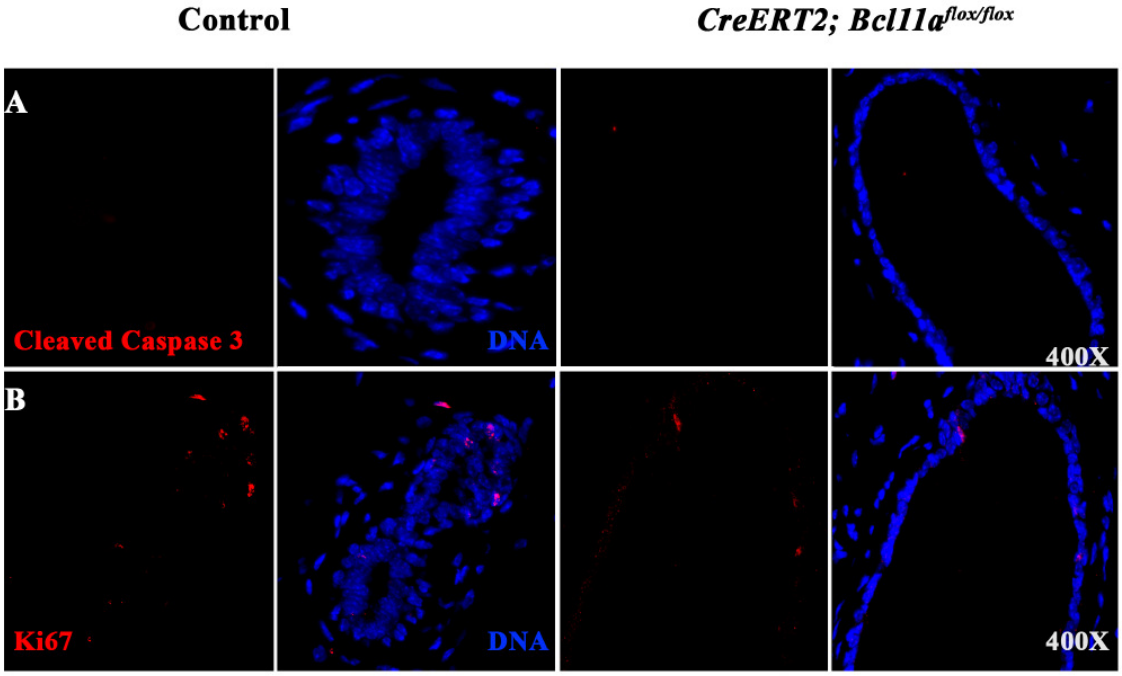


Figure 5.7. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands. Immunostaining of sections using antibodies against **(A)** Cleaved Caspase 3 and **(B)** Ki67. Control: Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$.

5.2.4 Loss of *Bcl11a* in the mammary epithelium results in a loss of Gata-3$^+$ and an increase in ERα$^+$ epithelial cells

Gata-3 is the most highly enriched transcription factor in the mammary ductal epithelium and has recently been shown to be essential for the maintenance of luminal cell fate (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). Interestingly, I found that loss of *Bcl11a* in the virgin glands resulted in a dramatic reduction in the number of Gata-3$^+$ luminal cells, suggesting a general loss of luminal cells (Figure 5.8A).

Estrogen receptor (ER) is required for ductal proliferation and differentiation (Korach, 2000). In the wild-type virgin gland, up to 37% of the luminal cells in the ducts and alveoli are ERα-positive (ERα$^+$), and most of these ERα$^+$ cells seem to be non-dividing cells that instruct adjacent epithelial cells to proliferate via paracrine regulatory effects (Clarke et al., 1997; Mallepell et al., 2006; Russo et al., 1999; Seagroves et al., 1998). It has been shown that the basal mammary stem cell (MaSC)-enriched population is ERα and progesterone receptor (PR) negative, and that ERα is localized to the more differentiated luminal cells (Asselin-Labat et al., 2006; Asselin-Labat et al., 2007; Lindeman and Visvader, 2006). The luminal population is composed of two phenotypically distinct lineages (Stingl and Watson, manuscript in preparation) (Sleeman et al., 2007). The Sca1$^+$ lineage is characterized by high levels of ERα expression whereas the Sca1$^-$ lineage is characterized by high levels of milk proteins expression. Hence, the Sca1$^+$ luminal progenitors were hypothesized to give rise to ERα$^+$ differentiated cells whereas the Sca1$^-$ luminal progenitors to function as alveolar precursors (Personal communication with John Stingl). Since *Bcl11a* is expressed in both the Sca1$^+$ and the Sca1$^-$ luminal progenitors, I examined the ERα status in the *Bcl11a*-deficient virgin glands. In the *Bcl11a*-deficient virgin glands, there was an apparent increase in the number of ERα$^+$ cells (Figure 5.8B). Quantification of ERα$^+$ luminal cells showed that 49.0 $\pm$ 8.8% of luminal cells were ERα$^+$ in the *Bcl11a*-deficient virgin glands (n=2) compared to 25.0 $\pm$ 7.1% in control glands (n=2). Since ERα was not increased in RT-PCR (See below, Figure 5.11), this result indicated a selective loss of the ERα$^-$ cells in the *Bcl11a*-deficient virgin glands. Consistent with this result, only the Sca1$^-$ (ERα$^-$) population showed decreased mammary colony-forming cell (Ma-CFC) capabilities upon *Bcl11a* deletion (Appendix A.8). The Sca1$^+$ (ERα$^+$) luminal progenitors, on the other

hand did not show any significant reduction in Ma-CFC capabilities (Appendix A.8). Taken together, these results suggest that deletion of *Bcl11a* results in the loss of $Sca1^-$ ($ER\alpha^-$) luminal cells, hence leading to a relative increase in the percentage of $ER\alpha^+$ luminal cells and dysregulating the balance between the $ER\alpha^+$ and $ER\alpha^-$ luminal cells.

In the *Gata-3* conditional knockout mice, deletion of *Gata-3* in the mammary glands resulted in a loss of $ER\alpha^+$ luminal cells (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). This suggests that there may be an $ER\alpha^+$ luminal population which is not affected by loss of *Bcl11a* and *Gata-3*. Alternatively, this suggests that Gata3 might not the primary effector of Bcl11a in the virgin gland or that these *Bcl11a*-deficient $ER\alpha^+$ luminal cells no longer depend on *Gata-3*.

Over-expression of Notch intracellular domains impairs mammary ductal growth and causes dilated ducts in virgin glands (Hu et al., 2006; Jhappan et al., 1992). Our lab had previously showed that the T-cell leukaemia phenotypes observed in the *Bcl11a* mutant lymphocytes might be caused by abnormal Notch signalling (Liu et al., 2003b). I thus examined whether there was altered expression of components of the Notch pathway in the *Bcl11a*-deficient virgin glands. Immunohistochemistry for Notch1 and its ligand, Jagged1 showed a clear increase of Notch1 and a modest increase of Jagged1 expression in the sections of the *Bcl11a*-deficient virgin glands (Figure 5.9A and 5.9B). Therefore deletion of *Bcl11a* in the virgin gland appeared to lead to aberrant activation of the Notch pathway.

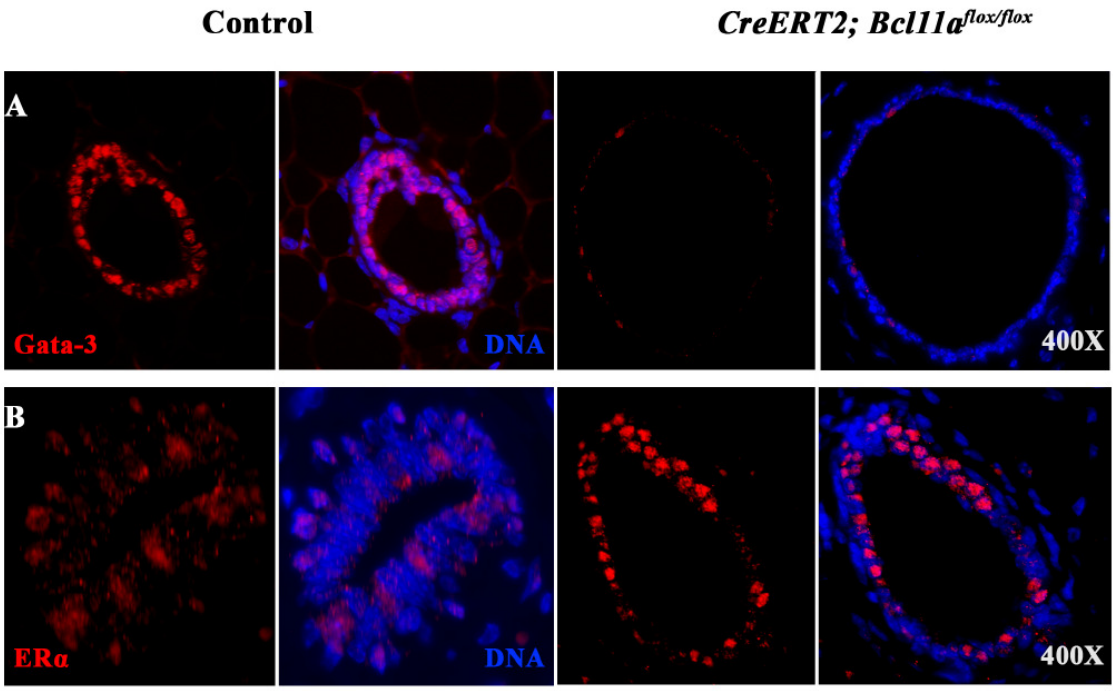


Figure 5.8. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands. Immunostaining of sections using antibodies against **(A)** Gata-3 and **(B)** ERα. Control: Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$.

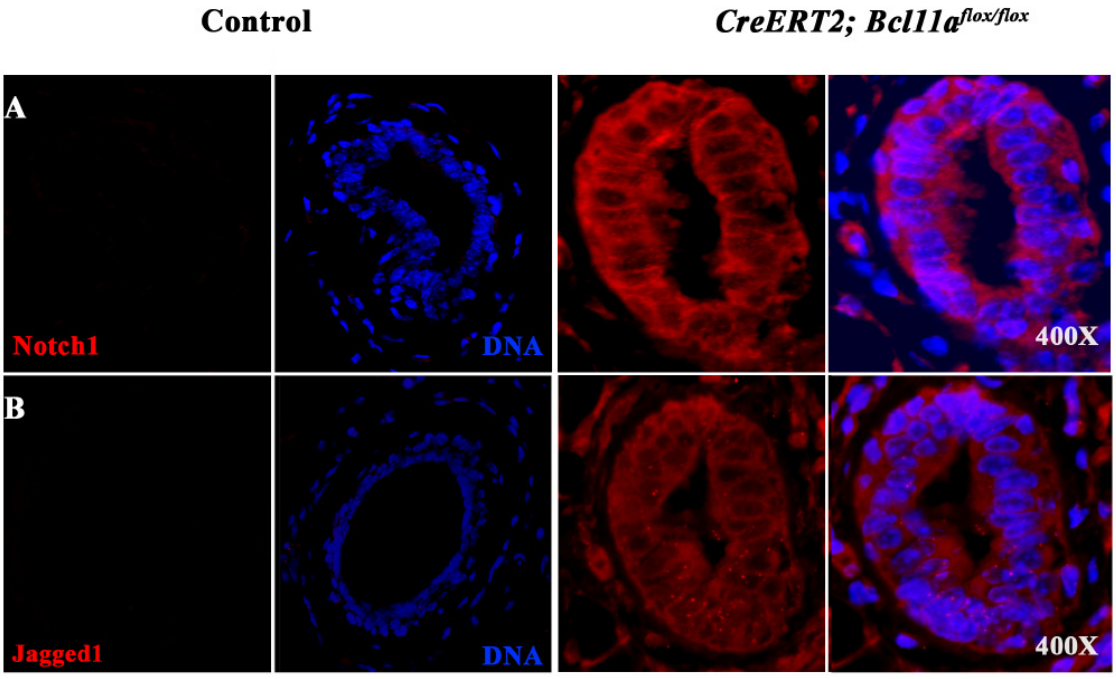


Figure 5.9. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands sections using additional markers. Immunostaining of sections using antibodies against **(A)** Notch1 and **(B)** Jagged1. Control: Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$.

5.2.5 *Bcl11a*-deficient virgin glands exhibit a shift towards luminal profile

Notch signalling has been suggested to influence lineage specification of the mammary epithelium (Buono et al., 2006; Smith et al., 1995). Given that the deletion of *Bcl11a* disrupted ductal architecture and resulted in ectopic activation of Notch signalling, I stained the mammary epithelial cells with antibodies to cell surface markers CD24 and CD49f and analyzed them by flow cytometry. The total percentage of epithelial cells ($CD24^+$) in *Bcl11a*-deficient mammary glands was slightly decreased (12.99 $\pm$ 0.45% compared to 15.65 $\pm$ 0.35% in the control glands; n=3). In the control glands (n=3), luminal ($CD24^{hi}CD49f^+$) and basal cells ($CD24^+CD49f^{hi}$) accounted for about 5.0 $\pm$ 1.6% and 5.7 $\pm$ 1.8% of total Lin^- cells respectively. Loss of *Bcl11a* resulted in an increase in the percentage of $CD24^{hi}CD49f^+$ luminal cells (8.80 $\pm$ 1.0%), and a decrease in percentage of $CD24^+CD49f^{hi}$ basal cells (2.56 $\pm$ 0.2%) (n=3) (Figure 5.10). Since most of the Gata-3-expressing cells were lost in the *Bcl11a*-deficient gland, the higher percentage of $CD24^{hi}CD49f^+$ luminal cells most likely reflected the relative increase of more differentiated $ER\alpha^+$ luminal cells rather than a general increase of total luminal cells.

To further interrogate the molecular changes, I analyzed gene expression changes in the *Bcl11a*-deficient virgin glands using RT-PCR. As shown in Figure 5.11, there was reduction in expression of luminal (*NKCC1*, *Muc1*) and basal (*CK14*) markers in the *Bcl11a*-deficient mammary glands, indicating the important roles of Bcl11a in both luminal and basal lineages. Additionally, consistent with the immunohistochemistry results, *Gata-3* expression was significantly reduced and *Bcl11b* and *Notch1* expression were clearly increased in the *Bcl11a*-deficient mammary glands. *ERα* expression however, was not significantly reduced compared to *Muc1* and *NKCC1*, further highlighting the selective depletion of the $ER\alpha^-$ cells. Elf5 is a key player in alveolar cell fate specification and is expressed at low levels in the virgin gland. *Bcl11a* deletion caused a near complete loss of *Elf5* expression. This result, and the selective depletion of $ER\alpha^-$ cells, suggested a likely loss of alveolar progenitors when *Bcl11a* was deleted.

These phenotypes in the virgin glands have clearly established that *Bcl11a* is a critical regulator in normal mammary development.

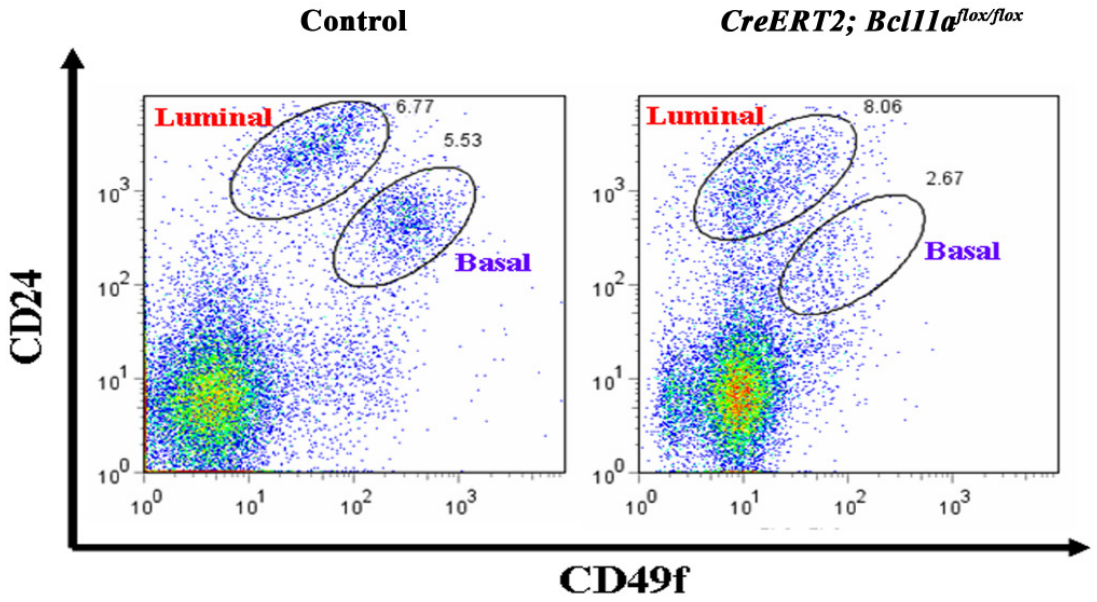


Figure 5.10. Mammary FACS profile of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands. Flow cytometric analysis of isolated epithelial cells using CD24 and CD49f markers. There is an increase in the percentage of luminal cells (CD24^{hi}CD49f^{+}) in the *Bcl11a*-deficent glands compared to the control glands (n=3). Control: Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$.

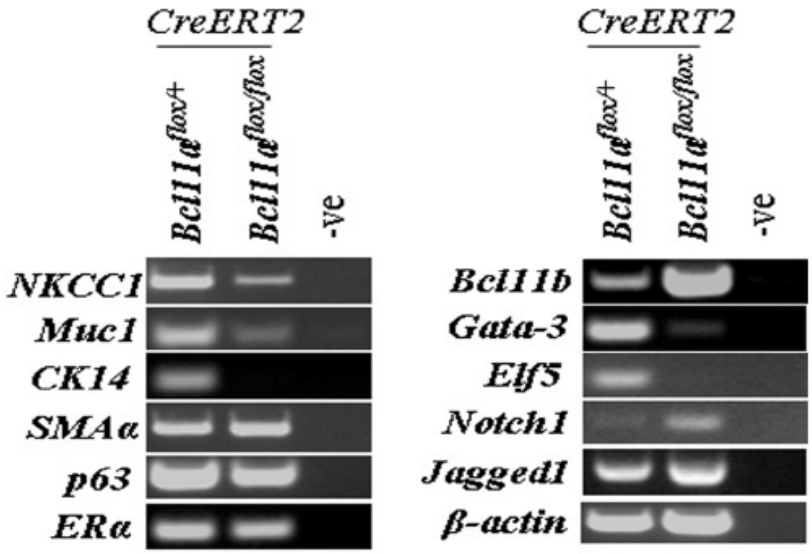


Figure 5.11. Analysis of tamoxifen treated control and *Cre-ERT2; Bcl11a*$^{flox/flox}$ whole mammary glands using semi-quantitative RT-PCR. RT-PCR shows a decrease in levels of both luminal and basal markers and up-regulation of Notch receptors in the *Bcl11a*-deficient glands. β-actin is used as a control. Control: Wild-type/*CreERT2*; *Bcl11a*$^{flox/+}$. –ve indicates no template control.

5.2.6 Loss of *Bcl11b* in virgin glands results in precocious alveolar development

To investigate the function of *Bcl11b* in the virgin mammary gland, I administered TAM to 4-8-week-old *CreERT2; Bcl11b*$^{flox/flox}$ and control (*Bcl11b*$^{flox/+}$ heterozygous and wild-type) females to induce *Bcl11b* deletion. The injected mice were analyzed three weeks post injection. Genomic DNA was extracted from TAM-injected mammary glands and Cre-mediated excision was detected using PCR primers as shown in Figure 5.12A. Cre-mediated excision of *Bcl11b* was confirmed by the presence of deletion bands in the mammary glands (Figure 5.12B) and quantitative real time PCR (qRT-PCR) amplification of genomic DNA derived from these mammary glands showed that the deletion efficiency was ~60%. In addition, mammary epithelial cells were sorted into luminal ($CD24^{hi}CD49f^{+}$) and basal ($CD24^{+}CD49f^{hi}$) fractions using FACS and genomic DNA was extracted and Cre-mediated excision of *Bcl11b* was again confirmed by the presence of deletion bands (Figure 5.12C).

Strikingly, deletion of *Bcl11b* resulted in pregnancy-like alveologenesis in the virgin glands (Figure 5.13A). Numerous alveolar structures developed along the primary ducts of the *Bcl11b*-deficient virgin glands and histological examination revealed that these alveoli resembled that of a wild-type day 10 gestation mammary gland (Figure 5.13B-C). To confirm that the alveoli in the *Bcl11b*-deficient virgin glands had indeed undergone secretory differentiation, I stained sections of the glands with an antibody to β-casein. As shown in Figure 5.14A, β-casein producing cells were present in the *Bcl11b*-deficient virgin glands but not in the control glands, indicating that loss of *Bcl11b* resulted in proliferation and differentiation of luminal (ductal and alveoli) mammary epithelial cells normally only found in the pregnant gland. Immunostaining with anti-Aquaporin 5 antibody showed that Aquaporin 5 was not detected in *Bcl11b*-deficient virgin glands even in ducts (Figure 5.14B), a phenomenon similar to the wild-type mid-gestation mammary glands (Shillingford et al., 2003). To further investigate the mammary defects, I examined additional markers of luminal (CK18) and basal (p63 and SMA) cells using immunohistochemistry. The alveolar-like structures in the *Bcl11b*-deficient glands expressed basal markers p63 and SMA in the outer layer, as well as CK18 in the luminal cells (Figure 5.14B-E), suggesting that the mammary architecture of

Bcl11b-deficient glands were similar to that wild-type mid-gestation glands. Interestingly, some of the p63 positive cells appeared to be in the luminal layer (Figure 5.14C).

Activation of Stat5-mediated transcriptional activities by phosphorylation (p-Stat5) is essential for normal lobulo-alveolar cell proliferation and differentiation. Phosphorylation of Stat5 is very low in virgin and during early pregnancy but rises sharply after day 14 of pregnancy (Liu et al., 1997). Immunostaining with antibody to p-Stat5 revealed the absence of p-Stat5 protein in *Bcl11b*-deficient glands (Figure 5.15A), demonstrating that these abnormal glands were earlier than day 14 of pregnancy. In addition, immunostaining with anti-ERα antibodies did not reveal any significant change in the number of $ER\alpha^+$ cells in the mammary ducts (Figure 5.15B). It was noted that there were very few $ER\alpha^+$ cells in the alveolar-like structures in the *Bcl11b*-deficient virgin glands (Figure 5.15B). Next, to determine whether deletion of *Bcl11b* led to abnormal proliferation or cell death, the mammary sections were stained using antibodies to Ki67 and cleaved Caspase 3. No obvious differences between the mutant and the control glands were observed (Figure 5.16A and 5.16B), suggesting that loss of *Bcl11b* did not result in abnormal proliferation or apoptosis of mammary epithelial cells in the virgin glands. However, it was likely that the initial proliferation of epithelial cells to generate the alveolar-like structures had probably occurred before the analysis that was performed 3 weeks after TAM administration.

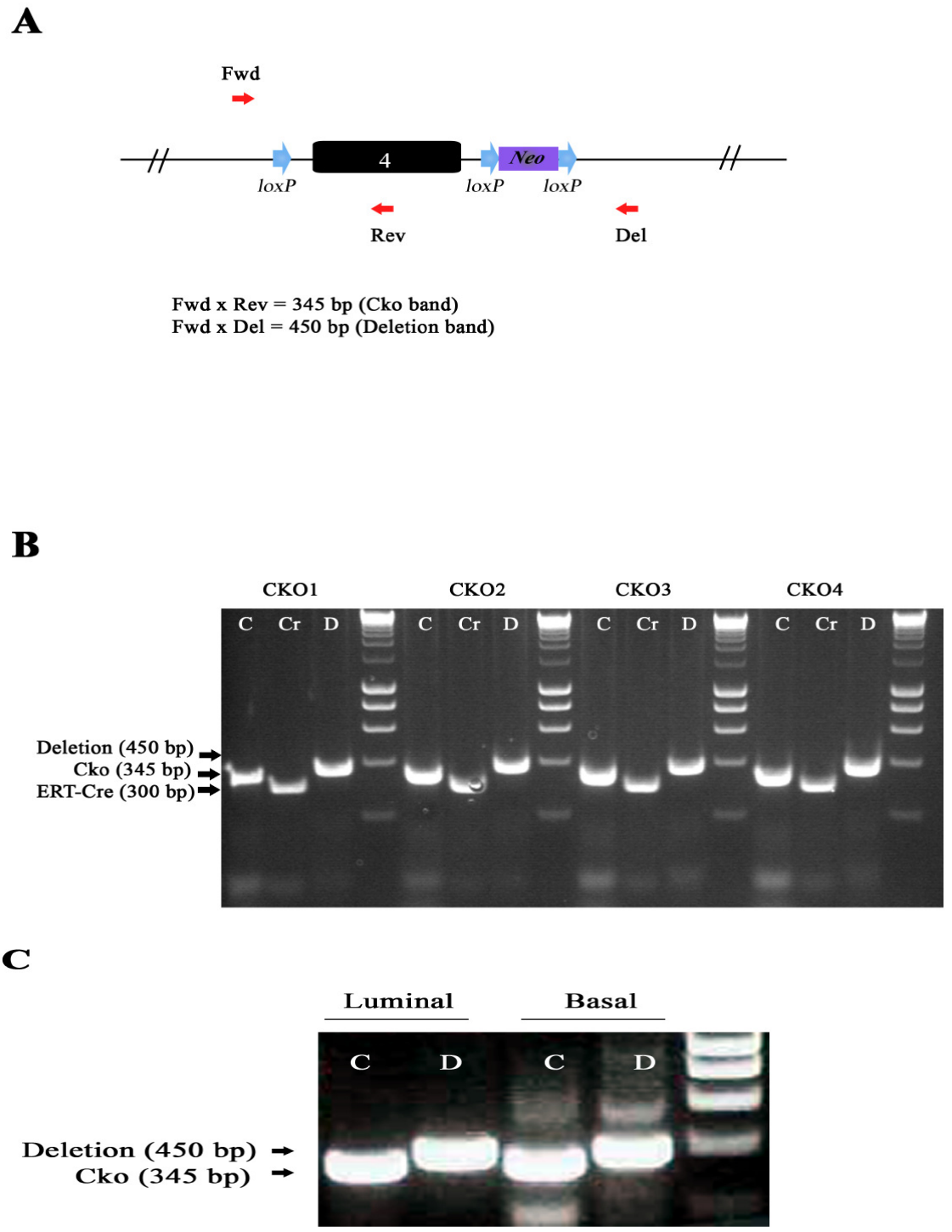


Figure 5.12. Detection of deletion of *Bcl11b* after Cre expression. **(A)** Schematic diagram showing relative position of primers in the *Bcl11b* locus used for detection of deletion band after Cre expression. **(B)** Gel image showing PCR products obtained with primers using genomic DNA extracted from mammary glands of tamoxifen injected *CreERT2*; *Bcl11b*$^{flox/flox}$ mice. **(C)** Gel image showing PCR products obtained from genomic DNA extracted from sorted luminal (CD24^{hi}/CD49f^{+}) and basal (CD24^{+}/CD49f^{hi}) epithelial cells. CKO1-4 represents mammary glands from four independent mice. C: cko PCR; Cr: Cre PCR; D: Deletion PCR.

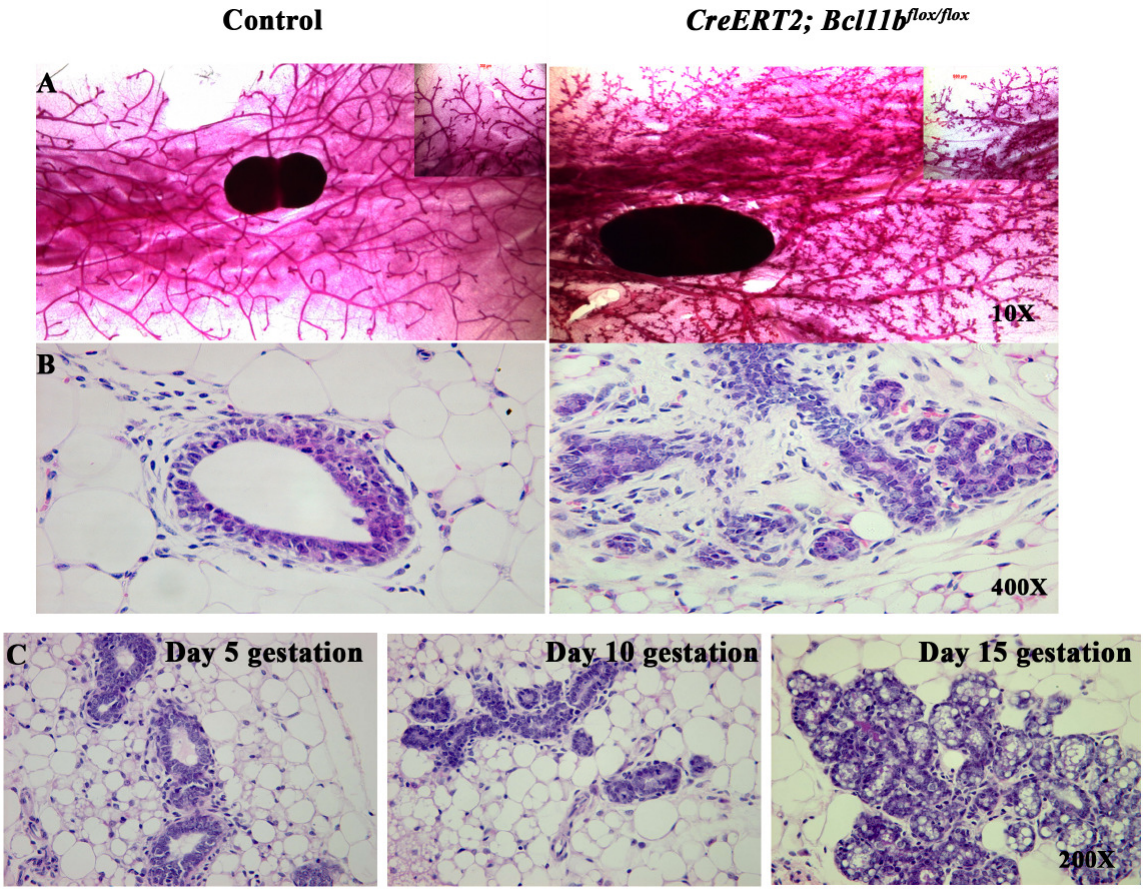


Figure 5.13. Morphological analysis of tamoxifen treated control and *Cre-ERT2; Bcl11b*$^{flox/flox}$ mammary glands. (A) Whole mount carmine alum-stained and **(B)** H&E stained sections of TAM injected control (Wild-type/*CreERT2*; *Bcl11b*$^{flox/+}$) and *CreERT2*; *Bcl11b*$^{flox/flox}$ mammary glands. **(C)** H&E stained sections of day 5/10/15 gestation wild-type mammary glands. The *Bcl11b*-deficient virgin glands resemble that of a day 10 gestation gland.

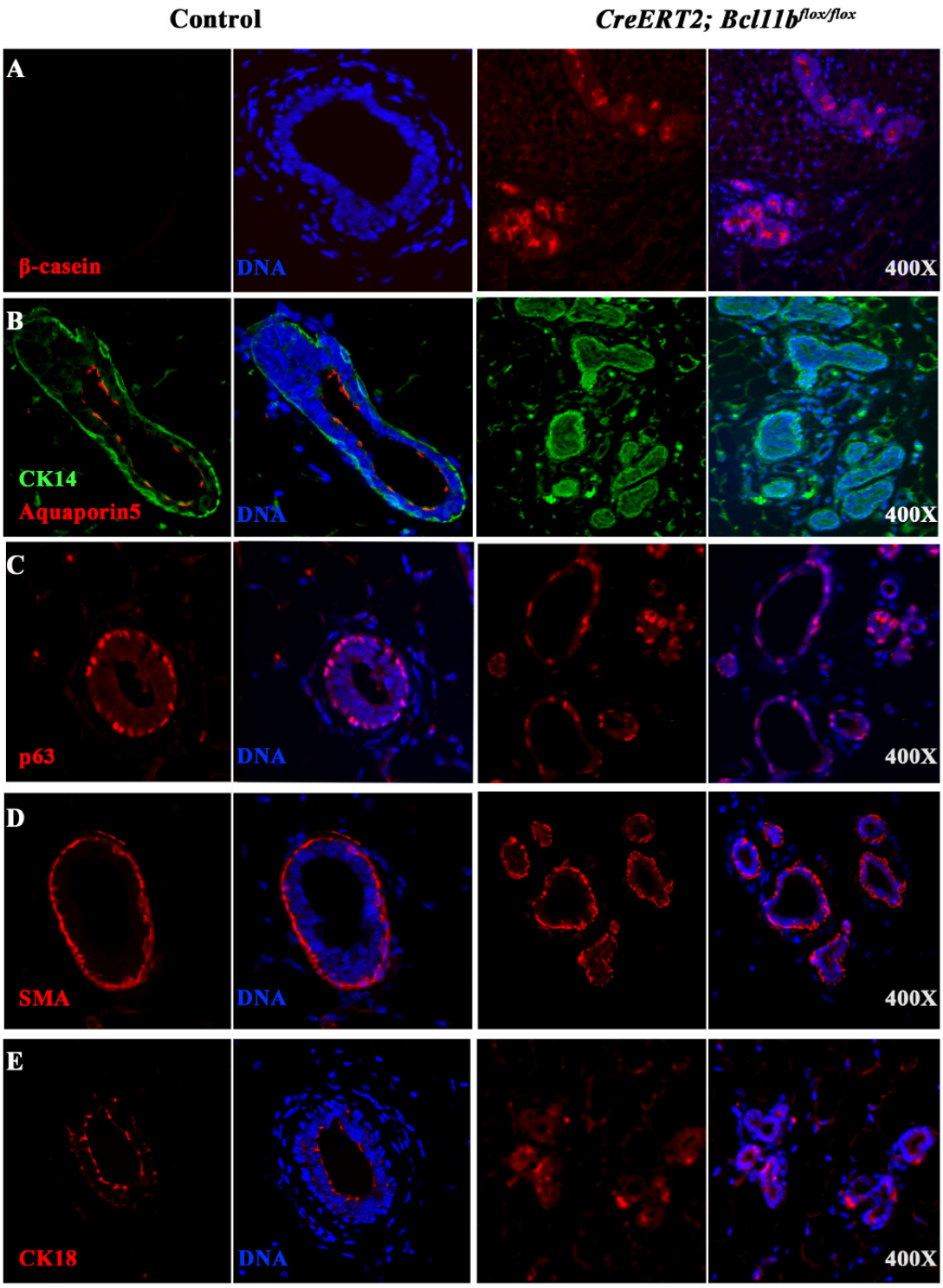


Figure 5.14. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11b*$^{flox/flox}$ mammary glands sections using luminal/basal markers. Immunostaining of sections using antibodies against **(A)** β-casein, **(B)** Aquaporin 5, CK14, **(C)** p63, **(D)** SMA and **(E)** CK18. Control: Wild-type/*CreERT2*; *Bcl11b*$^{flox/+}$. Note: Images from control panels were the same as that used for Figure 5.6.

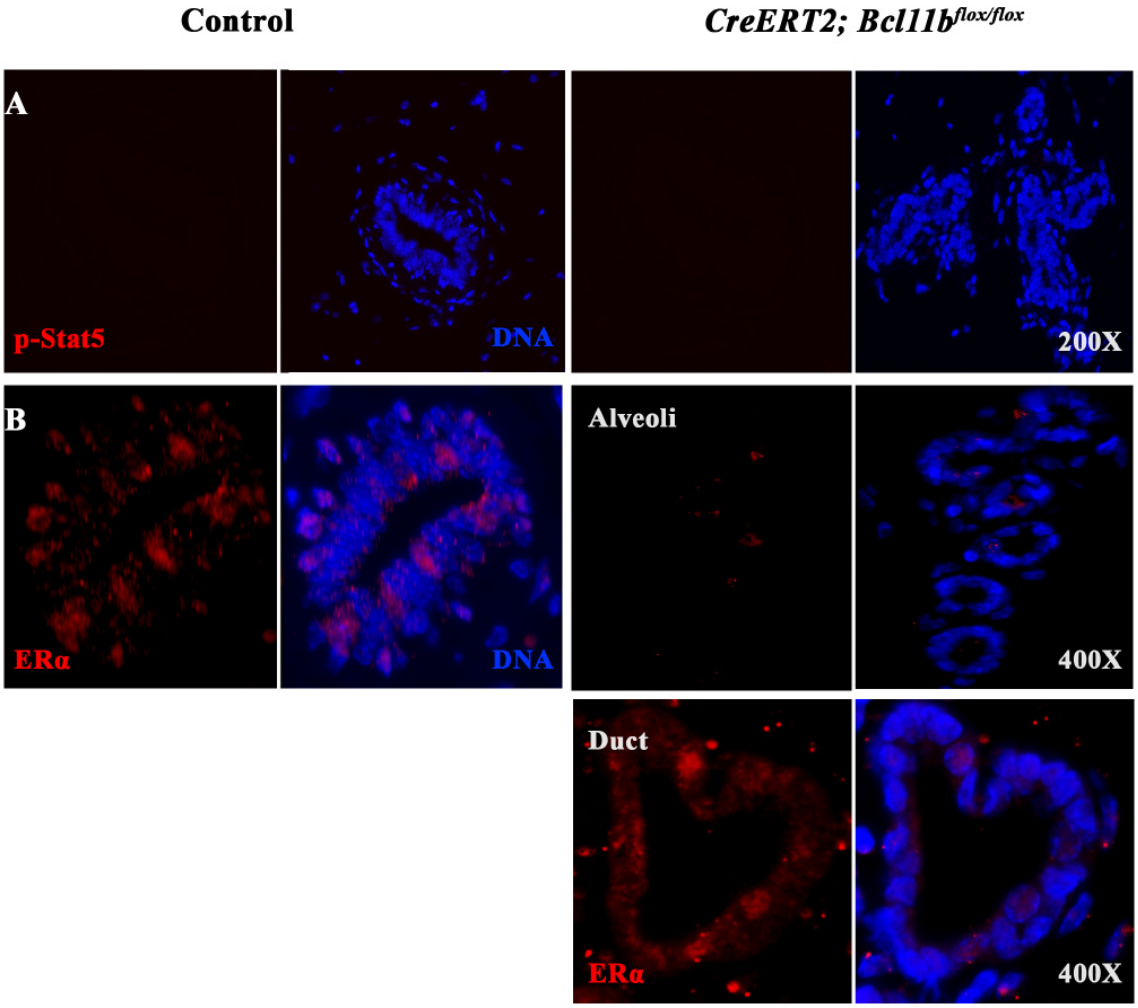


Figure 5.15. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11b*$^{flox/flox}$ mammary glands. Immunostaining of sections using antibodies against **(A)** p-Stat5 and **(B)** ERα. Control: Wild-type/*CreERT2*; *Bcl11b*$^{flox/+}$. Note: Images from control panels were the same as that used for Figure 5.8B.

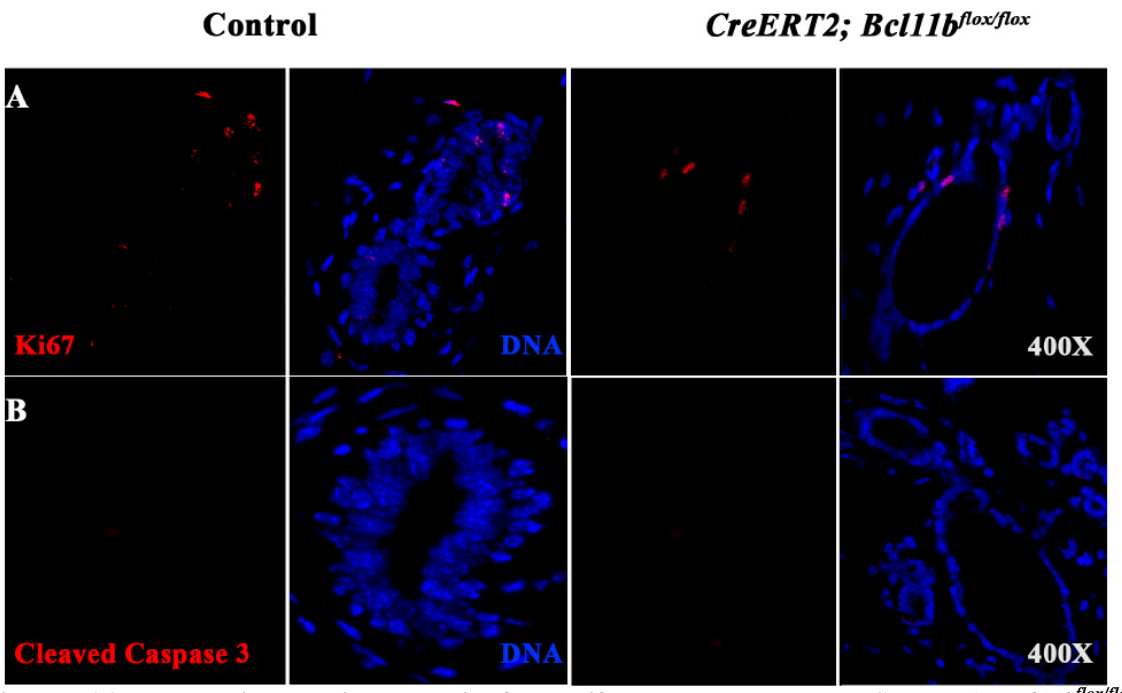


Figure 5.16. Immunohistochemical analysis of tamoxifen treated control and *Cre-ERT2; Bcl11b*$^{flox/flox}$ mammary glands sections using additional markers. Immunostaining of sections using antibodies against **(A)** Ki67 and **(B)** Cleaved Caspase 3. Control: Wild-type/*CreERT2*; *Bcl11b*$^{flox/+}$. Note: Images from control panels were the same as that used for Figure 5.7.

5.2.7 Basal fractions of *Bcl11b*-deficient virgin glands express luminal markers

To determine if loss of *Bcl11b* altered the proportion of luminal and basal cells within the mammary epithelium, mammary epithelial cells were stained using the antibodies to CD24 and CD49f and analyzed by flow cytometry. There was an increase in the total percentage of epithelial cells ($CD24^+$) in the *Bcl11b*-deficient mammary glands (22.5 $\pm$ 5.48% compared to 15.65 $\pm$ 0.35% in the control glands; n=3). Interestingly, the *Bcl11b*-deficient glands (n=3) showed an increase in the percentage of $CD24^+CD49f^{hi}$ (basal) cells (15.3 $\pm$ 5.4% compared to 5.7 $\pm$ 1.8% in the control glands), accompanied with a drastic reduction in the percentage of $CD24^{hi}CD49f^+$ (luminal) cells (1.8 $\pm$ 0.5% compared to 5.0 $\pm$ 1.6% in the control glands) (Figure 5.17). The expansion of basal fraction was unexpected as phenotypic analyses showed that loss of *Bcl11b* resulted in luminal differentiation and development of a pregnancy-like mammary gland (Figure 5.13); hence an increase in the luminal fraction was expected. In order to examine the identity of the basal population of *Bcl11b*-deficient glands, I sorted this basal fraction ($CD24^+CD49f^{hi}$) and performed RT-PCR to determine expression of various lineage markers. Strikingly, I found that these basal cells expressed typical luminal markers such as *NKCC1* along with milk protein genes *α-casein* and *β-casein*, in addition to the basal genes *CK14*, *SMA* and *p63* (Figure 5.18). Moreover, *Gata-3*, *Notch1* and *Notch3* which were normally only expressed in luminal cells, were now detected in the basal fractions of *Bcl11b*–deficient mammary glands (Figure 5.18). Since *Bcl11b* was predominantly expressed in basal cells, these results demonstrated that these *Bcl11b*–deficient basal cells had acquired luminal properties and that the CD24 and CD49f flow cytometric profile no longer faithfully represented basal or luminal cells under this particular condition. The switch from basal cells to luminal alveolar cells was not complete since these cells did not express *Elf5*, a key gene in specifying alveolar cell fate, nor did they have increased *Bcl11a* expression (Figure 5.18). Taken together, these results suggest that *Bcl11b* maintains basal identity and suppresses luminal cell fate.

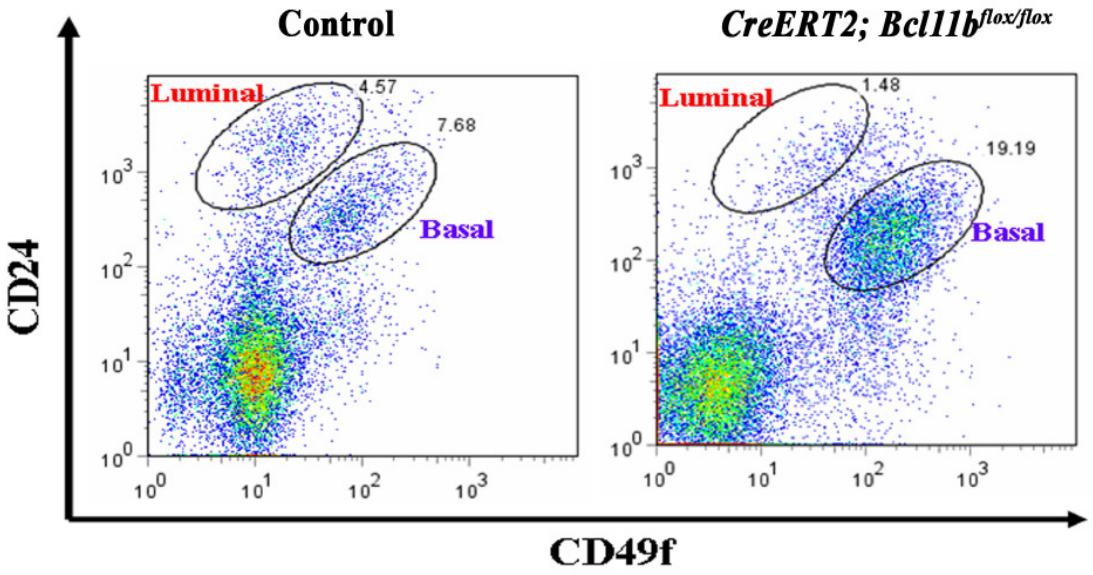


Figure 5.17. Mammary FACS profile of tamoxifen treated control and *Cre-ERT2; Bcl11b*$^{flox/flox}$ mammary glands. Flow cytometric analysis of isolated epithelial cells using CD24 and CD49f markers. There is an increase in the percentage of basal cells ($CD24^{+}CD49f^{hi}$) in the *Bcl11b*-deficent glands compared to the control glands (n=3). Control: Wild-type/*CreERT2*; *Bcl11b*$^{flox/+}$.

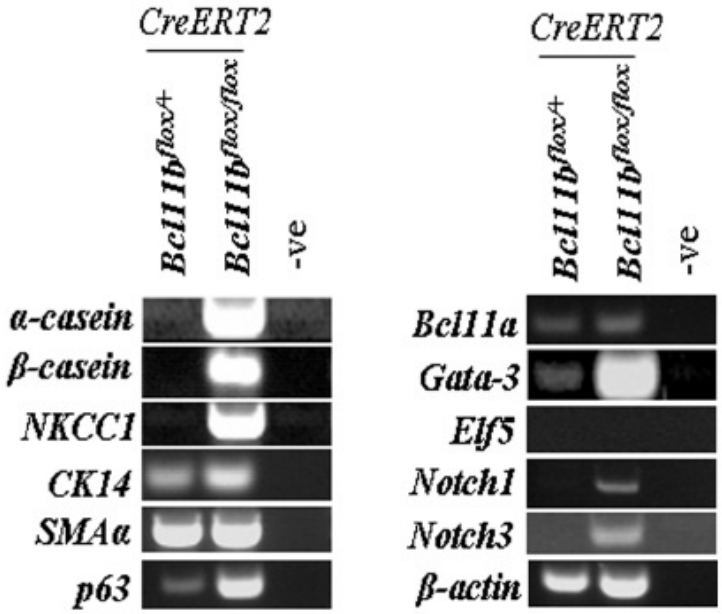


Figure 5.18. Analysis of sorted basal cells from tamoxifen treated control and *Cre-ERT2; Bcl11b*$^{flox/flox}$ mammary glands using semi-quantitative RT-PCR. The $CD24^{+}CD49f^{hi}$ basal cells are sorted out and RT-PCR shows the expression of luminal markers within this fraction of the *Bcl11b*-deficient glands. In addition, Notch receptors which are normally expressed in luminal cells are also detected in the basal fraction. β-actin is used as a control. Control: Wild-type/*CreERT2*; *Bcl11b*$^{flox/+}$. –ve indicates no template control.

5.2.8 *Bcl11b* promotes basal identity in mammary epithelial cells

Expression of *Bcl11b* was detected predominantly in basal cells and in a small number of $Sca1^-$ luminal progenitors in the virgin glands. The results from the loss-of-function studies demonstrated that *Bcl11b* maintains basal identity and suppresses luminal identity and differentiation. Deletion of *Bcl11b* led to the abnormal proliferation and/or differentiation, resulting in a "pregnancy-like" phenotype. To determine whether *Bcl11b* promotes the basal lineage, I over-expressed *Bcl11b* using the over-expression vectors (as described in Appendix A.7) in the murine mammary epithelial cell line, KIM2 (maintained by Dr Walid Khaled) (Gordon et al., 2000). As shown in Figure 5.19, transient over-expression of *Bcl11b* in KIM2 resulted in up-regulation of basal markers such as *CK14*, *p63* and *SMA* but not luminal markers *Gata-3* and *Elf5*, indicating that expression of *Bcl11b* alone induced expression of basal markers in epithelial cells.

In collaboration with Dr Walid Khaled (who performed the experiments), we analyzed the *Stat6*$^{-/-}$ knockout mice that have a delay in alveologenesis during early gestation (Khaled et al., 2007). An increase in the levels of *Bcl11b* in the luminal ($CD24^{hi}CD49f^+$) cells of the virgin and day 5 gestation glands from *Stat6*$^{-/-}$ knockout mice compared to wild-type controls was observed (Figure 5.20A). We quantified the levels of *Bcl11* genes in the luminal ($CD24^{hi}CD49f^+$) and basal ($CD24^+CD49f^{hi}$) fractions using qRT-PCR and found that there was a 74-fold increase in *Bcl11b* expression in the luminal cells of the $Stat6^{-/-}$ virgin mice compared to wild-type controls (Figure 5.20B). When the mammary colony-forming cell (Ma-CFC) capabilities of the luminal progenitors were assessed and compared to the wild-type controls, a significant reduction in Ma-CFC capabilities of *Stat6*$^{-/-}$ luminal progenitors was observed (Figure 5.20C). These data indicated that the increased expression of *Bcl11b* was associated with the decrease in Ma-CFC capabilities of luminal progenitors and the alveologenesis defects in *Stat6*$^{-/-}$ mammary glands. Taken together, these data demonstrate that *Bcl11b* maintains basal cell fate and acts as a suppressor to luminal differentiation and/or luminal cell fate.

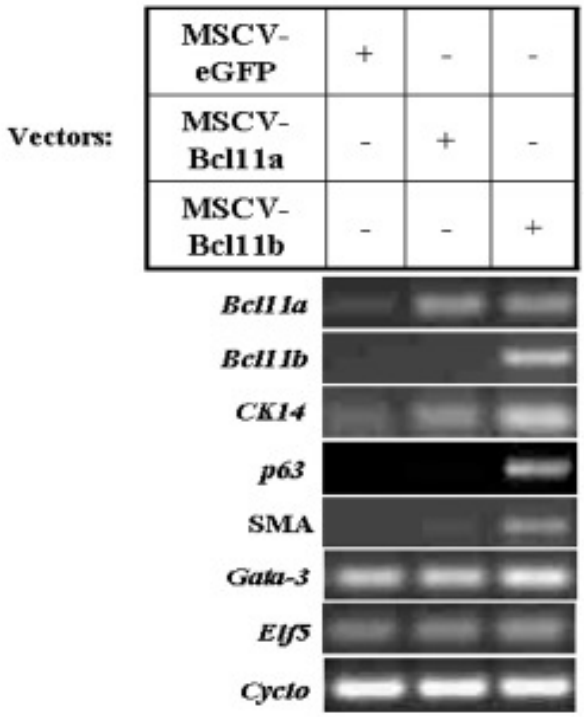


Figure 5.19. Over-expression of *Bcl11* genes in mammary epithelial cell line, KIM2. Over-expression of *Bcl11b* in KIM2 cells results in up-regulation of basal markers (*p63*, *CK14* and *SMA*) as determined by RT-PCR. In contrast, expression of *Gata-3* and *Elf5* remains unchanged. No changes in the expression of the examined markers are observed in KIM2 cells transfected with MSCV-Bcl11a. MSCV-eGFP vector is used as the transfection control. + indicates vector used for each transfection; - indicates vector not used for transfection. Cyclophilin A is used as a control for RT-PCR.

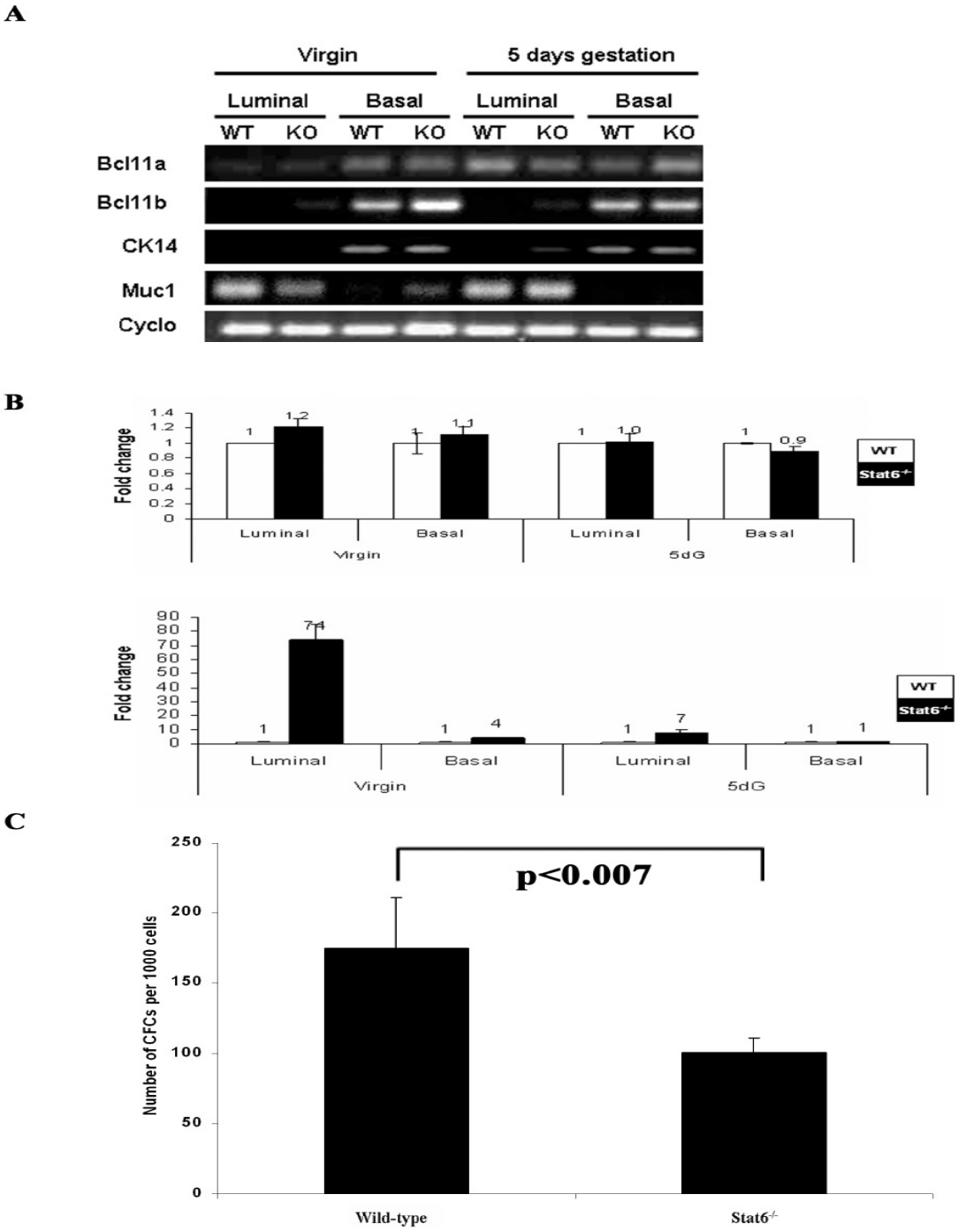


Figure 5.20. Analysis of mammary glands from *Stat6*$^{-/-}$ virgin and day 5 gestation female mice. (A) Semi-quantitative RT-PCR analysis of the luminal ($CD24^{hi}CD49f^{+}$) and basal ($CD24^{+}CD49f^{hi}$) fractions of wild-type and *Stat6*$^{-/-}$ mammary glands. Up-regulation of *Bcl11b* expression is observed in the luminal fractions of *Stat6*$^{-/-}$ mammary glands. **(B)** Quantitative real time PCR analysis of *Bcl11a* and *Bcl11b* expression in sorted luminal ($CD24^{hi}CD49f^{+}$) and basal ($CD24^{+}CD49f^{hi}$) cells of *Stat6*$^{-/-}$ and wild-type mammary glands. A 74-fold increase in *Bcl11b* expression is observed in luminal cells of the *Stat6*$^{-/-}$ virgin gland. **(C)** Mammary colony-forming cell (Ma-CFC) assay of sorted $CD24^{hi}CD49b^{+}$ cells (luminal progenitors) from either wild-type or *Stat6*$^{-/-}$ virgin glands. A 40% reduction in Ma-CFCs (per 1000 cells) is observed in the *Stat6*$^{-/-}$ virgin glands. Error bars denote standard deviation (n=3) and $p<0.007$. WT: Wild-type; KO: *Stat6*$^{-/-}$.

5.3 Discussion

In this Chapter, I demonstrated that *Bcl11a* and *Bcl11b* play important roles in development of the mammary gland. In the embryo, loss of *Bcl11a* resulted in defective mammary buds formation and failure of mammary bud regression in a male embryo. In contrast, loss of *Bcl11b* caused the absence of the third pair of mammary buds and defects in other mammary buds. As *Bcl11* homozygous pups died within 24 hours of birth, conditional knockout alleles of both genes were used to study their roles in postnatal mammary gland development. Condition deletion of *Bcl11a* in the virgin glands disrupted normal mammary bi-layer architecture and resulted in a general loss of mammary epithelial cells. In addition, there was also a loss of $Gata3^+$ cells and a relative increase in the number of $ER\alpha^+$ epithelial cells in the *Bcl11a*-deficient virgin glands. Dysregulation in Notch signalling was also observed in the virgin glands following deletion of *Bcl11a*. Loss of *Bcl11b* on the other hand resulted in precocious alveologenesis and lineage alteration in the virgin glands. Gain-of-function studies also indicated the essential role of *Bcl11b* in promoting basal lineage. Increased levels of *Bcl11b* in luminal progenitors were associated with a dramatic decrease in Ma-CFCs capabilities. Over-expression of *Bcl11b* in mammary epithelial cells resulted in up-regulation of basal markers. Taken together, these results demonstrate the essential roles of *Bcl11* genes in normal mammopoiesis and lineage commitment, suggesting that the interplay between the two genes is critical for the regulation of mammary epithelial cell fate.

5.3.1 *Bcl11* genes are important for embryonic mammary development

The embryonic mammary gland develops via extensive epithelial-mesenchymal interactions, similar to other skin appendages such as the hair follicles and teeth (Mikkola and Millar, 2006). Mammary bud formation was detected in the $Bcl11a^{lacZ/lacZ}$ homozygous mutant embryos using the antisense *Fgfr2TK* probe. No missing buds were observed in all the mutants analysed. This was consistent with the expression of *Bcl11a* that was only detected in the mammary anlage from 13.5 dpc onwards. This indicated that loss of *Bcl11a* did not affect the inductive signals required for the initial formation of

the mammary placodes. Further analyses using other *in situ* markers showed that the mammary buds formed in the *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant 13.5-14 dpc embryos had a flattened morphology whereas those in the wild-type embryos were visible as elevated knob-liked structures. Expression of *Bcl11a* was detected in both the mammary epithelium and the surrounding mesenchyme. Epithelial-mesenchymal interactions are important for the proper formation of the mammary anlage. The mesenchyme may at least have a permissive function for the induction of mammogenesis (Veltmaat et al., 2003). Mesenchyme from the mammary region of a 13 dpc mouse embryo induces functional mammary epithelium in rat midventral or dorsal epidermis (Cunha and Hom, 1996). Furthermore, the mesenchyme still retains its inductive capacity for placode formation even after the stage of placode induction. Thus the mammary mesenchyme also provides an instructive signal for mammary placode formation and suggests that placode formation is not an intrinsic property of the epithelium of the mammary region. Therefore, the phenotype observed in *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant 13.5-14 dpc embryos might be a result of defective epithelial-mesenchymal interactions and/or signalling, though further studies are needed to ascertain this.

Consistent with a previous study which showed that by 14.5-15.5 dpc, *Lef1* expression becomes undetectable in the epidermis destined to become nipple sheath (Foley et al., 2001), expression of *Lef1* was not detected in all the buds with the exception of the fifth pair of the wild-type 14.5 dpc embryos. In contrast, expression of *Lef1* was observed in all the mammary epidermis of *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant embryos at 14.5 dpc. Furthermore, presence of rudimentary mammary buds in a *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant male embryo at 16.5 dpc was detected using *in situ* hybridization with antisense *Fgfr2TK* probe, suggesting that loss of *Bcl11a* resulted in failure of regression of mammary buds. Whether the continued *Lef1* expression in the mammary epidermis of *Bcl11a*$^{lacZ/lacZ}$ homozygous mutant 14.5 dpc embryos (which might serve as a survival signal), and the failure of regression of mammary buds in the male embryo are causatively linked remains to be determined (Figure 5.21). Determination of the proliferative and apoptotic status of mammary epithelial cells of male embryos between 14.5-16.5 dpc should provide an answer.

Mammary gland development begins at approximately 10.5 dpc with the appearance of milk lines that appear in response to mesenchymal signals (Mailleux et al., 2002; Veltmaat et al., 2004). Interestingly, *Bcl11b* expression was also detected around 10.5 dpc along stripes of ventral lateral surface ectoderm between the fore- and hindlimbs and the expression became localized specifically within the mammary buds from 12.5 dpc. Therefore *Bcl11b* is one of the earliest known mammary lineage markers. Mammary bud formation occurs asynchronously, with the third pair of mammary bud being the first to appear (Mailleux et al., 2002). Studies have also shown that different signals are required for the induction of each pair of placodes (Eblaghie et al., 2004; Mailleux et al., 2002; van Genderen et al., 1994). The importance of *Bcl11b* in embryonic development was underscored by the absence of the third pair of mammary buds in the *Bcl11b* homozygous mutant embryos as indicated by the absence of *Wnt10b* and *Lef1* expression in the presumptive region of the third bud and the reduced *Wnt10b* and *Lef1* expression in other buds at 12.5 dpc. Further examinations of $Bcl11b^{lacZ/lacZ}$ homozygous mutant 14.5 dpc embryos showed that expression of *Fgfr2TK* was reduced in other buds and completely absent in the presumptive third bud region. Hence *Bcl11b* is a specification signal for formation of the third pair of mammary buds. This result also confirms that formation of the third pair of mammary buds is independent of the other buds and requires a different inductive signal.

The genetic hierarchy of *Fgf10* and *Tbx3* remains to be completely elucidated, but the severity of phenotypes in the mutant embryos of either gene where most mammary buds are missing suggests that both of these genes probably lie upstream of *Bcl11b*. Expression of *Fgf10* and *Tbx3*, markers of the mesodermal mammary line, appeared normal in $Bcl11b^{lacZ/lacZ}$ homozygous mutant 10.5 dpc embryos, suggesting that specification of the mammary line itself was not defective. In contrast, expression of *Wnt10b* and *Lef1* was altered in $Bcl11b^{lacZ/lacZ}$ homozygous mutant 12.5 dpc embryos. Hence, it is likely that *Bcl11b* functions upstream of these genes and downstream of *Fgf10* and/or *Tbx3* to relay signals to the precursor mammary epithelial cells (which appear to express *Wnt10b*) (Veltmaat et al., 2004) in order to induce the formation of the third pair of mammary buds (Figure 5.21). Alternatively, *Bcl11b* might also function in a

separate pathway to induce formation of the third pair of mammary buds. Thus, absence of *Bcl11b* resulted in aberrant placode formation off the mammary line.

Loss of the third pair of mammary buds was also observed in the mouse scaramanga (*ska*) mutation (Howard et al., 2005). This mutation has been mapped to the Neuregulin3 (*Nrg3*) gene. In addition, supernumerary glands were formed at high frequency in *ska* mice. *Nrg3* belongs to the neuregulin/heregulin cytokine family and is a direct ligand for the Erbb4 tyrosine kinase receptor. Both *Nrg3* and *Erbb4* are expressed in the presumptive mammary region at 10.75 dpc (Howard et al., 2005). The binding of Nrg3 to Erbb4 results in tyrosine phosphorylation and activation of the receptor leading to activation of downstream signalling cascade. It would be interesting to determine if *Bcl11b* is a target of Nrg3 induced Erbb4 activation in the presumptive third bud region which would explain the observed phenotype. Future experiments can be carried out to determine changes in *Nrg3* expression in $Bcl11b^{lacZ/lacZ}$ homozygous mutant embryos as well as changes in *Bcl11b* expression in *ska* embryos to establish the genetic hierarchy of *Bcl11b* and *Nrg3* in induction of the third pair of mammary buds (Figure 5.21).

In summary, both *Bcl11a* and *Bcl11b* are essential for normal embryonic mammary development.

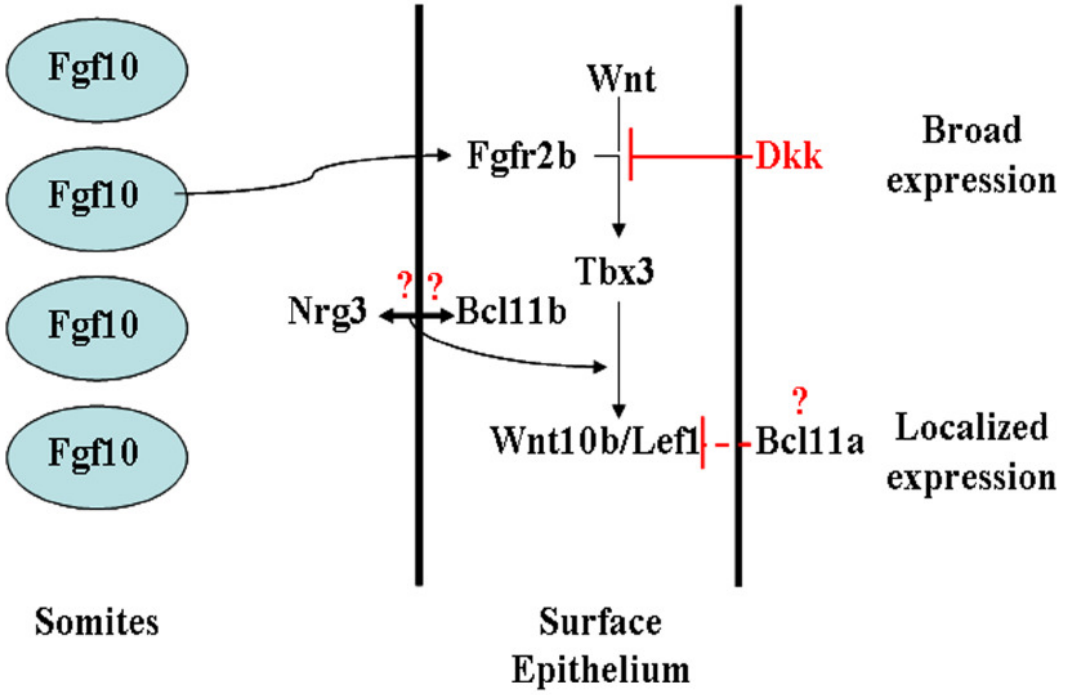


Figure 5.21. Model of possible regulatory factor interactions in the mammary gland. Proposed interactions of *Bcl11* genes within known regulatory networks in embryonic mammary development.

5.3.2 *Bcl11* genes play essential roles in the virgin mammary gland

In the virgin gland, loss of *Bcl11a* caused profound defects in the ductal structure where the mammary architecture was severely disrupted and the luminal and basal cells appeared to form a thin cellular layer in many areas of the ducts. This structural defect was associated with loss of *Gata-3* and *Elf5* expression. Gata-3 and Elf5 are considered to be the key players in luminal specification and alveologenesis. Loss of expression of both genes clearly shows a critical role of Bcl11a in the luminal lineage. Moreover, *Bcl11a* deletion resulted in reduction in expression of several luminal and basal markers, and the *Bcl11a*-deficient epithelium had a lineage shift towards $CD24^{hi}CD49f^{+}$ region in flow cytometric profile, normally considered to be the luminal fraction in the wild-type gland. Relative increase of luminal cells was unexpected since Gata-3 is considered to be the crucial factor in promoting luminal differentiation. It is also surprising that there was a relative increase of the $ER\alpha^{+}$ cells even though Gata-3 expression was lost. Gata-3-expressing cells are mainly $ER\alpha^{+}$ and Gata-3 deletion leads to loss of differentiated cells including the $ER\alpha^{+}$ cells (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). This apparent discrepancy is likely to be caused by multiple roles of Bcl11a in several epithelial compartments, which requires further studies. In addition, Bcl11a and Gata-3 are likely to function in distinct epithelial populations or they may act through different pathways. *Bcl11a* is expressed and necessary in the $Sca1^{-}$ luminal progenitors, and is also likely to be expressed in the alveolar progenitors together with *Elf5*. Loss of *Bcl11a* would thus lead to depletion of these progenitors, which are likely to be $ER\alpha^{-}$. Consequently, there would be accumulation of $ER\alpha^{+}$ luminal cell. Alternatively, Bcl11a may play a role in the proposed asymmetric division of $ER\alpha^{-}$ mammary stem cells to produce the $ER\alpha^{+}$ stem/progenitor cells (Asselin-Labat et al., 2006; Booth and Smith, 2006; Dontu et al., 2004; LaMarca and Rosen, 2007; Sleeman et al., 2007). In addition, loss of *Bcl11a* also resulted in decrease in expression of basal markers, suggesting that *Bcl11a* also plays a role in the basal lineage.

Our lab has shown previously that Notch signalling is highly expressed in the *Bcl11a* deficient T-cell leukemia (Liu et al., 2003b). *Bcl11b*, in co-operation with *Notch1* and *Gata3*, is proposed to be the key gene in T cell progenitor commitment to T cell lineage (Rothenberg, 2007a, b), although the definitive evidence is still lacking. In the

virgin mammary glands, loss of *Bcl11a* resulted in increased Notch1 expression. Although *Bcl11a* was expressed primarily in the putative luminal progenitor cells ($CD49b^{+}$), once *Bcl11a* was deleted in the virgin gland, most of the luminal cells appeared to be positive for Notch1, indicating that these luminal cells were likely to be the descendents of the *Bcl11a*-expressing luminal progenitors. Alternatively, non-cell autonomous effects of Notch signalling resulting in the observed phenotypes were also possible. Notch signalling is known to be involved in lineage specification and cell fate determination (Buono et al., 2006; Rothenberg, 2007a). Further studies are needed to investigate the roles of Notch signalling in the mammary gland. Notably, up-regulation of *Bcl11b* was also observed in the *Bcl11a*–deficient mammary glands. The increased level of *Bcl11b* could be partially responsible for the phenotypes in the *Bcl11a* mutant glands. The increased levels of *Bcl11b* could also suggest that there might be a luminal to basal lineage switch as *Bcl11b* was shown to promote basal lineage in this study. Alternatively, the increased levels of *Bcl11b* could be partly responsible for the loss of luminal cells. These possibilities could be investigated using *Bcl11a* and/or *Bcl11b*-over-expressing transgenic mice. Taken together, these results clearly demonstrate that *Bcl11a* is a critical regulator in normal mammary development.

Loss of *Bcl11b* in the virgin gland resulted in precocious alveologenesis and the development of a 'pregnancy-like' state. Histological examination showed the alveolar-like structures developed along primary ducts, and the functional differentiation of luminal cells was confirmed by positive β-casein staining. One remarkable phenotype was that *Bcl11b*-deficient mammary epithelial cells were found within the $CD24^{+}CD49f^{hi}$ (basal) region. It has been observed that differentiated luminal cells from the pregnant glands are fragile and many do not survive the cell dissociation protocol, and are preferentially killed off (Personal communication with John Stingl). This would cause a relative increase in the percentage of $CD24^{+}CD49f^{hi}$ basal cells, resulting in the disparity in the mammary FACS profile observed in the *Bcl11b*-deficient glands. However, expression of luminal (*NKCC1*, *Gata-3*) and milk protein (*α-casein*, *β-casein*) markers was observed in the sorted $CD24^{+}CD49f^{hi}$ (basal) cells of the *Bcl11b*-deficient mammary glands, indicating that these basal cells had acquired luminal properties. Since *Bcl11b* was predominantly expressed in the basal cells, therefore loss of *Bcl11b* in the basal cells

resulted in a lineage switch of some of the basal cells, suggesting that *Bcl11b* normally acts as suppressor to the luminal lineage. These results indicate that *Bcl11b* normally acts as a suppressor to luminal differentiation and blocks precocious alveolar development. Loss of *Bcl11b* would release this brake on luminal differentiation and result in abnormal proliferation and differentiation of luminal cells in the absence of pregnancy hormones. Interestingly, development of these alveolar-like cells was Elf5-independent and did not require increased *Bcl11a* expression, further highlighting the critical role of Bcl11b in maintaining the basal identity. Notably, *Bcl11b* is also expressed in some Sca1$^-$ luminal progenitors. Deletion of *Bcl11b* might also allow the abnormal proliferation and differentiation of these progenitors and eventually contributing to the pregnancy-like phenotype observed. Transplantation of the *Bcl11b*-expressing Sca1$^-$ luminal progenitors into cleared fat-pads and treating the mice with tamoxifen to delete *Bcl11b* will help to clarify this possibility.

Bcl11b not only maintains the basal identity but also promotes the basal lineage. Transient over-expression of *Bcl11b* was sufficient to induce expression of basal cell specific genes such as *CK14*, *p63* and *SMA* in KIM2 mammary epithelial cells (Figure 5.19). In addition, increased levels of *Bcl11b* expression in the luminal progenitors of *Stat6*$^{-/-}$ knockout mice were associated with a dramatic reduction in Ma-CFC capabilities of luminal progenitors (Figure 5.20). In summary, the loss-of-function and gain-of-function analyses clearly demonstrate that *Bcl11b* promotes and maintains the basal lineage and at the same time suppresses luminal cell fate and blocks precocious alveolar development in the virgin mammary gland. Thus the levels of *Bcl11b* are critical to regulate proliferation and differentiation of luminal cells.

The phenotypes following the loss of *Bcl11* genes observed in this study could have arisen due to systemic effects or local stromal effects. This is unlikely as the injected mice were kept within the same cage to minimise the effects of individual female being at a different stage of the estrus cycle. However, to further eliminate any stromal or systemic effects, *Bcl11*-deficient epithelial cells could be transplanted into cleared wild-type mammary fat pads to confirm epithelial contributions to the phenotypes.

5.3.3 Putative roles of *Bcl11* genes in mammary cell fate determination and lineage commitment

Specification, proliferation, differentiation, death and survival of the mammary cells are regulated by networks of information mediated by external signals such as steroids, simple peptide hormones, and by intrinsic cellular factors of various signal pathways(Hennighausen and Robinson, 2005; Watson and Khaled, 2008). Several transcription factors that have been implicated in hematopoietic lineage specification have now been shown to play essential roles in mammary epithelial cell fate decisions (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006). I have now shown that *Bcl11a* and *Bcl11b*, two transcription factors essential in specifying B and T lineages respectively, have critical roles in the mammary epithelium.

The unique expression patterns, phenotypic analysis, and alterations in molecular pathways, have established the roles of both genes at various stages of mammary development (Figure 5.22). *Bcl11a* and *Bcl11b* could be expressed in the multi-potent mammary progenitors since they are among the earliest genes that are specifically expressed in the mammary gland. *Bcl11a* is also expressed in TEBs, in a small number of basal cells of the virgin gland, and deletion of *Bcl11a* results in severe defects in both luminal and basal lineages. From this multi-potent progenitor pool, Bcl11a, working together with Gata-3, promotes the luminal lineage. Although both Bcl11a and Gata-3 are expressed and functional in the luminal progenitors, they have distinct roles. Bcl11a maintains luminal progenitors since *Bcl11a* deletion results in loss of Gata-3 expression and selective depletion of $Sca1^-$ luminal progenitors *in vitro* and $ER\alpha^-$ cells *in vivo*, while Gata-3 primarily promotes progenitor differentiation. *Bcl11a* deletion leads to loss of *Elf5* expression in the virgin gland and reduction of Ma-CFCs. On the other hand, reduction of *Elf5* expression results in accumulation of luminal progenitors, and forced expression of *Elf5* causes reduction of Ma-CFCs. Therefore, it appears that Bcl11a maintains the alveolar progenitors while Elf5 promotes differentiation of this progenitor pool. Similar to *Bcl11a*; *Bcl11b* is also likely to be expressed in the multi-potent progenitors as well. From this progenitor pool, Bcl11b promotes and maintains the basal lineage, and prevents excessive luminal commitment and abnormal differentiation in the absence of pregnancy hormones.

In summary, I have shown that *Bcl11a* and *Bcl11b* are essential for the development, lineage specification and maintenance of the mammary epithelium. In Chapter 6, I will discuss the roles of *Bcl11* genes in maintenance of terminally differentiated luminal secretory cell fate using *Bcl11*$^{flox/flox}$ mice that have been bred to the BLG-Cre mice where Cre recombinase is expressed specifically during lactation (Selbert et al., 1998).

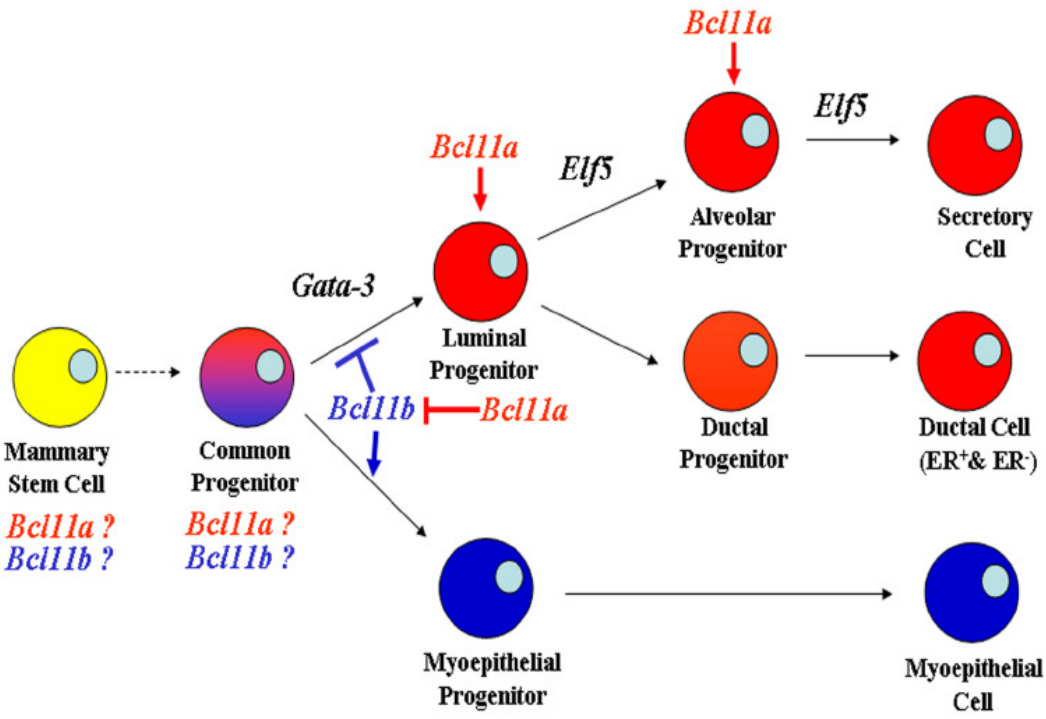


Figure 5.22. Proposed working model of the roles of *Bcl11* genes in mammary lineages. Schematic representation of a model for the roles of *Bcl11* genes in determination of mammary epithelial cell fate determination in relationship to the proposed epithelial hierarchy in the mammary gland.

CHAPTER 6:
BCL11A IS ESSENTIAL FOR MAINTENANCE OF LUMINAL SECRETORY CELL FATE

6.1 Introduction

6.1.1 Mammary gland development during pregnancy

The mammary gland undergoes extensive proliferation and attains terminal differentiation only after pregnancy. The functional differentiation state is suggested to be defined predominantly by the hormonal status of the animal (Hennighausen and Robinson, 1998). Embryonic development in the mice sets up a rudimentary network of ducts which undergoes extensive growth and arborisation during puberty, stimulated by hormones such as estrogen and progesterone. This generates the mature mammary ductal network that will ultimately serve as channels for milk transport during lactation. These ducts are lined by a single layer of luminal epithelial cells and are surrounded by a sleeve of myoepithelial cells. During pregnancy, additional ductal branching occurs and extensive lobulo-alveolar proliferation which eventually fills the fat pad completely at parturition. Cell division occurs in both the ductal and alveolar cell populations throughout the duration of gestation period and persists through the early phase of lactation. In addition, the luminal alveolar cells undergo functional terminal differentiation to attain a secretory phenotype that produces and secretes milk into the lumen of ducts and the contractile forces generated by the myoepithelium enable the transport of milk to the nipple to feed the newborn pups.

Systemic hormones such as estrogen, prolactin and progesterone play a key role in the regulation of the mammary gland during pregnancy and lactation. Estrogen is required for ductal elongation and arborisation during postnatal mammary development and lobulo-alveolar development during pregnancy (Mallepell et al., 2006). Similarly, progesterone is also required for side branching and alveolar differentiation in the mice (Lydon et al., 1995). During pregnancy, additional systemic local and intracellular signals are required for alveolar proliferation and differentiation. Prolactin signalling is essential for the proliferation and functional differentiation of lobulo-alveolar structures during

pregnancy (Horseman et al., 1997; Liu et al., 1997; Ormandy et al., 1997; Topper and Freeman, 1980). In addition, colony-stimulating factor-1 (Pollard and Hennighausen, 1994) and oxytocin (Nishimori et al., 1996; Young et al., 1996) are also required for alveolar development and milk secretion. Therefore, regulation of mammary gland development during pregnancy is of utmost importance to ensure that it is functionally ready for lactation.

6.1.2 Differentiation of luminal epithelial cells

With the recent publications on roles of transcription factors in determining luminal cell fate in the mammary gland (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006), it is now clear that functional differentiation of the mammary gland is not defined solely by systemic hormones and local growth factors. It is apparent that transcription factors are also key mediators of the differentiation process. Gata-3 is the first transcription factor shown to be a critical determinant of luminal cell fate. Loss-of-function analyses showed that *Gata-3* is essential for differentiation and maintenance of luminal cell fate (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). In addition, luminal differentiation and alveolar morphogenesis were reduced in both *Stat6* and *IL4/IL13* doubly deficient mice during pregnancy, suggesting that Stat6 signalling is also important for luminal differentiation during pregnancy (Khaled et al., 2007). Recently, the prolactin-regulated Ets transcription factor Elf5 was shown to specify alveolar cell fate (Oakes et al., 2008). Even though both *Gata-3* and *Elf5* deficiencies led to accumulation of $CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitor populations, some distinct differences exist. In addition to the accumulation of luminal progenitors, conditional deletion of *Gata-3* also resulted in disruptions to the ductal epithelium in virgin mice. In contrast, the effects of *Elf5* deficiency were restricted to alveolar cell fate specification during pregnancy. It was also observed that *Gata-3* and *Elf5* were expressed in different mammary luminal epithelial cells (Oakes et al., 2008) and the $CD29^{lo}CD24^{+}CD61^{+}$ luminal progenitor population may be a heterogenous population consisting of *Gata-3*- and *Elf5*-responsive progenitors. These studies imply that the combinatorial effects of different transcription factors are required to specify and maintain luminal cell fate and

highlight the complexity of the gene regulatory networks required to establish mammary cell fate.

As discussed in Chapter 5, *Bcl11* genes play important roles in the virgin gland development. Expression of *Bcl11a* was up-regulated during gestation and was detected in luminal ductal cells and differentiated alveolar cells (Figure 4.14). In addition, expression of *Bcl11a* was detected in the terminally differentiated secretory luminal cells during lactation (Figure 4.15). In contrast, expression of *Bcl11b* was detected only in the basal layer of ducts but not in the differentiated luminal alveolar cells (Figure 4.14). These differential expression patterns imply that *Bcl11a* may also have important functions in maintenance of terminally differentiation secretory luminal cell fate. In this Chapter, the main question that I will address is whether *Bcl11* genes are essential for the maintenance of terminally differentiated secretory luminal cell fate. I crossed both $Bcl11a^{flox/flox}$ and $Bcl11b^{flox/flox}$ conditional knockout mice to the BLG-Cre mice where Cre is specifically expressed in the lactation mammary gland (Selbert et al., 1998) and examined the effects of Cre-mediated excision of *Bcl11* genes in the lactation glands.

6.2 Results

6.2.1 *BLG-Cre;* $Bcl11a^{flox/flox}$ females have severe lactational defects

BLG-Cre expression was found to closely reflect the kinetics of BLG expression (Selbert et al., 1998). Low levels of recombination were detected in the virgin glands (7%), with increasing levels observed throughout gestation and parturition (20-30%). I first analyzed mammary glands from *BLG-Cre*; $Bcl11^{flox/flox}$ and control (*BLG-Cre*; $Bcl11^{flox/+}$) females at late gestation (16.5-17.5 dpc). No obvious histological defects were detected in the mammary glands from either *BLG-Cre*; $Bcl11a^{flox/flox}$ (Figure 6.1A2 and A4) or *BLG-Cre*; $Bcl11b^{flox/flox}$ (Figure 6.1B2 and B4) females compared to their respective *BLG-Cre*; $Bcl11^{flox/+}$ heterozygous control glands (Figure 6.1A1 and A3; B1 and B3). Lobulo-alveolar structures were abundant and sections showed that these structures contained lipid droplets and secretory vesicles, suggesting that lobulo-alveolar development was not significantly affected during gestation when low levels of Cre recombinase was expressed.

The highest levels of BLG-Cre expression were detected during lactation and quantification at mid-lactation determined the recombination efficiency to be approximately 70-80% (Selbert et al., 1998). Hence the effects of BLG-Cre mediated deletion of *Bcl11* genes were studied during lactation. I found that *BLG-Cre; Bcl11a*$^{flox/flox}$ females were unable to nurse their pups since no pups were successfully weaned. This result suggests a severe lactational insufficiency or impairment of the *BLG-Cre; Bcl11a*$^{flox/flox}$ females. To confirm this observation, I decided to monitor the weights of the pups from control *BLG-Cre; Bcl11a*$^{flox/+}$ and *BLG-Cre; Bcl11a*$^{flox/flox}$ females from postnatal day 0 (P0). Pups from *BLG-Cre; Bcl11a*$^{flox/flox}$ dams had little or no milk in their stomach from postnatal day 1 (P1) and displayed a significant reduction in weights over the observation period compared to pups from *BLG-Cre; Bcl11a*$^{flox/+}$ females (Figure 6.2). These pups were severely dehydrated with loose skin and were much smaller in size compared to age-matched pups nursed by control *BLG-Cre; Bcl11a*$^{flox/+}$ females. Most of these pups died by postnatal day 7 (P7) if not rescued (Figure 6.3B).

To rule out possible suckling defects of pups contributing to their inability to feed and hence resulting in a reduction in body weights, P1 and postnatal day 6 (P6) pups from *BLG-Cre; Bcl11a*$^{flox/flox}$ females were fostered to the control *BLG-Cre; Bcl11a*$^{flox/+}$ lactation females. The body weights of the fostered pups were monitored over a period of five days. All pups that were fostered to control *BLG-Cre; Bcl11a*$^{flox/+}$ females thrived with milk in their stomach and their weights increased constantly over the period observed (Figure 6.3A and B). This is in contrast to the original observation that the pups that remained with the *BLG-Cre; Bcl11a*$^{flox/flox}$ lactation females are usually dead by P7 (Figure 6.3A and B). These observations demonstrated that the pups from *BLG-Cre; Bcl11a*$^{flox/flox}$ females suckled normally. Hence the reduction in body weights of the pups were not due to suckling defects of the pups. Therefore loss of *Bcl11a* led to lactational deficiency in the *BLG-Cre; Bcl11a*$^{flox/flox}$ females, resulting in the lactating females being unable to feed their pups, suggesting that *Bcl11a* is critical to the maintenance of lactation state of the mammary glands.

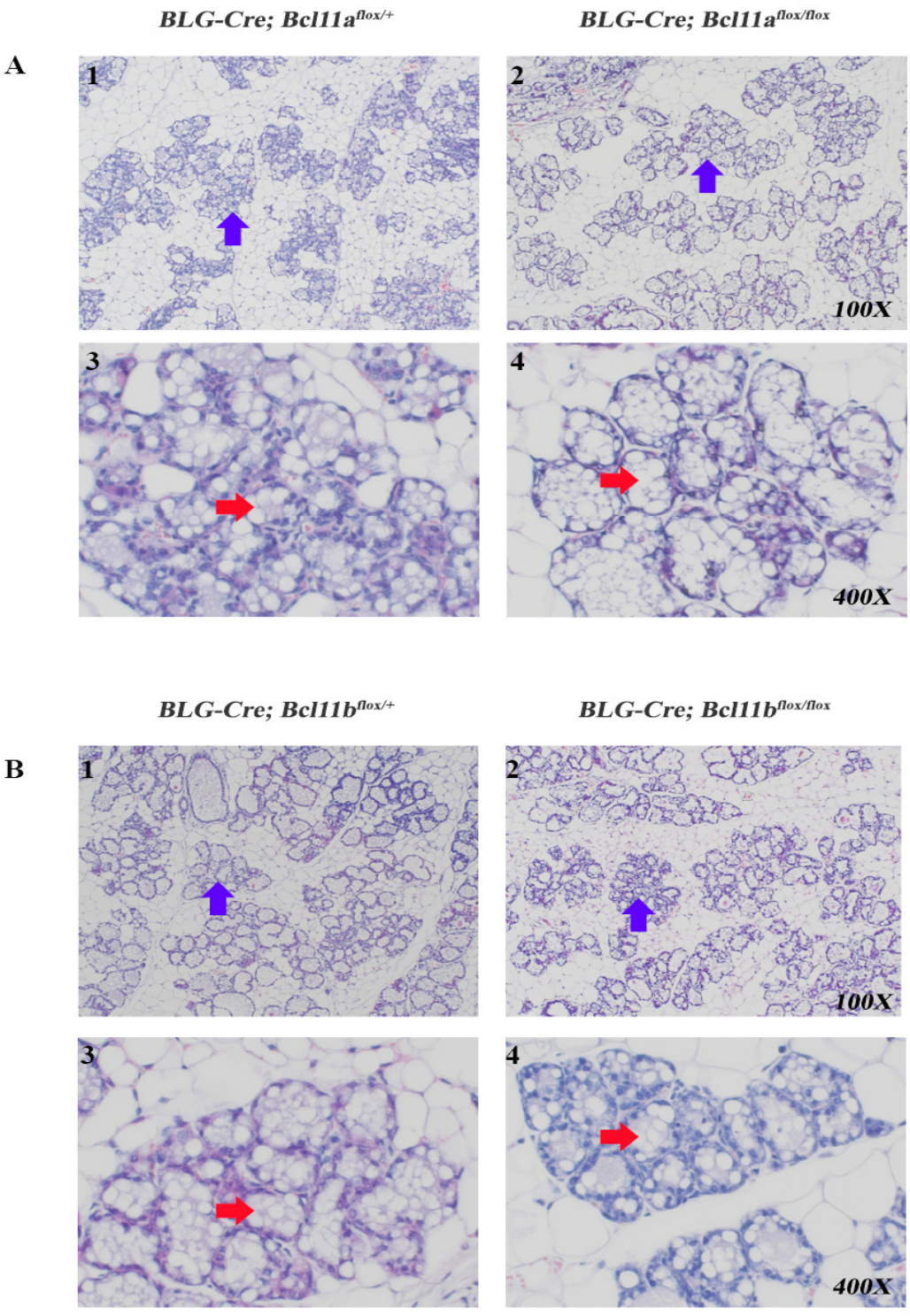


Figure 6.1. Histological analysis of mammary glands from *BLG-Cre; Bcl11*$^{flox/flox}$ and control females. H & E stained sections of day 16.5-17.5 gestation mammary glands from **(A1 and A3)** *BLG-Cre*; *Bcl11a*$^{flox/+}$; **(A2 and A4)** *BLG-Cre*; *Bcl11a*$^{flox/flox}$; **(B1 and B3)** *BLG-Cre*; *Bcl11b*$^{flox/+}$ and **(B2 and B4)** *BLG-Cre*; *Bcl11b*$^{flox/flox}$ females. Panels **A3**, **A4**, **B3** and **B4** are higher magnifications of panels **A1**, **A2**, **B1** and **B2** respectively. Blue arrows indicate lobulo-alveolar structures. Red arrows indicate fatty droplets.

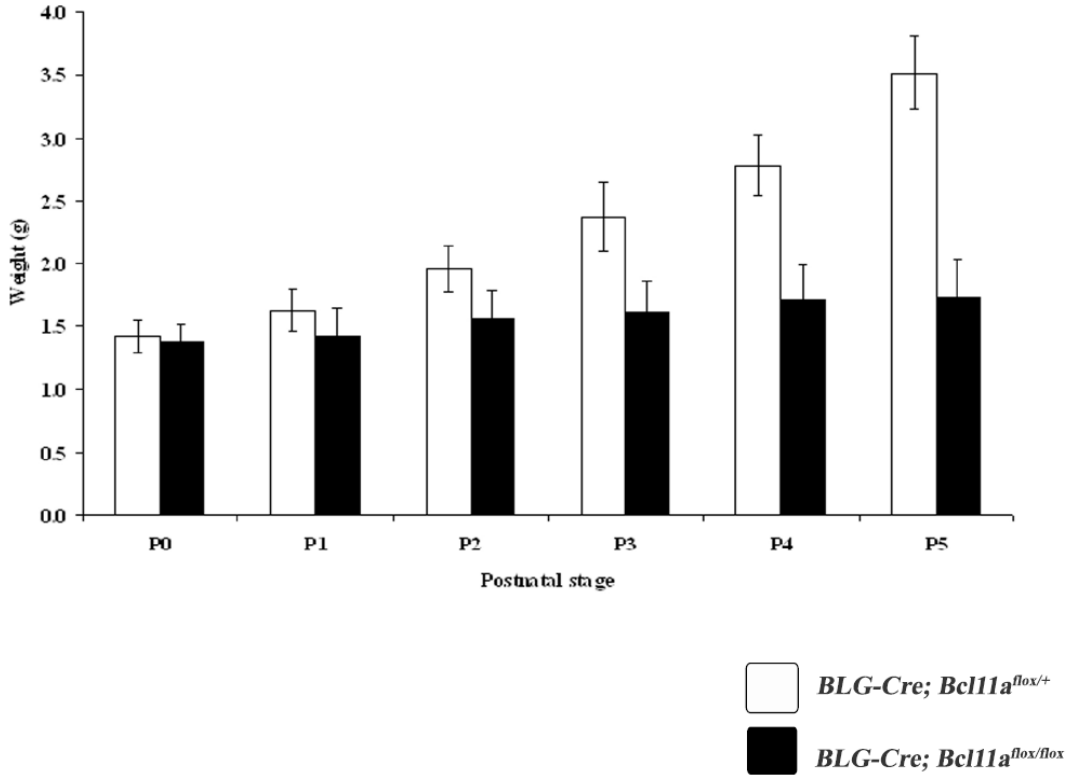
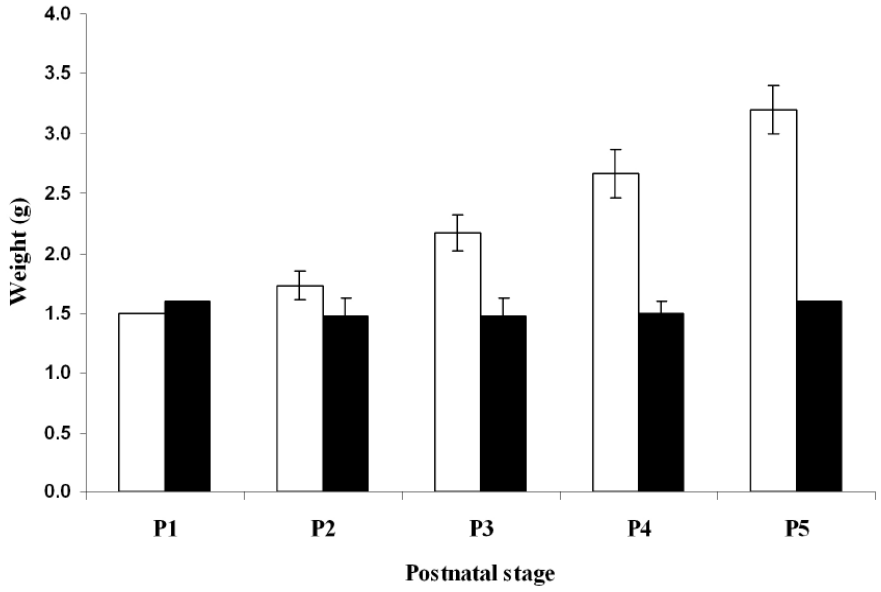
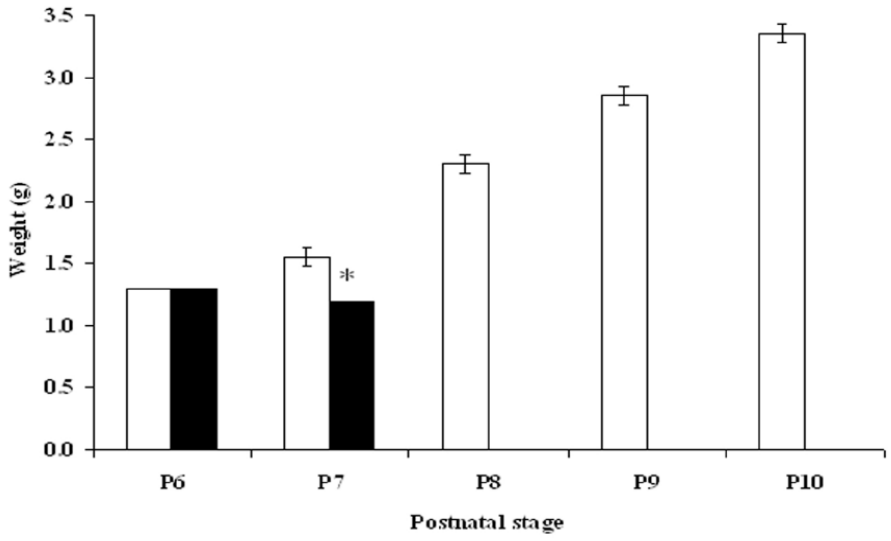


Figure 6.2. Graph showing average weights of pups from *BLG-Cre; Bcl11a*$^{flox/+}$ and *BLG-Cre; Bcl11a*$^{flox/flox}$ females. White and black bars represent average weights of pups from *BLG-Cre; Bcl11a*$^{flox/+}$ and *BLG-Cre; Bcl11a*$^{flox/flox}$ females respectively. Errors bar denote standard deviation obtained from analyzing 71 and 45 pups of *BLG-Cre; Bcl11a*$^{flox/+}$ and *BLG-Cre; Bcl11a*$^{flox/flox}$ females respectively.

A



B



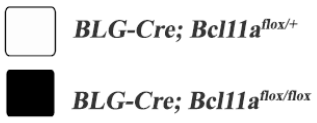
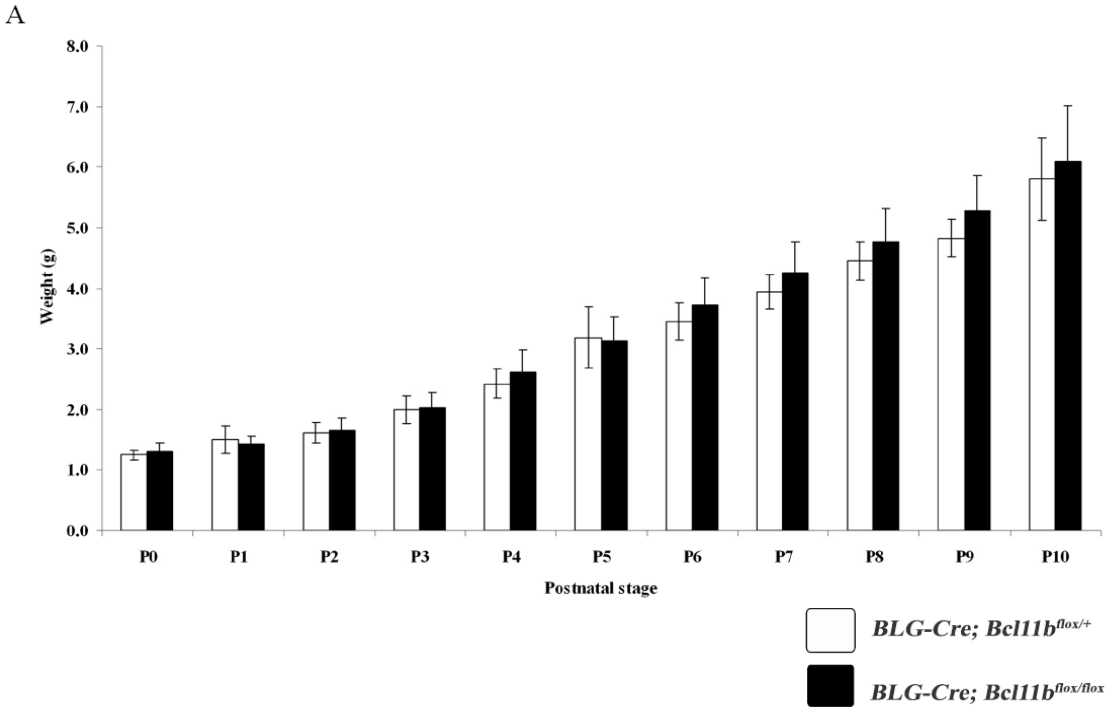


Figure 6.3. Graphs showing average weights of pups fostered from *BLG-Cre; Bcl11a$^{flox/flox}$* to *BLG-Cre; Bcl11a$^{flox/+}$* females. Pups from *BLG-Cre; Bcl11a$^{flox/flox}$* females are fostered to *BLG-Cre; Bcl11a$^{flox/+}$* females from **(A)** P1 and **(B)** P6. Their weights are monitored for 5 days and compared to littermates which remained with *BLG-Cre; Bcl11a$^{flox/flox}$* females. White and black bars represent average weights of pups fostered to *BLG-Cre; Bcl11a$^{flox/+}$* and *BLG-Cre; Bcl11a$^{flox/flox}$* females respectively. Error bars denote standard deviation obtained from analyzing 9 pups. * Pups found dead at P7.

6.2.2 *BLG-Cre; Bcl11b$^{flox/flox}$* females are able to nurse their pups

In contrast to the *BLG-Cre; Bcl11a$^{flox/flox}$* females, *BLG-Cre; Bcl11b$^{flox/flox}$* females were able to nurse their pups. Pups from *BLG-Cre; Bcl11b$^{flox/flox}$* females were healthy and there were no significant differences in weights of these pups compared to those from the control *BLG-Cre; Bcl11b$^{flox/+}$* females (Figure 6.4A). All the pups from *BLG-Cre; Bcl11b$^{flox/flox}$* females were weaned successfully. As shown in Chapter 4.2.6.2, expression of *Bcl11b* in the mammary glands was undetectable during lactation. Thus deletion of *Bcl11b* in the lactation gland did not affect the lactational capabilities of the glands. Whole mount carmine alum and histological analysis of day 2 lactation mammary glands showed no obvious differences between *BLG-Cre; Bcl11b$^{flox/flox}$* and the control *BLG-Cre; Bcl11b$^{flox/+}$* females (Figure 6.4B-C). The lobulo-alveoli in the *BLG-Cre; Bcl11b$^{flox/flox}$* lactation glands were stretched with large lumens and contained fat droplets and secretory vesicles (Figure 6.4C). In addition, molecular analysis also failed to reveal any significant differences in these *Bcl11b*-deficient glands from the wild-type controls (See below, Figures 6.10-6.12). These results suggest that *Bcl11b* is not critical in the lactation gland.



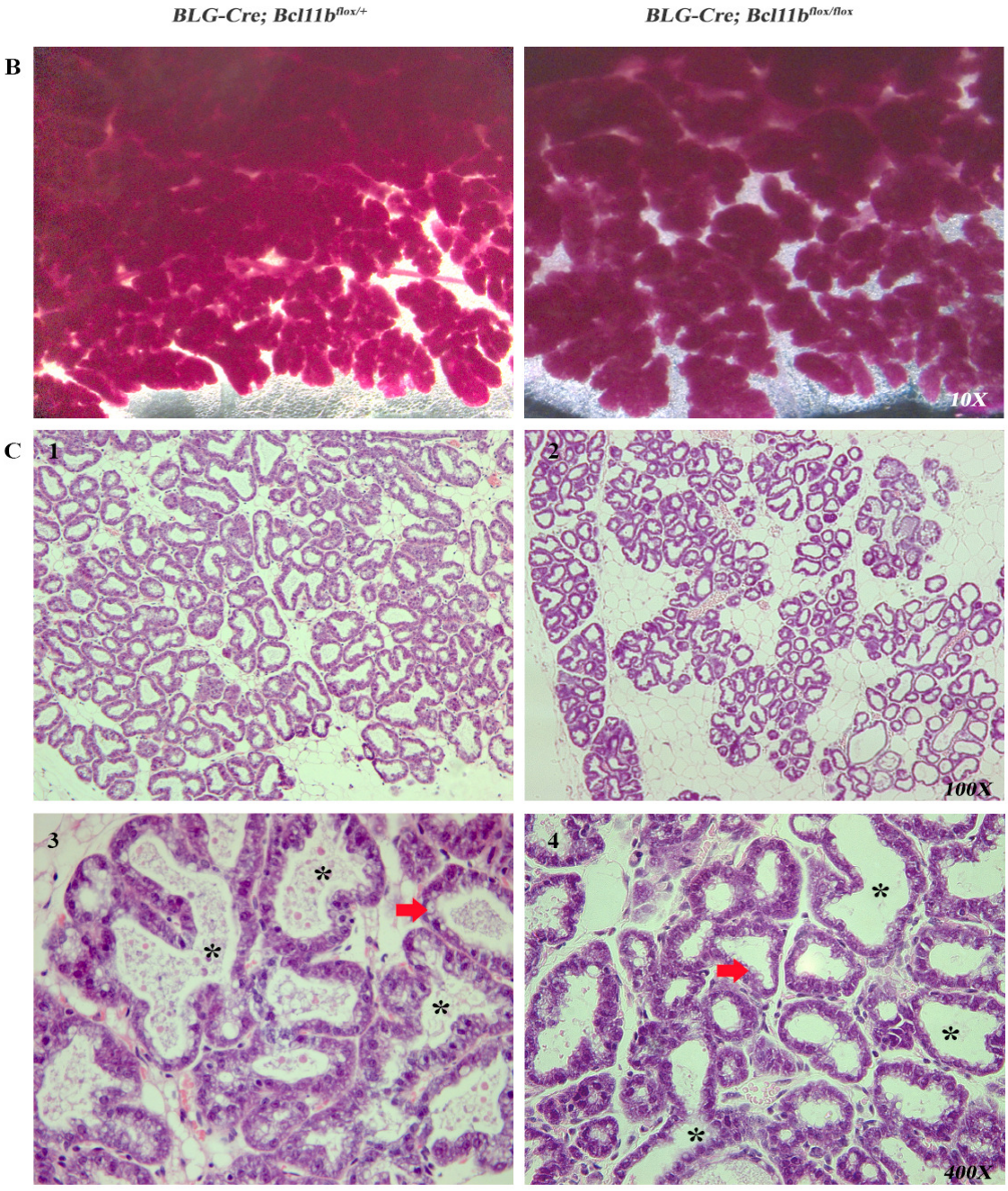


Figure 6.4. Analysis of *BLG-Cre; Bcl11b*$^{flox/flox}$ and control females. **(A)** Graph showing average weights of pups from *BLG-Cre; Bcl11b*$^{flox/+}$ and *BLG-Cre; Bcl11b*$^{flox/flox}$ females. White and black bars represent average weights of pups from *BLG-Cre; Bcl11b*$^{flox/+}$ and *BLG-Cre; Bcl11b*$^{flox/flox}$ females respectively. Errors bar denote standard deviation obtained from analyzing 67 and 60 pups of *BLG-Cre; Bcl11b*$^{flox/+}$ and *BLG-Cre; Bcl11b*$^{flox/flox}$ females respectively. **(B)** Whole mount carmine staining of day 2 lactation mammary glands isolated from *BLG-Cre*; *Bcl11b*$^{flox/+}$ and *BLG-Cre*; *Bcl11b*$^{flox/flox}$ females. **(C1-4)** H & E stained sections of day 2 lactation mammary glands isolated from *BLG-Cre*; *Bcl11b*$^{flox/+}$ and *BLG-Cre*; *Bcl11b*$^{flox/flox}$ females. **C3** and **C4** are higher magnification images of lobulo-alveolar structures of **C1** and **C2** respectively. Red arrows indicate fatty droplets within alveolar structures. * indicates examples of stretched alveolar structures.

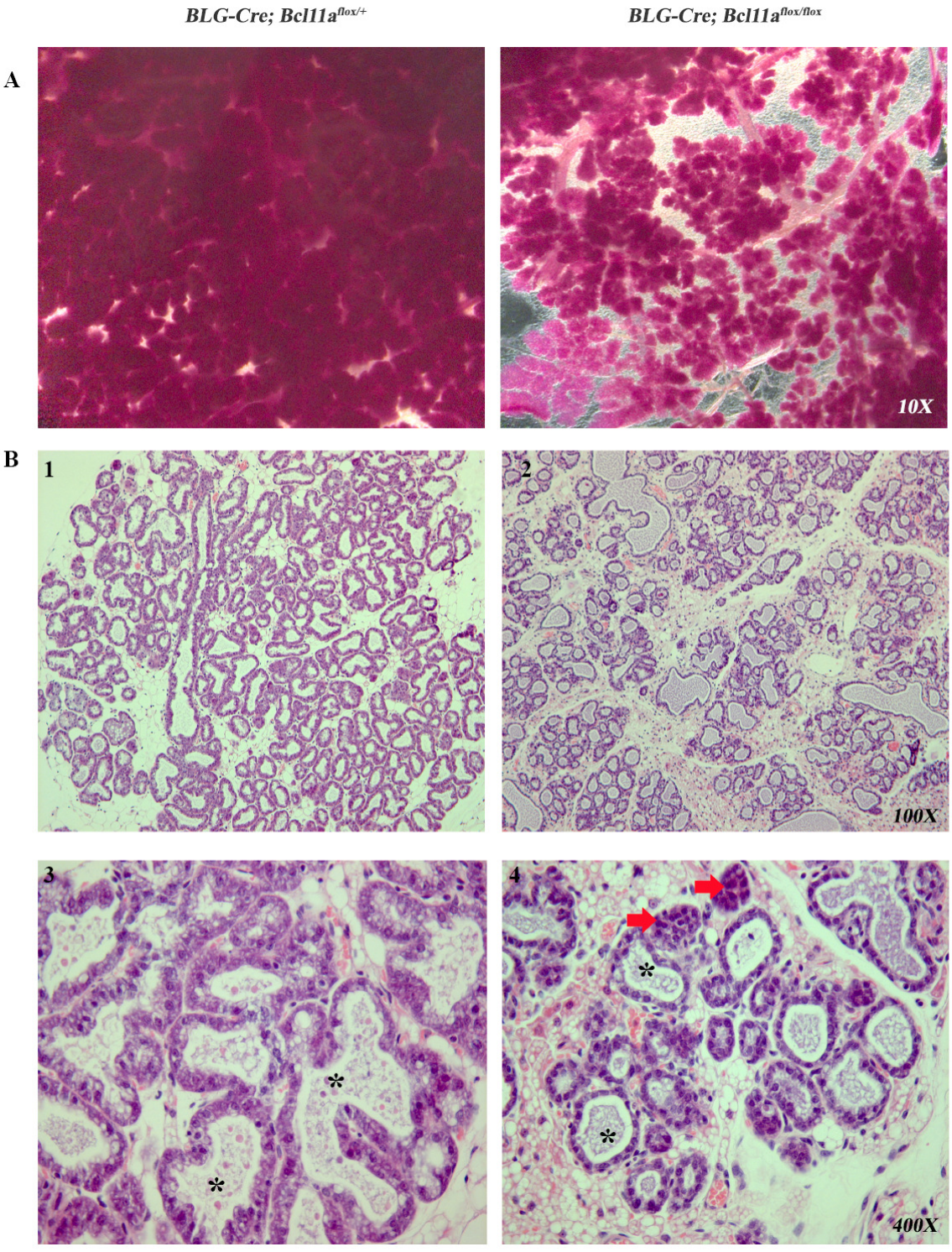
6.2.3 *Bcl11a* is essential in the lactation gland

As discussed in Chapter 6.2.1, *BLG-Cre; Bcl11a*$^{flox/flox}$ dams were unable to nurse their pups. Therefore, in order to confirm that this lactational defect was a functional consequence of the loss of *Bcl11a* in the lactation gland, whole mount and histological analyses of the mammary glands of *BLG-Cre; Bcl11a*$^{flox/flox}$ dams were carried out. As shown in Figure 6.5A, the mutant *BLG-Cre; Bcl11a*$^{flox/flox}$ glands had a substantial reduction in the number and size of milk-producing lobulo-alveoli compared to control *BLG-Cre; Bcl11a*$^{flox/+}$ mammary glands during early lactation (day 2) (Figure 6.5A). Histological analysis also showed that the control mammary glands were populated with numerous alveolar structures (Figure 6.5B1). In contrast, the mutant mammary glands had much fewer alveolar structures (Figure 6.5B2). Closer examination revealed that the alveoli in the control glands were stretched with large lumens (Figure 6.5B3). However the alveoli in the mutant glands contained lesser fat droplets and secretory vesicles (Figure 6.5B4). Numerous collapsed alveolar structures in the *Bcl11a*-deficient glands were also visible (Figure 6.5B4). By day 5 lactation, the alveolar structures in the control females had expanded and populated the entire mammary gland (Figure 6.5C1). In contrast, the mutant glands displayed sparse number of lobulo-alveoli with a large area of fat cells still visible in the mammary gland (Figure 6.5C2). In addition, the alveolar cells of the control glands had a thick luminal layers filled with secretory vesicles and fat droplets (Figure 6.5C3). However, most of alveoli in the mutant glands had thin luminal layers and reduced fat droplets and secretory vesicles (Figure 6.5C4). As development of lobulo-alveolar structures was unaffected in the *BLG-Cre; Bcl11a*$^{flox/flox}$ glands during pregnancy (Figure 6.1A), these observations demonstrated that loss of *Bcl11a* in the lactation glands affected the terminal differentiation state of the lobulo-alveoli, resulting in the observed lactational defects and hence a failure of *BLG-Cre; Bcl11a*$^{flox/flox}$ females to nurse their pups.

BLG-Cre-mediated excision of *Bcl11* genes was detected by PCR amplification of genomic DNA of the mammary glands (Figure 6.6A) and quantification using quantitative real time PCR (qRT-PCR) showed the deletion efficiency to be ~75-78%. Analysis of these glands using qRT-PCR showed ~95% reduction in the levels of *Bcl11a* mRNA in *BLG-Cre*; *Bcl11a*$^{flox/flox}$ mammary tissues relative to that in control glands

(Figure 6.6B). Thus the lactational defects observed in *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females were a functional consequence of loss of *Bcl11a* in the differentiated luminal cells. To confirm the lobulo-alveolar defects at the molecular level, I examined the RNA levels of milk protein genes using RT-PCR. As shown in Figure 6.7A, there were slight reductions in the mRNA levels of α-*casein*, *β-casein* and *α-lactalbumin* but a substantial reduction in the mRNA levels of *whey acidic protein* (WAP) in the *Bcl11a*-deficient mammary tissues (Figure 6.7A). This observation was confirmed in the Western analysis as WAP protein was undetectable in the *Bcl11a*-deficient mice (Figure 6.7B). In addition, immunostaining with an antibody to β-casein showed a significant reduction in the levels of β-casein in the *Bcl11a*-deficient glands (Figure 6.7C). No changes in the milk transcripts or protein levels were observed in the *Bcl11b*-deficient lactation glands compared to control glands (Figure 6.7A-B). Therefore, loss of *Bcl11a* but not *Bcl11b* in the lactation glands resulted in severe reduction in milk proteins, leading to the lactational defects and affected the ability of *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females to nurse their pups.

During early lactation, mammary epithelial cells undergo rapid proliferation. In the *Bcl11a*-deficient day 2 lactation mammary glands, staining with antibody to Ki67 showed a complete absence of proliferating cells (Figure 6.8A). Thus, deletion of *Bcl11a* in the terminally differentiated luminal cells also resulted in proliferation defects. Immunostaining with antibody to smooth muscle actin (SMA) on the other hand did not detect any obvious changes in the basal layer of the *Bcl11a*-deficient lactation glands (Figure 6.8B). This was expected as BLG-Cre-mediated excision should occur only in the luminal cells. In summary, whole mount, histological and molecular analyses demonstrated that loss of *Bcl11a* in the lactation glands resulted in severe lactational defects that affected milk production and proliferation of the luminal alveolar cells.


BLG-Cre; Bcl11a flox/+
BLG-Cre; Bcl11a flox/flox
A
10X
B
1
2
100X
3
4
400X

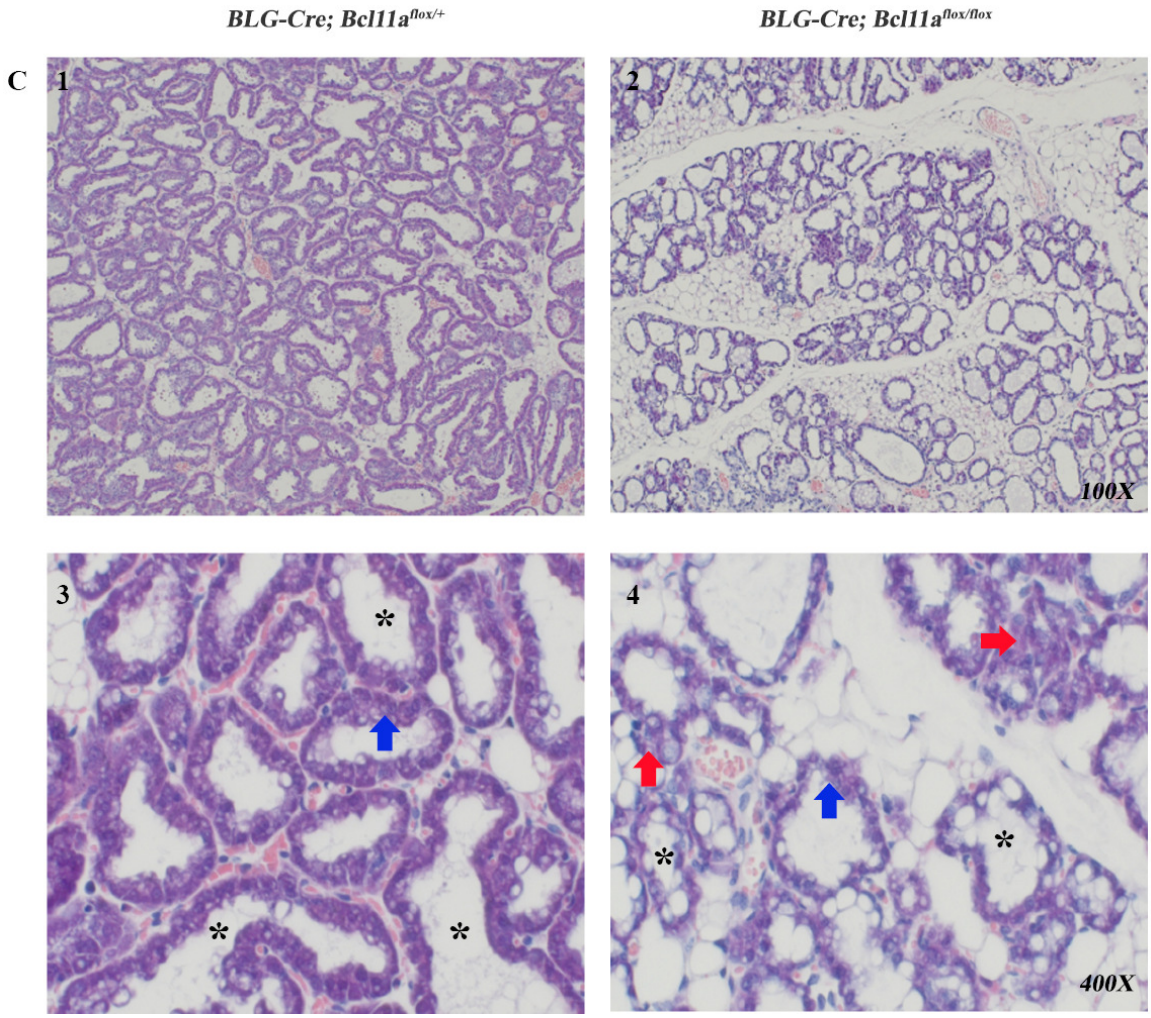


Figure 6.5. Histological analysis of *BLG-Cre; Bcl11a*$^{flox/flox}$ and control females. **(A)** Whole mount carmine staining of day 2 lactation mammary glands isolated from *BLG-Cre*; *Bcl11a*$^{flox/+}$ and *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females. **(B-C)** H & E stained sections of day 2 **(B1-4)** and day 5 **(C1-4)** lactation mammary glands isolated from *BLG-Cre*; *Bcl11a*$^{flox/+}$ and *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females. **B3** and **B4** are higher magnification images of lobulo-alveolar structures of **B1** and **B2** respectively; **C3** and **C4** are higher magnification images of lobulo-alveolar structures of **C1** and **C2** respectively. Red arrows indicate collapsed alveolar structures. Blue arrows show the thick luminal layer of alveolar structures in *BLG-Cre*; *Bcl11a*$^{flox/+}$ compared to that of *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands. * indicates the lumen of alveolar structures. The lumens of alveoli in *BLG-Cre*; *Bcl11a*$^{flox/+}$ lactation glands are large and distended with secretory vesicles while lumens of alveoli in *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands are small.

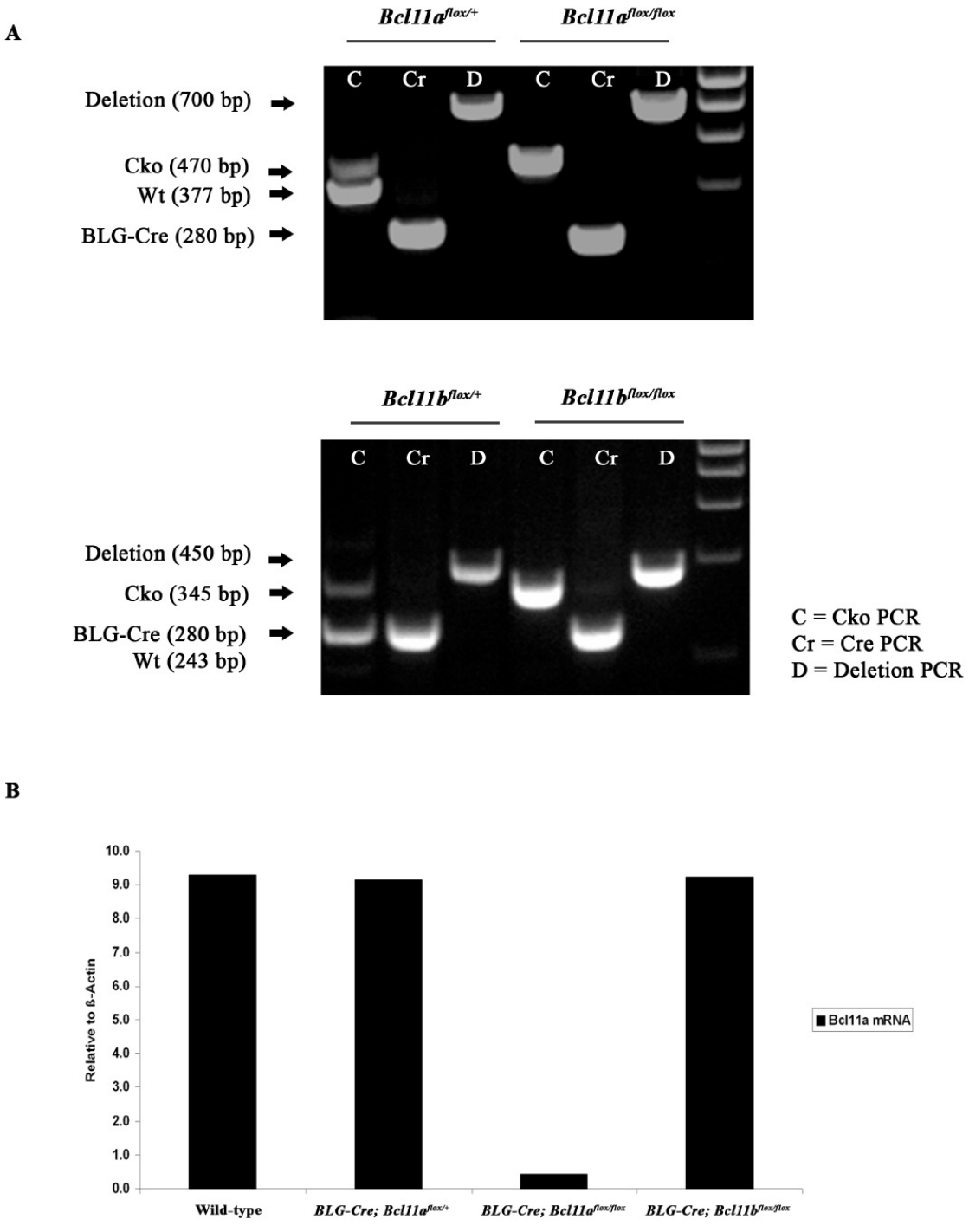


Figure 6.6. Detection of deletion of *Bcl11a* and *Bcl11b* after *BLG-Cre* expression. **(A)** Gel image showing PCR products obtained with primers using genomic DNA extracted from day 2 lactation mammary glands. Primers used are as described in Figure 5.4A and 5.12A. Cko: conditional band; Wt: Wild-type band; BLG-Cre: Cre band; Deletion: Deletion band. C: Cko PCR; Cr: Cre PCR; D: Deletion PCR. **(B)** Quantification of levels of *Bcl11a* transcripts in day 2 lactation mammary glands using quantitative real time PCR. Levels are normalized to β-actin.

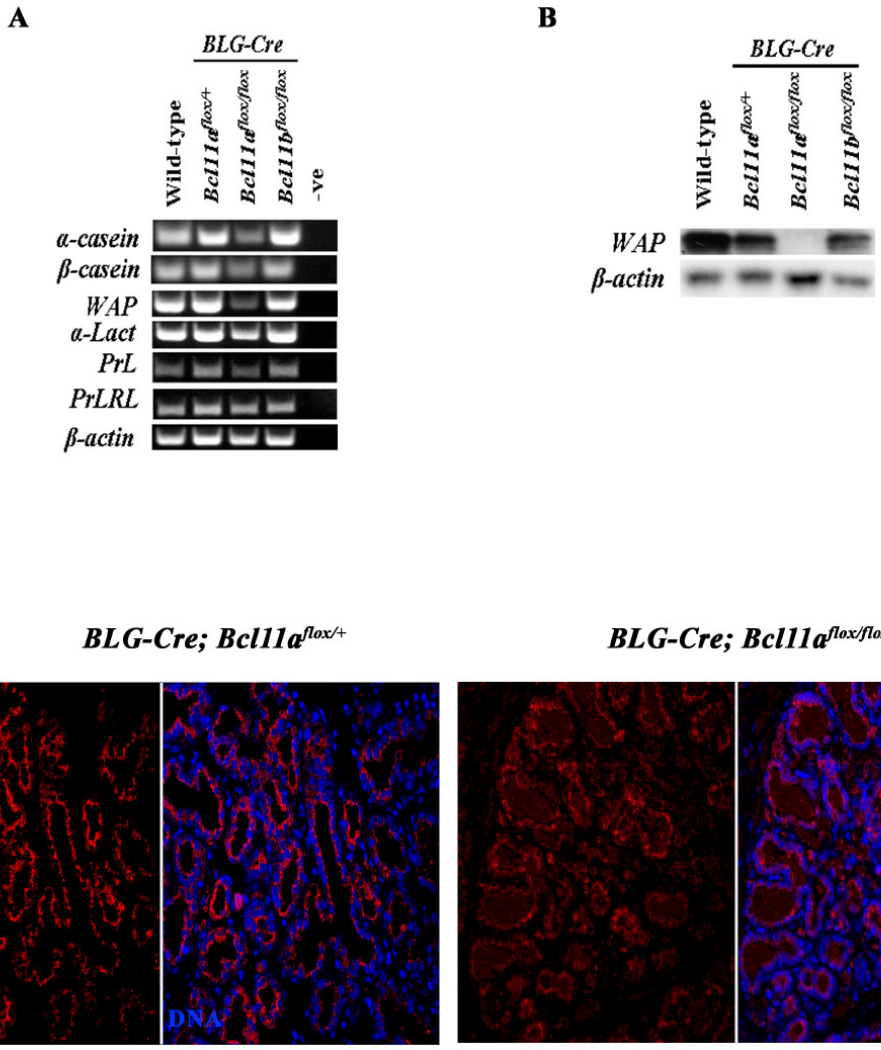


Figure 6.7. Analysis of milk transcripts and protein levels in *BLG-Cre; Bcl11*$^{flox/flox}$ and control females. (A) RT-PCR analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11*$^{flox/+}$) and mutant (*BLG-Cre*; *Bcl11*$^{flox/flox}$) females using milk protein primers. β-actin is used as a control. –ve indicates no template control. **(B)** Immunoblot analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11*$^{flox/+}$) and mutant (*BLG-Cre*; *Bcl11*$^{flox/flox}$) females with antibody to WAP. β-actin is used as a loading control. **(C)** Immunostaining of day 2 lactation mammary glands of *BLG-Cre*; *Bcl11a*$^{flox/+}$ and *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females with antibody to β-casein.

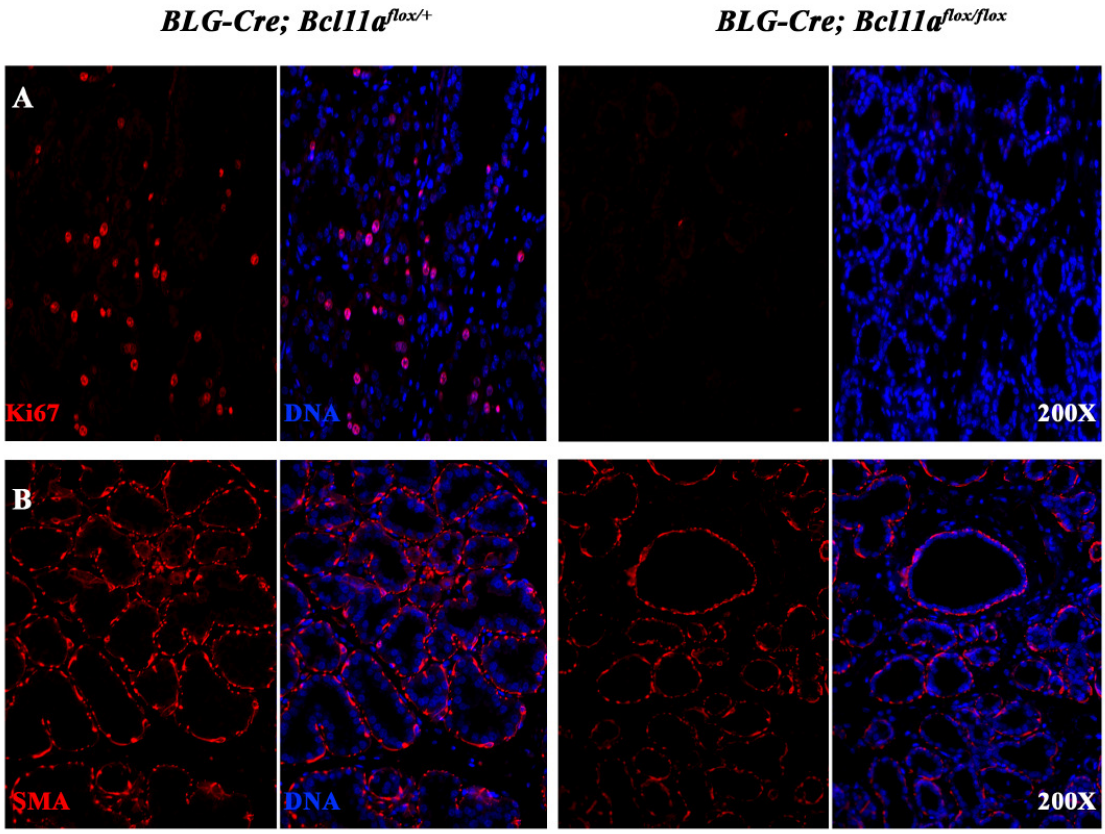


Figure 6.8. Immunostaining of day 2 lactation mammary glands of *BLG-Cre; Bcl11a*$^{flox/+}$ and *BLG-Cre; Bcl11a*$^{flox/flox}$ females. Immunostaining using antibodies to **(A)** Ki67 and **(B)** smooth muscle actin (SMA).

6.2.4 Loss of *Bcl11a* in lactation glands results in premature onset of involution

The JAK-Stat pathway plays an important role in development of the mammary gland during gestation and lactation. Activation of Stat5 transcriptional activities by phosphorylation is essential for normal lobulo-alveolar cell proliferation and differentiation (Liu et al., 1997). *Stat5*-deficient female mice showed curtailed mammary lobulo-alveolar outgrowth during pregnancy and a failure to lactate due to failed terminal differentiation of secretory luminal cells. To determine whether deletion of *Bcl11a* affected phosphorylation of Stat5 (p-Stat5), immunostaining with antibody to p-Stat5 was performed. Abundant p-Stat5 positive epithelial cells were detected in the control *BLG-Cre*; *Bcl11a*$^{flox/+}$ day 2 lactation glands, characteristics of alveolar differentiation (Figure 6.9A). Strikingly, no p-Stat5 positive epithelial cells were observed in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ day 2 lactation glands (Figure 6.9A), indicating a likely loss of lobulo-alveolar cells. Stat3 is an important mediator of the switch between survival and death signalling in mammary epithelial cells and activation of Stat3 is pivotal to the normal induction of involution (Abell et al., 2005; Chapman et al., 1999). In contrast to the loss of p-Stat5, phospho-Stat3 (p-Stat3, activated form of Stat3) was detected in many alveolar cells of the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ day 2 lactation glands (Figure 6.9B). Moreover, an increase in the number of cleaved Caspase-3 positive cells was observed in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ day 2 lactation glands (Figure 6.9C), indicating that loss of *Bcl11a* in the lactation glands resulted in activation of Stat3-mediated apoptosis. These results suggest that loss of *Bcl11a* resulted in premature onset of involution in the lactation glands. Despite the severe lactation defects, there was no change in the expression of prolactin and prolactin receptor, suggesting that *Bcl11a* may regulate Stat5 and Stat3 activities via different pathways in the lactation gland (Figure 6.6B).

To further study the premature onset of involution at the molecular level in the *Bcl11a*-deficient lactation glands, I first used RT-PCR to examine expression changes in key lactation and involution genes. I found that expression of *Stat5*, *Stat6*, *p85* and *Akt1* was down-regulated, and a dramatic up-regulation of *Lif*, *Stat3* and the *p55α/p50α* regulatory subunits of phosphoinositide-3-OH kinase [PI(3)K] in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ day 2 lactation glands (Figure 6.10A). Loss of *Bcl11a* in the lactation gland

therefore resulted in expression of genes that were normally only expressed during involution. To further probe the molecular changes resulting from the premature onset of involution in the *Bcl11a*-deficient glands, immunoblotting was performed to examine changes in components of the JAK-Stat pathway. As shown in Figure 6.10B, there was a decrease in Stat5 and a complete absence of p-Stat5 in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands. No changes in the levels of Stat6 and p-Stat6 (phospho-Stat6) were found. As expected, a significant increase in the levels of Stat3 and p-Stat3 was observed in the *Bcl11a*-deficient lactation glands (Figure 6.10B). During involution, expression of *p55α* and *p50α* is induced by Stat3 to down-regulate PI(3)K-Akt-mediated survival signalling (Abell et al., 2005). The premature onset of involution in the *Bcl11a*-deficient lactation glands was confirmed by increase in levels of cleaved Caspase 3, p55α and p50α together with a decrease in total Akt (Figure 6.10B). Consistent with no obvious phenotypes in the *Bcl11b*–deficient lactation glands, I did not find any noticeable changes in the transcript or protein levels of the components of the JAK-Stat pathway (Figure 6.10). Taken together, these results strongly suggest that *Bcl11a* plays an essential role in maintaining the identity of the terminally differentiated secretory cells as loss of *Bcl11a* in secretory luminal cells results in premature activation of Stat3-mediated apoptosis in the lactation glands.

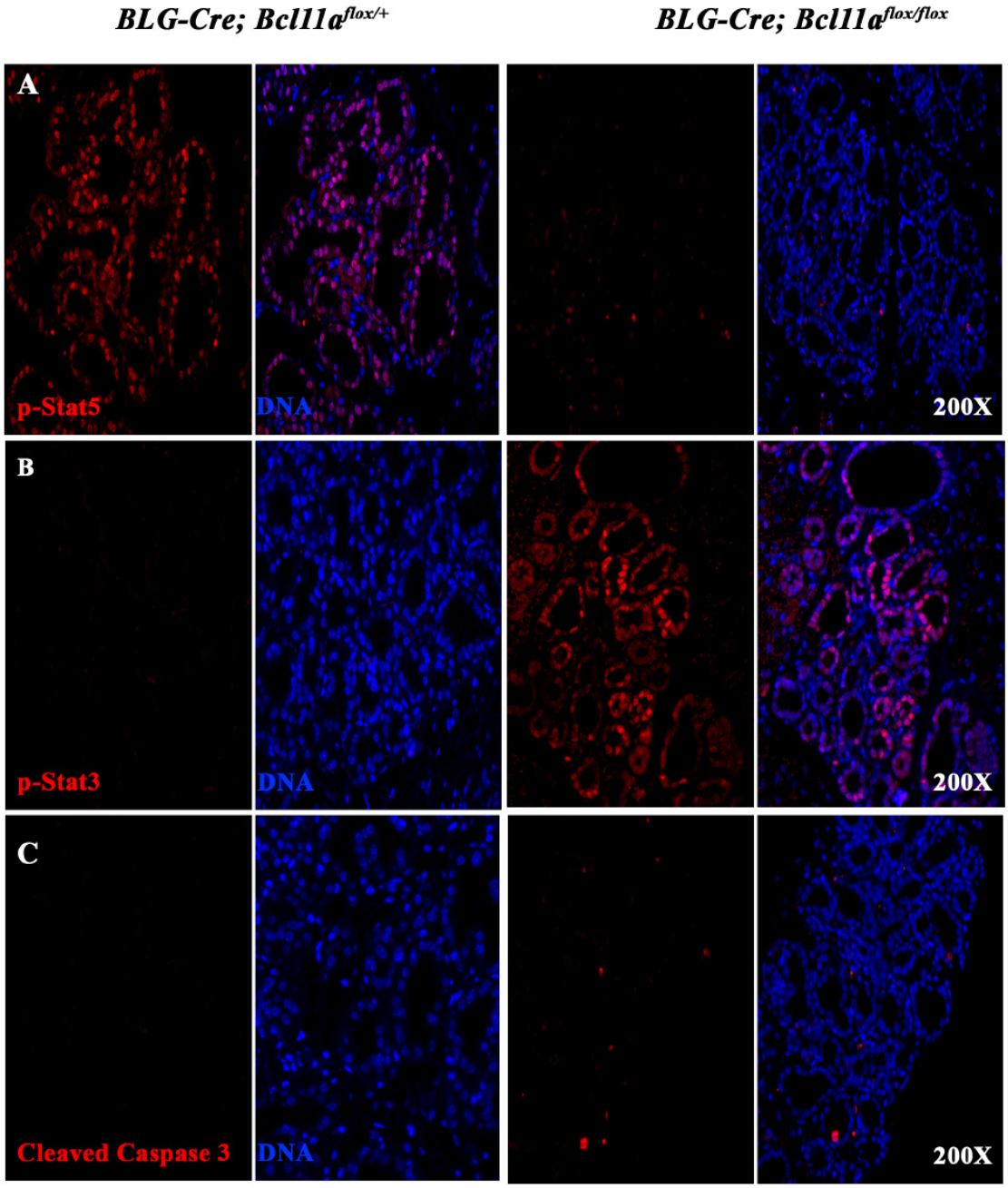


Figure 6.9. Loss of *Bcl11a* results in activation of Stat3-mediated apoptosis. Immunostaining using antibodies to **(A)** phospo-Stat5, **(B)** phospo-Stat3 and **(C)** cleaved Caspase-3.

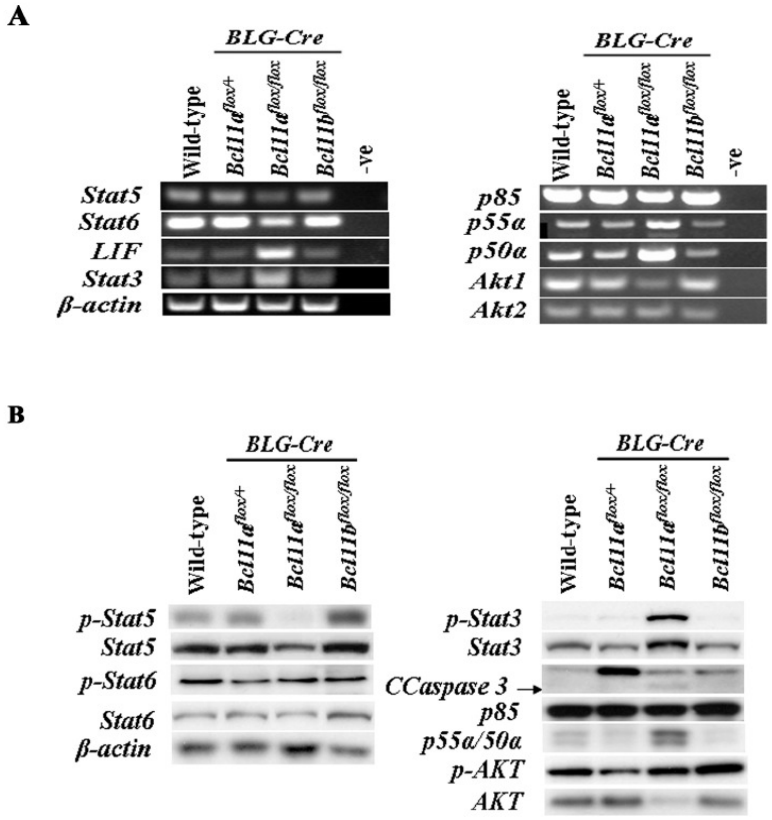


Figure 6.10. Analysis of Stat transcripts and protein levels in *BLG-Cre; Bcl11*$^{flox/flox}$ and control females. (A) RT-PCR analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11*$^{flox/+}$) and mutant (*BLG-Cre*; *Bcl11*$^{flox/flox}$) females. β-actin is used as a control. –ve indicates no template control. **(B)** Immunoblot analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11*$^{flox/+}$) and mutant (*BLG-Cre*; *Bcl11*$^{flox/flox}$) females with antibodies to components of JAK-Stat pathway. β-actin is used as a loading control. Note: The loading control panel for the immunoblot was the same as that in Figure 6.7B as experiments were carried out using the same protein samples.

6.2.5 Lactational defects in *Bcl11a*-deficient glands are not compensated by physiological levels of Gata-3

Gata-3 has been demonstrated to be essential for generation, differentiation and maintenance of luminal cells (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). Over-expression of *Gata-3* is sufficient to drive primary luminal progenitors to produce milk *in vitro* (Asselin-Labat et al., 2007). However, in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands, Gata-3 mRNA and protein levels appeared unchanged (Figure 6.11A and 6.11B), indicating that loss of *Bcl11a* in the terminally differentiated luminal cells was not compensated by the physiological amounts of Gata-3. Like *Bcl11a*, *Gata-3* is highly expressed in the lactation gland (Asselin-Labat et al., 2007), the fact that 95% reduction in *Bcl11a* expression in the lactation gland (Figure 6.6B) did not significantly alter *Gata-3* expression shows that Bcl11a and Gata-3 function in distinct luminal cell populations in the lactation gland. The Ets transcription factor, *Elf5* has recently been shown to specify mammary alveolar cell fate (Oakes et al., 2008). *Elf5* expression was reduced in the *Bcl11a*-deficient lactation glands (Figure 6.11A). Interestingly, *Gata-3* and *Elf5* have been shown to be mostly expressed in different luminal cell populations, with *Elf5*-expressing cells predominantly being ERα-negative and *Gata-3*-expressing cells mainly ERα-positive. Interestingly, I noticed that loss of a single copy of *Bcl11a* resulted in up-regulation of *Bcl11b* expression but this increase was apparently insufficient to cause an obvious mammary phenotype in the heterozygous females (Figure 6.11A). Taken together, these data demonstrate that *Bcl11a* plays an essential role in maintaining the identity of the terminally differentiated secretory cells.

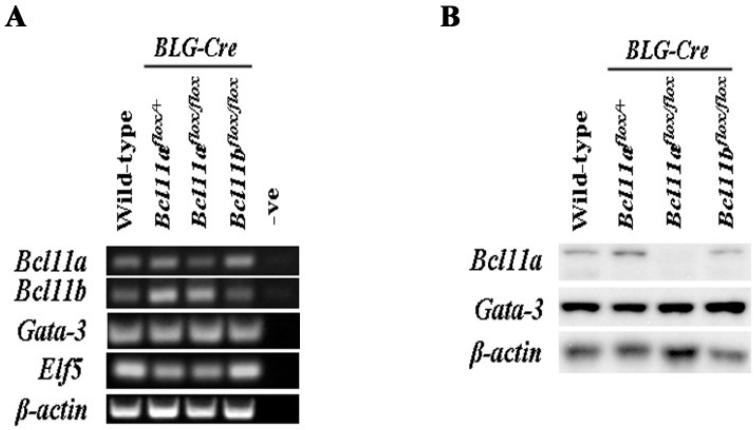


Figure 6.11. Analysis of transcription factors transcripts and protein levels in *BLG-Cre; Bcl11*$^{flox/flox}$ and control females. (A) RT-PCR analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11*$^{flox/+}$) and mutant (*BLG-Cre*; *Bcl11*$^{flox/flox}$) females. β-actin is used as a control. –ve indicates no template control. **(B)** Immunoblot analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11*$^{flox/+}$) and mutant (*BLG-Cre*; *Bcl11*$^{flox/flox}$) females with antibodies to Bcl11a and Gata-3. β-actin is used as a loading control. Note: The loading control panel for the immunoblot was the same as that in Figure 6.7B as experiments were carried out using the same protein samples.

6.2.6 Loss of *Bcl11a* in lactation glands also results in dysregulation of Notch signalling pathways

Besides the JAK-Stat pathway, several molecular pathways such as Notch signalling pathway are known to be involved in mammary lobulo-alveolar development. Constitutive activation of Notch signalling by expression of the intracellular domains of Notch1, Notch3, and Notch4 inhibits alveolar terminal differentiation and causes lactational defects (Gallahan et al., 1996; Hu et al., 2006; Jhappan et al., 1992; Smith et al., 1995). However, the roles of different Notch ligands and receptors in mammary development and cell fate determination remained unclear. Loss of *Bcl11a* in the virgin glands resulted in increased Notch1 and Jagged1 expression (Chapter 5). This is similar to the loss of *Bcl11a* in lymphocytes which showed up-regulation in *Notch1* transcripts (Liu et al., 2003b). To interrogate whether other components of the Notch signalling pathway were altered in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands, RT-PCR was performed to detect changes in gene expression. All the known Notch ligands, except Delta-like ligand 3 (*Dll3*), were expressed in the lactation gland (Figure 6.12A). Remarkably, loss of *Bcl11a* led to dramatic up-regulation of *Dll3* expression in the lactation glands (Figure 6.12A). In contrast, the mRNA levels of *Jagged1* and *Dll1* (Delta-like ligand 1) were reduced while expression of *Dll4* (Delta-like ligand 4) was completely ablated (Figure 6.12A). No changes in the transcript levels of *Jagged2* were observed. Using immunoblotting (Figure 6.12C) and immunostaining (Figure 6.12D) with antibody to Jagged1, a dramatic reduction in Jagged1 protein levels in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands was observed. This observation differs from that of the virgin gland where loss of *Bcl11a* resulted in an increase in Jagged1, suggesting that Jagged1 probably plays different roles in the virgin and lactation glands.

All four Notch receptors were expressed in the wild-type day 2 lactation gland (Figure 6.12A). Deletion of *Bcl11a* resulted in an increase in *Notch1* expression (Figure 6.12A). Interestingly, no obvious increase in the protein level of Notch1 was detected using immunoblotting but activation of Notch1 signalling pathway was confirmed by a slight increase in activated Notch1 (Notch1 intra-cellular domain, ICN) (Figure 6.12C) and up-regulation of downstream targets of Notch1 as discussed below (Figure 6.12B). Additionally, an obvious increase in Notch1 was detected using immunostaining with

antibody to Notch1 (Figure 6.12E). Surprisingly, loss of *Bcl11a* led to an absence of Notch3 expression at both transcript and protein levels (Figure 6.12A and 6.12C). Interestingly, Notch3 has been shown to promote lobulo-alveolar differentiation during late pregnancy (Hu et al., 2006) and could act as a repressor by blocking the ability of the Notch1 intracellular domain to activate expression through the Hairy enhancer of split (Hes) family of transcriptional repressors, *Hes1* and *Hes5* promoters (Beatus et al., 1999). In addition, a recent report proposed that in the human mammary epithelium, Notch3 is critical for luminal differentiation program *in vitro* (Raouf et al., 2008). No apparent changes in the expression of *Notch2* and *Notch4* were observed (Figure 6.12A).

To determine whether deletion of *Bcl11a* affected the downstream target genes of the canonical Notch pathway in the lactation mammary gland, I analyzed expression of *Hes* genes using RT-PCR. Most of the *Hes* genes, including *Hes1*, *Hes2*, *Hes3*, *Hes5* and *Hes6* were expressed at low levels in the control lactation glands (Figure 6.12B). However, significant increases in the expression of these *Hes* genes were detected in the *BLG-Cre*; *Bcl11a*$^{flox/flox}$ lactation glands, suggesting activation of the canonical Notch signalling pathway (Figure 6.12B). In contrast, *Hes7*, which was expressed at high levels in the control lactation glands, was completely absent in the *Bcl11a* mutant glands (Figure 6.12B). It is not clear at this moment whether the loss of expression of *Notch3*, *Dll4, Jagged1* and *Hes7*, and the gain of *Dll3* and *Notch1* expression in the *Bcl11a*-deficient glands are directly implicated in the phenotypes of the *Bcl11a*-deficient lactation glands. Nevertheless, these results demonstrate that loss of *Bcl11a* in the lactation mammary glands results in dysregulation of Notch signalling pathway, and indicate that different Notch receptors and ligands play different roles in mammary development and epithelial differentiation. In contrast, no changes in the transcript or protein levels of the components of the Notch pathway were observed in the *Bcl11b*–deficient lactation glands compared to control glands (Figure 6.12).

Taken together, deletion of *Bcl11a* but not *Bcl11b* in the lactation gland ablated the function of terminally differentiated secretory luminal cells and resulted in premature onset of involution. These data clearly demonstrate that *Bcl11a* is critical during lactation and maintains the differentiated functional state of the lactation gland at least partly through modulating Notch signalling.

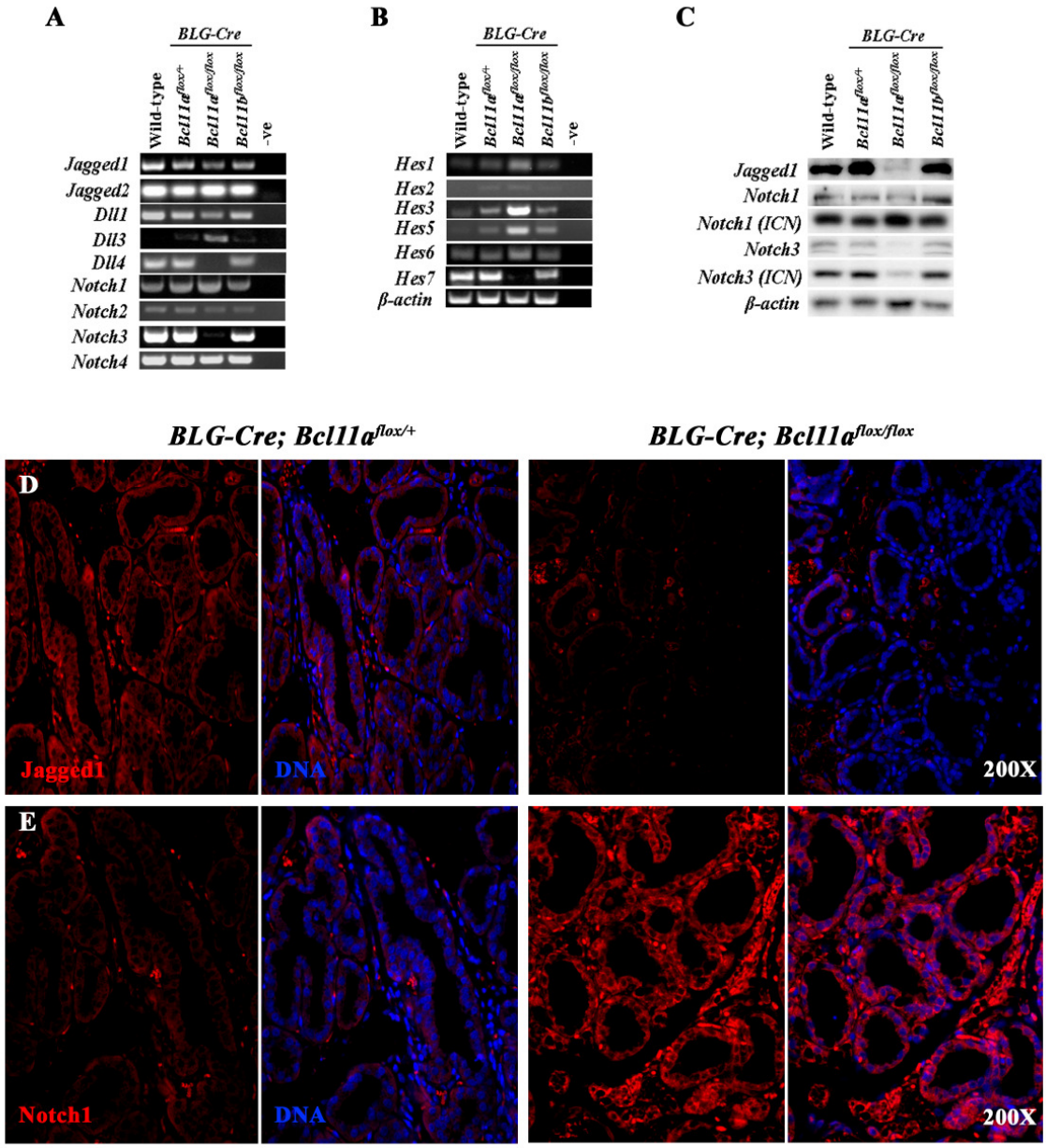


Figure 6.12. Analysis of Notch signalling pathway transcripts and protein levels in *BLG-Cre; Bcl11$^{flox/flox}$* and control females. (A-B) RT-PCR analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11$^{flox/+}$*) and mutant (*BLG-Cre*; *Bcl11$^{flox/flox}$*) females using primers to **(A)** *Notch* receptors/ligands and **(B)** *Hes*. β-actin is used as a control. –ve indicates no template control. **(C)** Immunoblot analysis of day 2 lactation mammary glands from control (*BLG-Cre*; *Bcl11$^{flox/+}$*) and mutant (*BLG-Cre*; *Bcl11$^{flox/flox}$*) females using antibodies to Notch1, Notch3 and Jagged1. β-actin is used as a loading control. **(D-E)** Immunostaining of day 2 lactation mammary glands from *BLG-Cre; Bcl11a$^{flox/+}$* and *BLG-Cre; Bcl11a$^{flox/flox}$* females using antibodies to **(D)** Jagged1 and **(E)** Notch1. Note: The loading control panel for the immunoblot was the same as that in Figure 6.7B as experiments were carried out using the same protein samples.

6.3 Discussion

In this Chapter, I showed that *Bcl11a* is critical for the maintenance of terminally differentiated luminal secretory cells. Expression of *Bcl11a* was detected in the luminal layer of lobulo-alveoli while expression of *Bcl11b* was virtually undetected during lactation. To delete *Bcl11a* and *Bcl11b* specifically in the luminal layers during lactation, I used the BLG-Cre mice (Selbert et al., 1998). Expression of BLG-Cre is low in the virgin mice but increases throughout gestation and is highest during lactation. Histological analyses of mammary glands from *BLG-Cre*; *Bcl11*$^{flox/flox}$ females at late gestation showed the presence of lobulo-alveolar structures with lipid droplets and secretory vesicles, suggesting that development and differentiation of alveoli during gestation was not significantly affected.

6.3.1 *Bcl11a* maintains terminally differentiated luminal secretory cells

Deletion of *Bcl11a* during lactation resulted in severe lactational defects in the mammary glands. *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females were unable to nurse their pups. These pups had little or no milk in their stomach from P1, were severely dehydrated and usually died by P7 if not rescued. Consistent with the undetectable *Bcl11b* expression during lactation, the *BLG-Cre*; *Bcl11b*$^{flox/flox}$ females were able to nurse their pups and histological analyses showed no obvious lactational defects in the mammary glands. Histological analyses of lactation mammary glands from *BLG-Cre*; *Bcl11a*$^{flox/flox}$ females showed a drastic reduction in the number and size of lobulo-alveoli compared to the control females. The mutant lobulo-alveoli had small lumens and contained fewer fat droplets and secretory vesicles. Moreover, numerous collapsed alveolar structures were present within the *Bcl11a*-deficient glands. Severe lactational defects were exemplified by the decrease in transcripts of milk protein genes and the complete absence of p-Stat5 and WAP in the lactation glands. I found that a premature onset of involution had occurred in the *Bcl11a*–deficient lactation gland, which was confirmed by expression of involution genes and activation of Stat3-mediated apoptosis. Surprisingly, the levels of the key luminal differentiation transcription factor, *Gata-3* were not affected in the *Bcl11a*-deficient lactation glands, suggesting that physiological levels of *Gata-3* was

unable to compensate for loss of *Bcl11a*. This also suggests that Bcl11a and Gata-3 may function in distinct luminal progenitor populations. On the other hand, deletion of *Bcl11a* did result in a reduction in *Elf5* expression. This, together with the earlier observations that *Bcl11a* was expressed and necessary in the $Sca1^-$ progenitors (Chapter 4.19B and Appendix A.8), indicate that *Bcl11a* is likely to be expressed in the alveolar progenitors together with *Elf5* but not *Gata-3*. Loss of *Bcl11a* would thus lead to depletion of these alveolar progenitors and their differentiated derivatives. In addition, Bcl11a appeared to suppress *Bcl11b* in the lactation gland as loss of a single copy of *Bcl11a* resulted in up-regulation of *Bcl11b* expression. These results thus demonstrate that Gata-3 is not a key player in specifying and maintaining the alveolar secretory cell fate and that *Bcl11a* is essential for the maintenance of terminally differentiated luminal secretory cells.

6.3.2 Implications of different Notch signalling pathways

The Notch signalling pathway plays a pivotal role in several cell functions; such as cell fate decision, proliferation, differentiation, and cell death during development (Artavanis-Tsakonas et al., 1999; Radtke et al., 2004). Notch signalling is thought to be required for luminal cell development and lineage maintenance (Buono et al., 2006). However over-expression of activated *Notch1*, *Notch3* or *Notch4/Int3* in the mammary epithelium blocked or delayed alveologenesis and resulted in mammary tumour formation (Gallahan et al., 1996; Hu et al., 2006). This apparent discrepancy regarding the role of Notch in mammary development can be explained by the fact that Notch signalling is dosage-sensitive. For example, Notch is essential for T-cell development. However, abnormal high levels of Notch signalling block normal thymic T-cell development (Rothenberg, 2007a). Thus, these results suggest that regulating the levels of Notch signalling is critical to the balance between normal mammary development and tumorigenesis.

Our lab has shown that loss of *Bcl11a* in lymphocytes resulted in up-regulation of *Notch1* transcripts (Liu et al., 2003b). This study herein revealed that loss of *Bcl11a* in the lactation gland resulted in an increase in the levels of *Notch1* transcripts and activated Notch1 intracellular domain (ICN) as well as up-regulation of downstream targets of *Notch1* such as the *Hes* genes. An unexpected finding is that Bcl11a differentially

regulates Notch signalling in the mammary gland. In contrast to the up-regulation of Notch1, deletion of *Bcl11a* caused the loss of Notch3 in the lactation gland. Notch3 has been shown to promote lobulo-alveolar development during late pregnancy (Hu et al., 2006) and also plays a distinct role from NOTCH1 in ERBB2 negative breast cancer (Yamaguchi et al., 2008). In addition, Notch3 has been shown to be critical for the commitment of bi-potent progenitors to the luminal lineage *in vitro* (Raouf et al., 2008). In addition, expression of *Dll4* and *Hes7*, were completely absent in the *Bcl11a* mutant lactation mammary epithelium. Further work is required to address the function of each of these Notch components in normal mammary gland development. Nevertheless, my current data clearly implies that Notch1 and Notch3 may play distinct roles in luminal differentiation and maintenance of terminal differentiated luminal cell state and also highlight the diverse roles played by different components of Notch signalling pathway in mammary development.

6.3.3 Proposed working model of *Bcl11* genes during lactation

A precise regulation of the spatial and temporal control of *Bcl11a* and *Bcl11b* is required for normal mammary development. Up-regulation of *Bcl11a* and down-regulation of *Bcl11b* is required for terminal differentiation of secretory luminal cells during gestation and lactation. The JAK-Stat pathway plays critical roles during mammary gland development. During pregnancy and lactation, Stat5 is a key mediator of prolactin signalling, which is required for lobulo-alveolar development and lactogenesis (Cui et al., 2004; Liu et al., 1997). In contrast, Stat3-mediated apoptosis is essential for the involution process that removes redundant alveolar structures (Chapman et al., 1999). Thus Stat5 and Stat3 play reciprocal functions in the mammary gland and the balance of phosphorylated Stat5 and Stat3 levels is critical to the normal function of the gland. This study has demonstrated that *Bcl11a* prevents premature onset of involution in the lactation gland. Loss of *Bcl11a* thus led to the loss of p-Stat5 and gain of p-Stat3 positive luminal cells. My data also revealed that *Bcl11a* maintains secretory luminal cell fate by differentially regulating the Notch1 and Notch3 signalling pathways which may in turn regulate the levels of Stat3 and Stat5 during lactation (Figure 6.13). Loss of *Bcl11a* resulted in the absence of *Jagged1/Dll4/Notch3/Hes7* and ectopic expression of

Dll3/Notch1/Hes1/3/5/6 that cumulatively caused the activation of Stat3-mediated apoptosis in the lactation gland. Formal proof of the direct regulation and the functional consequences require further molecular studies and analyses of mutant mice that have gain or loss of different components of the Notch pathway. Unexpectedly, Gata-3, a key luminal transcription factor is not affected by loss of *Bcl11a*. This suggests that *Bcl11a* and *Gata-3* are likely to be expressed in distinct alveolar cell population. In conclusion, the results in Chapter 6 demonstrate that *Bcl11a* is essential for the maintenance of secretory luminal cell fate in the lactation mammary gland.

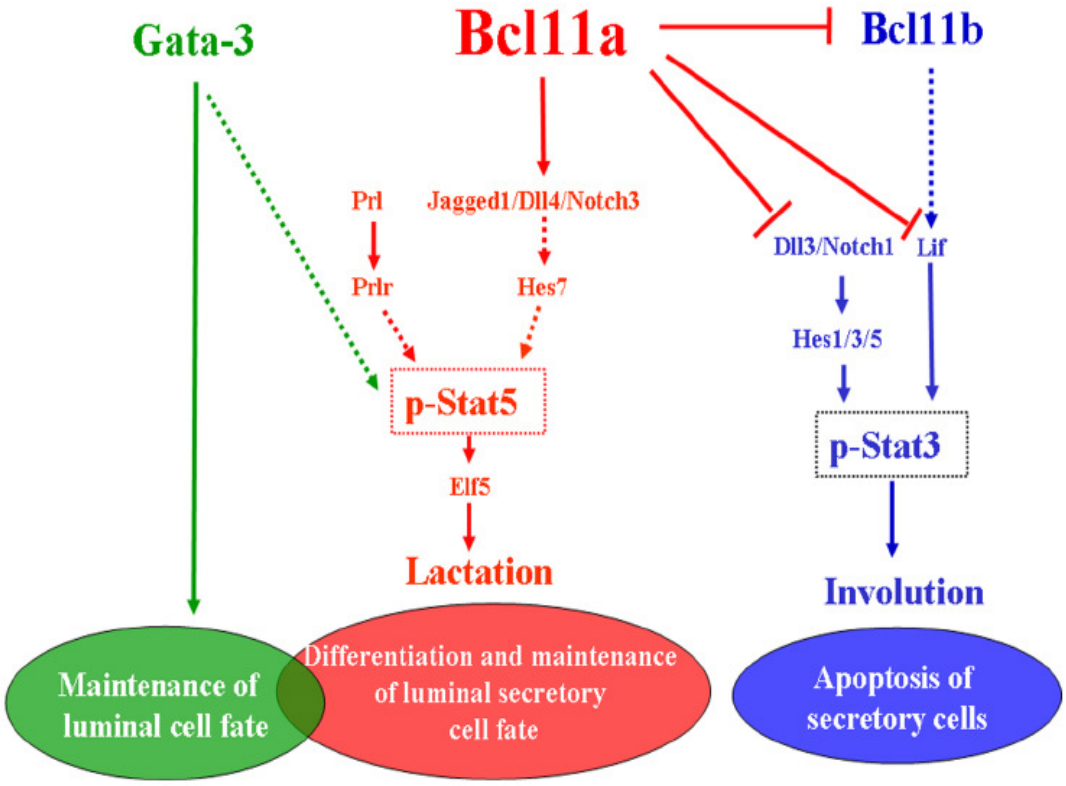


Figure 6.13. Proposed working model of the roles of *Bcl11* genes in mammary lineages. Putative model for role of *Bcl11a* in maintenance of secretory luminal cell fate in relationship to Notch and JAK-Stat pathways.

CHAPTER 7:
GENERAL DISCUSSION

7.1 Summary

7.1.1 Expression patterns of *Bcl11* genes in embryonic and adult tissues

Bcl11a and *Bcl11b* exhibited specific spatial and temporal expression patterns during mouse embryogenesis. Using *lacZ*-tagged alleles, *Bcl11a* and *Bcl11b* expression was detected. At 10.5 dpc, both genes showed overlapping expression patterns in the pharyngeal arches. From 12.5 dpc, expression of both genes was detected in the craniofacial regions of the embryos. In addition, both genes were highly expressed in the CNS from 12.5 dpc and this overlapping expression pattern was maintained to adulthood. Taken together, the overlapping expression patterns of both genes in the craniofacial regions and CNS of the mouse suggest that these two genes may play similar and complementary roles in the development of these specific regions. Alternatively, they may also have cross-antagonistic roles in these regions and the relative levels of each Bcl11 transcription factor ultimately determine the phenotype of the region.

Interestingly, several regions of differential *Bcl11* expression were observed during embryonic development. *Bcl11a* was expressed in the heart and not in the lungs at 14.5 dpc while expression of *Bcl11b* was detected in the lung and not in the heart. Intriguingly, at 18.5 dpc, expression of *Bcl11a* was detected in the lungs whereas expression of *Bcl11b* was absent in the lungs. These dynamic and reciprocal expression patterns of *Bcl11* genes in the lungs suggest that they may be involved in different phases of lung development: while *Bcl11b* may be required at the early pseudoglandular stage (initial generation of lung structures), *Bcl11a* may then be required at the late terminal sac stage (terminal differentiation of lung alveolar). In addition, differential *Bcl11* expression was also detected in the fetal liver, thymus, developing bones, cartilage and mammary gland, suggesting that *Bcl11* genes might play different roles in development of these tissues.

Taken together, whole mount X-gal staining revealed that *Bcl11* genes are expressed in a plethora of tissues with regions of overlapping expressions (craniofacial

regions and CNS) and regions of differential expression (lungs, fetal liver, thymus and mammary gland).

7.1.2 Reciprocal expression of *Bcl11* genes in hematopoietic lineages

The generation of *Bcl11-lacZ*-tagged mice made it possible to determine *Bcl11* expression at a single cell level. By using $Bcl11^{lacZ/+}$ cells with Fluorescein di-β-D-galactopyranoside (FDG) in combination with other cell surface markers, expression of *Bcl11* genes in hematopoietic cells was determined. Expression of *Bcl11a* was detected in hematopoietic stem cells (HSCs) and in all other blood lineages such as myeloid, erythroid and lymphoid cells with the exception of T lymphocytes. In contrast, expression of *Bcl11b* was only detected in T lymphocytes and not in other blood lineages.

Maturation of T lymphocytes occurs in distinct stages characterized by expression of different cell surface markers (Rothenberg, 2007a). Examination of the immature CD4/CD8 double-negative (DN) thymocytes with additional markers (CD44 and CD25) revealed that *Bcl11a* was only expressed at the DN1 stage in ~16% of immature thymocytes but not in other T cells. In contrast, *Bcl11b* was expressed from the DN1 stage in ~24% of immature thymocytes and became highly expressed in all thymocytes from DN2 stage. The dynamic, reciprocal, and possibly mutually exclusive expression patterns of *Bcl11* genes during T cell development are summarised in Figure 4.19 (Chapter 4). These expression patterns are consistent with *Bcl11b*'s role in T cell development as demonstrated in the knockout mouse where T cell development was blocked at the DN3 stage (Wakabayashi et al., 2003b). Taken together, these results demonstrate that while *Bcl11a* is expressed in all blood lineages including HSCs, expression of *Bcl11b* is restricted only to T cell lineages. Importantly, expression patterns of *Bcl11b* in early T cell development imply that it may be the key transcription factor for T cell lineage commitment and maturation. Indeed, it has been suggested in a recent review that *Bcl11b* could be the primary regulator of T cell lineage commitment (Rothenberg, 2007a).

7.1.3 Dynamic expression patterns of *Bcl11* genes during mammary development

Bcl11a and *Bcl11b* exhibited unique and dynamic expression patterns in the mammary gland. Expression of both genes was detected in early embryonic mammary development. *Bcl11b* was expressed in the milk line at 10.5 dpc and from 12.5 dpc, its expression became localized in the mammary placodes and mammary buds. In contrast, expression of *Bcl11a* was detected in the mammary buds and mesenchyme only from 13.5 dpc. In the virgin gland, *Bcl11a* was expressed in terminal end buds (TEBs) and histological analysis revealed that this expression was specifically located within both the cap and body cells of TEBs. Conversely, *Bcl11b* was only detected in the neck region of TEBs within the cap cell layer which goes on to form the basal/myoepithelial layer of mature ducts. In the mature virgin glands, *Bcl11a* was expressed in both the luminal and basal layers, while expression of *Bcl11b* was detected predominantly in the basal layer. In the luminal compartments of the virgin gland, *Bcl11a* expression was found primarily in the luminal progenitors ($CD24^{hi}CD49b^{+}$) and majority of these *Bcl11a*-expressing cells were the $Sca1^{-}$ ($ER\alpha^{-}$) luminal progenitors. On the other hand, only a small number of $Sca1^{-}$ ($ER\alpha^{-}$) luminal progenitors expressed *Bcl11b*.

During gestation, both *Bcl11* genes were up-regulated and *Bcl11a* was detected in all luminal lineages while *Bcl11b* was restricted to the basal layer of ducts. During lactation, only expression of *Bcl11a* was detected in luminal secretory cells. These results suggest that *Bcl11a* may be important for the luminal lineages while *Bcl11b* may be important for maintaining basal cell identity and/or suppressing luminal cell fate. In addition, the differential expression patterns of both genes during involution suggest that the *Bcl11* genes might play different roles during the involution phase. Collectively, these unique and dynamic expression patterns of *Bcl11* genes during mammary gland development suggest the possible important roles of these genes at different stages of mammary gland development.

7.1.4 *Bcl11* genes are critical regulators of lineage commitment in the mammary epithelium

Loss of *Bcl11a* and *Bcl11b* in the mouse embryos resulted in embryonic mammary defects. Deletion of *Bcl11a* led to defective embryonic bud formation and failure of regression of mammary buds in the male embryo while loss of *Bcl11b* caused the absence of the third pair of mammary buds and significantly affected formation of the other buds. These data demonstrated that *Bcl11* genes play important roles in normal embryonic mammary development. Conditional deletion of *Bcl11a* in the virgin glands led to profound defects in the bi-layer ductal epithelium where luminal and basal cells appeared to form a thin cellular layer within many areas of the ducts. Flow cytometric analysis showed that the *Bcl11a*-deficient epithelium exhibited a $CD24^{hi}CD49f^{+}$ profile (luminal fraction in the wild-type virgin gland). At the cellular level, loss of *Bcl11a* also led to a relative increase in $ER\alpha^{+}$ luminal cells and a dramatic decrease in Gata-3^{+} luminal cells. RT-PCR also demonstrated that loss of *Bcl11a* resulted in a decrease in luminal and basal markers. Furthermore, *Bcl11a* was shown to be essential for the maintenance of terminally differentiated luminal secretory cells. Conditional deletion of *Bcl11a* in the lactation gland led to severe lactational defects and the premature onset of involution. Taken together, these results show that *Bcl11a* is a key gene in mammary lineage specification and function.

Conditional deletion of *Bcl11b* caused precocious alveologenesis in the virgin gland, resulting in the formation of alveolar-like structures that produced β-casein milk protein. Interestingly, flow cytometric analysis of the *Bcl11b*-deficient mammary epithelium showed that these epithelial cells exhibited a $CD24^{+}CD49f^{hi}$ profile (basal fraction in the wild-type virgin gland). However, further analysis of these sorted *Bcl11b*-deficient $CD24^{+}CD49f^{hi}$ cells showed that they expressed luminal markers such as *β-casein*, *NKCC1*, *Notch1*, *Notch3* and *Gata-3*, indicating that there was a basal to luminal lineage switch in these basal cells. These data demonstrate that *Bcl11b* maintains basal identity and suppresses the luminal lineage. *Bcl11b* also promotes the basal lineage as over-expression of *Bcl11b* was sufficient to induce the expression of basal cell specific genes such as *CK14*, *p63* and *SMA* in KIM2 mammary epithelial cells. Additionally, increased levels of *Bcl11b* in luminal progenitors were associated with a dramatic

reduction in luminal mammary colony-forming-cell (Ma-CFCs) capabilities in *Stat6*$^{-/-}$ knockout mice. Taken together, *Bcl11b* is critical for the maintenance of basal identity in the mammary gland.

These results clearly demonstrate that both *Bcl11a* and *Bcl11b* are critical regulators of lineage commitment in the mammary gland and that levels of each of these transcription factors are crucial for ensuring proper mammary development and functionality.

7.1.5 *Bcl11* genes are connected to networks of mammary regulators

Mammary gland development is a highly co-ordinated series of events that is regulated by both systemic hormones and local growth factors (Hennighausen and Robinson, 1998, 2005). Recent studies have shown that transcription factors are also critical for lineage commitment in the mammary epithelium (Asselin-Labat et al., 2007; Khaled et al., 2007; Kouros-Mehr et al., 2006; Miyoshi et al., 2002). Developmentally important pathways such as Notch, Wnt and Hedgehog (Hh) have critical roles in the mammary gland. Our lab has previously shown that the canonical Notch signalling pathway was over-expressed in the *Bcl11a*-deficient T cell leukaemia (Liu et al., 2003b). Consistent with high levels of *Notch1* in the *Bcl11a* mutant T-cell leukemia, loss of *Bcl11a* in the virgin and lactation mammary glands resulted in increased *Notch1* expression. Activation of Notch signalling was confirmed by up-regulation of downstream targets of Notch signalling such as hairy-enhancer of split genes (*Hes*) in the *Bcl11a*-deficient lactation glands. In the *Bcl11a*-deficient lactation gland, Notch activation might be responsible for the activation of Stat3, and loss of phosphorylated Stat5 (p-Stat5) based on recent biochemical studies (Kamakura et al., 2004; Nie et al., 2008). Interestingly, in the *Bcl11a*-deficient mammary epithelium in both the virgin and lactation glands, *Bcl11b* expression was significantly increased. The increased level of *Bcl11b* could be partially responsible for the phenotypes in the *Bcl11a* mutant glands. This could be investigated using *Bcl11a*- and/or *Bcl11b*-over-expressing mice.

It has been suggested in a previous study that Notch1 and Notch3 function in a similar way as over-expression of Notch1 and Notch3 using the mouse mammary tumour virus (MMTV) provirus results in impaired ductal and lobulo-alveolar development

during pregnancy (Hu et al., 2006). However, the results herein clearly demonstrate that Notch1 and Notch3 do not have similar functions in the lactation gland. The *Bcl11a*–deficient lactation mammary epithelium did not have Notch3 expression. In addition, a dramatic reduction in the transcript levels of *Jagged1*, *Delta-like-ligand 4* (*Dll4*) and *Hes7* was found in the *Bcl11a*-deficient lactation gland. Whether gene expression changes in these Notch components are functionally related or directly caused the *Bcl11a* mutant phenotypes remained to be determined. Nevertheless, these data revealed that not all the Notch pathway components behave in a similar way, which is consistent with recent studies (Raouf et al., 2008; Stylianou et al., 2006; Yamaguchi et al., 2008).

The JAK-Stat pathway plays essential roles during mammary gland development, particularly during gestation and lactation. Stat5 and Stat3 display reciprocal functions during gestation, lactation and involution. Stat5 is a key mediator of prolactin signalling and plays critical roles in the differentiation of lobulo-alveolar cells during gestation and lactation (Cui et al., 2004; Liu et al., 1997). Conversely, activation of Stat3 is a critical event during involution, resulting in the activation of the apoptotic pathway to facilitate the removal of redundant alveolar cells (Chapman et al., 1999). Hence, the physiological levels of phosphorylated Stat5 and Stat3 are critical determinants of the functional state of the mammary gland. As discussed in Chapter 6, deletion of *Bcl11a* resulted in premature onset of involution in the lactation gland where a loss of p-Stat5 positive cells and a dramatic increase in p-Stat3 positive cells were observed. As deletion of *Bcl11a* also led to dysregulation of Notch and JAK-Stat signalling pathways, and Hes binding Stat3 was shown to mediate crosstalk between Notch and JAK-Stat signalling (Kamakura et al., 2004), therefore it is possible that Bcl11a regulate the levels of Stat5 and Stat3 via the Notch signalling pathway that in turns determine the functional state of the mammary gland.

Gata-3 has recently been shown to be essential for the maintenance of luminal cell fate in the mammary gland (Asselin-Labat et al., 2007; Kouros-Mehr et al., 2006). Conditional loss of *Gata-3* in the mammary epithelium resulted in loss of luminal cells due to the block in differentiation of the luminal progenitors. However, the results herein indicate that loss of *Bcl11a* in the virgin gland resulted in an increase in $ER\alpha^+$ luminal cells concomitant with a loss of *Gata-3*$^+$ luminal cells, indicating a selective loss of $ER\alpha^-$

luminal cells. Additionally, conditional deletion of *Bcl11a* in the lactation glands did not affect levels of Gata-3, suggesting that Bcl11a and Gata-3 are likely to function in distinct epithelial populations. Taken together, these results suggest that Bcl11a but not Gata-3 is the key player in alveolar specification and maintenance.

7.1.6 *Bcl11* genes in mammary progenitors and lineage commitment

Bcl11a and *Bcl11b* were among the earliest genes that were specifically expressed in the mammary placodes, and were only expressed in small numbers of epithelial cells in the virgin gland. Loss of either gene in the virgin gland caused profound defects in the mammary epithelium. These results suggest that at least some of the *Bcl11a*- or *Bcl11b*-expressing cells might represent the multi-potent progenitors or even the mammary stem cells (Figure 7.1). Loss-of-function analysis in the virgin gland showed that *Bcl11a* was required for maintenance of both luminal and basal lineages, as deletion of *Bcl11a* resulted in down-regulated expression of both luminal and basal markers. In addition, *Bcl11a* was also required to maintain luminal progenitors; in particular, the alveolar progenitors as loss of *Bcl11a* in the virgin and lactation glands resulted in a dramatic reduction in *Elf5* levels (Figure 7.1). *Bcl11b*, on the other hand maintained basal identity, suppressed luminal cell differentiation, and also prevented precocious alveologenesis in the virgin gland. Loss- and gain-of-function analyses demonstrated that *Bcl11b* also promoted the basal lineage. *Bcl11a* continued to be necessary in the lactation gland where it maintained terminally differentiated luminal secretory cell fate. The phenotypes in the *Bcl11a*-deficient lactation gland led me to hypothesize that *Bcl11a* maintains secretory luminal cell fate by positively regulating Notch3 and negatively regulating Notch1 and Lif (Figure 6.14). Loss of *Bcl11a* would thus lead to Notch3 deficiency, *Lif* expression and activation of the canonical Notch signalling, resulting in Stat3-mediated apoptosis.

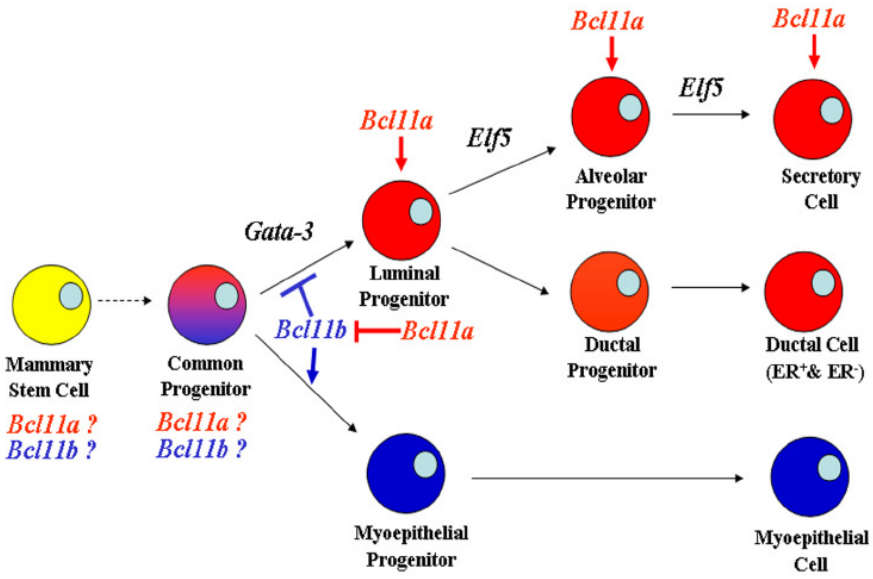


Figure 7.1. *Bcl11a* and *Bcl11b* are critical regulators for the development of mammary epithelial hierarchy. A working model of *Bcl11* genes in mammary epithelial lineage specification and maintenance. *Bcl11a* and *Bcl11b* could be expressed in the proposed multi-potent mammary progenitors or stem cells as deletion of these two genes in the virgin gland results in dramatic phenotypes. *Bcl11a* is also essential for the maintenance of the terminally differentiated lobulo-alveolar cells. *Bcl11b* promotes the basal lineage and maintains the basal identity by inhibiting the commitment of progenitors to the luminal lineage. Red – luminal cells; Blue – basal cells.

7.2 Significance

7.2.1 Novel roles of *Bcl11* genes in lineage commitment

The mammary phenotypes observed in the *Bcl11* knockout mice are striking and identify *Bcl11a* and *Bcl11b* as two new key transcription factors implicated in mammary development and lineage commitment. The data presented here indicate that Bcl11a and Bcl11b, together with Gata-3 and Stat6 are transcription factors that are required for mammary lineage commitment. Most importantly, *Bcl11a* and *Bcl11b* not only provide key functions in various epithelial progenitor compartments, they also serve as additional molecular markers for these progenitors. *Bcl11a* and *Bcl11b* are also required in lineage specification in lymphocyte development, suggesting that genetic control of mammary and lymphocyte development is conserved. Characterization of additional transcription factors implicated in lymphocyte development should uncover additional transcription factors that are critical to mammary lineage commitment.

7.2.2 Implications for tumour development

Human breast cancers have heterogeneous pathologies with diverse molecular profiles, therefore no single dominant pathway or histological presentation has emerged (Stingl and Caldas, 2007). Recent gene expression profiling using microarray analysis has offered an alternative to the classical histopathology for the classification of human breast cancers which clusters breast tumours into at least five reproducible subtypes: luminal A, luminal B, ERBB2, basal and normal-like (Perou et al., 2000; Sorlie et al., 2001). Comparing this molecular classification with the traditional histopathological analyses revealed some correlations and discrepancies between the two methods in the classification of breast tumours (Perou et al., 2000). In order to appreciate the new molecular classification, a thorough understanding of the cellular hierarchy of the normal mammary epithelium is required to establish the cellular origin of the tumours. Given the critical roles of *Bcl11a* and *Bcl11b* in the mammary gland, it is conceivable that they, or the pathways or networks that they are involved in, should have a causal role in breast cancer. Two *BCL11A* mutations have been identified in human breast cancer samples recently (Wood et al., 2007). However, no information on the histopathological or molecular subtype of these tumours is reported. The data presented in this thesis provide

evidence that *Bcl11a* is an important gene involved in breast cancer. In addition, loss of *Bcl11a* resulted in dysregulation of the Notch signalling pathway, a key signalling pathway that has been implicated in tumorigenesis, further linking *Bcl11a* and breast cancer (Efstratiadis et al., 2007).

Interestingly, luminal A tumours, which have higher expression of ERα and Gata-3 than luminal B tumours, have a better prognosis. This study has shown that loss of *Bcl11a* led to a relative increment in ERα$^+$ luminal cells concomitant with an absence of Gata-3$^+$ luminal cells. As *Bcl11a* is probably not expressed and functional in the Gata-3$^+$ and ERα$^+$ cells, it would be interesting to determine the expression status of *Bcl11a* in luminal A tumours. In addition, the molecular classification of human breast cancer samples with *BCL11A* mutations can be studied and the differences in the gene expression profiles compared to establish if this represents a distinct class of breast tumours from luminal A and luminal B. Additionally, using gene expression profiling to classify breast cancer tumours can theoretically provide insights into the ontological history of the tumour. Molecular profiling could also determine if the breast tumours arising from *BCL11A* mutations originated from mammary progenitor cells as expression of *Bcl11a* was detected in luminal progenitors. At present, no known *BCL11B* mutations are implicated in human breast tumorigenesis. As *Bcl11b* is predominantly expressed in basal cells, it would be interesting to investigate expression status of *BCL11B* in the basal-like breast cancer subtypes which are ERα$^-$ and PR$^-$ (progesterone receptor) and generally more aggressive and resistant to therapies.

7.3 Future experiments

The results presented in this thesis demonstrate that *Bcl11a* and *Bcl11b* are critical regulators of lineage commitment in the mammary epithelium. Both genes are required for normal mammary gland development and *Bcl11a* is also essential for maintenance of differentiated luminal secretory cell fate. Hence regulating the levels of both genes is crucial to the development of a functional mammary gland. The expression and phenotypic analyses also suggest that both of these genes are expressed in mammary progenitors and possibly in mammary stem cells (MaSCs). Future work is required to address the roles of *Bcl11* genes in the mammary stem and/or progenitor cells. In

addition, the molecular mechanisms underlying the phenotypes in the *Bcl11* knockout mice involved dysregulation of the Notch and JAK-Stat pathways. Future experiments are required to confirm if various components of the Notch and JAK-Stat pathways are direct targets of Bcl11 transcription factors and to find out whether other molecular pathways also play a role.

7.3.1 *Bcl11* genes and mammary stem/progenitor cells

One of the major findings in this thesis is that *Bcl11a* is expressed in luminal progenitors (both $Sca1^+$ putative ductal and $Sca1^-$ putative alveolar progenitors) and a small percentage of basal cells (where the MaSCs are believed to be located). To further confirm the expression of *Bcl11a* in mammary stem and/or progenitor cells, *Bcl11a*-expressing cells from $Bcl11a^{lacZ/+}$ mice can be sorted and transplanted in limiting dilutions into cleared mammary fat pads. The transplanted mammary gland can then be analysed to assess the contribution of *Bcl11a*-expressing cells to various mammary lineages. Expression of *Bcl11b* was first detected in the milk line of the embryos and predominantly in the basal fractions of the mature gland. These expression data suggest that *Bcl11b* could also be expressed in mammary stem/progenitor cells. To address this question, *Bcl11b*-expressing cells from $Bcl11b^{lacZ/+}$ mice can be sorted and transplanted in limiting dilutions into cleared mammary fat pads. The transplanted mammary gland can then be analysed to determine contribution of *Bcl11b*-expressing cells to various mammary lineages.

One hypothesis is that mammary progenitor cells that co-express both *Bcl11a* and *Bcl11b* have bi-potent capabilities (both luminal and basal). To test this hypothesis, each *Bcl11* gene can be tagged with a fluorescent protein [for example green fluorescent protein (GFP) and red fluorescent protein (RFP)]. Transgenic mice containing these tagged alleles can be crossed to obtain double heterozygotes. Finally, the *Bcl11a* and *Bcl11b* double positive mammary cells (GFP and RFP double positive cells) can be sorted using FACS and transplanted at limiting dilutions into cleared fat pad. Subsequently, the transplanted mammary gland can be analysed to determine the contribution of *Bcl11a* and *Bcl11b* double positive epithelial cells to the mammary lineages. Technically, the fluorescent proteins are much better than FDG for sorting out

live cells. These mice are therefore useful for detecting *Bcl11* expression in progenitor compartments.

To further address the roles of *Bcl11* genes in mammary stem/progenitor cells, the *Bcl11* conditional knockout mice can be crossed with the transgenic mice expressing Cre recombinase under the control of the keratin 5 (K5)-promoter that directs Cre expression to the basal cells of mammary epithelial (where mammary stem/progenitor cells reside) (Taddei et al., 2008). Mammary glands from *K5-Cre; Bcl11*$^{flox/flox}$ females can then be analysed for effects of Cre-mediated excision of *Bcl11* genes in the basal layer of mammary gland.

The *BLG-Cre; Bcl11*$^{flox/flox}$ females described herein, could also potentially be used to study the effects of *Bcl11* deletion in mammary luminal progenitors. I anticipate the BLG-Cre transgene to be expressed at sufficient levels in some non-secretory luminal cells, including the alveolar progenitors in the mice, to allow us to assay for the effects of *Bcl11a* deletion in the luminal progenitors. Therefore *BLG-Cre; Bcl11*$^{flox/flox}$ females that have undergone one round of pregnancy-lactation-involution development cycle would have *Bcl11* genes deleted in their luminal cells (differentiated and possibly progenitor cells). The effects of *Bcl11* deletion in luminal progenitors could then be assessed by *in vitro* mammary colony-forming cell assay.

Over-expression vectors of *Bcl11a* and *Bcl11b* described in this study could also be used to over-express *Bcl11* genes in mammary progenitors and these cells can then be transplanted into cleared mammary fat pad to assess the effects of over-expression.

7.3.2 Mammary fat pad transplantation of *Bcl11*-deficient mammary cells

The Cre recombinase in the *Cre-ERT2; Bcl11*$^{flox/flox}$ mice used in this study is driven by the *Rosa26* promoter, hence Cre recombinase is expressed in other cell types in addition to mammary epithelium (Hameyer et al., 2007). Following TAM injection to induce deletion of *Bcl11* genes, some *Cre-ERT2; Bcl11*$^{flox/flox}$ mice were found dead presumably because of the vital roles that *Bcl11* genes play in other tissues, hence hampering mammary gland analysis. Therefore, to minimise the effects of *Bcl11* deletion in other tissues, mammary epithelial cells from *Cre-ERT2; Bcl11*$^{flox/flox}$ mice could be

transplanted into cleared fat pads of wild-type mice. Deletion of *Bcl11* genes in the mammary epithelium of these transplanted mice can then be mediated by administration of tamoxifen and the mammary glands analysed for effects of loss of *Bcl11* genes. Alternatively, different populations of mammary cells such as $Sca1^+$ and $Sca1^-$ luminal progenitors can be sorted from *Cre-ERT2; Bcl11*$^{flox/flox}$ mice and Cre-mediated excision of *Bcl11* genes can be carried out *in vitro* by addition of tamoxifen. These *Bcl11*-deficient cells can then be transplanted into cleared mammary fat pads to assess the effects of deletion of *Bcl11* genes *in vivo*.

7.3.3 Mammary tumour formation in *Bcl11*-deficient glands

Mammary cells from *Cre-ERT2; Bcl11*$^{flox/flox}$ mice can be transplanted into wild-type mammary fat pad and Cre-mediated excision of *Bcl11* genes carried out before monitoring these transplanted mice for the development of mammary tumours. *Bcl11* heterozygous mice could also be crossed to other tumour-prone strains such as *p53* heterozygous mice. These double heterozygous mice can then be monitored to assess if loss of one copy of *Bcl11* gene could accelerate mammary tumour formation.

7.3.4 Microarray analysis of *Bcl11*-deficient mammary glands

The results from this thesis demonstrate that *Bcl11a* is a key transcription factor required to maintain terminal differentiated secretory luminal cell fate. Loss of *Bcl11a* in the lactation glands resulted in a lactation to involution transcriptional switch. The molecular pathways mediated by Bcl11a during lactation are still poorly understood. In order to gain insight into the effects of Bcl11a-mediated signalling on the lactation-involution switch, a genome-wide approach using microarray analysis could be performed. Gene expression profiles between lactation glands from *BLG-Cre; Bcl11a*$^{flox/flox}$, *BLG-Cre; Bcl11a*$^{flox/+}$ and wild-type mice can be compared and putative downstream targets verified by quantitative real-time PCR. Similarly, microarray analysis on sorted luminal and basal fractions from *Bcl11a* and *Bcl11b*-deficient virgin glands could also be carried out to ascertain the gene expression changes following *Bcl11* deletion and therefore identify putative targets that might be essential for maintenance of epithelial cell fate. These identified gene expression changes following deletion of *Bcl11*

genes can then be compared to those identified in this thesis and validated using quantitative real-time PCR. Components of the Notch signalling pathway would be of particular interest as dysregulation of this pathway was identified following loss of *Bcl11a*. Luciferase reporter assays for various known downstream targets of Notch signalling pathway can be performed. In addition, immunohistochemistry of *Bcl11a*-deficient glands using specific antibodies to activated Notch1 ICN should be performed to confirm activation of the Notch signalling pathways.

7.3.5 Direct targets of Bcl11 transcription factors

Putative targets of Bcl11a transcription factor have been identified in this study, including Notch1, Notch3, Lif and Stat5. Validation of these targets could be carried out using chromatin immunoprecipitation (ChIP). Next-generation sequencing technology combined with ChIP provides an efficient method to identify transcription factor binding sites on a genome-wide scale (Robertson et al., 2007). Conventional ChIP protocol would first be performed on wild-type mammary tissues using Bcl11 antibodies and the DNA fragments released after ChIP would then be sequenced using the Solexa sequencing technology. The resulting sequences obtained would be mapped back to the reference genome, where the most frequently sequenced DNA fragments formed peaks at specific genomic regions, indicating an enrichment of these DNA sites bound by the transcription factors. These DNA sequences would therefore represent the target sites directly bound by Bcl11 transcription factors.

7.3.6 Human breast cancer cell lines and tissues

There are many human breast cancer cell lines that are widely used to study signalling pathways implicated in breast cancer. These different human breast cancer cell lines are derived from different sources and genomic approaches have demonstrated that the gene expression profiles of these cell lines can be clustered into luminal and basal subtypes of breast cancers (Sorlie et al., 2001). Given the results obtained in this study, it would be interesting to determine the expression of *BCL11* genes in these different cell lines and determine if different subtypes of breast cancer cell lines exhibit differential expression of *BCL11* genes. Correlation of *BCL11* expression with the various subtypes

of cell lines could then be made and this would provide an indication whether expression of *BCL11* genes could be used as prognostic markers for different subtypes of human breast cancers. Similarly, tumour and normal breast tissue arrays can be used to study tumour-specific cellular localization of BCL11 expression, allowing the comparison of BCL11 expression between normal and pathological samples.

7.3.7 Quantification and statistical analysis

Quantification of the reduction in secondary branches following loss of *Bcl11a* in the virgin glands should be performed using microscopic images of whole mount carmine alum stained glands. Glands from three independent *Bcl11a*-deficient glands should be assessed and the numbers analysed using the Mann-Whitney U test. This statistical analysis is a non-parametric test for assessing whether two samples of observations come from the same distribution. This statistical analysis is important so that a conclusion can be drawn on the effect of changes in the values. Quantification of the changes in levels of transcripts (RT-PCR) and proteins (Immunoblot) should also be quantified using available softwares such as Image J (http://rsb.info.nih.gov/ij/). In addition, independent appropriate controls should be carried out for each set of experiments and it is unacceptable to use the same control in separate experiments.

7.4 Conclusions

In this study, two *Bcl11-lacZ*-tagged mice were generated using gene targeting technologies. Characterization of the spatial and cellular expression patterns of *Bcl11* genes using these *lacZ*-tagged mice revealed that these two genes exhibited unique differential expression patterns in the mammary gland. Subsequently, novel roles of both Bcl11 transcription factors in mammary development were demonstrated using loss- and gain-of-function analyses. While *Bcl11a* is important for normal mammary development and maintenance of terminally differentiated luminal secretory cell fate, *Bcl11b* is essential for the maintenance of basal cell fate. This study has identified both Bcl11a and Bcl11b as critical regulators of lineage commitment in the mammary epithelium and demonstrated that transcription factors that specify lymphoid lineages play important roles in mammary cell fate decisions.

REFERENCES

Abell, K., Bilancio, A., Clarkson, R.W., Tiffen, P.G., Altaparmakov, A.I., Burdon, T.G., Asano, T., Vanhaesebroeck, B., and Watson, C.J. (2005). Stat3-induced apoptosis requires a molecular switch in PI(3)K subunit composition. Nat Cell Biol *7*, 392-398.

Alvi, A.J., Clayton, H., Joshi, C., Enver, T., Ashworth, A., Vivanco, M.M., Dale, T.C., and Smalley, M.J. (2003). Functional and molecular characterisation of mammary side population cells. Breast Cancer Res *5*, R1-8.

Amsen, D., Antov, A., Jankovic, D., Sher, A., Radtke, F., Souabni, A., Busslinger, M., McCright, B., Gridley, T., and Flavell, R.A. (2007). Direct regulation of Gata3 expression determines the T helper differentiation potential of Notch. Immunity *27*, 89-99.

Angrand, P.O., Daigle, N., van der Hoeven, F., Scholer, H.R., and Stewart, A.F. (1999). Simplified generation of targeting constructs using ET recombination. Nucleic Acids Res *27*, e16.

Ansel, K.M., Djuretic, I., Tanasa, B., and Rao, A. (2006). Regulation of Th2 differentiation and Il4 locus accessibility. Annual review of immunology *24*, 607-656.

Artavanis-Tsakonas, S., Rand, M.D., and Lake, R.J. (1999). Notch signaling: cell fate control and signal integration in development. Science (New York, NY *284*, 770-776.

Asselin-Labat, M.L., Shackleton, M., Stingl, J., Vaillant, F., Forrest, N.C., Eaves, C.J., Visvader, J.E., and Lindeman, G.J. (2006). Steroid hormone receptor status of mouse mammary stem cells. Journal of the National Cancer Institute *98*, 1011-1014.

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

Austin, C.P., Battey, J.F., Bradley, A., Bucan, M., Capecchi, M., Collins, F.S., Dove, W.F., Duyk, G., Dymecki, S., Eppig, J.T.*, et al.* (2004). The knockout mouse project. Nature genetics *36*, 921-924.

Auwerx, J., Avner, P., Baldock, R., Ballabio, A., Balling, R., Barbacid, M., Berns, A., Bradley, A., Brown, S., Carmeliet, P.*, et al.* (2004). The European dimension for the mouse genome mutagenesis program. Nature genetics *36*, 925-927.

Avram, D., Fields, A., Pretty On Top, K., Nevrivy, D.J., Ishmael, J.E., and Leid, M. (2000). Isolation of a novel family of C(2)H(2) zinc finger proteins implicated in transcriptional repression mediated by chicken ovalbumin upstream promoter transcription factor (COUP-TF) orphan nuclear receptors. J Biol Chem *275*, 10315-10322.

Avram, D., Fields, A., Senawong, T., Topark-Ngarm, A., and Leid, M. (2002). COUP-TF (chicken ovalbumin upstream promoter transcription factor)-interacting protein 1 (CTIP1) is a sequence-specific DNA binding protein. Biochem J *368*, 555-563.

Behrens, J., von Kries, J.P., Kuhl, M., Bruhn, L., Wedlich, D., Grosschedl, R., and Birchmeier, W. (1996). Functional interaction of beta-catenin with the transcription factor LEF-1. Nature *382*, 638-642.

Bocchinfuso, W.P., and Korach, K.S. (1997). Mammary gland development and tumorigenesis in estrogen receptor knockout mice. J Mammary Gland Biol Neoplasia *2*, 323-334.

Booth, B.W., and Smith, G.H. (2006). Estrogen receptor-alpha and progesterone receptor are expressed in label-retaining mammary epithelial cells that divide asymmetrically and retain their template DNA strands. Breast Cancer Res *8*, R49.

Boras-Granic, K., Chang, H., Grosschedl, R., and Hamel, P.A. (2006). Lef1 is required for the transition of Wnt signaling from mesenchymal to epithelial cells in the mouse embryonic mammary gland. Developmental biology *295*, 219-231.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Bradley, A., and Liu, P. (1996). Target practice in transgenics. Nature genetics *14*, 121-123.

Brinster, R.L. (1974). The effect of cells transferred into the mouse blastocyst on subsequent development. The Journal of experimental medicine *140*, 1049-1056.

Brinster, R.L., Chen, H.Y., Trumbauer, M., Senear, A.W., Warren, R., and Palmiter, R.D. (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell *27*, 223-231.

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

Bryder, D., Rossi, D.J., and Weissman, I.L. (2006). Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. The American journal of pathology *169*, 338-346.

Buchholz, F., Angrand, P.O., and Stewart, A.F. (1998). Improved properties of FLP recombinase evolved by cycling mutagenesis. Nature biotechnology *16*, 657-662.

Buono, K.D., Robinson, G.W., Martin, C., Shi, S., Stanley, P., Tanigaki, K., Honjo, T., and Hennighausen, L. (2006). The canonical Notch/RBP-J signaling pathway controls the balance of cell lineages in mammary epithelium during pregnancy. Developmental biology *293*, 565-580.

Cantor, A.B., and Orkin, S.H. (2001). Hematopoietic development: a balancing act. Current opinion in genetics & development *11*, 513-519.

Capecchi, M.R. (1989). Altering the genome by homologous recombination. Science (New York, NY *244*, 1288-1292.

Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., *et al.* (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell *122*, 33-43.

Chan, W., Costantino, N., Li, R., Lee, S.C., Su, Q., Melvin, D., Court, D.L., and Liu, P. (2007). A recombineering based approach for high-throughput conditional knockout targeting vector construction. Nucleic Acids Res *35*, e64.

Chapman, R.S., Lourenco, P., Tonner, E., Flint, D., Selbert, S., Takeda, K., Akira, S., Clarke, A.R., and Watson, C.J. (2000). The role of Stat3 in apoptosis and mammary gland involution. Conditional deletion of Stat3. Adv Exp Med Biol *480*, 129-138.

Chapman, R.S., Lourenco, P.C., Tonner, E., Flint, D.J., Selbert, S., Takeda, K., Akira, S., Clarke, A.R., and Watson, C.J. (1999). Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. Genes Dev *13*, 2604-2616.

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

Cismasiu, V.B., Adamo, K., Gecewicz, J., Duque, J., Lin, Q., and Avram, D. (2005). BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter. Oncogene *24*, 6753-6764.

Clarke, R.B., Howell, A., Potten, C.S., and Anderson, E. (1997). Dissociation between steroid receptor expression and cell proliferation in the human breast. Cancer research *57*, 4987-4991.

Collins, F.S., Rossant, J., and Wurst, W. (2007). A mouse for all reasons. Cell *128*, 9-13.

Copeland, N.G., Jenkins, N.A., and Court, D.L. (2001). Recombineering: a powerful new tool for mouse functional genomics. Nat Rev Genet *2*, 769-779.

Costa, R.H., Kalinichenko, V.V., and Lim, L. (2001). Transcription factors in mouse lung development and function. American journal of physiology *280*, L823-838.

Covarrubias, L., Cervantes, L., Covarrubias, A., Soberon, X., Vichido, I., Blanco, A., Kupersztoch-Portnoy, Y.M., and Bolivar, F. (1981). Construction and characterization of new cloning vehicles. V. Mobilization and coding properties of pBR322 and several deletion derivatives including pBR327 and pBR328. Gene *13*, 25-35.

Cui, Y., Riedlinger, G., Miyoshi, K., Tang, W., Li, C., Deng, C.X., Robinson, G.W., and Hennighausen, L. (2004). Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation. Mol Cell Biol *24*, 8037-8047.

Cunha, G.R., and Hom, Y.K. (1996). Role of mesenchymal-epithelial interactions in mammary gland development. J Mammary Gland Biol Neoplasia *1*, 21-35.

Daniel, C.W., Silberstein, G.B., and Strickland, P. (1987). Direct action of 17 beta-estradiol on mouse mammary ducts analyzed by sustained release implants and steroid autoradiography. Cancer research *47*, 6052-6057.

Darnell, J.E., Jr. (1997). STATs and gene regulation. Science (New York, NY *277*, 1630-1635.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America *97*, 6640-6645.

Datta, S., Costantino, N., and Court, D.L. (2006). A set of recombineering plasmids for gram-negative bacteria. Gene *379*, 109-115.

Davis, M.M. (2007). Blimp-1 over Budapest. Nat Immunol *8*, 445-447.

Deng, C., and Capecchi, M.R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol *12*, 3365-3371.

Deome, K.B., Faulkin, L.J., Jr., Bern, H.A., and Blair, P.B. (1959). Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer research *19*, 515-520.

Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery, C.A., Jr., Butel, J.S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature *356*, 215-221.

Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J., and Wicha, M.S. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev *17*, 1253-1270.

Dontu, G., El-Ashry, D., and Wicha, M.S. (2004). Breast cancer, stem/progenitor cells and the estrogen receptor. Trends in endocrinology and metabolism: TEM *15*, 193-197.

Eblaghie, M.C., Song, S.J., Kim, J.Y., Akita, K., Tickle, C., and Jung, H.S. (2004). Interactions between FGF and Wnt signals and Tbx3 gene expression in mammary gland initiation in mouse embryos. Journal of anatomy *205*, 1-13.

Eeckhoute, J., Keeton, E.K., Lupien, M., Krum, S.A., Carroll, J.S., and Brown, M. (2007). Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. Cancer research *67*, 6477-6483.

Efstratiadis, A., Szabolcs, M., and Klinakis, A. (2007). Notch, Myc and breast cancer. Cell Cycle *6*, 418-429.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fang, T.C., Yashiro-Ohtani, Y., Del Bianco, C., Knoblock, D.M., Blacklow, S.C., and Pear, W.S. (2007). Notch directly regulates Gata3 expression during T helper 2 cell differentiation. Immunity *27*, 100-110.

Feng, Y., Manka, D., Wagner, K.U., and Khan, S.A. (2007). Estrogen receptor-alpha expression in the mammary epithelium is required for ductal and alveolar morphogenesis in mice. Proceedings of the National Academy of Sciences of the United States of America *104*, 14718-14723.

Feng, Z., Marti, A., Jehn, B., Altermatt, H.J., Chicaiza, G., and Jaggi, R. (1995). Glucocorticoid and progesterone inhibit involution and programmed cell death in the mouse mammary gland. The Journal of cell biology *131*, 1095-1103.

Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D.I., Enver, T., Ley, T.J., and Groudine, M. (1995). Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. Genes Dev *9*, 2203-2213.

Foley, J., Dann, P., Hong, J., Cosgrove, J., Dreyer, B., Rimm, D., Dunbar, M., Philbrick, W., and Wysolmerski, J. (2001). Parathyroid hormone-related protein maintains mammary epithelial fate and triggers nipple skin differentiation during embryonic breast development. Development (Cambridge, England) *128*, 513-525.

Funaki, M., Katagiri, H., Inukai, K., Kikuchi, M., and Asano, T. (2000). Structure and function of phosphatidylinositol-3,4 kinase. Cellular signalling *12*, 135-142.

Gallahan, D., Jhappan, C., Robinson, G., Hennighausen, L., Sharp, R., Kordon, E., Callahan, R., Merlino, G., and Smith, G.H. (1996). Expression of a truncated Int3 gene in developing secretory mammary epithelium specifically retards lobular differentiation resulting in tumorigenesis. Cancer research *56*, 1775-1785.

Glaser, S., Anastassiadis, K., and Stewart, A.F. (2005). Current issues in mouse genome engineering. Nature genetics *37*, 1187-1193.

Godin, I., and Cumano, A. (2002). The hare and the tortoise: an embryonic haematopoietic race. Nat Rev Immunol *2*, 593-604.

Godin, I., Dieterlen-Lievre, F., and Cumano, A. (1995). Emergence of multipotent hemopoietic cells in the yolk sac and paraaortic splanchnopleura in mouse embryos,

beginning at 8.5 days postcoitus. Proceedings of the National Academy of Sciences of the United States of America *92*, 773-777.

Golonzhka, O., Leid, M., Indra, G., and Indra, A.K. (2007). Expression of COUP-TF-interacting protein 2 (CTIP2) in mouse skin during development and in adulthood. Gene Expr Patterns *7*, 754-760.

Gordon, K.E., Binas, B., Chapman, R.S., Kurian, K.M., Clarkson, R.W., Clark, A.J., Lane, E.B., and Watson, C.J. (2000). A novel cell culture model for studying differentiation and apoptosis in the mouse mammary gland. Breast Cancer Res *2*, 222-235.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proceedings of the National Academy of Sciences of the United States of America *89*, 5547-5551.

Gunnersen, J.M., Augustine, C., Spirkoska, V., Kim, M., Brown, M., and Tan, S.S. (2002). Global analysis of gene expression patterns in developing mouse neocortex using serial analysis of gene expression. Molecular and cellular neurosciences *19*, 560-573.

Hameyer, D., Loonstra, A., Eshkind, L., Schmitt, S., Antunes, C., Groen, A., Bindels, E., Jonkers, J., Krimpenfort, P., Meuwissen, R.*, et al.* (2007). Toxicity of ligand-dependent Cre recombinases and generation of a conditional Cre deleter mouse allowing mosaic recombination in peripheral tissues. Physiol Genomics *31*, 32-41.

Hanks, M., Wurst, W., Anson-Cartwright, L., Auerbach, A.B., and Joyner, A.L. (1995). Rescue of the En-1 mutant phenotype by replacement of En-1 with En-2. Science (New York, NY *269*, 679-682.

Hardy, R.R., and Hayakawa, K. (2001). B cell development pathways. Annual review of immunology *19*, 595-621.

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

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

Heinzel, T., Lavinsky, R.M., Mullen, T.M., Soderstrom, M., Laherty, C.D., Torchia, J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R.*, et al.* (1997). A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature *387*, 43-48.

Hennighausen, L., and Robinson, G.W. (1998). Think globally, act locally: the making of a mouse mammary gland. Genes Dev *12*, 449-455.

Hennighausen, L., and Robinson, G.W. (2005). Information networks in the mammary gland. Nature reviews *6*, 715-725.

Hennighausen, L., and Robinson, G.W. (2008). Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. Genes Dev *22*, 711-721.

Hens, J.R., and Wysolmerski, J.J. (2005). Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland. Breast Cancer Res *7*, 220-224.

Ho, I.C., and Pai, S.Y. (2007). GATA-3 - not just for Th2 cells anymore. Cellular & molecular immunology *4*, 15-29.

Hooper, J.E., and Scott, M.P. (2005). Communicating with Hedgehogs. Nature reviews *6*, 306-317.

Horseman, N.D., Zhao, W., Montecino-Rodriguez, E., Tanaka, M., Nakashima, K., Engle, S.J., Smith, F., Markoff, E., and Dorshkind, K. (1997). Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. The EMBO journal *16*, 6926-6935.

Howard, B., Panchal, H., McCarthy, A., and Ashworth, A. (2005). Identification of the scaramanga gene implicates Neuregulin3 in mammary gland specification. Genes Dev *19*, 2078-2090.

Hu, C., Dievart, A., Lupien, M., Calvo, E., Tremblay, G., and Jolicoeur, P. (2006). Overexpression of activated murine Notch1 and Notch3 in transgenic mice blocks mammary gland development and induces mammary tumors. The American journal of pathology *168*, 973-990.

Humphreys, R.C., Bierie, B., Zhao, L., Raz, R., Levy, D., and Hennighausen, L. (2002). Deletion of Stat3 blocks mammary gland involution and extends functional competence of the secretory epithelium in the absence of lactogenic stimuli. Endocrinology *143*, 3641-3650.

Humphreys, R.C., Krajewska, M., Krnacik, S., Jaeger, R., Weiher, H., Krajewski, S., Reed, J.C., and Rosen, J.M. (1996). Apoptosis in the terminal endbud of the murine mammary gland: a mechanism of ductal morphogenesis. Development (Cambridge, England) *122*, 4013-4022.

Jhappan, C., Gallahan, D., Stahle, C., Chu, E., Smith, G.H., Merlino, G., and Callahan, R. (1992). Expression of an activated Notch-related int-3 transgene interferes with cell differentiation and induces neoplastic transformation in mammary and salivary glands. Genes Dev *6*, 345-355.

Johnson, R.L., and Tabin, C.J. (1997). Molecular models for vertebrate limb development. Cell *90*, 979-990.

Jonkers, J., and Berns, A. (1996). Retroviral insertional mutagenesis as a strategy to identify cancer genes. Biochimica et biophysica acta *1287*, 29-57.

Kamakura, S., Oishi, K., Yoshimatsu, T., Nakafuku, M., Masuyama, N., and Gotoh, Y. (2004). Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signalling. Nat Cell Biol *6*, 547-554.

Kamimura, K., Ohi, H., Kubota, T., Okazuka, K., Yoshikai, Y., Wakabayashi, Y., Aoyagi, Y., Mishima, Y., and Kominami, R. (2007). Haploinsufficiency of Bcl11b for suppression of lymphomagenesis and thymocyte development. Biochemical and biophysical research communications *355*, 538-542.

Katso, R., Okkenhaug, K., Ahmadi, K., White, S., Timms, J., and Waterfield, M.D. (2001). Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer. Annual review of cell and developmental biology *17*, 615-675.

Kaufman, C.K., Zhou, P., Pasolli, H.A., Rendl, M., Bolotin, D., Lim, K.C., Dai, X., Alegre, M.L., and Fuchs, E. (2003). GATA-3: an unexpected regulator of cell lineage determination in skin. Genes Dev *17*, 2108-2122.

Kazansky, A.V., Raught, B., Lindsey, S.M., Wang, Y.F., and Rosen, J.M. (1995). Regulation of mammary gland factor/Stat5a during mammary gland development. Mol Endocrinol *9*, 1598-1609.

Kerr, J.F., Wyllie, A.H., and Currie, A.R. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. British journal of cancer *26*, 239-257.

Khaled, W.T., Read, E.K., Nicholson, S.E., Baxter, F.O., Brennan, A.J., Came, P.J., Sprigg, N., McKenzie, A.N., and Watson, C.J. (2007). The IL-4/IL-13/Stat6 signalling pathway promotes luminal mammary epithelial cell development. Development (Cambridge, England) *134*, 2739-2750.

Kleinsmith, L.J., and Pierce, G.B., Jr. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. Cancer research *24*, 1544-1551.

Koller, B.H., and Smithies, O. (1989). Inactivating the beta 2-microglobulin locus in mouse embryonic stem cells by homologous recombination. Proceedings of the National Academy of Sciences of the United States of America *86*, 8932-8935.

Korach, K.S. (2000). Estrogen receptor knock-out mice: molecular and endocrine phenotypes. Journal of the Society for Gynecologic Investigation *7*, S16-17.

Kordon, E.C., and Smith, G.H. (1998). An entire functional mammary gland may comprise the progeny from a single cell. Development (Cambridge, England) *125*, 1921-1930.

Kouros-Mehr, H., Kim, J.W., Bechis, S.K., and Werb, Z. (2008). GATA-3 and the regulation of the mammary luminal cell fate. Current opinion in cell biology *20*, 164-170.

Kouros-Mehr, H., Slorach, E.M., Sternlicht, M.D., and Werb, Z. (2006). GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell *127*, 1041-1055.

Kratochwil, K., and Schwartz, P. (1976). Tissue interaction in androgen response of embryonic mammary rudiment of mouse: identification of target tissue for testosterone. Proceedings of the National Academy of Sciences of the United States of America *73*, 4041-4044.

Kreitzer, A.C., and Regehr, W.G. (2001). Retrograde inhibition of presynaptic calcium influx by endogenous cannabinoids at excitatory synapses onto Purkinje cells. Neuron *29*, 717-727.

Kritikou, E.A., Sharkey, A., Abell, K., Came, P.J., Anderson, E., Clarkson, R.W., and Watson, C.J. (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*, 3459-3468.

Kubo, M., Hanada, T., and Yoshimura, A. (2003). Suppressors of cytokine signaling and immunity. Nat Immunol *4*, 1169-1176.

Lafontaine, D., and Tollervey, D. (1996). One-step PCR mediated strategy for the construction of conditionally expressed and epitope tagged yeast proteins. Nucleic Acids Res *24*, 3469-3471.

Lakso, M., Sauer, B., Mosinger, B., Jr., Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L., and Westphal, H. (1992). Targeted oncogene activation by site-specific recombination in transgenic mice. Proceedings of the National Academy of Sciences of the United States of America *89*, 6232-6236.

LaMarca, H.L., and Rosen, J.M. (2007). Estrogen regulation of mammary gland development and breast cancer: amphiregulin takes center stage. Breast Cancer Res *9*, 304.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lee, E.C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D.A., Court, D.L., Jenkins, N.A., and Copeland, N.G. (2001). A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics *73*, 56-65.

Leid, M., Ishmael, J.E., Avram, D., Shepherd, D., Fraulob, V., and Dolle, P. (2004). CTIP1 and CTIP2 are differentially expressed during mouse embryogenesis. Gene Expr Patterns *4*, 733-739.

Li, M., Liu, X., Robinson, G., Bar-Peled, U., Wagner, K.U., Young, W.S., Hennighausen, L., and Furth, P.A. (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*, 3425-3430.

Lim, K.C., Lakshmanan, G., Crawford, S.E., Gu, Y., Grosveld, F., and Engel, J.D. (2000). Gata3 loss leads to embryonic lethality due to noradrenaline deficiency of the sympathetic nervous system. Nature genetics *25*, 209-212.

Lin, T.P. (1966). Microinjection of mouse eggs. Science (New York, NY *151*, 333-337.

Lindeman, G.J., and Visvader, J.E. (2006). Shedding light on mammary stem cells and tumorigenesis. Cell Cycle *5*, 671-672.

Lindeman, G.J., Wittlin, S., Lada, H., Naylor, M.J., Santamaria, M., Zhang, J.G., Starr, R., Hilton, D.J., Alexander, W.S., Ormandy, C.J.*, et al.* (2001). SOCS1 deficiency results in accelerated mammary gland development and rescues lactation in prolactin receptor-deficient mice. Genes Dev *15*, 1631-1636.

Liu, H., Ippolito, G.C., Wall, J.K., Niu, T., Probst, L., Lee, B.S., Pulford, K., Banham, A.H., Stockwin, L., Shaffer, A.L.*, et al.* (2006). Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. Mol Cancer *5*, 18.

Liu, P., Jenkins, N.A., and Copeland, N.G. (2003a). A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res *13*, 476-484.

Liu, P., Keller, J.R., Ortiz, M., Tessarollo, L., Rachel, R.A., Nakamura, T., Jenkins, N.A., and Copeland, N.G. (2003b). Bcl11a is essential for normal lymphoid development. Nat Immunol *4*, 525-532.

Liu, X., Gallego, M.I., Smith, G.H., Robinson, G.W., and Hennighausen, L. (1998). Functional rescue of Stat5a-null mammary tissue through the activation of compensating signals including Stat5b. Cell Growth Differ *9*, 795-803.

Liu, X., Robinson, G.W., and Hennighausen, L. (1996). Activation of Stat5a and Stat5b by tyrosine phosphorylation is tightly linked to mammary gland differentiation. Mol Endocrinol *10*, 1496-1506.

Liu, X., Robinson, G.W., Wagner, K.U., Garrett, L., Wynshaw-Boris, A., and Hennighausen, L. (1997). Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes Dev *11*, 179-186.

Lloyd, R.G. (1974). The segregation of the SbcA and Rac phenotypes in an Escherichia coli recB mutant. Mol Gen Genet *134*, 249-259.

Loonstra, A., Vooijs, M., Beverloo, H.B., Allak, B.A., van Drunen, E., Kanaar, R., Berns, A., and Jonkers, J. (2001). Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America *98*, 9209-9214.

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

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

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

Mallepell, S., Krust, A., Chambon, P., and Brisken, C. (2006). Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. Proceedings of the National Academy of Sciences of the United States of America *103*, 2196-2201.

Marti, A., Feng, Z., Altermatt, H.J., and Jaggi, R. (1997). Milk accumulation triggers apoptosis of mammary epithelial cells. European journal of cell biology *73*, 158-165.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proceedings of the National Academy of Sciences of the United States of America *78*, 7634-7638.

Martin, G.R., and Evans, M.J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. Proceedings of the National Academy of Sciences of the United States of America *72*, 1441-1445.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H.*, et al.* (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nature genetics *39*, 1197-1199.

Metzger, D., and Chambon, P. (2001). Site- and time-specific gene targeting in the mouse. Methods (San Diego, Calif *24*, 71-80.

Mikkola, M.L., and Millar, S.E. (2006). The mammary bud as a skin appendage: unique and shared aspects of development. J Mammary Gland Biol Neoplasia *11*, 187-203.

Miyoshi, K., Meyer, B., Gruss, P., Cui, Y., Renou, J.P., Morgan, F.V., Smith, G.H., Reichenstein, M., Shani, M., Hennighausen, L.*, et al.* (2002). Mammary epithelial cells are not able to undergo pregnancy-dependent differentiation in the absence of the helix-loop-helix inhibitor Id2. Mol Endocrinol *16*, 2892-2901.

Miyoshi, K., Shillingford, J.M., Smith, G.H., Grimm, S.L., Wagner, K.U., Oka, T., Rosen, J.M., Robinson, G.W., and Hennighausen, L. (2001). Signal transducer and activator of transcription (Stat) 5 controls the proliferation and differentiation of mammary alveolar epithelium. The Journal of cell biology *155*, 531-542.

Mueller, S.O., Clark, J.A., Myers, P.H., and Korach, K.S. (2002). Mammary gland development in adult mice requires epithelial and stromal estrogen receptor alpha. Endocrinology *143*, 2357-2365.
Murphy, K.C. (1998). Use of bacteriophage lambda recombination functions to promote gene replacement in Escherichia coli. Journal of bacteriology *180*, 2063-2071.
Nagy, A. (2000). Cre recombinase: the universal reagent for genome tailoring. Genesis *26*, 99-109.
Nakamura, T., Yamazaki, Y., Saiki, Y., Moriyama, M., Largaespada, D.A., Jenkins, N.A., and Copeland, N.G. (2000). Evi9 encodes a novel zinc finger protein that physically interacts with BCL6, a known human B-cell proto-oncogene product. Mol Cell Biol *20*, 3178-3186.
Nguyen, A.V., and Pollard, J.W. (2000). Transforming growth factor beta3 induces cell death during the first stage of mammary gland involution. Development (Cambridge, England) *127*, 3107-3118.
Nie, L., Perry, S.S., Zhao, Y., Huang, J., Kincade, P.W., Farrar, M.A., and Sun, X.H. (2008). Regulation of lymphocyte development by cell-type-specific interpretation of notch signals. Mol Cell Biol *28*, 2078-2090.
Nishimori, K., Young, L.J., Guo, Q., Wang, Z., Insel, T.R., and Matzuk, M.M. (1996). Oxytocin is required for nursing but is not essential for parturition or reproductive behavior. Proceedings of the National Academy of Sciences of the United States of America *93*, 11699-11704.
O'Connor, M., Peifer, M., and Bender, W. (1989). Construction of large DNA segments in Escherichia coli. Science (New York, NY *244*, 1307-1312.
Oakes, S.R., Naylor, M.J., Asselin-Labat, M.L., Blazek, K.D., Gardiner-Garden, M., Hilton, H.N., Kazlauskas, M., Pritchard, M.A., Chodosh, L.A., Pfeffer, P.L.*, et al.* (2008). The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev *22*, 581-586.
Okazuka, K., Wakabayashi, Y., Kashihara, M., Inoue, J., Sato, T., Yokoyama, M., Aizawa, S., Aizawa, Y., Mishima, Y., and Kominami, R. (2005). p53 prevents maturation of T cell development to the immature CD4-CD8+ stage in Bcl11b-/- mice. Biochemical and biophysical research communications *328*, 545-549.
Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell *132*, 631-644.
Ormandy, C.J., Camus, A., Barra, J., Damotte, D., Lucas, B., Buteau, H., Edery, M., Brousse, N., Babinet, C., Binart, N.*, et al.* (1997). Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. Genes Dev *11*, 167-178.
Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A.*, et al.* (2000). Molecular portraits of human breast tumours. Nature *406*, 747-752.
Pollard, J.W., and Hennighausen, L. (1994). Colony stimulating factor 1 is required for mammary gland development during pregnancy. Proceedings of the National Academy of Sciences of the United States of America *91*, 9312-9316.
Przybylski, G.K., Dik, W.A., Wanzeck, J., Grabarczyk, P., Majunke, S., Martin-Subero, J.I., Siebert, R., Dolken, G., Ludwig, W.D., Verhaaf, B.*, et al.* (2005). Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-

TRDC fusion transcripts and is associated with the absence of wild-type BCL11B transcripts in T-ALL. Leukemia *19*, 201-208.

Pulford, K., Banham, A.H., Lyne, L., Jones, M., Ippolito, G.C., Liu, H., Tucker, P.W., Roncador, G., Lucas, E., Ashe, S.*, et al.* (2006). The BCL11AXL transcription factor: its distribution in normal and malignant tissues and use as a marker for plasmacytoid dendritic cells. Leukemia *20*, 1439-1441.

Radtke, F., Wilson, A., and MacDonald, H.R. (2004). Notch signaling in T- and B-cell development. Current opinion in immunology *16*, 174-179.

Rajewsky, K., Gu, H., Kuhn, R., Betz, U.A., Muller, W., Roes, J., and Schwenk, F. (1996). Conditional gene targeting. The Journal of clinical investigation *98*, 600-603.

Ramirez-Solis, R., Davis, A.C., and Bradley, A. (1993). Gene targeting in embryonic stem cells. Methods in enzymology *225*, 855-878.

Ramirez-Solis, R., Liu, P., and Bradley, A. (1995). Chromosome engineering in mice. Nature *378*, 720-724.

Raouf, A., Zhao, Y., To, K., Stingl, J., Delaney, A., Barbara, M., Iscove, N., Jones, S., McKinney, S., Emerman, J.*, et al.* (2008). Transcriptome analysis of the normal human mammary cell commitment and differentiation process. Cell stem cell *3*, 109-118.

Raught, B., Liao, W.S., and Rosen, J.M. (1995). Developmentally and hormonally regulated CCAAT/enhancer-binding protein isoforms influence beta-casein gene expression. Mol Endocrinol *9*, 1223-1232.

Raymond, C.S., and Soriano, P. (2007). High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. PLoS ONE *2*, e162.

Robertson, E., Bradley, A., Kuehn, M., and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature *323*, 445-448.

Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A.*, et al.* (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nature methods *4*, 651-657.

Rohr, O., Lecestre, D., Chasserot-Golaz, S., Marban, C., Avram, D., Aunis, D., Leid, M., and Schaeffer, E. (2003). Recruitment of Tat to heterochromatin protein HP1 via interaction with CTIP2 inhibits human immunodeficiency virus type 1 replication in microglial cells. Journal of virology *77*, 5415-5427.

Rothenberg, E.V. (2007a). Cell lineage regulators in B and T cell development. Nat Immunol *8*, 441-444.

Rothenberg, E.V. (2007b). Regulatory factors for initial T lymphocyte lineage specification. Current opinion in hematology *14*, 322-329.

Russo, J., Ao, X., Grill, C., and Russo, I.H. (1999). Pattern of distribution of cells positive for estrogen receptor alpha and progesterone receptor in relation to proliferating cells in the mammary gland. Breast cancer research and treatment *53*, 217-227.

Saiki, Y., Yamazaki, Y., Yoshida, M., Katoh, O., and Nakamura, T. (2000). Human EVI9, a homologue of the mouse myeloid leukemia gene, is expressed in the hematopoietic progenitors and down-regulated during myeloid differentiation of HL60 cells. Genomics *70*, 387-391.

Sakakura, T., Kusano, I., Kusakabe, M., Inaguma, Y., and Nishizuka, Y. (1987). Biology of mammary fat pad in fetal mouse: capacity to support development of various fetal epithelia in vivo. Development (Cambridge, England) *100*, 421-430.

Satokata, I., Ma, L., Ohshima, H., Bei, M., Woo, I., Nishizawa, K., Maeda, T., Takano, Y., Uchiyama, M., Heaney, S.*, et al.* (2000). Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. Nature genetics *24*, 391-395.

Satterwhite, E., Sonoki, T., Willis, T.G., Harder, L., Nowak, R., Arriola, E.L., Liu, H., Price, H.P., Gesk, S., Steinemann, D.*, et al.* (2001). The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. Blood *98*, 3413-3420.

Sauer, B., and Henderson, N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proceedings of the National Academy of Sciences of the United States of America *85*, 5166-5170.

Schlake, T., and Bode, J. (1994). Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. Biochemistry *33*, 12746-12751.

Schorr, K., Li, M., Krajewski, S., Reed, J.C., and Furth, P.A. (1999). Bcl-2 gene family and related proteins in mammary gland involution and breast cancer. J Mammary Gland Biol Neoplasia *4*, 153-164.

Seagroves, T.N., Krnacik, S., Raught, B., Gay, J., Burgess-Beusse, B., Darlington, G.J., and Rosen, J.M. (1998). C/EBPbeta, but not C/EBPalpha, is essential for ductal morphogenesis, lobuloalveolar proliferation, and functional differentiation in the mouse mammary gland. Genes Dev *12*, 1917-1928.

Selbert, S., Bentley, D.J., Melton, D.W., Rannie, D., Lourenco, P., Watson, C.J., and Clarke, A.R. (1998). Efficient BLG-Cre mediated gene deletion in the mammary gland. Transgenic Res *7*, 387-396.

Senawong, T., Peterson, V.J., Avram, D., Shepherd, D.M., Frye, R.A., Minucci, S., and Leid, M. (2003). Involvement of the histone deacetylase SIRT1 in chicken ovalbumin upstream promoter transcription factor (COUP-TF)-interacting protein 2-mediated transcriptional repression. J Biol Chem *278*, 43041-43050.

Senawong, T., Peterson, V.J., and Leid, M. (2005). BCL11A-dependent recruitment of SIRT1 to a promoter template in mammalian cells results in histone deacetylation and transcriptional repression. Arch Biochem Biophys *434*, 316-325.

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

Shillingford, J.M., Miyoshi, K., Robinson, G.W., Bierie, B., Cao, Y., Karin, M., and Hennighausen, L. (2003). Proteotyping of mammary tissue from transgenic and gene knockout mice with immunohistochemical markers: a tool to define developmental lesions. J Histochem Cytochem *51*, 555-565.

Shinbo, T., Matsuki, A., Matsumoto, Y., Kosugi, S., Takahashi, Y., Niwa, O., and Kominami, R. (1999). Allelic loss mapping and physical delineation of a region harboring a putative thymic lymphoma suppressor gene on mouse chromosome 12. Oncogene *18*, 4131-4136.

Silver, L.M. (1985). Mouse Genetics, Concepts and Applications (Oxford University Press).

Singh, H., Medina, K.L., and Pongubala, J.M. (2005). Contingent gene regulatory networks and B cell fate specification. Proceedings of the National Academy of Sciences of the United States of America *102*, 4949-4953.
Sleeman, K.E., Kendrick, H., Robertson, D., Isacke, C.M., Ashworth, A., and Smalley, M.J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. The Journal of cell biology *176*, 19-26.
Smalley, M.J., Titley, J., and O'Hare, M.J. (1998). Clonal characterization of mouse mammary luminal epithelial and myoepithelial cells separated by fluorescence-activated cell sorting. In vitro cellular & developmental biology *34*, 711-721.
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*, 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*, 681-687.
Smith, G.H., and Boulanger, C.A. (2003). Mammary epithelial stem cells: transplantation and self-renewal analysis. Cell proliferation *36 Suppl 1*, 3-15.
Smith, G.H., Gallahan, D., Diella, F., Jhappan, C., Merlino, G., and Callahan, R. (1995). Constitutive expression of a truncated INT3 gene in mouse mammary epithelium impairs differentiation and functional development. Cell Growth Differ *6*, 563-577.
Smith, G.H., and Medina, D. (1988). A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. Journal of cell science *90 (Pt 1)*, 173-183.
Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature *317*, 230-234.
Snell, G.D., and Reed, S. (1993). William Ernest Castle, pioneer mammalian geneticist. Genetics *133*, 751-753.
Solter, D., Skreb, N., and Damjanov, I. (1970). Extrauterine growth of mouse egg-cylinders results in malignant teratoma. Nature *227*, 503-504.
Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S.*, et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America *98*, 10869-10874.
Spangrude, G.J., Heimfeld, S., and Weissman, I.L. (1988). Purification and characterization of mouse hematopoietic stem cells. Science (New York, NY *241*, 58-62.
Sternlicht, M.D. (2006). Key stages in mammary gland development: the cues that regulate ductal branching morphogenesis. Breast Cancer Res *8*, 201.
Stingl, J., and Caldas, C. (2007). Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. Nat Rev Cancer *7*, 791-799.
Stingl, J., Eaves, C.J., Zandieh, I., and Emerman, J.T. (2001). Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast cancer research and treatment *67*, 93-109.

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

Stylianou, S., Clarke, R.B., and Brennan, K. (2006). Aberrant activation of notch signaling in human breast cancer. Cancer research *66*, 1517-1525.

Suzuki, M., Gerstein, M., and Yagi, N. (1994). Stereochemical basis of DNA recognition by Zn fingers. Nucleic Acids Res *22*, 3397-3405.

Taddei, I., Deugnier, M.A., Faraldo, M.M., Petit, V., Bouvard, D., Medina, D., Fassler, R., Thiery, J.P., and Glukhova, M.A. (2008). Beta1 integrin deletion from the basal compartment of the mammary epithelium affects stem cells. Nat Cell Biol *10*, 716-722.

Teglund, S., McKay, C., Schuetz, E., van Deursen, J.M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J.N. (1998). Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. Cell *93*, 841-850.

Thangaraju, M., Rudelius, M., Bierie, B., Raffeld, M., Sharan, S., Hennighausen, L., Huang, A.M., and Sterneck, E. (2005). C/EBPdelta is a crucial regulator of pro-apoptotic gene expression during mammary gland involution. Development (Cambridge, England) *132*, 4675-4685.

Thomas, K.R., and Capecchi, M.R. (1986). Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene. Nature *324*, 34-38.

Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. Cell *44*, 419-428.

Timmons, L., Becker, J., Barthmaier, P., Fyrberg, C., Shearn, A., and Fyrberg, E. (1997). Green fluorescent protein/beta-galactosidase double reporters for visualizing Drosophila gene expression patterns. Developmental genetics *20*, 338-347.

Tonner, E., Barber, M.C., Allan, G.J., Beattie, J., Webster, J., Whitelaw, C.B., and Flint, D.J. (2002). Insulin-like growth factor binding protein-5 (IGFBP-5) induces premature cell death in the mammary glands of transgenic mice. Development (Cambridge, England) *129*, 4547-4557.

Topper, Y.J., and Freeman, C.S. (1980). Multiple hormone interactions in the developmental biology of the mammary gland. Physiological reviews *60*, 1049-1106.

Trainor, P.A., Zhou, S.X., Parameswaran, M., Quinlan, G.A., Gordon, M., Sturm, K., and Tam, P.P. (1999). Application of lacZ transgenic mice to cell lineage studies. Methods in molecular biology (Clifton, NJ *97*, 183-200.

Udy, G.B., Towers, R.P., Snell, R.G., Wilkins, R.J., Park, S.H., Ram, P.A., Waxman, D.J., and Davey, H.W. (1997). Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proceedings of the National Academy of Sciences of the United States of America *94*, 7239-7244.

Vaillant, F., Asselin-Labat, M.L., Shackleton, M., Lindeman, G.J., and Visvader, J.E. (2007). The emerging picture of the mouse mammary stem cell. Stem cell reviews *3*, 114-123.

van Deursen, J., and Wieringa, B. (1992). Targeting of the creatine kinase M gene in embryonic stem cells using isogenic and nonisogenic vectors. Nucleic Acids Res *20*, 3815-3820.

van Genderen, C., Okamura, R.M., Farinas, I., Quo, R.G., Parslow, T.G., Bruhn, L., and Grosschedl, R. (1994). Development of several organs that require inductive epithelial-

mesenchymal interactions is impaired in LEF-1-deficient mice. Genes Dev *8*, 2691-2703.
Veltmaat, J.M., Mailleux, A.A., Thiery, J.P., and Bellusci, S. (2003). Mouse embryonic mammogenesis as a model for the molecular regulation of pattern formation. Differentiation; research in biological diversity *71*, 1-17.
Veltmaat, J.M., Van Veelen, W., Thiery, J.P., and Bellusci, S. (2004). Identification of the mammary line in mouse by Wnt10b expression. Dev Dyn *229*, 349-356.
Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A.*, et al.* (2001). The sequence of the human genome. Science (New York, NY *291*, 1304-1351.
Visvader, J.E., and Lindeman, G.J. (2006). Mammary stem cells and mammopoiesis. Cancer research *66*, 9798-9801.
Wagner, E.F., Stewart, T.A., and Mintz, B. (1981a). The human beta-globin gene and a functional viral thymidine kinase gene in developing mice. Proceedings of the National Academy of Sciences of the United States of America *78*, 5016-5020.
Wagner, K.U., Young, W.S., 3rd, Liu, X., Ginns, E.I., Li, M., Furth, P.A., and Hennighausen, L. (1997). Oxytocin and milk removal are required for post-partum mammary-gland development. Genes and function *1*, 233-244.
Wagner, T.E., Hoppe, P.C., Jollick, J.D., Scholl, D.R., Hodinka, R.L., and Gault, J.B. (1981b). Microinjection of a rabbit beta-globin gene into zygotes and its subsequent expression in adult mice and their offspring. Proceedings of the National Academy of Sciences of the United States of America *78*, 6376-6380.
Wakabayashi, Y., Inoue, J., Takahashi, Y., Matsuki, A., Kosugi-Okano, H., Shinbo, T., Mishima, Y., Niwa, O., and Kominami, R. (2003a). Homozygous deletions and point mutations of the Rit1/Bcl11b gene in gamma-ray induced mouse thymic lymphomas. Biochemical and biophysical research communications *301*, 598-603.
Wakabayashi, Y., Watanabe, H., Inoue, J., Takeda, N., Sakata, J., Mishima, Y., Hitomi, J., Yamamoto, T., Utsuyama, M., Niwa, O.*, et al.* (2003b). Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. Nat Immunol *4*, 533-539.
Walton, K.D., Wagner, K.U., Rucker, E.B., 3rd, Shillingford, J.M., Miyoshi, K., and Hennighausen, L. (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*, 281-293.
Wang, J., Sarov, M., Rientjes, J., Fu, J., Hollak, H., Kranz, H., Xie, W., Stewart, A.F., and Zhang, Y. (2006). An improved recombineering approach by adding RecA to lambda Red recombination. Molecular biotechnology *32*, 43-53.
Warren, A.J., Colledge, W.H., Carlton, M.B., Evans, M.J., Smith, A.J., and Rabbitts, T.H. (1994). The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. Cell *78*, 45-57.
Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P.*, et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.
Watson, C.J. (2006a). Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. Breast Cancer Res *8*, 203.
Watson, C.J. (2006b). Post-lactational mammary gland regression: molecular basis and implications for breast cancer. Expert Reviews in Molecular Medicine *8*, 1.

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

Welm, B.E., Tepera, S.B., Venezia, T., Graubert, T.A., Rosen, J.M., and Goodell, M.A. (2002). Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population. Developmental biology *245*, 42-56.

Weniger, M.A., Pulford, K., Gesk, S., Ehrlich, S., Banham, A.H., Lyne, L., Martin-Subero, J.I., Siebert, R., Dyer, M.J., Moller, P.*, et al.* (2006). Gains of the proto-oncogene BCL11A and nuclear accumulation of BCL11A(XL) protein are frequent in primary mediastinal B-cell lymphoma. Leukemia *20*, 1880-1882.

Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjoblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J.*, et al.* (2007). The genomic landscapes of human breast and colorectal cancers. Science (New York, NY *318*, 1108-1113.

Wormald, S., and Hilton, D.J. (2004). Inhibitors of cytokine signal transduction. J Biol Chem *279*, 821-824.

Wysolmerski, J.J., Philbrick, W.M., Dunbar, M.E., Lanske, B., Kronenberg, H., and Broadus, A.E. (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*, 1285-1294.

Yamaguchi, N., Oyama, T., Ito, E., Satoh, H., Azuma, S., Hayashi, M., Shimizu, K., Honma, R., Yanagisawa, Y., Nishikawa, A.*, et al.* (2008). NOTCH3 signaling pathway plays crucial roles in the proliferation of ErbB2-negative human breast cancer cells. Cancer research *68*, 1881-1888.

Young, W.S., 3rd, Shepard, E., Amico, J., Hennighausen, L., Wagner, K.U., LaMarca, M.E., McKinney, C., and Ginns, E.I. (1996). Deficiency in mouse oxytocin prevents milk ejection, but not fertility or parturition. Journal of neuroendocrinology *8*, 847-853.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000). An efficient recombination system for chromosome engineering in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America *97*, 5978-5983.

Zhang, Y., Buchholz, F., Muyrers, J.P., and Stewart, A.F. (1998). A new logic for DNA engineering using recombination in Escherichia coli. Nature genetics *20*, 123-128.

Zheng, W., and Flavell, R.A. (1997). The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. Cell *89*, 587-596.

Zhou, J., Chehab, R., Tkalcevic, J., Naylor, M.J., Harris, J., Wilson, T.J., Tsao, S., Tellis, I., Zavarsek, S., Xu, D.*, et al.* (2005). Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. The EMBO journal *24*, 635-644.

APPENDIX

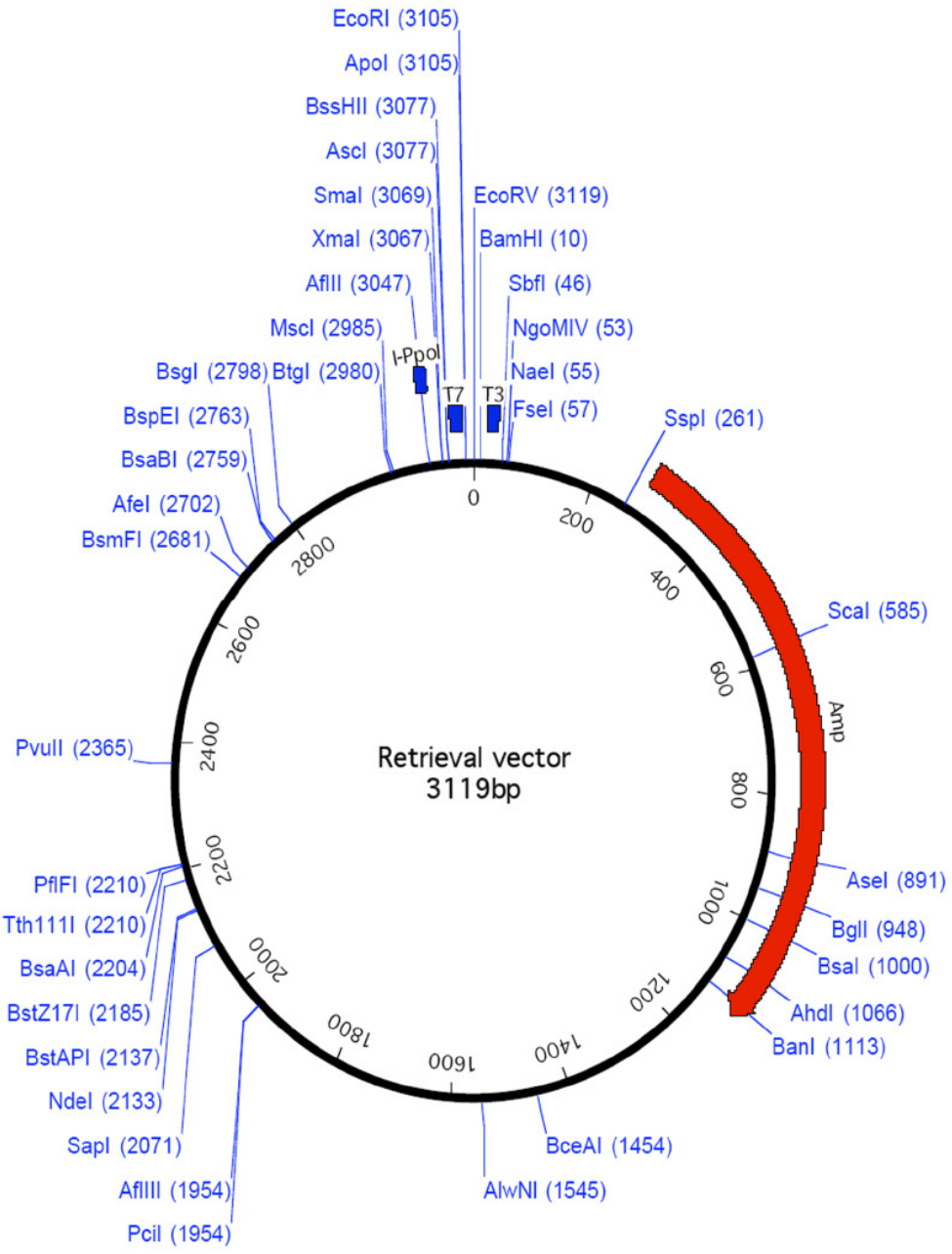


Figure A.1. Plasmid map of PL611 retrieval vector.

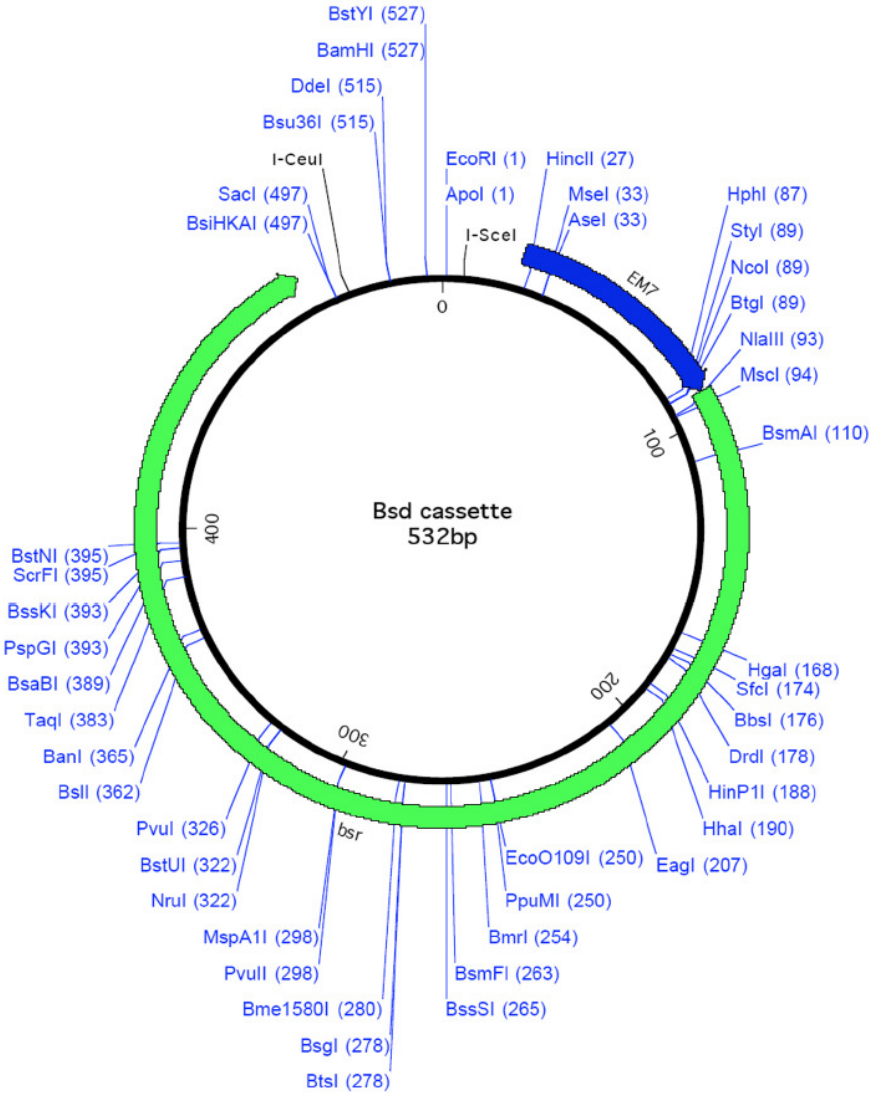


Figure A.2. Plasmid map of *Bsd* cassette.

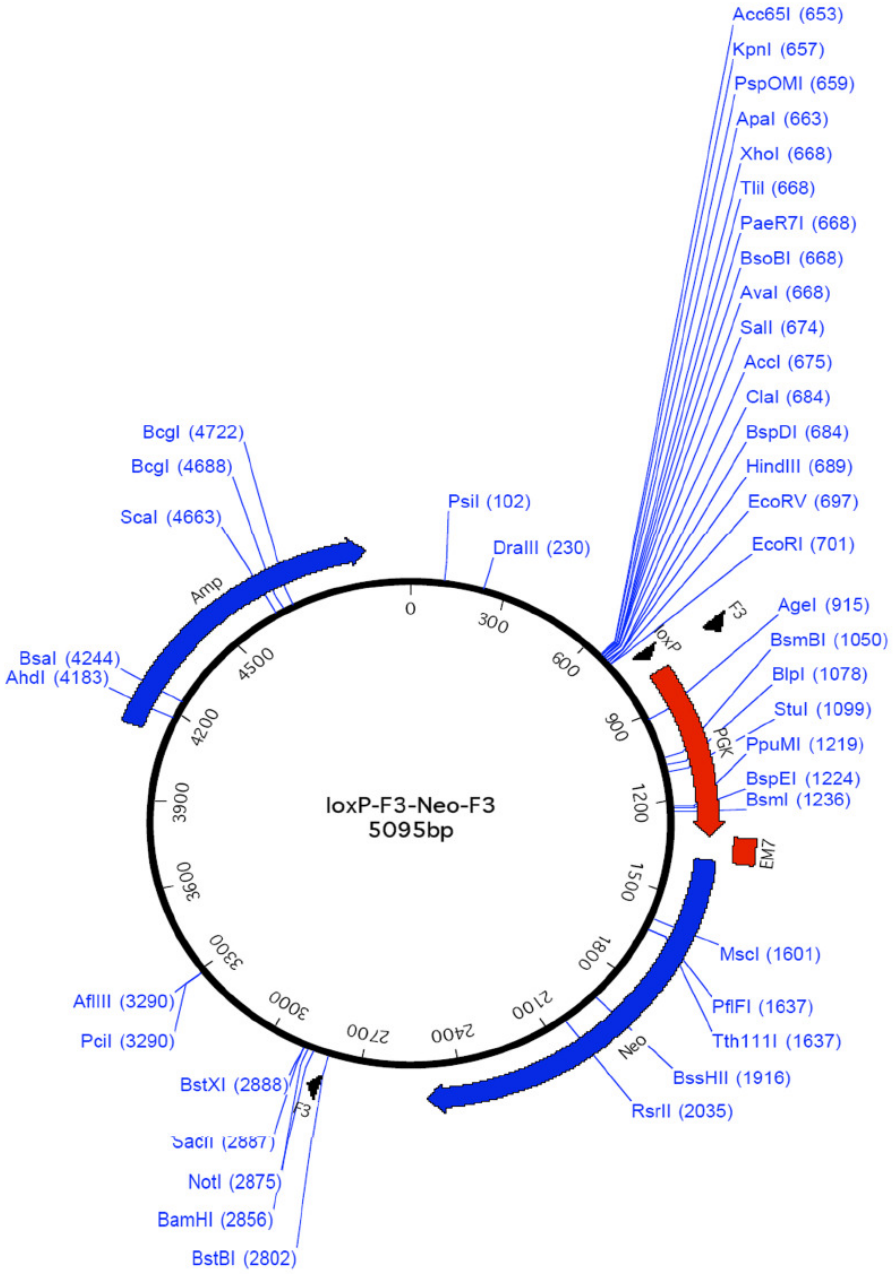


Figure A.3. Plasmid map of *Neo* cassette.

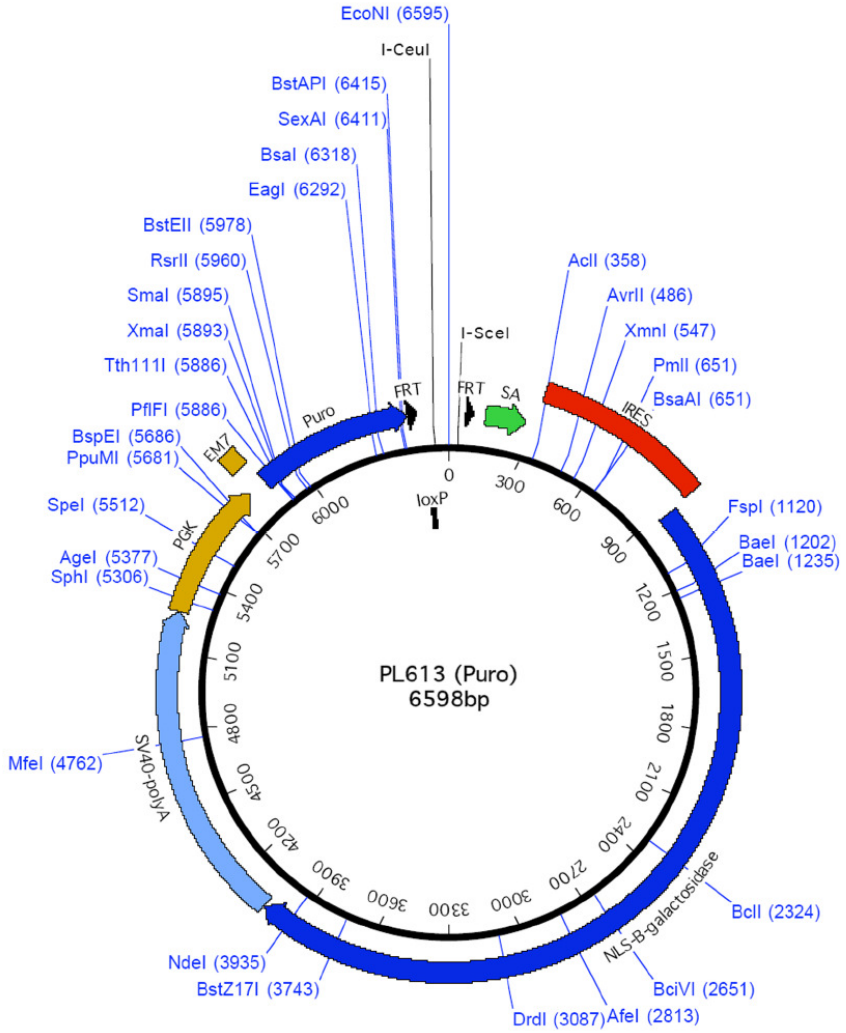


Figure A.4. Plasmid map of PL613 *lacZ* reporter cassette.

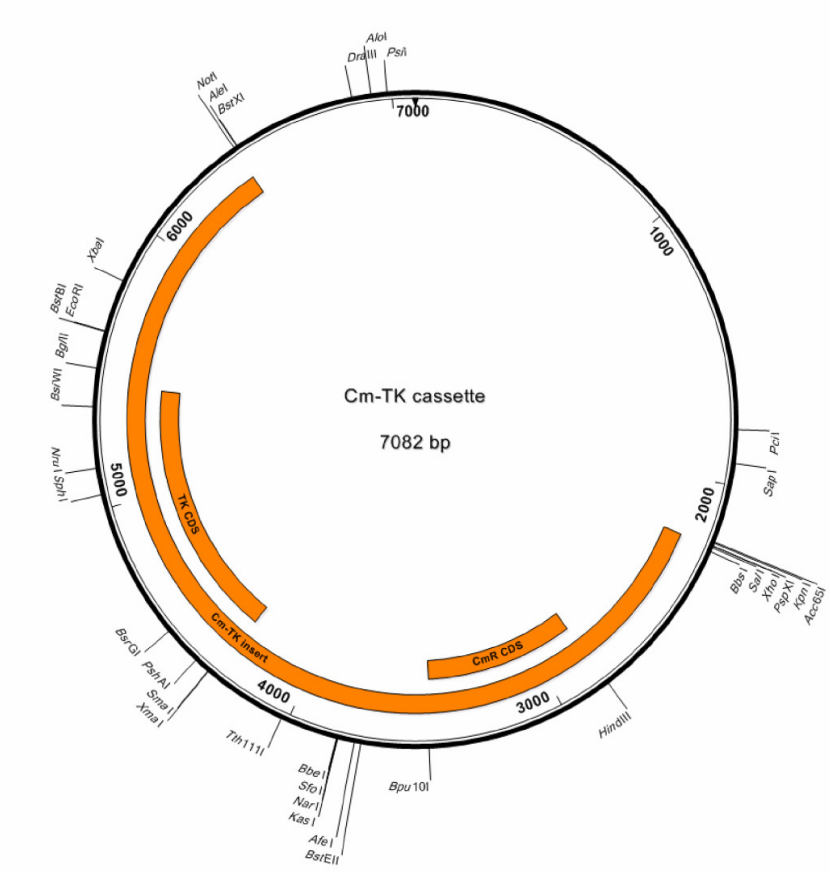


Figure A.5. Plasmid map of *Cm-TK* cassette.

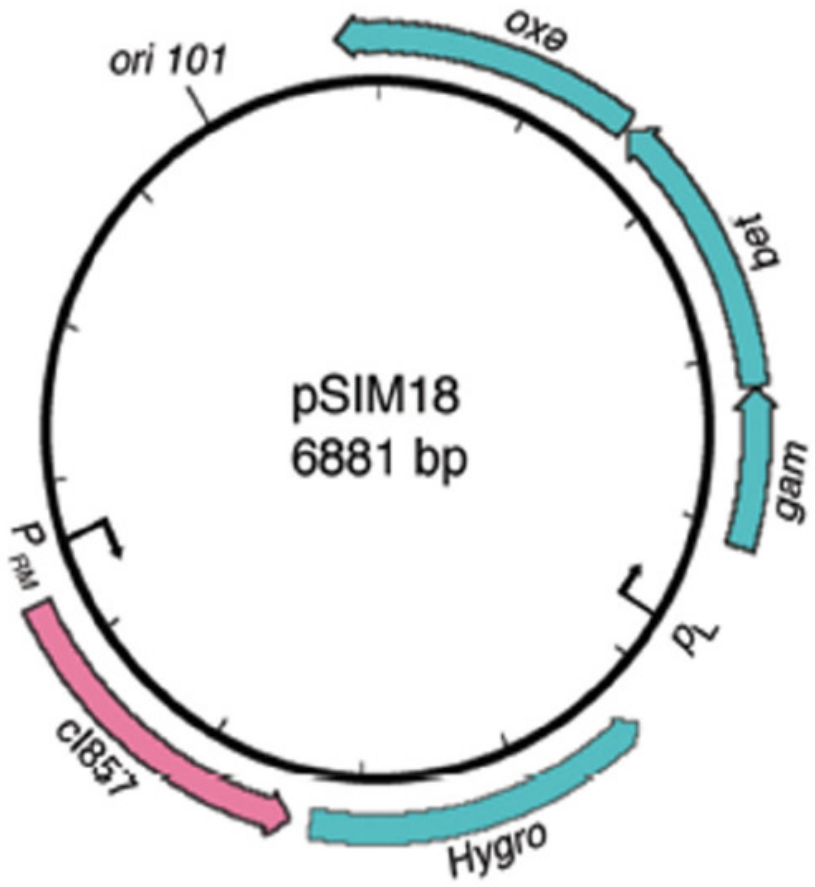


Figure A.6. Plasmid map of pSim18. Obtained from (Chan et al., 2007).

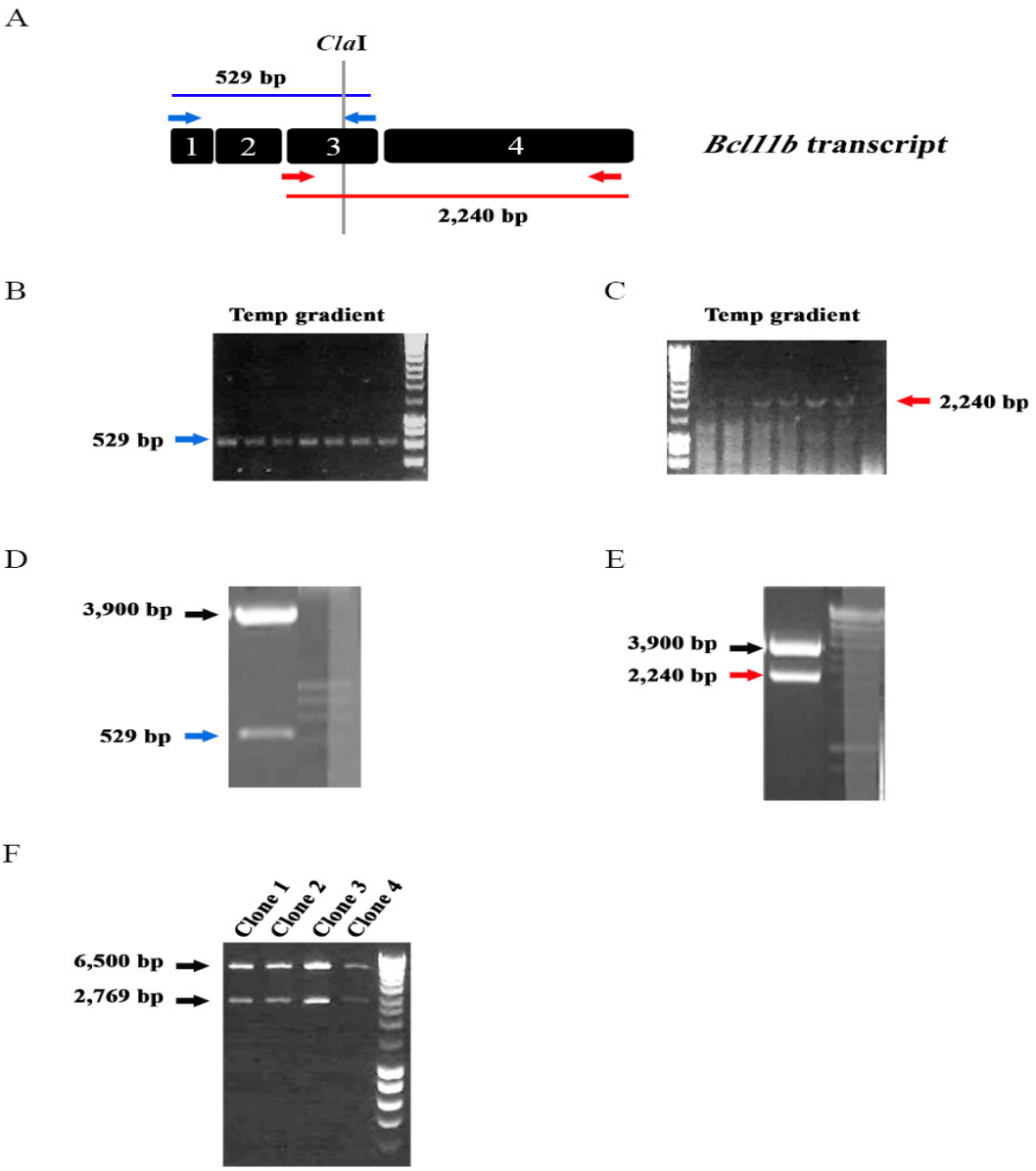


Figure A.7. Cloning of *Bcl11b* cDNA and construction of *Bcl11b* over-expression vector. (A) Positions of primer pairs used to amplify *Bcl11b* exons 1-2-3 (Blue arrows) and exons 3-4 (Red arrows). Gel images showing PCR products from the respective primer pairs. Fragments corresponding to **(B)** 529 bp and **(C)** 2,240 bp are purified and cloned into pCR-BluntII-TOPO vector (3,900 bp). Gel images showing restriction digestion products of **(D)** *Bcl11b* exons 1-2-3-pCR-BluntII-TOPO and **(E)** *Bcl11b* exons 3-4-pCR-BluntII-TOPO plasmids following *Bgl*II + *Cla*I and *Cla*I + *Eco*RI digestion respectively. Fragments corresponding to **(D)** 529 bp and **(E)** 2,240 bp are purified and cloned into *MSCV-IRES-eGFP* over-expression plasmid in a 3-way ligation reaction. **(F)** Gel images showing restriction digestion products of *MSCV-Bcl11b* cDNA-*IRES-eGFP* plasmid following *Bgl*II + *Eco*RI digestion. All four clones show the expected fragment sizes (*Bcl11b* cDNA insert – 2,769 bp and *MSCV-IRES-eGFP* backbone – 6,500 bp).

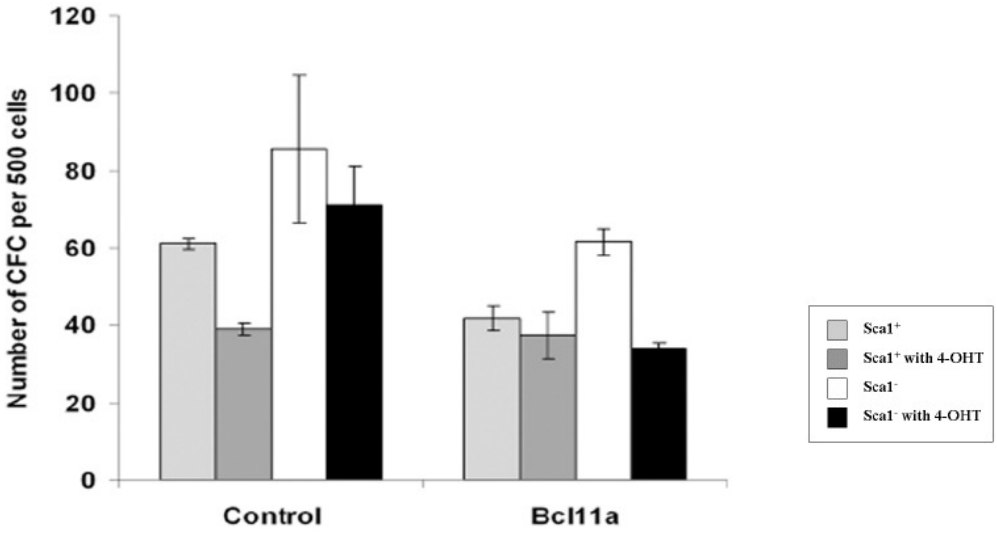


Figure A.8. ***In vitro*** **mammary colony-forming cells (Ma-CFCs).** Graphs showing number of Ma-CFCs per 500 cells after deletion of *Bcl11a* in sorted $Lin^{-}CD24^{hi}CD49b^{+}Sca1^{+/-}$ luminal progenitors. $CD24^{hi}CD49b^{+}Sca1^{+/-}$ luminal progenitors from *Cre-ERT2; Bcl11a*$^{flox/flox}$ mammary glands are sorted and plated with irradiated feeders in NSA media for 24 hours before 1 μM of 4-hydroxytamoxifen (4-OHT) is added to induce deletion of *Bcl11a*. After 2 hours, fresh NSA media is replaced and cells are maintained at $37^{o}C/5\%CO_2$ for another 6 days before the number of Ma-CFCs is enumerated.