



**BINDING SERVICES**  
Tel +44 (0)29 2087 4949  
Fax +44 (0)29 20371921  
e-mail [bindery@cardiff.ac.uk](mailto:bindery@cardiff.ac.uk)



Book + DNA Repair - Abou Ghar

# **Transgenic Analysis of Smad4.**

**Vasilios Kotsikoris**

**PhD**

**Cardiff University**

**2005**

UMI Number: U584696

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U584696

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## **Declaration**

**I declare that this thesis and the work presented in it is my own, except where specifically acknowledged.**

**Signature.....**

A handwritten signature in black ink, appearing to be 'VK', written over a dotted line.

**Vasileios Kotsikoris**

**2005**

## **Acknowledgements**

*To family, friends (Andreas Ballis, James Matthews, Mark Bishop, Owen Sansom), lovers and Professor Alan R Clarke.*

## Abstract

Mammary gland involution proceeds through massive, highly controlled epithelial cell apoptosis and tissue remodelling. Thus, mammary gland represents an ideal physiological environment in which to study apoptosis. Recently, the Smad gene family has been identified as mediators of TGF- $\beta$  superfamily signaling and has been implicated in mediating epithelial cell apoptosis. The Smad family of signal transduction proteins focuses around a central mediator, Smad4, and this report presents a number of analyses which have been undertaken to investigate the potential apoptotic role of Smad4 in the mammary gland.

To investigate the role of Smad4 I have utilised an over-expressing Smad4 transgenic mouse. This transgenic mouse has been designed to increase Smad4 protein in the secretory epithelial cells of the mammary gland during pregnancy and lactation by using the biological properties of the BLG promoter. Histological investigation of Smad4 slides indicated accelerated involution and significant increase of apoptosis at day 2 and 3 of involution. Molecular analysis through Western blots for STAT3 and STAT5a levels showed a deregulation of both proteins at the same time points. The mechanism with which transgenic mice regulated apoptosis was prove by Western blots to be via the *p27<sup>Kip1</sup>* molecular pathway and independent of *p21<sup>waf1</sup>* or *Bax*. Microarray analysis for day 3 involuting mice showed an over expression of Vitamin D Receptor. A result which was confirmed by both semi quantitative RT-PCR and Western blot analysis. Raw microarray data showed a down regulation of Methyl Binding Domain 2 (MBD2) in Smad4 transgenic mice indicating that MBD2 could play a role in mammary gland involution. However, analysis of conditional mammary knockout MBD2 mice at day 3 involution did not show any involvement of MBD2 in the regulation mammary gland.

Another way of investigating Smad4 was by passing the Smad4 construct to Stat3 mammary knockout animals with the hypothesis that Smad4 could restore STAT3's delayed involution. Based on the deregulation of STAT3 protein levels in the smad4 transgenics the above hypothesis was investigated at day 3 of involution. Phenotypic investigation showed that over expression of smad4 is not able to restore the delayed

phenotype of STAT3 knockouts. My final investigation was to identify molecules whose expression is altered by a conditional deletion of STAT3 in the mammary gland. This investigation was approached through microarrays which demonstrated an over expression of the Bone Morphogenic Protein receptor, a result which demonstrates once more the synergistic cooperation of Smads and STATS.

## Abbreviations

ActRII/IIB	Activin Type II and type IIB receptor
ALK1-6	Activin receptor-like kinase 1 to 6
APC	<i>Adenomatous Polyposis Coli</i>
ARIP	Activin Receptor Interacting Protein
BLG	Beta lactoglobulin
BMP	Bone Morphogenic Protein
BMPRII	BMP type II receptor
bp	base pairs
CBP	CREB Binding Protein
CDK	Cyclin Dependant Kinase
cDNA	Complementary DNA
Co-Smad	Collaborating Smad
CRE	cAMP Response Element
CRM1	Chromosomal region maintenance 1
DDW	Deionised Distilled Water
DPC	Deleted in Pancreatic Carcinomas
DPP	Decapentaplegic Gene
DPX	Distyrene, Plasticizer, Xylene
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
EMSA	Electrophoretic Mobility Shift Assay
ES	Embryonic Stem cells
EST	Expressed Sequence Tag
FAST	Forkhead Activin Signal Transducer
FDR	False Discovery Rate
FHA	Forkhead Associated Domain
GH	Growth Hormone
GDNF	Glial Cell Derived Neurotrophic Factor
H&E	Haematoxylin & Eosin
HAT	Histone Acetylase
HDAC	Histone Deacetylase
HPRT	hypoxanthine guanine phosphoribosyl transferase
ICF	Immunodeficiency-Centromeric instability-Facial anomalies
INF- $\gamma$	Interferon- $\gamma$
I-Smads	Inhibitory Smads
IVT	In Vitro Transcription
JAK	Janus Activated Kinase
kD	Kilo-Dalton
LEF/TCF1	Lymphoid Enhancer Factor 1/Transcription Factor 1
LOXP	Locus of Crossover of P1
MAD-H	Mothers against decapentaplegic-Homologues
MAD	Mothers against decapentaplegic
MBD	Methyl Binding Domain
MH1,2	Mad Homology 1,2
MMP	Matrix Metalloproteinase
NES	Nuclear Export Signal

NLS	Nuclear Localisation-like Sequence
PAI-1	Plasminogen Activator Inhibitor 1
PBS	Phosphate Buffered Saline
PCD	Programmed Cell Death
PCR	Polymerase Chain Reaction
PEBP2/CBF	Polyoma Virus Enhancer Binding Protein/Core Binding Factor
PIAS	Protein inhibitor of activated STAT
PRL	Prolactin
PVDF	Polyvinylidene Difluoride
R-Smads	Receptor Regulated Smads
SAD	Smad activation domain
SAM	Significance Analysis of Microarrays
SARA	Smad Anchor for Receptor Activation
SBD	Smad Binding Domain
SBE	Smad Binding Element
SD	Standard Deviation
SIP1	Smad Interacting Protein-1
SKI	Sloan Kettering avian retrovirus
SMAD	Sma-Mad
SMURF	Smad-ubiquitin Regulatory Factor
Sno-N	Ski-related Novel Gene
SOM	Self Organising Map
STAT	Signal Transducer and Activator of Transcription
STAT3 <sup>FL/FL</sup>	Signal Transducer and Activator of Transcription homozygous for FL
TBE	Tris Borate EDTA
TβRII	TGF-β type II receptor
TEB	Terminal End Bud
TF3	Transcription Factor 3
TGIF	Transcription Growth Factor Interacting Factor
TRAP1	TGF-β receptor- associated protein-1
TSP-1	Thrombospondin-1
TTM	TBS/Tween/Marvel
TUNEL	Terminal Deoxyuridine Triphosphate Nick End Labelling
VDR	Vitamin D Receptor
VEGF	Vascular Endothelial Growth Factor
VDRE	Vitamin D Receptor Element

## Contents

<b>Chapter 1-TGF-<math>\beta</math> and Smads</b>	<b>1</b>
<b>1.1 Introduction</b>	<b>1</b>
<b>1.1.1 The Transforming Growth Factor-<math>\beta</math> (TGF-<math>\beta</math>) Superfamily</b>	<b>1</b>
<b>1.1.2 Biological Actions of TGF- <math>\beta</math></b>	<b>2</b>
<b>1.2 TGF-<math>\beta</math> Signal Transduction Pathway and Components</b>	<b>3</b>
<b>1.2.1 Smad Dependent and Independent Pathways</b>	<b>3</b>
<b>1.2.2 TGF-<math>\beta</math> Signaling Receptors</b>	<b>5</b>
<b>1.2.3 TGF-<math>\beta</math> Receptor Types and Structure</b>	<b>6</b>
<b>1.2.4 Receptor Activation</b>	<b>7</b>
<b>1.3 Smad Proteins</b>	<b>8</b>
<b>1.3.1 Classification of Smads</b>	<b>8</b>
<b>1.3.2 Structural Features of Smads</b>	<b>11</b>
<b>1.4 The Core Smad Pathway</b>	<b>12</b>
<b>1.4.1 Smad Recognition, Access and Activation by the             Activated Receptor Complex</b>	<b>12</b>
<b>1.4.2 Nuclear import and export of Smads</b>	<b>14</b>
<b>1.4.3 Smads in the Nucleus</b>	<b>15</b>
<b>1.4.4 Direct Binding of Smads in the DNA</b>	<b>15</b>
<b>1.4.5 Interaction with DNA Binding Proteins</b>	<b>16</b>
<b>1.4.6 Interaction with Transcriptional co-Activators and             co-Rrepressors</b>	<b>17</b>
<b>1.4.7 Ending Smad Signaling</b>	<b>18</b>
<b>1.5 Mammary Gland Structure and Function</b>	<b>20</b>
<b>1.5.1 TGF-<math>\beta</math> and the Mammary Gland</b>	<b>23</b>
<b>1.6 The Cre-Lox Recombination System</b>	<b>24</b>
<b>1.7 Aims</b>	<b>29</b>

<b>Chapter 2-Materials and Methods</b>	<b>31</b>
<b>2.1 Mice</b>	<b>31</b>
<b>2.1.1 Maintenance of Mouse Colonies</b>	<b>31</b>
<b>2.1.2 Generation of Smad4 Transgenics</b>	<b>31</b>
<b>2.2 Genotyping of Mice</b>	<b>32</b>
<b>2.2.1 DNA Extraction from Mouse tails</b>	<b>32</b>
<b>2.2.2 Genotyping of Mice via PCR</b>	<b>33</b>
<b>2.2.3 BLG-SMAD4 PCR</b>	<b>33</b>
<b>2.2.4 BLG Cre Genotyping</b>	<b>34</b>
<b>2.2.5 Stat3 Flox PCR</b>	<b>34</b>
<b>2.2.6 Stat3 Null PCR</b>	<b>35</b>
<b>2.2.7 MBD2 PCR</b>	<b>35</b>
<b>2.3 Harvesting Mammary Glands</b>	<b>36</b>
<b>2.4 Mammary Gland Analysis</b>	<b>36</b>
<b>2.4.1 Histological Analysis</b>	<b>36</b>
<b>2.4.2 Wholemout Analysis</b>	<b>36</b>
<b>2.4.3 Peroxidase Staining of paraffin Embedded Tissue                 (Tunnel Analysis)</b>	<b>37</b>
<b>2.4.4 P21 Staining</b>	<b>38</b>
<b>2.5 Protein Extraction from Mammary Gland</b>	<b>40</b>
<b>2.5.1 Determination of protein concentration</b>	<b>41</b>
<b>2.5.2 Western Analysis</b>	<b>41</b>
<b>2.6 Primary Mammary Cell Culture</b>	<b>44</b>
<b>2.7 Total RNA extraction from mammary gland</b>	<b>45</b>
<b>2.7.1 RNA Extraction from Monolayered Cell Cultures</b>	<b>47</b>
<b>2.7.2 RNA Cleanup using the Qiagen RNeasy Kit</b>	<b>48</b>
<b>2.7.3 Target Preparation For Affymetrix GeneChip</b>	<b>49</b>
<b>2.8 RT-PCR. Primers and PCR Programs</b>	<b>56</b>

## **Chapter 3 - *In Vivo* Characterisation of a BLG-Smad4 Transgenic Line**

<b>in the Mammary Gland</b>	<b>58</b>
<b>3.1 Introduction</b>	<b>58</b>
<b>3.1.1 TGF-<math>\beta</math> and Breast Cancer. A Dual Role</b>	<b>58</b>
<b>3.1.2 Smad4 and Apoptosis</b>	<b>59</b>
<b>3.1.3 The Transgenic Approach</b>	<b>60</b>
<b>3.1.4 Aims</b>	<b>61</b>
<b>3.2 Results</b>	<b>62</b>
<b>3.2.1 Time points and requirements</b>	<b>62</b>
<b>3.2.2 BLG promoter expresses human Smad4 in the murine mammary gland throughout involution</b>	<b>62</b>
<b>3.2.3 Over expression of Smad4 accelerates involution of mammary epithelial cells</b>	<b>63</b>
<b>3.2.4 Over expression of Smad4 increases apoptosis in mammary epithelial cells</b>	<b>67</b>
<b>3.2.5 Molecular analysis of Stat3 and Stat5a in an over expressing Smad4 mammary gland environment</b>	<b>69</b>
<b>3.2.6 Over expression of Smad4 increases expression of p27<sup>Kip1</sup> and is not influencing expression of p21 and Bax</b>	<b>72</b>
<b>3.2.7 BLG promoter is not expressed in mammary epithelial cell cultures</b>	<b>75</b>
<b>3.3 Discussion</b>	<b>76</b>

## **Chapter 4- Microarray Analysis of Day 3 involuting mammary glands**

<b>in a Smad4 over expressing environment</b>	<b>80</b>
<b>4.1 Introduction</b>	<b>80</b>
<b>4.1.1 Microarray Analysis of Smad4</b>	<b>80</b>
<b>4.1.2 Affymetrix Technology</b>	<b>81</b>
<b>4.1.3 Normalization</b>	<b>82</b>

<b>4.2</b>	<b>Results</b>	<b>84</b>
4.2.1	RNA preparation and Affymetrix chip	84
4.2.2	Affymetrix gene controls and run consistency between chips	84
4.2.3	Data Filtering	86
4.2.4	Data Normalization and Linearity	86
4.2.5	Parametric Tests; T-Test and Significance	
	Analysis of Microarrays (SAM)	88
4.2.6	Expression profile changes in Smad4 transgenic mice	90
4.2.7	Smad4 over expression induces Vitamin D receptor	95
4.2.8	Raw Data Analysis and the MBD2 case	96
4.2.8.1	Depletion of MBD2 from the mammary gland does not alter mammary gland architecture in virgin mice	98
4.2.8.2	Depletion of MBD2 from the mammary gland does not alter involution nor apoptosis of mammary epithelial cells at day 3 involution	99
<b>4.3</b>	<b>Discussion</b>	<b>101</b>
4.3.1	Introduction	101
4.3.2	Affymetrix Technology; Filtering, Data processing and Statistical Analysis	101
4.3.3	Smad4's Expression profile	104
4.3.4	Smad4 over expresses Vitamin D Receptor	105
4.3.5	MBD2	106
 <b>Chapter 5- Smad-Stat Interactions and Microarray Analysis of Stat3<sup>+/-</sup></b>		<b>108</b>
<b>5.1</b>	<b>Introduction</b>	<b>108</b>
5.1.1	STATs and the mammary gland	108
5.1.2	Role of STAT3 in Mammary Gland Development	109
5.1.3	STAT-SMAD interaction	111
5.1.4	STAT3 Knockout	113
5.1.5	Aims	113
<b>5.2</b>	<b>Results</b>	<b>114</b>

<b>5.2.1</b>	<b>Analysis of Over Expressing Smad4 in a Stat3 Null Environment</b>	<b>114</b>
<b>5.2.1.1</b>	<b>Smad4 over expression does not influence involution in a conditional Stat3 knockout mouse</b>	<b>114</b>
<b>5.2.1.2</b>	<b>Over expression of Smad4 does not alter apoptosis in a conditional Stat3 knockout mouse</b>	<b>116</b>
<b>5.2.2</b>	<b>Stat3 Microarrays</b>	<b>117</b>
<b>5.2.2.1</b>	<b>Primary Mammary Epithelial cell cultures</b>	<b>117</b>
<b>5.2.2.2</b>	<b>The Stat3<sup>-/-</sup> deficient cultures were not viable</b>	<b>118</b>
<b>5.2.2.3</b>	<b>Affymetrix chips and Data Processing</b>	<b>118</b>
<b>5.2.2.4</b>	<b>Data Clustering</b>	<b>122</b>
<b>5.3</b>	<b>Discussion</b>	<b>124</b>
<b>5.3.1</b>	<b>Introduction</b>	<b>124</b>
<b>5.3.2</b>	<b>Smad-Stat Interactions</b>	<b>124</b>
<b>5.3.3</b>	<b>Microarray Analysis of Stat3<sup>+/-</sup> Primary Mammary Cell Cultures</b>	<b>125</b>
 <b>Chapter 6- Conclusion</b>		 <b>128</b>
<b>6.1</b>	<b>Introduction</b>	<b>128</b>
<b>6.2</b>	<b>Smads and Apoptosis</b>	<b>129</b>
<b>6.2.1</b>	<b>Transgenic Analysis of Smad4</b>	<b>130</b>
<b>6.2.2</b>	<b>Future work</b>	<b>131</b>
<b>6.3</b>	<b>Microarray analysis of Smad4</b>	<b>132</b>
<b>6.3.1</b>	<b>Future Work</b>	<b>133</b>
<b>6.3.2</b>	<b>The MBD2 mouse</b>	<b>133</b>
<b>6.4</b>	<b>Smad4-Stat3 Interactions at Day 3 Involution</b>	<b>133</b>
<b>6.5</b>	<b>Stat3 Microarrays</b>	<b>134</b>
<b>6.6</b>	<b>Conclusions</b>	<b>135</b>



# Chapter 1-TGF- $\beta$ and Smads.

## 1.1 Introduction

### 1.1.1 The Transforming Growth Factor- $\beta$ (TGF- $\beta$ ) Superfamily.

The Transforming growth factor  $\beta$  (TGF- $\beta$ ) family consists of more than 35 members of structurally related polypeptide growth factors which include the Bone Morphogenic Proteins (BMPs), Activins, Inhibins, the Nodals, anti-Mullerian hormone, and many other related factors in vertebrates, insects and nematodes (Massague 1998). Each of these factors is capable of regulating a fascinating array of cellular responses including cell-cycle control, regulation of early development, differentiation, extracellular matrix formation, hematopoiesis, angiogenesis, chemotaxis, and immune functions (Bottner 2000, Dunker and Krieglstein 2000, Lawrence 1996, Mummery 2001, Saltis 1996). Collectively, these factors account for a substantial portion of the intracellular signals governing cell fate.

One basic concept concerning the role of the TGF- $\beta$  family as hormonally active agents warrants mention at the outset. Unlike classical hormones, whose actions are few and clearly defined, the members of the TGF- $\beta$  family have many different effects depending on the type and state of the cell. For example, in the same healing wound TGF- $\beta$  may stimulate or inhibit cell proliferation depending on whether the target is a fibroblast or a keratinocyte (Ashcroft 1999); in mammary epithelial cells TGF- $\beta$  will cause growth arrest or metastatic behaviour depending on the level of oncogenic *Ras* activity present in the cell (Oft 1996). Finally, human BMP4 and its *Drosophila* ortholog, *dpp* (decapentaplegic gene), can signal dorsalization in the fly (Padgett 1993) yet bone formation in a vertebrate (Sampath 1993). TGF- $\beta$  family members are multifunctional hormones, the nature of their effects depending on what has been called “the cellular context”.

### **1.1.2 Biological Actions of TGF- $\beta$ .**

TGF- $\beta$  and related factors occupy a central position in the signaling networks that control cell fate, by regulating proliferation, differentiation and apoptosis and are therefore important for both development and maintenance of nearly all tissues (Hogan 1996). TGF- $\beta$  has a multitude of actions, most important of which is the regulation of genes whose products contribute to the growth and stability of the extracellular matrix (ECM). TGF- $\beta$  regulates these processes by stimulating the production of collagen, fibronectin and other ECM components and by decreasing the production of proteases that degrade the ECM. This cytokine was originally discovered as an autocrine growth factor of tumor cells that allowed these cells to grow in soft agar, reflecting a transformed phenotype. It is produced by a variety of cell types but CD4 T-cells are a particularly significant source. The actions of TGF- $\beta$  are very pleiotropic, in that this cytokine inhibits the growth of many cell types (e.g. epithelial, endothelial, neuronal, haematopoietic, and lymphoid origins) and stimulates the growth of others (e.g. mesenchymal cells). The growth inhibitory effect of TGF- $\beta$  is thought to be the molecular basis for many more of its global effects including the regulation of immune responses, development and cellular differentiation. In addition to its effects on cell proliferation and differentiation, TGF- $\beta$  also regulates many biological processes through its ability to transcriptionally regulate a number of cellular genes (Massague 1990).

As stated above, TGF- $\beta$  is able to stimulate cell proliferation of some cells and inhibit the growth of others. Although, this ability to affect different cell types in opposite ways is puzzling, it is now clear that the initial cellular responses to TGF- $\beta$  are similar and there are no fundamental differences in the types of receptors that get activated or downstream regulatory proteins that transduce signals in the nucleus. Instead, the different types of cellular responses occur because TGF- $\beta$  initiates a number of changes in all responsive cells and depending on the individual responses of specific cell types and on the environment of the cells, the balance of the effects of these changes could result either in cellular proliferation or inhibition of proliferation.

TGF- $\beta$  is a cytokine commonly implicated with differentiation of cells. For example in mammals it is found at highest concentrations in the tip of the intestinal villus where it is associated with the most differentiated cells (Murphy 1998). This observation has given rise to the postulate that it is involved in an “autocrine axis” that regulates the orderly progression of differentiated cells not only in the intestine but also in the skin, perhaps explaining its vital role in wound healing (Wahl 1989). Other members of the family, especially the BMPs, are also very important in the healing processes that occur in bone and cartilage (Bostrom and Asnis 1998).

Despite the fact that one might predict numerous and complex transduction pathways regulating this diversity of responses, surprisingly, a disarmingly simple system has recently been elucidated involving a family of membrane receptor protein kinases and a family of receptor substrates, the Smad proteins.

## **1.2 TGF- $\beta$ Signal Transduction Pathway and Components.**

### **1.2.1 Smad Dependent and Independent Pathways.**

The transforming growth factor- $\beta$  pathway occupies a central position in the signaling networks that control growth, differentiation and final fate of metazoan cells. Over the past few years, remarkable progress has been made in identifying the central components of this pathway, defining their interactions, and deciphering how a cell interprets its signals. The TGF- $\beta$  pathway involves a family of transmembrane receptor protein kinases and a family of receptor substrates, the Smad proteins, which translocate into the nucleus where they act as transcription factors. The ligand, TGF- $\beta$ , assembles a receptor complex that activates Smads, and in turn, the Smads assemble multisubunit complexes that regulate transcription (reviewed in Derynck 2003, Shi 2003) (Figure 1.1). Two general steps thus suffice to carry the hormonal stimulus to target genes.

How can such a simple system mediate a variety of cell specific responses? An incoming Smad complex is met in the nucleus by a set of partner proteins that are specific to a particular cell type in each particular set of conditions. These partners determine the DNA sequences that the Smad complex will bind, the transcriptional co-activators or co-repressors it will recruit, the other transcription factors it will cooperate with, and how long this process will last. The mixture of Smad partners and regulators present in a given cell at the time of TGF- $\beta$  stimulation thus determines the outcome of the response and at the same time defines, in molecular terms, the cellular context. Identifying these partners and regulators is, therefore, critical for understanding TGF- $\beta$  action.

### **1.2.2 TGF- $\beta$ Signaling Receptors.**

TGF- $\beta$  and related factors signal through a family of transmembrane protein serine/threonine kinases known as the TGF- $\beta$  receptor family. Protein kinases play an important role in signal transduction by phosphorylating specific amino acids in downstream components. Catalytic domains can predict what kind of residues will be phosphorylated by a given kinase. Many receptors for hormones and growth factors are membrane bound tyrosine kinases with reasonably well-understood signal transduction pathways. In contrast, until relatively recently, serine/threonine kinase activity had been detected only in cytoplasmic proteins downstream of receptor signaling.

This family of receptors came to light with the cloning of an Activin receptor (Mathews 1991), now referred to as Activin Type II Receptor (ActR-II), with properties similar to those of TGF- $\beta$  receptors identified in cross-linking studies (Cheifetz 1987) and genetically implicated in TGF- $\beta$  signal transduction (Laiho 1990). The cloning of ActR-II also revealed a striking similarity between this molecule and *Daf-1*, a previously identified orphan receptor from *Caenorhabditis elegans* (Georgi 1990). Exception to the above scenario is the glial cell-derived

neurotrophic factor (GDNF), which signals through the tyrosine kinase Ret receptor (Massague 1998). GDNF is the most divergent family member and shows very little sequence similarity to other members. These findings provided the basis and impulse for the rapid identification of many other members of this receptor family.

### 1.2.3 TGF- $\beta$ Receptor Types and Structure

Based on their structural and functional properties, TGF- $\beta$  receptor family is divided into two subfamilies: type I and type II receptors. Both of the receptors are glycoproteins of approximately 55kDa and 70kDa, respectively, containing an extracellular, transmembrane and kinase domain. A unique feature of the type I receptor is a highly conserved 30-amino acid region preceding the kinase domain region. This domain is called the GS domain because of the characteristic SGSGSGSG sequence it contains (Wrana 1994). This region has been proved to be a key regulatory region able to control the catalytic activity of type I receptor kinase. The function of this domain in TGF- $\beta$  signaling will be explained in detail below.

The extracellular domains of both type I and II receptors consists of approximately 150 amino acids which can be N-glycosylated and contain 10 or more cysteine residues that may determine the general fold of this region (Wells 1997). Three of this cysteines form a characteristic cluster near the transmembrane sequence (Wrana 1994) while the spacing of the remaining varies and seems more conserved in type I receptors. No main differences exist in the transmembrane and cytoplasmic juxtamembrane regions. However, Ser213 in this region of type II receptor is phosphorylated by the receptor kinase in a ligand independent manner and is required for signaling activity (Luo 1997). Ser165 in the juxtamembrane region of type I receptor is phosphorylated by the type II receptor in a ligand dependent manner, and this appears to selectively modulate the intensity of different TGF- $\beta$  responses.

The kinase domain in both type I and II receptors is a typical serine/threonine receptor kinase (Mathews 1991, Franzen 1993). Type I receptors have been shown to

phosphorylate their substrates, the Smads, on serine residues, while the type II receptors are able to phosphorylate themselves and type I receptors on serine and threonine residues.

As previously mentioned the GS domain is a unique feature of the type I receptor. Ligand induced phosphorylation of the serines and threonines in this region of type I receptors by the type II is required for activation of signaling (Souchelnytskyi 1996, Wrana 1994, Wieser 1995). Following this specific SGSGSG sequence, all type I receptors have a LEU-Pro motif that serves as binding site for the immunophilin FKBP12 (Chen Y 1997). FKBP12 may act as a negative regulator of the receptor signaling function. Furthermore, at the boundary of the GS and the kinase domain, there is either a threonine or glutamine residue. Mutation of this residue to aspartate or glutamate in type I receptors has been shown to endow the receptor with elevated kinase activity in vitro and constitutive signaling activity in the cell (Wiersdorff 1996). Thus, the GS domain is a key regulator region that may control the catalytic activity of the type I receptor kinase or its interaction with substrates.

#### **1.2.4 Receptor Activation.**

TGF- $\beta$  and related family members activate signaling by binding and bringing together pairs of type I and type II receptors. Primarily, ligand binding occurs to the type II receptor (T $\beta$ R-II) (Massague 1998), which exists in the cell membrane in an oligomeric form (Lu 1996) with activated kinase (Henis 1994, Chen 1994). In turn the type I receptor, which may also occur in an oligomeric form (Lu 1996) and cannot bind TGF- $\beta$  in the absence of type-II receptor, is recruited into the complex. Activation of the type I receptor from the type II occurs in the GS domain through transphosphorylation at threonine and serine domains (Franzen 1995). The assembly of the receptor complex is triggered by ligand binding, but the complex is also stabilised by direct interaction between the cytoplasmic parts of the receptors (Feng 1996). This model predicts that the type II and type I receptors act in sequence, which is supported by the finding that a constitutively active type I receptor (Thr 204

replaced by an aspartate residue) is able to exert TGF- $\beta$  signals in the absence of type II (Wieser 1995). This model of binding was discovered using TGF- $\beta$  resistant cell mutants: TGF- $\beta$ 1 can bind to type II receptors in cell mutants lacking a type I receptor but cannot bind to a type I receptor in mutants lacking type II (Laiho 1990, Massague 1998).

## 1.3 Smad Proteins.

### 1.3.1 Classification of Smads.

The Smad protein family are the only known substrates of type I receptor kinases and play a central role in the transduction of receptor signals to target genes in the nucleus. These proteins were first identified as the products of the *Drosophila Mad* (Sekelsky 1995) and *C.elegans Sma* genes (Derynck 1996), which lie downstream of the BMP-analogous ligand-receptor systems (Patterson and Padgett 2000, Whitman 1998). The human genome encodes eight Smad family members (Smad 1-8) [Mad-homologues (MAD-H)], and related proteins are known in the rat, mouse, *Xenopus*, Zebrafish, *Drosophila* and *C.elegans*. Smads are ubiquitously expressed in all adult tissues (Fladers 2001, Luukko 2001), and many of them (Smad2, 4, 5, 6 and 8) are produced from alternatively spliced mRNAs.

Functionally, Smads fall into three subfamilies (Figure 1.2). a) receptor-regulated Smads (R-Smads: Smad1, Smad2, Smad3, Smad5 and Smad8), which become phosphorylated by the type I receptors. b) common mediator Smads (Co-Smad: Smad4), which oligomerise with activated R-Smads and directs them to the nucleus, and c) the inhibitory Smads (I-Smads: Smad6 and Smad7), which are induced by TGF- $\beta$  family members and provide a negative feedback loop to TGF- $\beta$  signaling.

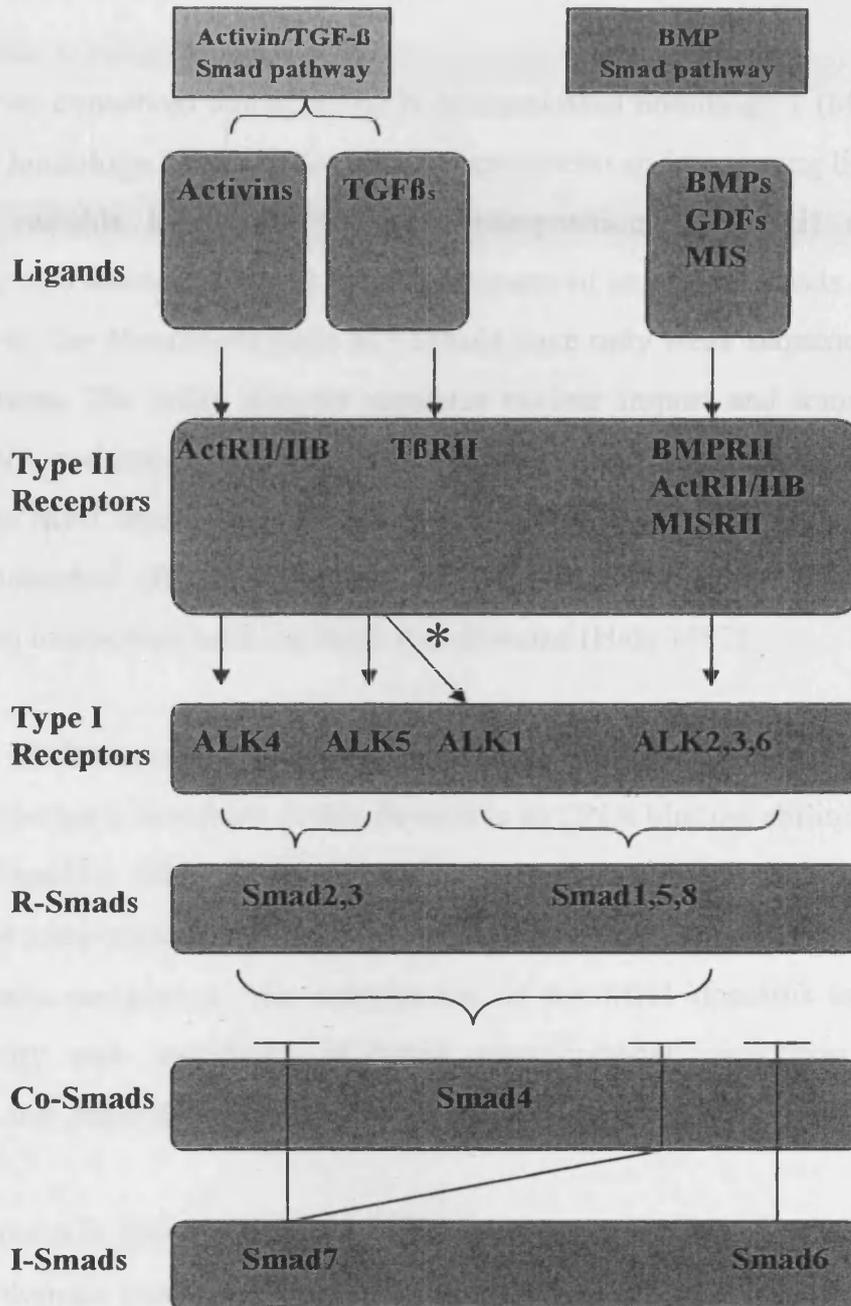
Amongst the receptor-regulated Smads, Smad1 and its close homologues Smad5 and Smad8 get activated upon BMP ligand binding in their corresponding receptors. On

the other hand, Smad2 and Smad3 are substrates of TGF- $\beta$  and Activin signal transduction.

Signaling by receptor-regulated Smads requires the participation of a collaborating Smad. The only known member of this group in vertebrates is Smad4. Smad4 associates with receptor-regulated Smads after their activation from their corresponding receptors. Although Smad4 is similar to the R-Smads in overall structure (see below), it fails to become phosphorylated in response to agonists. Smad4 is required for Smad2 and/or Smad3 dependent growth inhibitory responses in mammalian cells, and a dominant-negative Smad4 construct interferes with Smad1 and Smad2 signaling in frog embryos and mammalian cells (Zhang 1996, Lagna 1996). Smad4, therefore, participates in TGF- $\beta$ , Activin and BMP signaling pathways as a shared partner of receptor-regulated Smads.

I-Smads exert a negative feedback effect by competing with R-Smads for receptor interaction and by marking the receptors for degradation. More specifically, Smad6 preferentially inhibits BMP signaling (Imamura 1997, Hata 1998) and Smad7 inhibits TGF- $\beta$  and BMP signaling. Smad7 inhibits phosphorylation of R-Smads by occupying the type I receptors for BMPs, Activins, and TGF- $\beta$ . Smad6 preferentially inhibits BMP signaling by competing with Smad4 for binding to receptor activated Smad1 and forms an inactive Smad1-Smad6 complex. Smad6 and Smad7 levels are increased in response to Activin, BMP or TGF- $\beta$  signaling, suggesting that these Smads function as negative feedback controls for different pathways.

## Signalling Specificity in the TGF- $\beta$ superfamily



**Figure 1.2** Classification of the mammalian Smad signalling cascade into Activin, TGF- $\beta$  and BMP pathways. Representative examples of typeII, I, R-Co, I-Smads are depicted in pathways linked by arrows or signs of inhibition. Bifurcation of the TGF- $\beta$  pathway at the level of type I receptors towards both TGF- $\beta$  and BMP Smads is marked by the asterisk.

### 1.3.2 Structural Features of Smads.

Smads have two conserved domains, the N-terminal Mad homology 1 (MH1) and C-terminal Mad homology 2 (MH2) domains. There is also an intervening linker region, which is of variable length and sequence composition. The MH1 domain has approximately 130 amino acids and is highly conserved among R-Smads and the Co-Smad; however, the N-terminal parts of I-Smads have only weak sequence similarity to MH1 domains. The MH1 domain regulates nuclear import and transcription by binding to DNA and interacting with nuclear proteins (will be discussed later). In the basal state, the MH1 domain has been shown to inhibit MH2's transcriptional (Liu F 1996) and biological (Baker 1996) activities and this inhibitory effect has been attributed to an interaction between these two domains (Hata 1997).

However, the MH1 domain does not have a purely inhibitory function. As mentioned above one of the basic functions of this domain is its DNA binding ability. It has been shown that Smad4's MH1 domain contributes to the DNA-binding activity of a Smad2-Smad4 transcriptional complex (Liu 1997). The same appears to be true for all Smad4-R-Smads complexes. The contribution of the MH1 domains to the DNA-binding affinity and specificity of Smad transcriptional complexes may vary depending on the particular target gene.

The MH2 domain is about 200 amino acids long and is highly conserved among all Smads. This domain contains receptor phosphorylation sites (in R-Smads) (Macias-Silva 1996), which regulate Smad oligomerisation, has effector function (Liu F 1996, Baker 1996), and is involved in several important protein-protein interactions (discussed later) with various cytoplasmic adaptors and transcription factors. More specifically, interactions between the MH2 domains support Smads homo-oligomeric complexes, which exist in all subfamilies at the basal state (Lagna 1996, Hata 1997, Hata 1998, Wu 1997). Furthermore, MH2 is the domain which mediates the association of R-Smads with type I receptors (Macias-Silva 1996) and Smad4 upon activation (Hata 1997) and with DNA binding factors. It has been shown that Smad1 and Smad2 require the presence of the MH2 domain in order to activate transcription

(Liu F 1996, Liu F 1997). In the case of inhibitory Smads, the MH2 domain is sufficient for their inhibitory effect (Hata 1998, Topper 1997).

The crystal structure of Smad4 MH2 domain has provided insights into the basis for some of the above interactions. Its structure contains several  $\alpha$ -helices and loops, which surround a  $\beta$ -sandwich (Shi 2001), and it resembles the forkhead associated (FHA) domain, a phosphopeptide-binding domain common in transcription and signaling factors (Li 2000). Compared with other Smad proteins, Smad4 has a characteristic insertion in the MH2 domain and lacks the SSXS motif, the site of phosphorylation by the type I receptors. Due to this sequence alteration, Smad4 is not a substrate of the type I receptors (Zhang 1996). Furthermore, Smad4's essential role in Smad-mediated transcriptional activation is partly due to a unique Smad activation domain (SAD), a 48 amino acid proline-rich regulatory element in its linker region (Itoh 2000).

The linker region is highly variable in size and sequence. Despite this variability all linker regions contain multiple phosphorylation sites, which allow specific crosstalk with other signaling pathways. Additionally, this region contributes to the formation of Smad homo-oligomers (Hata 1997, Wu 1997).

## **1.4 The Core Smad Pathway.**

### **1.4.1 Smad Recognition, Access and Activation by the Activated Receptor Complex.**

Phosphorylated type I receptor exhibit a significantly enhanced binding affinity for the R-Smads. This binding selectivity has been attributed to a positively charged surface patch on the R-Smads's MH2 domain which is absent in Smad4 (Wu 2000). It has been demonstrated that mutation of His331, on the basic patch of Smad2, leads to a reduction of its affinity for, and phosphorylation by, type I receptor (Huse 2001).

The first step in the intracellular TGF- $\beta$ /Smad pathway is the recruitment of R-Smads to the TGF- $\beta$  receptor complex (Figure 1.1). Several proteins with anchoring, scaffolding, and/or chaperone activity have now been identified in regulating and facilitating this process. Smad anchor for receptor activation (SARA) has been shown to regulate the subcellular distribution of Smad2 and Smad3 (Tsukazaki 1998). SARA is associated to the inner leaflet of the plasma membrane via its FYVE phospholipids binding domain, (Tsukazaki 1998). In addition, SARA can simultaneously interact with (non activated) R-Smads and the TGF- $\beta$  receptor complex through its Smad binding domain (SBD) and its C-terminal region, respectively. After type I receptor activation, two serine residues in the C-terminal sequence of Smad2 and Smad3 are phosphorylated by the type I receptor kinase, followed by the dissociation from SARA and type I receptor. Activated Smad2/3 forms subsequently a heteromeric complex with Smad4. Overall, TGF- $\beta$  receptor activation is required for R-Smad presentation by SARA to the receptor (Tsukazaki 1998).

It has also been demonstrated that the TRAP-1 protein (TGF- $\beta$  receptor-associated protein-1) can specifically interact with Smad4 and aid the recruitment of the co-Smad to the TGF- $\beta$  and Activin receptor complex (Wurthner 2001). In the basal state TRAP-1 associates with the inactive TGF- $\beta$  receptor complex, and, upon receptor activation dissociates from the complex and associates with Smad4. The TRAP-1/Smad4 interaction is not continuous but is disrupted by activated R-Smads. It has been suggested that TRAP-1 may function as a chaperone to reduce the auto-inhibitory interactions between the MH1 and MH2 domains of Smad4 and facilitate in this way the interaction of Smad4 with activated R-Smads (Wurthner 2001).

Several other factors with the same properties have been demonstrated to play various possible roles in Smad anchoring. Microtubules can anchor inactive Smads in the cytoplasm (Dong 2000). More specifically, ligand induction results in dissociation of Smads from the microtubule network. Sasaki (2001) demonstrated that filamin, an actin cross linking factor with scaffolding abilities, regulates transduction of Smad signals. Other examples of Smad associating scaffolding protein are the ARIPs

(Activin receptor interacting proteins), which associate with Smad2 and enhance Smad2-mediated signaling in response to Activin (Tsuchida 2001).

The available data support the notion that interactions between Smads and proteins with scaffolding/adaptor proteins are an important regulatory mechanism. Proper receptor localisation in plasma membrane, their proximity to cytoplasmic anchors that hold the Smads and the ability of such compartments are exciting new aspects of the regulation of Smad signaling. Such mechanisms could provide cell-context specificity, allowing differential regulation of the basic Smad pathway.

#### **1.4.2 Nuclear import and export of Smads.**

R and Co-Smads primarily exist in the cytoplasm, where upon ligand stimulation they subsequently translocate to the nucleus (Heldin 1997, Massague and Wotton 2000). This is accomplished through a nuclear localisation-like sequence (NLS-like; Lys-Lys-Leu-Lys), which exists in the MH1 domain of both R-and Co-Smads (Xiao 2000a, Kurisaki 2001). This motif has been reported to be important for Smad3 translocation to the nucleus after TGF- $\beta$  stimulation. Smad3 directly binds to importin- $\beta$  via this region but not to importin- $\alpha$  (Xiao 2000a, Kurisaki 2001), and Ran GTPase promotes nuclear translocation of activated Smad3 to the nucleus (Kurisaki 2001). Smad2 has a unique insert of exon3 in the same domain, which prevents its association with importin- $\beta$  (Kurisaki 2001). Nuclear import of non-phosphorylated Smad2 is blocked by SARA, which masks the intrinsic nuclear import activity of Smad2, which was reported to have an NLS in its MH2 domain (Xu 2000).

In contrast with Smad2 and Smad3, Smad4 has a leucine-rich nuclear export signal (NES) in its linker region (Pierreux 2000). The nuclear export of Smad4 is mediated by the general export factor CRM1; inhibition of CRM1 by treatment with leptomycin B results in rapid accumulation of Smad4 (but not Smad2 or 3) in the nucleus (Pierreux 2000). Smad4, which in the basal state is located in the cytoplasm, is proposed to rapidly and constitutively shuttle between the nucleus and the cytoplasm. It has also been proposed that upon heteromeric complex formation between R-and

Co-Smads, the nuclear accumulation of the complex may be stimulated by the shielding of the NES and/or unmasking NLS on R-and/or Co-Smad (Pierreux 2000).

Smad1 (but not Smad2 and 3) has also been found to contain a functional NES in its MH2 domain (Xiao 2001). Thus, Smad1 may also constantly shuttle between nucleus and cytoplasm, with receptor-induced Smad1 phosphorylation shifting the balance towards nuclear accumulation of Smad1.

### **1.4.3 Smads in the Nucleus.**

Once formed, the heteromeric R-Smad/Smad4 complex translocates to the nucleus where it regulates transcription of target genes by directly binding to consensus DNA sequences, interacting with other DNA binding proteins, and by recruiting several transcriptional co-activators or co-repressors. The nature of the biological readout in a given cell is highly dependent on the particular R-Smad that is activated, and on the cell-specific expression of Smad interactors. These interactions affect Smad availability, Smad activity, the selection of target genes and so help determine which particular constellation of genes is up- or down-regulated.

### **1.4.4 Direct Binding of Smads in the DNA.**

The key function of Smad proteins in TGF- $\beta$  family signal transduction is transcriptional regulation. Biochemical evidence indicates that the Smads can bind DNA. Thus, the R-Smads (except for Smad2) and Smad4, recognise a specific DNA sequence (AGAC) or its complementary (GTCT), termed Smad binding elements (SBEs) (Dennler 1998, Zawel 1998). This binding is facilitated via an interaction between Smad's MH1 domains and the SBE within target gene promoter. Smad3 and Smad4 bind directly within the major DNA groove but with low affinity, through a conserved beta-hairpin loop in the MH1 domain. Although Smad2 is structurally

similar to Smad3, and both Smad2 and Smad3 are activated by TGF- $\beta$  or Activin type I receptors, Smad2 does not bind to DNA. In contrast to Smad3, Smad2 has a short region composed of 30 amino acids in the MH1 domain encoded by its unique exon-3, which prevents its DNA binding (Yagi 1999). Additional MH1 sequences, such as  $\alpha$ -helix 2, contribute to SBE DNA-binding by Smad3 (Kusanagi 2001). I-Smads cannot bind to DNA.

#### 1.4.5 Interaction with DNA Binding Proteins.

Smads bind directly to DNA sequences with relatively low affinity ( $K_d = 1.14 \times 10^{-7}$  mol/L), suggesting that the Smads need to interact with other cellular factors to form transcriptional complexes and regulate transcription of target genes (Shi 1998).

The first and most extensively characterised Smad transcriptional partners are members of the FAST (Forkhead Activin Signal Transducer) family. FAST-1 is a member of the winged-helix family of DNA-binding proteins. *Xenopus*, mouse and human FASTs have been demonstrated to be required for TGF- $\beta$  or Activin dependant activation of the *Xenopus mix2* gene. The homeobox gene *mix2* is a response gene responsible for early *Xenopus* development (Chen 1997). The mammalian homologue (alternative named FAST-2) can mediate the activation of the homeobox gene *gooseoid* following gastrulation (Labbe 1998, Zhou 1998). FAST can bind constitutively to specific elements in the promoters of these genes, but cannot activate transcription on their own. However, upon activation of Activin or TGF- $\beta$  signaling pathways, a complex containing FAST, R-Smads and Smad4 assembles on the DNA and transcription is strongly activated (Labbe 1998, Zhou 1998). More specifically, FAST interacts with Smad2-Smad4 or Smad3-Smad4 complexes, but not with BMP-activated Smad complexes.

Consistent with the idea that Smad binding alone is not sufficient for gene activation, there are now a growing number of examples in which Smads are found to cooperate with DNA binding partners to regulate transcription.

Furthermore, at TGF- $\beta$  responsive AP-1 binding sites, Smad3 interacts through its MH1 domain with *c-jun* and through its MH2 domain with *c-Fos*, and thus forms a multiprotein complex (Yamamura 2000). The cooperativity between Smad3-Smad4 and *c-Jun/c-Fos* then results in TGF- $\beta$ -induced transcription for these promoter sequences. In a conceptually similar way, Smad3-Smad4 cooperates with the basic helix-loop-helix protein TFE3 to induce transcription from the PAI-1 promoter in response to TGF- $\beta$  although no physical interaction between TFE3 and Smad3-Smad4 has been demonstrated (Hua 1998).

In addition, the responsiveness of the promoters for the *p15* and *p21* CDK cell cycle inhibitors to TGF- $\beta$  has been localised to Sp1-binding sites, suggesting that Sp1 is required for their transcriptional activation by Smads in response to TGF- $\beta$ . It has also been demonstrated that Smads can associate with ATF2, PEBP2/CBF, vitamin D receptor, *Gli3* and *Hoxc8* (Denryck 1998). Smad4 was found to interact and functionally cooperate with  $\beta$ -catenin and LEF1/TCF, which are components of the *Wnt* signaling pathway, in inducing *Xenopus twin* (*Xtwn*) gene expression (Nishita 2000). Smad3 has also been reported to functionally interact with LEF1/TCF in a TGF- $\beta$ -dependant manner independent of  $\beta$ -catenin (Labbe 2000).

#### **1.4.6 Interaction with Transcriptional co-Activators and co-repressors.**

Transcriptional activity of Smads also depends on interactions with co-activators within the complex. It has been shown that R-Smads bind directly to p300/CBP. Smad3 interacts with p300 with high affinity. Smad1 and Smad2 can also interact with p300 (Nishihara 1998, Feng 1998). Interaction of p300 with Smad3 leads to the enhancement of transcriptional activation by Smad3 of target genes, including p3TP-Lux. In turn, p300 binds to various transcription factors, including STAT, and positively regulates the transcription of target genes in combination with Smads. Although Smads and STATs do not directly interact with each other they are bridged

by p300 in the nucleus resulting in efficient activation of translation (Nakashima 1999).

Smad signaling can also lead to repression of gene expression. Smad3 has been reported to associate with histone deacetylase (HDAC) activities through its MH1 domain, but whether Smads interact directly with HDACs remains unclear (Liberati 2001). Alternatively, Smads interact with various transcriptional co-repressors that recruit HDACs. These co-repressors include the homeodomain DNA-binding protein TGIF (Transcription Growth Factor Interacting Factor) (Wotton 1999). TGIF interacts with activated Smad2/3 and may limit the magnitude of TGF- $\beta$  responses in cells. Furthermore, R-Smad/Smad4 complexes interact with the proto-oncogene products *c-Ski* and *Sno-N*, which are important transcriptional co-repressors. Smad2/3 as well as Smad4 interacts with *c-ski* through multiple regions. Neither BMP-activated R-Smads nor I-Smads bind to *c-ski*. Smad3 interacts with *c-ski* through its MH2 domain and since p300 also interacts with the same domain of Smad3, there is a competition between *c-ski* and p300 for Smad3 binding. Smad3 interacts with *c-ski*, which in turn recruits HDAC (Akiyoshi, 1999). Thus, histone acetylation is induced by Smad3 and p300, and abolished by the addition of *c-ski*.

Additional inhibitors of the nuclear functions of Smads include SIP1 (Smad Interacting Protein-1) and Evi-1. SIP1 interacts with the MH2 domain of R-Smads but not with Smad4, in both yeast and mammalian cells (Verschueren 1999).

Such co-repressors appear to modulate the nuclear activity of Smads, and their levels of expression define the level of Smad transcriptional activity.

#### **1.4.7 Ending Smad Signaling.**

Termination of Smad signaling has been attributed so far to two mechanisms. Smad signaling termination can be achieved by dephosphorylation caused by as yet unidentified phosphatases (Randall 2002), and ubiquitination and proteasome-

mediated degradation of activated R-Smads. In unstimulated cells, the Smad-ubiquitin E3 ligase (Smurf1) targets Smad1 and Smad5 for destruction (Zhu 1999). This action is considered important for the possible maintenance of the basal state in unstimulated cells. Activated Smad2 is ubiquitinated in the nucleus and undergoes proteasome-mediated degradation in a process which includes Smurf2, which has extensive sequence similarity to Smurf1 (Lo 1999, Lin 2000, Zhang 2001). Ubiquitination of activated Smad3 appears to be mediated by a different E3, the SCF/Roc1 complex (Fukuchi 2001).

Smurf1 and 2 also mediate ubiquitination of activated TGF- $\beta$  receptors, leading to their degradation in the proteasome (Ebisawa 2001, Tajima 2003). I-Smad 7 facilitates this process by moving from the nucleus to the plasma membrane upon TGF- $\beta$  or BMP stimulation (Itoh 1998) where it forms a complex with both Smurf1 and 2. This complex binds directly to the activated type I TGF- $\beta$  receptor and inhibits phosphorylation of R-Smads (Kavsak 2000, Suzuki 2002). In the membrane complex, Smurf1 subsequently directly mediates the ubiquitination and turnover of the receptors (Ebisawa 2001, Tajima 2003). Smad7 itself also undergoes ubiquitination and degradation in this process. However, this negative feedback loop is countered by the ability of TGF- $\beta$  (and BMPs) to transcriptionally activate Smad7, thus ensuring a steady supply of the protein as it undergoes degradation.

In the nucleus, Smad7 associates with p300, which acetylates lysine residues 64 and 70 in Smad7, protecting it against Smurf-mediated ubiquitination of the same residues (Gronroos 2002). Smad7 acetylation is lost as the protein leaves the nucleus in response to TGF- $\beta$  stimulation and separates from p300. Although the R- and co-Smads also associate with p300 and the related acetyl transferase CBP in the context of assembling transcriptional complexes, no evidence has been found that they can undergo acetylation.

These observations underscore the complex mechanisms controlling TGF- $\beta$  signaling and suggest the involvement of additional factors that may regulate the stability of Smad proteins at both the basal and the activated states.

## 1.5 Mammary Gland Structure and Function.

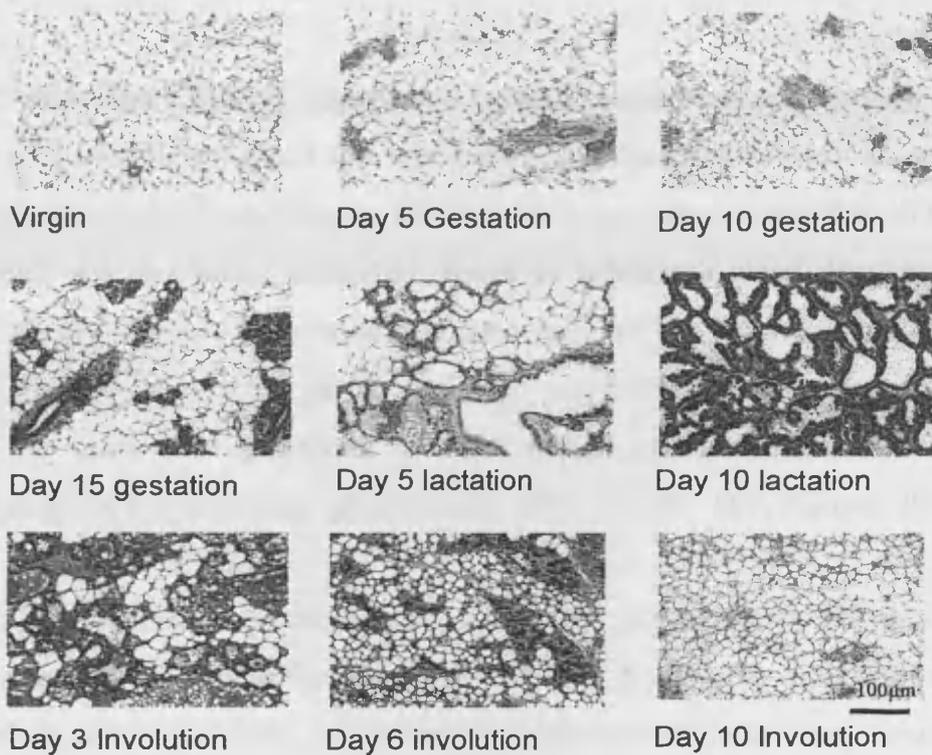
The mammary gland is characterised by a unique dependence of hormonal signals for terminal differentiation, which is attained only after pregnancy. At the time of birth, the glands consist of a few rudimentary ducts in the vicinity of the nipple. Pronounced ductal outgrowth and branching commences at puberty, and during pregnancy an expanded lobulo-alveolar compartment develops. Functional differentiation of the secretory epithelium coincides with parturition and large amounts of milk are produced and secreted during lactation. The fully functional gland is 90% epithelial. After weaning of the young, the entire alveolar epithelial compartment is remodelled to resemble a virgin like state, a process termed involution. With each pregnancy, a new round of lobulo-alveolar development occurs.

During involution a coordinated process of alveolar programmed cell death (PCD) and lobulo-alveolar remodelling restructures the mammary gland. Simple removal of the suckling stimulus triggers this process. Mammary gland involution proceeds through two distinct phases. In the first stage, alveolar cells undergo programmed cell death, but there is no remodelling of the lobulo-alveolar structure. During the second stage, the lobulo-alveolar structure of the gland is obliterated as proteinases degrade the basement membrane and extracellular matrix. The two stages exhibit characteristic changes in gene expression or activity. First stage changes include up-regulated expression of sulphated glycoprotein, tissue inhibitor of metalloproteinases, interleukin-1 $\beta$  converting enzyme, cell cycle control proteins (*c-jun*, *c-Fos*, *JunB*, *JunD*, and *c-Myc*) and decreased expression levels of milk protein genes. Second stage changes include increased expression levels of matrix metalloproteinases (MMPs), gelatinase A, stromelysin-1 and serine protease urokinase-type plasminogen activator.

Programmed cell death of individual alveolar cells during the first days of involution is correlated with increased expression levels of the death inducers *bax* and *bcl-x*.

Furthermore, changes in activity of two STAT family members accompany mammary gland involution; decreased activity of the prolactin signaling molecules STAT5a and STAT5b and activation of STAT3.

This repetitive cycling of the gland from full lactation to involution and back again is a phenomenon that occurs in few other tissues and thus elects the mammary gland as an ideal model to study highly controlled in vivo cellular apoptosis (Figure 1.3). Systemic hormones and the influence of the stroma on mammary epithelial cells have been characterised for some time (Sakakura 1991). However, only now through the availability of knockout mice have we been able to dissect individual steps in the pathways of the translation of hormonal signals into morphogenetic and developmental events. Two unique aspects of mammary gland development have greatly aided in exploiting these knockout animals and elucidate the specific roles of the epithelium and the stroma. First, the mammary gland develops predominantly in postpartum animals. Therefore, an entire developmental program, mimicking embryonic development of other organs, can be viewed and followed in postpartum animals. This characteristic has several ramifications; the tissue can be easily manipulated, and reasonable amounts of tissue are available for analysis. Second, genetic manipulations whose consequences in other tissues would result in lethality can now be studied due to the fact that mammary cells transplanted into the cleared fat pad continue to grow and divide very effectively. Thus, normal virgin ductal or lobulo-alveolar cells will readily grow and fill the fat pad with mammary ducts in virgin animals (or differentiate into alveolar cells if the recipient animal becomes pregnant).



**Figure 1.3. Mammary Gland Development (H&E staining).**

The development of the gland in mice as a model system proceeds in distinct phases; in newborn mice a rudimentary system of small ducts is present which grows slowly until the onset of puberty when pronounced ductal growth occurs. Development of the ducts continues in cycling virgins leading to the formation of a ductal tree which fills the entire mammary fat pad. Extensive ductal branching and alveolar growth occurs during pregnancy and is largely completed at parturition. Terminal differentiation of the alveolar epithelium is completed at the end of gestation with the onset of milk secretion at parturition. After weaning the entire alveolar epithelium undergoes apoptosis and the gland is then remodeled. Within a few weeks the gland has the appearance of that of a mature virgin. Magnification x40. Scale bar 100µm.

### 1.5.1 TGF- $\beta$ and the Mammary Gland.

Programmed cell death (PCD) (apoptosis) occurs during normal growth and development of the mammary gland and is perhaps one of the most dramatic examples of apoptosis driven remodelling (Strange 1995). It has been demonstrated that all three TGF- $\beta$  isoforms are expressed during all stages of mammary gland development except lactation (Robinson 1991). Mouse studies indicate key roles for TGF- $\beta$ s in establishing proper mammary gland structure, regulating stem cell kinetics, maintaining the mammary epithelium in functionally undifferentiated state, and inducing apoptosis in the involuting gland (Daniel 1996, Joseph 1999, Nguyen 2000).

Transgenic mouse models have demonstrated that over expression of polypeptides such as transforming growth factor alpha (TGF- $\alpha$ ) and insulin like growth factor I (IGF-I) can block this remodelling, suggesting that these growth factors may be acting as survival factors for the mammary epithelium (Jhappan 1990, Weber 1998). In contrast, transgenic mice that over express the growth inhibitor TGF- $\beta$  show increased apoptosis in the mammary epithelium throughout mammary development, suggesting a mechanism working to counterbalance the survival factors (Pierce 1993, Jhappan 1993). Furthermore, Gorska (1998) showed that over expression of a dominant negative receptor type II for TGF- $\beta$  in mammary epithelial cells disrupts normal TGF- $\beta$  signaling and leads to mammary hyperplasia.

Many members of the TGF- $\beta$  super family seem to be involved in mammapoiesis. Inhibins and Activins which are members of the TGF- $\beta$  family are important regulators. A specific function for the Activin and inhibin( $\beta$ )B subunit was identified for the mammary gland as mice deficient for the inhibin ( $\beta$ )B gene exhibit retardation of ductal elongation and alveolar morphogenesis during puberty and pregnancy respectively. Further, mammary transplantation experiments demonstrated a localised defect in the gland that was associated with the stroma. This resulted in ductal elongation and alveolar proliferation, with differentiation severely curtailed in inhibin( $\beta$ )B-deficient stroma (Robinson and Henninghausen 1997). These

experiments also showed abnormal morphology of terminal bud ends (TEBs) implying a disturbance in the balance of cell proliferation and cell death in this structure as one cause of the observed developmental defect.

## 1.6 The Cre-Lox Recombination System.

Recent advances in mouse embryology and molecular genetics have achieved the specific mutation of defined genes in the germline of mice creating a plethora of models for biological research. Protocols using transgenesis (Gordon 1980) and homologous recombination in embryonic stem (ES) cells (Capecchi 1989, Capecchi 1989a) permit over expression, inactivation and modification of genes at will. Our understanding of some of the important issues regarding the mechanisms controlling cell division, differentiation and death has dramatically advanced in recent years through exploitation of these techniques to generate transgenic mice. Mutations, identical to those observed in inherited or somatically acquired diseases such as in cancer, can now be mimicked in mouse model systems. These technologies are valuable in assessing the role of genes in complex processes such as tumorigenesis, embryonic development and functioning of the immune system. In particular, the generation of mice with targeted mutations in genes encoding proteins of interest has proved to be a useful way of elucidating the function of these gene products *in vivo*.

These technologies are invaluable in assessing the role of genes in biological processes as those mentioned above. However, there are a number of limitations as well e.g. nullizyosity appears to be lethal in many instances or causes complex pleiotropic effects and therefore do not permit the development of an *in vivo* model system in which gene inactivation is restricted to a defined subset of cells (Copp 1995). To overcome these limitations, strategies for conditional, cell type-specific gene targeting (Gu 1994) and inducible gene disruptions (Kuhn 1995) have been developed. These systems utilize site-specific recombinases such as the Cre/Loxp recombination system of bacteriophage P1.

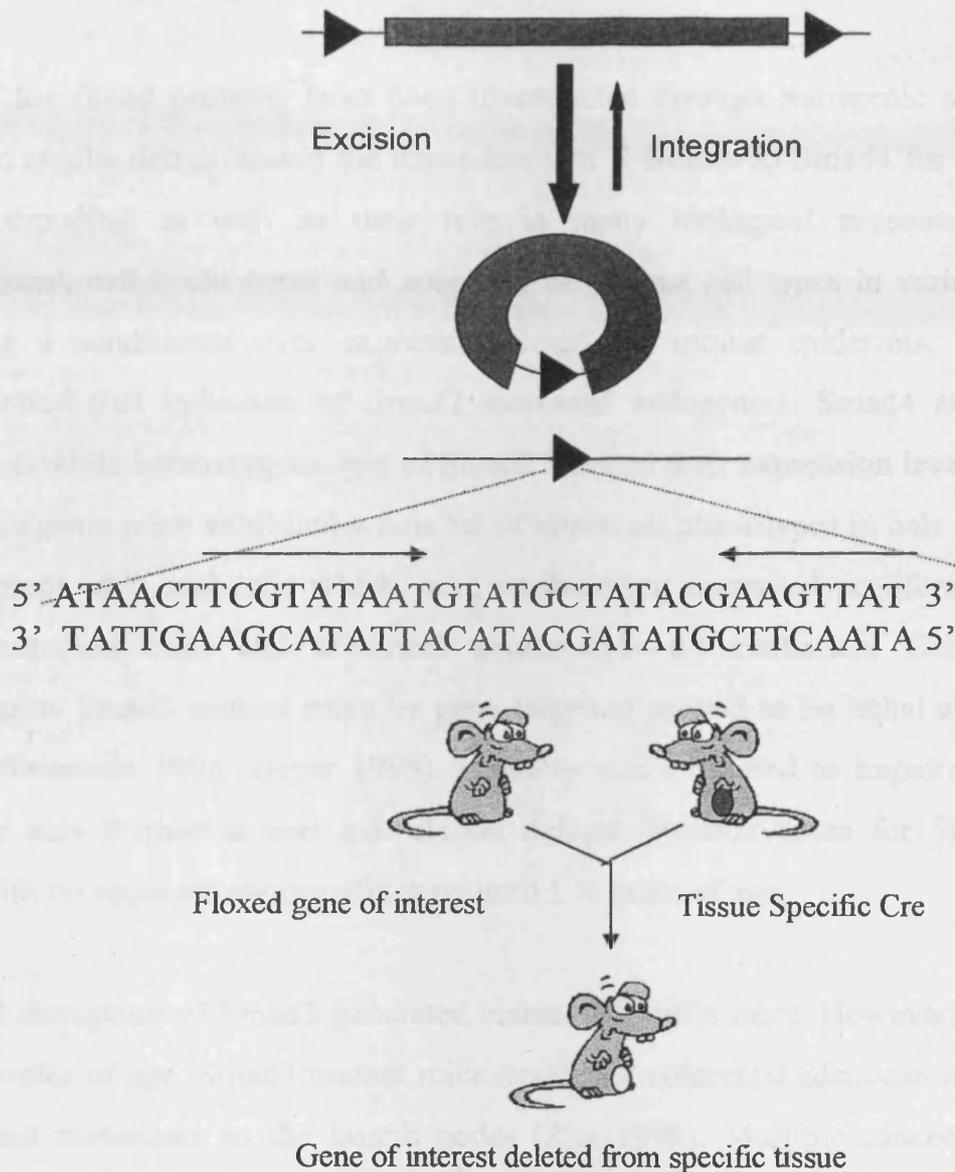
Cre is the 38kDa product of the *cre* gene of bacteriophage P1 (Sternberg, 1978) and is a site-specific DNA recombinase. Cre recombinase has the ability to recognise and mediate site-specific recombination between 34bp sequences referred as loxP (locus of crossover (x) in P1) (Sternberg, 1981). The loxP sequence consists of two 13 bp inverted repeats interrupted by an 8 bp nonpalindromic sequence which dictates the orientation of the overall sequence (Figure 1.4) (Hoess, 1984). When two loxP sites are placed in the same orientation on a linear DNA molecule, a Cre-mediated intramolecular recombination event results in the excision of the loxP flanked, or “floxed”, sequence as a circular molecule with one loxP site left on each reaction product (Figure 1.4). The reverse reaction, an intermolecular recombination event, will result in the integration of the circular DNA molecule into the linear molecule, each possessing a loxP site. Unlike many other recombinases no accessory factors or DNA topological requirements are needed for efficient Cre-mediated DNA recombination and this makes this system uniquely suited to genomic manipulation of eukaryotic cells (Sauer, 1988).

In general, the experimental outline of Cre-Lox mediated gene targeting consists of the targeted incorporation of loxP sites into the genome of ES cells via a corresponding targeting vector. After identification of homologous recombinants the final genetic modification is obtained by a Cre-mediated gene recombination event, either stably in cultured ES cells and the germline of mice derived from these ES cells or in a conditional manner in mice by the expression of Cre as transgene.

A particularly powerful feature of a conditional gene inactivation strategy using Cre is that the same loxP-tagged mouse can be used for gene ablation independently in a large number of different tissues, or at different developmental times, by simply mating it with a corresponding Cre transgenic that displays the desired tissue or temporal specificity of expression. Thus, the same genetically modified animal can be used to answer a variety of different questions relating to the expression and function of the target gene. Crucial to the success of this strategy is the prior careful evaluation of the pattern of expression of the Cre transgenic animal. Mosaic expression or exogenous tissue expression can defeat the purpose of the experiment and so careful choice of promoter must be considered. In our case the promoter of choice is the ovine

beta lactoglobulin (BLG) (Whitelaw 1992, Farini 1995, Clark 1992) which has shown that promotes over expression of target genes in the mammary gland from mid pregnancy to lactation.

An attractive alternative is the use of a viral-mediated gene transfer of Cre recombinase (Anton and Graham, 1995). This uses a basic strategy where sections of the wild type virus genome, such as sequences encoding the E1 and E3 genes are replaced with a Cre expression construct. Such recombinant viruses are replication deficient but still capable of host cell infection. With this approach, cultured cells can be exposed directly to the virus and very high efficiencies of infections can be achieved.



**Figure 1.4. Conditional Gene Targeting**

Cre recognises a 34 base pair site on the P1 genome called loxP (locus of X-over of P1) and efficiently catalyses reciprocal conservative DNA recombination between pairs of loxP sites. The loxP site consists of two 13 base pair inverted repeats flanking an 8 base pair non palindromic core region that gives the loxP site an overall directionality that, by convention, is depicted as an arrow. Cre mediated recombination between two directly repeated loxP sites results in excision of the DNA between them as a covalently closed circle. Correctly targeted ES cells are injected into mice and animals containing the modified gene (which are at this point phenotypically wild type), are then crossed with mice expressing Cre in the desired target tissue and Cre-mediated excision results in a tissue-specific gene ablation.

## 1.7 Smad Transgenic and Knockout mice.

Most of the Smad proteins have been investigated through transgenic or knockout mice and results demonstrated the dependency of R-Smads to Smad4 for appropriate TGF- $\beta$  signaling as well as their role in many biological processes such as development, cell cycle arrest and apoptosis of various cell types in various tissues. By using a conditional over expressing Smad2 in mouse epidermis, Ito (2001), demonstrated that induction of Smad2 increased endogenous Smad4 and TGF- $\beta$ 1 expression while heterozygous loss of Smad2 reduced their expression levels. Further, these transgenic mice exhibited a number of abnormal phenotypes in hair growth, ear development, skin and tails which were attributed to increased proliferation of the basal epidermal cells and abnormal keratinocyte differentiation. Generation of homozygous Smad2 mutant mice by gene targeting proved to be lethal at day E7.5-E10.5 (Weinstein 1998, Heyer 1999). Lethality was attributed to impaired anterior-posterior axis formation and gastrulation defects. Heterozygotes for Smad2 were fertile with no apparent abnormalities up until 1 ½ years of age.

Targeted disruption of Smad3 generated viable and fertile mice. However, between 4 and 6 months of age, Smad3 mutant mice developed colorectal adenocarcinomas with subsequent metastasis to the lymph nodes (Zhu 1998). Multiple cancers including (pancreatic, colorectal, lung, breast, prostate, ovarian, head and neck, esophageal, gastric bladder, hepatocellular and renal cell) have been shown to occur in mice with somatic mutations for Smad4 while homozygous Smad4 mutant mice die before day E7.5 due to failure to gastrulate and abnormal visceral endoderm development (Sirard 1998).

Finally, over expression of Inhibitory Smads leads to a number of abnormalities. Smad6 over expression in chondrocytes showed postnatal dwarfism and inhibition of proper phosphorylation in Smad 1, 5 and 8 (Horiki 2004). Further, over expression of Smad7 in mice epidermis caused a number of abnormalities in multiple epithelial tissues and caused death 10 days after birth (He 2002).

## 1.8 Aims.

The broad aim of this project is to further understand the role of Smad4 during mammary gland involution. Stat3 will be also utilised in analyzing the interplay with Smads as well as analyzed on its own. Finally, because the Methyl Binding Domain (MBD2) protein was implicated as a modulated target in preliminary studies I also analysed its potential role in controlling mammary gland development and involution. In order to achieve this goal I undertook the following approaches.

The first approach was to characterise a mammary gland-specific Smad4 transgenic mouse. This mouse contains a stable transgene consisting of the BLG promoter which drives expression of a human Smad4 cDNA. This transgene should therefore over express Smad4 in the mammary gland, and should be useful in addressing any possible role Smad4 plays in the lactation to involution progression and to what extent it is involved in the control of apoptosis in the gland. **This analysis is based on the hypothesis that “over expression of Smad4 will promote the anti-proliferative effects of TGF- $\beta$  by accelerating apoptosis in the involuting mammary gland”.**

As part of the same analysis, I performed a transcriptome microarray analysis from day 3 involuting mice. This time point was selected based on my phenotypic and molecular analysis as shown in Chapter 3. My aim in this analysis was to identify molecular targets with altered expression as a consequence of Smad4 over expression (Chapter 4). Identification of these molecular targets should then lead to a clearer picture of the interplay between the TGF- $\beta$  and other signaling pathways during involution and Programmed Cell Death.

Western blot analysis of Stat3 in the Smad4 transgenic mice identified an interaction of the two proteins. Based on this data I also expanded the mammary gland analysis by over expressing Smad4 in a Stat3 null environment. Conditional deletion of Stat3 from the mammary gland is achieved through use of the Cre-Lox recombination system. This model allowed an assessment of the potential interaction between Smad4 and Stat3 in involution of the mammary gland (Chapter 5). In the same chapter I also

analyzed Stat3 heterozygous mice by performing microarray experiments. This approach will be in a scaling manner for identifying differences in gene expression from Stat3 wild type and Stat3 heterozygous mice. For the purpose of these experiments I performed primary mammary cell cultures in order to avoid mammary epithelial cell contamination by adipocytes.

MBD2 was identified as a candidate target of Smad4 in preliminary studies. As a small project I also analyzed MBD2 knockout mice during involution in order to investigate any potential role played by this protein in normal involution (Chapter 4).

## **Chapter 2 -Materials and Methods.**

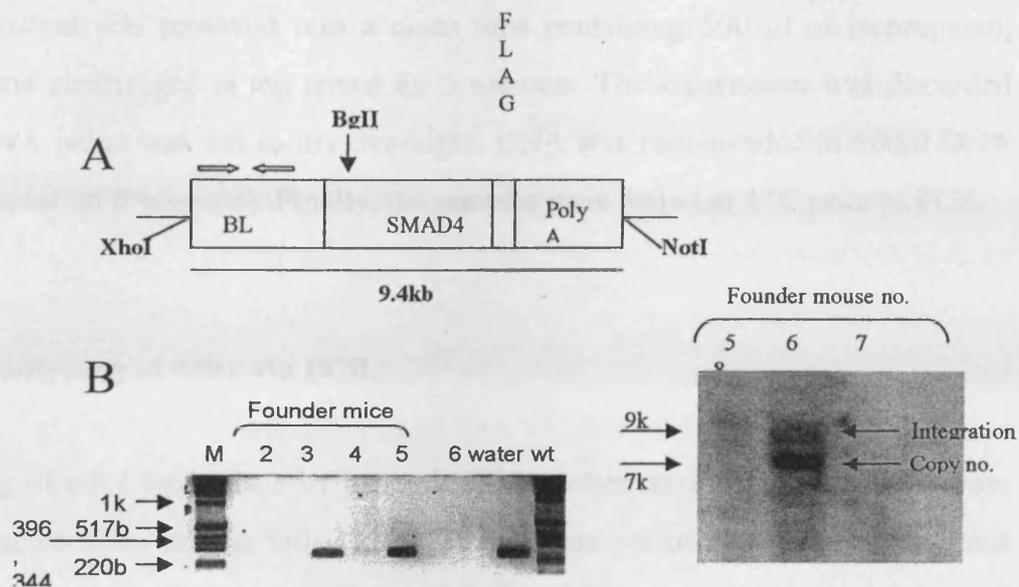
### **2.1 Mice.**

#### **2.1.1 Maintenance of Mice Colonies.**

Mice were a mixed colony of C57 B6/S129 (50%-50%) and maintained under non-barrier conditions.

#### **2.1.2 Generation of Smad4 Transgenics.**

The Smad4 transgenics were generated in Edinburgh by Eleonor Duff in the following way. cDNA coding for the human Smad4 gene was expressed under the control of a BLG promoter (Figure 2.1A). Two F1 founder mice were detected by PCR analysis for the presence of the BLG promoter (Figure 2.1B). One founder mouse was further tested for the number of integration sites and transgene copy number by Southern blotting (Figure 2.1C). Mice expressing BLG-Smad4 and wild types reached adulthood with no apparent abnormalities.



**Figure 2.1.** The generation of Smad4 transgenic mice. **A:** a schematic of the BLG-FLAG-Smad4 transgene. A 4.2kb BLG promoter fragment was fused to a FLAG-Smad4 expression construct. Restriction sites were used to linearize or excise the insert and prepare genomic DNA for Southern blot analysis. **B:** PCR analysis for the presence of the BLG promoter **C:** Southern blot analysis of transgenic genomic DNA. Transgenic mice were identified by Southern blot analysis using the 1682bp BLG promoter sequence as a probe (arrows).

## 2.2 Genotyping of Mice.

### 2.2.1 DNA Extraction from mouse tails.

DNA was extracted from tails using the PUREGENE DNA extraction kit. Tails were lysed overnight in 500  $\mu$ l of cell lysis solution (Puregene) and 10  $\mu$ l of proteinase K (20 mg/ml, Sigma), shaken at 37°C. Tails were left to cool at room temperature, 200  $\mu$ l of protein precipitation solution (Puregene) was added to each tube. These were vortexed and centrifuged at top speed for 5 minutes in a microfuge.

The supernatant was removed into a clean tube containing 500 µl of isopropanol, vortexed and centrifuged at top speed for 5 minutes. The supernatant was discarded and the DNA pellet was left to dry overnight. DNA was resuspended in 500µl DNA hydration solution (Puregene). Finally, the samples were stored at 4 °C prior to PCR.

### **2.2.2 Genotyping of Mice via PCR.**

Genotyping of mice took place by genomic PCR (Polymerase Chain Reaction) from DNA extracted from mouse tails. All PCR reactions performed in 50 µl volumes using 2 µl of the tail DNA preparation. All PCR reactions were carried out using MJ Research PCR machines (PTC-100). 15 µl of PCR products were analysed on 3% TBE agarose gel stained with ethidium bromide (10 mg/ml) (Sigma) and viewed under UV light.

### **2.2.3 BLG-SMAD4 PCR.**

The BLG-Smad4 PCR was designed in our lab. The protocol for a 50 µl reaction mix contained 5 µl PCR buffer (Sigma), 1 µl of each primer (10 pmoles per µl) (OSWELL), 1 µl of each dNTP (40 mM), 3 µl of MgCl<sub>2</sub> (Sigma), 2 µl of Taq polymerase (Sigma), 33 µl of autoclaved DDW (deionised distilled water), and 2 µl of DNA template. Reaction conditions were as follows: 94 C° 2:30 minutes, then 30 cycles of 94 C° 30 seconds (denaturation), 64 C° 30 seconds (annealing) and 72 C° 1 minute (extension) and 5 minutes of 72 C°.

The BLG-Smad4 transgene generated a 350bp band.

Primer1: 5' TCGTGCTTCTGAGCTCTGCAG 3'

Primer2: 5' GCTTCTGGGGTCTACCAGGAAC 3'

#### **2.2.4 BLG Cre Genotyping.**

The protocol for a 50 µl BLG Cre PCR reaction mix was as follows: 5 µl PCR buffer (Sigma), 0.1 µl of each primer (10 pmoles per µl) (OSWELL), 0.5 µl of each dNTP (40 mM), 3 µl of MgCl<sub>2</sub> (Sigma), 2 µl of Taq polymerase (Sigma), 37.3 µl of autoclaved DDW (deionised distilled water), and 2 µl of DNA template. Reaction conditions were as follows: 94C° 2 minutes, then 34 cycles 94C° 1 minute (denaturation), 55 C° 1 minute (annealing) and 72 C° 2 minutes (extension) and 10 minutes of 72 C°.

Wild type fragment generated no bands whilst the Cre fragment a 1Kb band.

Primer 1: 5'TGACCGTACACCAAATTTG 3'

Primer 2: 5 ATTGCCCTGTTTCACTATC 3'

#### **2.2.5 Stat3 Flox PCR.**

The protocol for a Stat3 Floxed 50 µl reaction mix contained 5 µl PCR buffer (Sigma), 1 µl of each primer (10pmoles per µl) (OSWELL), 1µl of each dNTP (40mM), 5µl of MgCl<sub>2</sub> (Sigma), 2µl of Taq polymerase (Sigma), 33µl of DDW, and 2µl of DNA template. Reaction conditions were as follows: 95 C° 2:30 minutes, then 34 cycles of 95 C° 30 seconds (denaturation), 67 C° 1 minute (annealing) and 72 C° 1 minute (extension) and 10 minutes of 72 C°.

The Stat3 Floxed PCR does not recognise null alleles. Floxed allele gives a 350bp band whilst wild type allele a 250bp band.

Primer1: 5`CACACAAGCCATCAAACCTCTGGTCTCC 3`

Primer2: 5`CCTGAAGACCAAGTATCCTGTGTGAC 3`

### 2.2.6 Stat3 Null PCR.

The protocol for a Stat3 Null 50 µl reaction mix contained 25 µl PCR buffer (EPICENTRE), 0.5 µl of each primer (10 pmoles per µl) (OSWELL), 0.5 µl of Taq polymerase (EPICENTRE), 21.5 µl of autoclaved DDW) deionised distilled water), and 2 µl of DNA template. Reaction conditions were as follows: 95 C° 3 minutes, then 37 cycles of 95 C° 30 seconds (denaturation), 72 C° 1 minute (annealing) and 72 C° 1 minute (extension) and 10 minutes of 72 C°.

The Stat3 Null PCR gives a 1Kb band.

Primer 1: 5' ATCGCCTTCTATCGCCTTCTTGACGAG 3'

Primer 2: 5' AGCAGCTGACAACGCTGGCTGAGAAGCT 3'

### 2.2.7 MBD2 PCR.

The protocol for the MBD2 wild type and null PCR reactions is as follows: 50 µl reaction mix contained 25 µl PCR buffer (EPICENTRE), 0.5 µl of each primer (10 pmoles per µl) (OSWELL), 0.5 µl of Taq polymerase (EPICENTRE), 21.5 µl of autoclaved DDW) deionised distilled water), and 2 µl of DNA template. Reaction conditions were as follows: 95 C° 3 minutes, then 35 cycles of 94 C° 30 seconds (denaturation), 56 C° 30 seconds (annealing) and 72 C° 1 minute (extension) and 10 minutes of 72 C°.

The MBD2 Wild type PCR gives a 428 bp product while null PCR gives a 292 bp product.

Primer 1 (wild type): 5' ACGCTGGCCTAGTGCCGTGC 3'

Primer 2 (wild type): 5' AAGAACAAGCAGAGAACTCCG 5'

Primer 1 (null): 5' TTGTGGTTGTGCTCAGTTC 3'

Primer 2 (null): 5' TCCGCAAACCTCCTATTTCTG 3'

### **2.3 Harvesting Mammary Glands.**

Mice were culled by cervical dislocation and mammary glands were taken by an incision made in the midline through the skin but not the body wall. Skin was peeled back on either side and pinned to reveal the mammary glands. The glands were removed and either fixed in formalin for paraffin embedding and histological analysis or snap-frozen in liquid nitrogen in an Eppendorf microfuge tube for protein analysis.

### **2.4 Mammary Gland Analysis.**

#### **2.4.1 Histological Analysis.**

Formalin fixed mammary glands transferred to 70% ethanol before embedded in paraffin wax. 3µm thick sections cut and placed onto histobond-coated slides. Slides stained with haematoxylin and eosin (H&E) for histological analysis.

#### **2.4.2 Wholemound Analysis.**

Mammary glands were dissected and placed onto glass slides. They were fixed in 2% Paraformaldehyde for a minimum of 2 hours. After fixation glands washed with PBS and dH<sub>2</sub>O for 5 minutes respectively. Slides stained with carmine aluminium overnight (1gr carmine, 2.5 gr Aluminium Potassium sulphate in 500mls of dH<sub>2</sub>O).

Stained slides washed with dH<sub>2</sub>O and dehydrated with 15 minute washes in 70%, 95% and 2x 100% alcohol respectively. After dehydration, slides were cleaned with Benzyl Benzate-Benzyl Alcohol (2:1) and covered with cover slips in DPX.

### 2.4.3 Peroxidase Staining of paraffin Embedded Tissues (Tunnel Analysis).

For measuring apoptotic cells I used a Peroxidase *In Situ* Apoptosis Detection Kit (Intergen) as follows.

- Tissue sections were deparaffinized by washes in:
  - 3x Xylene (5 min each)
  - 2x 100% Ethanol (5 min each)
  - 1x 95% Ethanol (3 min)
  - 1x 70% Ethanol (3 min)
  - 1x PBS (5 min)
  
- Following deparaffinization tissue sections were pretreated as follows:
  - Apply 60  $\mu$ l Proteinase K (20 ug/ml). Leave for 15 minutes at room temperature.
  - Wash 2x dH<sub>2</sub>O (2 min each)
  
- Endogenous Peroxidase was quenched as follows:
  - Apply 60  $\mu$ l Hydrogen Peroxide (3% in PBS) for 5 minutes at room temperature.
  - Wash 2x PBS (5 min)
  
- Gently tap off Excess liquid and add 75  $\mu$ l of Intergen's Equilibration Buffer directly on the specimen for 15 seconds.
- Gently tap off Excess liquid and add 55  $\mu$ l of Working Strength TdT enzyme for 1 hour at room temperature.
- Stop the reaction by adding the specimens into coplin jars containing STOP/Wash Buffer (1:34 dilution in dH<sub>2</sub>O). Agitate specimens for 15 seconds and incubate for 10 minutes at room temperature.

- Wash Specimens in 3 changes of PBS for 1 minute each wash. Tap off excess liquid and add 65  $\mu$ l of Anti-Digoxigenin Peroxidase conjugate directly to the specimens. Incubate for 30 minutes in a humidified chamber at room temperature.
- Following incubation wash specimens in 4 changes of PBS (2 minutes each wash), in a coplin jar at room temperature.
- Develop colour by adding 75  $\mu$ l Peroxidase substrate for 4 minutes at room temperature.
- Wash specimens in 3 changes of dH<sub>2</sub>O (1 minute each wash) in a coplin jar.
- Counterstain with 60  $\mu$ l of 0.5% Methyl Green for 4 minutes at room temperature.
- Wash specimens in 3 changes of dH<sub>2</sub>O by dipping the slides 10 times each in the first and second washes, followed by 30 seconds without agitation in the third wash.
- Wash specimens in 3 changes of 100% N-BUTANOL in a coplin jar, dipping the slides 10 times each in the first and second washes, followed by 30 seconds without agitation in the third wash.
- Dehydrate specimens by moving the slides through three coplin jars of XYLENE (2 minutes each wash).
- Mount specimens under a coverslip in DPX (Distyrene Plasticizer Xylene) and view under a microscope in a 100x magnification.

#### 2.4.4 P21 Staining.

P21 immunohistochemistry was performed by facilitating the “ABC” kit (Vector Laboratories). This procedure employs biotinylated antibody and a preformed Avidin: Biotinylated enzyme Complex and has been termed the “ABC” technique. The first step of the procedure is to incubate the section with a primary antibody (in this case rabbit anti-p21) raised against the antigen of interest. Next, a biotin-labelled secondary antibody is added, which in this case is biotinylated anti-rabbit IgG. This introduces many biotins into the section at the location of the primary antibody. The avidin:

biotinylated enzyme complex (ABC) is then added and binds to the biotinylated secondary antibody. In the last step of the procedure, the tissue antigen is localized by incubation with a substrate for the enzyme.

More specifically:

- Tissues were deparaffinized by washes in:
  - 3x Xylene (5 min each)
  - 2x 100% Ethanol (5 in each)
  - 1x 95% Ethanol (3 min)
  - 1x 70% Ethanol (3 min)
  - 1x PBS (5 min)
  
- Block peroxidase 15 minutes at room temperature.

Note: Use a 20 L stock containing:

- a) 83.2gr Citric acid (Sigma)
- b) 215.2gr diSodium hydrogen Phosphate (Sigma)
- c) 20gr Na Azide (Sigma)

to whatever volume required to this stock add H<sub>2</sub>O<sub>2</sub> to a final 1.5% solution.

- Retrieve the antigen by boiling the sections for 20 minutes in Na Citrate 10Mm, pH 6.0.
- Once boiled transfer slides to a preheated plastic coplin jar containing same buffer and cool it down by submerging under cold running water.
- Wash 2-3x PBS
- Block sections for 30 minutes with 5% goat serum in PBS
- Wash 2x PBS
- Incubate slides with Santa Cruz anti-p21 (M19) diluted 1/500 in 5% goat serum/PBS for 1 hour at room temperature.
- Wash PBS 3x for 5-10 minutes each.
- Block 10 minutes with 5% goat serum in PBS

- Incubate slides for 30 minutes with biotinylated secondary antibody from rabbit ABC kit (Vector Laboratories)
- Wash PBS 3x for 5-10 minutes each.
- The staining was visualised with an ABC-HRP (horse radish peroxidase) kit (Vectalabs) and then DAB (3,3 diaminobenzedene) (DAKO), a chromogen which makes a chemical reaction leading to brown colouration in positive areas.

## **2.5 Protein Extraction from Mammary Gland.**

- Remove mammary glands from sacrificed animals and snap frozen at  $-70^{\circ}\text{C}$  until extraction. Crush a frozen piece of tissue using a mortar pellet in liquid Nitrogen until it becomes a fine powder.
- Use liquid nitrogen and pour this powder into a universal and leave liquid nitrogen to evaporate off.
- Add  $400\mu\text{l}$  of RIPA buffer and shear the tissue with a 21G needle. Centrifuge for 15 minutes at full speed at  $4^{\circ}\text{C}$ . Remove supernatant, aliquot, snap frozen into liquid nitrogen and store at  $-70^{\circ}\text{C}$  until use.

### **RIPA Buffer**

500mM Tris HCL pH 7.5 (Fisher)

150mM NaCl (Fisher)

1% Nonidet p40 (Sigma)

0.5% sodium deoxycholate (Sigma)

0.1% SDS (Fisher)

To 10ml of RIPA buffer add 1 mini-cocktail tablet (Roche) of protease inhibitors.

### **2.5.1 Determination of protein concentration.**

- For estimation of protein concentration a microscale variant of Bradford's dye of binding method was used (Bradford 1976). Add 5-10 $\mu$ l of protein sample to 1ml of Bradford reagent (BioRad), mix and allow to stand for 10 minutes before measurement at  $A_{595}$ . In each case perform absorbance measurements against blanks containing an equal volume of the Bradford reagent.
- Correlate absorbance measurements with a freshly generated calibration curve (0-25  $\mu$ g of bovine serum albumin in 1ml of bovine serum albumin in 1 ml Bradford Reagent) to estimate the protein concentration of the unknown sample.

### **2.5.2 Western Analysis.**

- For Western analysis, run proteins on denaturing polyacrylamide gels (10% gels for > 40 kDa proteins and 15% gels for < 40 kDa proteins).
- Equalise protein samples with RIPA buffer so that all samples are 20  $\mu$ g and of equal volume (20  $\mu$ l). Boil samples for 5 minutes in 4 x Loading Buffer containing  $\beta$ -mercaptoethanol, quench on ice, centrifuge briefly and load onto the gel. Run gels for 2 hours at 125V in running buffer or until the protein markers (Gibco) separate.
- Blot gels on PVDF (Polyvinylidene Difluoride) membrane (Millipore) in transfer buffer overnight at 15 mA or for 2 hours at 200 mA. Prior to transfer, soak PVDF membrane in methanol (Fisher) for 30 minutes.

- After transfer, block membranes in TBS/0.1% Tween/10% Marvel (TTM) for one hour. All blots were incubated with primary antibodies for one hour at room temperature. The Stat3 antibody SC-482 (rabbit polyclonal) (Santa Cruz) was used at a concentration of 1 in 500 whilst the Stat5a antibody (rabbit polyclonal) SC-1081 (Santa Cruz) was used at a concentration of 1 in 1000. P21 (Pharmigen 556431) (rabbit polyclonal) used at a concentration of 1:50. P27 (Neo Marker) (rabbit polyclonal) used at a concentration of 1:500. Bax (Santa Cruz) SC-493 (rabbit polyclonal) used at a concentration of 1:500. Vitamin D receptor (Santa Cruz) SC-1008 used at a concentration of 1:250. Actin (Santa Cruz) H-196 used at a concentration of 1:1000.
- Wash blots 3 times in TTM (10 minutes each) and incubate in secondary antibody (Santa Cruz): anti-rabbit-HRP at the following concentrations: for Stat3 1:1000, for Stat5a 1:2000, for p21 1: 1000, for p27 1: 1000, for Bax 1: 1000, for Vitamin D receptor 1:2000, Actin 1:1000.
- Wash blots with TBS (10 minutes each) and visualise using ECL plus (Amersham) on ECL film (Amersham). To confirm equal loading after blotting, stain blots with Ponso Red (Sigma).

## **Reagents**

15% Loading Gel	10% loading Gel	10% Stacking Gel
3.33ml DDW	6.65ml DDW	3.57ml DDW
11.69ml 30% acrylamide (1:29)	8.35ml 30% acrylamide (1:29)	1.70ml 30% acrylamide (1:29)
9.37ml 1M Tris HCL pH 8.8	9.37ml 1M Tris HCL pH 8.8	0.62ml 1M Tris HCL pH6.8
250µl of 10% SDS (Fisher)	250µl of 10% SDS	50µl of 10% SDS
72µl of 25% APS (Fisher)	72µl of 25% APS	33µl of 25% APS
13.2µl of Temed (Sigma)	13.2µl of Temed	3.6µl of Temed

### **4x Loading Buffer**

200 mM Tris HCl pH 6.8

400 mM Dithiothreitol (DTT)

8% SDS

0.4% Bromophenol blue

40% Glycerol

### **10x Running Buffer: For 1 L**

30.2 g Tris

188 g Glycine (Fisher)

### **Transfer Buffer For 1L:**

800 ml DDW

200 ml methanol

2.9 g Tris

14.5 g Glycine

## 2.6 Primary Mammary Cell Culture.

- Dissect mammary glands from 16-18 day pregnant mice and place them in a sterile 50 ml falcon tube on ice.
- Finely chop mammary glands with sterile scalpels and resuspend in fresh collagenase/trypsin digestion mix (980 mg of F10 powered medium-Sigma, 150 mg of trypsin-Gibco, 300 mg collagenase A-Gibco, and 5 ml of Foetal Calf Serum in a final volume of 100 mls with distilled water).
- Filter the collagenase digestion mixture through a 0.2  $\mu\text{m}$  sterile filter before use and store at 4 °C until required.
- Place minced glands in autoclaved, sterile 50 ml polypropylene tubes and digest for 90 minutes on a rotary shaker, 250 rpm at 37 °C. Use 4 mls of digestion mixture per gram of tissue.
- Remove the digested cell suspension to a 50 ml tube and spin for 30 seconds at 100 rpm. Remove the supernatant on a new 50 ml tube on ice. Resuspend the resulting undigested pellet in the remaining digestion mixture for a further 30 minutes. After 30 minutes, spin the tube for 30 seconds at 100 rpm and transfer the supernatant to the same 50 ml tube on ice.
- Centrifuge the supernatants at 800 rpm for 3 minutes and discard the aqueous solution.
- Resuspend the pellet in 5 ml of F12 medium (containing gentamicin) and make the volume in the tube to 50 ml with F12 medium/gentamisin (50  $\mu\text{g}/\text{ml}$ ).

- Spin the tubes for 3 minutes at 800 rpm and discard the supernatant.
- Resuspend the pellet in F12 medium containing gentamisin (50 µg/ml) and 10% foetal calf serum with the following hormones; hydrocortisone 1µg/ml (Sigma stock solution: 1 mg/ml in 100% ethanol), insulin 5 µg/ml (Sigma stock solution: 5 mg/ml in 5mM HCL), epidermal growth factor 5 ng/ml (Promega stock solution: 5 µg/ml in F12 medium) and fetuin (Sigma 1 mg/ml).
- Seed cells on tissue culture plastic at a density of  $2.4 \times 10^6$  cells/ml and let grow for 48 hours at 37 °C in a humidified atmosphere of 5% CO<sub>2</sub>.

## **2.7 Total RNA extraction from mammary gland.**

- Dissect mammary gland and immediately transfer to 5 mls of RNA later solution (Sigma). Just before homogenisation remove RNA later solution, weight sample and add 1 ml of Trizol (Invitrogen) per 100 mgs of sample.
- Homogenise mammary gland in a homogeniser until no large pieces of tissues are visible. Between homogenisations, clean homogeniser with a) 4M NAOH, b) 75% Ethanol (treated with di-ethylPyrocarbonate (DPC)) and c) with nuclease free water (DPC treated) for 15 seconds respectively.
- If necessary aliquot the homogenised sample in 1 ml volumes in sterile, RNase free 2 ml centrifuge tubes and keep on ice.
- Spin samples for 10 minutes at maximum speed and transfer the supernatants to new 2 ml RNase free centrifuge tubes.

- Add 200  $\mu$ l of cold chloroform to each sample. Shake samples vigorously for 10 seconds for mixing and chill in ice for 5-10 minutes.
- Spin samples for 15 minutes at maximum speed. After spinning, transfer the top aqueous phase which contains the RNA to a new RNase free 2 ml tube.
- To precipitate the RNA, add an equal volume of cold isopropanol, mix by inversion and leave in ice for 45-60 minutes.
- Following incubation, spin samples for 15 minutes at maximum speed. Discard the supernatant and wash the RNA pellet with 1ml of cold 75% ethanol (DPC treated).
- Spin samples for 15 minutes at maximum speed.
- After spinning, discard the supernatant and leave the pellets to air dry for 5-10 minutes.
- Resuspend RNA in 100  $\mu$ l of nuclease free water and clean it with a Qiagen RNeasy clean up mini kit (see below).
- Evaluate total RNA on a 2% agarose, 1X TBE gel for the presence of a 28S and 18S rRNA bands. Samples with an almost double intensity of the 28S band compared to the intensity of the 18S band are characterised as useful samples and can be kept for quantitative analysis.
- Quantitate RNA concentration by UV spectrophotometry at an absorbance of 260 and 280 nm. RNA samples with a 260/280 ratio of 1.7-2.1 can be considered useful and used for RT-PCR and In Vitro Transcription (IVT).

### **2.7.1 RNA Extraction from Monolayered Cell Cultures.**

- Lyse cells directly on Petri dishes by adding 1ml of Trizol (Invitrogen) reagent per  $10^7$  cells and pipette up and down until no clumps are visible. Lyse cells further through a 21G (or smaller) needle. Transfer cells to sterile 14 ml round-bottom polypropylene tubes and incubate for 5 minutes before splitting into 1 ml aliquots in 2 ml sterile round-bottom tubes.
- Add 0.2 ml of chloroform to each tube and mix vigorously by hand for 15 seconds. Incubate tubes at room temperature for 3 minutes.
- After incubation, centrifuge tubes at 12000g for 15 minutes at 4C°.
- Transfer the upper aqueous phase (~0.6ml) to a fresh tube and add 0.5 ml of isopropanol. Leave for 10 minutes at room temperature.
- After incubation, centrifuge samples at 12000g for 15 minutes at 4C°.
- Remove the supernatant and wash tubes with 1ml of 75% ethanol in DEPC treated sterile water. Mix tubes by vortexing and centrifuge for 5 minutes at 7.500g at 4C°.
- Remove the supernatants and leave the pellets to air dry for 5-10 minutes.
- Resuspend RNA pellets in 50  $\mu$ l DEPC treated sterile water.
- Analyse RNA quantitatively by spectrophotometry and qualitatively by electrophoresis of 2  $\mu$ g of RNA on a 1% agarose 1X TBE gel.

- Clean RNA samples by using the Qiagen RNeasy clean up kit (see below).

### 2.7.2 RNA Cleanup using the Qiagen RNeasy Kit.

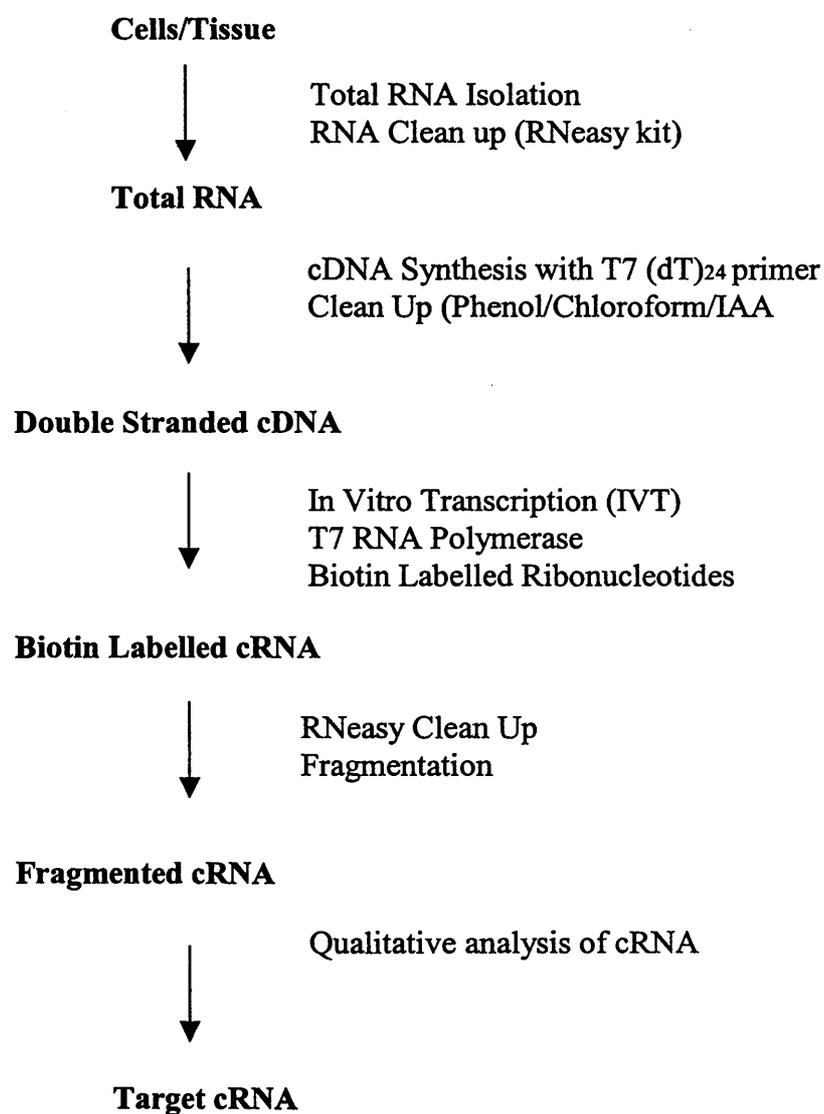
Note: Use a maximum of 100 µg of RNA for this cleanup protocol. This amount corresponds to the binding capacity of the RNeasy mini columns.

- Add 10 µl of β-mercaptoethanol per 1ml of RTL buffer.
- Adjust RNA volume to 100 µl with RNase-free water. Add 350 µl of RTL buffer and mixed thoroughly.
- Add 250 µl of absolute ethanol to the diluted RNA and mix thoroughly.
- Apply the 700 µl sample to an RNeasy mini column placed in a 2 ml collection tube. Close tubes gently and centrifuge for 15 seconds at 10000rpm. Discard the flow-through and collection tube.
- Transfer the RNeasy mini column to a new collection tube and add 500 µl of RPE buffer. Close tubes gently and centrifuge for 15 seconds at 10000 rpm. Discard the flow- through.
- Add another 500 µl of buffer RPE to the RNeasy mini column. Close tube gently and centrifuge for 2 minutes at 10000rpm to dry the RNeasy silica-gel membrane. Discard the flow-through and collection tube.
- Transfer the RNeasy mini column to a 1.5 ml (RNase free) collection tube. Add 50 µl of RNase-free water directly onto the RNeasy silica-gel membrane. Close tubes gently and centrifuge for 1 minute at full speed.

- Measure new RNA concentrations by UV spectrophotometry as described above.

### 2.7.3 Target Preparation for Affymetrix GeneChip.

#### OUTLINE OF PROCEDURE



- After incubation add 2  $\mu\text{l}$  of Superscript II Reverse Transcriptase (200 U/ $\mu\text{l}$ ) in each reaction tube.
- Incubate tubes for 1 hour at 42C°.

### Second Strand cDNA Synthesis

- Prepare a master mix per reaction tube on ice with the following reagents:

91  $\mu\text{l}$  DEPC treated sterile water

30  $\mu\text{l}$  (5x) Second Strand Buffer

3  $\mu\text{l}$  (10mM) dNTPs

1  $\mu\text{l}$  *E.coli* DNA Ligase (10U/ $\mu\text{l}$ )

4  $\mu\text{l}$  *E.coli* DNA Polymerase I (10U/ $\mu\text{l}$ )

1  $\mu\text{l}$  *E.coli* RNase H (2U/ $\mu\text{l}$ )

**130  $\mu\text{l}$  Total**

<p><b>Superscript Double Stranded cDNA Synthesis Kit (Invitrogen)</b></p>
---

- Mix the master by vortexing. Add 130  $\mu\text{l}$  of master mix to each first strand reaction tubes and mix by pipetting.
- Incubate tubes at 16 C° for 2 hours.
- After the incubation period, add 2  $\mu\text{l}$  (10U) of T4 DNA Polymerase to each reaction tube and incubate for 5 minutes.
- Finally, add 10  $\mu\text{l}$  of 0.5M EDTA (Sigma) to each reaction tube.
- Place tubes on ice and proceed to “Clean Up of double stranded cDNA” or store at -80C°.

## **Clean Up of Double Stranded cDNA.**

Note: This method uses Phase-Lock tubes (Eppendorf) which contain a gel that separates the aqueous and organic phases of the phenol extraction allowing better recovery of the aqueous phase.

- Spin phase-lock tubes in a microfuge at maximum speed for 30 seconds.
- Add 160  $\mu$ l of room temperature buffer saturated phenol/chloroform/IAA (Ambion) to the cDNA reaction. Vortex tubes briefly.
- Place mixture into the phase-lock tube and spin at full speed for 2 minutes.
- Transfer the upper (aqueous) phase to a new tube and then add the following reagents:

0.5 volumes (approximately 80  $\mu$ l) 7.5M Ammonium acetate (Sigma)

4 $\mu$ l Glycogen (5 mg/ml) (Ambion)

2.6 volumes (approximately 600  $\mu$ l) 100% room temperature ethanol

- Mix tubes by tapping and spin at room temperature at maximum speed for 20 minutes.
- Remove the supernatants from the tubes without dislodging the pellets.
- Add 160  $\mu$ l of cold 80% ethanol (DPC treated) to the pellets and spin tubes at maximum speed for 5 minutes in room temperature.
- After spinning, discard the supernatants and repeat the 80% ethanol washing step.
- Discard the supernatants and leave the pellets to air dry for 5-10 minutes.

- After drying, resuspend pellets in 12  $\mu$ l of DEPC treated sterile water.
- Place tubes on ice in order to proceed to the in vitro transcription reaction or store at  $-80^{\circ}\text{C}$ .

### **Synthesis of cRNA-In vitro transcription.**

Note: All reagents used for the in vitro transcription are from the Enzo BioArray High Yield Transcript Labelling Kit (Affymetrix).

- Thaw all reagents and double stranded cDNA at room temperature.
- A master mix for each reaction tube was prepared at room temperature as follows:

10  $\mu$ l DEPC treated sterile water  
 4  $\mu$ l (10x) HY Reaction Buffer  
 4  $\mu$ l Biotin-labelled ribonucleotides  
 4  $\mu$ l DTT  
 4  $\mu$ l RNase Inhibitor Mix  
 2  $\mu$ l T7 RNA Polymerase  
**28  $\mu$ l Total**

- Add 28  $\mu$ l of the master mix to each of the 12  $\mu$ l cDNA previously prepared and mix by pipetting.
- Incubate reaction tubes at  $37^{\circ}\text{C}$  for 5 hours. Gently mix by tapping every hour.
- After the incubation, clean up reaction tubes or store to  $-80^{\circ}\text{C}$ .

## Clean Up of cRNA.

Note: The clean up process utilizes the Qiagen RNeasy Mini Kit.

- To the in vitro reaction tubes add the following:

60  $\mu$ l DEPC treated sterile water

350  $\mu$ l RTL Buffer (with 10 $\mu$ l  $\beta$ -mercaptoethanol/ml)

250  $\mu$ l 100% ethanol.

- Mix samples by pipetting and apply them to the RNeasy mini columns.
- Spin columns at maximum speed for 15 seconds.
- Discard the flow-through from the collection tubes.
- Add 500  $\mu$ l of RPE to each of the columns.
- Spin columns for 15 seconds at maximum speed.
- Discard the flow-through from the collection tubes and add 500  $\mu$ l of RPE buffer.
- Spin columns for 2 minutes at maximum speed.
- Transfer columns to fresh tubes and spin for 1 minute at maximum speed.
- Transfer spin columns to new collection tubes in order to prevent any possible RPE carry over.

- Elute cRNA by placing 50  $\mu$ l of DEPC treated sterile water on the middle of the membrane of the spin column and incubate at room temperature for 4 minutes.
- Spin columns for 1 minute at maximum speed.
- Take the eluate and re-apply it to the same spin column. Incubate columns for 4 minute at room temperature.
- After the incubation period, spin columns for 1 minute at maximum speed.
- Collect cRNA and measure its quantity by UV spectrophotometry as described above.

### **cRNA Fragmentation.**

Note: 25  $\mu$ g of cRNA is required for fragmentation.

- Prepare the following master mix for each fragmentation reaction:

X  $\mu$ l cRNA (25 $\mu$ g)

Y  $\mu$ l DEPC treated sterile water

10 $\mu$ l (5x) Fragmentation Buffer\*

**50  $\mu$ l Total**

- Incubate tubes at 94C° for 35 minutes in a heated block.
- After the incubation period, place tubes on ice and store at -80 C°.

**\*(5X) FRAGMENTATION BUFFER**

4 ml 1M Tris-acetate pH 8.1 (Trizma base. pH adjusted with glacial acetic acid.)

0.64 gr Magnesium acetate (Sigma)

0.98 gr Potassium acetate (Sigma)

DEPC treated sterile water to a total of 20 ml.

- Mix solutions thoroughly and filter through a 0.2  $\mu\text{m}$  vacuum filter unit. Store at room temperature.

**2.8 RT-PCR. Primers and PCR Programs.**

- The following were placed in a microfuge tube:

X  $\mu\text{l}$  (1 $\mu\text{g}$ ) total RNA in DEPC treated sterile water

Y  $\mu\text{l}$  DEPC treated sterile water

**9  $\mu\text{l}$  Total**

- Mix the above gently by pipetting and incubated for 10 minutes at 70 C<sup>o</sup> in a PCR machine.

- Prepare a master mix per reaction tube on ice with the following reagents:

2  $\mu\text{l}$  Primers (Random Hexamers)

4  $\mu\text{l}$  (5x) First Strand Buffer

2  $\mu\text{l}$  DTT (0.1M)

1  $\mu\text{l}$  dNTPs (10mM)

**9  $\mu\text{l}$  Total**

<b>Superscript Double Stranded cDNA Synthesis Kit (Invitrogen)</b>
--

- Mix gently by pipetting and add 9  $\mu\text{l}$  of master mix to each reaction tube.
- Incubate reaction tubes at 42 C<sup>o</sup> for 2 minutes in a heated block.

- After 2 minutes, add 1  $\mu$ l of Superscript II Reverse Transcriptase (200 U/ $\mu$ l) in each reaction tube.
- Incubate tubes for 1 hour at 42C<sup>o</sup>.
- Inactivate enzyme at 70C<sup>o</sup> for 10 minutes.
- Use 2  $\mu$ l of cDNA for PCR by utilising primer sequences and Annealing Temperatures as given in Table 2.1

Gene	Primer Sequence	Annealing Temperature	PCR Cycles	Product Size (bp)
<b>ADRB</b>	5' AAGAATAAGGCCCGAGTGGT 3' 5' GTCTTGAGGGCTTTGTGCTC 3'	59 <sup>o</sup> C	27	212
<b>ALS</b>	5' ACTCAGTTTGGGCAACAACC 3' 5' CGGTCCAGGTAGAGCTTCTG 3'	59 <sup>o</sup> C	27	168
<b>CLDN1</b>	5' GATGTGGATGGCTGTCATG 3' 5' CGTGGTGTTGGGTAAGAGGT 3'	59 <sup>o</sup> C	27	250
<b>GLP2</b>	5' AAGCAGCCAGAAGCAGACTC 3' 5' TAAGACCGTCCTGGAGCACT 3'	58 <sup>o</sup> C	27	228
<b>NUPRI</b>	5' TCTGCTTCTTGCTCCCATCT 3' 5' CAGACCACAGACACCACACC 3'	60 <sup>o</sup> C	27	155
<b>TBL3</b>	5' CACAGCCTCTGCTGATGGTA 3' 5' CTTCTGCCTGCTCAGCTTCT 3'	59 <sup>o</sup> C	27	317
<b>VDR</b>	5' GAGGTGTCTGAAGCCTGGAG 3' 5' ACCTGCTTTCCTGGGTAGGT 3'	59 <sup>o</sup> C	27	155
<b>P21</b>	5' CCCGTGGACAGTGAGCAGT 3' 5' GGGCACTTCAGGGTTTTCTCT 3'	59 <sup>o</sup> C	30	450
<b>HPRT</b>	5' GTCAAGGGCATATCCAA 3' 5'CGTGCTGGATTACATTAAGCA 3'	57 <sup>o</sup> C	35	350

**Table.2.1.** Primer sequence, PCR annealing temperatures, PCR cycles and product sizes for RT-PCRs.

## **Chapter 3 – *In Vivo* Characterisation of a BLG-Smad4 Transgenic Line in the Mammary Gland.**

### **3.1 Introduction.**

#### **3.1.1 TGF- $\beta$ and Breast Cancer. A Dual Role.**

TGF- $\beta$ s are characterised as tumour suppressors in the mammary gland based on the following observations; TGF- $\beta$ s are expressed during all stages of mammary development except lactation (Robinson 1991), and they show potent growth inhibitory effects on mammary epithelial cells both *in vivo* and *in vitro* (Hosobuchi 1989). TGF- $\beta$  appears to regulate normal ductal and alveolar development in the mammary gland (Pierce 1993, Jhappan 1993). It has also been demonstrated that many breast cancer cell lines exhibited absent or decreased TGF- $\beta$  responsiveness (Reiss 1997). Finally, early premalignant breast lesions showed a decreased expression of TGF- $\beta$  type II receptors and therefore increased probability of an invasive disease (Gobbi 1999).

The growth inhibition by TGF- $\beta$  on epithelial cells has been characterised extensively and, depending on the cell system and experiments, has been shown to be concomitant with apoptosis (Rotello 1991, Bursch 1993). Some of TGF- $\beta$  actions are the induction of *CD95/Fas* (Ashley 1998), down-regulation of *c-Myc* and up-regulation of *p27* (Kim 1998), down regulation of *NF-kB* (Azuma 1999), down regulation of *Bcl-2* (Tsukada 1995), down regulation of *p21* and increased expression of *p53* (Yan and Sage 1998), induction of *Bax* (Teramoto 1998), induction of G1 arrest (Motyl 1998) and practically every other permutation of genes involved in cell cycle or cell death. Thus, TGF- $\beta$ - mediated growth inhibition and apoptosis can be correlated with a general function of a tumour suppressor within this tissue.

However, contrary to the above, it has also been demonstrated that TGF- $\beta$ s can have oncogenic activities under certain circumstances. During multistage tumorigenesis, TGF- $\beta$  growth inhibitory and apoptotic effects are lost, frequently by subversion of the normal signaling pathway due to activation of other signaling molecules including *PI3K* and *Ras* (Derynck 2001, Akhurst 2001, Wakefield 2002). Moreover, as tumour cell progress, they secrete ever-increasing quantities of TGF- $\beta$ 1 and become more migratory and invasive (Derynck 2001, Akhurst 2001, Wakefield 2002, Oft 2002).

### 3.1.2 Smad4 and Apoptosis.

Smad4 or DPC4 (Deleted in Pancreatic Carcinoma, locus 4) plays a key role in TGF- $\beta$  signaling thereafter participates in the regulation of cell proliferation, differentiation and apoptosis. It has been characterised as a candidate tumor suppressor gene due to its frequent inactivation/mutation in pancreatic and colorectal cancers (Hahn 1996, Riggins 1996). Results from several studies have suggested that defects in Smad4 expression play a significant role in the malignant progression of tumors (Wilentz 2000, Takaku 1998) while expression of Smad4 in Smad4-defective tumor cell lines has been shown to restore TGF- $\beta$  signaling and induce cell cycle arrest and apoptosis (Dai 1999a). In addition to Smad4's antiproliferating activity it has been demonstrated that expression of Smad4 in tumor cell lines inhibits expression of the vascular endothelial growth factor (VEGF) and enhance the levels of the angiogenesis inhibitor thrombospondin-1 (TSP-1), causing cells to switch from potentially angiogenic to antiangiogenic both *in vivo* and *in vitro* (Schwarte-Waldhoff 1999). Further evidence for Smad4 significance has been demonstrated in the following studies. Expression of Smad4 in Smad4-null breast carcinoma cell lines restored TGF- $\beta$  signaling and showed reduced efficiency of colony formation in soft agar and a modest reduction of cells in the S phase without induction of apoptosis under normal culture conditions. Interestingly, induction of apoptosis by Smad4 expression was demonstrated in cell suspension in a process termed anoikis (Ramachandra 2002). In another paper, Yoo (2003), demonstrated that GADD45b is a positive mediator of TGF- $\beta$  induce apoptosis and that Smad2, 3 and Smad4 are responsible for its activation. In a paper

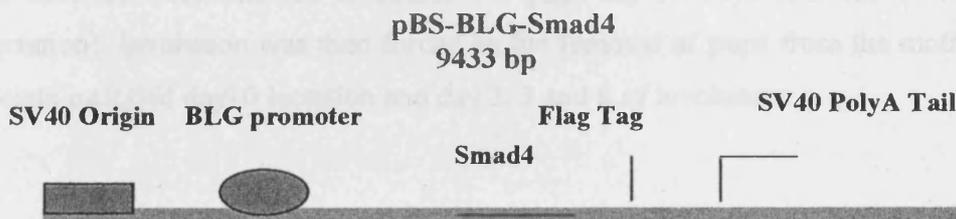
by Narula (2002), conditional overexpression of Smad4 in the testis demonstrated amongst other apoptosis of germ cells and spermatogenic arrest illustrating once more Smad4 as an important regulator of apoptosis. Takaku (1998) demonstrated that Smad4/APC double mutant heterozygote mice developed tumours while single heterozygotes did not. Finally, Lee (2001) demonstrated that decreased Smad4 mRNA expression is associated with defective TGF- $\beta$  response and growth inhibition and that activation of PAI-1 by TGF- $\beta$  is Smad4 dependent.

### **3.1.3 The Transgenic Approach.**

In order to investigate the role of Smad4 in mammary epithelial cell apoptosis it seemed appropriate to attempt a transgenic approach to the problem. I decided to address this question via tissue specific over expression of the human Smad4 gene through a transgenic model, termed BLG-Smad4. These mice express higher than normal levels of human Smad4 protein during mammary gland lactation and involution and were previously constructed in Edinburgh by other workers. To achieve over expression of Smad4 in mouse mammary epithelial cells we chose the BLG (beta-lactoglobulin gene) promoter.

The BLG gene has been extensively characterised and its promoter has been successfully used in the past to target gene expression to secretory epithelial cells of the mammary gland (Whitelaw 1992, Farini 1995, Clark 1992). The gene encoding the milk protein BLG in sheep is expressed in the mammary gland in a tissue-specific manner during pregnancy and lactation. Expression is induced mid-way through pregnancy, peaks during lactation and begins to decline at the onset of involution. The unmodified sheep gene behaves appropriately in transgenic mice, and it has been shown that many of the cis-acting elements that mediate this pattern of expression are located in the proximal 400bp of the promoter (Clark 1998). Expression patterns of the BLG are also well characterised and it has been demonstrated that BLG-driven transcription is correctly initiated in mice and that BLG-driven synthesis is restricted to the secretory epithelial cells of the mammary gland (Harris 1991). The BLG-

Smad4 transgene vector used to generate our transgenic mice is illustrated in Figure 3.1.



**Figure 3.1** The complete transgene vector of the ovine beta-lactoglobulin promoter upstream of a human FLAG-tagged human Smad4.

### 3.1.3 Aims.

The hypothesis that Smad4 is an essential component of the initiation of epithelial apoptosis and thereby in the normal involution of the gland implies that over-expression of Smad4, either before or during this period, “may lead to acceleration in the involution programme”. In this part of the project I chose to identify any possible alterations in mammary epithelial contribution during involution by histological comparison of BLG-Smad4 mice against wild types. Levels of apoptosis were analysed by performing TUNEL analysis and finally, molecular targets were analysed through Western blotting in order to identify any possible alterations in their expression following Smad4 over expression.

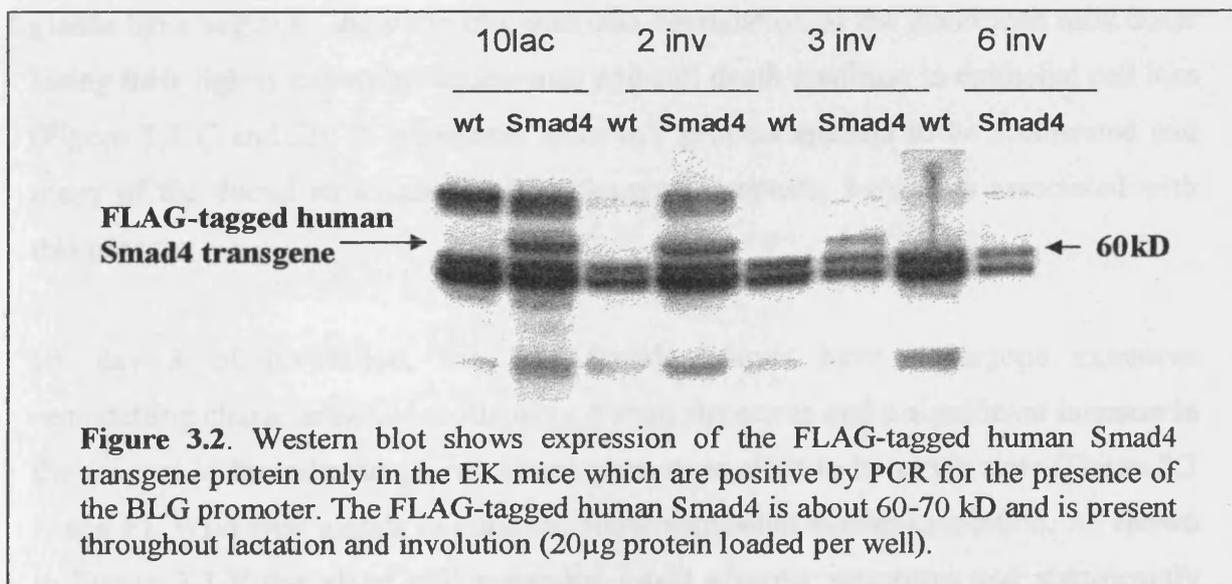
## 3.2 Results

### 3.2.1 Time points and requirements.

All samples used in the following experiments were generated from the same cohort. Mice were genotyped by PCR (chapter 2) for the presence of a part of the BLG-Smad4 construct and named BLG-Smad4 (experimentals) while wild type littermates were used as controls. In order to minimize any alterations in mammary gland stimuli, all mothers were allowed to suckle 6-8 pups for 10 days (defined as day 10 of lactation). Involution was then forced by the removal of pups from the mother. Time points included day 10 lactation and day 2, 3 and 6 of involution.

### 3.2.2 BLG promoter expresses human Smad4 in the murine mammary gland throughout involution.

Mammary tissue from our breeder mice was subjected to anti-FLAG western blot analysis, in order to determine expression of the transgene. The FLAG-tagged human Smad4 transgene has a predicted size of 60-70 kD and it was shown to be constitutively expressed in the mammary gland throughout involution exclusively in mice carrying the BLG promoter (Figure 3.2).



**Figure 3.2.** Western blot shows expression of the FLAG-tagged human Smad4 transgene protein only in the EK mice which are positive by PCR for the presence of the BLG promoter. The FLAG-tagged human Smad4 is about 60-70 kD and is present throughout lactation and involution (20 $\mu$ g protein loaded per well).

### **3.2.3 Over expression of Smad4 accelerates involution of mammary epithelial cells.**

To investigate any possible alterations in the contribution of epithelial cells in the mammary gland during lactation and involution between BLG-Smad4 and wild type animals, a histological phenotypic analysis was undertaken by sectioning and staining with hematoxylin- and -eosin (H&E). The number of each genotype for each timepoint analysed was at least 3. Following our observation regarding FLAG-tagged transgene expression from day 10 of lactation up to day 6 Involution, we generated sections from day 10 Lactation and day 2, 3, 6 Involution. Figure 3.3 shows representative glands from BLG-Smad4 mice (A, C, E, G) and wild type (B, D, F, H) mice.

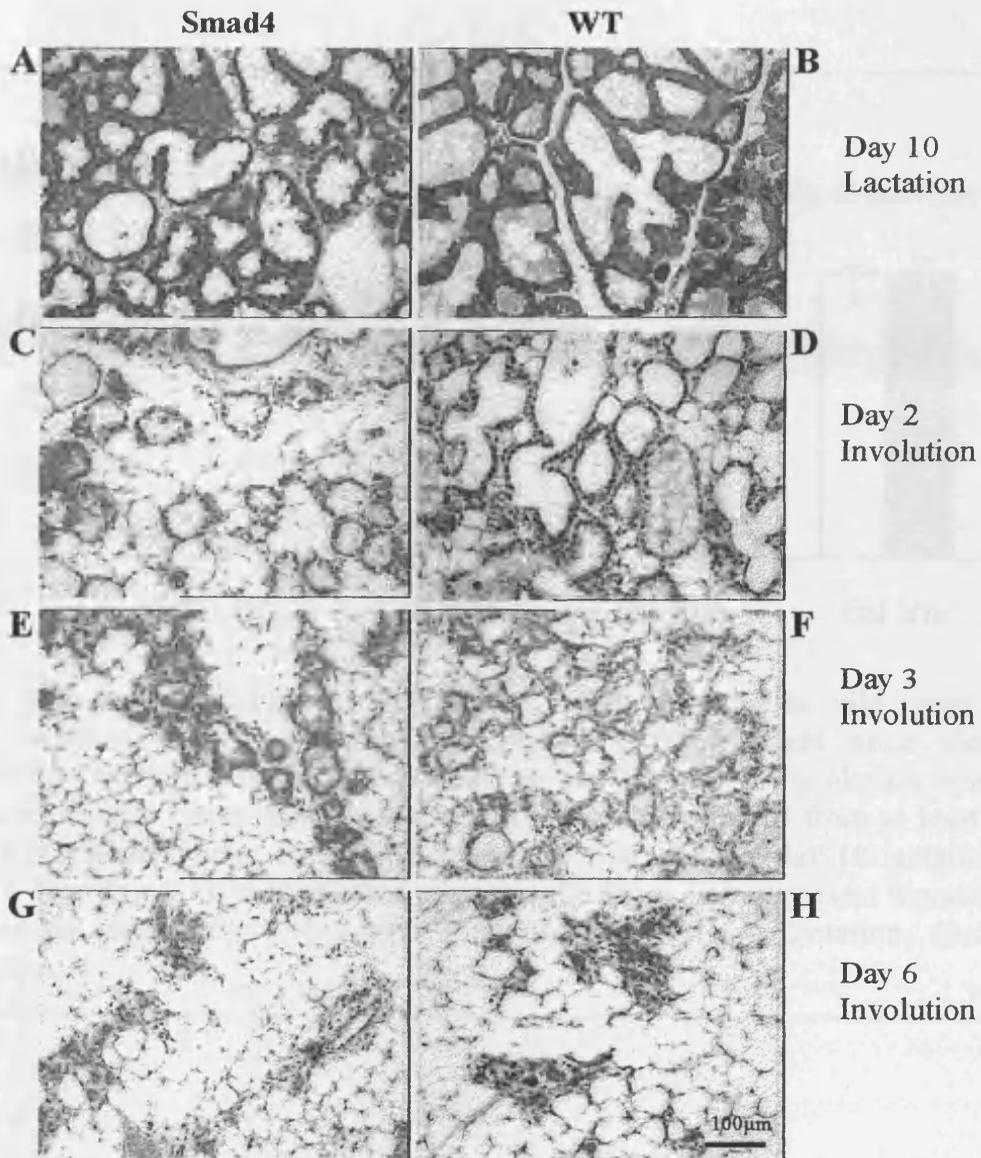
At day 10 of lactation, the majority of the gland was composed of alveoli lined by epithelial cells that secrete milk components into the alveolar lumina. No phenotypic differences were detected between BLG-Smad4 (Figure 3.3A) and wild type (Figure 3.3B) mice at this stage. Furthermore, no differences were observed in the ability of BLG-Smad4 and wild type mice to feed and maintain their litters.

Forty-eight hours following removal of the pups (day 2 of involution), wild type glands have begun to show the characteristic degradation of the gland with milk ducts losing their tightly expanded appearance and cell death resulting in epithelial cell loss (Figure 3.3 C and D). In transgenic mice this process appears to be accelerated and many of the ductal structures contain sloughed apoptotic fragments associated with this process.

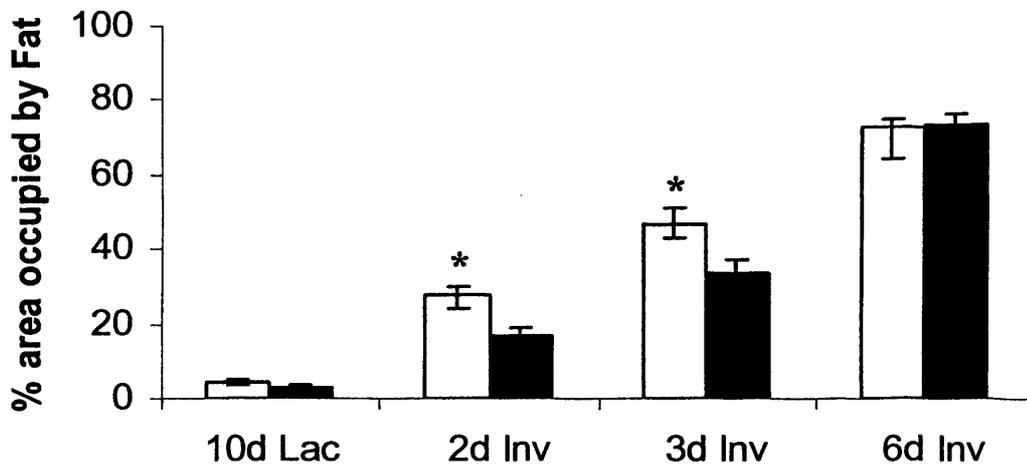
By day 3 of involution, the BLG-Smad4 glands have undergone extensive remodelling characterised by collapsing alveoli structures and a significant increase in the amount of fat indicating a mammary phenotype close to its virgin state (Figure 3.3 E and F). Wild type glands in contrast, show somewhat slower involution. As shown in Figure 3.3 F the gland still possesses intact alveolar structures and significantly fewer fat cells compared to BLG-Smad4 mice. At this point wild type glands are similar to BLG-Smad4 glands at their second day of involution.

Finally, by day 6 of involution in both wild type and transgenic mice glands have nearly all collapsed. Glands have been remodelled and are mainly occupied by fat cells and only occasional epithelial cords and ducts can be seen (Figure 3.3 G, H).

To quantify the degree of involution occurring in the mammary over expressing Smad4 mice compared to wild types, the area of the gland occupied by fat was measured as previously described (Chapman 1999) and is shown in Figure 3.4. At the tenth day of lactation less than 5% of the area of the gland was occupied by fat cells in both Smad4 and wild type animals. By day 2 and 3 of involution, Smad4 animals have a significantly higher fat cell contribution when compared to wild types. Specifically, at day 2, Smad4 glands showed 28% fat cell composition while wild types showed 17%. At day 3 of involution, Smad4 glands had a fat cell contribution of 47% compared to 33% in wild types. Taken together, these results show an accelerated involution in mice over expressing human Smad4 protein. Finally at day 6, both glands have remodelled to resemble to a virgin like state characterised by a high fat cell contribution. At this point the programme of involution in wild type glands has apparently “caught up” with the Smad4 transgenics probably during day 4 and 5 of involution.



**Figure 3.3.** Accelerated involution in BLG-Smad4 mammary glands (A, C, E, G) compared to wild types (B, D, F, G). Mammary glands at day 10 of lactation (A,B); Day 2 Involution (C,D); Day 3 Involution (E,F); Day 6 Involution (G,H). Microscope scale: X40. Scale bar 100 $\mu$ m.



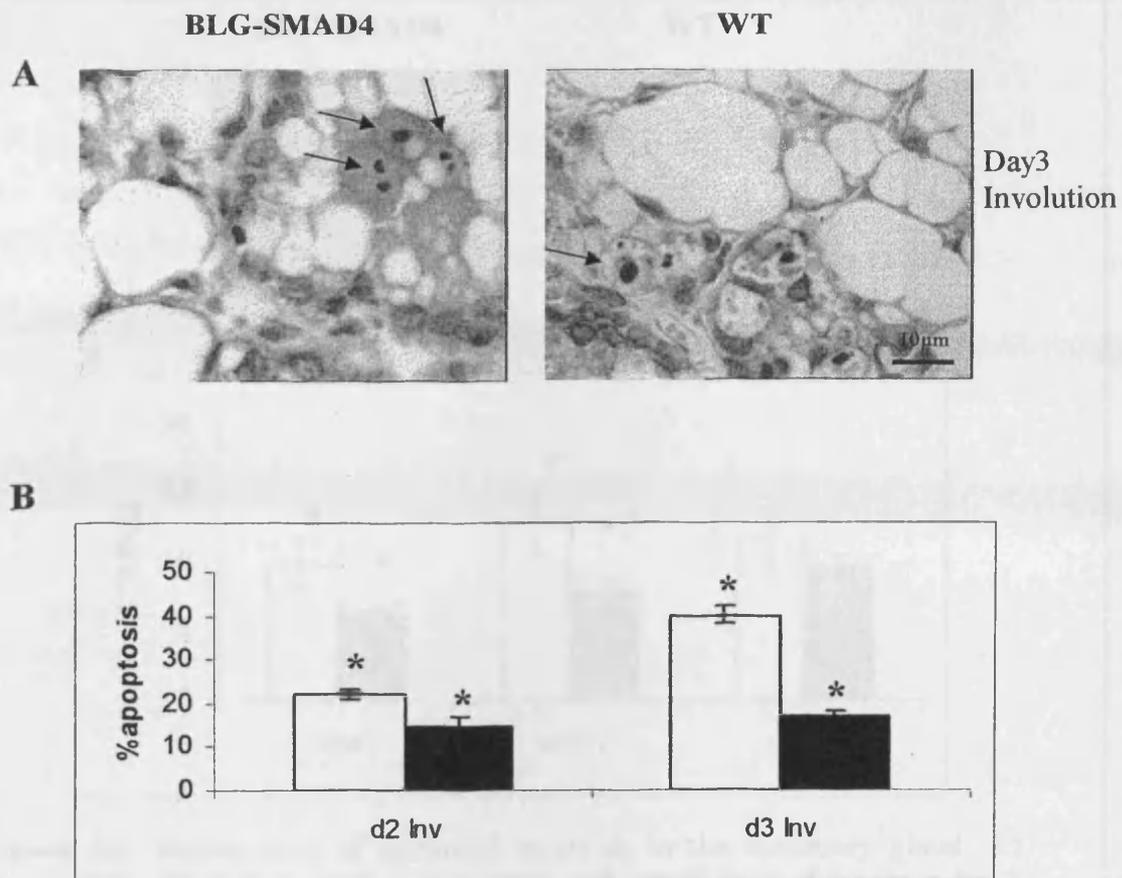
**Figure 3.4.** Fat contribution in BLG-Smad4 mice compared to wild types at day10 lactation and 2, 3 and 6 of involution. BLG-Smad4 mice show significant accelerated reappearance of fat at day 2 and 3 of involution when compared to wild types. Each bar represents the data collected from at least 3 mice. BLG-Smad4=open bars, Wild types=solid bars.d10lac=day 10 lactation, d2, 3, 6 Inv=day 2, 3, 6 Involution respectively. Error bars represent standard error of the mean. \* $p < 0.05$  Mann-Whitney U test. (Lac)- lactation, (inv)-Involution

### **3.2.4 Over expression of Smad4 increases apoptosis in mammary epithelial cells.**

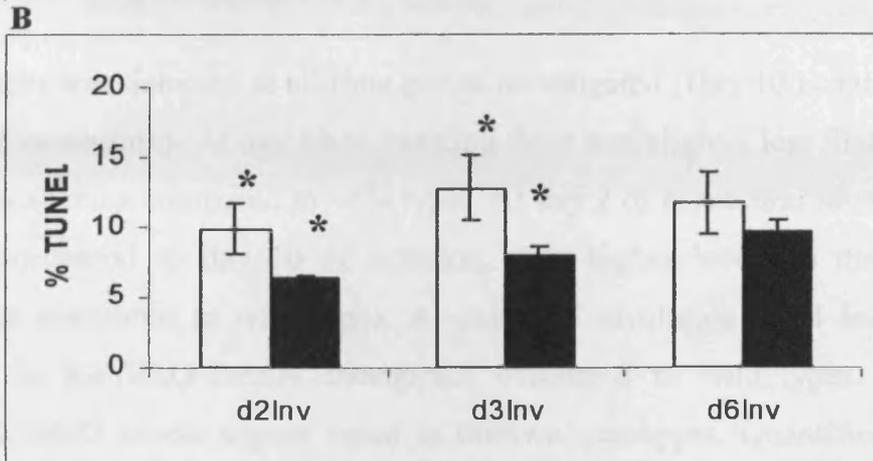
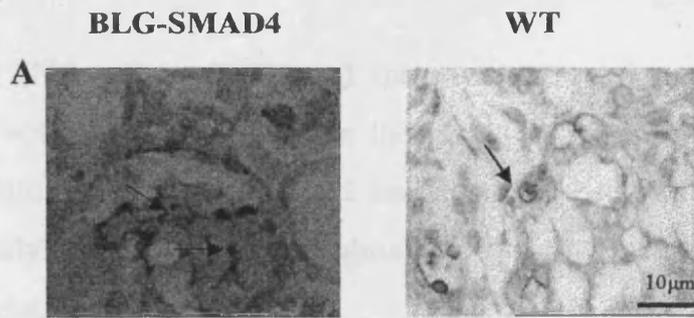
During mammary gland involution, epithelial cells undergo apoptosis and are shed from the epithelial wall (Strange 1995). In this way the gland resembles a virgin like stage composed of adipocytes. Following my observation regarding accelerated involution at day 2 and 3 in the Smad4 transgenics, I decided to investigate whether these changes were caused by increased apoptosis. Apoptosis was measured by morphological criteria, such as condensed chromatin (Figure 3.5), and also by TUNEL analysis of DNA strand breaks (Figure 3.6).

Morphological assessment of mammary epithelial cell apoptosis at day 2 and 3 of involution showed a significant increase of apoptosis in the BLG-Smad4 mice compared to wild types ( $p < 0.05$ ) (Figure 3.5).

This increase in apoptosis in the BLG-Smad4 mice was confirmed by TUNEL analysis (Figure 3.6). A significant increase in epithelial apoptosis was seen at day 2 and day 3 of involution in Smad4 transgenic mice when compared to wild types. Specifically, at day 2 Smad4 mice showed a 3% increase in apoptosis compared to wild types while at day 3 this increase is 4.8%. At day 6 of involution, both transgenic and wild type glands showed similar levels of apoptosis.



**Figure 3.5.** Acceleration of epithelial apoptotic cells in BLG-Smad4 mice mammary glands. (A) Morphological assessment of apoptotic epithelial cells at Day 3 of Involution. Arrows indicate apoptotic cells. (B) Graph represents apoptotic data from day 2 and 3 of involution. BLG-Smad4 mice shows significant increased levels of apoptotic cells when compared to wild types at both day 2 and 3 of involution. BLG-Smad4=open bars, Wild types=solid bars. Each bar represents the mean of data collected from at least three mice. N=number of mice used.d2 Inv and d3 Inv=day 2, 3 involution respectively. Error bars represent standard error of the mean \* $p < 0.05$ . Scale bar 10µm.



**Figure 3.6.** Acceleration of epithelial apoptosis in the mammary gland. A) Photographs (BLG-Smad4) and (wild type) represent TUNEL slides from day 3 involuting mice. Arrows indicate apoptotic cells. B) Graph illustrates apoptosis as counted by TUNEL. Each bar represents the mean of data collected from at least 3 mice. Significant increased levels of apoptosis can be seen at day 2 and 3 of involution in the BLG-Smad4 mice compared to wild types. BLG-Smad4=open bars, Wild types=solid bars. d2, 3, 6 Inv=Day 2, 3, 6 involution respectively. Error bars represent standard error of the mean. \* $p < 0.05$  Mann-Whitney U test. Scale bar  $10\mu\text{m}$ .

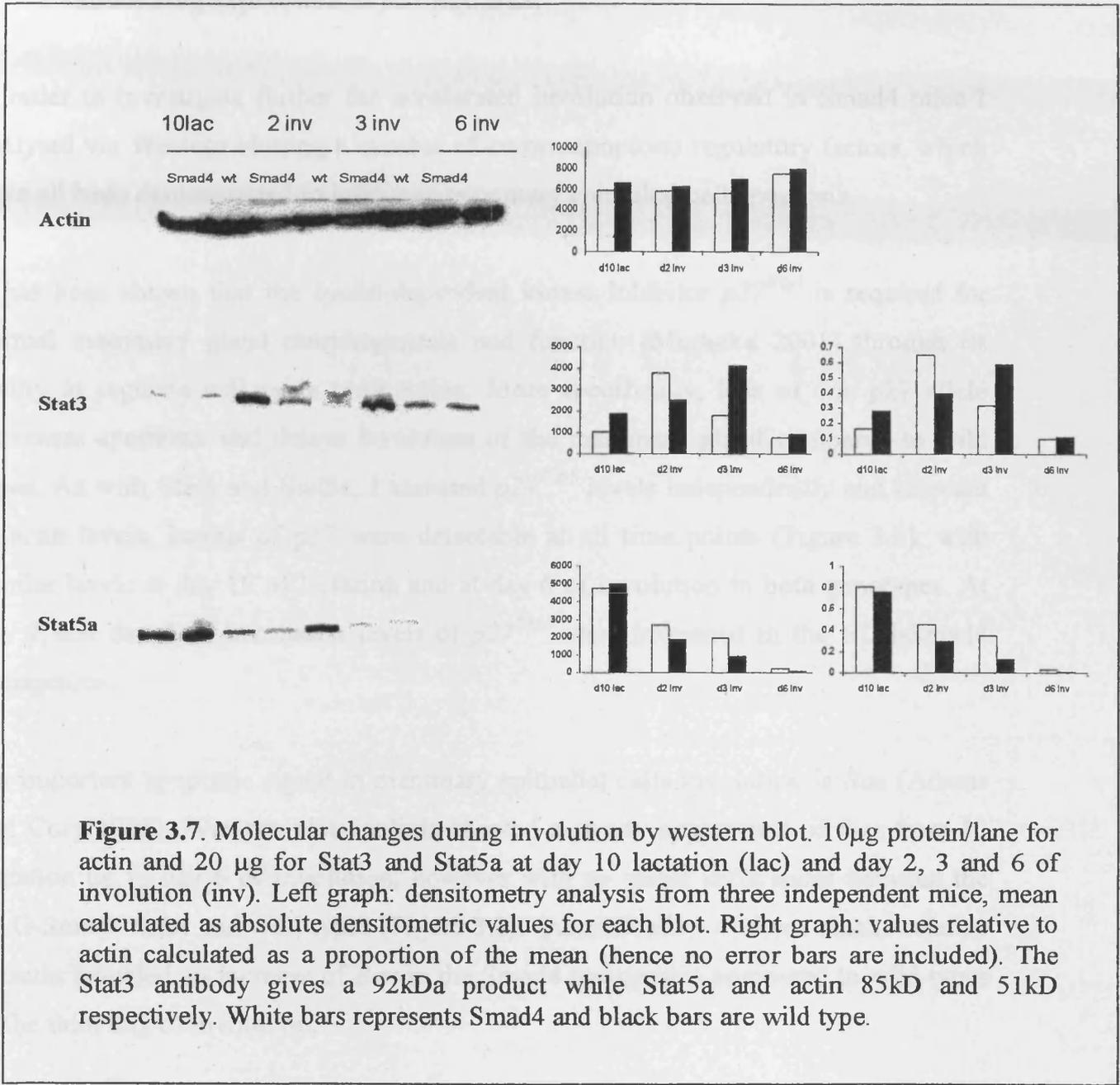
### 3.2.5 Molecular analysis of Stat3 and Stat5a in an over expressing Smad4 mammary gland environment.

It is well known that Stat3 and Stat5a are important mediators of mammary gland development. Stat5a was originally described as a regulator of milk protein gene expression and was subsequently shown to be essential for mammary development and lactogenesis (Liu 1996). In contrast, Stat3 is an essential mediator of apoptosis and post-lactational regression (Philip 1996). Furthermore, it has been shown that Smads and Stats can act in synergy by binding to the transcriptional co-activator p300

(see also 4.1.3) (Nakashima 1999) and that over expression of TGF- $\alpha$  results in deregulation of both Stat3 and Stat5a in the involuting mammary gland (Schroeder 2001). Taking into account the above, I investigated the levels of both proteins by Western blot analysis (Figure 3.7). Absolute levels of Stat3 were assessed as well as levels of Stat3 relative to actin.

Stat3 protein was detected at all time points investigated (Day 10 lactation and day 2, 3 and 6 of involution). At day 10 of lactation there was slightly less Stat3 in the BLG-Smad4 transgenics compared to wild types. At day 2 of involution levels of Stat3 are elevated compared to day 10 of lactation, with higher levels in the BLG-Smad4 transgenics compared to wild types. At day 3 of involution Stat3 levels are down regulated in the BLG-Smad4 transgenics compared to wild types. At day 6 of involution, Stat3 levels appear equal in the two genotypes. Quantification of these results relative to actin levels indicate that the above results are not confounded by changes in cellular composition, which was a possibility given that BLG-Smad4 mice have less epithelial cells compared to wild types.

Stat5a levels were also measured and again quantified against actin levels (Figure 3.7). Stat5a was again detected throughout involution. High and equal levels of Stat5a were present at day 10 of lactation. During involution Stat5a levels dropped significantly, however, Stat5a remained relatively higher in the BLG-Smad4 transgenics compared to controls of both day 2 and 3 of involution. Quantification of these results relative to actin indicates that Stat5a is induced in BLG-Smad4 mice on a "per cell" basis compared to wild types at both day 2 and 3 of involution.



**Figure 3.7.** Molecular changes during involution by western blot. 10µg protein/lane for actin and 20 µg for Stat3 and Stat5a at day 10 lactation (lac) and day 2, 3 and 6 of involution (inv). Left graph: densitometry analysis from three independent mice, mean calculated as absolute densitometric values for each blot. Right graph: values relative to actin calculated as a proportion of the mean (hence no error bars are included). The Stat3 antibody gives a 92kDa product while Stat5a and actin 85kD and 51kD respectively. White bars represents Smad4 and black bars are wild type.

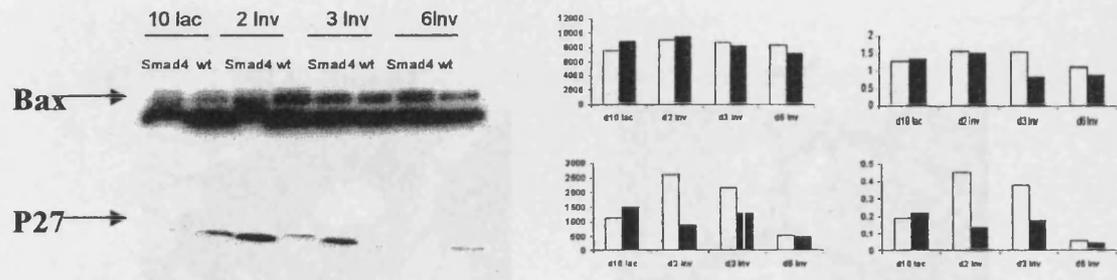
### 3.2.6 Over expression of Smad4 increases expression of $p27^{Kip1}$ and is not influencing expression of $p21$ and *Bax*.

In order to investigate further the accelerated involution observed in Smad4 mice I analysed via Western blotting a number of known apoptosis regulatory factors, which have all been demonstrated to influence mammary epithelial cell apoptosis.

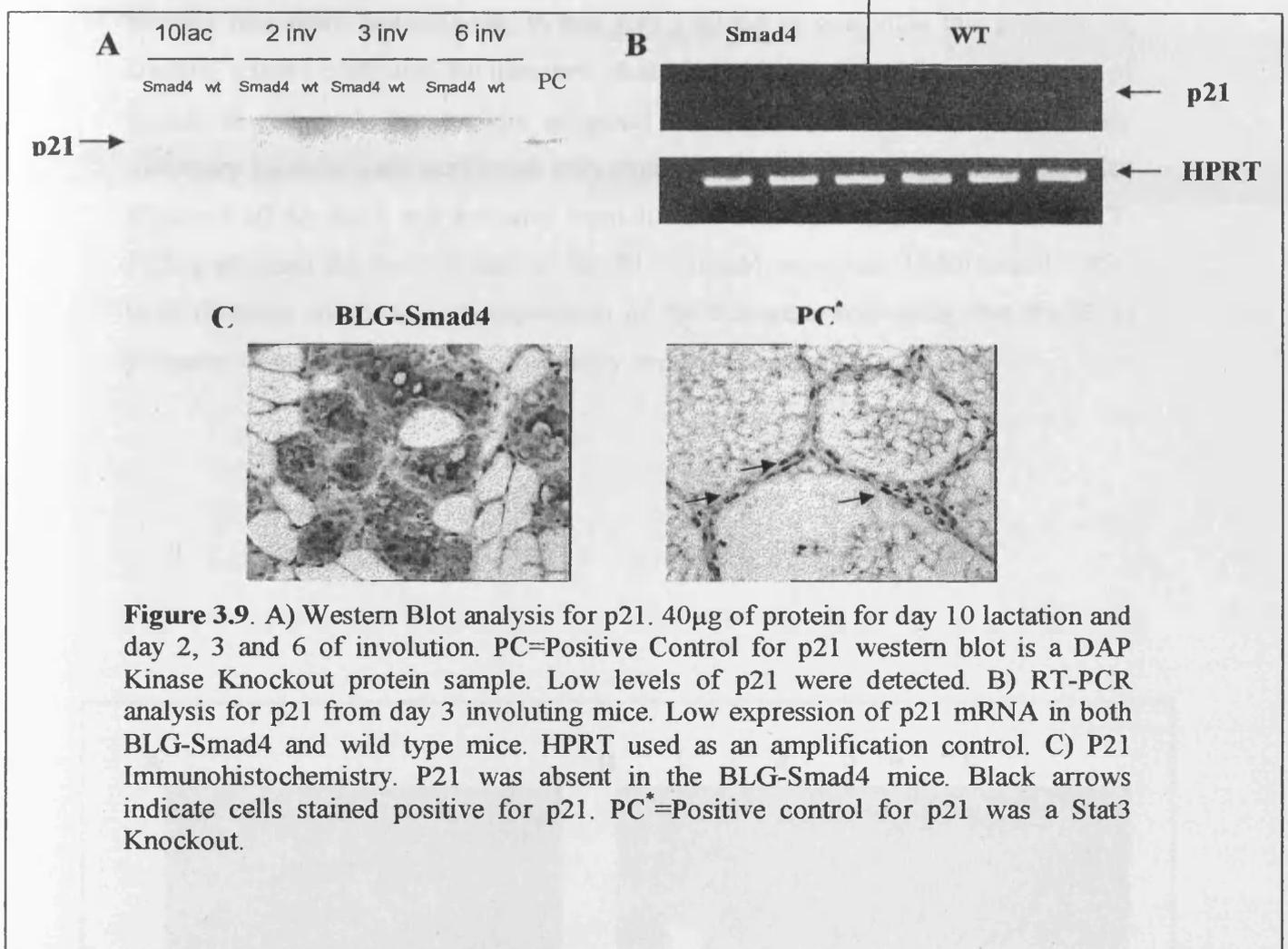
It has been shown that the cyclin-dependent kinase inhibitor  $p27^{Kip1}$  is required for normal mammary gland morphogenesis and function (Muraoka 2001) through its ability to regulate cell cycle progression. More specifically, loss of one p27 allele decreases apoptosis and delays involution of the mammary gland compared to wild types. As with Stat3 and Stat5a, I assessed  $p27^{Kip1}$  levels independently and relevant to actin levels. Levels of p27 were detectable at all time points (Figure 3.8), with similar levels at day 10 of lactation and at day 6 of involution in both genotypes. At day 2 and day 3 of involution levels of  $p27^{Kip1}$  were increased in the BLG-Smad4 transgenics.

An important apoptotic signal in mammary epithelial cells involution is *Bax* (Adams and Cory 1998). Western blot analysis showed a steady appearance of *Bax* from 10 lactation up to day 6 of involution, however with no major differences between the BLG-Smad4 mice and wild types (Figure 3.8). Quantification of these results relative to actin revealed an increase of *Bax* in the Smad4 transgenics compared to wild types at the third day of involution.

Another important cyclin-dependent kinase inhibitor whose mRNA levels have been demonstrated to increase during mammary gland involution is  $p21^{Waf1}$  (Jerry 1998). Smad3 and Smad4 have been shown to up regulate transcription of the  $p21^{Waf1}$  gene by cooperating with Sp1 (Pardali 2000). However, low levels of  $p21^{Waf1}$  were detected by Western blot analysis and no increase was observed in the Smad4 transgenics (Figure 3.9A).  $P21^{Waf1}$  levels were also investigated via semi-quantitative RT-PCR from day 3 involuting mice (Figure 3.9B). Very low and equal levels of  $p21^{Waf1}$  were observed in the BLG-Smad4 transgenics as well as in the wild type animals.  $P21^{Waf1}$

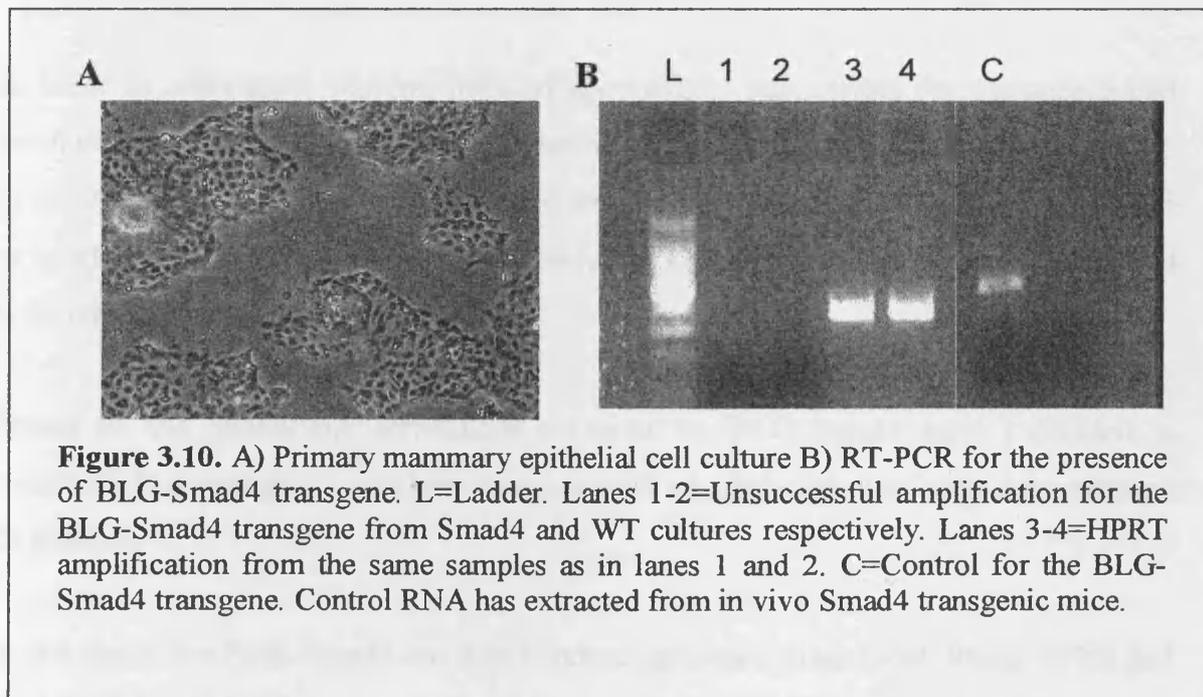


**Figure 3.8.** Molecular changes during involution by Western blot. 10 $\mu$ g protein/lane for *Bax*, 20 $\mu$ g protein/lane for P27 at day 10 lactation (lac) and day 2, 3 and 6 of involution (Inv). Left graph: densitometry analysis from 3 independent mice. Mean calculated as absolute densitometric values for each blot. Right graph values relative to actin calculated as a proportion of the mean. *Bax* antibody produces a 23 kD product, p27 shows as a 27kD product.



### 3.2.7 BLG promoter is not expressed in mammary epithelial cell cultures.

In order to overcome problems arising from the examination of a mixed population of cells within a dynamic tissue such as the mammary gland I next decided to perform primary mammary cell cultures. In this way I aimed to overcome this problem by creating a more controlled environment where I could investigate the contribution of Smad4 to cell cycle progression, apoptosis and possibly cell attachment. Primary mammary cultures were performed with glands taken from 15-17 day pregnant mice (Figure 3.10 A). RNA was extracted from 50-60% confluent epithelial cells and RT-PCR performed for the presence of the BLG-Smad4 construct. Unfortunately, RT-PCR revealed an absence of expression of the transgene indicating that the BLG promoter was not functioning in mammary explant cultures.



**Figure 3.10.** A) Primary mammary epithelial cell culture B) RT-PCR for the presence of BLG-Smad4 transgene. L=Ladder. Lanes 1-2=Unsuccessful amplification for the BLG-Smad4 transgene from Smad4 and WT cultures respectively. Lanes 3-4=HPRT amplification from the same samples as in lanes 1 and 2. C=Control for the BLG-Smad4 transgene. Control RNA has extracted from in vivo Smad4 transgenic mice.

### **3.3 Discussion.**

Here I report the analysis of a transgenic strain over expressing Smad4 within the mammary gland. Tissue specific Smad4 expression was driven by the BLG promoter, to deliver expression during mid-pregnancy, lactation and involution. I have confirmed the expression of human flag tagged Smad4 transgene by western blot analysis.

Smad4 has been characterised as a tumour suppressor gene due to its ability to arrest cell division and promote apoptosis (Dai 1999). In agreement with this observation BLG-Smad4 mice exhibited accelerated epithelial cell death at day 2 and 3 of involution. The level of difference in epithelial cell contribution in the gland was measured indirectly by measuring fat contribution (Chapman 1999). At day 2 involution BLG-Smad4 mice showed an 1.6 fold increase in the level of cell death compared to wild types whilst at day 3 this increase was 1.4 fold. At day 6 of involution the majority of the gland has been remodelled to resemble a virgin like state.

In order to investigate whether induced apoptosis is responsible for the accelerated involution I scored apoptosis by the criterion of chromatin condensation and TUNEL (Li 1997). These two approaches yielded similar data indicating a significant increase in apoptosis at both day2 and day3 of involution in the BLG-Smad4 mice compared to wild types.

Based on the phenotypic differences exhibited by BLG-Smad4 mice I decided to molecularly investigate a number of targets with which Smad4 has been demonstrated to interact.

It is known that both Smad4 and Stat3 induce apoptosis (Dai 1999, Philip 1996) and that are bridged by the transcriptional co-activator p300 (Nakashima 1999). Previous studies have shown that Stat3 deficiency in the mammary gland resulted in delayed involution with decreasing levels of apoptosis (Chapman 1999). Here, western blot analysis revealed an up-regulation of Stat3 in the BLG-Smad4 transgenic mice at

day2 of involution compared to wild types. This induction of Stat3 becomes even greater after quantification relevant to actin. Contrary, at day3 of involution the opposite scenario takes place. Stat3 shows a down regulation in the BLG-Smad4 transgenic animals when compared to controls. It has been demonstrated that Stat3 becomes activated at the start of involution (Li 1997). This activation has been enhanced at day 2 involution in the over expressing Smad4 mice indicating a mechanism where a co-operation between Smad4-Stat3 may promote apoptosis in mammary epithelial cells. However, down regulation of Stat3 in the over expressing Smad4 mice may indicate a possible compensatory mechanism where Stat3 tries to balance this accelerated cell death observed during day 3 involution. Down regulation of Stat3 at day 3 of involution in BLG-Smad4 mice indicates a reciprocal relationship between Stat3 and Smad4 and may go some way to explain the observed phenotype at day 6 of involution.

Stat5a is a tyrosine phosphorylated signaling protein that lies downstream of the prolactin-signaling pathway in normal mammary epithelial cells (Hennighausen 1997a). It has been demonstrated that Stat5a's physiological role is to promote cell survival and its loss results in increased apoptosis during mammary gland involution (Humphreys and Hennighausen 1999). Here I report a steady induction of Stat5a in the BLG-Smad4 mice compared to controls throughout involution. Induction of Stat5a in the BLG-Smad4 mice possibly reflects the induction of a cell defense mechanism towards survival or may reflect a dynamic balance operating in the mammary which can restore a normal phenotype at day 6 of involution in the BLG-Smad4 glands.

The *Bcl-2* gene family regulates tissue development and homeostasis through the interplay of survival and death factors (Schorr 1999). *Bax* is a member of the *Bcl-2* family, levels of which had been shown to increase during lactation and remain high during involution (Merto 1997). *Bax* levels were investigated by Western blotting and in agreement with Merto's (1997) paper, I found high and equal levels throughout involution in wild type mice and also in the BLG-Smad4. Quantification relevant to actin however, revealed an increase of *Bax* levels in the BLG-Smad4 transgenics

compared to controls at day 3 of involution. These data suggest once more the involvement of *Bax* in apoptosis during involution.

To investigate further the molecular events underlying the accelerated involution in the BLG-Smad4 transgenic mice, I analysed the expression levels of the cell cycle regulators *p21<sup>Waf1</sup>* and *p27<sup>Kip1</sup>* which are known to play important roles in cell cycle progression (Sherr and Roberts 1999). The *p21<sup>Waf1</sup>* protein has been shown to interact with Smad3 and Smad4 through the cooperation with Sp1 (Pardali 2000). Low levels of *p21<sup>Waf1</sup>* were detected by Western blot analysis and no increase was observed in the Smad4 transgenics. Similarly, low levels of *P21<sup>Waf1</sup>* were detected by RT-PCR analysis at day 3 involution, indicating an independency of mammary epithelial cell apoptosis from this molecule. *P21<sup>Waf1</sup>* was not detected by immunohistochemical analysis in either genotype. Based on the above results I therefore conclude that in my model of study Smad-dependent expression of the *P21<sup>Waf1</sup>* is absent.

Another important cyclin-dependent kinase inhibitor studied was *p27<sup>Kip1</sup>*. *P27<sup>Kip1</sup>* has been shown to be important in mouse mammary gland morphogenesis and function (Muraoka 2001). In the same paper it has been demonstrated that haploid *p27<sup>Kip1</sup>* gene inactivation causes amongst others increased proliferation and delayed involution. Cooperation between TGF- $\beta$ /Smads and *p27<sup>Kip1</sup>* has been shown to exist through *Jab1*, a co-activator of *c-Jun*. *Jab1* induces degradation of both *p27<sup>Kip1</sup>* and Smad4 and antagonises TGF-beta function by inducing degradation of Smad4 through a distinct degradation pathway (Wan 2002). *P27<sup>Kip1</sup>* levels were analysed through western blotting and found to be significantly increased at day 2 and 3 of involution. The data presented herein support the notion that *p27<sup>Kip1</sup>* levels regulate proliferation and comes in agreement with the accelerated involution as well as increased apoptosis observed at the same time points.

In an attempt to simplify the analysis I decided to also investigate Smad4's over expression by using primary mammary epithelial cell cultures. However, RT-PCR analysis indicated that expression of the BLG-Smad4 construct failed *in vitro*.

In conclusion, over expression of Smad4 driven by the BLG promoter in the mammary gland accelerates involution of epithelial cells by inducing apoptosis at day 2 and 3 of involution through a  $p27^{Kip1}$  mechanism which is independent of  $p21^{Waf1}$ . Targets such as Stat3 and Stat5a showed a differential expression at day3 of involution in our transgenic mice compared to wild types, indicating once more the interplay between the Smad and Stat pathway. Considering the numerous potential targets of Smad signaling further studies are clearly necessary both *in vivo* and *in vitro* in order to address whether any other mechanisms are responsible for this acceleration of involution. To directly approach this question, I adopted a microarray approach described in the following chapter.

## **Chapter 4- Microarray Analysis of Day 3 involuting mammary glands in a Smad4 over expressing environment.**

### **4.1 Introduction**

#### **4.1.1 Microarray Analysis of Smad4.**

The use of microarray technologies to monitor gene expression in model organisms, cell lines, and human tissues has become an important part of biological research over the last several years. Teasing apart biochemical pathways, identifying genes responsible for a particular phenotype, and assessing the effect of a drug compound on the expression levels of any number of genes have all benefited from expression array technology. Taking into account the usefulness of this technology, I decided to undertake a mouse DNA microarray approach.

It has been demonstrated that TGF- $\beta$  and Smads regulate, and are regulated by, a vast number of biological pathways. My goal in this part of the project was to identify, categorize and link to known pathways, a number of genes whose expression is altered by Smad4. In this way I am aiming to gain a broader view of any possible alterations in gene expression caused by over expression of Smad4 in the mammary gland during involution.

Based upon the phenotypic and molecular differences presented in Chapter 3, I decided to compare and analyze gene expression profiles of BLG-Smad4 transgenic mice against wild types from day 3 involution. Furthermore, in this chapter I will discuss briefly Affymetrix technology and describe key issues that relate to the preprocessing and analysis of data.

#### **4.1.2 Affymetrix Technology.**

There are several publications discussing the fundamentals of the oligonucleotide expression microarray technology (Lockhart 1996). However, for the purpose of this chapter it will be useful to review some of the basics of this technology. Affymetrix chips are microscope slides that contain a series of samples (DNA, RNA, protein tissue). As mentioned above, we have utilised cDNA printed slides in order to check expression of genes in this level. Each gene is represented on the chip as multiple probe sets each consisting of 10-25 oligonucleotide pairs. The oligonucleotide pair (probe pair) is composed by a series of oligonucleotides which are unique to a gene or show less similarity to other genes. These oligonucleotides are chosen as perfect matches (PM) (i.e. perfectly complementary to the mRNA of that gene). In addition, Affymetrix has generated the same oligonucleotide probes, which are identical to the PM oligonucleotides except for the central position 13, where one nucleotide has been changed to its complementary nucleotide. This probe set is termed Mismatch oligos (MM) and their aim is to detect non-specific and background hybridizations. Probes are designed within 500 base pairs of the 3' end of each gene in order to hybridize uniquely in the same, predetermined hybridization conditions.

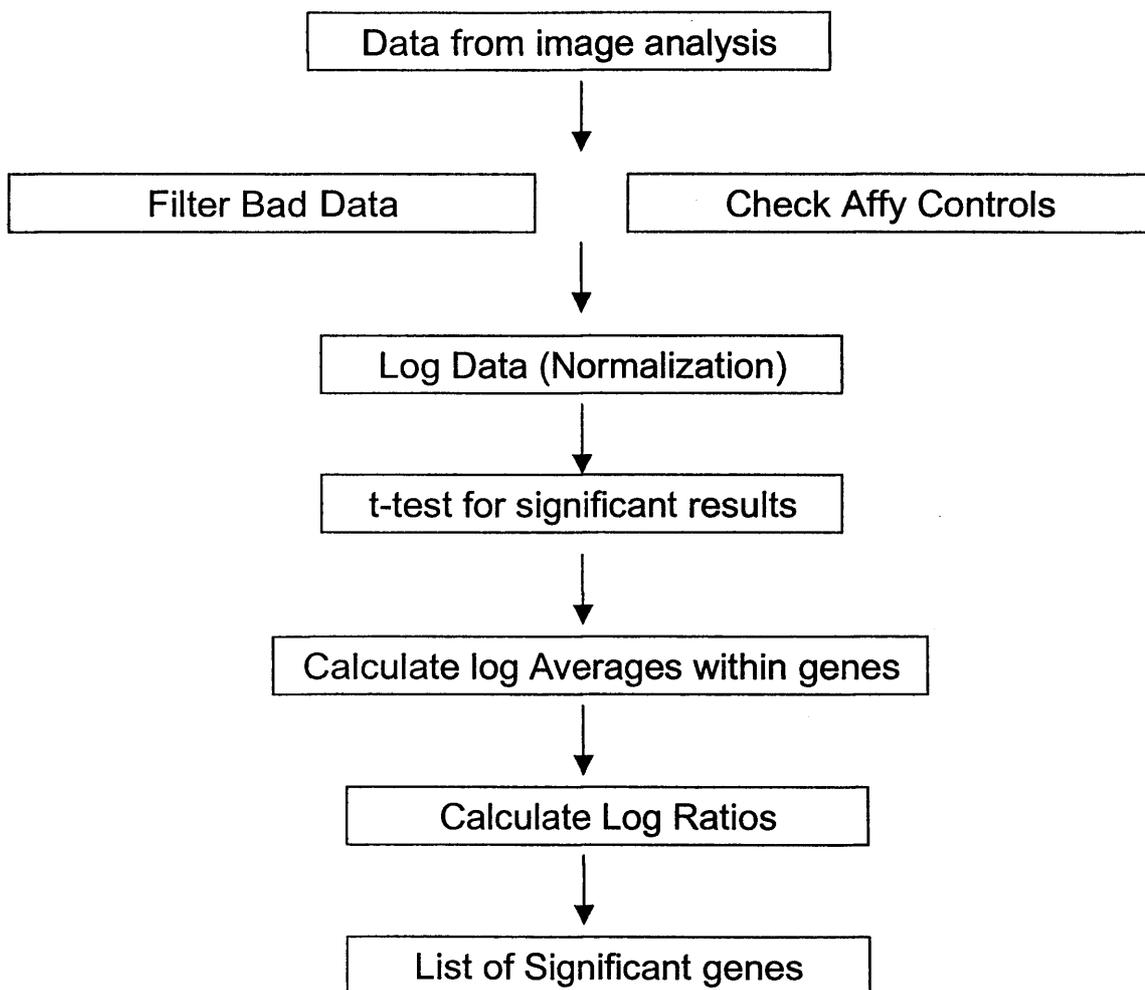
RNA is extracted, converted to cDNA, labelled with biotin and fragmented (see Materials and Methods), before the actual hybridization takes place. Biotin labelled mRNA will hybridise to its corresponding perfect match. The technology is based on the simple hypothesis that the greater the abundance of the gene in the sample mRNA the more expression signal and the other way around. The biotin labelled RNA is stained with phycoerythrin conjugated streptavidin after washing and scanned with a Gene array scanner. A grid is automatically laid over the array image and the intensities of each probe pair are used to calculate expression measurements with the Affymetrix Microarray suite.

Analysis generates qualitative and quantitative values from one gene expression experiment and provides initial data required to perform comparisons between

experiments. A quantitative value, a Detection call, indicates whether a transcript is reliably detected (Present) or not detected (Absent) in the array. Based on the Present-Absent calls, preprocessing, normalization and statistical analysis are the three main steps before a reliable indication for altered gene expression. Figure 4.1 illustrates the representative approach which I followed for performing the above steps.

### **4.1.3 Normalization.**

Affymetrix oligonucleotide arrays simultaneously measure the abundances of thousands of mRNAs in biological samples. There are many sources of systematic variation in microarray experiments that affect gene expression levels. To reliably compare data from multiple chips we need to minimize non biological differences that may exist. Typically, the first transformation applied to expression data, referred to as normalization, adjusts the individual hybridization intensities to balance them appropriately so that meaningful biological comparisons can be made (Yang 2001). There are two broad characterizations used for the type of variation that can be seen when comparing arrays: interesting variation (true biological variations in expression levels of genes between control and experimental arrays) and obscuring variations. Examples of obscuring variations arise due to differences in sample preparation (for example labeling differences), production of the arrays and preprocessing of the arrays (for example scanner differences) (Hartemink 2001). The purpose of normalization is to deal with obscuring variations and bring the data from the different experiments onto a level baseline field. Depending on the normalization method of choice, statistical analysis could be directed towards parametric or non-parametric mode. For the purpose of this analysis I chose to normalize by logging the data and thus create a normally distributed (linear) shape data. Linear data can be analyzed by means of parametric tests (in this analysis t-test ) and will be explained in more detail below.



**Figure 4.1.** A representative method for preprocessing (filtering), normalizing and statistically analyzing Affymetrix data.

## **4.2 Results.**

### **4.2.1 RNA preparation and Affymetrix chip.**

Samples collected from day 3 involuting mice. Samples were subjected to RNA extraction and cleaning, double stranded RNA production, nucleotide labeling (one color) and fragmentation following all procedures and materials as written in Chapter 2 (2.7.3).

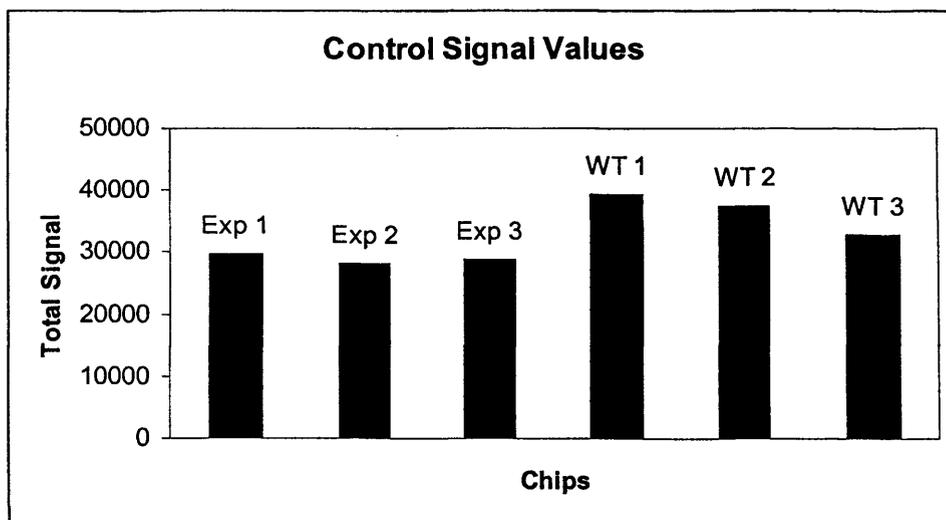
Labeled RNA samples were supplied to University of Wales College of Medicine Affymetrix facility where 3 Affymetrix U74v2 chips used for each genotype (6 chips in total). Chips were hybridized and analyzed with Affymetrix Gene Chip Hybridization and Analysis system.

### **4.2.2 Affymetrix gene controls and run consistency between chips.**

Affymetrix chips also contain various mouse housekeeping genes as three probe sets. Each set is designed to the 5' end, to the middle of the gene and the 3' end respectively. In addition to mouse specific genes, spiked-in control probe sets exist originating from organisms such as *E.Coli*, Bacteriophage and *B.Subtilis*. Control gene sequences do possess signal intensity values despite their inability to show differential expression to mouse RNA. Reasons for this signal values are due to a number of systematic variations (labelling, scanner differences etc.) in microarray experiments that can affect measured gene expression levels. Signal values of gene controls provide a good indication for checking consistency of runs between different Affymetrix chips i.e. checking whether different chips possess the same error levels or not.

As stated above I utilized 6 Affymetrix chips in total. I therefore checked chip consistency by summing all signal intensity values from 67 control gene sequences that exist in the Affymetrix data sheet. Figure 4.2 illustrates the comparison of control gene

errors between genes. From this graph I concluded that all experimental and wild type chips show similar levels of error signal intensities and only slightly altered levels of errors between the two groups. Similarity in the control error intensities indicated similar patterns of hybridization procedures for each chip, thereafter indicating that these chips could be used further for gene expression analysis.



**Figure 4.2.** Summation of signal intensities from control genes. Each bar represents a different chip. Experimental (Exp) and wild type (WT) chips have a similar level of error signal intensities between group chips as well as between groups.

### **4.2.3 Data Filtering.**

Affymetrix analysis generates qualitative and quantitative values from one gene expression experiment and provides initial data required to perform comparisons between experiments. A quantitative value, a detection call, indicates whether a transcript is reliably detected (Present) or not detected (Absent) in the array. The detection call is determined by comparing the Detection  $p$ -value generated in the analysis against user-defined cut-offs. This cut-off by default is 0.015.

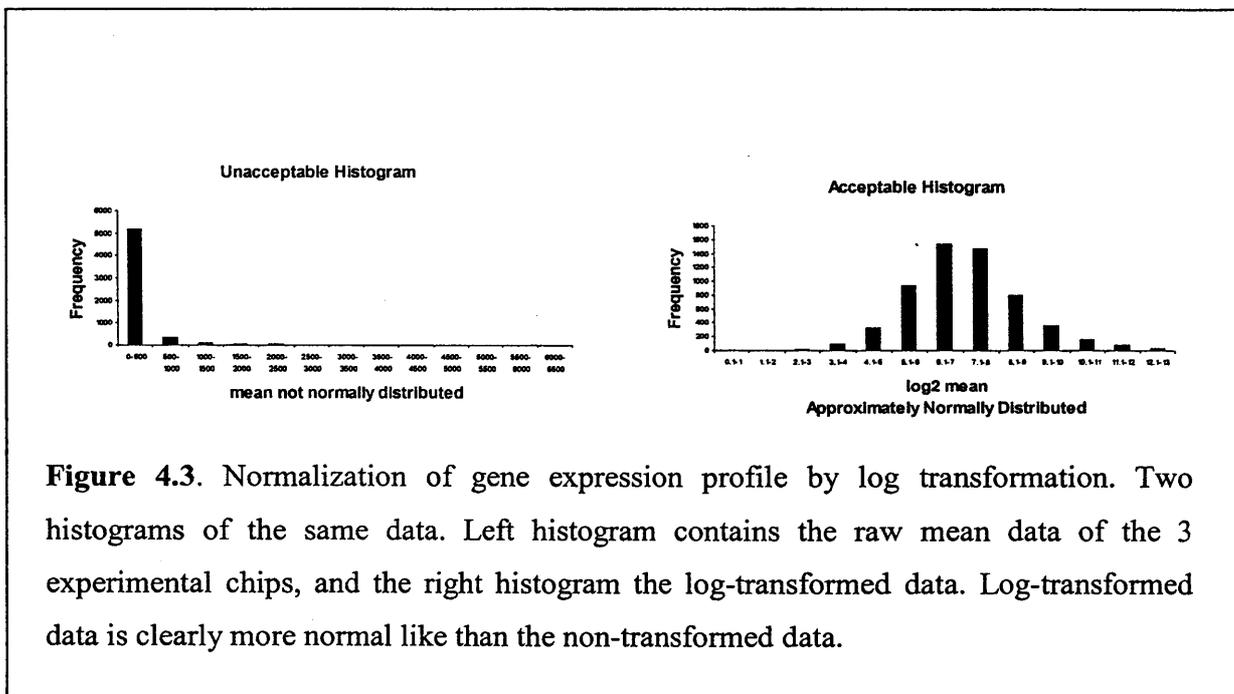
For the purpose of this data analysis I have filtered out 'Absent' calls as such that a transcript can be present at least once in the control and experimental pools respectively. Filtering drastically reduced the available data from 12000 transcripts to 5800. Absent calls reflects sequences which are not expressed either in the control or experimental samples and which are removed by the "Filtering" process. P27, a target which its protein was over expressed in the BLG-Smad4 mice at day 3 involution (Chapter 3) was identified as absent in its mRNA expression at the same time point. A result which indicates that p27 mRNA and protein expression levels have an expression time difference of at least one day.

### **4.2.4 Data Normalization and Linearity.**

As discussed in 4.1.3 many sources of systematic variations exist in microarray experiments affecting each time gene expression levels differently. Normalization is the term used to describe the process of removing such variation (Yang 2001). Normalization can also be thought of as an attempt to remove the non-biological influences on biological data. Removal of variations between microarray chip data brings the data from different experiments to a common baseline from which valuable and statistically correct comparisons can be made.

For the purpose of data pre-processing (normalization) I have utilised a common way of data transformation; log transformation. Log transformation has several important effects on the data (Speed 2000), one of them being to decrease signal errors (in an absolute sense) as the signal expression values decrease by logging.

Additionally, log transformation serves microarray data from another important point of view. Log transformation makes data more symmetrical (linear), one of the important assumptions of normality. Furthermore, log transformation reduces the influence of a single measurement. Figure 4.3 illustrates a representative graph of the mean values of experimental chip data before and after log transformation.



**Figure 4.3.** Normalization of gene expression profile by log transformation. Two histograms of the same data. Left histogram contains the raw mean data of the 3 experimental chips, and the right histogram the log-transformed data. Log-transformed data is clearly more normal like than the non-transformed data.

#### 4.2.5 Parametric Tests; T-Test and Significance Analysis of Microarrays (SAM).

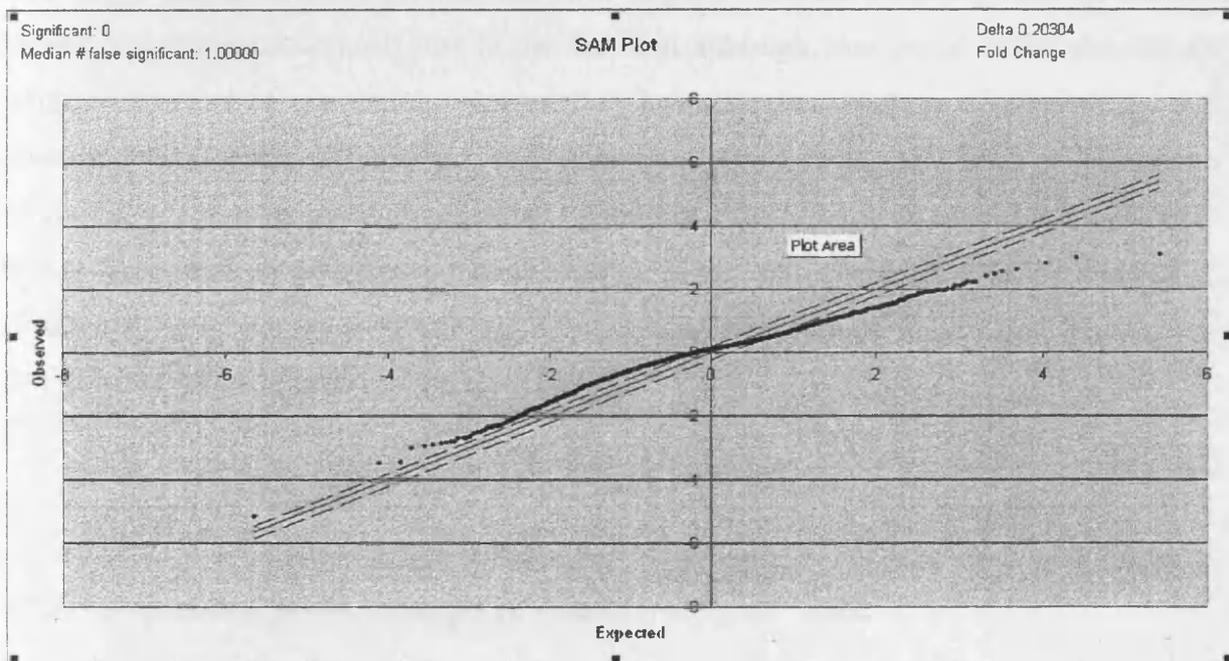
Parametric statistics test the hypothesis that one or more treatments have no effect on the mean of a chosen variable. As mentioned earlier, these tests are based on the assumption that data is taken from a normally distributed population, a distributed population that can be created by logging the data.

For identifying differentially expressed genes I utilised the SAM (Significance Analysis of Microarrays) software (Tusher 2001). SAM is publicly available software and can be located at <http://www-stat-class.stanford.edu/SAM/SAMServlet>. SAM has been developed to identify differentially expressed genes which show significant changes between them by assimilating a set of gene-specific  $t$  tests. For achieving its purpose SAM assigns a score in each gene on the basis of its change in expression relative to the standard deviation of repeated measurements for that gene. Genes with scores above a specified threshold are considered significant. From the group of genes which have been characterised as significant, there is a number of genes which have been identified by chance; this is what termed as the False Discovery Rate (FDR). SAM estimates the FDR by identifying nonsense genes through analysing permutations of the measurements. The cut off threshold can be adjusted to identify smaller or larger sets of genes, and FDRs are calculated for each set.

5800 Present transcripts were applied to the SAM software according to designer's instructions. After the software logged the data in the base of 2, differential analysis created a plot which is presented in Figure 4.4. This graph created a number of questions because it failed to identify any significant differentially expressed genes (should have appeared as coloured dots). Furthermore I observed that the distribution of dots were presented graphically as a mirror image of what a normal plot should be like, showing that the observed significant targets are less than the expected ones (see below).

In order to reduce the stringency of the statistical analysis I decided to use another parametric test, the Student's  $t$ -test. The  $t$ -test assesses whether the means or medians (in

our case the means) of two groups (experimental against control) are statistically different from each other. More specifically after logging the data I have checked for significant results between the two groups. To test for significance I have used a probability level of 0.05 (acceptable false rate) in a two tail paired t-test. A paired t-test is very powerful in that it compares the same gene between two treatments within an experiment, and so, variations (outliers due to noise) in the baseline and experimental values between experiments are mitigated.



**Figure 4.4.** SAM plot analysis. SAM failed to identify any significant differentially expressed genes despite the various adjustments of the delta error area (dotted lines). Furthermore, the graph itself is presented as a mirror image of a normal SAM plot, showing that the observed significant targets are less than the expected ones.

Following log transformation and two tail paired t-test of 5800 transcripts with a probability level of 5% I have identified 114 targets as statistically significant potential targets. Although  $p=0.05$  is significant in the context of experiments designed to evaluate small numbers of genes, a microarray experiment for 5800 genes would identify 290 genes by chance (FDR). This data indicates that the observed significant targets are less than the expected ones and can therefore explain the failure to observe any significant altered gene expressions from the SAM analysis as well as the mirror image which presented in the same graph (Figure 4.4).

In order to find whether these targets show altered signal expression values I calculated their log averages between groups. For each target I calculated the log ratio of the two means (experimental/control) due to the fact that although true signal ratios provide an intuitive measure of expression changes, they have the disadvantage of treating up- and down-regulated genes differently. Genes up regulated by a factor of 2 have an expression ratio of 2, whereas those down regulated by the same factor have an expression ratio of (0.5). I have used an alternative transformation of the ratio by logging in the base of 2, which has an advantage of producing a continuous spectrum of values and treating up- and down-regulated genes in a similar fashion.

#### **4.2.6 Expression profile changes in Smad4 transgenic mice.**

I have utilized DNA microarrays to analyze gene expression profiles of human Smad4 over expressing mouse in mammary epithelial cells. Smad4 over expressing epithelial cells have been compared to wild types. All samples have been harvested from mice at their 3<sup>rd</sup> day of involution. The Affymetrix gene chip used is the U74v2 consisting of 12488 transcripts. Following preprocessing and statistical analysis as described above, 114 clones and ESTs showed statistically significant signal values in both the experimental and control pools. Based on both raw signals and Log transformed values no sample showed a more than 2 fold up- or down-regulation. More specifically 58

clones and ESTs were up-regulated in Smad4 transgenic samples (Figure 4.5) and 56 were down regulated (Figure 4.6).

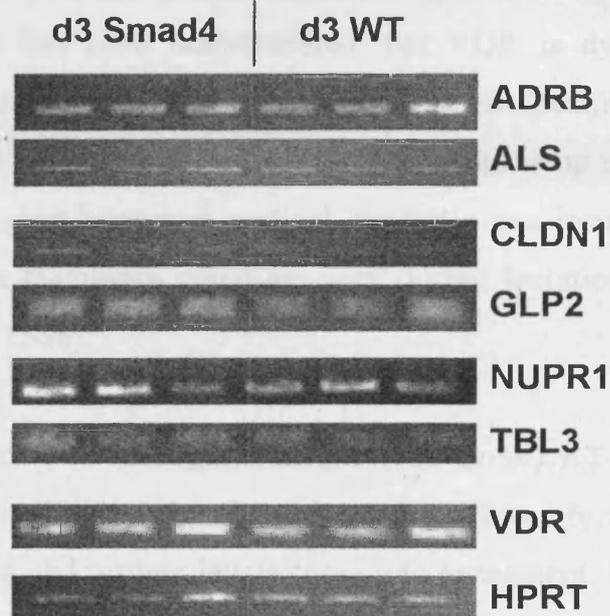
Clones were identified by NCBI'S UNIGENE database (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene>) for sequence similarity with known genes. Genes identified as significant in the array with known literature association to TGF- $\beta$ , apoptosis and/or p27<sup>Kip1</sup> were selected as candidates for further analysis. In total 7 genes were subjected to semi quantitative RT-PCR analysis: Adrenergic receptor beta 2 (ADRB2), Acid labile subunit of insulin like growth factor (ALS), Claudin 1 (CLDN1), Interferon-alpha inducible protein (Glp2), Nuclear Protein 1 (NUPR1), Vitamin D Receptor (VDR) and Transducin beta like 3 (TBL3). Representative photos from RT-PCRs are illustrated in Figure 4.7. For primer sequences, PCR programs and product sizes see materials and methods (2.11).

Rank	GI Number	Name	Log Ratio (exp/control)	Fold Difference
<b>Induced Genes</b>				
1	5244427	unknown	1.291642	1.65
2	5357450	G1p2: interferon, alpha-inducible protein	1.213956	1.68
3	4787196	unknown	1.1391243	1.75
4	50100	ADRB2: Adrenergic receptor, beta 2	1.1260933	1.58
5	2660687	lysosomal alpha-N-acetylglucosaminidase	1.1219273	1.71
6	6098665	Blvra: biliverdin reductase A	1.1193939	1.58
7	6008672	Mrps18a: mitochondrial ribosomal protein S18A	1.1136907	1.70
8	4832500	Ndubfab1: NADH dehydrogenase (ubiquinone) 1	1.0986256	1.62
9	5488686	unknown	1.0950076	1.50
10	5477614	unknown	1.0921962	1.58
11	6099931	unknown	1.0870913	1.39
12	2765361	Acyl-CoA thioesterase	1.0838415	1.56
13	3426168	Ly86: lymphocyte antigen 86	1.0758751	1.37
14	5469733	Psm5: proteasome 26S subunit, non-ATPase, 5	1.0746457	1.47
15	4441612	unknown	1.0741596	1.49
16	5497270	Fatty acid desaturase (61% identity)	1.072327	1.51
17	5103141	Mpd1: mannose-P-dolichol utilization defect 1	1.0716086	1.43
18	1621612	Insulin-like growth factor binding protein	1.0704348	1.52
19	1213012	GOS2-like protein	1.0697198	1.42
20	6096137	unknown	1.0674593	1.39
21	6096364	Nola2: nucleolar protein family A, member 2	1.0664326	1.40
22	6008767	Vdr: vitamin D receptor	1.065815	1.38
23	2253712	Ras-related protein (DEXRAS1)	1.0639931	1.39
24	5490578	unknown	1.0605533	1.26
25	3127048	cGMP phosphodiesterase (PDE9A*1)	1.0591679	1.23
26	2892281	Fbxw5: F-box and WD-40 domain protein 5	1.058297	1.28
27	6096512	Mfap5: microfibrillar associated protein 5	1.0572589	1.33
28	6098645	Rtn1: reticulon 1(68.06% identity)	1.0572367	1.24
29	5474052	unknown	1.0526707	1.39
30	3860028	FK506 binding protein (Fkbp63)	1.0515857	1.31
31	200231	Hydroxymethylbilane synthase	1.0506994	1.25
32	5493772	Pip5k2c: phosphatidylinositol-4-phosphate 5-kinase	1.0495584	1.25
33	6098472	Sec61b: Sec61 beta subunit	1.0486744	1.33
34	6291225	Entpd8: ectonucleoside 3P diphosphohydrolase 6	1.0453636	1.25
35	5931566	Cks1: CDC28 protein kinase 1	1.0447385	1.25
36	5469365	unknown	1.0432319	1.24
37	6097524	Vps4a: vacuolar protein sorting 4a	1.0432191	1.25
38	5492428	unkown	1.041302	1.20
39	1044898	RAD23b homolog	1.0401101	1.20
40	5496217	Pex11b: peroxisomal biogenesis factor 11b (38.22%)	1.0382703	1.17
41	5771446	C7-1 protein	1.0379301	1.25
42	6253284	unknown	1.0378962	1.20
43	2592911	unknown	1.0361223	1.20
44	5910615	unknown	1.0350326	1.18
45	5498807	unknown	1.0341911	1.22
46	6098125	unknown	1.0319457	1.21
47	309180	Col5a2: procollagen, type V, alpha 2	1.0308961	1.20
48	6100432	unknown	1.0308267	1.21
49	5497337	unknown	1.0292498	1.14
50	1872796	unknown	1.0285726	1.19
51	6100870	Ormdl2: ORM1-like 2	1.0258752	1.14
52	55216	Valosin containing protein (S. cerevisiae) (83.01%)	1.0256666	1.17
53	6099582	unknown	1.0251573	1.15
54	5470060	unknown	1.0249225	1.15
55	2104688	alpha glucosidase II alpha subunit	1.0240113	1.13
56	5469805	Mitochondrial 28S ribosomal protein S21	1.0231315	1.14
57	6098503	unknown	1.0199849	1.11
58	6099715	unknown	1.0131259	1.06

Figure 4.5. List of significant up regulated gene transcripts in the BLG-Smad4 mice. Ranked is based on their log ratio. Raw Fold changes are also included. Percentages in brackets (wherever applicable) represents protein similarity to mouse.

Rank	GI Number	Name	Log Ratio (exp/control)	Fold Difference
<b>Repressed Genes</b>				
1	2911371	Igk-V22:Immun kappa light chain region	0.76857741	0.34
2	2232095	S107V1: mutated immunoglobulin heavy chain	0.85269757	0.45
3	4613476	Cldn1: claudin 1	0.85875699	0.70
4	3800794	unknown	0.86752001	0.63
5	984941	Il15: interleukin 15	0.88006573	0.76
6	202451	Zinc finger protein 35	0.88523575	0.66
7	5311002	Psmb5: proteasome subunit, beta type 5	0.89122078	0.63
8	5492352	Btbd1: BTB (POZ) domain containing 1	0.89288354	0.69
9	6182775	unknown	0.90569688	0.67
10	473406	Hsp70-related NST-1 (hsr.1)	0.90759725	0.72
11	5476888	unknown	0.90885746	0.68
12	2991396	unknown	0.91045998	0.67
13	3417385	Mtap7: microtubule-associated protein 7	0.91929666	0.69
14	5496547	Nupr1: nuclear protein 1	0.91936337	0.58
15	2692014	Cdan1: (90% identity)	0.91959959	0.73
16	5905762	Nup50: nucleoporin 50 (99.57% identity)	0.92627494	0.71
17	1209719	B-cell leukemia/lymphoma 6	0.92941388	0.69
18	2979018	unknown	0.93016049	0.67
19	6097421	Spin: spindlin (99.56% identity)	0.93128578	0.73
20	5869933	Rab6ip1: Rab6 interacting protein 1 (98.36%)	0.93182718	0.72
21	6096384	unknown	0.93331829	0.96
22	6172102	unknown	0.9338621	0.78
23	2652035	unknown	0.93450206	0.76
24	4720470	Cggbp1: CCG triplet repeat binding protein 1	0.9391079	0.77
25	520479	loricrin gene	0.94417112	0.77
26	4602807	unknown	0.94511179	0.74
27	5498535	unknown	0.945216	0.83
28	5475706	unknown	0.94801975	0.81
29	192166	IgE-binding factor	0.94805197	0.76
30	5496643	unknown	0.94853873	0.80
31	1002423	yolk sac permease-like molecule 1 (YSPL-1)	0.94941527	0.87
32	897846	Ubiquitin-conjugating enzyme E2H	0.95087262	0.96
33	4320863	Tbl3: transducin (beta)-like 3 (25.62 identity)	0.95111102	0.79
34	191743	adipose fatty acid binding protein (422)	0.95123297	0.81
35	6098931	D11Moh35: DNA segment, Chr 11 (42.77%)	0.9512788	0.86
36	199142	Milk fat globule-EGF factor 8 protein	0.95698024	0.81
37	220404	Casein gamma	0.95706324	0.85
38	4615467	unknown	0.95732447	0.83
39	5184330	unknown	0.95769219	0.85
40	5469257	unknown	0.95980886	0.86
41	191574	Casein alpha	0.96008203	0.82
42	2642628	cysteine string protein mRNA	0.9603419	0.91
43	6514944	unknown	0.96144054	0.81
44	5491016	unknown	0.96147292	0.88
45	200991	Small nuclear ribonucleoprotein B	0.96314075	0.89
46	200949	Stearoyl-coenzyme A desaturase 1	0.96455125	0.86
47	194417	Mouse germline IgH chain gene	0.96525624	0.80
48	52848	mRNA for 21 kd polypeptide	0.96689535	0.86
49	988280	Y-box binding protein MYB-1a mRNA	0.97026808	0.89
50	1938401	ribosomal protein L41 mRNA	0.97063618	0.86
51	5907730	Whsc2: Wolf-Hirschhorn syndrome 2(25.22%)	0.97252304	0.88
52	1199649	thymic shared antigen-1 (TSA-1)	0.97260459	0.86
53	1044896	RAD23a homolog (S. cerevisiae)	0.9750189	0.85
54	1360011	Artificial mRNA for single chain antibody scFv	0.97671671	0.88
55	5489762	unknown	0.97794677	0.95
56	5473355	Mapk1: mitogen activated protein kinase 1 (84.74%)	0.98818655	0.99

**Figure 4.6.** List of significant down regulated gene transcripts in the BLG-Smad4 mice. Ranked is based on their log ratio. Raw Fold changes are also included. Percentages in brackets (wherever applicable) represents protein similarity to mouse.



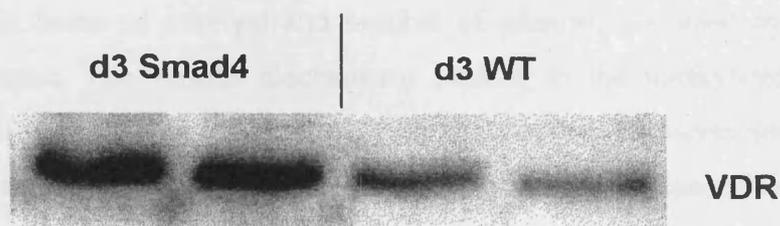
**Figure 4.7.** RT-PCRs for ADRB, ALS, CLDN1, GLP2, NUPR1, TBL3, VDR and HPRT as our control gene of choice. d3= day 3 involution. HPRT=Loading control.

Based on the RT-PCRs no major differences identified between Smad4 and wild type products except from the vitamin D receptor (VDR). VDR RT-PCR indicates an up-regulated gene expression profile in Smad4 compared to wild types, a result which is in agreement with the results obtained from microarray analysis.

#### 4.2.7 Smad4 over expression induces Vitamin D receptor.

Vitamin D receptor (VDR) is a receptor protein, member of the nuclear receptor super family (Carlberg 1995). VDRs are present in normal breast and many other epithelial tissues (Berger 1988). It has been demonstrated that VDR is dynamically regulated during pregnancy and lactation, but little is known about its specific functions. VDR is expressed at low levels in mammary glands of virgin rats and is up regulated in response to the differentiation inducing hormones cortisol, prolactin and insulin (Mezzetti 1987). Highest levels of VDR in mammary gland are seen during lactation, being maximal at day 3 involution (Colston 1988).

In order to confirm the mRNA up regulation observed in the RT-PCR I undertook a western blot analysis. This will also give the opportunity to identify whether a direct link between mRNA expression and protein levels come into agreement. Figure 4.8 illustrates a representative VDR western blot from day 3 involuting mice. There is an up-regulation of VDR's protein levels in the Smad4 transgenics compared to wild types at day 3 of involution.



**Figure 4.8.** Western blot analysis for vitamin D receptor protein. VDR is a 51kD protein. A clear up regulation of Vitamin D receptor at day 3 Smad4 mice compared to day 3 wild type. D3=Day 3 involution

#### 4.2.8 Raw Data Analysis and the MBD2 case.

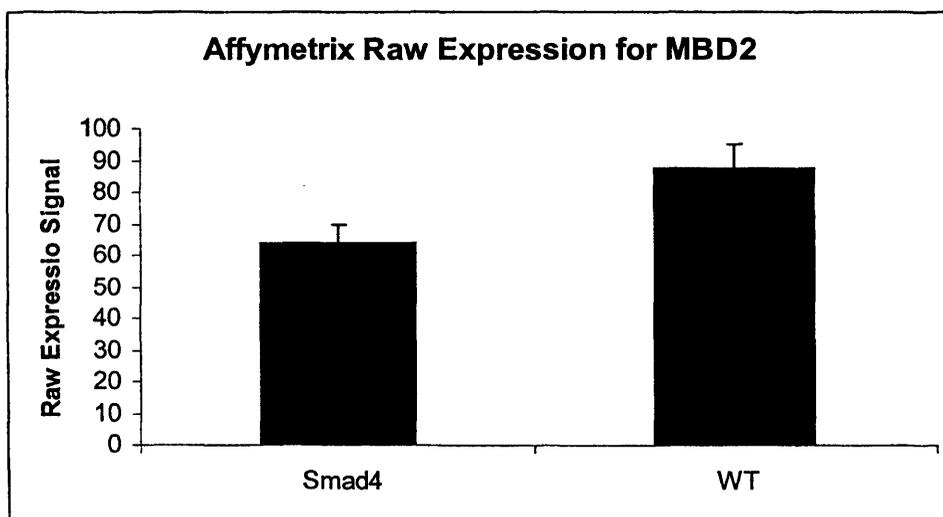
Methylation of CpG dinucleotides is an important epigenetic mechanism used by vertebrate cells to control transcription of many tissue specific genes (Bird 2002). To date, a direct link correlating DNA hypermethylation, hypoacetylation of histones, tightly packed chromatin, and transcriptional repression exists. CpGs in the genome are methylated as 5 methyl cytosine (approximately 85%) apart from large CpG islands in the promoters of genes which are unmethylated. The best example of this is the dense methylation of CpG promoters on the inactivated X chromosome and the *Xist* gene on the active X chromosome (Hendrich 2000). Effects of DNA methylation are mediated through proteins, which bind to symmetrically methylated CpGs. Such proteins contain a specific domain, the Methyl-CpG-Binding-Domain (MBD). Five MBD (MBD 1-4 and MeCP2) proteins have been identified to date all sharing the same functional Methyl Binding Domain (MBD domain) (Hendrich 1998).

It has been recently shown that, in colon cancer cell lines, MBD2 is associated with the aberrantly methylated promoters of silent *p14/p16* genes and this methylation-dependant association seems to be responsible for their silencing (Magdinier and Wolffe 2001). Furthermore, recently Sansom (2003) demonstrated that MBD2 deficient mice resisted intestinal cancer in terms of survival and number of adenomas caused by the tumour prone  $Apc^{Min/+}$  mouse. The various mechanisms leading to the methylation-dependant down regulation of transcription remain to be fully determined. However several lines of evidences indicate that, among them, the targeting of histone deacetylases (HDACs) complexes mediated by methyl-CpG binding proteins plays a major role (Bird and Wolffe 1999).

Abnormal methylation is a cause of human genetic diseases, including ICF (Immunodeficiency-Centromeric Instability-Facial anomalies syndrome) (Hansen 1999), and is involved in carcinogenic processes primarily through aberrant hypermethylation of tumor suppressor's promoter regions (Jones 2002). MBDs are important constituent of

the DNA methylation machinery, since they are directly involved in the mediation of the epigenetic signal (Bird 1999).

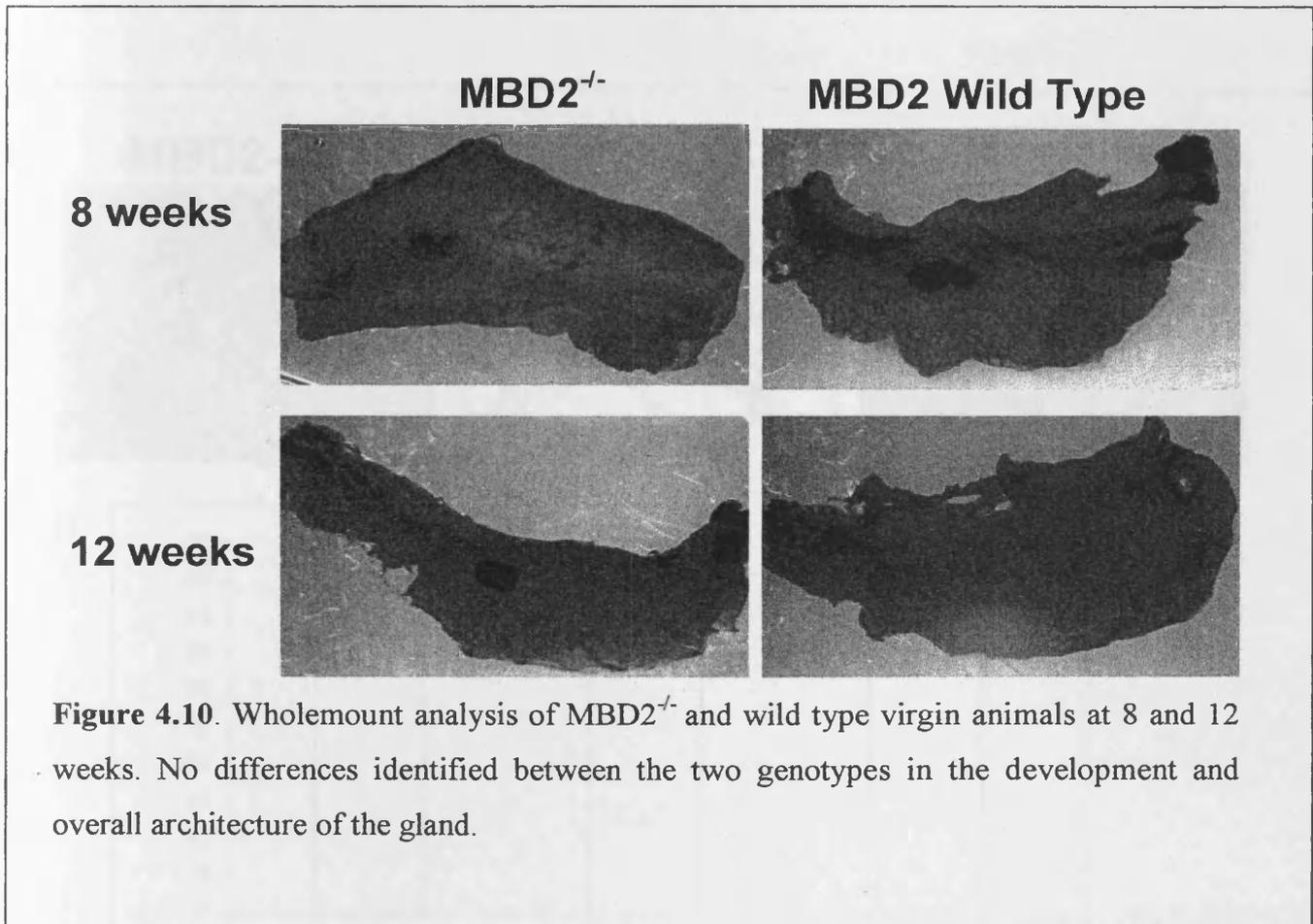
Raw data analysis indicated a down regulation of the MBD2 transcript (Figure 4.9). Therefore, it seemed appropriate to utilize conditional knockout MBD2 mice located in our animal house and try to investigate the consequences of MBD2 depletion in the mammary gland. My hypothesis was that the observed accelerated involution in the Smad4 transgenic mice can be correlated with the down regulation of MBD2.



**Figure 4.9.** Raw expression values for MBD2 indicates a significant down regulation in the BLG-Smad4 transgenic mice compared to wild types at day 3 involution.

#### 4.2.8.1 Depletion of MBD2 from the mammary gland does not alter mammary gland architecture in virgin mice.

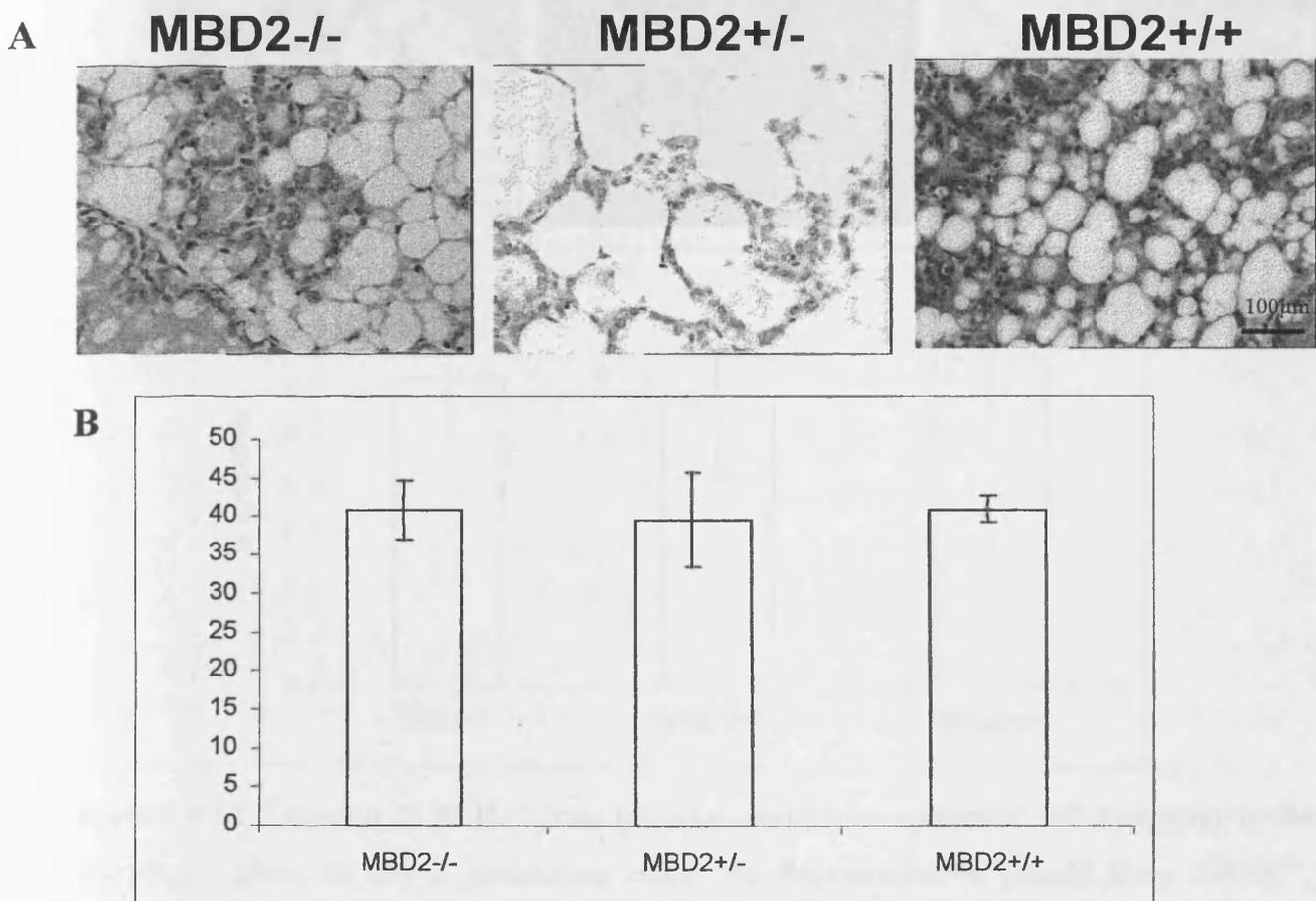
In order to investigate whether MBD2 depletion can create differences in mammary gland development, wholemounts were generated between MBD2<sup>-/-</sup> and wild type virgin mice. Figure 4.10 illustrates representative photos from 8 and 12 week old virgin mice.



**Figure 4.10.** Wholemount analysis of MBD2<sup>-/-</sup> and wild type virgin animals at 8 and 12 weeks. No differences identified between the two genotypes in the development and overall architecture of the gland.

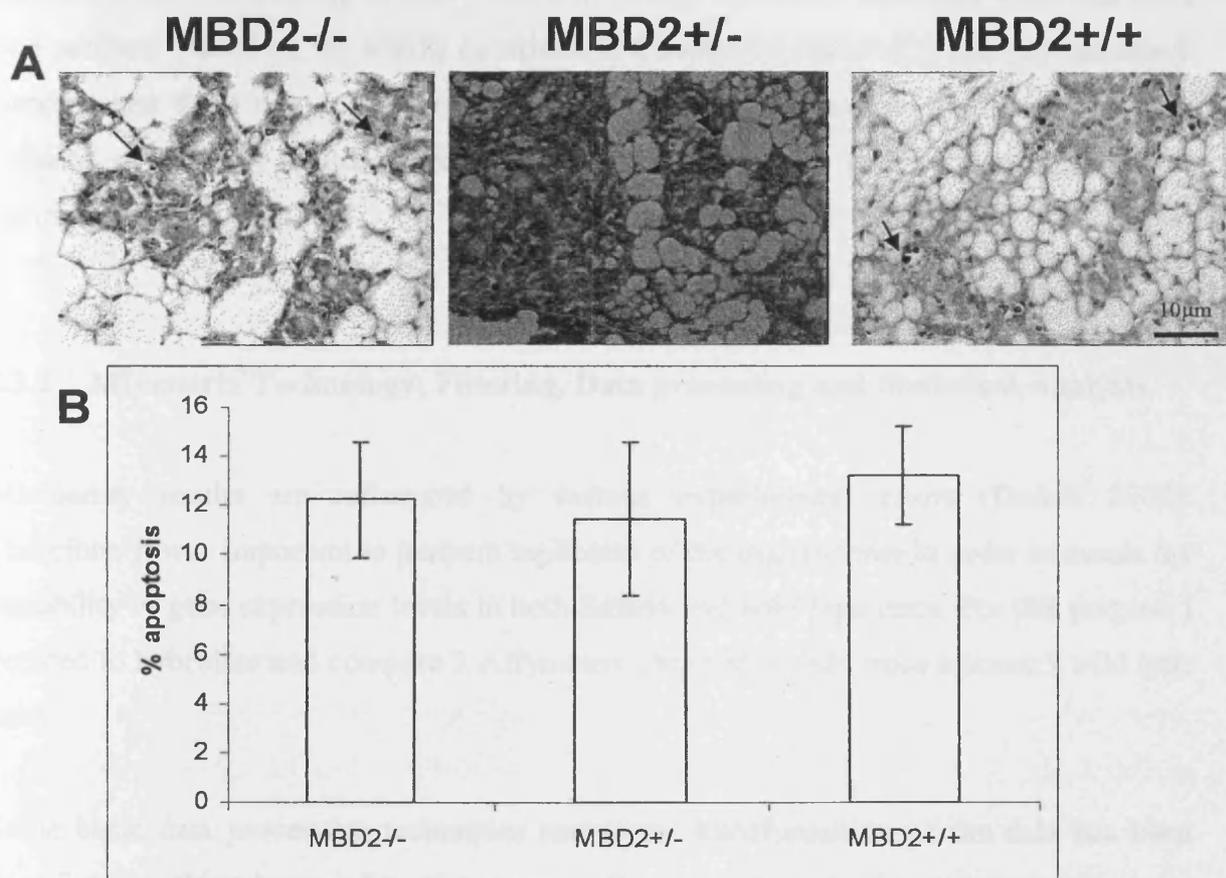
#### 4.2.8.2 Depletion of MBD2 from the mammary gland does not alter involution nor apoptosis of mammary epithelial cells at day 3 involution.

To investigate any possible alterations in the contribution of epithelial cells in mammary gland between MBD2<sup>-/-</sup>, MBD2<sup>+/-</sup> and wild type animals a histological phenotypic analysis by sectioning and staining with hematoxylin-and-eosin (H&E) was undertaken. The number of mice for this analysis was at least 3 (n=3). Time point used was day 3 involution. Quantification of slides as well as representative photos of the glands are illustrated in Figure 4.11.



**Figure 4.11.** Deletion of MBD2 does not alter mammary epithelial cell contribution in the mammary gland at day 3 involuting mice. A) Representative glands from MBD2<sup>-/-</sup>, MBD2<sup>+/-</sup> and MBD2 wild types. B) Quantification of slides for fat contribution showed no significant changes between the three genotypes. Scale bar 100µm.

To investigate any possible alterations in the regulation of apoptosis by depletion of MBD2, TUNEL analysis performed (Figure 4.12). In agreement with the above result MBD2 deletion from the mammary gland did not alter the apoptotic index.



**Figure 4.12.** Deletion of MBD2 does not alter mammary epithelial cell apoptosis in the mammary gland at day 3 involuting mice. A) Representative glands from MBD2<sup>-/-</sup>, MBD2<sup>+/-</sup> and MBD2 wild types. Arrows indicate apoptotic cells. B) Quantification of slides for apoptotic cells showed no significant changes between the three genotypes. Scale bar 10µm.

## **4.3 Discussion.**

### **4.3.1 Introduction.**

The technique of expression profiling by means of hybridization to cDNA arrays offers a new tool for investigating the expression levels of thousands of genes simultaneously. I have utilized this technology for comparing gene expression profiles between conditional transgenic mice expressing human Smad4 in mouse mammary epithelial cells and wild type controls. Based on the results as written in Chapter 3 I decided to analyze mammary glands taken from day 3 involuting mice. My aim in this part of the project was to validate or identify genes differentially expressed in an over expressing Smad4 environment.

### **4.3.2 Affymetrix Technology; Filtering, Data processing and Statistical Analysis.**

Microarray results are influenced by various experimental errors (Dudoit 2002). Therefore, it was important to perform replicates of the experiments in order to assess the variability of gene expression levels in both Smad4 and wild type mice. For this purpose I decided to hybridize and compare 3 Affymetrix chips of Smad4 mice against 3 wild type ones.

Some basic data processing techniques and linear transformation of the data has been described in this chapter. In microarray analysis, exclusion of non-informative gene expression values (Absents) before normalize and statistically validate the data is of great importance. Currently, various filtering procedures exist and a comprehensive list of them can be found in Saviozzi (2003). The stringency of the filtering procedure can strongly affect (in a positive or a negative manner) the final results, as it can cause the loss of differentially expressed genes or increase the number of false positives contaminating the final results. For the purpose of this analysis I decided to filter the data by subtracting gene expression values which were considered as 'Absents' by the

Affymetrix software. Filtering through the method described above reduced drastically the data from 12800 transcripts to 5808. This is a considerable reduction of ~65%, however, filtering of this kind is a necessity for removing false positives. By filtering out transcripts characterized as 'Absents' I reduce the possibility of identifying false positives but at the same time I exclude valuable targets that I wish to analyze.

A characteristic example is  $p27^{Kip1}$ . Protein levels of  $p27^{Kip1}$  were found to be significantly increased at day 3 involuting BLG-Smad4 transgenic mice compared to controls (Chapter 3). Unfortunately,  $p27^{Kip1}$  transcript was identified as 'Absent' in all 6 Affymetrix chips, and therefore was filtered out from further analysis despite my interest in analyzing mRNA levels of this molecule. From the observed absence of  $p27^{Kip1}$  mRNA expression at day 3 involution I can hypothesize that its RNA expression probably occurs at an earlier time point (day 1 or 2 of involution) and regulate  $p27^{Kip1}$ 's protein over expression as observed in the day 2 and day 3 involution western blot samples.

Array experimental conditions can strongly affect microarray hybridization intensities. It has been demonstrated that sources of error are multiplicative and can strongly affect true expression levels (Hartemink 2001), especially if genes are moderately expressed (Rocke 2001). Therefore, normalization of gene expression data is a crucial preprocessing procedure essential for nearly all gene expression studies in which data from one array (or set) is compared to data from another array (or set). A number of normalization approaches may be taken into account (Kim 2000, Colub 1999), each of which utilize different software and algorithms, however, a gold standard method for microarray data normalization has not been defined. For a representative review on different normalization methods see Wolkenhauer (2002).

For the purpose of this data normalization I utilized a method of data transformation which is suggested by Affymetrix; log transformation. Log transformation serves data in a dual role. Firstly, it accomplishes reduction of signal errors (in an absolute sense) and furthermore, it transforms data in a normal like shape. Transformation of the data to a

linear form allows statisticians to perform parametric statistical tests. Data transformation has been achieved by logging in the base of 2.

I have utilized a publicly available software SAM (Significant Analysis of Microarrays) (Tusher 2001) which can be located at <http://www-stat-class.stanford.edu/SAM/SAMServlet> for identifying differentially expressed genes. SAM is a software specifically designed for normalizing (by logging) and analyzing microarray data. Analysis is performed in logged microarray values by assimilating a set of gene-specific *t* tests. Genes with scores over a specified by the user cut off value are considered as significant. SAM has also the ability to identify false differentially expressed genes (genes which have been identified by chance) amongst the significant pool (False Discovery Rate) by analysing permutations of the measurements.

SAM was unable to identify differentially expressed genes due to the fact that the observed significant targets were lower than the expected ones. This was demonstrated when statistical analysis was undertaken manually by a student's *t-test*. In this test the significant differentially expressed genes identified were 114 (observed) when the expected False Discovery Rate (FDR) out of 5800 transcripts in a 5% probability risk level is 290 (see below).

Following data transformation by logging, I performed a parametric *t-test*. Parametric statistics test the hypothesis that one or more treatments (in our case BLG-Smad4) have no effect on the mean (or median) of a chosen variable. 114 ESTs and clones were identified as statistically significant in a 5% probability risk level. Statistically significant targets show identical expression values between experimental and control groups and exclude the possibility of data "contamination" by outliers generated by experimental noise.

The number of statistically significant identified targets is "relatively few" compared to the thousands analyzed. This failure to identify many changes may reflect the dynamics and composition of the mammary gland itself. As stated at chapter 1 the mammary gland

is a very dynamic tissue. Epithelial cells undergo programmed cell death and are substituted by adipocytes during involution. The quantity of epithelial cells as well as their quality regarding RNA integrity and expression stage cannot be precisely controlled in this *in vivo* approach. The above factors may well combine in contributing to the generation of biological outliers between samples i.e. samples with different expression values for a gene in the same pool, and therefore reduce the number of statistically significant targets. To overcome this potential problem I attempted to pursue microarray analysis by performing primary mammary cell cultures where cell quantity, quality and RNA expression status can be controlled. As I illustrate in Chapter 5 this is the method of choice for extracting RNA and is a method of producing more statistically significant results than a pure *in vivo* approach.

#### **4.3.3 Smad4's Expression profile.**

No major up- or down- regulated genes were identified following calculation of the ratios of the mean logs between experimental and control groups. Ratio of the log mean values above 1 were considered as 2 fold up-regulated in the BLG-Smad4 mice compared to experimentals whereas ratios below 1 to 0 considered as down regulated (Note:  $\text{Log}_2 2=1$ ,  $\text{Log}_2 0.5=-1$ ). All statistically significant targets showed ratios between 1.29 to 0.76. More specifically 58 clones and ESTs were induced in Smad4 transgenic samples and 56 were repressed.

The observed ratio of log means indicate close mRNA expression profiles at day 3 involution between experimental and control groups despite the phenotypic and molecular changes as observed and written in chapter 3. Possible explanations for this phenomenon might be the time point itself or that these sets of data need more replicates. I hypothesize that at day 3 of involution large numbers of epithelial cells have already commit apoptosis therefore mRNA signal for the genes responsible for this programmed cell death probably may have occurred at earlier time points (e.g. the  $p27^{Kip1}$  case). Therefore, it may be interesting to repeat the same experiments during day 2 or even day

demonstrated in transient experiments a synergistic activation of the osteocalcin gene promoter by binding of a Smad3-Smad4 heterodimer and Vitamin D receptor in their cognate DNA recognition elements (SBE and Vitamin D Receptor Element-VDRE respectively) located in close proximity between them (Subramaniam 2001). Smad3 acts as a co-activator of Vitamin D receptor and positively regulates the VDR signaling pathway. Negative regulation of VDR function has also been demonstrated both *in vivo* and *in vitro* by Smad7 inhibition in the formation of the VDR-Smad3 complex, whereas Smad6 had no effect (Yanagi 1999).

Following all data handling and RT-PCR analysis I identified up regulation of the Vitamin D receptor mRNA level. Western blot analysis also indicated an up-regulation of VDR protein levels in the Smad4 transgenics compared to wild types at the 3<sup>rd</sup> day of involution. This result agrees with previous findings correlating increased levels of TGF- $\beta$  with VDR induction. What remains to be elucidated is whether VDR's induction is due to increased levels of Smad3 brought to the nucleus by Smad4 or by Smad4 alone.

#### 4.3.5 MBD2.

Affymetrix raw data indicated a down regulation of MBD2 mRNA levels in day 3 involuting Smad4 transgenic mice. MBD2 belongs to the family of the methylation binding proteins which responsible for the methylation of CpG dinucleotides. DNA has been characterized as an important epigenetic mechanism for controlling transcription of many tissue specific genes (Bird 2002). It has been demonstrated that promoter methylation by MBD2 causes gene silencing and colon cancer (Magdinier and Wolffe 2001). In contrast Sansom (2003) showed that MBD2 depletion suppressed and resistance in intestinal cancer in terms of survival and number of adenomas caused by the tumour prone Apc<sup>Min/+</sup> mouse. Taking this into account it seemed appropriate to hypothesize that MBD2 repression in the Smad4 over expressing mice could be one of the reasons for the accelerated involution and the increased apoptotic index and that MBD2 could regulate mammary gland development or involution.

In order to investigate the above hypothesis I decided to undertake a wholemount analysis of virgin knockout MBD2 mice as well as analysis at day 3 of involution. Wholemount analysis in the MBD2 knockout mice could provide valuable information regarding its role in mammary gland development while investigation of the same mice at day 3 of involution could shed light in the role of MBD2 in mammary epithelial cell apoptosis. Both analyses indicated that MBD2<sup>-/-</sup> and MBD2<sup>+/-</sup> did not have an altered phenotype or apoptotic differences when compared to wild type glands.

## **Chapter 5-Smad-Stat Interactions and Microarray Analysis of Stat3<sup>+/-</sup>.**

### **5.1 Introduction.**

#### **5.1.1 STATs and the mammary gland.**

Signal Transducers and Activators of Transcription (STATs) proteins comprise a family of transcription factors latent in the cytoplasm that participate in normal cellular events, such as differentiation, proliferation, cell survival, apoptosis and angiogenesis. STATs get activated in response to many cytokines and growth factors (Schindler 1995) and have been shown to be capable of both activating and suppressing gene transcription (Fukada 1998) as well as act as growth suppressors (Kaplan 1998) or growth promoters (Catlett-Falcone 1999) depending on the particular STAT and the promoter sequence to which they bind.

The STAT signaling pathway has well been defined as an archetypal membrane to nucleus signal transduction pathway, which utilizes the Janus kinase (JAK) proteins to achieve a cascade of phosphorylation signals leading to regulation of gene transcription. The classic activation pathway involves the cytokine ligand Interferon- $\gamma$  (IFN $\gamma$ ) which upon binding to its corresponding IFN $\gamma$ - receptor induces the oligomerisation of its subunits, leading to phosphorylation and activation of the pre-associated JAKs. The phosphorylated JAKs in turn phosphorylate the IFN $\gamma$  receptor, allowing STATs to bind, get phosphorylated and form a homodimer that migrates to the nucleus (Figure 5.1). The specificity of STATs for receptors appears to depend on the cell type and a single ligand can activate a number of different STATs. To date seven STAT proteins have been identified and divided into two functional groups. STAT2, 4 and 6 comprise the first group due to their activation by a small number of cytokines and their role in the development of T-cells. STAT 1, 3 and 5 (Stat5a and Stat5b), comprise the second group due to their activation by a series of ligands and their involvement in the development of

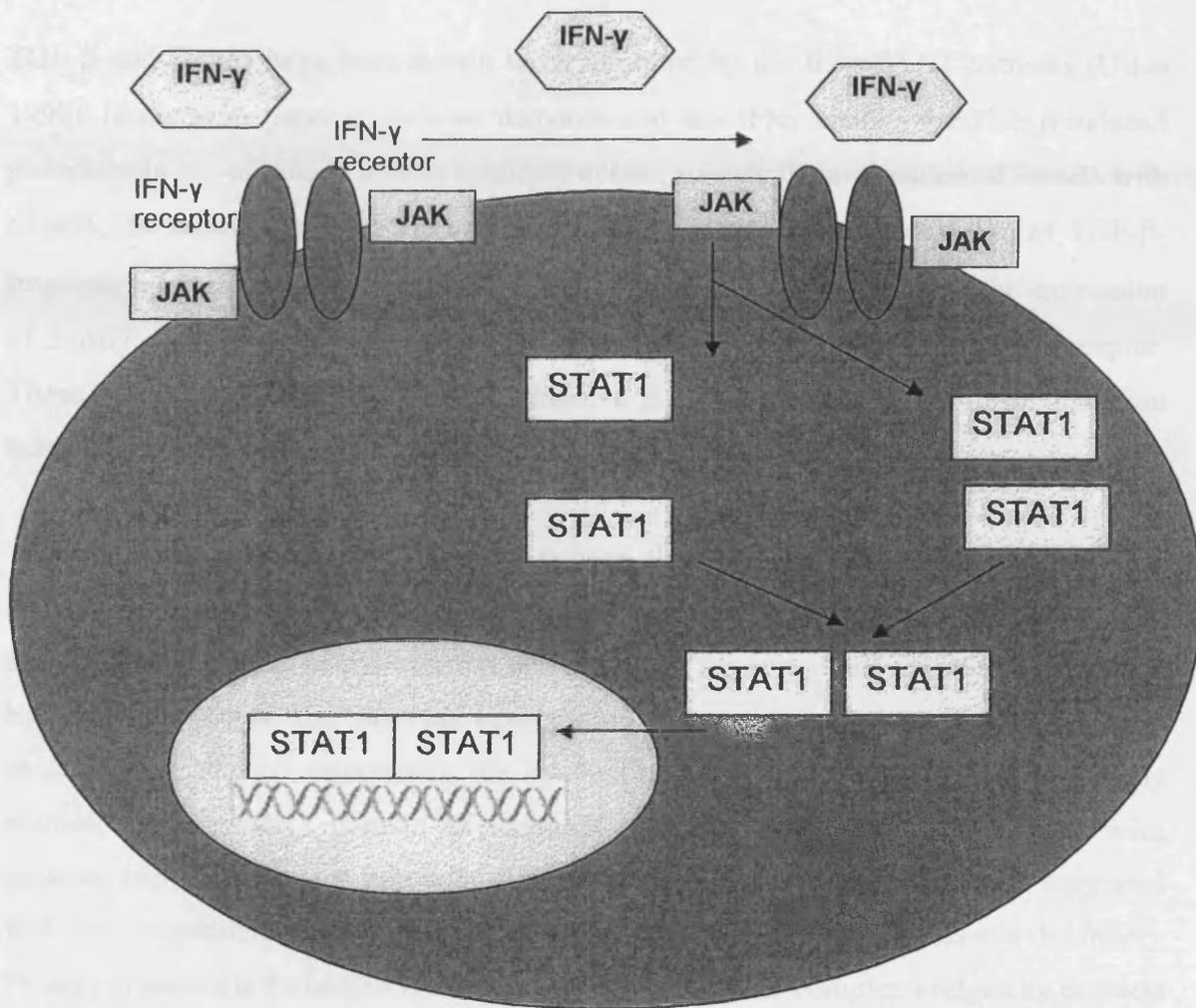
mammary gland and embryogenesis. This later group of STATs plays an important role in controlling cell-cycle progression and apoptosis (Bromberg 2002). It has been demonstrated that STATs regulate transcription of a number of genes including: *IRF-1*, *Bcl-XL*, *cyclin D1*, *PIM-1*, *c-Myc* and *p21*.

### **5.1.2 Role of STAT3 in Mammary Gland Development.**

Analysis of phosphorylated protein levels and DNA binding activities have shown that STAT3 is activated at day 5 of pregnancy (Philip 1996). Furthermore, initiation of involution is characterized by changes in the phosphorylation/activity of STAT3 and STAT5a (Liu X 1996). STAT5a and STAT3 have reciprocal patterns of phosphorylation with decreasing STAT5a levels and increasing STAT3 at the beginning of involution (Li 1997). In the mammary gland, STAT5 activation can be induced by Prolactin (PRL), Epidermal Growth Factor (EGF) and Growth Hormone (GH) (Gallego 2001), whereas factors that activate STAT3 during involution are unknown.

Conditional STAT3 knockout mice have demonstrated that in the absence of STAT3, involution is not initiated for at least 3 days following the removal of pups and that extent of apoptosis in mammary epithelial cells is significantly reduced. Mammary glands in the conditional knockout animals do eventually undergo remodeling, demonstrating that a compensatory mechanism exist. However, remodeling is not complete and alveolar structures remain, indicating that a disruption of the remodeling pathways involving STAT3 leads to the survival of cells otherwise destined to undergo apoptosis (Chapman 1999).

Molecularly, STAT3 knockout animals have demonstrated possible mechanisms responsible for initiating delayed involution and apoptosis by up regulating *p21* and *p53* levels as well as levels of STAT1, a dimerization partner for STAT3. No changes occurred in the *Bcl-xl* gene while levels of IGFBP-5 (Insulin Growth Factor Binding protein-5) that normally increase at the 2<sup>nd</sup> day of involution, remained unchanged.



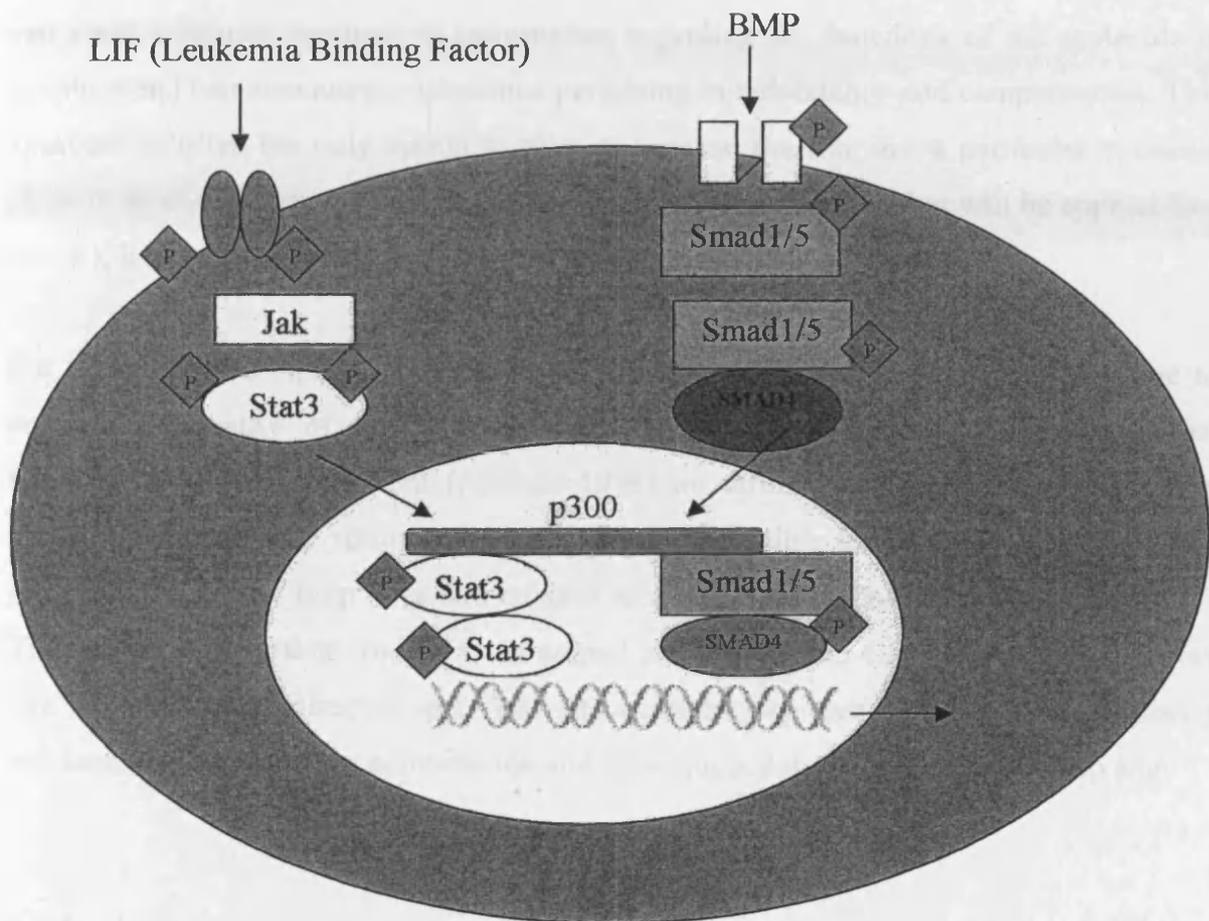
**Figure 5.1.** STAT signaling pathway. The classic activation pathway involves the cytokine ligand IFN $\gamma$  receptor subunits, leading to phosphorylation and activation of the preassociated JAKs, which in turn phosphorylate the IFN $\gamma$  receptor allowing STAT1 to dock. STAT1 is then phosphorylated and forms a homodimer that migrates to the nucleus where regulates transcription.

### 5.1.3 STAT-SMAD interaction.

TGF- $\beta$  and Smads have been shown to be inhibited by the IFN $\gamma$ /STAT pathway (Ulloa 1999). In the same paper it has been demonstrated that IFN $\gamma$  inhibits the TGF- $\beta$  induced phosphorylation of Smad3 and its attendant events, namely the association of Smad3 with Smad4, the accumulation of the complex in the nucleus, and the activation of TGF- $\beta$ -responsive genes. Furthermore, through JAK1 and STAT1, IFN $\gamma$  induces the expression of Smad7, an I-Smad which prevents the interaction of Smad3 with the TGF- $\beta$  receptor. These observations were the first to indicate a possible mechanism of transmodulation between the STAT and Smad signal-transduction pathways.

Transcriptional activation by R-Smads has been shown to occur, in part at least, by their ability to recruit the general coactivators p300 and CBP (Feng 1998, Shen 1998). p300 and CBP have histone acetyl transferase (HAT) activity, suggesting that their recruitment by a Smad complex may increase transcription of target genes by altering nucleosome structure and thereby remodeling the chromatin template. This interaction is directly mediated by the MH2 domain of R-Smads. p300 and CBP are large proteins with separate regions for interaction with different transcription factors. It has been suggested that the cooperative signaling of BMP2 and the cytokine LIF (Leukemia Inhibitory Factor) in astrocyte formation is mediated by a Smad1-Stat3 complex bridged by contacts with separate regions of p300 (Nakashima 1999). In addition, recently it was reported that PIAS3, a member of the protein inhibitor of activated STAT (PIAS) family interacts with Smad proteins, most strongly with Smad3. This interaction is accomplished via Smad3's MH2 domain both *in vivo* and *in vitro* through p300/CBP and leads to Smad induced transcriptional activity (Long 2004). Figure 5.2 illustrates a representative diagram of the Smad-STAT interplay.

## Smad-Stat3 Interaction



**Figure 5.2.** Cooperativity between Smad and Stat signaling pathways via p300/CBP.

#### 5.1.4 STAT3 Knockout.

In order to dissect the functions of individual components of complex biological systems it has become almost routine to engineer some type of transgenic mouse. This approach can yield immense amounts of information regarding the functions of the molecule in question and can also answer questions pertaining to redundancy and compensation. This approach is often the only option to fully understand the role that a particular molecule plays in an *in vivo* context and, combined with the other analyses that will be applied (see below), it is an approach that seems entirely appropriate for this project

For the purpose of this chapter I utilized a conditional Stat3 knockout mouse. Due to embryonic lethality of the STAT3 knockout mouse (Takeda 1997), a conditional knockout was been generated (Akaishi 1998) by utilizing the Cre-lox recombination system. A floxed Stat3 strain, in which an exon responsible for tyrosine phosphorylation has been flanked by loxp sites and crossed to a null Stat3 animal to generate a Stat3<sup>FL/-</sup>. This animal was further crossed to an animal carrying a BLG-Cre transgene. In this way Cre recombinase is directed specifically to mammary epithelial cells during pregnancy and lactation and causes recombination and subsequent deletion of STAT3's loxp site.

#### 5.1.5 Aims.

This chapter is divided into two parts. The first part will address the interacting Smad4-Stat3 role in the mammary gland by crossing the BLG-Smad4 transgenic construct onto a mammary conditional Stat3 'floxed' background carrying a BLG-Cre construct. Deletion of Stat3 from the mammary gland delayed involution (Chapman 1999), while Smad4 over expression in the mammary gland results in accelerated involution with a subsequent Stat3 down regulation at day 3 involution (Chapter 3). Here a comparison of BLG-Smad4-BLG-Cre-Stat3<sup>FL/-</sup> against BLG-Cre-Stat3<sup>FL/-</sup> mice will be investigated at day 3 of involution.

The second part of this chapter will investigate the consequences of a conditional deletion of Stat3 in the mammary gland by undertaking a microarray approach. In this way I aimed to identify alterations in the expression profile of known STAT3 targets and indicate new ones. For this attempt a comparison of mammary gene expressions between Stat3<sup>+/-</sup> and Stat3<sup>+/+</sup> mice will be carried out. In contrast to the Smad4 microarray experiments (see chapter 4), it was decided to approach this question by utilizing RNA from primary mammary epithelial cell cultures. Primary mammary cell cultures were used in order to avoid problems created by adipose tissue contamination.

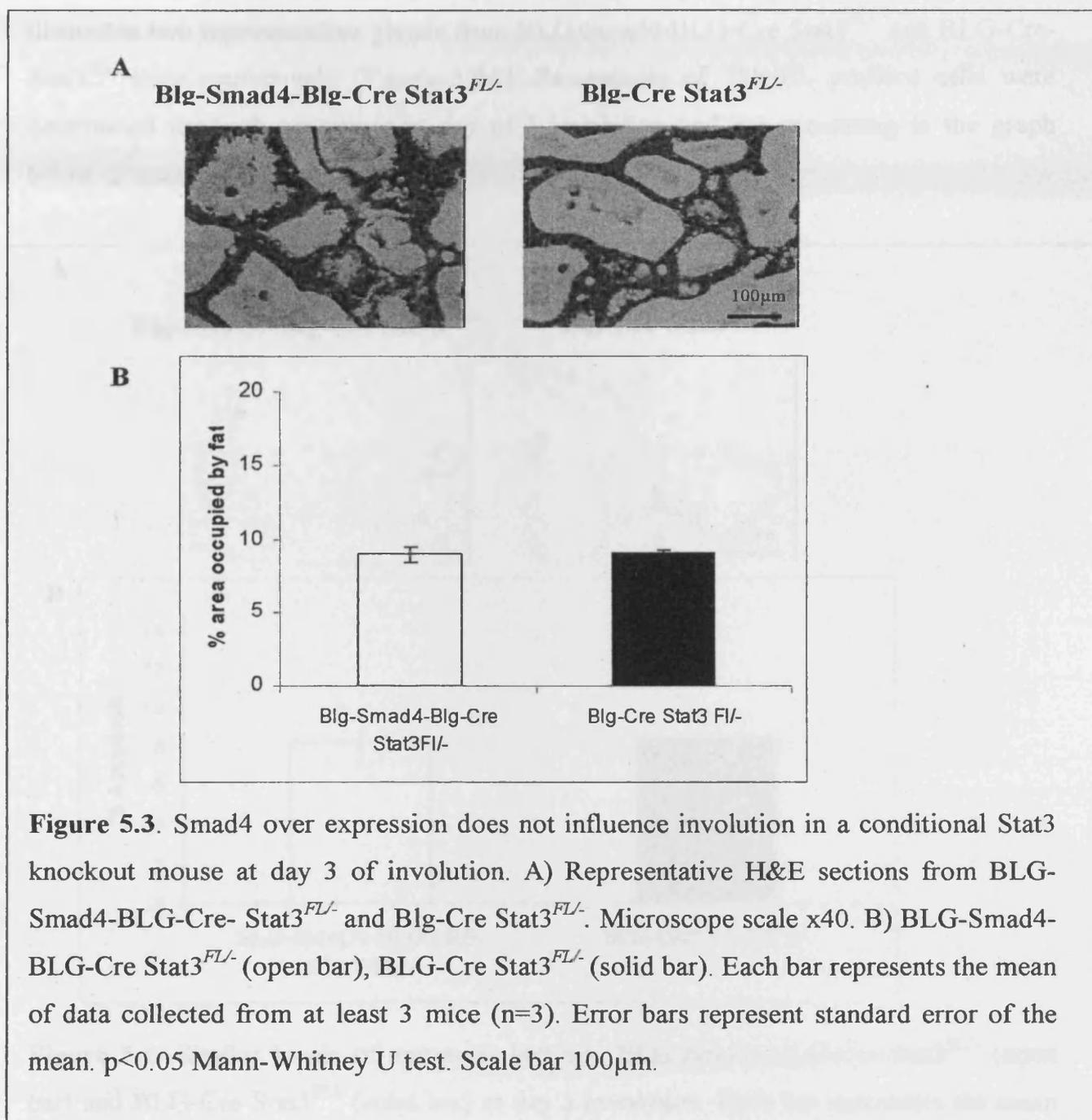
## 5.2 Results

### 5.2.1 Analysis of Over Expressing Smad4 in a Stat3 Null Environment.

#### 5.2.1.1 Smad4 over expression does not influence involution in a conditional Stat3 knockout mouse.

To investigate any possible alterations in the contribution of epithelial cells in the mammary gland between BLG-Smad4-BLG-Cre- Stat3<sup>FL/-</sup> (experimentals) and BLG-Cre-Stat3<sup>FL/-</sup> (controls) a histological phenotypic analysis by sectioning and staining with hematoxylin-and-eosin (H&E) was undertaken. The number of mice for this analysis was at least 3 (n=3). Following the observation of decreased STAT3 expression at day 3 of involution in the BLG-Smad4 mice I generated sections from the same time point. As in Chapter 3, mice were allowed to lactate for 10 days (6-8 pups) and then pups were withdrawn to initiate involution. Figure 5.3A illustrates representative glands from both experimental and control mice. Both genotypes exhibited a similar phenotype. Alveoli remained intact and distended and although a small number of apoptotic cells seen, glands retained the general appearance of a lactating gland. To quantify the amount of

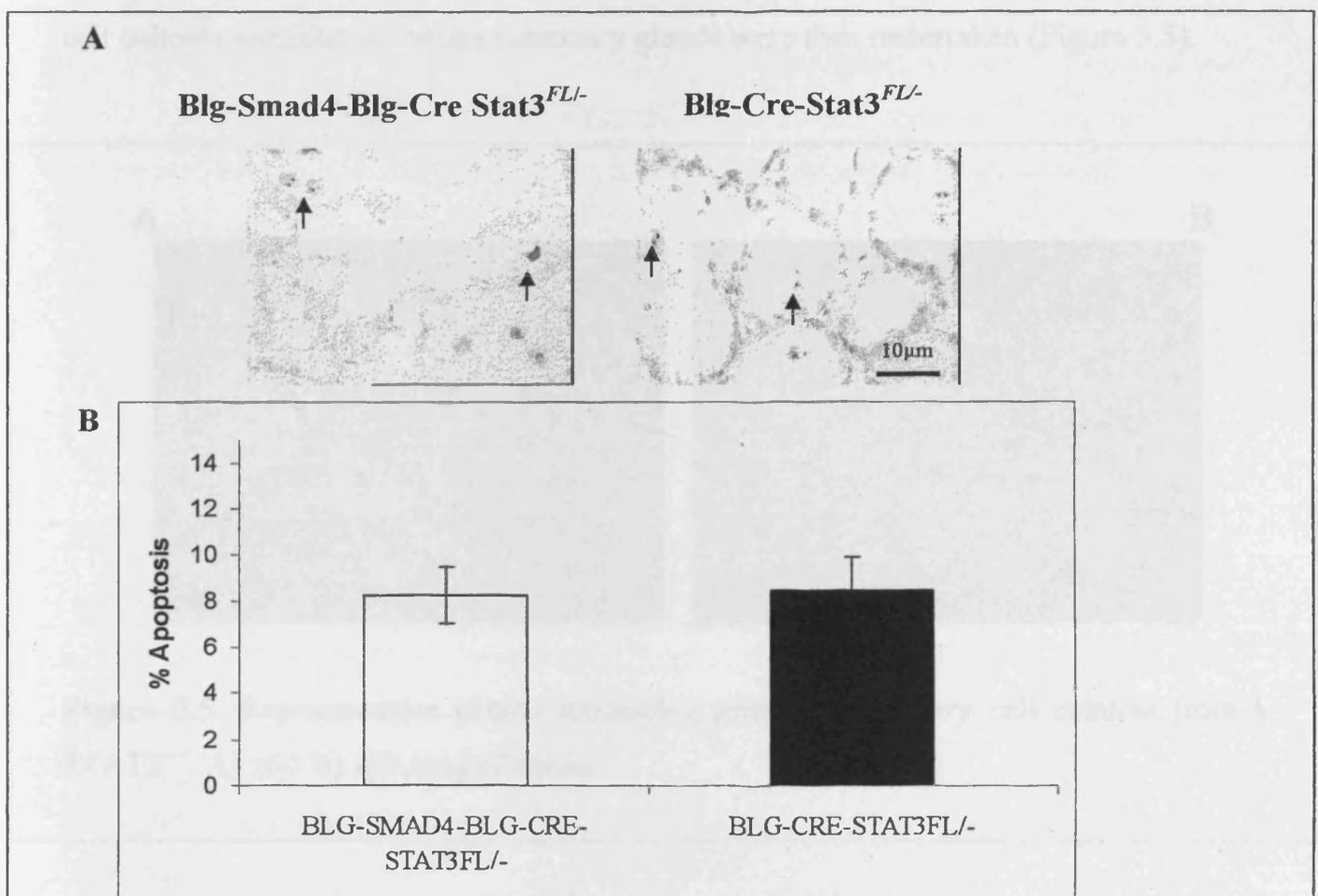
involution occurred the areas of the glands occupied by adipocytes were measured and illustrated in Figure 5.3B.



**Figure 5.3.** Smad4 over expression does not influence involution in a conditional Stat3 knockout mouse at day 3 of involution. A) Representative H&E sections from BLG-Smad4-BLG-Cre- Stat3<sup>FL/-</sup> and Blg-Cre Stat3<sup>FL/-</sup>. Microscope scale x40. B) BLG-Smad4-BLG-Cre Stat3<sup>FL/-</sup> (open bar), BLG-Cre Stat3<sup>FL/-</sup> (solid bar). Each bar represents the mean of data collected from at least 3 mice (n=3). Error bars represent standard error of the mean. p<0.05 Mann-Whitney U test. Scale bar 100µm.

### 5.2.1.2 Over expression of Smad4 does not alter apoptosis in a conditional Stat3 knockout mouse.

In order to confirm the above result it seemed appropriate to examine levels of apoptosis in the mammary tissues of both genotypes by performing TUNEL analysis. Figure 5.4A illustrates two representative glands from BLG-Smad4-BLG-Cre Stat3<sup>FL/-</sup> and BLG-Cre-Stat3<sup>FL/-</sup> mice respectively (Figure 5.4A). Percentages of TUNEL positive cells were determined for both genotypes at day 3 involution and are presented in the graph below (Figure 5.4 B).



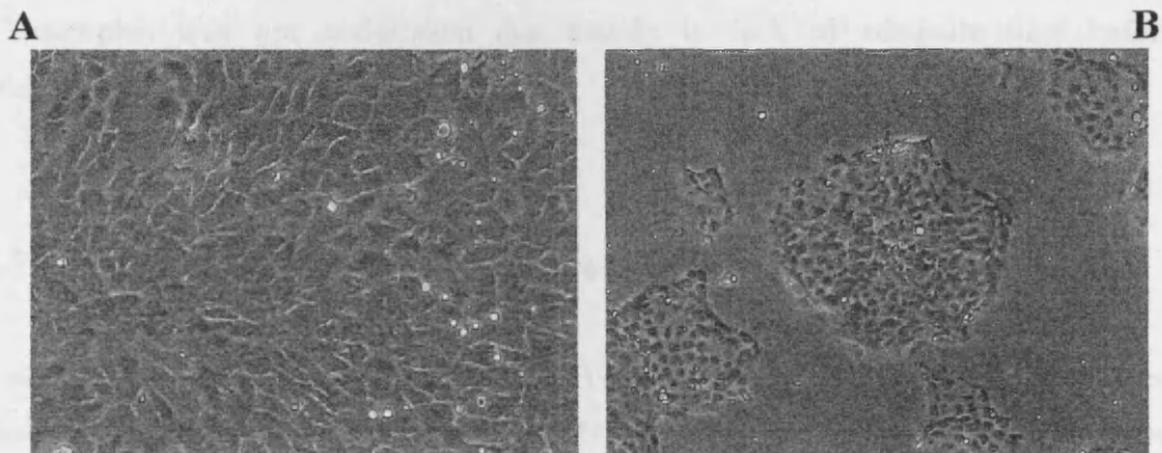
**Figure 5.4.** Similar levels of apoptosis between BLG-Smad4-BLG-Cre-Stat3<sup>FL/-</sup> (open bar) and BLG-Cre-Stat3<sup>FL/-</sup> (solid bar) at day 3 involution. Each bar represents the mean data collected by at least 3 mice. Error bars represent standard error of the mean.  $p < 0.05$  Mann-Whitney U test. Scale bar 10µm.

Taken together the above data I conclude that the BLG-Smad4 transgene does not modify the STAT3<sup>-/-</sup> phenotype at day 3 of involution.

## 5.2.2 Stat3 Microarrays.

### 5.2.2.1 Primary Mammary Epithelial cell cultures.

Stat3<sup>+/-</sup> and Stat3<sup>+/+</sup> mice (Takeda 1997) were mated in order to derive pregnant females, which were then sacrificed at their 16-18 day of pregnancy. Primary mammary epithelial cell cultures utilizing all mouse mammary glands were then undertaken (Figure 5.5).



**Figure 5.5.** Representative photos illustrating primary mammary cell cultures from a STAT3<sup>+/-</sup>. A) x60 B) x10 magnification.

### 5.2.2.2 The Stat3<sup>-/-</sup> deficient cultures were not viable.

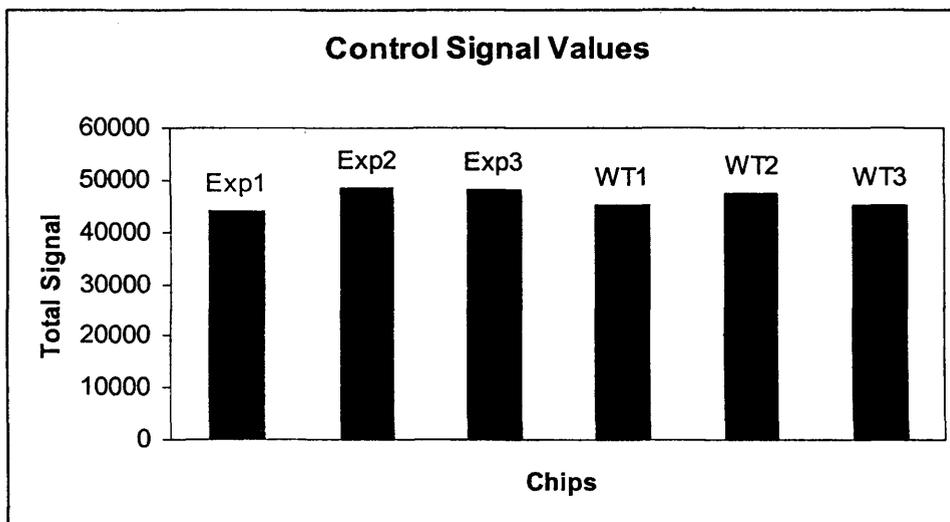
The original objective were to investigate the consequences of Stat3 depletion in the mammary gland via a scaled approach in which Affymetrix chip arrays could identify changes in mRNA levels from Stat3<sup>-/-</sup> and Stat3<sup>+/-</sup> compared to wild type Stat3 (Stat3<sup>+/+</sup>). Both wild and type and heterozygous Stat3 mammary epithelial cell cultures had a lifespan of 6-7 days before commit apoptosis. I found that sufficient RNA for performing chip hybridization without RNA amplification was available after 5 days of culturing in 35mm Petri dishes. Although RNA was extracted successfully from Stat3<sup>+/-</sup> and Stat3<sup>+/+</sup> at this time point, Stat3<sup>-/-</sup> epithelial cells exhibited a failure to reach adequate cell quantities before committing apoptosis (no picture has been taken). Therefore, I was unable to extract adequate amounts of RNA and perform chip hybridization. RNA amplification for Stat3<sup>-/-</sup> samples was not undertaken due mainly to lack of adequate time before completing this PhD.

### 5.2.2.3 Affymetrix chips and Data Processing.

Following *in vitro* RNA extraction and RNA processing as described in 2.7, labeled fragmented oligonucleotides hybridized to 6 U74v2 Affymetrix chips (3 experimentals: Stat3<sup>-/-</sup> and 3 controls: Stat3<sup>+/+</sup>) and analyzed with Affymetrix Gene Chip Hybridization and Analysis system (University of Wales College of Medicine).

Chip consistency of errors was investigated through summation of Affymetrix's control gene raw data as described in Chapter 4.2.2. Figure 5.6 illustrates the comparison of control gene errors between the 6 Affymetrix chips. From this graph I can conclude that all experimental and wild type chips show similar levels of error signal intensities amongst them. Similarity of control error intensities indicate similar patterns of hybridization procedures for each chip, thereafter indicate that all of our chips can be used further for gene expression analysis without creating noise in the data.

Filtering out “Absent” calls performed as described in 4.2.3. Filtering reduced the data from 12808 to 5312 transcripts. As in 4.2.4 data normalized by logging raw values in the base of 2. A two tail Student t-test identified 98 transcripts as significant potential targets at a 1% risk level. Comparison of Stat3<sup>+/-</sup> mice to wild type Stat3 indicated 46 targets with a significantly induced expression (Figure 5.7) and 52 with significantly reduced expression (Figure 5.8). Contrary to the Smad4 microarray experiments the FDR is almost half (53 targets) than the significant altered identified targets indicating that extraction of RNA from primary mammary gland tissues can produced better quality of results probably due to the avoidance of adipose contamination.



**Figure 5.6.** Summation of signal intensities from control genes. Each bar represents a different chip. Experimental (Exp) and wild type (WT) chips have a similar level of error signal intensities between group chips as well as between groups.

Rank	GI Number	Name	Log Ratio (exp/control)	Fold Difference
<b>Induced Genes</b>				
1	5125856	Cav: caveolin, caveolae protein	1,173589	2,19
2	192705	Crabp2: cellular retinoic acid binding protein II	1,171748	1,72
3	1907082	hexokinase 2	1,143682	1,92
4	5777875	ubiquitin specific protease 49	1,125872	1,59
5	5489704	Pla2g12a: phospholipase A2, group XIIA	1,12541	1,90
6	6099674	Plekhh2: pleckstrinB (evectins) member 2 (40.87%)	1,105862	1,78
7	722340	Fgfr1: fibroblast growth factor receptor 1	1,104163	1,67
8	1027499	Zfx1a: zinc finger homeobox 1a	1,102586	1,36
9	6390211	unknown	1,088566	1,41
10	2605641	Rgs16: regulator of G-protein signaling 16	1,087794	1,55
11	577633	Bmpr1a: bone morphogenetic rec type 1A (94.66%)	1,085905	1,40
12	4103990	Spata6: spermatogenesis associated 6	1,083911	1,46
13	4586867	Ggps1: geranylgeranyl diphosphate synthase 1	1,08346	1,36
14	2516393	unknown	1,075543	1,38
15	3171941	Diap2: diaphanous homolog 2 (56.58%)	1,074847	1,25
16	454834	peroxisome membrane protein (PMP70)	1,066383	1,39
17	193563	Ggt1: glycoprot galactosyltransferase alpha (96.80)	1,064984	1,34
18	436576	Cpox: coproporphyrinogen oxidase	1,064174	1,32
19	5907418	unknown	1,05737	1,23
20	3493348	Ampd3	1,050648	1,19
21	46195799	Rgl2: ral guanine nucleotide dissociation stimulator,-like 2	1,046892	1,21
22	3750037	unknown	1,046159	1,26
23	4218069	Tbpl1: TATA box binding protein-like 1	1,045515	1,15
24	5490939	unknown	1,042793	1,14
25	2062606	Kif5b: kinesin family member 5B	1,04007	1,22
26	5473499	unknown	1,038824	1,25
27	3372490	Tulp3: tubby-like protein 3	1,038494	1,22
28	1151214	Lamb3: laminin, beta 3	1,03667	1,20
29	198776	Ldh1: lactate dehydrogenase 1, A chain (95.73%)	1,03454	1,30
30	2959869	Syng2: synaptogyrin 2	1,033133	1,19
31	4602583	unknown	1,032163	1,18
32	5477102	Atp6v0b: ATPase, H+ transporting, V0 subunit B	1,031113	1,21
33	2952521	Golga4: golgi autoantigen golgin subtype a4	1,030251	1,16
34	5488570	RB4B_HUMAN Ras-related prot Rab4B(85.23%)	1,029713	1,15
35	5493810	unknown	1,029672	1,15
36	3805948	ST4 oncofetal trophoblast glycoprotein gene	1,027733	1,13
37	5469572	Ssr3: signal sequence receptor, gamma	1,026454	1,19
38	5475602	phosphopyruvate hydratase (94.06%)	1,025939	1,21
39	6095958	unknown	1,025113	1,14
40	5488417	unknown	1,024496	1,13
41	1771285	Hdac1: histone deacetylase 1	1,024239	1,13
42	1405932	Pkm2: pyruvate kinase M2 isozyme	1,022168	1,19
43	6467220	type I peroxiredoxin	1,021917	1,18
44	5931552	Morf4l2: mortality factor 4 like 2 (73.66%)	1,021106	1,15
45	50986	Itgb1: integrin beta 1 (fibronectin receptor beta)	1,011812	1,09
46	2645736	Ubl1: ubiquitin-homology domain protein	1,0067	1,04

**Figure. 5.7.** List of significant up regulated gene transcripts in the Stat3<sup>+/-</sup> mice. Ranked is based on the log ratio (Stat3<sup>+/-</sup>/wild types). Raw Fold changes are also included. Percentages in brackets (wherever applicable) represents protein similarity to mouse.

Rank	GI Number	Name	Log Ratio (exp/control)	Fold Difference
<b>Repressed Genes</b>				
1	3108056	Cipp: channel-interacting PDZ domain protein	0,784376	0,50
2	6375851	unknown	0,812498	0,54
3	2074486	Ncoa6: nuclear receptor coactivator 6	0,83668	0,59
4	3411010	Pparbp: peroxisome proliferator activated receptor binding protein	0,872563	0,53
5	4729959	Uxs1: UDP-glucuronate decarboxylase 1 (26.84%)	0,897974	0,77
6	2857305	unknown	0,904367	0,67
7	4616082	heterogeneous nuclear ribonucleoprotein H1 (62.12%)	0,912549	0,67
8	5469973	unknown	0,915006	0,73
9	6096519	S12207 hypothetical protein (68.57%)	0,916512	0,68
10	4663775	unknown	0,918826	0,74
11	533330	Top2b: topoisomerase (DNA) II beta	0,920983	0,70
12	5477783	unknown	0,923339	0,67
13	2125862	Sfrs3: splicing factor, arginine/serine-rich 3 (SRp20)	0,926954	0,68
14	2893204	unknown	0,927234	0,73
15	2687851	Zhx1: zinc fingers and homeoboxes protein 1(99.54%)	0,929368	0,76
16	1903415	Eif4g2: eukaryotic translation initiation factor4, gamma 2	0,929711	0,73
17	404930	Snta1: syntrophin, acidic 1	0,930577	0,68
18	3733840	Hnrpa1: heterogeneous nuclear ribonucleoprotein A1 (79.09%)	0,934669	0,76
19	5470226	Fem1a: feminization 1 homolog a (36.40%)	0,940826	0,77
20	5871950	unknown	0,941204	0,78
21	1871224	Rbl1: retinoblastoma-like 1 (p107)	0,941338	0,78
22	5496771	unknown	0,941762	0,78
23	5497628	unknown	0,941794	0,75
24	1098540	Osp94: osmotic stress protein (61.97%)	0,941949	0,76
25	2755261	unknown	0,942182	0,74
26	6099890	unknown	0,945984	0,77
27	54158	Sp1 gene (3' end)	0,946736	0,77
28	1545956	Ebp: phenylalkylamine Ca <sup>2+</sup> antagonist (emopamil) binding protein	0,951864	0,79
29	6095959	unknown	0,952589	0,82
30	4092839	Ptprj: protein tyrosine phosphatase, receptor type, J	0,953031	0,83
31	6340759	unknown	0,958507	0,79
32	191603	Chrnbl: cholinergic receptor, nicotinic, beta polypeptide 1 (99.80%)	0,959579	0,79
33	5907081	unknown	0,959798	0,84
34	5476816	unknown	0,960562	0,81
35	6100589	EGL nine homolog 2	0,960622	0,79
36	1546828	Idh3g: isocitrate dehydrogenase 3 (NAD <sup>+</sup> ), gamma	0,961855	0,80
37	6100992	unknown	0,963955	0,82
38	199762	Mod1: malic enzyme	0,964275	0,81
39	6096493	unknown	0,965816	0,85
40	4249732	low density lipoprotein B	0,966350	0,84
41	6097084	unknown	0,969016	0,83
42	5492751	unknown	0,969466	0,85
43	4262394	Cldn3: claudin 3	0,971961	0,81
44	2506077	Hiat1: hippocampus abundant gene transcript 1	0,974607	0,88
45	1550786	Dgcr2: DiGeorge syndrome critical region gene 2 (99.64%)	0,976625	0,89
46	471977	Ppp1cb: protein phosphatase 1, catalytic subunit, beta isoform (99.69%)	0,976997	0,86
47	2719590	unknown	0,977711	0,86
48	5471741	Jak1: Janus kinase 1	0,979612	0,88
49	5492041	Trpc4ap: transient receptor potential cation channel, subfamily C, member 4 associated protein	0,982929	0,90
50	5489559	unknown	0,983669	0,93
51	5476173	Snx6: sorting nexin 6 (66.67%)	0,986389	0,91
52	6098087	unknown	0,988999	0,93

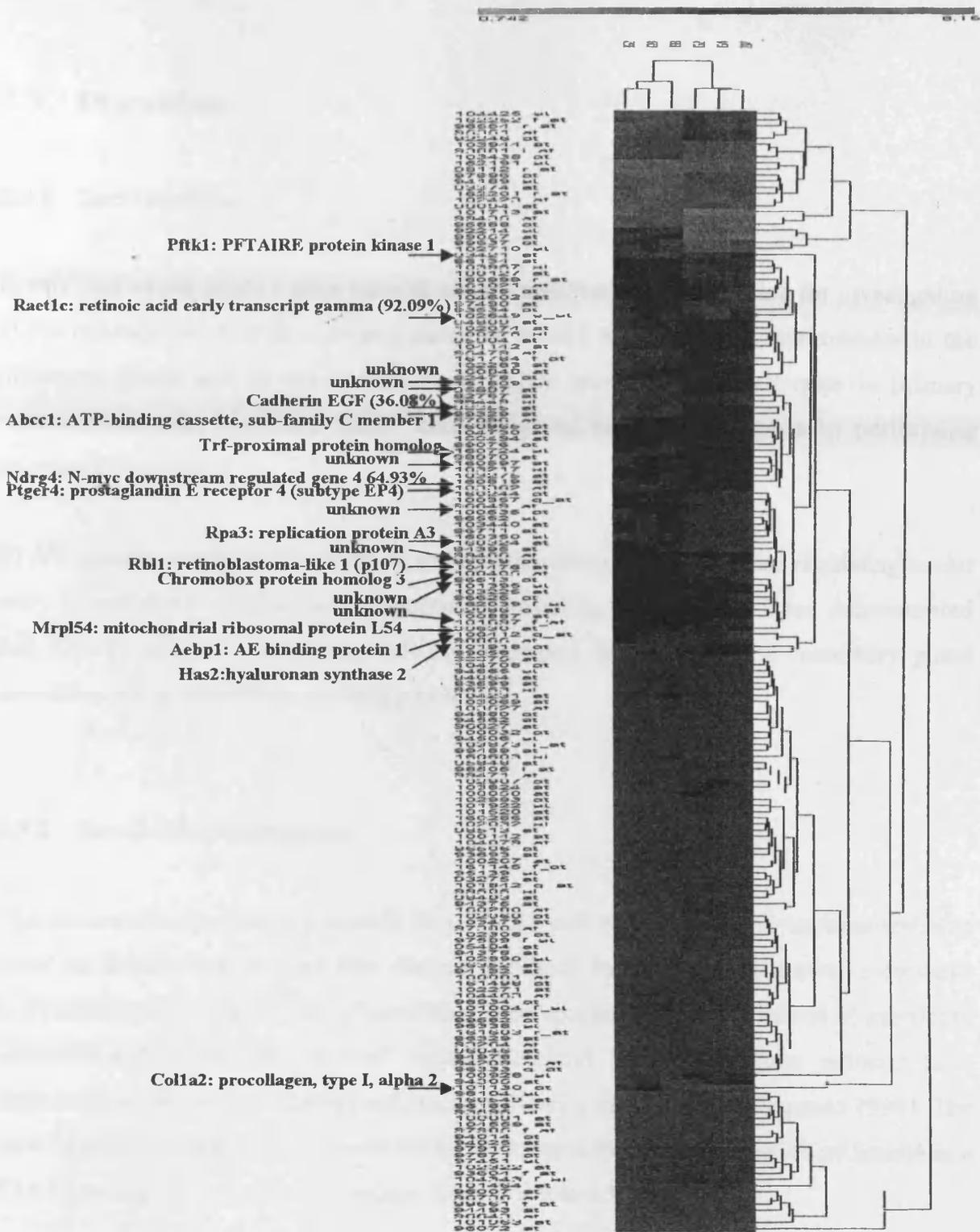
**Figure 5.8.** List of significant down regulated gene transcripts in the Stat3<sup>+/-</sup> mice. Ranked is based on the log ratio (Stat3<sup>+/-</sup>/wild type) Raw Fold changes are also included. Percentages in brackets (wherever applicable) represents protein similarity to mouse.

#### 5.2.2.4 Data Clustering.

Another way of visualizing alterations in gene expression profiles is via data clustering. The basic concept in clustering is to identify and group co-regulated and functionally related genes together and then try to correlate the observations to biology. There are several methods of clustering like Hierarchical clustering (Eisen 1998), K-means (Tavazoie 1999) and self organizing maps (SOMs) (Tamayo 1999). For the purpose of this analysis I have chosen to cluster my data in a Hierarchical mode.

Hierarchical clustering is a statistical method for identifying relatively homogeneous clusters. The hierarchical clustering algorithm either iteratively joins the two closest clusters starting from single clusters (bottom up approach) or iteratively partitions clusters starting from the complete set (top down approach). After each step, a new distance matrix between the newly formed clusters and the other clusters is recalculated. The complete linkage, average linkage and single linkage methods use maximum, average and minimum distances between the members of the two clusters respectively.

Average Hierarchical data clustering performed between 3 Stat3<sup>+/-</sup> and 3 wild type Stat3 samples. In average linkage, the distance between two clusters is considered to be equal to the average distance from any member of one cluster to any member of the other cluster. The hierarchical clustering can be represented as a tree, or a dendogram. Branch lengths represent the degree of similarity between the genes. Average linkage analysis at a 1% cut off produced the histogram illustrated in Figure 5.9.



**Figure 5.9.** Hierarchical clustering of gene expression matrices. The image shows an average linkage clustering of 99 genes constructed between Stat3<sup>+/-</sup> (352, 188, 198) and Stat3 wild types (162, 164, 237). The color image on the left shows the numerical values encoded by color according to the method introduced by the EASI software. Red is used to represent up-regulated values and green down-regulated.

## 5.3 Discussion.

### 5.3.1 Introduction.

In this part of the thesis I have utilized conditional Stat3 knockout mice for investigating a) the consequences of an over expression of Smad4 in a Stat3 null environment in the mammary gland and b) the alteration of mRNA levels of various targets in primary cultured mammary epithelial Stat3<sup>+/-</sup> cells compared to Stat3 wild types by performing microarray analysis.

STAT proteins comprise a family of transcription factors important for regulating a vast array of cellular events including apoptosis (Bromberg 2002). It has been demonstrated that STAT3 plays an important role in promoting apoptosis during mammary gland involution by up regulating *p21* and *p53* levels.

### 5.3.2 Smad-Stat Interactions.

The decision to over express Smad4 in a STAT3 null mammary gland environment was based on the fact that western blot analysis in Smad4 transgenic mice showed a decrease in STAT3 levels at the 3<sup>rd</sup> day of involution accompanied by an acceleration of mammary epithelial cell death. By contrast, mammary gland STAT3 knockout animals have demonstrated delayed involution and decreased levels of apoptosis (Chapman 1999). The investigation I undertook was based on the hypothesis that over expression of Smad4 in a STAT3 null environment could restore STAT3's delayed involution.

Following quantification of H&E slides between Blg-Cre-Smad4-Blg-Cre-Stat3<sup>fl/-</sup> and Blg-Cre- Stat3<sup>fl/-</sup> at day 3 involution, no differences were observed in their mammary epithelial cell contribution. TUNEL investigation for the apoptotic index of these mice at the same day agreed with the above result and showed no significant differences.

Although, cooperation between Smads and STATs has been demonstrated in a number of occasions (Ulloa 1999, Nakashima 1999) and over expression of Smad4 decreased STAT3 levels at day 3 involuting mice, Smad4 over expression in STAT3 null mammary glands did not restore STAT3's phenotype. Further experimentation is necessary for identifying relative levels of various cellular co-activators and transcription factors that may govern this net effect.

### 5.3.3 Microarray Analysis of Stat3<sup>+/-</sup> Primary Mammary Cell Cultures.

In this part of the project I decided to undertake a microarray analysis by utilizing mammary STAT3 knockout animals. My aim was to investigate mRNA alterations of various molecules by comparing Stat3<sup>-/-</sup> and Stat3<sup>+/-</sup> to wild type mammary glands.

In order to overcome problems such as fat contamination, mammary glands were subjected to primary cell culture. Although adequate amounts of RNA were possible to extract from both Stat3<sup>+/-</sup> and Stat3 wild types that was not the case for Stat3<sup>-/-</sup>. Mammary epithelial cells of this genotype demonstrated inability to create clusters and remain viable for more than 2-3 days. Inability of Stat3<sup>-/-</sup> cells to reach adequate confluency in this time resulted in extraction of low amounts of RNA which was insufficient for performing *in vitro* transcription and labeling for Affymetrix hybridization. The option of performing RNA amplification was abandoned for fear that amplification may create differences in the overall experiment when compared to un-amplified samples as well as due to time constraints for the completion of this thesis.

Following data filtering, normalization and statistical analysis 98 transcripts were identified as significant at a 1% risk interval. The number of significant targets identified at the 1% risk interval was almost double than the False Discover Rate (FDR). This can be attributed to lack of mammary cell contamination from adipocytes that is achieved under culture conditions. Therefore, I concluded that primary mammary epithelial cell

cultures serves microarray experiments better than utilizing *in vivo* (Chapter 4) glands and should be seriously taken into account in the experimental design of investigators.

All significant differentially expressed transcripts are presented in Figure 5.7 and 5.8. Molecules such as Jak1, bone morphogenic protein receptor type 1a (BMP1a), fibroblast growth factor receptor 1 (FGFr1), phospholipase A2 (Pla2g12a), and Caveolin (Cav) show altered mRNA expression levels and have been linked to STAT signaling in the past.

Interaction between TGF- $\beta$  and STATs is illustrated in this thesis once more in the level of Bone morphogenic protein receptor type 1a (BMP1a) mRNA expression. BMP1a showed an induction in its mRNA in the Stat3<sup>+/-</sup> mice compared to wild types. Kawamura et al (2000) showed that BMP2 caused cell-cycle arrest and subsequent apoptosis of myeloma cells by up regulating levels of *p21<sup>cip1/waf1</sup>* and *p27<sup>kip1</sup>*. In the same paper it was demonstrated that BMP2 treatment caused the inactivation of STAT3. Taken together with the above I hypothesize that over expression of BMP1a may results in activation of the BMP signaling pathway and induction of deregulation of cell development as observed in the Stat3<sup>-/-</sup> mammary epithelial cells.

The same microarray data was analysed by means of Hierarchical clustering. Hielustering is producing some areas where there seems to be alteration of mRNA expression between Stat3<sup>+/-</sup> and wild types with most important been the Retinoblastoma-like protein (*p107*). Hierarchical clustering indicated a down regulation of *p107* in the Stat3<sup>+/-</sup> mice compared to wild types. P107 together with Retinoblatoma 1 and retinoblastoma-like 2 (p130) play a pivotal role in regulating eukaryotic cell cycle progression, apoptosis, and terminal differentiation and inactivation of these molecules has been associated to abrogation of the cell cycle restriction point, leading to increased cell proliferation and apoptosis, and predisposing to cancer (MacPherson 2003, MacPherson 2004).

Affymetrix chip analysis and clustering are producing indications, facilitating a biologist to visualize alterations in gene expression profiles. As in Chapter 4, targets should be

checked through laboratory experiments both in the RNA and protein level. Unfortunately, due to time constraints no further experimentation for confirmation of some of the above targets was undertaken. However, this analysis can provide valuable information for other investigators working on the STAT-signaling pathway and also indicate the usefulness of primary mammary cell cultures compared to in vivo RNA experiments.

## Chapter 6-Conclusion.

### 6.1 Introduction.

Apoptosis is an important and well-documented form of cell death that occurs in a variety of physiological and pathological conditions. This process has been known to play a key role in autoimmune diseases, carcinogenesis, cancer progression, and killing of cancer cells induced by chemotherapeutic drugs (Friesen 1996). Apoptosis represents a physiological event that is essential for homeostasis, and inappropriate apoptosis seems to be involved in many diseases (Carson 1993). Consequently, since control of apoptosis has been proposed as an important target for therapeutic intervention, the elucidation of the molecular pathways and mechanisms underlying this process is of primary interest.

Understanding of the molecular pathways responsible for controlling cell division, differentiation and apoptosis has been greatly facilitated by the use of transgenic animal technology. Transgenic technology offers the possibility of creating targeted gene mutations or over expression in genes of choice via germline manipulations. In particular, the generation of mice with targeted mutations in genes encoding proteins of interest has proved to be a useful way of elucidating the function of these gene products *in vivo*. Two of the most common uses of transgenic mice have been for (1) studies of tissue-specific and developmental stage-specific gene regulation and (2) for experiments of the phenotypic effects of transgene expression. Transgenic mouse models have provided some insight into the complex events contributing to cellular deregulation and the loss of growth control that can lead to tumorigenesis. They have also helped elucidate crucial roles for genes in development and normal tissue homeostasis.

Analysis of a conditional Smad4 transgenic mouse and how over expression of Smad4 influences apoptosis in mammary epithelial cells during involution is the primary goal of this thesis. Based on results in the protein and mRNA level via western blots and microarrays respectively this project was expanded in the analysis of conditional knockout animals such as Stat3 and mbd2.

Mammary gland was chosen as the tissue of choice for the analysis of the above named transgenic and knockout animals. Mammary gland possesses the unique phenomenon of repetitive cycling from full lactation to involution and back and thus elects it as an ideal model to study highly controlled *in vivo* cellular apoptosis.

## 6.2. Smads and Apoptosis.

Smad4 is the central mediator of the TGF- $\beta$  superfamily signaling and the initial idea of addressing its role in mammary gland epithelial apoptosis (involution) was based on published data indicating the following; all TGF- $\beta$  isoforms are expressed during all stages of mammary gland development except lactation (Robinson 1991). TGF- $\beta$  expression plays many important roles in mammary gland physiology including apoptosis during involution (Daniel 1996, Joseph 1999, Nguyen 2000). Furthermore, Smads have been shown in a number of different cellular environments to be capable of initiating apoptosis.

Smad4, the main molecule in question in this thesis, has been characterised as a tumor suppressor gene due to its frequent inactivation/mutation in pancreatic and colorectal cancers. Smad4 expression in Smad4-defective tumor cell lines has been shown to restore TGF- $\beta$  signaling and induce cell cycle arrest and apoptosis (Dai 1999a). Schwarte-Waldhoff (1999) demonstrated the antiangiogenic role of Smad4 in tumor cell lines whilst Ramachandra (2002) demonstrated the role of Smad4 in the induction of apoptosis in a process termed anoikis. Conditional overexpression of Smad4 in the testis demonstrated amongst other apoptosis of germ cells and spermatogenic arrest indicating once more its role in apoptosis (Narula 2002). Finally, only recently Li (2004), demonstrated that lack of Smad4 in mouse mammary glands gradually induced cell proliferation, alveolar hyperplasia and transdifferentiation of mammary epithelial cells into squamous epithelial cells.

### 6.2.1 Transgenic Analysis of Smad4.

Over expression of Smad4 has been directed specifically in mammary epithelial cells during involution. Protein induction, specificity of tissue and time duration has been achieved by the utilization of the ovine beta lactoglobulin promoter (BLG) (Whitelaw 1992, Farini 1995, Clark 1992).

The initial analysis was to identify whether there is an altered epithelial cell contribution in mammary gland composition during involution between Smad4 transgenic and wild type mice. For this purpose I generated H&E slides at Day 10 of lactation and Day 2, 3 and 6 of involution. Following quantification I identified a significant increase in epithelial cell death at both day 2 and 3 in the Smad4 transgenics compared to wild types. TUNEL as well as quantification of apoptotic cells revealed that Smad4 transgenics are significantly inducing apoptosis at day 2 and 3 of involution. This result comes into agreement with the altered phenotype observed in the same time points.

In an effort to elucidate which molecular mechanism(s) may facilitate this apoptosis I examined a number of known mediators of mammary gland development as well as some apoptosis regulatory factors. Western blot analysis for Stat3 indicated a significant up-regulation at day 2 involution in the Smad4 transgenic mice while at day 3 Stat3 levels were significantly down regulated. This result supports the Stat-Smad cooperation which has also been illustrated by others (Dai 1999, Philip 1996, Nakashima 1999). Based on these findings, I hypothesize that Stat3 gets induced by Smad4 at day 2 involution, a time point where there is a great remodelling of the gland. Stat3 facilitates cell death and remodelling and demonstrates its tumor suppressor activity. However, at day 3 involution Stat3 levels are down regulated in Smad4 transgenics compared to wild types. This changing mode of Stat3 can be attributed to the survival factors influencing epithelial cells probably by generalised gland stimuli that remain to be identified.

Stat5a is a protein responsible for regulation of milk protein gene levels. Furthermore, it has been shown that Stat5a promotes cell survival and its loss results in increased apoptosis during mammary gland involution. Thereafter, it comes to no surprise that

levels of Stat5a are induced in both genotypes at day 10 lactation compared to involution. During lactation epithelial cells are fully functional in producing milk and Stat5a is a protein that facilitates this process. However, at day 2 and 3 of involution Stat5a levels are significantly induced in the Smad4 mice compared to wild types, a result which supports again Smad-Stat cooperation and possibly can be explained as a self defence mechanism of epithelial cells to survive.

Western blots indicated  $p27^{kip1}$  as the cyclin-dependant kinase responsible for accelerated cell death during day 2 and 3 of involution in the Smad4 transgenic animals compared to controls. Levels of  $p27^{kip1}$  were increased during this time points in the transgenics animals indicating that increased apoptosis influenced by Smad4 is a retinoblastoma dependent pathway. It was surprising to see that  $p21$  levels were absent both in the protein and mRNA levels. Finally, *Bax*, an important mammary apoptotic signal was investigated via western blotting. *Bax* levels were present throughout involution without indicating significant changes in its expression between Smad4 and wild type animals except of day 3 of involution.

I have performed primary mammary cell cultures in order to investigate a) the consequence of Smad4 over expression in epithelial cell cycle and apoptosis and b) the influences that TGF- $\beta$  administration can have in a Smad4 over expressing environment *in vitro*. Primary cell cultures performed with glands taken from day 15-17 old pregnant mice. Unfortunately, I discovered that the BLG-Smad4 construct expression in the mRNA level was absent.

### 6.2.2 Future work.

Future work with this model can be centre on a biochemical analysis of Smad proteins (R-Smads 1, 2 and phospho-2 and the I-Smads 6 and Smad7). Smad4 has shown that can play an important role in inducing mammary epithelial cell apoptosis. The detailed investigation of the Smad-Stat interplay appears to be of great importance. Electrophoretic Mobility Shift Assays (EMSA) should be performed in an effort to understand the binding properties and the level of collaboration between the Smad4/3 and Stat3 in different developmental stages of the mammary gland. Furthermore, a

very interesting experiment will be to “boost” the whole TGF- $\beta$  signaling pathway by crossing/passing the BLG-Smad4 construct to conditionally transgenic mice for TGF- $\beta$ . Another interesting experiment will be through the facilitation of siRNA technology in primary mammary epithelial cell cultures. In this way Smad4 can be knocked out and a series of analysis in both RNA and protein level can be undertaken.

### 6.3 Microarray analysis of Smad4.

In an attempt to investigate levels of molecules known to interact with the Smads and furthermore identify new targets of the Smad signaling pathway I performed a microarray analysis from day 3 involuting mammary glands. Day 3 was chosen as the time point of choice based on the phenotypic and molecular changes observed in my initial analysis. Following data normalisation and statistical analysis, Vitamin D receptor (VDR) was indicated as a molecule that gets induced by Smad4 over expression. Levels of VDR were analysed both in the mRNA and protein level by semi-quantitative RT-PCR and western blot analysis respectively at day 3 of involution. I was able to confirm Affymetrix’s result by identifying an up-regulation of VDR levels at day 3 involution in the Smad4 transgenics compared to controls at both RNA and protein levels.

Cooperation between VDR and Smads comes to no surprise as this interaction has been illustrated by others (Yanagisawa 1999, Yanagi 1999). More specifically, Smad3 acts as a co-activator of VDR and positively regulates its pathway which is characterised by growth inhibition of cancer cells as well as inhibition of breast tumorigenesis (Issa 1998).

### **6.3.1 Future Work.**

It will be of great importance if microarray experiments could be performed in earlier time points e.g. day 1 or 2 of involution. This will create a better understanding about the influence of Smad4 during involution. Experimentation should also be undertaken in elucidating the detailed mechanism of VDR activation by Smad4. Issa (1998) demonstrated a VDR induction by Smad3. EMSA can identify whether Smad3 or Smad4 binding can induce VDR expression.

### **6.3.2 The MBD2 mouse.**

Raw Affymetrix data indicated a down regulation of MBD2 gene expression at day 3 of involution in the Smad4 transgenic mice. I utilised conditional knockout MBD2 mice in an attempt to investigate the consequences of MBD2 depletion in the mammary gland during its development and involution. No phenotypic differences were identified in any case between knockouts and wild types. Despite the lack of phenotypic differences between MBD2 knockouts and wild types it would be interesting to analyse the consequences of an MBD2 over expressing mouse due to the fact that this molecule has been associated to various cancers (Magdinier and Wolffe 2001, Sansom 2003).

## **6.4 Smad4-Stat3 Interactions at Day 3 Involution.**

In Chapter 3, I demonstrated a down regulation of Stat3 in the Smad4 transgenic mice at the 3<sup>rd</sup> day of involution. It has been shown by others that Smads and Stats are two cytokines that exert antagonistic effects on fibroblasts (Ghosh 2001) and that Stat1 induces Smad7 (I-Smad) up-regulation. Synergistic signaling between Smads and Stats is indirect facilitated through the p300 co-activator (Ulloa 1999). Furthermore, Stat3 knockout mice in the mammary gland have shown a delay in involution (Chapman 1999).

Based on the above, it seemed appropriate to “pass” the BLG-Smad4 construct in mice deficient for Stat3 in the mammary gland. My hypothesis was that over expression of Smad4 could rescue Stat3’s delayed involution. H&E slides were generated and fat contribution between Blg-Cre-Smad4-Blg-Cre Stat3<sup>fl/-</sup> and Blg-Cre Stat3<sup>fl/-</sup> showed no differences. Furthermore, apoptotic levels were also counted between these two genotypes and revealed no differences between the two genotypes.

I can conclude from the above result that over expression of Smad4 in a Stat3 null environment cannot restore physiological levels of apoptosis and rescue Stat3’s delayed involution. Further experimentation is necessary for identifying relative levels of various cellular co-activators and transcription factors that may govern this net effect. Of particular interest will be the performance of an Electrophoretic Mobility Shift Assay (EMSA), investigating the binding of Smad3/4 and Stat3 on the p300 co-activator. Binding levels of these two molecules will facilitate the elucidation of the overall interaction and provide valuable information on how over expression of Smad4 can influence Stat3’s physiological role in the mammary gland.

## 6.5 Stat3 Microarrays.

A scaled microarray approach was undertaken in an attempt to identify new and check known targets influenced by Stat3 depletion in the mammary gland. Contrary to Smad4’s microarray experiments I performed primary mammary epithelial cell cultures in an attempt to avoid fat contamination. Stat3<sup>-/-</sup> mammary epithelial cells exhibited inability to cluster and to reach adequate confluency levels for RNA extraction.

Stat3<sup>+/-</sup> mammary epithelial cells were subjected to Affymetrix hybridization and compared to Stat3 wild types. Data analysis following filtering and normalization indicated a number of significant altered targets like the Bone Morphogenetic receptor 1a (BMP1a) and Jak1 whilst average hierarchical clustering identified amongst others a down regulation of the retinoblastoma-like 1 expression. For BMP1a, further experimentation identifying levels and distribution should be undertaken via western

blotting and immunohistochemistry. Based on the observation that Stat3<sup>-/-</sup> mice exhibited inability to colonise time lapse microscopy as well as cell cycle studies should be undertaken in order to investigate and provide valuable information about the modulators of this phenomenon. Further experimentation should be undertaken regarding the down regulation of the retinoblastoma-like 1 gene and a possible strategy would be to utilise tools such as siRNA.

## **6.6 Conclusions.**

In this thesis I have attempted to describe an analysis of the role of Smad4 in the murine mammary gland involution. This was undertaken by utilizing conditional transgenic animals for Smad4. Based on my data this investigation was directed towards Smad-Stat interactions as well as MBD2 by the utilization of conditional knockout animals. Characterisation of the animal models is a lengthy process and was never going to be completed within the realms of a 3-year Ph.D. However, in the near future I hope that information can be gained from all models which will be constructive in determining what role Smad4, Stat3 and MBD2 plays in mammary gland lactation and involution.

## References

- Adams, J.M. and S. Cory. 1991. Transgenic models of tumor development. *Science* **254**:1161-1167.
- Akaishi H, Takeda K, Kaisho T, Shinela R, Satomi S, Takeda J, Akira S. (1998). Defective IL-2-mediated IL-2 receptor  $\alpha$  chain expression in STAT3 deficient T lymphocytes. *Int. Immunol* **10**: 1747-1751.
- Akhurst RJ, Denryck R. (2001). TGF- $\beta$  signalling in cancer: a double edged sword. *Trends Cell Biol* **11**: S44-S51.
- Akiyoshi S, Inoue H, Hanai J, Kusanagi K, Nemoto N, Miyazono K, Kawabata M. (1999). c-Ski acts as a transcriptional co repressor in transforming growth factor- $\beta$  signalling through interaction with Smads. *J Biol Chem* **274**: 35269-35277.
- Anton M, Graham FL. (1995). Site-specific recombination mediated by an adenovirus vector expressing the Cre recombinase protein: a molecular switch for control of gene expression. *J Virol* **69**: 4600-4606.
- Ashcroft GS, Yang X, Glick AB, Weinstein M, Letterio JJ, Mizel DE, Anzano M, Greenwell-Wild T, Wahl SM, Deng C & Roberts AB. (1999). Mice lacking Smad3 show accelerated wound healing and an impaired local inflammatory response. *Nature Cell Biol* **1**: 260-266.
- Ashley DM, Kong FM, Bigner DD, Hale LP. (1998). Endogenous expression of transforming growth factor- $\beta$ 1 inhibits growth and tumorigenicity and enhances Fas-mediated apoptosis in murine high-grade glioma model. *Cancer Res* **58**: 302-309.
- Azuma M, Motegi K, Aota K, Yamashita T, Hoshida H, Sato M. (1999). TGF- $\beta$ 1 inhibits NF-KappaB activity through induction of IkappaB-alpha expression in human salivary gland cells: a possible mechanism of growth suppression by TGF- $\beta$ 1. *Exp Cell Res* **250**: 213-222.

Baker JC, Harland RM. (1996). A novel mesoderm inducer, *Madr2*, functions in the activin signal transduction pathway. *Genes Dev* **10**: 1880-1889.

Berger U, Wilson P, McClelland RA, Colston K, Haussler MR, Pike JW, Coombes RC. (1988). Immunocytochemical detection of 1,25-dihydroxyvitamin D3 receptor in normal human tissues. *J Clin Endo Metab* **67**: 607-613.

Bird A, Wolffe A. (1999). Methylation-induced repression-belts, braces and chromatin. *Cell* **99**: 451-454.

Bird A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* **1**: 6-21.

Bostrom, M.P. and P. Asnis. 1998. Transforming growth factor beta in fracture repair. *Clin Orthop* **S124-S131**.

Bottner M, Kriegstein K, Unsicker K. (2000). The transforming growth factor-betas: structure, signalling, and roles in nervous system development and functions. *J Neurochem* **75**: 2227-2240.

Bradford M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248-254.

Bromberg JF. (2002). Stat proteins and oncogenesis (review). *J Clin Invest* **109**: 1139-1142.

Bursch W, Oberhammer F, Jirtle RL, Askari M, Sedivy R, Grasl-Kraupp B, Purchio AF, Schulte-Hermann R. (1993). Transforming growth factor- $\beta$ 1 as a signal for induction of cell death by apoptosis. *Br J Cancer* **67**: 531-536.

Capecchi M. (1989). The new mouse genetics: altering the genome by gene targeting. *Trends Genet* **5**: 70-76.

Capecchi M. (1989a). Altering the genome by homologous recombination. *Science* **244**: 1288-1292.

Carlberg C. (1995). Mechanisms of nuclear signalling by vitamin D3. Interplay with retinoid and thyroid hormone signalling. *European J Biochem* **231**: 517-527.

Carson D, Ribeiri J. (1993). Apoptosis and disease. *Lancet* **341**: 1251-1254.

Catlett-Falcone R, Landowski TH, Oshiro MM, Turkson J, Levitzki A, Savino R, Ciliberto G, Moscinski L, Fernandez-Luna JL, Nunez G, Dalton WS, Jove R. (1999). Constitutively activation of Stat3 signaling confers resistance to apoptosis in human Y266 myeloma cells. *Immunity* **10**: 105-115.

Chapman RS, Lourenco PC, Tonner E, Flint DJ, Selbert S, Takeda K, Akira S, Clarke AR, Watson CJ. (1999). Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. *Genes Dev* **13**: 2604-2616.

Cheifetz S, Weatherbee JA, Tsang ML-S, Anderson JK, Mole JE, Lucas R, Massague J. (1987). Multiple type-beta transforming growth factors and their receptors. *J Cell Physiol Suppl.* **5**:43-7.

Chen RH, Derynck R. (1994). Homomeric interactions between type II transforming growth factor- $\beta$  receptors. *J Biol Chem* **269**: 22868-22874.

Chen X, Weisberg E, Fridmacher V, Watanabe M, Naco G. (1997). Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature* **389**: 85-89.

Chen YG, Liu F, Massague J. (1997). Mechanism of TGF beta receptor inhibition by FKBP12. *EMBO J* **16**: 3866-76.

Clark AJ, Cowper A, Wallace R, Wright G, Simons JP. (1992). Rescuing transgene expression by co-integration. *Biotechnology (NY)* **10**: 1450-1454.

Clark AJ. (1998). Gene expression in the mammary glands of transgenic animals. *Biochem Soc Symp* **63**: 133-140.

Colston K, Berger U, Wilson P, Hadcocks L, Naeem I, Earl H, Coombes R. (1988). Mammary gland 1,25-dihydroxyvitamin D3 receptor content during pregnancy and lactation. *Mol Cell Endoc* **60**: 15-22.

Colub TR, Slonim DK, Tamayo P. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**: 531-537.

Copp A. (1995). Death before birth: clues from gene knockouts and mutations. *Trends Genet* **11**: 87-93.

Dai JL, Schutte M, Bansal RK, Wilentz RE, Sugar AY, Kern SE. (1999). Transforming growth factor- $\beta$  responsiveness in DPC4/Smad4-null cancer cells. *Mol Carcinog* **26**: 37-43.

Dai JL, Bansal RK, Kern SE. (1999a). G1 cell cycle arrest and apoptosis induction by nuclear Smad4/Dpc4: phenotypes reversed by a tumorigenic mutation. *Proc Natl Acad Sci USA* **96**: 1427-1432.

Daniel CW, Robinson S, Silberstein GB. (1996). The role of TGF- $\beta$  in patterning and growth of the mammary ductal tree. *J Mamm Gland Biol Neoplasia* **1**: 331-341.

Dennler, S., S. Itoh, D. Vivien, P. ten Dijke, S. Huet, and J.M. Gauthier. (1998). Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J* **17**:3091-3100.

Derynck R, Akhurst RJ, Balmain A. (2001). TGF- $\beta$  signalling in tumor suppression and cancer progression. *Nat Genet* **29**: 117-129.

Derynck R, Gelbart WM, Harland RM, Heldin CH, Kem SE, Massague J, Melton DA, Miodzik M, Padgett RW, Roberts AB. (1996). Nomenclature: vertebrate mediators of TGF- $\beta$  family signals. *Cell* **87**: 173.

Derynck R, Zhang EY. (2003). Smad dependent and Smad independent pathways in TGF- $\beta$  family signalling. *Nature* **425**: 577-584.

Derynck R. (1998). SMAD proteins and mammalian anatomy [news; comment]. *Nature* **393**: 737-739.

Dong C, Li Z, Alvarez Jr R, Feng XH, Goldschmidt-Clermont PJ. (2000). Microtubule binding to Smads may regulate TGF- $\beta$  activity. *Mol Cell* **5**: 27-34.

Dudoit S, Yang YH, Callow MJ, Speed TP. (2002). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica* **12**: 111-139.

Dunker K, Krieglstein K. (2000). Targeted mutations of transforming growth factor-beta genes reveal important roles in mouse development and adult homeostasis. *Eur J Biochem* **267**: 6982-6988.

Ebisawa T, Fukuchi M, Murakami G, Chiba T, Tanaka K, Imamura T, Miyazono K. (2001). Smurf1 interacts with transforming growth factor- $\beta$  type I receptor through Smad7 and induces receptor degradation. *J Biol Chem* **276**: 12477-12480

Eisen M, Spellman P, Botstein D, Brown P. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95** (25): 14863-14867.

Farini E, Whitelaw CB. (1995). Ectopic expression of beta-lacyoglobulin transgenes. *Mol Gen Genet* **246**: 734-738.

Feng XH, Derynck R. (1996). Ligand-independent activation of transforming growth factor (TGF)  $\beta$  signalling pathways by heteromeric cytoplasmic domains of TGF- $\beta$  receptors. *J Biol Chem* **271**: 13123-13129.

Feng XH, Zhang Y, Wu RY, Derynck R. (1998). The tumour suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for Smad3 in TGF- $\beta$  induced transcriptional activation. *Genes Dev* **14**: 2153-2163.

Fladers KC, Kim ES, Roberts AB. (2001). Immunohistochemical expression of Smads1-6 in the 15-day gestation mouse embryo: signalling by BMPs and TGF- $\beta$ s. *Dev Dyn* **220**: 141-154.

Franzen P, ten Dijke P, Ichijo H, Yamashita H, Schulz P, Heldin CH, Miyazono K. (1993). Cloning of a TGF beta type I receptor that forms a heteromeric complex with the TGF beta type II receptor. *Cell* **75**: 681-692.

Franzen, P., C.H. Heldin, and K. Miyazono. (1995). The GS domain of the transforming growth factor-beta type I receptor is important in signal transduction. *Biochem Biophys Res Commun* **207**:682-689.

Friesen C, Herr I, Krammer P, Debatin K. (1996). Involvement of the CD95 (APO-1/Fas) receptor/ligand system in drug-induced apoptosis in leukaemia cells. *Nat Med* **2**: 574-577.

Fukada T, Ohtani T, Yoshida Y, Shirogane T, Nishida K, Nakajima K, Hibi M, Hirano T. (1998). STAT3 orchestrates contradictory signals in cytokine-induced G1 to S cell-cycle transition. *EMBO J* **17**: 6670-6677.

Fukuchi M, Imamura T, Chiba T, Ebisawa T, Kawabata M, Tanaka K, Miyazono K. (2001). Ligand-dependent degradation of Smad3 by a ubiquitin ligase complex of ROC1 and associated proteins. *Mol Biol Cell* **12**: 1431-1443.

Gallego MI, Binart N, Robinson GW, Okagaki R, Coschigano KT, Perry J, Kopchick JJ, Oka T, Kelly PA, Hennighausen L. (2001). Prolactin, growth hormone, and epidermal growth factor activate STAT5 in different and overlapping developmental effects. *Dev Biol* **229**: 163-175.

Georgi LL, Albert PS, Riddle DL. (1990). Daf-1, a *C. elegans* gene controlling dauer larva development, encodes a novel receptor protein kinase. *Cell* **61**: 635-645.

Ghosh AK, Yuan W, Mori Y, Chen Sj, Varga J. (2001). Antagonistic regulation of type I collagen gene expression by interferon-gamma and transforming growth factor-beta. Integration at the level of p300/CBP transcriptional coactivators. *J Biol Chem* **276**: 11041-11048.

Gobbi H, DuPont W, Simpson J. (1999). Relationship between TGF- $\beta$  type II receptor expression and breast cancer risk in women with epithelial hyperplasia lacking atypia. *J Natl Cancer Inst* **91**: 2096-2101.

Gordon J, Scangos G, Plotkin D, Barbosa J, Ruddle F. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc Natl Acad Sci USA* **77**: 7380-7384.

Gorska AE, Joseph R, Deryneck R, Moses HC, Serra R. (1998). Dominant-negative interference at the TGF- $\beta$  type II receptor in mammary gland epithelium results in alveolar hyperplasia and differentiation in virgin mice. *Cell Growth Differ* **3**: 229-238.

Gronroos E, Hellman U, Heldin C, Ericsson J. (2002). Control of Smad7 stability by competition between acetylation and ubiquitination. *Mol Cell* **10**: 483-493.

Gu H, Marth J, Orban P, Mossmann H, Rajewsky K. (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265**: 103-106.

Hahn, S.A., A.T. Hoque, C.A. Moskaluk, L.T. da Costa, M. Schutte, E. Rozenblum, A.B. Seymour, C.L. Weinstein, C.J. Yeo, R.H. Hruban, and S.E. Kern. (1996). Homozygous deletion map at 18q21.1 in pancreatic cancer. *Cancer Res* **56**:490-494.

Hansen R, Wijmenga P, Luo A, Stanek T, Canfield C, Gartler S. (1999). The Dnmt3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci USA* **96**:14412-14417.

Harris S, McClenaghnam M, Simons JP, Ali S, Clark AJ. (1991). Developmental regulation of the sheep beta-lactoglobulin gene in the mammary gland of transgenic mice. *Dev Genet* **12**: 299-307.

Hartemink AJ, Gifford DK, Jaakkola TS, Young RA. (2001). Using graphical models and genomic expression data to statistically validate models of genetic regulatory networks. *Pac Symp Biocomput* :422-33.

Hata A, Lagna G, Massague J, Hemmati-Brivanlou A. (1998). Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes Dev* **12**: 186-197.

Hata A, Lo RS, Wotton D, Lagna G, Massague J. (1997). Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4. *Nature* **388**: 82-87.

He W, Li A, Wang D, Han S, Zheng B, Goumans MJ, ten Dijke P, Wang XJ. (2002). Overexpression of Smad7 results in severe pathological alterations in multiple epithelial tissues. *EMBO* **21**: 2580-2590.

Heldin CH, Miyazono K, ten Dijke P. (1997). TGF- $\beta$  signalling from cell membrane to nucleus through Smad proteins. *Nature* **390**: 465-471.

Hendrich B, Bird A. (2000). Mammalian methyltransferases and methyl-CpG-binding domains: Proteins involved in DNA methylation. *Curr. Top. Microbiol: DNA methylation and cancer* **249**: 55-74.

Hendrich B, Bird A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18**: 6538-6547.

Henis YI, Moustakas A, Lin HY, Lodish HF. (1994). The types II and III transforming growth factor- $\beta$  receptors form homo-oligomers. *J Cell Biol* **126**: 139-154.

Hennighausen L, Robinson GW, Wagner KU, Liu X. (1997a). Developing a mammary gland is a stat affair. *J Mammary Gland Biol Neoplasia* **2**: 365-372.

Heyer J, Escalante-Alcalde D, Lia M, Boettinger E, Edelmann W, Stewart C, Kucherlapati R. (1999). Postgastrulation Smad2-deficient embryos show defects in embryo turning and anterior morphogenesis. *Proc Natl Acad Sci USA* **96**: 12595-12600.

Hoess R, Abremski K, Sternberg N. (1984). The nature of the interaction of the P1 recombinase Cre with the recombining site loxP. *Cold Spring Harb Symp Quant Biol* **49**: 761-768.

Hogan BL. (1996). Bone morphogenic proteins; multifunctional regulators of vertebrate development. *Genes Dev* **10**: 1580-1594.

Horiki M, Imamura T, Okamoto M, Hayashi M, Murai J, Myoui A, Ochi T, Miyazono K, Yoshikawa H, Tsumaki N. (2004). Smad6/Smurf1 over expression in cartilage delays chondrocyte hypertrophy and causes dwarfism with osteopenia. *J Cell Biol* **165**: 433-445.

Hosobuchi M, Stampfer MR. (1989). Effects of transforming growth factor  $\beta$  on growth of human mammary epithelial cells in culture. *In vitro Cell Dev Biol* **25**: 705-713.

Hua X, Liu X, Ansari DO, Lodish HF. (1998). Synergistic cooperation of TFE3 and smad proteins in TGF-beta-induced transcription of the plasminogen activator inhibitor-1 gene. *Genes Dev* **12**: 3084-3095.

Humphreys RC, Hennighausen L. (1999). Signal transducer and activator of transcription 5a influences mammary epithelial cell survival and tumorigenesis. *Cell Growth Differ* 10:685-94.

Huse M, Muir TW, Xu L, Chen YG, Kuriyan J, Massague J. (2001). The TGF- $\beta$  receptor activation process: an inhibitor- to substrate binding switch. *Mol.Cell* 8: 671-682.

Imamura T, Takase M, Nishihara A, Oeda E, Hanai J. (1997). Smad6 inhibits signalling by the TGF- $\beta$  superfamily. *Nature* 389: 622-626.

Issa I, Leong GM, Eisman JA. (1998). Molecular mechanism of vitamin D receptor action. *Inflammation Res* 47: 451-475.

Ito Y, Sarkar P, Mi Q, Wu N, Bringas P Jr, Liu Y, Reddy S, Maxson R, Deng C, Chai Y. (2001). Overexpression of Smad2 reveals its concerted action with Smad4 in regulating TGF-beta-mediated epidermal homeostasis. *Dev Biol* 1: 181-194.

Itoh S, Ericsson J, Nishikawa J, Heldin C, ten Dijke P. (2000). The transcriptional co-activator P/CAF potentiates TGF- $\beta$ /Smad signalling. *Nucleic Acids Res* 28: 4291-4298.

Itoh S, Landstrom M, Hermansson A, Itoh F, Heldin C, Heldin N, ten Dijke. (1998). Transforming growth factor  $\beta$ 1 induces nuclear export of inhibitory Smad7. *J Biol Chem* 273: 29195-29201.

Jerry DJ, Kuperwasser C, Downing SR, Pinkas J, He C, Dickinson E, Marconi S, Naber SP. (1998). Delayed involution of the mammary epithelium in BALB/c-p53null mice. *Oncogene* 17: 2305-2312.

Jhappan C, Geiser AG, Kordon EC, Bagheri D, Hennighausen L, Roberts AB, Smith GH, Merlino G. (1993). Targeting expression of a transforming growth factor- $\beta$ 1

transgene to the pregnant mammary gland inhibits alveolar development and lactation. *EMBO J* **12**: 1835-1845.

Jhappan C, Stahle C, Harkins RN, Fausto N, Smith GH, Merlino GT. (1990). TGF- $\alpha$  over expression in transgenic mice induces liver neoplasia and abnormal development of the mammary gland and pancreas. *Cell* **61**: 1137-1146.

Jones A, Baylin B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet* **6**:415-428.

Joseph H, Gorska AE, Sohn P, Moses HL, Serra R. (1999). Over expression of kinase-deficient transforming growth factor- $\beta$  type II receptor in mouse mammary stroma results in increased epithelial branching. *Mol Biol Cell* **10**: 1221-1234.

Kaplan MH, Daniel C, Schindler U, Grusby MJ. (1998). Stat proteins control lymphocyte proliferation by regulating p27Kip1 expression. *Mol Cell Biol* **18**: 1996-2003.

Kavsak P, Rasmussen R, Causing C, Bonni S, Zhu H, Thomsen G, Wrana J. (2000). Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF- $\beta$  receptor for degradation. *Mol Cell* **6**: 1365-1375.

Kawamura C, Kizaki M, Yamato K, Uchida H, Fukuchi Y, Hattori Y, Koseki T, Nishihara T, Ikeda Y. (2000). Bone morphogenic protein-2 induces apoptosis in human myeloma cells with modulation of STAT3. *Blood* **96** (6): 2005-2011.

Kim S, Dougherty ER, Chen Y. (2000). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Genomics* **67**: 201-209.

Kim, J.W., H.S. Kim, I.K. Kim, M.R. Kim, E.Y. Cho, H.K. Kim, J.M. Lee, and S.E. Namkoong. (1998). Transforming growth factor-beta 1 induces apoptosis through down-regulation of c-myc gene and overexpression of p27Kip1 protein in cervical carcinoma. *Gynecol Oncol* **69**:230-236.

Kuhn R, Schwenk F, Aguet M, Rajewsky K. (1995). Inducible gene targeting in mice. *Science* **269**: 1427-1429.

Kurisaki A, Kose S, Yoneda Y, Heldin CH, Moustakas A. (2001). Transforming growth factor- $\beta$  induces nuclear import of Smad3 in an importin- $\beta$ 1 and Ran dependant manner. *Mol Biol Cell* **12**: 1079-1091.

Kusanagi K, Kawabata M, Hiromu K M, and Miyazono K (2001).  $\alpha$ -Helix 2 in the Amino-terminal Mad Homology 1 Domain Is Responsible for Specific DNA Binding of Smad3. *J Biol Chem* **276**: 28155-28163.

Labbe E, Letamendia A, Attisano L. (2000). Association of Smads with lymphoid enhancer binding factor 1/T cell-specific factor mediates cooperative signalling by the transforming growth factor- $\beta$  and Wnt pathways. *Proc Natl Acad Sci USA* **97**: 8358-8363.

Labbe E, Silvestri C, Hoodless PA, Wrana JL, Attisano L. (1998). Smad2 and Smad3 positively and negatively regulate TGF- $\beta$ -dependant transcription through the forkhead DNA-binding protein FAST-2. *Mol Cell* **1**: 109-120.

Lagna G, Hata A, Hemmati-Brivanlou A, Massague J. (1996). Partnership between DPC4 and SMAD protein in TGF- $\beta$  signalling pathways. *Nature* **383**: 832-836.

Laiho M, De Caprio JA, Ludlow JW, Livingston DM, Massague J. (1990). Growth inhibition by TGF-beta linked to suppression of retinoblastoma protein phosphorylation. *Cell* **62**: 175-185.

Lawrence DA. (1996). Transforming growth factor-beta: a general review. *Eur Cytokine Netw* **7**: 363-374.

Lee S, Cho YS, Shim C, Kim J, Choi J, Oh S, Kim J, Zhang W, Lee J. (2001). Aberrant expression of Smad4 results in resistance against the growth-inhibitory

effect of transforming growth factor-beta in the SiHa human cervical carcinoma cell line. *Int J Cancer* **94**: 500-507.

Li W, Qiao W, Chen L, Xu X, Yang X, Li D, Li C, Brodie SG, Meguid MM, Hennighausen L and D CX. (2004). Squamous cell carcinoma and mammary abscess formation through squamous metaplasia in Smad4/Dpc4 conditional knockout mice. *Development* **130**: 6143-6153.

Li J, Lee GI, Van Doren SR, Walker JC. (2000). The FHA domain mediates phosphoprotein interactions. *J Cell Sci* **113**: 4143-4149.

Li M, Liu X, Robinson G, Bar-Peled U, Wagner KU, Young WS, Hennighausen L, Furth PA. (1997). Mammary-derived signals activate programmed cell death during the first stage of mammary gland involution. *Proc Natl Acad Sci USA* **94**: 3425-3430.

Liberati NT, Moniwa M, Borton AJ, Davie JR, Wang XF. (2001). An essential role for MAD homology domain 1 in the association of Smad3 with histone deacetylase activity. *J Biol Chem* **276**: 22595-22603.

Lin X, Liang M, Feng X. (2000). Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent degradation of Smad2 in TGF- $\beta$  signalling. *J Biol Chem* **275**: 36818-36822.

Liu F, Hata A, Baker JC, Doody J, Carcamo J. (1996). A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature* **381**: 620-623.

Liu F, Poupponnot C, Massague J. (1997). Dual role of the Smad4/DPC4 tumor suppressor in TGF beta-inducible transcriptional complexes. *Genes Dev* **11**: 3157-3167.

Liu X, Robinson GW, Hennighausen L. (1996). Activation of STAT5a and STAT5b by tyrosine phosphorylation is tightly linked to mammary gland differentiation. *Mol Endocrinol* **10**: 1496-1506.

Lo R, Massague J. (1999). Ubiquitin-dependent degradation of TGF- $\beta$ -activated Smad2. *Nat Cell Biol* 1: 472-478.

Long J, Wang G, Matsuura I, Liu F. (2004). Activation of Smad transcriptional activity by protein inhibitor of activated STAT3 (PIAS3). *Proc Natl Acad Sci USA* 101: 99-104.

Lu KX, Lodish HF. (1996). Signalling by chimeric erythropoietin-TGF- $\beta$  receptors: Homodimerisation of the cytoplasmic domain of the type I TGF- $\beta$  receptor and heterodimerisation with the type II receptor are both required for intracellular signal transduction. *EMBO J* 15: 4485-4496.

Luo KX, Lodish HF. (1997). Positive and negative regulation of type II TGF-beta receptor signal transduction by autophosphorylation on multiple serine residues. *EMBO J* 16: 1970-1981.

Luukko K, Ylikorkala A, Makela TP. (2001). Developmentally regulated expression of Smad3, Smad4, Smad6 and Smad7 involved in TGF- $\beta$  signalling. *Mech Dev* 101: 209-212.

Macias-Silva M, Abdollah S, Hoodless PA, Pirone R, Attisano L. (1996). MADR2 is a substrate of the TGF- $\beta$  receptor and its phosphorylation is required for nuclear accumulation and signalling. *Cell* 87: 1215-1222.

MacPherson D, Sage J, Crowley D, Trumpp A, Bronson RT, Jacks T. (2003). Conditional mutation of Rb causes cell cycle defects without apoptosis in the central nervous system. *Mol Cell Biol* 23: 1044-1053.

MacPherson D, Sage J, Kim T, Ho D, McLaughlin ME, Jacks T. (2004). Cell type-specific effects of Rb deletion in the murine retina. *Genes Dev* 18: 1681-1694.

Magdinier F, Wolffe A. (2001). Selective association of the methyl-CpG binding protein MBD2 with the silent p14/p16 locus in human neoplasia. *Proc Natl Acad Sci USA* **98**: 4990-4995.

Massague J, Wotton D. (2000). Transcriptional control by the TGF $\beta$ /Smad signalling system. *EMBO J* **19**: 1745-1754.

Massague J. (1990). The transforming growth factor- $\beta$  family. *Annu Rev Cell Biol* **6**: 507-641.

Massague J. (1998). TGF- $\beta$  signal transduction. *Annu Rev Biochem* **67**: 753-791.

Mathews LS, Vale WW. (1991). Expression cloning of an activin receptor, a predicted transmembrane serine kinase. *Cell* **65**: 973-982.

Mercier T, Chaumontet C, Gaillard-Sanchez I, Martel P, Heberden C. (1996). Calcitriol and lexicalcitol (KH1060) inhibit the growth of human breast adenocarcinoma cells by enhancing transforming growth factor- $\beta$  production. *Biochemical Pharmacology* **52**: 505-510.

Merto GR, Cella N, Hynes NE. (1997). Apoptosis is accompanied by changes in Bcl-2 and Bax expression, induced by loss of attachment, and inhibited by specific extracellular matrix proteins in mammary epithelial cells. *Cell Growth Differ* **8**:251-60.

Mezzetti G, Barbiroli B, Oka T. (1987). 1,25-Dihydroxycholecalciferol receptor regulation in hormonally induced differentiation of mouse mammary gland in culture. *Endocrinology* **120**: 2488-2493.

Motyl T, Grzelkowska K, Zimowska W, Skierski J, Wareski P, Ploszaj T, Trzeciak L. (1998). Expression of Bcl-2 and Bax in TGF- $\beta$ 1-induced apoptosis of L1210 leukemic cells. *Eur J Cell Biol* **75**: 367-374.

Mummery CL. (2001). Transforming growth factor beta and mouse development. *Microsc Res Tech* **52**: 374-386.

Muraoka RS, Lenferink AE, Simpson J, Brantley DM, Roebuck LR, Yakes FM, Arteaga CL. (2001). Cyclin-dependent kinase inhibitor p27(Kip1) is required for mouse mammary gland morphogenesis and function. *J Cell Biol* **153**: 917-932.

Myrphy MS. (1998). Growth factors and gastrointestinal tract. *Nutrition* **14**: 771-774

Nakashima K, Yanagisawa M, Arakawa H, Kimura N, Hisatsune T, Kawabata M, Miyazono K, Taga T. (1999). Synergistic signalling in fetal brain by STAT3-Smad1 complex bridged by p300. *Science* **284**: 479-482.

Narula A, Kilen S, Ma E, Kroeger J, Goldberg E, Woodruff T. (2002). Smad4 overexpression causes germ cell ablation and leydic cell hyperplasia in transgenic mice. *Am J Path* **161**: 1723-1734.

Nguyen AV, Pollard JW. (2000). Transforming growth factor  $\beta$ 3 induces cell death during the first stage of mammary gland involution. *Development* **127**: 3107-3118.

Nishihara A, Hanai JI, Okamoto N, Yanagisawa J, Kato S, Miyazono K, Kawabata M. (1998). Role of p300, a transcriptional coactivator, in signalling of TGF- $\beta$ . *Genes Cells* **3**: 613-623.

Nishita M, Hashimoto MK, Ogata S, Laurent MN, Ueno N, Shibuya H, Cho KWY. (2000). Interaction between Wnt and TGF- $\beta$  signalling pathways during formation of Spemann's organiser. *Nature* **403**: 781-785.

Oft M, Akhurst RJ, Balmain A. (2002). Elevated levels of activated Smad2 and H-ras control epithelial-mesenchymal transformation, tumor cell extravasation and metastasis. *Nat. Cell. Biol* **4**: 587-494.

Oft M, Peli J, Rudaz C, Schwarz H, Beug H, Reichmann E. (1996). TGF- $\beta$ 1 and Ha-Ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumour cells. *Genes Dev* **10**: 2462-2477.

Padgett RW, Wozney JM, Gelbart WM. (1993). Human BMP sequences can confer normal dorsal-ventral patterning in the Drosophila embryo. *Proc Natl Acad Sci USA* **90**: 2905-2909.

Pardali K, Kurisaki A, Moren A, ten Dijke P, Kardassis D, Moustakas A. (2000). Role of Smad proteins and transcription factor Sp1 in p21 (Waf1/Cip1) regulation by transforming growth factor-beta. *J Biol Chem* **275**: 29244-29256.

Patterson GI, Padgett RW. (2000). TGF- $\beta$  related pathways; roles in Caenorhabditis elegans development. *Trends Genet* **16**: 27-33.

Philip J, Burdon T, Watson C. (1996). Differential activation of Stat3 and Stat5 during mammary gland development. *FEBS lett.* **396**: 77-80.

Pierce DFJ, Johnson MD, Matsui Y, Robinson SD, Gold LI, Purchio AF, Daniel CW, Hogan BL, Moses HL. (1993). Inhibition of mammary duct development but not alveolar outgrowth during pregnancy in transgenic mice expressing active TGF- $\beta$ 1. *Genes Dev* **7**: 2308-2317.

Pierreux CE, Nicolas FJ, Hills CS. (2000). Transforming growth factor- $\beta$ -independent shuttling of Smad4 between the cytoplasm and nucleus. *Moll Cell Biol* **20**: 9041-9054.

Ramachandra M, Atencio I, Rahman A, Vaillancourt M, Zou A, Avanzini J, Wills K, Bookstein R, Shabram P. (2002). Restoration of Transforming Growth Factor  $\beta$  Signalling by functional expression of Smad4 Induces Anoikis. *Cancer Res* **62**: 6045-6051.

Randall R, Germain S, Inman G, Bates P, Hill C. (2002). Different Smad2 partners bind a common hydrophobic pocket in Smad2 via a defined proline-rich motif. *EMBO J* **21**: 145-156.

Reiss M, Barcellos-Hoff M. (1997). Transforming growth factor- $\beta$  in breast cancer: a working hypothesis. *Breast Cancer Res Treat* **45**: 81-95.

Riggins, G.J., S. Thiagalingam, E. Rozenblum, C.L. Weinstein, S.E. Kern, S.R. Hamilton, J.K. Willson, S.D. Markowitz, K.W. Kinzler, and B. Vogelstein. 1996. Mad-related genes in the human. *Nat Genet* **13**:347-349.

Robinson GW, Hennighausen L. (1997). Inhibins and activins regulate mammary epithelial cell differentiation through mesenchymal-epithelial interactions. *Development* **124**: 2701-2708.

Robinson SD, Silberstein GB, Roberts AB, Flanders KC, Daniel CW. (1991). Regulated expression and growth inhibitory effects of transforming growth factor- $\beta$  isoforms in mouse mammary gland development. *Development* **113**: 867-878.

Rocke DM, Durbin B. (2001). A model for measurement error for gene expression arrays. *J Comput Biol* **8**: 557-569.

Rotello RJ, Lieberman RC, Purchio AF, Gerschenson LE. (1991). Coordinated regulation of apoptosis and cell proliferation by transforming growth factor- $\beta$ 1 in cultured uterine epithelial cells. *Proc Nat Acad Sci USA* **88**: 3412-3415.

Sabramaniam N, Leong GM, Cock TA, Flanagan JL, Fong C, Eisman JA, Kouzmenko AP. (2001). Cross-talk between 1,25-Dihydroxyvitamin-D<sub>3</sub> and Transforming growth factor- $\beta$  signalling requires binding of VDR and Smad3 proteins to their cognate DNA recognition Elements. *J Biol Chem* **276**: 15741-15746.

Sakakura T. (1991). New aspects of stroma-parenchyma relations in mammary gland differentiation. *Int Rev Cytol* **125**: 165-202.

Saltis J. (1996). TGF-beta: receptors and cell cycle arrest. *Mol Cell Endocrinol* **116**: 227-232.

Sampath TK, Rashka KE, Doctor JS, Tucker RF, Hoffman FM. (1993). Drosophila transforming growth factor  $\beta$  induce endochondral bone formation in mammals. *Proc Natl Acad Sci USA* **90**: 6004-6008.

Sansom OJ, Berger J, Bishop SM, Hendrich B, Bird A, Clarke AR. (2003). Deficiency of MBD2 suppresses intestinal tumorigenesis. *Nat Genet* **34**: 145-147.

Sasaki A, Masuda Y, Ohta Y, Ikeda K, Watanabe K. (2001). Filamin associates with Smads and regulates transforming growth factor- $\beta$  signalling. *J Biol Chem* **276**: 17871-17877.

Sauer B, Henderson N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci USA* **85**: 5166-5170.

Saviozzi S, Iazzetti G, Caserta E, Guffanti A, Calogero RA. (2003). Microarray data analysis and mining. In *methods in Molecular Medicine: Molecular Diagnosis of Infectious Diseases*, Decker J, Reischl U (eds). Humana: Totowa, MA.

Schindler C, Darnell JEJ. (1995). Transcriptional responses to polypeptide ligands: the JAK-STAT pathway. *Annu Rev Biochem* **64**: 621-651.

Schorr K, Li M, Krajewski S, Reed JC, Furth PA. (1999). Bcl-2 gene family and related proteins in mammary gland involution and breast cancer. *J Mammary Gland Biol Neoplasia* **4**:153-64.

Schroeder MD, Rose-Hellekant TA, Sandgren EP, Schuler LA. (2001). Dysregulation of mammary Stats 1, 3 and 5 and PRL receptors by overexpression of TGF $\alpha$ . *Mol Cell Endocrinol* **175**:173-83.

Schwaller J, Koeffler HP, Niklaus G, Loetscher P, Nagel S, Fey MF. (1995). Post-transcriptional stabilization underlies p53-independent induction of p21<sup>WAF1/CIP1/SDI1</sup> in differentiating human leukemic cells. *J Clin Invest* **95**: 973-979.

Schwarte-Waldhoff I, Klein S, Blass-Kampmann S, Hintelmann A, Eilert C, Dreschers S, Kalthoff H, Hahn S, Schmiegel W. (1999). DPC4/SMAD4-mediated tumor suppression of colon carcinoma cells is associated with reduced urokinase expression. *Oncogene* **18**: 3152-3158.

Sekelsky JJ, Newfeld SJ, Raftery LA, Chartoff EH, Gelbart WM. (1995). Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in *Drosophila melanogaster*. *Genetics* **139**: 1347-1358.

Shen X, Hu PP, Liberati NT, Datto MB, Frederick JP, Wang XF. (1998). TGF- $\beta$ -induced phosphorylation of Smad3 regulates its interaction with co-activator p300/CREB-binding protein. *Mol Biol Cell* **9**: 3309-3319.

Sherr CJ, Roberts JM. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev* **13**:1501-12.

Shi Y, Hata A, Lo RS, Massague J, Pavletich NP. (1997). A structural basis for mutational inactivation of the tumour suppressor Smad4. *Nature* **388**: 87-93.

Shi Y, Massague J. (2003). Mechanisms of TGF- $\beta$  Signalling from cell membrane to the nucleus. *Cell* **113**: 685-700.

Shi Y, Wang YF, Jayaraman L. (1998). Crystal structure of a MH1 domain bound to DNA: insights on DNA binding in TGF- $\beta$  signalling. *Cell* **94**: 585-594.

Shi Y. (2001). Structural insights on Smad function in TGF- $\beta$  signalling. *Bioessays* **23**: 223-232.

Sirard C, de la Pompa JL, Elia A, Itie A, Mirtsos C, Cheung A, Hahn S, Wakeham A, Schwartz L, Kern SE, Rossant J, Mak TW. (1998). The tumor suppressor gene Smad4/Dpc4 is required for gastrulation and later for anterior development of the mouse embryo. *Genes Dev* **12**: 107-119.

Souchelnytskyi S, ten Dijke P, Miyazono K, Heldin CH. (1996). Phosphorylation of Ser165 in TGF-beta type I receptor modulates TGF-beta1-induced cellular responses. *EMBO J* **15**: 6231-6240.

Speed T. (2000) Microarray Data Analysis Group. "Always log spot intensities and ratios". <http://www.stat.berkeley.edu/users/terry/zarray/Html/log.html>

Sternberg N, Austin S, Hamilton D, Yarmolinsky. (1978). Analysis of bacteriophage P1 immunity by using lambda-P1 recombinants constructed *in vitro*. *Proc Natl Acad Sci USA* **75**: 5594-5598.

Sternberg N, Hamilton D. (1981). Bacteriophage P1 site-specific recombination. Recombination between loxP sites. *J Mol Biol* **150**: 467-486.

Strange R, Friis RR, Bemis LT, Geske FJ. (1995). Programmed cell death during mammary gland involution. *Methods Cell Biol* **46**: 355-368.

Suzuki C, Murakami G, Fukuchi M, Shimanuki T, Shikauchi Y, Imamura T, Miyazono K. (2002). Smurf1 regulates the inhibitory activity of Smad7 by targeting Smad7 to the plasma membrane. *J Biol Chem* **277**: 39919-39925.

Tajima Y, Goto K, Yoshida M, Shinomiya K, Sekimoto T, Yoneda Y, Miyazono K, Imamura T. (2003). Chromosomal region maintenance 1 (CRM1)-dependent nuclear export of Smad ubiquitin regulatory factor 1 (Smurf1) is essential for negative regulation of transforming growth factor  $\beta$  signalling by Smad7. *J Biol Chem* **278**: 10716-10721.

Takaku K, Oshima M, Miyoshi H, Matsui M, Seldin MF, Taketo MM. (1998). Intestinal tumorigenesis in compound mutant mice of both Dpc4 (Smad4) and Apc genes. *Cell* **92**:645-56.

Takeda K, Noguchi K, Shi W, Tanaka T, Matsumoto M, Yoshida N, Kishimoto T, Akira S. (1997). Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proc Natl Acad Sci USA* **94**: 3801-3804.

Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander E, Golub T. (1999). Interpreting patterns of gene expression with self-organising maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* **96**: 2907-2912.

Tavazoie S, Hughes D, Campbell M, Cho R, Church G. (1999). Systematic determination of genetic network architecture. *Nature Genet* **22**: 281-285.

Teramoto T, Kiss A, Thorgeirsson SS. (1998). Induction of p53 and Bax during TGF- $\beta$ 1 initiated apoptosis in rat liver epithelial cells. *Transgenic Res* **6**: 349-356.

Topper JN, Cai J, Qai Y, Anderson KR, Xu YY. (1997). Vascular MADs: two novel MAD-related genes selectively inducible by flow in human vascular endothelium. *Proc Natl Acad Sci USA* **94** : 9314-9319.

Tsuchida K, Matsuzaki T, Yamakawa N, Liu Z, Sugino H. (2001). Intracellular and extracellular control of activin function by novel regulatory molecules. *Mol Cell Endocrinol* **180**: 25-31.

Tsukada T, Eguchi K, Migita K, Kawabe Y, Kawakami A, Matsuoka N, Takashima H, Mizokami A, Nagataki S. (1995). Transforming growth factor- $\beta$ 1 induces apoptotic cell death in cultured human umbilical vein endothelial cells with down-regulated expression of Bcl-2. *Biochem Biophys Res Commun* **210**: 1076-1082.

Tsukazaki T, Chiang TA, Devison AF, Attisano J, Wrana JL. (1998). SARA, a FYVE domain protein that recruits Smad2 to the TGF- $\beta$  receptor. *Cell* **95**: 779-791.

Tusher VG, Tibshirani R, Chu G. (2001). Significant analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* **98**: 5116-5121.

Ulloa L, Doody J, Massague J. (1999). Inhibition of transforming growth factor- $\beta$ /SMAD signalling by the interferon-gamma/STAT pathway. *Nature* **397**: 710-713.

Verschueren K, Remacle JE, Collart C, Kraft H, Baker BS, Tylzanowski P, Nelles L, Wuytens G, Su MT, Bodmer R, Smith JC, Huylebroeck D. (1999). SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT-3' sequences in candidate target genes. *J Biol Chem* **274**: 20489-20498.

Wahl, S.M., N. McCartney-Francis, and S.E. Mergenhagen. 1989. Inflammatory and immunomodulatory roles of TGF-beta. *Immunol Today* **10**:258-261.

Wakefield LM, Roberts AB. (2002). TGF- $\beta$  signalling: positive and negative effects on tumorigenesis. *Curr. Opin. Genet. Dev.* **12**: 22-29.

Wan M, Cao X, Wu Y, Bai S, Wu L, Shi X, Wang N, Cao X. (2002). Jab1 antagonizes TGF-beta signalling by inducing Smad4 degradation. *EMBO Rep* **3**:171-6.

Wang QM, Luo X, Studzinski GP. (1997). Cyclin-dependent kinase 6 is the principal target of p27<sup>kip1</sup> regulation of the G1 phase traverse in 1,25-dihydroxyvitamin D3 treated HL60 cells. *Cancer Res* **57**: 2851-2855.

Weber MS, Boyle PS, Corl BA, Wong EA, Gwazdauskas FC, Akers RM. (1998). Expression of ovine insulin-like growth factor-1 (IGF-1) stimulates alveolar bud development in mammary gland of transgenic mice. *Endocrine* **8**: 251-259.

Weinstein M, Yang X, Li C, Xu X, Gotay J, Deng C. (1998). Failure of egg cylinder elongation and mesoderm induction in mouse embryos lacking the tumor supressor Smad2. *Proc Natl Acad Sci USA* **95**: 9378-9383.

Wells RG, Yankelev H, Lin HY, Lodish HF. (1997). Biosynthesis of the type I and type II TGF-beta receptors. Implications for complex formation. *J Biol Chem* **272**: 11444-11451.

Whitelaw CB, Harris S, McClenaghham M, Simons JP, Clark AJ. (1992). Position-independent expression of the ovine beta-lactoglobulin gene in transgenic mice. *Biochem J* **286 (Pt1)**: 31-39.

Whitman M. (1998). Smads and early developmental signalling by the TGF- $\beta$  superfamily. *Genes Dev* **12**: 2445-2462.

Wiersdorff V, Lecuit T, Cohen SM Mlodzik M. (1996). Mad acts downstream of Dpp receptors, revealing a differential requirement for Dpp signalling in initiation and propagation of morphogenesis in the *Drosophila* eye. *Development* **122**: 2153-2162.

Wieser R, Wrana JL, Massague J. (1995). GS domain mutations that constitutively activate T beta R-I, the downstream signaling component in the TGF-beta receptor complex. *EMBO J* **14**: 2199-21208.

Wilentz R, Iacobuzio-Donahue A, Argani P, McCarthy M, Parsons L, Yeo J, Kern E, Hruban H. (2000). Loss of expression of Dpc4 in pancreatic intraepithelial neoplasia: evidence that DPC4 inactivation occurs late in neoplastic progression. *Cancer Res* **60**: 2002-2006.

Wolkenhauer O, Moller-Levet C, Sanchez-Cabo F. (2002). The curse of normalization. *Comparative and Functional Genomics* **3**: 375-379.

Wotton D, Lo RS, Lee S, Massague J. (1999). A Smad transcriptional corepressor. *Cell* **97**: 29-39.

Wrana JL, Attisano L, Wieser R, Ventura F, Massague J. (1994). Mechanism of activation of the TGF-beta receptor. *Nature* **370**: 341-347.

Wu G, Chen YG, Ozdamar B, Gyuricza CA, Chong PA, Wrana JL, Massague J, Shi Y. (2000). Structural basis of Smad2 recognition by the Smad anchor for receptor activation. *Science* **287**: 92-97.

Wu RY, Zhang Y, Feng XH, Derynck R. (1997). Heteromeric and homomeric interactions correlate with signaling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Mol Cell Biol* 17: 2521-2528

Wurthner JU, Frank DB, Felici A, Green HM, Cao Z, Schneider MD, McNally JG, Lechleider RJ, Roberts AB. (2001). Transforming growth factor- $\beta$  receptor-associated protein 1 is a Smad4 chaperone. *J Biol Chem* 276: 19495-19502.

Xiao Z, Liu X, Henis YI, Lodish HF. (2000a). A distinct nuclear localisation signal in the N terminus of Smad3 determines its ligand-induced nuclear translocation. *Proc Natl Acad Sci USA* 97: 7853-7858.

Xiao Z, Watson N, Rodriguez C, Lodish HF. (2001). Nucleocytoplasmic shuttling of Smad1 conferred by its nuclear localisation and nuclear export signals. *J Biol Chem* 276: 39404-39410.

Xu L, Chen YG, Massague J. (2000). The nuclear import function of Smad2 is masked by SARA and unmasked by TGF- $\beta$ -dependant phosphorylation. *Nat Cell Biol* 2: 559-562.

Yagi K, Goto D, Hamamoto T, Takenoshita S, Kato M, Miyazono K. (1999). Alternatively Spliced Variant of Smad2 Lacking Exon 3. Comparison with wild type Smad2 AND Smad3. *J Biol Chem* 274: 703-709

Yamamura Y, Hua X, Bergelson S, Lodish HF. (2000). Critical role of Smads and AP-1 complex in transforming growth factor-beta-dependent apoptosis. *J Biol Chem* 275: 36295-36302.

Yan Q, Sage EH. (1998). Transforming growth factor- $\beta$ 1 induces apoptotic cell death in cultured retinal endothelial cells but not pericytes: association with decreased expression of p21waf/cip1. *J Cell Biochem* 70: 70-83.

Yanagi Y, Suzawa M, Kawabata M, Miyazono K, Yanagisawa J, Kato S. (1999). Positive and Negative modulation of Vitamin D receptor function by transforming growth factor- $\beta$  signalling through Smad proteins. *J Biol Chem* **274**: 12971-12974.

Yanagisawa J, Yanagi Y, Masuhiro Y, Suzawa M, Watanabe M, Kashiwagi K, Toriyabe T, Kawabata M, Miyazono K, Kato S. (1999). Convergence of Transforming growth factor- $\beta$  and Vitamin D signalling pathways on SMAD transcriptional coactivators. *Science* **283**: 1317-1321.

Yang H, Dudoit S, Luu P, Lin D, Peng V, Ngai J, Speed T. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Research* **30**: e15

Yoo J, Ghiassi M, Jirmanova L, Balliet A, Hoffman B, Fornace A, Liebermann D, Bottinger E, Roberts A. (2003). Transforming Growth Factor- $\beta$ -induced apoptosis is mediated by Smad-dependent expression of GADD45b through p38 activation. *J Biol Chem* **278**: 43001-43007.

Zawel L., J.L. Dai, P. Buckhaults, S. Zhou, K.W. Kinzler, B. Vogelstein, and S.E. Kern. (1998). Human Smad3 and Smad4 are sequence-specific transcription activators. *Mol Cell* **1**:611-617.

Zhang Y, Chang C, Gehling D, Hemmati-Brivanlou A, Derynck R. (2001). Regulation of Smad degradation and activity by Smurf2, an E3 ubiquitin ligase. *Proc Natl Acad Sci USA* **98**: 974-979.

Zhang Y, Feng XH, Wu RY, Derynck R. (1996). Receptor associated Mad homologues synergize as effectors of the TGF- $\beta$  response. *Nature* **383**: 168-172.

Zhou S, Zawel L, Lengauer C, Kinzler KW, Vogelstein B.(1998). Characterization of human FAST-1, a TGF- $\beta$  and activin signal transducer. *Mol Cell* **1**: 121-127.

Zhu H, Kavsak P, Abdollah S, Wrana J, Thomsen G. (1999). A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. *Nature* **400**: 687-693.

Zhu Y, Richardson A, Parada F, Graff J (1998). Smad3 mutant mice develop metastatic colorectal cancer. *Cell* **94**: 703-714.

