Transgenic approaches for the investigation of putative airway stem cells as potential targets for gene correction therapy

Duncan Andrew Whitaker

Thesis Presented for the Degree of Doctor of Philosophy
University of Edinburgh
2004

Declaration

I declare,

(a) that this thesis is composed by myself,
(b) that this work is my own except where otherwise stated, and
(c) that this work has not been submitted for any other degree or professional qualification.

Duncan Whitaker

December 2003

Acknowledgements

While it hardly seems fair to taint the reputations of some otherwise fine people by associating them with this work, it would be rude not to acknowledge their contributions, both direct and indirect, to the completion of this thesis.

Thanks to Julia and the rest of the Dorin lab, particularly Sheila and Fiona for their technical assistance and willingness to help when I had too much to do and not nearly enough time in which to do it. Thanks to Heather for her attempts to instill me with her dogged optimism, to Alison M for providing the perfect antidote, and to both of them for tolerating (and occasionally participating in) some frankly outrageous conversations, for which Simon "Master Chief" Heaney is partly to blame. Karen and Pam: good luck.

Thanks also to David Porteous and the gene therapy group, particularly Peter, Barbie, and Chris for looking after me during the first half of my project. I'm very grateful to David Melton for providing transgenic mice, showing a genuine interest in my progress, and doing his best to make me a published author. The staff of the transgenic facility, especially Heather, Scott, and Janice, deserve thanks for helping out when I was knee-deep in mice.

Several fellow C3-ers deserve thanks for putting up with me when I came nagging for help and advice, including Laura, Shirley, and Lee: despite Dr Spraggon's apparent attempts to sabotage my work, he took an inexplicable interest in reading my thesis chapters and provided many valuable comments for which I will always be grateful. Allyson Ross deserves thanks for her consistently obliging and helpful histology advice, and I am very grateful to Paul Perry for being so approachable and genuinely caring about the plight of the final-year students. I am deeply indebted to Alison W who, shortly after emerging from "Thesis Hell" herself, provided invaluable help in formatting this thesis, and probably spared me the fate of spending Christmas 2003 alone in front of a computer – I will never forget your generosity.

On a personal note, thanks to Darren for sharing my sense of humour for the last four years, and to him and Rachel for putting a roof over my head. Cheers to the Haldane Boys for guaranteed good times, particularly Gordon for the classic Pizza Hut/movie combo, and the late-night Opium sessions. Essential to me completing this work was the constant support of my mum, dad, and bro (I'll listen to you next time), who I will reward with my company for the foreseeable future.

Finally, immeasurable thanks to Sapna for her unique ability to cheer me up without even trying, and for not forgetting about the miserable, whining sod back in Edinburgh. You were all that kept me going when things were going badly, and the reason I'll never regret embarking on this project – tha-yoo-soh-mush!

Abstract

Since the discovery of the *CFTR* gene over a decade ago, Cystic Fibrosis (CF) has been regarded as amenable to intervention by gene therapy. The ultimate aim of gene therapy must be the correction, within cells capable of repopulating the tissue, of the genetic defect in its chromosomal context. Towards that end, a mouse model designed to evaluate the efficiency of gene correction was created, and a transgenic approach was taken to the investigation of a putative progenitor cell population in the adult murine respiratory tract.

Before gene correction systems can be considered as valid therapeutic agents, their utility in the cells and tissues of living animals must be demonstrated. Thus, an *in vivo* system permitting the simple quantification of correction frequency in a wide range of tissues would be a valuable resource for the gene correction community. The generation and analysis of a transgenic mouse carrying an inactivated, but potentially correctable, reporter transgene is described.

The full potential of a gene correction strategy to provide a single-dose, permanent solution to a genetically-diseased tissue will only be realised once the therapy is able to target resident stem cells. For CF lung disease, this will require the prior identification of stem cells in the respiratory epithelium. Previous work has indicated that potential stem cells are spatially coincident with small groups of cells expressing high levels of keratin 5 (K5) protein in the proximal murine trachea. In order to investigate lineage arising from this putative stem cell niche, transgenic mice have been generated which express an inducible form of Cre recombinase from the K5 promoter. Preliminary experiments demonstrate recombination of a conditional reporter gene after induction of Cre activity in K5-expressing tissue. Comparison of the inducible system with a constitutive K5 promoter-driven Cre line validated the choice of the former, as the clarity of data obtained from the conventional system was undermined as a result of K5 promoter activity causing reporter gene activation prior to the onset of the experiment.

In the course of these studies it became evident that the conventional, constitutive Cre line gave rise to segregating patterns of reporter gene activation. While some mice displayed the expected K5-derived expression profile, other animals demonstrated ubiquitous expression. Universal activation of the conditional reporter was detected only in animals derived from females carrying the Cre transgene, and was found to be the result of unanticipated production of Cre protein in the maternal germline. This transgenic line is unusual and valuable in offering a choice of tissue-specific and generalised recombination of floxed alleles.

Table of Contents

Declaration ..... *i*
Acknowledgements ..... *ii*
Abstract ..... *iii*
Table of Contents ..... *iv*
List of Figures ..... *xii*
List of Tables ..... *xiv*
List of Abbreviations ..... *xv*
Transgenic nomenclature ..... *xviii*
Chapter 1 Introduction *1*
1.1 Why is it important to identify airway epithelial stem cells? *2*
1.2 Stem cell theory *3*
1.2.1 Adult tissue-specific stem cells **3**
1.2.2 The stem cell niche **6**
1.2.3 The evolving concept of the stem cell **9**
1.3 Epithelial stem cells *12*
1.3.1 Identification of stem cells **12**
1.3.2 Epithelial stem cells share common features **14**
1.4 Stem cells of the respiratory epithelium *15*
1.4.1 Development of the mammalian respiratory system **16**
1.4.2 Anatomy of the mammalian respiratory system **17**
1.4.3 Cellular composition and kinetics of the mammalian respiratory system **18**
1.4.3.1 Alveolar cells 18
1.4.3.2 Ciliated cells 19
1.4.3.3 Basal cells 19
1.4.3.4 Clara cells 23
1.4.3.5 Pulmonary neuroendocrine cells 24
1.4.4 Submucosal glands **24**
1.4.5 Zonal organisation of airway epithelial stem cell systems **29**
1.4.5.1 The proximal airway 29
1.4.5.2 The distal airway 30
1.4.5.3 Airway stem cell niches? 31
1.4.5.4 A single multipotent stem cell for the entire respiratory tract? 33

1.4.6 How do current stem cell concepts apply to the airway epithelium? **33**

1.5 Aims *36*

1.5.1 Investigation of a putative stem cell niche in murine tracheal submucosal gland ducts **36**

1.5.2 An in vivo model for gene correction **36**

Chapter 2 A transgenic mouse for inducible tracheal lineage analysis *37*

2.1 Introduction *38*

2.1.1 Keratins as stem cell markers **38**

2.1.2 Keratin 5 and the human K5 promoter **40**

2.1.3 Lineage analysis **42**

2.1.4 Temporally controlled activation of tissue-specific lineage marking **44**

2.1.5 Aims **49**

Results *50*

2.2 The transgenic K5 promoter-driven lacZ control line, K5A1 ..... *52*

2.2.1 Microinjection of constructs and identification of transgenic founders **52**

2.2.2 Analysis of K5A1 expression **55**

2.2.3 Embryonic expression of eGFP **56**

2.2.4 Embryonic expression of βgal **56**

2.2.4.1 Transgenic line K5A1100 57

2.2.4.2 Transgenic line K5A1200 61

2.2.5 Adult expression of K5A1 ..... **64**

2.2.6 Tracheal expression of K5A1 ..... **69**

2.3 The transgenic K5 promoter-driven inducible Cre line, K5CE *75*

2.3.1 Microinjection of constructs and identification of transgenic founders **75**

2.3.2 Identification of K5CE lines demonstrating inducible Cre-mediated recombination **75**

2.3.3 Immunodetection of CreER^{T2} in embryonic epidermis **77**

2.3.4 Analysis of K5CE expression in adult skin **79**

2.3.5 A functional assay for inducible Cre-mediated recombination **79**

2.3.6 Determination of relative transgene copy number **87**

2.3.7 Analysis of K5CE expression in the trachea **89**

2.4 Discussion *92*

2.4.1 The transgenic K5 promoter-driven control line, K5A1 ..... **92**

2.4.1.1 Expression of eGFP 92

2.4.1.2 Expression of βgal 93

2.4.1.3 Endogenous βgal expression in tracheal submucosal glands 96

2.4.2 Identification of K5CE lines demonstrating inducible Cre-mediated recombination **98**
2.4.2.1 Epidermal expression of $CreER^{T2}$ transcript 100
2.4.2.2 Tamoxifen-inducible recombination in adult epidermis 101
2.4.2.3 Functional expression of $CreER^{T2}$ did not correlate with transgene copy number 102
2.4.2.4 Tracheal expression of $CreER^{T2}$ transcript 104

Chapter 3 A preliminary investigation of cell lineage in regenerating tracheal epithelium *106*

3.1 Introduction *107*

3.1.1 A model for putative tracheal stem cells in submucosal gland ducts **107**

3.1.2 Do these cells contribute to repair of the tracheal epithelium? **110**

3.1.3 Aims **112**

Results *114*

3.2 Evaluation of the bK5Cre mouse for monitoring cell lineage in regenerating tracheal epithelium *114*

3.2.1 βgal expression in $bK5Cre^{tg}$ $R26R^{tg}$ compound transgenic mice **114**

3.2.2 Intratracheal instillation of polidocanol to damage tracheal epithelium **117**

3.2.3 Analysis of βgal expression in repairing $bK5Cre^{tg}$ $R26R^{tg}$ tracheal epithelium **120**
3.2.3.1 Unanticipated variation in tracheal βgal expression between individual $bK5Cre^{tg}$ $R26R^{tg}$ mice 122
3.2.3.2 Clonal expansion from bK5Cre-expressing progenitors? 123
3.2.3.3 Variation in the staining and distribution of submucosal glands between individual $bK5Cre^{tg}$ $R26R^{tg}$ mice 127
3.2.3.4 Absence of lineage patterns in the distal trachea 128
3.2.3.5 βgal expression in both columnar and basal epithelial cells 128

3.3 Evaluation of the K5CE mouse for monitoring cell lineage in regenerating tracheal epithelium *131*

3.3.1 Analysis of βgal expression in repairing $K5CE^{tg}$ $R26R^{tg}$ tracheal epithelium **131**

3.3.2 Tamoxifen-induced xgal staining results from excision within the R26R allele **136**

3.4 Discussion *139*

3.4.1 Constitutive lineage marking: bK5Cre x R26R **140**

3.4.2 Inducible lineage marking: K5CE x R26R **145**

3.4.3 Validation of an inducible approach to tracheal epithelial lineage marking **147**

3.4.4 Further evidence of progenitorial capacity amongst tracheal basal cells 149

3.4.5 Are submucosal gland duct progenitors involved in steady-state maintenance of the tracheal epithelium? 150

3.4.6 Future directions 151

3.4.6.1 Further evaluation of the inducible Cre transgenic lines 151

3.4.6.2 Inducible lineage analysis of other airway populations 152

3.4.6.3 Further experiments using the constitutive Cre transgenic line 153

3.4.6.4 Do circulating stem cells respond to tracheal injury? 153

Chapter 4 The ability of the bK5Cre mouse to mediate both tissue-specific and generalised recombination *155*

4.1 Introduction *156*

4.1.1 Aims 157

4.2 Results *158*

4.2.1 Alternate patterns of recombination in bK5Cre^{tg} R26R^{tg} mice correlate with the sex of the bK5Cre transgene donor parent 158

4.2.2 Cre activity in bK5Cre oocytes causes universal recombination in the subsequent generation 162

4.2.3 Cre protein mediates recombination of R26R within oocytes 166

4.3 Discussion *170*

4.3.1 Murine oogenesis in brief 170

4.3.2 Accumulation of Cre protein precedes the first meiotic division 172

4.3.3 Nucleolar accumulation of Cre protein does not prevent access to genomic loci 173

4.3.4 Does expression from the bK5 promoter in oocytes reflect endogenous K5 production? 174

Chapter 5 A functional, in vivo model for gene correction *177*

5.1 Introduction *178*

5.1.1 Gene correction 178

5.1.2 RNA/DNA oligonucleotides 179

5.1.3 small DNA fragments 182

5.1.4 Evaluating gene correction in vivo: limitations of current systems 184

5.1.5 Aims 186

Results *187*

5.2 The transgenic mutated reporter line, R26X *187*

5.2.1 Site-directed mutagenesis of lacZ 187

5.2.2 in vitro validation of the R26X construct 191

5.2.3 Microinjection of constructs and identification of transgenic founders 193
5.3 Analysis of R26Z lines *196*
5.3.1 Expression of eGFP 196
5.3.2 Expression of βgal 198
5.3.3 Adult tissues 201
5.4 Analysis of R26X lines *203*
5.4.1 Expression of eGFP 203
5.4.2 Expression of βgal 203
5.4.3 Adult tissues 203
5.5 Discussion *208*
5.5.1 Site-directed mutagenesis of lacZ 208
5.5.2 The R26 promoter 209
5.5.3 R26Z lines 210
5.5.4 R26X lines 211
5.5.5 Low levels of eGFP expression from the IRES-eGFP 213
5.5.6 Non-ubiquitous expression of R26X and R26Z transgenes 214
5.5.7 Alternative approaches 216
5.5.7.1 The CAGG promoter *216*
5.5.7.2 The genomic ROSA26 locus *216*
5.5.7.3 A new reporter system for gene correction strategies *217*
5.5.8 Future Work 218
5.5.8.1 Further analysis of R26X lines *219*
Chapter 6 Concluding Remarks *221*
6.1 A transgenic mouse for the investigation of putative tracheal stem cells *222*
6.2 The molecular basis of alternate recombination patterns mediated by the bK5Cre transgenic mouse *223*
6.3 An in vivo model for functional analysis of gene correction *223*
6.4 Airway stem cells as targets for gene correction therapy *224*
Chapter 7 Materials and Methods *226*
7.1 Manipulation of nucleic acids *227*
7.1.1 General molecular biology reagents 227
7.1.2 Agarose gel electrophoresis 228
7.1.2.1 Solutions 228
7.1.3 DNA quantification 228

7.1.4 Restriction enzyme digestion **229**

7.1.5 Making cohesive ends blunt **229**
7.1.5.1 T4 DNA polymerase 230
7.1.5.2 Mung Bean Nuclease 230

7.1.6 Removal of phosphorylated 5' termini **230**

7.1.7 Ligations **231**

7.1.8 Removal of buffer salts **231**

7.1.9 Extraction of genomic DNA for Southern blot hybridisation **231**

7.1.10 Southern blot hybridisation **232**
7.1.10.1 Preparation of radioactively labelled DNA probes 232
7.1.10.2 Preparation of radioactively labelled oligonucleotide probes 233
7.1.10.3 Hybridisation of radioactive DNA probes 233
7.1.10.4 Hybridisation of radioactive oligonucleotide probes 234

7.2. Polymerase chain reaction (PCR) *234*

7.2.1 Reagents **234**

7.2.2 Amplification programmes **235**

7.3 DNA sequencing *236*

7.4 Site-directed mutagenesis *237*

7.5 Microbiology *238*

7.5.1 Growth media for bacterial cultures **238**
7.5.1.1 Antibiotic selection 239
7.5.1.2 Xgal/IPTG indicator plates 239

7.5.2 Production of electrocompetent cells **239**

7.5.3 Transformations **240**

7.5.4 Preparation of bacterial colony hybridisation filters **240**

7.5.5 Isolation of plasmid DNA **241**

7.6 Production of transgenic animals *241*

7.6.1 Preparation of recombinant DNA for microinjection **241**

7.6.2 Reagents and equipment used during the manipulation of preimplantation embryos **242**
7.6.2.1 Equipment 242
7.6.2.2 Solutions 243

7.6.3 Microinjection of recombinant DNA into fertilised eggs **243**

7.6.4 Oviducal transfers **244**

7.6.5 Genotyping of breeding mice **245**

7.6.6 Animal husbandry **245**

7.6.7 Harvesting of postimplantation embryos **246**

7.6.8 Genotyping of embryos ... 246
7.7 Wholemount expression analysis ... 247
7.7.1 Visualisation of eGFP ... 247
7.7.2 Fixing tissue ... 247
7.7.3 Xgal staining ... 248
7.7.3.1 Solutions ... 248
7.7.4 Wholemount microscopy ... 249
7.8 Histology ... 249
7.8.1 Tissue processing for paraffin wax sectioning ... 249
7.8.2 Microtome sectioning ... 250
7.8.3 Immunohistochemistry and immunofluorescence ... 251
7.8.3.1 Anti-Cre immunofluorescence ... 251
7.8.3.2 Anti-βgal immunofluorescence ... 251
7.8.3.3 Anti-K14 immunohistochemistry ... 252
7.8.4 Microscopic analysis of tissue sections ... 252
7.9 Analysis of mRNA expression ... 253
7.9.1 RNA extraction ... 253
7.9.2 cDNA synthesis ... 254
7.10 Induced injury to tracheal epithelium ... 254
7.10.1 Preparation and administration of polidocanol ... 254
7.11 Ligand-inducible Cre recombinase activity ... 255
7.11.1 Preparation and administration of tamoxifen ... 255
7.11.1.1 Ear-painting assay ... 255
7.1.11.2 Systemic administration ... 256
7.12 in vitro validation of R26X and R26Z constructs ... 256
7.12.1 G-418-mediated translational readthrough assay ... 256
7.12.2 Evaluation of R26 promoter activity and inactivity of the R376X lacZ sequence ... 257
7.13 Statistical analysis ... 257
7.14 Computational methods ... 258
Appendices ... 259
A1 Construction of transgenic lines K5A1 and K5CE ... 260
A1.1 Cloning the human K5 promoter into the promoter-less pIRES2-eGFP ... 260
A1.2 Incorporating the pCI intron from pK5LH14 ... 261
A1.3 Cloning lacZ to generate pK5A1 ... 262

A1.4 Cloning CreER^{T2} into pK5-eGFP to create pK5CE 262
A2 Construction of transgenic line R26X 267
A2.1 Exchanging the CMV promoter for the R26 promoter 267
A2.2 Addition of IRES-eGFP to make a bicistronic construct 267
References 272

List of Figures

1.1	Independent axes of proliferation capacity and differentiation	5
1.2	Invariant and regulative mechanisms for generating differentiating cells	5
1.3	Niche-mediate stem cell regulation through control of mitotic orientation	8
1.4	The stem cell niche of the mammalian intestinal crypt	8
1.5	The "stem cell highway"	11
1.6	Xenograft lineage model of airway epithelial differentiation	22
1.7	Structure and composition of human submucosal glands	26
1.8	Location and appearance of murine submucosal glands	26
1.9	Stem cell niches in murine proximal and distal airways	32
2.1	Ligand-induced recombinase-mediated ubiquitous reporter activation	46
2.2	Cloning strategy to generate *K5A1* and *K5CE* transgenic mice	51
2.3	PCR analysis of integrated *K5A1* transgenes	54
2.4	Expression of βgal and eGFP in *K5A1100* embryos	58
2.5	Expression of eGFP in *K5A1200* embryos	59
2.6	Expression of βgal in *K5A1100* embryonic epidermis	60
2.7	Expression of βgal in *K5A1200* embryos	62
2.8	Expression of βgal in *K5A1200* embryonic epidermis	63
2.9	Expression of eGFP in *K5A1100* and *K5A1200* pinna tissue	65
2.10	Expression of βgal in adult *K5A1100* pinna	67
2.11	Expression of βgal in adult *K5A1200* tissues	68
2.12		
(i)	Expression of βgal in *K5A1100* tracheas	72
(ii)	Expression of βgal in *K5A1200* tracheas	73
(iii)	Endogenous xgal staining in tracheas of non-transgenic *K5A1* littermates	74
2.13	PCR analysis of integrated *K5CE* transgenes	76
2.14	Anti-Cre immunofluorescence on *K5CE* embryonic skin	78
2.15	RT-PCR for expression of $CreER^{T2}$ in skin of seven *K5CE* lines	80
2.16	Tamoxifen ear-painting regime for identifying functional *K5CE* lines	83
2.17	Results of tamoxifen ear-painting assay	85
2.18	Relative copy number analysis of seven *K5CE* lines	88
2.19	RT-PCR for expression of $CreER^{T2}$ in trachea of three *K5CE* lines	91
3.1	Three zone model for label retention around the submucosal gland duct	109
3.2	βgal expression in $bK5Cre^{tg}$ $R26R^{tg}$ transgenic mice	116
3.3	PCR-based discrimination of hemizygosity and homozygosity for *R26R*	118
3.4	Intratracheal instillation apparatus	118
3.5	Polidocanol-induced damage of tracheal epithelium after 24 hours	119
3.6	βgal expression in repairing $bK5Cre^{tg}$ $R26R^{tg}$ tracheal epithelium	124
3.7	Further observations from $bK5Cre^{tg}$ $R26R^{tg}$ lineage marking	130
3.8	Timeline of *K5CE* tracheal lineage marking experiment	133
3.9	Systemic tamoxifen-induced βgal expression in $K5CE^{tg}$ $R26R^{tg}$ mice	134
3.10	Tamoxifen-induced βgal activity correlates with *R26R* recombination	138
3.11	Observed lineage patterns are consistent with the three zone model	143

4.1	Alternate patterns of Cre-mediated recombination	159
4.2	Universal recombination of R26R independent of *bK5Cre* inheritance	161
4.3	Immunofluorescence detection of Cre protein in *bK5Cre* oocytes	163
4.4	Immunofluorescence analysis of serial *bK5Cre* oocyte sections	165
4.5	βgal production in $bK5Cre^{tg}$ $R26R^{tg}$ oocytes	167
4.6	Zygotic recombination of paternal *R26R* alleles by maternal Cre protein	169
4.7	Development of the mouse oocyte	171
5.1	RNA/DNA oligonucleotides and single-stranded oligonucleotides	180
5.2	Cloning strategy to generate *R26X* transgenic mice	188
5.3	Site-directed mutagenesis of *lacZ*	190
5.4	*in vitro* validation of pR26X	192
5.5	PCR analysis of integrated *R26X* transgenes	195
5.6	eGFP expression in *R26Z* embryos	197
5.7	βgal expression in *R26Z100* embryos	199
5.8	βgal expression in *R26Z300* embryos	200
5.9	βgal expression in *R26Z* adult brain	202
5.10	eGFP and βgal expression in *R26X100*	204
5.11		
(i)	RT-PCR analysis of *lacZ* expression in 12 *R26X* lines	206
(ii)	*R26X* lines express mutated *lacZ* transcript	207
A1.1	Cloning the K5 promoter into the promoter-less pIRES2-eGFP vector	263
A1.2	Cloning the pCI intron into pK5-eGFP	264
A1.3	Cloning *lacZ* into pK5CI-eGFP	265
A1.4	Cloning the $CreER^{T2}$ fusion gene into pK5-eGFP	266
A2.1	Cloning the R26 promoter into pCMVβ	269
A2.2		
(i)	Addition of an IRES-eGFP component	270
(ii)	Addition of an IRES-eGFP component	271

List of Tables

2.1	Summary of microinjection sessions for *K5A1* and *K5CE* constructs	53
2.2	Predicted frequencies of transgene hemizygosity and homozygosity	82
2.3	Summary of expression analysis on seven *K5CE* lines	99
3.1	Summary of *bK5Cre* tracheal lineage analysis experiment	121
3.2	Summary of *K5CE* tracheal lineage analysis experiment	133
4.1	Genotype-phenotype correlations in $bK5Cre^{tg/wt}$ x $R26R^{tg/wt}$ offspring	160
4.2	Numbers of DAPI- and anti-Cre-positive oocytes in *bK5Cre* ovaries	160
5.1	Summary of microinjection sessions for *R26X* and *R26Z* oocytes	194

List of Abbreviations

°C	degrees centigrade
[^{3}H]TdR	Tritiated thymidine
A	absorbance
ABC	avidin biotin complex
BAC	bacterial artificial chromosome
BADJ	bronchoalveolar duct junction
βgal	β-galactosidase
bluogal	5-bromo-3-indolylβ-D-galactopyranoside
bp	base pairs
BM	bone marrow
BrdU	bromodeoxyuridine
BSA	bovine serum albumin
CCSP	Clara cell secretory protein
cDNA	complementary deoxyribonucleic acid
CF	Cystic Fibrosis
CFE	colony-forming efficiency
CFTR	cystic fibrosis transmembrane conductance regulator
CGRP	calcitonin gene-related peptide
CNS	central nervous system
CMV	cytomegalovirus
DAB	diaminobenzidine
DAPI	4'6-diamidino-2-phenylindole-2HCl
DEPC	diethyl pyrocarbonate
dH_2O	distilled water
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	dinucleotide triphosphate
ds	double stranded
E	embryonic day
ECM	extracellular matrix
E. coli	Escherichia coli
EDTA	ethyldiaminetetra-acetic acid di-sodium salt
eGFP	enhanced green fluorescent protein
ER	oestrogen receptor
ES	embryonic stem
g	gram
G-418	Geneticin
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
hCG	human chorionic gonadotrophin
HCl	hydrogen chloride
hPAP	human placental alkaline phosphatase
HSC	haematopoietic stem cell
HSP	heat shock protein
IP	intraperitoneal

IRES	internal ribosome entry site
K	keratin
kb	kilobase pairs
l	litres
lacZ	β-galactosidase gene
LBD	ligand-binding domain
LEF-1	lymphoid enhancer factor-1
LRC	label-retaining cell
M	molar
m	prefix *milli*
μ	prefix *micro*
mRNA	messenger RNA
myoG	myogenin
n	prefix *nano*
NEO	neomycin phosphotransferase
NEB	neuroepithelial body
NO_2	nitrogen dioxide
NSN	non-surrounded nucleolus
nt	nucleotides
O_3	ozone
OD	optical density
OHT	4-hydroxytamoxifen
ORF	open reading frame
PBS	phoshpate buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
pmoles	pico moles
PMS	pregnant mare serum
PNEC	pulmonary neuroendocrine cell
POL	poloxethylene-9-lauryl-ether (polidocanol)
PR	progesterone receptor
RDO	RNA/DNA oligonucleotide
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcription-polymerase chain reaction
SDF	short DNA fragment
SDS	sodium dodecyl sulfate
SFHR	short fragment homologous replacement
SMG	submucosal gland
SN	surrounded nucleolus
SO_2	sulphur dioxide
SP	side population
SP-C	surfactant protein C
SPU	structural-proliferative unit
ss	single stranded
SSC	saline sodium citrate
SSO	single-stranded oligonucleotide

TA	transit amplifying
TAE	tris, acetic acid, EDTA
TAM	tamoxifen
TBE	tris, boric acid, EDTA
TD	terminally differentiated
TE	tris, EDTA
TPA	12-*O*-tetra-decanolyphorbol-13-acetate
UV	ultraviolet
xgal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Transgenic nomenclature

Throughout this thesis, transgenic mice are referred to after the name of the transgene carried. For example, mice carrying the *K5CE* transgene are known as *K5CE* mice. Individual transgenic lines are referred to by the transgene name suffixed with a number, for example *K5CE100* and *K5CE200*. This may be abbreviated to *K5CE100,200*.

For clarity in experiments involving compound transgenic and mono-transgenic mice, the suffixes "^{tg}" and "^{wt}" are used to signify transgene-positive and transgene-negative, respectively. For example, a $K5CE^{tg}$ $R26R^{tg}$ mouse carries both *K5CE* and *R26R* transgenes, whereas a $K5CE^{tg}$ $R26R^{wt}$ mouse harbours only the *K5CE* transgene. Where appropriate, this nomenclature also signifies transgene hemizygosity (for example $K5CE^{tg/wt}$) or homozygosity (for example $K5CE^{tg/tg}$).

Chapter 1 Introduction

1.1 Why is it important to identify airway epithelial stem cells?

Contrary to the situation in certain tissues, progenitor-progeny relationships and patterns of cell lineage have not been definitively established within the mammalian respiratory epithelium. Other than out of fundamental biological interest, there are several clinical justifications for efforts aimed at elucidating airway epithelial lineages and the identification of multipotent progenitor cells.

The airway epithelium is highly labile in its cellular composition and function, and deviations from the steady-state likely contribute to the deterioration in lung function of patients with chronic airway diseases such as asthma, bronchitis and cystic fibrosis (CF). Alterations in the composition of the surface epithelium can reflect ongoing cycles of injury and repair, or may result from atypical differentiation profiles amongst epithelial cell types in response to aberrant cell signalling. An understanding of epithelial cell lineage relationships is a prerequisite to potentially intervening in disease processes. For example, it is possible that a "failure" of stem cells may underlie the loss of epithelium in airway allografts characteristic of bronchiolitis obliterans (Neuringer *et al.*, 2002; Schoch *et al.*, 2003), the single greatest cause of morbidity and mortality following lung transplantation (Hertz *et al.*, 2002). Gaining an insight into this process might suggest approaches to improving patient outcome. Also, adult stem cells have been implicated in carcinogenesis, with malignant transformation often preceded by progressive phenotypic changes (Trosko, 2003). An appreciation of the location and characteristics of stem cells in the respiratory epithelium would benefit efforts to understand the mechanisms involved, to potentially therapeutic ends.

Of most relevance within the context of this thesis, stem cells of the respiratory epithelium would be the ideal cellular targets of gene therapy for airway diseases such as CF, which has been the longstanding interest of the research group in which this work was undertaken. Complementation or correction of a genetic defect within

airway progenitors should be transmitted to their differentiated progeny, potentially allowing a permanent, single-dose gene therapy to be realised.

1.2 Stem cell theory

1.2.1 Adult tissue-specific stem cells

Although the concept of a stem cell has begun to evolve in recent years, the existence of tissue-specific adult stem cells active in both homeostatic maintenance and injury-induced regeneration of their host tissue, is well established (Blau *et al.*, 2001). While useful molecular markers have been identified for the resident stem cells in a minority of tissues, such as keratin (K) 15 for human hair follicle stem cells (Lyle *et al.*, 1998), and stem cell enrichment is possible using the side population (SP) protocol in several tissues (Goodell *et al.*, 1997; Gussoni *et al.*, 1999; Dunnwald *et al.*, 2001), no universal molecular identifier for stem cells has been discovered (Ivanova *et al.*, 2002; Ramalho-Santos *et al.*, 2002; Fortunel *et al.*, 2003). As such, stem cells can only truly be defined by their future potential, that is, the number and nature of their cellular progeny. Inherent in this concept is that the functional capabilities of a cell must be tested before it can be considered a stem cell. Analogous to the uncertainty principle in physics, the experimental manipulation required to evaluate the function of a cell may alter its properties. Nevertheless, criteria have been proposed which, to greater or lesser degrees, candidate stem cells should fulfil. In addition to lacking markers of differentiation, they should be: capable of proliferation; able to self-maintain; able to produce a large number of differentiated, functional progeny; able to regenerate the tissue after injury, and able to exercise a flexible control over these attributes (Potten and Loeffler, 1990). Cells or populations of cells fulfilling all of these criteria at a given instance comprise the "actual stem cells", while others harbouring such latent capabilities, for example the non-dividing stem cells of quiescent tissues, are considered "potential stem cells" (Potten and Loeffler, 1990).

If proliferation and differentiation are considered independent characteristics, then the cells of self-renewing tissues such as epithelia can be perceived as existing at varying positions on an axis where proliferative capacity is inversely proportional to the degree of differentiation (Potten and Loeffler, 1990) (Figure 1.1). At one extreme lie the undifferentiated stem cells with an apparently infinite capacity for proliferation and self-renewal, while at the opposite extent exist the functionally mature end point cells, exhibiting little or no proliferative potential. The latter population are known as terminally differentiated (TD) cells. Situated between the stem and TD compartments are the transit amplifying (TA) cells, exhibiting properties intermediate to those of the stem cells and TD cells: while TA cells are developmentally committed and normally show the onset of differentiation marker expression, they retain a significant but finite proliferative capability, and are likely capable of limited self-renewal. Thus the key difference between stem cells and early TA cells is the former's ability for unlimited self-renewal. The TA compartment can therefore be considered to represent a spectrum of declining stem cell properties, with early TA cells (or "progenitors") possibly representing potential stem cells capable of re-occupying a vacant stem cell compartment (Loeffler and Potten, 1997). However under normal circumstances, the TA cell population exists essentially to amplify the number of output functional cells generated from a low number of input stem cells, thereby permitting a smaller stem cell pool and a reduced target for genetic and carcinogenic damage. Most of the proliferation occurring in a steady-state tissue is performed by TA cells, while the stem cells cycle slowly. This affords the stem cell population a measure of protection from the accumulation of genetic errors which result from high levels of DNA replication and mitosis.

There are two general strategies by which a stem cell can both self-maintain, and give rise to daughter cells committed to terminal differentiation. One mechanism is invariant and results in the predetermined asymmetric production of one stem cell and one destined-to-differentiate daughter cell (Figure 1.2a). There are numerous examples of such lineages in unicellular organisms and invertebrates. In the *Drosophila* peripheral nervous system, asymmetric divisions of the sensory organ precursor cell are controlled by a hierarchy of genes which coordinate the

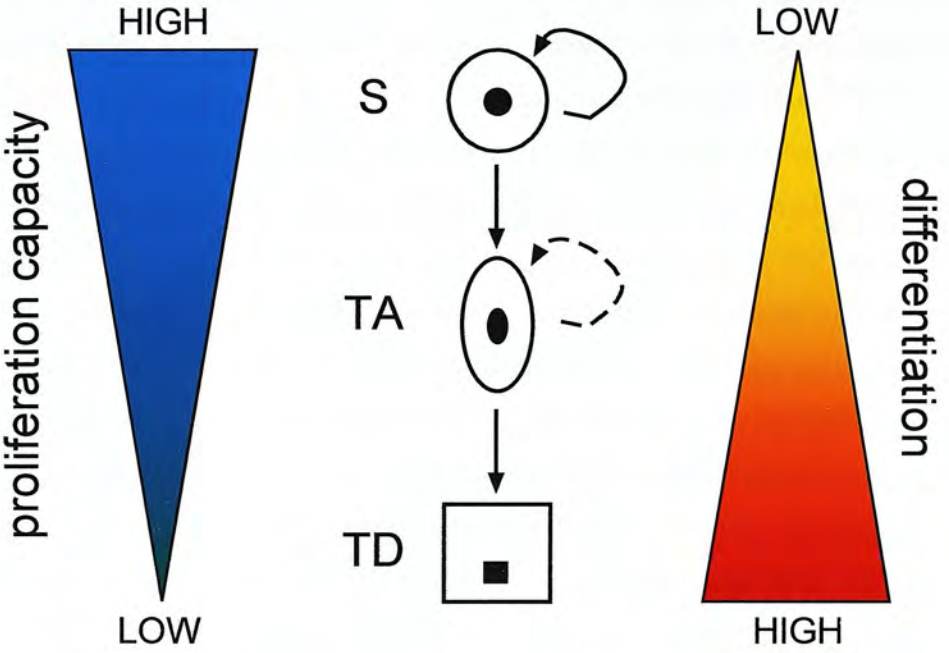


Figure 1.1 Stem (S), transit amplifying (TA), and terminally differentiated (TD) cells can be considered to reside on independent axes of proliferation capacity and differentiation.

The dashed arrow represents finite self-renewal capability amongst TA cells.

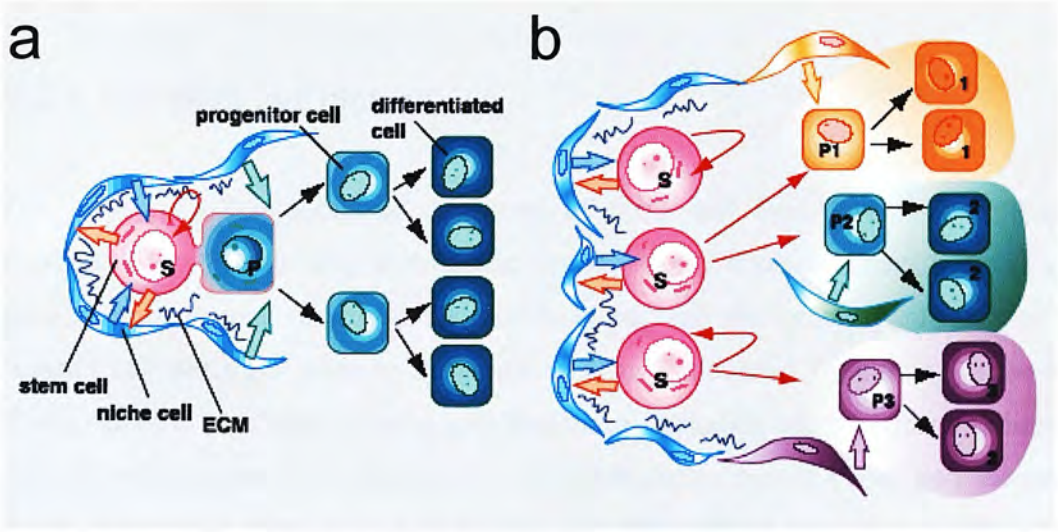


Figure 1.2 The generation of differentiated daughter cells from stem cells can occur by invariant or regulative mechanisms (modified from Watt and Hogan, 2000).

(a) Invariant asymmetry. A stem cell (S) gives rise to another stem cell and a non-stem "progenitor" (P) daughter which differentiates in response to extrinsic signals (thick green arrows). The stem cell state is regulated by reciprocal interactions with niche cells (thick blue and red arrows). The concept of the stem cell niche is discussed in section 1.2.2. (b) Regulative (populational) asymmetry. In response to tissue demand, stem cells can undergo symmetric division, to produce only stem cells (top) or only progenitor cells (middle), or asymmetric division (bottom). Daughter cells differentiate along different pathways (1, 2 and 3) depending on the combination of extrinsic factors to which they are exposed. ECM = extracellular matrix.

asymmetric localisation of cell fate determinants and influence the orientation of mitotic spindles such that, following cell division, the two daughter cells occupy specific positions which likely dictate their exposure to different extracellular signals (Jan and Jan, 1998). In strictly asymmetric stem cell lineages, no regulation of stem cell number is possible (Morrison *et al.*, 1997), yet there is ample evidence for changes in the size of stem cell populations in mammals, for example the amplification of foetal primordial germ cells during their migration into the genital ridge (Tam and Snow, 1981). This is consistent with the view that most mammalian self-renewing tissues are produced through highly regulative mechanisms in which a stem cell gives rise to daughter cells capable of becoming either stem cells or committed progenitors (Watt and Hogan, 2000) (Figure 1.2b). In such systems, steady-state stem cell divisions give rise to, on average, one stem cell and one committed cell, with asymmetry achieved on a population level through symmetric stem cell divisions. Populational asymmetry facilitates the response to variable physiological needs, such as regeneration following injury, and can be mediated by stochastic events, heterogeneity amongst the stem cell population, and environmental regulation through intercellular and extracellular signalling (Hall and Watt, 1989).

1.2.2 The stem cell niche

The intrinsic cell-autonomous regulators of stem cell fate, including proteins responsible for establishing asymmetric cell divisions, nuclear factors controlling gene expression and chromosomal modifications, and timing mechanisms which "count" cell divisions prior to differentiation, are modulated by external signals. These extrinsic regulators of stem cell behaviour, including secreted factors, direct cell-cell interactions, and adhesion to the extracellular matrix (Watt and Hogan, 2000), comprise a microenvironment which is often defined by a discrete physical location within a tissue. Indeed, it has been speculated that since many of the specialised functions required to ensure proper stem cell function are manifest in neighbouring differentiated cells, stem cells themselves may be relatively unspecialised, and may be characterised more by their localisation within physically defined compartments rather than distinctive patterns of gene expression (Spradling

et al., 2001). Units comprising a subset of tissue cells and extracellular substrates and signals, which create a distinct microenvironment able to house and regulate stem cells, are known as niches. While the location of the niche remains obscure in most stem cells systems, it is believed to have been identified in several tissues including the epidermis (hair follicle bulge) (Taylor *et al.*, 2000), gut (crypt) (Booth and Potten, 2000), and the blood (osteoblasts within the bone marrow) (Calvi *et al.*, 2003).

The mode of stem cell regulation, asymmetric or symmetric, within a given tissue system can be considered to be a consequence of interactions between the stem cells and extracellular niche components. In asymmetric stem cell systems, so-called "lineage niches" specify precisely individual stem cell divisions and daughter cell fates (Spradling *et al.*, 2001). This might occur through contact of the stem cell with the underlying stroma: by appropriately orientating its division plane, the stem cell can ensure that only one daughter inherits adhesive contacts with the basement membrane (Figure 1.3). While this daughter will be retained in the niche, the other daughter is free to move away from the stem cell-maintaining stromal signals, and will thus begin to differentiate. Precisely such a mechanism operates in the *Drosophila* male germline stem cell niche (Yamashita *et al.*, 2003). In stem cell systems regulated on a population basis, local cues could stimulate differentiation, or retention of stem cell characteristics in both daughters by influencing the orientation of the division plane, or by other means. While some niches may only operate by a single mechanism, others, such as the *Drosophila* female germline stem cells, are able to switch between invariant and regulative mechanisms (Xie and Spradling, 2000).

One of the best characterised examples of the classic stem-TA-TD model of tissue maintenance is the vertebrate small intestine, which is composed of outwardly projecting ciliated villi surrounded by embedded crypts (Figure 1.4). Lineage tracing experiments have identified that approximately 4-5 multipotent stem cells reside near the bottom of each crypt (Bjerknes and Cheng, 1999), though no markers are known to identify these cells. The rapidly proliferating TA population migrate up the walls

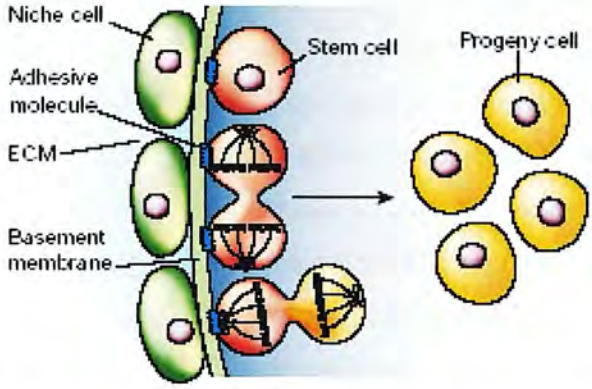


Figure 1.3 Niches can mediate stem cell regulation through control of mitotic orientation (from Spradling *et al*, 2001).

Niche cells (green) underlying the basement membrane signal to stem cells (red) to block differentiation and regulate division. Invariant asymmetric division (lower mitotic cell) can result from the orientation of mitosis permitting only one daughter cell to remain attached to the niche. The unattached daughter (yellow) moves away from niche signals and differentiates. The attachment of both daughters (upper mitotic cell) or neither daughter (not shown) causes symmetric division in systems which are regulated on a populational, asymmetric basis. ECM = extracellular matrix.

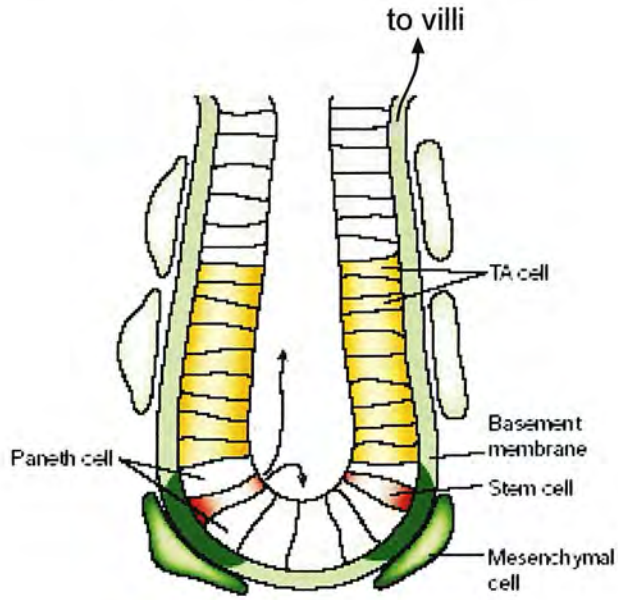


Figure 1.4 The stem cell niche of the mammalian intestinal crypt (modified from Spradling *et al*, 2001).

Stem cells (red) are located in the basal region of the crypt, and give rise to transit amplifying (TA) cells (yellow) which migrate up the walls of the crypt towards the villi. Secretory Paneth cells are also produced, but remain in the basal compartment. Signals emanating from mesenchymal cells (green) under the basement membrane are believed to regulate stem cell activity.

of the crypt and differentiate into the absorptive brush-border enterocytes, mucous-secreting goblet cells, or enteroendocrine cells of the villi. As the TD cells die they are shed from the villi tips into the gut lumen. Thus the stem, TA and TD compartments occupy distinct spatial domains. Paneth cells are also produced and remain in the crypt where they secrete antimicrobial peptides, digestive enzymes, and growth factors. Studies using an inducible lineage marking system have revealed a homogenisation of marking amongst the crypt cells, suggesting a population mode of stem cell regulation (an asymmetric mechanism would not be expected to affect the number of marked stem cells) (Wong *et al.*, 2000).

Intestinal crypts have also revealed the involvement of extracellular factors which likely constitute elements of the niche microenvironment. Signals emanating from the mesenchyme underlying the crypt cells are thought to maintain the activity of crypt stem cells, and include the intestinal-specific Lef/Tcf family member Tcf4. Tcf4 functions as part of the Wnt pathway, and mice null for Tcf4 lack intestinal stem cells (Korinek *et al.*, 1998). The product of Wnt signalling is the accumulation of β-catenin, which is thought to play a role in regulating the "release" of stem cell daughters from the niche, and their subsequent proliferation and differentiation (Fuchs and Segre, 2000). Further evidence for the role of Wnt pathway factors in stem cell maintenance/commitment regulation is the finding that stabilised Tcf/Lef/β-catenin complexes can induce *de novo* hair follicle production, which follows from the observation that Tcf3 is expressed in the hair follicle bulge, known to represent an epidermal stem cell niche (DasGupta and Fuchs, 1999).

1.2.3 The evolving concept of the stem cell

The importance of the niche in regulating stem cell characteristics has been reinforced by recent studies indicating an unanticipated degree of plasticity of stem cells when removed from their usual residence and transferred to foreign physiologic microenvironments. While adult stem cells have traditionally been thought to be restricted in their differentiative and regenerative potential to their tissues of origin, numerous reports now document the "transdifferentiation" of bone marrow (BM)-

derived cells (haematopoietic stem cells, HSCs) into a range of other somatic tissues, including across embryonic germ layers. Such cells have now been shown to contribute to both chemically-damaged and dystrophic skeletal muscle (Ferrari *et al.*, 1998; Gussoni *et al.*, 1999), and can differentiate into cardiac myocytes, smooth muscle cells and endothelial cells in animal models of myocardial infarction (Orlic *et al.*, 2001; Kocher *et al.*, 2001). BM-derived cells were shown to reconstitute 30-50% of hepatocytes in enzymatically-deficient liver, rescuing the otherwise lethal phenotype (Lagasse *et al.*, 2000), and form glomeruli and renal tubules in both mouse and human kidneys (Poulsom *et al.*, 2001). Differentiation of a single BM-derived cell into mature epithelial cells of the liver, lung, gastrointestinal tract and skin, as well as HSCs, has also been described (Krause *et al.*, 2001). This plasticity is apparently not confined to BM-derived cells: neural and skeletal muscle cells are both able to repopulate the BM of irradiated mice (Bjornson *et al.*, 1999; Jackson *et al.*, 1999), and dermal cells can differentiate into a range of other cell types (Toma *et al.*, 2001).

These studies, and many others, have prompted a re-evaluation of how stem cells are conceptually envisioned, with Blau and colleagues proposing that a stem cell might usefully be considered a biological activity, rather than a discrete entity, that under the right circumstances can be induced in many distinct cell types (Blau *et al.*, 2001). They suggested a model whereby cells from diverse organs transit systemically through the circulation, responding to "homing" signals from sites of local damage or tissue demand (Figure 1.5). Once a stem cell has entered a foreign tissue, microenvironmental elements within the local niche would induce the cell to participate in the functions of that tissue.

Much of the interest in the phenomenon of adult pluripotent stem cells derives from their potentially significant impact on cell-based therapies and regenerative medicine. The ability to culture some of these BM-derived cells prior to transplantation (Jiang *et al.*, 2002) should in principle facilitate *ex vivo* gene therapy protocols. Even without modification, enrichment of adult BM-derived pluripotent

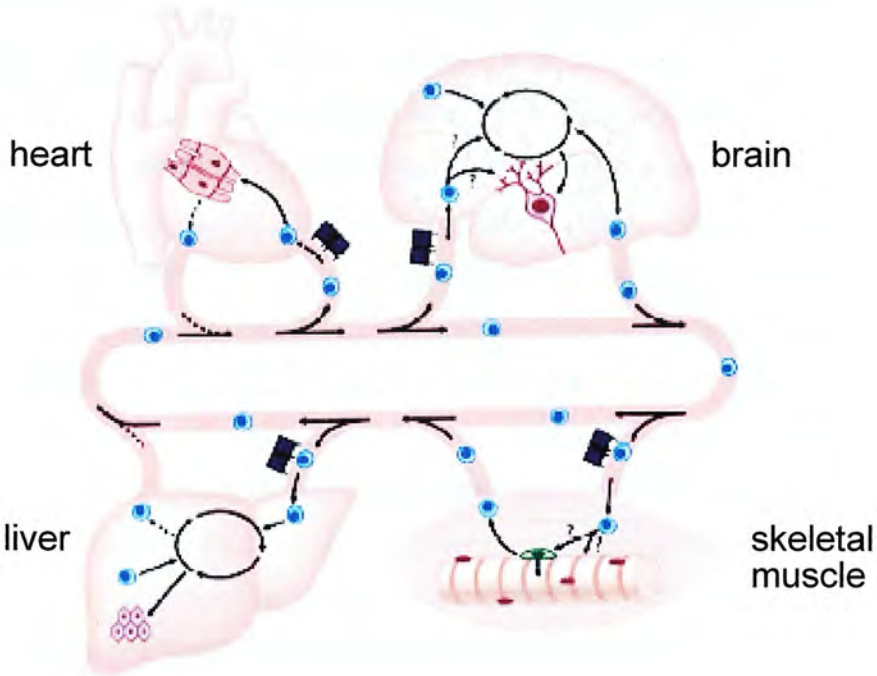


Figure 1.5 The "stem cell highway" (modified from Blau *et al*, 2001).

A model showing documented (solid arrows) and hypothesised (dashed arrows) transitions in stem cell identity and differentiation. In addition to tissue-specific stem cells, some cells may transit systemically through the circulation. If this is considered a "highway" for stem cells, then homing signals and growth factors are depicted on "billboards" near "on-ramps", which stimulate stem cells to enter the tissue and participate in its functions, respectively. The model suggests that stem cells can contribute to the existing pool of stem cells, or directly generate differentiated cells.

stem cells using cell surface markers allows autologous cell transplantation, circumventing problems of immunogenicity.

However, despite the promise of the newly recognised plasticity of adult stem cells, the field remains controversial (Wagers *et al.*, 2002), and much work remains to establish its physiological relevance. Reported frequencies of engraftment of donor cells within recipient tissues are low, particularly in the absence of injury, and there has been little evidence of functional integration (Lemischka, 2002). Alternative explanations for such results include fusion of donor and recipient cells (Ying *et al.*, 2002; Weimann *et al.*, 2003; Alvarez-Dolado *et al.*, 2003), or phagocytosis of donor cells by host macrophages, creating "cells" containing both donor- and tissue-specific markers (Lemischka, 2002). To address these points and make a credible case for the clinical potential of adult pluripotent stem cells, further studies should aim to demonstrate the abilities of single cells to make robust, functional contributions to the expected and unexpected tissues (Lemischka, 2002). From a fundamental biological point of view, further investigation of the role of circulating stem cells in steady-state tissue maintenance, and the relationship between these pluripotent cells and tissue-resident stem cells populations, is required.

1.3 Epithelial stem cells

1.3.1 Identification of stem cells

It is thought that most, if not all, epithelial tissues contain stem cells responsible for normal tissue renewal or for regeneration from damage (Slack, 2000). In the absence of known molecular markers for the stem cells of most tissues, their predicted characteristics have been exploited in attempts to define the locations of stem cell populations. The long lifetime and slow-cycling phenotype of stem cells in steady-state conditions render them amenable to identification through their prolonged retention of a nucleotide analogue such as bromodeoxyuridine (BrdU) or tritiated thymidine ([^{3}H]TdR). The rationale behind this approach is that dividing cells will

incorporate the nuclear label, which will become progressively diluted amongst the DNA of daughter cells during subsequent proliferation. Thus, after a pulse of sufficient duration to ensure uptake of the label by infrequently dividing stem cells, and following a chase period of appropriate length such that the label becomes detectably less abundant in highly proliferative TA cells, the remaining highly-labelled cells (label-retaining cells, LRCs) should represent the slow-cycling, stem cell compartment. This technique has been used extensively in the search for stem cells, including the epidermis and oral mucosa (Bickenbach, 1981), cornea (Cotsarelis *et al.*, 1989), and hair follicle (Cotsarelis *et al.*, 1990).

That hair follicle stem cells exist in the bulge region, rather than in the matrix which represents the TA compartment, was revealed by seven days of repeated subcutaneous injections of [^{3}H]TdR into neonatal mice (Cotsarelis *et al.*, 1990). After a four-week chase period, LRCs were confined to the bulge region, now known to represent a niche of stem cells feeding both the follicle and neighbouring epidermis (Taylor *et al.*, 2000). Attempts to identify LRCs in the relatively quiescent corneal epithelium required prolonged [^{3}H]TdR labelling combined with wounding to stimulate stem cell recruitment and division (Cotsarelis *et al.*, 1989). After a two-week pulse and four-week chase period, 30% of cells in the corneal limbus were classed at LRCs, while no LRCs were detected in the corneal epithelium. This indicated that, following induced proliferation and consequent labelling, the stem cells had reverted back to their slow-cycling state. Application of the tumour promoter TPA revealed a significantly elevated proliferative capacity in limbal epithelial cells compared to corneal epithelia, further supporting the notion that corneal epithelial stem cells reside in the limbus.

Another experimental approach to identifying potential stem cells is based on the notion that they should exhibit considerable proliferative potential, and that this should be reflected by the extent of clonal expansion observed *in vitro*: stem cells form large, self-renewing clones, whereas their TA daughters form abortive clones due to their finite capacity for proliferation (Braun *et al.*, 2003). Clonal analysis has been used to confirm that, within the hair follicle, cells with a high proliferative

capability (clonogenic cells) are confined to the bulge region, and are absent from the matrix-containing hair bulb (Rochat *et al.*, 1994), consistent with the LRC data of Cotsarelis and colleagues (Cotsarelis *et al.*, 1990). In the epidermis, the level of β_1-integrin expression, and the rapidity of adhesion to extracellular matrix proteins, both predicted characteristics of basal epidermal stem cells attached to the basal lamina, were found to correlate with *in vitro* colony-forming efficiency (Jones and Watt, 1993). While the removal of cells from their native *in vivo* environments is likely to alter their characteristics, and may therefore misrepresent their physiologic behaviour, studies in which label-retaining and clonogenic cells have been directly compared revealed that stem cells do share both of these properties (Morris and Potten, 1994; Oshima *et al.*, 2001).

1.3.2 Epithelial stem cells share common features

Following the identification of the LRC compartment within the corneal limbus, Cotsarelis and co-workers noted that various epithelial tissues displayed similarities with respect to the location and biological properties of their resident stem cell populations (Cotsarelis *et al.*, 1989). They are often located in physically protected compartments separate from, but spatially proximal to, their differentiated progeny. For example, the intestinal stem cells are confined to the crypt (Booth and Potten, 2000), removed from the contents of the gut lumen, and epidermal stem cells are located at the bottom of deep rete ridges (Lavker and Sun, 1982; Lavker and Sun, 1983). Corneal limbal stem cells are confined to the edge of the cornea and, unlike the remaining corneal epithelium, are protected from solar damage by heavy pigmentation (Cotsarelis *et al.*, 1989). Both the corneal limbus and the follicular bulge region are more resistant to physical damage than neighbouring corneal epithelium or follicular matrix (bulb), respectively (Cotsarelis *et al.*, 1989; Cotsarelis *et al.*, 1990). Limbal stem cells are found to be associated with nourishing vasculature, as are the stem cells of the palm and trunk epidermis (Lavker and Sun, 1982; Lavker and Sun, 1983), the tongue epithelium (Hume and Potten, 1976), and the hair follicle (Cotsarelis *et al.*, 1990). Given the dependency of epithelial tissues on their stem cells for regeneration following damage, the observation that the stem

cells are usually protected, either intrinsically (for example, pigmentation) or extrinsically (physical sequestration), is intuitive, and this characteristic may provide a useful anatomical clue as to where stem cells may reside.

The discrete spatial positioning of the stem cell compartment with respect to differentiated cells, and the consequences of such an arrangement on histological structure, has led to the concept of "structural-proliferative units" (SPUs), in which one or a few stem cells feed a defined differentiated compartment (Potten, 1978). The intestinal crypt-villus is a good example of such a unit, though many other epithelia exhibit similar histological structures, for example the glands of the stomach, the lobules of the liver, and the nephrons of the kidney (Slack, 2000). Although little evidence supports the notion, studies of the clonal composition of epithelia suggest that these structures may also represent units of cell renewal. Although initially polyclonal, intestinal crypts become monoclonal over time (Schmidt *et al.*, 1988). While more than one stem cell is thought to operate in each crypt (Bjerknes and Cheng, 1999), these could be genetically identical as a result of symmetric stem cell division (Loeffler *et al.*, 1993) which, along with crypt fission (Cheng and Bjerknes, 1985), could also result in loss of genetic diversity within the stem cells of a crypt. According to Slack, gastric glands and hair follicles likely comprise SPUs, but their existence in other tissues is not yet clear as a result of slow cell turnover prolonging the polyclonal phase (Slack, 2000).

1.4 Stem cells of the respiratory epithelium

Stem cell research into the respiratory system has progressed slowly for several reasons, including the anatomical and functional complexities associated with numerous distinct cell types, and differences in the composition and distribution of cells between rodents and primates. Great plasticity in growth capacity amongst resident airway cells in response to injury, often deemed necessary to stimulate stem cell proliferation in this largely quiescent tissue, has reduced the clarity of experimental data and often led to apparently contradictory interpretations. Also, the relationship between stem cells involved in reparative proliferation and those active

in the steady-state and during development, is not clear. The recent discovery of non-tissue-specific, circulating stem cells has further complicated the field. However, several studies in recent years have been informative with regard to lineage and potential progenitors in the airways, and will form the main focus of this discussion. As this thesis is primarily concerned with the prospect of stem cells in the conducting airways, specifically the trachea, only a brief mention will be made of cell lineage in the alveolar regions.

1.4.1 Development of the mammalian respiratory system

The mammalian respiratory system develops from E9.5 foregut in two parts, the unbranched trachea and the remainder of the lung (bronchi plus respiratory tree). The former is created by the division of the foregut into two tubes, the oesophagus and the trachea, and is controlled by mesenchymal expression of Gli1 and Gli2, transcription factors that act downstream of sonic hedgehog (Hogan, 1999). The lung arises by branching morphogenesis of two primary buds that appear on the ventrolateral wall of the foregut, and is thought to be stimulated by Fgf10 expression in the distal mesenchyme around the budding endoderm tips (Hogan and Yingling, 1998). Numerous other morphogenetic signalling factors have been implicated in the process, including Bmp4, Egf, Fox, and Gata-family members (Hogan, 1999; Perl *et al.*, 2002). While it has been speculated that the Gli gene family, and Nkx2.1, might act as "master genes" to impart specificity on the system, deficiencies in these genes have consequences beyond the respiratory system, and no factors specific to morphogenesis of the lung have yet been identified (Warburton *et al.*, 2000).

Using a doxycycline-dependent system to induce permanent lineage marking from surfactant protein C (SP-C)-expressing cells, cell lineages lining bronchioles, acini and alveoli were distinguished from conducting airway lineages lining the trachea and bronchi early in development, prior to formation of the lung buds (Perl *et al.*, 2002). Interestingly, cell marking was never observed in neuroepithelial bodies (NEBs) or tracheal glands, both putative stem cell niches (discussed in section

1.4.5.3), indicating that the cells of these structures do not share a common lineage with other airway cells.

Embryonic development of the lung has been divided into three major stages (Hogan and Yingling, 1998): the pseudoglandular stage, during which most of the branching morphogenesis occurs; the canalicular stage, when the columnar epithelium becomes cuboidal or flattened; and the terminal sac stage, when septation results in the formation of distal saccules. The differentiation of alveolar type I and type II pneumocytes, and the formation of an efficient gas exchange interface, occurs postnatally.

1.4.2 Anatomy of the mammalian respiratory system

The mammalian respiratory system is a complex structure that serves as the site for gas exchange between the pulmonary capillaries and the external environment. At its most basic structure, the respiratory system is composed of two distinct compartments. The first is the conducting portion, comprising a series of bifurcating tubes of ever-decreasing lumen, the function of which is to deliver the inspired air to the alveolar spaces. The trachea is the most proximal, and therefore the largest (in terms of the luminal capacity) of these tubes, connecting the lung to the nose and mouth through the larynx. Distally the trachea splits at the carina into two bronchi, serving the two lobes of the lung. While the trachea and bronchi are supported by cartilage rings, the further distal bronchioles which branch off the bronchi are non-cartilaginous. Ultimately, the bronchioles split to form the alveolar sacs within the lung, the second compartment of the respiratory system, which presents a large surface area to facilitate efficient gaseous exchange. The epithelial layer of cartilaginous airways is pseudostratified, becoming a simple, monolayered epithelium in the distal bronchioles and alveoli.

1.4.3 Cellular composition and kinetics of the mammalian respiratory system

1.4.3.1 Alveolar cells

The alveolar epithelium is simple, being composed primarily of two morphologically distinct cell types, known as type I and type II pneumocytes. Type I cells are large squamous cells with a protruding nucleus and complex cytoplasmic processes, and cover 97% of the alveolar surface, despite representing only 33% of the total cells (Weibel *et al.*, 1976; Hogan and Yingling, 1998). They form a tight sheet of cells covering the alveolar septum to form the pulmonary capillary, by fusing to the basal lamina of vascular endothelial cells. In this way type I cells generate the gas exchange barrier. Type II cells, constituting the remaining 67% of alveolar epithelial cells (though covering only approximately 3% of the surface) are cuboidal and contain abundant mitochondria, endoplasmic reticulum, and Golgi complexes, consistent with their role as secretory cells. The surfactant proteins which reduce the surface tension in alveolar sacs are produced by type II cells. Type I cells are believed to be incapable of cell division, and type II cells are considered bipotential progenitors able to give rise to both type I and type II cells (Magdaleno *et al.*, 1998). In the normal lung, nearly all type II cells are quiescent. Cell division can only occur following cell growth and DNA synthesis, and results in the production of two relatively undifferentiated daughter cells. These cells have the potential to rapidly revert to the type II phenotype, or differentiate into type I cells. It is now thought that either phenotype can transform to the other without cell division, though only type II cells are thought able to re-enter the cell cycle (Uhal, 1997).

While the type II cell is generally considered to be the stem cell responsible for alveolar repair, its sheer frequency renders this notion inconsistent with observations from other tissues, in which the true stem cell pool is represented by a small subpopulation of cells. Mason and colleagues speculated that type II cells may not be functionally identical, and that a subset may act as stem cells, with a candidate population being type II cells located at the junction of the terminal bronchiole and

alveolar duct, which are known to proliferate after injury (Mason *et al.*, 1997). Interestingly, a pollutant-resistant population of airway epithelial cells which contribute to the restoration of injured bronchiolar epithelium, have been identified at this location, discussed in section 1.4.5.2 (Giangreco *et al.*, 2002). However, the existence or location of alveolar epithelial stem cells has not been demonstrated.

In contrast to the alveolar epithelium, the airway epithelium is complex, containing a range of cell types which vary in abundance and distribution between species (Plopper *et al.*, 1983). The major cell types are the ciliated cells, basal cells and Clara (secretory) cells. Other, less abundant cell types include mucous, serous and neuroendocrine cells.

1.4.3.2 Ciliated cells

The most distinctive cell type, with approximately 38% abundance in murine tracheobronchiolar epithelium (Pack *et al.*, 1981) is the ciliated cell, which is common in the larger airways but less frequent in bronchioles. As the name implies, ciliated cells contain abundant cilia at the apical end of the cell, which project into the airway lumen and act to clear mucous and inhaled particles. Consistent with their functionally mature phenotype, ciliated cells are generally considered to be TD, and have a very low proliferative capacity, as evidenced by their failure to incorporate a [^{3}H]TdR label (Donnelly *et al.*, 1982). However, exemplifying the plasticity demonstrated by airway cell types in response to injury, selective damage to Clara cells in the mouse airway has been shown to induce squamation of neighbouring ciliated cells, with subsequent proliferation and re-differentiation into Clara cells (Van Winkle *et al.*, 1995).

1.4.3.3 Basal cells

Comprising approximately 10% of murine tracheobronchial cells (Pack *et al.*, 1981), basal cells are small and contact the basement membrane, but do not have an interface with the airway lumen. Significant evidence exists to support a role for the

basal cell as a multipotent airway progenitor. Basal cells were shown to be the progenitors of rat primary tracheal epithelial cell cultures (Ford and Terzaghi-Howe, 1992a), and were found to give rise, via an intermediate cell stage (discussed later in this section), to non-dividing ciliated cells by assessing [^{3}H]TdR incorporation in steady-state rat tracheal epithelium (Donnelly *et al.*, 1982). Steady-state analysis of [^{3}H]TdR incorporation and retention in hamster bronchial epithelium attributed epithelial renewal activity to both basal and secretory cells, though only basal cells were observed to show a decreasing labelling index, indicating they were the most proliferative cells (Breuer *et al.*, 1990). Although the notion of high proliferative rate is not consistent with current views on the slow-cycling nature of stem cells, the study generated kinetic data which was interpreted to suggest that basal cells were both self-renewing and gave rise to other cell types.

It had been previously shown that purified suspensions of rabbit tracheal basal cells could give rise to all tracheal cell types after inoculation into denuded tracheal xenografts (Inayama *et al.*, 1988). However, the fact that the basal cell suspensions were only 83% to 94% pure meant the involvement of a contaminating cell in epithelial reconstitution could not be excluded. Phenotypic markers of normal airway epithelium were used to assess the cell types involved in regeneration of rat tracheal epithelium following mechanical injury (Shimizu *et al.*, 1994). Adjacent non-wounded epithelium was found to flatten and migrate into the denuded region, where cells became "poorly differentiated" and expressed basal cell-specific markers. These cells gave rise to secretory and ciliated cells which initially continued to express markers typical of basal cells, implying a role for basal cell-like intermediates in regeneration. When high purity basal and secretory/ciliated cell fractions of rat tracheal epithelial cells were inoculated into denuded tracheal grafts, both fractions were able to reconstitute a complete mucociliary epithelium (Randell *et al.*, 1991). However, in a subsequent study aimed at identifying airway progenitors involved in *de novo* submucosal gland (SMG) development (discussed in section 1.4.4), Engelhardt and colleagues studied airway reconstitution by retrovirally-labelling human bronchial cells before seeding them into denuded rat tracheal xenografts (Engelhardt *et al.*, 1995). Following examination of the cellular composition of

labelled clonal clusters, the absence of clones comprising certain combinations of cell types led the authors to conclude that two basal pluripotent progenitor cell types must exist, one of which is an airway stem cell that gives rise to all other cell types (Figure 1.6).

Examination of normal human airway with the anti-Ki-67 proliferation marker revealed that in the upper airway, 51% of the proliferative compartment was comprised of basal cells, despite only 31% of cells reacting with basal cell-specific anti-K5 and -K14 antibodies (Boers *et al.*, 1998). In the lower airways where only 6% of cells were classified as basal, these comprised 30% of the proliferating fraction. This study also revealed a significant proliferative contribution from "parabasal" cells, which correspond to the nebulous "intermediate" (Breuer *et al.*, 1990; Engelhardt *et al.*, 1995) or "poorly differentiated" (Randell, 1992) cells described by others. The parabasal cell type forms a layer located just above the basal cells, though like basal cells they do not contact the airway lumen. They lack the morphological characteristics of other airway cell types, but are immunoreactive with anti-K5 and -K14 antibodies, suggesting they may be a direct descendent of basal cells. A consensus on the legitimacy of the parabasal cell as a genuine epithelial compartment has yet to be reached, however Boers and colleagues considered the parabasal cell to represent the TA cells, with the basal cells acting as stem or progenitor cells (Boers *et al.*, 1998).

Despite the many reports in support of basal cells as airway progenitors, there are several lines of evidence to challenge this concept. In the rhesus monkey, intense proliferation of secretory cells occurs prior to extensive mitotic activity in basal cells, where pre-ciliated cells were observed to contain both secretory granules and basal bodies, suggesting a secretory to ciliated cell pathway (Plopper *et al.*, 1986). The observation that ciliated and secretory cells appear before basal cells during development (McDowell *et al.*, 1985) has been used to counter suggestions that basal cells are airway progenitors. This argument appears less convincing in the light of recent speculation that adult stem cells are not necessarily the first cells present in a

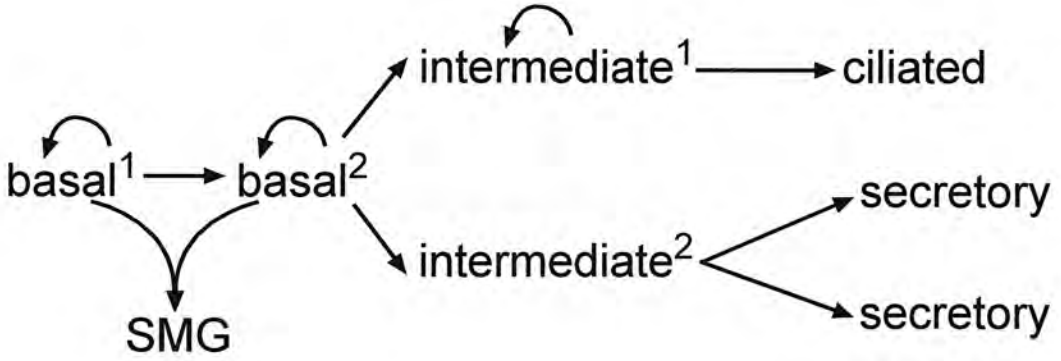


Figure 1.6 Lineage model of airway epithelial differentiation based on human bronchial xenograft experiments (modified from Engelhardt *et al*, 1995).

The model shows that two basal progenitors exist. Basal^1 is a stem cell from which all other cell types are derived, including the progenitor cell, basal^2. Submucosal glands (SMGs) were found to originate from more than one progenitor cell, and were always with complex clones of all four cell types, indicating an association between SMGs and multipotent progenitors. The model hypothesises that basal^2 cells must be activated to divide into intermediate^1 and intermediate^2 cells to initiate the process of SMG formation. The absence of surface epithelial clones containing (i) only intermediate, secretory, and ciliated cells (ii) two cell clones containing secretory cells and (iii) three cell clones containing basal, secretory and ciliated cells, can be explained by the model. While intermediate^1 cells divide to self-renew and produce ciliated cells, intermediate^2 cells can only divide to produce two secretory cells. The curved arrows indicate that only basal^1, basal^2, and intermediate^1 cells have the capacity to divide in culture and therefore integrate retroviral lineage markers.

developing tissue, but appear later to maintain the tissue during adulthood (van der Kooy and Weiss, 2000). The apparent morphological simplicity of basal cells has contributed to the notion that they represent "primitive" progenitor cells. However, electron microscopic analysis of basal cells from several species has revealed the presence of a complex system of junctional adhesion structures including keratin filaments, desmosomes, and hemidesmosomes (Evans *et al.*, 1989; Evans *et al.*, 1990). This indicated that basal cells do exhibit differentiated properties, and have a function in attaching columnar cells to the basal lamina.

1.4.3.4 Clara cells

Clara cells are columnar, non-ciliated, secretory cells, and the most common cell in the mouse trachea, constituting approximately 50% of cells, but diminishing slightly distally (Pack *et al.*, 1981). In this sense, the mouse trachea resembles the human bronchiole, in which approximately 35% of cells are Clara cells, though the human bronchiole has few if any basal cells (Mercer *et al.*, 1994). By contrast, Clara cells comprise only approximately 3% of human bronchial cells. Clara cells produce significant amounts of Clara cell secretory protein (CCSP, otherwise known as CC10), which has several functions, including an anti-inflammatory role (Magdaleno *et al.*, 1998). Several studies have implicated Clara cells in airway regeneration after injury. Keenan and co-workers observed that, following mild mechanical injury to the hamster tracheal epithelium, Clara cells mount the main proliferative response (Keenan *et al.*, 1982). Oxidant-induced injury specifically to ciliated cells (using NO_2), followed by [^{3}H]TdR labelling, revealed that the Clara cell was the primary progenitor cell giving rise to labelled ciliated cells, and that basal cells did not play a role in this process (Evans *et al.*, 1986). When denuded tracheal grafts were inoculated with highly purified basal or secretory rat tracheal epithelial cells, only the secretory fraction was able to reconstitute a complete epithelium: the basal fraction gave rise to an epithelium consisting only of basal and ciliated cells (Johnson and Hubbs, 1990).

Examination of normal human airway with the anti-Ki-67 proliferation marker revealed that Clara cells, which were found to be confined to the bronchioles, make a substantial contribution to airway maintenance (Boers *et al.*, 1999). While Clara cells constituted 11% and 22% of cells in the terminal and respiratory bronchioles, respectively, their corresponding contributions to the proliferative compartments of these regions were 15% and 44%. Clara cells contain P450 enzymes capable of catalysing the metabolic activation of carcinogens or toxicants (Magdaleno *et al.*, 1998). As discussed further in section 1.4.5.2, this property renders Clara cells specifically susceptible to ablation using naphthalene, permitting the analysis of progenitors cells active following depletion of the Clara cell compartment.

1.4.3.5 Pulmonary neuroendocrine cells

Pulmonary neuroendocrine cells (PNECs) are small, infrequent cells distributed throughout the airway epithelium as solitary cells or in clusters called neuroepithelial bodies (NEBs). These cells secrete a variety of neuropeptides that affect the growth and differentiation of airway epithelial cells, and may have a role in foetal lung development (Hoyt, Jr. *et al.*, 1993). PNECs can be identified using antibodies to one of the secreted neuropeptides, calcitonin gene-related peptide (CGRP) and, both independently and within NEBs, have been shown to undergo hyperplasia and hypertrophy in response to Clara cell ablation (Peake *et al.*, 2000). Evidence to suggest that PNECs may be a component of an airway stem cell niche is discussed in section 1.4.5.

1.4.4 Submucosal glands

SMGs consist of a network of secretory tubules and associated ducts, which act in the production and delivery of mucous and antimicrobial agents to the epithelium of cartilaginous airways. In humans, they are found in the trachea, bronchi and proximal bronchioles, lying most commonly between the epithelium and plates of supporting cartilage (Tos, 1966). The structure and composition of the glands is illustrated in Figure 1.7. The ciliated duct, which connects the airway luminal epithelium to the

collecting duct, is lined with ciliated cells and small numbers (approximately 6%) of another type of cell containing numerous large mitochondria and a well developed Golgi apparatus (Meyrick *et al.*, 1969; Meyrick and Reid, 1970). The collecting duct is lined with tall non-ciliated columnar cells and ends abruptly where it meets the secretory tubules, which are composed entirely of mucous and serous cells. The secretory tubules end in spherical acini composed of one or both cell types (Meyrick *et al.*, 1969).

SMGs are considered to be involved in the development of CF lung disease, for several reasons. They are the predominant region of cystic fibrosis transmembrane conductance regulator (*CFTR*) expression in the human bronchus, with serous cells and a subpopulation of "flask-like" cells in the ciliated ducts identified as the sites of expression (Engelhardt *et al.*, 1992). Histologic evidence of SMG luminal dilation is apparent in CF patients before any other signs of lung disease (Oppenheimer and Esterly, 1975), and hypertrophy and hyperplasia of SMGs is associated with disease progression (Sturgess and Imrie, 1982). When chloride transport in SMGs is blocked *in vitro*, mucous production is severely reduced, and mucous obstructions resembling those observed in CF occur (Inglis *et al.*, 1997).

Until recently, mice were considered to be "essentially void" of SMGs (Engelhardt *et al.*, 1992). Since then, the presence of SMGs in murine airways has been confirmed, where they exist at the same density as in the human airway, though in mice they are confined to the proximal trachea just below the larynx (Borthwick *et al.*, 1999) (Figure 1.8). Murine SMGs were observed to be similar in structure to human SMGs, and were found to produce mucous in both mucous and serous cells, and lysozyme and Cftr protein in serous cells. Interestingly, SMGs were noted to be more distally distributed in two mouse models of CF compared to congenic wild-type animals (Borthwick *et al.*, 1999), possibly reflecting the SMG hyperplasia observed in CF individuals (Oppenheimer and Esterly, 1975). Subsequently, significant variation in SMG distribution has been observed between inbred strains of mice, with glands extending more distally in BALB/CanN and FVB/N animals than in A/J and

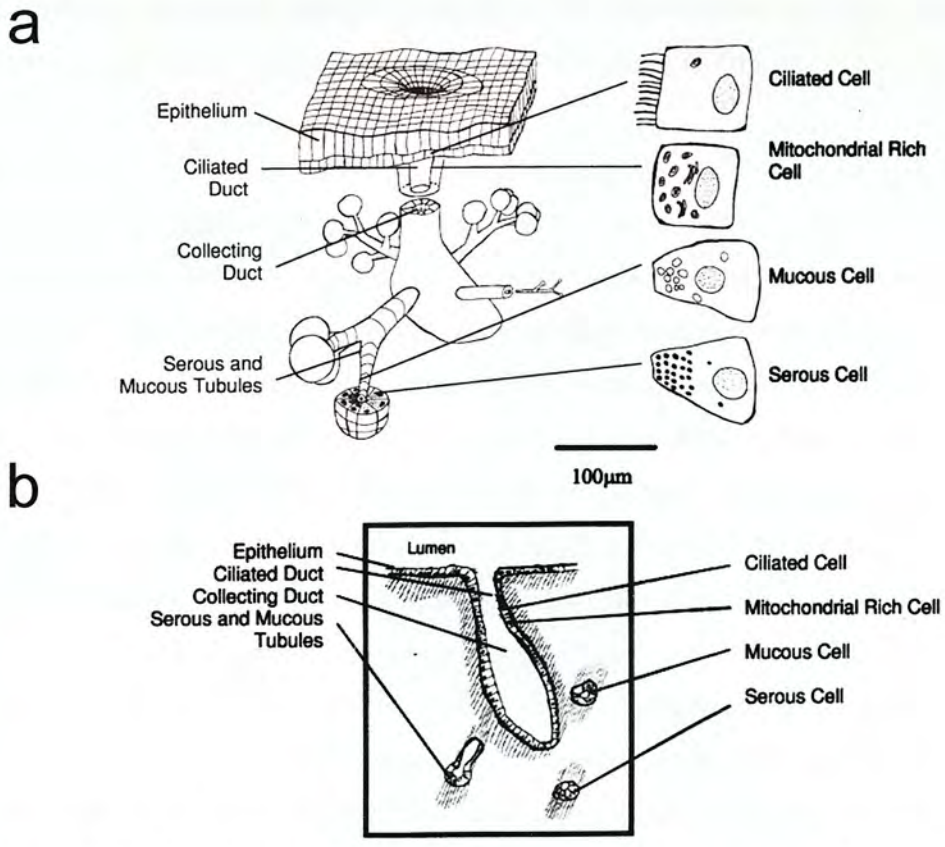


Figure 1.7 Structure and cellular composition of human submucosal glands (SMGs) (from Borthwick, 1999).

(a) Three dimensional representation of a human SMG based on the observations of Meyrick and colleagues (Meyrick *et al*, 1969). (b) Annotated illustration showing the appearance of a SMG in cross-section.

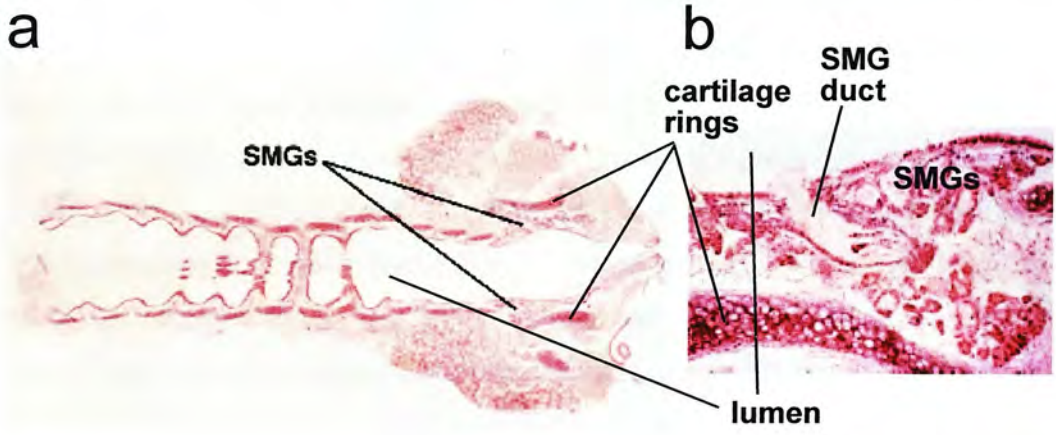


Figure 1.8 Location and appearance of murine submucosal glands (SMGs) (modified from Borthwick *et al*, 1999).

(a) Longitudinal section through a mouse trachea showing SMGs situated between proximal cartilage ring and the luminal surface. (b) Higher magnification of SMG region showing an SMG duct opening onto the luminal surface.

C57BL/6N animals. A genetic determinant on chromosome nine was found to be significantly linked to this phenotype (Innes and Dorin, 2001).

In humans, SMG development initiates during the 15th week of gestation and continues for approximately 10 weeks (Filali *et al.*, 2002). However, *de novo* production of SMGs is thought to be possible in adult airways, as suggested by apparent gland hyperplasia in CF patients (Oppenheimer and Esterly, 1975). As SMGs are thought to arise from buds derived from surface airway epithelium during foetal development (Thurlbeck *et al.*, 1961; Tos, 1968), it was thought that the epithelium of adult airways might contain progenitors capable of forming SMGs. Engelhardt and colleagues demonstrated SMG production by seeding retrovirally-labelled human bronchial epithelial cells into denuded rat tracheas, which were implanted subcutaneously into athymic mice (Engelhardt *et al.*, 1995). In addition to generating a fully differentiated surface airway epithelium, the isolated bronchial airway cells were able to develop SMGs. Labelled glands were commonly spatially associated with neighbouring labelled clonal clusters of airway surface epithelial cells. While heterogeneity in the composition of labelled clones was observed throughout the surface epithelium, clones associated with adjacent labelled SMGs were consistently found to contain all four of the cell types examined for (basal, secretory, intermediate, and ciliated). The same series of experiments, involving the simultaneous use of two different cellular labels, revealed that glands were often composed of differently-labelled cells, indicating the action of multiple progenitors in gland development. Likewise, partially transgenic glands were observed in experiments using a single cell label. The authors concluded that all glands formed from expansion of complex four cell type surface epithelial clones, and described a model involving a basal stem and progenitor cell to explain these observations, which incorporated the apparent absence of certain clone compositions (Figure 1.6).

Subsequently, analysis of mice chimaeric for detectable transgenes suggested SMGs were clonally derived (Borthwick *et al.*, 1999), contrary to the report of Engelhardt and colleagues (Engelhardt *et al.*, 1995). Also, the polyductal glands described by Engelhardt and colleagues (Engelhardt *et al.*, 1995) did not appear to be present in

mice (Borthwick *et al.*, 1999). The conflicting results obtained in these investigations could be the result of differences in the experimental system, for example mouse vs human, developmental vs adult SMG production, and *in vivo* vs xenograft context.

The high mobility transcription factor Lef1 has recently been shown to be expressed specifically in a subset of surface airway epithelial cells which give rise to gland-forming buds, in ferret airway xenografts (Duan *et al.*, 1998). *In situ* hybridisation revealed *Lef1* expression during the earliest stages of gland development, and within the distal tips of the primordial tubules for up to five weeks. While SMG development occurs postnatally in both the ferret (Leigh *et al.*, 1986) and the mouse (Borthwick, 1999), the ferret was used because it shows the greater similarity in SMG distribution and airway epithelial cell composition to the human. The impaired development of several organs including the kidney, teeth, and mammary glands in the Lef1 "knockout" mouse revealed the transcription factor is involved in inductive epithelial-mesenchymal interactions (van Genderen *et al.*, 1994), and can activate a variety of genes including keratins (Zhou *et al.*, 1995) and the cell adhesion molecule E-cadherin (Huber *et al.*, 1996). The absence of nasal and tracheal SMGs in Lef1-deficient mice, and an 8-fold decrease in SMG morphogenesis following antisense oligonucleotide inhibition of *Lef1* expression in ferret tracheal xenografts, revealed a requirement for Lef1 in the promotion of SMG development (Duan *et al.*, 1999). Also, expression of human *LEF1* from the CCSP promoter was shown to rescue the glandless phenotype of Lef1-deficient animals. Lef1 is thus considered to define a distinct subpopulation of surface epithelial cells with progenitorial capacity to form SMGs. However, the obligatory involvement of additional, undefined factors in SMG development was demonstrated by the inability of overexpressed *LEF1* to increase the extent of gland development in human bronchial xenografts, or when ectopically expressed in transgenic mouse airways. Interestingly, Lef/Tcf/β-catenin complexes have been identified as important extrinsic factors in controlling stem cell maintenance and/or commitment to differentiation in other epithelial tissues such as the epidermis and intestine (Korinek *et al.*, 1998; DasGupta and Fuchs, 1999).

1.4.5 Zonal organisation of airway epithelial stem cell systems

The recognised differences in structure and cellular composition between the various segments of the conducting airways (trachea, bronchus, bronchioles), and between the airways and the alveolar regions, have led to a general consensus that resident stem cell populations are likely partitioned between these portions of the respiratory system (Borthwick *et al.*, 2001; Engelhardt, 2001). The concept of local reserves of stem cells able to respond rapidly to injury is intuitive, and has been supported by most work to date (Otto, 2002), though alternative notions exist, including the concept of adult multipotent stem cells (discussed in section 1.4.6). Recent experiments aimed at identifying putative stem cells in the proximal (trachea) and distal (bronchiolar) airway have provided further evidence of separate progenitor populations existing in discrete niches both between and within the proximal and distal segments, while also revealing similarities in potential niche elements.

1.4.5.1 The proximal airway

As discussed previously, the slow-cycling nature of stem cells can be exploited in their identification, since they will retain a pulsed nucleotide label longer than the more proliferative TA cells. Borthwick and colleagues used this technique to identify slow-cycling cells within the mouse trachea (Borthwick *et al.*, 2001). Because of the low rate of cell division and turnover of cells in the steady-state airway, the tracheal epithelium was purposely injured to stimulate the proliferation of reparative stem cells, as cell division is required for incorporation of a nuclear label. The two methods used to damage the epithelium, inhalation of SO_2 and instillation of the detergent polidocanol, produced similar results, and were combined with repetitive administration of BrdU. Following repair of the epithelial surface, and an appropriate chase period, LRCs were found to be confined to the SMG ducts in the proximal trachea, and arranged in systematically arrayed foci along the surface epithelium in the glandless distal trachea. These foci of primarily basal cells appeared to be localised at cartilage-intercartilage junctions, where gland rudiments were occasionally found. The pattern of LRCs throughout the trachea was reminiscent of

the expression of *lacZ* from a human K5 promoter-driven transgene, which led to the discovery of cells rich in basal- and non-basal-specific keratins within the gland ducts. The prospect that differential expression of certain keratins within tissue sub-compartments may distinguish stem cells from their more differentiated progeny is discussed in chapter 2. Antibodies to CGRP revealed that PNECs appeared to be systematically distributed with LRCs in the glandless lower trachea, but that these phenotypes defined separate cell populations.

To further evaluate whether gland ducts may contain progenitors capable of regenerating the surface airway epithelium, tracheas in which the surface epithelium had been completely removed by protease digestion were implanted subcutaneously into athymic mice (Borthwick *et al.*, 2001). After 28 days, a surface epithelium resembling that present in mouse bronchi was present, supporting the notion of a stem or progenitor population within tracheal gland ducts.

1.4.5.2 The distal airway

Clara cells, the major TA population of the rodent conducting epithelium (Hong *et al.*, 2001), are susceptible to specific ablation using naphthalene. Using this technique, it had previously been shown that the proliferative response to bronchiolar Clara cell loss occurred in two subsets of cells within NEBs, and resulted in regeneration of the epithelium with associated hyperplasia of NEBs (Reynolds *et al.*, 2000a). These [^{3}H]TdR label-retaining candidate stem cell populations were found to be either CCSP-positive or CGRP-positive, (a pollution-resistant subset of Clara cells, or PNECs, respectively). In subsequent experiments employing a transgenic mouse expressing herpes simplex virus thymidine kinase from the CCSP promoter in all Clara cells, regeneration of the distal airway epithelium was monitored after the elimination of all Clara cells through administration of ganciclovir (Reynolds *et al.*, 2000b; Hong *et al.*, 2001). Despite the occurrence of PNEC proliferation and hyperplasia, indicating that CCSP-positive cells within NEBs were not necessary for NEB hypertrophy, no epithelial regeneration was observed suggesting that PNECs, while self-renewing, were unable to differentiate into Clara cells. These results

indicated the importance of a subpopulation of naphthalene-resistant CCSP-expressing cells associated with the NEB microenvironment in regenerating the bronchiolar epithelium after injury. The authors postulated that the NEB may represent an airway stem cell niche, though could not exclude the possibility that naphthalene-resistant CCSP-expressing cells represented an accessory cell essential for the proliferation and differentiation of a CGRP-positive (PNEC) progenitor. The same authors have since reported the existence of a putative, NEB-independent stem cell niche of pollutant-resistant CCSP-expressing cells situated at bronchoalveolar duct junction (BADJ), which serves to regenerate terminal bronchiolar epithelium after naphthalene injury (Giangreco *et al.*, 2002).

1.4.5.3 Airway stem cell niches?

The findings of Borthwick and colleagues (Borthwick *et al.*, 2001) and Hong and co-workers (Hong *et al.*, 2001) are summarised in Figure 1.9. These experiments, with the others, indicate that, even within anatomical segments of the airway, different lineage systems likely operate. In the trachea, putative stem cells appear localised to gland ducts proximally, and in regularly distributed foci near cartilage-intercartilage junctions more distally. Similarly, possible bronchiolar stem cells are generally located within NEBs, but a population also exists at the BADJ for renewal of the most distal terminal bronchioles. Although no hyperplasia of PNECs was detected by Borthwick and colleagues (Borthwick *et al.*, 2001), the association of PNECs, which are known to secrete neuropeptides thought to influence growth and differentiation of airway cells (Hoyt, Jr. *et al.*, 1993), with putative stem cell populations in both the proximal and distal airways, may be indicative of a role for PNECs in creating a stem cell-maintaining microenvironment. Given the characteristics of epithelial stem cell niches discussed in section 1.3.2, the SMG duct would seem a highly appropriate compartment, providing direct access to the overlying surface epithelium while remaining protected from inhaled toxins. By contrast, the bronchiolar stem cells resident in the epithelial surface appear to employ an intrinsic defence against environmental insult, as evidenced by their apparent lack of immunoreactivity with

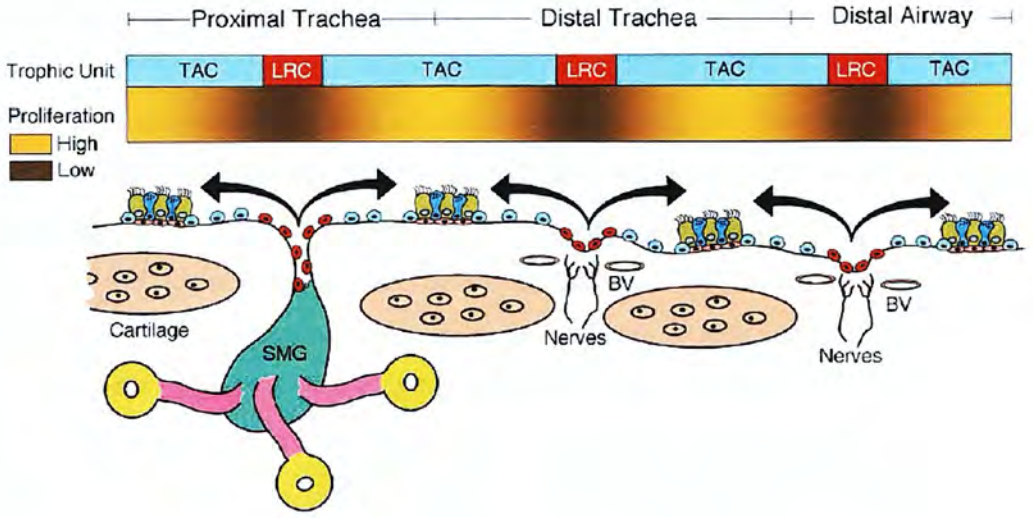


Figure 1.9 Schematic representation of stem cell niches in murine proximal and distal airways (after Borthwick *et al*, 2001 and Hong *et al*, 2001) (from Engelhardt, 2001).

For simplicity, two "trophic units" have been classified as label-retaining cells (LRCs, red), representing candidate slow-cycling stem cells, or highly proliferative transit amplifying (TA) cells (TAC, blue). Following injury, LRCs were found in submucosal gland (SMG) ducts and the immediate surrounding surface airway epithelium in the proximal glandular trachea. In the non-glandular distal trachea and distal bronchiolar airways, ablation of TA cells revealed LRC niches were found associated with innervated neuroepithelial bodies (NEBs) containing calcitonin gene-related peptide (CGRP)-expressing pulmonary neuroendocrine cells (PNECs, not shown). In the distal trachea, LRCs were mainly confined to inter-cartilaginous junctions associated with vascularised (BV) and innervated regions of the airway, while LRCs co-localised with NEBs in the non-cartilaginous bronchioles. NEBs, or their constituent PNECs, may serve as stem cell niches in the distal trachea and bronchioles.

antibodies against the cytochrome P450 isoenzyme responsible for producing the toxic metabolite of naphthalene (Giangreco *et al.*, 2002).

1.4.5.4 A single multipotent stem cell for the entire respiratory tract?

Despite general acceptance of the concept of region-specific stem cell niches in the respiratory epithelia, re-examination of earlier work employing incompletely pure "basal" and "secretory" cell isolates, which often led to conflicting results, and which was often hampered by a lack of consensus regarding the "intermediate" cell type, has led Emura to speculate that there is in fact a single multipotent stem cell for the entire respiratory system (Emura, 1997). Observations that a cloned epithelial cell line derived from a female foetal Syrian hamster lung could assume functional features of other respiratory cells types in response to manipulation of the *in vitro* environment was cited as support for the existence of CCSP-expressing cells with pre-Clara phenotypes as candidate stem cells for the whole airway (Emura, 2002). Engelhardt also speculated that, because Clara cells are present throughout murine airways, CCSP-expressing cells could represent a similar stem cell phenotype in both the proximal and distal respiratory tract (Engelhardt, 2001). This could also be relevant to the human airway: although morphologic counterparts to Clara cells are confined to human bronchioles, CCSP-expressing cells are also found in the bronchus (Engelhardt *et al.*, 1994). However, in a study correlating proliferation (using anti-Ki-67 antibody) with expression of CCSP in normal, hyperplastic and metaplastic human bronchial samples, CCSP-expressing bronchial cells demonstrated no proliferative activity and were thus considered to have no precursor function (Barth *et al.*, 2000).

1.4.6 How do current stem cell concepts apply to the airway epithelium?

Several studies demonstrating plasticity of BM-derived stem cells have revealed that the injured or irradiated lung is a site of frequent stem cell engraftment. In a study revealing contribution of a single BM-derived cell to liver, lung, gastrointestinal tract

and skin, Krause and colleagues labelled male BM-derived cells depleted of lineage markers and transplanted them into lethally-irradiated female recipients (Krause *et al.*, 2001). After 48 hours, donor-derived BM cells were found to be enriched for the HSC markers CD34 and SCA-1. Single cells were obtained by limiting dilution and transplanted into female lethally-irradiated hosts, where engraftment was detected by Y chromosome-specific fluorescence *in situ* hybridisation in the bronchi and pneumocytes of the respiratory epithelium. Anti-keratin antibodies were used to demonstrate that the engrafted cells were appropriately differentiated, and not haematopoietic cells. High-level engraftment in the lung was thought to be a result of its particular susceptibility to radiation damage. In a subsequent study, radiation induced pneumonitis was shown to promote engraftment of unfractionated BM as type II pneumocytes (Theise *et al.*, 2002). In some cases entire alveoli were donor-derived, suggesting that type I pneumocytes differentiated from type II pneumocytes. The prospect for genetic modification of donor cells was demonstrated by Grove and co-workers, who retrovirally transduced enhanced green fluorescent protein (eGFP) into BM-derived cells prior to serial transplantation in irradiated recipients (Grove *et al.*, 2002). eGFP-expressing type I and type II pneumocytes were detected in recipients up to 11 months later, comprising an average of 3% of all alveolar cells in the lung tissue analysed.

Using bleomycin-induced oxidative injury, rather than irradiation, of the respiratory epithelium, Kotton and colleagues observed donor-derived type I, but not type II, pneumocytes (Kotton *et al.*, 2001). Although donor-derived cells were found in clusters, they did not express proliferation-specific antigens. A low level of engraftment was observed in control treated mice, indicating that injury is not an absolute prerequisite for contribution from BM-derived stem cells. Differences in the overall frequency and distribution of engrafted cells between this study and that of Krause and colleagues (Krause *et al.*, 2001) likely resulted from several differences in the experimental protocol, including mode of lung damage, presence or ablation of host BM, and the nature of the transplanted cells: while Krause and co-workers used uncultured BM-derived cells (Krause *et al.*, 2001), Kotton and colleagues transplanted a selectively expanded plastic-adherent population (Kotton *et al.*, 2001).

Despite the demonstrated ability of BM-derived, pluripotent stem cells to contribute to the regeneration of the injured respiratory system, the involvement of endogenous BM-derived stem cells in steady-state tissue maintenance, and whether such cells enter the local niche and adopt a tissue-specific stem cell phenotype in advance of contributing to the tissue, has yet to be clearly established. Regardless of stem cell origin, defined compartments of stem, TA, and TD cells have been identified in several epithelial tissues, and most cell turnover within a tissue is likely to result from transit through these populations. Thus, investigation of cell lineage patterns in adult tissues remains an appropriate means of locating and identifying tissue-specific stem cells.

1.5 Aims

1.5.1 Investigation of a putative stem cell niche in murine tracheal submucosal gland ducts

It has previously been observed that LRCs within the SMG ducts of the proximal trachea, possibly representing slow-cycling stem cells, were spatially coincident with cells expressing high levels of keratin protein (Borthwick *et al.*, 2001). That the pattern of keratin protein was reflected by the activity of a K5 promoter-driven reporter transgene provided the opportunity to use the promoter to investigate lineage arising from the SMG duct in response to injury. The aim of this project was to use a transgenic approach to determine whether K5 expression could be used as a marker to verify the presence of a stem cell niche within SMG ducts which was capable of contributing to repair of the airway epithelium.

1.5.2 An in vivo model for gene correction

The positive identification of stem cells in any tissue raises the prospect of being able to specifically target them in gene therapy protocols. In principle, this should facilitate a single-dose application, with propagation of treated cells throughout the tissue as a result of normal tissue renewal or in response to induced injury. Gene correction is a relatively new concept in gene therapy, and describes the permanent, site-specific correction of chromosomal defects by a range of small oligonucleotide-based reagents. Comparative evaluation of different gene correction strategies *in vivo* is difficult using existing systems, such as mouse disease models. The aim of this project was to generate a transgenic mouse designed to permit the simple and rapid evaluation, development, and optimisation of *in vivo* gene correction protocols.

Chapter 2 A transgenic mouse for inducible tracheal lineage analysis

2.1 Introduction

The observation that the human keratin (K) 5 promoter directed expression of a *lacZ* reporter transgene to a subset of keratin-rich cells within murine submucosal gland (SMG) ducts ultimately led to the discovery of slow-cycling label-retaining cells (LRCs) at this location (Borthwick *et al.*, 2001). In the K5-*lacZ* mice, βgal activity was visible only in a circumferential band in the upper trachea, in addition to infrequent small patches distributed along a descending posterior stripe. This pattern was consistent with the location of the tracheal glands, and upper tracheal sections revealed βgal positive basal cells scattered in the surface epithelium, and more frequently in small groups within SMG ducts. Further analysis using anti-keratin antibodies indicated a correlation with high-level keratin protein production. Both basal-specific K14 (the co-expressed binding partner of K5), and K18, which is otherwise expressed in supra-basal, columnar surface cells in the trachea, were found to be expressed in basal gland duct cells.

2.1.1 Keratins as stem cell markers

Keratins are the intermediate filament proteins of epithelial cells, and are subdivided into two groups, type I and type II, on the basis of their amino acid composition, either acid or neutral-basic, respectively (Moll *et al.*, 1982). Type I keratins (K9-20) are co-expressed in pairs with type II keratins (K1-8), the proteins forming obligate heterodimers which constitute cytoskeletal subunits. These keratin pairs are differentially expressed in epithelial tissues at various stages of development and differentiation, and their cell type specificity is such that they are frequently used as markers for epithelial cell identification. For example, K8/K18 are expressed in simple epithelia, K1/K10 are expressed in the supra-basal layers of the epidermis, K4/K13 are expressed supra-basally throughout internal stratified epithelia, while K6/K16 are found in the supra-basal cells of the hair follicle outer root sheath (Byrne and Fuchs, 1993). Differentiating cells of the cornea express K3 (Schermer *et al.*, 1986), together with its binding partner K12, and K15 is expressed mainly in internal

epithelia during neonatal development, where it dimerises with K5 (Lloyd *et al.*, 1995).

Earlier studies had indicated that the differential expression of a keratin within the cells of an epithelial tissue might distinguish stem cells from their more differentiated progeny. Schermer and colleagues used a monoclonal antibody to show that the basal cells of the corneal limbal region were negative for a 64K keratin protein (K3), while limbal supra-basal cells, and both the basal and supra-basal epithelia of the central cornea, expressed the protein (Schermer *et al.*, 1986). The authors interpreted this to suggest that the basal cells of the central cornea were more differentiated than their limbal counterparts, and likely represented the transit amplifying (TA) population which gave rise to the supra-basal layer. Evidence that limbal basal cells constitute the corneal stem cell compartment was provided by a label-retaining assay, in which corneal stem cells were stimulated to proliferate in response to induced corneal wounding in the presence of a labelled nucleotide analogue. Significant numbers of LRCs, representing putative stem cells, were found in the limbal region, while none were detected in corneal epithelium (Cotsarelis *et al.*, 1989). Clonal analysis in chimaeric adult corneas has demonstrated directly that marked cells arise in the basal limbal epithelium and migrate centripetally towards the centre of the cornea (Collinson *et al.*, 2002).

More recently, LRCs were localised to the bulge region of human hair follicles, for which K15 was found to be a specific marker. K19, which had previously been purported to be a stem cell marker for murine bulge LRCs (Michel *et al.*, 1996) was also found to be expressed in the human bulge region, though K19 expression extended downward within the basal cell layer towards the bulb. The authors speculated that, while K15 defined the undifferentiated cells of the LRC-containing bulge region, K15-negative/K19-positive cells might represent the early TA population of the human hair follicle (Lyle *et al.*, 1998). Interestingly, K15 is an alternative binding partner for K5 (Lloyd *et al.*, 1995), which normally pairs with K14 in basal epithelia. K5 has also been found to be expressed within the stem cell-containing basal limbal region of cornea (Byrne and Fuchs, 1993).

2.1.2 Keratin 5 and the human K5 promoter

K5 is normally associated with K14, which together are the most broadly expressed keratins, and are found in mitotically active basal cells in all stratified epithelia, regardless of anatomical location (Nelson and Sun, 1983). In the epidermis, K5/K14 expression in basal cells switches to K1/K10 as cells become committed to terminal differentiation and migrate supra-basally. Given the apparent insensitivity of K5/K14 expression to the varied environments in which basal keratinocytes are found, and the fact that the marked restriction of K5/K14 expression to basal epithelial layers appeared to be mediated at the transcriptional level (Lersch *et al.*, 1989; Stellmach *et al.*, 1991), their respective promoters have been used as tools to study cell type-specific regulation of gene expression (Leask *et al.*, 1990; Byrne and Fuchs, 1993; Byrne *et al.*, 1994). While both human promoters have been shown to drive *in vivo* expression of reporter genes in a manner faithful to their respective endogenous counterparts (Vassar *et al.*, 1989; Byrne and Fuchs, 1993), the K14 promoter, which was characterised first (Marchuk *et al.*, 1985), or bovine K5 promoter sequence (Ramirez *et al.*, 1994), have tended to be used when basal epithelial-specific transgene expression has been required (Cheng *et al.*, 1992; Brakebusch *et al.*, 2000). Also, one group has used a 14kb human K5 promoter (Tarutani *et al.*, 1997). The following discussion is concerned only with the 6kb human K5 promoter fragment isolated and characterised by Fuchs and colleagues (Lersch *et al.*, 1989; Byrne and Fuchs, 1993).

The activity of the human K5 gene promoter has been analysed both *in vitro* and *in vivo*. When 6kb of the human K5 promoter was used to drive expression of reporter genes after transient transfection into a range of cell lines, activity was detected in a human epidermal keratinocyte line, but not in hepatocytes or fibroblasts (Byrne and Fuchs, 1993). This suggested that 6kb of upstream sequence was sufficient to confer cell type-specific expression *in vitro*, and agreed with the results of previous experiments in which considerably shorter sequences were shown to drive keratinocyte-preferred expression (Jiang *et al.*, 1991; Ohtsuki *et al.*, 1992).

In order to evaluate the specificity of the 6kb sequence *in vivo*, transgenic mice were generated carrying the *lacZ* reporter under the transcriptional control of the K5 promoter, and the pattern of βgal activity was compared with that of endogenous K5 protein, as detected using anti-K5 antiserum (Byrne and Fuchs, 1993). In two similarly expressing transgenic lines, βgal activity in most tissues was found in a pattern indistinguishable from that of endogenous K5, including the basal layer of tail skin epidermis (including the sebaceous and eccrine glands) and tongue, and the corneal limbus. The profile of βgal deviated from that of endogenous protein in the oesophagus, forestomach, and thymus, in which only a small minority of basal cells expressed the transgene. The existence of keratin-positive basal cells lacking the corresponding mRNAs had previously been reported (Schweizer *et al.*, 1988), and the authors speculated that the shorter half-life of βgal compared to the durable K5 filaments may explain some of these observed disparities. Unanticipated expression in non-epithelial tissues was confined to a subset of neurons in the brain, where no K5 protein was detected. Parallel experiments revealed that while a truncated 90bp K5 promoter maintained keratinocyte-specific expression of βgal, it was not sufficient to restrict transgene expression to basal epithelial layers, implying the presence within the remaining 5.9kb of elements controlling basal specificity and suppressing supra-basal expression (Byrne and Fuchs, 1993). Subsequently, the profile of K5 promoter-driven βgal expression in developing epidermis was characterised in detail (Byrne *et al.*, 1994), and their observations are briefly described in section 2.2.4.

Given the described precedents for the presence or absence of certain keratins to define stem cell compartments in some epithelial tissues, it is conceivable that the observed co-localisation of LRCs in the mouse tracheal glands to regions containing cells expressing high levels of K5, might indicate that K5 could be a specific marker for these putative stem cells.

2.1.3 Lineage analysis

Putative tissue-specific stem cells have been identified in a range of tissues by their retention of a pulsed label following an appropriate chase period (Bickenbach, 1981; Cotsarelis *et al.*, 1989; Cotsarelis *et al.*, 1990), indicative of a slow-cycling phenotype, which is regarded as a characteristic of epithelial stem cells. However, a stem cell can only truly be defined on the basis of its cellular progeny. Since candidate stem cells should ideally be analysed in their native context, to prevent potentially artefactual observations arising from their manipulation in a foreign environment, strong evidence to support a claim of stem-ness can be obtained by monitoring cell lineage relationships either during development, or in steady-state or regenerating adult tissue. This necessarily requires that the putative stem cell population be specifically and detectably tagged with a marker which is transmitted to its cellular descendants. Retrospective clonal analysis has been achieved in the mouse myotome (Bonnerot and Nicolas, 1993) and brain (Mathis and Nicolas, 2000) using an inactivated *lacZ* reporter carrying an internal sequence duplication ("*laacZ*") (Sanes, 1994). Spontaneous intragenic recombination occurred infrequently (10^{-6}) (Bonnerot and Nicolas, 1993), reconstituting functional *lacZ* sequence which was transmitted to cellular progeny. *lacZ* encodes the histochemically-detectable β-galactosidase enzyme (βgal), allowing cell lineages to be unambiguously visualised.

One major limitation of approaches in which expression of the reporter is under the direct control of a promoter active within the putative progenitor population, is that the progeny of these cells may not maintain expression of the reporter as they differentiate. In recent years, techniques have been developed which permit tissue-specific activation of a ubiquitously expressed reporter. In this way, initiation of lineage marking can be confined to a subset of cells expressing a common promoter, but cell fate can be tracked outside this compartment, regardless of the final phenotype or destination of the daughter cells. Such approaches require the use of site-specific recombinases, such as the Cre recombinase, an integrase protein encoded by the bacteriophage P1 (Sauer and Henderson, 1988). Cre catalyses

recombination between 34bp target sequences called loxP sites, which are composed of a core 8bp sequence and two 13bp palindromic flanking sequences. When two loxP sites are arranged in *cis* in the same orientation, as defined by the core sequence, Cre-mediated recombination between them results in the precise excision of intervening (flanked by loxP, "floxed") DNA (Tronche *et al.*, 2002). Since the first demonstrations that Cre could be used to effect site-specific modifications within the mouse genome (Lakso *et al.*, 1992; Orban *et al.*, 1992), numerous transgenic lines have been created, expressing Cre from a variety of promoters for a range of purposes ((Lobe and Nagy, 1998; Nagy, 2000) for review). In many such experiments, the Cre-expressing line is crossed with another transgenic line in which the gene of interest has been modified through gene targeting to include a pair of loxP sites such that, in the presence of Cre protein, part or all of the gene is deleted from the genome. These conditional knockouts, in which gene inactivation is restricted to tissues expressing the *Cre* transgene promoter, have several advantages over conventional knockouts, including the ability to bypass embryonic lethal phenotypes (Tarutani *et al.*, 1997).

When a novel Cre-expressing line is generated, it is usual to validate its stage- and tissue-specific recombination profile by crossing with a Cre reporter line in which the pattern of Cre activity is represented by excision-dependent reporter gene activation. Several such lines exist, though all are similar in design, and are comprised of a reporter gene separated from a ubiquitous promoter by a floxed cassette containing transcriptional and/or translational termination signals (Figure 2.1). Ubiquitous expression of the reporter gene is dependent on Cre activity. The floxed cassette can include a selectable marker (Araki *et al.*, 1995; Soriano, 1999; Mao *et al.*, 1999) or a second reporter gene (Lobe *et al.*, 1999; Novak *et al.*, 2000). In examples of the latter, the Cre reporter line expresses βgal prior to Cre-mediated recombination, and human placental alkaline phosphatase (hPAP) (Lobe *et al.*, 1999) or enhanced green fluorescent protein (eGFP) (Novak *et al.*, 2000) following excision of the *lacZ* sequence, permitting discrimination between a lack of Cre excision and the absence of reporter expression. Despite the apparent advantages of the double reporter system, the most widely used Cre reporter remains the *R26R*

mouse, in which a floxed neomycin phosphotransferase (NEO) cassette including a triple polyadenylation sequence was positioned upstream of *lacZ*, and targeted to the ubiquitously active *ROSA26* locus (Soriano, 1999). The *ROSA26* locus was identified in a promoter trap experiment, where it was observed to drive ubiquitous and robust expression of a retrovirally integrated *lacZ*-NEO fusion gene throughout development (Friedrich and Soriano, 1991). Subsequent characterisation of the chromosome 6 locus revealed that the provirus trapped, and thus eliminated expression of, two non-coding transcripts, but did not affect the expression of a third, highly conserved transcript arising from the complementary strand (Zambrowicz *et al.*, 1997). *ROSA26* mice have been widely used when ubiquitous expression of βgal is required, for example in aggregation chimaeras (Borthwick *et al.*, 1999) and cell transplantation experiments (Jiang *et al.*, 2002). In the *R26R* Cre reporter line, robust expression of βgal is activated in Cre-expressing cells, and is maintained in all generations of daughter cells as a result of the permanence of the genetic alteration and the ubiquity of expression. Thus, transgenic lines used for evaluating expression in novel, tissue-specific Cre lines are ideal tools for monitoring lineages arising from Cre-expressing populations. This approach has been used to fate map the mid-hindbrain border (Zinyk *et al.*, 1998), neural-crest derivatives (Chai *et al.*, 2000; Jiang *et al.*, 2000; Epstein *et al.*, 2000), pancreatic islet cells (Herrera, 2000) and memory T cells (Jacob and Baltimore, 1999).

2.1.4 Temporally controlled activation of tissue-specific lineage marking

Although the combination of tissue-specific Cre expression with ubiquitous but excision-dependent reporter activity is sufficient for many lineage studies, the activation of the reporter at the onset of expression from the *Cre* transgene promoter could be problematic in certain circumstances. For example, analysis of cell lineage patterns in adult tissues could be complicated or rendered futile if the *Cre* transgene had been expressed, and thus the reporter activated, in cells of the same tissue earlier in development. Consequently, the utility of this strategy for adult fate mapping has been greatly enhanced by overlaying the tissue-specificity, conferred by the promoter

of the *Cre* transgene, with temporal control. This has been achieved by fusing the *Cre* open reading frame with the ligand-binding domain (LBD) of the human (Indra *et al.*, 1999) and mouse (Danielian *et al.*, 1998) oestrogen receptors (ERs). The LBD acts by preventing the Cre portion of the fusion protein from interacting with the nuclear DNA through heat shock protein (HSP)-mediated cytoplasmic sequestration (Scherrer *et al.*, 1993). In the presence of exogenous inducer, a synthetic oestrogen agonist, the interaction with HSPs is disrupted and the fusion protein translocates to the nucleus where it is able to effect site-specific recombination (Figure 2.1).

The principle of conferring regulation on protein function through the use of nuclear receptor fusions had previously been established (Picard, 1994) when the first fusion of Cre with the human ER (CreER) was shown to mediate oestradiol- or 4-hydroxytamoxifen (OHT, a synthetic oestrogen analogue)-induced recombination of a floxed selectable marker in cultured cells (Metzger *et al.*, 1995). The authors recognised that the sensitivity of the system to endogenous oestrogens would undermine its use *in vivo*, so the first inducible Cre mouse carried the CreER^{T} fusion, containing a mutation predicted to abolish activation by oestradiol but retain binding to synthetic ligands such as tamoxifen (TAM) and its metabolite OHT (Danielian *et al.*, 1993). The *CreER^{T}* transgene was shown to mediate OHT-induced recombination of a floxed allele *in vivo*, with no uninduced activity detectable by polymerase chain reaction (PCR) for the presence or absence of the floxed cassette (Feil *et al.*, 1996). In mice expressing a similar transgene (*CreER^{TAM}*), in which Cre was fused with the murine ER containing the equivalent mutation to that in *CreER^{T}*, and driven by the K14 promoter, administration of TAM to adult mouse skin efficiently activated the *R26R* reporter in the basal epidermis, including stem cells (Vasioukhin *et al.*, 1999). When a comparable construct (*CreER^{TM}*) was expressed from the *Wnt1* promoter, *in utero* administration of TAM caused approximately 25% activation of a βgal reporter in the embryonic nervous system. However, these experiments revealed that the high concentration of TAM required to activate CreER^{TM} was close to levels that interfere with pregnancy (Danielian *et al.*, 1998).

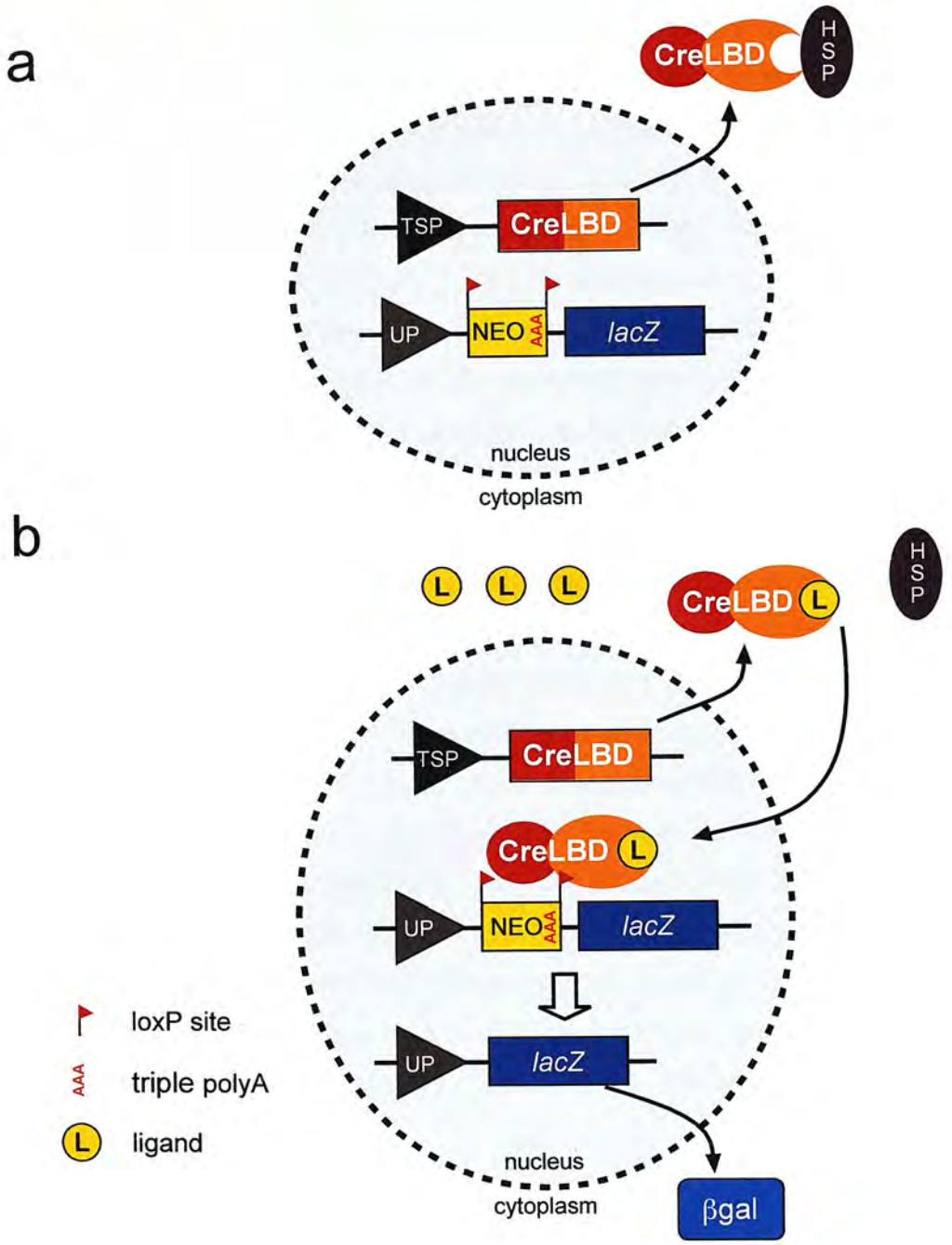


Figure 2.1 Schematic showing ligand-induced recombinase-mediated activation of ubiquitous reporter expression.

In the example shown, the cell carries two transgenes: a ligand (L)-inducible Cre recombinase transgene (*CreLBD*) driven by a tissue-specific promoter (TSP), and a ubiquitous promoter (UP)-driven Cre-dependent reporter gene, in which transcription of *lacZ* is prevented by the triple polyadenylation (polyA) sequence at the 3' end of the upstream floxed NEO cassette. (a) Under normal circumstances, CreLBD protein is produced in cells expressing the tissue-specific promoter, but is cytoplasmically sequestered by heat-shock proteins (HSP). Transcription of the Cre-dependent reporter gene does not proceed into the *lacZ* sequence. (b) In the presence of exogenous ligand, the interaction between CreLBD and the HSP is disrupted, promoting nuclear translocation of the fusion protein, where the Cre recombinase causes excision of the floxed NEO cassette. In the absence of the transcriptional termination sequence, βgal protein is ubiquitously produced as a result of *lacZ* transcription from the ubiquitous promoter.

It had previously been noted by Schwenk and colleagues that the mutation used to preferentially lower the sensitivity of their Cre^{EBD} construct to endogenous ligands, identical to the mutation used by Metzger and co-workers ($CreER^{T}$) (Metzger *et al.*, 1995), and equivalent to the murine LBD mutation used by Danielian and colleagues ($CreER^{TM}$) (Danielian *et al.*, 1998), also caused a 100-fold reduction in affinity to synthetic agonists (Schwenk *et al.*, 1998). To address this, the original human ER mutation was replaced by three new mutations, and the resulting $CreER^{T2}$ exhibited an approximately 4-fold increase in affinity to OHT compared with $CreER^{T}$ (Feil *et al.*, 1997). *In vivo*, the difference in sensitivity to OHT between $CreER^{T}$ and $CreER^{T2}$ when expressed from the bovine K5 promoter, both in terms of nuclear translocation and recombinase activity, was shown to be approximately 10-fold (Indra *et al.*, 1999). This experiment was elegant in demonstrating induced recombination in proliferative basal epidermal cells, detected after supra-basal differentiation into cells of the granular epidermal layer which specifically expressed the Cre reporter transgene. $CreER^{T}$ and $CreER^{T2}$ have since been used to effect temporally-regulated site-specific recombination in a range of tissues including skin (Brocard *et al.*, 1997), liver (Imai *et al.*, 2001b), nervous system (Doerflinger *et al.*, 2003), internal epithelia (Wen *et al.*, 2003), adipocytes (Imai *et al.*, 2001a), and brain (Weber *et al.*, 2001). While most of these studies did not involve lineage analysis *per se*, many of them were proof of principle experiments for novel promoter-driven Cre fusion constructs, necessarily using Cre reporter lines to evaluate patterns of induced expression, and thereby demonstrating the potential of the technique for the temporal activation of lineage marking.

One example of the use of an inducible Cre transgene for fate mapping was demonstrated by Kimmel and colleagues, who used *in utero* administration of TAM at various stages of embryonic development to characterise the location and dynamics of the multiple borders influencing apical ectodermal ridge formation (Kimmel *et al.*, 2000). Given the extended time period over which the promoter driving $CreER^{T}$ expression was active, this experiment would have been significantly less informative had the authors not been able to control the time at which lineage marking was initiated.

In addition to the CreER fusions described, a mutated progesterone receptor (PR) sensitive to the synthetic steroid RU 486, has been used to regulate Cre activity (Kellendonk *et al.*, 1996). The CrePR1 fusion was shown to induce site-specific recombination in mouse brain, albeit with a low level of uninduced recombinase activity. In cell culture, it was shown to perform with a similar efficiency to ER-regulated Cre constructs (Kellendonk *et al.*, 1999). Inducible recombination has since been demonstrated in the skin (Arin *et al.*, 2001; Zhou *et al.*, 2002) and heart (Minamino *et al.*, 2001), though the CrePR1 fusion has not been as widely used as the various CreER constructs.

2.1.5 Aims

The observation that LRCs identified in murine SMG ducts after tracheal epithelial regeneration were spatially coincident with keratin-rich cells expressing a K5 promoter-driven *lacZ* transgene (Borthwick *et al.*, 2001), suggested that K5 may be a marker for this putative stem cell population. If this were the case, then analysis of lineage arising from this K5-expressing population in response to epithelial injury would directly reveal whether a relationship existed between these putative stem cells and the differentiated cells of the tracheal epithelium. To this end, transgenic mice were created which carried the $CreER^{T2}$ inducible *Cre* fusion gene (Feil *et al.*, 1997) under the control of the human K5 promoter. While mice expressing the $CreER^{T2}$ transgene from the bovine K5 (bK5) promoter exist (Indra *et al.*, 1999), the pattern of bK5 promoter-directed transgene expression in the trachea has not been characterised. Since the aim was to recapitulate the human K5 promoter-driven tracheal expression described by Borthwick and colleagues (Borthwick *et al.*, 2001), the $CreER^{T2}$ transgene was cloned downstream of the human K5 promoter (Byrne and Fuchs, 1993).

The transgenic mouse was generated with the intention that administration of exogenous ligand would activate permanent lineage marking, by means of a ubiquitous conditional reporter transgene, from K5 promoter-expressing cells. An inducible Cre recombinase was used to prevent lineage marking occurring as a result of possible developmental or steady-state tracheal expression from the K5 promoter, which could potentially lead to uninterpretable patterns of marked cells. Mice carrying a K5 promoter-directed *lacZ* reporter transgene were created in parallel to verify appropriate expression from the subcloned K5 promoter.

This chapter describes (1) the construction of both transgenic lines, (2) the analysis of βgal expression directed by the K5 promoter, and (3) the identification of inducible Cre lines potentially useful for marking the lineage of the K5 promoter-expressing SMG duct cells. A pilot experiment employing one of the inducible Cre lines is described in chapter 3.

Results

The K5 promoter-driven *lacZ* (*K5A1*) and inducible Cre (*K5CE*) constructs, and the cloning strategies used to create them, are shown schematically in Figure 2.2. Briefly, both constructs contain the human K5 promoter subcloned into a promoter-less vector containing an internal ribosome entry site (IRES)-enhanced green fluorescent protein (*eGFP*) sequence. The IRES-*eGFP* was included with the intention that it would provide a simple means of detecting transgene expression. In *K5A1*, *lacZ* was positioned 3' of the K5 promoter, with the IRES-*eGFP* component further downstream. The intron which was positioned 3' of the K5 promoter in the parental pK5LH14 plasmid was incorporated into the *K5A1* construct between the K5 promoter and *lacZ*. To make *K5CE*, a fragment containing the inducible Cre fusion gene $CreER^{T2}$ and an upstream intron, was subcloned 3' of the K5 promoter in the pK5-eGFP intermediate from *K5A1* construction. The *K5A1* mouse was used as a simple means to confirm appropriate expression from the subcloned K5 promoter prior to use of the *K5CE* mouse, designed for inducible lineage analysis in the adult trachea.

The DNA manipulations involved in the generation of these constructs are described in detail in appendix A1. The intermediates in vector construction are named in Figure 2.2 to allow referencing.

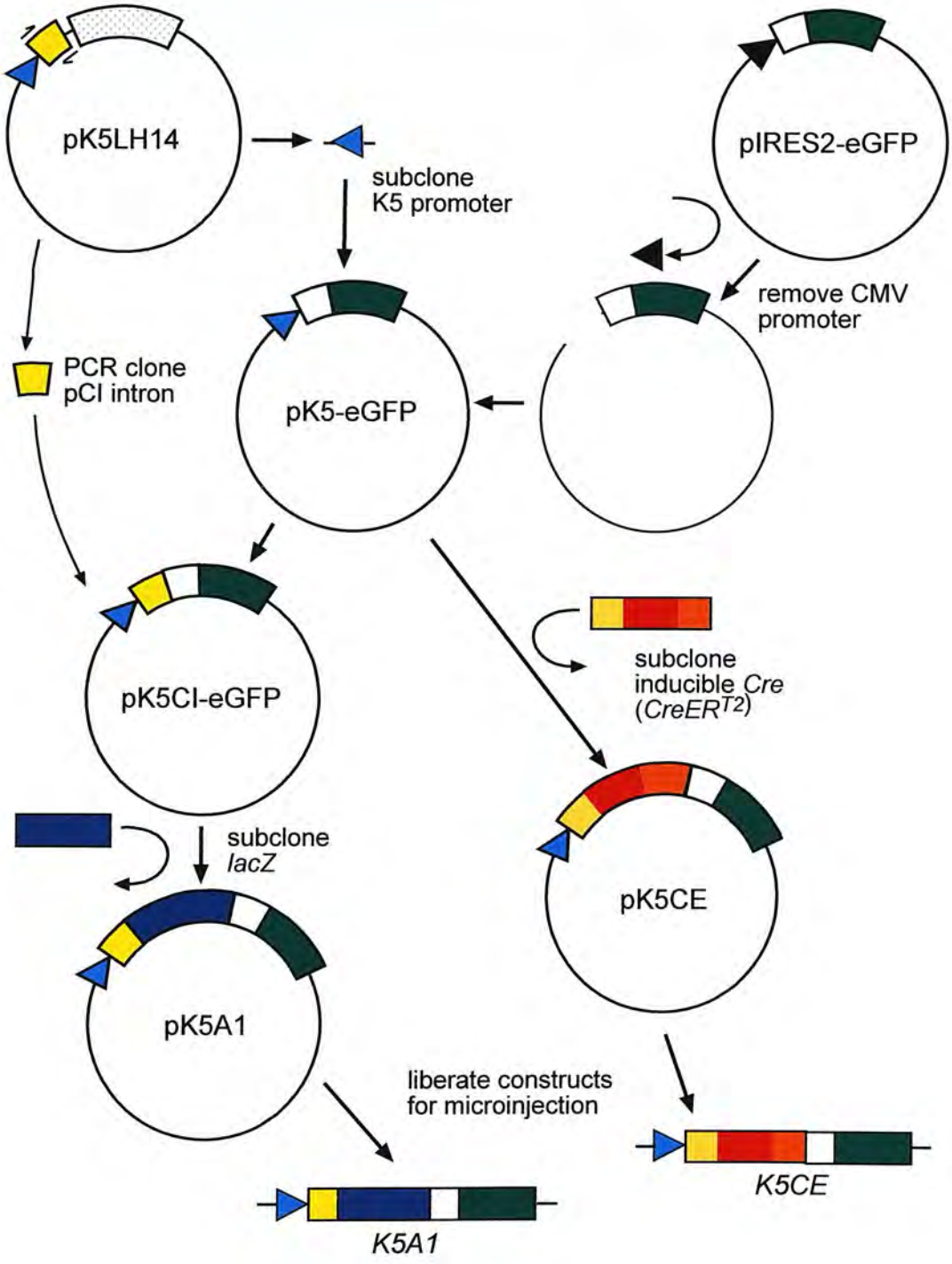


Figure 2.2 Cloning strategy to generate *K5A1* and *K5CE* transgenic mice.

Construct elements are represented by the following: CMV promoter (black triangle); K5 promoter (blue triangle); pCI intron (yellow rectangle); *lacZ* gene (blue rectangle); IRES-*eGFP* (white-green rectangles); *CreER*^{T2} fusion gene with 5' intron (tan-red-orange rectangle). Intermediates in vector construction are named to allow reference in main text.

2.2 The transgenic K5 promoter-driven *lacZ* control line, *K5A1*

The DNA manipulations required to generate the *K5A1* construct are described in appendix A1.1-3.

2.2.1 Microinjection of constructs and identification of transgenic founders

In the absence of available cell lines known to express K5 for *in vitro* validation of the construct, such as the human squamous cell carcinoma line SCC-13 (Wu and Rheinwald, 1981; Byrne and Fuchs, 1993), pK5A1 was linearised using BspHI to cut twice in the pIRES2-eGFP-derived backbone, and microinjected into F1 fertilised oocytes as described in chapter 7 section 7.6. The microinjection sessions are summarised in Table 2.1. Genomic DNA prepared from ear-clip biopsies was used as template for genotyping. A transgene-specific PCR, using primers based on *lacZ* sequence, was combined with a PCR which produced a smaller product from the genomic *myogenin* (*myoG*) gene, the latter serving as a control to ensure the quality of the DNA preparations for PCR (Figure 2.3a,b).

Additional PCRs were conducted on founders identified by the *lacZ*-specific reaction in order to check the 5' and 3' integrity of the transgenes, and to ensure that genomic DNA preparations were not contaminated with the plasmids from which the transgenes were derived. As illustrated in Figure 2.3a, an upstream-oriented primer ("5'K5") within the K5 promoter at the 5' end of the construct could generate two PCR products of different length depending on whether it was used in conjunction with primers "5'pos" or "5'neg". 5'pos lay in flanking vector sequence upstream of the K5 promoter, but within the construct as defined by the BspHI sites, while 5'neg was situated in the backbone fragment removed prior to microinjection. An equivalent situation existed at the 3' end of the construct, with a primer in eGFP ("3'eGFP") opposing primers "3'pos" and "3'neg". Generation of a product from PCRs of founder DNA using primer pairs 5'K5 - 5'pos and 3'eGFP - 3'pos indicated

Construct, Session date	Total no. embryos	Total no. injected	Total no. cultured	Total no. transferred	Total no. born	Total no. transgenic
K5A1						
05/10/01	116	76	57	34	0	-
08/10/01	150	79	52	46	0	-
05/12/01	187	126	121	98	17	4
Total	453	281	230	178	17	4
K5CE						
?	?	?	?	?	10	4
15/10/01	200	121	76	69	7	1
01/11/01	323	209	137	90	15 (4 died)	4
05/11/01	322	166	142	104	16	7
12/11/01	269	168	147	131	17	5
15/11/01	121	70	67	62	4 (1 died)	3
19/11/01	149	95	83	72	6	5
Total	>1384	>829	>652	>528	70	29

Table 2.1 Summary of microinjection sessions for *K5A1* and *K5CE* constructs.

"?" indicates that no records were available for one of the *K5CE* microinjection sessions, which produced four founders from 10 offspring.

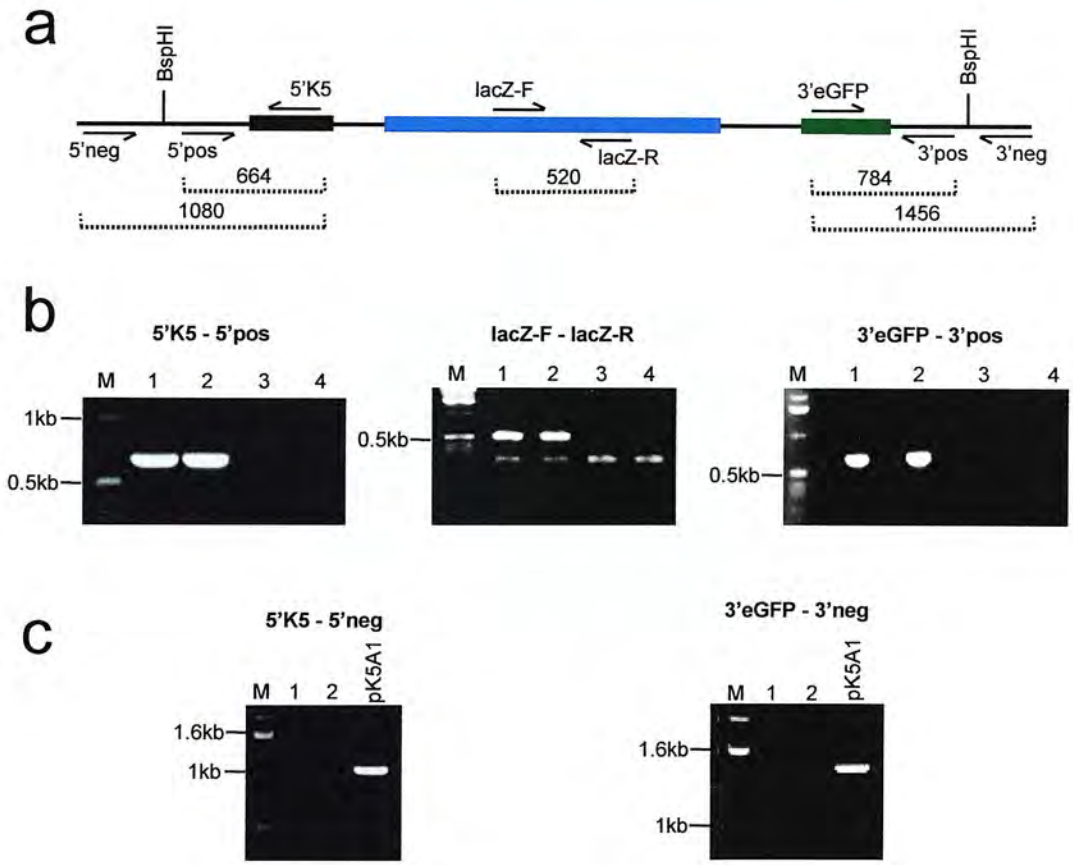


Figure 2.3 PCR analysis of integrated *K5A1* transgenes.

(a) Schematic showing the position and orientation of PCR primers on the *K5A1* construct, including the sizes of the amplification products. K5 promoter, *lacZ* and *eGFP* represented by black, blue and green bars, respectively. The BspHI sites used to linearise the construct for microinjection, are shown. (b) Genomic DNA samples from two positive (1,2) and two negative (3,4) potential *K5A1* founders used as template in the *lacZ* genotyping reaction (lacZ-F - lacZ-R, centre) and 5' and 3' integrity-checking reactions (left and right, respectively) are shown. The *lacZ* genotyping reaction also includes *myoG* PCR primers, which amplify a 245bp product in all four samples, indicating all samples are of PCR quality. (c) PCR involving primers lying outside the BspHI sites, showing that both founders 1 and 2 are genuinely positive, and not contaminated with pK5A1. M = 1kb ladder. Map not to scale, fragment sizes in bps.

that the 5' and 3' ends of the transgene, respectively, were intact, and had not undergone degradation prior to integration (Figure 2.3a,b). All four *K5A1* founders were found to contain intact transgenes and failed to amplify a PCR product with the 5'neg and 3'neg primers (Figure 2.3a,c). These PCRs demonstrated that both of the reporter construct termini were present in the transgenic mice, suggesting integrity of the transgene, and that there was no plasmid contamination of the genomic samples.

2.2.2 Analysis of K5A1 expression

The *K5A1* mice were generated to ensure the subcloned human K5 promoter would express according to published reports (Byrne and Fuchs, 1993; Byrne *et al.*, 1994; Borthwick *et al.*, 2001). It was recognised that a *lacZ* reporter mouse would be easier to analyse for its expression pattern than the inducible Cre mouse. Although both constructs were designed with an IRES-*eGFP* sequence, with the aim of being able to directly observe transgene expression in each line, it was recognised that expression of the cistron downstream of the IRES is often downregulated with respect to the first cistron, potentially limiting the usefulness of the second gene (Mizuguchi *et al.*, 2000). Regardless of the activity of the IRES in constructs *K5A1* and *K5CE*, βgal activity in the former line would provide a suitable means to assess K5 promoter activity. K5 promoter-mediated expression was assessed by eGFP fluorescence and xgal (or bluogal, an alternative βgal substrate), staining in both embryos and adult tissues.

To generate embryos for analysis of transgene expression, male founders, or male progeny produced from crosses between female founders and C57BL/6 males, were set up in timed matings with CD1 females. Embryos were collected at stage E15.5, as K5 promoter activity is known to be robust at this stage, while the skin is still permeable to fixing and staining reagents (Byrne *et al.*, 1994). At least three E15.5 litters from each line were examined, in addition to embryos at other stages. Genotyping was performed on genomic DNA prepared from yolk sacs or from amputated limb or tail tissue. Both methods allowed embryos to be identified with respect to their genotype, to confirm that phenotypically-detected transgene

expression was associated with inheritance of the transgene, and indicated transmission of the *K5A1* transgene by all founders.

2.2.3 Embryonic expression of eGFP

Although all four *K5A1* founders transmitted the transgene, analysis of eGFP by fluorescence microscopy suggested that only two of the four lines (*K5A1100,200*) expressed the transgene (Figure 2.4k-m, Figure 2.5 respectively). The pattern of fluorescence was similar in both lines, with the strongest signal evident in the vibrissae follicles. The follicles of the sensory vibrissae, the pinna, and the area around the mouth also exhibited greater fluorescence than the rest of the body. The transgene appeared more active in line *K5A1200*, as evidenced by a general but subtle increase in eGFP signal over the entire embryo surface. The relative differences in eGFP intensity predicted the presence, but not the extent, of the variation in βgal activity subsequently detected by xgal staining of fixed embryos.

2.2.4 Embryonic expression of βgal

The embryonic xgal staining patterns were compared with the pattern described by Byrne and colleagues (Byrne *et al.*, 1994). The authors used K5 mRNA *in situ* hybridisation to show faithful expression of βgal from the same human K5 promoter as used in these studies, in three independent lines. They first detected expression in dorsolateral ectoderm at E9.5. Prior to the differentiation of the epidermis, K5 expression at E12.5 was most notable in the maxillary region, and subsequently in the whisker pads. By E13.5, the K5 promoter was active in most lateral areas, though it appeared to lag dorsally and ventrally, while expression was highest in the vibrissae (whisker) follicles. At this stage, the first pelage (coat) follicles became apparent as foci of βgal production in the upper dorsolateral region. The activity of the K5 promoter increased markedly between E13.5 and E14.5, with intense βgal expression in pelage follicles and in the innermost layer of differentiating interfollicular epidermis. However, the intensity of expression remained visibly reduced over the spine region, and also ventrally. This dorsal stripe of decreased

promoter activity appeared to close during development to E16.5, with staining apparent in follicles and interfollicular epithelia over the entire embryo.

2.2.4.1 Transgenic line K5A1100

E15.5 expression in line *K5A1100* was observed to be predominantly in pelage follicles, except for the face, where strong staining was evident extending dorsally from the vibrissae follicles to the developing ear, and in patches of interfollicular epithelia distributed over the body (Figure 2.4a-d). While uniform expression in the vibrissae follicles and face was observed in all embryos, the extent of interfollicular expression varied between individual embryos in the same litter. While this observation could be related to differences in the developmental stage of littermates, it may be a function of epigenetic modulation of transgene expression (Palmiter and Brinster, 1986). The generally low level of interfollicular expression, and relatively wide region of reduced dorsal staining, resembled the pattern observed by Byrne and colleagues at E14.5 (Byrne *et al.*, 1994) (Figure 2.4i). At E16.5, staining appeared intense and uniform in the head region, and while βgal activity was more widespread than at E15.5, staining in the truck remained spotted, indicating the presence of unstained interfollicular epidermis (Figure 2.4e,f). Staining was still observed to lag ventrally, and dorsally (data not shown).

To determine the localisation of βgal within the layers of the skin, stained E15.5 embryos were paraffin embedded and sectioned sagittally (Figure 2.6). As expected, bluogal staining was found localised to the single layer of cells at the junction between the inner epidermis and the dermis. Even in non-transgenic embryos, the boundary between the epidermis and dermis was easily distinguished on the basis of the nuclear counterstain, which revealed significantly lower cell density in the dermis compared to the epidermis (see Figure 2.8a). While bluogal staining was detectable in the skin on all regions of the embryo surface (Figure 2.6b), expression in the basal layer was not continuous. Figure 2.6b,c shows a region of the forelimb in close proximity to the ventral trunk. βgal activity appeared uniform in the basal layer of the forelimb epidermis, while in the ventral body skin staining was patchy and

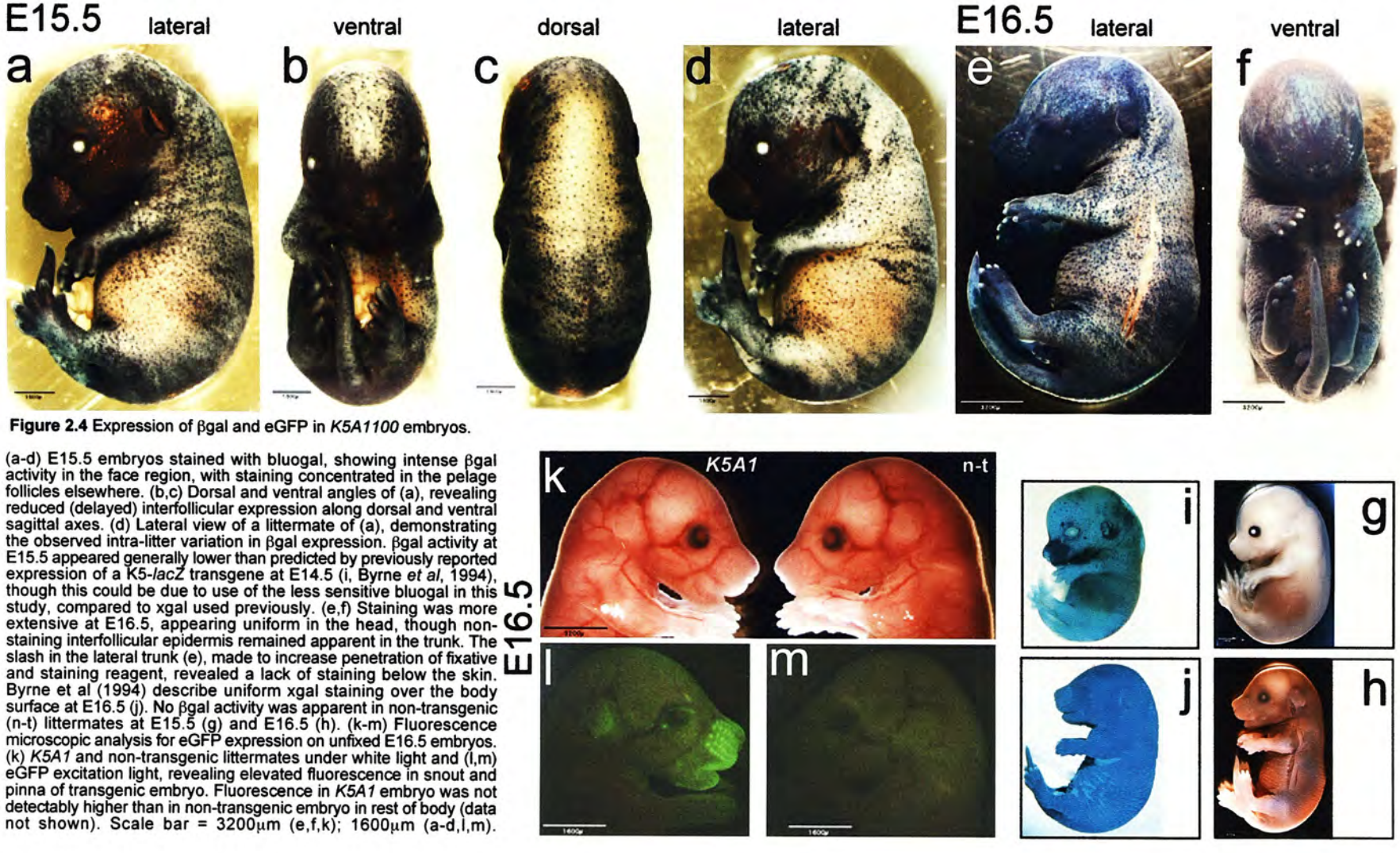


Figure 2.4 Expression of βgal and eGFP in *K5A1100* embryos.

(a-d) E15.5 embryos stained with bluogal, showing intense βgal activity in the face region, with staining concentrated in the pelage follicles elsewhere. (b,c) Dorsal and ventral angles of (a), revealing reduced (delayed) interfollicular expression along dorsal and ventral sagittal axes. (d) Lateral view of a littermate of (a), demonstrating the observed intra-litter variation in βgal expression. βgal activity at E15.5 appeared generally lower than predicted by previously reported expression of a K5-*lacZ* transgene at E14.5 (i, Byrne *et al*, 1994), though this could be due to use of the less sensitive bluogal in this study, compared to xgal used previously. (e,f) Staining was more extensive at E16.5, appearing uniform in the head, though non-staining interfollicular epidermis remained apparent in the trunk. The slash in the lateral trunk (e), made to increase penetration of fixative and staining reagent, revealed a lack of staining below the skin. Byrne et al (1994) describe uniform xgal staining over the body surface at E16.5 (j). No βgal activity was apparent in non-transgenic (n-t) littermates at E15.5 (g) and E16.5 (h). (k-m) Fluorescence microscopic analysis for eGFP expression on unfixed E16.5 embryos. (k) *K5A1* and non-transgenic littermates under white light and (l,m) eGFP excitation light, revealing elevated fluorescence in snout and pinna of transgenic embryo. Fluorescence in *K5A1* embryo was not detectably higher than in non-transgenic embryo in rest of body (data not shown). Scale bar = 3200μm (e,f,k); 1600μm (a-d,l,m).

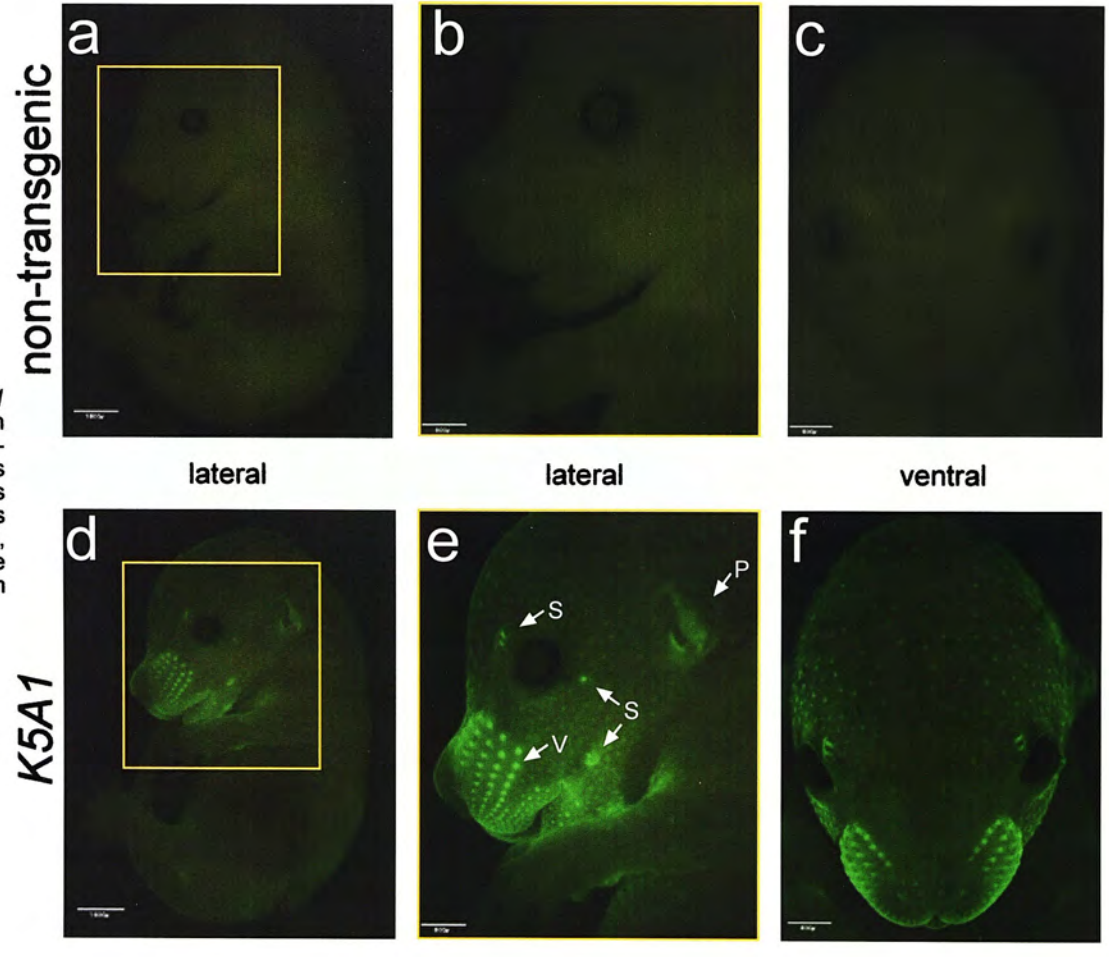


Figure 2.5 Expression of eGFP in *K5A1200* embryos.

Fluorescence microscopy of non-transgenic (a-c) and *K5A1* (d-f) E15.5 embryos, revealing a subtle overall increase in fluorescence in transgenic embryos compared to non-transgenic littermates. (b,e) Magnification of boxed regions in (a,d), respectively. Significant eGFP production was observed in the face region, particularly the vibrissae follicles of the snout, the pinna, and sensory vibrissae follicles (V, P and S, respectively, e). Fluorescing pelage follicles gave a spotted appearance to the skin (e,f). Scale bar = 1600μm (a,d); 800μm (b,c,e,f).

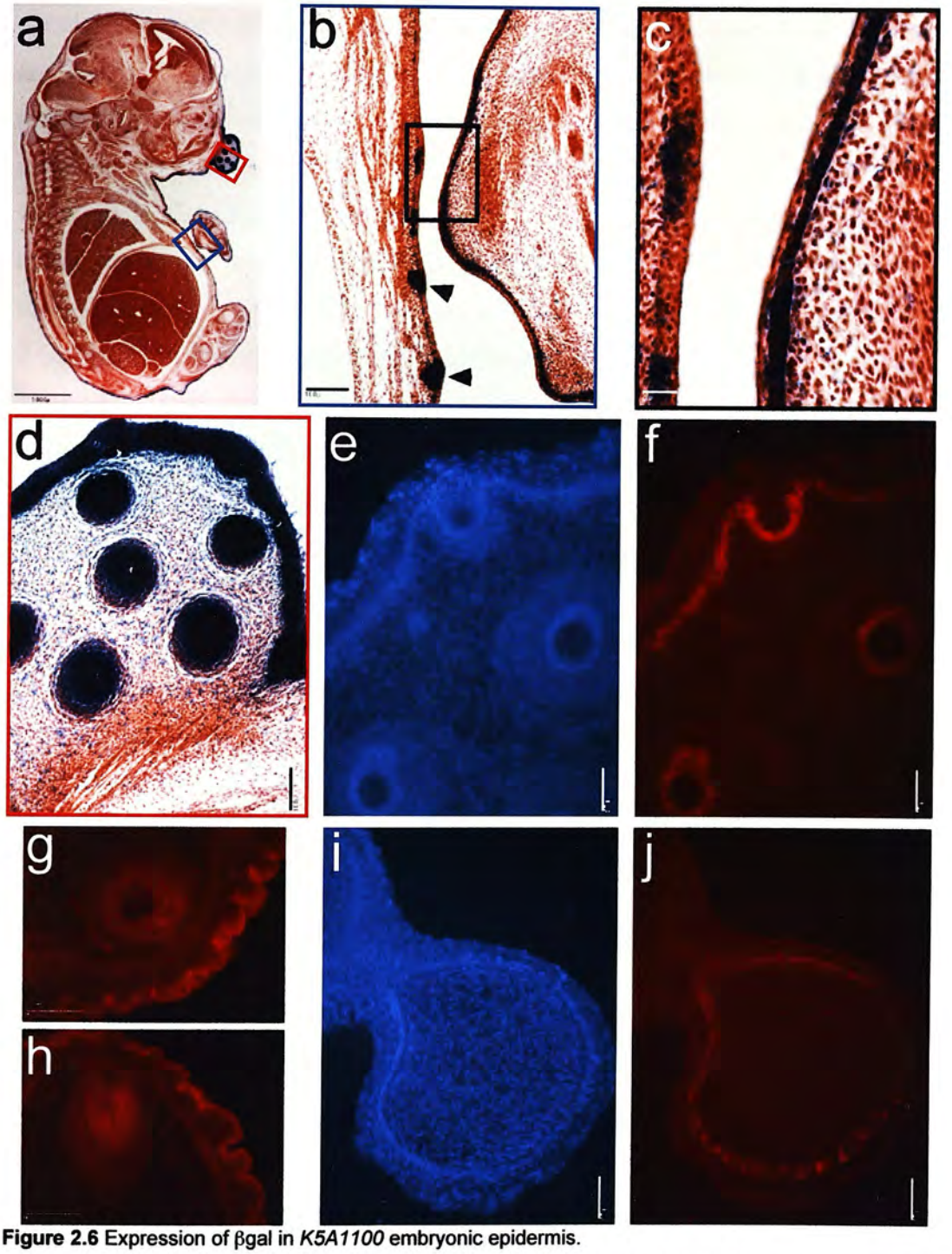


Figure 2.6 Expression of βgal in *K5A1100* embryonic epidermis.

(a) Sagittal section through bluogal stained E15.5 embryo. βgal expression was visible as thin layer around edge of embryo, though it was difficult to visualise in the upper dorsal region, consistent with reduced dorsal and ventral expression evident in wholemount embryos (Figure 2.4). βgal production was most conspicuous in epidermis of snout and vibrissae follicles. (b) Magnification of blue-boxed ventral region in (a) showing adjacent ventral trunk epidermis and developing forelimb epidermis, exhibiting patchy and continuous βgal expression, respectively. Intense staining foci descending into trunk dermis are pelage follicles (arrowheads, (b)). (c) Magnification of boxed region in (b) showing bluogal staining confined to the basal layer of both ventral trunk and forelimb epidermis (compare with anti-K14 immunostained embryonic skin, Figure 2.8a). (d) Magnification of red-boxed snout region in (a) showing bluogal staining throughout layers of epidermis and vibrissae follicles. However, closer examination revealed staining was most intense in basal epidermal layer and outer layer of follicle. Anti-βgal immunofluorescence indicated βgal expression was confined to basal epidermis in limb (i,j), and in the basal layer of snout epidermis which was continuous with the outer vibrissae follicle layer (e,f). (e,i) DAPI fluorescence revealed cell nuclei. (f-h,j) Anti-βgal-specific fluorescence. No specific immunofluorescence was detected in the absence of primary (anti-βgal) antibody (g) or on non-transgenic skin (h). Wax sections (a-d)/(e-j) were 10/7μm, counterstained with haematoxylin and eosin (a-d). Scale bar = 1600μm (a); 100μm (b,d); 50μm (e-j); 20μm (c).

occasionally observed in foci which extended into the dermal layer. These structures are likely to be developing pelage follicles, with the lack of intervening βgal activity reflecting the low level of interfollicular staining apparent on the ventral (and dorsal) surface in wholemount (Figure 2.4a-c). The region of most conspicuous βgal activity, when the sectioned embryo was viewed in its entirety (Figure 2.6a), was the skin lining the snout, and the associated vibrissae follicles. On initial examination, bluogal staining is apparent throughout the epidermal and follicular layers (Figure 2.6d), rather than being confined to basal and outer layers, respectively (Byrne *et al.*, 1994). To address whether this observation was the result of localised misexpression of βgal, or an artefact of the bluogal staining method, an anti-βgal antibody was used on unstained *K5A1100* embryo sections. The antibody proved to be less sensitive than the staining protocol, revealing the presence of βgal only in the regions corresponding to the most intense bluogal staining, including the snout region. However, the immunofluorescence technique indicated that high levels of βgal production were restricted to the basal layer of the epidermis and outer follicle layer (Figure 2.6e,f). The antibody also detected patchy βgal production in the basal layer of limb epidermis (Figure 2.6i,j), despite continuous bluogal staining in this region (Figure 2.6b,c). The broader band of bluogal staining in the snout may have been the result of bluogal reaction product leaching into neighbouring cells, or persistence of βgal protein into non-basal progeny of K5-expressing cells, or ectopic expression of βgal in non-basal layers, at a level below the threshold of detection by the antibody. Close re-examination of the bluogal staining follicular region reveals the darkest staining is found in a pattern consistent with the anti-βgal immunofluorescence. Elevated staining in surrounding non-epithelial cells, compared to that observed in other body regions (compare Figure 2.6d and c), may suggest leaching of the stain in this region of high βgal expression.

2.2.4.2 *Transgenic line* K5A1200

Examination of E15.5 embryos from line *K5A1200* revealed a more uniform expression pattern than observed in *K5A1100*, with the decreased dorsal and ventral expression resembling more closely that described by Byrne and co-workers

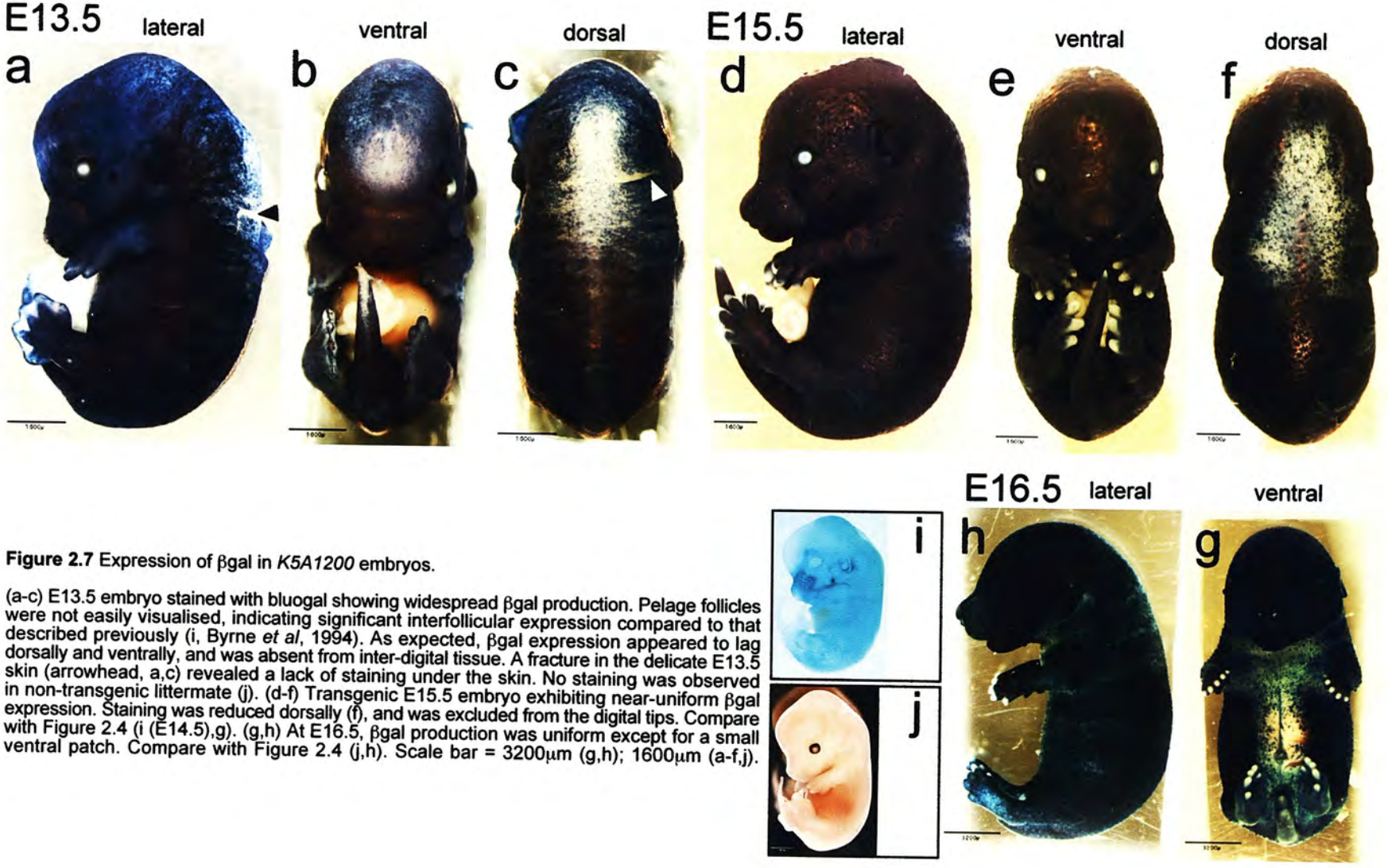


Figure 2.7 Expression of βgal in *K5A1200* embryos.

(a-c) E13.5 embryo stained with bluogal showing widespread βgal production. Pelage follicles were not easily visualised, indicating significant interfollicular expression compared to that described previously (i, Byrne *et al*, 1994). As expected, βgal expression appeared to lag dorsally and ventrally, and was absent from inter-digital tissue. A fracture in the delicate E13.5 skin (arrowhead, a,c) revealed a lack of staining under the skin. No staining was observed in non-transgenic littermate (j). (d-f) Transgenic E15.5 embryo exhibiting near-uniform βgal expression. Staining was reduced dorsally (f), and was excluded from the digital tips. Compare with Figure 2.4 (i (E14.5),g). (g,h) At E16.5, βgal production was uniform except for a small ventral patch. Compare with Figure 2.4 (j,h). Scale bar = 3200μm (g,h); 1600μm (a-f,j).

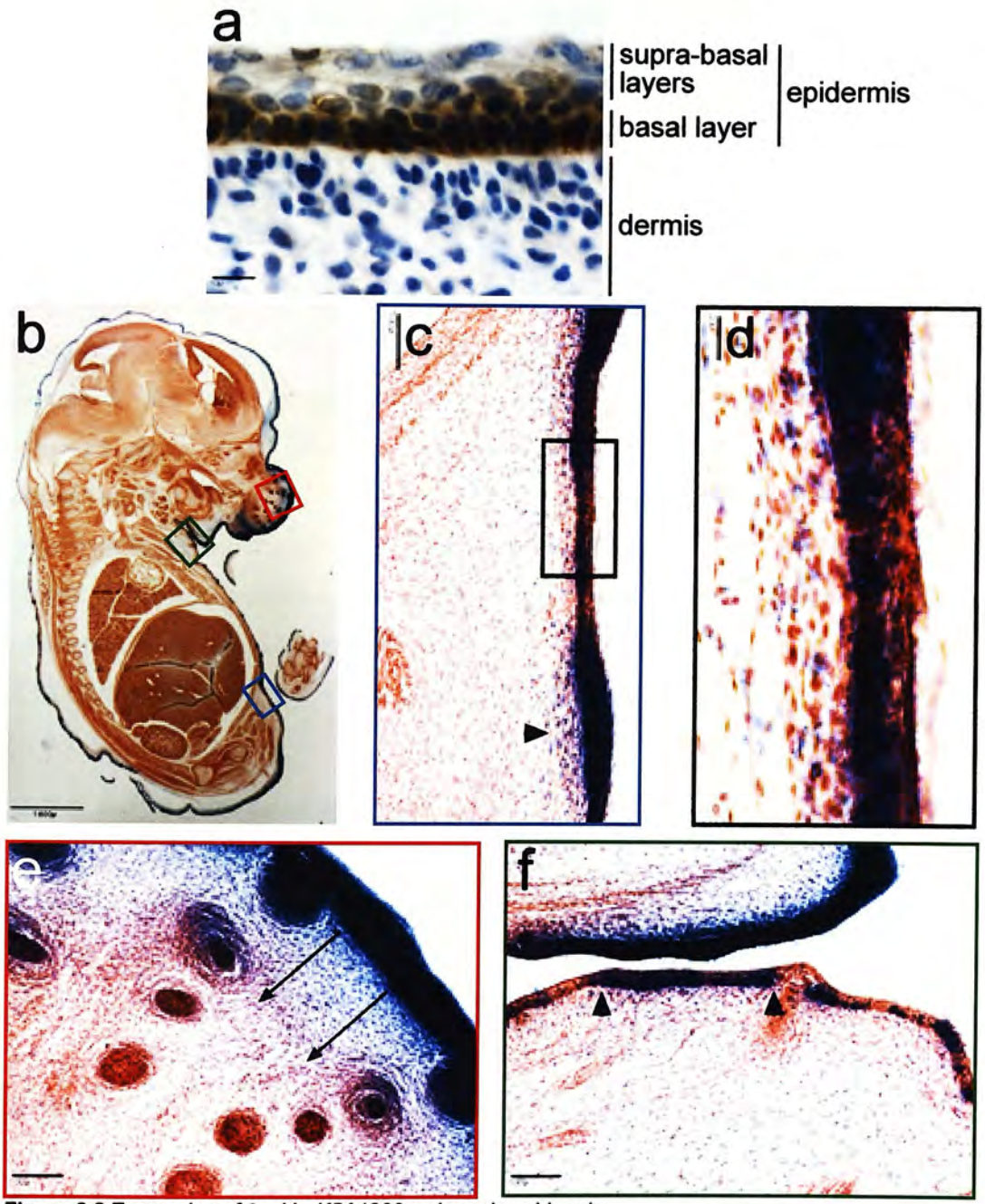


Figure 2.8 Expression of βgal in *K5A1200* embryonic epidermis.

(a) Non-transgenic embryonic skin illustrating the appearance of the dermal and epidermal layers, and immunostained with anti-K14 antibody to identify the basal epidermis (brown staining). (b) Sagittal section through bluogal stained E15.5 embryo. βgal expression was visible as thin layer around edge of embryo, though it was difficult to visualise in the upper dorsal region, consistent with reduced dorsal and ventral expression evident in wholemount embryos (Figure 2.7). βgal activity was highest in epidermis and vibrissae follicles of the snout. (c) Magnification of blue-boxed ventral region in (b) showing continuous epidermal βgal expression. Arrowhead shows region of apparent stain leaching into dermis. (d) Magnification of boxed area in (c) including a section of reduced epidermal βgal which appeared to exhibit basally-restricted transgene expression. (e) Magnification of red-boxed snout region of (b) revealing intensely stained epidermis and surface follicles. Non-staining follicles were likely internal and inaccessible to stain. Arrows represent extent of apparent stain leaching. (f) Magnification of green-boxed region in (b), showing apparent staining of all epidermal layers in head region (top of image), and less intense and continuous staining in ventral trunk. Area between arrowheads in (f) showed stronger expression basally than in supra-basal layers, possibly suggesting apparent supra-basal βgal expression was artefactual. Wax sections cut at 10μm and counterstained with haematoxylin and eosin (b-f), except (a), eosin and anti-K14 antibody. Scale bar = 1600μm (b); 100μm (c,e,f); 20μm (d); 10μm (a).

(Byrne *et al.*, 1994) (Figure 2.7i(d-f)). Even at E13.5, widespread staining was evident to the extent that pelage follicles were not obvious, indicating significant interfollicular staining (Figure 2.7a-c). By late E16.5, staining was uniform except for a ventral patch, which may have been due to localised reduction in epidermal permeability at this developmental stage (Figure 2.7h,g). Sagittal sections of E15.5 embryos (Figure 2.8b) revealed that epidermal expression, which was interrupted by non-staining patches less frequently than in *K5A1100*, resembled that observed in the snout region of *K5A1100* (Figure 2.8c-e). While βgal production was visibly most intense in the basal layer, staining was also present in the supra-basal layers. Apparent diffusion of bluogal product into the dermis (Figure 2.8c), and tissue underlying the snout epithelium (Figure 2.8e), where the K5 promoter is unlikely to be active, suggested the metabolised stain could diffuse beyond its site of production. Hence, diffusion upwards into the supra-basal epidermis could explain the apparent epidermal-wide staining observed in *K5A1200* embryonic skin. Unfortunately, *K5A1200* embryos were not available at the time of anti-βgal antibody staining.

2.2.5 *Adult expression of* K5A1

Although eGFP expression was insufficiently intense to permit genotyping of live neonate or adult mice using a hand-held UV lamp (Okabe *et al.*, 1997), transgenic *K5A1100,200* mice could be differentiated from non-transgenic littermates on the basis of eGFP fluorescence in ear-clip biopsies (Figure 2.9). However, in practice, genotyping was always confirmed by PCR.

Byrne and colleagues demonstrated that the 6.2kb human K5 promoter directed βgal expression in all examined stratified epithelial tissues, consistent with the expression profile of the endogenous K5 gene (Byrne and Fuchs, 1993). A subset of these tissues were taken from *K5A1* animals (n = four per line) for detection of βgal activity by wholemount xgal staining. Although a minority of the animals may have been homozygous for *K5A1*, no significant differences in staining were observed between mice of the same line. While staining was observed in pinna from both lines (Figure 2.10, Figure 2.11d), back skin, tail skin and tongue from *K5A1200* also

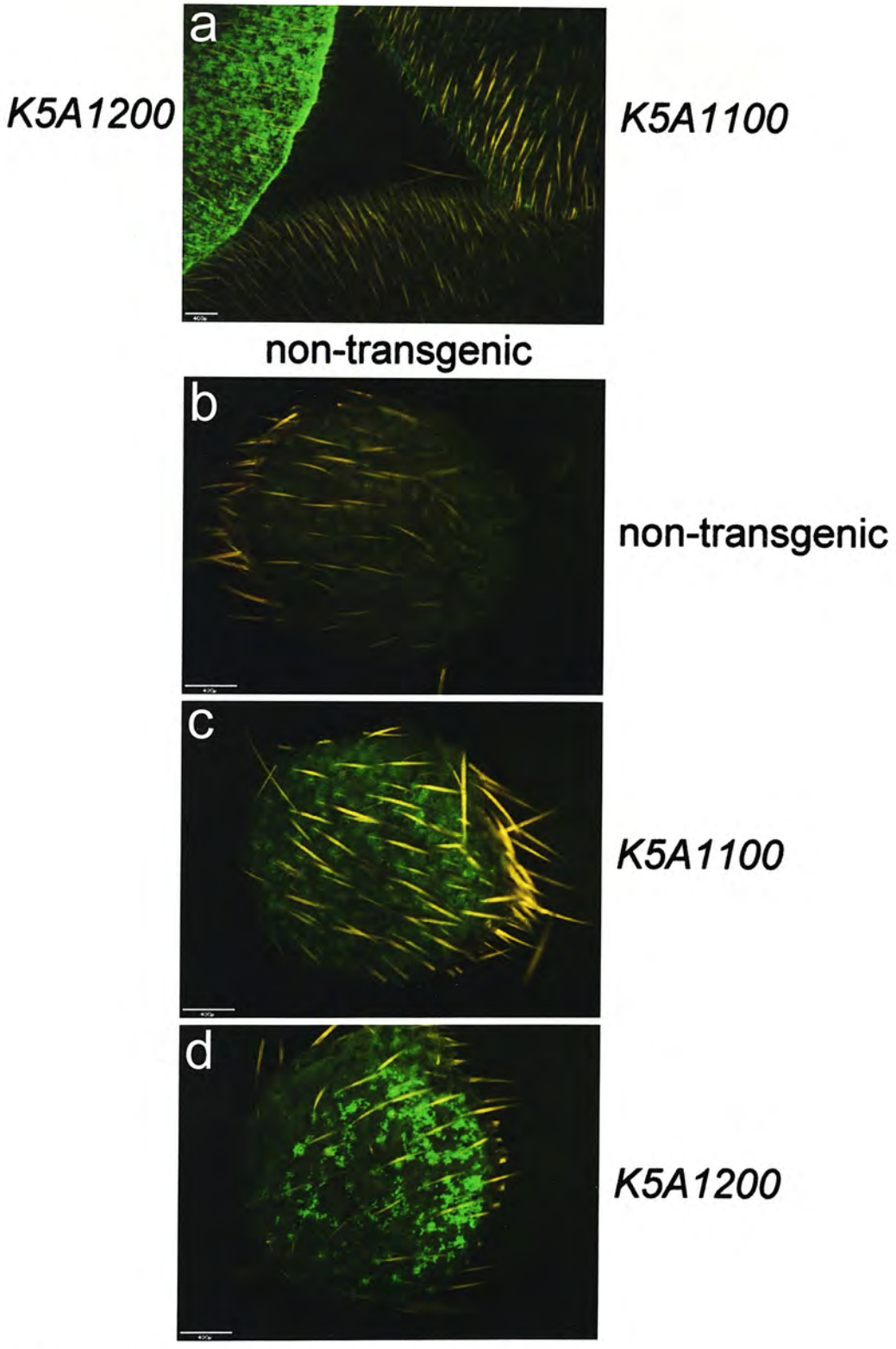


Figure 2.9 Expression of eGFP in *K5A1100* and *K5A1200* pinna tissue.

Unfixed ear-clips and whole pinnas were examined using fluorescence microscopy for eGFP expression. (a) Whole pinnas from *K5A1200* were easily distinguished from non-transgenic pinnas, while *K5A1100* pinnas were only subtly more fluorescent. (b-d) Ear-clips at higher magnification. Although *K5A1200* ear-clips (d) were always unambiguously fluorescent, confident detection of eGFP in *K5A1100* ear-clips (c) required comparison with a non-transgenic sample (b). Scale bar = 400μm.

stained (Figure 2.11a-c), albeit in a non-uniform and apparently variegated manner. The poor staining in these adult tissues is likely due in large part to the fact that they are all epidermal tissues possessing a stratum corneum, which constitutes a significant barrier to penetration by liquids. Such is the relative impermeability of the epidermal layer that the skin of intact post-E16.5 homozygous *ROSA26* embryos, which express characteristically ubiquitous and high levels of βgal, fails to stain significantly after incubation with xgal (data not shown). The reasonable amount of staining observed in back skin (Figure 2.11) was attributed to unintentional damage of the epidermal surface by the razor used to depilate the sample. Back skin samples from transgenic littermates shaved using an electric trimmer showed significantly reduced staining, with most positive cells clustered along the cut edge. Improved detection of βgal expression in epidermal tissues by Byrne and colleagues is probably a consequence of their methodology, which involved cryosectioning before staining, thus avoiding the problem of stain penetration (Byrne and Fuchs, 1993). The lack of detectable staining in the oesophagus in this study (data not shown), which is known to express endogenous K5 and is readily accessible to the stain, is consistent with the previous report, in which only a few basal cells were observed to express the transgene (Byrne and Fuchs, 1993).

The observation that βgal was produced in significant levels in the pinna of both lines, and was detectable by wholemount xgal staining, indicated its potential usefulness as a surrogate tissue in which to evaluate K5 promoter-driven expression in the *K5A1* lines, as well as in other lines discussed subsequently. The low density of hair follicles in the pinna, and the thin profile of the tissue which presents both dorsal and ventral surfaces, facilitates efficient stain penetration. Stained pinna were paraffin embedded and sectioned to determine the localisation of the *lacZ* signal. The pinna in cross-section comprises two layers of skin (epidermis and dermis) separated by a central band of cartilage (Figure 2.10b). Depending on the plane of section, patches of skeletal muscle can be seen in the dorsal surface. The stratum corneum of keratin-rich dead squames tended to separate from the epidermal surface during sectioning, and could normally be seen lying parallel to the surface, to which it often remained attached at hair follicles.

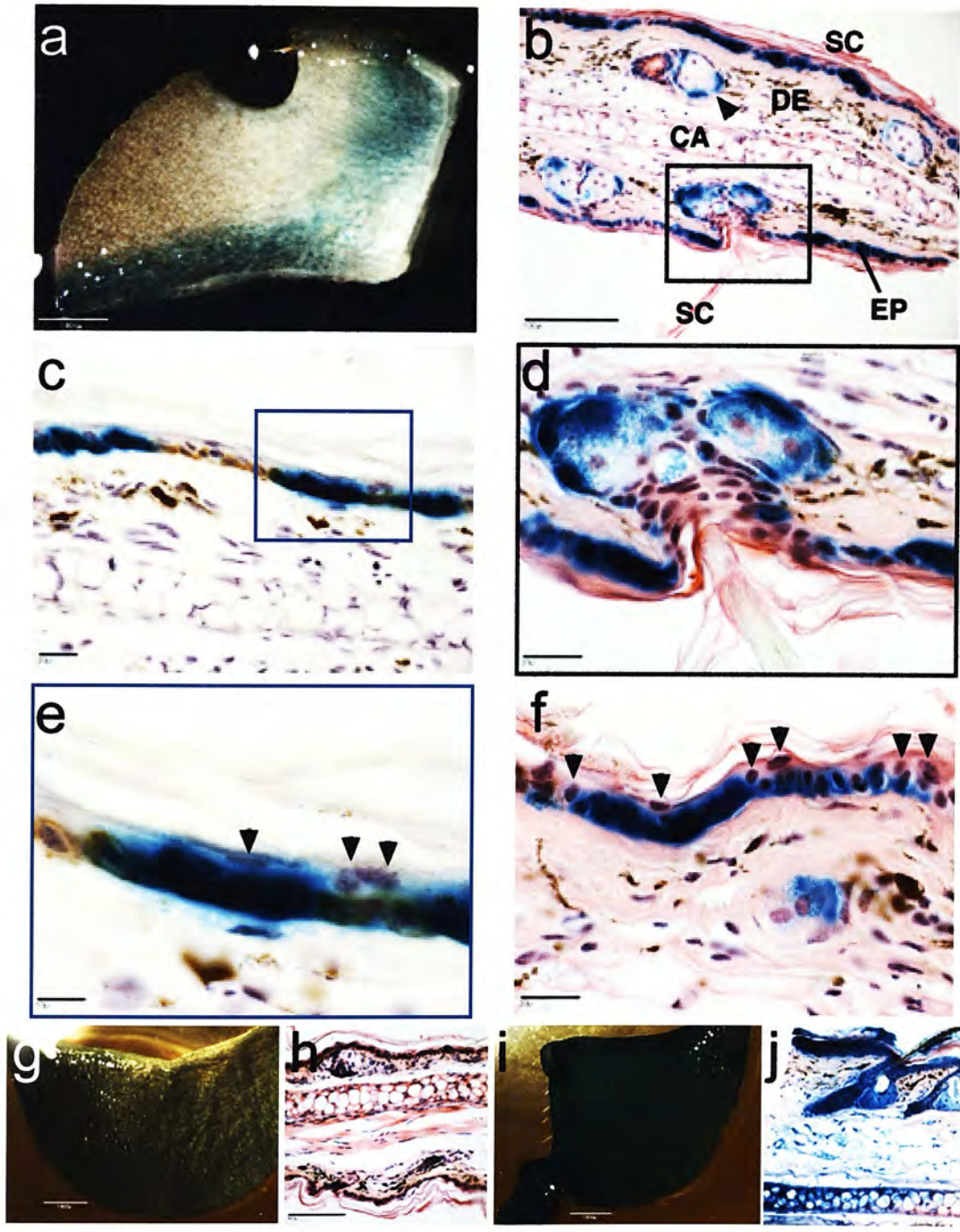


Figure 2.10 Expression of βgal in adult *K5A1100* pinna.

(a) Wholemount xgal treated pinna showing staining parallel to cut edges where stain can most easily penetrate tissue. If the level of βgal is sufficient, staining can result from the reagent entering the dorsal and ventral surfaces, as in the case of the *ROSA26* pinna (i). (b) Section through the stained portion of the pinna, revealing βgal activity in the epidermis, hair follicles (boxed) and associated sebaceous glands (arrowhead). Staining was mostly continuous, but non-uniform in intensity. CA = cartilage, DE = dermis, EP = epidermis, SC = stratum corneum. (d) Magnification of boxed region in (b) showing epidermal expression continuous with the hair follicle. (f) Arrowheads indicate unstained supra-basal cells, suggesting expression was basal epidermis-specific. (c) Patches of discontinuous staining co-localise to layer defined by anti-K14 antibody (brown staining). (e) Magnification of boxed region in (c), showing βgal- and K14-negative supra-basal cells (arrowheads). Confinement of staining to epidermis was not the result of incomplete stain penetration, as shown by sectioned *ROSA26* pinna (j). No staining was observed in non-transgenic littermate (g,h). Wax sections were cut at 10μm and counterstained with haematoxylin and eosin (b,d,f,h,j), or eosin and anti-K14 antibody (c,e). Scale bar = 1600μm (a,g,i); 100μm (b,h,j); 20μm (c,d,f); 10μm (e).

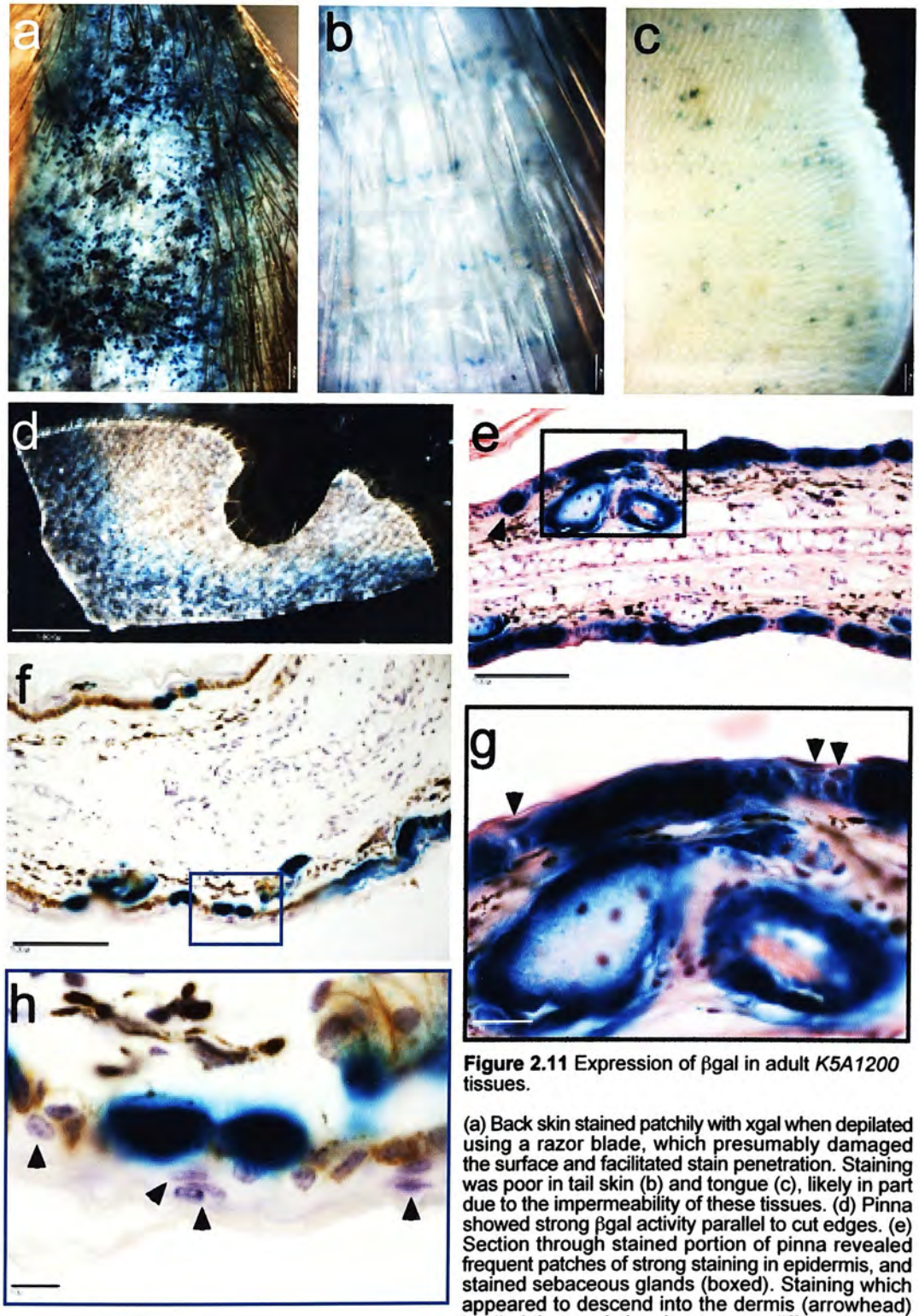


Figure 2.11 Expression of βgal in adult *K5A1200* tissues.

(a) Back skin stained patchily with xgal when depilated using a razor blade, which presumably damaged the surface and facilitated stain penetration. Staining was poor in tail skin (b) and tongue (c), likely in part due to the impermeability of these tissues. (d) Pinna showed strong βgal activity parallel to cut edges. (e) Section through stained portion of pinna revealed frequent patches of strong staining in epidermis, and stained sebaceous glands (boxed). Staining which appeared to descend into the dermis (arrowhead) implied that xgal signal was more diffuse than the site of expression it represented, suggesting that apparent βgal activity in supra-basal layers may have been a visual artefact of high levels of metabolised xgal in basal layer. (g) Magnification of boxed region in (e) showing unstained supra-basal cells (arrowheads). (f) Region of patchy expression showing co-localisation of xgal staining with anti-K14 signal (brown staining). (h) Magnification of boxed region in (f) clearly showing basal location of βgal-positive cells and the presence of non-xgal staining supra-basal cells (arrowheads). For staining controls, see Figure 2.10(g-j). Wax sections were cut at 10μm and counterstained with haematoxylin and eosin (e,g), or eosin and anti-K14 antibody (f,h). Scale bar = 1600μm (d); 400μm (a); 200μm (b,c); 100μm (e,f); 20μm (g); 10μm (h).

While neither line exhibited uniform production of βgal in the basal layer, *K5A1100* showed *lacZ* expression in a high proportion of basal epidermal cells, in addition to the hair follicles and sebaceous glands, known sites of significant K5 expression (Byrne *et al.*, 1994) (Figure 2.10). Expression in *K5A1200* appeared more variegated than in *K5A1100*, with more frequent and extensive non-staining regions (Figure 2.11e), in apparent contrast to the embryonic situation, where interfollicular staining was markedly more continuous than in the latter line. However, the high intensity of xgal staining observed in *K5A1200* embryos was reflected in adult pinna, with βgal activity not obviously restricted to the basal layer of the epidermis (Figure 2.11g). The use of an antibody directed against the K5 binding partner K14 demonstrated that small groups of *lacZ* expressing cells were basal (Figure 2.11f-h). It is clear that the subset of epidermal cells expressing the transgene do so at a high level. It seems likely that, as shown using the anti-βgal antibody in the vibrissae pad region of *K5A1100* embryos (Figure 2.6e,f), strong βgal production in the single cell basal layer of the epidermis, which is in itself only a few cells thick, results in a slightly diffuse xgal signal not precisely faithful to the pattern in which βgal is expressed. The occasional observation that staining can appear to extend into the dermal layer (Figure 2.8e), in a region sparsely populated with cells, supports the idea that apparent supra-basal expression in *K5A1200* is an xgal artefact resulting from high levels of βgal.

2.2.6 *Tracheal expression of* K5A1

Although the *K5A1* lines were generated primarily to confirm that the subcloned K5 promoter was driving transgene expression broadly consistent with published descriptions (Byrne and Fuchs, 1993; Byrne *et al.*, 1994), tracheas from *K5A1100,200* mice (n = four *K5A1* and three non-transgenic tracheas per line) were examined for βgal activity in SMGs, as reported by Borthwick and colleagues (Borthwick *et al.*, 2001) (no eGFP fluorescence was detectable in tracheas). Wholemount xgal stained tracheas were cut dorsally along the gaps in the cartilage C-rings, and pinned out to allow visualisation of the ventral surface. Staining was apparent in a band across the proximal (upper) trachea, consistent with the location

of tracheal glands (Figure 2.12i,ii(a,b)). Endogenous galactosidase activity was detected in the thyroid glands, and has been described previously (Borthwick, 1999). The thyroid glands flank the proximal trachea, resulting in thyroid staining being visible through the glandular region of pinned out tracheas. This gave the superficial appearance of intense staining in the tracheas of non-transgenic littermates. While closer examination revealed thyroid staining was partially responsible for this observation, stained tracheal glands were apparent in non-transgenic mice (Figure 2.12iii(a,b)). Occasional, weak background staining of SMGs in wild-type tracheas was also noted by Borthwick (Borthwick, 1999).

Tracheas were sectioned parallel to the proximal-distal axis and the SMGs, most of which are located between one of the proximal cartilage rings and the luminal surface, were examined for xgal staining (Figure 2.12i-iii(c-e)). *K5A1100* tracheas demonstrated robust staining in cells lining the majority of glands. Staining appeared stronger and more widespread than previously reported (Figure 2.12i(h) (Borthwick *et al.*, 2001). Consistent with the previous study, stained surface epithelial cells were observed in the proximal trachea (distal to the glandular region), but not in the distal trachea (Figure 2.12i(a,f,g)). Xgal staining in *K5A1200* SMGs was generally at a low level, though a minority of strongly staining glands were consistently apparent (Figure 2.12ii(c-e)). Staining in surface epithelial cells was faint and infrequent proximally, and absent distally (Figure 2.12ii(a,f,g)). The extent of genuine βgal activity in *K5A1* tracheas became unclear after examination of sectioned non-transgenic tracheas. While the intensity of staining varied between non-transgenic tracheas (strong and weaker examples are shown in Figure 2.12iii(c,d and e, respectively)), it was consistently detected, appearing to localise to the lining of glands. Unlike in the *K5A1* tracheas, no surface epithelial cells were observed to stain in either proximal or distal regions of non-transgenic tracheas (Figure 2.12iii(f,g)). While staining in *K5A1* glands was typically more intense than in non-transgenic glands, the strongest staining was often found in the basal portions of gland structures (Figure 2.12i(e),ii(c,d)), while Borthwick and colleagues noted xgal stained cells primarily in gland ducts just below the tracheal luminal surface (Figure 2.12i(h)) (Borthwick *et al.*, 2001). An anti-βgal antibody was used on sectioned

K5A1 tracheas in an attempt to visualise only the bacterial galactosidase. However, as illustrated previously in Figure 2.6, where only the regions showing the highest level of xgal staining were positive with the antibody, immunofluorescence on wax sections was considerably less sensitive than xgal staining, and no signal was apparent from tracheas of either line (data not shown). The SMG staining patterns are discussed further in section 2.4.1.3.

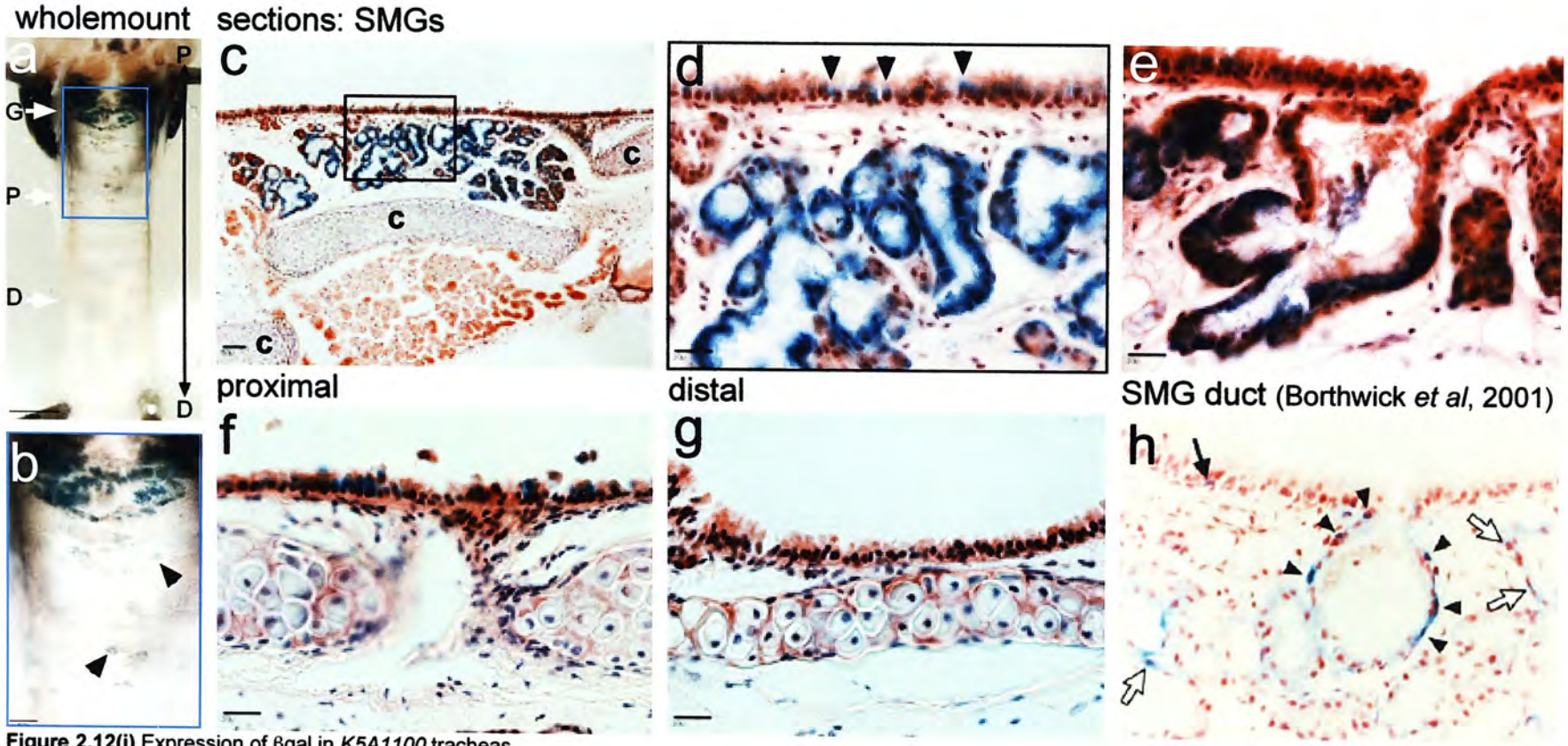


Figure 2.12(i) Expression of βgal in *K5A1100* tracheas.

K5A1200 and non-transgenic littermate tracheas are shown in Figure 2.12(ii) and (iii), respectively, on following pages. (a) A representative xgal treated trachea showing staining in proximal trachea in region corresponding to SMGs. Endogenous galactosidase activity could be detected in thyroid glands, at either side of the proximal trachea. Proximal (P) – distal (D) axis is indicated. (b) Magnification of boxed region in (a) showing stained glands and faint staining foci more distally, which may have been minor gland structures or stained surface epithelial cells (arrowheads). (c) Section of trachea at level of glands (G in (a)). The majority of glands exhibited robust βgal activity which appeared more widespread than described previously (h). C = cartilage rings. (d) Magnification of boxed area in (c) revealing staining in basal cells lining glands, and in a subset of surface epithelial cells (arrowheads), which may correspond to faint staining foci in (b). (e) SMG section from another xgal treated *K5A1100* trachea showing staining in lower gland duct cells. (f) Surface epithelia from proximal region corresponding approximately to P (a). Several small clusters of stained cells were evident. (g) More distally, approximately D (a), no βgal-positive cells were observed. (h) Stained gland from K5-*lacZ* mouse (Borthwick *et al*, 2001), original magnification x400, showing scattered surface basal cells (solid arrow); gland myoepithelial cells (open arrows); gland duct cells (arrowheads). Although staining in *K5A1100* tracheas was reproducibly robust, less intense but consistent staining was detected in non-transgenic littermates Figure 2.12(iii). Wax sections were cut at 10μm, counterstained with haematoxylin and eosin. Scale bar = 1600μm (a); 400μm (b); 50μm (c); 20μm (d-g).

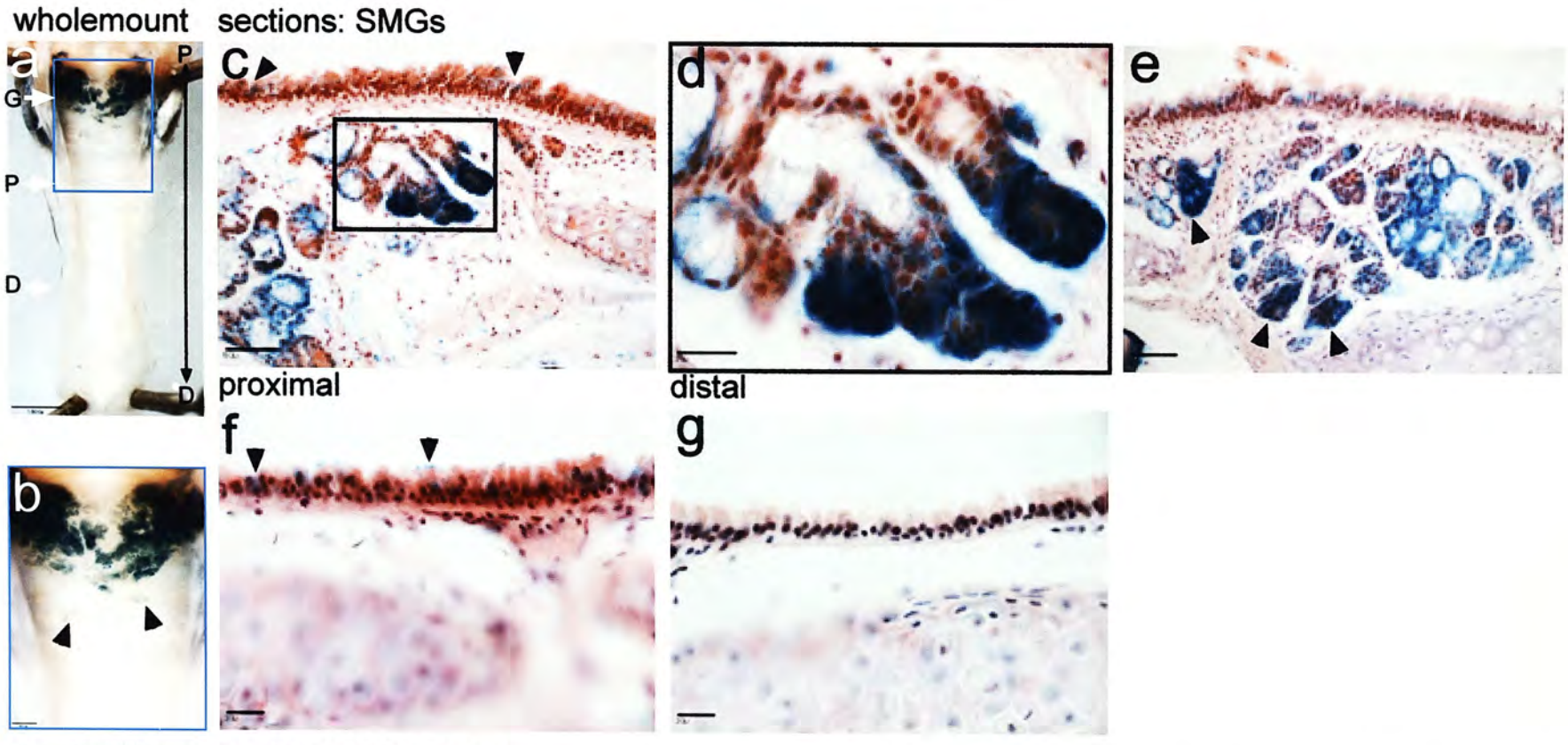


Figure 2.12(ii) Expression of βgal in *K5A1200* tracheas.

(a) A representative xgal treated trachea showing staining in proximal trachea in region corresponding to SMGs. Proximal (P) – distal (D) axis is indicated. (b) Magnification of boxed region in (a) showing more intense staining of glands, but fewer faint distal staining foci (arrowheads) relative to *K5A1100*, see Figure 2.12(i). (c) Section of trachea at level of glands (G in (a)). Consistent with observed epidermal βgal activity, βgal expression in *K5A1200* was stronger but patchier and less widespread than in *K5A1100*, see Figure 2.12(i). Except for the intensely stained gland (boxed), staining in neighbouring glands was not significantly elevated with respect to non-transgenic littermate, see Figure 2.12(iii). Staining in surface epithelia was faint and infrequent (arrowheads). (d) Magnification of boxed region in (c) showing glands with intense βgal activity. (e) SMG section from another xgal treated *K5A1200* trachea showing strongly staining glands (arrowheads) surrounded by glands exhibiting weaker staining. (f) Surface epithelia from proximal region corresponding approximately to P (a). Faint staining was apparent in discrete cells (arrowheads). (g) More distally, approximately D (a), no βgal-positive cells were observed. Compare *K5A1200* tracheal xgal staining with that observed in non-transgenic littermates, see Figure 2.12(iii). Wax sections were cut at 10μm, counterstained with haematoxylin and eosin. Scale bar = 1600μm (a); 400μm (b); 50μm (c,e); 20μm (d,f,g).

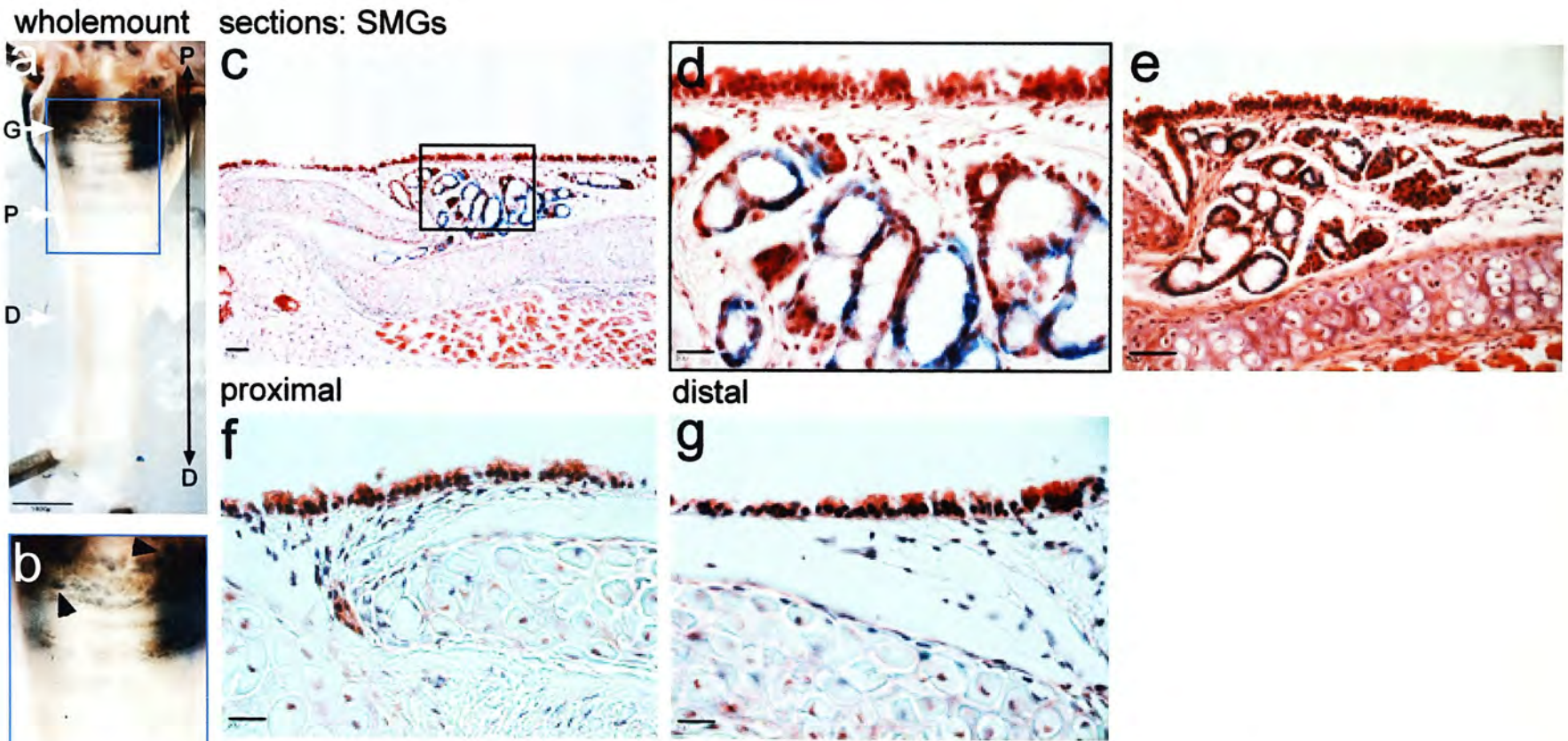


Figure 2.12(iii) Endogenous xgal staining in tracheas of non-transgenic *K5A1* littermates.

(a) A representative xgal treated trachea from a non-transgenic littermate of *K5A1* animals. The majority of apparent staining was due to visualisation of the endogenously staining thyroid glands, which lay behind the pinned-out trachea on either side of the glandular region. Proximal (P) – distal (D) axis is indicated. (b) Magnification of boxed region in (a). Although staining was apparent in the glands (arrowheads), it appeared less extensive and of reduced intensity by comparison with *K5A1100,200*, see Figure 2.12(i,ii(b)). (c) Section of trachea at level of glands (G in (a)). Endogenous galactosidase activity appeared widespread though less intense than in *K5A1100*, see Figure 2.12(i). No glands were observed to stain as strongly as a subset of *K5A1200* glands, see Figure 2.12(ii). (d) Magnification of boxed region in (c). No staining was detected in surface epithelial cells. (e) SMG section from another xgal treated non-transgenic trachea showing a low level of galactosidase activity. (f,g) Staining was absent from tracheal luminal epithelia at positions corresponding approximately to P and D (a), respectively. Wax sections were cut at 10μm, counterstained with haematoxylin and eosin. Scale bar = 1600μm (a); 400μm (b); 50μm (c,e); 20μm (d,f,g).

2.3 The transgenic K5 promoter-driven inducible Cre line, *K5CE*

The DNA manipulations required to generate the *K5CE* construct are described in appendix A1.1 and A1.4.

2.3.1 Microinjection of constructs and identification of transgenic founders

Since no restriction enzymes were predicted to cut pK5CE twice within flanking vector sequences but not within the construct, as BspHI did in pK5A1, the single cutter SfiI was used to linearise the vector for microinjection. The microinjection sessions are summarised in Table 2.1. Founders were identified using primers specific for the *Cre* portion of the fusion gene (Figure 2.13). PCRs using *myoG* primers were performed separately to ensure the ear-clip genomic DNA preparation was of sufficient quality (data not shown). The SfiI site separated the 3'eGFP primer sequence from the 3'neg primer sequence, thus allowing this primer pair to be used to check genomic genotyping samples for plasmid contamination (the 5'neg and 5'K5 primer sequences were not separated by SfiI digestion and therefore could not be used for this purpose, Figure 2.13a). No contamination was detected (data not shown). As for the *K5A1* lines, 5'K5 – 5'pos and 3'eGFP – 3'pos reactions demonstrated the presence of intact transgene termini in all 29 founder animals (data not shown).

2.3.2 Identification of K5CE *lines demonstrating inducible Cre-mediated recombination*

Although the trachea was the site of interest for *K5CE* expression, it was expected to be confined to a small subset of SMG and luminal epithelial cells, as observed for the K5 promoter-driven *lacZ* transgene (section 2.2.6 and (Borthwick *et al.*, 2001)). Rather than examine tracheas for low levels of transgene expression, the robust activity of the K5 promoter in basal epidermis was exploited to identify lines in

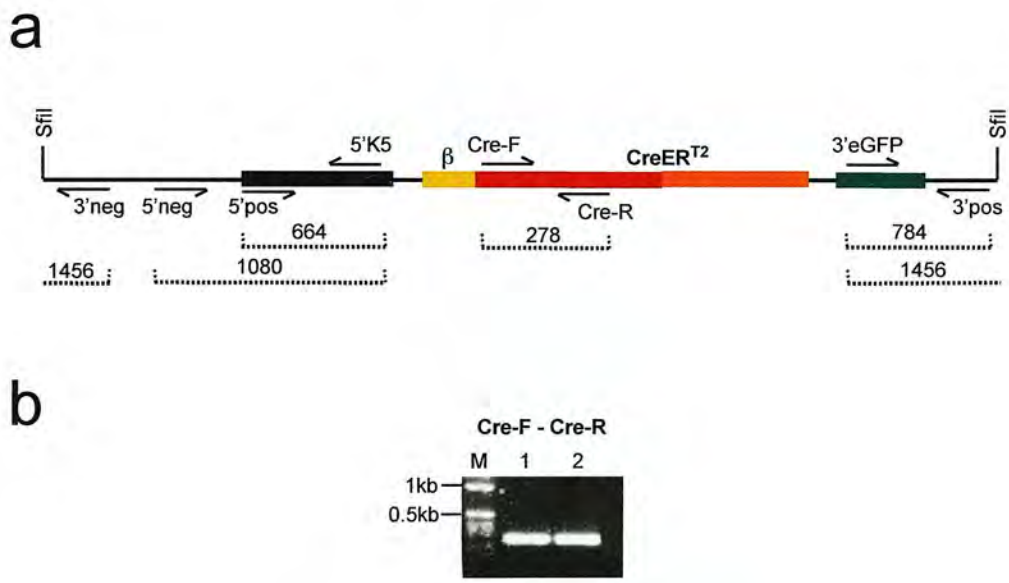


Figure 2.13 PCR analysis of integrated *K5CE* transgenes.

(a) Schematic showing the position and orientation of PCR primers on the *K5CE* construct, including the sizes of the amplification products. K5 promoter and *eGFP* represented by black and green bars, respectively. The integrity-checking PCRs and plasmid contamination control PCRs are as shown for the *K5A1* transgene in Figure 2.3, except for the 5'K5 - 5'neg plasmid contamination control PCR, which is uninformative in the SfiI linearising digest used. (b) Genomic DNA samples from two positive *K5CE* founders used as template in the *Cre* genotyping reaction (Cre-F - Cre-R) are shown. M = 1kb ladder. Map not to scale, fragment sizes in bps.

which the transgene was active, and thus worthy of further examination. Based on analysis of the two *K5A1* lines previously described, it was anticipated that eGFP expression would allow expressing lines to be easily identified by ear-clip biopsy and/or in embryos. However, fluorescence microscopic analysis of ear-clips from each of the founders did not reveal the presence of eGFP protein (data not shown). Furthermore, no fluorescence was detected in embryos from 12 independent lines, despite the fact that litters from at least eight of the lines were confirmed as containing *K5CE*-positive embryos (data not shown).

The absence of a visual marker of transgene activity necessitated the use of approaches to directly evaluate the expression (transcription and/or translation) of the *CreER^{T2}* fusion gene. Since the techniques required would not be suitable for screening all 29 lines generated, seven known transmitting lines were chosen for further analysis (*K5CE100,300,500,600,700,800,1100*).

2.3.3 Immunodetection of CreER^{T2} in embryonic epidermis

To evaluate the production of CreER^{T2} protein, hemizygous E15.5 embryos from each of the lines were sectioned and immunostained with a rabbit polyclonal anti-Cre antibody. Embryos were used in preference to adult skin because the thin and hairless skin is easily sectioned, and K5 expression is known to be confined to the basal epidermis by E15.5, thus providing an expected pattern of expression to compare against the antibody signal. An embryo known to express the conventional, non-inducible form of Cre from the bovine K5 promoter in basal epithelia (line *bK5Cre*, A. Ramirez, unpublished results, courtesy of D. Melton) was used as a positive control, since it was recognised that, despite the polyclonal origin of the antibody, the presence of the ER component may affect its immunoreactivity. While the *bK5Cre* embryo exhibited strong fluorescence specific to the basal epidermal layer, no signal was evident in any of the *K5CE* embryos (Figure 2.14).

Several explanations, in addition to the obvious interpretation that no protein was present, could have accounted for this result. As mentioned, the epitopes recognised

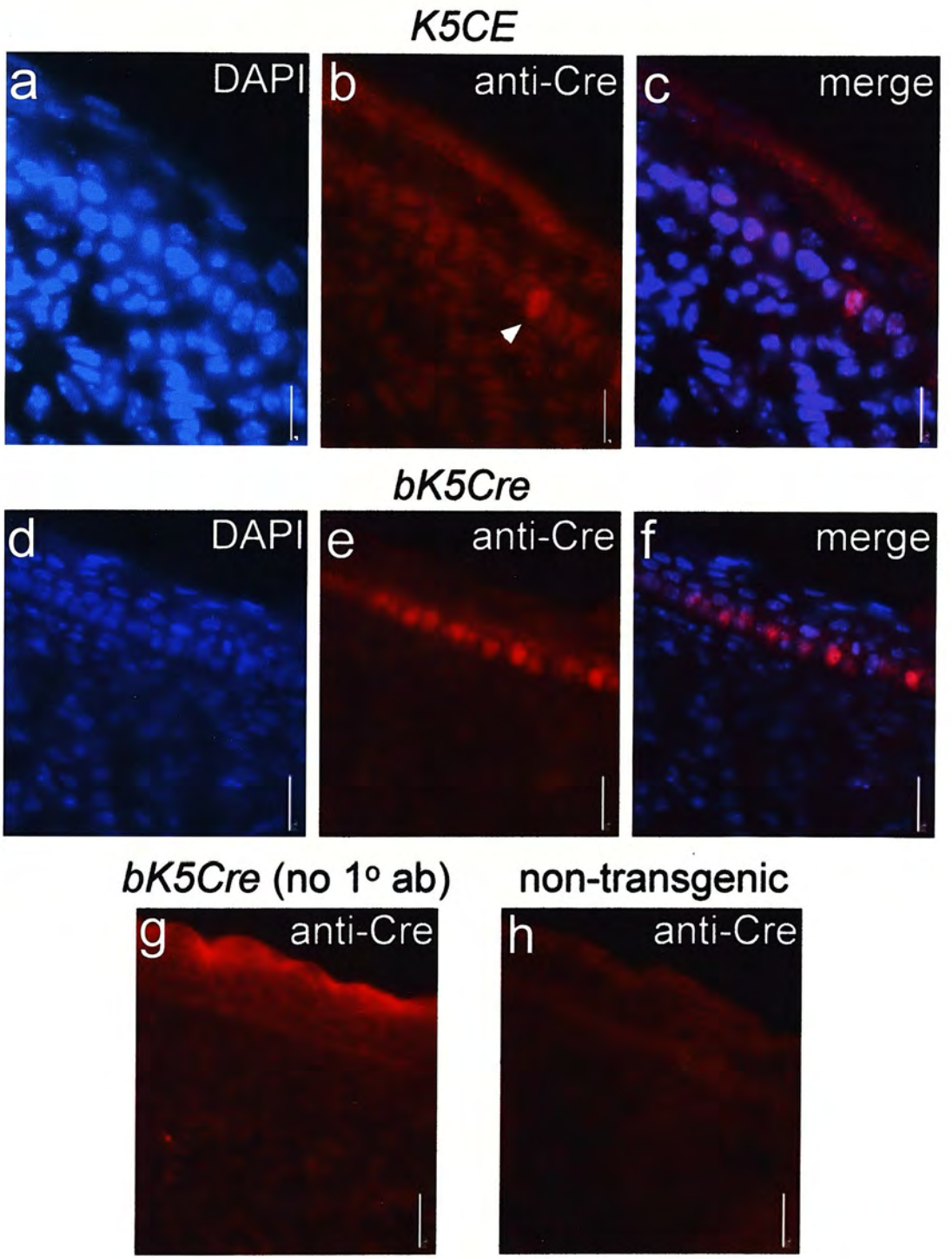


Figure 2.14 Anti-Cre immunofluorescence on *K5CE* embryonic skin.

E15.5 embryos from seven *K5CE* lines were sectioned and stained with a polyclonal anti-Cre primary (1°) antibody and fluorescent secondary antibody. Staining distinguishable from non-transgenic embryos (h) or transgenic embryos treated only with secondary antibody (g) was not detected in any of the lines. *K5CE1100* is shown to represent the *K5CE* lines (a-c), although it was the only line showing any potential expression. A single cell appears positive (arrowhead, b). Failure to detect Cre protein in *K5CE* lines was not the result of the technique or reagents, since Cre was detected in identically processed embryos from line *bK5Cre*, which expresses conventional (non-fusion) Cre protein from the bovine K5 promoter (d-f). Wax sections were cut at 7μm. DAPI fluorescence shows cell nuclei. Scale bar = 20μm.

by the anti-Cre antibody may have been altered or inaccessible on the CreER^{T2} fusion protein. In the absence of TAM, which stimulates nuclear accumulation of the fusion protein, CreER^{T2} would have remained cytoplasmic, potentially resulting in a weak signal compared to nuclear-localised, conventional Cre protein. Indeed, it has previously been shown using anti-Cre immunofluorescence on sectioned tails from CreER^{T2} mice before and after systemic TAM administration that, while anti-Cre signal was low in uninduced tail skin, robust fluorescence was evident in cells in which TAM-induced nuclear translocation had occurred (Indra *et al.*, 1999). Alternatively, or additionally, CreER^{T2} production in the *K5CE* lines may have occurred at a significantly lower level than that of Cre in *bK5Cre*, possibly below the threshold of immunodetection. S. Webb provided some technical assistance with establishing the anti-Cre immunoassay.

2.3.4 Analysis of K5CE expression in adult skin

In order to determine whether the transgene was being transcribed in any of the seven lines, RNA was extracted from adult hemizygous backskin and reverse transcribed to complementary DNA (cDNA) for use as template in PCRs specific for *Cre* (Figure 2.15). Robust products were evident from lines *K5CE100,500,600,700,1100*, and a faint signal was apparent from *K5CE300*. No product was visible from *K5CE800*, or in non-transgenic controls. The identity of the *Cre* amplification products was confirmed by hybridisation with a radiolabelled internal *Cre* oligonucleotide (Figure 2.15c). Thus, of seven lines analysed, only one appeared detectably devoid of *CreER^{T2}* transcript in skin samples.

2.3.5 A functional assay for inducible Cre-mediated recombination

Given the presence of detectable *Cre* transcript in six of seven lines tested, a functional assay was used to evaluate whether this translated into levels of protein capable of effecting induced recombination at loxP sites. The Cre-dependent *lacZ* reporter locus *R26R* (Soriano, 1999) was introduced into the genome of *K5CE* mice from all seven lines by crossing with *R26R* mice. The breeding schemes were not

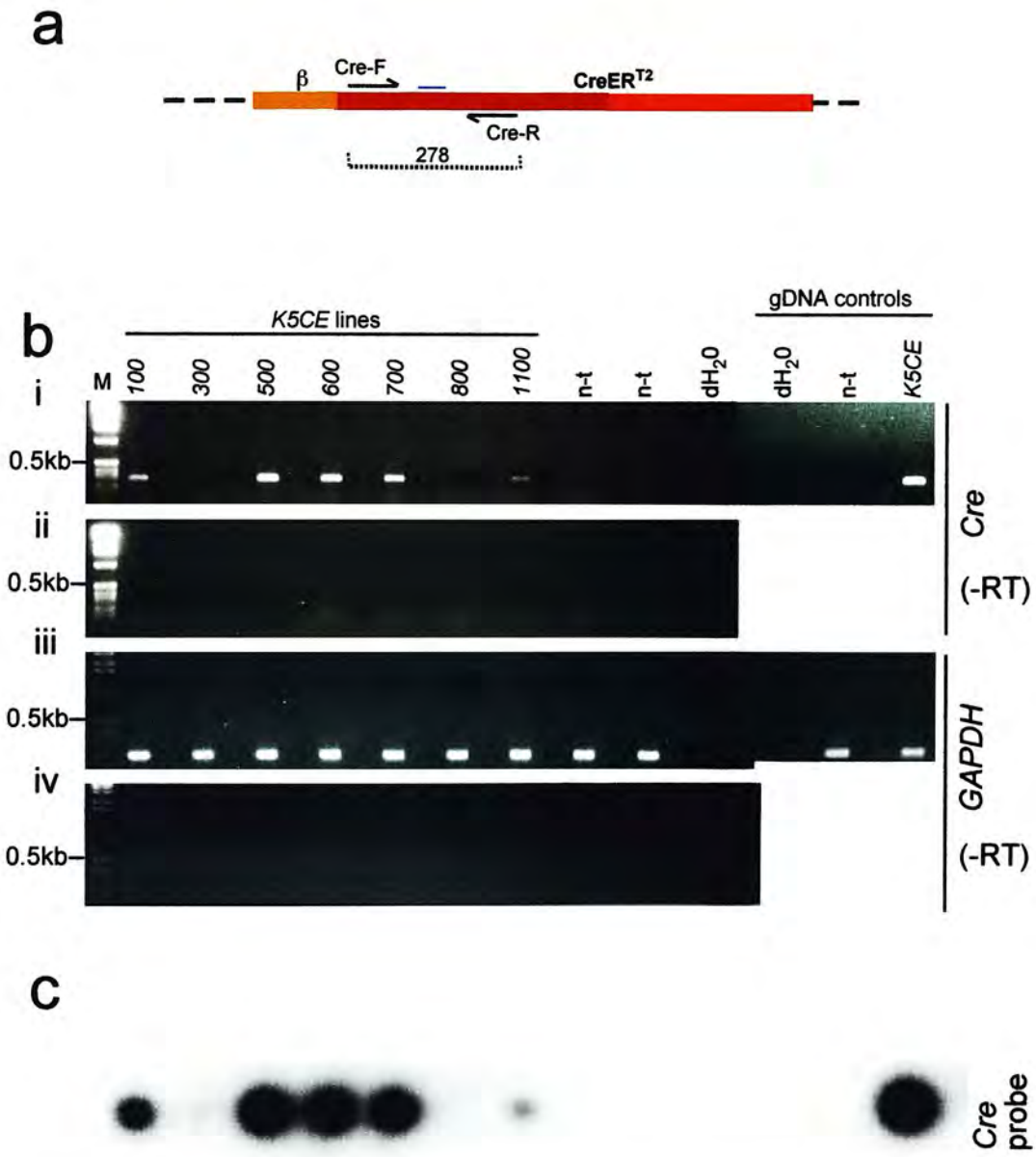


Figure 2.15 RT-PCR for expression of *CreER^{T2}* in skin of seven *K5CE* lines.

RNA was isolated from the back skin of mice representing seven *K5CE* lines. RNA was DNase-treated before cDNA was prepared and equal amounts used as template in *Cre*-specific PCRs. (a) Schematic of *CreER^{T2}* portion of the *K5CE* transgene, showing position of internal *Cre* oligonucleotide probe (blue bar) with respect to *Cre*-specific PCR primers. (b(i)) After 30 cycles, PCR product was generated from all lines except *K5CE800*, though signal was very weak in *K5CE300*. No product was generated by two non-transgenic littermates (n-t) or by the dH$_2$0 control subjected to reverse transcription. Non-transgenic and dH$_2$0 samples used only in PCR step were also negative, while product was generated from *K5CE* genomic DNA (gDNA). (ii) RNA and control samples put through the reverse transcription protocol in absence of reverse transcriptase (RT) enzyme (-RT) were all negative, indicating the absence of contaminating gDNA. (iii) Using *GAPDH* primers to control for RNA quality generated comparable PCR products from all samples, while dH$_2$0 controls remained negative. (iv) *GAPDH* -RT reactions were negative. (c) RT-PCR products in (b(i)) were transferred to a nylon membrane and hybridised with a labelled 18mer internal *Cre* oligonucleotide (a) which confirmed the identity of the PCR product and the absence of low levels of product in negative controls. M = 1kb ladder. Map not to scale, fragment sizes in bps.

equivalent for all lines, with some involving $R26R^{tg/wt}$, and others $R26R^{tg/tg}$, animals. Also, intercrosses were used to increase the proportion of compound $K5CE^{tg}$ $R26R^{tg}$ animals obtained, with a concomitant increase in the incidence of homozygosity at both loci. As a randomly integrated transgene, PCR could not be used to distinguish animals hemizygous and homozygous for *K5CE*. While a PCR assay was subsequently developed to discriminate hemizygosity and homozygosity for the *R26R* transgene (chapter 3, Figure 3.3), it was not available at the time of these experiments.

Since the number of *K5CE* and *R26R* loci could influence the frequency of Cre-mediated activation of βgal production from *R26R*, it was important to be aware of the differences in potential genotype of the animals used in the assay. While this was less crucial for the *K5CE* transgene, where hemizygous copy number was likely to vary by more than two-fold between lines (confirmed in section 2.3.6), doubling the number of *R26R* loci would in principle double the cellular targets for Cre action. For this reason, the probability of hemizygosity and homozygosity for each transgene, of the animals used from each of the lines, was calculated from known genotypes of their ancestors (Table 2.2).

Inducible Cre activity was assayed using a method of topical administration modified from Vasioukhin and colleagues (Vasioukhin *et al.*, 1999) (B. Stripp, personal communication), exploiting the pinna as an epidermal tissue more accessible to soluble reagents than the skin, and which was amenable to biopsy analysis. Figure 2.16 shows a schematic of the experiment. Two $K5CE^{tg}$ $R26R^{tg}$ compound transgenic animals from each line received topically applied TAM in dimethyl sulphoxide (DMSO) to both surfaces of the pinna. Two doses of TAM were evaluated, with the two pinna of each mouse receiving approximately 10mg and 5mg, respectively, every day for five consecutive days. Ear-clips were taken from both pinna of each animal on the day following the final treatment, and stained with xgal. In each experiment, at least one $K5CE^{wt}$ $R26R^{tg}$ single transgene-positive control animal was processed identically. One compound transgenic mouse from each line received

K5CE line	*K5CE*		*R26R*	
100	tg/tg	0.33	tg/tg	0.33
			tg/wt	0.66
	tg/wt	0.66	tg/tg	0.33
			tg/wt	0.66
300	tg/tg	0.33	tg/tg	0.33
			tg/wt	0.66
	tg/wt	0.66	tg/tg	0.33
			tg/wt	0.66
500	tg/tg	0.33	tg/tg	0.5
			tg/wt	0.5
	tg/wt	0.66	tg/tg	0.5
			tg/wt	0.5
600	tg/tg	0.33	tg/tg	0.33
			tg/wt	0.66
	tg/wt	0.66	tg/tg	0.33
			tg/wt	0.66
700	tg/wt	1.0	tg/wt	1.0
800	tg/tg	0.33	tg/tg	0.33
			tg/wt	0.66
	tg/wt	0.66	tg/tg	0.33
			tg/wt	0.66
1100	tg/wt	1.0	tg/tg	0.5
			tg/wt	0.5

Table 2.2 Predicted frequencies of hemizygosity and homozygosity for the *K5CE* and *R26R* transgenes in compound transgenic mice from seven *K5CE* lines used in the tamoxifen ear-painting assay.

The breeding strategies were not equivalent for each of the seven *K5CE* lines, with the result that mice homozygous for one or both of the transgenes would be more common for some lines than others. Since the number of transgenes could theoretically affect the frequency of recombination observed in response to tamoxifen (TAM) treatment, it was desirable to know the hemizygosity/homozygosity status of the mice used from each line in the TAM ear-painting assay. However, no simple means of assessing the hemizygosity/homozygosity status of individual mice was available at the time of the experiment. Therefore, breeding records were used to predict probabilities of transgene inheritance (assuming Mendelian inheritance, and that the transgenes were unlinked) at each of the crosses used in the generation of the litters from each line used in the assay. For each line, final probabilities of hemizygosity and homozygosity for the *K5CE* transgene are shown, where 1.0 is 100% probability of the genotype shown. For each *K5CE* genotype (tg/wt or tg/tg), the probability of tg/wt or tg/tg for *R26R* is shown. Since only compound transgenic mice should show Cre-dependent recombination, mice carrying only one of the transgenes, and non-transgenic mice, are excluded.

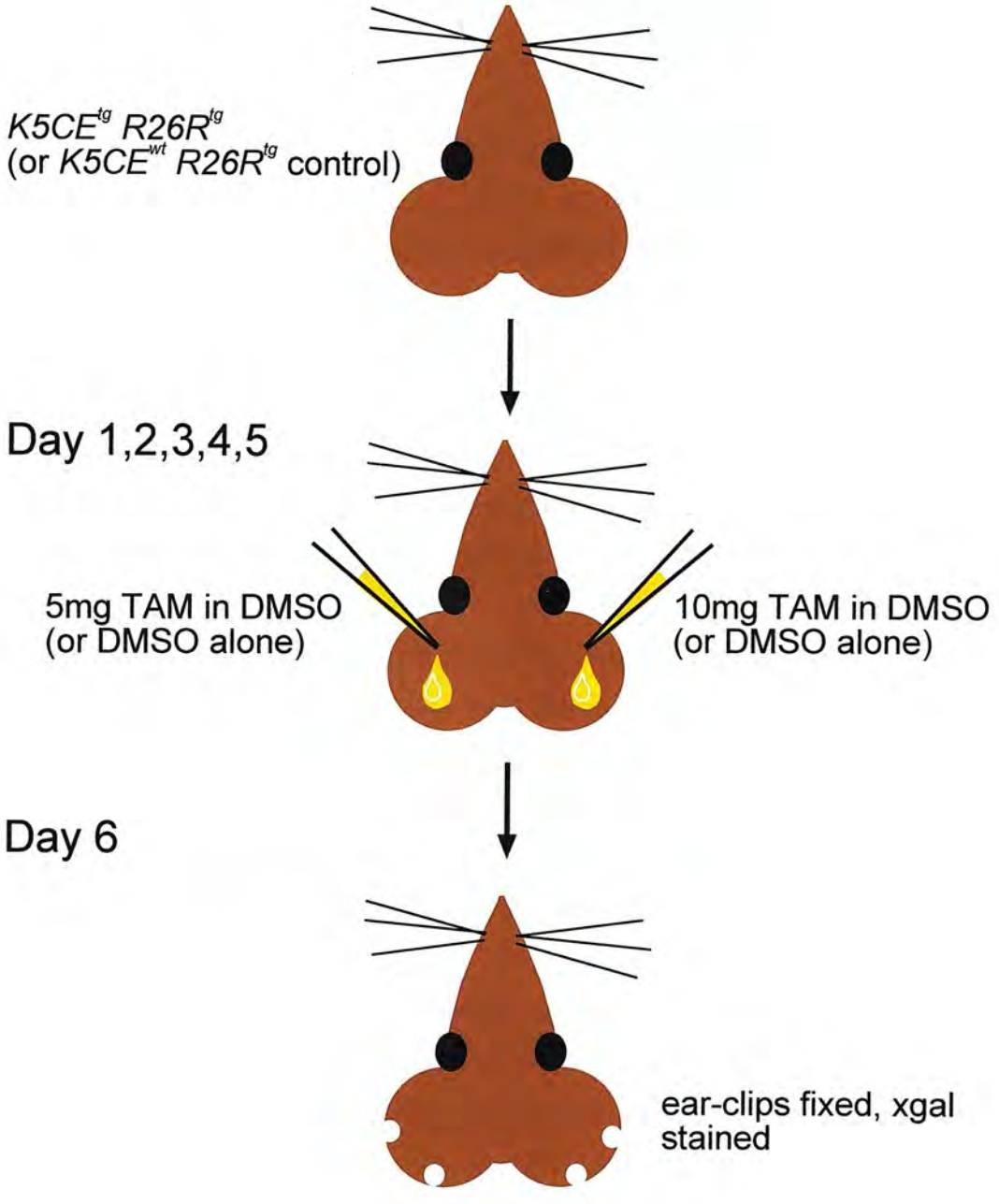


Figure 2.16 Tamoxifen ear-painting regime used to identify functionally expressing *K5CE* lines.

Of the three *K5CE*^{tg} *R26R*^{tg} compound transgenic mice from each of seven *K5CE* lines used in the ear-painting assay, two received two different concentrations of topically administered tamoxifen (TAM) in dimethyl sulphoxide (DMSO) to each pinna for five consecutive days. The third animal received DMSO alone to both pinna. In each experiment, at least one *K5CE*^{wt} *R26R*^{tg} mouse received TAM as per the *K5CE*^{tg} *R26R*^{tg} animals. 50µl of solution was applied using a micropipette, and a cotton bud stripped of most of the cotton was used to "paint" the solution over dorsal (outside) and ventral (inside) surfaces of the pinna. Two ear-clips were taken from each pinna on the day following the final treatment, and fixed before xgal staining. Stained ear-clips were examined wholemount prior to sectioning (Figure 2.17).

DMSO in the absence of TAM. No adverse effects on the health of the animals were observed.

Wholemount stained earclips were inspected for the presence of βgal positive cells (Figure 2.17). Stained cells were evident in ear-clips from all *K5CE* lines except *K5CE800*, the only line which appeared completely negative for *K5CE* transcription by RT-PCR (see Figure 2.15). *K5CE100,300,500* all exhibited low numbers of stained cells, which appeared marginally more frequent in *K5CE300* than in *K5CE100,500* (Figure 2.17a-d(i), *K5CE500* not included). However, approximately equal numbers of stained cells were present in both DMSO and TAM treated ear-clips from these lines, indicating insignificant inducibility of recombinase activity. Contrastingly, significant numbers of xgal stained cells were visible in ear-clips from *K5CE600,700,1100* animals treated with TAM (Figure 2.17e-j). Both ears from each mouse showed a comparable amount of staining, though difficulties in accurate TAM administration prevented conclusions being drawn regarding dosage (discussed in section 2.4.2.2), and the extent was similar in both TAM treated mice from each line (data not shown). Table 2.2 shows that the observed activity in *K5CE600,700,1100* compared to *K5CE100,300,500*, which all expressed $CreER^{T2}$ transcript (see Figure 2.15), was not the result of increased probability of homozygosity of either transgene in the former lines. Overall, the frequency of staining observed was *K5CE600<1100<700*. While significantly more βgal positive cells were present after TAM treatment than after application of DMSO alone, more stained cells were apparent in ear-clips from the DMSO control mice of lines *K5CE600,700,1100* than in the TAM or DMSO treated ear-clips from the remaining lines. Given the complete absence of detectable xgal staining in the TAM treated $K5CE^{wt}$ $R26R^{tg}$ controls (Figure 2.17k(i)), and in all ear-clips from *K5CE800* (not shown), this indicated a "background leakiness" of TAM-independent Cre activity which correlated with the potential for inducible activity. Although the ER component of $CreER^{T2}$ had been modified to reduce its sensitivity to endogenous oestrogen (Feil *et al.*, 1997), the $CreER^{T2}$ fusion protein, and analogous constructs, are known to mediate a low frequency of recombination in the absence of exogenous inducing agents (Vasioukhin *et al.*, 1999; Kellendonk *et al.*, 1999; Weber *et al.*, 2001).

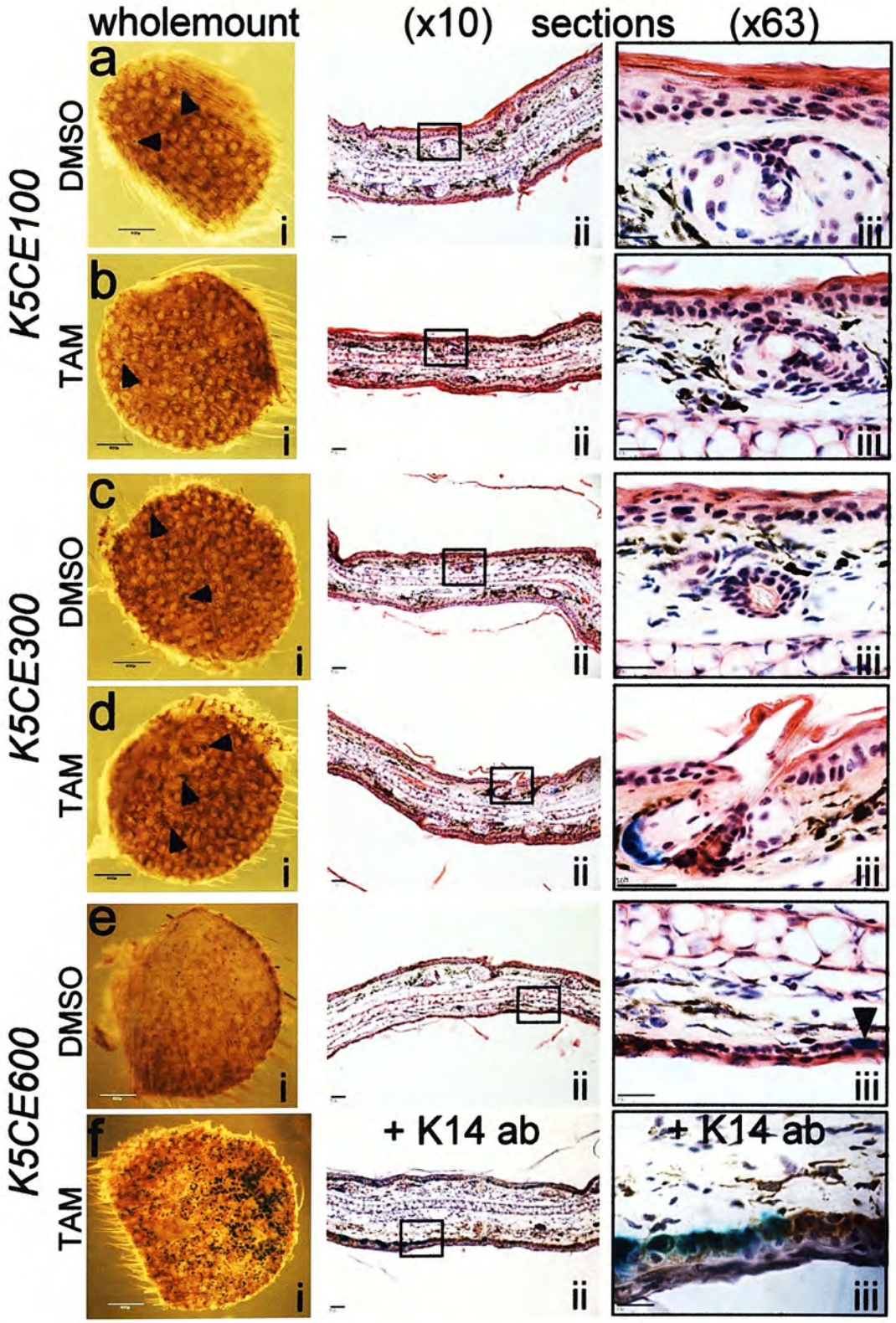


Figure 2.17 Results of functional assay for Cre-mediated recombination after topical application of tamoxifen to mouse pinna (continued over page).

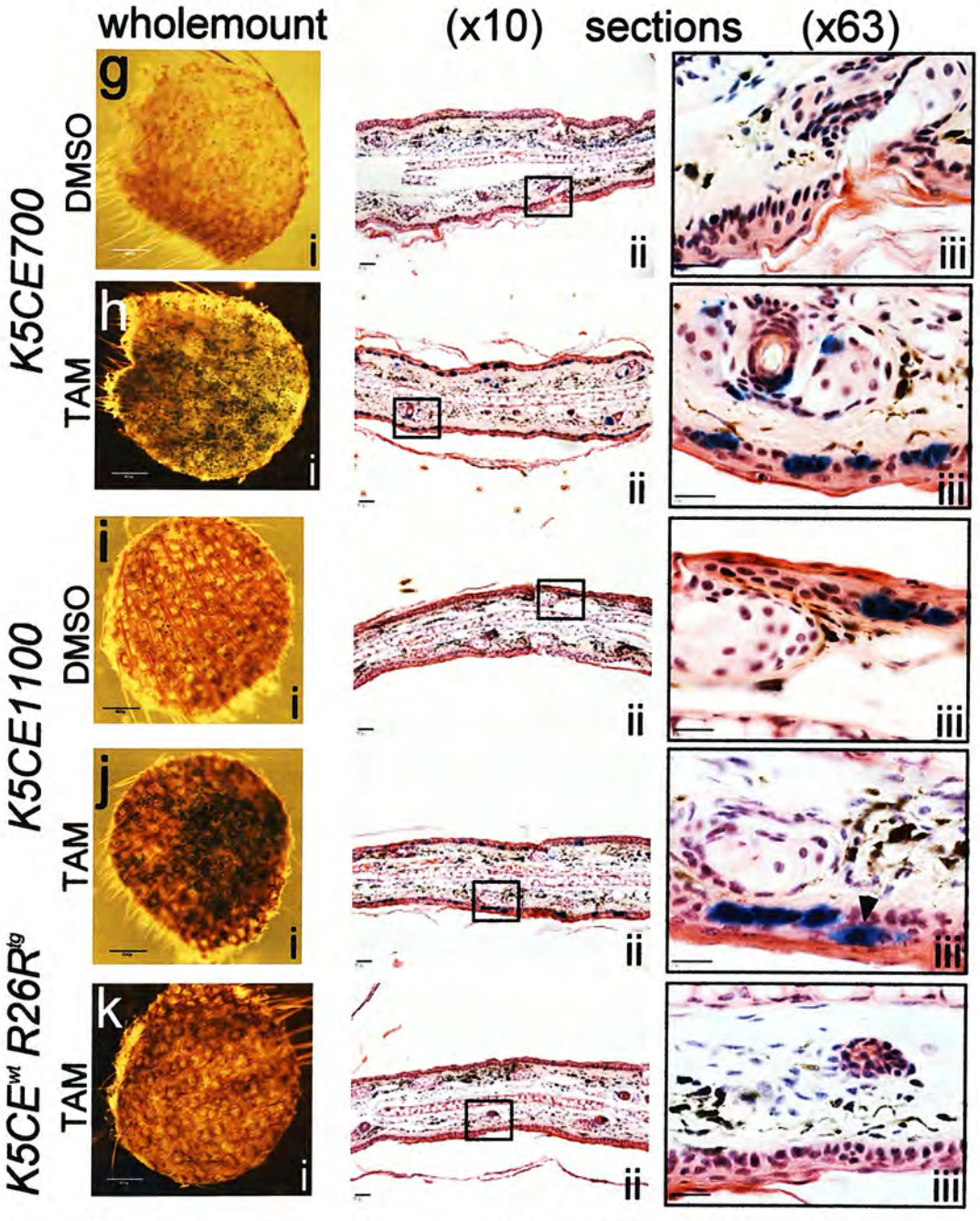


Figure 2.17 Results of functional assay for Cre-mediated recombination after topical application of tamoxifen to mouse pinna (continued from previous page).

Representative wholemount and sectioned ear-clips are shown for *K5CE^{tg} R26R^{tg}* animals from five *K5CE* lines (a-j(i-iii)), plus a *K5CE^{wt} R26R^{tg}* control (k(i-iii)). Sections are shown at x10 magnification (a-k(ii)), which reveals the general frequency of staining, with boxed regions magnified to x63 (a-k(iii)) to demonstrate epidermal localisation of βgal-positive cells. Staining was equivalent after treatment with 5mg or 10mg tamoxifen (TAM). For each line, the TAM and DMSO treated ear-clips exhibiting the most extensive staining are shown. *K5CE500,800* are not included because *K5CE500* showed very similar staining to *K5CE100*, and no staining was evident in *K5CE800*. Infrequent stained cells were visible on whole mount ear-clips from *K5CE100,300* (arrowheads, a-d(i)). While slightly more cells were evident on the latter, there was no obvious difference between the TAM and DMSO treated ear-clips of either line. No βgal-positive cells were encountered during sectioning of *K5CE100* earclips (a,b(i,ii)), though stained hair follicle cells were detected on a TAM treated *K5CE300* section (d(i,ii)). DMSO treated ear-clips from *K5CE600,700,1100* showed similar background of uninduced recombination (e,g,i(i)) with stained cells visible in sections from *K5CE600,1100* (e,i(ii,iii)). After TAM treatment, significantly elevated numbers of stained cells were observed in all three lines (f,h,j(i-iii)). As confirmed by anti-K14 antibody staining on *K5CE600* sections, βgal-positive cells were generally confined to the basal epidermis, however an apparently supra-basal stained cell was evident (arrowhead, j(iii)). No βgal activity was detected in *K5CE^{wt} R26R^{tg}* ear-clips treated with TAM (k(i-iii)). Wax sections were cut at 10μm and counterstained with haematoxylin and eosin, except (f(ii,iii)), haematoxylin and anti-K14 antibody. Scale bar = 400μm (a-k(i)); 50μm (a-k(ii)); 20μm (a-k(iii)).

Sectioning of the stained ear-clips revealed βgal to be mostly confined to the basal epidermal layer in *K5CE600,700,1100* (Figure 2.17e-j(ii-iii)), indicating that differentiation of recombined basal cells into supra-basal descendants had not occurred in during the 5-day experiment. Due to the low frequency of stained cells on *K5CE100,500,800* ear-clips, and on DMSO-treated ear-clips from lines *K5CE600,700,1100*, few or no stained cells were encountered upon sectioning of these tissues. Despite a low incidence of βgal expression in *K5CE300*, xgal stained hair follicles were apparent in one ear-clip section (Figure 2.17d(iii)). Anti-K14 immunostaining of a TAM treated *K5CE600* ear-clip provided additional confirmation of the basal location of βgal expressing cells (Figure 2.17f(ii,iii)).

The unambiguous correlation between TAM treatment and the number of cells exhibiting βgal activity detected after pinna biopsy, strongly suggested that xgal staining was the result of induced Cre activity facilitating transcription of *lacZ* from recombined *R26R* loci. An association between βgal production following TAM administration, and excision of the floxed component of the *R26R* transgene, was subsequently demonstrated using a PCR assay (chapter 3, Figure 3.10).

2.3.6 Determination of relative transgene copy number

Although six of seven *K5CE* lines expressed the *CreER*^{T2} transcript, only three lines demonstrated inducible recombinase activity. Given this apparent lack of a clear correlation between transcription and translation of *CreER*^{T2} (assuming that the inability to detect recombinase activity reflected an absence of CreER^{T2} protein), variation in *K5CE* transgene copy number between the lines was assessed to determine whether this was a more useful predictor of functional *K5CE* lines. This was achieved by Southern hybridisation of BglII digested genomic DNA (gDNA) with a radiolabelled *Cre*-specific PCR probe (Figure 2.18). Densitometric quantification of signal intensities was achieved using a phosphorimager. As a result of the difference in specific activities of the *Cre* and *CFTR* probes, evident from Figure 2.18d(ii), accurate determination of absolute copy number, by assuming the *CFTR* signal was representative of two gene copies and comparing with *Cre* signal,

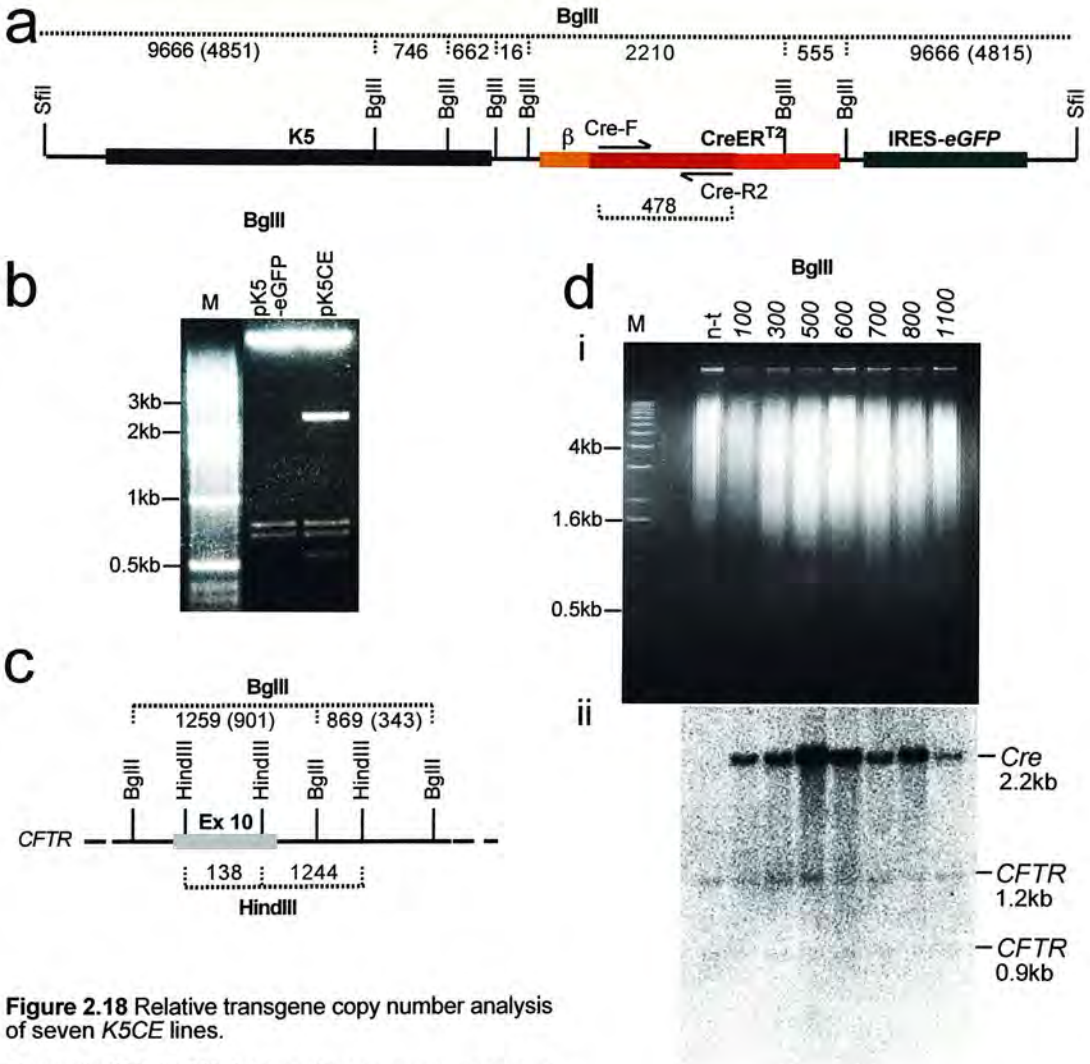


Figure 2.18 Relative transgene copy number analysis of seven *K5CE* lines.

Genomic DNA (gDNA) from mice representing seven *K5CE* lines, plus non-transgenic (n-t) littermate gDNA, was digested with BglII, transferred to a nylon membrane, and hybridised with probes specific for *Cre* and *CFTR*. *CFTR* was used as a known single-copy gene with which to normalise *Cre* hybridisation results for differences in DNA loading. (a) Map of SfiI linearised *K5CE* construct showing position of Cre-F – Cre-R2 PCR primers, within the *Cre* portion of *CreER^{T2}*, with respect to BglII restriction sites. The *Cre* PCR product was predicted to hybridise specifically to the 2210bp BglII fragment (shown in green). (b) BglII digest of pK5CE, confirming predicted restriction pattern (16bp fragment too small to visualise). Digestion of pK5-eGFP confirms that 2.2kb and 0.6kb fragments arise from subcloned *CreER^{T2}* element. (c) Map of BglII and HindIII sites surrounding exon 10 (Ex 10) of *CFTR* on chromosome 6. The *CFTR* probe was a 1244bp HindIII fragment (shown in blue), which is bisected by a BglII site. Thus, the probe was predicted to hybridise to 1259bp and 869bp BglII fragments, with which it contains 901bp and 343bp of complementarity, respectively (brackets). (d) BglII digested gDNA was separated on 1.8% gel, which revealed comparable loading of samples (i). (ii) Phosphorimager representation of membrane after hybridisation with radiolabelled *Cre* PCR probe and *CFTR* restriction fragment probe. *Cre* signal was robust in all *K5CE* samples (no signal evident in n-t control), but both *CFTR* fragments appeared faint. (e) Densitometric analysis was performed by dividing each *Cre* value by the corresponding 1.2kb *CFTR* value. The ratio was then simplified by setting the lowest value (*K5CE1100*) to 1, adjusting the others proportionally, and rounding off (Adj.). This does not indicate absolute transgene copy numbers, but reflects the fold difference in relative numbers between lines. M = 1kb ladder. Map not to scale, fragment sizes in bps. Bracketed fragment sizes indicate proportion of total fragment size.

e Densitometry values

K5CE Line	*Cre*	*CFTR*	*Cre/ CFTR*	Adj.
n-t	6887.68	8955.67	0.8	-
100	19281.21	10921.46	1.8	1
300	31591.99	13480.22	2.3	1
500	1595074.8	14434.65	11.1	7
600	53724.28	12641.39	4.2	3
700	21862.52	10635.87	2.1	1
800	48663.77	8902.74	5.5	3
1100	13991.98	8466.66	1.7	1

was not valid. However, after normalisation of the *Cre* hybridisation signal from each line with the corresponding *CFTR* signal, which controlled for discrepancies in DNA loading and lane-specific background hybridisation, values representing approximate fold differences in copy number were obtained (Figure 2.18e). Four of the lines (*K5CE100,300,700,1100*) were observed to have approximately equal numbers of transgenes, with two lines exhibiting 3-fold higher numbers (*K5CE600,800*). The remaining line (*K5CE500*) was calculated to carry 7-fold more copies than the four lowest copy number lines. F. Kilanowski provided some technical assistance with the radioactive work.

2.3.7 Analysis of K5CE expression in the trachea

Having identified three *K5CE* lines in which epidermal activation of a ubiquitous reporter transgene could be induced through the administration of exogenous ligand, it was desirable to determine whether the *K5CE* transgene was expressed in the trachea. RNA samples were prepared from *K5CE600,700,1100* tracheas, as well as from non-transgenic control and *bK5Cre* tracheas. The *bK5Cre* line had previously been used in anti-Cre immunofluorescence experiments as a positive control known to express Cre in basal epithelia, Figure 2.14. The bovine K5 promoter driving expression of Cre in *bK5Cre* had been reported to drive weak expression of a bK5-*lacZ* construct in embryonic trachea, though no βgal had been detected in adult trachea (Ramirez *et al.*, 1994). While a low level of Cre-dependent reporter activity had previously been reported in the trachea of *bK5Cre* mice, the pattern of recombination or Cre expression had not been characterised (Brakebusch *et al.*, 2000).

Oesophagus and thyroid tissue was carefully removed from tracheas to minimise contamination of the tracheal RNA, as at least oesophagus is known to express endogenous K5, though it has been shown that oesophageal βgal activity was poor in two independent K5-*lacZ* transgenic lines (Byrne and Fuchs, 1993). RT-PCR was performed using *Cre* primers as described in Figure 2.15 (Figure 2.19). *Cre* transcript was detected in tracheas from all three lines after 45 cycles of PCR, though notably

more product was generated from *bK5Cre* tracheal RNA. No product was visible in non-transgenic controls. The identity of the *Cre* amplification product was confirmed by hybridisation with a radiolabelled internal *Cre* oligonucleotide (Figure 2.19c).

Detectable expression of *CreER*^{T2} in the trachea of each of the three lines, *K5CE600,700,1100*, which demonstrated TAM inducible activity of CreER^{T2} protein, suggested their potential usefulness in tracheal lineage analysis.

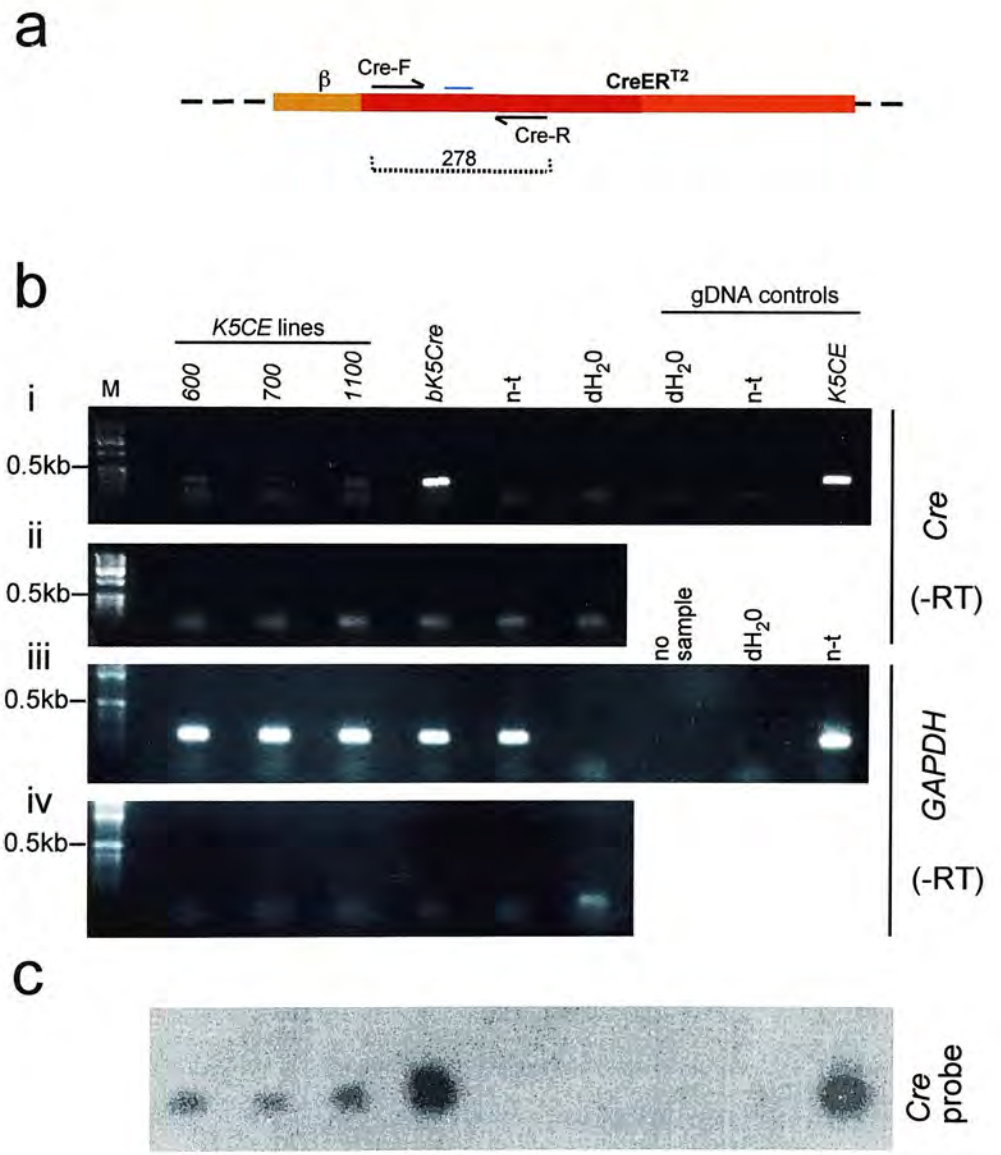


Figure 2.19 RT-PCR for expression of *CreER^{T2}* in trachea of three *K5CE* lines.

RNA was obtained from tracheas of the three *K5CE* lines which demonstrated significant inducible Cre activity in the tamoxifen ear-painting assay (*K5CE600,700,1100*). Care was taken to remove thyroid and oesophagus tissue from tracheas prior to RNA extraction. RNA was DNase-treated before cDNA was prepared and equal amounts used as template in *Cre*-specific PCRs. *bK5Cre* tracheal RNA was included to assess whether this line expressed the transgene in the trachea. (a) Schematic of *CreER^{T2}* portion of the *K5CE* transgene, showing position of internal *Cre* oligonucleotide probe (blue bar) with respect to *Cre*-specific PCR primers. (b(i)) At 45 cycles, PCR product was visible in three *K5CE* lines, though the *bK5Cre* sample produced a stronger product. No product was generated by non-transgenic (n-t) littermates or by the dH_20 control subjected to reverse transcription. Non-transgenic and dH_20 samples used only in PCR step were also negative, while product was generated from *K5CE* genomic DNA (gDNA). (ii) RNA and control samples put through reverse transcription in absence of reverse transcriptase (RT) enzyme (-RT) were all negative, indicating the absence of contaminating gDNA. (iii) Using *GAPDH* primers to control for RNA quality generated comparable PCR products from all samples, while dH_20 controls remained negative. (iv) *GAPDH* -RT reactions were negative. (c) RT-PCR products in (b(i)) were transferred to a nylon membrane and hybridised with a labelled 18mer internal *Cre* oligonucleotide (a) which confirmed the identity of the PCR product and the absence of low levels of product in negative controls. M = 1kb ladder. Map not to scale, fragment sizes in bps.

2.4 Discussion

The generation and analysis of a novel transgenic line (*K5CE*) designed to permit temporal induction of lineage marking within a subset of tracheal basal cells, thereby circumventing ambiguities arising from potential developmental or adult steady-state transgene expression, has been described. A range of techniques has been employed in order to select potentially useful *K5CE* lines. This was preceded by a description of reporter expression in two *K5A1* lines, produced to allow the evaluation of the subcloned human K5 promoter which drives expression of a fusion gene encoding an inducible Cre protein in *K5CE*.

2.4.1 The transgenic K5 promoter-driven control line, K5A1

2.4.1.1 Expression of eGFP

Two of four *K5A1* lines demonstrated detectable transgene expression. In the expressing lines, patterns of eGFP production, in embryonic skin and adult pinna, and βgal activity, in embryonic and adult basal epidermis, were broadly consistent with reports in which the transgenic behaviour of the 6.2kb human K5 promoter was well characterised, and shown to be largely faithful to that of the endogenous K5 gene (Byrne and Fuchs, 1993; Byrne *et al.*, 1994). The absence of detectable eGFP in adult tissues other than the pinna is likely the result of insufficient translation of eGFP from the IRES sequence to permit visualisation of eGFP signal above high levels of autofluorescence. Problems associated with the expression of downstream sequences in bicistronic constructs containing IRESs are not uncommon, and are discussed further in chapter 5 section 5.5.5. However it has recently been reported that, even in mice where eGFP is expressed directly from the K5 promoter, sensitive and consistent detection of specific fluorescence requires the use of anti-eGFP antibodies (Schoch *et al.*, 2003).

2.4.1.2 Expression of βgal

Subtle differences in the timing of the onset of K5-driven βgal expression, where *K5A1100* and *K5A1200* appeared "late" and "early", respectively, by comparison with the description of Byrne and colleagues (Byrne *et al.*, 1994), may have been simply due to several hours difference in developmental age of embryos representing the same day of embryonic development. Alternatively, epigenetic factors responsible for the mosaic expression, evident to varying extents in both lines, may also have perturbed the precise temporal regulation of transgene activation. Variegated expression of transgenes in a population of equivalent transgene-bearing cells is not uncommon, though it is not well understood, and has been variously attributed to integration within pericentromeric chromatin (Festenstein *et al.*, 1996; Milot *et al.*, 1996; Dobie *et al.*, 1996), epigenetic effects on high copy number arrays (Garrick *et al.*, 1998; Lau *et al.*, 1999), and the prokaryotic origin of certain transgenes (Clark *et al.*, 1997), particularly *lacZ* (Cui *et al.*, 1994). Unanticipated expression of *lacZ* constructs in transgenic mice is discussed further in chapter 5 section 5.5.6.

Of particular interest is the observation that a *lacZ* transgene driven by the bK5 promoter, also used in the *bK5Cre* mouse, was expressed in a consistently variegated fashion, in a manner similar to lines *K5A1100,200* (Ramirez *et al.*, 2001). The reproducibly mosaic expression pattern in animals bearing this transgene provided the authors with a means to investigate the effects of several parameters on the extent of observed variegation. While there was some evidence that pericentromeric integration sites could mediate variegation, it was apparent that mosaic expression also occurred from transgenes integrated far from the centromeric heterochromatin. Using a Cre/loxP approach to reduce the number of transgene copies within an integrated array did not have an effect on the proportion of expressing cells, suggesting that, at least in that system, repeat induced gene silencing was unnecessary for transgene repression. Given the non-mosaic expression of cDNAs from the bK5 promoter, including *cyclin D1* (Robles *et al.*, 1996), *IGF-1* (DiGiovanni *et al.*, 2000), *leptin*, (Larcher *et al.*, 2001), and *E2F1/4* (Wang *et al.*,

2000), the authors speculated that the prokaryotic *lacZ* sequence was responsible for the variegated expression (Ramirez *et al.*, 2001). While this is consistent with the observations of Clark and colleagues regarding transgenic expression of prokaryotic reporter sequences, the same group also noted a similar effect using mammalian cDNAs, concluding that such intron-less sequences were unable to attenuate the formation of local repressive chromatin structures (Clark *et al.*, 1997). Reports of non-mosaic bK5-driven expression of the prokaryotic, intron-less *Cre* gene ((Brakebusch *et al.*, 2000; Berton *et al.*, 2003) and A. Ramirez, unpublished results) indicate that susceptibility to variegation is not a feature of all prokaryotic sequences. Severe repression of a *lacZ* transgene expressed from the human K18 promoter (Thorey *et al.*, 1993) may imply that the mechanisms resulting in mosaic expression are particularly stimulated by keratin promoter-*lacZ* constructs. Despite the observations of Ramirez and colleagues (Ramirez *et al.*, 2001), no variegation was reported for the human K5 promoter-driven *lacZ* mice characterised by Byrne and co-workers (Byrne and Fuchs, 1993; Byrne *et al.*, 1994): however, the fact that only two of nine βgal-producing lines were selected for further analysis may suggest sub-optimal βgal expression was observed in some of the remaining lines. Depending on the relative influence of the *lacZ* sequence and position effect on expression, non-mosaic *K5A1* lines may have been obtained had more than four founders been created. Nevertheless, differences in construction of the K5-*lacZ* (Byrne and Fuchs, 1993) and *K5A1* transgenes, including the amount of flanking vector sequence, cannot be discounted as factors that may have influenced the uniformity of expression.

While βgal activity appeared confined to the basal epidermal layer in *K5A1100*, bluogal stain was noted to be present throughout *K5A1200* embryonic and adult pinna epidermis, though staining often appeared more intense basally, and non-stained supra-basal cells were occasionally visible. Despite a similar pattern of staining in *K5A1100* vibrissae pad epidermis, anti-βgal immunofluorescence detected βgal only in the basal layer. Ramirez and colleagues reported precisely the same observation in their most highly expressing bK5-*lacZ* lines, and suggested that persistence of βgal below the threshold of immunodetection, in cells migrating

upwards from the basal layer, might account for the disparity in histochemical and immunofluorescence data (Ramirez *et al.*, 1994). An analogous process is thought to explain the appearance of basal cell-specific keratin filaments in supra-basal cells lacking the corresponding mRNAs (Schweizer *et al.*, 1988), though keratin filaments are likely to be significantly more durable than βgal (Lersch *et al.*, 1989). The authors conceded that, due to an absence of supra-basal xgal staining in some highly expressing lines, a low level of inappropriate transgene expression in supra-basal layers of those lines exhibiting supra-basal xgal staining, could not be discounted (Ramirez *et al.*, 1994). This could occur if the promoter component of the microinjected transgene had undergone degradation or re-arrangement prior to integration. It has been shown that a truncated 90bp version of the 6.2kb human K5 promoter is still competent to direct largely keratinocyte-specific expression, but that differentiation specificity switches from basal to supra-basal cells, implying the presence of elements promoting basal specificity, and suppressing supra-basal specificity, in the remainder of the promoter fragment (Byrne and Fuchs, 1993). Loss of the latter regulatory sequences would result in pan-epidermal expression. Although intact *K5A1* transgene 5' termini were shown to be present in both *K5A1100,200* founders, this does not preclude 5' degradation in a subset of transgene copies. While the persistence of basally-produced βgal, or misexpression of βgal production, could underlie this observation in either this study or that of Ramirez and colleagues (Ramirez *et al.*, 1994), examination of bluogal stained and sectioned embryos and adult pinna from *K5A1* lines suggested leaching of metabolised bluogal precipitate from the basal epidermal layer, and outer follicle layer, into surrounding tissues including the supra-basal epidermal layers, is the most likely explanation. It would be desirable to probe *K5A1200* expression with the anti-βgal antibody to determine whether the immunofluorescence signal was basal layer-specific. The occasional difficulty in designating stained cells as basal or supra-basal is a consequence of the fine profile of embryonic, and adult pinna, epidermis, the same property which renders them conveniently more permeable to fixing and staining reagents than other adult epidermal tissues.

2.4.1.3 Endogenous βgal expression in tracheal submucosal glands

Upon examination of *K5A1* tracheas to determine whether βgal expression in these lines recapitulated the SMG duct expression described by Borthwick and co-workers (Borthwick *et al.*, 2001), xgal staining was consistently localised to the proximal trachea. However, a similar pattern, albeit less intense, was observed in stained non-transgenic littermate tracheas. The results of SMG staining are difficult to interpret. While low levels of endogenous staining were observed during analysis of *ROSA26*/wild-type chimaeric SMGs (Borthwick, 1999), non-specific glandular staining was not reported when K5-*lacZ* expression was detected in SMG ducts (Borthwick *et al.*, 2001). However, it has since become apparent that endogenous tracheal xgal staining can be problematic, and results from a reaction between mucous and the strongly negatively-charged xgal (S. Randell, personal communication). Similarly, endogenous activity in thyroid glands is due to thyroglobulin. While Borthwick and colleagues reported K5 promoter-directed βgal expression primarily in gland duct cells proximal to the tracheal luminal surface (Borthwick *et al.*, 2001), staining in *K5A1* glands appeared more distal (that is, further from the tracheal luminal surface). Particularly in *K5A1200*, patches of staining which were more intense than observed in non-transgenic glands appeared localised to regions previously shown to produce mucous (Borthwick, 1999), suggesting it may represent endogenous rather than transgene-specific activity. Alternatively, the more intense and widespread staining in *K5A1* SMGs could indicate that tracheal activity of the K5 promoter is higher in these lines than in the single line examined by Borthwick and colleagues (Borthwick *et al.*, 2001), which was one of the well characterised lines produced by Byrne and co-workers (Byrne and Fuchs, 1993), though the latter group did not report any tracheal expression.

Although Borthwick and colleagues confirmed the presence of K5 protein-expressing cells in tracheal SMGs by immunofluorescence (Borthwick *et al.*, 2001), until recently the possibility remained that the tracheal βgal activity they observed was confined to the single transgenic line used in that study (though the fact that this line was one of two previously shown to express βgal in an identical, K5-faithful pattern

(Byrne and Fuchs, 1993), might suggest that the tracheal expression would have been shared between the lines). However, Schoch and colleagues recently generated a K5-*eGFP* mouse, using the same configuration of promoter and intron elements used in the *K5A1* lines, which expressed eGFP in tracheal gland duct and surface epithelial cells near gland duct openings in the upper trachea, and in systematically arrayed foci in the lower trachea centred over cartilage-intercartilage junctions (Schoch *et al.*, 2003). This pattern is reminiscent of the distribution of βgal-positive cells in the K5-*lacZ* trachea, and the subsequently discovered BrdU LRCs (Borthwick *et al.*, 2001). That K5 promoter-directed transgene expression to SMG ducts has now been observed in two independent lines carrying different expressed sequences suggests that a subset of cells in the proximal trachea represent a *bona fide* site of K5 promoter activity. If this is the case, then, given the apparently K5-specific basal epidermal βgal expression observed in the *K5A1* lines, they are likely to also demonstrate the appropriate SMG expression. While some of the SMG staining observed was thus likely to have been the result of *K5A1* activity, the unexpectedly high levels of glandular staining probably reflected a degree of endogenous staining.

Attempts to specifically detect βgal by reducing the xgal incubation temperature either had no effect or completely eliminated staining, and anti-βgal antibodies were insufficiently sensitive on wax sections to be informative (data not shown). A better approach may have been to increase the pH of the xgal staining solution into the alkaline range, though this technique was designed to minimise the activity of endogenous lysosomal βgal which has an acidic pH optimum (Weiss *et al.*, 1999), whereas some of the mucous found in SMGs are neutral in pH. At any rate, these observations underlined the importance of including non-transgenic littermate control samples in all subsequent tracheal xgal staining experiments.

Although apparent endogenous xgal staining undermined the certainty with which K5 promoter-directed βgal expression in SMGs could be identified, the *K5A1* lines were nevertheless useful in demonstrating appropriate activity of the subcloned K5 promoter at least in epidermis, from which SMG expression might reasonably be extrapolated on the basis of similar transgenic lines employing the K5 promoter

(Byrne and Fuchs, 1993; Borthwick *et al.*, 2001; Schoch *et al.*, 2003). While anti-βgal immunodetection would be required to definitively confirm basal cell-specific expression in *K5A1200*, the potential lack of fidelity for basal cells may not affect SMG expression, if it is the presence of basal layer-specific promoter elements that drive tracheal expression. Of course, transgene expression in SMGs of the two *K5A1* lines was only important in its capacity to predict equivalent expression of the $CreER^{T2}$ fusion gene in the *K5CE* lines where, in the absence of visible eGFP fluorescence, the expression profile of the transgene is more difficult to ascertain. Given that considerably more *K5CE* than *K5A1* lines were produced (29 to 4, respectively), the frequency of *K5CE* lines showing appropriate K5 promoter-directed expression should be correspondingly greater.

While it was recognised that the occurrence of significant levels of endogenous staining in tracheas lacking the *K5A1* transgene could potentially complicate the results of βgal-based lineage analysis in experiments using $K5CE^{tg}$ $R26R^{tg}$ mice, it was envisioned that a marked lineage would be identifiable on the basis of the strength, consistency, and pattern of staining from the ubiquitous promoter of the *R26R* conditional reporter. It has previously been shown that *ROSA26*-derived SMGs were unambiguously distinguishable from neighbouring, endogenously staining wild-type-derived SMGs in chimaeric tracheas (Borthwick, 1999).

2.4.2 Identification of K5CE lines demonstrating inducible Cre-mediated recombination

Expression analysis data on the seven *K5CE* lines selected for further study is summarised in Table 2.3. That such an extensive analysis was warranted, using several approaches, was partly due to the absence of eGFP signal either in embryos or adult tissues. IRES-dependent gene expression in bicistronic constructs is known to be unpredictable, and is discussed further in chapter 5 section 5.5.5. While, in principle, a functioning IRES-*eGFP* might have permitted the tissue-specific activity of the transgene to be visually assessed, abrogating the need for RT-PCR analysis, experience from the *K5A1* lines indicated that eGFP levels would unlikely have been

K5CE line	Relative transgene copy number	*CreER*^{T2} mRNA (back skin)	Inducible Cre activity (pinna)	*CreER*^{T2} mRNA (trachea)
100	1	+	-	N/D
300	1	+	-	N/D
500	7	+	-	N/D
600	3	+	+	+
700	1	+	+	+
800	3	-	-	N/D
1100	1	+	+	+

Table 2.3 Summary of expression analysis on seven *K5CE* lines.

N/D = not determined.

sufficient to be visually detectable above background autofluorescence in adult tissues, including the trachea. Nevertheless, eGFP in embryonic skin would have proved a useful initial screen for expressing lines.

Despite the absence of eGFP expression, the fact that the easily accessible epidermis is the primary site of K5 expression could still be exploited in the identification of lines expressing the *K5CE* transgene. It was reasoned that, if the tracheal SMGs were a genuine site of K5 promoter activity, then the most effective strategy to select useful lines would be to identify those demonstrating functional expression of the inducible Cre protein in the epidermis, before evaluating tracheal expression in this smaller subset of lines.

2.4.2.1 Epidermal expression of CreER^{T2} *transcript*

To that end, transcription of *CreER*^{T2} in back skin was evaluated by RT-PCR. Since it was not feasible to analyse mice derived from all 29 founders, seven known transmitting lines derived from female founders crossed with *R26R* males, were examined. All but one of the lines generated a *Cre*-specific PCR product. Although the product generated from line *K5CE300* appeared significantly less than that from the five remaining expressing lines, and despite using equal amounts of cDNA template in each PCR, the use of a saturating number of thermal cycles in the *GAPDH* positive control reaction masked any evidence of technical inaccuracies, preventing the confident inference of relative transcript levels between the lines. Thus, the lines were simply considered either to express the transgene (*K5CE100,300,500,600,700,1100*) or not (*K5CE800*). While it had been anticipated that the RT-PCR experiment would have excluded several lines from evaluation in the subsequent, significantly more laborious, functional assay, only *K5CE800* was shown to be negative for *CreER*^{T2} transcription at the number of PCR cycles used. It was decided to include *K5CE800* as an additional negative control.

2.4.2.2 Tamoxifen-inducible recombination in adult epidermis

The functional assay for TAM-inducible recombination involved a modification of the topical administration of TAM described by Vasioukhin and colleagues (Vasioukhin *et al.*, 1999) (B. Stripp, personal communication). TAM was used, rather than its significantly more expensive metabolite OHT, which have both been shown to be similarly effective at inducing Cre activity (Danielian *et al.*, 1998; Metzger and Chambon, 2001). DMSO was used as the solvent for its known penetration enhancing properties (Caspers *et al.*, 2002). The assay demonstrated an association between the application of TAM and production of basal cell-localised βgal, as detected by xgal staining. Formal proof that βgal activity after TAM administration resulted from recombination at *R26R* loci was subsequently shown by PCR (chapter 3, Figure 3.10). Although the results of this assay were unambiguous in indicating inducible Cre-mediated recombination in three of the seven lines evaluated, problems with TAM solubility prevented inferences being confidently made into the relative recombination efficiencies of the different lines, and any differences resulting from the two TAM concentrations used. Although not mentioned in reports describing TAM inducible activation of Cre fusion proteins, many researchers have anecdotally reported difficulties in obtaining solutions of TAM or OHT in oil (with or without ethanol), which is normally used as the carrier for the more common systemic administration (L. Bugeon, G. McLean, personal communications). Similar problems were encountered with DMSO in this experiment. At the concentrations used, TAM in DMSO was an opaque, paste-like suspension (previously noted by B. Stripp, personal communication). It proved impossible to dispense the entire volume, or prevent some of the suspension from remaining on the cotton bud used to "paint" the suspension over the pinna. Thus administered doses were not accurate, though approximately equivalent staining between mice, and between the pinna of individual mice, of the same line suggested that the general differences in staining between lines, as detected by visual inspection, might reflect genuine differences in protein activity. Because of the solubility issue, it is unclear whether the degree of recombination observed in lines *K5CE600,700,1100* was maximal, or was limited by levels of ligand or by expression

from the K5 promoter. In the absence of informative anti-Cre immunofluoresence data, it is not known whether the *K5CE* lines share the variegated expression of the *K5A1* lines.

Low levels of recombination mediated by CreER fusions in the absence of inducer, as observed in this experiment, have previously been reported (Vasioukhin *et al.*, 1999; Hayashi and McMahon, 2002). In response to speculation that this background activity might result from proteolytic cleavage of the linker joining the Cre and ER domains, leading to unregulated Cre activity, a fusion protein was created in which Cre was flanked at both amino- and carboxy-terminal ends by mutated mouse ERs (MerCreMer, (Zhang *et al.*, 1996)). This resulted in negligible ligand-independent recombinase activity, though there was disagreement over whether this reflected an overall reduction in inducible activity of the bulky protein (Verrou *et al.*, 1999; Kellendonk *et al.*, 1999). More recently, an enhanced PR-regulated Cre was produced in which a reduction in the length of the Cre-PR linker, and removal of a cryptic splice donor site in the 3' of the Cre portion, effectively eliminated background activity without affecting the maximum induced activity (Cre*PR, (Wunderlich *et al.*, 2001)). When Cre*PR was compared *in vitro* with CrePR1, the former demonstrated approximately 50-fold lower background activity, a significantly expanded range of inducibility, and enhanced dose-response characteristics. If the activity profile of Cre*PR is recapitulated in transgenic mice, then it will be the construct of choice for studies requiring inducible site-specific recombinase activity. However, because of the relatively low number of SMG cells expected to express the transgene, the observed level of uninduced recombination appeared unlikely to be problematic in this study.

2.4.2.3 Functional expression of CreER^{T2} did not correlate with transgene copy number

It is intriguing that, while expression of *CreER^{T2}* was detected in six of seven lines, only three of the expressing lines demonstrated inducible Cre-mediated recombination. It was shown that differences in recombinational activity were not

likely to be due to increased levels of *R26R*, or *K5CE*, homozygosity in certain lines. Due to the non-quantitative RT-PCR protocol used, observed recombination rates could not be confidently related to transcript levels. Thus, it may simply have been that differences in transcriptional activity not accurately reflected by the RT-PCR data were responsible for proportionally different levels of $CreER^{T2}$ protein. It would be desirable to assess transcription levels by quantitative or semi-quantitative RT-PCR to address this. One means by which apparently equal transgene expression, as determined by quantitative RT-PCR, could translate into variation in protein activity would be if deletions or other deleterious re-arrangements had occurred within transgene loci in a subset of transgenic lines. Multimeric transgene arrays can form inverted repeat structures if transgenes are arranged in head-to-head or tail-to-tail configurations. Such structures have been shown to be unstable and prone to high levels of re-arrangement, including partial deletions (Collick *et al.*, 1996). Since the RT-PCR primers used amplified only a part of the *Cre* portion of the $CreER^{T2}$ transcript, it would not reveal deletions outside this region, such as in the ER domain. In this way, apparently similar levels of $CreER^{T2}$ transcription could result in markedly different levels of functional protein. Since higher copy number arrays are more likely to contain inverted repeat structures (Collick *et al.*, 1996), and are thus more susceptible to this form of structural instability than lower copy number arrays, this would be consistent with the apparent absence of a proportional relationship between transgene copy number and functional protein in the *K5CE* lines. The highest copy number line (*K5CE500*) harboured approximately seven-fold more copies than the lowest copy number lines (*K5CE100,300,700,1100*), yet showed no increase in recombination above background levels in response to TAM, despite demonstrating transcription of $CreER^{T2}$. Two of the lines which demonstrated significant inducible recombination carried the least transgene copies (*K5CE700,1100*), while the remaining functional line carried approximately three-fold more (*K5CE600*), equal to another line which showed neither transcription of $CreER^{T2}$, nor induced or uninduced Cre activity (*K5CE800*).

The proposed mechanism of transgene array instability would not explain an observed inverse correlation between copy number/transcription and protein levels in

transgenic lines carrying the MerCreMer inducible Cre fusion protein (Verrou *et al.*, 1999). A line carrying many transgene copies and expressing high levels of mRNA did not generate detectable protein until some copies (or a transgene locus) were lost through segregation, which coincided with high levels of protein production (L. Bugeon, personal communication). Whether the lack of inducible recombinase activity in some lines which expressed *CreER^{T2}* was a result of low or absent protein levels, or was indicative of mutations in the *CreER^{T2}* gene, could be investigated by *in situ* immunodetection or Western blotting of tissue extracts using an anti-Cre antibody known to bind the CreER^{T2} fusion protein. Such an antibody exists (Doerflinger *et al.*, 2003), but failed to detect protein in the skin of *K5CE* mice in preliminary experiments (data not shown). Conventional Cre protein in *bK5Cre* skin was also not detected, indicating that a technical issue, rather than the absence of CreER^{T2} protein or the inability of the antibody to bind the fusion protein, was responsible. While the antibody could indicate the presence or absence of the protein, it could also be used to assess its inducibility by immunostaining after TAM administration, either in primary cultures derived from transgenic mice, or in tissue sections from animals recently treated with TAM (Indra *et al.*, 1999), and determining the sub-cellular localisation of the protein.

2.4.2.4 *Tracheal expression of* CreER^{T2} *transcript*

Evaluation of tracheal *CreER^{T2}* expression by RT-PCR was limited to the three lines showing activity of CreER^{T2} in pinna epidermis (*K5CE600,700,1100*). While a product was generated from the *bK5Cre* sample at 30 PCR cycles (data not shown), 45 cycles were required to detect transcript in the *K5CE* lines. Although tracheal expression of bK5 promoter-driven transcripts has been alluded to in several reports (Ramirez *et al.*, 1994; Brakebusch *et al.*, 2000), it has not been characterised. The *bK5Cre* mouse is described further in chapters 3 and 4. While the RT-PCR was not quantitative, the absence of detectable amplification product from all three *K5CE* tracheal samples at 30 cycles, compared to robust products from equivalent amounts of skin cDNA template, was consistent with the expectation that the K5 promoter is active at a low level overall in the trachea, which could be a consequence of high-

level expression in a small subset of tracheal cells. Preliminary attempts to detect $CreER^{T2}$ protein using anti-Cre antibodies on tracheal sections were unsuccessful, as predicted by the failure of these antibodies to bind the fusion protein in skin sections, despite the presence of functional protein as revealed by the TAM ear-painting assay. However, the K5-specific expression pattern demonstrated by the subcloned K5 promoter in the *K5A1* lines, in combination with the functional expression data and apparent tracheal transcription of $CreER^{T2}$, suggested that lines *K5CE600,700,1100* were likely to be suitable for their intended purpose of investigating lineage arising from the K5 promoter-expressing cells of the adult tracheal SMG ducts.

Chapter 3 A preliminary investigation of cell lineage in regenerating tracheal epithelium

3.1 Introduction

The generation of a transgenic mouse designed to facilitate monitoring of the lineage arising from a putative population of tracheal stem cells was described in chapter 2. The identification of the submucosal gland (SMG) duct as a possible stem cell niche was based on observations that, following long-term bromodeoxyuridine (BrdU) labelling and an appropriate washout period, label-retaining cells (LRCs) were preferentially localised to SMG ducts in the proximal trachea, and to systematically arrayed foci more distally (Borthwick *et al.*, 2001). Unlike in rapidly renewing systems such as the epidermis (Bickenbach, 1981; Lavker and Sun, 1983), where stem cells are constantly dividing to replenish the layers lost through desquamation, the basal rate of mitosis in steady-state airways is extremely low (Evans and Shami, 1989). Since cell division is a prerequisite for nuclear incorporation of a pulsed label, the identification of LRCs in the similarly quiescent cornea had previously been achieved by induced epithelial injury to stimulate stem cell recruitment (Cotsarelis *et al.*, 1989).

3.1.1 A model for putative tracheal stem cells in submucosal gland ducts

Borthwick and colleagues applied the same approach to the trachea, using two sources of injury in parallel experiments (Borthwick *et al.*, 2001): intratracheal instillation of polidocanol (POL), a detergent clinically used as a sclerosing agent and to facilitate gene transfer to airways (Parsons *et al.*, 1998), and SO_2, a well-studied model of pulmonary epithelial injury (Lamb and Reid, 1968). While the two methods differed in the nature of the damage caused to the epithelial layer, the BrdU labelling index of viable surface epithelial cells 24 hours after treatment exceeded 50% in each case, compared to 1.3% in undamaged trachea. Repair was rapid, with a columnar layer evident within 3 days of both treatments, and re-establishment of a fully ciliated epithelium by 7 days. After four weeks of repeated tracheal damage combined with frequent BrdU injections, tracheas from mice killed 3, 6, 21, and 95 days after the final injury were examined for incorporated label. At the 3 and 6 day

timepoints, many labelled cells were detected throughout the tracheal surface and gland system, though densely stained cells appeared to be preferentially localised to the SMG ducts and surface clusters in both the proximal and distal trachea. By 21 days after damage, the overall BrdU staining was markedly weaker, though the labelling index still exceeded 30%. Cells exhibiting intense staining were localised to the SMG ducts in upper trachea, whereas in the glandless lower trachea, they were predominantly basal and were arranged in foci, which tended to correspond with cartilage-intercartilage junctions.

Closer examination of the proximal trachea from 21 and 95 day timepoint animals revealed that labelled cells were concentrated in gland ducts and the immediately adjacent tracheal luminal surface epithelium (Borthwick, 1999). Three zones of LRCs could be recognised on the basis of staining intensity and morphology. The most heavily labelled cells were morphologically simple (by light microscopy) and found at the base of the SMG duct (zone one). Slightly less-heavily stained basal cells, extending from the top of the ciliated duct into the adjacent tracheal epithelium, comprised zone two. Zone three described cells in a finite region of luminal epithelium beyond zone two. These cells exhibited weaker staining than those of zones one and two, and were of intermediate morphology, lacking contact either with the luminal surface or the basal lamina. This situation appeared to show parallels with the corneal epithelium, in which stem cells in the adjacent limbal region give rise to transit amplifying (TA) cells which migrate through the basal layer, dividing to produce further TA cells and supra-basal, differentiated daughter cells (Lehrer *et al.*, 1998). Since the intensity of retained label within a cell should, in principle, correlate inversely with the frequency with which it has divided, a model was proposed whereby the most densely staining cells of the SMG duct represent tracheal stem cells (Figure 3.1). In response to tracheal epithelial injury, these cells would give rise to a TA population, corresponding to the less intensely staining cells of zones two and three, which would migrate into, and repopulate, the neighbouring epithelium.

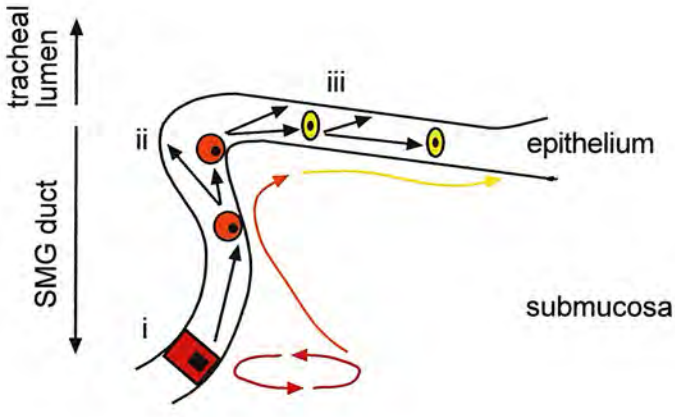


Figure 3.1 Three zone model for label retention around the submucosal gland duct (modified from Borthwick, 1999).

(i) Zone one - intensely-labelled stem cells in the submucosal gland duct divide to self-renew and produce early transit amplifying (TA) cells. (ii) Zone two - early basal TA cells migrate towards the surface epithelium, dividing to produce late TA cells and differentiated cell types. (iii) Zone three - late TA cells of intermediate morphology migrate further from ductal opening, dividing to produce further late TA cells and differentiated cell types.

3.1.2 Do these cells contribute to repair of the tracheal epithelium?

While LRCs have been shown to correspond to tissue-specific stem cells in several other systems such as the epidermis (Bickenbach, 1981), corneal epithelium (Cotsarelis *et al.*, 1989), and hair follicle (Cotsarelis *et al.*, 1990), and the localisation of LRCs to tracheal gland ducts is consistent with observations from other epithelia in which stem cells have been found in specialised, generally well protected niches spatially proximal to their more differentiated progeny (Cotsarelis *et al.*, 1989), these combined observations remain circumstantial, and do not directly indicate a relationship between the SMG duct LRCs and the cells of the reconstituted epithelial layer. Label-retention studies are often combined with *in vitro* clonogenic assays, which measure the capacity of candidate stem cells to form large, self-renewing clones (Morris and Potten, 1994; Oshima *et al.*, 2001). While these studies are useful in revealing the potential potency of subpopulations of cells, they do not indicate progenitorial behaviour of those cells in the context of a renewing or regenerating tissue.

An elegant means to test the model would be to mark the putative stem cells of the SMG ducts and follow their descendents, if any, in response to induced injury. As discussed in chapter 2, due to the spatial coincidence of keratin-rich cells with tracheal LRCs, in particular the ability of the human K5 promoter to drive βgal expression in a pattern apparently consistent with these cells (Borthwick *et al.*, 2001), and given the precedents for the presence of specific keratins to distinguish stem cells from their differentiated derivatives (Schermer *et al.*, 1986; Lyle *et al.*, 1998), it was hypothesised that K5 might represent a marker for this putative stem cell niche (Borthwick *et al.*, 2001). To test this hypothesis, and thus to potentially test the model, transgenic mice carrying a human K5 promoter-driven fusion gene encoding an inducible Cre protein were generated (*K5CE*, chapter 2). When combined with the ubiquitously expressed Cre reporter allele *R26R*, permanent marking of lineages arising from K5 expressing cells can be temporally regulated through the administration of the exogenous ligand tamoxifen (TAM).

This bipartite genetic approach to lineage analysis was used for two reasons pertinent to the context in which lineage was to be tracked. Firstly, the permanence, in principle, of the marking that should be achieved using a ubiquitously expressed reporter, appeared to be a necessary feature of the strategy. This is because the small number of high K5 promoter-expressing, keratin-rich cells within the tracheal epithelium and gland ducts implies that, if these cells are indeed stem or progenitor cells, then their immediately differentiated progeny have ceased K5 expression. As a result, alternative strategies, for example the retrospective *laacZ* technique (Bonnerot and Nicolas, 1993), in which the reporter remains under the transcriptional regulation of the promoter that drives its expression within the progenitors, would not allow lineage to be followed outside the putative stem cell population. Secondly, an inducible, rather than a constitutively active, form of Cre was used to circumvent problems arising from the potential precocious activation of the reporter prior to the onset of the experiment. Since ubiquitous, permanent marking of lineage would be activated in any cell in which the K5 promoter became switched on, K5 expression in developmental precursors of the adult trachea, should it occur, would result in an entirely marked, and thus uninformative, adult tissue. Similarly, if the high K5 expressing population has a role in the maintenance of the steady-state adult tracheal epithelium, the extent of marked cells resulting from constitutive Cre activity could be uninterpretable. The inducible system would prevent lineages being marked until the desired moment, permitting cell fates to be tracked through time by examining tissue at various timepoints after the known induction of reporter activation.

3.1.3 Aims

As a preliminary investigation into the ability of the *K5CE* mouse to facilitate inducible marking of the high K5-expressing population within adult SMG ducts, *K5CE*^{tg} *R26R*^{tg} animals were injected with TAM to stimulate systemic Cre-mediated activation of the ubiquitous reporter. This was preceded by intratracheal instillation of POL as described previously (Borthwick, 1999; Borthwick *et al.*, 2001), to stimulate epithelial repair with the concomitant recruitment of stem cells. Tracheas were removed at various timepoints and xgal stained to visualise patterns of βgal-positive cells.

In order to validate the requirement for an inducible recombinase, the pattern of βgal expression from *R26R* was determined in tracheas from mice in which the bovine K5 promoter directs expression of constitutive (non-inducible) Cre (*bK5Cre*, A. Ramirez, unpublished results). A low level of Cre-dependent reporter activity had previously been reported in the trachea of *bK5Cre* mice (Brakebusch *et al.*, 2000), but the pattern of recombination or Cre expression had not been characterised. The use of *bK5Cre* tracheal RNA in a Cre-specific RT-PCR suggested Cre transcript may be present at higher levels than in corresponding samples from human K5 promoter-driven *K5CE* lines (chapter 2, section 2.3.7). Thus it was anticipated that, if the bK5 promoter mirrored its human counterpart's ability to direct expression within the putative stem cell niche, and indeed if the K5-expressing cells of the SMG duct were in fact airway progenitors, then constitutive Cre activity in the adult trachea would lead to an uninformatively high level of reporter activation.

In fact, as shown in Figure 3.2, while significant βgal activity was apparent in *bk5Cre*^{tg} *R26R*^{tg} tracheas, staining was not as widespread as anticipated, and certainly suggested Cre-mediated recombination resulting from potential bK5 promoter activity in tracheal precursors during development had not occurred. The patches of βgal expression appeared sufficiently small and discrete that available *bK5Cre*^{tg} *R26R*^{tg} mice were used in a pilot study to evaluate their suitability in monitoring lineage patterns during tracheal repair. It was anticipated that, if the bK5-

expressing cells were progenitors capable of repopulating the damaged airway, more βgal-positive cells would be present in a repaired, than in a steady-state, epithelium.

Results

3.2 Evaluation of the *bK5Cre* mouse for monitoring cell lineage in regenerating tracheal epithelium

3.2.1 βgal expression in bK5Cre^{tg} R26R^{tg} *compound transgenic mice*

Since it was first shown to drive appropriate cell type-specific expression *in vitro* (Casatorres *et al.*, 1994), and in a pattern broadly consistent with endogenous K5 *in vivo* (using *lacZ*, (Ramirez *et al.*, 1994)), the 5.3kb bovine K5 promoter has been used to drive basal-specific epithelial expression of a range of transgenes, including *Myc* (Rounbehler *et al.*, 2001), *E2F4* (Wang *et al.*, 2000), *IGF-1* (DiGiovanni *et al.*, 2000), and *Cre* (Brakebusch *et al.*, 2000; Berton *et al.*, 2003).

To verify that the *bK5Cre* transgene would effect appropriate basal epithelial activation of the *R26R* allele, *bK5Cre* mice were crossed with *R26R* mice. Compound transgenic progeny were analysed, and *bK5Cre^{tg} R26R^{tg}* males were crossed with non-transgenic CD1 females to obtain compound transgenic embryos. Such embryos showed widespread activation of βgal expression, which increased markedly from E14.5 to E15.5 (Figure 3.2a,b). A lag in the extent of βgal-activity compared to equivalent stage *K5A1* embryos (chapter 2, section 2.2.4), in which βgal was expressed directly from the human K5 promoter, was likely due to the time taken for Cre protein to effect recombination at *R26R* alleles, and for βgal to subsequently accumulate. Sectioned xgal stained E15.5 skin revealed intense staining in supra-basal layers, while staining in the basal layer, the site of Cre production (shown previously, Figure 2.14), was variegated (Figure 3.2e). Although the bK5 promoter has been reported to drive mosaic expression of a *lacZ* transgene (Ramirez *et al.*, 2001), expression of Cre has not been reported to be mosaic ((Brakebusch *et al.*, 2000; Berton *et al.*, 2003) and A. Ramirez, unpublished observations), and immunodetected Cre protein in E15.5 epidermis was not noticeably variegated

(chapter 2, Figure 2.14). Interestingly, a similar heterogeneity in basal epidermal staining was noted in adult *bK5Cre*^{tg} *R26R*^{tg} skin (A. Ramirez, unpublished results), and was interpreted as indicating that, while Cre production is specific to basal cells, these cells have differentiated into supra-basal derivatives by the time Cre has caused activation of βgal expression from *R26R*. Despite these observations, xgal stained adult pinna in this study demonstrated intense and continuous βgal activity throughout all epidermal layers, including the outermost stratum corneum, and in hair follicles and associated sebaceous glands (Figure 3.2f).

Unexpectedly, a proportion of *bK5Cre*^{tg} *R26R*^{tg} compound transgenic embryos and adults were found to exhibit a strikingly different βgal expression profile, which did not reflect K5 production. Further investigation, described in chapter 4, revealed the cause of this phenomenon, and such animals were excluded from the studies discussed in this chapter. However, the necessary but unanticipated exclusion of these animals meant that a limited number of suitable *bK5Cre*^{tg} *R26R*^{tg} mice were available for the experiment described in section 3.2.3.

In reports of mice expressing βgal (Ramirez *et al.*, 1994), cyclin D1 (Robles *et al.*, 1996) and Cre (Brakebusch *et al.*, 2000) from the bK5 promoter, expression of the transgene (in the case of *Cre*, expression of a Cre-dependent reporter) in the trachea was mentioned but not described. That *Cre* transcript was detected in *bK5Cre* adult tracheal RNA at 30 cycles of RT-PCR (data not shown), compared to 45 cycles required for its detection in three *K5CE* lines (chapter 2, Figure 2.3.7), supported these observations of bK5 promoter activity within the trachea. Regardless of whether the transgene was expressed in tracheal precursors during development, it was thought that if the detected expression of constitutive Cre in adult, steady-state trachea was occurring in stem cells, proposed to express the human K5 promoter, then the degree of pre-existing reporter activation would render the system useless for following specific lineages. To demonstrate this, and with the intention of validating the choice of a temporally regulated system, adult *bK5Cre*^{tg} *R26R*^{tg} tracheas (n = 3) were xgal stained. Significant βgal activity was observed throughout the luminal surface of all three tracheas, though minor variation was apparent in the

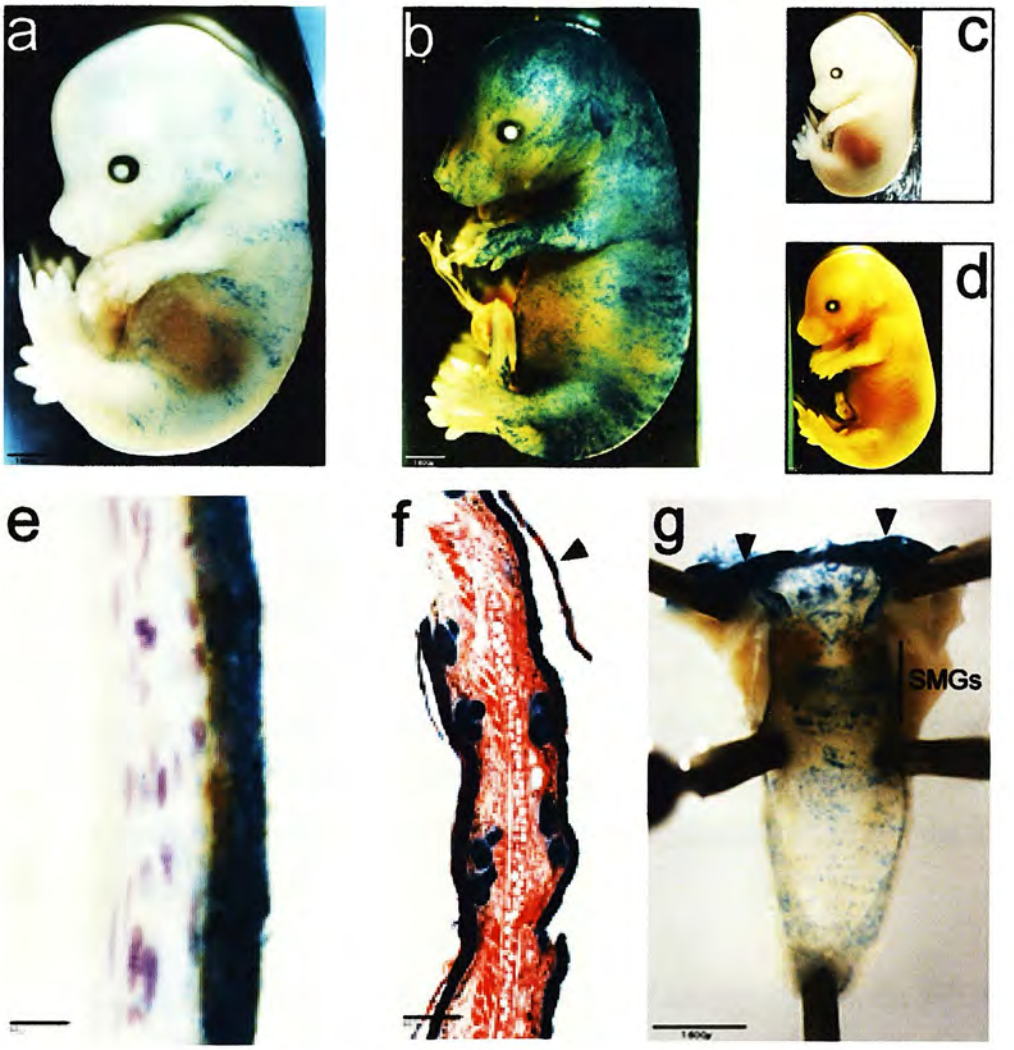


Figure 3.2 βgal expression in *bK5Cre^{tg} R26R^{tg}* transgenic mice.

(a,b) βgal activity in *bK5Cre^{tg} R26R^{tg}* E14.5 and E15.5 compound transgenic embryos, respectively. (c,d) No staining was observed in single transgene-positive *bK5Cre^{wt} R26R^{tg}* embryos. (e) βgal activity in E15.5 epidermis. Staining was uniform throughout the supra-basal layers, though was patchier in the basal layer, stained brown with anti-K14 antibody. (f) Xgal stained adult pinna showing intense staining throughout the epidermis, including the dead squames of the stratum corneum (arrowhead). Hair follicles and associated sebaceous glands showed strong βgal expression. (g) Trachea from *bK5Cre^{tg} R26R^{tg}* mouse showing distinctive, strong staining in the larynx at the proximal end of the trachea (arrowheads), with stained cells distributed over the luminal surface for the length of the trachea. Excluding the larynx, staining was typically more intense in the region corresponding to the submucosal glands (SMGs). No significant endogenous, background glandular staining was apparent in the three tracheas examined. Scale bar = 1600μm (a-d,g); 100μm (f); 10μm (e).

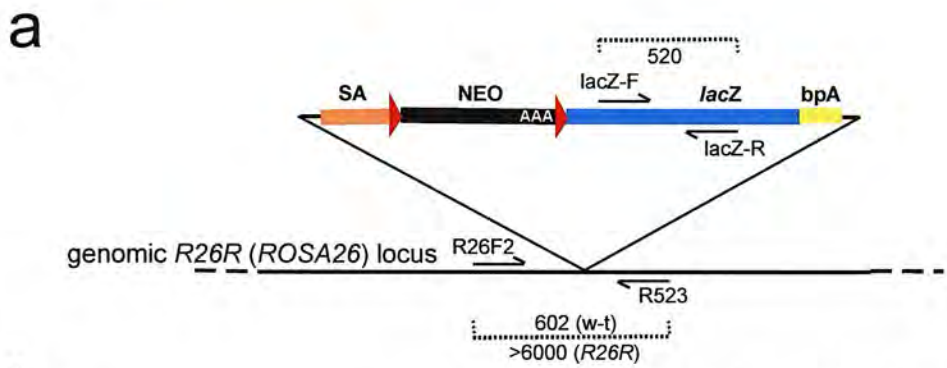
overall intensity of staining, representing differences in the frequency of βgal-positive cells. The trachea exhibiting staining intermediate between the three examined is shown in Figure 3.2g. The larynx was consistently observed to stain heavily. On the luminal surface, the most intense staining corresponded to the region of the tracheal glands.

Although *bK5Cre*^{tg} *R26R*^{tg} tracheas demonstrated significant xgal staining which was clearly not endogenous background activity (described previously, chapter 2, section 2.4.1.3), it was evidently not the result of activation of *R26R* activity in tracheal precursors during development, which would have been expected to result in extensive or ubiquitous βgal expression. Indeed, the moderate level of staining did not obviously exclude the *bK5Cre*^{tg} *R26R*^{tg} mouse from being informative with regard to tracheal lineage patterns. Particularly, the fact that the glandular proximal trachea showed the highest levels of xgal staining within the luminal surface, suggested that the bK5 promoter may have been most active in the SMGs, as is the case for its human counterpart (Borthwick *et al.*, 2001). For these reasons, the usefulness of *bK5Cre*^{tg} *R26R*^{tg} mice for tracing lineage in regenerating tracheal epithelium was evaluated in parallel with analogous experiments involving the *K5CE* mouse (section 3.3).

The *R26R* status (hemizygous or homozygous) of the animals used in the following studies was determined using a PCR specific to the wild-type *R26R* (*ROSA26*) locus in conjunction with the *lacZ*-specific PCR described in chapter 2, Figure 2.3 (Figure 3.3).

3.2.2 Intratracheal instillation of polidocanol to damage tracheal epithelium

Prior to undertaking the tracheal injury protocol on the limited number of suitable *bK5Cre*^{tg} *R26R*^{tg} animals available for the lineage study, a 24 hour POL damage (and PBS control) timepoint was performed on wild-type animals ($n = 3$ per treatment) to check that results consistent with those described by Borthwick (Borthwick, 1999)



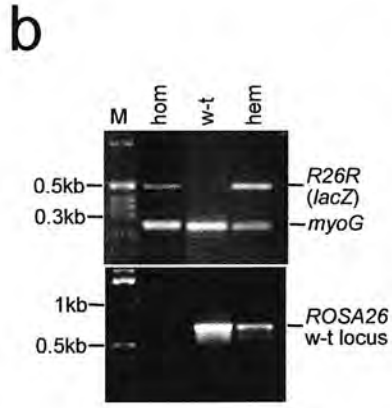


Figure 3.3 PCR assay for the discrimination of mice hemizygous and homozygous for *R26R*.

(a) Schematic of the *R26R* (*ROSA26*) locus showing the position of PCR primers (R26F2 – R523, Soriano, 1999) within the genomic sequence flanking the *R26R* construct. The primers are only able to amplify a product (602bp) from the wild-type (w-t) locus. *R26R* (*lacZ*)-specific primers (lacZF – lacZR) are also shown. SA (splice acceptor), AAA (triple polyadenylation sequence), bpA (polyadenylation sequence).M = 1kb ladder. Map not to scale, fragment sizes in bps. (b) The *R26R* hemizygote sample (hem) is positive in both the *lacZ*-specific and *R26R* locus-specific PCRs. Amplification only in the *lacZ*-specific PCR indicates a homozygote (hom, no wild-type *R26R* loci). Wild-type samples typically produced a stronger signal in the *R26R* locus-specific PCR. *myoG*-specific primers in the *lacZ*-specific PCR served as a positive control for DNA quality.

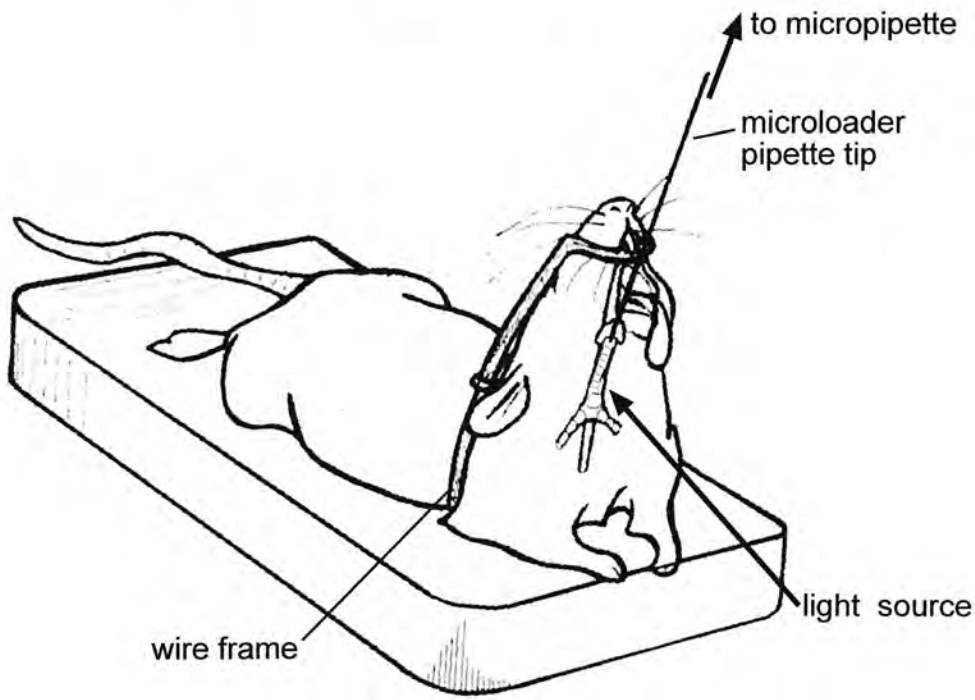


Figure 3.4 Apparatus for tracheal instillation of PBS or polidocanol into anaesthetised mice (diagram modified from Davidson, 2000).

See chapter 7 (section 7.10) for details of method.

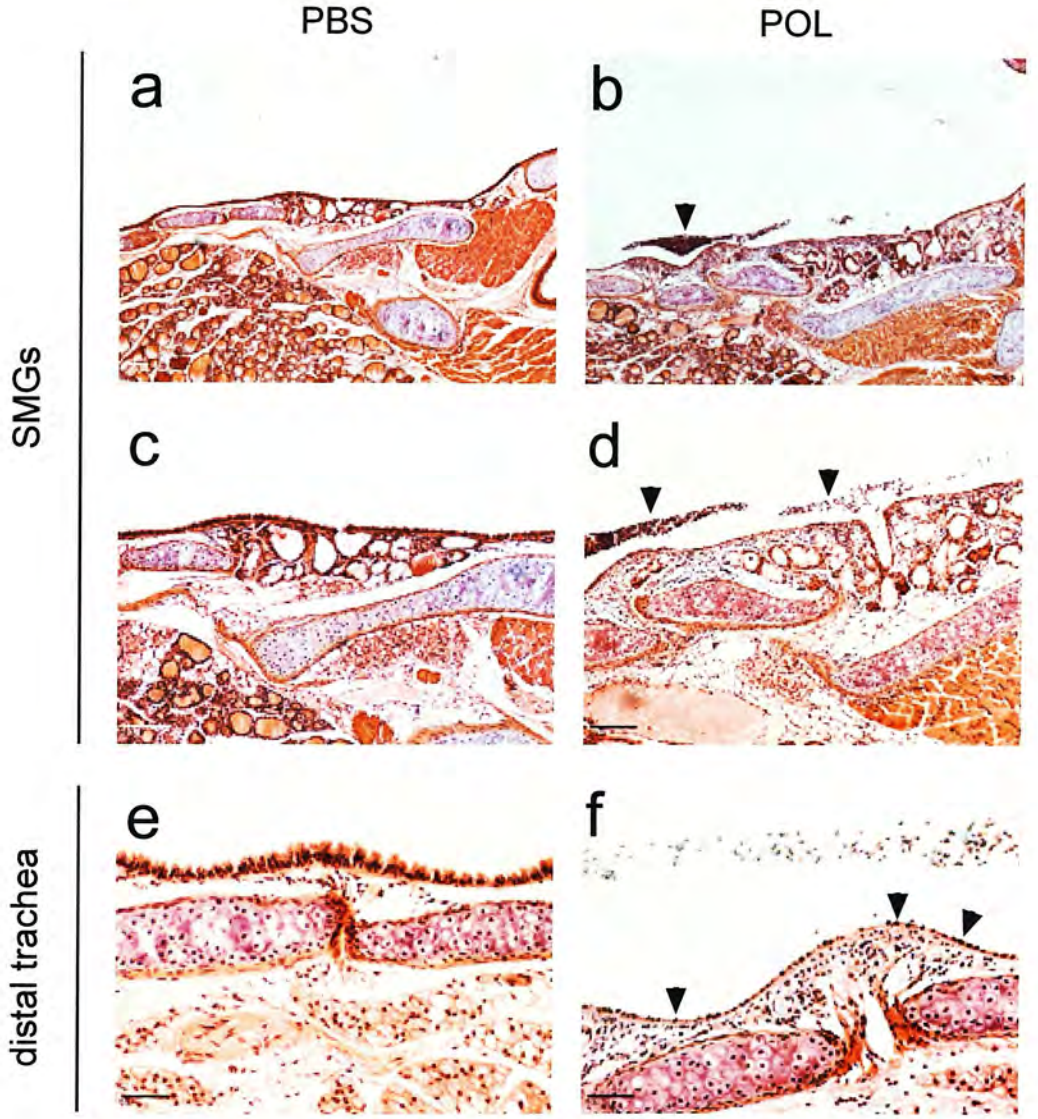


Figure 3.5 Polidocanol-induced damage of tracheal epithelium after 24 hours.

10μl 2% polidocanol (POL) in PBS, or PBS alone, was administered to wild-type mice by direct tracheal instillation. Tracheas were removed after 24 hours and sectioned for analysis of damage. (a,c) Submucosal gland (SMG) regions shown at two different magnifications after treatment with PBS alone. The epithelial layer was intact. (b,d) Equivalent images to (a,b) for trachea receiving POL. The epithelium had been removed, and detached material could be seen in the tracheal lumen (arrowheads). (e,f) Distal tracheal luminal surface following PBS or POL treatment, respectively. Arrowheads in (f) indicate remaining basal epithelial cells. Scale bar = 100μm (a-d); 50μm (e,f).

were obtained. Intratracheal instillation was achieved by placing an anaesthetised mouse in the apparatus depicted in Figure 3.4. A narrow-gauge, flexible micropipette tip was used to dispense 10µl of 2% POL in PBS (or PBS alone) to the top of the trachea. S. Webb provided technical assistance with establishing the reproducibility of this technique.

After approximately 24 hours the mice were culled and the tracheas sectioned for analysis of epithelial injury (Figure 3.5). As expected, while PBS treated tracheas displayed an intact epithelial layer, POL had stripped superficial columnar cells from the full length of the tracheal surface, though some basal cells closely adherent to the underlying lamina were seen to persist. Material presumably consisting of detached epithelial cells was occasionally observed in the tracheal lumen. This study confirmed that the intended epithelial injury could be reproducibly obtained using the apparatus and technique described.

3.2.3 Analysis of βgal expression in repairing bK5Cre^{tg} R26R^{tg} *tracheal epithelium*

To investigate whether the *bK5Cre* line could be used to monitor lineage in regenerating tracheal epithelium, 13 *bK5Cre^{tg} R26R^{tg}* adult compound transgenic mice, and four adult *bK5Cre^{wt} R26R^{tg}* mice, were split into four groups to be culled at different timepoints after intratracheal instillation of POL or PBS. Table 3.1 shows the distribution of genotypes and treatments between the groups. The limited availability of useable *bK5Cre^{tg} R26R^{tg}* mice (see section 3.2.1) prevented the use of equal numbers between the groups. The timepoints used were (after treatment), +1 day (d), +3d, +7d and +20d. Few stained cells were expected to be seen in animals receiving POL at +1d due to removal of the epithelial cells. The epithelial layer should be largely restored by +3d, and completely by +7d (Borthwick *et al.*, 2001), thus these timepoints were expected to be informative if the *bK5Cre* transgene was active in epithelial progenitors. At +20d, the epithelium should have returned to the steady-state.

Timepoint	Mouse number	Genotype		Treatment (POL/PBS)	Ave. stained cells		Ave. cells per block	
		bK5Cre	*R26R*		SMGs	distal	SMGs	distal
+1d	100	tg/wt	tg/wt	PBS	36.0	7.0	2.6	2.3
	130			POL	0	0	-	-
	616	wt/wt	tg/tg		0	0	-	-
+3d	102	tg/wt	tg/wt	PBS	13.5	2.0	2.5	2.0
	127				63.5	7.0	5.1	2.3
	103			POL	36.0	6.5	3.0	1.6
	128				71.0	21.0	4.4	3.5
	613	wt/wt	tg/tg		0	0	-	-
+7d	104	tg/wt	tg/wt	PBS	47.0	8.0	4.5	2.0
	129				83.5	36.5	5.1	3.8
	105			POL	81.5	65.0	5.6	5.7
	131				75.0	75.0	10.0	8.8
	794	wt/wt	tg/tg		0	0	-	-
+20d	134	tg/wt	tg/wt	PBS	24.5	6.5	3.5	2.5
	118			POL	87.0	4.5	7.9	2.3
	135				8.5	6.5	2.4	2.6
	615	wt/wt	tg/tg		0	0	-	-

Table 3.1 Summary of experiment to evaluate the *bK5Cre* mouse for tracing lineage in the regenerating tracheal epithelium.

For each trachea, stained cells were counted in submucosal gland (SMG) and distal tracheal regions on a total of four microscopic fields per region, on two different sections. Averaged values are shown for the number of stained cells and the number of cells per continuous stained block of cells, and reveal significant variation between tracheas, including those of equivalent genotype, receiving the same treatment, and taken at the same timepoint. tg/wt = hemizygous, tg/tg = homozygous, wt/wt = transgene-negative. POL = polidocanol.

3.2.3.1 Unanticipated variation in tracheal βgal expression between individual bK5Cre^{tg} R26R^{tg} *mice*

Examples of wholemount and sectioned xgal stained tracheas are presented in Figure 3.6. The range of variation in tracheal expression of βgal between individual animals, irrespective of treatment, was considerably greater than suggested by the initial analysis of three untreated *bK5Cre^{tg} R26R^{tg}* tracheas (section 3.2.1). The extent of variation, represented by the number of xgal stained cells counted, is apparent in Table 3.1. As a result of this variation, any increase in the number of βgal-positive cells during the period of epithelial regeneration following POL treatment was masked by pre-existing differences in βgal expression inherent to the individual animals. The most informative POL and PBS treated tracheas from each timepoint, usually those exhibiting the most staining, are shown in Figure 3.6. Many of the tracheas that are not shown displayed low levels of βgal activity, with xgal stain appearing faint on sections, though they were consistently distinguishable from *bK5Cre^{wt} R26R^{tg}* tracheas. This heterogeneity in βgal expression was intriguing, as it did not correlate with the number of transgene copies (all compound transgenic animals were hemizygous for both transgenes), or the age or sex of the animals. Fixing and staining was performed identically, using the same xgal preparation, making it unlikely that the differences were due to technical factors. Also, variation in staining distribution and intensity was not an artefact of endogenous galactosidase activity, as determined by its appearance, and by comparison with control (*bK5Cre^{wt} R26R^{tg}*) tracheas stained concurrently (for example, Figure 3.6i).

In the absence of a visually discernible correlation between the extent of xgal staining observed in tracheas, and the treatments and timepoints pertaining to each trachea, the numbers of stained cells and blocks of continuous stained cells were counted on two SMG sections and two distal tracheal sections, for each trachea. Averaged counts, represented as the total number of stained cells and the number of cells per block, are recorded in Table 3.1 and indicate the wide variation in βgal expression apparent even between transgenically-identical mice which received the same treatment and were killed at the same timepoint. Standard analysis of variance

(ANOVA) confirmed that variation in the numbers of stained cells and blocks observed between individual animals was significantly greater than the residual variation between replicate observations from each trachea, allowing averaged values to be used in the subsequent analysis. This indicated an overall marginally significant ($p<0.05$) increase in the total number of stained cells and stained cells per block after POL treatment compared to PBS treatment. However, the importance of this result is unclear given the absence of statistically significant differences in the numbers of stained cells observed in POL treated tracheas between the various timepoints at which tracheas were examined (n = two per +3d, +7d, and +20d timepoint). Statistical analysis was performed by A. Carothers.

3.2.3.2 Clonal expansion from bK5Cre-expressing progenitors?

Despite the lack of clear evidence, in terms of total numbers of marked cells, for clonal expansion originating in *bK5Cre*-expressing cells during re-epithelialisation, examination of sectioned tracheas revealed intriguing staining patterns (Figure 3.6). The pattern of xgal staining within SMG ducts, indicative of bK5 promoter expressing cells and their descendents, suggested the bK5 promoter was active in the cells previously shown to express the human K5 promoter (Borthwick *et al.*, 2001). Particularly, the SMG duct shown in Figure 3.6a(iii) is reminiscent of the SMG duct containing cells expressing βgal from the human K5 promoter (chapter 2, Figure 2.12(i)h). This is supported by the fact that staining within the tracheal lumen was consistently significantly more widespread and intense in the region corresponding to the SMGs, than more distally. The majority of stained cells were found in continuous stained blocks of cells interspersed with non-stained blocks (for example, Figure 3.6c(ii)), consistent with the notion of clonal expansion of marked cells. The most encouraging observation with regard to the hypothesis that K5-expressing epithelial progenitors are present within SMG ducts, was numerous examples of apparent continuous marked lineage extending from gland duct cells to the overlying surface epithelium (Figure 3.6d(iii),f(iii,vi,viii),h(v)). Interestingly, two of the best examples of this were located in discrete glands, one or two intercartilaginous spaces distal to the main block of SMGs (Figure 3.6f(v and vii, respectively)). Although direct

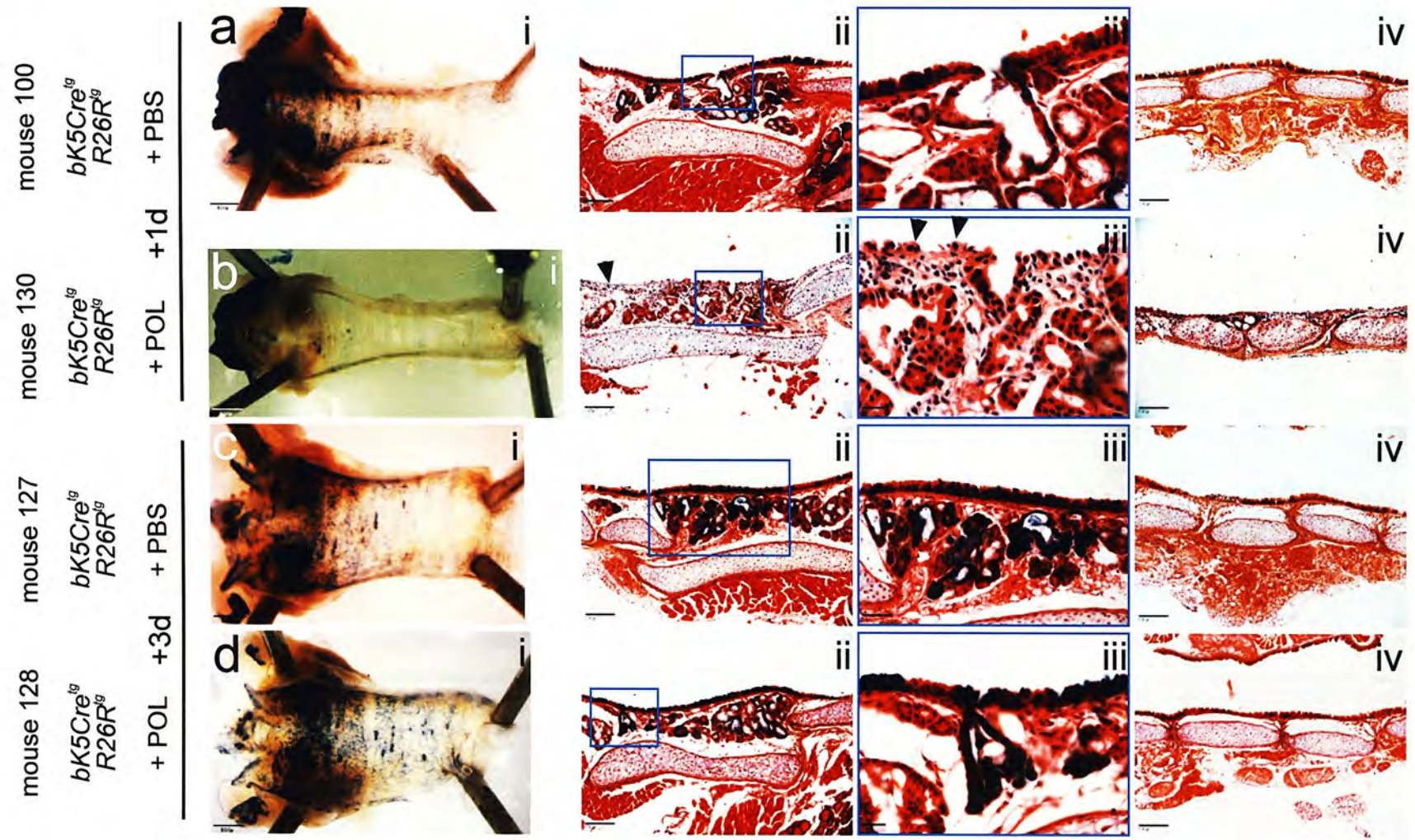


Figure 3.6 Analysis of βgal expression in repairing *bK5Cre*^{tg} *R26R*^{tg} tracheal epithelium (continued over page).

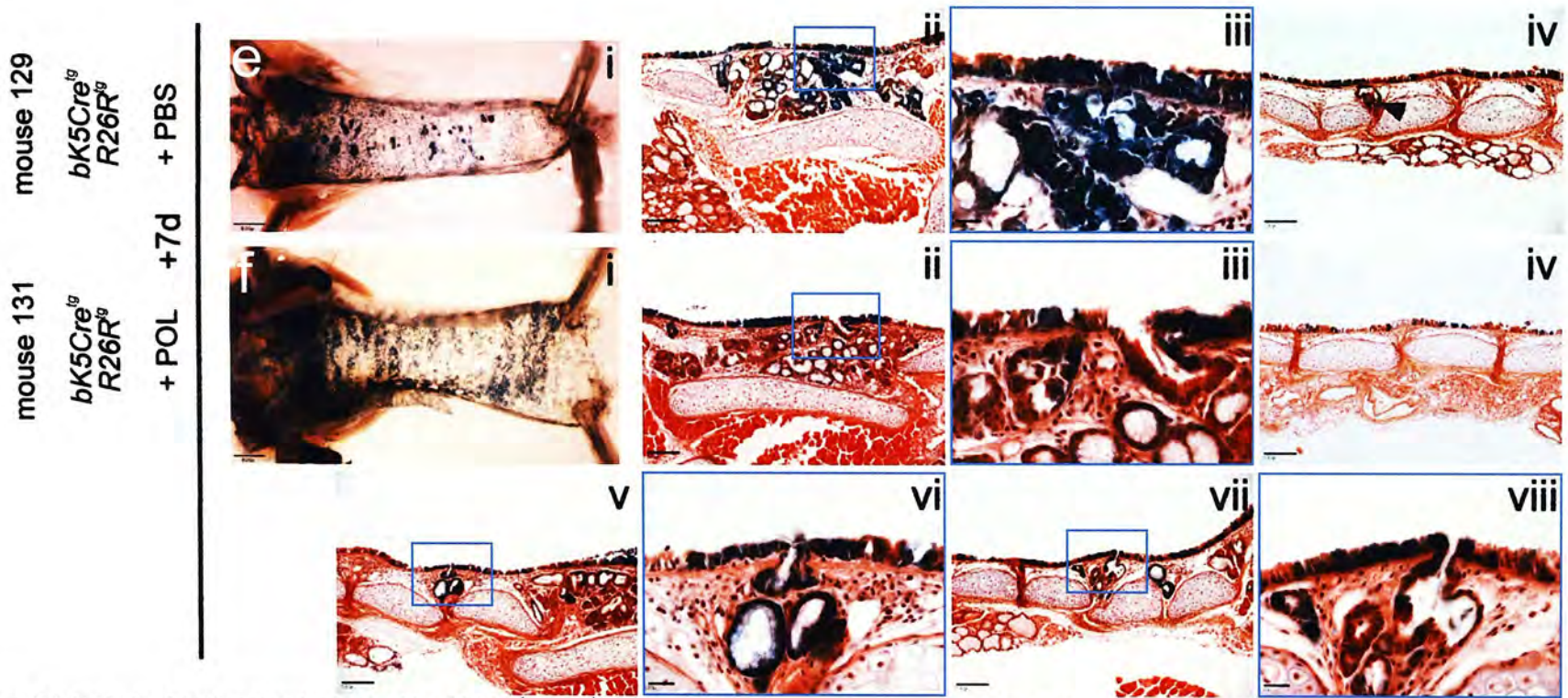


Figure 3.6 Analysis of βgal expression in repairing *bK5Cre*^{tg} *R26R*^{tg} tracheal epithelium (continued from previous page, continued over page).

The most informative, as determined by the extent of xgal staining, polidocanol (+POL) and PBS (+PBS) treated tracheas from each time point are shown. (a-i(i)) Wholemounts cut to reveal the ventral luminal surface. Two different magnifications of the sectioned submucosal gland (SMG) region are shown in (a-i(ii,iii)), and additionally in (f(v-viii)). (a-i(iv)) shows a section of distal, non-glandular trachea. (a,b) +1d. (a(i)) The PBS treated trachea showed the typical staining pattern. Blocks of stained cells were evident in the epithelium overlying the SMGs (a(ii)), and stained cells were present within the SMG duct (a(iii)). Discrete staining was observed in the distal trachea (a(iv)). After POL treatment, no staining was apparent in the tracheal surface (b(i)), and this corresponded with removal of columnar surface cells both proximally and distally (b(ii,iv)). Although some surface regions appeared cleanly denuded (arrowhead, b(ii)), persisting basal cells could be seen in other areas (arrowheads, b(iii)). (c,d) +3d. (c(i)) The PBS treated trachea showing expected staining pattern. Sectioned SMGs (c(ii,iii)) revealed heavy glandular staining and continuous staining in overlying epithelium. Low-level staining was visible in the distal surface. (c(iv)) Widespread staining was apparent in the POL treated trachea (d(i)), which coincided with a re-established epithelial layer showing large blocks of dense staining (d(ii)). A possible lineage relationship between a gland duct and neighbouring epithelium is shown (d(iii)). Distal epithelia had also been repaired and showed low levels of staining (d(iv)). (e,f) +7d. (e(i)) The PBS treated trachea showed numerous, distinctive foci of intense staining. Sectioned SMGs showed dense staining in a subset of glands, and a possible spatial relationship between these glands and heavily stained overlying epithelium (e(ii,iii)). (e(iv)) Blocks of staining were evident in the distal trachea, where gland structures were also noted (arrowhead), and likely corresponded to the distal staining foci evident in (e(i)). (f(i)) The POL treated trachea showed extensive staining throughout the tracheal surface. Sectioning of the SMG region revealed several examples of possible lineage relationships between gland duct openings and surrounding luminal epithelial cells (f(ii,iii,v-viii). Stained cells were apparent in the distal trachea (f(iv)).

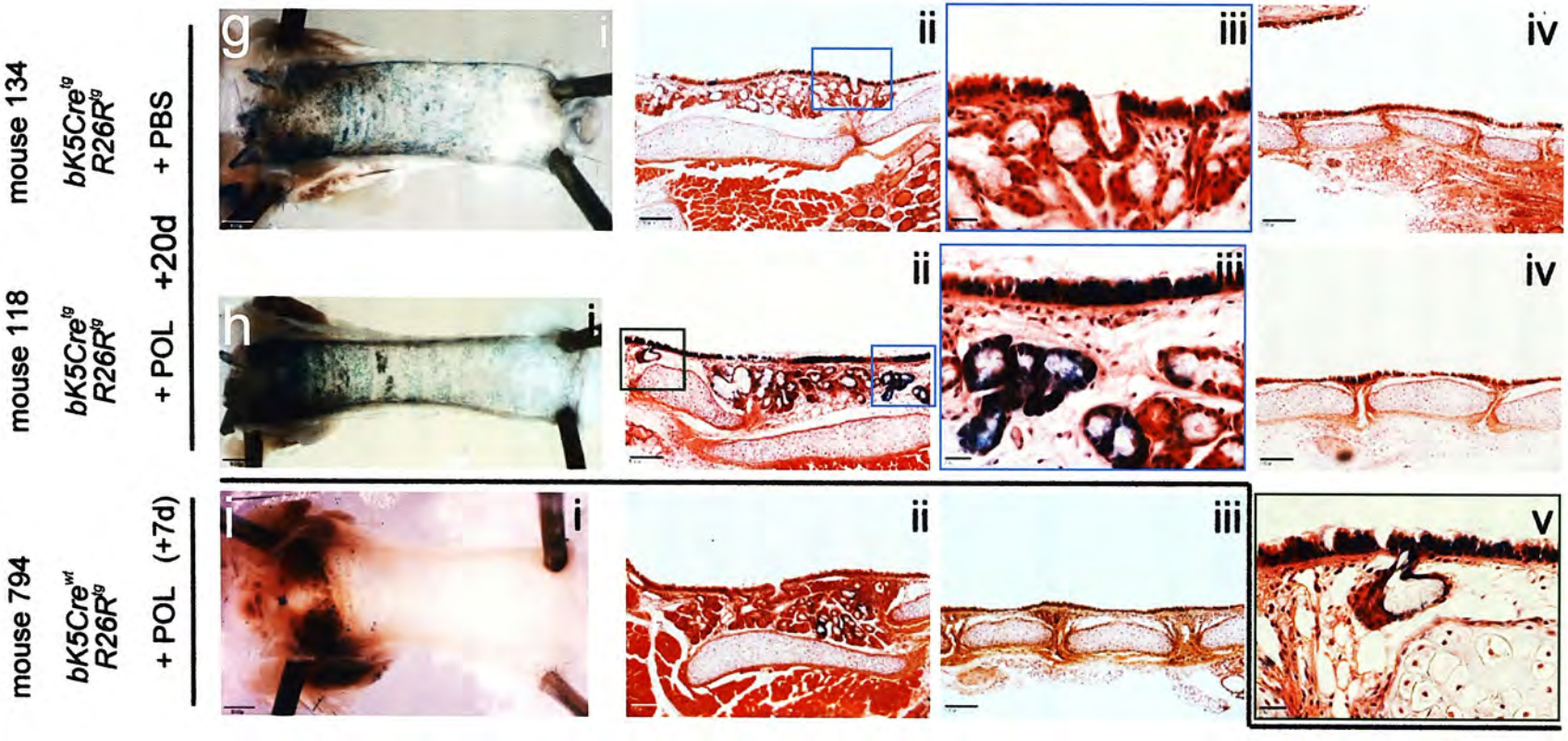


Figure 3.6 Analysis of βgal expression in repairing *bK5Cre^{tg} R26R^{tg}* tracheal epithelium (continued from previous page).

(g,h) +20d. (g(i)) Staining appeared moderate in the PBS treated trachea, though sectioning of the SMG region showed low levels of staining clustered around a gland duct opening (g(ii,iii)). Infrequent staining was detected in distal epithelium (g(iv)). The POL treated trachea showed staining throughout the tracheal surface, which was particularly intense proximally (h(i)). This corresponded with large blocks of stained epithelium overlying the SMGs (h(ii)), and examples of possible spatial (h(iii)) and lineage (h(v)) relationships between stained glands and surface cells are shown. Weak staining was apparent in the distal trachea (h(iv)). A representative single transgene-positive *bK5Cre^{wt} R26R^{tg}* trachea is shown (+7d, i(i)). The tracheal surface was devoid of staining, with apparent proximal staining due to visualisation of endogenous thyroid staining from "behind" the trachea. (i(ii)) A sectioned SMG revealed a single stained gland, likely a result of endogenous galactosidase activity. No stained epithelial cells were evident proximally or distally (i(ii,iii)). Scale bar = 800μm (a-i(i)); 100μm (a-h(ii,iv), f(v,vii), i(ii,iii)); 50μm (c(iii)); 20μm (a,b,d-h(iii), f(vi, viii)), h(v)).

continuity of stained cells connecting the gland duct to the tracheal surface was only observed in POL treated tracheas, the significance of this was unclear, and may simply have been the result of a chance increase in the frequency of sections passing directly through gland ducts in POL compared to PBS treated tracheas. Close spatial correlations between staining in glands and in a discrete portion of directly overlying surface epithelium were frequently observed in both POL and PBS treated tracheas (Figure 3.6h(iii) and c,e(ii), respectively), and were interpreted as likely representing direct gland-surface lineage relationships existing on another plane of sectioning. Serial sectioning would be required to determine whether this was the case.

3.2.3.3 Variation in the staining and distribution of submucosal glands between individual bK5Cre^{tg} R26R^{tg} *mice*

The distribution of βgal activity within the glandular region of *bK5Cre^{tg} R26R^{tg}* tracheas was intriguing. While in some tracheas, staining was apparent throughout many of the glands within the glandular region (for example, Figure 3.6c(ii)), staining appeared to be confined to few glands in the glandular region of other tracheas (for example, Figure 3.6h(ii)). While this variation in the extent of staining within the glandular regions of different tracheas may seem to raise the issue of endogenous galactosidase activity, it must be stressed that such non-specific staining as described in chapter 2 section 2.4.1.3 did not appear to occur at appreciable levels within this experiment. Furthermore, staining in surface epithelial cells, which as described was often continuous with stained gland cells, was not observed to result from non-specific sources in any of the experiments performed for this thesis. Thus, it may be that K5-expressing cells do not exist within all glands, and that the proportion of glands containing these cells may vary between individual animals. The existence of variation in tracheal characteristics between individual animals used in the study, as previously described in terms of marked differences in staining frequency in genotypically equivalent and identically treated animals, was further supported by the appearance of numerous, intensely staining foci in the wholemount trachea of (PBS treated) mouse 129 (Figure 3.6e(i)). While similar foci were evident in wholemount tracheas from many of the mice, and corresponded to SMGs, the

mouse 129 trachea was distinctive due the high number and widespread distribution of the foci, which extended into the distal trachea. Figure 3.7a confirms the presence of distal glands in the trachea of mouse 129, with a gland positioned in the seventh intercartilage space distal to the main gland block. However, no obvious relationship between the stained distal gland and the local epithelial layer was apparent.

3.2.3.4 Absence of lineage patterns in the distal trachea

The patterns of staining in the generally glandless distal trachea were less informative than in proximal regions. Staining was consistently less intense and appeared more scattered, with fewer and smaller blocks of continuous staining. If the bK5 promoter recapitulates the expression from the human K5 promoter, which was shown to express in a pattern consistent with LRCs, found to be localised to cartilage-intercartilage junctions in the distal trachea, then it might be expected that staining would be concentrated in these regions. While small blocks of staining were occasionally detected at these sites (Figure 3.6h,e(iv)), they were not the exclusive sites of distal βgal activity (Figure 3.6e,f(iv)). It may be that the proposed progenitorial role of high K5-expressing basal cells is diminished in the distal trachea, as discussed in section 3.4.1.

3.2.3.5 βgal expression in both columnar and basal epithelial cells

Stained epithelial blocks in the proximal glandular region were examined at high power magnification to determine whether they contained multiple cell types, based on morphology, which would indicate multipotency of the progenitors. Due to the relative thickness of the sections (10μm), the presence of several layers of cells reduced the clarity with which individual cells could be visualised. Staining corresponding to basal cells could be detected at the edges of, and separate from, blocks of stained columnar cells (Figure 3.7b). However, the POL treatment is known to leave some basal cells attached to the luminal surface, and it is known that a subset of surface basal cells, particularly near gland duct openings, show high levels of (human) K5 promoter activity (Borthwick *et al.*, 2001; Schoch *et al.*, 2003).

For these reasons, it could not be said with certainty whether the basal cells observed were part of, and thus potentially clonally related to, an adjacent stained block also containing columnar cells, or whether they were bK5 promoter-expressing cells which may themselves have given rise to neighbouring blocks.

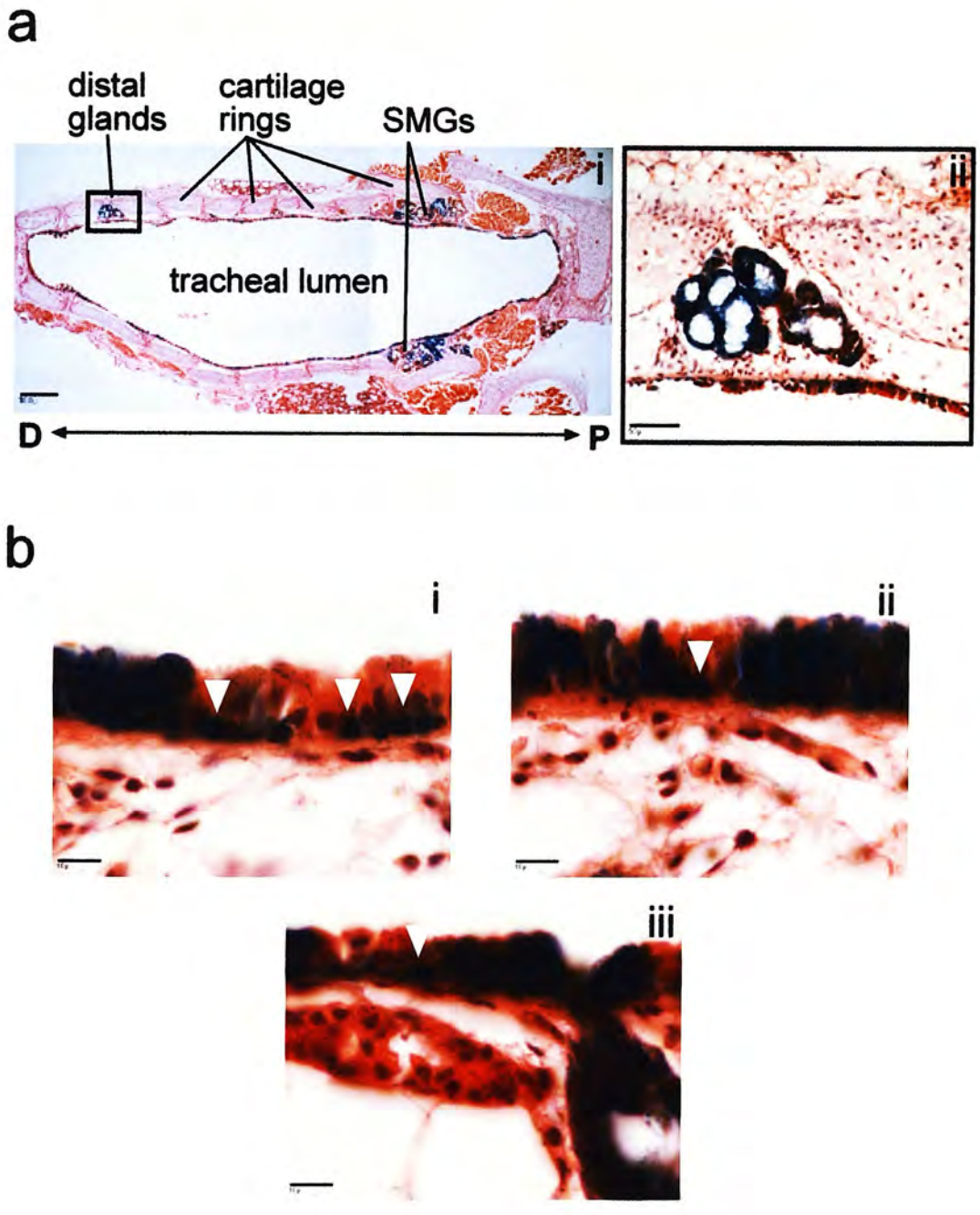


Figure 3.7 Further observations from tracheal epithelial lineage marking using *bK5Cre*^{tg} *R26R*^{tg} mice.

(a(i)) Longitudinal section through mouse 129 trachea showing xgal stained proximal submucosal glands (SMGs), stained epithelial cells, and stained distal glands. The proximal (P) - distal (D) axis is shown. (ii) Higher magnification of boxed region in (a(i)) showing stained distal gland overlying cartilage-intercartilage junction. Apparent epithelial damage over the gland, possibly resulting from incomplete repair of distal epithelium at +7d timepoint, or mechanical damaged caused when the trachea was opened for wholemount inspection, prevented the detection of a potential correlation between staining in the gland and overlying surface epithelium. (b) High magnification of proximal tracheal epithelium from mouse 105 (i,ii) and 128 (iii), showing staining corresponding to the expected position of basal epithelial cells (arrowheads). Scale bar = 100μm (a(i)); 50μm (a(ii)); 10μm (b(i-iii)).

3.3 Evaluation of the *K5CE* mouse for monitoring cell lineage in regenerating tracheal epithelium

3.3.1 Analysis of βgal expression in repairing K5CE^{tg} R26R^{tg} *tracheal epithelium*

Although three potentially useful *K5CE* lines had been identified in the experiments described in chapter 2 section 2.3, time constraints and mouse availability dictated that only line *K5CE1100* would be assessed for its ability to facilitate inducible marking of the putative tracheal stem cell population. Nine *K5CE^{tg} R26R^{tg}* adult compound transgenic mice, and three each of adult mice carrying either the *K5CE* or *R26R* transgenes, were split into three timepoint groups. All mice carrying the *R26R* allele were confirmed to be homozygous by PCR (data not shown). Although a similar assay was not available for the *K5CE* transgene, breeding records were used (as described previously, chapter 2, Table 2.2) to calculate that all *K5CE^{tg}* animals had a 0.61:0.39 probability of being hemizygous:homozygous. The timepoints (+7d, +20d, +30d) differed from those used in the previous experiment, since it was known that induction of lineage marking by repetitive TAM administration would take several days, and that damage had to be effected prior to reporter activation (B. Stripp, personal communication). A timeline for the experiment is shown in Figure 3.8. All animals received intratracheal instillation of POL followed immediately with an intraperitoneal injection of 4mg TAM in corn oil, or oil alone (+0d). The mice received further injections of TAM or oil on +1d and +2d, and were killed on +7d, +20d, and +30d.

This protocol unexpectedly resulted in significant mortality, with three animals failing to emerge from anaesthesia. Two more mice died subsequently, on +3d and +5d. Several of the remaining mice appeared "hunched", and all animals were monitored closely for any further deterioration in their condition for the duration of the study. The unexpected loss of one third of the experimental animals prior to the first timepoint prompted a reallocation of animals between timepoint groups in order to minimise the effect on the results. The final timepoint groups are shown in Table

3.2, including the treatment administered to each mouse, and the prematurely deceased mice. As a result of the reallocation, no control animals (genetic or treatment) were present in the +30d group, the *K5CE^{tg} R26R^{wt}* control was absent from the +20d group, and only one TAM treated *K5CE^{tg} R26R^{tg}* mouse was available in the +7d group.

At each timepoint, pinna and tail skin, in addition to the trachea, were taken from mice as positive controls for systemic action of administered TAM. Representative stained tissues are shown in Figure 3.9, and the results are summarised in Table 3.2. No staining above endogenous levels (which were minimal in this experiment) was observed in the tracheas of any of the animals, as shown in TAM treated *K5CE^{tg} R26R^{tg}* tracheas from each timepoint (Figure 3.9a-c(i)), and genetic and treatment control tracheas (+20d, Figure 3.9d,e(i)). Sectioning of the tracheas revealed that, while some degree of epithelial repair had occurred, it was variable between animals and appeared impaired relative to that observed in the previously described experiment involving the *bK5Cre* mice (compare Figure 3.9a-e(ii) with Figure 3.6d,f,h,i(ii)).

Examination of the pinna and tail skin indicated that recombination as a result of systemic TAM activity had occurred. At +7d (Figure 3.9a(iii,iv)), low-level staining in pinna was apparent, and had increased markedly by +20d (Figure 3.9b(iii-v)), with staining evident in both basal and supra-basal cells, indicating differentiation of cells in which the recombination event had occurred. Scattered βgal expression was also apparent in tail skin (Figure 3.9b(vi-v)). Large blocks of staining at were apparent in pinna at +30d (Figure 3.9c(iii-v)), and in some cases stained supra-basal cells lay over βgal-negative basal cells, suggesting that some CreER^{T2}-producing cells had begun to differentiate before βgal expression had been activated from *R26R*. An increased frequency of stained cells within tail skin was noted at +30d, with βgal activity evident in all epidermal layers (Figure 3.9c(vi-vii)). Although stained cells were observed in oil treated *K5CE^{tg} R26R^{tg}* pinna, they were extremely infrequent and not visible at the magnification used (Figure 3.9d(iii)). No stained cells were

Timepoint	Mouse number	Genotype		Treatment (TAM/OIL)	Xgal staining		
		K5CE	*R26R*		Trachea	Pinna	Tail skin
+7d	787^{+5d}	tg	tg/tg	TAM	N/A	N/A	N/A
	818				-	+	-
	824			OIL	-	LI	-
	784	tg	wt/wt	TAM	-	-	-
	904	wt/wt	tg/tg		-	-	-
+20d	788	tg	tg/tg		-	+++	++
	814				-	+++	+
	828			OIL	-	LI	-
	783^{+3d}	tg	wt/wt	TAM	N/A	N/A	N/A
	905	wt/wt	tg/tg		-	-	-
+30d	819	tg	tg/tg		-	+++	++
	820				-	+++	+++
	829^{+0d}			OIL	N/A	N/A	N/A
	812^{+0d}	tg	wt/wt	TAM	N/A	N/A	N/A
	898^{+0d}	wt/wt	tg/tg		N/A	N/A	N/A

Table 3.2 Summary of experiment to evaluate the *K5CE* mouse for tracing lineage in the regenerating tracheal epithelium.

All animals received polidocanol on +0d. Superscript after mouse numbers indicates day of premature death. +++,++,+ indicates high, moderate, and low level staining, respectively, compared to equivalent tissue from other animals. LI indicates very low-level staining resulting from ligand-independent activity of $CreER^{T2}$. tg = transgene-positive, tg/tg = homozygous, wt/wt = transgene-negative. TAM = tamoxifen. N/A = not applicable.

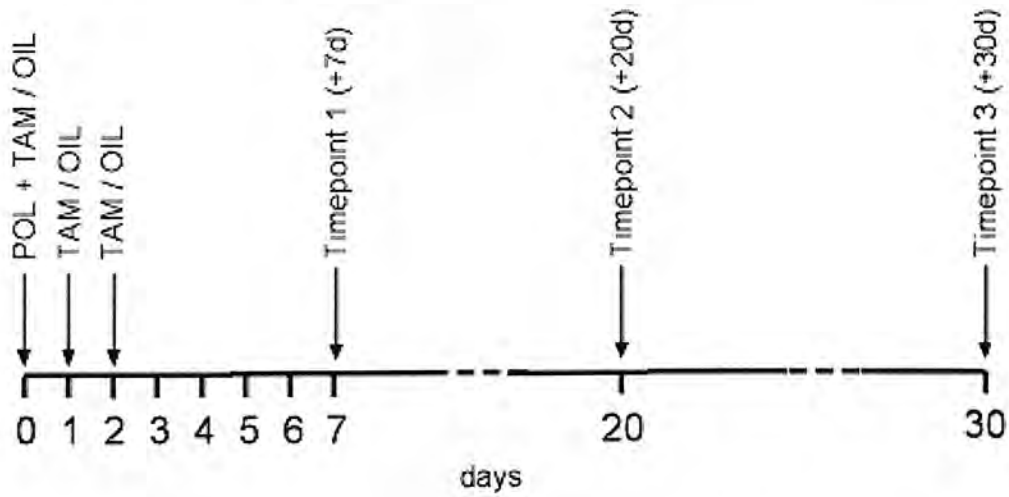


Figure 3.8 Timeline of experiment to evaluate the *K5CE* mouse for tracing lineage in the regenerating tracheal epithelium.

POL = polidocanol, TAM = tamoxifen.

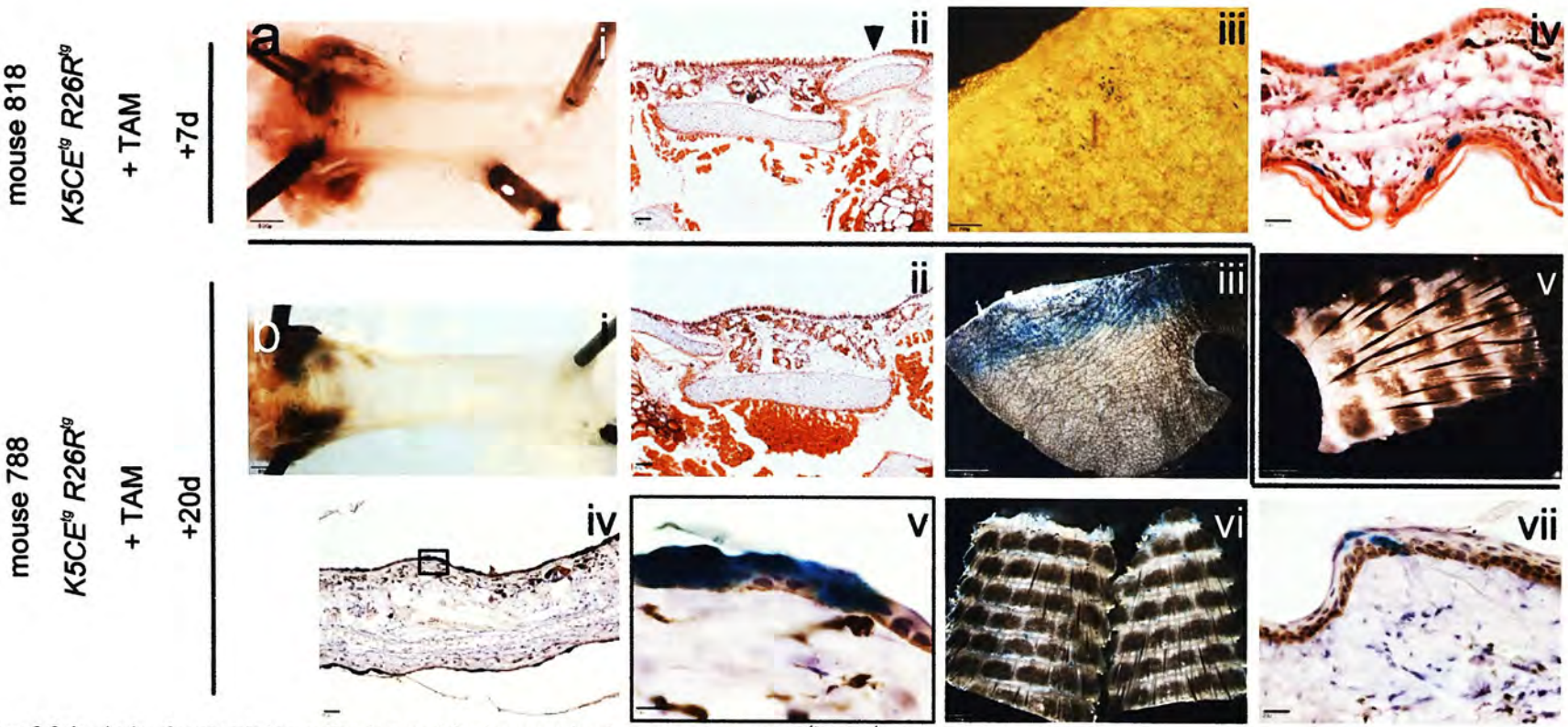


Figure 3.9 Analysis of systemic tamoxifen-induced βgal expression in tissues from *K5CE^{tg}R26R^{tg}* mice (continued over page).

Tissues from one tamoxifen (TAM) treated *K5CE^{tg} R26R^{tg}* mouse per timepoint are shown (a-c), plus one oil treated *K5CE^{tg} R26R^{tg}* mouse (d) and one TAM treated *K5CE^{wt} R26R^{tg}* mouse (e). (a) +7d. (a(i)) An absence of staining in the wholemount trachea was confirmed by sectioning (a(ii)), which revealed a ragged epithelial layer with persisting patches of denuded lamina (arrowhead, a(ii)). A low level of xgal staining, but above that due to ligand-independent Cre activity, was noted on pinna (a(iii)), and was observed to be primarily in the basal epidermis (a(iv)). No staining was observed in tail skin at this timepoint (a(v)). (b) +20d. No staining was observed in wholemount (b(i)) or sectioned (b(ii)) trachea, the latter revealing improved epithelial repair compared to the +7d timepoint. (b(iii)) Strong staining was apparent in a band parallel to the cut edge of the pinna. Anti-K14 stained pinna sections revealed staining to be both basal and supra-basal (b(iv and magnified inset, v)). Stained cells were observed between the scales of tail skin, and were seen to be both basal and supra-basal in anti-K14 stained section (b(vi,vii)). (c) +30d. (c(i)) Wholemount trachea showed no staining. (c(ii)) Epithelia appeared patchy over the submucosal gland region. Pinna showed significant staining along cut edges (c(iii)), with staining apparent throughout epidermal layers (c(iv,v)), including stratum corneum (arrowhead, c(v)). Numerous stained cells were visible in tail skin, in both basal, supra-basal, and cornified layers (c(vi,vii)). (d) (+20d). (d(i,ii)) No staining was present in wholemount or sectioned trachea, the latter displaying a ragged epithelial layer. An extremely low level of ligand-independent recombination was present in pinna but was not detectable at low magnification (d(iii)), and no staining was observed in tail skin (d(iv)). (e) (+20d). No staining was observed in trachea (e(i,ii), pinna (e(iii)) or tail skin (e(iv)). Scale bar = 1600μm (b-e(iii)); 800μm (a-e(i)); 400μm (a(v), b,c(vi), d,e(iv)); 200μm (a(iii)); 50μm (a-e(ii), b,c(iv)), 20μm (a(iv), b,c(vii)); 10μm (b,c(v)).

Figure 3.9 Analysis of systemic tamoxifen-induced βgal expression in tissues from *K5CE^{tg}R26R^{tg}* mice (continued from previous page).

observed in tail skin from these mice, or in pinna or tail skin of TAM treated *K5CE*^{wt} *R26R*^{tg} mice (Figure 3.9e).

The observed βgal activity in epidermal tissues suggested that the apparent absence of Cre-mediated βgal expression from *R26R* within tracheas was not due to a failure in achieving systemic TAM administration. Alternative explanations are considered in section 3.4.2.

3.3.2 Tamoxifen-induced xgal staining results from excision within the R26R *allele*

Although the βgal activity in *K5CE*^{tg} *R26R*^{tg} tissue, observed to correlate with TAM administration in experiments described in this chapter and in chapter 2 (section 2.3.5), had been assumed to be the result Cre-mediated site-specific recombination at *R26R* alleles, this had not been demonstrated directly. To this end, a PCR assay was developed in which a primer pair separated by the floxed NEO cassette of the *R26R* construct, could only amplify a product after excision of the intervening sequence (Figure 3.10a). The lacZ4 primer was designed by Rijnkels and Rosen, who also noted that DNA extracted from fixed tissue, as used here, can only support PCR amplification of fragments of up to approximately 800bp (Rijnkels and Rosen, 2001). This dictated that only the post-recombination amplification product could be obtained. The PCR assay was validated using DNA templates containing known recombined *R26R* alleles: *bK5Cre*^{tg} *R26R*^{tg} pinna DNA in which the epidermal cells carry recombined copies, and DNA carrying entirely recombined copies (the two right-most lanes on Figure 3.10b, respectively), the source of which is detailed in chapter 4.

DNA extracted from fixed and stained pinna tissue from mice 819, 820, 788, 828 and 905 (see Table 3.2) was used as template in the reaction (Figure 3.10b). Both stained and non-stained pinna tissue from mice 819 and 820 was used, to check that incomplete staining of pinna from these mice (Figure 3.9a-c(iii)) was the result of incomplete stain penetration, rather than regional activation of βgal expression.

Though very faint, a product corresponding to that seen in the positive controls was observed in both 819 samples, the "stained" 820 sample, and sample 788. No product could be visualised from samples 828 and 905, or from *K5CE*^{wt} *R26R*^{tg} mice, or *K5CE*^{tg} *R26R*^{tg} mice not exposed to TAM. Since the PCR products were difficult to visualise directly, hybridisation with a loxP sequence-specific oligonucleotide probe confirmed the presence of all products, plus a faint product from the "non-stained" 820 sample (Figure 3.10c). This indicated that Cre-mediated recombination had occurred throughout TAM treated *K5CE*^{tg} *R26R*^{tg} pinna, despite staining being restricted to regions of high stain accessibility. Other than the positive controls, no additional samples yielded PCR products.

PCR on template mixtures containing varying proportions of unrecombined and recombined *R26R* alleles were performed to assess the sensitivity of the PCR assay (Figure 3.10d). A product was faintly visible when 10% of *R26R* loci were recombined, but not when only 5% were recombined, suggesting that recombination in pinna resulting from systemic TAM administration exceeded 5%.

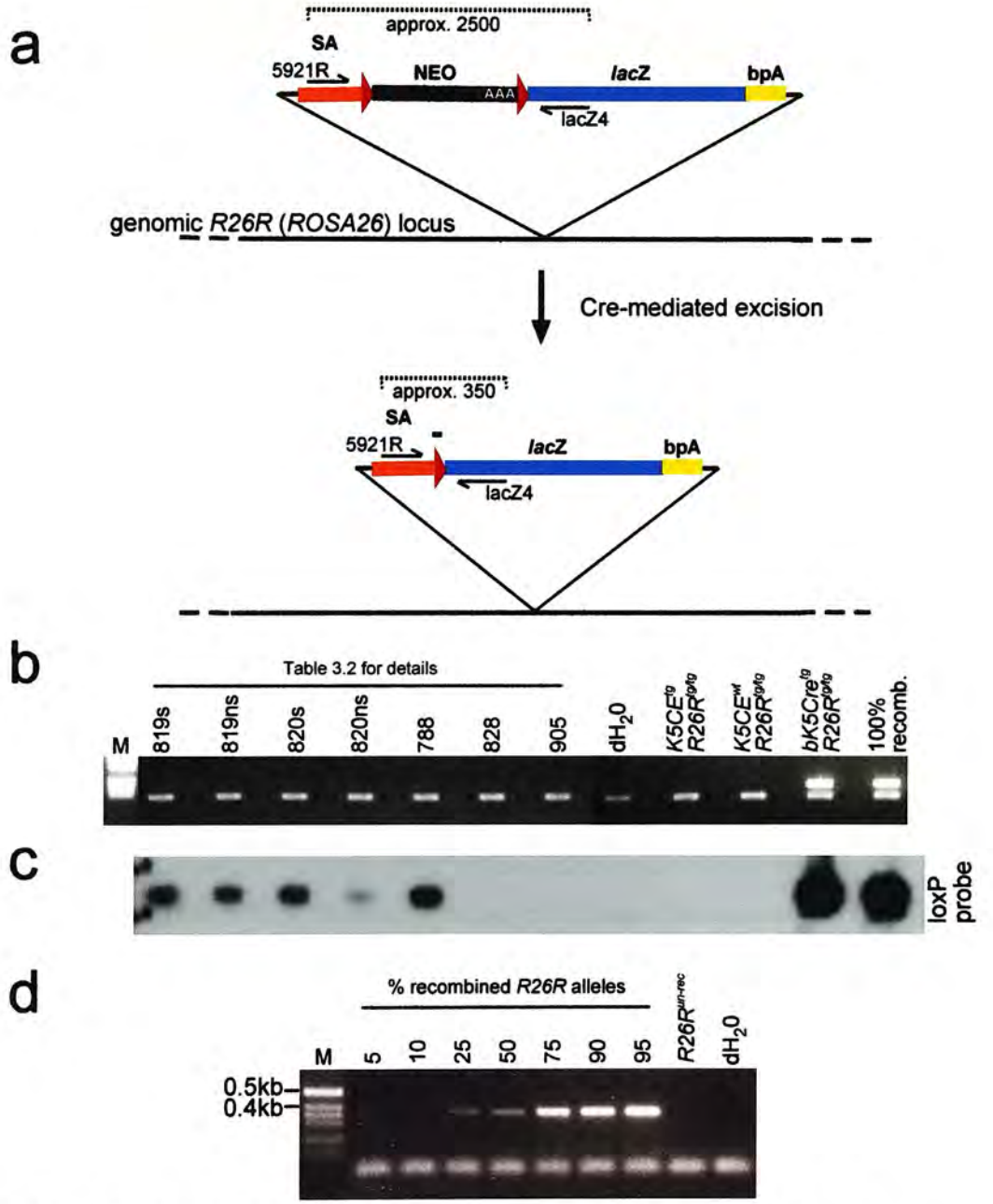


Figure 3.10 Tamoxifen-induced βgal activity correlates with recombination at *R26R* alleles.

(a) Schematic of the *R26R* allele before and after Cre-mediated recombination, which results in the loss of the floxed NEO cassette. Primers for detecting the recombined allele are positioned in the splice acceptor (SA) and the 5' of *lacZ*. PCR amplification only occurred from the recombined allele, when the region between the primers is reduced from approximately 2500bp to approximately 350bp. The loxP sites are represented by red triangles. The position of an oligonucleotide probe consisting of the 34bp loxP sequence is shown by a green bar. bpA (polyadenylation sequence), AAA (triple polyadenylation sequence). Map not to scale, fragment sizes in bps. (b) 5921R - lacZ4 PCR on pinna DNA of five mice following tracheal damage and tamoxifen treatment. The band apparent in every lane is the PCR primers. Although barely visible in the figure, a PCR product was observed on the original photograph in lanes corresponding to mice 819, 820, 788. "s" and "ns" after the mouse number indicates DNA from stained and non-stained region of pinna, respectively. No visible product was generated from mice 828 and 905, or dH_20, or untreated *K5CE*^{tg} *R26R*$^{tg/tg}$ or *K5CE*^{wt} *R26R*$^{tg/tg}$ controls. Robust product was amplified from DNA from pinna showing epidermal staining (*bK5Cre*^{tg} *R26R*$^{tg/tg}$), and ubiquitous staining ("100% recomb.", all *R26R* alleles recombined). (c) To verify the presence and identity of the weak PCR products, the bands were transferred to a nylon membrane and hybridised with a labelled oligonucleotide specific for the loxP sequence. Only samples 819, 820, 788 and the positive controls hybridised with the probe. (d) Genomic DNA containing 100% recombined *R26R* loci was mixed with DNA in which all *R26R* loci were unrecombined (*R26R*$^{un\text{-}rec}$) at the ratios shown (5%-95%), and used as template in 5921R - lacZ4 PCR. 10% recombined *R26R* loci was the lowest ratio which produced a visible product. M = 1kb ladder.

3.4 Discussion

Putative airway stem cells have previously been identified, on account of their long-term label-retaining characteristics, to reside in and around SMG ducts in the upper trachea, and at positions corresponding to cartilage-intercartilage junctions in the lower trachea (Borthwick *et al.*, 2001). In common with known stem cell niches in several other epithelial tissues, the SMG duct represents a well protected compartment distinct from, but spatially proximal to, the differentiated cell population (the luminal surface epithelium) (Cotsarelis *et al.*, 1989). In addition to appearing morphologically undifferentiated, consistent with the expected "primitive" phenotype (Cotsarelis *et al.*, 1989), the LRCs were observed to be spatially coincident with cells expressing high levels of keratin proteins, and which expressed a *lacZ* transgene driven by the human K5 promoter (Borthwick *et al.*, 2001). The expression, or otherwise, of specific keratins has previously been found to distinguish stem cells from their more differentiated progeny (Schermer *et al.*, 1986; Lyle *et al.*, 1998). To test the hypothesis that the high K5 promoter-expressing population represented tracheal progenitors, a transgenic mouse designed to permit temporally controllable activation of permanent lineage marking within the high K5-expressing cells was created (*K5CE*, chapter 2). A preliminary experiment to evaluate the usefulness of one of the *K5CE* lines in marking the lineage arising from the putative stem cell niche in response to tracheal epithelial injury, has been described. An inducible system was used to prevent unanticipated precocious activity of the Cre transgene resulting in widespread, uninterpretable reporter expression. In attempting to validate this decision, it was observed that a constitutive (non-inducible) system (*bK5Cre*) effected less extensive reporter activity than had been anticipated, and thus might be of use for the same purpose. A pilot study using the *bK5Cre* mouse has been described.

3.4.1 Constitutive lineage marking: bK5Cre x R26R

Prior to the studies described in section 3.2, the pattern of tracheal expression from the bovine K5 promoter was not known. In fact, the precise distribution of bK5 promoter-expressing cells remains unknown, since the xgal staining of cells observed in *bK5Cre^{tg} R26R^{tg}* mice did not distinguish between such cells and their progeny. However, it is clear that the bK5 promoter is active in a subset of cells throughout the tracheal luminal epithelium, and in the SMGs. The presence of staining in untreated and PBS treated tracheas suggested that, if the bK5 promoter was active in progenitor cells, then these cells had a role in the steady-state trachea, regardless of any involvement in epithelial regeneration. Support for the notion that bK5 promoter-expressing cells had progenitorial capacity derived from the observation on tracheal sections that βgal-positive cells were arranged into continuous stained blocks, consistent with clonal expansion. Despite the moderate level of staining displayed by untreated tracheas (Figure 3.2g), it was anticipated that recruitment of stem cells during repair of induced epithelial damage would result in a detectable overall increase in the number of stained cells. However, when stained wholemount POL and PBS treated tracheas from the various timepoints were compared, although large differences in the intensity and extent of staining were apparent, the variation did not obviously correlate with the treatment received or the timepoints at which the tracheas were taken. Although a marginally statistically significant correlation between POL treatment and an increase in the frequency of stained cells and the size of stained blocks was detected, consistent with the stimulation of bK5 promoter-expressing progenitors, there was no significant effect relating to the time at which POL-treated tracheas were collected (after the +1d timepoint, when no epithelial cells were present in the POL treated trachea). This may have been because, since reconstitution of the columnar layer was largely complete by the +3d timepoint, no further expansion of stained blocks occurred. Interestingly, in a recent report by Hong and colleagues (Hong *et al.*, 2003), discussed in section 3.4.3, the size of stained groups of tracheal epithelial cells was noted to continue to increase following the return of the epithelium to its normal composition after damage.

In the absence of a correlation between the individual-specific variation in tracheal staining and factors such as age or sex, or the likely influence of technical issues or endogenous enzyme activity, it may be that genetic background differences underlie the inconsistencies in the frequency of βgal-positive cells. This could occur by modulation of the number or distribution of bK5 promoter-expressing putative progenitor cells. The unusual frequency and distribution of SMGs in the trachea of mouse 129 (Figure 3.6e and Figure 3.7a) may argue in favour of segregating genetic elements affecting patterns of xgal staining, since SMG distribution in the mouse is known to differ between inbred strains (Borthwick, 1999), and the location of at least one genetic determinant has been elucidated (Innes and Dorin, 2001). Since complete breeding records were not available for the *bK5Cre* or *R26R* mice, the possible sources of genetic variability cannot be identified. However, the combining of genetic backgrounds from two separate transgenic lines will undoubtedly have resulted in an outbred genotype. Given the time required to make the animals congenic through backcrossing, the use of large numbers of animals per treatment and timepoint group may be the only feasible means of revealing statistically meaningful trends.

One other point to consider as regards detecting differences in the frequency of marked cells in repaired compared to undamaged epithelia, is that is that there is no internal control for tracheal damage. Since, as determined by earlier work, a columnar layer is expected to be reconstituted within three days of POL treatment (Borthwick *et al.*, 2001), an efficiently regenerated epithelium should appear superficially indistinguishable from an unsuccessfully damaged one. While instances of poor or negligible epithelial injury in this experiment cannot be excluded, the tracheal damage protocol had previously been shown to effective and reproducible (section 3.2.2).

Regardless of actual numbers of stained cells, the patterns of apparent lineage revealed on tracheal sections, irrespective of treatment, were intriguing. Numerous examples of continuously stained epithelial blocks extending from gland ducts to the epithelial surface, occasionally on either side of the gland duct opening (for example,

Figure 3.6f(vi)), were noted. This pattern is precisely what would have been predicted by the model based on SMG duct LRC data (Borthwick, 1999) (Figure 3.11), assuming that the K5-expressing cells correspond to the LRCs (and that LRCs do, in fact, represent progenitors), or that an overlap between these populations exists. Serial sectioning would have been useful in allowing indirect spatial relationships between stained glands and local epithelial blocks to be clarified. The apparent involvement of (bK5-expressing) SMG duct progenitors in steady-state epithelial maintenance is consistent with observations from the tracheas of aggregation chimaeric *ROSA26*:wild-type mice (Borthwick, 1999). SMGs were found to be either entirely derived from *ROSA26* cells, or wild-type cells, indicative of their clonal origin. Interestingly, whenever a *ROSA26*-derived gland was found in a tracheal glandular region also containing βgal-negative wild-type glands, a finite region of surface epithelium surrounding the *ROSA26* gland shared its xgal stained phenotype, suggestive of a lineage relationship.

While high magnification examination of proximal tracheal stained cell blocks appeared to reveal the presence of basal cells as well as columnar cells, the intensity of staining meant that basal cells could only be discriminated when they occurred against a background of unstained columnar cells, either at the edge of, or separate from, stained blocks. Certain cells within groups of stained basal cells not associated with columnar stained blocks or SMG duct lineages were likely bK5 promoter-expressing cells, possibly with unipotent differentiation potential to give rise to adjacent stained basal cells, as has been described (Hong *et al.*, 2003). When stained basal cells were detected adjacent to stained columnar blocks, it was unclear whether they represented the furthest extent of a lineage block, whether they gave rise to the block, or whether they were unrelated to the block. The use of cell type-specific markers such as *Griffonia simplicifolia* isolectin B_4 (GSI-B_4) for basal cells (Shimizu *et al.*, 1991), and antibodies against Clara cell secretory protein (CCSP) for columnar secretory cells and acetylated tubulin for ciliated cells (Hong *et al.*, 2003), would be an obvious next step in evaluating the cellular composition of stained blocks. However, inferences regarding the multipotency of the progentitors from which these blocks arose could only be made once the clonality of the blocks was confidently

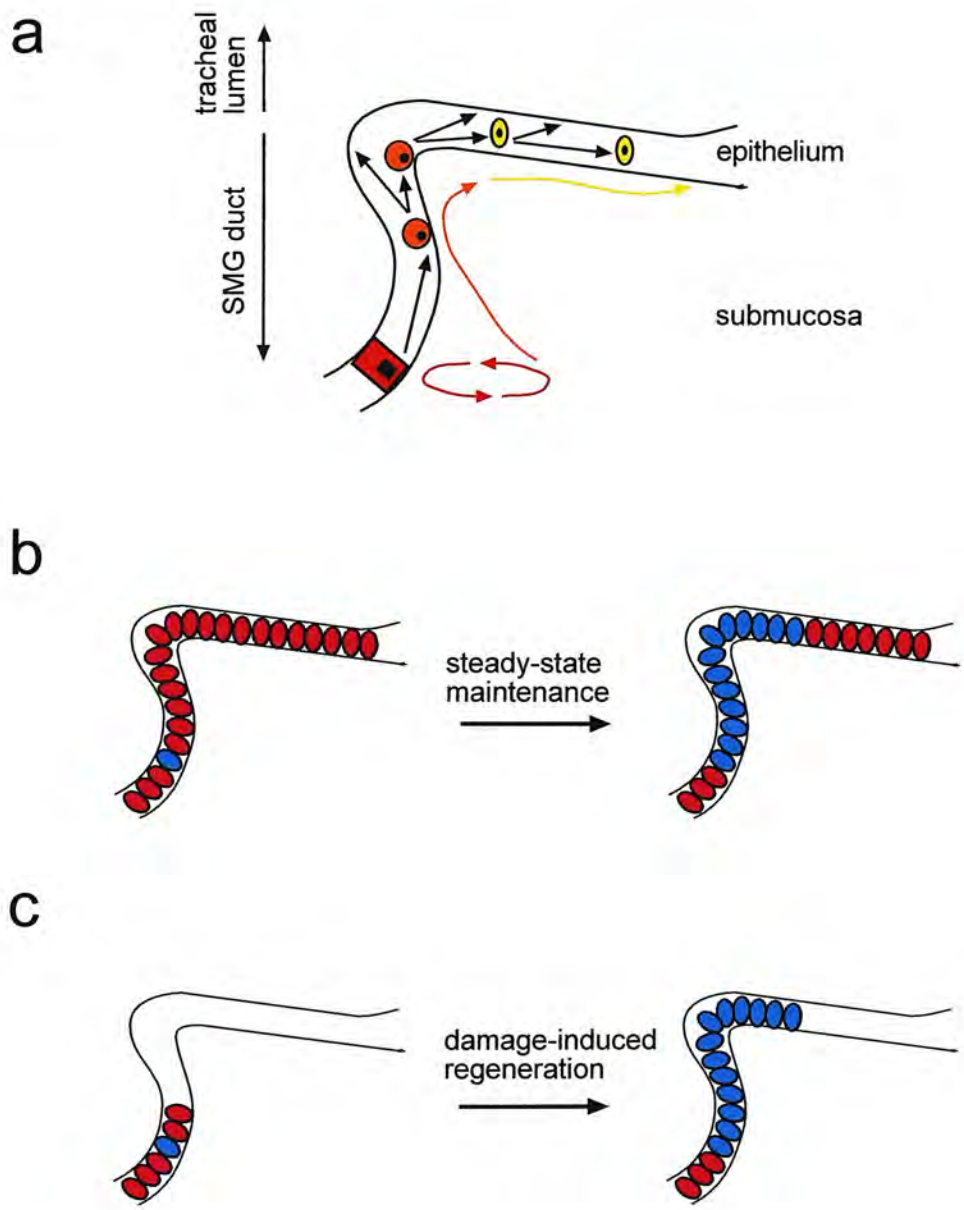


Figure 3.11 Observed apparent lineage patterns are consistent with predictions of the three zone label retention model.

Apparent patterns of cell lineage connecting the submucosal gland (SMG) duct with the surface epithelium resemble the migration of cells described by the model based on tracheal label retention data (Borthwick, 1999). (a) The three zone model of label retention (annotated in Figure 3.1). (b,c) Cre-mediated activation of permanent βgal expression in a bovine K5 promoter-expressing SMG duct cell could result in observed xgal staining patterns (for example, Figure 3.6f(iii,vi,viii)) during steady-state replenishment (b), or after damage-induced repair (c), of tracheal surface epithelium.

established. Examination of serial sections would assist greatly in this regard, and would be essential in determining the likely origin of many stained blocks which, on the limited number of sections analysed in this pilot study, did not display a clear relationship with the underlying SMGs. The use of anti-Cre with anti-βgal antibodies may prove an elegant means of identifying the cellular source of stained blocks (preliminary attempts at this were unsuccessful, data not shown). Examination of xgal stained and/or immunostained wholemount tracheas using Optical Projection Tomography (OPT) (Sharpe *et al.*, 2002) would allow visualisation of clones in three dimensions, and could be used to produce virtual tissue sections. Also, given that a columnar layer is largely reconstituted within three days of POL damage, but is still absent after 24 hours, the inclusion of a +2d timepoint would provide a valuable "snapshot" of lineage patterns during the early repair phase, and may yield clues as to the origin of marked blocks. These techniques and modifications to the experimental protocol would also be relevant within the context of a successful inducible lineage marking study.

While the *bK5Cre*^{tg} *R26R*^{tg} mice undoubtedly provided tantalising hints as to possible lineage relationships, at least in the upper trachea, it may be that the sheer frequency of stained cells would continue to prove an impediment to the confident identification of clonal lineages. Elimination of ambiguities of this sort was the rationale behind the construction of the inducible *K5CE* mouse, the preliminary evaluation of which is discussed in section 3.4.2.

Despite previous observations that (human) K5 promoter-driven transgene activity (Schoch *et al.*, 2003) was localised to epithelial foci corresponding to underlying cartilage-intercartilage junctions in the distal trachea, an analogous pattern of xgal staining was not observed in this study. The absence of large stained blocks of cells, and of discernible patterns in the distribution of those cells which did stain, in the distal trachea is consistent with the notion that related, but likely different, lineage systems operate in the proximal and distal tracheal zones. It has been speculated that cell renewal in the upper trachea relies more on gland duct and surface basal cell progenitors, whereas the glandless lower trachea, which contains slightly more Clara

cells than the upper trachea (Pack *et al.*, 1981), may utilise a Clara cell based lineage system (Schoch *et al.*, 2003).

3.4.2 Inducible lineage marking: K5CE x R26R

The preliminary experiment permitted by time on one of the three functional *K5CE* lines described in chapter 2 section 2.3 was disappointing in that no induction of βgal expression was evident in the tracheal epithelium, despite activation of the *R26R* allele in epidermal tissues. While not informative with regard to tracheal lineage patterns, the experiment is included to demonstrate the general experimental design which would be followed if time was available to repeat the study, as this result does not invalidate the strategy, or the use of *K5CE* mice for this purpose, which merit further evaluation. Moreover, the experiment was valuable in suggesting useful modifications to a range of parameters.

Given the rapid and generally complete re-epithelialisation observed after POL-induced damage of $bK5Cre^{tg}$ $R26R^{tg}$ tracheas, the variable and apparently impaired epithelial restoration which occurred in the $K5CE^{tg}$ $R26R^{tg}$ tracheas was presumed to relate to the poor condition of many of the animals following consecutive treatment with POL and TAM or oil. While such a compromised epithelial repair context is clearly not an ideal situation in which to monitor lineage patterns, there is no obvious reason why induced lineage marking could not have been activated in the repair that did occur. As a result of the unexpected mortality and morbidity observed (both in TAM and oil alone treated animals), a repeat of this experiment would necessitate a revised protocol, likely separating the tracheal damage and the TAM administration onto different days. Such a strategy was used successfully, albeit with a different damaging agent, by Hong and co-workers (Hong *et al.*, 2003), discussed in section 3.4.3.

Before any repeat of this experiment is undertaken, further efforts should be made to establish whether lines *K5CE600,700,1100* express the $CreER^{T2}$ transgene in the SMG ducts and surface basal epithelial foci, as expected of the human K5 promoter.

A preliminary investigation of cell lineage in regenerating tracheal epithelium

In a preliminary experiment, an antibody previously described to bind the CreER^{T2} fusion protein (Doerflinger *et al.*, 2003) showed no binding to sectioned *K5CE* or *bK5Cre* tracheas. Its failure to bind the latter, which clearly contains Cre protein, rendered the experiment uninformative. It may prove useful to induce nuclear localisation of the CreER^{T2} protein by TAM treatment prior to immunodetection, since this has been shown to increase the intensity of the signal (Indra *et al.*, 1999). Previous efforts to demonstrate SMG duct expression from the *lacZ*-expressing *K5A1* lines were hindered by apparent endogenous galactosidase activity, levels of which were high in that experiment (chapter 2, section 2.2.6). *In situ* hybridisation to the *CreER^{T2}* transcript would be an alternative means of evaluating expression of the *K5CE* transgene. If it can be unambiguously shown that *K5CE* tracheas do not express CreER^{T2} appropriately, then the experiment is clearly destined to fail. However, all three functional lines need to be rigorously analysed to determine whether the tracheal *CreER^{T2}* expression detected by RT-PCR is reflected in an appropriate pattern of CreER^{T2} protein. The need to test all three lines is emphasised by the experience of Schoch and colleagues, who found levels of K5 promoter-driven eGFP to vary between tracheas of different lines (Schoch *et al.*, 2003).

While the successful induction of CreER^{T2} activity in examined epidermal tissues validated the TAM preparation and route of administration, and could be interpreted as indicating the absence of CreER^{T2} protein in the trachea, it may simply have been that the dosage, or the timing or frequency of administration, was not appropriate for targeting the relatively few cells expected to produce the protein in the trachea. These parameters should be experimentally investigated, as should alternative modes of delivery, such as topical application by intratracheal instillation or nebulisation. In addition to TAM, its metabolite 4-hydroxytamoxifen (OHT) could be used. While several reports describe similar results using both OHT and TAM (Danielian *et al.*, 1998; Metzger and Chambon, 2001), OHT has appeared more effective in at least one study (Zhang *et al.*, 1996). Comparative evaluation of inducible Cre activity in each line may be more easily achieved *ex vivo*, administering ligand to injured tracheas maintained in a tracheal culture system (Scott *et al.*, 2000).

3.4.3 Validation of an inducible approach to tracheal epithelial lineage marking

A recent study by Hong and colleagues revealed multipotent differentiation potential of tracheal basal cells in response to naphthalene-induced ablation of secretory (Clara) cells (Hong *et al.*, 2003). This experiment also demonstrated proof of principle for inducible recombinase-mediated lineage analysis in the regenerating tracheal epithelium. Systemic naphthalene had previously been shown to stimulate proliferation of neuroendocrine cells (Hong *et al.*, 2001), and of a pollutant-resistant subpopulation of Clara cell secretory protein (CCSP)-expressing cells associated with either the neuroepithelial body (NEB) (Reynolds *et al.*, 2000a) or the bronchoalveolar duct junction (BADJ) (Giangreco *et al.*, 2002), in the distal airways. More recently, bronchial epithelial cells that assumed a K14-expressing phenotype following injury were shown to exhibit multipotent differentation potential (Hong *et al*, unpublished results). In the most recent study the authors determined that in the steady-state trachea, a subset of tracheal basal cells were immunoreactive for K14, and were usually found in small clusters associated with epithelial regions above cartilage rings (exactly how this distribution differs from the cartilage-intercartilage junction distribution described for K5 promoter-driven transgene expression, is unclear) (Hong *et al.*, 2003). However, in response to naphthalene-induced ablation of Clara cells, basal cell hyperplasia was observed, and was associated with expression of K14 in virtually all basal cells. The correlation between restoration of a Clara cell-containing pseudostratified epithelium, and regression of the hyperplastic K14-expressing basal cells, was suggestive of a progenitor-progeny relationship.

To determine the differentiation potential of tracheal basal cells, the authors used a bipartite transgenic lineage marking approach analogous to the *K5CE R26R* strategy attempted in this study (Hong *et al.*, 2003). Using mice harbouring a K14 promoter-driven inducible Cre transgene (*CreER*^{TAM}, (Vasioukhin *et al.*, 1999)) in combination with the *R26R* transgene, naphthalene treatment was followed two days later by TAM administration, resulting in apparently clonal clusters of xgal stained cells throughout the tracheal epithelium by day six. The βgal-positive clusters were

enlarged in tracheas taken at 21 days, at which time the epithelium had returned to its normal composition, and were observed to have increased further in size by day 43. Analysis of the phenotypes of cells present within the stained clusters revealed that while some clusters contained only basal cells, indicative of a unipotent progenitor, the majority of clusters harboured multiple (basal, columnar, ciliated) cell types, revealing the existence of K14-positive basal cells capable of multipotent differentiation following secretory cell depletion. The authors did not conduct lineage marking on undamaged mice, so the contribution of these cells to steady-state maintenance of the tracheal epithelium is unclear.

The basal phenotype of the progenitors identified after Clara cell ablation is consistent with several previous studies suggesting that basal cells represent airway progenitors (Inayama *et al.*, 1988; Breuer *et al.*, 1990; Ford and Terzaghi-Howe, 1992b; Shimizu *et al.*, 1994; Engelhardt *et al.*, 1995). Since K14 is generally co-expressed with K5 (Byrne and Fuchs, 1993), marked cells might have been expected within SMGs, however none were observed (Hong *et al.*, 2003). The authors speculated that this may have resulted from the low frequency with which lineage marking was activated, or the lack of injury among epithelial cells of the SMG. The latter explanation is consistent with the established notion that the selectivity of injury towards specific airway epithelial cell types is an important determinant of progenitor cell activation. It has been shown that secretory cells of both the proximal and distal airway epithelium of rats act as the main progenitors for replacement of the epithelium in response to oxidant-induced injury to terminally differentiated ciliated cells, while basal cells were implicated in the steady-state maintenance of this tissue (Evans *et al.*, 1976; Evans *et al.*, 1986). As mentioned, naphthalene-induced ablation of secretory cells was accompanied by proliferation of neuroendocrine cells and a naphthalene resistant population of Clara cells in the bronchioles (Reynolds *et al.*, 2000a; Hong *et al.*, 2001; Giangreco *et al.*, 2002), and K14-positive basal cells in the tracheobronchial epithelium ((Hong *et al.*, 2003) and Hong *et al*, unpublished results). Epithelial damage resulting from exposure to POL or SO_2, both of which remove columnar cells and, to varying extents, basal cells, stimulates slow-cycling basal cells to take up a nuclear label (Borthwick *et al.*,

2001). Thus, it is conceivable that the mode of damage effected by naphthalene was not appropriate for the stimulation of the putative SMG duct progenitors.

3.4.4 Further evidence of progenitorial capacity amongst tracheal basal cells

Despite the absence of marked lineage within SMG ducts in the study of Hong and co-workers (Hong *et al.*, 2003), recent experiments by Schoch and colleagues have provided additional support for the notion that the K5-positive basal cells postulated to equal tracheal LRCs (Borthwick *et al.*, 2001), have progenitor-like characteristics (Schoch *et al.*, 2003). Examination of a transgenic mouse expressing eGFP from the human K5 promoter revealed a distribution of eGFP-positive basal cells which corresponded with the pattern of LRCs noted previously (Borthwick *et al.*, 2001; Schoch *et al.*, 2003): eGFP expression was focussed on tracheal gland ducts and the surrounding surface epithelium in the upper airway, and to systematically arrayed foci centred over cartilage-intercartilage junctions in the lower airway. The K5-*eGFP* mice were crossed with *ROSA26* mice such that the resulting compound transgenic animals expressed eGFP in a subset of tracheal basal cells, and βgal in all cells. Tracheal epithelial cell preparations were generated from K5-*eGFP*^{tg} *ROSA26*^{tg} mice, and the eGFP-expressing subset purified by flow cytometry. The colony forming efficiency (CFE) of the eGFP-positive fraction compared to the eGFP-negative fraction was compared by seeding the cells in an air-liquid interface cell culture system which supported secretory and ciliated cell differentiation, and staining with xgal to reveal *ROSA26*-derived colonies. The eGFP-positive population showed a consistently greater CFE than the eGFP-negative population. While the average difference in CFE was 4.5-fold, the eGFP-positive population exhibited a 12-fold higher CFE specifically for large colonies (>100 cells).

These results demonstrated a diverse progenitorial capacity amongst different tracheal epithelial cell types, and even within the basal compartment. Of course, while *ex vivo* clonogenic assays have proved useful for identifying putative stem cells (Jones and Watt, 1993; Rochat *et al.*, 1994), they represent an inherently

artificial context in which to analyse the potency of cell populations, and should ideally be interpreted as indicative of a given cell's proliferative potential, rather than its behaviour or function within its native physiological situation. Nevertheless, the enhanced clonogenic properties of the eGFP-positive cells is consistent with the notion that a high K5 promoter-expressing subset of tracheal basal cells, thought to correspond to LRCs, may represent the stem or early TA progenitor population. These findings support further attempts to mark the lineage of the high K5-expressing population *in vivo* using the *K5CE* mouse, which has the potential to indicate the physiological role of these cells, including those of the SMG duct. Additionally, the K5-*eGFP* mice could be used in a repeat of the original BrdU label-retaining experiment (Borthwick *et al.*, 2001) to determine whether the high K5-expressing population are in fact LRCs.

3.4.5 Are submucosal gland duct progenitors involved in steady-state maintenance of the tracheal epithelium?

While the lineage marking experiment of Hong and colleagues employed a selective mode of airway damage to reveal multipotency of a subset of tracheal basal cells, which may otherwise have been obscured by contributions from the secretory cell pool (Hong *et al.*, 2003), the tracheal injury used in the label-retention experiments of Borthwick and co-workers was simply a means to stimulate proliferation of stem cells which might normally act in steady-state maintenance of the tracheal epithelium (Borthwick, 1999; Borthwick *et al.*, 2001). Borthwick cites several pieces of evidence for the involvement of the SMG duct LRCs in steady-state epithelial renewal (Borthwick, 1999), including the observation that cells from all three of the LRC zones (Figure 3.1) undergo mitosis in steady-state human bronchial SMGs (Leigh, 1995), the presence of (a low number of) LRCs in the corresponding regions of undamaged tracheas (Borthwick, 1999), and the correlation between SMG and overlying epithelial origin in chimaeric tracheas (Borthwick, 1999). Preliminary lineage analysis using the *bK5Cre* mouse (section 3.2.3) would appear to further support the argument for a role of (b)K5 promoter-expressing SMG duct cells in steady-state tracheal maintenance, consistent with the concept of a functional

relationship between the glands and the surface epithelium resulting from the developmental derivation of SMGs from overlying epithelial cells (Thurlbeck *et al.*, 1961; Engelhardt *et al.*, 1995). However, more confident identification of clonal relationships within renewing or regenerating tracheal epithelium should be obtained through the use of the inducible lineage marking system, an approach recently validated (Hong *et al.*, 2003).

3.4.6 Future directions

3.4.6.1 Further evaluation of the inducible Cre transgenic lines

Given the preliminary nature, necessitated by time, of the experiments described in this chapter, further evaluation of inducible Cre lines *K5CE600,700,1100* is a priority, and efforts to this end are underway. The demonstration of human K5 promoter-driven transgene expression within SMG ducts (Schoch *et al.*, 2003) is positive with regard to the prospects for $CreER^{T2}$ expression within the same cells of *K5CE* mice. The recent report describing inducible lineage marking of K14 promoter-expressing cells in the trachea (Hong *et al.*, 2003) is encouraging from a proof of principle perspective, but could also be considered to render the *K5CE* mouse redundant, since K14 is generally regarded as the co-expressed partner of K5 (Byrne and Fuchs, 1993). However, the issue of K5 promoter-expressing progenitors within the SMG duct was not resolved using the K14-$CreER^{TAM}$ mouse, as no marked cells were evident within SMGs (Hong *et al.*, 2003). While this could have been an effect of the selectivity of the injury induced by naphthalene, or could have resulted from the overall low frequency of marked cells, it is possible that the K14 promoter did not reflect endogenous K14 activity in the gland ducts. Alternatively, it may be that the previously identified K5 promoter-expressing cells of the SMG duct are a separate population from those shown to express K14 protein (Borthwick *et al.*, 2001): K5 is known to partner with K15 in certain epithelia (Lloyd *et al.*, 1995) and in oocytes (Gallicano, 2001). Intriguingly, K15 is considered a marker for human hair follicle stem cells (Lyle *et al.*, 1998), and it may be interesting to check for its presence in tracheal epithelia. Since K5 and K14 are therefore not exclusively co-

expressed, the *K5CE* mouse may prove useful in revealing lineages arising from K5 promoter-expressing cells in which the K14 promoter is not active. Also, the *K5CE* mouse would be necessary to demonstrate an *in vivo*, physiologic progenitor role for the K5 promoter-expressing cells of the SMG duct, since the demonstration of an elevated *in vitro* CFE amongst K5 promoter-expressing tracheal basal cells did not distinguish between SMG duct and surface epithelial basal cells (Schoch *et al.*, 2003).

If further evaluation of the *K5CE* lines demonstrates their suitability for inducible marking of lineages arising from the high K5-expressing cells of SMG ducts, then it would be interesting to monitor their contribution to airway restoration in a range of injury contexts. In addition to POL and SO_2, NO_2- and O_3-induced oxidative damage has been used to specifically injure ciliated cells (Evans *et al.*, 1976), and bleomycin is known to cause damage to both epithelial airway cells and endothelial alveolar cells (Kotton *et al.*, 2001). Naphthalene has been used to specifically ablate Clara cells (Stripp *et al.*, 1995), while portions of airway epithelium have been completely denuded using mechanical techniques (Shimizu *et al.*, 1994). Observing lineage patterns arising in response to a variety of airway damage scenarios should provide valuable clues as to functional redundancies and potential hierarchies of progenitor or stem cell utilisation in the respiratory epithelium.

3.4.6.2 Inducible lineage analysis of other airway populations

The successful demonstration of the inducible lineage marking strategy in adult airways ((Hong *et al.*, 2003) and Hong *et al*, unpublished results) could be usefully extended to other cell types known to be involved in airway kinetics. For example, Clara cells and the naphthalene-resistant subset of Clara-like cells are defined by their expression of Clara cell secretory protein (CCSP) (Hong *et al.*, 2001). Likewise, pulmonary neuroendocrine cells (PNECs) are identified by their expression of calcitonin gene-related peptide (CGRP) (Hong *et al.*, 2001). Using the promoters of the CCSP and CGRP genes to drive expression of an inducible recombinase should

facilitate further elucidation of the potency of these cell types, and their contribution to airway maintenance and regeneration.

3.4.6.3 Further experiments using the constitutive Cre transgenic line

The bovine K5 promoter-driven constitutive Cre line (*bK5Cre*) appeared to reflect the SMG duct expression of the human K5 promoter, and could therefore be useful for following the lineage of these cells, particularly in the event that the *K5CE* lines prove refractory to attempts at inducing lineage marking. Although, as anticipated, the generally high frequency of marked cells prevented the unambiguous identification of clonally marked blocks, potentially interesting patterns of cell lineage were observed. Further investigation of lineage patterns arising from *bK5Cre*-expressing cells might be more informative in the context of aggregation chimaeras containing $bK5Cre^{tg}$ $R26R^{tg}$ compound transgenic cells and wild-type cells. This should reduce the frequency of marked lineages, increasing the confidence with which blocks could be identified as being of clonal origin. Alternatively, should further analysis reveal that the human K5 promoter does not drive expression of the $CreER^{T2}$ transgene in the tracheas of the three potentially useful *K5CE* lines, expression of the fusion gene from the bovine K5 promoter would be a suitable alternative.

3.4.6.4 Do circulating stem cells respond to tracheal injury?

While current research continues to support the notion of region-specific stem cells within the various compartments of the airway (Borthwick *et al.*, 2001; Hong *et al.*, 2001; Giangreco *et al.*, 2002; Schoch *et al.*, 2003; Hong *et al.*, 2003), it is important to consider the potential contribution from circulating bone marrow (BM)-derived stem cells. These cells could be useful for *ex vivo* gene therapy, since they offer the prospect of genetic manipulation before engraftment in a range of tissues. This would circumvent potential difficulties in targeting therapies at tissue-resident airway stem cells, which may prove inaccessible to luminally-delivered gene therapy vectors because of their proposed localisation to basal positions in the proximal tracheal

surface epithelium and gland ducts. Efforts are currently underway to evaluate the involvement of BM-derived cells after airway epithelial injury.

Chapter 4 The ability of the *bK5Cre* mouse to mediate both tissue-specific and generalised recombination

4.1 Introduction

In the course of studies into cell lineage in the mouse airway (chapter 3), marked inconsistencies in βgal expression were detected in the tracheas and other tissues of *bK5Cre*^{tg} *R26R*^{tg} compound transgenic mice (section 3.2.1). This is a separate observation from that regarding the described variation in frequency and intensity of xgal staining in tracheas from *bK5Cre*^{tg} *R26R*^{tg} animals (discussed in section 3.4.1). While some mice displayed, as expected, βgal expression in basal epithelial cells and their supra-basal daughter cells, other animals exhibited universal βgal production, indicating constitutive recombination of the *R26R* allele.

4.1.1 Aims

This chapter demonstrates the alternate expression patterns observed and investigates the molecular basis for this phenomenon.

4.2 Results

4.2.1 Alternate patterns of recombination in bK5Cre^{tg} R26R^{tg} *mice correlate with the sex of the* bK5Cre *transgene donor parent*

As demonstrated in chapter 3, the *bK5Cre* transgene, when combined with the *R26R* conditional reporter allele, causes expression of βgal in the basal cells of epithelial tissues (Figure 3.2). This is due to a permanent genetic rearrangement within the *R26R* allele and, because of the ubiquitously active nature of the *ROSA26* locus, all descendants of bK5 promoter-expressing cells express the reporter. In the skin, the entire epidermis including hair follicles and sebaceous glands becomes marked, as K5-positive basal epidermal cells divide to produce daughters which populate the supra-basal layers. This is shown in pinna section (Figure 4.1d), which also reveals that the dermis remains entirely unstained after xgal treatment. As shown previously in chapter 3, βgal is consistently and strongly expressed in the larynx at the proximal extent of the trachea, and to a more variable degree in the epithelial layer of the tracheal lumen, and the SMGs (Figure 4.1a).

However in some first generation litters inheriting both the *bK5Cre* transgene and the *R26R* locus, all animals staining after xgal treatment exhibited ubiquitous expression of βgal (Figure 4.1b,e), reminiscent of the *ROSA26* phenotype (Figure 4.1c,f), which would depend upon universal recombination of *R26R*. Intercrosses between *bK5Cre*$^{tg/wt}$ *R26R*$^{tg/wt}$ compound hemizygous animals that displayed K5-derived βgal expression consistently resulted in litters composed of ubiquitously staining animals, plus a minority of non-staining animals: such litters were devoid of animals exhibiting a K5-derived pattern of recombination.

In an effort to elucidate the pattern of segregation of the ubiquitously staining phenotype, embryonic litters were collected from crosses between several *bK5Cre*$^{tg/wt}$ and *R26R*$^{tg/wt}$ single transgene-positive, hemizygous mice (Table 4.1).

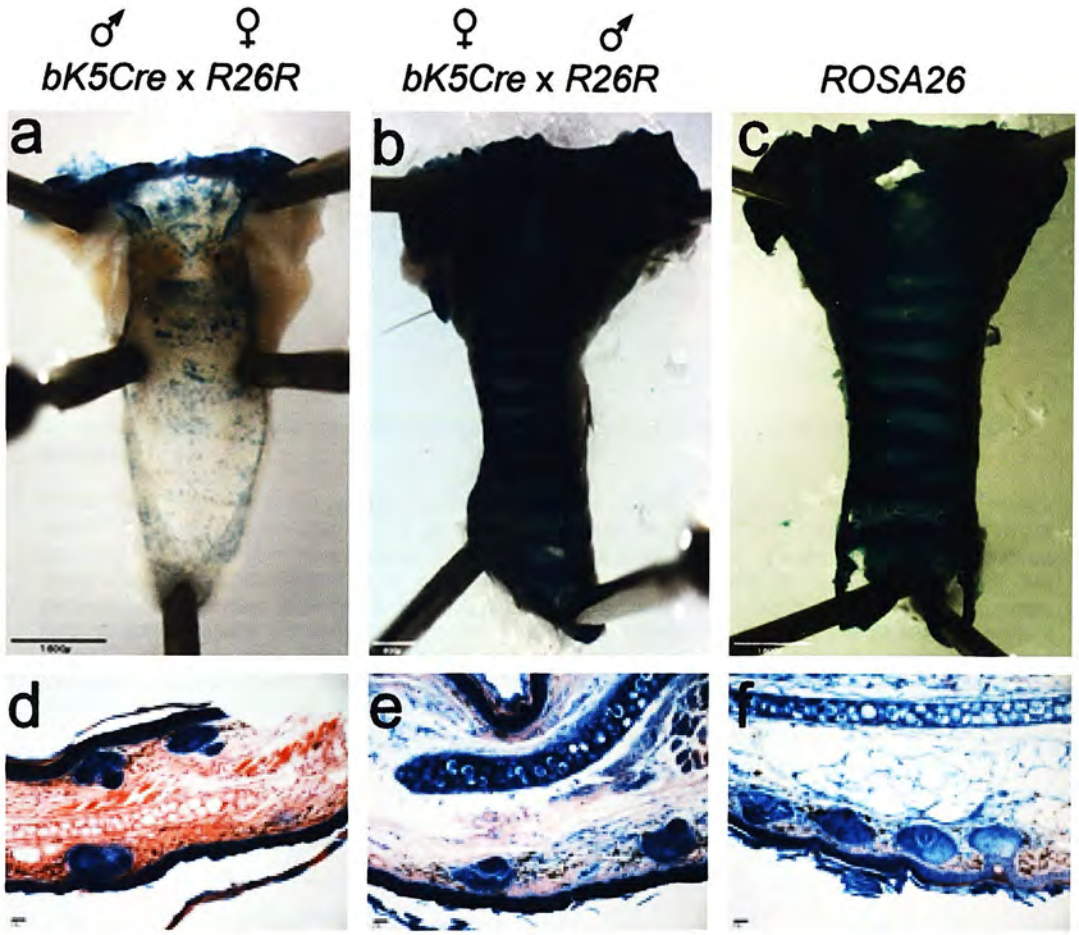


Figure 4.1 The pattern of Cre-mediated recombination differed depending on the sex of the *bK5Cre* transgene-transmitting parent.

Xgal-stained wholemount tracheas (a,b) and sectioned pinna (d,e) of mice resulting from reciprocal *bK5Cre* x *R26R* crosses. Corresponding *ROSA26* tissues are included for comparison (c,f). (a,d) *bK5Cre*^{tg} *R26R*^{tg} tissue from *bK5Cre* male x *R26R* female cross. (a) *R26R* recombination was strongest at the proximal end of the trachea, with less intense staining in the luminal surface. (d) Staining was confined to the epidermis of the pinna, including in the hair follicles which descend into the dermis. (b,e) *bK5Cre*^{tg} *R26R*^{tg} tissue from *bK5Cre* female x *R26R* male cross. (b) The trachea showed complete recombination of *R26R*, mirrored in the pinna (e), where an intensely stained layer of cartilage is flanked by the more weakly staining dermis. (b) and (e) exhibit comparable staining to the *ROSA26* tissues (c,f). Scale bar = 1600μm (a,c); 800μm (b); 20μm (d-f).

Crosses	Phenotype\ Genotype	K5-derived	Ubiquitous	None
bK5Cre$^{tg/wt}$ ♂ x *R26R*$^{tg/wt}$ ♀	*bK5Cre*^{wt} *R26R*^{wt}	0,0,0	0,0,0	2,2,1
	bK5Cre^{tg} *R26R*^{wt}	0,0,0	0,0,0	2,2,2
	bK5Cre^{wt} *R26R*^{tg}	0,0,0	0,0,0	3,1,2
	bK5Cre^{tg} *R26R*^{tg}	2,2,1	0,0,0	0,0,0
	Total	2,2,1 (5)	0 (0)	7,5,5 (17)
bK5Cre$^{tg/wt}$ ♀ x *R26R*$^{tg/wt}$ ♂	*bK5Cre*^{wt} *R26R*^{wt}	0,0,0	0,0,0	3,2,1
	bK5Cre^{tg} *R26R*^{wt}	0,0,0	0,0,0	3,1,3
	bK5Cre^{wt} *R26R*^{tg}	0,0,0	2,3,1	0,0,0
	bK5Cre^{tg} *R26R*^{tg}	0,0,0	3,2,0	0,0,0
	Total	0 (0)	5,5,1 (11)	6,3,4 (13)

Table 4.1 Genotype-phenotype correlations in embryonic offspring of reciprocal *bK5Cre*$^{tg/wt}$ x *R26R*$^{tg/wt}$ crosses.

Three litters from both types of cross were obtained: 9-, 7-, and 6-embryo litters from *bK5Cre*$^{tg/wt}$ male x *R26R*$^{tg/wt}$ female crosses and 11-, 8-, and 5-embryo litters from *bK5Cre*$^{tg/wt}$ female x *R26R*$^{tg/wt}$ male crosses. For each of the four possible embryonic genotypes, the number of mice from each of the three litters displaying a K5-derived or a ubiquitous βgal expression pattern, or no βgal expression, are recorded. The number of embryos showing each phenotype were totalled and revealed an expected ratio of approximately 1:3 between K5-derived expression:no expression from the *bK5Cre*$^{tg/wt}$ male x *R26R*$^{tg/wt}$ female crosses, with xgal staining embryos corresponding to those genotyped as compound transgenic. The *bK5Cre*$^{tg/wt}$ female x *R26R*$^{tg/wt}$ male crosses generated no embryos exhibiting a K5-derived xgal staining pattern, and the ratio of ubiquitous expression:no expression was approximately 1:1. Amongst the embryos showing ubiquitous βgal expression, approximately half were compound transgenic, while the remainder carried only the *R26R* allele.

Genotype	Fluorescence\ Follicle stage	Nuclear	Nucleolar	None
bK5Cre$^{tg/wt}$	Primary	1	0	47
	Mature	15	21	7
	Antral	15	20	5
non-transgenic	Primary	0	0	39
	Mature	0	0	21
	Antral	0	0	16

Table 4.2 Numbers of DAPI-positive oocytes displaying anti-Cre signal in *bK5Cre*$^{tg/wt}$ and non-transgenic ovaries.

DAPI-positive oocytes were scored for anti-Cre fluorescence (nuclear or nucleolar), and categorised by follicle stage. "Primary" follicles were small and surrounded by a single layer of follicle cells (Figure 4.7). "Mature" follicles were all follicles developmentally intermediate between primary follicles and "antral" follicles (not equal to the "mature follicle" annotated in Figure 4.7). "Antral" follicles were defined as those displaying large antral spaces (approximately corresponding to all stages subsequent to the "secondary oocyte" pictured in Figure 4.7). However, due to potential misrepresentation of follicle stages as a result of the plane of section viewed, these categorisations were not strict, particularly between the mature and antral follicles. Also, it is possible that the single *bK5Cre* anti-Cre-positive primary follicle was an incorrectly classified, more developed oocyte.

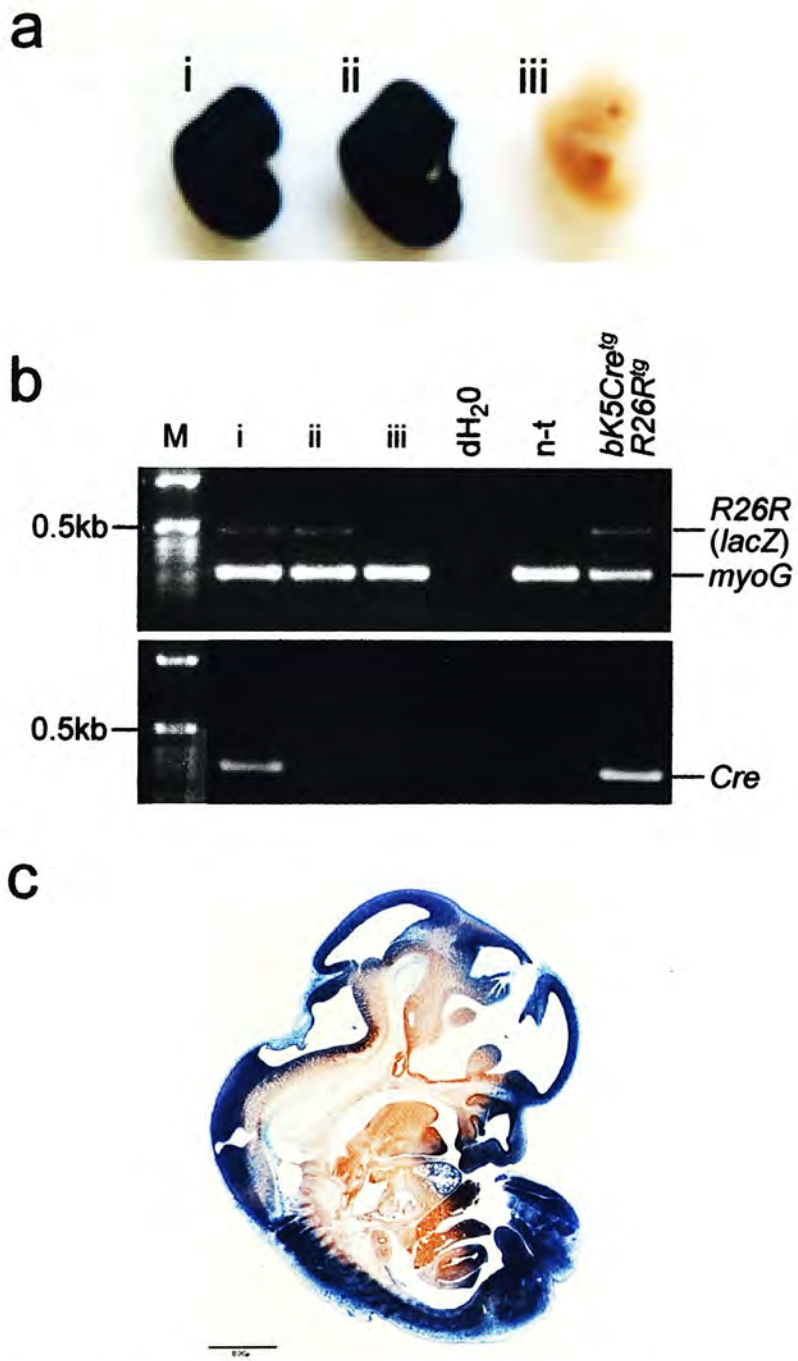


Figure 4.2 Universal recombination of *R26R* in offspring from *bK5Cre* females does not require transmission of *bK5Cre*.

(a) Xgal stained E13.5 embryos from a cross between a *R26R*$^{tg/wt}$ male and *bK5Cre*$^{tg/wt}$ female. (b) Genotyping of embryos for the *R26R* and *bK5Cre* transgenes revealed that ubiquitously stained embryo (ii) did not carry the *bK5Cre* transgene. Genotypes are *bK5Cre*^{tg} *R26R*^{tg} (i), *bK5Cre*^{wt} *R26R*^{tg} (ii), *bK5Cre*^{wt} *R26R*^{wt} (non-transgenic, iii). *Myogenin* (*myoG*) primers were used as an internal PCR control. (c) A sagittal section through an xgal stained embryo showing ubiquitous expression of βgal (lack of internal staining was due to incomplete xgal stain penetration). Compare with epidermal-specific staining in an *bK5Cre*^{tg} *R26R*^{tg} embryo from a cross between a *R26R* female and *bK5Cre* male, chapter 3, Figure 3.2a,b,e. M = 1kb ladder, n-t = non-transgenic. Scale bar = 800μm (c).

Xgal staining of resulting embryos revealed that the universal activation of *R26R* occurred only when the *bK5Cre* transgene was carried by the female parent. Furthermore, genotyping of the embryos revealed that this staining pattern was not dependant on inheritance of *bK5Cre* (Figure 4.2).

4.2.2 Cre activity in bK5Cre *oocytes causes universal recombination in the subsequent generation*

That universal recombination of *R26R* was apparent only in progeny descended from *bK5Cre* females, combined with the observation that the *bK5Cre* transgene was not required to mediate this event in the offspring, suggested that recombination was occurring at the beginning of embryonic development, and being propagated through subsequent cell division. Accumulation of Cre protein within oocytes would explain these findings, with persistence of protein into the early zygote necessary to effect the observed recombination of paternally-derived *R26R* alleles in *bK5Cre*-negative offspring.

To investigate this, immunofluorescence detection of Cre protein was performed on sectioned ovaries from three hemizygous *bK5Cre* mice (Figure 4.3). As it was anticipated that Cre protein would be abundant immediately preceding fertilisation, the animals were first superovulated to elevate the proportion of mature oocytes. Highly fluorescing foci were observed in *bK5Cre* oocytes (Figure 4.3d-f), but were absent in non-transgenic oocytes (Figure 4.3l). Serial sections revealed a correlation between DAPI fluorescent staining of chromatin and anti-Cre fluorescence, indicating that Cre was localised to the oocyte nucleus (Figure 4.4). The spatial relationship between the Cre signal and the chromatin conformation of a subset of oocytes revealed the protein to be particularly abundant in nucleoli (Figure 4.3c,f). S. Webb provided some technical assistance with establishing the anti-Cre immunoassay.

The fact that all embryos inheriting a *R26R* allele from crosses between $R26R^{tg/wt}$ males and $bK5Cre^{tg/wt}$ females stained ubiquitously after xgal treatment suggested

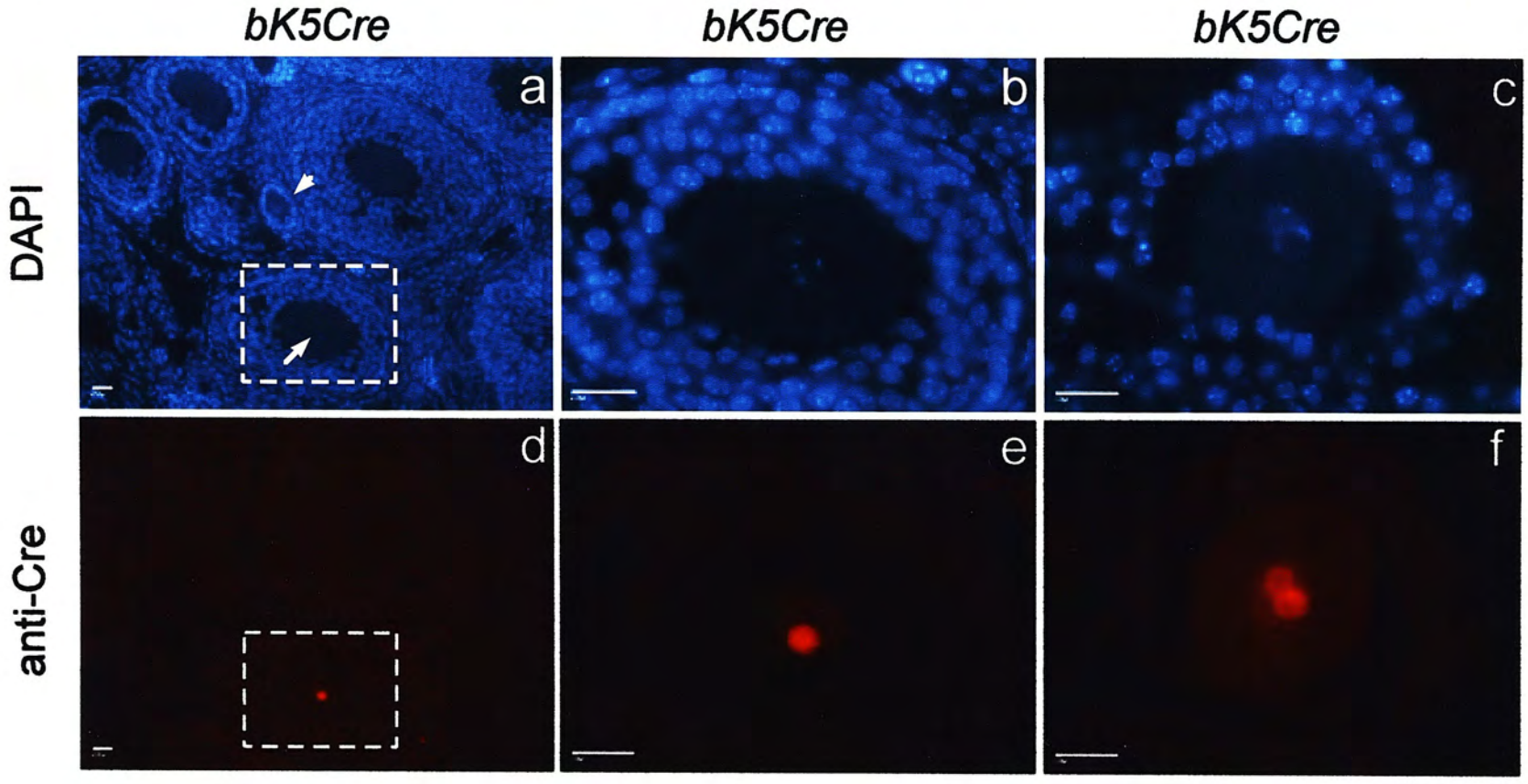


Figure 4.3 Immunofluorescence detection of Cre protein in *bK5Cre* oocytes (continued over page).

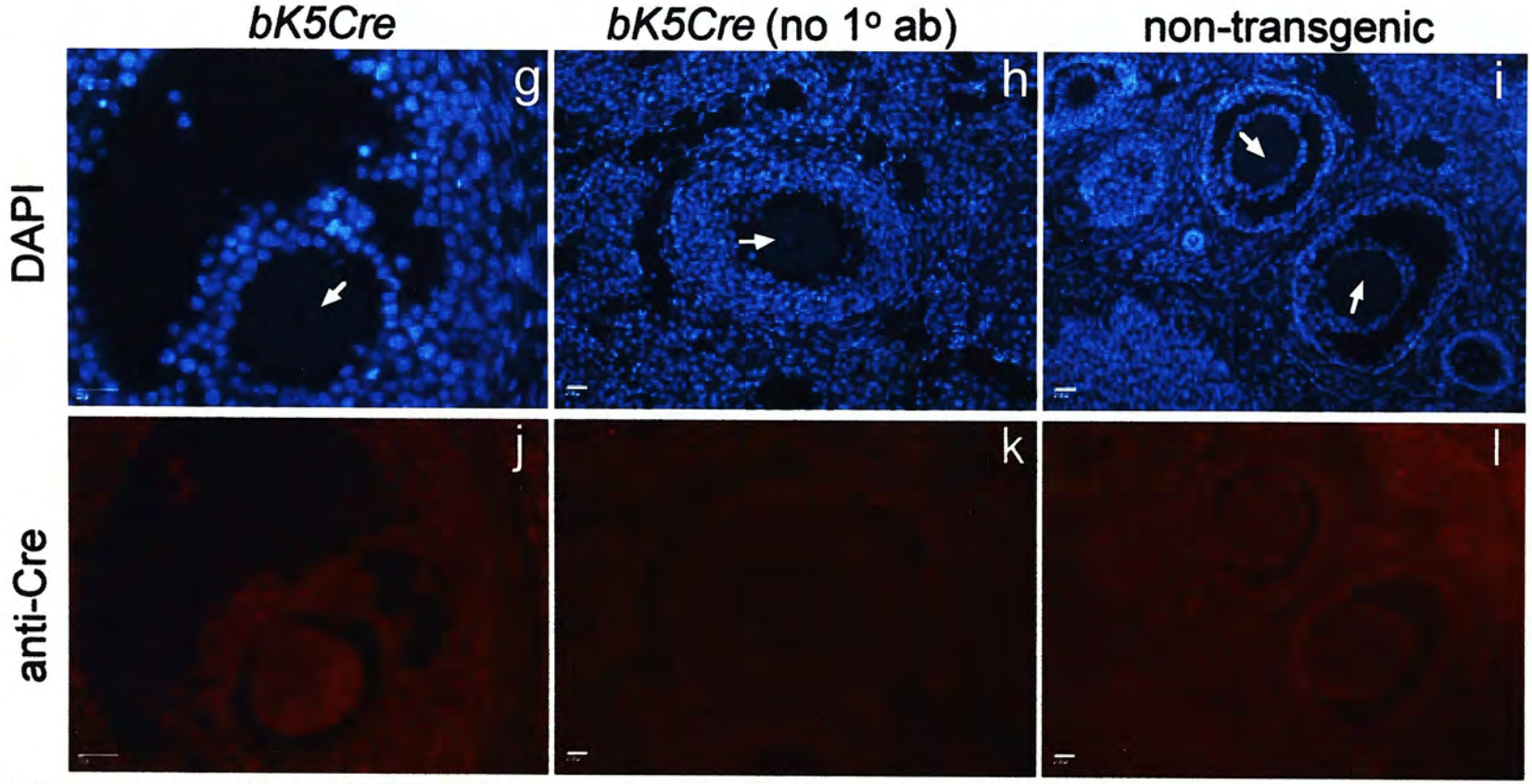


Figure 4.3 Immunofluorescence detection of Cre protein in *bK5Cre* oocytes (continued from previous page).

Sectioned ovaries from superovulated hemizygous *bK5Cre* and non-transgenic mice were sectioned and stained with anti-Cre antibody, visualised with a fluorescent secondary antibody, and counterstained with DAPI. DAPI staining revealed oocytes through which the plane of sectioning had encountered the nucleus (a-c, g-i). (a,d) Five follicles from a *bK5Cre* ovary. The smaller central follicle is a single-layered primary follicle (see Figure 4.7). The nucleus is visible in only two of the five oocytes, the primary follicle (arrowhead) and a mature follicle (arrow). Cre protein was present in the mature oocyte, but not in the primary follicle. (b,e) Higher magnification of the positive oocyte boxed in (a,d), revealing that Cre protein was localised to the nucleus, and appeared particularly concentrated within a sub-nuclear compartment. (c,f) *bK5Cre* oocyte showing two rings of DAPI-stained chromatin, believed to outline nucleoli, corresponding to the edges of strong anti-Cre fluorescence. (g,j) A single spot of DAPI staining (arrow) in a *bK5Cre* oocyte lacking anti-Cre signal suggested the plane of section had entered the edge of the nucleus where Cre levels may be low. (h,k) Oocyte from a *bK5Cre* ovary in the absence of the primary (1°) anti-Cre antibody. No fluorescence was detected from the nucleus, though the nucleolus is clearly visible under DAPI staining (arrow). (i-l) Non-transgenic ovary lacking Cre signal from two mature oocyte nuclei (arrows). Scale bar = 20μm.

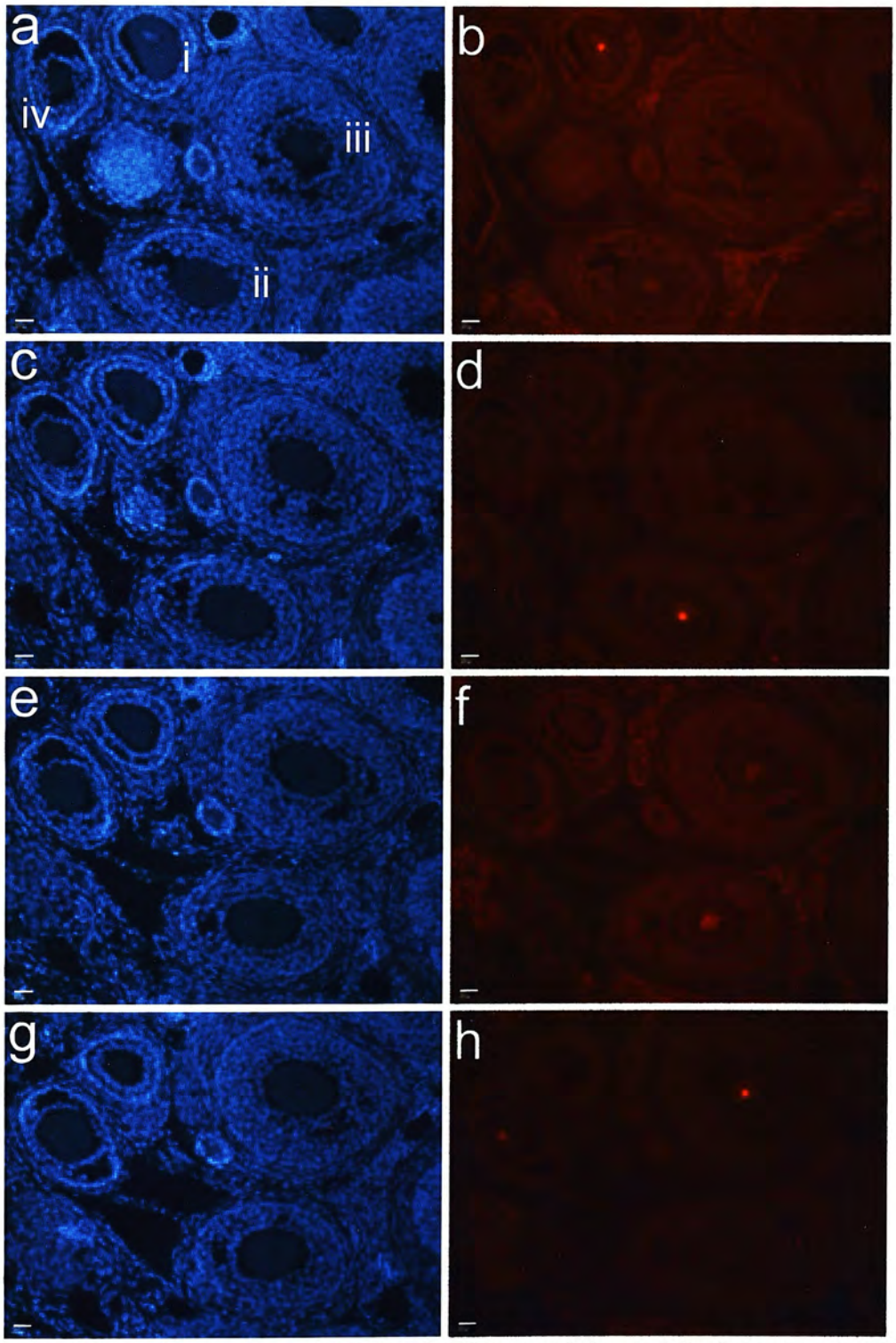


Figure 4.4 Immunofluorescence analysis of serial sections revealed a correlation between nuclear DAPI staining and anti-Cre signal.

Corresponding fields of view from four consecutive serial sections of a *bK5Cre* ovary, imaged for DAPI (a,c,e,g) and anti-Cre signal (b,d,f,h). Excluding the central primary follicle, four mature follicles are visible, numbered (i-iv). As the sections are examined in turn for anti-Cre fluorescence from (b) to (h), the nuclei of all four oocytes are revealed as positive, in the order that they have been numbered. Careful examination of the DAPI images reveals an association between the respective peaks of anti-Cre fluorescence and the appearance of DAPI-stained chromatin in the same oocyte. This indicated that Cre protein was localised to oocyte nuclei. Thus, DAPI signal was used as a control, indicating in which oocytes the plane of section had bisected the nucleus. Accordingly, scoring of Cre-positive nuclei was limited to these oocytes. Scale bar = 20μm.

that all fertilisable oocytes within *bK5Cre* ovaries produced Cre protein. To quantify this, Cre-positive oocytes were counted on 31 ovary sections from three hemizygous *bK5Cre* animals (Table 4.2): since the plane of section through an oocyte often did not pass through its nucleus, Cre-positive oocytes were scored per DAPI-positive oocyte (Figure 4.4). Excluding the immature oocytes of primary follicles, the nuclei of 86% of oocytes contained detectable amounts of Cre protein. In 58% of Cre-positive oocytes, the protein was noted to be concentrated in the nucleoli, with a reduced intensity of fluorescence in the surrounding nucleus. However it is likely that in many or all of the remaining 42%, the plane of section had simply not entered nucleoli. In most cases where DAPI-positive oocytes appeared to lack Cre protein, only a small number of DAPI staining foci were observed, suggesting the plane of section was through the edge of the nucleus where Cre levels may be low, and did not bisect nucleoli (Figure 4.3g,j). Of 48 DAPI-positive primary follicles examined, only one contained an apparently Cre-positive oocyte. However, classification of this follicle as primary, based on its relatively small size and single layer of surrounding follicle cells, may have been incorrect, as both follicle size and structure can be misrepresented by the plane of section. Alternatively, this oocyte may have been at the transition between the primary follicle and double-layered stages, during which Cre protein begins accumulating prior to the first meiotic division (see section 4.3.2).

As a control, 16 ovary sections from three non-transgenic mice were examined for fluorescent nuclei. Of 37 non-primary follicles observed, none exhibited nuclear fluorescence.

4.2.3 Cre protein mediates recombination of R26R *within oocytes*

To further confirm the presence of Cre protein in *bK5Cre* oocytes, and to demonstrate that the levels of Cre produced were sufficient to effect recombination of *R26R*, ovaries from *bK5Cre*$^{tg/wt}$ *R26R*$^{tg/wt}$ compound hemizygous mice were stained with xgal. In order to firstly demonstrate transcriptional activity of the *R26R* allele in oocytes, ovaries from mice produced by crossing a *R26R*$^{tg/wt}$ male with a *bK5Cre*$^{tg/wt}$ female, and thus being universally recombined at *R26R*, were xgal

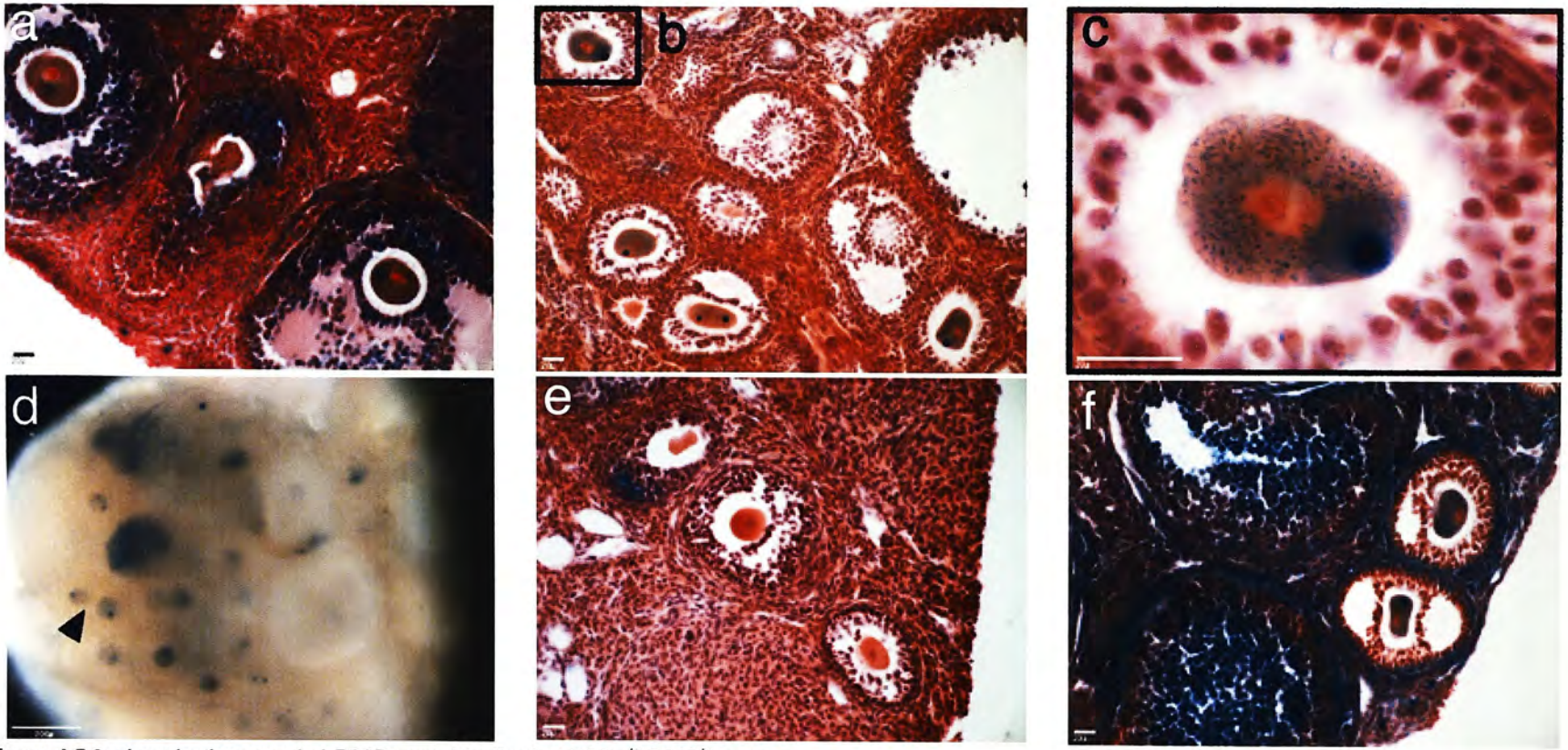


Figure 4.5 βgal production revealed *R26R* recombination in *bK5Cre*^{tg} *R26R*^{tg} oocytes.

Ovaries from *bK5Cre*^{tg} *R26R*^{tg} compound transgenic (a-d), and *R26R* (e) and *ROSA26* (f) single transgene-positive mice after xgal staining and sectioning (except wholemount, d). (a) *bK5Cre*$^{tg/wt}$ *R26R*$^{tg/wt}$ ovary from *bK5Cre* female x *R26R* male cross. Ubiquitous, though varyingly intense, staining resulting from Cre activity in maternal germline. Inconsistent staining was due partly to incomplete stain penetration, though variation in *R26R* transcription between different cell types must have been partially responsible, as a similar effect was observed in *ROSA26* ovaries (f). Stained oocytes in (a) confirmed transcriptional activity of the *R26R* allele in oocytes. Many oocytes contained discrete, intense foci of xgal precipitate. (b-d) *bK5Cre*$^{tg/wt}$ *R26R*$^{tg/wt}$ ovary from *bK5Cre* male x *R26R* female cross. (b) βgal production was limited to oocytes, except for patches of stained follicle cells, also observed in *bK5Cre*^{wt} *R26R*^{tg} ovaries (e), and assumed to represent non-specific staining. Oocyte-specific production of βgal confirmed Cre was present at levels sufficient to mediate activation of *R26R*, and that the genomic DNA was accessible to the nucleolar-accumulated Cre protein. (c) Higher magnification of boxed region in (b), revealing that spots of intense staining did not co-localise to the nucleus, also visible. (d) Stained oocytes with foci of βgal visible in wholemount (arrowhead). (e) In the absence of the *bK5Cre* transgene, no βgal was produced in *R26R* oocytes. (f) Intensely staining spots were also detected in oocytes of *ROSA26* ovaries. Scale bar = 20μm (a-c,e,f); 200μm (d).

stained and sectioned (Figure 4.5a). As expected, the ovary exhibited widespread, albeit non-uniform, production of βgal. While all oocytes demonstrated expression from the recombined *R26R* allele, they differed in ooplasmic staining intensity, and many contained a discrete spot of intense xgal staining.

In ovaries from compound hemizygous mice in which the *bK5Cre* transgene was inherited paternally, xgal staining was confined to oocytes (Figure 4.5b). Although general ooplasmic staining was less intense than in ovaries carrying a universally recombined *R26R* allele, many oocytes contained a similar spot of intense xgal staining (Figure 4.5c). This distinctive staining pattern, which was also visible in wholemount ovaries (Figure 4.5d), was not observed in single transgene-positive $bK5Cre^{wt}$ $R26R^{tg}$ mice (Figure 4.5e), or in non-transgenic mice. Staining foci did not appear to coincide with known sub-cellular structures (J. McNeilly, personal communication) and, as is shown clearly on Figure 4.5c, were distinct from the nucleus.

Figure 4.6 shows schematically how Cre activity in the female germline could result in the genotype-phenotype correlation presented in Table 4.1 for a cross between a $R26R^{tg/wt}$ male and a $bK5Cre^{tg/wt}$ female, assuming Mendelian inheritance.

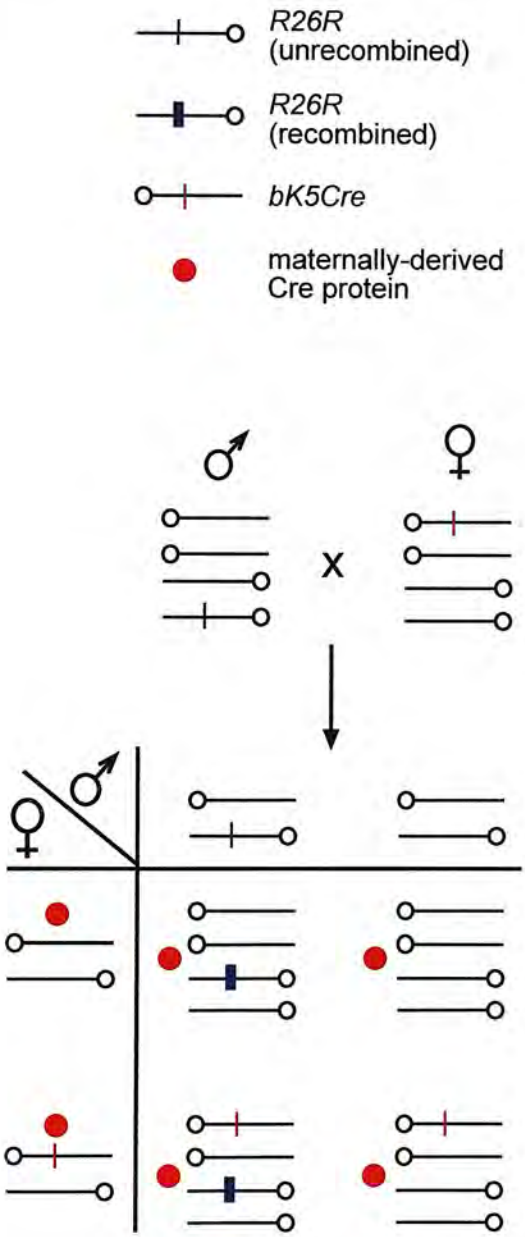


Figure 4.6 Cre activity in the female germline causes zygotic recombination of paternally-derived *R26R* alleles.

A cross between a $R26R^{tg/wt}$ male and a $bK5Cre^{tg/wt}$ female is represented. As shown, it is assumed that the *bK5Cre* allele had integrated onto a chromosome other than chromosome 6, where *R26R* is located (however for this cross, the genotype-phenotype correlation should be the same if the transgenes are located on homologous chromosomes). The potential male and female gametes are shown, as are the genotypes of the four theoretical progeny, representing four genotypic classes: $bK5Cre^{tg}R26R^{tg}$: $bK5Cre^{tg}R26R^{wt}$:$bK5Cre^{wt}$ $R26R^{tg}$:$bK5Cre^{wt}R26R^{wt}$ (non-transgenic), at a ratio of 1:1:1:1. The Cre protein within both female gametes, including that not transmitting the *bK5Cre* transgene, is thought to persist into the early zygote, resulting in the recombination-mediated activation of βgal expression in all zygotes inheriting the *R26R* allele (50%). In these animals, this permanent genetic change is inherited by all cells, resulting in ubiquitous βgal expression. In a cross where the female also carried the *R26R* allele, recombination would occur in all *R26R*-positive oocytes prior to ovulation (not shown).

4.3 Discussion

The *bK5Cre* line of mice used in chapter 3 to study cell lineage in the tracheal epithelium displayed alternate patterns of recombination of floxed alleles which was found to be determined by the mode of inheritance of the transgene: the paternally-inherited transgene caused basal epithelia-specific recombination typical of the bK5 promoter, while its unanticipated activity in the female germline resulted in universal excision of floxed sequences in the subsequent generation. Since accumulation of Cre protein in the maternal oocytes was apparently responsible for this phenomenon, further discussion of these results would be usefully informed by a basic review of oocyte development in the mouse.

4.3.1 Murine oogenesis in brief

Figure 4.7 shows a composite drawing of the mouse ovary including all the stages of oogenesis, from the primordial germ cells which appear at 8 days gestation, to the mature antral follicles which are first seen at approximately 17 days after birth. Primary oocytes are present in the adult ovary, and are initially distinguished by their single, loose layer of squamous follicle cells, and slightly enlarged nucleus compared to adjacent cells. As the oocyte develops, the nucleus continues to grow, and the single layer of follicle cells becomes cuboidal, producing the distinctive primary follicle. Over time, the follicle cells surrounding the primary oocyte become multi-layered, fluid-filled antral spaces begin to form, and single or multiple nucleoli appear as RNA production increases markedly. As the antra start to appear more distinct, recombination and segregation of alleles occurs as the oocyte undergoes the first meiotic division, resulting in the formation of the first polar body and leaving a secondary oocyte. The second meiotic division begins immediately in the maturing follicle with the oocyte, coated in cumulus cells, suspended in the single fused antrum. However in the swollen, mature follicle, the secondary oocyte nucleus remains arrested in metaphase, which does not resume until fertilisation. Finally, a corpus luteum is left following ovulation. Superovulation hormones are used to

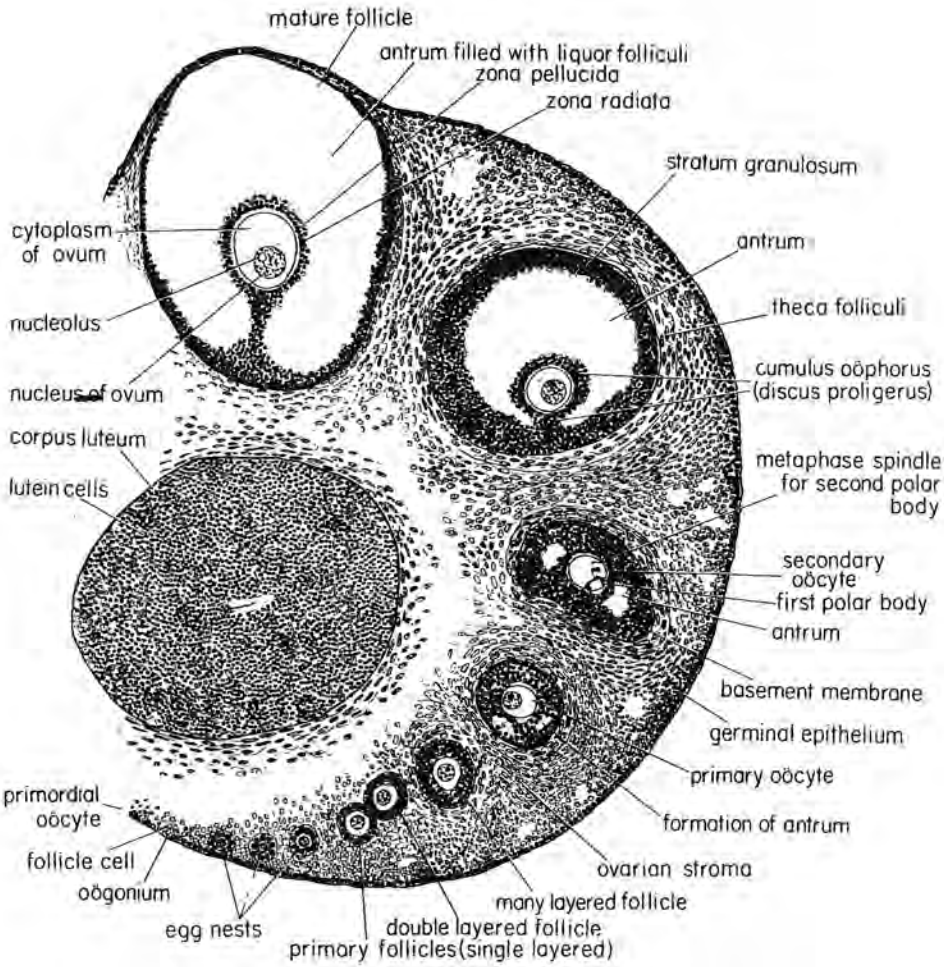


Figure 4.7 Development of the mouse oocyte (modified from Rugh, 1990).

Composite drawing showing all stages of oocyte development from primordial oocyte to mature antral follicle and corpus luteum.

See section 4.3.1.

stimulate a large number of follicles to synchronously mature, before triggering their release. In experiments described here, the ovaries were removed after administration of human chorionic gonadatrophin (hCG), but before the predicted onset of ovulation.

4.3.2 Accumulation of Cre protein precedes the first meiotic division

The observation that all *R26R*-positive offspring derived from *bK5Cre* females stained ubiquitously with xgal, regardless of whether they also inherited the *bK5Cre* transgene, suggested that Cre protein was being produced in all oocytes, rather than in merely a subset. Since the female parents were hemizygous for *bK5Cre*, only approximately half of fertilised oocytes would be expected to carry the transgene: in the remainder, *bK5Cre* would have been lost to the polar body during the first meiotic division. Thus, the presence of Cre protein in all fertilised oocytes would necessitate its synthesis prior to the first meiotic division, when all oocytes contain the *bK5Cre* transgene. This is supported by anti-Cre immunofluorescence data and xgal detection of *R26R* recombination, revealing the presence and activity of Cre in relatively small oocytes lacking antra (Figure 4.4b, Figure 4.5b). The alternative explanation, that Cre protein is synthesised in follicle cells and transported into the oocyte, is not supported, as no Cre-positive follicle cells were detected.

That the paternally-derived *R26R* allele is universally recombined by Cre protein produced in the female germline indicates that maternally-derived Cre must persist until male and female pronuclei fuse following fertilisation. Although this has not been directly demonstrated, the use of ovaries from superovulated mice revealed that Cre protein was clearly abundant in the oocytes of late stage follicles, and thus would likely be present at fertilisation. Xgal staining of two-cell embryos has previously been used to demonstrate zygotic recombination of paternally-derived alleles by Cre produced in oocytes from mice carrying a ubiquitously-expressed *Cre* transgene (Sakai and Miyazaki, 1997). This technique, or anti-Cre immunofluorescence, could be usefully employed on two-cell embryos resulting from crosses between *R26R*

males and *bK5Cre* females to prove that Cre protein is present and active in the early embryo.

4.3.3 Nucleolar accumulation of Cre protein does not prevent access to genomic loci

The immunofluoresence data presented reveals a marked sub-nuclear accumulation of Cre protein. In a proportion of oocytes, a ring of DAPI fluorescent staining encircled the region of high anti-Cre fluorescence. Oocytes in which the nucleoli are surrounded by a ring of Hoechst-positive chromatin ("surrounded nucleolus, SN"-type) have previously been described and distinguished from "non-surrounded nucleolus (NSN)" -type oocytes, in which the chromatin is organised into discrete spots which do not delineate the nucleolar border (Debey *et al.*, 1993). Both chromatin configurations are represented in Figure 4.3c,f and b,e, respectively: the spatial relationship between the SN-type oocyte's rings of DAPI fluorescence and the foci of intense anti-Cre signal indicated that the sub-nuclear compartment exhibiting elevated levels of Cre protein was the nucleolus.

The significance of this observation, and the means by which it occurs, is unclear. The Cre gene within the *bK5Cre* construct bears an exogenous nuclear localisation signal (NLS) (A. Ramirez, personal communication). NLS-Cre has been reported to mediate a 2-fold increase in recombinational excision of floxed sequences, compared to wild-type Cre (Gu *et al.*, 1993). However, a subsequent study found that the same NLS had no effect on nuclear accumulation of Cre protein, attributing this to the efficient import facilitated by the protein's endogenous nuclear targeting determinants (Le *et al.*, 1999). Targeting of Cre to nucleoli was not reported, though the nuclear activity of the somatic cells in which those experiments were conducted is likely to have been considerably different from the situation in the developing oocyte. Despite this, the concentration of Cre within the DNA-poor nucleolus is counter-intuitive. In the absence of an obvious role within, or means of targeting, the nucleolus, the apparent accumulation of Cre there may be an artefact resulting from high expression of the prokaryotic protein.

Despite its apparent accumulation within nucleoli, Cre protein clearly had access to the *R26R* locus of genomic DNA, as demonstrated by xgal staining of oocytes from compound transgenic mice carrying a paternally-derived *bK5Cre* transgene. The staining pattern was characterised by faint, generalised ooplasmic staining and single or, less commonly, multiple intense, discrete foci of staining within the ooplasm. These staining spots could be seen on wholemount ovaries, and sectioning revealed they did not coincide with obvious sub-cellular structures. Similar spots, albeit against a background of stronger ooplasmic staining, were also observed in homozygous *ROSA26* ovaries, and in ovaries containing universally recombined *R26R* alleles from mice with *bK5Cre* mothers, but not in single transgene-positive mice carrying entirely unrecombined *R26R* alleles. Similar foci of xgal staining have been reported in *ROSA26* preimplantation embryos, and were not found to be associated with distinct cytoplasmic organelles by electron microscopy (Neganova *et al.*, 2001). The authors speculated that the sub-cellular localisation of xgal precipitate might reflect an involvement of the *ROSA26* locus in RNA or protein distribution. Alternatively, the staining spots may represent βgal aggregates produced as the oocyte attempts to remove the foreign protein, as they have been observed in other transgenic lines expressing βgal in oocytes (G. Wei, personal communication).

4.3.4 Does expression from the bK5 promoter in oocytes reflect endogenous K5 production?

It was initially imagined that the activity of the *bK5Cre* transgene in oocytes was ectopic, the likely result of position effect on the transgene in the particular line used. However, analysis of another *bK5Cre* line has revealed universal recombination of floxed alleles following *bK5Cre* transmission through the female germline, albeit with a lower penetrance than the line used in these studies (A. Ramirez, unpublished results). In the other *bK5Cre* line, 21% of offspring inheriting only the reporter gene showed complete recombination, with no recombination detected in the remainder, while 17% inheriting both transgenes showed complete recombination, the remaining 83% exhibiting tissue-specific recombination (A. Ramirez, unpublished results). This suggests that expression of bK5 promoter-driven transgenes in oocytes might be

representative of endogenous K5 transcription. Although K5 activity is generally regarded as being restricted to epithelia (Byrne and Fuchs, 1993), and contrary to reports describing the absence of intermediate filaments from mammalian embryos prior to the 8-cell stage (Chisholm and Houliston, 1987), several keratins including K5 have been detected in mouse oocytes and early embryos (Lehtonen, 1987; Gallicano *et al.*, 1994). It appears that K5, and its apparent oocyte partner K15, are maternally synthesised during oogenesis and stored in cytoskeletal sheets until required at the morula or blastula stage when they would function by binding to desmosomes (Gallicano, 2001). If expression of the *bK5Cre* transgene in oocytes is a consequence of endogenous K5 promoter activity, then the apparent absence of K14 protein in oocytes (Gallicano, 2001) may explain why constitutive recombination of floxed alleles in the female germline has not been reported for mice in which Cre has been introduced to the genomic K14 locus (Huelsken *et al.*, 2001). Mice expressing Cre from the K15 promoter have not been described.

Despite this, other studies using *Cre* transgenes driven by K5-derived promoters (both bovine and human) make no mention of constitutive recombination of floxed alleles in progeny born of K5-*Cre* mothers: although only male mice carrying Cre downstream of a 14kb human K5 promoter were used in matings by Tarutani and colleagues (Tarutani *et al.*, 1997) and Gao and co-workers (Gao *et al.*, 2002), while the sex of the K5-*Cre* transgene-bearing mice used by Suzuki and colleagues (Suzuki *et al.*, 2003) was not specified, no evidence of expression within the female germline was apparently observed (J. Takeda, personal communication). In another study, the absence of such a phenomenon in experiments using the same 5.2kb bovine K5 promoter as used here is not informative, as no attempt was made to activate the inducible version of Cre in transgenic mothers (Zhou *et al.*, 2002). The detection of Cre activity in the oocytes of two independent lines of *bK5Cre* mice suggests that genomic position effect on the randomly integrated transgenes is unlikely to be entirely responsible. That such an expression profile is representative of the endogenous K5 gene is supported by reports of K5 protein within oocytes (Gallicano *et al.*, 1994). In light of this, it is surprising that the more extensive promoter region (14kb of human sequence, compared to 5.2kb of bovine sequence used here), which

one would expect to drive a more faithful pattern of expression, apparently fails to emulate the female germline-specific phenomenon described herein. This effect may therefore be a characteristic feature of the *bK5Cre* construct used in these studies.

Although several transgenic mice have been generated which express Cre in their oocytes, using either strong ubiquitous promoters (Schwenk *et al.*, 1995; Lakso *et al.*, 1996; Sakai and Miyazaki, 1997; Lallemand *et al.*, 1998) or regulatory elements from known oocyte-expressed genes (Lewandoski *et al.*, 1997; de Vries *et al.*, 2000) it would appear that only one other line exists in which paternal transmission results in a tissue-specific recombination pattern and maternal transmission produces constitutive recombination in offspring, independent of inheritance of the *Cre* transgene. However the unanticipated expression of Cre within the oocytes of this line, in which the transgene is driven by elements of the rat prostate-specific *probasin* gene promoter, only achieved constitutive recombination in 47% of offspring (Wu *et al.*, 2001). Thus the *bK5Cre* line described here, which effects universal recombination in 100% of offspring when *bK5Cre* is present in the female parent, offers both basal epithelia-specific and ubiquitous recombination of floxed alleles, simply by changing the parent carrying the *Cre* transgene.

Chapter 5 A functional, *in vivo* model for gene correction

5.1 Introduction

5.1.1 Gene correction

"Gene correction", also known as "targeted gene repair", describes the ability of a range of small molecule reagents to permanently and site-specifically effect small-scale sequence alterations at genomic loci, and is a conceptually more elegant approach to the established complementation approach of gene therapy. The direct correction of a genetic defect at its native chromosomal location offers several advantages over the addition of an exogenous functional copy of the mutant gene. The repaired gene would remain under the regulation of its endogenous control elements, including 5', 3' and intronic sequences typically absent from cDNA-based gene therapy vectors, restoring the precise physiological expression characteristics of the normal gene. Since gene correction reagents act site-specifically and introduce no exogenous DNA into the genome, there is little risk of insertional mutagenesis as has been observed with integrating viral vectors (Hacein-Bey-Abina *et al.*, 2003). Also, the permanent nature of the sequence alteration should circumvent the need for repeated administration, which in the case of adenoviral vectors can lead to severe inflammatory responses (Thomas *et al.*, 2003). Unlike complementation strategies, gene correction approaches are applicable to dominant-negative, gain-of-function mutations, since repair should simultaneously reduce the amount of deleterious protein and increase the proportion of functional protein. If the permanent genetic "fix" promised by gene repair techniques could be delivered to the stem cells of the target tissue, and then by propagation to all the differentiated cells, a single-dose gene therapy could be realised.

Gene correction has been achieved through the use of several types of small molecule reagents. The most prolific of these are modified oligonucleotides pioneered by Kmiec and colleagues, which have evolved from RNA/DNA oligonucleotides (RDOs or "chimeraplasts") to single-stranded oligonucleotides (SSOs) (Liu *et al.*, 2003). Other systems include small DNA fragments (SDFs)

(Gruenert *et al.*, 2003), triplex-forming oligonucleotides (TFOs) (Seidman and Glazer, 2003), and a recently recognised property of adeno-associated virus (AAV) (Inoue *et al.*, 2001). The two systems which were under investigation in the laboratory when this project was undertaken, RDOs and SDFs, will be briefly introduced.

5.1.2 RNA/DNA oligonucleotides

The original RDO design was a single-stranded, self-complementary oligonucleotide of approximately 70-80 bases, which adopted a double-hairpin conformation through intra-molecular base-pairing (Kmiec, 2003) (Figure 5.1a). The sequence of the paired strands was complementary to the target sequence, except for the position corresponding to the mutation, to facilitate homologous pairing. One strand contained RNA bases, modified to resist nuclease digestion, which were believed to enhance pairing of the molecule with the target sequence (Kotani *et al.*, 1996). Additional stability was provided by 4-thymine loops at either end of the molecule, and one of the phosphodiester bonds was left unligated to allow the molecule to become intertwined with the target helix. Although the mechanism of action has not been determined, the molecule was postulated to act by mimicking a mismatched base pair at the site of the mutation, inducing the recruitment of cellular DNA repair proteins to resolve the mismatch and thus alter the target sequence (Kmiec, 1999) (Figure 5.1b).

The range of experimental systems, genetic targets, molecule designs, and associated experimental parameters used in different studies have meant that meaningful comparisons between the correction frequencies achieved have been difficult to make. However, there has clearly been great variation in reported efficiencies. While many groups have persistently failed to achieve any RDO-mediated correction (van der Steege *et al.*, 2001), there have been some notable successes. Kren and colleagues reported a single nucleotide conversion efficiency of 43% in cultured hepatocytes (Kren *et al.*, 1997), and subsequently achieved conversion of 40% of *factor IX* loci in rat liver *in vivo* (Kren *et al.*, 1998). The single base pair mutation

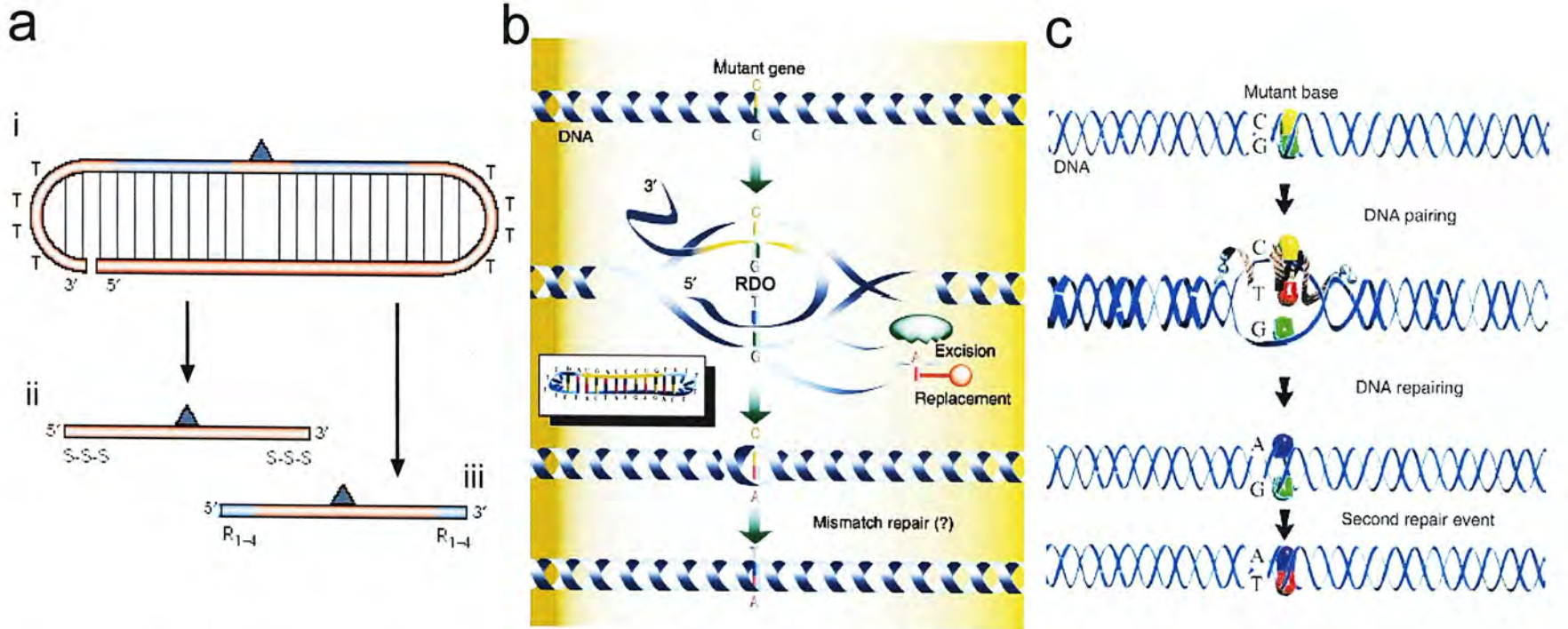


Figure 5.1 RNA/DNA oligonucleotides (RDOs) and single-stranded oligonucleotides (SSOs).

(a) The evolution of oligonucleotide gene repair vectors (modified from Liu *et al*, 2003). (i) Prototype RDO design. The self-complementary, single-stranded molecule adopts a double-hairpin conformation through intra-molecular base-pairing (vertical lines). 4-thymine (T) loops confer additional stability. DNA bases are shown in red, modified RNA bases in grey. The grey triangle signifies the base pair which is mismatched with the target sequence. (ii) All-DNA SSO capped with phosphorothioate linkages (S). (iii) SSO capped with protective modified RNA bases (R). (b) Possible mode of action of RDO (modified from Kmiec, 2003). The molecule (inset) anneals to the target site, with the RNA section (yellow) hybridising with 100% complementarity to the one strand, and the DNA section annealing with the opposite strand. The T/G mismatch is recognised by cellular DNA repair systems. In this example, the G residue is excised and replaced by an A residue to complement the T base of the RDO. After dissociation of the RDO, the C/A mismatch is resolved to a T/A (correction, as shown), or a C/G (no correction, not shown) base pair. (c) Possible mode of action of SSO (modified from Kmiec, 2003). The SSO anneals to the complementary target strand which will be dictated by the polarity (sense or antisense) of the SSO, while the opposite strand becomes displaced. In this example, a C/T mismatch forms and is repaired to a A/T, leaving a A/G mismatch after dissociation of the SSO. Correction results from resolution of this mismatch to A/T.

introduced into the *factor IX* gene was associated with prolonged blood coagulation. Also in rat liver, insertion of a single nucleotide was achieved at a frequency of approximately 20% at the UDP-glucuronosyltransferase gene, resulting in reduced serum bilirubin levels (Kren *et al.*, 1999). In a mouse over expressing human apoE2, the dysfunctional version of the apoE protein responsible for premature artherosclerosis, approximately 25% of *APOE2* was converted to the *APOE3* form (Tagalakis *et al.*, 2001), and such a change was shown to have effects at the protein level *in vitro*. Other reports of functional *in vivo* correction, albeit demonstrating more modest efficiencies, include the detection of wild-type dystrophin fibres in the *mdx* mouse (Rando *et al.*, 2000) and *GRMD* dog (Bartlett *et al.*, 2000) models of muscular dystrophy, and the production of melanin pigment in albino melanocytes (Alexeev *et al.*, 2000).

Several explanations have been forwarded for the variability in observed correction rates, including the differential recombination and repair activities of different cell types (Santana *et al.*, 1998) and the efficiency with which targeting molecules are delivered to the nucleus (Thorpe *et al.*, 2002b). Another issue has been the quality and purity of RDO syntheses (Liu *et al.*, 2003; Manzano *et al.*, 2003). The latter factor, combined with the high cost of producing complex RDO molecules, led to the experimental deconstruction of RDO functionality. It was discovered that single-stranded oligonucleotides (SSOs) comprising 25-90 nucleotides of DNA protectively capped with modified RNA nucleotides (Igoucheva *et al.*, 2001) or phosphorothioate linkages (Gamper *et al.*, 2000) could effect sequence-specific modifications with similar frequencies to RDOs (Figure 5.1a(ii,iii),c). Moreover, in experiments demonstrating both episomal and chromosomal correction of a defective *lacZ* gene by two different lengths of sense and antisense SSOs, no correction was effected by a corresponding RDO (Nickerson and Colledge, 2003).

Recent work has concentrated on the use of SSOs, though reported efficiencies of correction are modest compared to the previously described *in vivo* work using RDOs. Correction of the α-D-glucosidase gene was achieved in cultured cells at frequencies of between 1.5-5%, reflected by increases of approximately 0.5-4% in

enzyme activity, with sense SSOs showing higher corrective activities than antisense versions (Lu *et al.*, 2003). The latter observation contrasts with reports of higher conversion rates being mediated by antisense than sense SSOs (Igoucheva *et al.*, 2001; Nickerson and Colledge, 2003). While the binding of antisense SSOs to the nontranscribed strand would not be destabilised by transcription, as might occur in the case of sense SSOs, the conflicting results might reflect an influence of DNA replication, with the polarity of the process favouring sense or antisense SSOs at different loci (Liu *et al.*, 2003). Attempts to identify cellular factors which influence correction frequencies are now being conducted in yeast. For example overexpression of *RAD51*, which codes for a strand transferase protein known to be involved in recombination reactions, has been shown to enhance both episomal and chromosomal SSO-mediated gene repair (Liu *et al.*, 2002). This is consistent with the results of Thorpe and colleagues using mammalian cell lines, who observed significant levels of RDO-mediated correction only in a cell line overexpressing *RAD51* (Thorpe *et al.*, 2002b). Further dissection of the endogenous mechanisms involved in mediating the activities of SSOs should enhance the utility of this technique.

5.1.3 small DNA fragments

SDFs are typically 400-800bp single- or double-stranded PCR products derived from the wild-type version of the mutated target sequence, and have been shown to mediate gene correction through a poorly-understood process termed small fragment homologous replacement (SFHR). It has been postulated that the absence of sequence non-homologous to the target (except for the mismatched base(s)) is responsible for the levels of recombination observed (Goncz *et al.*, 1998), in contrast to the infrequent homologous replacement achieved using much longer classical targeting vectors which contain significant regions of non-homology, such as selectable markers (Deng and Capecchi, 1992).

Unlike RDOs, the action of which has so far been limited to single base conversions and insertions, the ability of SDFs to mediate the insertion or deletion of multiple

bases has been demonstrated: both insertion and deletion of 3bp corresponding to the ΔF508 mutation in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene has been achieved in cultured epithelial cells (Kunzelmann *et al.*, 1996; Goncz *et al.*, 1998). These studies also revealed that an independent single base pair conversion could be introduced concomitantly with the primary insertion or deletion. In addition to DNA- and RNA-based methods of detection, correction within 1-10% of transfected cells was confirmed functionally using electrophysiological measurements (Kunzelmann *et al.*, 1996).

In primary myoblast cultures from the *mdx* mouse model of Duchenne muscular dystrophy, multiple applications of wild-type SDF resulted in conversion at 15-20% of loci as revealed by PCR-restriction fragment length polymorphism (RFLP), but no corresponding transcript or protein was detected (Kapsa *et al.*, 2001). However, SDF-mediated correction of episomally-maintained mutant *eGFP* (Thorpe *et al.*, 2002b) or *lacZ* (Nickerson and Colledge, 2003) constructs has been functionally demonstrated in a range of cell lines. Reduced levels of SDF-mediated correction were apparent at chromosomally-integrated versions of the same mutant reporters (Thorpe *et al.*, 2002a; Nickerson and Colledge, 2003), the *eGFP* study implicating poor nuclear delivery using polyethylenimine transfection (Thorpe *et al.*, 2002a). The influence of the delivery method on the level of correction achieved has been revealed by other studies, and emphasises the importance of factors other than the molecule design in affecting the experimental outcome (Colosimo *et al.*, 2001; Sangiuolo *et al.*, 2002).

To date, only two studies have demonstrated SFHR *in vivo*. Intratracheal delivery of mutagenic SDFs to wild-type mice was shown to result in the expression of ΔF508 *Cftr* mRNA within airway and lung cells (Goncz *et al.*, 2001), though levels of conversion were not quantified. Injection of SDFs into dystrophic *mdx* muscle achieved low levels of correction which lasted for three weeks, though no dystrophin fibres were evident (Kapsa *et al.*, 2001). The elucidation of the molecular basis of SFHR, for example through the use of cell lines deficient in recombination,

replication or repair proteins, should permit the further development of SDF technology.

5.1.4 Evaluating gene correction in vivo*: limitations of current systems*

The development of gene correction modalities has taken place throughout a range of systems, both *in vitro* and *in vivo*, and has employed a range of experimental readouts, both sequence-based and functional. Cultured cells undoubtedly provide a technically simpler means of delivering, evaluating the performance of, and ultimately gaining an insight into the actions of gene correction molecules. Efficient sequence-specific modification in cultured cells would be of significant value in a range of applications, for example *ex vivo* gene therapy of cultured progenitor cells. However, an *ex vivo* approach may not be applicable to all genetic diseases, therefore gene correction must aim to effect repair in somatic tissue. Animal models are thus essential for addressing issues specific to *in vivo* gene correction, such as mode of delivery, dosage, and route of administration, in addition to the general question of molecule design.

Many studies have relied on DNA-based means of detecting and measuring sequence alterations, including sequencing of the target site, allele-specific PCR, and diagnostic restriction digests. Indeed, some reports have been criticised for employing PCR-based methods of detection found to be potentially artefactual (Zhang *et al.*, 1998). Far more elegant and convincing are functional assays of gene correction, of which there are several examples, where the induced sequence alteration results in an experimentally detectable change in protein activity. These have often targeted specific genes in animal models of disease, such as the dystrophin gene in the *mdx* mouse (Rando, 2002; Bertoni and Rando, 2002) and the *GRMD* dog (Bartlett *et al.*, 2000), both of which are deficient in dystrophin protein. In these studies, dystrophin-positive fibres were detected by immunohistochemistry after injection of RDOs into muscle. While a demonstration that correction achieved at the genomic level translates into the production of protein is convincing evidence

that the gene correction phenomenon is real, the most compelling data are those revealing physiological consequences of gene repair. In the Gunn rat model of Crigler-Najjar syndrome type I, correction of the causative single base deletion within the *UGT1A1* gene was, in addition to a battery of DNA-based methods, demonstrated by a persistent drop in serum bilirubin levels and the appearance of glucuronidated bilirubin species (Kren *et al.*, 1999).

While undoubtedly elegant, and demonstrative of the potential therapeutic applicability of gene correction strategies, one limitation of these functional and physiologic rescue-type approaches is that the efficacy of the gene correction system is assayed only in the target tissue. In the previous example, the evidence of functional correction applied only to the liver, where *UGT1A1* is expressed. Likewise, prolonged coagulation indicative of reduced coagulant activity after targeted mutagenesis of the *factor IX* gene in rats is informative only about the activity of RDOs in hepatocytes. Another important study, in which correction of the tyrosinase gene by topical and intradermal application of RDOs to the skin of albino mice was confirmed by the production of pigmented hairs (Alexeev *et al.*, 2000), is indicative solely of their ability to mediate gene repair in melanogenic cells.

In addition to the anatomically restricted nature of data derived from these types of functional assays, many of the techniques involved, including immunohistochemistry, western blotting, and blood sampling for high performance liquid chromatography (HPLC) analysis or enzymatic assays, are relatively laborious and time-consuming, in common with some of the DNA and RNA based methods of detection, such as DNA sequencing, RT-PCR, southern blotting, colony hybridisation and RFLP analysis. As a result, these methods are not ideal for large empirical studies evaluating the effects of a range of qualitative and quantitative parameters.

5.1.5 Aims

An *in vivo* model providing a functional, yet technically simple readout of correction frequency, facilitating the further development and optimisation of gene correction molecules in a physiological context, would be a major asset to gene correction researchers. This chapter describes the generation and analysis of a transgenic mouse carrying a mutated, correctable reporter gene, with the intention that it should permit the rapid histochemical detection and quantification of correction events occurring upon administration of gene correction molecules. While the respiratory tissues are of most relevance to CF research, the main interest in the laboratory in which this work was undertaken, there was no advantage in restricting the utility of such a model to particular anatomical regions. Thus, a ubiquitously active promoter was chosen to drive expression of the transgene, such that the model would be suitable for the investigation of factors influencing gene correction efficiency in all tissues.

Results

5.2 The transgenic mutated reporter line, *R26X*

The inactivated reporter transgene *R26X*, and the cloning strategy used to create it, is depicted in Figure 5.2. Briefly, the construct consists of a mutagenised bacterial *lacZ* gene positioned downstream of the R26 promoter fragment, which has been reported to drive ubiquitous expression (Kisseberth *et al.*, 1999). An internal ribosome entry site (IRES) 3' of the *lacZ* sequence is juxtaposed with the gene encoding the enhanced green fluorescent protein (*eGFP*). While βgal activity should be dependent on correction of the mutation within *lacZ*, eGFP should be produced constitutively, and should serve to indicate expression of the transgene. It was recognised that the designed non-functionality of the truncated βgal produced by *R26X* would hinder efforts to validate the construct. In order to better evaluate promoter activity in cell culture and predict the behaviour of the *R26X* construct in transgenic mice, a non-mutated version of the transgene, *R26Z*, was created. While only the production of *R26X* is described, both transgenes were constructed identically and in parallel, except for the initial mutagenesis step, which was omitted in the production of *R26Z*. The DNA manipulations involved in the generation of these constructs are described in detail in appendix A2. The intermediates in vector construction are named in Figure 5.2 to allow referencing.

5.2.1 Site-directed mutagenesis of lacZ

The commercial mammalian reporter plasmid pCMVβ (Clontech) containing the *E. coli lacZ* gene downstream of the human cytomegalovirus (CMV) promoter was selected as the parental plasmid for the generation of *R26X*. A PCR-based site-directed mutagenesis protocol (see chapter 7, section 7.4) was used to introduce a C to T susbstitution at position 1126 of *lacZ*, designed to result in an arginine (CGA) to stop (opal, UGA) mutation at amino acid position 376 (R376X) (Figure 5.3a). Interestingly, colonies which were found to carry the mutation appeared faintly blue

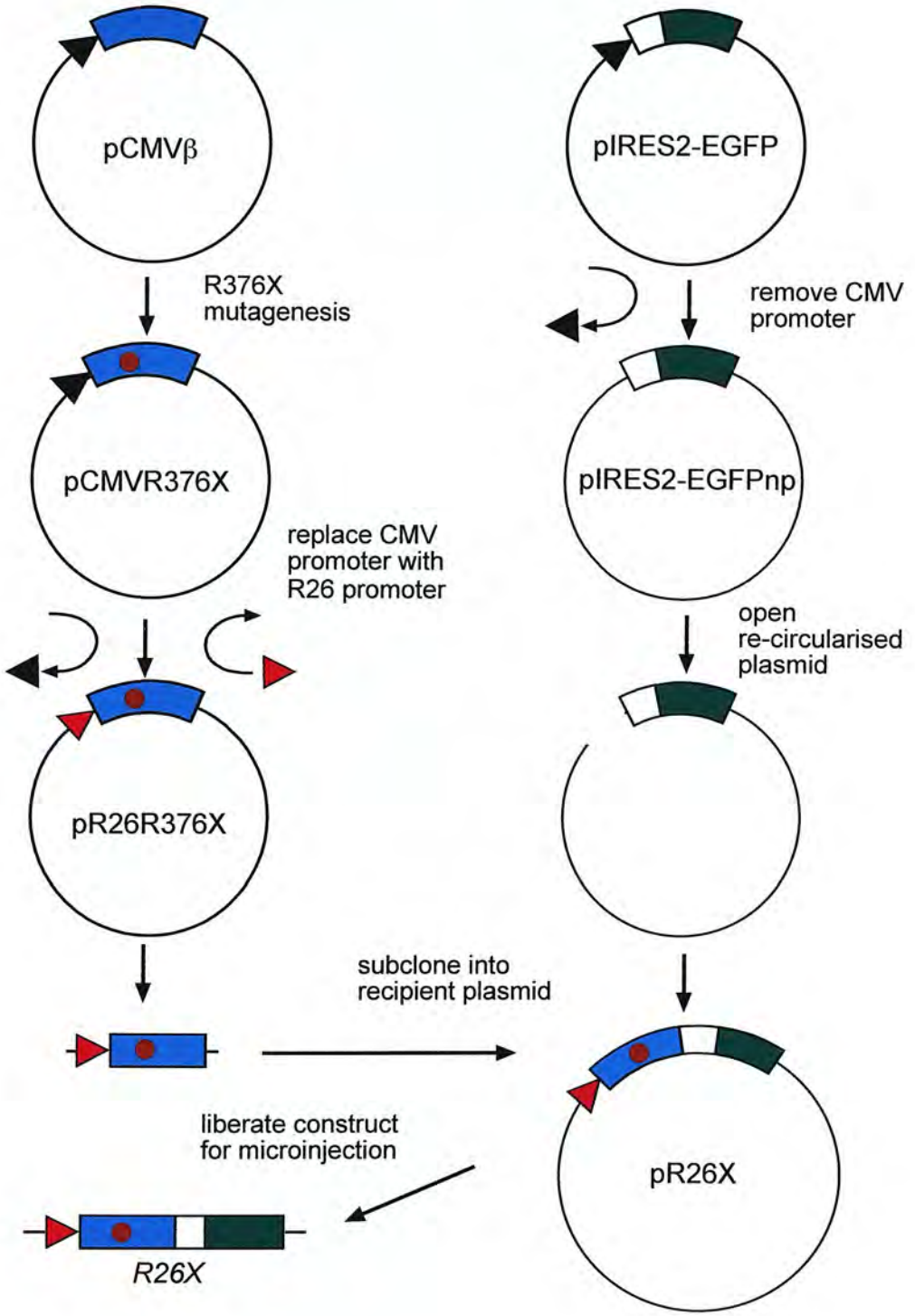


Figure 5.2 Cloning strategy to generate *R26X* transgenic mice.

Construct elements are represented by the following: CMV promoter (black triangle); R26 promoter (pink triangle); *lacZ* gene (blue rectangle); IRES-*eGFP* (white rectangle-green rectangle); R376X nonsense mutation (red spot). Intermediates in vector construction are named to allow reference in main text. *R26Z* mice were created identically, except for omission of the mutagenesis step.

on xgal transformation plates, while every white colony analysed had undergone deletions in the promoter region. This suggested that the mammalian CMV promoter is active, and that the induced stop codon undergoes partial suppression, in bacteria. The presence of the mutation was checked by RFLP, as the R376X mutation produces a novel endonuclease product owing to the loss of a TfiI recognition site (Figure 5.3a,c). The mutation was confirmed by DNA sequencing (Figure 5.3b "R376X").

After transfection into COS-7 cells, the mutated pCMVR376X plasmid (see Figure 5.2) showed no βgal activity by xgal staining (demonstrated subsequently, Figure 5.4), confirming the functional consequences of the mutation. In an attempt to induce readthrough of the premature stop codon, thereby validating the integrity of the mutagenised sequence, COS-7 cells transfected with pCMVR376X were incubated with varying concentrations of G-418. This aminoglycoside antibiotic has previously been shown to suppress premature stop codons in *CFTR*, with full-length protein produced to 35% of wild-type levels (Howard *et al.*, 1996). Translational readthrough was assessed by visual inspection of xgal stained cells, though no βgal activity was detected, the apparent condition of the cells suggesting a toxic effect of G-418 (data not shown).

To confirm the absence of additional *lacZ* mutations in pCMVR376X, the gene was sequenced. Eight sense and eight antisense primers were used to span the approximately 3.1kb gene, with no section receiving less than 2× coverage, and a contig was assembled from overlapping reads. Regions of ambiguous sequence were repeated, and the entire sequence was compared to Clontech's published pCMVβ sequence. Other than the designed mutation, two other mismatches were detected. However, the mismatches were again found after sequencing of the relevant regions of non-mutated pCMVβ, indicating either errors in the Clontech data, or accumulated base changes since the plasmid was obtained (Figure 5.3b). Both sequence alterations are predicted to be silent, and they do not ablate βgal enzyme activity as shown by the production of functional βgal from pCMVβ in transfections (see Figure 5.4).

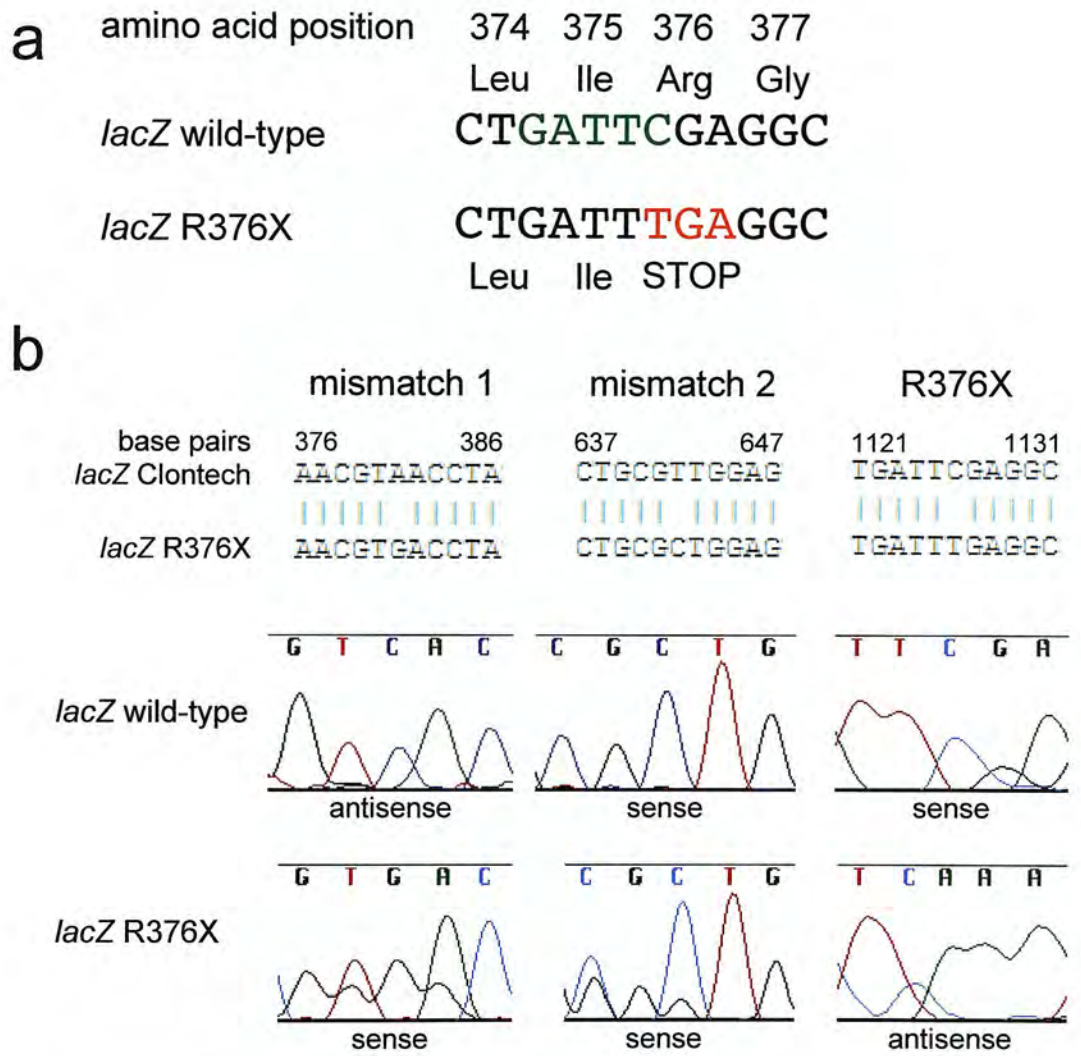


Figure 5.3 Site-directed mutagenesis of *lacZ*.

(a) C to T substitution at the first position of codon 376 converts an arginine to a stop with a concomitant elimination of a TfiI restriction site (GAWTC). (b) Aligning the mutagenised *lacZ* (*lacZ* R376X) contig with the *lacZ* sequence from Clontech's pCMVβ data (*lacZ* Clontech) revealed three mismatches, including the induced R376X mutation. Sequencing the corresponding regions of wild-type *lacZ* in pCMVβ revealed the two unanticipated mismatched base pairs to also be present in pCMVβ, indicating either errors in the Clontech data, or accumulated base changes since the plasmid was acquired. In the latter case, neither change is predicted to alter the amino acid sequence. The mismatched bases are shown at the centre of the chromatograms. (c) Diagnostic TfiI digest of pCMVβ and pCMVR376X. The elimination of a TfiI site created a novel 351bp fragment from the fusion of 219bp and 132bp fragments. The loss of the 132bp fragment is visible as a reduction in intensity of the upper of two bands (the lower band being 112bp) which is a doublet with a 140bp fragment present after digestion of both plasmids.

5.2.2 in vitro *validation of the* R26X *construct*

Because of doubts over the performance of the CMV promoter *in vivo* (Norris *et al.*, 1998), particularly in the lung (Schmidt *et al.*, 1990), it was replaced with a 0.8kb fragment derived from the genomic *ROSA26* locus (R26 promoter, provided by P. Soriano) (see appendix, section A2.1). The R26 promoter, in combination with two different reporter genes, had previously been shown to recapitulate the ubiquitous βgal expression pattern of *ROSA26* (Friedrich and Soriano, 1991) in mice and rats generated by conventional transgenesis (Kisseberth *et al.*, 1999). A fragment containing the R26 promoter and mutated *lacZ* gene were then subcloned upstream of an IRES-*eGFP* sequence to produce pR26X (see appendix, section A2.2).

In order to confirm the activity of the R26 promoter, and further assess the affect of the induced stop mutation, pR26X and its wild-type counterpart pR26Z were transfected into four cell lines representing different tissues (Figure 5.4): COS-7, SV40-transformed African green monkey kidney fibroblast; A-549, human lung epithelial; Huh7, human hepatoma; HT-1080, human acetabulum epithelial. pCMVβ, included as a positive control, demonstrated levels of βgal visually detectable without magnification, in each of the cell lines. Xgal staining was less intense following transfection with pR26Z, and was detectable by eye only in COS-7 and Huh7 cells. However, magnification revealed the presence of βgal-positive cells in all lines, indicating the competence of the R26 promoter to drive expression in cells from a range of tissues.

In each of the cell lines, pR26Z consistently produced lower levels of βgal than pCMVβ, despite cell line-specific variation of βgal expression from the latter plasmid. This suggested that differential transfection efficiencies between cell lines may have been responsible for the apparent low level of expression driven by the R26 promoter, particularly in A-549 and HT-1080 cells. However, it is clear that the CMV promoter drives stronger, if not more widespread, expression than the R26 promoter. The extent of this difference is under-represented in COS-7 cells, which express the SV40 T antigen. pR26Z contains the SV40 origin of replication from the

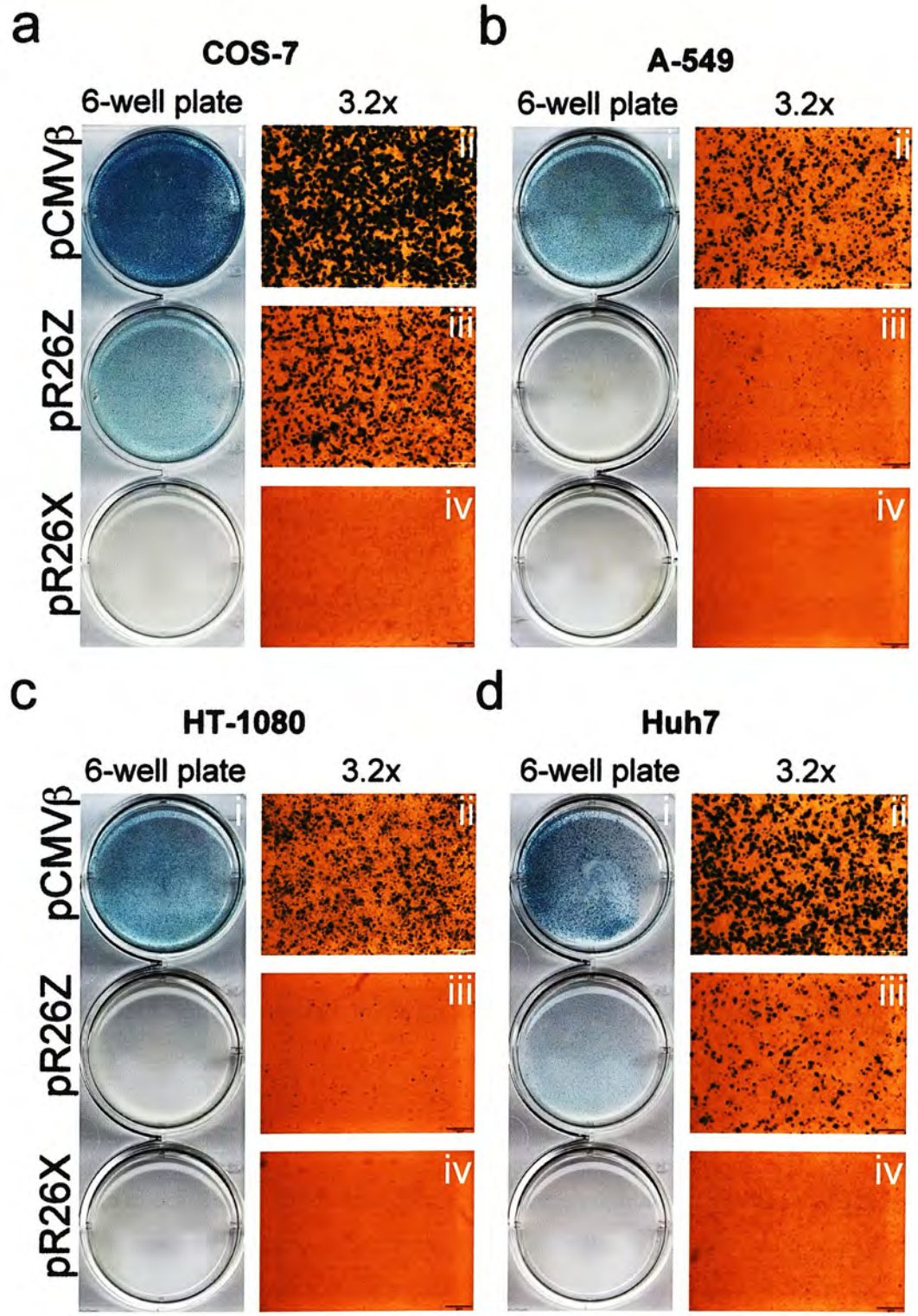


Figure 5.4 *in vitro* validation of pR26X.

Four different cell lines (a-d) were transfected with pR26X, pR26Z and pCMVβ. The cells were then fixed and xgal stained, and analysed by eye (a-d(i)) and microscopically (a-d (ii-iv)). The CMV promoter expressed βgal at levels detectable without microscopy in all four lines, while cells stained after transfection with pR26Z were visible unmagnified only in COS-7 and Huh7 cells. Magnification revealed βgal production from pR26Z in all lines, though at consistently lower levels than from pCMVβ. No stained cells were detectable after transfection with pR26X. Scale bar = 400μm.

pIRES-eGFP vector, absent in pCMVβ, resulting in amplification of transfected molecules. Thus the observed level of βgal production from pR26Z is exaggerated with respect to pCMVβ in COS-7 cells.

No βgal activity was detected after microscopic examination of pR26X-transfected cells. pR26X was assumed to have entered cells to an extent comparable with pR26Z, suggesting that the induced stop mutation had rendered βgal non-functional. That both plasmids had been successfully transfected was supported by fluorescence microscopic examination of duplicate transfected cells, which revealed a low level of eGFP expression from both pR26Z- and pR26X-transfected cells, but not from pCMVβ-transfected cells (data not shown).

5.2.3 Microinjection of constructs and identification of transgenic founders

Transgenic mice were produced as described in chapter 7 section 7.6. Constructs pR26X and pR26Z were linearised prior to microinjection with BspHI, which cut the vector backbone twice, and microinjected into F1 fertilised oocytes. The microinjection sessions are summarised in Table 5.1. Genomic DNA prepared from ear-clip biopsies was used as template for genotyping. A transgene-specific PCR, using primers based on *lacZ* sequence, was combined with a *myogenin* (*myoG*)-specific DNA quality control PCR, as shown for the *K5A1* transgene, chapter 2, Figure 2.3.

Additional PCRs were conducted on founders identified by the *lacZ*-specific reaction in order to check the 5' and 3' integrity of the transgenes, and to ensure that genomic DNA preparations were not contaminated with the plasmids from which the transgenes were derived (Figure 5.5). The 3' integrity-checking and plasmid contamination control PCRs were as shown for the *K5A1* transgene (chapter 2, Figure 2.3). The corresponding 5' PCRs, using a primer within the R26 promoter sequence, are shown in Figure 5.5a,b. These PCRs demonstrated that both of the *R26X* or *R26Z* reporter construct termini were present in all founders, suggesting

Construct, Session date	Total no. embryos	Total no. injected	Total no. cultured	Total no. transferred	Total no. born	Total no. transgenic
R26Z						
05/12/01	187	102	87	73	21 (1 died)	5
R26X						
17/10/01	290	192	146	128	10 (1 died)	2
21/11/01	204	151	130	100	13 (1 died)	4
28/11/01	207	133	120	99	13	2
03/12/01	345	191	145	119	28	5
10/12/01	345	228	148	111	27	7
Total	1391	895	689	557	89	20

Table 5.1 Summary of microinjection sessions for *R26X* and *R26Z* constructs.

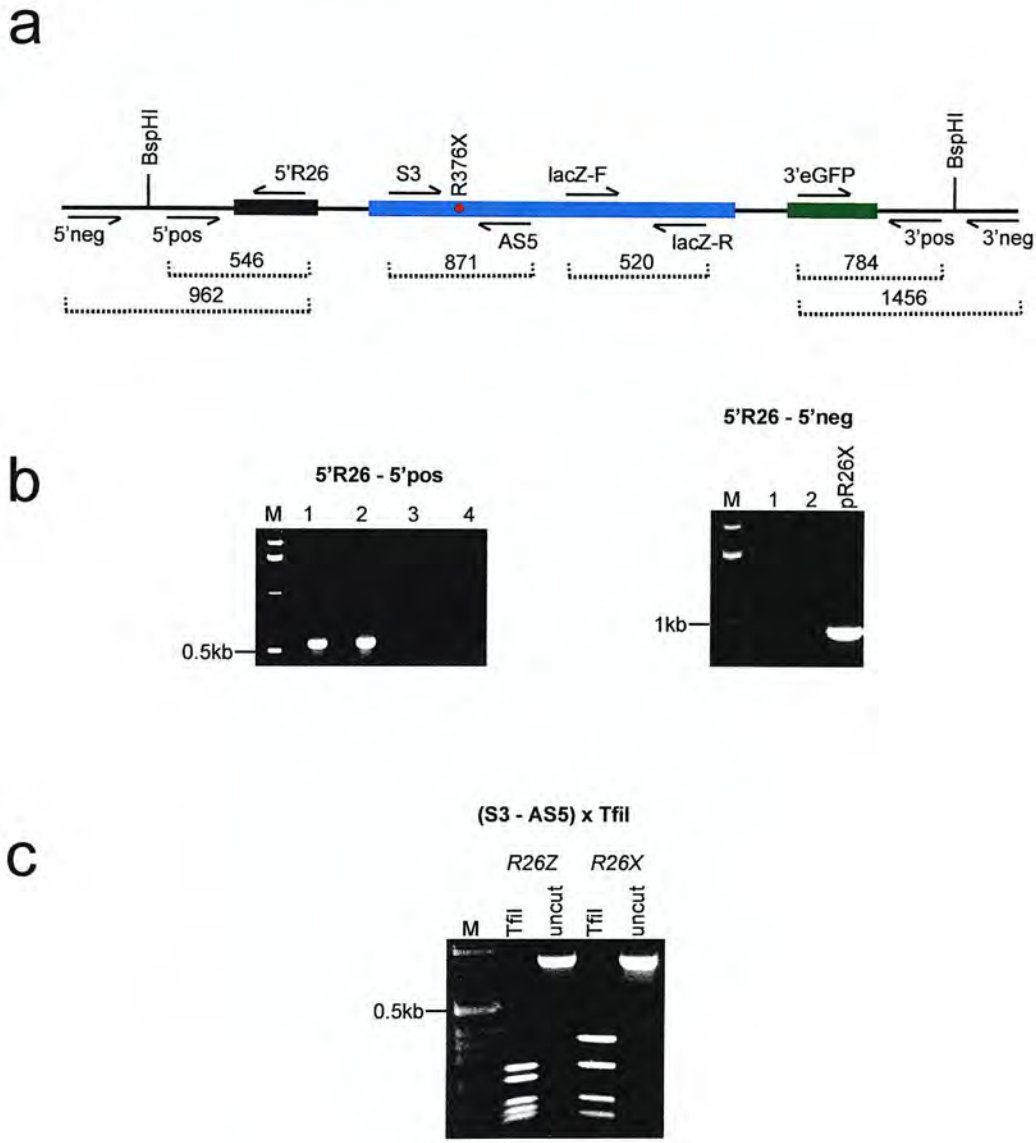


Figure 5.5 PCR analysis of integrated *R26X* transgenes.

(a) Schematic showing the position and orientation of PCR primers on the *R26X* construct, including the sizes of the amplification products. R26 promoter, *lacZ* and *eGFP* represented by black, blue and green bars, respectively. The BspHI sites used to linearise the construct for microinjection, are shown. *lacZ* genotyping (lacZ-F - lacZ-R) and 3' integrity-checking and plasmid contamination PCRs are as shown for the *K5A1* construct in chapter 2, Figure 2.3. (b) 5' integrity-checking PCR for two positive (1,2) and two negative (3,4) *R26X* founders (left) and 5' plasmid contamination PCRs showing that both founders 1 and 2 are genuinely positive, and not contaminated with pR26X. (c) PCR-RFLP of product amplified from genomic DNA using *lacZ* sequencing primers (S3 - AS5), subsequently digested with TfiI. A novel *R26X* 351bp fragment is generated from the loss of a TfiI site between 219bp and 132bp *R26Z* fragments. M = 1kb ladder. Map not to scale, fragment sizes in bps.

integrity of the transgenes, and that there was no plasmid contamination of the genomic samples.

Although both *R26X* and *R26Z* mice were genotyped using the same *lacZ*-specific PCR, if necessary they could be distinguished by PCR-RFLP. The product of another *lacZ*-specific PCR, employing primers situated on either side of the R376X mutation, when digested with TfiI gave a pattern of fragments indicative of whether the mutation was present in the amplified sequence (Figure 5.5a,c).

5.3 Analysis of *R26Z* lines

Before attempting to analyse expression from *R26X* lines carrying the mutated reporter, the five *R26Z* lines were examined for expression of the functional *lacZ* gene in E12.5 embryos, to allow comparison with ubiquitously expressed human alkaline phosphatase (hPAP) and eGFP in embryos reported by Kisseberth and colleagues (Kisseberth *et al.*, 1999). Timed matings were set up between founder males, or male progeny produced from crosses between female founders and C57BL/6 males, and non-transgenic CD1 females. The first embryonic litters obtained from each line were genotyped from genomic DNA prepared from yolk sacs or from amputated limb or tail tissue. Both methods allowed embryos to be identified with respect to their genotype, to confirm that phenotypically-detected transgene expression was associated with inheritance of the transgene. In lines producing no genotypically transgene-positive embryos in the first litter, a second litter was genotyped. One of the five *R26Z* lines, line *R26Z500*, failed to demonstrate transmission of the transgene.

5.3.1 Expression of eGFP

Prior to fixing the embryos for xgal staining, they were examined by fluorescence microscopy for expression of eGFP. Litters from two of the lines, *R26Z100,300*, contained embryos which were visibly fluorescent compared to littermates. *R26Z100*

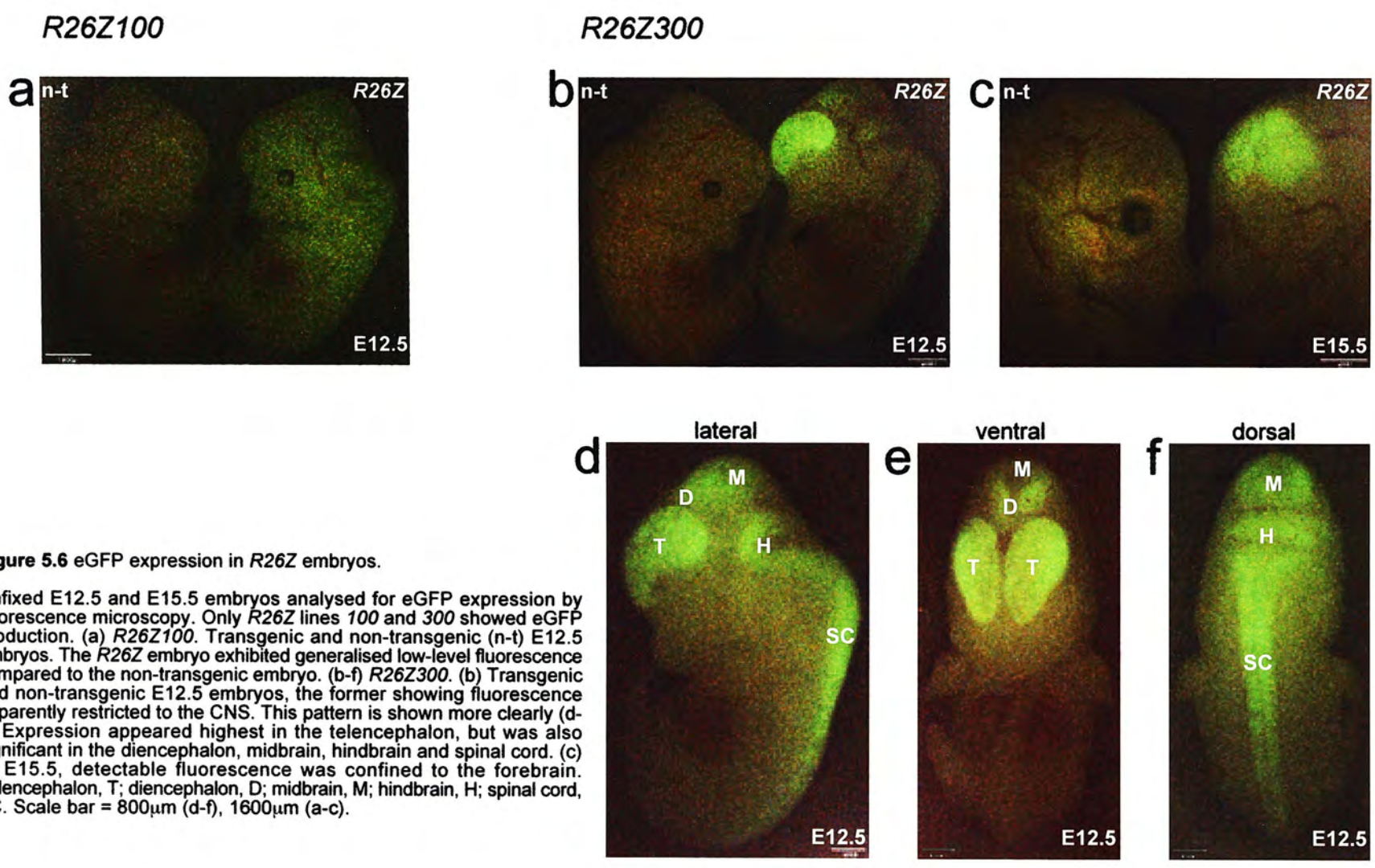


Figure 5.6 eGFP expression in *R26Z* embryos.

Unfixed E12.5 and E15.5 embryos analysed for eGFP expression by fluorescence microscopy. Only *R26Z* lines *100* and *300* showed eGFP production. (a) *R26Z100*. Transgenic and non-transgenic (n-t) E12.5 embryos. The *R26Z* embryo exhibited generalised low-level fluorescence compared to the non-transgenic embryo. (b-f) *R26Z300*. (b) Transgenic and non-transgenic E12.5 embryos, the former showing fluorescence apparently restricted to the CNS. This pattern is shown more clearly (d-f). Expression appeared highest in the telencephalon, but was also significant in the diencephalon, midbrain, hindbrain and spinal cord. (c) At E15.5, detectable fluorescence was confined to the forebrain. Telencephalon, T; diencephalon, D; midbrain, M; hindbrain, H; spinal cord, SC. Scale bar = 800μm (d-f), 1600μm (a-c).

(Figure 5.6a) exhibited generalised, low-level fluorescence, with no structures appearing particularly prominent. Contrastingly, transgenic embryos from *R26Z300* (Figure 5.6b-f) demonstrated marked fluorescence in a pattern consistent with the central nervous system (CNS): eGFP expression appeared most intense in the telencephalon, but was also visible in the diencephalon, the midbrain, and the hindbrain and spinal cord. eGFP levels outside these regions did not appear appreciably higher than in non-fluorescent embryos. At E15.5, visually detectable fluorescence was confined to the forebrain region in *R26Z300*, and was absent from line *R26Z100* (data not shown).

5.3.2 Expression of βgal

Fixing and xgal staining of whole embryos revealed eGFP expression to accurately indicate those embryos expressing βgal, while PCR genotyping confirmed eGFP and βgal expression was limited to embryos carrying the transgene (transgene-positive embryos from lines *R26Z200* and *R26Z400* did not show eGFP or βgal activity). Although eGFP expression predicted those embryos staining after xgal treatment, the enzymatic βgal assay provided a more informative expression pattern than the fluorescent signal. At E12.5 (Figure 5.7a-c), line *R26Z100* showed widespread staining, with regions of apparently elevated βgal expression, not evident from eGFP visualisation, including the developing ear, muscle groups in the limbs and, most notably, the forebrain and midbrain, though not in the hindbrain or spinal cord. βgal expression was low or absent in the forelimbs and hindlimbs, and the ventral trunk. At E15.5 (Figure 5.7f-i), staining was more homogenous but was excluded from the distal limbs, the tail, and the snout. An E15.5 embryo bisected along the sagittal midline prior to xgal treatment revealed widespread internal staining, though no βgal expression was apparent in the developing lungs or liver (Figure 5.7i).

In *R26Z300* at E12.5 (Figure 5.8a-e), βgal expression reflected the strong CNS-specific pattern of eGFP production, with a low level of xgal staining elsewhere, appearing slightly more intense in the follicles of the vibrissae. A sagittal section indicated that expression may be higher in the telencephalon and spinal cord than in

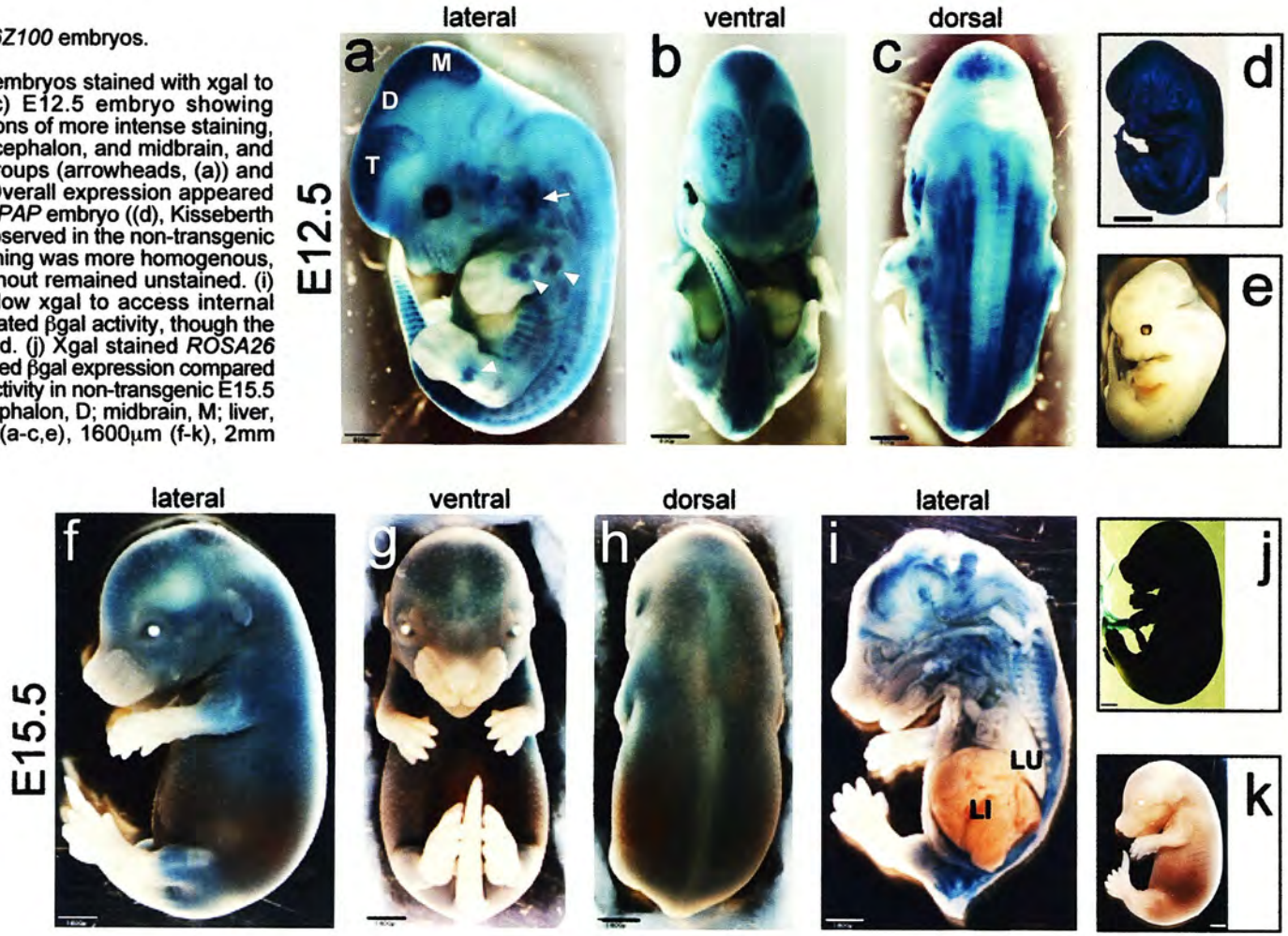


Figure 5.7 βgal expression in *R26Z100* embryos.

Fixed E12.5 (a-c) and E15.5 (f-i) embryos stained with xgal to visualise βgal expression. (a-c) E12.5 embryo showing widespread βgal activity, with regions of more intense staining, including the telencephalon, diencephalon, and midbrain, and what appear to be limb muscle groups (arrowheads, (a)) and the developing ear (arrow, (a)). Overall expression appeared significantly less than in the R26-*hPAP* embryo ((d), Kisseberth *et al,* 1999). (e) No staining was observed in the non-transgenic E12.5 embryo. (f-h) At E15.5, staining was more homogenous, though the distal limbs, tail and snout remained unstained. (i) Embryo bisected sagittally to allow xgal to access internal tissues. Most structures demonstrated βgal activity, though the liver and lung remained unstained. (j) Xgal stained *ROSA26* embryo showing significantly elevated βgal expression compared to *R26Z100*. (k) Absence of βgal activity in non-transgenic E15.5 embryo. Telencephalon, T; diencephalon, D; midbrain, M; liver, LI; lung, LU. Scale bar = 800μm (a-c,e), 1600μm (f-k), 2mm (d).

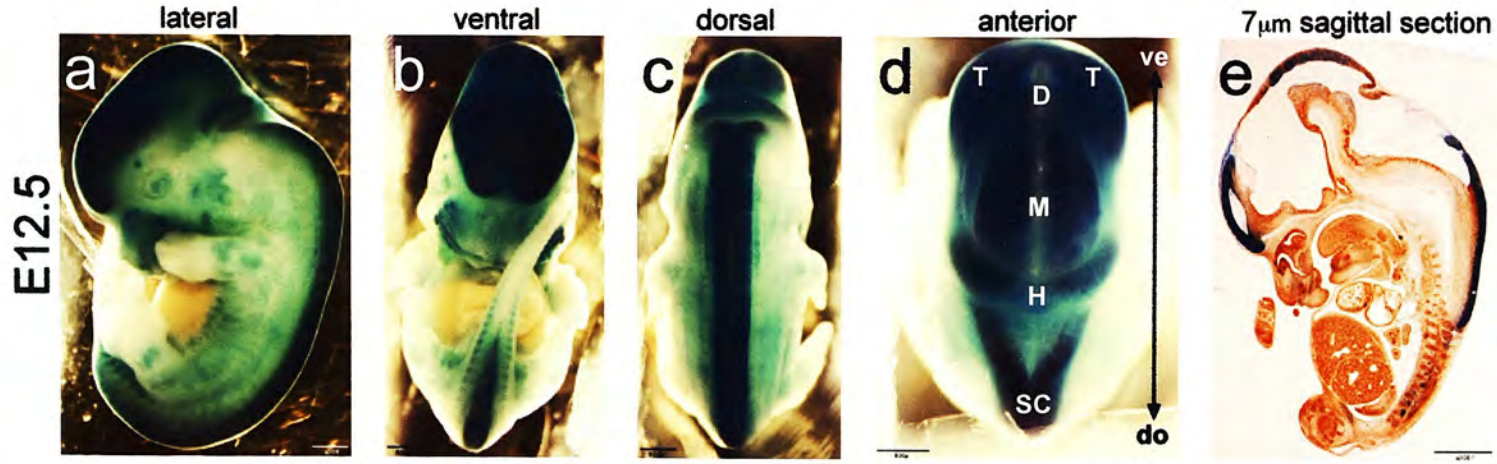


Figure 5.8 βgal expression in *R26Z300* embryos.

Fixed E12.5 (a-e) and E15.5 (f-i) embryos stained with xgal to visualise βgal expression. (a-c) E12.5 embryo showing intense βgal activity in the CNS, reflecting the eGFP pattern (Figure 5.6). Staining was markedly reduced in non-CNS tissues and was absent from the distal limbs and ventral trunk. (d) Anterior view showing βgal activity in the brain and spinal cord. (e) Sagittal section, counterstained with eosin, showing that expression in the midbrain and diencephalon appeared less intense than in the telencephalon and spinal cord. Although widespread, low-level staining was generally lost in thin section, close examination revealed patches of expression distributed throughout internal structures, including the heart, nasal passages, branchial arches, and dorsal root ganglia. Reduced staining on section compared to wholemount may also have been due to incomplete xgal penetration into intact stained embryos. (f-h) Staining at E15.5 appeared more uniform, though was excluded from the maternally-derived umbelicus (arrowhead, (g)). (i) Anterior view revealing that βgal activity remained elevated in the brain. Telencephalon, T; diencephalon, D; midbrain, M; hindbrain, H; spinal cord, SC; dorsal, do; ventral, ve. Scale bar = 800μm (a-d), 1600μm (e-i).

the diencephalon and midbrain region. At E15.5 (Figure 5.8f-i), expression appeared ubiquitous, but not as strong as in the *ROSA26* embryo (Figure 5.7j). Unlike in *R26Z100*, expression of βgal extended into the distal limbs, the tail, and the snout. βgal production remained detectably elevated in the telencephalic vesicles (the future cerebral hemispheres), and the midbrain.

5.3.3 Adult tissues

Attempts to discriminate between transgenic and non-transgenic live adult littermates using a hand-held UV lamp, as has been described for mice ubiquitously expressing eGFP (Okabe *et al.*, 1997), was not successful. This was due the apparent lack of expression in adult skin, as discovered subsequently.

A range of tissues including brain, heart, kidney, liver, spleen, trachea, lung, skin and gonads were taken from four adult hemizygous *R26Z100* and *R26Z300* mice (including both sexes) and examined for eGFP fluorescence. No tissues demonstrated expression of *eGFP*. For xgal staining, larger tissues were cut into two or more pieces to increase penetration of the fixative and the stain into the tissue. Corresponding tissues were taken from homozygous *ROSA26* mice and non-transgenic littermates as positive and negative staining controls, respectively. Except for the brain, none of the adult tissues from either line demonstrated βgal activity above background levels, as determined by comparison with non-transgenic tissues. In the brain (Figure 5.9), βgal expression in *R26Z100* appeared restricted to the cerebrum and pineal gland, derivatives of the telencephalon and diencephalon forebrain structures, respectively, both of which were observed to express the transgene at E12.5 (Figure 5.7a). While the embryonic midbrain also stained with xgal, the corresponding adult structures are mainly internal. *R26Z300* adult brain additionally showed βgal activity in the cerebellum and medulla, albeit at lower levels than the forebrain-derived structures, reflecting embryonic expression in the hindbrain (Figure 5.8d).

Figure 5.9 βgal expression in *R26Z* adult brain.

Dorsal view of xgal stained brains from *R26Z100,300* and non-transgenic mice. (a) Annotated non-transgenic brain showing no staining. (b) *R26Z100* brain exhibiting βgal expression in the cerebrum and pineal gland, both derived from the embryonic forebrain (telencephalon and diencephalon, respectively). (c) *R26Z300* brain was more intensely, but less uniformly, stained than that of *R26Z100*, with expression strongest in the cerebrum and pineal gland, but also evident in the cerebellum and medulla (both hindbrain-derived). Cerebrum, CR, pineal gland, P; cerebellum, CL; medulla, M; olfactory lobes, OL. Scale bar = 3200μm.

5.4 Analysis of *R26X* lines

5.4.1 Expression of eGFP

Of the 20 *R26X* founders generated, four male founders and male offspring from three of the female founders, were crossed with non-transgenic CD1 females to obtain E12.5 embryos. eGFP fluorescence was detected in only one of the seven lines examined (line *R26X100*, Figure 5.10a,b). While fluorescing embryos appeared generally, but not greatly, brighter than non-fluorescing littermates, significantly elevated eGFP expression appeared confined to the midbrain. Genotyping revealed the presence of transgenic embryos in litters from all lines, indicating that all seven founders were able to transmit the transgene, with eGFP fluorescence correlating with inheritance of *R26X* in *R26X100*. eGFP was undetectable at E15.5 (data not shown).

5.4.2 Expression of βgal

Xgal staining of fixed E12.5 embryos revealed all lines to be negative for βgal production. While this may have reflected a lack of transgene activity in some of the lines tested, expression of eGFP in *R26X100* indicated transcription was occurring and suggested that the absence of xgal staining was a result of the premature stop codon in *lacZ* (Figure 5.10c,d).

5.4.3 Adult tissues

Tissues corresponding to those taken from *R26Z* mice (section 5.3.3) were taken from three adult *R26X100* mice and were examined for eGFP and βgal expression. None of the tissues exhibited detectable eGFP fluorescence or xgal staining. The xgal stained brain, which exhibited eGFP embryonically, is shown (Figure 5.10g).

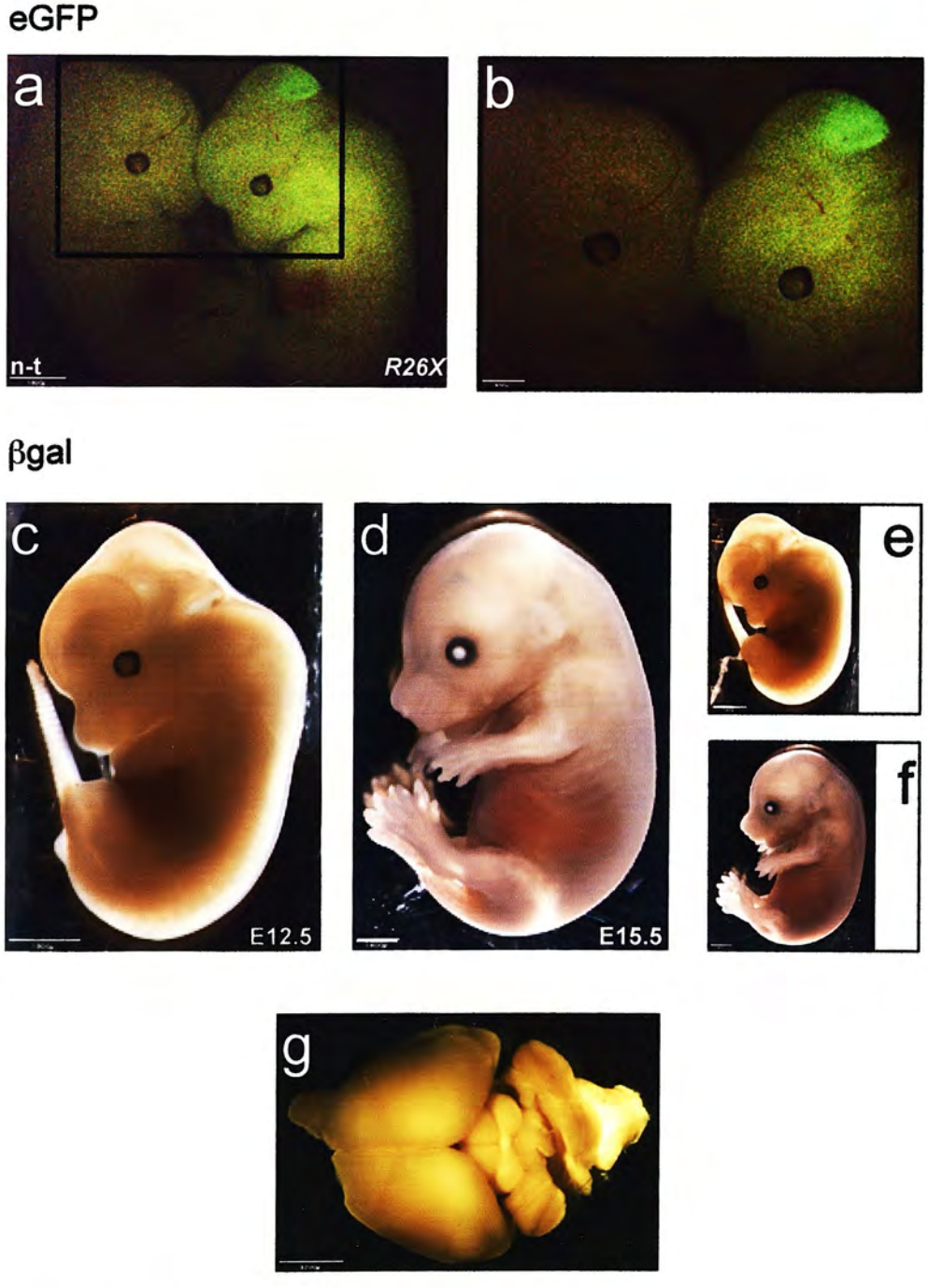


Figure 5.10 eGFP and βgal expression in *R26X100*.

(a) Unfixed transgenic and non-transgenic (n-t) E12.5 embryos analysed for eGFP expression by fluorescence microscopy. The *R26X* embryo showed a generalised low-level increase in fluorescence compared to the non-transgenic embryo, with significantly elevated eGFP confined to the dorsal midbrain. The boxed region in (a) is magnified in (b). (c,d) Fixed, xgal stained E12.5 (c) and E15.5 (d) transgenic embryos showed a complete absence of staining, and appeared indistinguishable from non-transgenic E12.5 (e) and E15.5 (f) embryos. (g) Transgenic adult brain showing no staining after xgal treatment. Scale bar = 800μm (b), 1600μm (a,c-f), 3200μm (g).

In the absence of visible eGFP expression in the adult tissues of the only line in which embryonic eGFP fluorescence was observed (*R26X100*), RT-PCR was used to detect transcription of *lacZ* in selected tissues from a range of *R26X* lines (Figure 5.11(i)). Lungs and trachea were chosen for their relevance as targets for CF therapy, while liver was included as hepatocytes have been reported to be efficiently corrected with RDOs (Kren *et al.*, 1997; Kren *et al.*, 1998). *lacZ*-specific RT-PCRs with RNA from 12 *R26X* lines, including *ROSA26* RNA as a positive control, revealed that while half of the lines did not show transgene expression in any of the tissues analysed, three lines exhibited detectable expression in all three tissues (*R26X100,500,1700*). The remaining three lines demonstrated *lacZ* transcription either in the lung (*R26X400*), or both the lung and trachea (*R26X1300,1800*). B. Stevenson provided some technical assistance with RNA extraction and RT-PCR.

To confirm that the amplification product was specific to *lacZ*, a gel containing the RT-PCR products was blotted onto a nylon membrane. This blot was probed with a radiolabelled internal *lacZ* oligonucleotide (Figure 5.11(ii)). As a demonstration that the RT-PCR product amplified from the *R26X* lines carried a deviation from the wild-type *lacZ* sequence, the blot was stripped and re-probed with a radiolabelled internal *lacZ* oligonucleotide containing the R376X mutation. The differences in intensity between the *R26X* signals and the *ROSA26* signal increased markedly relative to the initial probing, indicating differential strengths of hybridisation to the mutated probe.

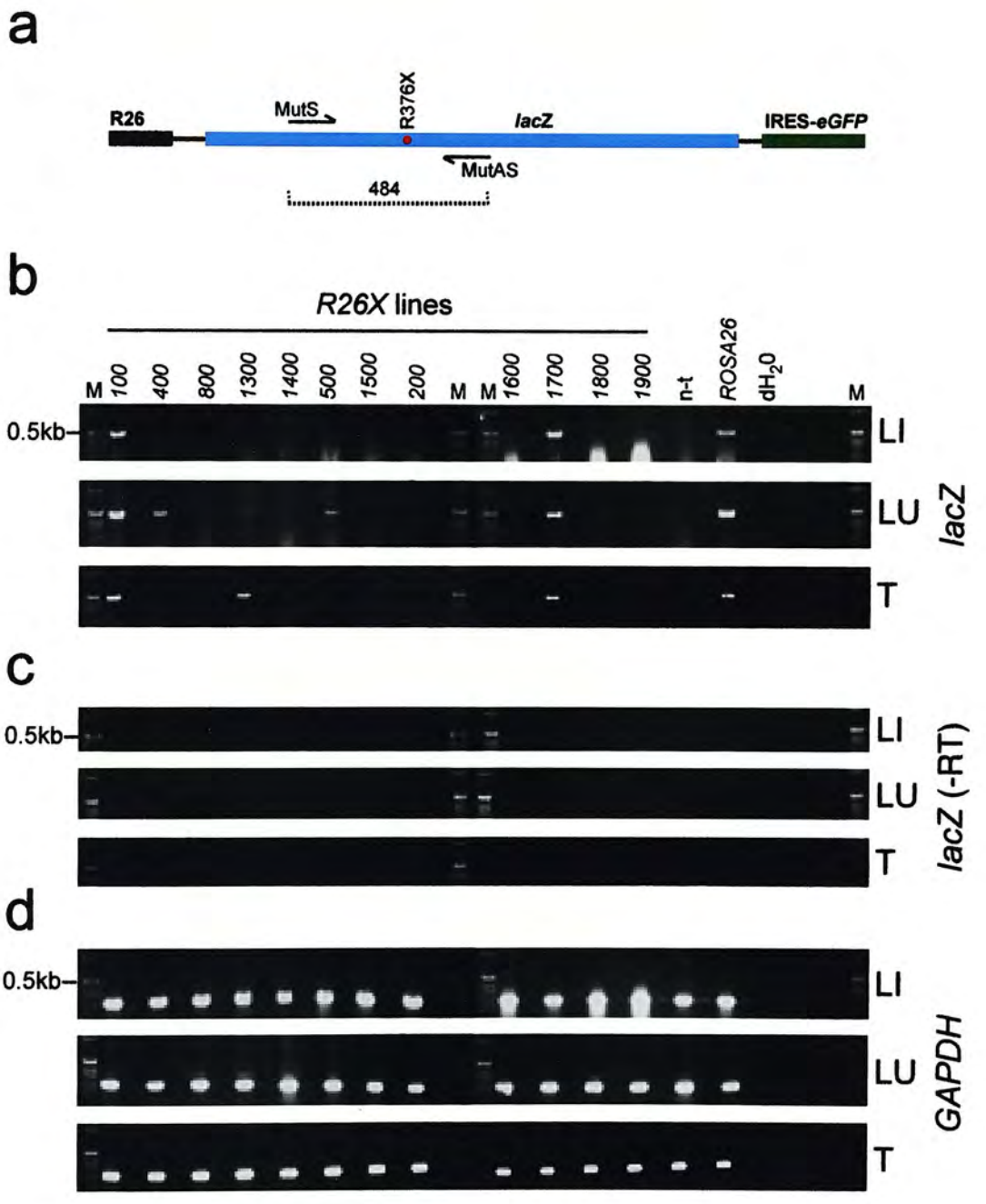


Figure 5.11(i) RT-PCR analysis of *lacZ* expression in 12 *R26X* lines.

(a) Schematic showing the position of PCR primers on the *R26X* transgene with respect to the R376X mutation. (b) RNA from liver (LI), lung (LU) and trachea (T) of 12 *R26X* lines plus non-transgenic (n-t) and *ROSA26* controls was subjected to 30 cycles of RT-PCR with *lacZ*-specific primers (MutS - MutAS) amplifying a 484bp product spanning the R376X mutation site. Six lines showed *lacZ* expression in at least one tissue (*R26X100,400,1300,500,1700,1800*), with *R26X100,500,1700* exhibiting expression in all three tissues (expression in *R26X500* trachea and liver is difficult to visualise on photograph). As expected, *ROSA26* showed expression in all tissues, while no product was generated from non-transgenic samples. (c) The *lacZ* -RT controls confirmed that RNA samples were not contaminated with genomic DNA. (d) Primers amplifying a 209bp product from *GAPDH* cDNA revealed the presence of RNA in all samples. M = 1kb ladder, map not to scale.

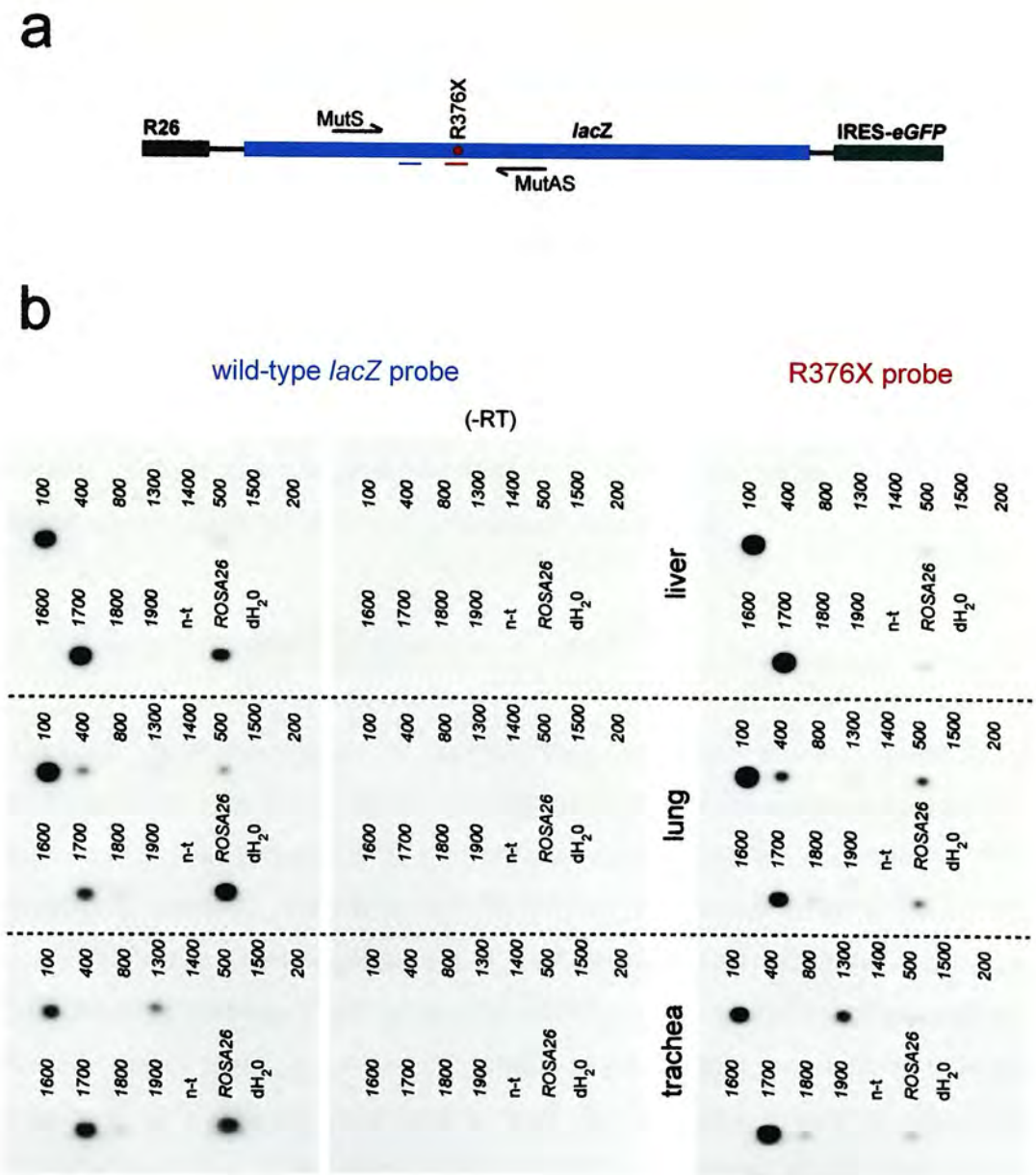


Figure 5.11(ii) *R26X* lines express mutated *lacZ* transcript.

(a) Schematic showing position of oligonucleotide probes on the *R26X* transgene with respect to the R376X mutation and the PCR primers used in (i). (b) *lacZ* RT-PCRs of 12 *R26X* lines plus controls were blotted and probed with a 19mer radiolabelled oligonucleotide complementary to wild-type *lacZ* sequence 5' of R376X, within the region amplified by the primers used in (i) ("wild-type *lacZ* probe", shown in blue). Specific hybridisation confirmed the identity of the RT-PCR amplification products. No hybridisation was observed on the -RT blot, or in the non-transgenic control. The blots were then stripped of the wild-type *lacZ* probe and hybridised with a 17mer probe corresponding to the mutated region and containing the R376X base pair ("R376X probe", shown in red). While the *R26X* samples showed a similar hybridisation intensity with both probes, the *ROSA26* sample showed significantly reduced hybridisation to the R376X probe, indicative of a deviation from wild-type *lacZ* sequence within the *R26X* RNA. No hybridisation was observed on the -RT blot using the R376X probe (not shown).

5.5 Discussion

The generation and analysis of a transgenic mouse designed to facilitate the efficient and functional evaluation of gene correction methodologies *in vivo*, has been described. Disadvantages in using mouse disease models for this purpose are that (1) the functional assays required to detect and quantify correction events are often time-consuming and labour-intensive and (2) the functional data they provide is generally tissue-specific in nature. An approach employing a ubiquitously expressed mutation, the functional correction of which being detectable by routine histochemistry, would counter both of these weaknesses and be an invaluable reagent for the further investigation and development of gene correction protocols.

5.5.1 Site-directed mutagenesis of lacZ

The mutation introduced into the *lacZ* sequence of pCMVβ, a C to T substitution at base pair 1126, was chosen for several reasons. Firstly, since complete abrogation of βgal activity is a prerequisite for the confident detection of rare correction events, a premature nonsense mutation was desirable. This mutation at position 1126 corresponds to a CGA (arginine, R) to UGA (opal stop, X) at amino acid residue 376. Since the version of βgal protein in pCMVβ consists of 1047 residues, mutation R376X should result in a truncated peptide approximately one third of wild-type length. While a premature stop codon which also induced a frameshift would have provided protection in the event of unanticipated readthrough of the stop codon, a base insertion or deletion would have been required to effect this. Although both RDOs and SDFs are known to be capable of correcting such mutations, at least RDOs do so with a lower frequency than base substitutions (Liu *et al.*, 2001). Thus, a frameshift mutation would be potentially more refractory to the intended correction methods. No translational suppression of the R376X mutation was observed, other than faint xgal staining in bacterial transformation colonies.

Secondly, the R to X substitution was selected with the intention that it would provide confirmation of the "correctability" of *lacZ*. It has been reported that aminoglycoside antibiotics such as G-418 are able to suppress premature nonsense mutations through the misincorporation of an amino acid following the pairing of a near-cognate animoacyl tRNA with the stop codon (Manuvakhova *et al.*, 2000). The ability to effect translational readthrough of the mutated *lacZ* gene in cell culture would provide an elegant means of validating the potential functionality of the gene, demonstrating an absence of additional defects introduced during mutagenesis. Opal codons were found to be the most susceptible to antibiotic-induced suppression (Manuvakhova *et al.*, 2000), and two of four premature stop codons within the *CFTR* gene, which were suppressed by G-418 and gentamicin (Howard *et al.*, 1996; Bedwell *et al.*, 1997), were arginine (CGA) to opal mutations. However, the use of G-418 at concentrations previously described to cause suppression (Howard *et al.*, 1996) did not induce detectable levels of readthrough, though this failure was consistent with current, similar attempts by fellow researchers (B. Stevenson, personal communication). Integrity of the mutant *lacZ* gene was subsequently confirmed by sequencing.

Finally, the R376X-induced loss of a TfiI restriction enzyme recognition sequence provided a crucial means of checking plasmid or transgene identity. Simple RFLP analysis of digested plasmids, or of PCR products amplified from the genomic DNA of transgenic mice, permitted rapid and unambiguous detection of the mutated sequence. This assay prevented any potential confusion arising from the parallel production of *R26X* and *R26Z* transgenic mice.

5.5.2 The R26 promoter

In order that the model would be useful for evaluating gene repair efficiencies in all tissues, a ubiquitously active promoter was required to drive expression of the mutated *lacZ* gene. The plasmid in which the R376X mutation was inserted into *lacZ* contained the human CMV promoter, which is generally considered to be active in a broad array of tissue types, and is commonly used to drive transcription in

mammalian expression vectors. However, it is known that many constitutive viral promoters can undergo cytokine-mediated attenuation in a transgenic context (Qin *et al.*, 1997; Paillard, 1997). While the CMV promoter is broadly and strongly active *in vitro*, it becomes attenuated *in vivo* (Norris *et al.*, 1998), and there is evidence to suggest it performs poorly in the lung (Schmidt *et al.*, 1990), an important target tissue for gene correction research. For these reasons, the CMV promoter was replaced with the R26 promoter, a 0.8kb fragment derived from sequence 5' of exon 1 of the *ROSA26* locus, which had previously demonstrated promoter activity when fused to *lacZ* and electroporated into embryonic stem (ES) cells (Zambrowicz *et al.*, 1997). When R26-driven transgenes were used to make transgenic mice and rats by conventional pronuclear microinjection, the promoter was shown to drive expression of both hPAP and eGFP in a ubiquitous pattern which faithfully reflected that of the genomic *ROSA26* locus (Kisseberth *et al.*, 1999).

5.5.3 R26Z *lines*

To assess the performance of the R26 promoter, the *R26Z* control construct carrying non-mutated *lacZ* was generated. In cell culture, βgal production was observed in a range of cell types, though at consistently and markedly lower levels than produced by an equivalent construct containing the CMV promoter. However, it is known that the relative activity of various promoters observed *in vitro* is not necessarily reflected in transgenic mice (Palmiter and Brinster, 1986), this point being particularly pertinent given the acknowledged difference in CMV promoter activity between *in vitro* and *in vivo* contexts. Examination of E12.5 embryos from the two lines which expressed the *R26Z* transgene revealed widespread, if not ubiquitous, βgal production, though expression levels did not appear equal between these lines. One of the lines exhibited intense xgal staining in the CNS, while in the other line, βgal expression was elevated in the developing brain, but not in the spinal cord. In both lines, βgal expression in non-CNS tissues appeared significantly lower than reporter gene expression in E12.5 embryos described by Kisseberth and colleagues (Kisseberth *et al.*, 1999).

In that study, embryos stained for hPAP activity displayed broadly uniform transgene expression and were indistinguishable in photographs from xgal stained heterozygous *ROSA26* embryos (Kisseberth *et al.*, 1999). An R26-*eGFP* transgene revealed that the promoter was active in late four-cell embryos, and by E12.5 eGFP protein was detectable throughout embryos. Importantly, the authors report that hPAP expression is maintained in adult tissues, with lung being amongst the highest expressing organs when analysed visually for metabolised substrate. Transgene expression was sufficiently strong that foci of hPAP- or eGFP-positive cells were visible in recipient mouse liver after transplantation of transgenic liver cells. This finding suggested that infrequent cells expressing a corrected *lacZ* gene from the R26 promoter would be detectable against a background of non-corrected cells.

However in the *R26Z* lines, despite βgal expression appearing increasingly homogenous at E15.5, probably as a result of greater cell density and a reduced ability to visualise internal structures, the only adult tissue examined which exhibited βgal production was the brain. Xgal staining in the brain was more intense and widespread in *R26Z300* than *R26Z100*, reflecting the embryonic situation.

5.5.4 R26X *lines*

Given the presence of xgal stained cells after transfection with pR26Z, the lack of apparent βgal activity in cells transfected with pR26X could be taken to demonstrate the successful inactivation of the protein. However, due to the variable effects of integration site and copy number on transgene expression *in vivo*, it was desirable to monitor transgene expression directly in each line, without comparison to other lines. As a result of the premature stop codon engineered into the *lacZ* gene of construct *R26X*, βgal production could not be used as an assay for transgene expression in *R26X* animals. For this reason, the *eGFP* gene was incorporated such that *R26X* (and *R26Z*, produced in parallel) was a bicistronic construct in which *lacZ* and *eGFP* would be transcribed as one unit, but translated independently, through the use of an IRES sequence upstream of *eGFP*. Of seven lines analysed for eGFP fluorescence in E12.5 embryos, only one showed visually detectable expression. *R26X100* embryos

exhibited a generalised low-level fluorescence compared to non-transgenic littermates, with increased eGFP production in the midbrain. As in the case of the *R26Z* lines, eGFP expression was not apparent in adult tissues.

The complete absence of observed βgal activity in *R26X* embryos and adult tissues, including *R26X100* which evidently expressed the transgene embryonically, suggested that the stop codon was not inappropriately suppressed to produce full-length protein at a level likely to, in principle, impair use of the mice in detecting rare gene correction events.

Analysis of βgal expression in selected tissues from twelve *R26X* lines, four of which were previously examined for embryonic eGFP production (including *R26X100*), revealed six lines showing some degree of *lacZ* transcription. Three of these lines, including *R26X100*, showed expression in all three tissues analysed (lung, trachea, and liver). While βgal expression in *R26X100* was predicted by embryonic eGFP production, line *R26X500*, which exhibited less intense RT-PCR signals in all three tissues, showed no fluorescence. *R26X1700*, giving similar RT-PCR signal strengths to *R26X100*, was not analysed for eGFP. Although differences in the amount of RT-PCR product generated between lines may have been broadly reflective of relative transcription rates, a comparison of the strength of hybridisation of the *ROSA26* product and the *R26X100* and *R26X1700* product with the wild-type *lacZ* probe revealed the quantitative limits of non-limiting cycle RT-PCR. Although apparently similar levels of RT-PCR product were generated, it is highly unlikely that *R26X100* produced (mutant) βgal in quantities comparable to that produced by *ROSA26*, given the evidently reduced level of eGFP production from *R26X100* compared to *R26Z300* (Figure 5.10a,b and Figure 5.6b-f). *R26Z300* clearly produced lower levels of βgal than *ROSA26* (Figure 5.8f and Figure 5.7j). Inclusion of RNA samples from *R26Z100,300* in the *lacZ* RT-PCR, and the use of quantitative or semi-quantitative RT-PCR methods, would have been useful in clarifying relative expression levels.

5.5.5 Low levels of eGFP expression from the *IRES*-eGFP

Although sufficient to allow the visual (using fluorescence microscopy) discrimination of transgenic and non-transgenic embryos in certain lines, eGFP fluorescence was weak compared to βgal activity, and was not useful as a means of identifying live transgenic animals either directly or by ear biopsy (data not shown). More importantly, fluorescence could not be used as an indicator of transgene activity in adult tissues, as indicated by the strongly βgal positive, but eGFP negative, brain tissues of *R26Z300* mice. This demonstrated the value of generating the control *R26Z* lines: had the loss of detectable eGFP production, but maintenance of βgal activity, from embryo to adult brain in *R26Z300* not been observed, it may have been assumed that the absence of eGFP in adult *R26X100* brain indicated inactivity of the transgene.

Although eGFP signal is potentially less intense than the signal from histochemically-visualised reporters, which can be enhanced through prolonged incubation with substrate, the weak eGFP signal observed in this study was likely partly due to the use of the IRES. It has been shown that in biscistronic constructs of the type described here, expression of the IRES-dependent second gene is typically 20-50% that of the first gene (Mizuguchi *et al.*, 2000). The strength of IRES-driven translation appears to be dependent on both the composition and arrangement of the cistrons, with certain gene sequences, when in the upstream position, able to exert an inhibitory effect over expression of the downstream gene (Hennecke *et al.*, 2001). The behaviour of *lacZ* as the upstream cistron, in this regard, is not known. However, the IRES-*eGFP* sequence was incorporated into the constructs as a single unit, ensuring the crucial spacing between the IRES and *eGFP* components (Houdebine and Attal, 1999), presumably optimised in the production of the commercial parent vector (pIRES2-eGFP), was not disturbed.

5.5.6 Non-ubiquitous expression of R26X and R26Z transgenes

The discrepancy between the ubiquitous R26-mediated expression reported by Kisseberth and colleagues (Kisseberth *et al.*, 1999) and the weaker, apparently CNS-biased expression observed in this study, is intriguing. Since only two transgene-expressing *R26Z* lines were obtained, position effect on the transgene integration sites cannot be eliminated as a cause. However, it would seem that at least the CNS aspect of the expression profile, also observed by eGFP analysis of *R26X100* embryos, must be dictated by the promoter sequence. Support for position effect susceptibility of the promoter fragment comes from the previous study: although more than half of the R26-*hPAP* lines expressed highly, only about one in five of the R26-*eGFP* lines expressed highly enough to be suitable marker animals, and expression levels correlated with transgene copy number (E. Sandgren, personal communication). Thus, the nature of the coding sequence might influence the degree to which expression is epigenetically modulated. It is known that different genes, when expressed from the same promoter/enhancer, can generate markedly different levels of mRNA not explainable by variations in mRNA stability alone (Palmiter and Brinster, 1986). This suggests that different genes can influence the rate of transcription, through uncharacterised mechanisms.

lacZ in particular is prone to exhibiting unanticipated or variegated expression patterns (Ramirez *et al.*, 2001). In a previous study, tandem arrays of 8-9 copies of a ubiquitously expressed *lacZ* transgene were found to be entirely methylated after passage through the female germline, the methylation status correlating with complete inactivity of the transgene in adult tissues (Lau *et al.*, 1999). Paternally-inherited transgene arrays, which produced detectable levels of βgal, were methylated at all but one copy of the transgene. The unmethylated, active copy was found to reside adjacent to the junction with the host genome. It is believed that the tandem fashion in which transgenes tend to integrate is recognised by the host cell and becomes a target for methylation-mediated silencing. The methylated transgene copies may then become a source of spreading heterochromatin, with only the proximity of the terminal copy to cellular euchromatin maintaining its expression.

Downregulation of a *lacZ* transgene during development from embryo to adult, similar to that observed in this study, has previously been reported and may suggest that the *lacZ* sequence, rather than the R26 promoter, is responsible (Beddington *et al.*, 1989). Also, the presence of prokaryotic vector sequences have been shown to be inhibitory for the expression of some transgenes (Palmiter and Brinster, 1986), though whether this remains an issue given the prokaryotic origin of the transgene coding sequence is unclear. Although most of the vector sequence was removed from the *R26Z* and *R26X* constructs, approximately 1.4kb remained adjacent to the terminal BspHI sites.

With hindsight, clues to the unexpectedly weak expression driven by the R26 promoter fragment may have been present since its initial isolation. When fused to *lacZ* and electroporated into ES cells, expression from the R26 promoter was observed to be 8-fold less than that from the positive control phosphoglycerate kinase-1 (PGK) promoter, which itself produced βgal at levels 3-fold less than ES cells isolated from *ROSA26* mice (Zambrowicz *et al.*, 1997). The authors speculated that the difference between expression from the genomic *ROSA26* locus and the 0.8kb R26 promoter may have been due to position effects on the randomly integrated transgene. In a personal communication, E. Sandgren stated that he found *lacZ* technically difficult to work with, despite producing ubiquitously expressing lines with *hPAP* and *eGFP*. It may be that, when fused to *lacZ*, deleterious consequences of position effect manifest on the R26 promoter at any integration site other than the native *ROSA26* locus. It is notable that, in the four years since the publication of what appeared to be an excellent choice of promoter for ubiquitous transgene expression, no additional mouse lines harbouring R26-driven transgenes have been described. Anecdotal reports indicate that, while several groups may have had success using the R26 fragment, most have failed to mimic the expression described by Kisseberth and co-workers (Kisseberth *et al.*, 1999) (A. Schedl, K. Foley, personal communications).

5.5.7 Alternative approaches

5.5.7.1 The CAGG promoter

The CAGG promoter may be a better choice for driving ubiquitous and strong expression of randomly integrated transgenes. This promoter is a fusion between the CMV-immediate early enhancer and the chicken β-actin promoter, and was initially shown to demonstrate strong expression of *lacZ* in transfected cells (Niwa *et al.*, 1991). It has since been used to drive strong expression of eGFP in "green mice", which express the transgene in all cells except erythrocytes and hair (Okabe *et al.*, 1997). When analysing promoters to drive ubiquitous expression of the Z/AP conditional double reporter for Cre recombinase activity, Lobe and colleagues found the CAGG promoter to drive stronger and more consistent expression than the CMV, human β-actin, and pIIHMG-coenzyme A reductase promoters (Lobe *et al.*, 1999). More recently, the promoter has been used to express an inducible Cre recombinase, though the authors had to screen 17 lines to find two exhibiting ubiquitous expression (Hayashi and McMahon, 2002). It is possible that, had many more than five *R26Z* lines been available for analysis, a subset may have produced the expected ubiquitous expression, although this depends upon whether the *lacZ* sequence has unanticipated influence over the expression pattern. Interestingly, Hayashi and colleagues reported that in ten of the 17 lines examined, Cre activity was restricted to the heart, somites and a portion of the CNS (Hayashi and McMahon, 2002). Thus, restriction of transgene expression to a predictable subset of tissues in multiple independent lines may be a common feature of randomly integrated, supposedly ubiquitously expressed transgenes.

5.5.7.2 The genomic ROSA26 locus

The ubiquitous and robust transgene expression driven by the *ROSA26* locus, from which the R26 promoter fragment was derived, has been well characterised (Friedrich and Soriano, 1991; Zambrowicz *et al.*, 1997). A more reliable, though more complex and time-consuming approach to achieving robust and ubiquitous

expression of the mutated reporter gene, would be to insert ("knock-in") the construct into the *ROSA26* locus of mouse ES cells using homologous targeting. This has been done using a glutamic acid to lysine substitution at amino acid position 461 of *lacZ*, previously demonstrated to abolish βgal activity *in vitro* (Nickerson and Colledge, 2003). Expression of the mutated *lacZ* gene was comparable to that of the *ROSA26* mouse, and the mutant βgal protein was detected in all tissues analysed, however no *in vivo* correction has yet been achieved (H. Nickerson, personal communication).

5.5.7.3 A new reporter system for gene correction strategies

Recently, a novel correctable reporter system was shown to sensitively detect repair in cell culture, and could potentially be used as a non-invasive means of assessing *in vivo* correction rates (Bennett and Schaack, 2003). The luciferase genes from *Renilla* (sea pansy, RLuc) and the firefly (FLuc) were fused, before a single base pair in the downstream FLuc was deleted to introduce an ochre stop codon, and shift the downstream sequence out of frame. In initial tests, the "RLuc/Fluc" transgene, under the control of a ubiquitous viral promoter, was transfected into cells with known amounts of the non-mutated version of the construct. It was shown to generate a readout indicative of the degree of simulated repair, providing the "correction" rate was below 50%. The variable error rate is a function of the complex formula used to calculate correction efficiencies, which incorporates the internal control RLuc reading, the FLuc reading (both adjusted for background and dilution), the known ratio of FLuc/RLuc activity, and the RLuc activity difference (RAD, the unanticipated cell line-specific difference in RLuc activity between the non-mutated and mutated versions of the construct), to calculate a percent correction efficiency.

Since the system measures the total increase in light emitted as a result of substrate metabolism by a population of cells, it reveals the number of corrected, functional sequences (Bennett and Schaack, 2003). In this aspect it differs from conventional systems, such as mutated *lacZ* (Nickerson and Colledge, 2003) or *eGFP* (Thorpe *et al.*, 2002a), which depend on the discrimination of individual corrected cells, and

thus indicate the number of cells in which one or more conversion events have occurred. The fusion of the luciferase genes is crucial to the consistency of the system, preventing differential ratios of RLuc and FLuc transcription or translation between cell lines, as may have occurred had separate promoters, or an IRES, been used. The system was used to demonstrate SDF- and SSO-mediated correction both in cell lysates and non-invasively, using a highly sensitive camera to detect light emission after addition of the respective luciferase substrates to the cell culture medium. The authors believe the latter technique could be used to measure gene correction in transgenic mice carrying the RLuc/FLuc construct. The dual luciferase system appears to address the same limitations inherent to mouse disease models that prompted the development of the *R26X* mouse. However, the necessary input of the cell line-specific RAD into the formula may complicate attempts to assess *in vivo* correction rates from tissue lysates comprising multiple cell types.

5.5.8 Future Work

Further analysis of the *R26X* lines is necessary to determine whether one or more lines express the transgene in a manner appropriate for use in evaluating gene correction activity. However, several factors led to a decision to cease further study of the model, and to abandon plans to use the mice for *in vivo* gene correction within the course of the work undertaken for this thesis. The departure of Dr P. Thorpe, who was leading the gene correction project in the Porteous Cystic Fibrosis Gene Therapy laboratory (Thorpe *et al.*, 2002a; Thorpe *et al.*, 2002b), to undertake more basic research into recombinational processes in yeast, meant that gene correction experiments were not performed after September 2002. Also around this time, the Cystic Fibrosis Trust Gene Therapy Consortium, which was funding the gene correction research, became reluctant to continue this support in a climate of increasing scepticism regarding the efficiency and reproducibility of gene correction, particularly the use of RDOs (van der Steege *et al.*, 2001; Taubes, 2002). Combined with the time demanded by the work described elsewhere in this thesis, these circumstances prompted the decision to preserve the lines for investigation and possible use by others.

5.5.8.1 Further analysis of R26X lines

Despite the increasing scepticism surrounding RDOs and derivative technologies purported to operate via induction of host mismatch repair systems (Albuquerque-Silva *et al.*, 2001; Taubes, 2002; Diaz-Font *et al.*, 2003), reports continue to emerge which describe the successful application of this technique (Bertoni *et al.*, 2003; Lu *et al.*, 2003), and of SFD-mediated repair (Bruscia *et al.*, 2002). In fact, the first report of gene correction by oligonucleotides (SSOs) in ES cells appeared recently, and indicated that the mismatch repair machinery acts as a significant impediment to the process of homology-mediated gene repair (Dekker *et al.*, 2003). Given the persisting interest in the concept of gene correction (Sullenger, 2003; Liu *et al.*, 2003) as the ideal manifestation of gene therapy, it is conceivable that others may wish to utilise the *in vivo* gene correction model described here. However, this would only be feasible after further analysis of the lines to identify whether, within the 12 *R26X* lines preserved as frozen sperm stocks (corresponding to the 12 lines analysed by RT-PCR, section 5.4.3), a subset may exhibit ubiquitous and sufficiently strong expression of the transgene. This could involve the use of anti-βgal, or potentially anti-eGFP, antibodies on tissues sections or western blots. Quantitative RT-PCR could be employed to accurately determine relative levels of transcript for comparison with, for example, the *ROSA26* line.

It would be important to calculate, for any line believed to be suitable for the intended use, its transgene copy number. Since correction of a single copy of the *R26X* transgene should be sufficient to yield a βgal-positive cell, correction frequency may increase as the number of correctable transgenes increases. Thus, a comparison of correction frequency between different *R26X* lines, or between an *R26X* line and the *ROSA26* locus-targeted mouse mentioned previously (section 5.5.7.2), would have to take account of the relative number of target sites. The potential for multiple copies of the correctable reporter, compared to the maximum of two copies in a mouse homozygous for the *ROSA26*-targeted mutant *lacZ* construct, may be crucial in elevating the frequency of correction events to a detectable level. Also, the randomly distributed integration sites, which will differ

between independent *R26X* lines, could in principle be informative regarding the ability of gene correction to operate at various genomic locations.

Chapter 6 Concluding Remarks

6.1 A transgenic mouse for the investigation of putative tracheal stem cells

The main focus of the work presented in this thesis was the further investigation of a putative stem or progenitor cell niche within submucosal gland (SMG) ducts of the adult murine proximal trachea. Slow-cycling cells had previously been identified as label-retaining cells in SMG ducts, and these cells appeared spatially coincident with cells expressing high levels of keratin protein, also reflected by the expression of a reporter transgene driven from the human keratin 5 (K5) promoter. To further investigate the proposition that K5 might be a marker for these putative stem cells, transgenic lines were produced in which a fusion gene encoding an inducible Cre recombinase was linked to the human K5 promoter. The intention was that the onset of ubiquitous, permanent expression of a Cre-dependent reporter gene could be temporally activated within the K5 promoter-expressing population, resulting in marking of these cells and their cellular progeny. Induction of lineage marking after injury-stimulated stem cell recruitment would permit the evaluation of the contribution made by the K5 promoter-expressing cells to epithelial regeneration. A range of techniques was used to select lines in which detectable levels of reporter gene activation could be achieved through administration of an exogenous inducer.

The available time permitted only a preliminary experiment using one of these lines. While not informative with regard to the trachea, inducible reporter gene activity was detected in other tissues, and the experiment served to highlight parameters which might productively be adjusted in subsequent studies employing this system. Efforts are ongoing to further evaluate the use of these lines in marking tracheal cell lineage.

A parallel preliminary experiment with a transgenic mouse expressing constitutive (non-inducible) Cre from the bovine K5 promoter revealed potentially interesting lineage patterns consistent with the proposed existence of a stem cell niche within SMG ducts. However, as anticipated, the clarity of the lineage data obtained was undermined by the frequency of marking effected by the constitutive activity of Cre, partially validating the inducible strategy. Further studies using this system are

required to determine the significance of the patterns observed, and may prove informative with regard to the role of K5 promoter-expressing cells in both regenerative and steady-state contexts.

6.2 The molecular basis of alternate recombination patterns mediated by the *bK5Cre* transgenic mouse

In studies preceding the latter experiment, the constitutive bK5 promoter-driven Cre transgenic line revealed alternate patterns of expression from a Cre-dependent reporter transgene. Analysis of embryonic litters revealed a correlation between the pattern of reporter gene expression and the sex of the parent from which the *Cre* transgene was inherited: while the paternally-inherited transgene resulted in the expected K5-derived pattern of reporter gene expression, ubiquitous reporter expression was apparent in the progeny of *Cre* transgene-bearing females. Interestingly, the ubiquitously expressing phenotype was independent of *Cre* transgene inheritance. Immunofluorescence and histochemical analysis of *Cre* transgenic ovaries indicated the unanticipated expression of Cre protein in oocytes. The accumulation of Cre protein in oocytes caused recombination of the Cre-dependent reporter transgene within the female germline. Propagation of this permanent genetic rearrangement throughout development resulted in ubiquitous reporter gene expression in offspring. That ubiquitous recombination was detected in mice inheriting the Cre-dependent reporter transgene paternally, but failing to inherit the maternal *Cre* transgene, suggested the persistence of maternally-synthesised Cre protein into the early zygote. The *bK5Cre* line thus provides both tissue-specific and generalised recombination of floxed alleles, depending on the sex of the *bK5Cre* transgene-transmitting parent.

6.3 An *in vivo* model for functional analysis of gene correction

Gene correction is a recent concept in gene therapy, and describes the site-specific repair of mutations at their chromosomal loci. Before gene correction systems and

protocols can be employed to clinical benefit, they must be extensively evaluated and developed *in vivo*. While animals models have provided elegant demonstrations of gene correction, the functional readouts are often time-consuming and labour-intensive, and thus unsuitable for the large-scale evaluation of experimental parameters. To address this, a transgenic mouse was generated which carried a mutated, but potentially correctable, reporter gene under the control of a ubiquitously expressing promoter. Analysis of control transgenic lines carrying a non-mutated version of the reporter transgene suggested the promoter did not reliably drive ubiquitous expression. In a subset of the mutated reporter transgenic lines, transcript corresponding to the mutated reporter sequence was detected in several adult tissues. While this suggested that these mice may be usefully employed in the evaluation of gene correction strategies, further work is required to confirm this.

6.4 Airway stem cells as targets for gene correction therapy

Both the inducible and constitutive Cre transgenic lines described in this thesis have the potential to be valuable reagents in the further investigation of putative airway stem cells. The identification of stem cells within the respiratory epithelium would facilitate attempts to target them in gene therapy strategies. Because of their potential for repopulating entire tissues, stem cells are obvious targets for gene complementation or correction. With respect to cystic fibrosis (CF) gene therapy, the existence of a putative stem cell compartment in the SMG duct is intriguing given that certain cell types within SMGs represent the primary sites of cystic fibrosis transmembrane conductance regulator (*CFTR*) expression in the human airway (Engelhardt *et al.*, 1992). *Cftr* has also been detected in the mouse SMG (Borthwick *et al.*, 1999). The relationship between the K5-expressing, label-retaining, and *Cftr*-expressing cells of the SMG is not clear. Murine SMGs are thought to be clonally-derived, and observations from *ROSA26*/wild-type chimaeric mice (Borthwick, 1999) and *bK5Cre*^{tg} *R26R*^{tg} mice (chapter 3), suggests a relationship between SMGs and the local surface epithelium. It may be that SMGs act as structural-proliferative units (Slack, 2000), in which keratin-rich, label-retaining stem cells are responsible

for populating gland cells and a finite region of neighbouring epithelium. In this way, assuming that the *CFTR*-expressing cells are not themselves stem cells, a CF gene therapy targeted at the SMG duct stem cell might be transmitted to the other gland cells including those expressing *CFTR*, with therapeutic consequences.

While the delivery of gene therapy reagents to stem cells is a prerequisite for a single-dose therapy, many conventional gene complementation protocols provide only transient therapeutic benefit, necessitating repeated administration. Combined with stem cell targeting, the potential of gene correction strategies to offer heritable restoration of physiologic gene expression may facilitate the realisation of a single-dose "cure". However, much work is required to advance current gene correction technologies to a level of therapeutic efficacy. The mutated reporter transgenic lines described in this thesis may prove valuable in the further development of such strategies.

Chapter 7 Materials and Methods

7.1 Manipulation of nucleic acids

7.1.1 General molecular biology reagents

All chemicals were analytical grade and were supplied by Sigma, Promega, Gibco BRL, Invitrogen, BDH, and Fisher Scientific. Restriction enzymes were supplied by Roche and New England BioLabs (NEB). Nucleic acid manipulations were performed in 1.5ml or 0.5ml microfuge tubes. General solutions were prepared by technical staff, autoclaved and stored at room temperature.

Tris.HCl

Tris base (tris[hydroxymethyl]aminomethane was dissolved in sterile water. HCl was used to adjust the pH to the required value.

EDTA

EDTA (ethyldiaminetetra-acetic acid di-sodium salt) was dissolved in sterile water. The solution was adjusted to pH8.0 by adding solid NaOH.

TE buffer

10mM Tris.HCl (pH7.5); 1mM EDTA.

TBE buffer, 20× stock

Tris base	216g
Boric acid	110g
0.5M EDTA	80ml

Distilled water was added to a final volume of 1 litre. Stock was diluted to 1× with distilled water.

TAE buffer, 50× stock

Tris base	242g
Glacial acetic acid	57.1ml
0.5M EDTA	100ml

Distilled water was added to a final volume of 1 litre. Stock was diluted to 1× with distilled water.

7.1.2 Agarose gel electrophoresis

For preparing gels, the required amount of agarose (BioGene) was dissolved in either 1× TBE or TAE by heating. Molten agarose was cooled and ethidium bromide added to a final concentration of 50μg/100ml agarose. DNA size markers were run alongside experimental samples to estimate the size and amount of DNA in the samples. The size markers used routinely were 1kb DNA ladder (Invitrogen) and bacteriophage lambda DNA digested with HindIII (Invitrogen). Samples were run in 1× loading buffer. Electrophoresis was used when DNA fragments of a particular size needed to be purified. To do this, samples were electrophoresed using low melting point agarose (Flowgen), excised from the gel using a razor blade, and the DNA was recovered using a Gel Extraction Kit (Qiagen), according to the manufacturer's instructions. Ethidium bromide stained gels were visualised by UV transillumination using a digital camera (BioRad).

7.1.2.1 Solutions

Gel loading buffer, 10×

Reagents were added to distilled water to obtain a final concentration of 100mM EDTA, 20% Ficoll and 1% Orange G.

7.1.3 DNA quantification

DNA concentrations were determined in one of two ways: by agarose gel electrophoresis or by measuring the absorbance in a spectrophotometer at 260nm (A_{260}). To determine the concentration by electrophoresis, several different volumes of the DNA were run alongside Low or High DNA Mass Ladder (Invitrogen). An

estimate of the concentration was made by visual comparison of the samples with the known amounts of DNA under UV illumination.

To determine the concentration with a spectrophotometer, the DNA or RNA sample was diluted 1:40 with dH_2O. The spectrophotometer was calibrated using a dH_2O blank sample. The samples were placed in clean quartz cuvettes and the absorbance at 260nm (A_{260}) was measured. The concentration of the original sample in mg/ml was calculated as follows:

Concentration (mg/ml) = $A_{260} \times 40$ (dilution factor) $\times$ 50.

7.1.4 Restriction enzyme digestion

Digests were ideally carried out in a large volume (usually 50-100µl) to minimise effects of evaporation. Minimum reaction volume was dictated by the requirement that the total volume of enzyme never exceeded 10% of the total reaction volume, to prevent inhibition of digestion by glycerol. Enzyme was added at a concentration of 5-10units/µg DNA, depending on the duration of the digests. The manufacturer's guidelines were followed to determine the correct buffer and incubation temperature for each enzyme. Double digests were performed where enzyme buffer requirements and incubation temperatures were compatible.

7.1.5 Making cohesive ends blunt

When cohesive ends needed to be made blunt for subsequent ligations, T4 DNA polymerase and Mung Bean Nuclease (both Roche) were used for filling in 3' recessed ends and removing 3' protruding ends, respectively.

7.1.5.1 T4 DNA polymerase

For every 1μg of DNA, 1× T4 DNA polymerase buffer, 1μl bovine serum albumin (BSA), 100μm dNTPs, 1μl T4 DNA polymerase, and dH_2O to final volume was used in the incubation (20 minutes at 11°C), and the reaction terminated by heating at 75°C for 10 minutes.

7.1.5.2 Mung Bean Nuclease

For every 1μg of DNA, 1× Mung Bean Nuclease buffer, 1 unit of Mung Bean Nuclease, dH_2O to final volume was used in the incubation (30 minutes at 30°C), and the reaction terminated by addition of SDS to 0.01%.

After every enzymatic reaction (restriction/nuclease/polymerase), samples were cleaned using the PCR Purification Kit or Nucleotide Removal Kit (both Qiagen), following the manufacturer's instructions.

7.1.6 Removal of phosphorylated 5' termini

Prior to ligation, vector DNA was treated with calf intestine alkaline phosphatase (CIP) (Roche) to dephosphorylate the 5' termini of vector DNA, preventing self-ligation. The linearised, purified vector DNA was made up to a volume of 90μl with sterile dH_2O followed by the addition of 10μl 10× buffer and 1μl CIP. Following 2×15 minute incubations, the first at 37°C and the second at 56°C, a further 1μl CIP was added and the 2×15 minute incubations repeated. The reaction was terminated by the addition of 2μl EDTA and the CIP enzyme heat inactivated by incubation for 20 minutes at 65°C. The dephosphorylated vector DNA was purified using a PCR Purification Kit (Qiagen), following the manufacturer's instructions.

7.1.7 Ligations

Fragments with cohesive complementary ends were ligated overnight at 16°C, using 1 unit of T4 DNA ligase (Roche). Blunt-ended ligations were performed overnight at room temperature using 5 units of high concentration ligase (Roche). In order to maximise the ligation efficiency, ligations were usually set up with 1:3, 1:1, 3:1 vector:inset molar ratios, usually using 50ng vector. The supplied ligation buffer was used, normally in 10µl reaction volume.

7.1.8 Removal of buffer salts

Prior to electroporation of ligation reactions, buffer salts were removed by drop dialysis. A nitrocellulose filter (Millipore) was placed in a petri dish containing filter sterilised dH_2O. The ligation reaction was carefully placed on the filter and salts were left to diffuse through the filter for at least 1 hour at room temperature.

7.1.9 Extraction of genomic DNA for Southern blot hybridisation

Tissue samples were re-suspended in 500µl of tail mix buffer (1M Tris pH8.0; 0.5M EDTA; 10% SDS; 5M NaCl). 5µl of proteinase-K (10 mg/ml) (Sigma) was added with gentle mixing. The samples were incubated at 55°C overnight, followed by incubation for 1 hour at 37°C with 10µl of 10mg/ml DNase-free RNase (Roche). The DNA was cleaned by 2× phenol:chloroform extractions, plus one chloroform extraction. DNA was precipitated from the aqueous layer with 0.6 volumes of 100% ethanol and centrifuged at 13,000 rpm for 20 minutes at 4°C. The supernatant was removed and the pellet was washed with 70% ethanol. The supernatant was discarded and the pellet was allowed to dry at room temperature and then re-suspended in dH_20.

7.1.10 Southern blot hybridisation

10μg of genomic DNA was digested overnight with BglII. The sample was then mixed with 5μl of loading buffer, and loaded onto an ethidium bromide stained agarose (0.8% 1× TAE) gel. Samples were electrophoresed until appropriate separation of 1kb ladder fragments had occurred, as detected by UV transillumination, and the position of the ladder fragments were noted with a ruler. The gel was immersed in denature buffer (0.5M NaOH; 1.5M NaCl) with agitation for 30 minutes and neutraliser buffer (0.5M Tris; 1.5M NaCl) for 15 minutes, before being placed onto 20× SSC-soaked 17mm filter paper (Whatman) in the blotting apparatus. The following was then placed on top of the gel: 1 piece of Hybond-N+ membrane (Amersham pharmacia biotech) soaked in 20× SSC, 2 pieces of 20× SSC-soaked 3mm filter paper (Whatman), followed by a 5cm stack of dry paper towels, a glass plate, and two 500ml bottles half-filled with water (weight approximately 0.5kg). The blotting apparatus was left overnight. The apparatus was dismantled and the position of the wells marked on the blot with a pencil. The blot was rinsed briefly in 2× SSC before the DNA was UV cross-linked using the auto cross-link feature on a Stratalinker 1800 (Stratagene). The same technique was used to blot electrophoresed PCR products for oligonucleotide hybridisation.

7.1.10.1 Preparation of radioactively labelled DNA probes

To label double stranded DNA probes, approximately 50ng of purified DNA in a total volume of 12μl was boiled for 10 minutes to ensure denaturation and then labelled using 4μl of High Prime (Roche) and 30mCi [α-^{32}P]-dCTP. The reaction was incubated at 37°C at 30 minutes and labelled DNA was separated from unincorporated nucleotides by passing the labelling reaction through TE equilibrated spin columns (Amersham pharmacia biotech). The probe was denatured by boiling for 10 minutes with 1mg sonicated salmon sperm DNA (Sigma) and incubating at 68°C for 45 minutes. The probe was then added to the pre-hybridised blot.

7.1.10.2 Preparation of radioactively labelled oligonucleotide probes

Labelled probes from oligonucleotides were made by end-labelling. Approximately 50ng of oligonucleotide were labelled in a total volume of 20µl containing 2µl of 10× polynucleotide kinase (PNK) buffer, 10 units of PNK (Roche), and 30mCi [γ-^{32}P]-dATP. The reaction was incubated at 37°C for 1 hour and then added directly to the pre-hybridised blot.

Sequences of oligonucleotide probes (5'-3'):

AseI-MluI bridging probe (section A1.1)	ATTAGTTATTACGCGTGTAAAA
pCI-as probe (section A1.2)	GTTAGTCGACCCAGAGAGCTCCTTAAGAG
lacZ probe (section A1.3)	GATTTCCTTACGCGAAATACG
Wild-type *lacZ* probe (Figure 5.11(ii))	CAGCGCGGCTGAAATCATC
R376X probe (Figure 5.11(ii))	TGCTGATTTGAGGCGTT
Cre probe (Figure 2.15, 2.19)	CAACGAGTGATGAGGTTC
loxP probe (Figure 3.10)	ATAACTTCGTATAGCATACATTATACGAAGTTAT

7.1.10.3 Hybridisation of radioactive DNA probes

Blots were pre-hybridised at 68°C for a minimum of 1 hour in rotating hybridisation bottles (Hybaid) with 20-50ml hybridisation solution (6× SSC; 10% (w/v) Dextran sulphate; 0.5% (w/v) SDS; 0.1% (w/v) BSA; 0.1% (w/v) Ficoll; 0.1% (w/v) polyvinylpyrrolidone; 0.1% (w/v) disodium pyrophosphate). After pre-hybridisation, radiolabelled probe was added directly to the solution and the blots hybridised overnight at 68°C. Blots were washed with 2× SSC; 0.1% (w/v) SDS at 68°C for 15 minutes with monitoring of background radiation using a Geiger counter. The blots were wrapped in saran wrap and exposed to a Molecular Dynamics Phosphorimager screen. The image was then scanned into the Storm phosphorimager (Molecular Dynamics) and analysed using ImageQuant software.

7.1.10.4 Hybridisation of radioactive oligonucleotide probes

Blots were pre-hybridised at 48°C for a minimum of 1 hour in rotating hybridisation bottles with 20-50ml oligo-hybridisation solution (6× SSC; 0.5% (w/v) SDS; 0.1% (w/v) BSA; 0.1% (w/v) Ficoll; 0.1% (w/v) polyvinylpyrrolidone; 0.1% (w/v) disodium pyrophosphate) and 1mg sonicated salmon sperm. After pre-hybridisation, radiolabelled oligo probe was added directly to the solution and the blots hybridised overnight at 48°C. Blots were washed with 4× SSC; 0.1% (w/v) SDS for 15 minutes followed by a second wash of 2× SSC; 0.1% (w/v) SDS if required, with monitoring of background radiation using a Geiger counter. The blots were exposed to Kodak X-OMAT AR film at –70°C in cassettes containing intensifying screens. To strip the blot of the existing probe to allow hybridisation with a second probe, it was washed in 5× SSC; 1% SDS (w/v) at temperatures up to 80°C. The loss of the probe was monitored using a Geiger counter, and checked by exposing to film. Re-probing was performed as for the initial probe.

7.2. Polymerase chain reaction (PCR)

7.2.1 Reagents

<u>dNTPs</u>

dNTPs were purchased from ABgene as stocks of 100mM. Working stocks of 25mM were made by mixing 25µl of each of the dNTPs (dATP, dCTP, dGTP, dTTP) to a final volume of 100µl and stocks were stored at –20°C. dNTPs were used in PCRs at a final concentration of 0.2mM.

<u>Oligonucleotide primers</u>

PCR primers were designed by selecting a sequence, usually between 17 and 25 nucleotides in length, for which the [G+C]:[A+T] ratio was approximately equal. Care was taken to avoid the likelihood of primers self-annealing or annealing with each other. Primers were purchased from MWG Biotech as lyophilised desalted

compounds. Stocks were made up at 100μM using sterile dH_2O and primers were used in PCRs at a final concentration of 0.5-1μM.

Additional reagents

Ampli*Taq* DNA polymerase (5units/μl) (Roche) was used at 0.1μl per 25μl reaction. PCR buffer (Roche) was a 10× stock and therefore diluted 1:10 for reactions. Mg^{2+} (Roche) was used at a final concentration of 1.5mM. The reaction was made up to 25μl with sterile dH_2O.

7.2.2 Amplification programmes

PCRs were performed in 0.5ml centrifuge tubes or 96-well PCR plates in a MJ Research DNA Engine Tetrad. Throughout this chapter, PCR thermal cycles are represented by formulas resembling the following:

1×(94°C[7m]) – 30×(94°C[1m]57°C[1m]72°C[1m]) – 1×(72°C[10m])

The formula shown describes a reaction comprising a single cycle of 7 minutes at 94°C, followed by 30 cycles of 1 minute at 94°C, 1 minute at 57°C, 1 minute at 72°C, and finishing with 1 cycle of 10 minutes at 72°C.

Sequences of primers and amplification programmes used in non-genotyping PCRs (5'-3'):

Cre PCR probe (Figure 2.18)	Cre forward	CAATTTACTGACCGTACAC
	Cre reverse	CAGATTACGTATATCCTGG
	Programme	1×(94°C[7m]) – 30×(94°C[1m]57°C[1m]72°C[1m]) – 1×(72°C[10m])
R26R excision PCR (Figure 3.10)	LacZ4 forward	GTGCGGGCCTCTTCGCTAT
	5921R reverse	GGTTGAGGACAAACTCTTCGC
	Programme	1×(94°C[5m]) – 40×(94°C[40s]57°C[40s]72°C[1m30s]) – 1×(68°C[7m])
Transgene integrity- and contamination-PCRs (Figure 2.3) and (Figure 5.5)	5'neg forward	CCATAGCCTCAGGTTACTC
	5'pos forward	GGCGATAAGTCGTGTCTTAC
	5'R26 reverse	CACCATTGCACCGCTCTTGC
	5'K5 reverse	GCTGACAGGTGACTGTGATG
	3'neg reverse	CTACTTCTGGAATAGCTCAG
	3'pos reverse	GGTGATGGTTCACGTAGTG
	3'eGFP forward	CTGGAGTACAACTACAACAG
	Programme	1×(94°C[7m]) – 30×(94°C[1m]55°C[1m]72°C[1m30s]) – 1×(72°C[10m])
lacZ (R376X) PCR-RFLP (Figure 5.5)	S3 forward	CAGCGATTTCCATGTTGCCAC
	AS5 reverse	GGAAGGGCTGGTCTTCATC
	Programme	1×(94°C[7m]) – 30×(94°C[1m]57°C[1m]72°C[1m]) – 1×(72°C[10m])

7.3 DNA sequencing

Plasmid DNA used for sequencing was prepared using a Miniprep or Maxiprep Kit (both Qiagen), according to the manufacturer's instructions. Samples were cleaned using the PCR Purification Kit or Nucleotide Removal Kit (both Qiagen). DNA was sequenced using dye-labelled terminators for cycle sequencing. The dye used was either dRhodamine or Big Dye Terminator RR mix. Reagents were thawed on ice and dyes were protected from light as much as possible. Reactions were set up in 0.2ml centrifuge tubes as follows:

200-500ng plasmid DNA in dH_2O	11µl
dRhodamine or Big Dye Terminator RR mix	8µl
primer (3.2 pmoles)	1µl
	20µl

Cycle sequencing was performed using an MJ Research DNA Engine Tetrad. The sequencing program was as follows: 24×(96°C[30s]50°C[15s]60°C[4m]). Reactions were ethanol precipitated following transfer to a fresh 1.5ml centrifuge tube containing 50µl ethanol and 3M 2µl sodium acetate, pH 5.2. Reactions were left on ice to precipitate, then centrifuged at 13,000g at 4°C for 30 minutes. Pellets were washed with 200µl 70% ethanol. The supernatant was removed following a second centrifugation step and pellets were allowed to air dry at room temperature for approximately 20 minutes. Samples were submitted to the sequencing service to be run on an ABI machine.

Sequences of sequencing primers (5'-3'):

R26 promoter (section A2.1)	T3	AATTAACCCTCACTAAAGGG
	T7	GTAATACGACTCACTATAGGGC
K5 promoter (section A1.1)	K5 5'	CATAGAGGAACAGGCTGAG
	K5 3'	GTGCTCAGCATGTACCGTAC
pCI intron (section A1.2)	pCI-s	GGAACTCGAGCTTCTGCGTCCTGCTGAGC
	pCI-as	GTTAGTCGACCCAGAGAGCTCCTTAAGAG
lacZ/CreER^{T2} (section A1.3, 1.4)	3'check	CATATAGACAAACGCACAC
R376X check (section 5.2.1)	R376X-s	CACACCGCCGACGGCACGC
	R376X-as	GATGGTTCGGATAATGCG
lacZ (R376X) sequencing primers (section 5.2.1)	S1	CTAAAAGCTGCGGAATTGTAC
	S2	CAAACTGGCAGATGCACGG
	S3	CAGCGATTTCCATGTTGCCAC
	S4	CCGTCACGAGCATCATCCTC
	S5	CTGGATCAAATCTGTCGATC
	S6	GACCGCACGCCGCATCCAG
	S7	CATCCCGCATCTGACCACCAG
	S8	GAAGTGGCGAGCGATACAC
	AS1	GATTTCCTTACGCGAAATACG
	AS2	CTAATCCGAGCCAGTTTAC
	AS3	CCAATCCACATCTGTGAAAG
	AS4	GTATTCGCTGGTCACTTCGATG
	AS5	GGAAGGGCTGGTCTTCATC
	AS6	CATCGTCTGCTCATCCATGAC
	AS7	CAGCGCGGCTGAAATCATC
	AS8	GGTTACGTTGGTGTAGATG

7.4 Site-directed mutagenesis

Site-directed mutagenesis was achieved using the QuikChange Site-Directed Mutagenesis Kit (Stratagene), following the manufacturer's instructions. Briefly,

125ng of both mutagenic primers (shown below, mutation indicated in bold) were incubated with 50ng parental plasmid DNA, 5ng 10× supplied reaction buffer, 1µl supplied dNTP mix, 1µl supplied *PfuTurbo* DNA polymerase, and dH_2O to a final volume of 50µl. The reaction mixture was subjected to the following thermal cycle using a MJ Research DNA Engine Tetrad:

1×(95°C) – 16×(95°C[30s]55°C[1m]68°C[15m])

After completion of the thermal cycle, 1µl of DpnI enzyme was added and the reaction incubated at 37°C for 1 hour. 1µl of the reaction was transformed into supplied Epicurian Coli XL1-Blue competent cells, and mutated clones selected on xgal/IPTG indicator plates.

Primers for introducing R376X mutation into *lacZ* in pCMVβ (Clontech) (5'-3'):

(section 5.2.1)	mms	GCAAGCCGTTGCTGATTTGAGGCGTTAACCGTCACG
	mmas	CGTGACGGTTAACGCCTC**A**AATCAGCAACGGCTTGC

7.5 Microbiology

Aseptic technique was observed for all steps involving the growth of bacterial cells (setting up cultures, pouring agar plates, selecting single colonies and storing bacterial stocks). Liquid cultures were grown with vigorous shaking (220rpm) and dry cultures were grown on inverted agar plates at 37°C overnight.

7.5.1 Growth media for bacterial cultures

Technical staff prepared Luria-Bertani broth (L-Broth) and Luria-Bertani agar (L-agar) as detailed below:

L-Broth

Amount per litre:	
Tryptone	10.0g
Yeast extract	5.0g
NaCl	10.0g
Glucose	1.0g

L-Broth with Mg^{2+}

L-Broth with 2.5g magnesium sulphate per litre.

L-Agar

Amount per litre:	
Tryptone	10.0g
Yeast extract	5.0g
NaCl	10.0g
Agar	15.0g

7.5.1.1 Antibiotic selection

	Stock concentration	Working concentration
Ampicillin (Amp)	10mg/ml in dH_2O	50μg/ml
Kanamycin (Kan)	10mg/ml in dH_2O	50μg/ml

Aliquots were stored at -20°C.

7.5.1.2 Xgal/IPTG indicator plates

When attempting to select clones carrying mutated (R376X) pCMVβ plasmids (chapter 5, section 5.2.1), xgal (Melford) was added to the molten L-agar at a final concentration of 40μg/ml whilst IPTG (Melford) was added at a final concentration of 0.2mM, before pouring plates.

7.5.2 Production of electrocompetent cells

A single colony of DH5α *E. coli* cells from an agar plate was used to inoculate approximately 10ml L-Broth for overnight growth at 37°C. The overnight culture

was used to inoculate 2× 400ml fresh L-Broth the following morning and cells were grown to an OD_{600} of 0.5-1.0. Flasks were chilled on ice for 15-30 minutes and cells were centrifuged at 4°C, 4,000g (6,000rpm) for 15 minutes. Pellets were re-suspended in 800ml ice-cold sterile dH_2O. The centrifugation step was repeated and cells were re-suspended in 400ml chilled, sterile dH_2O. Following a further centrifugation step, cells were re-suspended in 20ml 10% cold, sterile glycerol. Cells were centrifuged once more and re-suspended in 2-3ml 10% glycerol. Aliquots of cells were frozen on dry ice and stored at -70°C.

7.5.3 Transformations

Competent cells were transformed with 10-100ng plasmid DNA by electroporation. Electrocompetent cells were thawed on ice and 50µl was added to each DNA sample in an ice cold centrifuge tube. The mixture was transferred to ice cold 0.2cm electroporation cuvettes (EquiBio) and allowed to sit for 1 minute on ice. Cells were electroporated using a BioRad Gene Pulser set at 25µF, 2.5kV, and 200Ω. Immediately after electroporation, 1ml L-Broth/Mg^{2+} was added to the cells, which were then allowed to recover for 1 hour at 37°C, shaking, prior to plating on selective media and overnight incubation at 37°C. Transformants carrying cloned insert were identified by restriction digestion or PCR analysis of plasmid DNA preparations. When low numbers of such transformants were anticipated, many colonies were screened by colony hybridisation.

7.5.4 Preparation of bacterial colony hybridisation filters

Transformation colonies were transferred using sterile toothpicks to the equivalent positions on two agar plates, and both plates were incubated at 37°C. After storage of one of these plates at 4°C for 5 minutes, a circular nitrocellulose filter (0.45µm, Schleicher and Schuell) was placed onto the agar plate for 1 minute, and marked for orientation. The filter was transferred for 5 minutes to each of a series of 3mm filter papers (Whatman) soaked in, respectively, 10% SDS, denature buffer, and

neutraliser buffer. The DNA was UV cross-linked using the auto cross-link feature on a Stratalinker 1800 (Stratagene). Filters were probed using oligonucleotide probes (section 7.1.10.4). Useful transformants were then picked from the duplicate agar plate and DNA prepared.

7.5.5 Isolation of plasmid DNA

Plasmid DNA was prepared using a QIAquick Spin Miniprep Kit (Qiagen). For the extraction of small amounts of plasmid DNA (miniprep), a single colony was selected to inoculate 2ml L-Broth containing a suitable antibiotic for overnight growth at 37°C. The following morning, plasmid DNA was extracted using the kit according to the manufacturer's instructions. Plasmid DNA was eluted in either 30µl or 50µl elution buffer. The DNA concentration was determined either by agarose gel electrophoresis or by spectrophotometry. For the extraction of larger amounts of plasmid DNA (maxiprep), a 2ml starter culture was established as described above and incubated at 37°C for 8 hours. The starter culture was then re-suspended in 400ml L-Broth, with antibiotic selection, in a large conical flask to allow aeration of the bacteria and the larger volume culture was incubated at 37°C overnight. Plasmid DNA was extracted the following morning using the Qiagen Maxi Kit according to the manufacturer's instructions. DNA was eluted in 400µl autoclaved TE, pH8.0, and the concentration determined as for miniprep DNA.

7.6 Production of transgenic animals

7.6.1 Preparation of recombinant DNA for microinjection

10µg plasmid DNA was digested with enzymes (BspHI for pR26X, pR26Z, pK5A1 and SfiI for pK5CE) to release or linearise the transgene, respectively. A sample of DNA was run on an agarose gel alongside an equivalent amount of undigested DNA to ensure the DNA was completely digested and products were the predicted size. Once the transgene had been released/linearised, the remainder of the reaction was

run on an agarose gel in TAE buffer. The transgene fragment was excised from the gel, and the DNA extracted using a QIAEXII Agarose Gel Extraction Kit (Qiagen). DNA was eluted in 60µl microinjection buffer (0.1mM EDTA/10mM Tris pH7.4) and the concentration determined by electrophoresis. DNA was stored at –20°C until the day of microinjection. For microinjection, DNA was diluted in microinjection buffer to a final concentration on 2ng/µl and spun through a Spin-X 0.22µm column (Corning).

7.6.2 Reagents and equipment used during the manipulation of preimplantation embryos

7.6.2.1 Equipment

Embryos were transferred between dishes using pulled pasteur pipettes and mouth pipette tubing. A Zeiss Stemi SV11 dissecting microscope was used to aid in manipulating embryos in sterile glass staining blocks and petri dishes.

Microinjection apparatus

The microinjection apparatus consisted of a Zeiss Axiovert 100 microscope with a variable temperature stage set to 37°C, and mounted holders for the microinjection and holding pipettes (Narashige, Eppendorf). The equipment rested on an anti-vibration table (Carl Zeiss). Embryo holding pipettes were either made in-house using a microforge (MF90) (Narashige) and capillaries (size GC100 T15, Harvard Apparatus) or purchased (Eppendorf). Microinjection pipettes were purchased (Eppendorf). Microinjection needles were filled with DNA at a concentration of 2ng/µl using pipette filler tips (Eppendorf). DNA was microinjected with the aid of an automated microinjector (Narashige IM 300) to allow a constant flow of DNA at variable pressure.

7.6.2.2 Solutions

Embryo growth media

H6 media was used for handling embryos outside of the incubator. T6 media was used for growing embryos in the incubator (5% CO_2, 37°C). Media was either purchased (Sigma) or prepared in-house by Transgenic Facility staff using embryo culture grade reagents. Media was made using embryo culture grade dH_2O, filter sterilised, and embryo tested prior to use.

Compound	**T6 (g/100ml)**	**H6 (g/100ml)**
NaCl	0.472	0.472
KCl	0.011	0.011
NaH_2PO_4	0.0047	0.0047
Sodium pyruvate	0.003	0.003
Glucose	0.1	0.1
$MgCl_26H_2O$	0.01	0.01
$CaCl_22H_2O$	0.026	0.026
Penicillin	0.006	0.006
Streptomycin	0.005	0.005
$NaHCO_3$	0.2106	0.2106
Sodium lactate	0.34ml	0.34ml
Hepes	-	0.50
Phenol red	0.001	0.001
BSA	0.4	0.4

7.6.3 Microinjection of recombinant DNA into fertilised eggs

Embryos used for microinjection were derived from [CBA × C57BL/6] F1 × [CBA × C57BL/6] F1 matings. [CBA × C57BL/6] F1 females were superovulated by intraperitoneal (IP) injection of 5 units of pregnant mare serum (PMS) (Intervet) at noon on the first day, followed by 5 units of human chorionic gonadotrophin (hCG) (Intervet) 48 hours later. They were then mated to F1 males. Plugged females were sacrificed the following morning (E0.5) and oviducts were removed and rinsed in warm saline. Oviducts were placed in H6 media and fertilised eggs were removed by tearing the swollen ampullae with fine forceps. Several drops of hyaluronidase (1mg/ml PBS) (Sigma) were added to the media for approximately three minutes to remove the cumulus cells from the outside of the eggs. Fertilised eggs were identified by the presence of two large pronuclei and often polar bodies. Eggs awaiting microinjection were kept at 37°C, 5% CO_2 in drops of T6 media under

paraffin oil in sterile petri dishes. Eggs were microinjected in batches of 30-50 in a large drop of H6 media under oil in a glass dimple slide (the injection dish).

During microinjection, embryos were picked up and held on the holding pipette using mouth pipette tubing. The microinjection needle on the automatic injector was set such that a constant stream of DNA left the pipette. Once a pronucleus was in focus, the injection needle was brought into the same plane of focus and used to pierce the *zona pellucida* and egg cell membrane, before entering the pronucleus. The needle was held until the pronucleus swelled slightly, and then removed quickly and cleanly. Following microinjection, eggs were returned to drops of T6 media under paraffin oil and kept at 37°C, 5% CO_2 overnight. For each experiment, five to ten uninjected embryos were retained and cultured overnight to control for problems with media, oil, and incubation conditions.

7.6.4 Oviducal transfers

Microinjected embryos were screened the following morning. Only embryos that had developed to the two-cell stage were transferred into recipient females. Pseudopregnant recipients were prepared by selecting CD1 females in oestrous and mating them to vasectomised CD1 males. Plugged E0.5 females were used for oviducal transfers. Recipients were anaesthetised with an IP injection of anaesthetic at 0.1ml/10g per mouse: [0.75ml of Hypnorm (Fentanyl 0.315mg/ml / fluanisone10mg/ml): 4.5ml water: 0.2ml Hypnovel (Midazolom 10mg/2ml)]. The fat pad adjacent to the ovary and uterus was isolated and held in autoclips (Brownes Diagnostics). The bursa surrounding the oviduct was broken carefully to expose the infundibulum of the oviduct. The end of the pulled pipette containing the embryos was inserted into the infundibulum and embryos were delivered by mouth pipette. The successful delivery of the embryos was confirmed when bubbles that had been placed at either end of the embryos in the pipette were seen in the swollen ampulla. Ten to 20 embryos were transferred to each side of the recipient. The fat pad was released and tucked back into the abdominal cavity and the peritoneum was closed with sutures (HM&S). The skin was closed using Dieffenbachs bulldog clips

(Holborn surgical and medical instruments, ltd.), and analgesia was administered at 0.1ml/10g mouse of [10µl of Vetergesic (Buprenorphine 0.3mg/ml): 990µl dH_2O]. The animal was allowed to recover in a warmed post-operative cage with solid drink pouches (BS&S). Embryos were allowed to develop to term.

7.6.5 Genotyping of breeding mice

Breeding animals were genotyped by PCR amplification of transgenic sequences from genomic DNA extracted from ear-clips. The DNA was extracted from each ear clip by incubating the tissue in 50µl of 25mM NaOH, 0.2mM EDTA (pH12) at 95°C for 20 minutes, followed by vortexing briefly to break up the tissue. The reactions were neutralised by the addition of 50µl of 40mM Tris.HCl (pH5). 2µl of extracted DNA solution was used in 25µl genotyping PCR assays. Extracted genomic DNA was stored at -20°C.

Sequences of primers and amplification programmes used in genotyping PCRs (5'-3'):

R26Z, *R26X*,	LacZF forward	GTGACTACCTACGGGTAACA
K5A1, *R26R*	LacZR reverse	ATTCATTGGCACCATGCCGT
(Figure 2.3)	MyoG forward	TTACGTCCATCGTGGACAGC
and	MyoG reverse	TGGGCTGGGTGTTAGTCTTA
(Figure 3.3)	Programme	1×(94°C[7m]) – 30×(94°C[1m]58°C[1m]72°C[1m]) – 1×(72°C[10m])
K5CE	Cre forward	CAATTTACTGACCGTACAC
(Figure 2.13)	Cre reverse	AGCTGGCCCAAATGTTGCTG
	Programme	1×(94°C[7m]) – 30×(94°C[1m]57°C[1m]72°C[1m]) – 1×(72°C[10m])
R26R wild	R523 forward	GGAGCGGGAGAAATGGATATG
-type allele	R26F2 reverse	AAAGTCGCTCTGAGTTGTTAT
(Figure 3.3)	Programme	40×(94°C[30s]58°C[30s]65°C[1m])

7.6.6 Animal husbandry

All animals were maintained in a specific pathogen free (SPF) environment and experiments were carried out under Home Office licence. Wild-type animals (CD1, C57BL/6, and CBA x C57BL/6 F1) were either bred in-house or obtained from Charles River Laboratories. Embryos for all experiments were generated from timed

matings, with the morning of vaginal plug detection being counted as embryonic day 0.5 (E0.5).

7.6.7 Harvesting of postimplantation embryos

Postimplantation embryos were harvested for wholemount xgal staining and sectioning for immunodetection. The mothers were sacrificed by cervical dislocation. The abdominal cavity was opened and uteri were removed to petri dishes containing PBS. Embryos were removed from the uteri and freed from extraembryonic membranes using scissors and forceps. For some experiments, extraembryonic membranes were retained to genotype embryos by PCR. Embryos were rinsed in fresh PBS prior to subsequent processing.

7.6.8 Genotyping of embryos

In some experiments, embryos were genotyped using DNA extracted from yolksacs. These were digested overnight at 55°C in 0.5ml yolksac lysis buffer containing proteinase K.

Yolksac lysis buffer

	Final concentration
Tris.HCl pH8.0	100mM
EDTA	50mM
NaCl	100mM
SDS	1%

Proteinase K (stock 10mg/ml) (Sigma) was added to a final concentration of 0.2mg/ml.

Following overnight digestion, tubes were centrifuged at 13,000rpm for 10 minutes to pellet debris. The supernatant was transferred to a clean tube, to which 1ml chilled 70% acetone/5% DMF (in 25% dH_2O) was added. The solution was mixed manually by inversion of tubes and DNA pelleted by centrifugation for 10 minutes at 13,000rpm. The supernatant was then carefully decanted and DNA washed in 70%

ethanol on a rotator for at least 30 minutes. Tubes were then centrifuged for 2 minutes at 13,000rpm, the supernatant discarded, and following air drying of samples, DNA was re-suspended in 100-500µl TE, pH8.0 depending upon the size of the pellet. Incubation at 55°C for 15 minutes aided re-suspension of the DNA.

Alternatively, DNA was extracted from small pieces of amputated limb or tail tissue. Unique combinations of tail and limb amputation allowed embryos to be identified, and phenotypes could be correlated with genotypes. DNA was extracted from amputated tissue using the ear-clip digestion method used to genotype adult mice (section 7.6.5).

For both methods, 1µl of DNA was used in genotyping PCR reactions.

7.7 Wholemount expression analysis

7.7.1 Visualisation of eGFP

eGFP was visualised in fresh, unfixed embryos and tissues using a stereo fluorescence microscope with Leica GFP1 and GFP2 filters (section 7.7.4)

7.7.2 Fixing tissue

Paraformaldehyde

4% paraformaldehyde (PFA) (Sigma), used to fix embryos and tissues, was made by dissolving 20g PFA in 500ml PBS at 65°C until dissolved. Aliquots were stored at –20°C.

Embryos were dissected in PBS and fixed in 4% PFA at 4°C for between 1.5 and 3 hours depending upon age (E13.5-E16.5). Tissues were rinsed in PBS before fixing in 4% PFA, and large tissues were cut into smaller pieces to aid penetration of

fixative. Duration of tissue fixation was dictated by sample size, but never exceeded 3 hours. Tracheas were routinely fixed for 1 hour.

7.7.3 Xgal staining

Following fixation, embryos and tissues were first rinsed in PBS and then washed in detergent wash 3 times for 30 minutes each at room temperature. Samples were stained overnight (approximately 15 hours) in xgal staining solution, protected from light, at room temperature. Ear tissue from *ROSA26* and non-transgenic CD1 mice served as positive and negative staining controls, respectively. The following morning the staining solution was removed and samples were rinsed in detergent wash twice for 20 minutes at room temperature. In certain experiments, samples were fixed again in 4% PFA at 4°C overnight. The fix was removed and tissues were washed, and stored, in PBS. In other experiments, samples were immediately embedded in wax.

7.7.3.1 Solutions

Detergent wash (in PBS)

Compound	Per 500ml	Final concentration
$MgCl_2$ (Riedel-de Haen)	0.406g	2mM
Nonidet P-40 (ICN Biomedicals Inc.)	200µl	0.02%

Xgal staining solution (in PBS)

Compound	Per 500ml	Final concentration
$K_3Fe(CN)_6$ (Sigma)	820mg	5mM
$K_4Fe(CN)_6$ (Sigma)	1050mg	5mM
$MgCl_2$ (Riedel-de Haen)	0.406g	2mM
Sodium deoxycholate (Sigma)	1g	0.1%
Nonidet P-40 (ICN Biomedicals Inc.)	200µl	0.02%

Prior to use, 250µl xgal substrate (40mg/ml in dimethylformamide, DMF) was added per 10ml staining solution required. In some experiments an alternative βgal

substrate, bluogal (Sigma), was used. It was prepared and used exactly as xgal. Xgal staining solution was pH7.4.

7.7.4 Wholemount microscopy

Dissections of post-implantation embryos were effected with the aid of a Leica Stereo MZFLIII stereo fluorescence microscope (Leica Microsystems, Milton Keynes, UK) fitted with a 100W Hg source for fluorescence imaging, and fibre optic cold light source and illuminated base for incident/transmission illumination. Images of wholemount stained embryos and tissues were captured with a Photometrics CoolSnap colour CCD camera (Roper Scientific, Tucson, Arizona) controlled by scripts written for IPLab Spectrum (Scanalytics Inc., Fairfax, VA). Prior to imaging, wholemount intact stained tracheas were cut dorsally and pinned out onto a wax bed under PBS to permit examination of the luminal surface. Embryos were oriented within wells cut into agarose-filled petri dishes, and imaged under PBS.

7.8 Histology

7.8.1 Tissue processing for paraffin wax sectioning

For histochemisty, immunohistochemistry, and immunofluorescence, tissues and embryos were processed for paraffin wax embedding, and microtome sectioned. Tissue processing was automated using a Tissue-tek VIP, using the following settings:

Stage	Temperature (°C)
70% ethanol/PBS	40
85% ethanol/PBS	40
95% ethanol/PBS	40
100% ethanol/PBS	40
100% ethanol/PBS	40
xylene	40
xylene	40
paraffin wax	60
paraffin wax	60
paraffin wax	60
paraffin wax	60

Each stage was of equal duration, but durations varied between experiments from 20 minutes to 2.5 hours, depending on sample size. Following the final wax stage, samples were transferred into a suitable mould, oriented using a pair of forceps, and the wax allowed to solidify. The wax block was trimmed to the desired size and the edges angled to improve capture of the sections.

7.8.2 Microtome sectioning

Wax-embedded samples were sectioned at 7-10μm on a microtome and ribbons of sections floated out in a 42°C waterbath. With the aid of a fine-tipped paintbrush, the sections were transferred to Superfrost Plus slides (BDH) and the slides incubated at 55°C overnight to allow sections to dry. To remove the wax, slides were washed 3 times in fresh xylene for 10 minutes each, followed by 3× 10 minute washes in 100% ethanol. Slides were then passed through a graded ethanol series, washed in dH_2O, and stained with haematoxylin and/or eosin, with staining performed according to a standard protocol. Typically, slides were immersed in eosin for two minutes, and haematoxylin for 10 seconds. Tissue sections were mounted in DePeX (BDH) or Histomount (National Diagnostics) mounting medium.

7.8.3 Immunohistochemistry and immunofluorescence

Typically, 7μm sections were used for immunodetection. After dewaxing, sections were microwaved on full power in 10mM citrate buffer (Na_3Cit, pH 6.0) for 20 minutes, allowed to cool, and washed in dH_2O. They were then washed in PBS containing 0.1% Tween 20 (PBST), and blocked in PBST with 10% normal goat serum (NGS) (Scottish Antibody Production Unit) for 1 hour.

7.8.3.1 Anti-Cre immunofluorescence

For immunofluorescence analysis of ovaries, *bK5Cre* females were superovulated with 5 units PMS, followed 48 hrs later with 5 units hCG. Ovaries were removed after a further 12 hours (prior to ovulation), processed, and sectioned. The sections were then incubated with the primary rabbit polyclonal anti-Cre antibody (Novagen) at 1:1000 dilution for 1.5 hours, washed in PBST for 30 minutes, and incubated with the secondary Alexa 594 fluorescent antibody (Molecular Probes Europe BV) at 1:400 dilution in darkness for 1 hour. Finally, slides were washed in PBST for 30 minutes in darkness and mounted in Vectashield (Vector Laboratories Ltd.), containing 1μg/ml DAPI. The Alexa 594 fluorescence was visualised using a compatible Texas Red filter set (absorbance and emission maxima of 595nm/615nm, respectively).

7.8.3.2 Anti-βgal immunofluorescence

Sections were incubated with the primary rabbit polyclonal anti-βgal antibody (Europa Bioproducts Ltd) at 1:2000 dilution for 1.5 hours, washed in PBST for 30 minutes, and incubated with the secondary Alexa 594 fluorescent antibody (Molecular Probes Europe BV) at 1:400 dilution in darkness for 1 hour. Finally, slides were washed in PBST for 30 minutes in darkness and mounted in Vectashield (Vector Laboratories Ltd), containing 1μg/ml DAPI. Signal was visualised as for the anti-Cre antibody (section 7.8.3.1).

7.8.3.3 Anti-K14 immunohistochemistry

Prior to antigen retrieval, slides were dewaxed to 95% ethanol, then washed in 1% H_2O_2 in methanol for 20 minutes to block peroxidase activity. After washing in dH_2O for 5 min, antigen retrieval and blocking was performed as described in section 7.8.3. Sections were incubated with the primary affinity purified rabbit anti-mouse anti-K14 antibody (courtesy of S. Randell) at 1:40 dilution (2.5μg/ml) at 4°C overnight. Slides were washed in PBST for 30 minutes then incubated with the secondary biotinylated swine anti-rabbit antibody (DAKO) at 1:400 for 30 minutes, before washing with PBST.

Immunostaining was visualised by immunoperoxidase assay. Vectastain ABC complex (Vector Laboratories Ltd) was prepared 30 minutes prior to use according to manufacturer's instructions. ABC complex was added for 30 minutes, before washing with PBST. The slides were then incubated with DAB substrate (Vector Laboratories Ltd) for 10 minutes, and washed in running water. Slides were counterstained with haematoxylin and mounted as described in section 7.8.2.

7.8.4 Microscopic analysis of tissue sections

Tissue sections were analysed for brightfield (using DIC optics) and fluorescence using a Zeiss Axioplan II fluorescence microscope (Carl Zeiss Ltd., Welwyn Garden City, UK) equipped with colour additive filters (Andover Corp, Salem, NH) for sequential colour imaging and a Photometrics CoolSnap HQ monochrome CCD camera (Roper Scientific, Tucson, Arizona). Image capture was by scripts written for IPLab Spectrum (Scanalytics Inc., Fairfax, VA) which controlled camera capture and filter selection via motorised filter wheels (Ludl Electronic Products Ltd., Hawthorne, NY).

7.9 Analysis of mRNA expression

7.9.1 RNA extraction

Extraction of RNA was performed prior to expression analysis of various transcripts by reverse transcription (RT)-PCR. During this procedure, extreme care was taken to avoid contamination of samples with RNases. Filter pipette tips were used for all RNA/cDNA work. All solutions were made with DEPC-dH_2O, separate reagents reserved for RNA work only, and surfaces of benches and glassware treated with RNaseZAP (Sigma) to inactivate RNases. DEPC-dH_2O was made by adding 1µl/ml DEPC (Sigma) to dH_2O, followed by gentle mixing using a magnetic stirrer for 1-2 hours, and solutions were allowed to sit at room temperature in a fumehood overnight prior to re-autoclaving. Reagents were kept on ice until required. RNA was extracted as quickly as possible after obtaining samples in order to minimise the degradation of crude RNA by limiting the activity of endogenous RNases.

Tissues (trachea, depilated back skin, lung, liver) were harvested into RNAlater (Ambion) and stored at –20°C. Total RNA was extracted by homogenisation using a rotor-stator homogeniser (Omni) fitted with disposable Omni-Tips (Fisher Scientific). The homogenate was processed using an RNeasy Mini Kit (Qiagen), using on-column DNase treatment (RNase-Free DNase Kit, Qiagen) and eluted in Rnase-free dH_2O. Eluates were further treated with amplification grade DNase I (Sigma) according to the manufacturer's instructions. RNA samples were electrophoresed for visual inspection of RNA quality. The concentration of RNA was determined by spectrophotometry as for DNA (section 7.1.3) but calculated as follows:

Concentration (mg/ml) = $A_{260} \times 40$ (dilution factor) $\times$ 40.

7.9.2 cDNA synthesis

Using a First Strand cDNA Synthesis Kit (Roche) with random primers, 1μg of RNA was reverse transcribed to cDNA according to the manufacturer's instructions. Reactions were consistently performed in duplicate, with one set of reactions lacking reverse transcriptase enzyme as a control to confirm the absence of contaminating genomic DNA. Following the cDNA synthesis reaction, 2μl of sample was used in PCRs as described in section 7.2.

Sequences of primers and amplification programmes used in RT-PCRs (*Cre*-specific RT-PCRs used *Cre*-specific *K5CE* genotyping PCR primers at either 30 or 45 cycles for skin or tracheal RNA, respectively) (5'-3'):

R26X	MutS forward	GATGAGCGTGGTGGTTATG
(Figure 5.11(i))	MutAS reverse	GTTACGCGTTCGCTCATC
	Programme	1×(95°C[5m]) – 30×(94°C[30s]55°C[30s]72°C[1m)
GAPDH	GAPDH forward	TGGAGCCAAACGGGTCATCA
(Figure 2.15, 2.19)	GAPDH reverse	GTGGCAGTGATGGCATGGAC
(Figure 5.11(i))	Programme	1×(94°C[3m]55°C[1m]72°C[1m])–29×(93°C[30s]55°C[1m]72°C[1m])

7.10 Induced injury to tracheal epithelium

7.10.1 Preparation and administration of polidocanol

Poloxethylene-9-lauryl-ether (polidocanol, POL) (Sigma) was weighed out and added to sterile PBS to make a 2% (w/v) solution. The mixture was heated at 37°C to aid dissolution.

Mice were anaesthetised with an IP injection of anaesthetic at 0.1ml/10g per mouse: [0.75ml of Hypnorm (Fentanyl 0.315mg/ml / fluanisone10mg/ml): 4.5ml water: 0.2ml Hypnovel (Midazolom 10mg/2ml)]. Mice were placed in a tracheal instillation frame as described by Ho and Furst (Ho and Furst, 1973), shown in Figure 3.4. The anaesthetised mouse was placed with its upper teeth hooked over a wire at the top of the support frame, with a spring pushing the thorax forward to position the pharynx,

larynx and trachea in a vertical straight line. The lower jaw was held open with a loop of thread around the lower teeth and weighted. The airway was illuminated using a fibre optic lamp placed against the mid sternum. The tongue was moved to one side and, using a cut-off microloader pipette tip (Eppendorf), 10µl of POL or PBS alone was slowly instilled into the proximal region of the trachea. The animal was maintained in an upright position for 1 minute following instillation to allow the solution to spread down the trachea. Following subsequent administration of tamoxifen or oil control (section 7.1.11.2), analgesia was administered (0.1ml/10g per mouse) by IP injection of [10µl of Vetergesic (Buprenorphine 0.3mg/ml): 990µl dH_2O]. Animals were allowed to recover in a warmed post-operative cage with solid drink pouches (BS&S).

7.11 Ligand-inducible Cre recombinase activity

7.11.1 Preparation and administration of tamoxifen

Tamoxifen (TAM) (free base form) was obtained from Sigma and ICN Biomedicals.

7.11.1.1 Ear-painting assay

For delivering 50µl volumes of 5mg and 10mg TAM solution to mouse pinna, 100mg/ml and 200mg/ml solutions of TAM in dimethyl sulphoxide (DMSO), respectively, were prepared. TAM has a limited solubility in DMSO and formed a suspension after vortexing. The suspensions were stored, protected from light, in aliquots at –80°C. Prior to use, suspensions were agitated in a sonicating water bath, protected from light, for 15 minutes.

TAM was administered by dispensing the 50µl aliquot onto the dorsal surface of pinna of an animal restrained by Transgenic Facility staff. A cotton bud stripped of most of the cotton to reduce absorbency, was used to spread the suspension over both inner and outer surfaces of the pinna. Animals received 50µl of 100mg/ml and

200mg/ml solution to respective pinna, corresponding to 5mg and 10mg TAM. Control animals received 50µl DMSO alone to each pinna. The protocol was repeated each day for 5 days, with ear-clip biopsies taken for xgal staining on day 6.

7.1.11.2 Systemic administration

For systemic delivery, 60mg TAM was added to 600µl 100% analytical grade ethanol (BDH) and vortexed briefly, though complete dissolution was not obtained. 5.4ml corn oil (Sigma) was added to make a 10mg/ml solution. The mixture, in a 15ml polyethylene tube (Corning), was placed in a beaker of ice and subjected to repeated 5 minute sonications at high power using a sonicating probe (MSE). Complete dissolution was usually achieved after 3-4× 5 minutes. This procedure was repeated to produce fresh TAM solution for each day of injections.

4mg (400µl) of the solution, or the ethanol/oil mixture alone, was administered by IP injection for 3 consecutive days following polidocanol treatment (section 7.10).

7.12 *in vitro* validation of *R26X* and *R26Z* constructs

7.12.1 G-418-mediated translational readthrough assay

A 12-well tissue culture plate (Corning) was seeded with COS-7 cells (Gluzman, 1981) at a density of 10^5 cells per well. COS-7 cells were grown in Dulbecco's Modified Eagles Medium (DMEM) with 10% foetal calf serum, 2mM L-glutamine and penicillin (10 units/ml) and streptomycin (10ug/ml) (all GibcoBRL). The LID transfection reagent (Parkes *et al.*, 2002), a kind gift of S. Hart, was prepared previously by cyclising the P6 peptide by dissolving it to 0.1mg/ml in OptiMEM (GibcoBRL), pipetting the media into a conical flask, and leaving loosely covered at 4°C overnight. After filter sterilisation, the media was stored at –70°C or –20°C depending on the duration of storage. When the cells had reached an appropriate density, the transfection mixture was prepared: 1.5µl 1mg/ml lipofectin (Gibco),

80µl (8µg) of P6 peptide, and 2µg of pCMVR376X or pCMVβ (positive control) plasmid DNA were mixed per well, and allowed to complex at room temperature for approximately 1 hour. The complex was diluted with OptiMEM to 1µg plasmid DNA per 1ml OptiMEM, and added to cells after the removal of supplemented media. Cells were incubated with the transfection mixture for 4 hours, before supplemented media was replaced, containing 0, 0.05, 0.1 or 0.2mg/ml G-418 sulphate (Geneticin, GibcoBRL). Normal supplemented media was replaced after approximately 12 hours, and 48 hours later the cells were fixed in 2% formaldehyde/0.2% glutaraldehyde at 4°C for 5 minutes and stained with xgal overnight in a humidor at 37°C.

7.12.2 Evaluation of R26 promoter activity and inactivity of the R376X lacZ sequence

COS-7, A549 (Lieber *et al.*, 1976), HT1080 (Rasheed *et al.*, 1974), and Huh7 (Health Science Research Resources Bank, Osaka, Japan; #JCRB0403) were grown as for COS-7 (section 7.12.1), though HT1080 cells were supplemented with 1% non-essential amino acids (Invitrogen). Cells were seeded at 2×10^5 cells per well of a 6-well plate (Corning) and transfected with pR26X, pR26Z, or pCMVβ by scaling-up the LID transfection method described in section 7.12.1. After 48 hours, half of the samples were fixed in 2% formaldehyde/0.2% glutaraldehyde at 4°C for 5 minutes for xgal staining.

Duplicate transfections were fixed with 4% PFA for 30 minutes and viewed under PBS by fluorescence microscopy using eGFP filters.

7.13 Statistical analysis

The data regarding the frequency of xgal stained cells in tracheas in the experiment described in chapter 3 section 3.2.3 were analysed by standard analysis of variance (ANOVA) methods using MINITAB Release 11 for Windows (1996, Minitab, Inc,

State College, Pennsylvania). The main effects were "treatment" (POL or PBS), "timepoint" (+3d, +7d, +20d), the "position" of stained cells within the trachea (proximal or distal), and the variation in the individual "tracheas" themselves. The outcome variables were total number of cells, total number of clusters, and average number of cells per cluster. Preliminary analysis revealed that the variation between individual tracheas was in all cases significantly greater than the residual variation between replicates. Thus all further analyses were carried out on a simplified data format, in which observations were averaged over the two replicates (per treatment, timepoint, position, and tracheas).

7.14 Computational methods

All vector construction was simulated *in silico* using Clone Manager 6, version 6.00 (Science and Educational Software). This allowed the exact prediction of restriction fragment and PCR product sizes shown throughout this thesis. The Emboss suite of programs (download via *http://www.hgmp.mrc.ac.uk/Software/EMBOSS*) was also used for general sequence manipulation.

The bioinformatics programmes and resources referred to in this thesis are listed below, together with the relevant world wide web (www) link.

Ensembl	*http://www.ensembl.org*
NCBI LocusLink	*http://www.ncbi.nlm.nih.gov/LocusLink/*
BLAST	*http://www.ncbi.nlm.nih.gov/BLAST/*

Appendices

A1 Construction of transgenic lines *K5A1* and *K5CE*

The DNA manipulations required to generate the transgenic lines *K5A1* and *K5CE* are described. The cloning scheme is summarised schematically in chapter 2, Figure 2.2.

A1.1 Cloning the human K5 promoter into the promoter-less pIRES2-eGFP

The human K5 promoter was provided (courtesy of S. Randell) in plasmid pK5LH14, where it was positioned upstream of the *eGFP* gene, from which it was separated by a short intron. Sequence information for pK5LH14 or the K5 promoter was not available. It was desirable to obtain sequence data for the K5 promoter in order to predict restriction enzymes suitable for use in the subcloning of the promoter. The NCBI LocusLink facility was used to find a genomic bacterial artificial chromosome (BAC) clone (GenBank accession AC055736) corresponding to the proposed K5 locus at chromosome 12q12-13. The 150kb sequence was analysed for restriction enzyme patterns resembling those known approximately for pK5LH14. Using this method, approximate 5' and 3' regions of the promoter were identified, and outward-facing sequencing primers were designed in anticipation that respective sequencing reactions would encompass the 5' and 3' limits of the cloned promoter fragment. The limits were identified by using the BLAST program to compare the resulting sequence data against the BAC clone (data not shown). In both cases, homology between the sequenced K5 promoter and the chromosome 12 clone ended within the sequenced region, indicating the 5' and 3' junctions between the promoter and flanking vector sequence. This revealed the K5 promoter fragment to be 6213bps in length. The intervening and flanking (vector) sequences were used to identify useful enzyme recognition sites for subsequent manipulations.

The 6.2kb K5 promoter was released from pK5LH14 by cutting 5' with MluI before liberating the fragment by cutting 3' with AflII. The recipient vector, pIRES2-eGFP (Clontech), was prepared by digestion with AseI and BglII to release the CMV

promoter. The incompatible MluI-AseI and AflII-BglII ends were blunted with T4 DNA polymerase before the K5 promoter insert was ligated into the vacated position. Transformed clones containing correctly-oriented inserts were identified by colony hybridisation using an oligonucleotide probe designed to recognise sequence corresponding to the filled-in and ligated MluI and AseI sites, and flanking bases (Figure A1.1a,b). These clones were further verified by diagnostic restriction digests (Figure A1.1c,d).

A1.2 Incorporating the pCI intron from pK5LH14

Immediately downstream of the K5 promoter in pK5LH14 was a chimaeric (human β-globin/immunoglobulin) intron from the pCI expression vector (Promega). Since the promoter-intron combination appeared to work well in lines derived from pK5LH14 ((Schoch *et al.*, 2003) and S. Randell, personal communication), it was desirable to maintain this configuration. The distribution of restriction enzyme sites in pK5LH14 did not permit subcloning of the entire K5 promoter-intron fragment in one step, necessitating the subsequent subcloning of the intron. The relatively short length of the intron meant that PCR-cloning was the most convenient strategy (Figure A1.2a,b), which also allowed suitable enzyme recognition sequences to be incorporated into the primer tails. Primers were designed from sequence information generated during the prior sequencing of the 3' of the K5 promoter, with the limits of the intron corresponding to AflII sites known to exist within pK5LH14. XhoI and SalI sites were included in the 5' and 3' primers, respectively, and the digested PCR fragment was cloned into the SalI site 3' of the K5 promoter in pK5-eGFP. Transformed clones containing the insert were identified by colony hybridisation with the 3' pCI intron PCR primer oligonucleotide ("pCI-as"), and three correctly-oriented inserts were identified by PCR using the 3' K5 promoter sequencing primer ("3'K5") and the 3' pCI intron PCR primer (Figure A1.2a,b). A subtle difference in size between PCR products generated from clones 2 and 3 was reflected in a BamHI diagnostic digest (the digest also revealed clone 7 to be aberrant, Figure A1.2d). Sequencing the intron, which confirmed no errors had been induced by PCR, also revealed that while clone 3 had been ligated as anticipated, clone 2 had undergone a

recombination event at the 5' ligation site between flanking sequence duplicated in the PCR product, reconstituting precisely the sequence present at the junction of the K5 promoter and the pCI intron in pK5LH14 (Figure A1.2c). An AflII digest confirmed the presence of the intron, detectable as an 197bp fragment (Figure A1.2c,d). Thus, clone 2 was used in subsequent cloning steps.

A1.3 Cloning lacZ to generate pK5A1

To prepare the pK5CI-eGFP plasmid to receive the *lacZ* gene, SacII was used to cut 3' of the pCI intron, then blunted using T4 DNA polymerase. SalI was then used to cut 10bp 5' of the blunted SacII site. *lacZ* was excised from plasmid pPB2 (courtesy of P. Budd) using SalI to cut 5' and DraI to cut at the 3' of the gene. DraI cut upstream of the *lacZ* polyadenylation signal, which if included in the subcloning would have reduced or prevented transcription proceeding into the downstream IRES-*eGFP* cistron. The SalI-DraI fragment was ligated into the SalI-SacII(blunt) interval in pK5CI-eGFP to create pK5A1 (Figure A1.3a). Transformed clones containing the insert were identified by colony hybridisation with a *lacZ*-specific oligonucleotide. A sequencing reaction originating in the 5' of the IRES verified the 3' integrity of the *lacZ* insert (data not shown), and diagnostic digests were used to detect successful cloning events (Figure A1.3b).

A1.4 Cloning CreER^{T2} into pK5-eGFP to create pK5CE

The *CreER^{T2}* fusion gene (courtesy of P. Chambon) was subcloned, along with the rabbit β-globin intron II, downstream of the human K5 promoter in pK5-eGFP. The donor plasmid was first digested with NotI to remove a 2.86kb fragment that would otherwise have been co-purified with the desired 2.89kb SalI-NotI fragment, which was subsequently isolated after blunting the NotI end. The recipient pK5-eGFP vector was prepared using SacII, which was blunted using Mung Bean Nuclease, before cutting 5' with SalI and ligating with the intron-*CreER^{T2}* insert (Figure A1.4a) to create pK5CE. Diagnostic digests (Figure A1.4b) and sequencing (data not shown) of the 3' of the insert identified an intact clone.

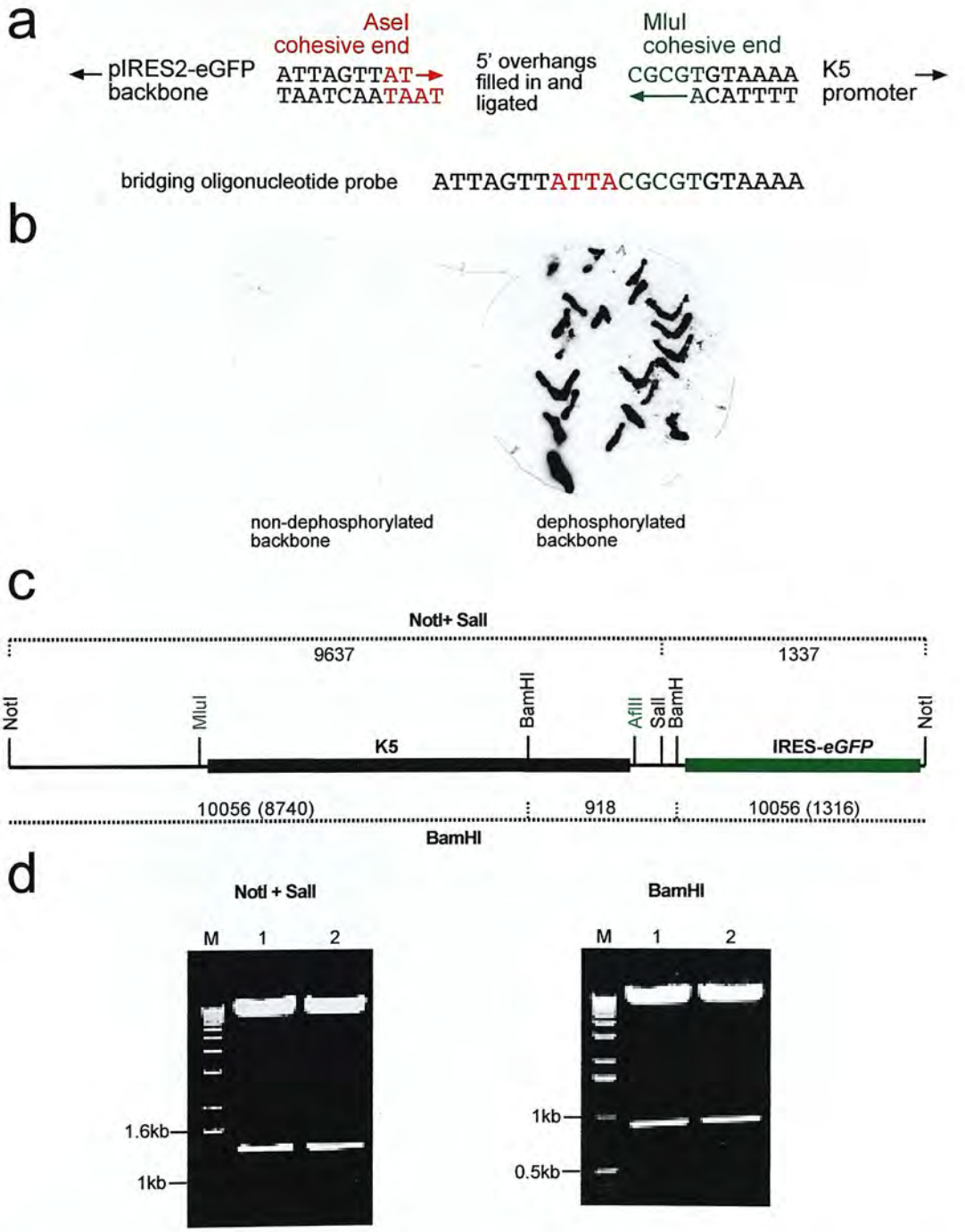


Figure A1.1 Cloning the K5 promoter into the promoter-less pIRES2-eGFP vector.

(a) pIRES2-eGFP was digested with AseI + BglII to remove the CMV promoter, and the cohesive ends were blunted to make them compatible with the blunted MluI and AflII sites used to release the K5 promoter from pK5LH14 (top). Correctly-oriented clones were identified by hybrisidisation to a radiolabelled 22mer bridging oligonucleotide containing sequence complementary to the ligated, blunted AseI and MluI sites (bottom). (b) Autoradiograph showing hybridising colony streaks resulting only from ligations involving dephosphorylated receipient vector blunt ends. In ligations performed without dephosphorylation, reannealing of vector blunt ends prevented insertion of the promoter fragment. (c) Map of cloned K5 promoter showing restriction sites used in diagnostic digests, including predicted fragment sizes. (d) Diagnostic digest of two clones which show restriction patterns predicted of successful cloning. Donor cloning sites are highlighted in green. Although both sites were blunted, both were reconstituted after blunt-end ligation. M = 1kb ladder. Map not to scale, fragment sizes in bps. Bracketed fragment sizes indicate proportion of total fragment size.

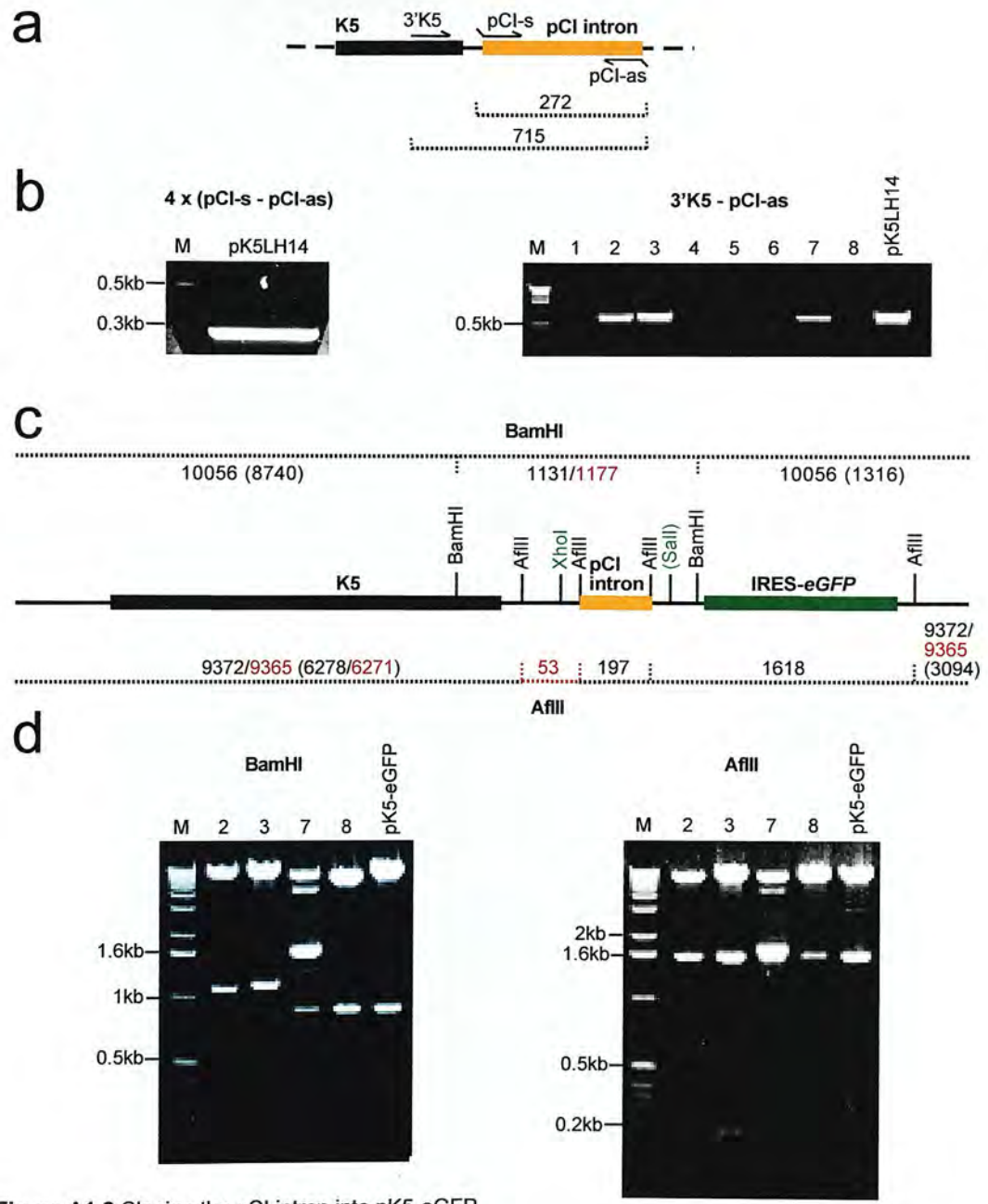


Figure A1.2 Cloning the pCI intron into pK5-eGFP.

(a) The pCI intron was amplifed by PCR from pK5LH14 using primers pCI-s - pCI-as which included 5' tails containing XhoI and SalI restriction sequences, respectively. The position of primer 3'K5 is shown. (b) The product from four PCR reactions was pooled and the 272bp fragment gel purified (left), before digestion with Xho + SalI. The pCI intron fragment was ligated into XhoI-cut pK5-eGFP. Once successful clonings were selected by probing with radiolabelled pCI-as oligonucleotide, correctly-oriented clones were identifed by PCR amplification using primers 3'K5 and pCI-as (right). Of eight pCI intron-containing clones, three (2,3,7) were correctly oriented. pK5LH14 included as control. (c) Restriction fragments predicted from diagnostic digests of a successful cloning, including predicted fragment sizes. (d) Diagnostic digests of three correctly- (2,3,7) and one incorrectly- (8) oriented clones from (b). Only clones 2 and 3 gave the predicted pattern, though they showed variation in size of the predicted 1.1kb fragment. Sequencing revealed clone 3 to be exactly as predicted, corresponding to additional fragments in red in (c), though 53bp fragment is too small to visualise in digests. Due to sequence duplication between pCI intron 5' amplification primer and sequence flanking the 5' cloning site, a recombination event occurred in clone 2, reconstituting the exact sequence from pK5LH14. Clone 2 does not contain the fragments in red in (c). The size difference between clones 3 (715bp, as predicted) and 2 (669bp) is subtly discernible in (b). The 197bp AflII fragment is faint but present in both clones 2 and 3 (d). Cloning sites in green, brackets indicate site was not reconstituted. M = 1kb ladder. Map not to scale, fragment sizes in bps. Bracketed fragment sizes indicate proportion of total fragment size.

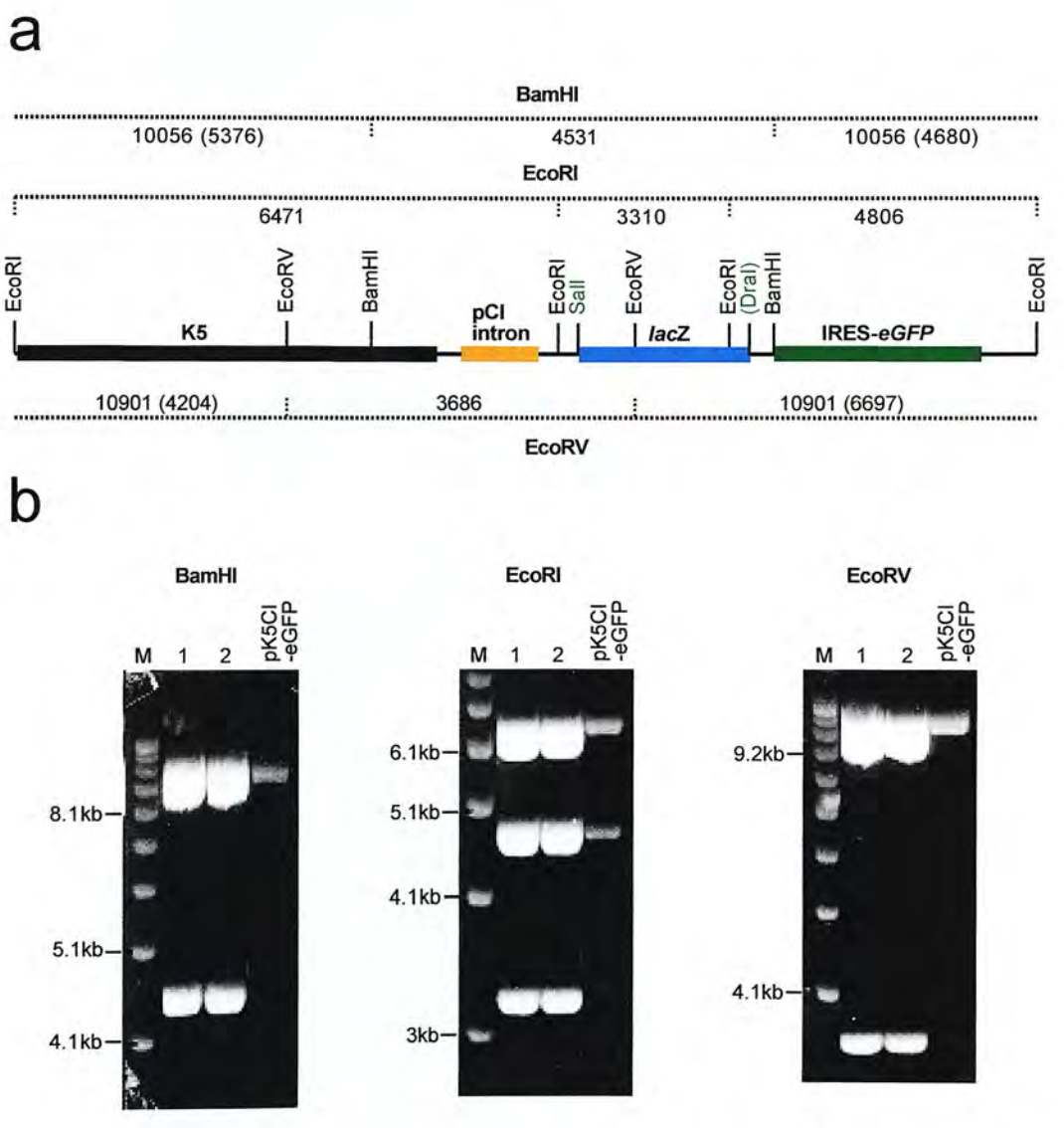
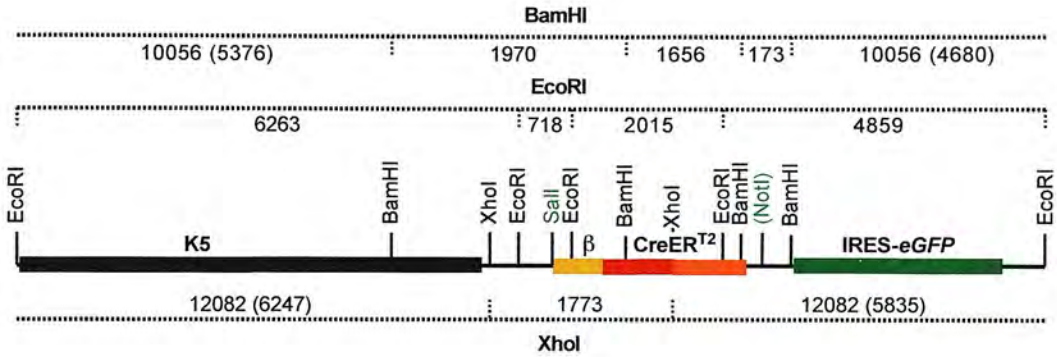


Figure A1.3 Cloning *lacZ* into pK5CI-eGFP.

(a) pK5CI-eGFP clone 2 (Figure A1.2) was prepared to receive *lacZ* (SalI-DraI fragment) by cutting with SacII 3' of the pCI intron, blunting the cohesive end, and cutting 10bp 5' with SalI. Restriction fragments diagnostic of a successful cloning are shown, including predicted fragment sizes. (b) Two clones showing restriction patterns predicted for a successful cloning in three different digests. The recipient pK5CI-eGFP vector was included in the digests, and gave a pattern expected in the absence of the *lacZ* sequence (the 1131bp BamHI fragment was too small for inclusion on BamHI gel). Donor cloning sites in green, brackets indicate site may not have been reconstituted. M = 1kb ladder. Map not to scale, fragment sizes in bps. Bracketed fragment sizes indicate proportion of total fragment size.

a



b

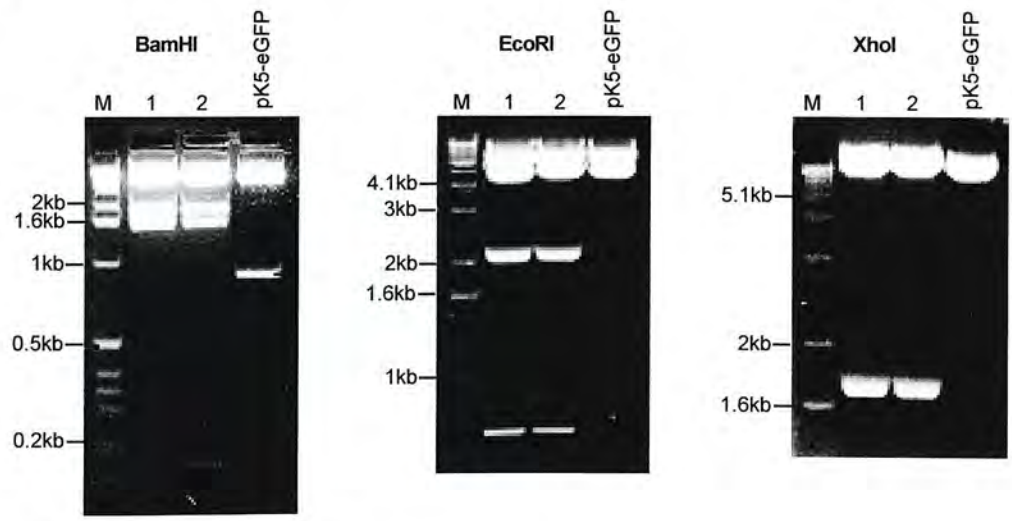


Figure A1.4 Cloning the *CreER^{T2}* fusion gene into pK5-eGFP.

(a) pK5-eGFP was prepared to receive *CreER^{T2}*, including the 5' rabbit β-globin intron II (β on map), by cutting with SacII 3' of the K5 promoter, blunting the cohesive end, and cutting 10bp 5' with SalI. The intron-*CreER^{T2}* fragment was liberated from the donor vector using NotI 3' (subsequently blunted) and SalI 5'. Restriction fragments diagnostic of a successful cloning are shown, including predicted fragment sizes. (b) Two clones showing restriction patterns similar to those predicted for a successful cloning in three different digests. However, only clone 2 exhibited the 173bp BamHI fragment which confirmed the presence of the BamHI site in, and thus the integrity of, the 3' intron-*CreER^{T2}* fragment. The sequence difference between clones 1 and 2 was also detected by sequencing, and additional EcoRV, HindIII, and EcoRI + NotI digests confirmed the orientation of the intron-*CreER^{T2}* fragment and overall integrity of the final pK5CE construct (not shown). The recipient pK5-eGFP vector was included in the digests, and gave patterns expect in the absence of the intron-*CreER^{T2}* fragment sequence. Donor cloning sites in green, brackets indicate site may not have been reconstituted. M = 1kb ladder. Map not to scale, fragment sizes in bps. Bracketed fragment sizes indicate proportion of total fragment size.

A2 Construction of transgenic line R26X

The DNA manipulations required to generate the transgenic mutated reporter line *R26X* are described. The cloning scheme is summarised schematically in chapter 5, Figure 5.2. The non-mutated control line *R26Z* was produced identically and in parallel, omitting the initial mutagenesis step.

A2.1 Exchanging the CMV promoter for the R26 promoter

The R26 promoter, provided in pBluescript (Stratagene), was sequenced using T3 and T7 primers to obtain 5' sequence to complement the available 524bp of 3' sequence (GenBank accession U83173). The composite sequence was subsequently confirmed using the BLAST program to align the 3' portion against the Ensembl mouse genome assembly.

The CMV promoter fragment was removed from pCMVβ by cutting at the 3' XhoI site, filling in the cohesive ends using T4 DNA polymerase, before cutting 5' with EcoRI. The 822bp R26 promoter fragment was ligated into the open vector following its liberation from pBluescript using BamHI to cut 3', which was filled in for ligation with the previously blunted XhoI site, before cutting 5' with EcoRI (Figure A2.1a). Diagnostic digests were as predicted for the correct construct (Figure A2.1b).

A2.2 Addition of IRES-eGFP to make a bicistronic construct

An IRES-*eGFP* component was incorporated into the construct by cloning the fragment containing the R26 promoter and R376X *lacZ* gene into pIRES2-eGFP (Clontech). pIRES2-eGFP was first prepared by removing the CMV promoter through a double digestion with AseI and BglII which cut 5' and 3' of the promoter, respectively. The cut ends were made blunt and ligated to re-circularise the plasmid (Figure A2.2(i)a). A double EcoRI/NotI digest gave a pattern of fragments consistent with loss of the promoter fragment. Digestion with SnaBI, which cut pIRES2-eGFP

uniquely within the CMV promoter, was shown to linearise the parental vector. Plasmids which had re-circularised after excision of the AseI-BglII CMV promoter fragment remained uncut (Figure A2.2(i)b).

The DNA fragment containing the R26 promoter and R376X *lacZ* gene was released from vector sequences by cutting initially with EcoRI 5' of the R26 promoter, then digesting with SacII 5' of the SV40 polyadenylation signal. An additional SacII site existed within the 5' region of *lacZ*, necessitating the use of a controlled partial digest (Figure A2.2(i)c). A SacII digest of pR26R376X was monitored to determine the incubation time at which the optimal proportion of the desired partial digest product was present (Figure A2.2(i)d). The 4.5kb EcoRI-SacII fragment containing the R26 promoter and R376X *lacZ* gene was gel purified from the four remaining products and ligated into the promoter-less pIRES2-eGFPnp after cutting the vector at the EcoRI and SacII polylinker sites (Figure A2.2(ii)a). A diagnostic NotI digest (supported by a linearising SalI reaction, not shown) was used to select successfully cloned pR26X plasmid (Figure A2.2(ii)b).

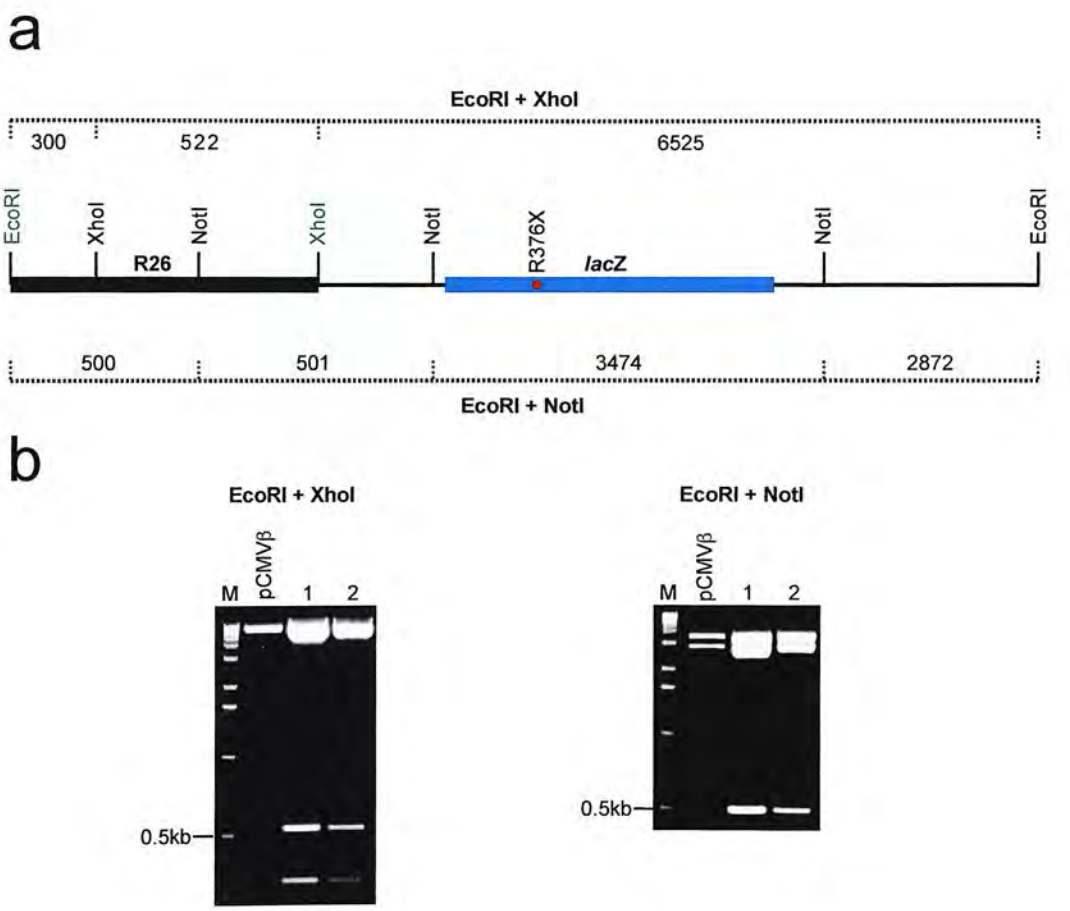


Figure A2.1 Cloning the R26 promoter into pCMVβ following the removal of the CMV promoter.

(a) The R26 promoter fragment was subcloned using EcoRI (5') and blunted BamHI/XhoI sites (3') into pCMVβ, following removal of the CMV promoter. Restriction sites used in diagnostic digests are shown, along with predicted fragment sizes. (b) Diagnostic digests of two clones which show fragments predicted for the correct construct, plus parental pCMVβ vector. 639bp fragment predicted in EcoRI + XhoI digestion of pCMVβ is too weak to visualise. Cloning sites are highlighted in green (digests reveal reconstitution of XhoI site after ligation of blunted XhoI site to blunted BamHI site). M = 1kb ladder. Map not to scale, fragment sizes in bps.

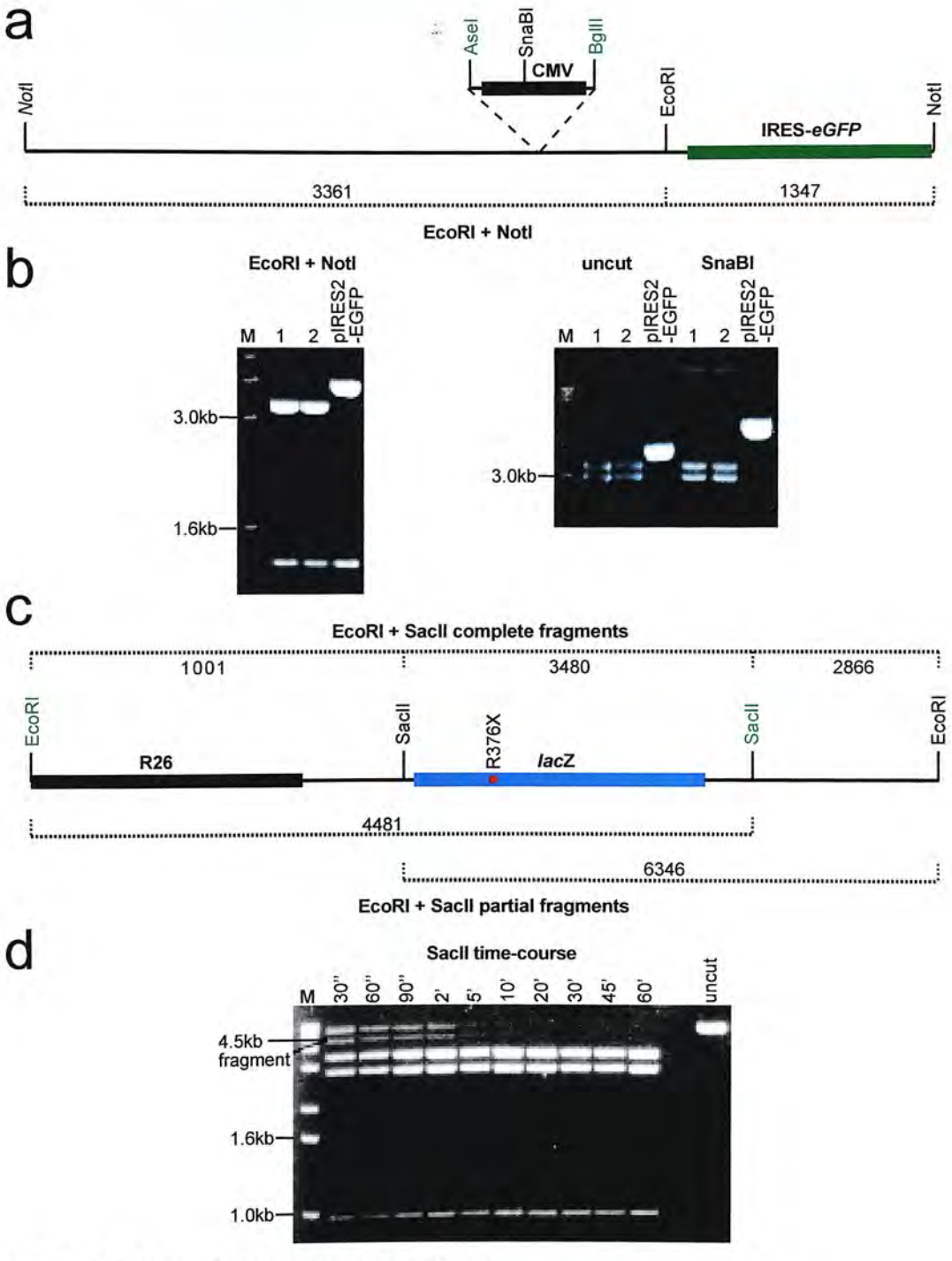


Figure A2.2(i) Addition of an IRES-*eGFP* component.

(a,b) Re-circularisation of pIRES2-eGFP after removal of the CMV promoter. (a) 602bp AseI-BglII fragment containing the CMV promoter was removed and the vector backbone re-ligated. Restriction sites used in diagnostic digests are shown, along with predicted fragment sizes. (b) Diagnostic digests of two clones which showed fragments predicted for the successful removal of the promoter, plus the parental pIRES2-eGFP. SnaBI, unique to the CMV promoter, cuts pIRES2-eGFP (evident as loss of supercoiling) but not clones 1 and 2. (c,d) Liberation of the R26-*lacZ* component. (c) After incubation with EcoRI, a SacII time-course was conducted to determine the digest duration which yielded the optimal proportion of the desired 4.5kb EcoRI-SacII fragment. The relevant restriction sites are shown, along with predicted complete and partial digest fragment sizes. (d) SacII time-course showing changing ratios of five EcoRI-EcoRI, SacII-SacII and EcoRI-SacII fragments (" = seconds, ' = minutes). A 40 second SacII incubation was used in the subcloning. Excision sites are highlighted in green. M = 1kb ladder. Map not to scale, fragment sizes in bps.

a

b

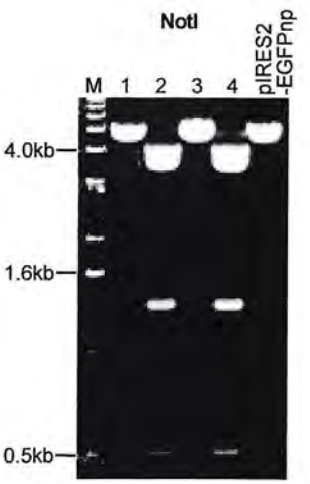


Figure A2.2(ii) Addition of an IRES-*eGFP* component.

Following re-circularisation of the promoter-less pIRES2-eGFPnp, the vector was re-cut with EcoRI + SacII to receive the R26-*lacZ* fragment. (a) The final pR26X construct. Restriction sites used in diagnostic digests are shown, along with predicted fragment sizes. (b) Diagnostic digest of four clones: clones 2 and 4 show the predicted pattern of four fragments (the larger fragments visible as a doublet), while clones 1 and 3 give the same pattern as the pIRES2-eGFPnp backbone. Cloning sites are highlighted in green. M = 1kb ladder. Map not to scale, fragment sizes in bps.

References

Albuquerque-Silva, J, Vassart, G, Lavinha, J, and Abramowicz, MJ. 2001. Chimeraplasty validation. *Nat. Biotechnol.* **19:** 1011-1011

Alexeev, V, Igoucheva, O, Domashenko, A, Cotsarelis, G, and Yoon, K. 2000. Localized in vivo genotypic and phenotypic correction of the albino mutation in skin by RNA-DNA oligonucleotide. *Nat. Biotechnol.* **18:** 43-47.

Alvarez-Dolado, M, Pardal, R, Garcia-Verdugo, JM, Fike, JR, Lee, HO *et al.* 2003. Fusion of bone-marrow-derived cells with Purkinje neurons, cardiomyocytes and hepatocytes. *Nature* **425:** 968-973.

Araki, K, Araki, M, Miyazaki, J, and Vassalli, P. 1995. Site-specific recombination of a transgene in fertilized eggs by transient expression of Cre recombinase. *Proc. Natl. Acad. Sci. U. S. A* **92:** 160-164.

Arin, MJ, Longley, MA, Wang, XJ, and Roop, DR. 2001. Focal activation of a mutant allele defines the role of stem cells in mosaic skin disorders. *J. Cell Biol.* **152:** 645-649.

Barth, PJ, Koch, S, Muller, B, Unterstab, F, von Wichert, P *et al.* 2000. Proliferation and number of Clara cell 10-kDa protein (CC10)-reactive epithelial cells and basal cells in normal, hyperplastic and metaplastic bronchial mucosa. *Virchows Arch.* **437:** 648-655.

Bartlett, RJ, Stockinger, S, Denis, MM, Bartlett, WT, Inverardi, L *et al.* 2000. In vivo targeted repair of a point mutation in the canine dystrophin gene by a chimeric RNA/DNA oligonucleotide. *Nat. Biotechnol.* **18:** 615-622.

Beddington, RS, Morgernstern, J, Land, H, and Hogan, A. 1989. An in situ transgenic enzyme marker for the midgestation mouse embryo and the visualization of inner cell mass clones during early organogenesis. *Development* **106:** 37-46.

Bedwell, DM, Kaenjak, A, Benos, DJ, Bebok, Z, Bubien, JK *et al.* 1997. Suppression of a CFTR premature stop mutation in a bronchial epithelial cell line. *Nat. Med.* **3:** 1280-1284.

Bennett, M and Schaack, J. 2003. Development of a dual-luciferase fusion gene as a sensitive marker for site-directed DNA repair strategies. *J. Gene Med.* **5:** 723-732.

Berton, TR, Matsumoto, T, Page, A, Conti, CJ, Deng, CX *et al.* 2003. Tumor formation in mice with conditional inactivation of Brca1 in epithelial tissues. *Oncogene* **22:** 5415-5426.

Bertoni, C, Lau, C, and Rando, TA. 2003. Restoration of dystrophin expression in mdx muscle cells by chimeraplast-mediated exon skipping. *Hum. Mol. Genet.* **12:** 1087-1099.

Bertoni, C and Rando, TA. 2002. Dystrophin gene repair in mdx muscle precursor cells in vitro and in vivo mediated by RNA-DNA chimeric oligonucleotides. *Hum. Gene Ther.* **13:** 707-718.

Bickenbach, JR. 1981. Identification and behavior of label-retaining cells in oral mucosa and skin. *J. Dent. Res.* **60 Spec No C:** 1611-1620.

Bjerknes, M and Cheng, H. 1999. Clonal analysis of mouse intestinal epithelial progenitors. *Gastroenterology* **116:** 7-14.

Bjornson, CR, Rietze, RL, Reynolds, BA, Magli, MC, and Vescovi, AL. 1999. Turning brain into blood: a hematopoietic fate adopted by adult neural stem cells in vivo. *Science* **283:** 534-537.

Blau, HM, Brazelton, TR, and Weimann, JM. 2001. The evolving concept of a stem cell: entity or function? *Cell* **105:** 829-841.

Boers, JE, Ambergen, AW, and Thunnissen, FB. 1998. Number and proliferation of basal and parabasal cells in normal human airway epithelium. *Am. J Respir. Crit Care Med.* **157:** 2000-2006.

Boers, JE, Ambergen, AW, and Thunnissen, FB. 1999. Number and proliferation of clara cells in normal human airway epithelium. *Am. J. Respir. Crit Care Med.* **159:** 1585-1591.

Bonnerot, C and Nicolas, JF. 1993. Clonal analysis in the intact mouse embryo by intragenic homologous recombination. *C. R. Acad. Sci. III* **316:** 1207-1217.

Booth, C and Potten, CS. 2000. Gut instincts: thoughts on intestinal epithelial stem cells. *J. Clin. Invest* **105:** 1493-1499.

Borthwick, DW. 1999. An Investigation of Murine Tracheal Cellular Kinetics: Implications for Stem Cell Gene Therapy (Thesis).

Borthwick, DW, Shahbazian, M, Krantz, QT, Dorin, JR, and Randell, SH. 2001. Evidence for stem-cell niches in the tracheal epithelium. *Am. J Respir. Cell Mol. Biol.* **24:** 662-670.

Borthwick, DW, West, JD, Keighren, MA, Flockhart, JH, Innes, BA *et al.* 1999. Murine submucosal glands are clonally derived and show a cystic fibrosis gene-dependent distribution pattern. *Am. J Respir. Cell Mol. Biol.* **20:** 1181-1189.

Brakebusch, C, Grose, R, Quondamatteo, F, Ramirez, A, Jorcano, JL *et al.* 2000. Skin and hair follicle integrity is crucially dependent on beta 1 integrin expression on keratinocytes. *EMBO J.* **19:** 3990-4003.

Braun, KM, Niemann, C, Jensen, UB, Sundberg, JP, Silva-Vargas, V *et al.* 2003. Manipulation of stem cell proliferation and lineage commitment: visualisation of label-retaining cells in wholemounts of mouse epidermis. *Development* **130:** 5241-5255.

Breuer, R, Zajicek, G, Christensen, TG, Lucey, EC, and Snider, GL. 1990. Cell kinetics of normal adult hamster bronchial epithelium in the steady state. *Am. J. Respir. Cell Mol. Biol.* **2:** 51-58.

Brocard, J, Warot, X, Wendling, O, Messaddeq, N, Vonesch, JL *et al.* 1997. Spatio-temporally controlled site-specific somatic mutagenesis in the mouse. *Proc. Natl. Acad. Sci. U. S. A* **94:** 14559-14563.

Bruscia, E, Sangiuolo, F, Sinibaldi, P, Goncz, KK, Novelli, G *et al.* 2002. Isolation of CF cell lines corrected at DeltaF508-CFTR locus by SFHR-mediated targeting. *Gene Ther.* **9:** 683-685.

Byrne, C and Fuchs, E. 1993. Probing keratinocyte and differentiation specificity of the human K5 promoter in vitro and in transgenic mice. *Mol. Cell Biol.* **13:** 3176-3190.

Byrne, C, Tainsky, M, and Fuchs, E. 1994. Programming gene expression in developing epidermis. *Development* **120:** 2369-2383.

Calvi, LM, Adams, GB, Weibrecht, KW, Weber, JM, Olson, DP *et al.* 2003. Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* **425:** 841-846.

Casatorres, J, Navarro, JM, Blessing, M, and Jorcano, JL. 1994. Analysis of the control of expression and tissue specificity of the keratin 5 gene, characteristic of basal keratinocytes. Fundamental role of an AP-1 element. *J Biol. Chem.* **269:** 20489-20496.

Caspers, PJ, Williams, AC, Carter, EA, Edwards, HG, Barry, BW *et al.* 2002. Monitoring the penetration enhancer dimethyl sulfoxide in human stratum corneum in vivo by confocal Raman spectroscopy. *Pharm. Res.* **19:** 1577-1580.

Chai, Y, Jiang, X, Ito, Y, Bringas, P, Jr., Han, J *et al.* 2000. Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* **127:** 1671-1679.

Cheng, H and Bjerknes, M. 1985. Whole population cell kinetics and postnatal development of the mouse intestinal epithelium. *Anat. Rec.* **211:** 420-426.

Cheng, J, Turksen, K, Yu, QC, Schreiber, H, Teng, M *et al.* 1992. Cachexia and graft-vs.-host-disease-type skin changes in keratin promoter-driven TNF alpha transgenic mice. *Genes Dev.* **6:** 1444-1456.

Chisholm, JC and Houliston, E. 1987. Cytokeratin filament assembly in the preimplantation mouse embryo. *Development* **101:** 565-582.

Clark, AJ, Harold, G, and Yull, FE. 1997. Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice. *Nucleic Acids Res.* **25:** 1009-1014.

Collick, A, Drew, J, Penberth, J, Bois, P, Luckett, J *et al.* 1996. Instability of long inverted repeats within mouse transgenes. *EMBO J.* **15:** 1163-1171.

Collinson, JM, Morris, L, Reid, AI, Ramaesh, T, Keighren, MA *et al.* 2002. Clonal analysis of patterns of growth, stem cell activity, and cell movement during the

development and maintenance of the murine corneal epithelium. *Dev. Dyn.* **224:** 432-440.

Colosimo, A, Goncz, KK, Novelli, G, Dallapiccola, B, and Gruenert, DC. 2001. Targeted correction of a defective selectable marker gene in human epithelial cells by small DNA fragments. *Mol. Ther.* **3:** 178-185.

Cotsarelis, G, Cheng, SZ, Dong, G, Sun, TT, and Lavker, RM. 1989. Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: implications on epithelial stem cells. *Cell* **57:** 201-209.

Cotsarelis, G, Sun, TT, and Lavker, RM. 1990. Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61:** 1329-1337.

Cui, C, Wani, MA, Wight, D, Kopchick, J, and Stambrook, PJ. 1994. Reporter genes in transgenic mice. *Transgenic Res.* **3:** 182-194.

Danielian, PS, Muccino, D, Rowitch, DH, Michael, SK, and McMahon, AP. 1998. Modification of gene activity in mouse embryos in utero by a tamoxifen- inducible form of Cre recombinase. *Curr. Biol.* **8:** 1323-1326.

Danielian, PS, White, R, Hoare, SA, Fawell, SE, and Parker, MG. 1993. Identification of residues in the estrogen receptor that confer differential sensitivity to estrogen and hydroxytamoxifen. *Mol. Endocrinol.* **7:** 232-240.

DasGupta, R and Fuchs, E. 1999. Multiple roles for activated LEF/TCF transcription complexes during hair follicle development and differentiation. *Development* **126:** 4557-4568.

Davidson, DJ. 2000. Lung Disease in the Cystic Fibrosis Mutant Mouse (Thesis).

de Vries, WN, Binns, LT, Fancher, KS, Dean, J, Moore, R *et al.* 2000. Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. *Genesis.* **26:** 110-112.

Debey, P, Szollosi, MS, Szollosi, D, Vautier, D, Girousse, A *et al.* 1993. Competent mouse oocytes isolated from antral follicles exhibit different chromatin organization and follow different maturation dynamics. *Mol. Reprod. Dev.* **36:** 59-74.

Dekker, M, Brouwers, C, and te Riele, H. 2003. Targeted gene modification in mismatch-repair-deficient embryonic stem cells by single-stranded DNA oligonucleotides. *Nucleic Acids Res.* **31:** e27

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

Diaz-Font, A, Cormand, B, Chabas, A, Vilageliu, L, and Grinberg, D. 2003. Unsuccessful chimeraplast strategy for the correction of a mutation causing Gaucher disease. *Blood Cells Mol. Dis.* **31:** 183-186.

DiGiovanni, J, Bol, DK, Wilker, E, Beltran, L, Carbajal, S *et al.* 2000. Constitutive expression of insulin-like growth factor-1 in epidermal basal cells of transgenic mice leads to spontaneous tumor promotion. *Cancer Res.* **60:** 1561-1570.

Dobie, KW, Lee, M, Fantes, JA, Graham, E, Clark, AJ *et al.* 1996. Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus. *Proc. Natl. Acad. Sci. U. S. A* **93:** 6659-6664.

Doerflinger, NH, Macklin, WB, and Popko, B. 2003. Inducible site-specific recombination in myelinating cells. *Genesis.* **35:** 63-72.

Donnelly, GM, Haack, DG, and Heird, CS. 1982. Tracheal epithelium: cell kinetics and differentiation in normal rat tissue. *Cell Tissue Kinet.* **15:** 119-130.

Duan, D, Sehgal, A, Yao, J, and Engelhardt, JF. 1998. Lef1 transcription factor expression defines airway progenitor cell targets for in utero gene therapy of submucosal gland in cystic fibrosis. *Am. J. Respir. Cell Mol. Biol.* **18:** 750-758.

Duan, D, Yue, Y, Zhou, W, Labed, B, Ritchie, TC *et al.* 1999. Submucosal gland development in the airway is controlled by lymphoid enhancer binding factor 1 (LEF1). *Development* **126:** 4441-4453.

Dunnwald, M, Tomanek-Chalkley, A, Alexandrunas, D, Fishbaugh, J, and Bickenbach, JR. 2001. Isolating a pure population of epidermal stem cells for use in tissue engineering. *Exp. Dermatol.* **10:** 45-54.

Emura, M. 1997. Stem cells of the respiratory epithelium and their in vitro cultivation. *In Vitro Cell Dev. Biol Anim* **33:** 3-14.

Emura, M. 2002. Stem cells of the respiratory tract. *Paediatr. Respir. Rev.* **3:** 36-40.

Engelhardt, JF. 2001. Stem cell niches in the mouse airway. *Am. J Respir. Cell Mol. Biol.* **24:** 649-652.

Engelhardt, JF, Schlossberg, H, Yankaskas, JR, and Dudus, L. 1995. Progenitor cells of the adult human airway involved in submucosal gland development. *Development* **121:** 2031-2046.

Engelhardt, JF, Yankaskas, JR, Ernst, SA, Yang, Y, Marino, CR *et al.* 1992. Submucosal glands are the predominant site of CFTR expression in the human bronchus. *Nat. Genet.* **2:** 240-248.

Engelhardt, JF, Zepeda, M, Cohn, JA, Yankaskas, JR, and Wilson, JM. 1994. Expression of the cystic fibrosis gene in adult human lung. *J. Clin. Invest* **93:** 737-749.

Epstein, JA, Li, J, Lang, D, Chen, F, Brown, CB *et al.* 2000. Migration of cardiac neural crest cells in Splotch embryos. *Development* **127:** 1869-1878.

Evans, MJ, Cox, RA, Shami, SG, and Plopper, CG. 1990. Junctional adhesion mechanisms in airway basal cells. *Am. J. Respir. Cell Mol. Biol.* **3:** 341-347.

Evans, MJ, Cox, RA, Shami, SG, Wilson, B, and Plopper, CG. 1989. The role of basal cells in attachment of columnar cells to the basal lamina of the trachea. *Am. J Respir. Cell Mol. Biol.* **1:** 463-469.

Evans, MJ, Johnson, LV, Stephens, RJ, and Freeman, G. 1976. Renewal of the terminal bronchiolar epithelium in the rat following exposure to NO2 or O3. *Lab Invest* **35:** 246-257.

Evans, M. J. and Shami, S. G. 1989. Lung Cell Kinetics *in Lung Cell Biology.* Marcel Dekker. New York

Evans, MJ, Shami, SG, Cabral-Anderson, LJ, and Dekker, NP. 1986. Role of nonciliated cells in renewal of the bronchial epithelium of rats exposed to NO2. *Am. J Pathol.* **123:** 126-133.

Feil, R, Brocard, J, Mascrez, B, LeMeur, M, Metzger, D *et al.* 1996. Ligand-activated site-specific recombination in mice. *Proc. Natl. Acad. Sci. U. S A* **93:** 10887-10890.

Feil, R, Wagner, J, Metzger, D, and Chambon, P. 1997. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochem. Biophys. Res. Commun.* **237:** 752-757.

Ferrari, G, Cusella-De Angelis, G, Coletta, M, Paolucci, E, Stornaiuolo, A *et al.* 1998. Muscle regeneration by bone marrow-derived myogenic progenitors. *Science* **279:** 1528-1530.

Festenstein, R, Tolaini, M, Corbella, P, Mamalaki, C, Parrington, J *et al.* 1996. Locus control region function and heterochromatin-induced position effect variegation. *Science* **271:** 1123-1125.

Filali, M, Liu, X, Cheng, N, Abbott, D, Leontiev, V *et al.* 2002. Mechanisms of submucosal gland morphogenesis in the airway. *Novartis. Found. Symp.* **248:** 38-45.

Ford, JR and Terzaghi-Howe, M. 1992a. Basal cells are the progenitors of primary tracheal epithelial cell cultures. *Exp. Cell Res* **198:** 69-77.

Ford, JR and Terzaghi-Howe, M. 1992b. Characteristics of magnetically separated rat tracheal epithelial cell populations. *Am. J. Physiol* **263:** L568-L574.

Fortunel, NO, Otu, HH, Ng, HH, Chen, J, Mu, X *et al.* 2003. Comment on " 'Stemness': transcriptional profiling of embryonic and adult stem cells" and "a stem cell molecular signature". *Science* **302:** 393b

Friedrich, G and Soriano, P. 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5:** 1513-1523.

Fuchs, E and Segre, JA. 2000. Stem cells: a new lease on life. *Cell* **100:** 143-155.

Gallicano, GI. 2001. Composition, regulation, and function of the cytoskeleton in mammalian eggs and embryos. *Front Biosci.* **6:** D1089-D1108.

Gallicano, GI, Larabell, CA, McGaughey, RW, and Capco, DG. 1994. Novel cytoskeletal elements in mammalian eggs are composed of a unique arrangement of intermediate filaments. *Mech. Dev.* **45:** 211-226.

Gamper, HB, Parekh, H, Rice, MC, Bruner, M, Youkey, H *et al.* 2000. The DNA strand of chimeric RNA/DNA oligonucleotides can direct gene repair/conversion activity in mammalian and plant cell-free extracts. *Nucleic Acids Res* **28:** 4332-4339.

Gao, XH, Kondoh, G, Tarutani, M, Hara, M, Inoue, S *et al.* 2002. Rapid compensation for glycosylphosphatidylinositol anchor deficient keratinocytes after birth: visualization of glycosylphosphatidylinositol-anchored proteins in situ. *J. Invest Dermatol.* **118:** 998-1002.

Garrick, D, Fiering, S, Martin, DI, and Whitelaw, E. 1998. Repeat-induced gene silencing in mammals. *Nat. Genet.* **18:** 56-59.

Giangreco, A, Reynolds, SD, and Stripp, BR. 2002. Terminal bronchioles harbor a unique airway stem cell population that localizes to the bronchoalveolar duct junction. *Am. J. Pathol.* **161:** 173-182.

Gluzman, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23:** 175-182.

Goncz, KK, Colosimo, A, Dallapiccola, B, Gagne, L, Hong, K *et al.* 2001. Expression of DeltaF508 CFTR in normal mouse lung after site-specific modification of CFTR sequences by SFHR. *Gene Ther.* **8:** 961-965.

Goncz, KK, Kunzelmann, K, Xu, Z, and Gruenert, DC. 1998. Targeted replacement of normal and mutant CFTR sequences in human airway epithelial cells using DNA fragments. *Hum. Mol. Genet.* **7:** 1913-1919.

Goodell, MA, Rosenzweig, M, Kim, H, Marks, DF, DeMaria, M *et al.* 1997. Dye efflux studies suggest that hematopoietic stem cells expressing low or undetectable levels of CD34 antigen exist in multiple species. *Nat. Med.* **3:** 1337-1345.

Grove, JE, Lutzko, C, Priller, J, Henegariu, O, Theise, ND *et al.* 2002. Marrow-derived cells as vehicles for delivery of gene therapy to pulmonary epithelium. *Am. J. Respir. Cell Mol. Biol* **27:** 645-651.

Gruenert, DC, Bruscia, E, Novelli, G, Colosimo, A, Dallapiccola, B *et al.* 2003. Sequence-specific modification of genomic DNA by small DNA fragments. *J. Clin. Invest* **112:** 637-641.

Gu, H, Zou, YR, and Rajewsky, K. 1993. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73:** 1155-1164.

Gussoni, E, Soneoka, Y, Strickland, CD, Buzney, EA, Khan, MK *et al.* 1999. Dystrophin expression in the mdx mouse restored by stem cell transplantation. *Nature* **401:** 390-394.

Hacein-Bey-Abina, S, von Kalle, C, Schmidt, M, Le Deist, F, Wulffraat, N *et al.* 2003. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* **348:** 255-256.

Hall, PA and Watt, FM. 1989. Stem cells: the generation and maintenance of cellular diversity. *Development* **106:** 619-633.

Hayashi, S and McMahon, AP. 2002. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev. Biol.* **244:** 305-318.

Hennecke, M, Kwissa, M, Metzger, K, Oumard, A, Kroger, A *et al.* 2001. Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. *Nucleic Acids Res.* **29:** 3327-3334.

Herrera, PL. 2000. Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. *Development* **127:** 2317-2322.

Hertz, MI, Taylor, DO, Trulock, EP, Boucek, MM, Mohacsi, PJ *et al.* 2002. The registry of the international society for heart and lung transplantation: nineteenth official report-2002. *J. Heart Lung Transplant.* **21:** 950-970.

Ho, W and Furst, A. 1973. Intratracheal instillation method for mouse lungs. *Oncology* **27:** 385-393.

Hogan, BL. 1999. Morphogenesis. *Cell* **96:** 225-233.

Hogan, BL and Yingling, JM. 1998. Epithelial/mesenchymal interactions and branching morphogenesis of the lung. *Curr. Opin. Genet. Dev.* **8:** 481-486.

Hong, KU, Reynolds, SD, Giangreco, A, Hurley, CM, and Stripp, BR. 2001. Clara cell secretory protein-expressing cells of the airway neuroepithelial body microenvironment include a label-retaining subset and are critical for epithelial renewal after progenitor cell depletion. *Am. J Respir. Cell Mol. Biol.* **24:** 671-681.

Hong, KU, Reynolds, SD, Watkins, S, Fuchs, E, and Stripp, BR. 2003. In vivo differentiation potential of tracheal basal cells: Evidence for multipotent and unipotent subpopulations. *Am. J. Physiol Lung Cell Mol. Physiol* (in press)

Houdebine, LM and Attal, J. 1999. Internal ribosome entry sites (IRESs): reality and use. *Transgenic Res.* **8:** 157-177.

Howard, M, Frizzell, RA, and Bedwell, DM. 1996. Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations. *Nat. Med.* **2:** 467-469.

Hoyt, RF, Jr., Sorokin, SP, McDowell, EM, and McNelly, NA. 1993. Neuroepithelial bodies and growth of the airway epithelium in developing hamster lung. *Anat Rec.* **236:** 15-22.

Huber, O, Korn, R, McLaughlin, J, Ohsugi, M, Herrmann, BG *et al.* 1996. Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. *Mech. Dev.* **59:** 3-10.

Huelsken, J, Vogel, R, Erdmann, B, Cotsarelis, G, and Birchmeier, W. 2001. beta-Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin. *Cell* **105:** 533-545.

Igoucheva, O, Alexeev, V, and Yoon, K. 2001. Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther.* **8:** 391-399.

Imai, T, Jiang, M, Chambon, P, and Metzger, D. 2001a. Impaired adipogenesis and lipolysis in the mouse upon selective ablation of the retinoid X receptor alpha mediated by a tamoxifen-inducible chimeric Cre recombinase (Cre-ERT2) in adipocytes. *Proc. Natl. Acad. Sci. U. S. A* **98:** 224-228.

Imai, T, Jiang, M, Kastner, P, Chambon, P, and Metzger, D. 2001b. Selective ablation of retinoid X receptor alpha in hepatocytes impairs their lifespan and regenerative capacity. *Proc. Natl. Acad. Sci. U. S. A* **98:** 4581-4586.

Inayama, Y, Hook, GE, Brody, AR, Cameron, GS, Jetten, AM *et al.* 1988. The differentiation potential of tracheal basal cells. *Lab Invest* **58:** 706-717.

Indra, AK, Warot, X, Brocard, J, Bornert, JM, Xiao, JH *et al.* 1999. Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen- inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res* **27:** 4324-4327.

Inglis, SK, Corboz, MR, Taylor, AE, and Ballard, ST. 1997. Effect of anion transport inhibition on mucus secretion by airway submucosal glands. *Am. J. Physiol* **272:** L372-L377.

Innes, BA and Dorin, JR. 2001. Submucosal gland distribution in the mouse has a genetic determination localized on chromosome 9. *Mamm. Genome* **12:** 124-128.

Inoue, N, Dong, R, Hirata, RK, and Russell, DW. 2001. Introduction of single base substitutions at homologous chromosomal sequences by adeno-associated virus vectors. *Mol. Ther.* **3:** 526-530.

Ivanova, NB, Dimos, JT, Schaniel, C, Hackney, JA, Moore, KA *et al.* 2002. A stem cell molecular signature. *Science* **298:** 601-604.

Jackson, KA, Mi, T, and Goodell, MA. 1999. Hematopoietic potential of stem cells isolated from murine skeletal muscle. *Proc. Natl. Acad. Sci. U. S. A* **96:** 14482-14486.

Jacob, J and Baltimore, D. 1999. Modelling T-cell memory by genetic marking of memory T cells in vivo. *Nature* **399:** 593-597.

Jan, YN and Jan, LY. 1998. Asymmetric cell division. *Nature* **392:** 775-778.

Jiang, CK, Epstein, HS, Tomic, M, Freedberg, IM, and Blumenberg, M. 1991. Functional comparison of the upstream regulatory DNA sequences of four human epidermal keratin genes. *J. Invest Dermatol.* **96:** 162-167.

Jiang, X, Rowitch, DH, Soriano, P, McMahon, AP, and Sucov, HM. 2000. Fate of the mammalian cardiac neural crest. *Development* **127:** 1607-1616.

Jiang, Y, Jahagirdar, BN, Reinhardt, RL, Schwartz, RE, Keene, CD *et al.* 2002. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* **418:** 41-49.

Johnson, NF and Hubbs, AF. 1990. Epithelial progenitor cells in the rat trachea. *Am. J Respir. Cell Mol. Biol.* **3:** 579-585.

Jones, PH and Watt, FM. 1993. Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. *Cell* **73:** 713-724.

Kapsa, R, Quigley, A, Lynch, GS, Steeper, K, Kornberg, AJ *et al.* 2001. In vivo and in vitro correction of the mdx dystrophin gene nonsense mutation by short-fragment homologous replacement. *Hum. Gene Ther.* **12:** 629-642.

Keenan, KP, Combs, JW, and McDowell, EM. 1982. Regeneration of hamster tracheal epithelium after mechanical injury. I. Focal lesions: quantitative morphologic study of cell proliferation. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **41:** 193-214.

Kellendonk, C, Tronche, F, Casanova, E, Anlag, K, Opherk, C *et al.* 1999. Inducible site-specific recombination in the brain. *J. Mol. Biol.* **285:** 175-182.

Kellendonk, C, Tronche, F, Monaghan, AP, Angrand, PO, Stewart, F *et al.* 1996. Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res.* **24:** 1404-1411.

Kimmel, RA, Turnbull, DH, Blanquet, V, Wurst, W, Loomis, CA *et al.* 2000. Two lineage boundaries coordinate vertebrate apical ectodermal ridge formation. *Genes Dev.* **14:** 1377-1389.

Kisseberth, WC, Brettingen, NT, Lohse, JK, and Sandgren, EP. 1999. Ubiquitous expression of marker transgenes in mice and rats. *Dev. Biol.* **214:** 128-138.

Kmiec, EB. 1999. Targeted gene repair. *Gene Ther.* **6:** 1-3.

Kmiec, EB. 2003. Targeted gene repair -- in the arena. *J. Clin. Invest* **112:** 632-636.

Kocher, AA, Schuster, MD, Szabolcs, MJ, Takuma, S, Burkhoff, D *et al.* 2001. Neovascularization of ischemic myocardium by human bone-marrow-derived angioblasts prevents cardiomyocyte apoptosis, reduces remodeling and improves cardiac function. *Nat. Med.* **7:** 430-436.

Korinek, V, Barker, N, Moerer, P, van Donselaar, E, Huls, G *et al.* 1998. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat. Genet.* **19:** 379-383.

Kotani, H, Germann, MW, Andrus, A, Vinayak, R, Mullah, B *et al.* 1996. RNA facilitates RecA-mediated DNA pairing and strand transfer between molecules bearing limited regions of homology. *Mol. Gen. Genet.* **250:** 626-634.

Kotton, DN, Ma, BY, Cardoso, WV, Sanderson, EA, Summer, RS *et al.* 2001. Bone marrow-derived cells as progenitors of lung alveolar epithelium. *Development* **128:** 5181-5188.

Krause, DS, Theise, ND, Collector, MI, Henegariu, O, Hwang, S *et al.* 2001. Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* **105:** 369-377.

Kren, BT, Bandyopadhyay, P, and Steer, CJ. 1998. In vivo site-directed mutagenesis of the factor IX gene by chimeric RNA/DNA oligonucleotides. *Nat. Med.* **4:** 285-290.

Kren, BT, Cole-Strauss, A, Kmiec, EB, and Steer, CJ. 1997. Targeted nucleotide exchange in the alkaline phosphatase gene of HuH-7 cells mediated by a chimeric RNA/DNA oligonucleotide. *Hepatology* **25:** 1462-1468.

Kren, BT, Parashar, B, Bandyopadhyay, P, Chowdhury, NR, Chowdhury, JR *et al.* 1999. Correction of the UDP-glucuronosyltransferase gene defect in the gunn rat model of crigler-najjar syndrome type I with a chimeric oligonucleotide. *Proc. Natl. Acad. Sci. U. S A* **96:** 10349-10354.

Kunzelmann, K, Legendre, JY, Knoell, DL, Escobar, LC, Xu, Z *et al.* 1996. Gene targeting of CFTR DNA in CF epithelial cells. *Gene Ther.* **3:** 859-867.

Lagasse, E, Connors, H, Al Dhalimy, M, Reitsma, M, Dohse, M *et al.* 2000. Purified hematopoietic stem cells can differentiate into hepatocytes in vivo. *Nat. Med.* **6:** 1229-1234.

Lakso, M, Pichel, JG, Gorman, JR, Sauer, B, Okamoto, Y *et al.* 1996. Efficient in vivo manipulation of mouse genomic sequences at the zygote stage. *Proc. Natl. Acad. Sci. U. S. A* **93:** 5860-5865.

Lakso, M, Sauer, B, Mosinger, B, Jr., Lee, EJ, Manning, RW *et al.* 1992. Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl. Acad. Sci. U. S. A* **89:** 6232-6236.

Lallemand, Y, Luria, V, Haffner-Krausz, R, and Lonai, P. 1998. Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. *Transgenic Res.* **7:** 105-112.

Lamb, D and Reid, L. 1968. Mitotic rates, goblet cell increase and histochemical changes in mucus in rat bronchial epithelium during exposure to sulphur dioxide. *J. Pathol. Bacteriol.* **96:** 97-111.

Larcher, F, Del Rio, M, Serrano, F, Segovia, JC, Ramirez, A *et al.* 2001. A cutaneous gene therapy approach to human leptin deficiencies: correction of the murine ob/ob phenotype using leptin-targeted keratinocyte grafts. *FASEB J.* **15:** 1529-1538.

Lau, S, Jardine, K, and McBurney, MW. 1999. DNA methylation pattern of a tandemly repeated LacZ transgene indicates that most copies are silent. *Dev. Dyn.* **215:** 126-138.

Lavker, RM and Sun, TT. 1982. Heterogeneity in epidermal basal keratinocytes: morphological and functional correlations. *Science* **215:** 1239-1241.

Lavker, RM and Sun, TT. 1983. Epidermal stem cells. *J. Invest Dermatol.* **81:** 121s-127s.

Le, Y, Gagneten, S, Tombaccini, D, Bethke, B, and Sauer, B. 1999. Nuclear targeting determinants of the phage P1 cre DNA recombinase. *Nucleic Acids Res.* **27:** 4703-4709.

Leask, A, Rosenberg, M, Vassar, R, and Fuchs, E. 1990. Regulation of a human epidermal keratin gene: sequences and nuclear factors involved in keratinocyte-specific transcription. *Genes Dev.* **4:** 1985-1998.

Lehrer, MS, Sun, TT, and Lavker, RM. 1998. Strategies of epithelial repair: modulation of stem cell and transit amplifying cell proliferation. *J. Cell Sci.* **111 (Pt 19):** 2867-2875.

Lehtonen, E. 1987. Cytokeratins in oocytes and preimplantation embryos of the mouse. *Curr. Top. Dev. Biol.* **22:** 153-173.

Leigh, MW, Gambling, TM, Carson, JL, Collier, AM, Wood, RE *et al.* 1986. Postnatal development of tracheal surface epithelium and submucosal glands in the ferret. *Exp. Lung Res.* **10:** 153-169.

Lemischka, I. 2002. A few thoughts about the plasticity of stem cells. *Exp. Hematol.* **30:** 848-852.

Lersch, R, Stellmach, V, Stocks, C, Giudice, G, and Fuchs, E. 1989. Isolation, sequence, and expression of a human keratin K5 gene: transcriptional regulation of keratins and insights into pairwise control. *Mol. Cell Biol.* **9:** 3685-3697.

Lewandoski, M, Wassarman, KM, and Martin, GR. 1997. Zp3-cre, a transgenic mouse line for the activation or inactivation of loxP-flanked target genes specifically in the female germ line. *Curr. Biol.* **7:** 148-151.

Lieber, M, Smith, B, Szakal, A, Nelson-Rees, W, and Todaro, G. 1976. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *Int. J. Cancer* **17:** 62-70.

Liu, L, Cheng, S, van Brabant, AJ, and Kmiec, EB. 2002. Rad51p and Rad54p, but not Rad52p, elevate gene repair in Saccharomyces cerevisiae directed by modified single-stranded oligonucleotide vectors. *Nucleic Acids Res.* **30:** 2742-2750.

Liu, L, Parekh-Olmedo, H, and Kmiec, EB. 2003. The development and regulation of gene repair. *Nat. Rev. Genet.* **4:** 679-689.

Liu, L, Rice, MC, and Kmiec, EB. 2001. In vivo gene repair of point and frameshift mutations directed by chimeric RNA/DNA oligonucleotides and modified single-stranded oligonucleotides. *Nucleic Acids Res.* **29:** 4238-4250.

Lloyd, C, Yu, QC, Cheng, J, Turksen, K, Degenstein, L *et al.* 1995. The basal keratin network of stratified squamous epithelia: defining K15 function in the absence of K14. *J. Cell Biol.* **129:** 1329-1344.

Lobe, CG, Koop, KE, Kreppner, W, Lomeli, H, Gertsenstein, M *et al.* 1999. Z/AP, a double reporter for cre-mediated recombination. *Dev. Biol.* **208:** 281-292.

Lobe, CG and Nagy, A. 1998. Conditional genome alteration in mice. *Bioessays* **20:** 200-208.

Loeffler, M, Birke, A, Winton, D, and Potten, C. 1993. Somatic mutation, monoclonality and stochastic models of stem cell organization in the intestinal crypt. *J. Theor. Biol.* **160:** 471-491.

Loeffler, M. and Potten, C. S. 1997. Stem cells and cellular pedigrees - a conceptual introduction *in Stem Cells*. Academic Press. London

Lu, IL, Lin, CY, Lin, SB, Chen, ST, Yeh, LY *et al.* 2003. Correction/mutation of acid alpha-D-glucosidase gene by modified single-stranded oligonucleotides: in vitro and in vivo studies. *Gene Ther.* **10:** 1910-1916.

Lyle, S, Christofidou-Solomidou, M, Liu, Y, Elder, DE, Albelda, S *et al.* 1998. The C8/144B monoclonal antibody recognizes cytokeratin 15 and defines the location of human hair follicle stem cells. *J Cell Sci.* **111 (Pt 21):** 3179-3188.

Magdaleno, SM, Barrish, J, Finegold, MJ, and DeMayo, FJ. 1998. Investigating stem cells in the lung. *Adv. Pediatr.* **45:** 363-396.

Manuvakhova, M, Keeling, K, and Bedwell, DM. 2000. Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system. *RNA.* **6:** 1044-1055.

Manzano, A, Mohri, Z, Sperber, G, Ogris, M, Graham, I *et al.* 2003. Failure to generate atheroprotective apolipoprotein AI phenotypes using synthetic RNA/DNA oligonucleotides (chimeraplasts). *J. Gene Med.* **5:** 795-802.

Mao, X, Fujiwara, Y, and Orkin, SH. 1999. Improved reporter strain for monitoring Cre recombinase-mediated DNA excisions in mice. *Proc. Natl. Acad. Sci. U. S. A* **96:** 5037-5042.

Marchuk, D, McCrohon, S, and Fuchs, E. 1985. Complete sequence of a gene encoding a human type I keratin: sequences homologous to enhancer elements in the regulatory region of the gene. *Proc. Natl. Acad. Sci. U. S. A* **82:** 1609-1613.

Mason, RJ, Williams, MC, Moses, HL, Mohla, S, and Berberich, MA. 1997. Stem cells in lung development, disease, and therapy. *Am. J Respir. Cell Mol. Biol.* **16:** 355-363.

Mathis, L and Nicolas, JF. 2000. Different clonal dispersion in the rostral and caudal mouse central nervous system. *Development* **127:** 1277-1290.

McDowell, EM, Newkirk, C, and Coleman, B. 1985. Development of hamster tracheal epithelium: II. Cell proliferation in the fetus. *Anat. Rec.* **213:** 448-456.

Mercer, RR, Russell, ML, Roggli, VL, and Crapo, JD. 1994. Cell number and distribution in human and rat airways. *Am. J. Respir. Cell Mol. Biol.* **10:** 613-624.

Metzger, D and Chambon, P. 2001. Site- and time- specific gene targeting in the mouse. *Methods* **24:** 71-80.

Metzger, D, Clifford, J, Chiba, H, and Chambon, P. 1995. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc. Natl. Acad. Sci. U. S. A* **92:** 6991-6995.

Meyrick, B and Reid, L. 1970. Ultrastructure of cells in the human bronchial submucosal glands. *J. Anat.* **107:** 281-299.

Meyrick, B, Sturgess, JM, and Reid, L. 1969. A reconstruction of the duct system and secretory tubules of the human bronchial submucosal gland. *Thorax* **24:** 729-736.

Michel, M, Torok, N, Godbout, MJ, Lussier, M, Gaudreau, P *et al.* 1996. Keratin 19 as a biochemical marker of skin stem cells in vivo and in vitro: keratin 19 expressing cells are differentially localized in function of anatomic sites, and their number varies with donor age and culture stage. *J. Cell Sci.* **109 (Pt 5):** 1017-1028.

Milot, E, Strouboulis, J, Trimborn, T, Wijgerde, M, de Boer, E *et al.* 1996. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* **87:** 105-114.

Minamino, T, Gaussin, V, DeMayo, FJ, and Schneider, MD. 2001. Inducible gene targeting in postnatal myocardium by cardiac-specific expression of a hormone-activated Cre fusion protein. *Circ. Res.* **88:** 587-592.

Mizuguchi, H, Xu, Z, Ishii-Watabe, A, Uchida, E, and Hayakawa, T. 2000. IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Mol. Ther.* **1:** 376-382.

Moll, R, Franke, WW, Schiller, DL, Geiger, B, and Krepler, R. 1982. The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell* **31:** 11-24.

Morris, RJ and Potten, CS. 1994. Slowly cycling (label-retaining) epidermal cells behave like clonogenic stem cells in vitro. *Cell Prolif.* **27:** 279-289.

Morrison, SJ, Shah, NM, and Anderson, DJ. 1997. Regulatory mechanisms in stem cell biology. *Cell* **88:** 287-298.

Nagy, A. 2000. Cre recombinase: the universal reagent for genome tailoring. *Genesis.* **26:** 99-109.

Neganova, I, Augustin, M, Eichenlaub-Ritter, U, and Jockusch, H. 2001. Maternal versus paternal expression of a LacZ transgene in preimplantation mouse embryos: effects of genetic background and 2-cell block. *Zygote.* **9:** 219-228.

Nelson, WG and Sun, TT. 1983. The 50- and 58-kdalton keratin classes as molecular markers for stratified squamous epithelia: cell culture studies. *J. Cell Biol.* **97:** 244-251.

Neuringer, IP, Aris, RM, Burns, KA, Bartolotta, TL, Chalermskulrat, W *et al.* 2002. Epithelial kinetics in mouse heterotopic tracheal allografts. *Am. J. Transplant.* **2:** 410-419.

Nickerson, HD and Colledge, WH. 2003. A comparison of gene repair strategies in cell culture using a lacZ reporter system. *Gene Ther.* **10:** 1584-1591.

Niwa, H, Yamamura, K, and Miyazaki, J. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108:** 193-199.

Norris, PS, Jepsen, K, and Haas, M. 1998. High-titer MSCV-based retrovirus generated in the pCL acute virus packaging system confers sustained gene expression in vivo. *J. Virol. Methods* **75:** 161-167.

Novak, A, Guo, C, Yang, W, Nagy, A, and Lobe, CG. 2000. Z/EG, a double reporter mouse line that expresses enhanced green fluorescent protein upon Cre-mediated excision. *Genesis.* **28:** 147-155.

Ohtsuki, M, Tomic-Canic, M, Freedberg, IM, and Blumenberg, M. 1992. Nuclear proteins involved in transcription of the human K5 keratin gene. *J. Invest Dermatol.* **99:** 206-215.

Okabe, M, Ikawa, M, Kominami, K, Nakanishi, T, and Nishimune, Y. 1997. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* **407:** 313-319.

Oppenheimer, EH and Esterly, JR. 1975. Pathology of cystic fibrosis review of the literature and comparison with 146 autopsied cases. *Perspect. Pediatr. Pathol.* **2:** 241-278.

Orban, PC, Chui, D, and Marth, JD. 1992. Tissue- and site-specific DNA recombination in transgenic mice. *Proc. Natl. Acad. Sci. U. S. A* **89:** 6861-6865.

Orlic, D, Kajstura, J, Chimenti, S, Jakoniuk, I, Anderson, SM *et al.* 2001. Bone marrow cells regenerate infarcted myocardium. *Nature* **410:** 701-705.

Oshima, H, Rochat, A, Kedzia, C, Kobayashi, K, and Barrandon, Y. 2001. Morphogenesis and renewal of hair follicles from adult multipotent stem cells. *Cell* **104:** 233-245.

Otto, WR. 2002. Lung epithelial stem cells. *J. Pathol.* **197:** 527-535.

Pack, RJ, Al Ugaily, LH, and Morris, G. 1981. The cells of the tracheobronchial epithelium of the mouse: a quantitative light and electron microscope study. *J. Anat.* **132:** 71-84.

Paillard, F. 1997. Promoter attenuation in gene therapy: causes and remedies. *Hum. Gene Ther.* **8:** 2009-2010.

Palmiter, RD and Brinster, RL. 1986. Germ-line transformation of mice. *Annu. Rev. Genet.* **20:** 465-499.

Parkes, R, Meng, QH, Elena, SK, McEwan, JR, and Hart, SL. 2002. High efficiency transfection of porcine vascular cells in vitro with a synthetic vector system. *J. Gene Med.* **4:** 292-299.

Parsons, DW, Grubb, BR, Johnson, LG, and Boucher, RC. 1998. Enhanced in vivo airway gene transfer via transient modification of host barrier properties with a surface-active agent. *Hum. Gene Ther.* **9:** 2661-2672.

Peake, JL, Reynolds, SD, Stripp, BR, Stephens, KE, and Pinkerton, KE. 2000. Alteration of pulmonary neuroendocrine cells during epithelial repair of naphthalene-induced airway injury. *Am. J Pathol.* **156:** 279-286.

Perl, AK, Wert, SE, Nagy, A, Lobe, CG, and Whitsett, JA. 2002. Early restriction of peripheral and proximal cell lineages during formation of the lung. *Proc. Natl. Acad. Sci. U. S. A* **99:** 10482-10487.

Picard, D. 1994. Regulation of protein function through expression of chimaeric proteins. *Curr. Opin. Biotechnol.* **5:** 511-515.

Plopper, CG, Alley, JL, and Weir, AJ. 1986. Differentiation of tracheal epithelium during fetal lung maturation in the rhesus monkey Macaca mulatta. *Am. J Anat* **175:** 59-71.

Plopper, CG, Mariassy, AT, Wilson, DW, Alley, JL, Nishio, SJ *et al.* 1983. Comparison of nonciliated tracheal epithelial cells in six mammalian species: ultrastructure and population densities. *Exp. Lung Res.* **5:** 281-294.

Potten, C. S. 1978. Epithelial Proliferative Subpopulations *in Stem Cells and Tissue Homeostasis*. Cambridge University Press. Cambridge

Potten, CS and Loeffler, M. 1990. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110:** 1001-1020.

Poulsom, R, Forbes, SJ, Hodivala-Dilke, K, Ryan, E, Wyles, S *et al.* 2001. Bone marrow contributes to renal parenchymal turnover and regeneration. *J. Pathol.* **195:** 229-235.

Qin, L, Ding, Y, Pahud, DR, Chang, E, Imperiale, MJ *et al.* 1997. Promoter attenuation in gene therapy: interferon-gamma and tumor necrosis factor-alpha inhibit transgene expression. *Hum. Gene Ther.* **8:** 2019-2029.

Ramalho-Santos, M, Yoon, S, Matsuzaki, Y, Mulligan, RC, and Melton, DA. 2002. "Stemness": transcriptional profiling of embryonic and adult stem cells. *Science* **298:** 597-600.

Ramirez, A, Bravo, A, Jorcano, JL, and Vidal, M. 1994. Sequences 5' of the bovine keratin 5 gene direct tissue- and cell-type- specific expression of a lacZ gene in the adult and during development. *Differentiation* **58:** 53-64.

Ramirez, A, Milot, E, Ponsa, I, Marcos-Gutierrez, C, Page, A *et al.* 2001. Sequence and chromosomal context effects on variegated expression of keratin 5/lacZ constructs in stratified epithelia of transgenic mice. *Genetics* **158:** 341-350.

Randell, SH. 1992. Progenitor-progeny relationships in airway epithelium. *Chest* **101:** 11S-16S.

Randell, SH, Comment, CE, Ramaekers, FC, and Nettesheim, P. 1991. Properties of rat tracheal epithelial cells separated based on expression of cell surface alpha-galactosyl end groups. *Am. J. Respir. Cell Mol. Biol.* **4:** 544-554.

Rando, TA. 2002. Oligonucleotide-mediated gene therapy for muscular dystrophies. *Neuromuscul. Disord.* **12 Suppl 1:** S55-S60.

Rando, TA, Disatnik, MH, and Zhou, LZ. 2000. Rescue of dystrophin expression in mdx mouse muscle by RNA/DNA oligonucleotides. *Proc. Natl. Acad. Sci. U. S A* **97:** 5363-5368.

Rasheed, S, Nelson-Rees, WA, Toth, EM, Arnstein, P, and Gardner, MB. 1974. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* **33:** 1027-1033.

Reynolds, SD, Giangreco, A, Power, JH, and Stripp, BR. 2000a. Neuroepithelial bodies of pulmonary airways serve as a reservoir of progenitor cells capable of epithelial regeneration. *Am. J Pathol.* **156:** 269-278.

Reynolds, SD, Hong, KU, Giangreco, A, Mango, GW, Guron, C *et al.* 2000b. Conditional clara cell ablation reveals a self-renewing progenitor function of pulmonary neuroendocrine cells. *Am. J Physiol Lung Cell Mol. Physiol* **278:** L1256-L1263.

Rijnkels, M and Rosen, JM. 2001. Adenovirus-Cre-mediated recombination in mammary epithelial early progenitor cells. *J. Cell Sci.* **114:** 3147-3153.

Robles, AI, Larcher, F, Whalin, RB, Murillas, R, Richie, E *et al.* 1996. Expression of cyclin D1 in epithelial tissues of transgenic mice results in epidermal hyperproliferation and severe thymic hyperplasia. *Proc. Natl. Acad. Sci. U. S. A* **93:** 7634-7638.

Rochat, A, Kobayashi, K, and Barrandon, Y. 1994. Location of stem cells of human hair follicles by clonal analysis. *Cell* **76:** 1063-1073.

Rounbehler, RJ, Schneider-Broussard, R, Conti, CJ, and Johnson, DG. 2001. Myc lacks E2F1's ability to suppress skin carcinogenesis. *Oncogene* **20:** 5341-5349.

Rugh, R. 1990. Reproductive Systems of Adult Mice *in The Mouse: It's Reproduction and Development.* Oxford University Press. Oxford

Sakai, K and Miyazaki, J. 1997. A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochem. Biophys. Res. Commun.* **237:** 318-324.

Sanes, JR. 1994. Lineage tracing. The laatest in lineaage. *Curr. Biol.* **4:** 1162-1164.

Sangiuolo, F, Bruscia, E, Serafino, A, Nardone, A, Bonifazi, E *et al.* 2002. In vitro correction of cystic fibrosis epithelial cell lines by small fragment homologous replacement (SFHR) technique. *BMC. Med. Genet.* **3:** 8

Santana, E, Peritz, AE, Iyer, S, Uitto, J, and Yoon, K. 1998. Different frequency of gene targeting events by the RNA-DNA oligonucleotide among epithelial cells. *J Invest Dermatol.* **111:** 1172-1177.

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

Schermer, A, Galvin, S, and Sun, TT. 1986. Differentiation-related expression of a major 64K corneal keratin in vivo and in culture suggests limbal location of corneal epithelial stem cells. *J Cell Biol.* **103:** 49-62.

Scherrer, LC, Picard, D, Massa, E, Harmon, JM, Simons, SS, Jr. *et al.* 1993. Evidence that the hormone binding domain of steroid receptors confers hormonal control on chimeric proteins by determining their hormone- regulated binding to heat-shock protein 90. *Biochemistry* **32:** 5381-5386.

Schmidt, EV, Christoph, G, Zeller, R, and Leder, P. 1990. The cytomegalovirus enhancer: a pan-active control element in transgenic mice. *Mol. Cell Biol.* **10:** 4406-4411.

Schmidt, GH, Winton, DJ, and Ponder, BA. 1988. Development of the pattern of cell renewal in the crypt-villus unit of chimaeric mouse small intestine. *Development* **103:** 785-790.

Schoch, KG, Lori, A, Burns, KA, Eldred, T, Olsen, JC *et al.* 2003. A subset of mouse tracheal epithelial basal cells generates large colonies in vitro. *Am. J. Physiol Lung Cell Mol. Physiol* (in press).

Schweizer, J, Rentrop, M, Nischt, R, Kinjo, M, and Winter, H. 1988. The intermediate filament system of the keratinizing mouse forestomach epithelium: coexpression of keratins of internal squamous epithelia and of epidermal keratins in differentiating cells. *Cell Tissue Res.* **253:** 221-229.

Schwenk, F, Baron, U, and Rajewsky, K. 1995. A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23:** 5080-5081.

Schwenk, F, Kuhn, R, Angrand, PO, Rajewsky, K, and Stewart, AF. 1998. Temporally and spatially regulated somatic mutagenesis in mice. *Nucleic Acids Res.* **26:** 1427-1432.

Scott, ES, Goddard, CA, Wiseman, JW, Evans, MJ, and Colledge, WH. 2000. A murine tracheal culture system to investigate parameters affecting gene therapy for cystic fibrosis. *Gene Ther.* **7:** 612-618.

Seidman, MM and Glazer, PM. 2003. The potential for gene repair via triple helix formation. *J. Clin. Invest* **112:** 487-494.

Sharpe, J, Ahlgren, U, Perry, P, Hill, B, Ross, A *et al.* 2002. Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science* **296:** 541-545.

Shimizu, T, Nettesheim, P, Mahler, JF, and Randell, SH. 1991. Cell type-specific lectin staining of the tracheobronchial epithelium of the rat: quantitative studies with Griffonia simplicifolia I isolectin B4. *J. Histochem. Cytochem.* **39:** 7-14.

Shimizu, T, Nishihara, M, Kawaguchi, S, and Sakakura, Y. 1994. Expression of phenotypic markers during regeneration of rat tracheal epithelium following mechanical injury. *Am. J. Respir. Cell Mol. Biol.* **11:** 85-94.

Slack, JM. 2000. Stem cells in epithelial tissues. *Science* **287:** 1431-1433.

Soriano, P. 1999. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21:** 70-71.

Spradling, A, Drummond-Barbosa, D, and Kai, T. 2001. Stem cells find their niche. *Nature* **414:** 98-104.

Stellmach, V, Leask, A, and Fuchs, E. 1991. Retinoid-mediated transcriptional regulation of keratin genes in human epidermal and squamous cell carcinoma cells. *Proc. Natl. Acad. Sci. U. S. A* **88:** 4582-4586.

Stripp, BR, Maxson, K, Mera, R, and Singh, G. 1995. Plasticity of airway cell proliferation and gene expression after acute naphthalene injury. *Am. J. Physiol* **269:** L791-L799.

Sturgess, J and Imrie, J. 1982. Quantitative evaluation of the development of tracheal submucosal glands in infants with cystic fibrosis and control infants. *Am. J. Pathol.* **106:** 303-311.

Sullenger, BA. 2003. Targeted genetic repair: an emerging approach to genetic therapy. *J. Clin. Invest* **112:** 310-311.

Suzuki, A, Itami, S, Ohishi, M, Hamada, K, Inoue, T *et al.* 2003. Keratinocyte-specific Pten deficiency results in epidermal hyperplasia, accelerated hair follicle morphogenesis and tumor formation. *Cancer Res.* **63:** 674-681.

Tagalakis, AD, Graham, IR, Riddell, DR, Dickson, JG, and Owen, JS. 2001. Gene correction of the apolipoprotein (Apo) E2 phenotype to wild-type ApoE3 by in situ chimeraplasty. *J Biol. Chem.* **276:** 13226-13230.

Tam, PP and Snow, MH. 1981. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J. Embryol. Exp. Morphol.* **64:** 133-147.

Tarutani, M, Itami, S, Okabe, M, Ikawa, M, Tezuka, T *et al.* 1997. Tissue-specific knockout of the mouse Pig-a gene reveals important roles for GPI-anchored proteins in skin development. *Proc. Natl. Acad. Sci. U. S. A* **94:** 7400-7405.

Taubes, G. 2002. Gene therapy. The strange case of chimeraplasty. *Science* **298:** 2116-2120.

Taylor, G, Lehrer, MS, Jensen, PJ, Sun, TT, and Lavker, RM. 2000. Involvement of follicular stem cells in forming not only the follicle but also the epidermis. *Cell* **102:** 451-461.

Theise, ND, Henegariu, O, Grove, J, Jagirdar, J, Kao, PN *et al.* 2002. Radiation pneumonitis in mice: a severe injury model for pneumocyte engraftment from bone marrow. *Exp. Hematol.* **30:** 1333-1338.

Thomas, CE, Ehrhardt, A, and Kay, MA. 2003. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4:** 346-358.

Thorey, IS, Meneses, JJ, Neznanov, N, Kulesh, DA, Pedersen, RA *et al.* 1993. Embryonic expression of human keratin 18 and K18-beta-galactosidase fusion genes in transgenic mice. *Dev. Biol* **160:** 519-534.

Thorpe, P, Stevenson, BJ, and Porteous, DJ. 2002a. Optimising gene repair strategies in cell culture. *Gene Ther.* **9:** 700-702.

Thorpe, PH, Stevenson, BJ, and Porteous, DJ. 2002b. Functional correction of episomal mutations withshort DNA fragments and RNA-DNA oligonucleotides. *J. Gene Med.* **4:** 195-204.

Thurlbeck, WM, Benjamin, B, and Reid, L. 1961. Development and distribution of mucous glands in the foetal human trachea. *Br. J. Dis. Chest* **55:** 54-64.

Toma, JG, Akhavan, M, Fernandes, KJ, Barnabe-Heider, F, Sadikot, A *et al.* 2001. Isolation of multipotent adult stem cells from the dermis of mammalian skin. *Nat. Cell Biol.* **3:** 778-784.

Tos, M. 1966. Development of the tracheal glands in man. Number, density, structure, shape, and distribution of mucous glands elucidated by quantitative studies of whole mounts. *Acta Pathol. Microbiol. Scand.* **68:** Suppl

Tos, M. 1968. Development of the tracheal glands in man. *Dan. Med. Bull.* **15:** 206-215.

Tronche, F, Casanova, E, Turiault, M, Sahly, I, and Kellendonk, C. 2002. When reverse genetics meets physiology: the use of site-specific recombinases in mice. *FEBS Lett.* **529:** 116-121.

Trosko, JE. 2003. The role of stem cells and gap junctional intercellular communication in carcinogenesis. *J. Biochem. Mol. Biol.* **36:** 43-48.

Uhal, BD. 1997. Cell cycle kinetics in the alveolar epithelium. *Am. J. Physiol* **272:** L1031-L1045.

van der Kooy, D and Weiss, S. 2000. Why stem cells? *Science* **287:** 1439-1441.

van der Steege, G, Schuilenga-Hut, PH, Buys, CH, Scheffer, H, Pas, HH *et al.* 2001. Persistent failures in gene repair. *Nat. Biotechnol.* **19:** 305-306.

van Genderen, C, Okamura, RM, Farinas, I, Quo, RG, Parslow, TG *et al.* 1994. Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes Dev.* **8:** 2691-2703.

Van Winkle, LS, Buckpitt, AR, Nishio, SJ, Isaac, JM, and Plopper, CG. 1995. Cellular response in naphthalene-induced Clara cell injury and bronchiolar epithelial repair in mice. *Am. J. Physiol* **269:** L800-L818.

Vasioukhin, V, Degenstein, L, Wise, B, and Fuchs, E. 1999. The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc. Natl. Acad. Sci. U. S A* **96:** 8551-8556.

Vassar, R, Rosenberg, M, Ross, S, Tyner, A, and Fuchs, E. 1989. Tissue-specific and differentiation-specific expression of a human K14 keratin gene in transgenic mice. *Proc. Natl. Acad. Sci. U. S. A* **86:** 1563-1567.

Verrou, C, Zhang, Y, Zurn, C, Schamel, WW, and Reth, M. 1999. Comparison of the tamoxifen regulated chimeric Cre recombinases MerCreMer and CreMer. *Biol. Chem.* **380:** 1435-1438.

Wagers, AJ, Sherwood, RI, Christensen, JL, and Weissman, IL. 2002. Little evidence for developmental plasticity of adult hematopoietic stem cells. *Science* **297:** 2256-2259.

Wang, D, Russell, JL, and Johnson, DG. 2000. E2F4 and E2F1 have similar proliferative properties but different apoptotic and oncogenic properties in vivo. *Mol. Cell Biol.* **20:** 3417-3424.

Warburton, D, Schwarz, M, Tefft, D, Flores-Delgado, G, Anderson, KD *et al.* 2000. The molecular basis of lung morphogenesis. *Mech. Dev.* **92:** 55-81.

Watt, FM and Hogan, BL. 2000. Out of Eden: stem cells and their niches. *Science* **287:** 1427-1430.

Weber, P, Metzger, D, and Chambon, P. 2001. Temporally controlled targeted somatic mutagenesis in the mouse brain. *Eur. J. Neurosci.* **14:** 1777-1783.

Weibel, E., Hehr, P., Haies, D., Gil, J., and Bachofen, M. 1976. The cell population of the normal lung *in Lung Cells in Disease.* Elsevier/North-Holland Biomedical Press. Amsterdam-New York

Weimann, JM, Johansson, CB, Trejo, A, and Blau, HM. 2003. Stable reprogrammed heterokaryons form spontaneously in Purkinje neurons after bone marrow transplant. *Nat. Cell Biol.* **5:** 959-966.

Weiss, DJ, Liggitt, D, and Clark, JG. 1999. Histochemical discrimination of endogenous mammalian beta-galactosidase activity from that resulting from lac-Z gene expression. *Histochem. J* **31:** 231-236.

Wen, F, Cecena, G, Munoz-Ritchie, V, Fuchs, E, Chambon, P *et al.* 2003. Expression of conditional cre recombinase in epithelial tissues of transgenic mice. *Genesis.* **35:** 100-106.

Wong, MH, Saam, JR, Stappenbeck, TS, Rexer, CH, and Gordon, JI. 2000. Genetic mosaic analysis based on Cre recombinase and navigated laser capture microdissection. *Proc. Natl. Acad. Sci. U. S. A* **97:** 12601-12606.

Wu, X, Wu, J, Huang, J, Powell, WC, Zhang, J *et al.* 2001. Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. *Mech. Dev.* **101:** 61-69.

Wu, YJ and Rheinwald, JG. 1981. A new small (40 kd) keratin filament protein made by some cultured human squamous cell carcinomas. *Cell* **25:** 627-635.

Wunderlich, FT, Wildner, H, Rajewsky, K, and Edenhofer, F. 2001. New variants of inducible Cre recombinase: a novel mutant of Cre-PR fusion protein exhibits enhanced sensitivity and an expanded range of inducibility. *Nucleic Acids Res.* **29:** E47

Xie, T and Spradling, AC. 2000. A niche maintaining germ line stem cells in the Drosophila ovary. *Science* **290:** 328-330.

Yamashita, YM, Jones, DL, and Fuller, MT. 2003. Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. *Science* **301:** 1547-1550.

Ying, QL, Nichols, J, Evans, EP, and Smith, AG. 2002. Changing potency by spontaneous fusion. *Nature* **416:** 545-548.

Zambrowicz, BP, Imamoto, A, Fiering, S, Herzenberg, LA, Kerr, WG *et al.* 1997. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl. Acad. Sci. U. S A* **94:** 3789-3794.

Zhang, Y, Riesterer, C, Ayrall, AM, Sablitzky, F, Littlewood, TD *et al.* 1996. Inducible site-directed recombination in mouse embryonic stem cells. *Nucleic Acids Res.* **24:** 543-548.

Zhang, Z, Eriksson, M, Falk, G, Graff, C, Presnell, SC *et al.* 1998. Failure to achieve gene conversion with chimeric circular oligonucleotides: potentially misleading PCR artifacts observed. *Antisense Nucleic Acid Drug Dev.* **8:** 531-536.

Zhou, P, Byrne, C, Jacobs, J, and Fuchs, E. 1995. Lymphoid enhancer factor 1 directs hair follicle patterning and epithelial cell fate. *Genes Dev.* **9:** 700-713.

Zhou, Z, Wang, D, Wang, XJ, and Roop, DR. 2002. In utero activation of K5.CrePR1 induces gene deletion. *Genesis.* **32:** 191-192.

Zinyk, DL, Mercer, EH, Harris, E, Anderson, DJ, and Joyner, AL. 1998. Fate mapping of the mouse midbrain-hindbrain constriction using a site-specific recombination system. *Curr. Biol.* **8:** 665-668.