

# **MOUSE MODELS IN CUTANEOUS BIOLOGY**

**A thesis submitted to Monash University  
for the degree of Doctor of Philosophy**

**by  
TIA MARIE DITOMMASO**

**Department of  
Biochemistry & Molecular Biology  
February 2014**

**Notice 1**

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.



*This thesis is dedicated to my parents,  
for their unwavering love and support.*

## CONTENTS

<b>SUMMARY .....</b>	<b>1</b>
<b>DECLARATION .....</b>	<b>3</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>5</b>
<b>CHAPTER ONE: LITERATURE REVIEW.....</b>	<b>7</b>
1.1 INTRODUCTION.....	7
1.2 MOUSE GENOMICS AND GENETIC SEQUENCING .....	8
1.3 TYPES OF MOUSE MUTAGENEIS SCREENS.....	9
1.3.1 Forward Genetic Screens .....	9
1.3.2 Reverse genetic screens.....	11
1.3.3 Significance and strategy towards functional genome annotation .....	13
1.4 SANGER-MGP AT THE WTSI .....	15
1.5 WTSI HIGH THROUGHPUT PRODUCTION OF MUTANTS.....	15
1.5.1 WTSI Knockout first design: expression and conditional ready targeting .....	17
1.5.2 High throughput gene targeting .....	18
1.5.3 Vector design .....	19
1.5.4 Recombineering: bacteriophage-λ encoded red system .....	21
1.5.5 Final Vector Assembly.....	25
1.5.6 ES cell production .....	26
1.5.7 Genotyping ES cells .....	27
1.6 HIGH THROUGHPUT GENERATION OF TRANSGENIC MICE FROM TARGETED ES CELLS.....	28
1.6.1 Microinjection into embryos.....	28
1.6.2 Embryo transfer .....	28
1.6.3 Chimeras and germ line transmission.....	28
1.6.4 Genotyping mice.....	29
1.7 PHENOTYPING FOR DEVELOPMENTAL DEFECTS, DYSFUNCTION, AND DISEASE MODELS.....	31
1.7.1 Generation of Phenotyping cohorts .....	31
1.7.2 Broad preliminary phenotyping pipeline .....	32
1.7.3 Disease phenotype focus .....	32
1.8 EXPRESSION ANALYSIS .....	33
1.9 INTEGRATING ORGAN SPECIFIC PHENOTYPING SCREENS .....	34
1.10 SKIN.....	34
1.10.1 Introduction .....	34
1.10.2 Skin developmental defects.....	35
1.10.3 Phenotyping skin as a model system .....	37
1.11 DEVELOPMENTAL SKIN DEFECTS.....	38
1.11.1 Blistering .....	39
1.11.2 Barrier function.....	39
1.12 ADULT SKIN DEFECTS .....	41
1.12.1 Pigmentation.....	42
1.12.2 Hair cycling.....	43
1.12.3 Wound healing.....	44
1.13 LINES SELECTED FOR FURTHER STUDY .....	45
1.13.1 <i>Mad2</i> .....	45
1.13.2 <i>KRT76</i> .....	46
1.14 RESEARCH AIMS .....	48

<b>CHAPTER TWO: HIGH THROUGHPUT CUTANEOUS PHENOTYPING OF KNOCKOUT MOUSE MODELS.....</b>	<b>51</b>
2.1 INTRODUCTION .....	51
2.2 DECLARATION .....	56
2.3 MANUSCRIPT .....	57
<b>CHAPTER THREE: SPINDLE CHECKPOINT DEFICIENCY IN SKIN .....</b>	<b>77</b>
3.1 INTRODUCTION .....	77
3.2 DECLARATION .....	79
3.3 MANUSCRIPT .....	80
<b>CHAPTER FOUR: THE ROLE OF <i>KRT76</i> IN THE EPIDERMIS.....</b>	<b>93</b>
4.1 INTRODUCTION .....	93
4.2 DECLARATION .....	97
4.3 MANUSCRIPT .....	98
<b>CHAPTER 5: DISCUSSION .....</b>	<b>115</b>
<b>CHAPTER 6: CONCLUSIONS .....</b>	<b>123</b>
<b>CHAPTER 7: EXTENDED MATERIALS AND METHODS.....</b>	<b>125</b>
7.1 STANDARD MOLECULAR TECHNIQUES.....	125
7.1.1 Bacterial cell maintenance.....	125
7.1.2 Cloning of DNA fragments.....	125
7.1.3 Transformation of bacterial cells .....	125
7.1.4 Purification of DNA from bacterial cells.....	125
7.1.5 Preparation of genomic DNA for genotyping.....	125
7.1.6 RNA isolation.....	126
7.1.7 DNA and RNA quantification.....	126
7.1.8 DNase treatment of RNA .....	126
7.1.9 cDNA synthesis.....	126
7.1.10 Amplification of DNA by Polymerase Chain Reaction (PCR) .....	127
7.1.11 PCR screening of ligations .....	127
7.1.12 DNA sequencing.....	127
7.1.13 Restriction endonuclease digestion of DNA.....	128
7.1.14 Separation of nucleic acid by gel-electrophoresis .....	128
7.1.15 Protein quantification .....	128
7.1.16 Sodium dodecyl sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE).....	129
7.1.17 Western blotting .....	129
7.1.18 Recombinant Protein Expression: Bacterial Expression.....	130
7.1.19 Recombinant Protein Expression: Mammalian Expression.....	130
7.1.20 Immunofluorescence (IF) .....	131
7.1.21 Statistical analysis .....	131
7.1.22 Animal housing .....	131
7.2 GENERATION AND ANALYSIS OF TRANSGENIC MICE.....	131
7.2.1 Expression analysis by Reverse-Transcriptase PCR (RT-PCR) .....	131
7.2.2 Expression analysis by quantitative RT-PCR .....	132
7.2.3 Genotyping <i>Krt76</i> alleles.....	132

7.2.4 LacZ staining of embryos and adult mice.....	133
7.2.5 Wound-healing experiment and macroscopic wound analysis .....	134
7.2.6 Re-epithelialisation analysis days post-wounding .....	134
7.2.7 Cornified Envelope isolation .....	134
7.2.8 Toluidine blue barrier function assays .....	134
7.2.9 Trans epidermal water loss (TEWL) assays .....	134
7.2.10 Biotin tracer assays .....	135
7.2.11 Protein Fractionation Assays .....	135

<b>REFERENCES.....</b>	<b>137</b>
------------------------	------------

<b>APPENDIX I SUPPLEMENTARY INFORMATION FOR CHAPTER 2 .....</b>	<b>147</b>
---	------------

<b>APPENDIX II OTHER PUBLISHED WORKS.....</b>	<b>187</b>
---	------------

## SUMMARY

The recent advent of large scale reverse genetics and phenotyping projects has signalled a new era in the application of mouse genetics to understand gene function. International efforts to elucidate gene function are exemplified in The Sanger Mouse Genetics Project. This project has created a genetically mutant mouse resource of unprecedented size for the phenotype-based investigation of gene function *in vivo*. The integration of organ-specific screening pipelines is critical for the continued existence and maximization of reverse screening efforts to ascribe gene function. Using this approach, the analysis of the genetic complement to skin biology is examined. Phenotype data gathered as a part of this screen initiated the in depth investigation of the contribution of two genes, *Krt76* and *Mad2*, to skin barrier function and skin stem cells, respectively. This is the first and largest tissue-specific reverse genetic screen ever to be carried out, and accordingly, has identified novel genes and novel roles for known genes required for normal skin function.

At the genome level, a multi-test, multi-parameter reverse genetic screen was applied to over 500 different knockout mouse strains with the aim of identifying defects in development or homeostasis of the skin. The dynamic, indispensable, and accessible nature of the skin make it a valuable model system in which to study genetic contributions to its function and model cell biology. This screen identified more than 20 different affected lines, many of which carry mutations in genes that have not previously been implicated in the biology of the organ. This work highlights the promise of high throughput reverse genetic screens to give critical insights into the genetics of skin biology.

The regenerative properties of the skin are regulated by multiple stem cell populations that reside there, making the skin an excellent organ in which to model stem cell biology. *Mad2* is a known component of the spindle assembly checkpoint (SAC) whose loss in mitotically active cells leads to aneuploidy. Because aneuploidy and disruptions in spindle assembly are incompatible with early mouse development, *Mad2* was conditionally inactivated in the skin. This allowed for the *in vivo* investigation of *Mad2*, and showed that defects in the checkpoint are tolerated in the dividing cells of the interfollicular epidermis, but not in the stem cell population in the hair follicle. This study

is an elegant example combining mouse genetics and skin biology to understand the differential response to aneuploidy within a single tissue.

The healing and protective properties of the skin were examined using germline inactivation of *Keratin76* (*Krt76*). It was shown that Krt76, a poorly characterized intermediate filament protein, is absolutely required to maintain the integrity of the skin and is essential for postnatal survival. The progressive development of cutaneous wounds occurs on the background of a primary defect in the maintenance of the skin barrier. Here, a novel mechanism by which intermediate filaments interact with tight junction components to maintain barrier function in the skin is described.

## DECLARATION

In accordance with Monash University *Doctorate Regulation 17.2 Doctor of Philosophy and Research Master's* regulations, the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis is based on conjointly published and unpublished work. Some of the work in this thesis has been published:

Spindle checkpoint deficiency is tolerated by murine epidermal cells but not hair follicle stem cells.

Foijer F, DiTommaso T, Donati G, Hautaviita K, Xie S, Heath E, Smyth I, Watt FM, Sorger PK and Bradley A

Proc Natl Acad Sci U S A. 2013 Feb 19;110(8):2928-33. doi: 10.1073/pnas.1217388110

Genome-wide Generation and Systematic Phenotyping of Knockout Mice Reveals New Roles for Many Genes.

White JK, Gerdin AK, Karp N, Ryder E, Buljan M, Bussell JN, Salisbury J, Clare S, Ingham NJ, Podrini C, Houghton R, Estabel J, Bottomley JR, Melvin DG, Sunter D, Adams NC The Sanger Institute Mouse Genetics Project\*, Tannahill D, Logan DW, MacArthur DG, Flint J, Mahajan VB, Tsang SH, Smyth I, Watt FM, Skarnes WC, Dougan G, Adams DJ, Ramirez-Solis R, Bradley A, Steel KP.

\*The Sanger Institute Mouse Genetics Project Phenotyping team: Baker L, Barnes C, Beveridge R, Cambridge E, Carragher D, Chana P, Clarke K, Hooks Y, Igosheva N, Ismail O, Jackson L, Kane L, Lacey R, Tino Lafont D, Lucas M, Maguire S, McGill K, McIntyre RE, Messenger S, Mottram L, Mulderrig L, Pearson S, Protheroe HJ, Roberson L, Salisbury G, Sanderson M, Sanger D, Shannon C, Thompson P, Tuck E, Vancollie VE. Genotyping and targeting team: Brackenbury L, Bushell W, Cook R, Dalvi P, Gleeson D, Habib B, Hardy M, Liakath KA, Miklejewska E, Price S, Sethi D, Trenchard E, von Schiller D, Vyas S, West AP, Woodward J, Wynn E. Mouse Informatics: Evans A, Gannon D, Griffiths M, Holroyd S, Iyer V, Kipp C, Lewis M, Li W, Oakley D, Richardson D, Smedley D. Tissue culture team: Agu C, Bryant J, Delaney L, Gueorguieva NI, Tharagannet H, Townsend AJ. Mouse Production team: Biggs D, Brown E, Collinson A, Dumeau C, Grau E, Harrison S, Harrison J, Ingle CE, Kundi H, Madich A, Mayhew D, Metcalf T, Newman S, Pass J, Pearson L, Reynolds H, Sinclair C, Wardle-Jones H, Woods M. Mouse Facility team: Alexander L, Brown T, Flack F, Frost C, Griggs N, Hrnčiarova S, Kirton A, McDermott J, Rogerson C, White G, Zieleszinski P. Collaborators: DiTommaso T, Edwards A, Heath E, Mahajan MA, Yalcin B.

Cell. 2013 Jul 18;154(2):452-64. doi: 10.1016/j.cell.2013.06.022.

Some of the work in this thesis has been prepared for publication. The ideas, development and writing up of the following 2 the papers in this thesis were the principal responsibility

of myself, the candidate, working within the Department of Biochemistry and Molecular Biology under the supervision of Associate Professor Ian Smyth:

Identification of genes important for cutaneous function revealed by a large-scale reverse genetic screen in the mouse.

**DiTommaso T.** Lynelle Jones, The Sanger Mouse Genetics Project, Anna-Karin Gerdin, Valerie E. Vancollie, Fiona M. Watt, Ramiro Ramirez-Solis, Allan Bradley, Karen P. Steel, John P. Sundberg, Jacqueline K. White, Ian Smyth  
Submitted to PLoS Gen

Keratin 76 is required for tight junction function and maintenance of the skin barrier.

**DiTommaso T.** Denny L. Cottle<sup>1</sup>, Helen B. Pearson, Holger Schlüter, Pritinder Kaur, Patrick O. Humbert, and Ian Smyth  
Submitted to PLoS Gen

The core theme of this thesis is Mouse Models of Cutaneous Biology. The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of Chapters 2-4, my contribution to the work involved the following:

Thesis chapter	Publication title	Publication status*	Nature and extent of candidate's contribution
1	Identification of genes important for cutaneous function revealed by a large scale reverse genetic screen in the mouse	Under Review (PLOS Genetics)	performed research, analyzed data, wrote manuscript
2	Spindle checkpoint deficiency is tolerated by murine epidermal cells but not hair follicle stem cells	published	performed research, contributed new reagents/analytic tools, analyzed data
3	Keratin 76 is required for tight junction function and maintenance of the skin barrier	Under Review (PLOS Genetics)	performed research, analyzed data, wrote manuscript

I have / have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Signed:

Date: 6/6/2014



## ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Ian Smyth for his support, guidance, and friendship over the last five years. Not only did Ian give me my first real science job, he let me go to Cambridge and allowed me to speak ‘off the record’ many times throughout my PhD. The opportunities that Ian gave me as a student have allowed me to go on to get a great job back home, and I’m very grateful to him for all of his support while still challenging me to be an independent researcher. I would also like to thank Ian for always sharing his knowledge, music, retrospective dating advice, and especially, for being my unofficial next of kin throughout my time in Australia and for always making sure I had somewhere to go on Christmas.

To all of the friends I made here in Australia, both in and out of the lab, you have all helped me so much. Cecilia, I cannot imagine doing this without your help, guidance, and encouragement along the way. Jelena and Holly, I will never forget the awesome times we had, of course I will always remember our nights out as some of the best and funniest times I had in Australia. Jelena, I’m so glad that I had you as a work and yoga buddy, especially in my last year. Finally, my sincerest thanks and so much love goes out to the Margaret Street team: David, Lauren, Leah, and Rosh. You guys have been my family here, and you have no idea how much you have done for me. I could not think of a better group of people to come home to. Thank you for looking out for me, giving me advice, listening to me, including me, and teaching me to be Australian.

To the rest of the Smythonians, Kieran, Fenny, Sally, Lynelle, Jenny, Olga and Denny; thank you for sharing your knowledge and advice, and for helping me to stay positive. It has been a pleasure to work with you all. Thank you to Lynelle for listening to me vent, Kieran for always fixing my computer, Fenny for being a great mentor and Olga for her honesty. Thank you Jenny for all of your support, kindness, and generosity. I would especially like to thank Denny for helping me so much when it counted—at the beginning in Cambridge, and at the end in Melbourne. I’d also like to thank all of the Smyth honours students, Mal, Syafiq, Gloria, and Wendy, for being great additions to the lab.

The following people have also helped in my project in one way or another and deserve a thanks: all members of the Cole lab, Moira O'bryan and her entire lab, Luc Furic, Matt Mangan, Xiangting, Anzari, ARL staff, the MMI staff, the purchasing team, The Imaging facility, and Ian McPherson. Also thanks to my friends from home, especially Kelly, Renee, and Emily for always making time to see me when I came back to the US. And thank you Vikram for giving me so much support, and encouragement, and ideas, and above all, for believing in me.

Finally, thank you to my parents and sisters, Gina and Jesse. This would have been impossible without you. I can’t thank you enough for always being there for me, even though while I was so far away. I love you.



## CHAPTER ONE: LITERATURE REVIEW

### 1.1 INTRODUCTION

Despite the extensive knowledge of our genetic makeup, we remain largely ignorant of the function of most genes and their contributions to disease. The study of genetically linked disease in animal models has proven to be a highly effective method by which to determine gene function. The sequencing of the mouse and human genomes, as well as the comparative analysis between the two, have provided the research community with confirmation of the validity of using mouse models to study human disease. Over 99% of mouse genes have a human homologue, making the mouse the most widely accepted model in which to characterize mammalian gene function *in vivo*. Furthermore, the importance of mouse models in studying genetically linked disease has led to the realization that a comprehensive mouse resource, in which all genes are systematically targeted in mouse embryonic stem (ES) cells, is required.

Leading the effort to systematically manipulate specific genes in mouse ES cells is the International Knockout Mouse Consortium (IKMC). IKMC is an international collaboration whose members are working together to mutate all protein-coding genes in the mouse genome using a combination of gene trapping and gene targeting in ES cells. By combining advances in ES cell techniques with high throughput genetic targeting technologies, the IKMC aims to generate targeted ES cell lines for each of the approximately 20,000 genes in the mouse genome. The Wellcome Trust Sanger Institute (WTSI) has taken this endeavour further by pledging to generate and phenotype knockout mice from the ES cell bank. This phenotyping project at the WTSI, hereafter referred to as the Sanger-MGP, aims to functionally annotate mouse genes, generate mouse models of disease and facilitate the distribution of knockout mouse lines to the scientific research community.

I began this research candidature with a year long placement at the WTSI during which I was embedded in the Sanger-MGP to screen knockout mouse lines for defects in skin development and homeostasis. To identify genes governing skin biology, I used a high throughput screening strategy on the genetically altered mice generated by the WTSI. This multi-test, multi-parameter screening strategy was designed to rapidly and robustly identify perturbations in skin function and skin development. In this thesis, I discuss the

findings of the largest reverse genetic screen to ever be reported in a single organ. I go on to examine the novel roles of 2 genes which I identified in the screen, *Keratin76* and *Mad2*, in maintaining normal skin function.

The skin is one of the best characterised and most accessible organs in the human body and is a highly suitable model organ in which to study genetic determinants of disease and model cellular biology. The paradigms established in epidermal development are recapitulated throughout postnatal life, allowing the skin to constantly regenerate and thereby maintain its necessary functional capacities. Many of these functions are indispensable and include providing a protective barrier, protecting from dehydration, enabling wound healing and facilitating immune surveillance. When these functions fail to be maintained, the result is often evident. Thus, the dynamic nature of this organ, together with its accessibility and indispensability, make it particularly suitable for examination in high throughput screens.

## **1.2 MOUSE GENOMICS AND GENETIC SEQUENCING**

Experimental organisms play a crucial role in the elucidation of gene function. In particular, the laboratory mouse, *Mus musculus*, is an important tool to study mammalian physiology and model human disease. The first ever demonstration of Mendelian genetics in mammals was carried out using inheritance of coat colours in mouse<sup>1,2</sup>. Nearly a century later, the sequencing of both the mouse and human genomes has been completed, and there is a general consensus about the protein coding sequences and conservation of synteny between the two species<sup>3,4</sup>. Currently there are some 5,000 mammalian genes whose function has been investigated using mouse models. The next challenge is functionally annotating the remainder of the genes. Mouse mutagenesis and mouse phenotyping are central to addressing this challenge.

The initial interrogation of the mouse genome began in 1990 as part of the Human Genome Project in which the mouse was included as one of the 5 central model organisms. A dense genetic map of the mouse genome was generated, by integrating simple sequence length polymorphisms (SSLPs) with restriction fragment length polymorphisms (RFLPs)<sup>5,6</sup>. After over 5 years, the coverage of the mouse genome reached near entirety identifying 7377 genetic markers,<sup>7</sup> thereby marking the first phase of the Human Genome Project. From here, both the mouse and human genomes were

sequenced. In generating the genome sequences, a valuable toolkit to annotate gene function was also generated. This toolkit includes collections of expressed sequence tags (ESTs), complementary DNAs and bacterial artificial chromosome (BAC) libraries, thereby setting the foundation for positional cloning, mapping of genetic mutations, and gene targeting. The application of such tools allows researchers to functionally annotate the genome by linking genes and phenotypes.

In addition to determining the function of each gene, identifying the genetic basis for human disease is an intended outcome of mouse mutagenesis screens<sup>8</sup>. Much of our current understanding of genetics and disease comes from linking a dysfunctional gene to a disease. Thus, the concept of intentionally disrupting a gene and observing the outcome has become an important strategy in elucidating the genetic contribution to disease. The mouse is an excellent model of human disease, thereby making it essential to this research. In addition to genetic similarities with humans, the mouse has the same tissues and organs and shares much of its physiology with humans. Changes in mouse genes leading to disease can closely mimic changes seen in human diseases.

### **1.3 TYPES OF MOUSE MUTAGENEIS SCREENS**

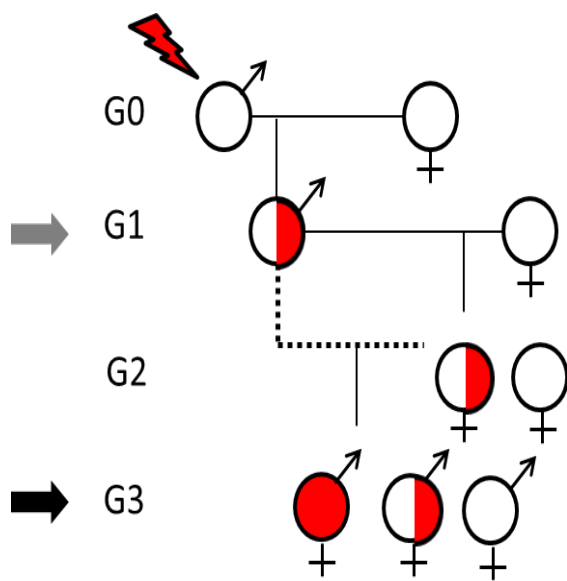
Mouse models are used to annotate gene function by disrupting a gene and in turn asking what effect this has on development and physiology. Mouse mutagenesis screens rely on both the rapid generation of mice and robust phenotyping practices. There are two approaches towards mouse mutagenesis screening: phenotype driven (forward) and genotype driven (reverse). Both types of screens have brought about new discoveries in gene function. This section will discuss the generation, assembly and execution of high throughput genetic screening in mice and demonstrate the value of mouse models in the investigation of human disease.

#### **1.3.1 Forward Genetic Screens**

Phenotype-driven or forward genetic screens are based on mutagenesis techniques which introduce random genome disruptions with the aim of generating a phenotype. The knowledge of the mouse and human genome sequences and the development of mapping and positional cloning techniques have made large, phenotype-driven mutagenesis screens possible. Such screens began in 1997 using N-ethyl-N-nitrosourea (ENU)<sup>9</sup>. ENU is a potent chemical mutagen which induces single-base pair substitution. However the

location of mutations induced with administration of ENU cannot be controlled, and may complicate the pinpointing of the responsible gene containing the base pair substitution<sup>10-12</sup>.

ENU screens can be carried out to detect dominant or recessive phenotypes. In both cases, the G0 male is injected with ENU, allowed to recover, and subsequently crossed with a wildtype G0 female. In dominant phenotype driven screens, it is preferable to cross the G0 male with a female from a different strain to enable linkage analysis. The G1 generation is screened for mutant phenotypes. Any animal showing a phenotype in the G1 generation becomes a mutant candidate and is used in subsequent matings to produce G2 mice. If the G1 potential mutant carries a heritable mutation, then the phenotype should be inherited at a 1:1 ratio in the G2 progeny. If the maternal G0 strain was different to the paternal strain, the G1 mutant is crossed to the G0 maternal strain, to produce the G2 strain. This makes mapping easier as the phenotype causing mutation should be segregated from the genetic markers of the maternal strain; ultimately allowing the use of genetic linkage information to narrow down the candidate genomic region containing the responsible mutation. In recessive phenotype screens, mutants are not identified until the G3 generation. The G1 males are crossed with wildtype females, producing a G2 generation. The G2 generation is expected to carry half the number of recessive mutations originally carried by the G1 males. G2 progeny can then be crossed with the G1 mutant candidate or crossed in brother-sister matings to produce the G3 generation.



**Figure 1.1 ENU breeding scheme.** A three-generation breeding scheme for recovering recessive traits in the third generation (G3, black arrow). Dominant traits can be detected in G1 (grey). G0 males are mutagenized and outcrossed to generate the G1 population. These are again outcrossed to generate G2 females, which may carry any given ENU mutation (red). These are backcrossed (dotted line) to the G1 male to create the G3 generation, which is then screened for phenotype(s). Solid Red indicates homozygosity for an ENU mutation, open indicates no mutation, solid/open indicates heterozygosity.

ENU is expected to induce 20-40 recessive mutations per G1 mouse, based on the specific-locus test<sup>13-15</sup>. Dominant phenotypes were reported in approximately 2-3% of G1 progeny by ENU screens in the UK<sup>8</sup>, and the value of ENU as a potent and effective mutagen encouraged researchers to conduct ENU screens. Although forward genetic ENU screens have produced many mutant lines, the difficulties associated with identification and cloning of the responsible gene can limit the efficiency of these screens<sup>9</sup>. The ability to identify phenotypes is a critical factor in successful mutant identification in ENU screens. Phenotyping strategies must therefore be reliable, thorough and reproducible.

Despite the challenges associated with forward genetic screens, ENU screens have generated valuable human disease models and accelerated the understanding of gene function. ENU screens not only validated the use of mutant mice in biomedical research, but also helped in the initial development of phenotyping procedures, mouse “clinics,” data management and mouse handling. Much of the infrastructure created as a result of forward genetic screens has been directly applied to, and incorporated in, reverse-genetic screens.

### **1.3.2 Reverse genetic screens**

Despite the usefulness of forward genetic screens employing ENU or other mutagens, the problem of systematic and functional annotation of the genome remains. Reverse genetic, or gene-driven screens allow for the directed, as opposed to random, disruption of gene function. Advances in genetic manipulation, together with BAC libraries and ES cell culturing techniques, have facilitated the implementation of reverse genetic screens. In this approach specific genes in ES cells are manipulated, thereby eliminating rigorous genome mapping required in forward genetic screens to identify the phenotype causing mutation. Additionally, reverse screens can be carried out systematically, knocking out particular genes in a specific and controlled manner.

Currently, the most widely used approach to knockout specific genes in the mouse is through homologous recombination in ES cells<sup>16</sup>. Using this technique, directed gene targeting can induce essentially any genome modification, from single base pair

mutations to the insertion of large cassettes<sup>17</sup>. Furthermore, the ability to target specific exons in a gene can circumvent issues with hypomorphic alleles, a common occurrence with random, single base pair disruptions such as those induced by ENU mutagenesis. Highly conserved regions of a gene are often necessary for function. The advent of predictive sequence tools to estimate domain function and highlight conserved elements helps researchers to determine the optimal location for targeting to fully disrupt gene function. The identification and removal of these highly conserved, functional domains through homologous recombination is an effective method to completely ablate gene function.

Not only does gene targeting allow for directed gene-specific manipulations, it also permits the temporal and spatial disruption of gene function with special targeting cassettes that give researchers the option to generate conditional mutants. This can be advantageous in investigating the function of genes with critical roles in embryonic development and whose global dysfunction results in embryonic lethality. The incidence of embryonic lethality in homozygous mutant mice is 30% (2,183 of 7,229 lines of mice; [ftp://ftp.informatics.jax.org/pub/reports/MGI\\_PhenoGenoMP.rpt](ftp://ftp.informatics.jax.org/pub/reports/MGI_PhenoGenoMP.rpt), February 2013), emphasizing the requirement for conditional gene inactivation in functional annotation of thousands of genes.

Gene targeting cassettes can also be engineered to contain elements that make it possible to track gene expression. This is important, especially in the study of novel genes, as conventional tools to investigate their function and track their expression have often not yet been developed. These tools include antibodies, validated TaqMan and *in situ* probes and published plasmids. Thus, expression data derived from 'built in' genetic manipulation can be used as a starting point to investigate gene expression and infer gene function.

To accelerate the knowledge of gene function, the WTSI and other consortia, have implemented visionary high throughput ES cell targeting pipelines, which in turn feed into reverse genetic mouse phenotyping screens. The ES cell bank, phenotyping data, and expression patterns generated at the WTSI are publically available to the research community to be used as a tool to functionally annotate the genome and to facilitate exploratory biomedical research on novel genes.



### 1.3.3 Significance and strategy towards functional genome annotation

The sequencing of both the mouse and human genomes has had a profound impact on both forward and reverse genetic screens, allowing for rapid mutation mapping and precise gene targeting, respectively. Surprisingly, a majority of research continues to focus on those genes whose function was known before the genomes were sequenced<sup>18</sup>. Increasing evidence suggests that these genes are not necessarily the most important and are not any more likely to be involved in disease. Many genes, despite being linked to disease, remain underrepresented in the published literature<sup>18</sup>. Multiple explanations have been put forth to justify this research bias, including reluctance and resistance from funding and peer reviewed systems to support research on unstudied proteins. Another deterrent for studying novel genes and proteins is the lack of tools, such as mouse models, available to investigate their function.

The most widely accepted way to determine gene function *in vivo* is through the generation of genetically modified mouse lines. The utility here lies in the supposition that disrupting the gene will indeed produce a phenotype, and that the phenotype can be detected. Unfortunately, our ability to predict the consequence of knocking out a gene is poor. Often, no effect is seen in response to loss of gene function, possibly owing to redundancy in gene function, insufficient characterization or a combination of both. Thus, individual laboratories may be hesitant to undertake the potentially laborious and expensive process required to generate a genetically mutant mouse line, especially when the outcome is unknown.

Since pleiotropic effects are a common occurrence in transgenic mice<sup>19</sup>, there is the probability that the phenotype will affect multiple organ systems. A specialized laboratory may not have the expertise to characterize or report unintended affected organs or associated secondary defects. Further, deleterious effects in one organ can mask the role of the gene in other organs by killing the animal before other organs can be assessed. The potential risks of producing a knockout mouse with no phenotype in the desired organ can be seen as a deterrent to generating a novel genetically modified mouse line.

To help shoulder some of the risks associated with studying new genes through knockout mouse models, a progressive and organized stance on generating a knockout mouse lines

as a resource must be implemented. The requirement for an unbiased annotation of all genes requires that genes be knocked out in a systematic manner, thereby making it necessary to use gene targeting and reverse genetic screening. Gene targeting allows for the use of versatile cassettes that can induce both spatial and temporal gene disruption. Creating a versatile knockout ES cell and mouse resource can maximize the genome, encourage the annotation of novel genes in mice, and generate new models in which to study disease.

Who bears the responsibility of promoting research on novel genes, and what is the best way to do so? For an endeavour of this magnitude, the responsibility lies with the greater scientific community as a whole, and can only be achieved by concerted effort across global biomedical science consortia. The effort is 2 fold, first generating the knockout mouse resource, and second, robustly reporting the consequence of gene knockout. The implementation of mouse mutagenesis screens has already generated new mouse models and continues to aid in the elucidation of gene function *in vivo*. Scaling up and broadening the focus of such mutagenesis screens is an unbiased way to encourage the study of novel and unappreciated genes. For the functional annotation of the entire genome, both the mouse resource as well as preliminary characterizations of the consequence of gene disruption should be publically available. For these reasons, maximizing the value of the genome sequence requires a global effort to mutate all coding genes and an effective way to publically and transparently report the resulting phenotypes.

A further benefit of large scale mutagenesis is streamlining the production of knockout lines, potentially limiting redundancy associated with having multiple labs mutating the same genes in the same manner. Genes can be mutated in many ways, some of which are more effective in ablating gene function than others. By disrupting genes, in a systematic manner, researchers can go deeper than gene annotation to identify specific exons that code for essential or non-essential elements in gene function. Conversely, when complete gene disruption does not elicit a phenotype, the availability of a large knockout mouse resource enables the generation of transgenic lines in which multiple genes and gene families can be simultaneously knocked out. This approach is especially important when redundancy in gene function is a contributing factor in undetectable or non-existent phenotypes.

Large scale mutagenesis screens have the potential to shoulder some of the risks involved in studying novel genes, allowing, at the very least, a broad idea of the consequence of knocking out a gene. Easy access to transgenic lines and preliminary phenotyping data also encourages expert labs to investigate gene function in this way. Accordingly, the generation of more mutant mice expands the pool of genes and gene families for which functional annotation is available, thereby promoting the study of understudied, but still important genes in development and disease.

#### **1.4 SANGER-MGP AT THE WTSI**

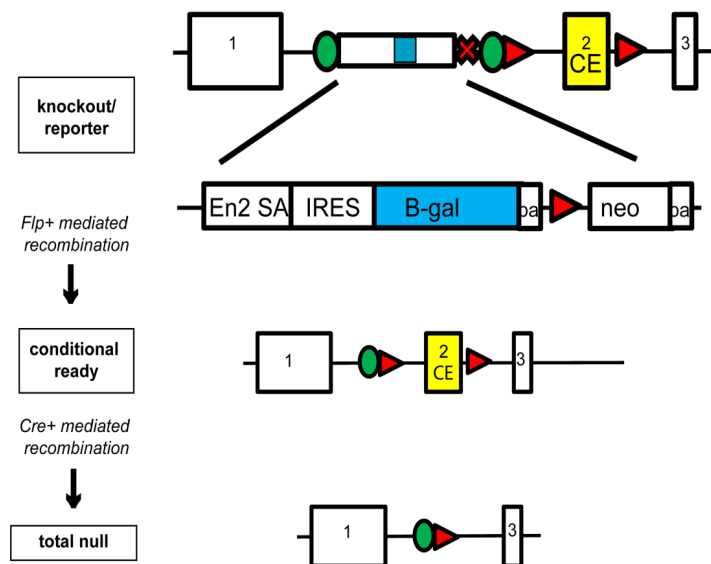
The overall aim of the Sanger-MGP is the high throughput generation and characterization of genetically altered mice. The first step in achieving this aim is to systematically disrupt gene function by adapting mutagenesis technologies to be applied in a high throughput manner. From here, the role of a gene in development, normal function, and disease is interrogated through methodical phenotypic characterization of these mice. These data are complimented by gene expression profiles in adults and at key developmental stages. Ultimately, the effects of a faulty gene are studied in detail to determine when, where and how the gene acts *in vivo*. Thus, the approach at the WTSI to functionally annotate the genome can be broken down into 3 parts: production of knockout mouse lines, phenotyping encompassing multiple organ systems, and expression analysis to determine where the gene acts. Each of the three parts is detailed in the following sections 1.5-1.7.

#### **1.5 WTSI HIGH THROUGHPUT PRODUCTION OF MUTANTS**

The Sanger-MGP is the major contributor in the global consortium to knockout genes in ES cells and derive phenotypic and expression data associated with the gene. To do this, multiple expert teams and a world class facility have been developed to create an assembly line-based method of generating knockout mice. To date, the Sanger-MGP together with other consortia, have generated knockout ES cells for over half of all known mouse genes<sup>17</sup>. As a research tool, these targeted ES cells will revolutionise the study of gene function, hopefully promoting new research by individual laboratories on the role of underrepresented genes. To further inspire the research community and to begin the functional annotation of the genome, the Sanger-MGP is converting significant numbers

of ES cell lines into mice for phenotypic analysis, implementing a high throughput reverse genetics screen in mice.

The first requirement in the production of large numbers of targeted mice at the WTSI is the systematic design of a targeting vector that contains necessary elements to disrupt, track and manipulate gene function. As discussed, the ability to track gene expression is important, as many of the targeted genes are indeed novel and conventional tools to characterize them are not available. Further, the phenotypic consequences of generating knockout (KO) mice are unpredictable, so the option to rapidly generate inducible and conditional alleles is critical to assess gene function *in vivo*. To address the requirements of a reporter-tagged and conditional competent allele, the WTSI employs a custom design, referred to as the ‘knockout first’ design<sup>20</sup>.



**Figure 1.2. Knockout-first targeting cassette with conditional potential.** Targeting cassettes are inserted in the non-coding regions upstream of exons predicted to encode elements critical for protein function (the critical exon, CE). The Engrailed 2a splice acceptor (En2a) together with the internal ribosome entry site (IRES) mediate the disruption of the gene transcription and instead allow the expression of the  $\beta$ -galactosidase ( $\beta$ -gal, blue box) reporter. Green circles and red triangles indicate frt and loxP sites for removal of the targeting cassette using flp recombinase and critical exon using cre recombinase, respectively.

### **1.5.1 WTSI Knockout first design: expression and conditional ready targeting**

The WTSI uses a multi-purpose targeting vector to knockout genes in ES cells. This vector delivers 3 important attributes to the resultant transgenic lines: transcript disruption, reporter function and conditional potential. In its basic form, a null allele with reporter capability is produced, allowing researches to determine the expression pattern of a gene while ablating its function (Fig. 1.2). Cre and FLP mediated recombination derives conditional alleles.

The conditional capability of the targeting vector allows researchers to circumvent the problem of embryonic lethality by enabling tissue specific and inducible KOs with simple cre-lox recombination. Knocking-out genes with critical roles in embryonic development can lead to embryonic lethality. One of the advantages of targeted mutagenesis is the ability to investigate lethal gene function by studying the gene's role in specific tissue or by inducing its disruption at a specific time. Thus, the importance of a versatile and easily manipulated allele was a requirement before endeavouring to generate KO mice on a large scale.

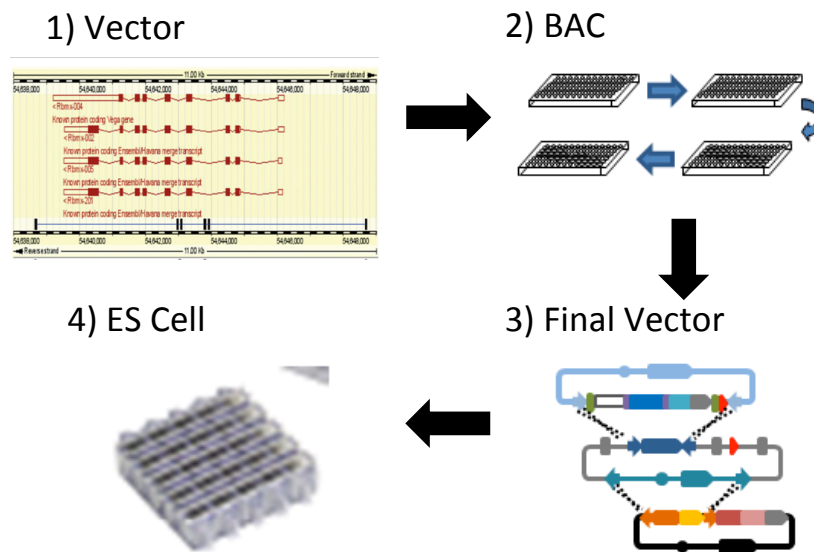
The knockout first allele has several elements to ensure the knockout of gene function, tracking of expression and conditional potential (Table 1.1). These elements include a strong splice acceptor site from the Engrailed-2 gene (En2 SA), Internal ribosome entry site (IRES), lacZ, neo and a polyadenylation site (pA). These elements are flanked by an frt site which when exposed to FLP recombinase, causes excision of the cassette to generate a 'wildtype' allele in function (Fig. 2). LoxP sites flanking a critical region for gene function allow for cre mediated excision of this region.

**Table 1.1 Breakdown of elements in targeting cassettes.**

Name	Abbreviation	Derived from	Function
Flippase Recognition Target	frt	<i>Saccharomyces cerevisiae</i>	Flippase enzyme (flp) mediated recombination between 2 Frt sites in the same orientation results in the deletion of the intervening DNA and leaves behind only one intact Frt site (the signature)
Locus of Crossover in P1	loxP	bacteriophage P1	Cyclisation Recombination enzyme (cre) mediated recombination between 2 loxP sites in the same orientation results in the deletion of the intervening DNA and leaves behind only one intact Cre site (the signature)
Strong splice acceptor site from the Engrailed-2 gene	En 2 SA	Engrailed gene -2 gene, exon 2	Mediates RNA splicing to remove intron and join targeting cassette
Beta-galactosidase	$\beta$ Gal	a bacterial hydrolase enzyme that catalyzes the hydrolysis of $\beta$ -galactosides into monosaccharides	One substrate is X-gal. X-gal is cleaved by $\beta$ -galactosidase yielding galactose and 5-bromo-4-chloro-3-hydroxyindole. The latter is then oxidized into 5,5'-dibromo-4,4'-dichloro-indigo, an insoluble blue product
neomycin phosphotransferase gene	Neo	transposon Tn5 present in the bacterium strain <i>Escherichia coli</i> K12	This confers resistance to the antibiotic G418, an analog of neomycin that works in eukaryotes
Internal ribosome entry site	IRES	Sequence initially discovered in poliovirus RNA	This a nucleotide sequence that allows for translation initiation in the middle of a messenger RNA (mRNA), thereby preventing concatenated synthesis of multiple proteins
polyadenylation site	pA	SV40 (Simian Vacuolating Virus 40 Tag)	The presence of this sequence tell the RNA polymerases to stop transcribing the DNA into mRNA and start adding in a non-template dependent manner a series of A nucleotides (adenine) to the 3' end of the messenger RNA
Beta-actin	$\beta$ act	chicken beta-actin gene	A ubiquitously expressed protein with a strong and effective promoter. This promoter is used to drive neomycin expression to ensure antibiotic resistance in correctly targeted events.

### 1.5.2 High throughput gene targeting

To apply the ‘knockout first allele’ to every gene, the WTSI had to first adapt modern gene targeting technologies to be suitable for high throughput targeting. Gene targeting technology has been used to elucidate the function of thousands of genes on an individual basis over the past 20 years. However, scaling the technology to the level required for targeting the remainder of the genome presents technical challenges and requires new methodologies. To solve these problems, the WTSI has implemented a pipeline approach to large scale gene targeting beginning with vector design, BAC recombineering, vector assembly, and finally, ES cell production and validation (Fig. 1.3). The details of each of these steps will be discussed in sections 1.5.3-1.5.6



**Figure 1.3. Overview of high throughput gene targeting.** Step one involves BAC selection and computer assisted vector design. Step two employs recombineering technology in 96 well plates to engineer gateway compatible BACs. In step 3 homologous recombination of the BACs results in final vector assembly. Electroporation in step four generates targeted ES cells.

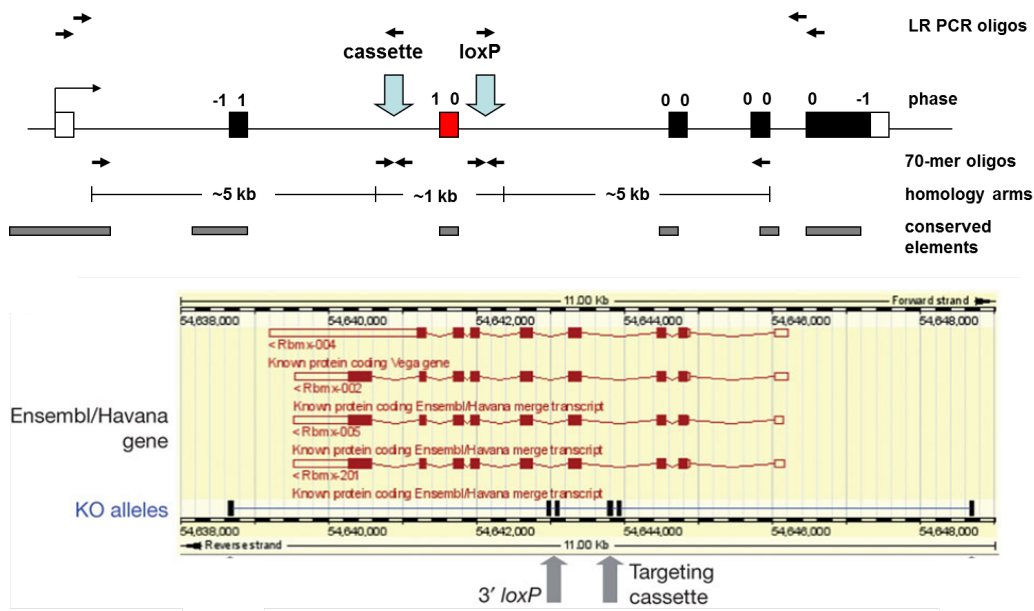
As previously discussed, sequencing the genome left behind a toolkit for gene targeting. One of these tools are BAC libraries, which were initially used to determine the genome sequence, but now serve as the template for gene targeting. Conveniently, extensive BAC libraries were generated as part of the initial genome sequencing effort. The WTSI manipulates existing BACs to serve as vectors for cloning and amplification of short segments of DNA in bacteria.

### 1.5.3 Vector design

Covering the genome multiple times, BAC libraries provide a powerful resource for gene targeting by making homologous sequences of DNA available for manipulation. Regions of homology increase gene specificity in targeting so the BAC libraries are critical in establishing gene specific targeting vectors. The vectors are designed to target genes using a combination of specialized software<sup>21</sup> to identify targetable regions and an expert team of bioinformatitions to validate the computer generated analysis. To ensure successful knockout, an exon critical for function must be selected and the targeting

cassette must be inserted in such a way as to induce a frame-shift mutation that, in turn, causes a premature stop codon and nonsense mediated decay (NMD) of the remaining transcript.

The cassette insertion point is determined by identifying an exon required for gene function—the critical exon (CE). The cassette must be inserted in the intron upstream of the CE, thereby producing a non-functional truncated protein and inducing a frame-shift mutation to disrupt the transcription of the CE and subsequent exons. The identification of the CE is reliant on computer automated gene prediction software. These software programs predict all variants of the gene and determine highly conserved regions. Because of their conservation, these regions are considered necessary for gene function. Although computer assisted gene prediction accelerates the vector design step, there is still a requirement for manual curation by bioinformaticians. The correct identification of a critical exon is essential in maximizing the likelihood of complete ablation of gene function.



**Figure 1.4 Computer assisted vector design.** Specialized software<sup>21</sup> is employed whereby critical exons (red) and oligonucleotide sequences for recombineering and sites are identified (A). Output from computer assisted vector design. Highly conserved exons (red) are identified and appropriate sequences of oligonucleotide (50-omers) suitable for insertion of loxP and targeting cassettes flanking conserved regions are indicated. NB, example shown is on reverse strand. Adapted from Skarnes et al<sup>17</sup>.



Once the critical region is determined, computer software identifies recombineering oligonucleotides (50-mers) within the intron upstream of the CE. These 50-mer sequences that are used to introduce mutations that allow for insertion of gateway recombination sites upstream of the CE and loxP sites downstream of the CE. These sequences are required for gene targeting. So, the accuracy of the cassette relies on the ability to predict the CE and find the appropriate 50-mer sequences.

Additional criteria for optimal vector design and placement are then evaluated against a standard set of criteria. The targeting cassette is inserted in such a way as to ensure that more than 50% of the protein coding sequence is disrupted thus, the CE should be in the first half of the coding DNA sequence to limit the possibility of partially functional protein production. It is also important that the introns flanking the critical exon are greater than 700 bases long to allow for insertion of the reporter cassette. If there are multiple recombineering oligonucleotides identified, then the most 5' conserved sequence is selected. Highly conserved intronic sequences are considered unsuitable for insertion of targeting elements as these may constitute non-coding regulatory elements. Despite the robust computer system used to initially identify regions suitable for targeting, there are still rigorous demands on bioinformaticians to manually analyse each vector design. Protein coding genes which do not meet selection criteria, for example, those containing only one coding exon, are set aside for customized targeting.

After the CE is defined and the 50mers are identified, the BACs are ready to undergo recombineering—a custom strategy designed at the WTSI to genetically manipulate the BAC into a targeting vector that can be used for ES cell targeting.

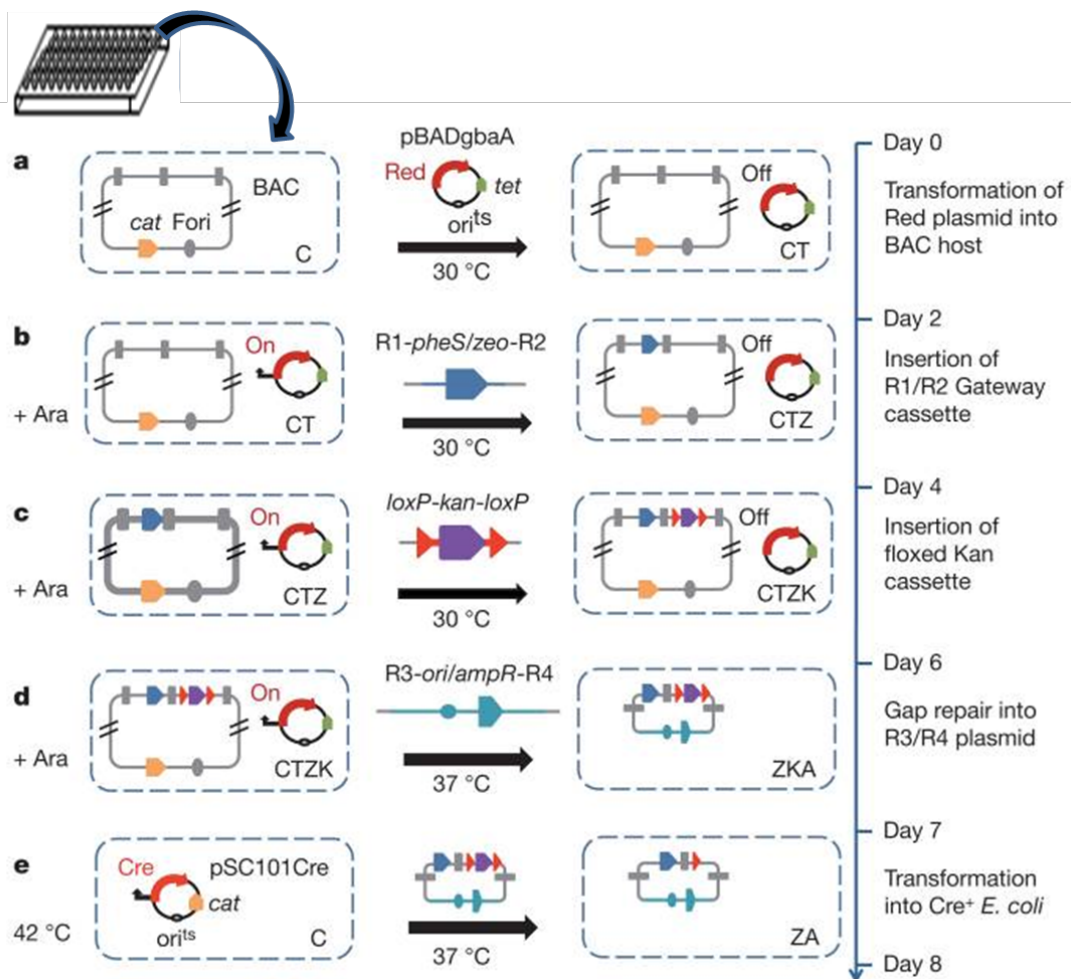
#### **1.5.4 Recombineering: bacteriophage- $\lambda$ encoded red system**

Insertion of targeting cassettes and loxP sites into BACs requires the BAC to be transformed into recombineering-competent *E.coli*. Recombinogenic engineering or recombineering, is critical for high throughput, standardized functional genomic screens. Prior to the development of recombineering techniques, generation of recombinant DNA molecules was limited by the presence and availability of restriction enzyme sites as well as the size of segment of DNA to be manipulated<sup>21</sup>. By taking advantage of endogenous homologous recombination systems in *E. coli* and phages, which inhabit bacteria and also

carry recombination functions, researchers can circumvent many of the challenges associated with vector generation<sup>21</sup>.

The goal of recombineering is to make the BACs compatible with the Gateway system by inserting certain required gateway elements, detailed below and in Figure 5. Only after recombineering will the BACs be gateway compatible (intermediate BACs) and able to undergo homologous recombination to generate final targeting vectors. The potential of recombineering and homologous recombination can be capitalized on by using it in combination with the knockout first allele design on a large scale as demonstrated by the WTSI<sup>22</sup>. To scale this technology to a level required for targeting the entire genome, a high throughput technique was developed to allow for recombineering of BACs in 96 well plates (i.e. 96 different genes simultaneously) thereby generating, on a massive scale, gateway compatible intermediate BAC vectors<sup>17</sup>.

To begin, the BAC host is transformed with the Red plasmid (Figure 1.5 a), a highly efficient phage based system that allows  $\lambda$ -phage-red-mediated recombination. Simple infection with  $\lambda$  phage delivers necessary recombination functions to BAC clones by encoding proteins whose functionality increases recombination efficiency.  $\lambda$ -phage allows the expression of the Exo, Beta, and Gam proteins<sup>23</sup>. Exo is an exonuclease which acts on linear, double stranded DNA leaving behind single stranded 3' DNA overhangs. Beta then promotes annealing of 3' ssDNA overhangs to the complementary strand. The third phage encoded protein is Gam, which inhibits linear DNA degradation activities by inhibition of RecBCD. Expression of Exo, Beta, and Gam is driven by the promoter pL from the natural operon. Thus, the molar ratios of the 3 proteins will be equivalent to those expressed in vivo<sup>23-26</sup>. These are all required for the integration of Gateway elements and subsequent homologous recombination.



**Figure 1.5 Steps and elapsed time required for high throughput recombineering.**

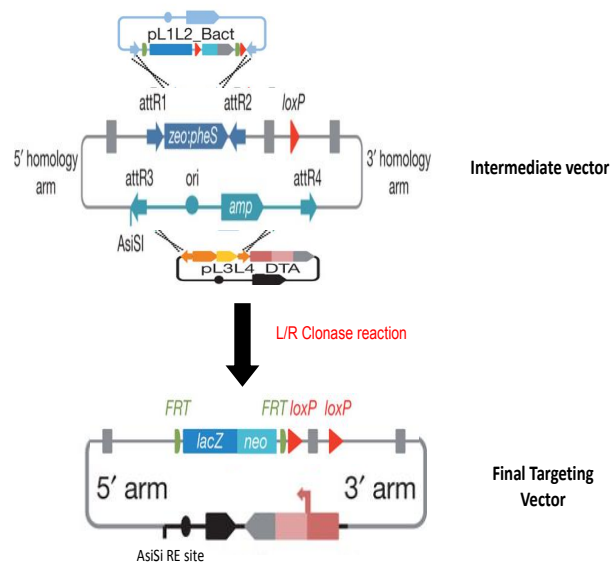
BAC clones, arrayed in 96-well format and electroporated with a plasmid expressing arabinose-inducible Red proteins (pBADgbaA)<sup>47</sup>. b–d, After arabinose induction, cells are electroporated with PCR fragments containing R1-*pheS*/zeo-R2 Gateway element (b), *loxP*-kan-*loxP* cassette (c) and R3-*ori*/ampR-R4 subcloning plasmid (d). After gap repair, plasmid DNA is prepared and transformed into Cre-expressing bacteria to remove the kanR cassette, leaving a single *loxP* site downstream of the critical exon (e). Antibiotics used at each step are: A, ampicillin; C, chloramphenicol; K, kanamycin; T, tetracycline; Z, zeocin. Schematic showing the structure of the Gateway-adapted intermediate plasmid (adapted from<sup>17</sup>)

After the Red plasmid is induced, PCR fragments, containing Gateway elements can then be introduced by electroporation. The first of these gateway elements is an attR1/attR2 *zeo-pheS* fragment, inserted upstream of a critical exon (Figure 1.5 b). This is followed by insertion of a floxed *KanR* cassette downstream of critical exon (Figure 1.5 c) which is responsible for delivering 2-3' loxP sites (one of which will be left behind in the ES cell, floxing the CE). Next, the attR3/attR4 *amp* sites are inserted (Figure 1.5 d). At each step, the listed BACs are grown under selection of the aforementioned antibiotics (italicized) to identify successfully recombineered BACs. At this point, these BACs are Gateway compatible, and are hereafter referred to as intermediate vectors. These vectors can be transformed and propagated in Cre<sup>+</sup> *E. coli* (Figure 1.5 e).

Transformation into cre<sup>+</sup> *E. coli* allow for the Kanamycin cassette to be removed, via homologous recombination, leaving behind its signature, a single loxP site downstream of the critical exon. This same loxP site can be used later when Cre facilitated removal of the critical exon is desired in the targeting vector or allele. Homologous recombination is a rare event, but is much more efficient by coupling with phage based delivery systems and selectable markers (antibiotic resistance)<sup>27</sup>. Cre is encoded by bacteriophage P1 and stands for Cyclisation Recombination enzyme. Its target sequence is loxP for the locus of crossing over (x) Phage1<sup>28</sup>. LoxP sites are 34 base pair segments of DNA sequences composed of two 13bp palindromic sequences flanking a highly conserved 8bp spacer region (ATAACTTCGTATAgcatatcTATACGAAGTTAT) and facilitate recombination<sup>29,30</sup> (see also Table 1.1). The spacer region defines the orientation of the loxP sites. Recombination between 2 loxP sites in the same orientation results in the deletion of the intervening DNA and leaves behind only one intact loxP site, whereas loxP sites in opposite directions causes inversion of DNA<sup>31</sup>. After completion of this step, the assembly of the intermediate vector is complete (Figure 1.6). It should be noted that at this point, the intermediate vectors can be further manipulated or reused, allowing for the addition of different reporters or even the knockin of cDNA sequences. They can also be combined with other targeting vectors with different selectable markers to knockout the second allele of a gene, allowing for in vitro functional studies of homozygous null ES cells. Thus, the intermediate vectors are an important bi-product of recombineering and are, in themselves, extremely useful.

### 1.5.5 Final Vector Assembly

Once the intermediate vector is assembled, the generation of the final targeting vector with the knockout first design can begin. Gateway equipped intermediate vectors are then assembled into final targeting vectors using L/R clonase and the att sites that were added in the recombineering steps. In its basic form, the intermediate BAC is converted to a targeting vector using L/R clonase facilitated recombination in combination with pL1L2 and pL3L4 plasmids. In the first recombination step the pL1L2 plasmid provides the lacZ neomycin cassette with L/R clonase driven recombination between the attR1 and attR2 sites. The att3 and att4 sites allow for the transfer of a selection marker (DTA) and a unique AsiSI restriction from the pL3L4 to the intermediate vector. The AsiSI site is rare, and thereby allows for the required linearization of the targeting vector for electroporation into ES cells with little chance of disrupting the BAC DNA.



**Figure 1.6. Assembly of final targeting vectors.** L/R clonase facilitated homologous recombination between the attR1/attR2 and attR3/attR4 introduces the pL1L2 and pL3L4 plasmids, respectively. The pL1L2 provides the essential targeting elements and the pL3L4 provides rare AsiSI restriction site for linearizing the final targeting vector before electroporation of ES cells. This assembly is made possible by the recombineering induced manipulation of the BAC, making it compatible with gateway technology (adapted from <sup>17</sup>).

The assembly of the final targeting vector is through a multi-Gateway reaction inserting the cassette *lacZ*-*neo* cassette and a DTA cassette that allows for selection and provides a rare AsiSI restriction site for linearization. The targeting constructs must be linearized before electroporation into ES cells.

### 1.5.6 ES cell production

In the 1980's, the capacity of ES to be isolated, propagated for prolonged periods in tissue culture, and colonized into the germ line after injection into blastocysts was discovered<sup>32-34</sup>. One of the greatest challenges of this project at the WTSI is the high throughput manipulation of ES cells. Previously, conditional ES cell methodologies were not robust enough to be scaled to the scope required for high throughput production of targeted ES cells. Advances in ES cell techniques together with the development of highly efficient, phage-based *E. coli* recombination systems (detailed above), have been instrumental in large scale generation of precise genetically manipulated mouse models.

Another setback to the initiation of this project was the mouse strain. Previously, genetic manipulation in mice was carried out predominately in ES cells derived from the 129 strain of mice. Despite the high usage of 129 ES cells, this strain is not entirely suitable for genetic studies due to background issues<sup>35</sup>. The C57BL/6 line is considered to be a better background on which to study mouse genetics. Thus, the C57BL/6 line was chosen for this screen as it is one of the best characterized inbred mouse strains, also making it appropriate for downstream phenotyping efforts<sup>4,35</sup>. Therefore, ES cell banks for the C57BL/6 line had to be generated and techniques had to be modified to fit this strain.

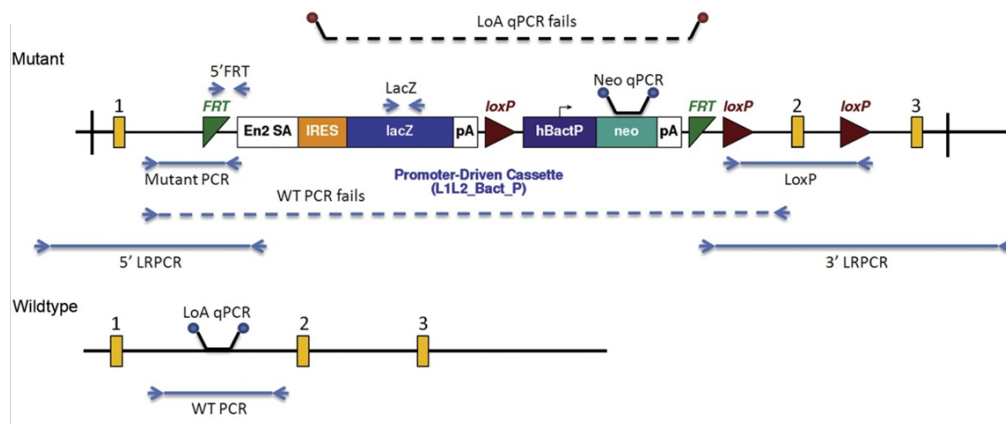
By using the targeting and ES cell culturing techniques described in this section, the Sanger-MGP and the KOMC have overcome many of the previous limitations of gene targeting. The global repository of targeted ES cells now covers over half of all known mouse genes. At the Sanger-MGP, the systematic introduction of targeting vectors into ES cells and characterization of targeted ES cells has allowed for the generation of hundreds of knockout mouse lines, each with a specific gene disrupted.

To introduce the targeting vector into ES cells, the vector is linearised using the AsiSI restriction enzyme site. ES cells are electroporated with linearized plasmid in multiwall plates to 'knockout' one allele via exchange of manipulated genetic material between the homology arms<sup>20,22</sup>. The cells are then grown under selection to determine successfully targeted ES cells. ES cell colonies are expanded, frozen until required for germ line colonization in mice. The advances in ES cells technologies are critical for these processes. After the generation of targeting vectors, homologous recombination induction via electroration, and selection of ES cells, proper ES cell targeting is confirmed. This is

done using molecular techniques, including sort and long range-PCR to confirm the desired genetic change.

### 1.5.7 Genotyping ES cells

Traditionally, the success of homologous recombination would be confirmed by Southern blot. However, this method was unsuitable for the scale required. So, instead, long range PCR (LR-PCR) using both universal and gene specific primer sets is used. The universal PCR primers are designed to detect the cassette and therefore can be used on all targeted ES cells to confirm cassette insertions. The gene specific primers ensure that the cassette has been integrated into the correct genome locus. Again, the gene specific primers were generated by computer software and this process was optimised to accommodate the large numbers of samples simultaneously.



**Figure 1.7 Summary of genotyping strategy.** Conformation of correct gene targeting is achieved in ES cells and in mice is done using a combination of universal and gene specific PCR's and qPCRs. For ES cells, LR-PCR, using a universal primer in the cassette and another outside of the homology arms of the allele design, was used to confirm the targeting on either the 3' or 5' side of the vector prior to micro-injection. To determine the genotype and confirm gene identity three short-range PCR assays were used: mutant allele-specific, wild-type allele-specific and to detect the *lacZ* gene. Targeting was confirmed by either LR-PCR, loss of the wild-type specific short-range PCR product in homozygotes or a qPCR assay confirming loss of the wild-type allele. Presence of the 3' LoxP site was detected by either qPCR or short-range PCR assays (adapted from [www.mouseportal.com](http://www.mouseportal.com)).

In summary, the optimization of gene targeting made the prospect of high throughput reverse genetic screening in mice a possibility. At each step, valuable resources have been

created and comprise libraries of gene specific conditional designs, manipulable intermediate vectors, targeting vectors, and most importantly in terms of mouse generation, LacZ conditional targeted ES cell clones for use in germ line transmission. These resources in turn allow for the high throughput generation of transgenic mice.

## **1.6 HIGH THROUGHPUT GENERATION OF TRANSGENIC MICE FROM TARGETED ES CELLS**

The following steps are necessary to convert targeted ES cells into transgenic mice.

### **1.6.1 Microinjection into embryos**

First, ES cells are subject to *in vitro* characterization to determine the change in the gene of interest. Once the defined change has been confirmed, the targeted ES cells are introduced, preimplantation, into host embryos at the morula or blastocyst stage using microinjection techniques. All targeted ES cells are from the feeder dependent substrain JM8.F6, which originates from C57BL/6N background. The albino strain, C57BL6J-Tyr, is used for blastocyst hosts of the targeted ES cells. This allows for an estimation of germline transmission in the chimeric progeny as black tissue originates from the targeted ES cells whereas white tissue originates from the wildtype blastocyst host strain.

### **1.6.2 Embryo transfer**

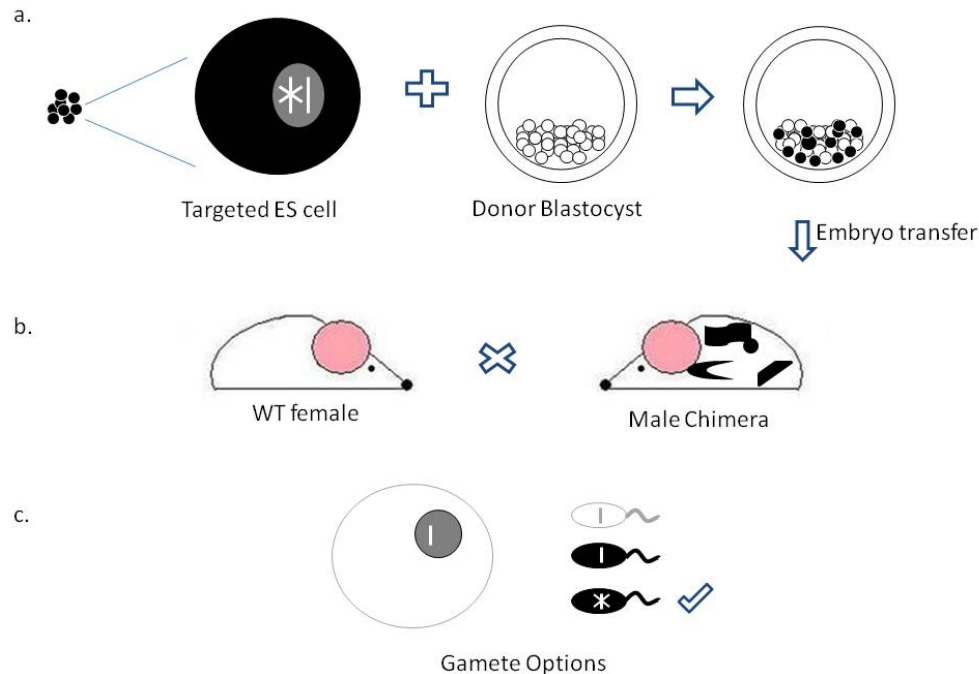
After ES cells have been successfully targeted and propagated in host embryos, they are transferred into the oviducts or uteri of pseudo-pregnant recipient foster females. To generate the pseudopregnant females, they are mated with sterile (vasectomized) stud males. This process prepares the female to accept the transfer of the embryos and carry them to term, ultimately generating the F0 generation.

### **1.6.3 Chimeras and germ line transmission**

Each mouse from the F0 generation is a genetic chimera, comprised of the genetic combination of the donor blastocysts (white) with the targeted ES cells (black). Male mice with a high proportion of black fur (and assumed high contribution from targeted ES cells) are bred with white wildtype females to create the F1 generation with germ line transmission of the targeted mutation. To assess this, the F1 generation is first screened on the basis of coat colour. As black coats are dominant, white pups are assumed to be wildtype. The initial electroporation of the ES cell only mutates 1 of the 2 alleles, thus each black pup has a 1 in 2 chance of receiving the targeted allele from the father. Black



pups are then genotyped to determine if the knockout allele was transmitted from the male chimera (Figure 1.8).

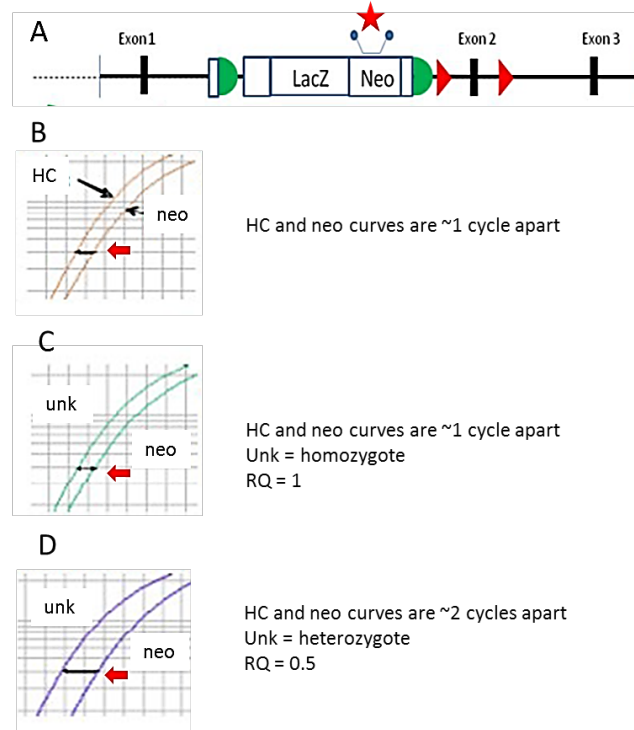


**Figure 1.8 ES cell to mouse.** Targeted ES cells are microinjected into donor blastocysts. ES cells are from black mice whereas donor blastocysts are from albino mice. Blastocysts are then transferred into pseudo-pregnant females, and the subsequent progeny are screened on the basis of coat colour(a). Coat colour is related to the contribution of ES cells. Mice with high percentages of black coats are mated with wildtype females(b). As black coats are dominant, white pups are assumed to be wildtype. Black pups are then genotyped to determine if the knockout allele was transmitted from the male chimera(c).

#### 1.6.4 Genotyping mice

To facilitate high throughput genotyping of mice, a qPCR neo counting based approach is taken at the Sanger-MGP. This approach allows for rapid genotyping to be completed in 96-well plates. A TaqMan qPCR reaction is used to count the neo gene marker present in the targeting cassette; heterozygotes will have one copy of the Neo gene, a homozygote will have two copies and a wild type will have no copies (Figure 1.9). The Neo count does not confirm the correct targeting of the allele, or that if the correct gene has been microinjected in the first place, so it is not suitable for genotyping ES cells. It can

however be used as a QC tool to detect multiple insertion events, problems with the internal structure of the cassette, or mixed colonies (for example if 2 copies are detected in the chimera progeny or more than two copies start being detected as the colony is established).



**Figure 1.9 Neo count qPCR assay.** To determine the number of copies of Neo cassette in a sample, qPCR using a neo specific TaqMan probe (star) is used (A). To normalise for variability in the reactions (eg DNA concentration or the presence of inhibitors) and allow comparison between two different samples, the target gene (Neo) is first compared to a homozygous calibrator (HC) reaction for each sample (B). The relative quantity (RQ) and thus copy number of the unknown sample against that of the calibrator can then be calculated. This translates to a one threshold cycle difference in amplification rate during the exponential phase of PCR and the RQ values = 1 for homozygotes (C). A heterozygote will have half the amount of template available for the Neo marker compared to that of a homozygote so the RQ = 0.5 (D). Wild type alleles will show no target amplification (not shown).

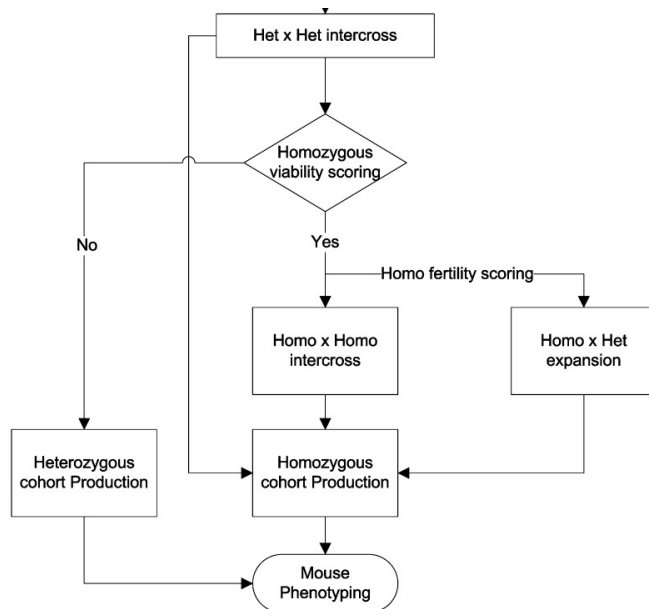
Neo count qPCR can be used as a routine genotyping tool once the rest of the QC has been performed. The non-gene specific nature of the assay makes it very conducive to high-throughput methods where many different genes may be present on a sample plate,

and thousands of samples can be easily genotyped using this method.

## 1.7 PHENOTYPING FOR DEVELOPMENTAL DEFECTS, DYSFUNCTION, AND DISEASE MODELS

### 1.7.1 Generation of Phenotyping cohorts

After confirmation of germline transmission, mice are bred for the generation of phenotyping cohorts. The generation of phenotyping cohorts causes considerable difficulties regarding space in animal houses as well as animal welfare issues. Logistical issues associated with phenotyping cohort generation and have proven to be the major rate-limiting factor for high-throughput phenotyping<sup>36</sup>. Mice are bred to assess fertility and viability initially. Once fertility and viability are determined, the appropriate crosses are set up to generate cohorts of age and sex matched mice for phenotypic screening. In cases of homozygous lethality, timed matings are set up to determine the approximate age at time of death during development. Heterozygous mice from embryonic lethal mutations are used for phenotyping in lieu of homozygous mice.



**Figure 1.10 Flowchart for mouse breeding strategies for cohort production.** Targeted ES cells were injected to generate chimeric animals. After germline transmission of the mutant allele, the heterozygous mice were bred for pedigree expansion, archiving, distribution, and cohort production. Depending on the fertility score, heterozygous or homozygous matings were used to generate cohorts of animals for mouse phenotyping.

### **1.7.2 Broad preliminary phenotyping pipeline**

After production of knockout lines, mice are available for phenotypic screens. At the WTSI, all mutant lines undergo a series of primary phenotypic screens that, collectively, contain over 200 different parameters relating to 28 physiological systems. The tests, including gross dysmorphology, hair follicle cycling, hair analysis, glucose tolerance, SHIRPA, x-ray and DEXA analysis, are designed to identify specific indicators of disease and allow for the immediate association of a phenotype to the genotype. Terminal necropsies are performed after testing is completed, and tissue from all mice is collected, fixed, embedded in paraffin wax and stored in a 'biobank'. Additionally, tissue is collected for expression analysis with x-gal staining. Stained tissue is imaged and archived in a similar biobank. The results of phenotype tests and gene expression data are being compiled in a database which will serve as an invaluable resource for researchers. Preliminary data is publicly available for some mutant lines and is available at <http://www.knockoutmouse.org/>. Further, a detailed analysis of the overall screen and the phenotype data from the first 250 lines generated at the WTSI was recently published<sup>19</sup>. Mice and targeted ES cells generated in this screen are also publically available.

### **1.7.3 Disease phenotype focus**

The Sanger-MGP characterization of mutant mouse lines is designed to identify genes that play a role in disease. Specific disease areas include, but are not limited to cancer, diabetes, kidney diseases, infertility, cardiovascular disease, obesity, developmental and musculo-skeletal problems.

#### *Cancer*

A combination of haematological and histological analyses may identify cancer phenotypes. FACS analysis on peripheral blood lymphocytes can be used to identify biomarkers of cancer such as deviations in resident immune cell populations or micronuclei formation. As tissues are stored in a biobank, FACS data can be further supported by immune-histochemical analysis of proteins commonly expressed or misexpressed in cancers.

#### *Diabetes and Obesity*

All mice in the Sanger-MGP phenotyping pipelines are maintained on a high fat diet from 4 weeks of age. Exposure to this high fat diet 'challenge' may identify genes whose loss

of function result in a predisposition towards obesity and other metabolic disorders. The simplest aspect in screening for these disorders begins with weekly weighing and dual emission X-ray absorptiometry densitometer analysis (DEXA) to determine fat and lean mass. To screen for more subtle anomalies in metabolism calorimetry and clinical chemistry analysis of the blood and urine is used. Findings here can be corroborated by observations made during terminal necropsies, such as fatty livers, or using histopathology from biobanked tissues.

### *Infertility*

Animals are assessed, by default, for infertility phenotypes as a part of generating homozygous cohorts for phenotyping. The most effective way of generating homozygous cohorts of animals is to set up matings between 2 homozygous animals. Initial flagging of infertility can happen at this stage and subsequent crosses can be set up to determine if the infertility phenotype is male or female specific. Again, the biobanking of tissues can aid in the description and characterization of infertility.

### *Defects in sensory acuity*

Sensory acuity tests comprise assessment of hearing, vision, response to acute thermal stimulus and histopathology. Hearing thresholds can be determined by auditory brainstem response (ABR) measurement during which anaesthetised mice are subjected to tones of varying levels of frequencies and intensities. Responses are measured by subcutaneous pin electrodes that are placed behind the ear. Eye and vision phenotypes are assessed using slit lamp bio-microscopy. This technology identifies defects in the iris, cornea, pupil and lens that may indicate susceptibility to vision impairment. Finally, the response to pain is measured using a hotplate to determine the reaction time to acute heat stimulus.

## **1.8 EXPRESSION ANALYSIS**

Phenotypic data is complimented and enhanced by expression analysis data. This data is derived from visualizing the reporter gene *lacZ*, which encodes  $\beta$ -gal. *LacZ* is expressed within a targeted allele by the endogenous promoter of the target gene. X-gal, a substrate for  $\beta$ -gal, is cleaved by  $\beta$ -gal, yielding galactose and an insoluble blue product. Hence staining with X-gal solution gives an overview of the expression profile of the target gene. Expression analysis is done on cohorts of heterozygous adults and embryos by histochemical analysis of  $\beta$ -gal activity. Not only does this data help to elucidate the

normal pattern of gene expression, it also acts as an indicator of homologous recombination and targeting cassette efficiency (ie ability of the splice acceptor to disrupt gene splicing patterns and instead transcribe the reporter gene).

Understanding the genetic basis of development and disease can greatly impact both the treatment and diagnosis of diseases. The ability to quickly identify the organ(s) where the gene is expressed can also give clues about systems that may potentially be affected by loss of function of the gene. This is valuable in disease diagnosis as a significant number of diseases have systemic effects and manifest themselves in multiple organ systems. For this reason, screening for disease phenotypes requires organ specific screening to identify defects in normal organ function. This has the added potential of providing new biomarkers for early disease diagnosis.

## **1.9 INTEGRATING ORGAN SPECIFIC PHENOTYPING SCREENS**

To maximize the efforts of the WTSI's preliminary phenotype screen, the integration of organ specific screens is required. This is because some organ specific phenotypes might be missed if they are not specifically screened for, and also because many of the live mutant mice are only available for phenotyping for a short window of time. Thus, we implemented a multi-parameter, multi-test skin screen aimed at identifying mouse models of skin dysfunction and elucidating genetic underpinnings of skin development and homeostasis. I spent one year embedded at the WTSI to see, first hand, the knockout mice generated there and select a new line for further characterization. To do this, I helped to implement and carryout tests designed to identify skin dysfunction.

## **1.10 SKIN**

### **1.10.1 Introduction**

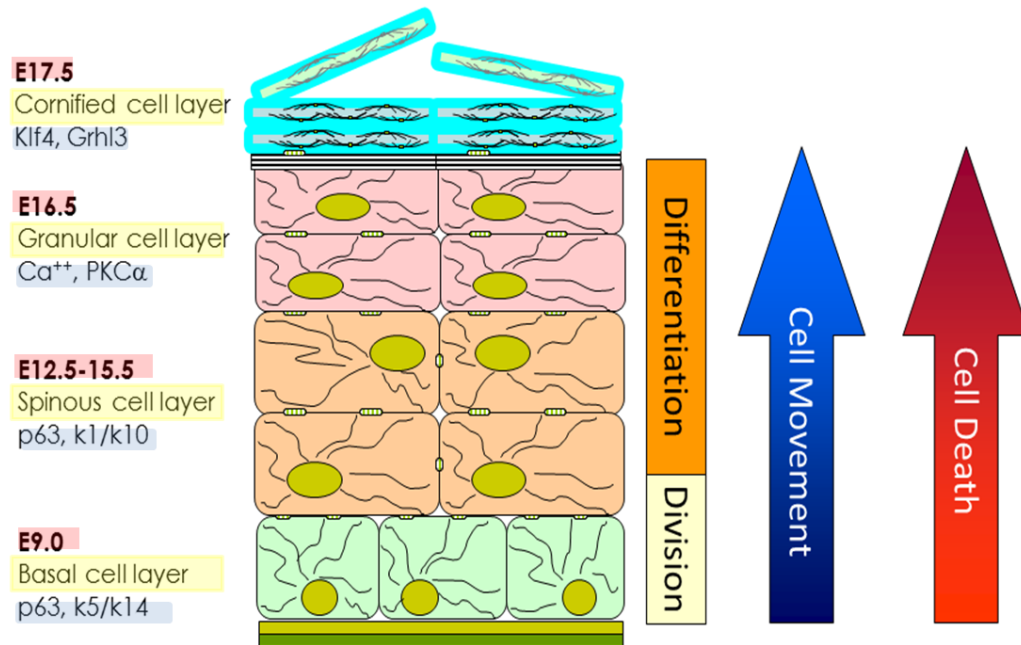
Developing tests for skin defects requires knowledge of skin development and homeostasis. This section provides an introduction into skin biology and the background required to understand the nature of the skin phenotyping tests. The skin is the largest organ in the human body and has been extensively studied and well characterised. It plays a fundamental role in immune surveillance, wound healing and hair generation. Furthermore it plays the critical role of providing a protective barrier between the other organs of the body and environmental assaults such as pathogens and mechanical stress.

The skin constantly responds to changing endogenous and exogenous environments to provide protection to other tissues and organs.

To better understand these processes and the genetics that govern them, the Sanger Skin Screen Consortium was developed. As a part of this consortium, I worked with members of the Sanger-MGP to carryout screens designed to identify skin defects in the mutant mice generated there. Although previous research has provided critical information about epidermal homeostasis and development, there are still many undiscovered genetic factors required for normal skin development and maintenance. This is indicated as there are thousands of dermatopathies for which the genetic underpinnings remain unknown. This section provides a brief summary of previous research findings regarding skin function, skin development, and homeostasis. The use of current knowledge of skin function is an important starting point in designing phenotypic tests to successfully detect epidermal phenotypes.

#### **1.10.2 Skin developmental defects**

The epidermis originates from the ectodermal germ layer in the developing embryo. Ectoderm progenitors give rise to skin epithelium or neural tissue, a process mediated by the presence or absence of Wnt signalling, respectively<sup>37</sup>. Epidermal fate acquisition arises when Wnt signalling blocks the ectoderm from responding to fibroblast growth factors (FGFs), which in turn allows the cells to express and respond to bone morphogenetic proteins (BMPs). In the absence of Wnt signalling, cells respond to FGF's, ultimately reducing responses to BMPs giving rise to neural tissue<sup>38,39</sup>. During embryonic days 9-12 in the mouse, a single layer of epithelial cells is covered by a layer of tightly connected squamous cells, known as the periderm. The periderm is eventually shed as the epidermis stratifies<sup>40</sup>. Skin specialization and skin stratification occur simultaneously during development, leading to the formation of a protective barrier and population of mini organs, including hair follicles and sebaceous glands.



**Figure 1.11 Epidermal development.** Skin development is initiated at 9 days of gestation. The basal layer of the epidermis proliferates and then terminally differentiates to populate the suprabasal layers of the epidermis. The gestational age is highlighted in red, cell layer in yellow and transcription factors and markers of each cell layer are highlighted in blue.

Skin stratification begins with a single layer (basal layer) of proliferative cells which give rise to terminally differentiating cells of the suprabasal layers (spinous and granular layers) and the stratum corneum (Figure 1.11). As cells migrate outwards and terminally differentiate, different keratins are expressed, marking the layers of the skin<sup>41</sup>. Basal cells are marked by the expression of KRT5 and KRT14. At embryonic day (ED) 12.5-15.5, a switch in expression from KRT5/14 to KRT1 and KRT10 is mediated by p63 and marks the spinous cell layer. Spinous cells synthesize extensive keratin filament networks and connect to desmosomes<sup>42</sup> forming the mechanical infrastructure of the skin. At ED16.5 the granular layer forms. Granular cells produce lamellar granules and keratohyalin granules that contribute to the intercellular epidermal barrier and aid in the aggregation of keratin filaments. As granular cells migrate outwards toward the stratum corneum, cytoplasmic organelles are lost and the cells become the cornified envelope which serves



as a scaffold for lipid bilayers and contributes to the epidermal barrier. Two key transcription factors in this process are *Klf4* and *Grhl3*, which control the expression of genes involved in cornification of keratinocytes and regulation of lipid secretion, respectively<sup>43,44</sup>. Thus, cell death in the skin occurs in a controlled and specific manner leading to a waterproof and protective barrier that is continually replenished as cells are sloughed off at the surface<sup>45,46</sup>.

This paradigm, established during development, is mimicked throughout postnatal life as the skin is constantly undergoing regeneration. The regenerative process is modified to accommodate internal and external changes, for example the upregulation of keratins in response to wounding or the recruitment of immune cells in response to infection. Understanding the development and regenerative capacities of the skin and the underlying genetics was a primary reason for implementing skin specific screens in the reverse genetic screen at the Sanger-MGP.

### **1.10.3 Phenotyping skin as a model system**

The Sanger-MGP skin screen was implemented to fully utilize the power of a knockout mouse resource by using the established preliminary phenotyping platforms that feed into secondary, skin specific platforms. The rate at which animals are generated limits the complexity and depth to which tests can be performed. It is impossible to provide a detailed phenotypic analysis of all organ systems. However, phenotyping of some organ systems can be done more rigorously, but still rapidly enough to encompass the high number of animals. The skin is one such organ, partially because it comprises the outer covering of the body.

As the most accessible organ, defects in skin biology can be readily assessed in a high throughput fashion histologically, macroscopically and functionally. The development and homeostasis of skin are dependent on stem cells which reside in basal keratinocytes, hair follicles and sebaceous glands. These stem cell populations work together to give differential responses to environmental, physiological and genetic changes. There are still many unanswered questions regarding how skin stem cell populations are maintained and how they control hair follicle cycling and response to wounding. The genetic mechanisms that signal the cells of the skin to migrate and differentiate to provide healing and protective qualities are also still largely unknown. Investigation of skin dysfunction

including inflammation, blistering, acne and cancer can be undertaken by linking dermopathic phenotypes with loss of function of a gene. Specific tests on mutant mice for defects in cutaneous biology can address all of these aspects of the skin and can identify genes which are not only responsible for skin defects, but also can provide insights into processes fundamental to cell biology and development.

The association of inheriting a faulty copy of a gene is the foundation of our understanding of some skin diseases, such as the deleterious skin condition, harlequin ichthyosis (HI), where mutations in a single gene, *Abca12*, results in defective lipid transport, and subsequent impairment of the epidermal barrier<sup>47</sup>. The *Abca12* knockout mouse phenocopies human HI disease and has allowed for investigation into the mechanism of this disease. Ultimately, this mouse model can be used to better understand the disease and give insight into therapeutic options for its treatment. For most genes and skin diseases, this understanding is lacking, and the role of most genes in human health and disease awaits discovery.

Skin disease constitutes a significant imposition on health expenditure in Australia and a large proportion (~15%) of visits to doctors relate to skin conditions, many of which have genetic bases. Furthermore, there are over 4000 described dermatopathies. Despite the large incidence of skin disease and the impact on the health system, the genetic basis of many skin associated diseases remains unknown. The identification of genes involved in the development and progression of diseases of the skin and in related skin conditions such as abnormal wound healing and impaired barrier function could lead to new, more effective gene based therapies to treat cutaneous disease. Reverse genetic screening has the potential to identify genes important for cutaneous development, homeostasis and maintenance, therefore, allowing for better understanding. In screening for skin phenotypes, there are a few key processes and phenotypes that are of considerable interest in both development and adulthood.

### **1.11 DEVELOPMENTAL SKIN DEFECTS**

Simple embryo dysmorphology screens for skin defects can be implemented to screen for epidermal phenotypes by harvesting late stage embryos. Genetic defects that lead to lethality just prior to or just after birth indicate that the gene may be required for the transition from the uterine environment to the outside world. Skin function is not fully

challenged until after birth, at which point the organ is faced with environmental assaults, mechanical stress and dehydration. Thus, assays in late stage development can identify causative developmental skin defects that are responsible for perinatal lethality.

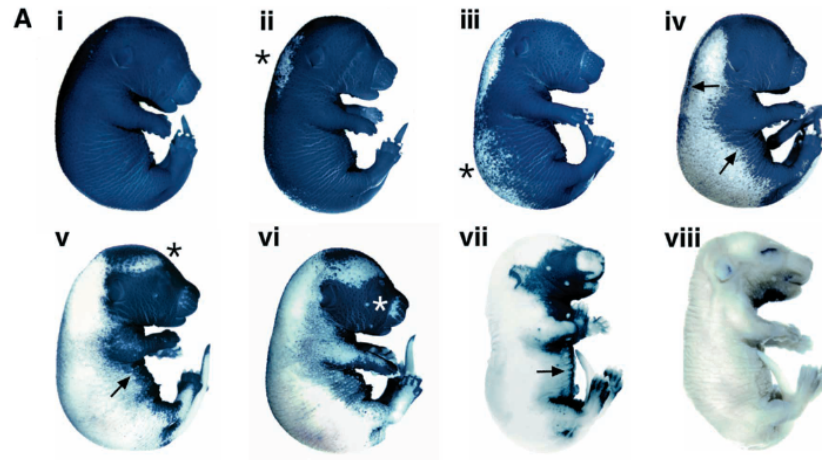
#### **1.11.1 Blistering**

Gross visual examination can easily identify blisters and blebs which are signs of skin development and adhesion defects. Mutations leading to congenital diseases affecting epithelial framework, structure and organization impact embryonic development. Tissue fragility disorders, such as epidermolysis bullosa, present skin blistering phenotypes. Mutations in genes expressed in the cutaneous basement membranes, which encode for structural proteins in the skin are responsible for epidermolysis bullosa disease phenotypes. To date, over 1000 different mutations across 13 different genes have been associated with EB, with varying degrees of tissue fragility<sup>48, 49,50</sup>. Another example of a disease affecting epidermal development is Fraser Syndrome. Fraser syndrome is a rare autosomal recessive disorder characterised by loss of embryonic epidermal adhesion during mid-gestation, defects in renal development and other developmental phenotypes<sup>51,52</sup>. A key to understanding Fraser Syndrome has been the family of blebs mouse mutants which model the disease in many different organ systems<sup>53</sup>. These mice; *blebbed*<sup>54</sup>, *head blebs*<sup>55</sup>, *eye blebs*<sup>56</sup> and *myelencephalic blebs*<sup>57</sup> have been used to link a dysfunctional gene to a specific disease, furthering our understanding of the disease pathology.

#### **1.11.2 Barrier function**

Complete skin development is essential for the acquisition of barrier function. Barrier function is critical to life outside the uterus, thus disruption of genes affecting the acquisition of this function can result in neonatal or early postnatal lethality. Defects in barrier function predominantly lead to death or disease by dehydration, infection or a combination of both in mice and in humans. Defects in skin development leading to barrier dysfunction can be assessed *in situ* through dye penetration assays<sup>58</sup>. At embryonic day 18.5, embryos with defects in barrier function take up the dye through their skin, while normal embryos exclude the dye (Figure 1.12). This relatively simple yet informative assay is well suited for high throughput phenotypic screens and can potentially identify genes with previously unknown roles in epidermal development and

disease. This method has already been validated for other models, such as the Abca12 mouse model of HI discussed in the previous section.



**Figure 1.12. Toluidine Blue barrier function assay.** Dye exclusion assay shows changes in ability of skin to exclude the toluidine blue dye with age. At embryonic day 16 (Ai-viii) the skin is almost completely penetrable by toluidine blue dye. Specific skin sites acquire impermeable characteristics (asterisks), followed by the directional dorsal to ventral (arrows) acquisition of barrier function as indicated by dye being excluded. Approximate average gestational ages for each barrier stage in ICR strain mice are: Ai, <E16; Av, E16; Aviii, >E17 (adapted from<sup>58</sup>)

Disruptions in the development of the skin barrier can indicate incomplete terminal differentiation. The stratum corneum, the outermost layer of the epidermis, is formed during terminal differentiation. Large, flattened, enucleated corneocytes make up the stratum corneum and are formed by crosslinking proteins such as keratins, filaggrin, involucrin and loricrin<sup>59-61</sup>. These highly insoluble corneocytes are enveloped in a lipid-enriched extracellular matrix to make up the ‘brick and mortar’ that is critical to skin barrier function<sup>62</sup>. Defects in the barrier have profound effects in the epidermis as a whole, including hyperproliferation and altered wounding response. This highlights the importance of complete terminal differentiation of the skin during development.

Other contributors to barrier function acquisition and maintenance of the epidermis are tight junction (TJ) structures. The TJ is an intercellular junctional structure that is

responsible for the adhesive properties of epithelial cells to one another. They play a critical role in selective permeability of stratified epithelial sheets like the epidermis. The main TJ protein family, the claudins, have been shown to be involved in paracellular permeability by regulating passage of both water and ions and thus establishing ion gradients across tissues<sup>63</sup>. TJ barrier dysfunction cannot be detected using the toluidine blue dye exclusion assay, as stratification, differentiation and lipid deposition are not compromised with the loss of TJ components. Instead, *ex vivo* trans-epidermal water loss (TEWL) assays can be used to confirm evaporation of water through the epidermis. The *Cldn-1* knockout mouse demonstrates the requirement of TJs in epidermal barrier function. *Cldn1* mutants die at birth from rapid water loss and dehydration, despite normal stratification and lipid deposition of the epidermis<sup>64</sup>. As opposed to dye exclusion assays; TEWL assays are too complex and time consuming to be integrated into a high throughput reverse genetic screen. However, because barrier dysfunction has profound effects on the epidermis as a whole, TJ mutants will likely demonstrate hallmarks of barrier dysfunction, making TEWL an effective tertiary screen for select mutant lines.

Thus, the normal development of skin involves complex signalling pathways and protein interactions. Many of the biological processes underlying skin development are not fully understood. Phenotyping and characterization of development in knockout embryos can lead to discovery of genes not previously known to be involved in development and congenital disorders. Defects originally identified in skin using relatively simple phenotypic assays can lead to further investigation of a gene and its roles in the development of other organ systems as well.

### **1.12 ADULT SKIN DEFECTS**

In postnatal life, the skin is constantly responding to changes and stress. The accessibility of skin allows for easy and quick assessment that provides insights and clues into biological pathways. For example, the assessment of coat coloration/presence, skin pigmentation, whisker and ear morphology, presence and pigmentation of footpads, length and morphology of claws and teeth are easy and informative. Going beyond gross dysmorphology screens, histological investigation of marker expression and localization patterns, chronologically tracking hair follicle cycling patterns, and wholemount analysis of hair follicle morphology can be modified to fit with high throughput phenotyping. Taken together, the examination of the skin and its dynamic processes serve to provide

insight into the ability of the skin to maintain homeostasis and provide necessary functions required in postnatal life.

### **1.12.1 Pigmentation**

The simplest aspect of skin screening is an overall assessment of skin and coat colour. The initial flagging of the mutant line for further analysis due to discovery of an obvious pigmentation defect demonstrate the insights to be gained from skin/coat specific phenotype assays. Themes in physiological and developmental pathways used by pigment cells are replicated in other, less accessible cell types<sup>65</sup>. Two informative skin phenotypes which have opened the door to more rigorous tests at the WTSI are white belly patches and pigmented footpads.

Absence of pigment along the ventral trunk resulting in a white belly spot or in a belt like pattern extending dorsally indicates disruptions in neural crest migration. Neural crest cells can give rise to neurons, glia, or pigment producing cells called melanocytes. Melanocyte precursors, called melanoblasts, are specified at the neural crest and then migrate along the dorsolateral pathway in a highly conserved and reproducible pattern. The unique migratory pattern of melanoblasts is governed by signalling molecules and receptors on their surface, both of which are genetically regulated and specified<sup>66</sup>. Ultimately, in mouse, melanoblasts reside deep in the hair follicles and are tightly coupled with hair cycling. Although melanocytes can be found in the dermis, their presence in the epidermis is uncommon<sup>66-68</sup>. The loss of pigmentation at the ventral portion of the trunk is likely due the failure of melanoblast progenitors to migrate. Furthermore, white spots on the ventral trunk are coincident with changes in cell survival and proliferation in those areas<sup>69</sup>. Investigation into the mechanisms involved in defining melanoblast number and subsequent cell survival could help in further understanding of neural crest derived stem cell populations.

Most mutations causing pigmentation phenotypes alter hair colour, while very few alter skin colour, partly because hair melanocytes are more numerous than skin melanocytes<sup>65</sup>. Whilst most coat colour mutations impact processes early in development (ie neural crest cell specification), mutations causing changes in skin colour occur after neural crest differentiation. After neural crest differentiation, subsequent melanoblasts can be found in the developing dermis/mesenchyme. Some melanoblasts remain in the dermis to become

dermal melanocytes, and others migrate into the stratifying epidermis to become follicular melanocytes or epidermal extrafollicular melanocytes. Mouse models of dark skin can be classified into two groups, based on the developmental pathway perturbed, which is coincidentally indicated by footpad pigmentation patterns and histological analysis of pigment accumulation in the footpads and tail. Phenotypically, Class I dark skin mutants show a prominent slanted strip of pigment beginning at the base of the first digit, and extending over the volar (palm) pad of the first digit only. In Class I dark skin mutants, histological analysis reveals normal pigment distribution in the epidermis, but increased accumulation in the dermis, suggesting defects in melanocyte differentiation and migration across the basement membrane. Because melanocyte differentiation occurs during development, Class I mutants are easily distinguishable by 2-3 days of age. Interestingly, Class I pigmentation mutants are an indicator in perturbations of developmental pathways, based solely on the pattern of hyperpigmentation. Class II mutants exhibit pigment on volar pads, intervolar pad scales and on phalangeal ridges. Histologically, all Class II mutants have excess pigmentation in the epidermis, as opposed to the dermis as in Class I mutants. Additionally, in all class II mutants, hyperpigmentation becomes more dramatic with age, suggesting involvement in epidermal melanoblast differentiation and/or proliferation.

### **1.12.2 Hair cycling**

Hair formation and hair follicle homeostasis is programmed by an inherent ‘clock’ mechanism within follicles. Hair follicles grow and eventually are shed from the surfaces of the skin in phases of growth (anagen), regression (catagen) and resting (telogen). In the first 2 phases of growth and regression after birth, follicles cycle synchronously<sup>70,71</sup>, providing an opportunity to study the factors driving this process. The anagen phase of the hair cycle is easily distinguishable by the presence of black skin. Shaving the fur and assessing the underlying skin colour at specific times is a simple and robust way to identify defects in cycling process. Internal data generated at the WTSI suggests that at postnatal day 43, wildtype mice are expected to have exited the anagen phase. Accordingly, black, as opposed to pink or grey skin at this time is an indication that the follicles have prematurely entered the next anagen or failed to exit the previous anagen. Previous studies suggest that the regulators of hair follicle cycling include important developmental regulators such as members of the BMP, hedgehog and Wnt signalling pathways<sup>72</sup>.

Analysis of follicle morphology through wholemount immunostaining of tail skin with directly conjugated monoclonal antibodies can provide information about folliculogenesis, in a high throughput manner. Folliculogenesis involves the specification of stem cell niches in a region of the follicle known as the bulge<sup>73</sup> and the differentiation of the accessory sebaceous gland. The hair follicle is a reservoir for stem cell populations necessary to direct the specification of cell layers which constitute the adult follicle. Tail epidermis can be peeled from the dermis after incubation in EDTA. The tissue can then be used for wholemount immunostaining<sup>74</sup>. A co-stain with KRT14, which marks basal keratinocytes, and KRT15, which labels bulge stem cells, provides information about the morphology of tail epidermis, hair follicle and sebaceous gland differentiation, spatial expression of bulge stem markers and the patterning and organisation of the hair follicles.

### **1.12.3 Wound healing**

Wound closure and repair is a complex process involving many different cell types. There are 3 overlapping phases of wound healing: the inflammatory phase, the proliferative phase, and the remodelling phase<sup>75,76</sup>. Immediately following injury, during the inflammatory phase, a fibrin clot is formed to stop blood flow and prevent the entry of pathogens. In addition, vasodilation allows mast cells, neutrophils and macrophages to infiltrate the wound, remove damaged tissue and release cytokines and growth factors necessary for re-epithelialisation. Following the inflammatory phase comes the proliferative phase which is hallmarked by the proliferation of keratinocytes and fibroblasts. During this phase, metalloproteinases release keratinocytes from the basal membrane by degrading adhesion molecules, allowing them to migrate to the wound site. Importantly, proliferation is inhibited during migration, so these two processes must work together for re-epithelialisation. After migration, keratinocytes are compressed under the fibrin clot by interacting with collagen V and the fibronectin-rich matrix via cell surface integrin expression. Keratinocytes located near the front of the emerging epithelial tongue begin to proliferate. TGF- $\beta$  and fibroblast growth factors attract fibroblasts to the wound site, stimulating proliferation and forming granulation tissue. Granulation tissue formation also involves angiogenesis, promoted by vascular endothelial cells and will eventually reconstruct the damaged dermis. Remodelling begins immediately in response to wounding with the formation of the clot and is resumed with the restoration of the dermis. Although remodelling is initiated immediately in response to wound healing, its



peak activity is 1-2 weeks after wounding and is therefore considered the last phase in wound-healing.

Although there has been much research regarding the wound healing process, the genetic underpinnings are still largely unknown. Wound healing studies are both costly and time consuming, precluding their administration in a primary, high throughput screen. However, loss of function of a gene that is normally upregulated in any of the 3 parts of the wound healing process might show a phenotype without formal wound healing experiments (ie everyday wear and tear associated with grooming, scratching etc). It can be anticipated that knockout mouse lines which are identified as abnormal in any of the primary screens are more likely to identify factors involved in other chronic aspects of skin disease such as wound healing, infection and other disorders associated with normal skin function.

### **1.13 LINES SELECTED FOR FURTHER STUDY**

After spending 1 year at as a part of the Sanger-MGP, identifying mouse models of skin dysfunction, I chose 2 different unrelated genes to study and characterize in depth. From the embryonic lethal and subviable screen, I chose to characterize a conditional knockout for *Mad2*, a mitotic spindle associated gene. From the adult pipeline, I identified a mutant carrying a mutation in KRT76, an intermediate filament protein.

#### **1.13.1 *Mad2***

During cell division, newly replicated chromosomes are distributed evenly between daughter cells. Failure of equal distribution of chromosomes results in an aneuploid, or unbalanced, chromosome content in the daughter cells. The consequences of aneuploidy at both the cellular and organism level are detrimental, with aneuploidy being one of the leading causes of mental retardation and miscarriage in humans<sup>77,78</sup>. The extreme consequences of aneuploidy make it difficult to study as it can be lethal in both cell lines and model organisms alike. In cells, aneuploidy disrupts cellular physiology and metabolism, decreases proliferative capacity and in severe cases, causes death. Conversely, in tumor cells, aneuploidy is oncogenic, thus making it a hallmark of cancer cells. Identifying the mechanism by which aneuploidy is tolerated in cancer cells might have a large impact in the treatment of aneuploid cancers through targeting the specific

pathways that allow some cells to tolerate aneuploidy. In this study, we identify the skin as a tissue that can cope with remarkably high levels of aneuploidy.

The spindle assembly checkpoint (SAC) is responsible for ensuring proper chromosome segregation and protecting against aneuploidy. In this study, SAC function was compromised through the conditional germline deletion of *Mad2*, a key signal transducer in the SAC pathway. Previous studies show that deletion of *Mad2* in cell lines causes apoptosis. Accordingly, global germ-line deletion of *Mad2*, results in severe aneuploidy and subsequent early embryonic lethality. However, conditional germline deletion of *Mad2* in the skin identified skin as a tissue that can cope with severe aneuploidy, allowing for the study of the mitotic checkpoint and the associated genes.

To study the effect of aneuploidy in the different cell lineages of the epidermis, mice carrying a LoxP floxed *Mad2* allele (*Mad2<sup>fl</sup>*) were intercrossed with *Keratin14-Cre* (*Krt14-Cre*) transgenic mice. *Mad2<sup>fl/fl</sup>; Krt14-Cre<sup>+</sup>* mice were born at normal Mendelian ratios, however, 80% of *Mad2<sup>fl/fl</sup>; Krt14-Cre<sup>+</sup>* neonates died within the first month of life. The cause of death in the neonates was initially thought to arise because of inadequate barrier formation. However, toluidine blue assays show normal acquisition of the barrier function of the epidermis, suggesting that the differentiation program in the epidermis is unaffected by loss of *Mad2*. The remaining 20% of *Mad2<sup>fl/fl</sup>; Krt14-Cre<sup>+</sup>* mice survive well into adulthood, albeit completely devoid of hair, also indicating that loss of *Mad2* does not alter IFE differentiation. The undisturbed IFE differentiation pattern together with the absence of hair was the first indication that aneuploidy might be tolerated in some proliferative cells of the IFE but not in the stem cell population of the hair follicle.

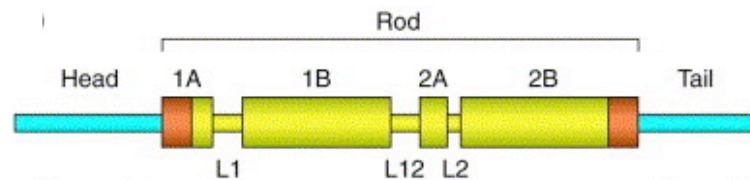
To better understand the effect of loss of *Mad2* in the epidermis and to explain the hairless phenotype, a histological examination of the skin using markers of skin differentiation was carried out. The methodology and results of this extensive skin characterization are detailed in results Chapter 3 of this thesis.

### **1.13.2 KRT76**

From the adult pipeline, I chose the knockout first *Krt76* allele to further characterize. Keratins are classically viewed as simple structural proteins, however, emerging research suggests that they are much more active participants in cellular homeostasis. Together

with microfilaments and microtubules, intermediate filaments (IFs) make up the major components of the epidermal cytoskeleton. IFs are the least understood component of the cytoskeleton and are developmentally, temporally, and spatially regulated. Keratins are the largest subgroup of the IF proteins and comprise the major structural proteins in epithelial cells<sup>79</sup>. Furthermore, keratins display a high degree of tissue specificity, indicating variable roles for keratins in the intricate architecture and physiology of different cell types and tissues.

Structurally, each keratin molecule is comprised of a central alpha helical rod domain, consisting of four alpha helical coils 1A, 1B, 2A, and 2B (Figure 1.12). The coils are separated by non-helical linker domains L1, L12, and L2. Highly conserved helix initiation and helix termination peptides flank the alpha helical rod and are involved in keratin intermediate filament interactions/assembly and filament elongation<sup>50</sup>. The N-terminal head domain is made of V1 and H1 subdomains, while the C-terminal tail domain is made up of V2 and H2 subdomains. The head and tail domains are involved in structural organization by acting as substrates for structural regulatory posttranslational modifications<sup>80</sup>. Keratins form parallel heterodimers in a coiled-coiled fashion between aligned rod domains of monomers. Lateral packing of parallel dimers into antiparallel tetramers leads to polymerization and elongation of keratin intermediate filaments<sup>81</sup>. By insertion at desmosomes and hemidesmosomes, keratin filaments play a role in cytoskeletal scaffolding, cell-cell stability and attachment to connective tissue<sup>49,82</sup>.



**Figure 1.12 Structure of keratin IFs.**

The tripartite domain structure of all keratin IFs comprises a head, rod and tail domain. The central rod domain is comprised by  $\alpha$ -helical subsegments (1A, 1B, 2A and 2B) and separated by short linker regions (L1, L12 and L2). The rod is flanked by nonhelical head and tail domains at the amino and carboxyl termini, respectively. The orange boxes depict the position of 15–20 amino acid segments that are highly conserved among keratins and other IF proteins.

*Krt76* characterization has been limited, likely because of the lack of tools available to study it (e.g. antibodies, animal models). With the generation and characterization of the *Krt76*<sup>tm1a(KOMP)Wtsi</sup> mouse, novel expression patterns and roles for this IF have been discovered and will be discussed, in depth, in this thesis. My preliminary studies at the Sanger-MGP demonstrated that *Krt76* is required for postnatal life, as all homozygous mice older than 12 weeks do not survive. Its expression is seen in a wide range of tissues, including the esophagus, vagina, palate, footpad and during cutaneous wound healing. The expression pattern alone sets this keratin apart from other intermediate filaments as usually, their expression is restricted to the hair or the skin, not both. Histologically, the loss of function of *Krt76* caused profound and unpredicted effects on the epidermis. Surprisingly these defects did not include the loss of mechanical stability or keratinocyte integrity. The unexpected expression pattern and deleterious, but unconventional phenotype indicated that *Krt76* might mediate a novel and necessary mechanism in epidermal biology. For these reasons, I chose to use the *Krt76* knockout mouse to investigate the mechanism by which this poorly understood IF regulates skin biology. The investigation into the function of *Krt76* comes from the examination of 3 alleles for *Krt76*: the knockout first reporter, *Krt76*<sup>tm1a(KOMP)Wtsi</sup> ; a global complete knockout, *Krt76*<sup>tm1d(KOMP)Wtsi</sup>, and a skin specific tamoxifen inducible allele whose expression was disrupted specifically in the epidermis using Cre-ERT expression driven by the *Krt14* promoter , *Krt76*<sup>tm1d(KOMP)Wtsi/Krt14CRE+</sup>. The knockout first conditionally ready allele allowed for the rapid generation of the 3 aforementioned different alleles for the same gene, further demonstrating the importance of generating a mouse resource targeted with this versatile cassette. The findings from studying these different lines are detailed in Chapter 4 of this thesis.

## 1.14 RESEARCH AIMS

The overall aim of this thesis is to identify genes necessary for normal cutaneous function. To address this question, I worked on a high throughput reverse mouse mutagenesis screen, looking for defects in specific aspects of skin biology in the mutant mice. The screening of knockout mice specifically for epidermal phenotypes to identify genes involved in skin development, homeostasis and disease comprises the first aim of this thesis. I have identified many unexpected skin phenotypes in the knockout mice. This

thesis investigates the genetic associations with the phenotypes I identified and explores their relevance in skin disease.

During my time at the WTSI, I worked with a cancer biologist to characterize the effect of aneuploidy in the skin. Thus, the second aim of this thesis is to use the skin as a model to study cell biology. To do this, I used *cK14* driven disruption *Mad2*, a mitotic spindle gene. Prior to this work, the consequence of chromosome missegregation and aneuploidy in skin development was unknown. Furthermore, the investigation of *Mad2* function in the skin is one of the only models that allows the characterization of the checkpoint as loss of *Mad2* is not tolerated in cell lines and other mammalian organs.

The third aim, to characterize an adult skin phenotype and potential model of human disease, was achieved using different mutant alleles for *Krt76*. *Krt76* is a very poorly studied gene whose expression is critical and necessary role in skin homeostasis. Further, the loss of function of this gene exhibits phenotypes that mimic aspects of human diseases such as psoriasis. The requirement of *Krt76* to stabilize the adult epidermis was unknown before the initiation of this thesis, as was the requirement for this intermediate filament's participation in tight junction assembly and maintenance. I will discuss the novel role of this IF in epidermal biology as well as its contribution to skin barrier function to fulfil the third research target in this thesis.

Ultimately the aim of this thesis is to highlight the promise of reverse genetic screening in mice to discover new genes involved in skin biology, study cell biology, and model genetically linked cutaneous disease.



## **CHAPTER TWO: High Throughput Cutaneous Phenotyping of Knockout Mouse Models**

### **2.1 INTRODUCTION**

This chapter reports a high throughput reverse genetic screen for skin phenotypes in mice. This is the largest organ-specific screen to be reported on mutant mice to date. I spent the first year of my research candidature working at the Sanger-MGP to screen mice for defects in skin biology. Throughout the course of this screen, I have assessed over 500 different genes and identified 23 genes causing disruptions in the development and maintenance of the skin, all of which are highlighted in this chapter. As a skin biologist, I was interested in identifying genes that may be involved in the development and homeostasis of the skin. Importantly, lines for preliminary phenotypic testing were selected only on the basis of availability (as opposed to predicted gene function); and thereby allowed for the discovery of novel genes as well as novel roles for known genes. The screen comprised multiple tests aimed at identifying defects in skin biology.

Spending a year screening the mutant mice as a part of the Sanger-MGP allowed me to access all of the mice generated there. Five years later, the data generated at the Sanger-MGP is still not updated in real time. I also was able to run my own, skin specific pilot screens. For these reasons, my placement there was essential, allowing me immediate access to mutant mice to generate phenotype data. To find genes related to skin biology, I focused on 3 main areas: adult phenotyping, recessive lethal phenotyping, and expression analysis. Each of these areas encompasses phenotyping tests that are detailed in the following manuscript.

In the adult pipeline, I worked alongside experts at the Sanger-MGP to modify tests for gross skin dysmorphology. In addition to giving an assessment of the gross skin/hair/nail morphology, this included modifying test parameters to robustly and reproducibly detect skin phenotypes. As the skin provides the outermost protective barrier, it is often subject to environmental assaults. A major challenge in this screen was distinguishing genetically linked phenotypes from incidental, background or environmentally induced phenotypes. To do this, wildtype baselines for phenotypes associated with inbred mice on different genetic backgrounds had to be established. Thus, there was a balance in

determining the sensitivity required to differentiate genetically linked phenotypes from those caused by other factors.

I also played a major role in screening mice for defects in hair follicle cycling in the adult pipeline. In addition to identifying 5 genes important for hair follicle cycling, much of the data from the hair follicle cycling screen helped to establish a baseline for normal wild type follicle cycling. This test, more so than the dysmorphology screen, was highly sensitive to gender. Initially, testing for hair follicle cycling seemed unproductive, as the data appeared to be extremely noisy. This was mainly due to a gender linked cycling difference. To determine this, I used the *Nsun2* allele as an internal positive as *Nsun2* is a well characterized and known regulator of skin biology, specifically in the areas of follicle cycling, epidermal stem cell self-renewal, and differentiation<sup>83</sup>. Accordingly, this allele was immediately connected to abnormalities in hair follicle cycling. Thus, this screen recapitulated alternate loss-of-function alleles for *Nsun2*, highlighting its role in hair and validating the testing strategy. Further, *Nsun2* results prove that the added skin specific tests increase confidence in the more subjective and general dysmorphology test, where even baseline levels of abnormal coat quality were nearly 40%. Expression data supported a role for *Nsun2* in the skin and hair follicles, as there is increased expression in the hair follicles specifically.

Another advantage to being at the Sanger was first hand access to tissue for LacZ expression analysis. In the expression pipeline, I LacZ stained and imaged wholemount tissues to determine the normal expression pattern of the targeted genes. Many of the alleles were available for expression analysis before the phenotyping could be completed. For this reason, working on the expression pipeline allowed me to look for expression patterns in the skin that might compliment preliminary phenotyping data in the phenotyping pipelines.

In addition to adult phenotyping pipelines, I also spent a lot of time working on the recessive lethal pipeline which incorporated both embryo dysmorphology and expression analysis at different developmental stages. The recessive lethal pipeline was still being established when I arrived, and the developmental stages assessed for phenotyping were still undetermined. I was able to access and generate mutant embryos at multiple stages in gestation. From a skin perspective, I was most interested in the genes whose



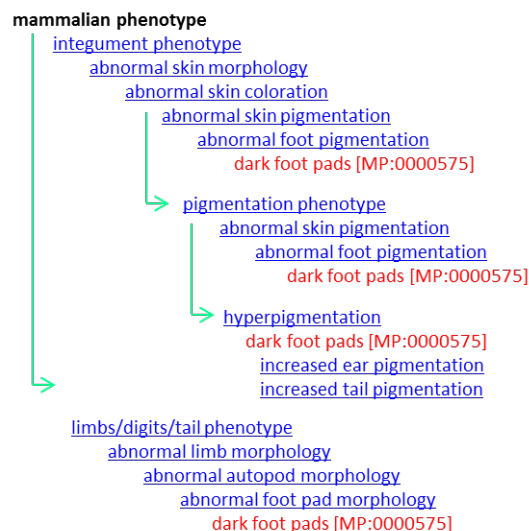
dysfunction lead to late stage lethality or lethality at weaning. This is because skin should be mostly functional by approximately 17 days of gestation and is not properly challenged *in utero*. Thus, when lethality is observed at birth, skin dysfunction is a likely contributor. For the late lethal lines, I implemented my own pilot test, a dye exclusion assay to test for defects in barrier function. Although I did not identify a novel genetic contributor to barrier function development, I still gained valuable experience in embryology. In my time at the Sanger-MGP, I had unprecedented access to targeted mutant mice. Thus, my involvement there was critical, not only for shaping my thesis, but also for the contribution to a deeper understanding of mouse genetics and skin developmental biology.

Because the skin and its associated mini-organs are dynamic and complex, multiple tests were used to examine different aspects of skin biology. The tests were implemented in such a way as to maximize indicators of skin disease associated with loss of function of a gene. To do this, each specific test comprised multiple parameters for assessment. The implementation of this testing style, in which a list of phenotypic features are cross checked, against a predetermined list of ‘normal’ parameters, was necessary for a screen as large as this. The generation of a list of normal skin features allowed people that are not necessarily skin experts to flag gross phenotypic features that may indicate skin dysfunction. This method also allows for standardization and reproducibility of tests. An example in which this proved useful was regarding pigmentation defects. As reviewed in Chapter 1 of this thesis, pigmentation in the foot pads is a recognized phenotype in mice and can provide clues to defects in a broad range of physiological processes. However, unless this particular feature is screened for, it is likely to get overlooked. A reliable reporting method had to be enacted so each line was screened in the same way for the same defects, allowing for accurate phenotype-to-gene linkage.

A major challenge in this screen was making meaningful observations that could be easily reported and searched. As the screen was in its initial stages when I arrived, I was able to give critical input regarding how skin phenotype data acquired and reported. My major contribution in this area was by modifying the parameters of the different skin screens to give searchable output in the form of Mammalian Phenotype (MP) terms. The MP ontology is a community effort to provide standard terms for annotating phenotypic data<sup>84</sup>. MP ontology allows for a more precise description of a phenotype and allows

easier access to phenotype-gene interactions. The terms are then presented in a hierarchal format, ranging from very broad, low level descriptive terms, to high level specific terms. MP ontology is structured as a directed acyclic graph (DAG), where each vocabulary term (node) may have both multiple parent term and multiple child term relationships<sup>84</sup>. This structured vocabulary aids in the standardization of phenotype annotation to unambiguously describe phenotypic observations, and makes annotation electronically accessible. Thus, every abnormal outcome for every parameter in every test can be linked to an MP term. These terms could then be immediately linked to the causative genes.

Using the example of paw pad hyperpigmentation, the MP term used is “dark foot pads [MP:0000575].” Many different textual explanations can be used to describe this phenotype. By using a standard vocabulary, this phenotype can be easily searched, both manually and computationally. In using a standard ontology, Figure 2.1 shows the different hierarchal paths leading to the dark foot pad MP term (red). As shown, there are 4 paths (green arrows) leading to this term, giving specificity and context to the term. These hierarchal maps can also be used to identify ‘parent’ and ‘child’ terms (blue). Finally, the terms, definitions, and term relationships can then be linked to genes.



**Figure 2.1 Mammalian phenotype hierarchy.** The dark foot pad MP term (red) can be identified and linked to other phenotypes. There are different paths (green arrows) that can be traversed to reach the term. The blue descriptors represent parent and child relationships to the term and can be used to find related terms.

With the implementation of this screen, the pool of available MP terms expanded rapidly, allowing for more specific and robust identification of phenotypes. This skin screen

provided major contributions to the MP term ontology, as many of the more specific tests (beyond gross skin morphology) required more specific terms. The screens for both histopathology and hair follicle cycling, for example, gave output that did not have designated MP terms, and therefore could not be easily reported or electronically linked to the causative gene when the screen was first initiated. Now however, many of these terms have been added to the ontology.

The ability to immediately link a phenotype with a specific gene represents the primary benefit of this screen. Accordingly, this chapter highlights the power of reverse genetic screening in identifying genes related to skin biology. Secondary and unexpected benefits include the generation and improvement of the MP term database with extensive skin descriptors. Another secondary benefit from this screen is the generation of ‘metadata’ that is amassed in conducting high throughput reverse genetic screens. I show that screening large numbers of genes, using a range of different parameters, is useful in identifying groups of genes that contribute to a phenotype. This metadata was further analysed to identify specific pathways that contribute to phenotypes and organ function, detailed in the below manuscript.

The global screen including the results of the preliminary phototyping tests has been published for the first 250 knockouts generated and phenotyped at the Sanger-MGP. As a collaborator, some of the work I did at the Sanger-MGP is represented in the manuscript, on which I am an author (White, et al., Appendix). Immediately following is my manuscript, written for PLoS Genetics that specifically reports the results of the overall skin screen, hair follicle cycling screen and histopathology screen.

## 2.2 DECLARATION

### Declaration for Thesis Chapter 2: Identification of genes important for cutaneous function revealed by a large scale reverse genetic screen in the mouse

In the case of Chapter 2, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Laboratory/Experimental <ul style="list-style-type: none"> <li>Primary Phenotyping of knockout mice (hair follicle cycling, dysmorphology, lacZ expression)</li> <li>Secondary phenotyping (collection of tissues, embryos)</li> </ul> Theoretical: <ul style="list-style-type: none"> <li>Designed and implemented secondary phenotyping tests</li> <li>Analysis of phenotyping data</li> <li>Facilitating data reporting</li> <li>Data analysis</li> </ul> Manuscript: <ul style="list-style-type: none"> <li>Wrote manuscript</li> <li>Made figures</li> </ul>	80%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Lynelle Jones	performed research	N/A
Sanger-MGP	performed research	N/A
Valerie Vancollie	performed research	N/A
Fiona M. Watt	Designed research	N/A
Ramiro Ramirez-Solis	Designed research, contributed reagents/analytic tools	N/A
Allan Bradley	Designed research, contributed reagents/analytic tools	N/A
Karen P. Steel	Designed research, contributed reagents/analytic tools	N/A
John P. Sundberg	Performed research, analysed data	N/A
Jacqueline K. White	Designed research, contributed reagents/analytic tools	N/A
Ian Smyth	Designed research, performed research, contributed reagents/analytic tools, analysed data, wrote manuscript	N/A

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work.

Candidate's  
Signature



Date 6/6/2014

Main  
Supervisor's  
Signature



Date

## 2.3 MANUSCRIPT

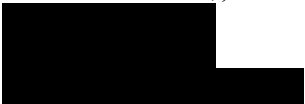
### **Identification of genes important for cutaneous function revealed by a large scale reverse genetic screen in the mouse**

Tia DiTommaso<sup>1</sup>, Lynelle Jones<sup>1</sup>, The Sanger Mouse Genetics Project<sup>2</sup>, Anna-Karin Gerdin<sup>2</sup>, Valerie E. Vancollie<sup>2</sup>, Fiona M. Watt<sup>3</sup>, Ramiro Ramirez-Solis<sup>2</sup>, Allan Bradley<sup>2</sup>, Karen P. Steel<sup>2,4</sup>, John P. Sundberg<sup>5</sup>, Jacqueline K. White<sup>2</sup>, Ian Smyth<sup>1,6\*</sup>

1. Department of Biochemistry and Molecular Biology, Monash University, Clayton, Melbourne 3800, Australia
2. Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom
3. Centre for Stem Cells and Regenerative Medicine King's College London, Guy's Hospital, London SE1 9RT, United Kingdom;
4. Wolfson Centre for Age-Related Diseases, King's College London, Guys Campus, London SE1 1UL, United Kingdom
5. The Jackson Laboratory, Bar Harbor, Maine 04609, USA
6. Department of Anatomy and Developmental Biology, Monash University, Clayton, Melbourne 3800, Australia.

\* To whom correspondence should be addressed

A/Prof Ian Smyth  
3<sup>Rd</sup> Floor Building 76  
Biomedical Sciences Precinct  
Monash University, Wellington Rd, Clayton, Melbourne, Australia



## Abstract

The skin is a highly regenerative organ which plays critical roles in protecting the body and sensing its environment. Consequently morbidity and mortality associated with skin defects represent a significant health issue. To identify genes important in the development and homeostasis of the skin we have applied a high throughput, multi-parameter, phenotype screen to the conditional targeted mutant mice generated by the Wellcome Trust Sanger Institute Mouse Genetics Project (Sanger-MGP). A total of 558 different mouse lines were subjected to a variety of tests assessing cutaneous expression, macroscopic clinical disease, histological change, hair follicle cycling, and aberrant marker expression. Cutaneous lesions were associated with mutations in 23 different genes. Many of these were not previously associated with skin disease in the organ (*Mysm1*, *Vangl1*, *Trpc4ap*, *Nom1*, *Sparc*, *Farp2*, and *Prkab1*), while others were ascribed new cutaneous functions on the basis of the screening approach (*Krt76*, *Lrig1*, *Myo5a*, *Nsun2*, and *Nf1*). The integration of these skin specific screening protocols into the Sanger-MGP primary phenotyping pipelines marks the largest reported reverse genetic screen undertaken in any organ and defines approaches to maximise the productivity of future projects of this nature.

## Author Summary

Recent developments in high throughput applications to manipulate and inactivate specific genes in mouse embryonic stem cells (ES cells) have allowed for the initiation of large scale reverse genetic screens in the mouse<sup>19,85</sup>. The immediate connection of a phenotype to a mutated (null) gene represents a paradigm shift in our ability to explore gene function. This study utilized such a screening approach to investigate the genetic contribution to skin development and homeostasis. Not only does this approach provide insight into the genetics of skin biology, it is also instrumental in generating novel models with which to study the genetic underpinnings of skin disease. Initial screening of 558 mutated genes in mice uncovered previously unrecognized genes involved in the biology of this organ and identified novel functions for previously studied genes associated with epidermal phenotypes. Taken together, these results highlight high throughput screening approaches that are valuable in reverse genetic screening and provide a pool of mouse

mutants, available to the scientific community, that will serve as the basis for further detailed investigations into skin function and skin disease.

## **Introduction**

Recent advances in the high throughput production of targeted mutant mice using gene targeting in ES cells <sup>16</sup> have made it possible to rapidly and effectively produce knockout mouse lines for use in phenotypic screening. Such approaches are now beginning to yield significant insights into mammalian gene function <sup>17,19</sup>. Central to recent community efforts in this regard has been the Wellcome Trust Sanger Institute Mouse Genetics Project (Sanger-MGP) which is undertaking a high throughput reverse genetic screen using the knockout ES cells generated by the International Mouse Knockout Consortium. To maximise the value of this resource, it is important to develop organ specific secondary screens to complement the broad phenotyping efforts applied in existing production pipelines. To this end, the study summarized here implemented a multi-parameter, multi-test skin screen to identify mouse models of skin dysfunction and to elucidate genetic pathways important for skin development and function.

The skin is the largest of the intermediate sized organs in the human body <sup>86</sup> and it plays integral roles in immune surveillance, wound healing and protection from environmental challenge <sup>87</sup>. Skin related conditions account for significant health expenditure and top the list of reasons for general practice appointments in the U.S. <sup>88</sup>. A considerable number of skin conditions have genetic underpinnings <sup>89</sup>. Therefore, understanding the genetic basis of skin development and skin disorders has the potential to greatly impact both the diagnosis and treatment of skin related conditions. Moreover, a significant number of other systemic diseases or those principally affecting other organs also manifest themselves in the skin. Such conditions include diabetes mellitus, chronic kidney disease, celiac disease, and rheumatoid arthritis. As a highly regenerative organ in which developmental gene expression is often recapitulated during adulthood, the skin also provides a window into many of the programs central to differentiation and homeostasis. The skin and its associated “mini organ”, the hair follicle, are particularly attractive targets for such a screen because of their ready accessibility and dynamic nature <sup>90</sup>. Furthermore, the clinical pathology and physiology based mouse phenotyping “pipelines” are not designed to find subtle lesions affecting the skin <sup>91</sup>. For these reasons a skin specific phenotyping program was implemented on the genetically modified mice

produced by the Sanger-MGP with a goal of defining new and previously unappreciated genetic determinants of skin disease.

In this paper 558 genes were investigated, when mutated, in new mouse lines (Appendix I, Supplementary Table 1, 2) and identified phenotypes caused by mutations in 23 unique genes (Appendix I, Supplementary Table 3). These studies identified genes that are important for maintenance of the skin not previously associated with skin phenotypes and ascribe further functions to genes which previously were associated with skin disease. We also used these tests to assign 25 *Mammalian Phenotype* (MP) terms, 2 of which are novel MP terms (Appendix I, Supplementary Table 3). The scale of the screen also provides the opportunity to assess the relative value of different phenotyping approaches, with a view to implementing these screens in other proposed and ongoing high throughput reverse genetic screens.

## **Results and Discussion**

### *The Skin Phenotyping Screen*

The vast majority of the mutant lines examined in this study (>95%) were derived from the EUCOMM/KOMP knockout first conditional-ready targeted ES cell resource on a C57BL/6NTac background<sup>17</sup>. Details of the genetic background strains are presented in Appendix I, Supplementary Tables 1 & 2. These lines are predicted to produce null alleles as a consequence of the insertion of the strong *Engrailed-2* splice acceptor upstream of a LacZ reporter cassette<sup>20</sup> (Appendix I, Supplementary Figure 1). Full details of the ES cells used and the alleles produced are available at <http://www.sanger.ac.uk/mouseportal/>. Primary screens were carried out either on viable homozygotes or, in the case of those alleles which exhibited embryonic lethality, on heterozygote animals (Appendix I, Supplementary Table 1,2). Over the course of testing, cohorts of mice (typically 7 males and 7 females) underwent visual tests to screen for obvious phenotypes (Supplementary Table 1). A number of these were designed specifically to assess cutaneous phenotypes. The appearance of the coat was initially assessed at 4 weeks of age, prior to the implementation of a high fat diet which often resulted in some disruption to the coat through incidental lipid transfer from food pellets. This screen was augmented by a visual assessment of the skin, nail units, pelage, and teeth at 10 weeks of age. At the time of necropsy (16 weeks of age) dorsal skin was collected for histology and immunohistochemical assays. Hair follicle cycling defects were evaluated at 6 weeks of



age by shaving a patch of hair at the dorsal thoraco-lumbar region. We found this time point to coincide with completion of the synchronous, second postnatal anagen phase of the hair cycle and entry into anagen<sup>92</sup>. The hair analysis and hair follicle cycling screens were included as a part of the internal primary pipeline at the Sanger-MGP, whereas the screen for skin histopathology was conducted externally. In total, we used 3 screens to examine 558 unique genes. We identified 23 genes and described 25 phenotypes (Supplementary Table 3) related to skin biology. The results of each of these specific screens are discussed below.

#### *Hair analysis and visual dysmorphology screening*

The visual assay for overt abnormalities of the pelage is a key primary screen in the Sanger-MGP pipeline and comprises the simplest aspect of the skin screen. Parameters relevant to skin in this screen include presence of an intact pelage, normal coloration, skin pigmentation (for pigmented mice), facial vibrissae morphology, presence and pigmentation of footpads, and length and morphology and number of nail units and teeth. Of the 263 phenotype parameters assessed in the Sanger-MGP pipeline (comprising 147 categorical and 116 continuous measurements), 43 (16.4%) were directly related to skin and hair<sup>19</sup> (Appendix I, Supplementary Table 4). In addition, 362 different genes were screened identifying 20 that exhibited abnormal phenotypes in one or more organ systems (Supplementary Table 1). Of these hits, 9 were directly related to skin and hair descriptors (Table 1) making skin defects the most common gross morphological finding together with craniofacial defects (n= 9) and above defects in the eye (n=4), tail/limbs (n=2), and genitalia (n=2). Of the 9 lines that exhibited a cutaneous phenotype, the most common reported phenotype was abnormal pigmentation (n=5/9; *Krt76*, *Myo5a*, *Mysm1*, *Vangl*, *Sparc*) followed by general integument phenotypes (scruffy coat, greasy coat, scaly skin n=3/9, *Krt76*, *Nsun2*, *Lrig1*) and nail phenotypes (n=1/9 *Prkab1*) (Table 1). We also examined the expression patterns of these genes in whole mount tail skin epidermis utilising the activity of the integrated beta galactosidase reporter gene and found that with the exception of two genes (*Myo5a* and *Prkab1*) all were expressed in the epidermis (Table 1). The latter genes may only be expressed at certain hair follicle stages or in a limited population of cells not evident in whole mount labelling.

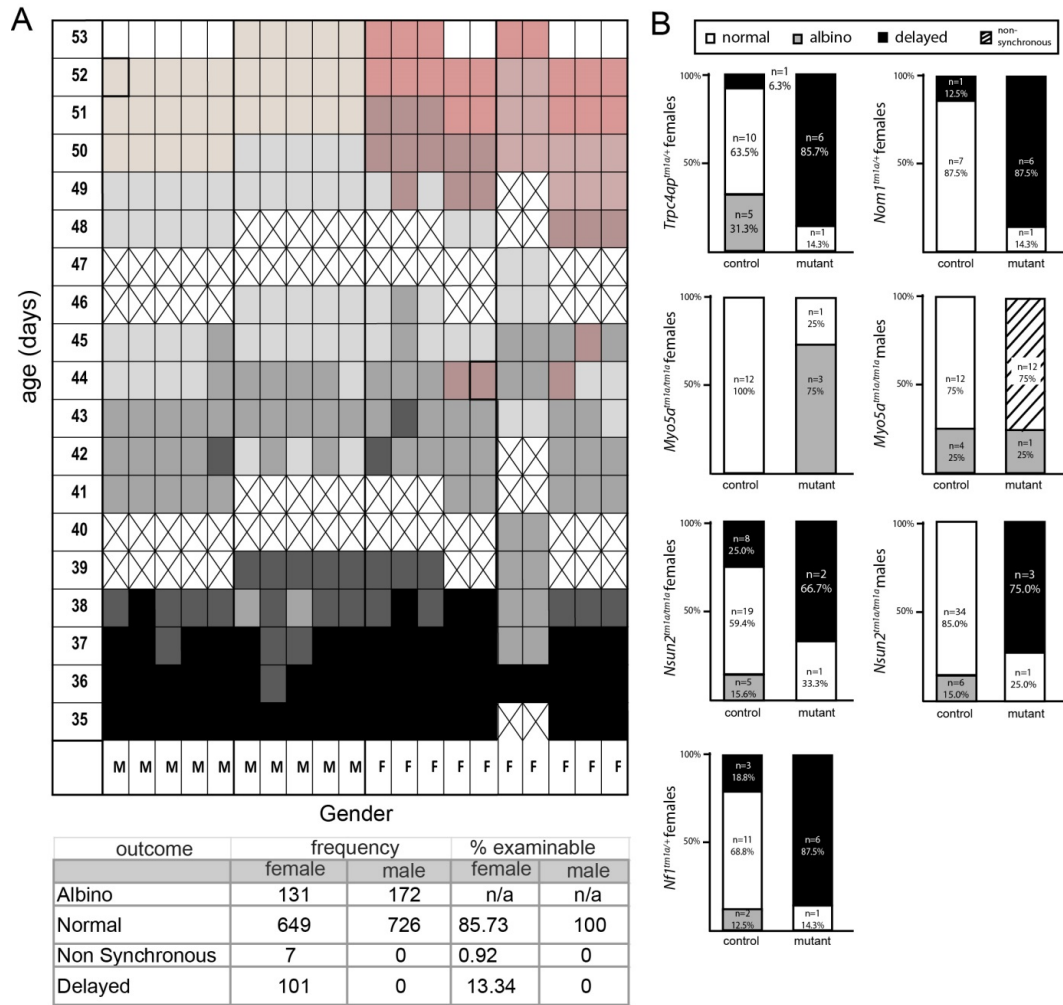
**Table1. Functions and phenotypes of genes identified in gross clinical skin evaluation screen.** Detailed gene phenotype information and genome ontology from genes identified in the gross clinical skin screen. \*=*expression detected using IHC*<sup>93</sup> as opposed to *lacZ* reporter.

Gene	Protein	Expression	Biological Function	Skin Function	Primary Phenotypic Features	Skin Associated MP term
Krt76	Keratin 76	palate, pawpad, oral epithelium, forestomach	intermediate filament, cytoskeleton, structural molecule activity	palatal keratin	behavior, integument	MP:0001510 - abnormal coat appearance MP:0010179 - rough coat MP:0000416 - sparse hair MP:0000575 - dark foot pads
Lrig1	leucine-rich repeats and immunoglobulin-like domains 1	most tissues including skin	Integral to membrane, tumor suppressor	Marker of hair follicle junctional zone stem cells	adipose, growth/size, hearing/vestibular/ear, homeostasis, immune, integument, limbs/digits/tail, skeleton, vision/eye	MP:0001191 - abnormal skin condition MP:0001192 - scaly skin
Myo5a	Myosin VA	brain, cartilage, pituitary gland, spinal cord, urinary system	ATP-dependent motor protein, actin filament based movement	melanocyte differentiation, pigmentation	integument, mortality/aging, pigmentation	MP:0000371 - abnormal coat/hair pigmentation MP:0002075 - diluted coat color
Myo6	Myosin VI	cochlear hair cells*	sensory preception of sound, auditory receptor cell differentiation, signal transduction	unknown	behavior, integument, skeletal, homeostasis/metabolism, immune	MP:0000367 - abnormal coat/hair morphology MP:0000418 - localized hair loss
Myo6l	Histone H2A deubiquitinase	most tissues including skin	histone deubiquitination, positive regulator of transcription	unknown	adipose, behavior, cellular, craniofacial, growth/size, hematopoietic, homeostasis, immune, integument, limbs/digits/tail, mortality/aging, other, pigmentation, skeleton, vision/eye	MP:0000373 - belly spot MP:0000574 - abnormal foot pad morphology MP:0000575 - dark foot pads
Nsun2	NOL1/NOP2/Sum domain family member 2	ubiquitous	RNA methyltransferase, cell cycle, cell division	Hair follicle stem cell self-renewal	adipose, behavior, craniofacial, growth/size, hematopoietic, homeostasis, integument, limbs/digits/tail, reproductive, skeleton, vision/eye	MP:0001510 - abnormal coat appearance,
Prkab1	protein kinase, AMP-activated, beta 1 non-catalytic subunit	brain, eye, liver, skull	fatty acid biosynthetic process, fatty acid metabolic process, lipid metabolic process	unknown	resistance to diet induced obesity, liver steatosis, and hyperinsulinemia	MP:0000579 - abnormal nail morphology
Sparc	secreted acidic cysteine rich glycoprotein	ubiquitous	bone and lung development, cell migration, tissue remodelling	unknown	craniofacial, skeleton, vision/eye	MP:0010096 - abnormal incisor color
Vangl1	vang-like protein 1	skin, brain, cartilage, colon, bladder, eye, kidney, bone, thyroid, intestine, heart, lung, oesophagus, parathyroid, spinal cord	development, integral to membrane, protein binding	unknown	hematopoietic system, homeostasis/metabolism, integument	MP:0003849 - greasy coat, MP:0000574 - abnormal foot pad morphology, MP:0000575 - dark foot pads

### *Hair Follicle Cycling*

An overt visual assay can be an informative indicator of a skin defect. However, the hair cycle cannot be evaluated this way. The hair follicle is a dynamic mini organ that continually undergoes a controlled program of growth (anagen), regression (catagen), and resting (telogen). Accordingly, hair follicle morphogenesis is maintained by a complex program of proliferation, differentiation, and death. Regulators of this process include developmentally important proteins including members of the bone morphogenetic protein (BMP), hedgehog (HH) and wingless-related MMTV integration site (WNT) signalling pathways, as well as a variety of cytokines, transcription factors, and adhesion molecules (reviewed in <sup>94</sup>). A simple way to examine the stage of hair follicle cycle at a gross level is to assess skin color in shaved animals; as follicles in anagen darken the skin <sup>95</sup>. This process is non-invasive and requiring of minimal expertise, making it highly suitable for application to high throughput screens. In newborn mice, the follicles of the animal undergo two synchronous cycles of growth and regression after birth <sup>70,71</sup>. This synchronicity affords a unique opportunity to study the factors driving this process. After 6 weeks of age hair follicles progress through the catagen stage of the cycle (3-4 days) to telogen<sup>96</sup>. This process correlates with a change in skin colour from black to pink. However, if the cycle is changed by either premature entry or delayed exit from anagen, it will be black. To generate a baseline dataset using the predominant background strain of the mice used in the screen we examined a cohort of C57BL/6NTac mice (n=10 females, 10 males) and assessed the colour of a shaved patch of skin daily for 3 weeks from 35 days post-partum. As expected, initially all of the mice were in the second postnatal anagen phase (and hence were black skinned) and all animals had exited anagen by 42 days of age (Figure 1A).

This pilot generated a reference baseline for hair cycling in wild type mice which was further refined through the addition of genotyped wild type control animals generated for weekly control cohorts for the phenotyping pipelines (a total of 1483 additional control mice (n=757 females, n=726 males)). We found a small but significant sex difference in cycling such that 13.3% of female animals had slightly delayed anagen exit (Figure 1A). Using these baseline measurements, a total of 362 mutant mouse lines (Supplementary Table 1) were then examined for differences in the hair cycling, typically using 7 male and 7 female mice for each strain (exact numbers for each strain are available at <http://www.sanger.ac.uk/mouseportal/>).



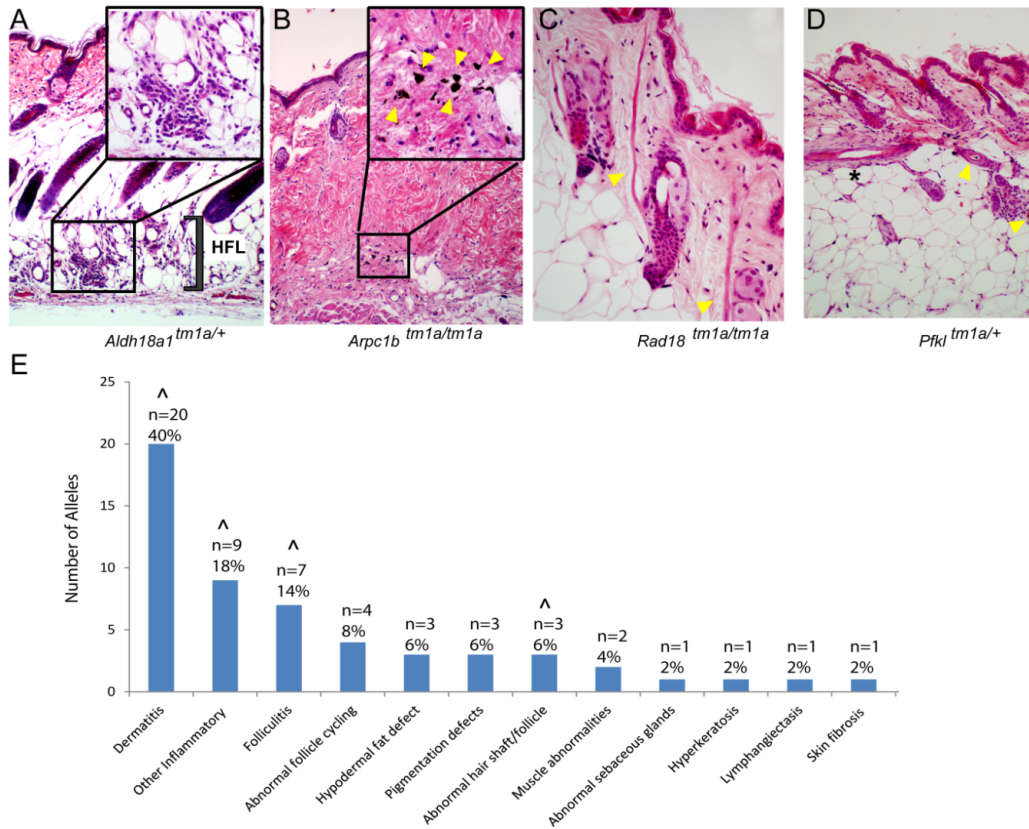
**Figure 1. Hair follicle cycling baseline and phenotypes.** Wildtype cohorts helped to establish a baseline for timing of hair follicle cycling. The results of a pilot study on 10 males and 10 females is depicted in a grid where the cell color represents the underlying skin colour after shaving. Pilot data was expanded on by incorporating the results of the wildtype controls (table) (A). Hair follicle cycling test flags 5 genes with cycling abnormalities (B). Full details of alleles are as follows: *Nf1*<sup>tm1a(KOMP)Wtsi</sup>, *Nsun2*<sup>tm1a(EUCOMM)Wtsi</sup>, *Myo5a*<sup>tm1e(KOMP)Wtsi</sup>, *Trp4ap*<sup>tm1a(KOMP)Wtsi</sup>, *Nom1*<sup>tm1a(KOMP)Wtsi</sup>.

This screen identified 5 genotypes with abnormalities in follicle cycling: *Trp4ap*+, *Nom1*+, *Nsun2*/*Nsun2*, *Nf1*+, and *Myo5a*/*Myo5a* (Figure 1B). With the exception of *Nsun2*, none of these genes had previously been reported to influence hair follicle cycling. Some lines carried a tyrosinase allele (C57BL/6JTyrc-Brd) which is why they were not included in this screen. As a consequence our screen most likely underestimates the incidence of these defects. We found two distinct types of cycling defects (Figure 1B) - a

delayed exit from anagen in which the skin remained black (*Trpc4*, *Nom1*, *Nsun2*, *Nf1*) and a non-synchronous hair cycling phenotype in which patches of the coat were at different stages of the follicle cycle (*Myo5a*). Interestingly, a number of these lines were heterozygotes, further illustrating the value in screening such animals for those lines which are homozygous lethal.

#### *Skin histopathology and marker screening*

An additional layer of screening was used to ascertain subtle phenotypes which might have been missed in other pipelines. This screen was undertaken at the point of terminal necropsy (age 16 weeks) at which point dorsal skin was collected for histopathology assessment. Tissue sections from the mutant mouse lines were stained with H&E and underwent blinded assessment by an expert pathologist. As a further resource, these samples were also immunolabeled for Keratin 14 and Keratin 10, markers of the basal and suprabasal keratinocyte layers respectively. A total of 351 different homozygous knockout lines and 166 heterozygotes from lethal and sub-viable strains, encompassing 508 unique genes, as well as 46 wild type control mice were subject to this screen (Supplementary Table 2). Both test and wild type mice exhibited a considerable incidence of mild, focal low-grade inflammation including folliculitis and dermatitis; confounding some of the results of the histopathology screen. We surmise that many of the abnormalities recorded in the control cohorts were a consequence of either the high fat diet challenge or group housing arrangement under which the animals were maintained. Despite the background infection observed in the wild type controls, we were still able to identify gene-specific phenotypes. In total, 13 different mutant lines were identified as presenting with phenotypes not observed in control animals (Figure 2). These included altered hypodermal fat levels (3 lines; *Aldh18a1*, *Ccdc57*, *Prmt3*) (Figure 2A), pigmentation (3 lines; *Anks4b*, *Krt76*, *Prkcz*) (Figure 2B), muscle (2 lines; *Scn3b2*, *Rad18*) (Figure 2C), sebaceous glands (*Kdm4c*), hyperkeratosis (*Krt76*) and aberrant follicle cycling (*Pfkl*, *Nsun2*, *Farp2*, *Lrig1*) (Figure 2D). We identified 12 different hits in the histopathology screen, 8 of which were only detected in mutant mice and 4 detected in both wildtype and mutant mice (Fig. 2E).



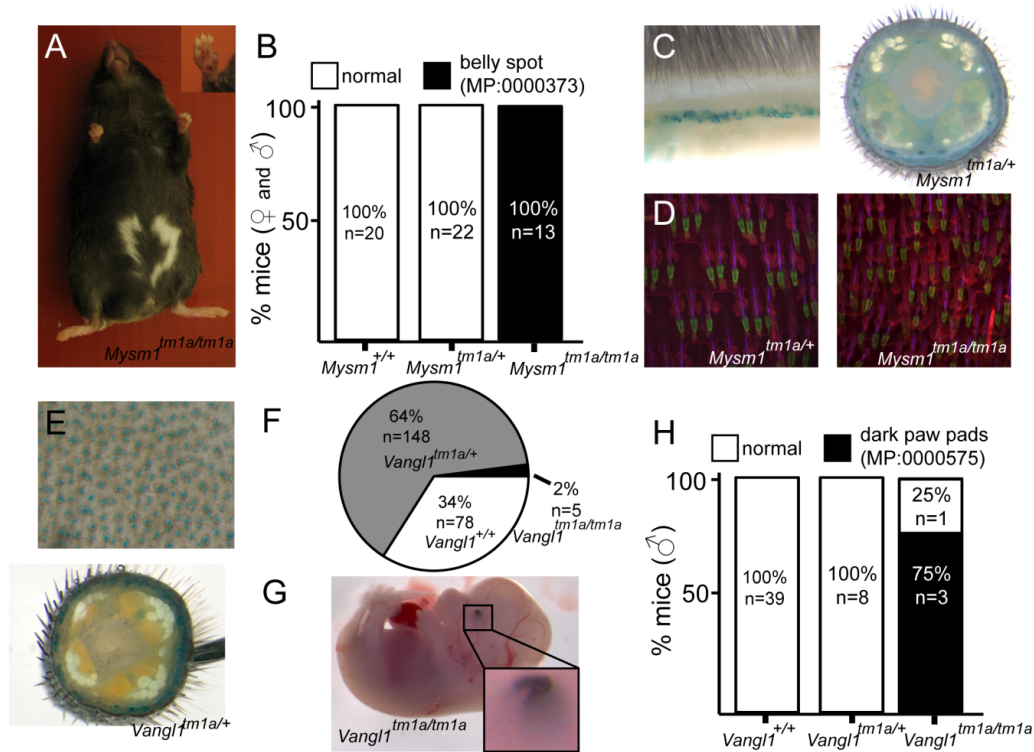
**Figure 2. Skin histopathology overview.** (A-D) Examples of various skin histopathology phenotypes are presented (arrowheads and asterisks indicate abnormalities). *Aldh18a1*<sup>tm1a/+</sup> mutants presented with minimal fibrosis and inflammation in the hypodermal fat layer (HFL) (A). Various pigmentation phenotypes represented 6% of the phenotypes identified in the histopathology analysis, with *Arpc1b*<sup>tm1a/tm1a</sup> mutants showing pigmentary incontinence (B). Examination of *Rad18*<sup>tm1a/tm1a</sup> skin showed prominent thick erector pili muscle (C). Different cycling phenotypes were identified with *Pfkfb1*<sup>tm1a/+</sup> mutants exhibiting hair follicle dystrophy of a disoriented follicle (\*) and mild inflammation (lower yellow arrow). A total of 50 targeted alleles showed abnormalities in one or more of the 12 phenotypic categories listed, and the percentage incidence as well as the number of alleles is presented. Phenotypes also identified in wild type cohorts are indicated with an ^ and therefore were not included in the overall phenotype analysis (E). Full details of alleles are as follows: *Aldh18a1*<sup>tm1a(KOMP)Wtsi</sup>, *Arpc1b*<sup>tm1a(EUCOMM)Wtsi</sup>, *Rad18*<sup>tm1a(EUCOMM)Wtsi</sup>, *Pfkfb1*<sup>tm1a(EUCOMM)Wtsi</sup>.

### Case studies

A number of the screens identified new players in cutaneous biology or identified new phenotypes not previously associated with these genes. Some individual examples are discussed.

### Pigmentation

Pigmentation defects were one of the principal phenotypes reported (Table 1). Mice carrying mutations in genes responsible for melanocyte migration or the melanin production/transport machinery have been instrumental in establishing the field of mouse genetics. The efforts of the “mouse fancy” in the late 19<sup>th</sup> century and the pioneering work of Cuenot and Castle in the early 20<sup>th</sup> century utilised coat colour to first demonstrate Mendelian inheritance in mammals<sup>97,98</sup>. Given the considerable period of time which has since passed it is remarkable then that our screen uncovered new genes and phenotypes associated with the biology of the melanocyte. The first of these is *Mysm1*, which encodes a histone 2A deubiquitinase that regulates stabilisation of linker histone H1 with nucleosomes<sup>99</sup> and which is required for bone marrow stem cell function and haematopoiesis<sup>100</sup>. Melanocyte migration and pigmentation were affected in *Mysm1* mutant mice that resulted in a white belly patch (Figure 3A). There were also defects apparent in the pigmentation of the foot pad (Figure 3A insert). Foot pads normally have very little pigmentation, but there is an emerging family of genes whose disruption results in their hyperpigmentation<sup>101-103</sup>. Beta-galactosidase reporter activity indicates that *Mysm1* is expressed in the hair follicles, principally in the lower region of the hair follicle (Figure 3C), a structure which is associated with a resident population of stem cells<sup>73</sup>. Additional parallel studies indicated that *Mysm1* is required for the normal patterning of hair follicles and sebaceous glands in the tail epidermis (Figure 3D). Taken together, two distinct and previously undescribed manifestations of cutaneous *Mysm1* function are reflected in defects in pigmentation and patterning respectively. While the mechanism by which *Mysm1* acts to regulate these phenotypes is unclear, its known functions do draw attention to potential roles for histone modification and stem cell niche activation in regulating these processes.



**Figure 3. Examples of pigmentation phenotypes and expression patterns in genes with novel roles in skin biology.** *Mysm1*<sup>tm1a/tm1a</sup> mice exhibited a range of pigmentation defects, including belly spots, and foot pad hyper-pigmentation (insert) (A, B). LacZ reporter expression of *Mysm1* is detected in the dorsal skin (left) and tail (right) whole mounts (C). Tail whole mount labelled with KRT14 (red) and KRT15 (green) indicate defects in hair follicle organization and associated structures in *Mysm1* knockouts (D). LacZ reporter expression of *Vangl1* in the ear skin (top) and tail whole mounts (lower) (E). *Vangl1*<sup>tm1a</sup> mice are neonatal lethal (F) and display abnormal eye development (insert) (G). Targeted mutagenesis of *Vangl1* reveals dark foot pads in 75% of knockout males (H). Full details of alleles are as follows: *Mysm1*<sup>tm1a(EUCOMM)Wtsi</sup>, *Vangl1*<sup>tm1e(KOMP)Wtsi</sup>.

Pigmentation differences were also noted in mice carrying mutations in *Vangl1*. VANGL1 is known to be involved in the planar polarity pathway<sup>104</sup> and mutations in *Vangl1* have been linked to neural tube defects<sup>105</sup>. A heterologous protein, VANGL2, is also known to be important for establishing hair cell orientation in the inner ear<sup>106</sup>. Consistent with these observations we observe extensive expression of *Vangl1* in the hair cells of the ear and tail skin (Figure 3E). Very few homozygous *Vangl1* knockouts survive to weaning (2% of live births from *Vangl1*<sup>tm1e(KOMP)Wtsi</sup>/+ intercrosses) (Figure 3F), and *Vangl1* knockouts have been identified with eye defects as early as ED14.5 (Figure 3G). In the small number of adult knockouts generated, we observed hyperpigmentation of the paw pads specifically in homozygous male animals (3/4

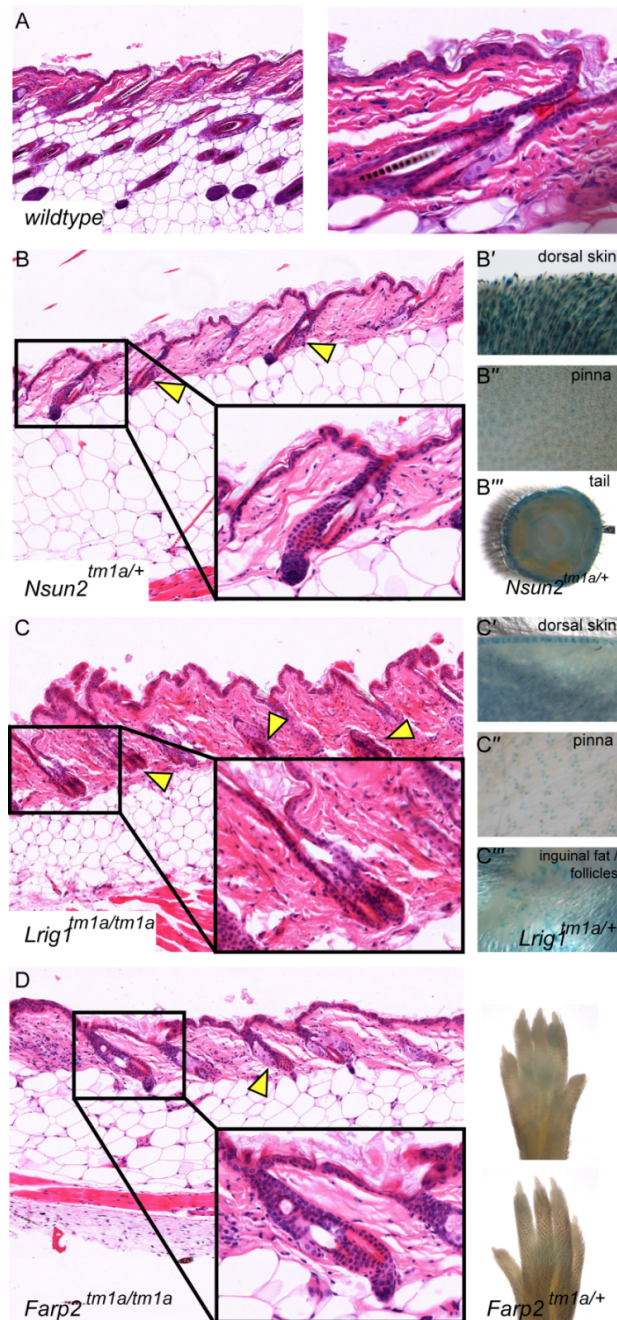


homozygotes) but not in females (0/3 homozygotes)(Figure 3H). Given the low animal numbers for this semi-lethal line we cannot discriminate between variable penetrance of the phenotype and a *bona fide* sexual dimorphism. However, as no pigmentation defects have previously been associated with loss of *Vangl1* function our study therefore adds a new member to the family of genes affecting cutaneous pigmentation.

## Exogen

Exogen describes the stage of the hair cycle during which club hair fibres are shed from the hair follicles, incorporating both the actual loss and the molecular events which promote this event<sup>107</sup>. Of the four genes identified in our histopathology screen that showed hair follicle cycling defects, three (*Nsun2*, *Lrig1*, and *Farp2*) were exogen specific (Figure 2, Figure 4) and none of them have previously been associated with this process. All of these genes show expression in the skin, with *Lrig1* and *Nsun2* expressed specifically in the hair follicles (Figure 4). Absence of distinguishable *Farp2* expression may be due to the follicle stage or skin type examined. Two of these genes, *Nsun2* (Figure 4B) and *Lrig1* (Figure 4C), are recognized for their involvement in cMYC (MYC) signalling. *Nsun2*, also known as MYC induced SUN domain containing protein 2 (*Misu2*), is a direct downstream target of MYC in the skin and is essential for keratinocyte proliferation<sup>108</sup>.

Mutations in *Nsun2* are known to affect the entry of hair follicle stem cells into the regenerative program and in regulating epidermal stem cell self-renewal and differentiation<sup>83</sup>. The gene was initially flagged in our screen as a result of coat abnormalities identified during primary visual assessments (Table 1) and as a factor in regulating hair follicle cycling (Figure 1B, Figure 4B). *Nsun2* is actively and dynamically transcribed in the hair follicle (Figure 4B). These observations correlate with other reports from studies of knockout mice<sup>109</sup>. Like *Nsun2*, *Lrig1* is also a MYC target gene<sup>108</sup> and it similarly helps to maintain the proliferative capacity of keratinocyte stem cells<sup>110</sup>. LacZ reporter studies highlight *Lrig1* expression in both the hair follicles and the inguinal fat pads surrounding the follicles (Fig. 4C''').

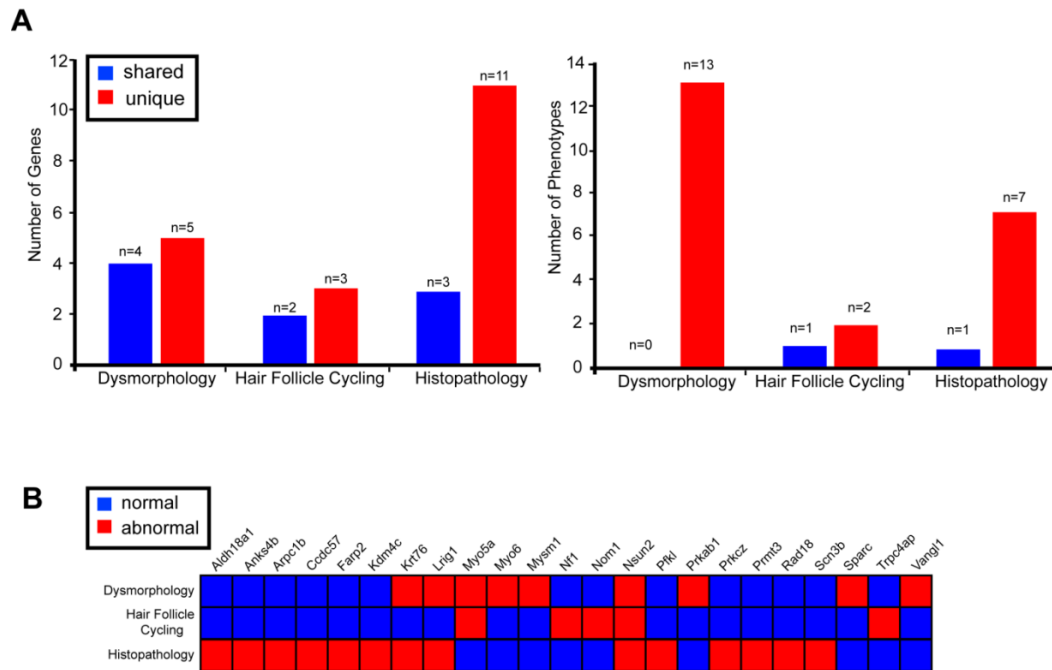


**Figure 4 High throughput screening provides insights into molecular mechanisms of exogen in hair follicle cycling.** In wildtype animals, the base of the telogen follicle shaft separates into two shafts. The early club formation between the 2 shafts is the first step of exogen. NB the club hair is still retained (A). We have identified 3 genes, *Nsun2* (B), *Lrig1* (C) and *Farp2* (D) with defects in the process, resulting in the premature shedding of the club hair (B-D). LacZ reporter expression indicates that all genes are highly expressed in skin (B-D). Full details of alleles are as follows: *Farp2*<sup>tm1a(KOMP)Wtsi</sup>; *Lrig1*<sup>tm1a(EUCOMM)Wtsi</sup>; *Nsun2*<sup>tm1a(EUCOMM)Wtsi</sup>.

Interestingly, recent studies have implicated intradermal adipocyte precursors in follicular stem cell activation through PDGF signalling<sup>111</sup>. The potential influence of extracutaneous adipocytes on the hair cycle is slowly being unravelled<sup>111</sup>. The third gene, *Farp2* (Figure 4D), is a guanine exchange factor known to be involved in osteoclast podosome dynamics<sup>112</sup> and axonal repulsion<sup>113</sup>. There have been no direct reports of the protein interacting with MYC signalling or hair follicle cycling; however, it has been found to be a specific activator of RAC1 signalling, a proposed regulator of MYC activity in the skin<sup>114</sup>. Although none of these genes changes significantly in expression as the club fibre nears release at the end of exogen<sup>18</sup> it could be that they (or MYC signalling generally) regulate this process through downstream signalling or by establishing a cellular phenotype that affects the transition to exogen. The signalling and adhesion factors mediating the active exogen process are considered to be regulated independently of normal hair follicle cycling<sup>115</sup>, potentially explaining why we did not see defects in our hair cycling screen in the *Lrig1/Lrig1*, *Nsun2/+*, and *Farp2/Farp2* lines.

### **Screen productivity**

A critical question for any screening approach is the relative value it brings in respect of identifying new phenotypes. While all skin screens identified some genes that were also identified in other tests, there were still sets of genes that were unique to each specific test and therefore not identified in any other screen (Figure 5A). For example three out of the five genes identified in the hair follicle cycling screens were not identified using any other screening approach in our pipeline (Figure 5A,B). This is also particularly true of the histopathology screen in which 11 of the 14 positive genotypes recorded no other hits in other screens – although the depth of screening by this approach was limited to 1-2 mice per line. More detailed screens at this level, employing larger number of mice will be required to accurately define these phenotypes. We also noted that each test identified MP terms that were not described by any other tests (Figure 5A). That is to say, the addition of extra screens for subtle phenotypes that were not reflected in gross changes to appearance, were reliably able to identify both new genes and new phenotypes (Figure 5A,B). Taken together this strongly argues for the inclusion of these more involved screens to flag histopathology and phenotype, particularly in light of the relatively low hit rates achieved using simple, gross observation, phenotypes.



**Figure 5. Multi parameter, multi test, organ specific screens add to the number of genes identified in skin biology.** Each of the 3 different tests identified genes that were represented in other screens (shared) as well as genes that were unique to the test (A, left). Histopathology screen identified the highest number of unique genes flagged with 79% (n=11/14) of genes not represented in any other skin screen. Different screens highlighted 25 unique MP terms, with only 1 MP term represented in multiple tests (A, right). Twenty three unique genes were identified across the 3 different skin tests with 4 represented in multiple tests (B).

## Observations

We detail one of the most comprehensive reverse genetic skin screens yet reported in mice. The gene targeting strategy, combined with a robust skin phenotyping protocol has allowed immediate establishment of a preliminary connection between a genotype and a phenotype. Such approaches have only become feasible through the recent advent of high throughput, targeted mouse mutagenesis screens. Clearly, more detailed follow-up studies are necessary to elaborate on these initial findings and to investigate in greater detail the contribution of a given gene to the biology of the skin. Nonetheless high throughput approaches of this type will be central to flagging genes for further study. An argument might be made that molecules important to something as visually evident as skin development and homeostasis would already have been identified through mutagenesis approaches, or historically as a result of sporadically occurring mutations. However, the identification of two new genes that cause defects in pigmentation (*Mysm1* and *Vangl1*)

and a further additional pigmentation phenotype in *Myo5a* mutants shows that even for the most obvious cutaneous feature, there remains considerable value in ongoing screening for new causative mutations. In particular, clinical and histological presentations vary between alleles based on the type of mutation (null, hypomorph) or inbred strain background (strain specific modifier genes), and so comparison between these “knockout ready” targeted mouse lines and existing spontaneous mutants present in various public repositories will be particularly informative in respect of gene function.

Similarly the skin might be considered an “easy” target for phenotypic screens, and that important genes could simply be identified based on a casual observation of the mutant animal in question. Our results argue that this is not the case. Instead we find that there is significant value in including secondary, organ specific screens in large scale projects of this nature. In deciding which phenotypic screens to apply to a given mutant mouse population one has to consider the likelihood that a given test will identify a phenotype not also apparent in other primary phenotyping screens. The object of such an assessment is therefore to reduce the costs and processes associated in examining such large cohorts of mice, but to maximise the value of the resource (i.e. in phenotypes flagged). Interestingly, based on our experience, each of the phenotypic screens utilised added significantly to the base level of hits in the screen (Fig 5). While gross skin abnormality assessments shared the greatest overlap with the other screened employed there were considerable numbers of genes identified in the more subtle screens that did not result in an obvious, visually assessable, phenotype in the mouse. The hair follicle cycling screen is a case in point. To all intents and purposes the visual phenotype of those animals which display abnormal hair follicle cycling is normal. However, a simple estimation of cycling stage uncovered a further class of mutants whose involvement in this process was never previously appreciated. The same is true of the section based histopathology screen, which was able to identify subtle but potentially valuable phenotypes. In both cases, whether through the subtle temporal regulation of follicle cycling or via the cellular diversity assessed in the pathology screen, looking a little harder, can be a particularly valuable exercise. Taken together our screen has identified a diverse collection of genes which will form the basis of further studies aimed at elucidating gene function in more detail. They also highlight the value of including detailed skin phenotyping approaches in any future efforts of this nature.

## Materials and Methods

### *Ethics statement*

All mouse studies were undertaken by Wellcome Trust Sanger Institute Mouse Genetics Project as part of the International Mouse Phenotyping Consortium and licensed by the UK Home Office in accordance with the Animals (Scientific Procedures) Act 1986.

### *Mouse resources*

Mice were generated from ES cells in the KOMP/EUCOMM pipeline and principally compromised conditional ready alleles in which a LacZ reporter cassette was integrated upstream of a floxed essential exon<sup>17</sup>. With a few exceptions (n=24/532) the mice were generated from C57BL/6N ES cells and maintained on this genetic background. A proportion of the lines examined (n=201), generally those that entered the pipeline in the initial stages of the project, were maintained on a C57BL/6N;C57BL/6Brd-Tyr<sup>c-Brd</sup> background (C57BL/6N-Tyr<sup>c-Brd</sup>). Details of the strain background, age of sampling, diet, sex, time of tissue harvest, allele name and heterozygosity/homozygosity are detailed in Supplementary Tables 1 and 2. Animals were maintained in a specific pathogen free environment with ad libitum access to food (typically Mouse Breeders Diet (LabDiets 5021-3, IPS, Richmond, USA)) and water. Specific details of housing and husbandry for individual lines are available from [mouseinterest@sanger.ac.uk](mailto:mouseinterest@sanger.ac.uk). For some lines (indicated in Table Supplementary Table 2) mice at 4 weeks of age were transferred to a high fat (21.4% fat by crude content; 42% calories provided by fat) dietary challenge (Special Diet Services Western RD 829100, SDS, Witham, UK).

### *Primary phenotype screen*

All mice generated by the Sanger-MGP underwent a broad primary phenotype screen (“dysmorphology screen”) which included the visual assessment of skin and hair at 4 weeks of age (before the implementation of a high fat diet in the event one was used, see Appendix I, Supplementary Tables 1 and 2) and a second phenotyping assessment just prior to termination at 16 weeks of age. Hair follicle cycling was assessed between 41-43 days of age by shaving a small patch of skin on the thoraco-lumbar region and assessing the underlying skin colour. Except for albino lines or those with reduced viability (see Supplementary Tables 1 and 2), a total of 7 male and 7 female mice were examined. In addition, tail skin (n=2) was collected at necropsy for whole mount immunohistochemistry (IHC) analysis using antibodies against keratins (KRT) 15 and 14

(as per <sup>116</sup>). The activity of the integrated beta-galactosidase cassette was assessed using heterozygous cohorts of mice (n=2 male, 2 female) generated specifically for expression analysis. The pattern of B-gal reporter gene activity was determined by LacZ labeling of whole mount tissue preparations from heterozygous knockout. Mice were perfused with fresh cold 4% paraformaldehyde (PFA). Tissues were fixed for a further 30 min in 4% PFA, rinsed in phosphate buffered saline (PBS) and labeled with 0.1% X-gal for 48 hours at 4°C. Samples were post-fixed overnight in 4% PFA at 4°C, cleared with 50% glycerol, transferred to 70% glycerol and imaged.

#### *Histopathology Analysis*

Dorsal skin at the thoraco-lumbar region/interscapular region was collected at the time of necropsy (at least 2 mice per sex), fixed in neutral buffered formalin solution, processed routinely, embedded in paraffin, sectioned at 6 µm, stained with hematoxylin and eosin (H&E). Samples from 1-2 male animals were reviewed by an experienced, board certified veterinary anatomic pathologist (JPS). Animals intended for histopathology were reserved for this purpose and not subjected to other phenotyping tests. Duplicate slides were also labelled by immunohistochemistry with keratin (KRT) KRT14 (Rabbit anti- Mouse @ 1:1000, Covance), KRT10 (Rabbit anti Mouse @ 1:1000, Covance) and Fillagrin (FLG) (Rb anti Ms @ 1:1000, Covance) and detected using diaminobenzidine (DAB, Dako, city, state) as per manufacturer's instructions. All slides were scanned and images archived using Zeiss Mirax Slide Scanner and associated software. In addition, representative photomicrographs were taken of all H&E stained slides). Images are available at the Sanger MGP mouse portal ([www.sanger.ac.uk/mouseportal](http://www.sanger.ac.uk/mouseportal)) and on the Mouse Genome Informatics webpage (<http://www.informatics.jax.org/genes.shtml>).

#### **Acknowledgements**

This work was supported by a Discovery Project grant from the Australian Research Council (DP1092723) to IS and in part by a grant from the National Institutes of Health (US; AR063781) to JPS. IS acknowledges fellowships support from the NHMRC, ARC and Monash University.





## CHAPTER THREE: Spindle Checkpoint Deficiency in Skin

### 3.1 INTRODUCTION

The use of skin genetics to better understand cell biology is exemplified in this chapter. The skin contains cell populations that are constantly dividing, therefore making it a suitable system in which to study mitosis and the mechanisms that regulate it. The spindle assembly checkpoint (SAC) is a surveillance mechanism that monitors kinetochore attachment to the mitotic spindle. The spindle assembly checkpoint is an essential cell cycle control mechanism in mammals that acts at each mitosis to prevent premature mitotic advance, thereby ensuring accurate chromosome segregation. The consequence of SAC disruption is compromised surveillance during mitosis. In progeny cells, this can lead to either no phenotype, if chromosome attachment to the spindle is efficient, or more often to aneuploidy, the condition of having an unequal complement of chromosomes. In such cases, the viability of the progeny is dependent on whether the unequal chromosome complement can be tolerated. One caveat to studying this mechanism *in vivo* is that it is absolutely required during development, since loss of function of genes involved in the SAC leads to early embryonic lethality<sup>117</sup>. One way to circumvent the problem of embryonic lethality when trying to study genes involved in cell cycle biology *in vivo* is to use tissue specific induction of gene disruption.

In this chapter, skin specific deletion of a key SAC gene, *Mad2*, is used to show the effect of checkpoint inactivation on the different epidermal lineages. The SAC is only required in cells that are mitotically active. Thus, the skin, with its highly regenerative properties, is a dynamic model system in which to study SAC associated genes. The occurrence of a highly aneuploid epidermis in response to loss of function of *Mad2* was not surprising. However, the phenotype, a functional but hairless epidermis, indicated that stem cells (hair follicles) are more sensitive to the checkpoint disruptions than the more committed lineages of the skin (interfollicular epidermis). The skin, with its heterogeneous population of cell types, proves to be an excellent organ to model differential responses to abrogated cell cycle control.

One of the fundamental gains from this research is identifying the skin as a tissue that can cope with remarkable levels of aneuploidy. In the broader sense, this study validates the argument for combining skin biology, mouse models and genetics to better understand cellular processes. Although the interfollicular cells of the epidermis can cope with aneuploidy, the loss of SAC function is not inconsequential as indicated by mRNA expression profiles. Since aneuploidy and chromosome instability are hallmarks of cancer cells, the *Mad2* skin conditional mouse could be a first step toward identifying currently unknown mechanisms and pathways that allow cells to survive with severe aneuploidy. This chapter highlights the importance of the skin as a system to model cell biology and to understand the genetics regulating cell biology. It also gives insights into the skin itself and the differential response of cells to aneuploidy.

I began working on the *Mad2* conditional mouse to supplement my work on the skin genetics screen at the Sanger-MGP and to demonstrate how the skin can be used as a system to study cell biology. When these mice were first generated, I was the only skin biologist working in the collaboration to describe the mutant and try to find a mechanism explaining the phenotype. This was the first opportunity for me to take the lead in an in depth characterization of the consequence of genetically induced skin dysfunction. It was not surprising that the inactivation of a critical gene such as *Mad2* in the skin showed a skin phenotype. However, the phenotype was surprising and proved yet another example in which our ability to predict the consequence of knockout is poor. I undertook an analysis of the postnatal skin to reveal an expansion of the epidermal layers, aberrant KRT14 and KRT10 expression and increased apoptosis and inflammation. Interestingly, these errors corrected themselves by the time the mice reached adulthood. The work I did to characterize this model was integral in providing insight into the mechanism by which aneuploidy is dealt with in the skin.

### 3.2 DECLARATION

#### Declaration for Thesis Chapter 3: Spindle checkpoint deficiency is tolerated by murine epidermal cells but not hair follicle stem cells

In the case of Chapter 3, the nature and extent of my contribution to the work was the following:

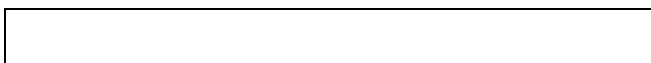
Nature of contribution	Extent of contribution (%)
Laboratory/Experimental <ul style="list-style-type: none"> <li>Macroscopic description of phenotype of <i>Mad2</i> conditional KO (cKO) mice in neonates and adults</li> <li>Collection of RNA and protein to confirm loss of function of <i>Mad2</i> in skin tissues</li> <li>Documentation of survival of <i>Mad2</i>-deficient mice within the first year</li> <li>Implementation and optimization of all assays to identify epidermal morphology, hyperkeratosis, and differentiation.</li> <li>Collection of cells for micro arrays</li> <li>Collection tails for follicle analysis</li> </ul> Theoretical: <ul style="list-style-type: none"> <li>Analysis of phenotype</li> <li>Analysis of microarrays</li> </ul> Manuscript: <ul style="list-style-type: none"> <li>Edits and comments</li> </ul>	30%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Floris Fojier	designed research, performed research, contributed new reagents/analytic tools, analysed data, wrote the paper	N/A
Giacomo Donati	performed research	N/A
Katta Hautaviita	performed research	N/A
Stephanie Z.Xie	performed research	N/A
Emma Heath	performed research	N/A
Ian Smyth	performed research, contributed reagents/analytic tools, analysed data	N/A
Fiona M. Watt	contributed reagents/analytic tools, analysed data	N/A
Peter K. Sorger	contributed reagents/analytic tools, analysed data, wrote paper	N/A
Allan Bradley	contributed reagents/analytic tools, analysed data, wrote paper	N/A

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work\*.

**Candidate's Signature**  **Date** 6/6/2014

**Main Supervisor's Signature**  **Date**

\*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

# Spindle checkpoint deficiency is tolerated by murine epidermal cells but not hair follicle stem cells

Floris Fojer<sup>a,b,1,2</sup>, Tia DiTommaso<sup>c</sup>, Giacomo Donati<sup>d</sup>, Katta Hautaviita<sup>a</sup>, Stephanie Z. Xie<sup>b,3</sup>, Emma Heath<sup>d</sup>, Ian Smyth<sup>c,e</sup>, Fiona M. Watt<sup>d</sup>, Peter K. Sorger<sup>b,4</sup>, and Allan Bradley<sup>a,4</sup>

<sup>a</sup>Mouse Genomics, Wellcome Trust Sanger Institute, Hinxton CB10 1SA, United Kingdom; <sup>b</sup>Center for Cell Decision Processes, Department of Systems Biology, Harvard Medical School, Boston, MA 02115; Departments of <sup>c</sup>Biochemistry and Molecular Biology and <sup>e</sup>Anatomy and Developmental Biology, Monash University, Clayton, VIC 3800, Australia; and <sup>d</sup>Centre for Stem Cells and Regenerative Medicine King's College London, Guy's Hospital, London SE1 9RT, United Kingdom

Edited by Douglas Koshland, University of California, Berkeley, CA, and approved January 11, 2013 (received for review October 9, 2012)

The spindle assembly checkpoint (SAC) ensures correct chromosome segregation during mitosis by preventing aneuploidy, an event that is detrimental to the fitness and survival of normal cells but oncogenic in tumor cells. Deletion of SAC genes is incompatible with early mouse development, and RNAi-mediated depletion of SAC components in cultured cells results in rapid death. Here we describe the use of a conditional KO of mouse *Mad2*, an essential component of the SAC signaling cascade, as a means to selectively induce chromosome instability and aneuploidy in the epidermis of the skin. We observe that SAC inactivation is tolerated by interfollicular epidermal cells but results in depletion of hair follicle bulge stem cells. Eventually, a histologically normal epidermis develops within ~1 mo after birth, albeit without any hair. *Mad2*-deficient cells in this epidermis exhibited abnormal transcription of metabolic genes, consistent with aneuploid cell state. Hair follicle bulge stem cells were completely absent, despite the continued presence of rudimentary hair follicles. These data demonstrate that different cell lineages within a single tissue respond differently to chromosome instability: some proliferating cell lineages can survive, but stem cells are highly sensitive.

mouse models | consequences of aneuploidy | whole chromosome instability | epidermal stem cell biology

At each cell division, newly replicated chromosomes are evenly distributed between daughter cells. If this process goes awry, daughters inherit an aneuploid chromosome content. Aneuploidy causes early developmental abnormalities in mice and is a leading cause of mental retardation and miscarriage in humans (1, 2). At the cellular level, imbalance in chromosome numbers disrupts metabolic homeostasis, decreasing the rate of proliferation (3). The mitotic spindle assembly checkpoint (SAC) is a key component of the machinery that guards against aneuploidy. The SAC functions by ensuring that all kinetochores are properly attached to spindle microtubules before the onset of anaphase (4). *Mad2* is a key signal transducer in the SAC pathway, shuttling between unattached kinetochores and the anaphase promoting complex. Germ-line deletion of *Mad2*, or of any other SAC gene, results in massive aneuploidy and early embryonic lethality. In cell lines, depletion of SAC components causes apoptosis (5). However, the vast majority of human cancers exhibit extensive aneuploidy, suggesting that aneuploidy can be tolerated by tumor cells in vivo (5–9). In this paper, we investigate the in vivo consequences of chromosomal instability by conditionally abrogating the SAC, using the epidermis as a model.

The mammalian skin is a highly organized and mitotic tissue consisting of layers of keratinocytes [interfollicular epidermis (IFE)] interspersed with hair follicles. In the IFE, keratinocytes divide in the basal layer and subsequently differentiate into stratified cells while migrating toward the surface of the skin. Hair follicles harbor several stem cell populations essential for hair formation and capable of repopulating the IFE after wounding (10). In

normal conditions, homeostasis within the skin is maintained by progenitor cells that reside in the basal layer of the IFE (11–13).

## Results

To study the consequences of induced aneuploidy in the skin, we generated a conditional allele of *Mad2* in which *loxP* sites flank exons 2 and 5, enabling Cre-mediated deletion of virtually the entire *Mad2* coding region. The structure of the deleted locus is identical to that of the conventional *Mad2* deletion we have previously shown to be embryonic lethal and eliminates detectable *Mad2* mRNA and protein (14). Mice carrying the *loxP-Mad2* allele (*Mad2*<sup>fl</sup>) were intercrossed with *Keratin14-Cre* (*Krt14-Cre*) transgenic mice (15) to generate animals in which *Mad2* was selectively deleted from multiple epidermal lineages. *Mad2*<sup>fl</sup>, *Krt14-Cre*<sup>+</sup> mice were born at normal Mendelian ratios, but unlike WT littermates, which acquired hair within 3–4 d after birth, mutant mice remained completely devoid of hair (Fig. 1A; Fig. S1A). Other than hair loss, the epidermis of mutant mice appeared largely normal despite the complete absence of *Mad2* RNA or protein in the epidermis (Fig. 1B and C). However, 80% of *Mad2*<sup>fl</sup>, *Krt14-Cre*<sup>+</sup> neonates died within the first month of life because of malnutrition that apparently arose from a combination of poor mothering and competition with control littermates for food. The remaining 20% of mice survived into adulthood (Fig. 1D).

To determine whether *Mad2* deficiency indeed resulted in aneuploidy, we harvested back skin epidermis from neonates (1 and 3 wk of age) and adults, generated single cell suspensions, and measured DNA content by flow cytometry. Whereas epidermal samples from control littermates exhibited a narrow G1 peak indicative of euploid DNA content, *Mad2*-deficient cells consistently exhibited a wider distribution with a significant increase in the number of cells with a 4n DNA content. This phenotype was most prominent at 3 wk of age and is suggestive of aneuploid cell division (Fig. 2A). To confirm this hypothesis we assayed chromosome number in the mutant epidermis using interphase FISH

Author contributions: F.F. designed research; F.F., T.D., G.D., K.H., and E.H. performed research; T.D., S.Z.X., and F.M.W. contributed new reagents/analytic tools; F.F., I.S., F.M.W., P.K.S., and A.B. analyzed data; and F.F., P.K.S., and A.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, [www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) (accession no. GSE42698).

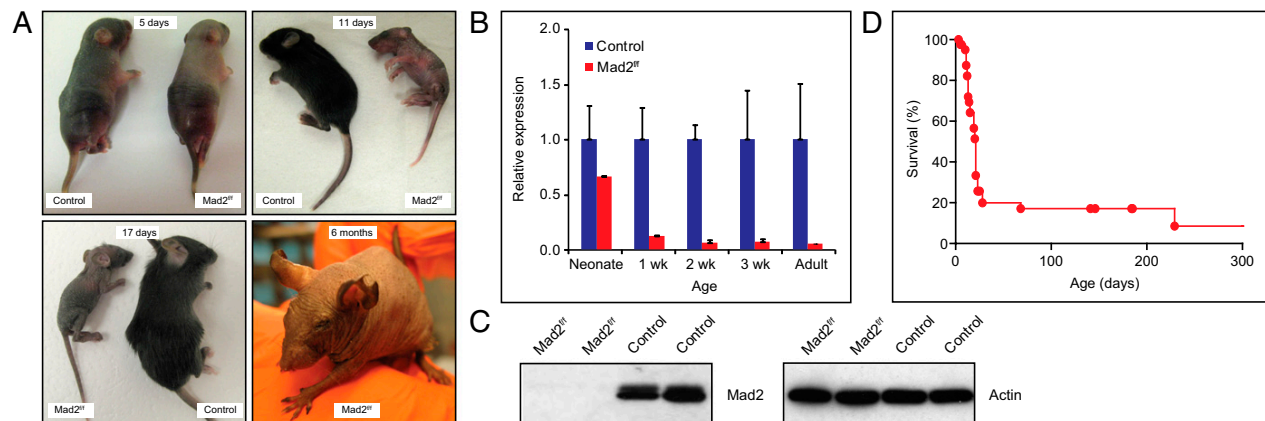
<sup>1</sup>Present address: European Institute for the Biology of Ageing (ERIBA), University Medical Center Groningen, University of Groningen, AV 9713, Groningen, The Netherlands.

<sup>2</sup>To whom correspondence should be addressed. E-mail: f.fojer@umcg.nl.

<sup>3</sup>Present address: Campbell Family Cancer Research Institute, Ontario Cancer Institute, Princess Margaret Hospital, University Health Network, Toronto, Ontario, Canada M5G2C1.

<sup>4</sup>P.K.S. and A.B. contributed equally to this work.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1217388110/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1217388110/-DCSupplemental).

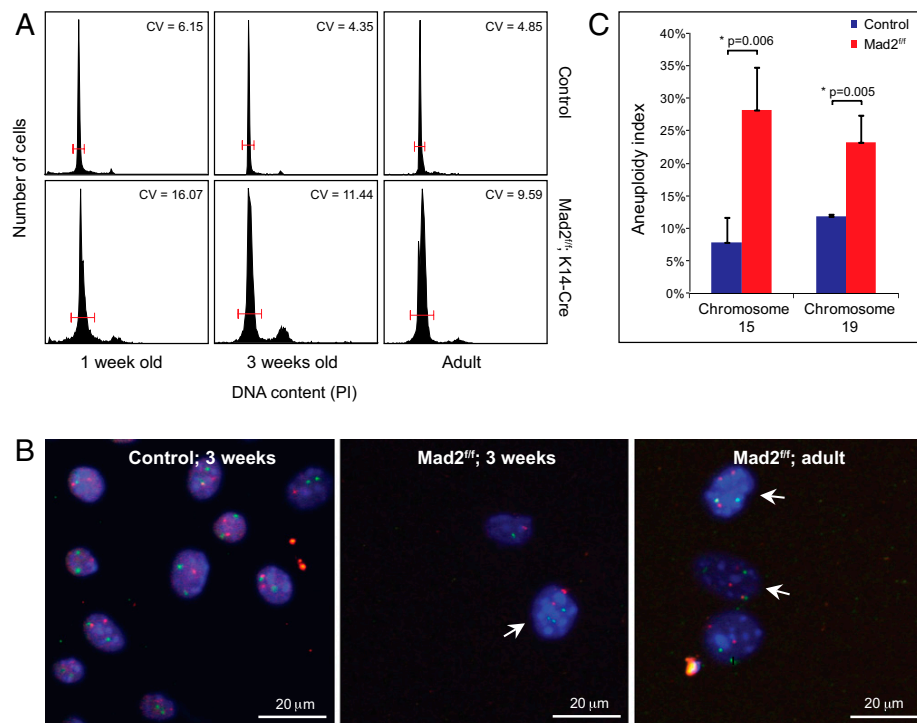


**Fig. 1.** *Mad2* deficiency provokes hair loss in mouse epidermis. (A) Macroscopic phenotype of *Mad2* conditional KO (cKO) mice in neonates and adults. (B) Quantitative PCR comparing *Mad2* RNA levels in *Mad2*<sup>fl/fl</sup>, *Krt14-Cre*<sup>+</sup> (*Mad2*-deficient), and control epidermis in neonates and adults. Error bars indicate SE of the mean (SEM) of at least two biological replicates. (C) Western blot for *Mad2* in *Mad2*-deficient and control epidermis. (D) Kaplan-Meier curve depicting survival of *Mad2*-deficient mice within the first year.

against chromosomes 4, 15, and 19 (Fig. 2*B* and *C*; Fig. S1*B*; Table S1). We observed a two- to threefold increase in the number of cells trisomic for the assayed chromosomes compared with WT controls, confirming widespread aneuploidy in *Mad2*-deficient epidermis. We measured up to 10% of the WT IFE cells to be aneuploid, which is similar to aneuploidy rates observed in WT mouse embryonic fibroblasts (MEFs) but higher than aneuploidy rates in for instance hematopoietic cells (6–8, 16), suggesting that IFE cells tolerate aneuploidy better than some other cell types.

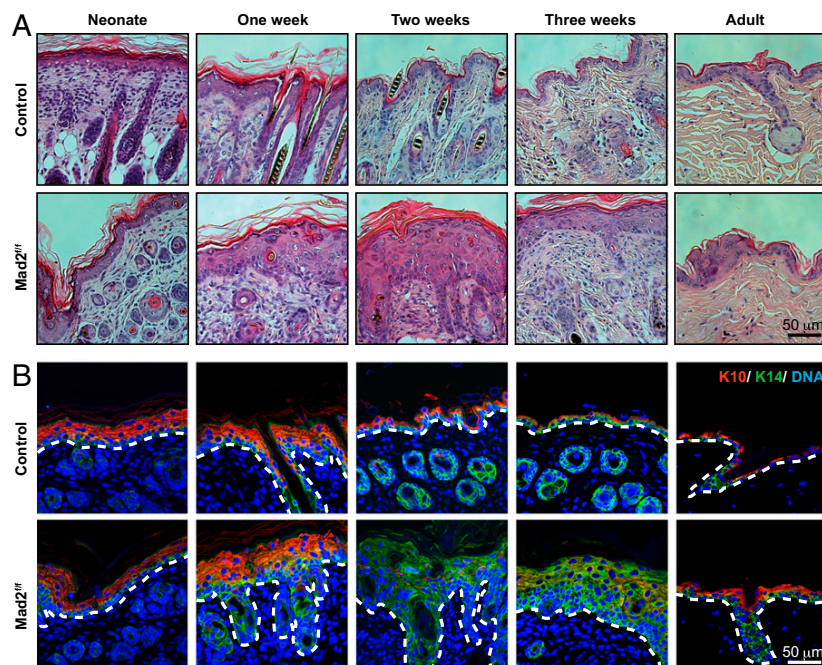
Histological examination of *Mad2*-null epidermis revealed hyperkeratosis during the first 3 wk of life, followed by the acquisition of a relatively normal morphology at later stages (Fig. 3*A*). Consistent with hyperkeratosis, we observed that some keratinocytes

in the *Mad2*-null epidermis in neonates coexpressed the basal marker *Krt14* along with the differentiation marker *Krt10*. Expression of these markers is normally mutually exclusive, with proliferating basal cells exhibiting *Krt14* staining (Fig. 3*B*), but coexpression can occur in hyperproliferative epidermis (17). Hair follicles appeared abnormal at all ages, explaining the persistent absence of hair. In agreement with the increased cellularity of the epidermis, we observed an increase in proliferation in the basal layers of the IFE and in hair follicles (Fig. S1*C*). However, both the pattern of *Krt10/Krt14* and levels of Ki67 staining returned to normal in the subset of *Mad2*<sup>fl/fl</sup>, *Krt14-Cre*<sup>+</sup> animals that survived until adulthood.



**Fig. 2.** *Krt14-Cre; Mad2* conditional mice display dramatic aneuploidy in the epidermis. (A) DNA content distribution in control and *Mad2*-deficient epidermal cells in neonates and adults. The CV is used as a measure of the variance of DNA content within the G1 peak. (B) Interphase FISH for chromosomes 15 (red) and 19 (green) on freshly isolated control and *Mad2*-deficient keratinocytes. (C) Percentage of cells showing more than two copies of chromosome 15 or 19 assessed by interphase FISH (aneuploidy index) in *Mad2*-deficient or control epidermis. Error bars indicate SE of the mean (SEM) values for at least three biological replicates. Individual values can be found in Table S1.





**Fig. 3.** Mad2 deficiency provokes a wave of hyperkeratosis in neonates, which largely normalizes in adults. (A) H&E staining of epidermal samples at various time points showing epidermal morphology with hyperkeratotic features in neonates. (B) Hyperkeratotic epidermis shows signs of increased differentiation (evidenced by thickened Krt10 positive epidermis).

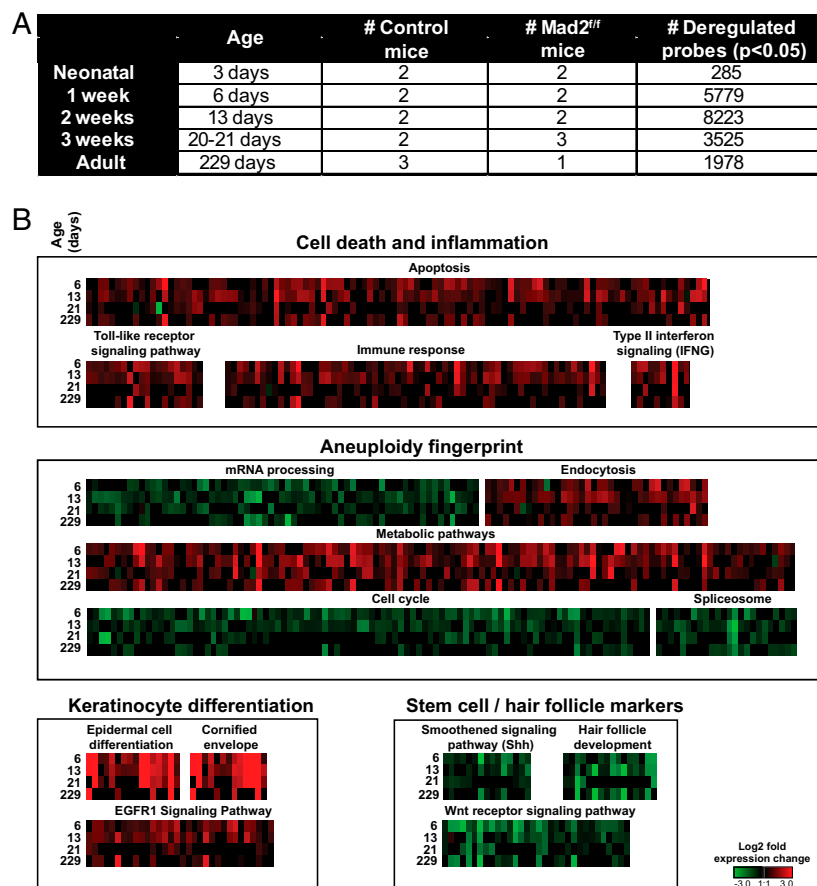
Following a wave of cell proliferation at postnatal week 2, we observed increasing levels of apoptosis in hair follicles and the IFE; this continued through postnatal week 3 (Fig. S1D). We also observed infiltration of immune cells into the underlying dermis, suggesting inflammation. Such inflammation can arise if the barrier function of the skin is compromised, thereby allowing microorganisms to penetrate into the dermis. However, when we tested epidermal barrier integrity in mutant mice, we observed that the skin barrier was intact (Fig. S2), implying that immune infiltration was a consequence of abnormal mitosis and cell death. We conclude that Mad2 deficiency provokes aneuploidy and apoptosis in IFE and hair follicle cells and that this promotes a temporary inflammatory reaction in the dermis, explaining the hyperkeratosis.

To analyze the responses of epidermal cells to Mad2 deletion, we isolated mRNA from samples of epidermis [including both hair follicle (HF) and IFE] of newborns at 3, 6, 13, and 21 d after birth and of adults and then compared the transcript profiles of Mad2-deficient skin with *Cre<sup>-</sup> Mad2*-positive littermates using Illumina microarrays (Fig. 4A). We next extracted all probes that were significantly deregulated ( $P < 0.05$ ) at two or more time points. To assess the quality of the microarray data, we also validated a selection of outlier genes (Fig. S3A) by qPCR, in all cases confirming our microarray results (Fig. S3B). Data from Mad2-null 3-d-old neonates were excluded from the analysis because transcript profiles were very similar to aged-matched controls, presumably because Mad2 transcripts had not completely disappeared yet [Fig. 1B, postnatal day (P)3]. The analysis of days 6–21 and adult mice generated a set of 1,619 down-regulated and 1,347 up-regulated genes (Dataset S1), which we processed to recover enriched biological pathways based on Wiki pathways and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (using Bonferroni-corrected  $P < 0.05$ ) and gene ontologies using Webgestalt (18) (Fig. 4B; Table S2; Datasets S2 and S3). This analysis revealed significant up-regulation of genes involved in apoptosis and inflammation (most clearly at days 6 and 13) and the appearance of an aneuploidy fingerprint involving enrichment for genes and pathways known to be altered by aneuploidy. This fingerprint was first described in yeast strains and MEFs carrying

supernumerary chromosomes and involves a substantial deregulation of genes that play a role in basic metabolic processes (19, 20). We now show that this response also occurs in vivo, because our dataset showed enrichment for similar pathways in Mad2-deficient skin: up-regulation of metabolic pathways and endocytosis and down-regulation of cell cycle genes, mRNA processing genes, and genes encoding spliceosome components. The latter pathways were recently associated with the response to aneuploidy in other organisms as well (plants, yeast, and humans) (21).

We also observed significant up-regulation of genes involved in epidermal differentiation and epidermal growth factor receptor (EGFR) signaling, and down-regulation of genes involved in Sonic Hedgehog (Shh), Wnt signaling, and hair follicle development. This down-regulation suggests that in skin, aneuploidy leads to depletion of the stem cell compartment. To test this hypothesis, we quantified the levels of stem cell lineage markers *Krt15*, *Lgr5*, *Lrig1*, and *Sox9* by qPCR. We observed substantial reductions in the levels of all four genes, showing that HF stem cells that normally reside in the bulge compartment of the hair follicle are depleted (Fig. 5A). We also labeled whole-mount tail epidermal samples with antibodies against Krt14 and  $\alpha$ -6-integrin, which label the IFE, and Krt15 and CD34, which label bulge HF stem cells. This confirmed depletion of the bulge stem cell compartment as a result of Mad2 depletion, whereas the IFE cells survived abrogation of the SAC. We conclude that HF stem cells are specifically depleted in the skin of Mad2-null animals (Fig. 5B; Fig. S3C). The differential response between these two cell lineages is not the result of differential Mad2 expression, because FACS-sorted WT HF stem cells and IFE cells express similar levels of Mad2 as determined by qPCR (Fig. S4A and B). We conclude that stem cells cannot survive Mad2 loss, whereas IFE cells can.

To test whether we could propagate Mad2 deficient keratinocytes in vitro, we transferred primary keratinocytes (not enriched for any epidermal cell lineage) from WT and Mad2-deficient back skin into tissue culture (22) and propagated these cells for 3 wk. Whereas control cultures contained many colonies that grew to confluence, no viable colonies were present in mutant cultures (Fig. 5C) despite having plated similar numbers of viable cells



**Fig. 4.** *Mad2* abrogation provokes a general aneuploidy response, keratinocyte differentiation, and depletion of HF stem cell markers. (A) Overview of microarray study with number of significantly deregulated probes and number of samples per time point. (B) Schematic overview of deregulated pathways grouped into four main responses: aneuploidy inflammation, aneuploidy fingerprint, keratinocyte differentiation, and stem cell and hair follicle markers.

from mutant and control animals (viability was assessed by trypan blue staining; Fig. S4 C and D). These findings are in line with earlier findings that *Mad2* deficiency provokes rapid cell death in vitro (5) and, even more important, emphasize that the in vivo response to chromosomal instability is fundamentally different from the in vitro response.

## Discussion

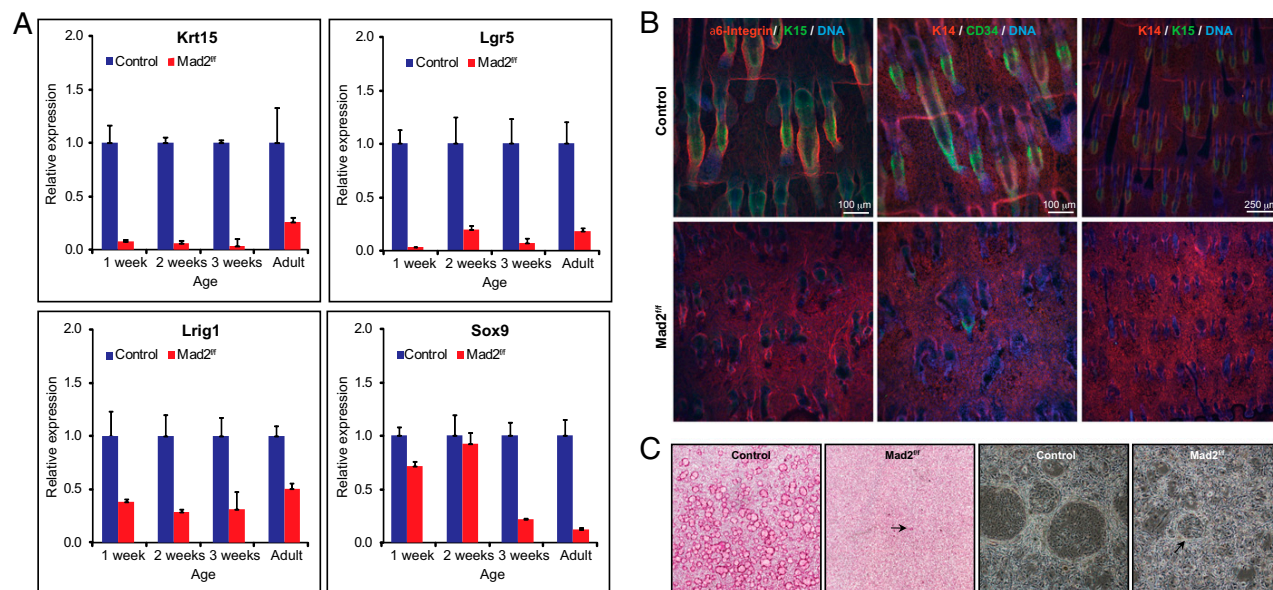
Chromosomal instability (CIN) and the resulting aneuploidy are hallmarks of cancer cells, yet they are detrimental to untransformed cells in vitro, where they result in decreased proliferation capacity, disrupted cell physiology, and ultimately when CIN is very severe, cell death (5, 19, 20, 23). In this study, we report the consequences of SAC abrogation and the resulting aneuploidy in vivo. To circumvent the embryonic lethality associated with SAC inactivation through the germ line (6–8), we created a conditional KO allele for *Mad2*, a key player in the SAC and used this allele to abrogate the SAC in the epidermis. Remarkably, we found that *Mad2* deletion is surprisingly well tolerated by IFE cells, resulting in mice with a functional *Mad2*-deficient epidermis, albeit without hair. The ability of IFE cells to tolerate *Mad2* loss is not a reflection of their limited proliferative potential and reduced requirement for mitosis, because basal cells that contribute to the IFE are at least equally proliferative in *Mad2* KO animals compared with controls, and they can generate all of the structures in the adult skin with the exception of functional hair follicles. Moreover, the majority of IFE cells divide asymmetrically, producing one proliferating and one differentiated cell (24), further arguing that a subpopulation of IFE cells is highly proliferative. Last, we show that *Mad2*-deficient IFE cells are highly aneuploid,

arguing that they must have undergone several rounds of cell division in the absence of a functional SAC.

However, cell survival does not simply imply that aneuploidy is inconsequential for IFE cells: we observe significant changes in gene transcription including up-regulation of metabolic pathways and down-regulation of mRNA processing, changes that have previously been proposed to be part of an aneuploidy signature in yeast and cultured murine cells. Presumably, this aneuploidy fingerprint is caused by a stress response induced by the burden of extra transcripts and proteins from the supernumerary chromosomes (19, 20), and our data provide evidence that this response also occurs in vivo.

However, although SAC deficiency is well tolerated in IFE cells, we found that it is incompatible with survival of HF stem cells, resulting in the complete absence of bulge stem cells in *Mad2*-deficient hair follicles. This observation suggests that different cell lineages exhibit different responses to aneuploidy in vivo. Indeed, there is increasing evidence that some somatic cell lineages can tolerate high levels of aneuploidy. For instance, in the healthy brain, 1 in 10 neurons is reported to be aneuploid, with increasing rates upon aging and in pathologies such as Alzheimer's disease (25–27). Furthermore, we observe up to 10% of aneuploid cells in normal mouse epidermis. On the other hand, other cell lineages (e.g., lymphocytes) show very little aneuploidy in vivo (6–8, 28).

What can explain this differential response toward *Mad2* depletion? First, different cell lineages might exhibit different sensitivities toward apoptosis, with stem cells being more sensitive to apoptosis than more differentiated dividing cells. Alternatively, the signaling pathways that transduce the aneuploidy-induced stress response are more active in stem cells, thus triggering a



**Fig. 5.** Aneploidy provokes bulge HF stem cell depletion leaving IFE cells unaffected. (A) Quantitative PCR reveals a dramatic decrease of established HF stem cell markers *Krt15*, *Lgr5*, *Lrig1*, and *Sox9*. (B) Whole mount tail epidermis labeled for  $\alpha$ -6 integrin or Krt14 (labeling IFE in red) and Krt15 or CD43 (both labeling bulge HF stem cells in green) showing HF stem cell depletion in vivo at P25. (C) *Mad2*-deficient primary keratinocytes do not form colonies in vitro. (Left) 1% Rhodamine-B-stained cultures, showing bright red colonies in control cultures. (Right) Phase contrast images, showing epidermal colonies in control cultures, but not in *Mad2*-deficient cultures. Note that cells were cultured on a MEF feeder layer. Arrows indicate one microcolony that emerged in a *Mad2*-deficient keratinocyte culture.

stronger stress response and thereby driving them into apoptosis. A third possible explanation is that IFE cells spend more time in prometaphase than HF stem cells, providing them with more time to biorient their chromosomes properly and making them less dependent on a functional SAC. In this scenario, SAC inactivation would provoke more dramatic aneploidy in HF stem cells than in IFE cells, specifically driving HF stem cells to cell death.

We speculate that aneploidy specifically induces apoptosis in stem cells through a specific unknown mechanism that is absent, or less active, in other proliferating cells in the epidermis. This stem cell apoptosis triggers the marked inflammatory response observed in the skin of *Mad2* KO mice, possibly as a consequence of the accumulation of cell debris. The inflammatory response then causes a wave of proliferation and apoptosis in neonates, explaining the hyperkeratosis and aberrant K14 expression (17). The inflammatory response dampens over time, ultimately allowing the appearance of a relatively normal adult epidermis. Although we were able to carefully map the response to aneploidy in IFE cells, our current model precluded us from performing a similar analysis for HF stem cells, as all bulge stem cells had disappeared before we could isolate sufficient numbers. Therefore, we are now developing a unique mouse model in which *Mad2* inactivation can be triggered in adults, which will allow us to isolate bulge stem cell and IFE cell populations directly following *Mad2* inactivation, enabling us to identify the mechanism underlying the differential responses to *Mad2* deletion-induced aneploidy.

Altogether, we show that full SAC abrogation can be tolerated in vivo, despite a robust change in cell metabolism and despite the resulting aneploidy. Up to two-thirds of all human cancers are aneuploid (29). It is likely that aneuploid cancer cells use similar pathways as IFE cells to cope with this aneploidy. Our findings are a first step toward identifying such pathways by identifying the skin as a tissue that can cope with remarkable levels of aneploidy and by carefully mapping the responses to aneploidy in this tissue. Importantly, we found that the surviving IFE cells exhibit an aneploidy-induced stress response in vivo, which

might have a large impact on the future treatment of aneuploid cancers, because this stress response could be exploited to specifically target aneuploid tumor cells (30).

## Materials and Methods

**Animals and Keratinocyte Culture.** Animal protocols were approved by the Home Office (UK) and specified in the Home Office Project Licence. *Mad2* conditional mice were generated by introducing LoxP sites surrounding exons 1 and 5, which when combined with Cre recombinase creates the same full KO as described previously (14). *Krt14-Cre* mice are described elsewhere (15). For genotyping, mice were ear clipped, and DNA was isolated using ear lysis buffer (Viagen). PCR primers are provided in Table S3, and primers were optimized for annealing temperatures at 60 °C. At selected time points, mice were killed, and back skin samples and tail samples were harvested for paraffin embedding or further processing. For further processing, subcutaneous fat was removed from the dermal side, and skin samples were incubated floating with the dermal side on Trypsin overnight at 4 °C. Epidermis was separated from the dermis using a scalpel, and epidermal samples were then prepared for interphase FISH, mRNA isolation, flow cytometry, or keratinocyte culture. In case of tail samples, skin was stripped from the bone and incubated in 5 mM EDTA for 1 h at 37 °C. The epidermis was then peeled from the dermis and fixed in 2% (vol/vol) paraformaldehyde for 1 h at room temperature. Keratinocytes for in vitro culture were harvested under sterile conditions and processed as described (22). Cell numbers were measured using an automated cell counter (Invitrogen) and viability by determining Trypan blue penetration.

**Interphase FISH.** For interphase FISH, epidermal samples were minced and filtered through a 40- $\mu$ m filter (Falcon) to obtain single cells. Cells were fixed in methanol/glacial acid after swelling in 0.075 M KCl for 30 min at 37 °C. Cells were dropped on slides, dehydrated, and treated with 0.1 mg/mL RNaseA (Sigma) in 2 $\times$  SSC for 30 min at 37 °C, followed by 0.01% Pepsin (Sigma) treatment in 0.01 M HCl for 3 min at room temperature. Slides were then baked at 65 °C for 1 h and denatured at 63 °C for 80 s in 70% (vol/vol) formamide in 0.6 $\times$  SSC. For probe preparation, the following bacterial artificial chromosomes (BACs) were used: RP23-341B14, RP23-32C12, RP23-463N9, RP23-24E11, RP23-67O11, and RP23-69G24 (recognizing chromosome 4); RP24-84H16, RP24-160M22, and RP24-160H6 (recognizing chromosome 15); and RP24-144K17, RP23-183J2, and RP23-292H23 (recognizing



chromosome 19). Probes were labeled with green dUTP (Abott) or Texas red chromotide (Invitrogen) and hybridized at 37 °C overnight in hybridization buffer [62.5% formamide (Sigma), 2× SSC, 10% (wt/vol) dextran sulfate (Sigma), 0.5 M phosphate buffer, pH 7.4] containing 1  $\mu$ L mouse Cot 1 DNA (Invitrogen) and 1.5  $\mu$ L of each probe. Slides were mounted in the presence of DAPI-containing mounting media (Vectashield) following washing steps in 50% (vol/vol) formamide and 1× SSC and analyzed on a IX81 Olympus widefield microscope. Chromosome numbers (events per cell) were counted on an IX81 Olympus microscope using Excellence Software and scored aneuploid in the case of three or more events per cell by hand. Statistical analysis was performed in Microsoft Excel to calculate the SD of the mean values, and GraphPad Prism was used to calculate *P* values (*t* test).

**Barrier Assay.** For barrier assays, neonates were killed and fixed in increasing series of ice-cold methanol in water. Next, carcasses were incubated in 0.1% toluidine blue (Sigma), and excess dye was washed away by three PBS washes.

**Immunohistochemistry, Flow Cytometry, and Antibodies.** For DNA content analysis, freshly isolated epidermal cells were fixed in 70% (vol/vol) ethanol overnight and stained in 10  $\mu$ g/mL propidium iodide (Invitrogen) in PBS in the presence of 0.1 mg/mL RNaseA (Sigma). For IFE and HF stem cells sorts, fresh keratinocytes were isolated from back skin and labeled with antibodies following sorting. Western blots were performed using standard protocols. For immunostaining of paraffin sections, slides were dewaxed, and epitopes were retrieved by boiling samples for 10 min in 10 mM sodium citrate. Samples were blocked using mouse serum (Vector Laboratories) and labeled overnight with primary antibodies in the presence of 1% BSA (Sigma), 10% (vol/vol) FCS (Invitrogen), and 0.01% TritonX100 (Sigma). Samples were then washed and incubated with secondary antibodies and mounted in the presence of DAPI-containing mounting media (Vectashield; Vector Laboratories) following confocal microscopy analysis (Zeiss). For whole mounts, epidermal samples were blocked [0.5% skimmed milk powder, 0.25% fish skin gelatin (Sigma), and 0.5% TritonX100 (Sigma) in 20 mM Hepes and 0.9% NaCl] for 1 h following overnight incubation at 4 °C and incubation with secondary antibodies. Samples were mounted in the presence of DAPI and analyzed by confocal microscopy (Leica). The primary antibodies used were Mad2 (BD), Actin (Cell Signaling), Krt14 (clone LL002; Abcam), Krt10 (Covance), Ki67 (Dako), CD34 (Ebioscience),  $\alpha$ -6-integrin (Abcam), and Krt15 (clone LHK15). Secondary antibodies were Alexa 555- and 488-labeled anti-mouse

and anti-rabbit antibodies (Invitrogen) and HRP-labeled goat anti-mouse (Cell Signaling). TUNEL assays on paraffin sections were performed using a Promega kit following the manufacturer's protocol.

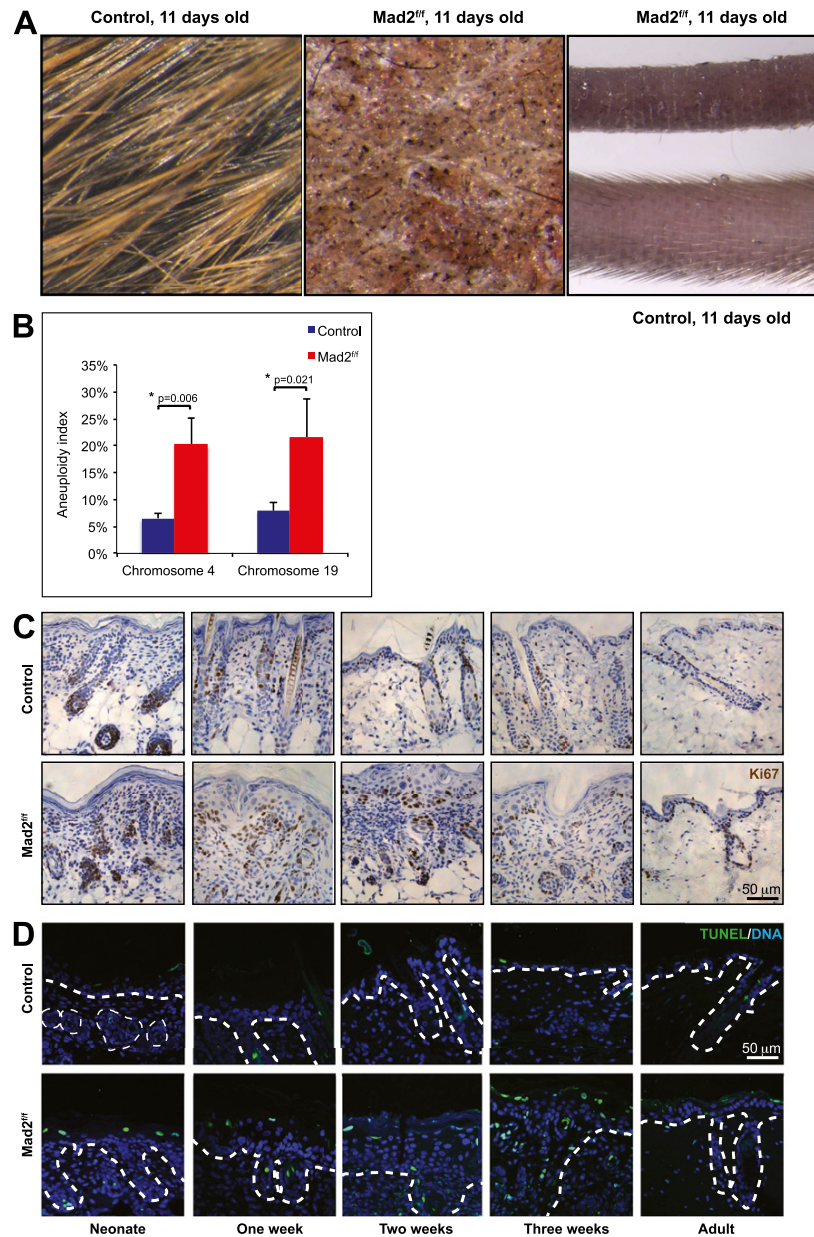
**qPCRs and Microarrays.** RNA was isolated using the RNeasy kit (Qiagen). Biotin-labeled cRNA was synthesized using the Illumina totalprep RNA amplification kit (Ambion). Biotinylated cRNA was hybridized to Illumina Sentrix Bead Chips and labeled with streptavidin-conjugated Cy3 (Amersham) according to the manufacturer's protocol. Data were quantile normalized (31) following scanning and analyzed using Bioconductor lumi (32) and limma packages (33). Data were *P* value adjusted to yield a sorted list of differentially expressed genes (34) and sorted in chromosomal position or frequency of deregulation in Microsoft Excel. Enrichment analyses were performed using Webgestalt software (18) using Bonferroni correction for multiple testing. For diagrams in Fig. 4B, genes in deregulated pathways were downloaded using Webgestalt, and matching expression Log2 ratios were extracted from Dataset S1 using Excel. Extracted values (values per category shown in Dataset S3) were then imported into Genesis software (35) to generate expression diagrams. All raw and normalized Illumina microarray data can be found at the NCBI Gene Expression Omnibus (GEO) under accession no. GSE42698. For qPCR reactions, cDNA was generated from 1  $\mu$ g of total RNA (Superscript II; Invitrogen) and used as a template for qPCR (ABI PRISM 7700 Sequence Detector) in the presence of SYBR green (Invitrogen) to label the products. Actin was used for normalization. The average expression values and SD of the mean were calculated as indicated in the figure legends and compared with the expression values in control mice (normalized to a value of 1). Primer sequences are provided in Table S3.

**ACKNOWLEDGMENTS.** We thank N. Park, R. Bautista, and R. Andrews for help with microarray processing and bioinformatical analysis; E. Langley, R. Banerjee, W. Benz, Y. Zhu, C. Rocchi, P. Bakker, and F. Yang for help with interphase FISH; R. Bronson for the initial pathology; the animal technicians for help with animal husbandry; and S. Bruggeman for critically reading the manuscript. This work was supported by the Dutch Cancer Society, the European Molecular Biology Organization (fellowships to F.F.), Australian Research Council and National Health and Medical Research Council fellowships (to I.S.), National Institutes of Health Grants CA084179 and CA139980 (to P.K.S.), and The Wellcome Trust (A.B.).

- Hassold T, et al. (1996) Human aneuploidy: Incidence, origin, and etiology. *Environ Mol Mutagen* 28(3):167–175.
- Brown S (2008) Miscarriage and its associations. *Semin Reprod Med* 26(5):391–400.
- Sheltzer JM, Amon A (2011) The aneuploidy paradox: Costs and benefits of an incorrect karyotype. *Trends Genet* 27(11):446–453.
- Musacchio A, Salmon ED (2007) The spindle-assembly checkpoint in space and time. *Nat Rev Mol Cell Biol* 8(5):379–393.
- Kops GJ, Foltz DR, Cleveland DW (2004) Lethality to human cancer cells through massive chromosome loss by inhibition of the mitotic checkpoint. *Proc Natl Acad Sci USA* 101(23):8699–8704.
- Folger F, Draviam VM, Sorger PK (2008) Studying chromosome instability in the mouse. *Biochim Biophys Acta* 1786(1):73–82.
- Schwartzman JM, Sotillo R, Benezra R (2010) Mitotic chromosomal instability and cancer: Mouse modelling of the human disease. *Nat Rev Cancer* 10(2):102–115.
- Holland AJ, Cleveland DW (2009) Boveri revisited: Chromosomal instability, aneuploidy and tumorigenesis. *Nat Rev Mol Cell Biol* 10(7):478–487.
- Baker DJ, Chen J, van Deursen JM (2005) The mitotic checkpoint in cancer and aging: What have mice taught us? *Curr Opin Cell Biol* 17(6):583–589.
- Watt FM, Jensen KB (2009) Epidermal stem cell diversity and quiescence. *EMBO Molec Med* 1(5):260–267.
- Pincelli C, Marconi A (2010) Keratinocyte stem cells: Friends and foes. *J Cell Physiol* 225(2):310–315.
- Jones PH, Simons BD, Watt FM (2007) Sic transit gloria: Farewell to the epidermal transit amplifying cell? *Cell Stem Cell* 1(4):371–381.
- Blanpain C, Fuchs E (2009) Epidermal homeostasis: A balancing act of stem cells in the skin. *Nat Rev Mol Cell Biol* 10(3):207–217.
- Dobles M, Liberal V, Scott ML, Benezra R, Sorger PK (2000) Chromosome mis-segregation and apoptosis in mice lacking the mitotic checkpoint protein Mad2. *Cell* 101(6):635–645.
- Dassule HR, Lewis P, Bei M, Maas R, McMahon AP (2000) Sonic hedgehog regulates growth and morphogenesis of the tooth. *Development* 127(22):4775–4785.
- Weaver BA, Cleveland DW (2006) Does aneuploidy cause cancer? *Curr Opin Cell Biol* 18(6):658–667.
- Gareus R, et al. (2007) Normal epidermal differentiation but impaired skin-barrier formation upon keratinocyte-restricted IKK1 ablation. *Nat Cell Biol* 9(4):461–469.
- Zhang B, Kirov S, Snoddy J (2005) WebGestalt: An integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 33(Web Server issue):W741–W748.
- Williams BR, et al. (2008) Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science* 322(5902):703–709.
- Torres EM, et al. (2007) Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science* 317(5840):916–924.
- Sheltzer JM, Torres EM, Dunham MJ, Amon A (2012) Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci USA* 109(31):12644–12649.
- Jensen KB, Driskell RR, Watt FM (2010) Assaying proliferation and differentiation capacity of stem cells using disaggregated adult mouse epidermis. *Nat Protoc* 5(5):898–911.
- Hanahan D, Weinberg RA (2011) Hallmarks of cancer: The next generation. *Cell* 144(5):646–674.
- Clayton E, et al. (2007) A single type of progenitor cell maintains normal epidermis. *Nature* 446(7132):185–189.
- Yurov YB, et al. (2007) Aneuploidy and confined chromosomal mosaicism in the developing human brain. *PLoS ONE* 2(6):e558.
- Iourov IY, Vorsanova SG, Liehr T, Yurov YB (2009) Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: Differential expression and pathological meaning. *Neurobiol Dis* 34(2):212–220.
- Thomas P, Fenech M (2008) Chromosome 17 and 21 aneuploidy in buccal cells is increased with ageing and in Alzheimer's disease. *Mutagenesis* 23(1):57–65.
- Peterson SE, et al. (2011) Normal human pluripotent stem cell lines exhibit pervasive mosaic aneuploidy. *PLoS ONE* 6(8):e23018.
- Duijff PH, Schultz N, Benezra R (2012) Cancer cells preferentially lose small chromosomes. *Int J Cancer*, 10.1002/ijc.27924.
- Tang YC, Williams BR, Siegel JJ, Amon A (2011) Identification of aneuploidy-selective antiproliferation compounds. *Cell* 144(4):499–512.
- Yang YH, et al. (2002) Normalization for cDNA microarray data: A robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30(4):e15.
- Du P, Kibbe WA, Lin SM (2008) lumi: A pipeline for processing Illumina microarray. *Bioinformatics* 24(13):1547–1548.
- Smyth GK (2005) *Limma: Linear Models for Microarray Data* (Springer, New York), pp 397–420.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57(1):289–300.
- Sturn A, Quackenbush J, Trajanoski Z (2002) Genesis: Cluster analysis of microarray data. *Bioinformatics* 18(1):207–208.

# Supporting Information

Foijer et al. 10.1073/pnas.1217388110

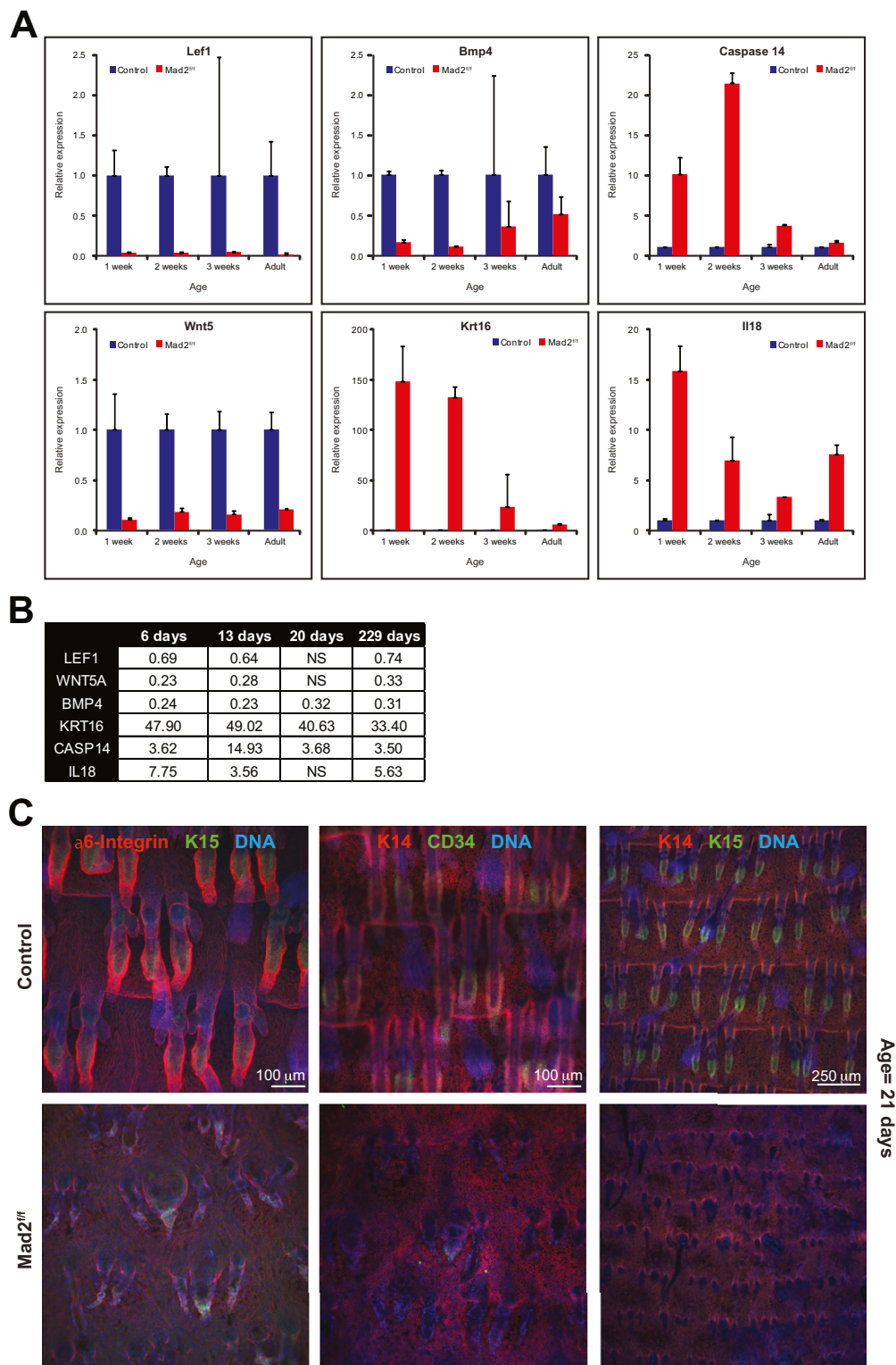


**Fig. S1.** Mad2 deficiency interferes with hair development and provokes aneuploidy in the epidermis. (A) Effect of Mad2 deletion on back skin (Center) and on the tail (Right) compared with control skin (Left) and tail (Right). (B) Aneuploidy rates in Mad2-deficient and control epidermis measured for chromosomes 4 and 19. (C) Increased proliferation in neonatal Mad2-deficient epidermis evidenced by increased Ki67 staining in hair follicle (HF) and basal interfollicular epidermis (IFE) cells most apparent at postnatal day (P)6 and P11. (D) TUNEL staining reveals increased apoptosis following the hyperkeratotic wave.

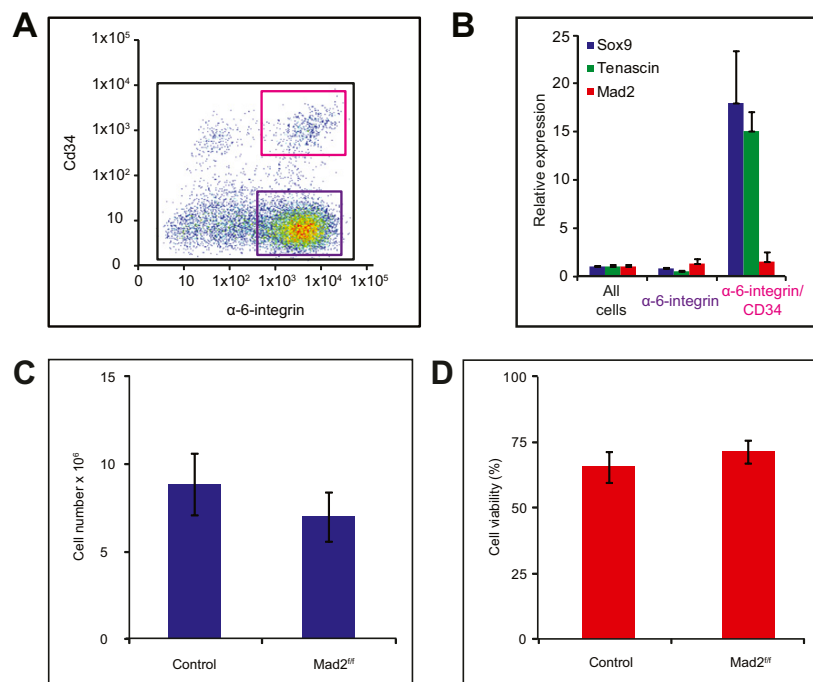


**Fig. S2.** *Mad2*-deficient mice show no indication for a barrier defect evidenced by lack of dye penetration in epidermis. Note that dye did penetrate at an injection site wound (arrow).





**Fig. S3.** Confirming expression array results (A and B) and stem cell depletion in situ (C). qPCR verification for a selection of six genes *Lef1*, *Wnt5a*, *Bmp4*, *Krt16*, *Caspase 14*, and *Il18* (A) and matching microarray values (B). Error bars indicate SEM of at least two biological replicates. (C) Whole mount tail epidermis labeled for  $\alpha$ -6 integrin or Krt14 (labeling IFE in red) and Krt15 or CD34 (both labeling bulge HF stem cells in green) showing HF stem cell depletion in vivo at P21.



**Fig. S4.** Characterization of Mad2 (loss) in vitro in epidermal cells. (A) Separating HF stem cells (red rectangle) and IFE cells (blue rectangle) by flow cytometry. (B) Quantitative PCR for Sox9 and Tenascin (HF stem cell markers) and Mad2 on IFE cells or HF stem cells showing similar expression of Mad2 in IFE cells and HF stem cells. (C) Cell number and (D) viability (assessed by Trypan blue penetration) are similar in freshly isolated control and Mad2<sup>fl/fl</sup>; Krt14-Cre<sup>+</sup> primary keratinocytes.

**Table S1.** Copy number changes for chromosomes 4, 15, and 19 in Mad2-deficient and control epidermis

Chromosome	Genotype	Age	More than two copies	Cells counted	Aneuploid (%)	Average (%)
Chromosome 15	<i>Mad2<sup>fl/fl</sup></i>	3 wk	26	88	30	28
		3 wk	21	94	22	
		3 wk	10	67	15	
		Adult	33	72	46	
	Control	3 wk	10	67	15	8
		3 wk	2	107	2	
		3 wk	4	61	7	
Chromosome 19	<i>Mad2<sup>fl/fl</sup></i>	3 wk	19	82	23	23
		3 wk	12	55	22	
		3 wk	8	56	14	
		Adult	23	68	34	
	Control	3 wk	6	53	11	12
		3 wk	7	57	12	
		3 wk	7	59	12	
Chromosome 4	<i>Mad2<sup>fl/fl</sup></i>	3 wk	12	99	12	20
		3 wk	34	111	31	
		3 wk	13	105	12	
		Adult	28	107	26	
	Control	3 wk	6	107	6	7
		Adult	5	67	7	
Chromosome 19	<i>Mad2<sup>fl/fl</sup></i>	3 wk	9	101	9	22
		3 wk	23	106	22	
		3 wk	16	108	15	
		Adult	44	107	41	
	Control	3 wk	10	105	10	8
		Adult	4	64	6	

Interphase experiments were performed in two independent series of hybridizations for mentioned genotypes. In the first set, we hybridized probes for chromosomes 15 and 19. In the second set, we hybridized chromosomes 4 and 19. Chromosome 19 was included in the second series to compare the results between the hybridizations.

Database	Description	Pathway ID	No. in category	No. observed	No. expected	Ratio enrichment	<i>P</i> value	Bonferroni-corrected <i>P</i> value
Wiki pathways	EGFR1 signaling pathway	WP572	153	29	11.1	2.6	1.64E−06	2.00E−04
Wiki pathways	Type II IFN signaling (IFN $\gamma$ )	WP1253	30	10	2.2	4.6	3.14E−05	2.90E−03
Wiki pathways	Toll-like receptor signaling pathway	WP1271	78	15	5.7	2.6	4.00E−04	3.72E−02
KEGG pathway	Endocytosis	4144	171	31	12.5	2.5	1.96E−06	3.00E−04
KEGG pathway	Metabolic pathways	1100	950	99	69.2	1.4	2.00E−04	2.80E−02
Biological process	Apoptosis	GO:0006915	632	85	47.1	1.8	5.00E−08	8.05E−05
Biological process	Immune response	GO:0006955	401	54	29.9	1.8	1.46E−05	2.35E−02
Biological process	Epidermal cell differentiation	GO:0009913	44	13	3.3	4.0	1.19E−05	1.92E−02
Cellular component	Cornified envelope	GO:0001533	14	10	1.1	9.5	4.16E−09	9.78E−07
Wiki pathways	mRNA processing	WP310	351	54	30.0	1.8	1.65E−05	1.60E−03
KEGG pathway	Spliceosome	3040	91	22	7.8	2.8	6.06E−06	9.00E−04
Biological process	Wnt receptor signaling pathway	GO:0016055	126	27	11.2	2.4	1.34E−05	2.37E−02
Biological process	Smoothened signaling pathway	GO:0007224	44	14	3.9	3.6	1.64E−05	2.90E−02
Biological process	Cell cycle	GO:0007049	553	83	49.3	1.7	1.24E−06	2.20E−03
Biological process	Hair follicle development	GO:0001942	43	14	3.8	3.7	1.22E−05	2.16E−02

0.79, 0.86, 0.91, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99

Primer name	Primer sequence
<b>Genotyping</b>	
Mad2f forward	AGGCTGAGCCGGGCTTAGGAC
Mad2f reverse	GTAAACCGTGTAATAACGTTTAAGTCTC
Cre forward	GCACGTTCAACCGGCATCAAC
Cre reverse	CGATGCAACGAGTGATGAGGTTTC
<b>qPCR</b>	
Mad2 forward	AAACTGGTGGTGGTCATCTC
Mad2 reverse	TTCTCTACGAACACCTTCCTC
Actin forward	CTAGGACCCAGGCTGTGATG
Actin reverse	GGCCTCGTCACCCACATAG
LRIG1 forward	GCAGGAAGTGTAACCTCAACAG
LRIG1 reverse	CCATCCACACTAAGGATCTTG
Sox9 forward	CTCTGGAGGCTGCTGAAC
Sox9 reverse	CGCTTGTCGGTTCTTCAC
Wnt5a forward	CCATGAAGAAGCCATTG
Wnt5a reverse	CCAAAGCCATTAGGAAGAAAC
Lef1 forward	CCCGAGGACATCAAATAAAG
Lef1 reverse	ATGCCTTGCTTGGAGTTG
Bmp4 forward	TCCATCACGAAGAACATCTG
Bmp4 reverse	GCTGAGGTTGAAGAGGAAAC
Lgr5 forward	CACCTCTACCTGGACCTC
Lgr5 reverse	CTTGGGAATGTGTGTCAAAG
Krt16 forward	GCCCACTGAGATCAAAGACTAC
Krt16 reverse	GGCAATAATGATCTTGCTCTTC
Casp14 forward	CATTCAGGAGGAAAGATATG
Casp14 reverse	CTTCATGTCTACTCGGAAC
Il18 forward	ATCAAAGTGCCAGTGAACC
Il18 reverse	CACAGCCAGTCTCTTACTTC
Krt15 forward	GCTCAGAACCCAGGATACAAC
Krt15 reverse	AATACCAGCCATCTTAGCATC

**Dataset S1. Significantly deregulated genes ( $P < 0.05$ ) in at least two time points in response to Mad2 deletion**

[Dataset S1](#)

This table contains all significantly deregulated genes in response to Mad2 deletion in at least two time points in mouse epidermis. Data was extracted from individual microarray analysis files and significantly deregulated genes were sorted on fold deregulation as described in *Materials and Methods*.

**Dataset S2. Complete list of significantly deregulated (Bonferroni  $P < 0.05$ ) pathways (Wiki pathways, KEGG pathways, and GO biological functions) in response to Mad2 abrogation**

[Dataset S2](#)

This table contains the Webgestalt analysis results described in *Materials and Methods* for genes that were either up-regulated or down-regulated in at least two time points in response to Mad2 depletion in mouse epidermis.

**Dataset S3. Complete list of extracted genes in their pathway as extracted from Table 1 using Webgestalt with matched expression values**

[Dataset S3](#)

This table was used to make Fig. 3B using Genesis software.





## CHAPTER FOUR: The Role of *Krt76* in the Epidermis

### 4.1 INTRODUCTION

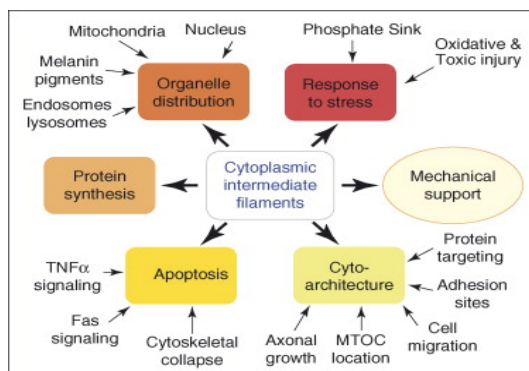
After identifying several genes involved in skin biology at the Sanger-MGP, I chose one of the lines to investigate for the remainder of my project. In line with the aim of this thesis, to understand the genetic complement of skin development and model skin disease, I chose to examine the consequence of germ line inactivation of Keratin 76 (*Krt76*), a poorly understood type II intermediate filament gene. *Krt76*, previously known as *K2p*, is strongly expressed in the palate<sup>118</sup>. Prior to this study, little was known about *Krt76* apart from a limited characterization of its expression pattern. I initially flagged this gene in the dysmorphology screen. It also received phenotype hits for skin histopathology. Although it is an extremely understudied gene, my preliminary studies at the Sanger-MGP indicated that this gene may have a novel role in regulating epidermal homeostasis. Furthermore, there is an increasing body of research suggesting that keratins play roles beyond structural support. Keratin disruption has previously been linked to human skin disease, making *Krt76* an attractive candidate gene to study, from both cutaneous biology and disease perspectives.

I am particularly interested in the regeneration of skin, especially with respect to wound repair. The *Krt76* model seemed to be a great system in which to study skin regeneration because these mice presented with multiple wounds and scabs, which did not seem to properly resolve. There are known ‘wounding keratin’ genes that are upregulated during wound healing, but *Krt76* had never been implicated in this process.

There are 54 known keratin genes in humans, of which, approximately 21 are known to be associated with disease<sup>101,102</sup>. Keratins are distinguished primarily by their roles in providing structural integrity to epithelial cells, enabling resistance to mechanical stress and trauma. Therefore, mutations in genes encoding keratin proteins commonly result in phenotypes characterized by compromised cellular integrity and tissue fragility. These phenotypes include, blistering and cytolysis which then lead to impaired ability of the skin to maintain a protective barrier and resist mechanical stress. Phenotypes are further exacerbated by hyperproliferation, as the organ attempts to compensate for fragility, resulting in hyperkeratosis. Epidermolysis bullosa (EB) comprises a heterogeneous group of heritable blistering diseases, classified by the severity and location of blistering.

Epidermolysis bullous simplex (EBS) was the first inherited keratin disorder and found to be caused by mutations of Keratin 5 and Keratin14<sup>94</sup>. Since then, more than 20 keratin associated disorders have been described, most of which involve epidermal fragility and or tissue overgrowth<sup>97</sup>. Other notable diseases involving keratin IF dysfunction include keratinopathic ichthyosis (KRT1, KRT2, KRT10), palmoplantar keratoderma (KRT9), pachyonychia congenita (KRT6a, KRT6b, KRT16, KRT17),<sup>7</sup> and monilethrix (KRT81, KRT83 and KRT86)<sup>49,101-103,109</sup>.

Dysfunctional IF proteins do not exclusively lead to cutaneous diseases. As discussed, keratins are expressed across a broad range of tissues and in both simple and stratified epithelia. Mutations in genes encoding IF proteins have been reported to directly cause or predispose individuals to diseases affecting many organ systems including lungs, liver, stomach, kidney, bladder, as well as skin<sup>119</sup>. The classical view that keratins provide a simple structural scaffold to preserve the integrity of the skin has been challenged by recent studies demonstrating their increasingly specialised and diverse functions in the epidermis<sup>120</sup> including protection from apoptosis<sup>119,121</sup>, injury<sup>122</sup>, regulation of epithelial polarity<sup>123,124</sup> and influence on cell size and protein translation<sup>121,125-127</sup>. This thesis contributes to the understanding of IF function and shows a novel mechanism by which IF's contribute to epidermal biology.



**Figure 4.1 Functions of cytoplasmic intermediate filaments.** There is conclusive evidence, derived in large part from genetically engineered mice, supporting the participation of keratin and other cytoplasmic intermediate filament proteins in processes beyond structural support. Adapted from<sup>128</sup>.

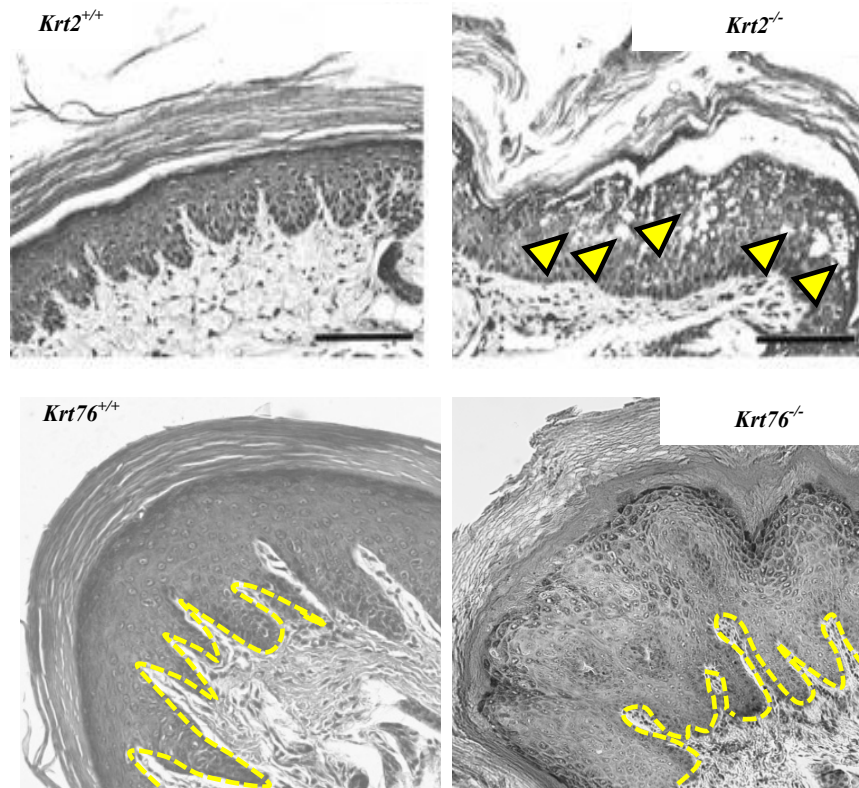
Preliminary screening at the Sanger-MGP indicated that *Krt76* might be an IF with a novel function in skin biology. *Krt76* is absolutely required to maintain the integrity of the skin and, because knockout mice develop extensive non-healing wounds, it is

essential for postnatal survival. Animals homozygous for *Krt76* mutations exhibit scaling and flaky, dry skin at birth. With the onset of hair growth, skin defects diminish and homozygous animals appear to develop normally. Between 4 to 6 weeks of age, hyperkeratotic areas on the footpads develop, presenting as brown, thick hyperkeratotic plaques. At sites of friction, including the footpads, tail, and facial areas; skin erosions, lesions and blistering can occur in response to trauma. In addition, homozygous mice displayed hyperpigmentation of the paw pads. Expression data in heterozygous mice complimented some of the phenotypic data, particularly with respect to the pawpad expression.

I started my investigation into *Krt76* with the paw pad hyperpigmentation phenotype, mainly because the strongest skin specific expression was reported in there. As reviewed in Chapter 1, very few mutations alter skin colour as compared to hair colour, partly because hair melanocytes are more numerous than skin melanocytes<sup>65</sup>. It was determined that *Krt76* mice were a Class II pigmentation mutant. Class II mutants exhibit pigment on volar pads, intervolar pad scales and on phalangeal ridges. Histologically, all Class II mutants have excess pigmentation in the epidermis, as opposed to the dermis as in Class I mutants. Additionally, in all class II mutants, hyperpigmentation becomes more dramatic with age, suggesting involvement in epidermal melanoblast differentiation and/or proliferation. Preliminary characterization of the *Krt76* knockout mouse show similarities to other Class II dark skin mutant models including late onset hyperpigmentation on pads and phalangeal ridges. Furthermore, *Krt76* is not the only keratin gene whose dysfunction induces Class II hyperpigmentation patterns. The gross pigmentation phenotype in *Krt76* mice is reminiscent to that of the *Dsk2* mutants. *Dsk2* was discovered in a phenotype driven screen where mutations have been mapped to *Krt2*. Like *Krt2*, *Krt76* is also expressed in the suprabasal layers of the epidermis. *Krt2/Dsk2* remains a model of melanocytosis in response to keratin scaffolding defects, and therefore was the basis for my initial investigation of *Krt76*. The first thing I did was compare the histological phenotypes of the two mutant mice.

Like *Krt76* mutants, *Krt2* mutants show excess accumulation of pigment at approximately 3-6 weeks of age. This delay in hyperpigmentation onset suggests that the accumulation of excess pigment is secondary to keratinocyte fragility. To investigate this, I used histology to look for evidence of cytolysis and intra-epidermal instability. Surprisingly,

the epidermis of the *Krt76* mice was completely intact. This was the first indication that the molecular mechanisms underlying the pigmentation phenotype and other phenotypes associated with loss of *Krt6* might be different from classical pathogenesis.



**Figure 4.2 Preliminary histological comparison of *Krt76*<sup>-/-</sup> and *Krt2*<sup>-/-</sup> paw pads.** *Krt2* mutants footpads (top) have a thickened epidermis and exhibit cytolysis (arrows) in suprabasal layers whereas *Krt76* mutants (bottom) also exhibit thickened epidermis, but without evidence of cellular lysis or keratinocyte fragility (Krt2 images adapted from Fitch et al.)

This preliminary data sparked a mechanistic investigation into the role of *Krt76* in skin biology. In this chapter, I have shown that a previously poorly characterised protein is essential for postnatal survival and that it is essential for maintaining epidermal barrier function and integrity. I show the upregulation of *Krt76* in wound healing and its contribution to the skin barrier. This study is part of the increasing body of research suggesting that there are roles for the intermediate filament network beyond maintaining tissue integrity. Finally, it again highlights the importance of using mouse models to investigate the mechanisms and genetics underlying skin pathology.

## 4.2 DECLARATION

### Declaration for Thesis Chapter 4: Keratin 76 is required for tight junction function and maintenance of the skin barrier

In the case of Chapter 2, the nature and extent of my contribution to the work was the following:

Nature of contribution	Extent of contribution (%)
Laboratory/Experimental <ul style="list-style-type: none"><li>Performed research</li></ul> Theoretical: <ul style="list-style-type: none"><li>Designed experiments</li><li>Data analysis</li></ul> Manuscript: <ul style="list-style-type: none"><li>Wrote manuscript</li><li>Made figures</li></ul>	90%

The following co-authors contributed to the work. If co-authors are students at Monash University, the extent of their contribution in percentage terms must be stated:

Name	Nature of contribution	Extent of contribution (%) for student co-authors only
Denny L. Cottle	Intellectual contributions, manuscript comments/edits	N/A
Helen B. Pearson	Performed research, contributed reagents	N/A
Holger Schlüter	performed research, contributed reagents	N/A
Pritinder Kaur	contributed reagents	N/A
Patrick O. Humbert	contributed reagents	N/A
Ian Smyth	Designed research, performed research, contributed reagents/analytic tools, analysed data, wrote manuscript	N/A

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the candidate's and co-authors' contributions to this work\*.

<b>Candidate's Signature</b>		<b>Date</b> 6/6/2014
------------------------------	---	----------------------

<b>Main Supervisor's Signature</b>		<b>Date</b>
------------------------------------	--	-------------

\*Note: Where the responsible author is not the candidate's main supervisor, the main supervisor should consult with the responsible author to agree on the respective contributions of the authors.

#### 4.3 MANUSCRIPT

### **Keratin 76 is required for tight junction function and maintenance of the skin barrier**

Tia DiTommaso<sup>1</sup>, Denny L. Cottle<sup>1</sup>, Helen B. Pearson<sup>2,3</sup>, Holger Schlüter<sup>2,3</sup>, Pritinder Kaur<sup>2,3,4</sup>, Patrick O. Humbert<sup>2,3,5,6</sup>, and Ian Smyth<sup>1,7</sup>

1. Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia, 3800.

2. Research Division, The Sir Peter MacCallum Cancer Centre, Melbourne, Australia, 3002.

3. The Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Australia, 3010.

4. Department of Anatomy & Neuroscience, University of Melbourne, Melbourne, Australia, 3010.

5. Department of Biochemistry and Molecular Biology, University of Melbourne, Melbourne, Australia, 3010.

6. Department of Pathology, University of Melbourne, Melbourne, Australia, 3010.

7. Department of Anatomy and Developmental Biology, Monash University, Melbourne, Australia, 3800.

*Classification:* BIOLOGICAL SCIENCES- genetics

Contact:

A/Prof Ian Smyth  
Monash University  
Wellington Rd, Clayton  
Melbourne, Australia 3800.

  
f. +61 (0)3 9902 9500

## Abstract

Keratins are cytoskeletal intermediate filament proteins that are increasingly being recognised for their diverse cellular functions. Here we report the consequences of germ line inactivation of *Keratin 76* (*Krt76*) in mice. Homozygous disruption of this epidermally expressed gene causes neonatal skin flaking, hyperpigmentation, inflammation, impaired wound healing, and death prior to 12 weeks of age. We show that this phenotype derives from functionally defective tight junctions that are characterised by mislocalization of the integral protein CLDN1. We further demonstrate that KRT76 interacts with CLDN1 and propose that this interaction is necessary to correctly position CLDN1 in tight junctions. The mislocalization of CLDN1 has been associated with various dermatopathies, including the inflammatory disease, psoriasis. Other features of psoriasis, including barrier dysfunction, hyperkeratosis and up-regulation of Keratin 6 and FABP5 are evident in the *Krt76* homozygous mice. These observations establish a previously unknown connection between the intermediate filament cytoskeleton network and tight junctions and showcase *Krt76* null mice as a possible model to study aberrant tight junction driven skin diseases.

## Significance Statement

The generation of knockout mice are a central approach to studying gene function. We have examined the consequences of the germ line inactivation of *Keratin 76* in mice and in doing so we reveal a previously undescribed mechanism by which keratin intermediate filaments regulate cellular interactions and tissue homeostasis. Our study supports an emerging body of evidence which challenges the classical view of the keratin intermediate filaments as simple structural proteins, highlighting *Krt76* as a gene whose function is indispensable for normal skin biology and postnatal survival as a result of its novel interactions with tight junction complexes. This study identifies a previously unknown and critical link between intermediate filaments and tight junctions and poses a novel mechanism by which intermediate filament dysfunction influences skin disease.

## Introduction

The epidermis provides a stable and selectively permeable barrier essential to terrestrial life. Together with microfilaments and microtubules, intermediate filaments (IFs) make up the major components of the epidermal cytoskeleton. Keratins are the largest subgroup of the IF proteins and comprise the major structural proteins in epithelial cells<sup>79</sup>. Keratins

are composed of a central, filament forming, alpha-helical rod domain of ~310 amino acids that is flanked by non-helical head and tail domains<sup>79,129-132</sup>. They act as a flexible scaffold enabling cells to resist physical stress. Consequently, defects in IFs can lead to cell fragility and are linked to a wide array of genodermatoses and cancers<sup>132,133</sup>. The classical view that keratins simply provide a structural scaffold has been challenged by recent studies demonstrating their increasingly specialised and diverse functions<sup>120</sup> which include protection from apoptosis<sup>119,134</sup> and injury<sup>122</sup>, regulation of epithelial polarity<sup>123,124</sup> and influence on cell size and protein translation<sup>121,122,125,127</sup>.

The functional integration of cytoskeletal elements and cellular junctions is critical for the establishment and maintenance of the epidermal barrier. Tight junctions (TJ) form a seal between cells which make up the layers of the epidermis<sup>135</sup>. This barrier is selectively permeable, allowing passage of small molecules, but restricting water loss, and allowing for antigen sampling by immune cells<sup>64,135,136</sup>. TJs are composed of scaffolding and adhesion molecules including claudins, junctional adhesion molecules and occludins. Defective tight junction organization has been linked to compromised barrier function<sup>64</sup> and the development of various dermatopathies including psoriasis<sup>137,138</sup>. The TJs are thought to interact with the IF network by binding of a number of integral or associated TJ proteins that complex to F-actin<sup>139</sup> but their associations, if any, with the keratin IF network are unclear.

Keratin 76 (KRT76, previously K2p) is a poorly studied member of the keratin family, although its down-regulation has recently been linked to oral carcinoma<sup>140</sup> in humans. We report the characterisation of murine *Krt76* gene disruption and show that *Krt76* is expressed in specific epithelia and is up-regulated during wound healing of the skin. KRT76 function is required for epidermal homeostasis and for survival beyond 12 weeks of age. We find that the protein is required to maintain epidermal barrier function and, unexpectedly, for the correct localisation of CLAUDIN1 (CLDN1) within tight junctions. These findings identify a critical relationship between the IF network and TJs which is essential for epidermal homeostasis *in vivo*.



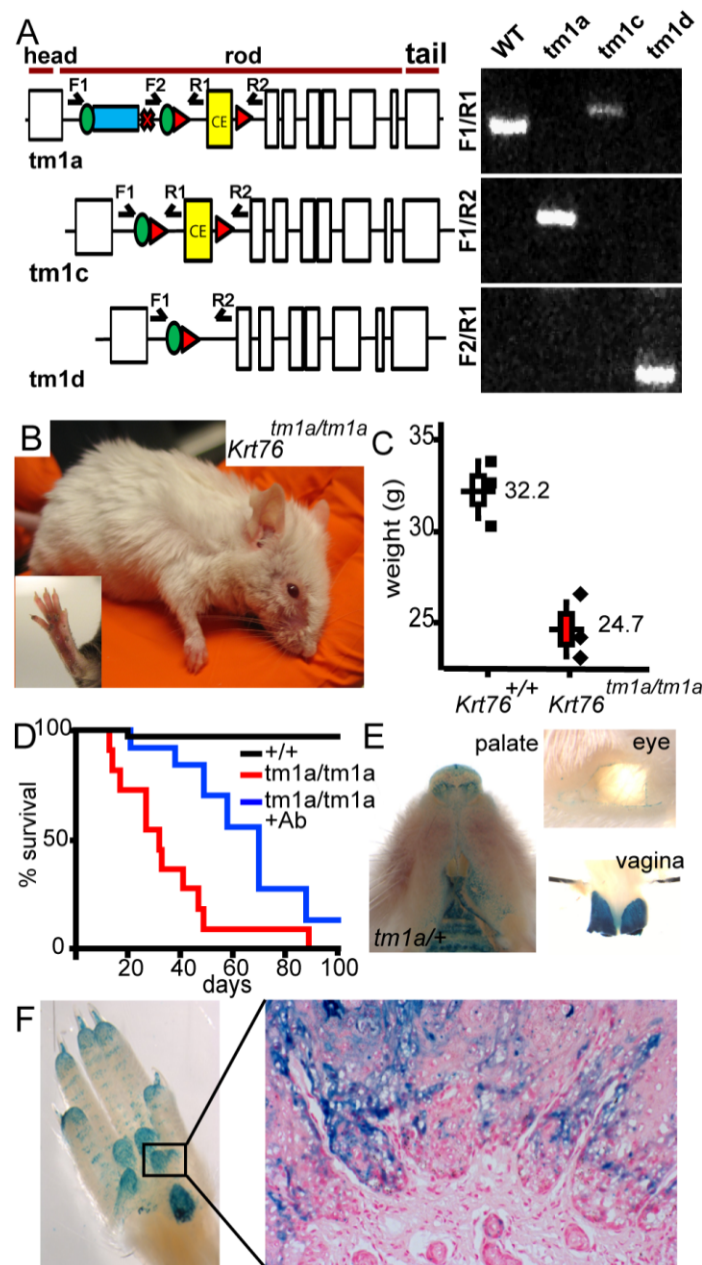
## Results

*Loss of Krt76 causes gross epidermal defects and results in lethality.*

As part of the Wellcome Trust Sanger Institute (WTSI) Mouse Genetics Programme<sup>19</sup> we identified cutaneous defects in the genetrapp “knockout first”<sup>17</sup> allele of *Keratin 76* (*Krt76*<sup>tm1a(KOMP)Wtsi</sup> hereafter *Krt76*<sup>tm1a</sup>). This mouse has a SA-β-gal reporter integrated upstream of floxed exon 2 (Fig. 1A), that allows gene expression to be traced whilst disrupting gene function. To generate a conditional-ready strain, these mice were bred with a germline flippase expressing line, thereby removing the reporter cassette (*Krt76*<sup>tm1c</sup>). This allele was then manipulated to achieve global and inducible epidermal gene deletion by crossing to CMV-Cre and K14CreERT transgenic strains respectively (*Krt76*<sup>tm1d/tm1d</sup> and *Krt76*<sup>tm1dK14/tm1dK14</sup>) (Fig.1A). Inducible deletion in the latter was achieved through the topical application of 4-hydroxytamoxifen (4OHT)).

*Krt76*<sup>tm1a/tm1a</sup> mutant mice on a mixed but predominantly C57BL/6 background, exhibit flaky skin at birth, although these defects diminish somewhat with the emergence of hair follicles. At weaning, mutant mice are distinguished by their unkempt, dull coats (Fig. 1B) and hyperpigmented paw pads (Fig. 1B insert). Ageing animals develop multiple skin lesions which are exacerbated in regions of active grooming such as the footpads, tail, and face (Fig.1B). *Krt76*<sup>tm1a/tm1a</sup> mutant mice also suffer from significant weight loss (Fig. 1C) and their survival is reduced compared to wildtype (WT) and heterozygous littermates (median = 32 days, Fig. 1D).

We observe a similar profile of disease development in *Krt76*<sup>tm1d/tm1d</sup>. To assess whether cutaneous bacterial infection may exacerbate morbidity, we treated *Krt76*<sup>tm1a/tm1a</sup> mice with a broad spectrum antibiotic (Baytril) and observed a considerable improvement in lifespan (median survival = 72 days, p = 0.04) (Fig. 1D). From this, we hypothesized that the progressive development of wounds and increased susceptibility to bacterial challenge, may be indicative of a primary skin barrier defect.



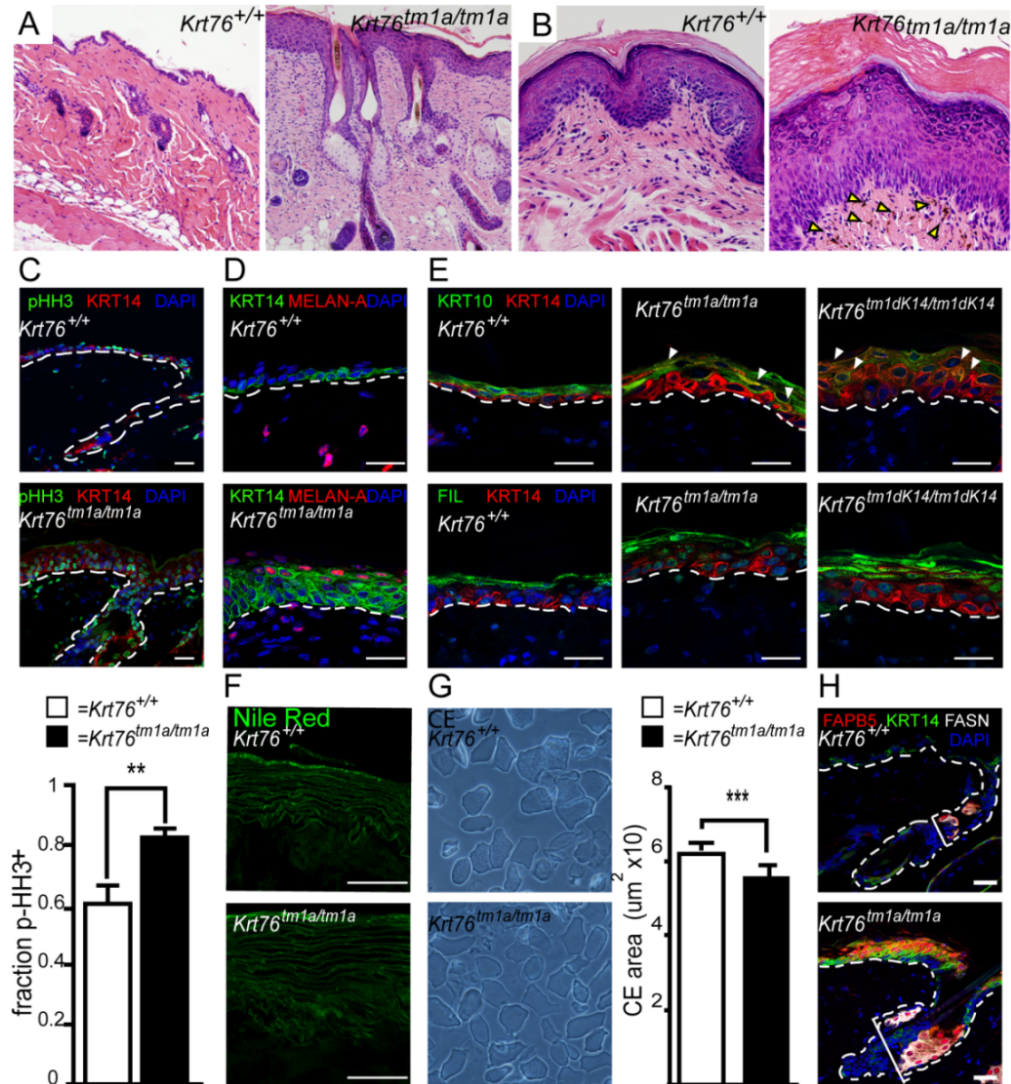
**Figure 1. Generation and phenotypic characterisation of Keratin 76 mouse models.** (A) Exon structure and domain prediction of mouse *Krt76*. Blue box represents insertion of  $\beta$ -galactosidase ( $\beta$ -gal) cassette in the *Krt76*<sup>tm1a</sup> reporter allele. Green circles and red triangles indicate frt and loxP sites. Validation of the mutant alleles was achieved using PCR amplification. (B) Mutagenesis of *Krt76* leads to the progressive development of skin lesions, rough coat and paw pad hyperpigmentation (insert). (C) *Krt76*<sup>tm1a/tm1a</sup> knockouts show significant reduction in weight (n=3 males, age 9 weeks). (D) *Krt76*<sup>tm1a/tm1a</sup> knockout mice die progressively from ~ 2 weeks after birth with no animals surviving beyond 12 weeks of age (n=32). Treatment with Baytril reduces morbidity and mortality. (E) Whole mount x-gal staining of *Krt76*<sup>tm1a/+</sup> mice shows *Krt76* expression in the palate (left panel), eyelid (upper middle), vagina (lower middle) and (F) paw pads. *Krt76* expression is detected throughout the layers of the paw pad epidermis and the exocrine glands.

*Krt76 is differentially expressed in epithelia*

*Krt76* expression has previously been reported in the palatal and gingival epithelium<sup>141</sup>. The dramatic consequences of *Krt76* disruption to the broader epidermis prompted us to evaluate its expression under homeostatic conditions; facilitated by the presence of an integrated  $\beta$ -gal reporter in our model. We confirmed expression of *Krt76* in oral epithelia (Fig. 1E palate) and detected previously unrecognised expression in the vagina and the eyelid (Fig. 1E). Furthermore, we noted expression throughout the stratified layers of the footpad and in the exocrine glands (Fig 1F right). The expression in the paw pad mimicked the pattern of hyperpigmentation observed in mutant mice (Fig 1B insert, Fig1F). Examination of embryos at E14.5, E16.5 and E18.5 detected only very modest  $\beta$ -gal detection in the paw pad and oral epithelium (data not shown), indicating that the defects in postnatal skin in homozygous mice were unlikely to be caused by defects established during development.

Although expression in the interfollicular epidermis (IFE) of healthy mice was low, histological characterization of *Krt76*<sup>tm1a/tm1a</sup> skin highlighted defects in this compartment. Shortly after birth both dorsal IFE and paw pads exhibited hyperplasia and hyperpigmentation (Fig. 2A, B). As the condition of these animals worsened, the cellularity of the dermis increased as a consequence of immune infiltration. The dorsal IFE developed keratinised nodules in the stratum spinosum/granulosum and hyperkeratosis was endemic (Fig. 2A, B).

Unlike other models of intermediate filament dysfunction (e.g. KRT5/14) we did not observe any evidence of intra-epidermal cytolysis, indicating a novel mechanism by which loss of an intermediate filament leads to skin dysfunction (Fig. 2A, B). Several of these features were reminiscent of the hyperproliferative skin disorder, psoriasis<sup>142-144</sup> and analysis using phospho-HistoneH3 (pHH3) antibody (Fig. 2C) confirmed a significant increase in basal keratinocyte proliferation in *Krt76*<sup>tm1a/tm1a</sup> mice ( $p=0.005$ , Fig. 2C). Hyperpigmentation of dorsal skin was assessed using an MELAN-A antibody (Fig. 2D) and melanocytes were abnormally detected in the IFE of *Krt76*<sup>tm1a/tm1a</sup> epidermis. Their location concurred with the increased incidence of pigment detected in H&E sections<sup>145</sup> (Fig. 2B arrows). MELAN-A<sup>+</sup> cells were also detected in the dermis at low frequency in both WT and *Krt76*<sup>tm1a/tm1a</sup> mice (Fig. 2C).



**Figure 2. A role for Krt76 in epidermal homeostasis.** (A) Back skin from *Krt76<sup>tm1a/tm1a</sup>* animals is hyperplastic and hyperkeratotic with immune infiltrate and enlarged sebaceous glands. (B) Paw skin from wild type (left) and *Krt76<sup>tm1a/tm1a</sup>* (right) shows thickened, hyperkeratotic, epidermis with pigment cells in the dermis (arrows) and no evidence of cell lysis. (C) Increased phospho-histone H3 (pHH3) +cells ( $p=0.005$ ) and (D) increased levels of MELAN-A staining are detected in *Krt76<sup>tm1a/tm1a</sup>* mice. (E) KRT14, FIL and KRT10 expression in *Krt76<sup>tm1a/tm1a</sup>* epidermis showing a relatively normal differentiation program, despite expansion of KRT14+ basal layer. *Krt76<sup>tm1dK14/tm1dK14</sup>* mice phenocopy the *Krt76<sup>tm1a/tm1a</sup>* knockout with expanded KRT14 + layer and normal induction of differentiation markers, FIL and KRT10. Arrows indicate increased coexpression of KRT10 and KRT14, a consequence of expanded KRT14+ basal layer. (F) Lipid deposition in intercellular lamellae is normal as assessed by Nile Red staining. (G) Corneocytes are grossly normal but exhibit a modest reduction in surface area ( $p=0.0004$ ). (H) Sebaceous glands, marked by FASN and FABP5, exhibit enlargement and FABP5 is upregulated in the suprabasal layers of the IFE. Error bars = S.E.M.

To examine the program of differentiation in keratinocytes we examined the ordinarily sequential expression of Keratin 14 (KRT14, stratum basale), Keratin 10 (KRT10, stratum spinosum) and Filaggrin (FIL, stratum granulosum). As an additional control we crossed the *Krt76*<sup>tm1c/tm1c</sup> conditional-ready mice into a K14-Cre-ERT background and applied 4-hydroxytamoxifen (4OHT) on the dorsal skin of 8 week old adult mice to inactivate *Krt76* specifically in the epidermis, thereby generating *Krt76*<sup>tm1dK14/tm1dK14</sup> mice. Both the germline *Krt76*<sup>tm1a/tm1a</sup> as well as the epidermis-specific *Krt76* KO mice (*Krt76*<sup>tm1dK14/tm1dK14</sup>) exhibited a thicker layer of KRT14<sup>+</sup> basal keratinocytes (Fig 2E), consistent with the hyperproliferation assessed by pHH3 (Fig 2D) in dorsal skin. Both the spinous (KRT10) and granular (FIL) cell layers appeared to differentiate in normal sequence (Fig. 2E), despite expanded KRT14 expression (Fig. 2E, arrows). We also surveyed lipid profiles with Nile Red, demonstrating that the deposition of extracellular lipid lamellae were unaffected in mutant animals (Fig. 2F). The terminal products of epidermal differentiation, the corneocytes, also appeared to form normally, albeit with a small but significant reduction in surface area which we propose derives from hypercellularity in the epidermis (p= 0.0004, Fig. 2G). Sebaceous glands of *Krt76*<sup>tm1a/tm1a</sup> mice exhibited enlargement (Fig. 2A) highlighted by sebocyte expression of FASN and FABP5<sup>146,147</sup> (Fig. 2H). Over-expression of the latter marker was also observed in the IFE and is associated with barrier dysfunction in psoriasis<sup>148</sup> and wound healing<sup>149</sup>.

#### *Krt76 is required for normal wound healing*

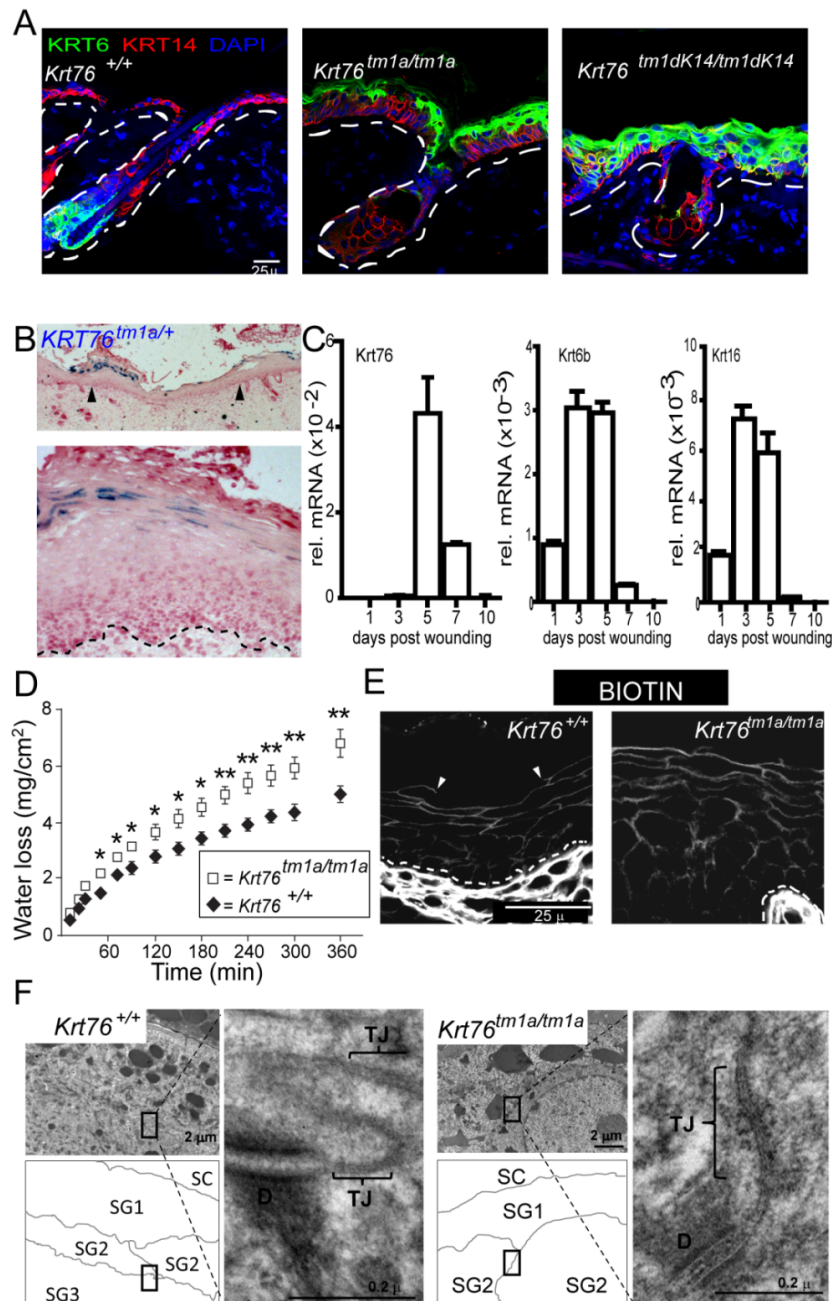
The hyper-proliferation and basal layer expansion in *Krt76*<sup>tm1a/tm1a</sup> skin were suggestive of a defective wound healing response in these mice. Indeed *Krt76*<sup>tm1a/tm1a</sup> and *Krt76*<sup>tm1dK14/tm1dK14</sup> animals displayed broad expression of the classic wounding KERATIN 6 in the IFE, in addition to its normal hair follicle expression<sup>150</sup> (Fig. 3A). To evaluate whether *Krt76* is also induced during the wound healing response, we took dorsal skin samples from *Krt76*<sup>tm1a/+</sup> and WT mice 1, 3, 5, 7, and 10 days after wounding by a punch biopsy.  $\beta$ -gal staining of *Krt76*<sup>tm1a/+</sup> wound sections highlighted expression of *Krt76* 7 days post wounding, specifically in suprabasal cells of the re-epithelised wound (Fig. 3B). Expression profiling by qRT-PCR in WT mice confirmed this expression in a profile slightly delayed from the “classical” wounding keratins *Krt6b* and *Krt16* (Fig. 3C). The spatial and temporal upregulation of *Krt76* in response to wounding correlates with the appearance of progressive unresolved cutaneous damage in *Krt76*<sup>tm1a/tm1a</sup> mice.

*Krt76<sup>tm1a/tm1a</sup> mice show barrier function defects.*

We noted that the hyperproliferation, induction of wounding keratins, unresolved wounds, and follicular dysmorphology observed upon *Krt76* inactivation are phenotypes associated with a loss of barrier function. To investigate whether this was the case *in utero* we examined fetal mice using a toluidine blue assay, but found no differences in dye exclusion (data not shown). However, examination of neonatal barrier function using a transepidermal water loss (TEWL) assay, identified a significant defect in the cutaneous barrier in *Krt76<sup>tm1a/tm1a</sup>* pups compared to their control littermates (Fig. 3D). Importantly, this defect (at P3) is apparent before skin lesions develop. These results show that while *Krt76* is not required for the development of fetal barrier function, it is essential for its maintenance after birth.

Our previous phenotypic characterisation indicated that these barrier function defects are not linked to overt defects in cell lysis/cell stability, epidermal stratification, lipid deposition or terminal differentiation. Loss of TJ function is known to result in barrier dysfunction independent of defects in lipid deposition or keratinocyte differentiation<sup>64,135</sup>. Furthermore, alterations in TJ proteins are an early event in psoriasis<sup>138</sup>, a disease with phenotypes that parallel some of those evident in *Krt76<sup>tm1a/tm1a</sup>* mice. Numerous studies describe the existence of TJ mediated barrier dysfunction in psoriasis while some go as far as to suggest that the response of epidermal cells to barrier dysfunction may exacerbate or even initiate such diseases (reviewed in<sup>151</sup>). On this basis we then investigated whether defects in TJ formation might contribute to the barrier defect and psoriasis-like phenotype observed in *Krt76* mutant animals. To investigate TJ integrity we employed a biotin tracer approach in which newborn mouse paw pads and dorsal skin were injected subcutaneously with membrane impermeable Sulfo-NHS-Biotin that was tracked using streptavidin immunohistochemistry. In WT epidermis, the diffusion of this high molecular weight compound was arrested prior to the interface of the granular and cornified layers (Fig. 3E). However in *Krt76<sup>tm1a/tm1a</sup>* littermates, diffusion of the tracer was detected through to the cornified layer, confirming defective TJ function in these mice (Fig. 3E). Ultrastructural investigation of P3 *Krt76<sup>tm1a/tm1a</sup>* skin by transmission electron microscopy (TEM) found that TJs displayed normal structural features (e.g. kissing points) and their number and position were comparable to their WT and heterozygote littermates (Fig. 3F).





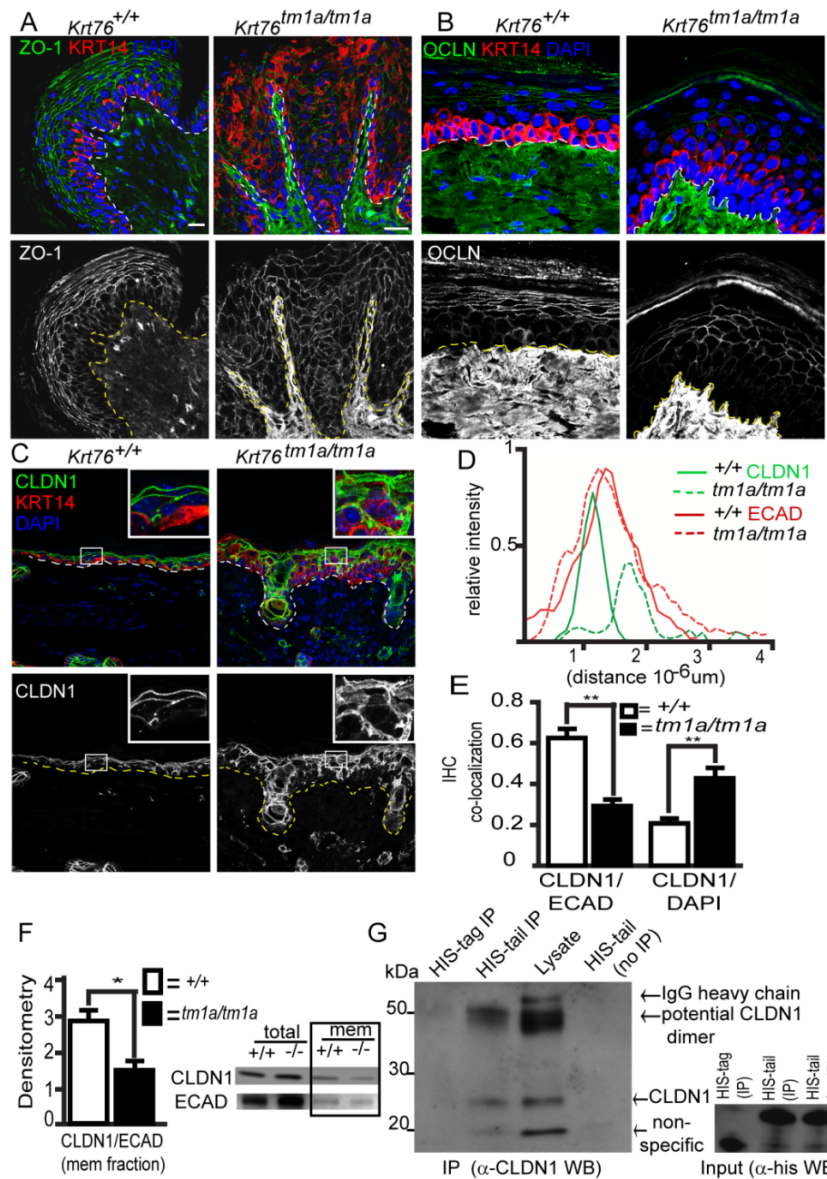
**Figure 3 Krt76 in wounding and barrier function.** (A) Loss of *Krt76* causes induction of KRT6 expression in the IFE in *Krt76*<sup>tm1a/tm1a</sup> and *Krt76*<sup>tm1dK14/tm1dK14</sup> knockout mice. (B) β-gal staining of *Krt76*<sup>tm1a/+</sup> wounds at 7 days post wound show expression of *Krt76* in the suprabasal layer of neoplastic epidermis. (C) *Krt76* is upregulated later than the classical wounding keratins, *Krt6b* and *Krt16*. (D) Transepidermal water loss assays highlight defects in epidermal barrier function in P3 mice. (E) A biotin tracer assay illustrates TJ mediated exclusion from the upper layers of the epidermis in P3 wild type skin (left panel, arrows) but not in *Krt76*<sup>tm1a/tm1a</sup> littermates (dotted line = basement membrane). (E) TJ ultrastructure revealed no defects in TJ incidence, localisation or structure. Low resolution images and schematics (left) show the position of TJs examined at high resolution (right). D= desmosome, Error Bars = S.E.M.

#### *Krt76 interacts with and stabilises Claudin1 at tight junctions*

Having ruled out gross structural defects as the basis for the TJ dysfunction, we next examined whether KRT76 is linked to the regulation of TJ composition by examining the TJ components ZO-1, OCCLUDIN (OCCLN), and CLAUDIN (CLDN1). The localisation of both ZO-1 and OCCLN (Fig. 4 A,B) was unaffected by loss of KRT76 however the subcellular distribution of CLDN1 was altered (Fig. 4C). Quantitative analysis of immunofluorescent intensity relative to the cell membrane confirmed CLDN1 showed a lower signal and broader distribution at the cell periphery. E-cadherin (ECAD) localisation remained unchanged (Fig. 4D). Further analysis also showed an increase in CLDN1 nuclear localization at the expense of membrane staining (Fig. 4E). To confirm our visual assessments, we fractionated P3 keratinocytes and observed significantly less protein at the cell surface relative to ECAD compared to WT littermates ( $p=0.01$ ) (Fig. 4F). No difference in total CLDN1 levels was observed in mutant skin relative to ECAD (Fig. 4F). Our observations using several different experimental approaches indicate that KRT76 is required for the normal membrane localization of CLDN1.

Given that KRT76 is required for normal CLDN1 association we next assessed a possible physical association between the proteins. Attempts to express soluble full length KRT76 in a number of cell systems were unsuccessful, so instead we tried separate domains. The tail domain of KRT76 proved to be soluble when overexpressed with a HIS-tag in *E.coli*. This HIS-KRT76 tail domain protein was purified and immobilised on a HIS-binding column and used to pull down interacting proteins from paw skin protein lysate. In doing so we found a specific interaction between the tail domain of KRT76 and endogenous CLDN1 which runs at 23 kDa, as well as a higher molecular weight (50 kDa) immunoreactive band which may represent previously reported CLDN1 dimers<sup>152,153</sup>. No such interactions were observed with the HIS-tag control protein (Fig. 4G). These bands were absent from samples containing bound HIS-tail domain protein not incubated with paw skin extracts, confirming specificity.





**Figure 4** Localization of CLDN1 is dependent on binding with KRT76 tail domain. (A) Loss of *Krt76* in the skin does not alter ZO-1 or (B) OCCLUDIN localization to the Tight Junction (TJ), (C) whereas localisation of CLDN1 in *Krt76<sup>tm1a/tm1a</sup>* mice is abnormal when compared with wild type (WT) mice. (D) Image analysis at the cellular surface shows a directional shift and a decrease in intensity of CLDN1 where red lines represent ECAD and green lines represent CLDN1 with the solid and dashed lines indicating WT and *Krt76<sup>tm1a/tm1a</sup>*, respectively (lower left). (E) Further quantification by image analysis reveals decreased levels of CLDN1 co-localising at the cell surface with ECADHERIN (ECAD) and a parallel increase in levels of the protein co-localising with DAPI in the nucleus. (F) These observations were confirmed by subcellular fractionation of p3 epidermal peels blotted for ECAD and CLDN1, which show a decrease in relative CLDN1 levels in the membrane fraction of the *Krt76<sup>tm1a/tm1a</sup>* mice. (G) Pull-down of CLDN1 following incubation of purified, recombinant HIS tagged tail domain and HIS tag alone with paw pad lysate indicates an interaction between KRT76- tail domain and CLDN1. Error Bars = S.E.M.

## DISCUSSION

The keratins are classically regarded as structural proteins whose role is to contribute to the fabric of the cytoskeleton and to stabilise epithelial cells. However, this somewhat simplistic view has increasingly been challenged through the description of their specialised and dynamic functions in a number of cellular and developmental contexts. The keratins are the most diverse class of intermediate filament proteins and in many cases their functions are poorly defined. In this study we describe the characterisation of KERATIN76, of one of the least understood of the protein family, delineating its essential role in the maintenance of the integrity of the skin. Under resting conditions, *Krt76* is expressed at its highest levels in the paw, oral epithelium and vagina. However, it is markedly induced during wounding in a profile distinct from other wounding keratins like *Krt6* and *Krt16*.

To examine the functional relevance of this expression and its role in epidermal homeostasis we inactivated the gene in mice globally and in a skin specific manner. Strikingly, loss of *Krt76* resulted in the rapid appearance of extensive non-healing wounds whose subsequent infection contributed significantly to morbidity and mortality. Unlike other knockout models of structural keratins in which extensive cytolysis and blistering is evident, assessment of the cutaneous phenotypes in these animals demonstrated that keratinocytes undergo a normal stepwise process of epidermal maturation. However, closer assessment did reveal an underlying hyperproliferative effect which, coupled with ectopic expression of proteins such as KRT6 and FABP5, indicated that these mice may have a defective epidermal barrier. This was confirmed through trans-epidermal water loss assays on neonatal skin. Having excluded the possibility that overt developmental defects were the cause of increased epidermal permeability we characterised a tight junction specific defect that we propose derives from mislocalisation of CLAUDIN1. There are striking parallels between our phenotype and those evident upon abrogated CLDN1 function in mouse skin<sup>64</sup>. In both cases barrier function defects are detectable by biotin tracer and TEWL assays (but not by dye exclusion), with apparently normal formation of TJ structures as assessed by EM. Overall, the phenotypes of *Krt76* null mice are somewhat milder than their CLDN1 counterparts, suggesting that, despite loss of KRT76, some CLDN1 can still contribute to partial TJ function. This correlates with the observation that the mislocalization of CLDN1 in *Krt76* mutant mice is not absolute. Similar mislocalization of CLDN1 has been previously shown in

cutaneous diseases such as psoriasis<sup>154</sup> and in metastatic colon cancer<sup>155</sup>, melanoma<sup>156</sup> and breast and ovarian cancers<sup>157 157</sup>.

We therefore wondered what the link might be between CLDN1 and KRT76? Utilising a recombinant tail domain of KRT76 we were able to pull down endogenous CLDN1 from skin samples, indicating that the two proteins interact. At present we do not know whether this interaction is direct, or whether the proteins exist in a larger complex. In either case, KRT76 is clearly required for normal tight junction function and for CLDN1 localisation. Although links between the tight junction and the cytoskeleton have been described for actin, this is the first report detailing an interaction with the keratin intermediate filaments.

Many of the hallmarks of psoriasis (eg. hyperproliferation, expanded expression of FAPB5, and barrier dysfunction) are recapitulated in the *Krt76* homozygotes. Our findings therefore may help to unravel the mechanisms behind TJ-IF mediated pathogenesis of psoriasis. Importantly, we establish a critical and unexpected role for KRT76 in regulating the skin stability. Although the exact mechanism by which KRT76 mediates CLDN1 localisation remains to be determined, our study identifies a gene required for the postnatal maintenance of skin integrity with potential relevance to the development of skin disease.

## **Materials and Methods**

### *Transgenic mice.*

Mice were generated in the Mouse Genetics Programme at the Wellcome Trust Sanger Institute<sup>17</sup>. Animals were bred and maintained on a mixed background of C57BL/6J-TyrcBrd, C57BL/6N.

### *Histological Preparation and Staining*

Staining for  $\beta$ -gal expression was performed as previously described<sup>158</sup> on frozen sections and counterstained with Nile Red. Immunofluorescence experiments were performed after citrate based antigen retrieval followed by blocking with M.O.M. (Vector Labs). Primary antibodies were ZO-1 (Invitrogen cat# 339100), Occludin (BD Transduction cat# 611090), Claudin-1 (ABCAM cat# ab15098), cytokeratin14 - LLOO2 (ABCAM cat # ab7800) keratin10 (Covance PRB-159P) keratin6 (Covance cat # PRB-169P), Ecadherin

(Invitrogen) phospho-histoneH3 (Cell Signalling). All secondary antibodies were Alexa conjugated (Invitrogen).

#### *Imaging and analysis*

Sections were imaged using confocal (Lecia SP5 5 Channel confocal microscope). Bright field images of wound healing experiments were taken with Olympus dotslide Brightfield microscope. Images for CLE assays were acquired with Olympus CKX41 and exported to FIJI software for cell analysis.

#### *Wound healing Experiments*

Mice (age-matched males; 6 weeks) were isoflurane-anaesthetized and 2 full-thickness excisional wounds were made with a 5 mm biopsy punch (Livingstone International). Wound tissue was harvested with an 8mm biopsy punch.

#### *qPCR*

One µg of DNase (Ambion) treated RNA was used for cDNA synthesis (SuperScript® VILO™). Multiplex quantitative PCR was performed using Taqman probes for *Gapdh* (VIC-primer limited labelled, cat# 4448484) and *Krt76* (FAM labelled, Cat#4351372) with TaqMan® Fast Advanced Master Mix Protocol (PN 4444605B). Gene specific primers were designed and used in conjunction with SYBR® Green PCR Master Mix (Applied Biosystems) for the detection and quantification of *Claudin1* (5'-ATTTCAAGTCTGGCGACATT-3' fwd, 5'-ACACCTCCCAGAAGGCAGAG-3- rev) , *Krt6b* (5'-CAGACCCCAGATACCCTGGC-3' fwd, 5'-GAGCAGAGATGGCATCATGTGAGCAACAGG-3' rev) , *Krt16* (5'-AACAGCCTAGAAGAGACCAAAGGC-3' fwd, 5'-GGTAGGGGAGACAGATGGGGAATGCGC-3' rev) mRNA as compared to *Gapdh* (5'- CTGCACCACCAACTGCTTAG-3' fwd, 5'- GTCTTCTGGGTGGCAGTGAT-3' rev).

#### *Protein Fractionation*

All fractionation experiments were performed on epidermis of P3 animals. Pups were euthanized (Pentobarbital) and skin was removed as previously described <sup>159</sup>. Skins were floated on 2.3U/mL Dispase (Gibco® Life Technologies) in PBS overnight at 4°C. The epidermis was separated from the dermis and protein fractionated using a Qproteome Cell

Compartment Kit (Qiagen). Western blots for E-cadherin (Invitrogen) and total Histone H3 (Cell Signaling) were performed on the nuclear and membrane fractions. Image Quant software was used to calculate densitometry and quantify protein levels. Claudin-1 levels in the membrane fraction were normalized to E-cadherin for each sample.

#### *Biotin tracer assays*

TJ permeability assays were undertaken as previously described<sup>64,160</sup>. Briefly, a solution of 10 mg/ml EZ-Link™ Sulfo-NHS-LC-Biotin (Pierce) in PBS containing 1 mM CaCl<sub>2</sub> was injected into the paw pads of P3 pups. Paw pads were incubated at room temperature for 30 minutes prior to frozen sectioning and IHC with conjugated Streptavidin antibody (Alexafluor).

#### *Ultrastructural analysis*

Tissue was fixed in Karnovsky's fixative (2% paraformaldehyde, 2.5% glutaraldehyde in 0.1 M Cacodylate buffer) for 2 hours. Then washed in 3 X 10min changes of 0.1 M Cacodylate buffer. Post-fixation was with 2% osmium tetroxide in 0.1 M Cacodylate buffer followed by dehydration through a graded series of alcohols, two acetone rinses and embedding in Spurr's resin. 80nm sections were cut with a diamond knife (Diatome, Switzerland) on an Ultracut-S ultramicrotome (Leica, Mannheim, Germany) and contrasted with uranyl acetate and lead citrate. Images were captured with a Megaview II cooled CCD camera (Soft Imaging Solutions, Olympus, Australia) in a JEOL 1011 transmission electron microscope.

#### *His pull-down from mammalian lysates*

Recombinant HIS-tagged proteins were produced by IPTG induction (0.4 mM) of T7 Express lysY/I<sup>q</sup> Competent *E.coli* (New England Biolabs C3013I) transformed with HIS-tag expressing control vector, pET-30a+ (Novagen) or HIS-tagged KRT76 domains in pDEST17 gateway backbone (Invitrogen), grown for 6-8 hours at 37°C in low salt LB, supplemented with 100 µg/ml ampicillin or 50 µg/mL kanamycin (as required). Recombinant protein was purified using 0.1 ml per 1 ml of culture of PopCulture lysis reagent (Novagen), 1 µl per mL of culture of 40U/ml of Lysonase bioprocessing reagent (Novagen), protease inhibitors (Sigma P8849), and His-Mag beads (Novagen) according to manufacturer's protocols. Bound recombinant HIS and HIS-KRT76 protein were

washed and stored at 4°C as a 1:2 resin slurry in Tris-saline pH 7.4 containing protease inhibitors. Paw pad skin of adult was collected in RIPA lysis buffer and incubated with HIS or HIS-KRT76 overnight at 4°C. HisMag bead-bound HIS and HIS-KRT76 + lysates were then washed four times in Tris-saline pH 7.4 including 1% Triton X-100 and immunoblotted for CLDN1 (Santa Cruz Biotechnology, clone XX7 ) and the HIS tag (Sigma-Aldrich, clone HIS-1).

#### *Statistical Analysis*

Statistical analysis was performed using unpaired students test, values of  $p < 0.05$  were deemed significant. A minimum of 3 mice were analysed per condition. Error bars represent Standard Error of the Mean (S.E.M).

#### *Cornified envelope assay*

Analysis of the size of corneocytes in the cornified lipid envelope (CLE) assay was performed as previously published <sup>116</sup>.

#### **Acknowledgements**

This work was supported by an Australian Research Council Discovery Project Grant to I.S. (DP1092723). I.S. acknowledges support from the NH&MRC and ARC Fellowship Schemes (R.Douglas Wright and Future Fellowship) and T.D. an International Studentship from Monash University. We thank the Monash Microimaging Platform for assistance in imaging. We thank the Wellcome Trust Sanger Institute Mouse Genetics Project (Sanger MGP) and its funders for providing the mutant mouse line (Krt76<sup>tm1a(KOMP)Wtsi</sup>). Funding and associated primary phenotypic information may be found at [www.sanger.ac.uk/mouseportal](http://www.sanger.ac.uk/mouseportal).

## CHAPTER 5: DISCUSSION

Using a high throughput reverse genetics mouse mutagenesis screen as a platform, this study identified 23 genes necessary for the normal development or function of the skin. As expected, some of these genes and associated skin phenotypes had been previously published, although none had been targeted in the knockout first conditional strategy used by the WTSI. The generation of alternate alleles independently confirms the initial findings regarding these genes' functions and also validates the accuracy (mutagenesis and phenotyping) of the screen itself. Nearly 600 different lines were screened for cutaneous dysfunction so, less than 4% of the genes appeared to play a role in skin biology. The anticipated percentage was slightly higher at ~10%, based on mining the previously published phenotypes in mutant mice. However, because the estimation was based only on published work, it can be expected that negative data was not published. Furthermore, some of the knockouts that would have been included in this estimation would be skin specific conditional knockouts, thereby increasing the skin associated phenotypes in the previously published literature. This screen is different, as all data, negative or not, is represented for every line.

Is the lack of published negative data the only reason for the discrepancy between the predicted and actual number of skin associated genes? It is generally accepted that gene targeting is a more effective way of ensuring total knockout of gene function. This is because of the components of the targeting constructs themselves, as well as the directed placement of the constructs into a specific location in the genome are designed to ensure total gene knockout. Early in the screen, there was a strong belief in the robustness of the *En2a* splice acceptor to prevent alternative splice isoforms of a gene and thereby inhibit partially functional transcripts and protein fragments. Because of the confidence in the *En2a* splice acceptor, validation of the gene trap and its position was done using both short and long range PCR. Theoretically, this approach is infallible, but experimental data taken from a small subset of alleles generated at the Sanger-MGP suggests not. To assess transcript levels on a select set of genes, the Sanger-MGP undertook rt-PCR on liver samples from 25 established lines, identifying 4 alleles that showed 'leakiness' as evidenced by ~20% of remaining mRNA transcript<sup>19</sup>. 'Leakiness' or splicing around the cassette, is a contributor to hypomorphic alleles. This could be a confounding factor in the low phenotypic hit rate. As indicated in chapter 4, this was a factor that was analysed in

depth with the generation of multiple *Krt76<sup>tml</sup>* alleles and subsequent comparison to the *Krt76<sup>tmla</sup>* allele.

As discussed, the *Krt76* mice have a strong, reproducible, fully penetrant phenotype. Further, LacZ expression can be detected, indicating that the splice acceptor was at least partially functional. Taken together, the properly functioning reporter, the fully penetrant phenotype, expected expression in the oral epithelium, and the successful sr and lr-PCRs indicated a high degree if not total knockout of gene function. However, subsequent rt-PCR analysis of *Krt76* mice also showed minor production of 3' mRNA transcript, indicating the potential production of full length KRT76 and suggesting the possibility that these mice are slightly hypomorphic in phenotype. Thus, it was important to further validate the model and confirm that the phenotype in the *Krt76* mice was the consequence of *Krt76* knockout and not due to alterations in splice isoform ratios or an accumulation of toxic products introduced by the targeting cassette.

Ideally, this validation would be done using antibodies to the protein; however, this was not possible due to crossreactivity of existing KRT76 antibodies to other keratins (namely KRT3). The main concern was that the allele was hypomorphic. As there is no phenotype detected in the heterozygous mice this effectively, rules out the possibility of a dominant negative or hypermorphic or toxic phenotype arising from excess N-terminally neo-LacZ fused protein. Another concern was that the large targeting cassette might be occupying a regulatory region for a neighbouring gene. So, to validate the phenotype, CRE and FLP mediated recombination was used generate mouse lines with the *tml* allele in which the critical exon and the reporter cassette are excised. Two *tml* alleles were generated, one a global, ubiquitously expressed *cmv* driven cre and another tamoxifen inducible *ck14* driven cre. The former was done to confirm that global loss of the critical exon was the same in both lines and ruled out LacZ or neomycin toxicity effect from the cassette. The latter, was necessary to confirm the skin barrier defect as the principle cause of death.

Preliminary data suggest that the global *cmv cre*- driven *tml* allele recapitulates the phenotype observed in the *tmla* mice (with 3 pups dying at between 2-10 weeks of age) however, the extent of validation has yet to be completed due to insufficient numbers at this time. However, the tamoxifen induced *Krt76<sup>tmlck14cre+</sup>* line was tested entirely and showed a stronger phenotype. This is likely the result of acetone exposure, as acetone is



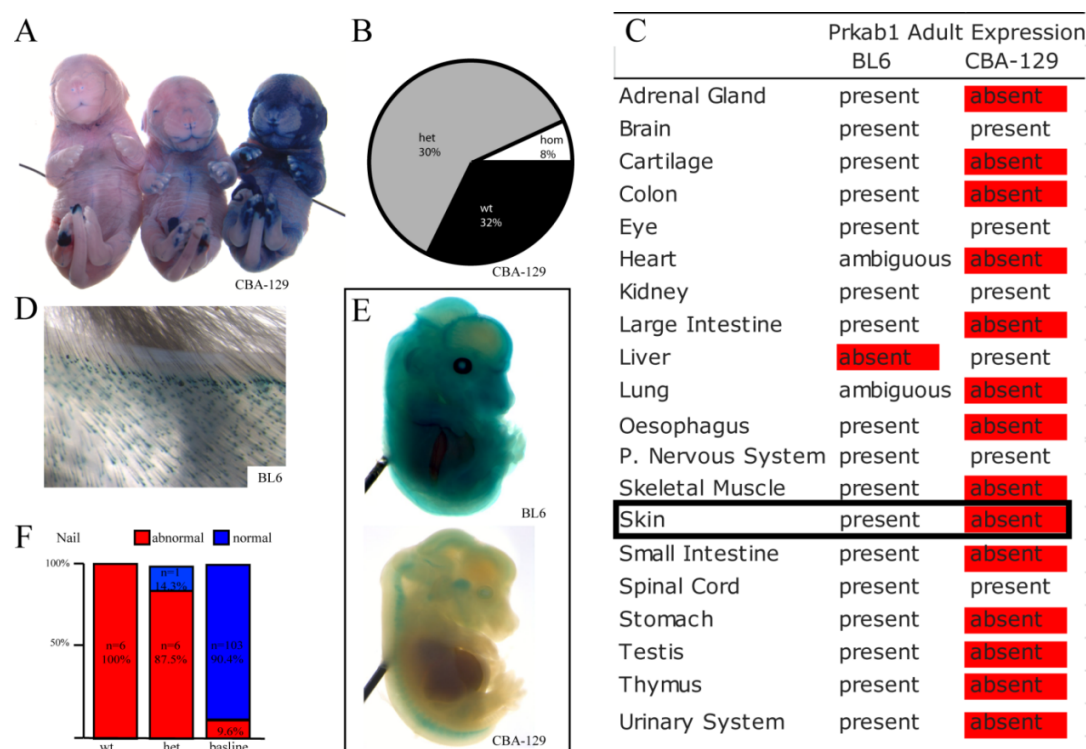
used to disrupt the skin barrier and deliver tamoxifen to the basal layers skin. Repeated exposure to tamoxifen causes a modest skin thickening reaction as a background defect and confounding the phenotype. This was indeed observed in WT tamoxifen-treated mice, however the recombined *tml1d* mice still showed the expected changes relative to this new baseline. Issues with skin damage associated with tamoxifen in acetone have not previously been reported. This study shows that the correct controls, especially vehicle only, are of paramount importance when characterizing skin phenotypes with the topical administration of tamoxifen. An alternative solution to circumvent the issues with topical acetone based administration of tamoxifen is intravenous injection of tamoxifen. Still, systemic exposure to tamoxifen might cause other, complicating phenotypes related to loss of gene function in other tissues, such as the oral epithelium, thereby not distinguishing this allele from the other *CMV-tml1d* variant.

Although the *tml1a* phenotype was validated as complete knockout, the question of how *Krt76* transcript is detected with rtPCR remains. The most obvious explanation is promiscuous primer binding elsewhere in the genome. An alternative explanation for the detection of this transcript is the existence of pseudogenes for this protein. Supporting this hypothesis is a published study screening the genome sequence for genes encoding keratins. This screen revealed 208 keratin-related DNA sequences across the genome, for which only 49 sequences reflect true keratin genes, making the majority inactive gene fragments and pseudogenes<sup>161</sup>. For the case of *Krt76*, the correct targeting was confirmed. However, most alleles have not received this rigorous interrogation. As such, we cannot completely eliminate incomplete gene disruption as a factor in low phenotype hits for the overall mutagenesis screen.

One of the aims of incorporating the skin specific screen was to increase the sensitivity of phenotypic tests to detect more skin phenotypes. Because the skin is required for postnatal survival, skin specific screens on the recessive lethal lines were implemented as pilot screens. The hypothesis was that by implementing recessive lethal screens for skin defects, we might identify genes whose dysfunction was tolerated until birth, but not after. As the skin is critical in the transition from life *un utero*, implementing skin specific screens on late lethal lines might identify more skin specific genes. To investigate this, I implemented a pilot recessive lethal screen for skin barrier defects, specifically.

Indeed, logistical issues with recessive lethal screening were one of the major drawbacks in this study. A pilot recessive lethal screen was undertaken in collaboration with the Sanger-MGP to identify skin developmental defects. Generally there are 3 broad developmental stage categories: 1) early lethals with pre- or peri-implantation defects, 2) midgestation lethals which die during organogenesis and 3) late lethals which die just prior to or immediately after birth. Late lethal lines are a good target for skin specific tests as they are often the result of failures in transition to the extra-uterine environment, a transition for which the skin is crucial. In an effort to address skin function in late recessive lethal lines, I implemented a barrier function test. In this screen, I was unable to detect a novel gene causative for barrier defect or any other organ anomaly. This is partially a reflection of the logistical difficulties associated with carrying out a recessive lethal screen. Despite the extensive facilities at the WTSI, there simply was not enough space or staff to assess each lethal line in depth. Firstly, assessment of the copulatory plug and subsequent separation of the mating pairs requires extensive technical expertise and time. Furthermore, the requirement for timed matings between heterozygous mice limited production of embryos, especially homozygous embryos as only  $\frac{1}{4}$  of embryos generated will be homozygous. In cases where embryos were generated at the appropriate stage, they underwent phenotypic testing before genotype was confirmed. So there was no way to know if the phenotyping was being completed on wildtype, heterozygous or homozygous animals. Often, entire litters would be assessed that did not contain a homozygous embryo. After screening 10 late lethal lines for 1 year, I was unable to generate enough homozygous animals to make any meaningful conclusions about skin or any other organ defect. This indicates that more rigorous tests, apart from necropsies, need to be done for embryonic phenotyping. It also highlights the complexity in undertaking these screens, particularly with having the facility and technical support required to generate large numbers of embryos at specific developmental stages.

Some of the problems associated with embryonic phenotyping that I encountered, are exemplified in the *Prkab1* line. *Prkab1* homozygotes appear at normal Mendelian ratios at ED14, but are found to be subviable at weaning (8% of animals at weaning are homozygous for *Prkab1* mutation). Furthermore, heterozygous males demonstrated an integument phenotype, exhibiting abnormal nail morphology. I chose to test this line for barrier function defects as the cause of neonatal lethality was unknown and could potentially be related to barrier dysfunction.



**Figure 5. *Prkab1* alleles on different genetic backgrounds show contradictory phenotypes.** Embryos at ED18.5 demonstrate varying degrees of skin barrier dysfunction (a) and postnatal subviability (n=438, 139 +/+, 263 +/-, 36 -/-) (b). LacZ staining differs in adult organs (c), the skin (d) and at ED14.5 (e). Controls and heterozygous mice on the CBA-129 show abnormal nail morphology (f).

Animals from this line showed varying levels of barrier dysfunction as demonstrated by dye penetration assays. Homozygous, heterozygous, and wildtype animals for *Prkab1* all came up positive for barrier dysfunction. However, the phenotype was not fully penetrant, robust nor reproducible. Furthermore, for an unknown reason, genotyping performed at the WTSI for this allele was unreliable causing concern about the effectiveness of the gene trap. The likely cause of all of the embryonic and genotyping problems for this allele was the background strain on which it was bred. Instead of the standard BL6 background, these mice were generated on a CBA-129 background. The 129 strain specifically has demonstrated negative impact on aspects of targeting technology<sup>162</sup>. The 129 strain has a complex history, with at least 3 instances of deliberate outcrossing and likely many more instances of accidental outcrossing. This complex genealogy has been further demonstrated using skin graft analysis. Skin grafts can be used as a measure of genetic disparity as the acceptance of a skin graft is highly dependent on homogeneity between

the donor and the recipient<sup>163</sup>. Skin grafts are rejected in 129 strains to a degree significantly higher than that found in other backgrounds such as the BL6<sup>162</sup>. Finally, the genetic identity of any 129 strain as a result of gene targeting is highly influenced by matching the ES cells origin with the parent strain. Indeed, Simpson et al show that for most pairings, animals will be segregating at a number of different loci, ultimately leading to a combination of alleles which may induce phenotypes, both independent of, or interacting with the targeted allele<sup>162</sup>.

However, to address potential background issues, a duplicate of this allele was generated on the standard BL6 background and compared to the allele on the CBA-129 background. This new line did not exhibit lethality or subviability. Furthermore, the same targeting construct was used on each strain, indicating that the phenotypic difference between the 2 strains is background dependant and can be explained by undiscovered background-unique modifiers. This comparison also showed profound differences in both adult and embryonic expression patterns of this gene. Taken together, this work stresses the importance of genetic background and highlights the problems with using the CBA-129 in genetic assays. Thus, concerns about the targeting strategy, genetic background and partially penetrant phenotype were a deterrent from further pursuing this strain, despite it being the only novel gene which showed barrier defects.

In parallel with collecting the screen data, I started my first in depth investigation into genetics and skin function at the Sanger-MGP. *Mad2*, a crucial gene whose expression is required for mitosis, was conditionally inactivated in the skin. Consistent with previous reports, significant aneuploidy was observed with loss of *Mad2* but surprisingly, the HF stem cells were more sensitive to this event than IFE cells. Before the generation of this knockout, the general dogma is that *Mad2* deficient/aneuploid cells cannot divide. Supporting this counter argument to our study are the difficulties associated with knocking out a gene completely. However, as opposed to *Krt76* no *Mad2* mRNA (or protein) could be detected in the epidermis beyond postnatal day 3, even though the epidermis can form adult skin. This indicates that the cells in the skin are incredibly tolerant to aneuploidy and in fact can still maintain a partially functional epidermis. If indeed aneuploid cells are unable to divide, the only way to explain our observations of a functional epidermis is that a small number of undetected *Mad2* proficient cells escaping *Cre* recombination go on to build a completely hairless skin that still bears the hallmark

of SAC deletion: aneuploidy. In this hypothesis an unidentified pool of IFE stem cells would have to escape *Mad2* deletion, recombine normally, and then become aneuploid. Currently, there are no specific, bonafide markers of IFE stem cells, making the counter argument both unlikely and untestable. Again, the doubts about complete *Mad2* inactivation go back to the general issues that are linked to gene targeting, even though the appropriate validation of *Mad2* deletion was done.

The phenotypic differences between the IFE and the HF could simply be due to the relative differences in proliferative behaviour in these two compartments at the time of *Mad2* recombination. The simplest explanation is that mammals can survive without hair, but not without the barrier provided by the IFE. Thus, the IFE may have evolved to accommodate more stress response than the HF. Still, there is no biological mechanism explaining the response of the IFE to *Mad2* loss compared to the HF lineage. One way to begin to elucidate this mechanism would be to take the HF and IFE separated FACS cell populations and further separate them to independently assess for proliferation (BrdU incorporation), cell cycle profile, and apoptosis (annexin V). Further, these separated populations could be analysed for colony forming efficiency and karyotyping/chromosome spreads. This separation, while sounding simple, is in fact quite complex. It has been done previously<sup>110</sup>, but the extreme incidence of aneuploid cells made it difficult to isolate enough cells for the initial separation of HF and IFE cells, thus making sub-separation and culturing extremely complex. Even if this analysis were successful, there was no real way to quantify the actual number of cell divisions that the individual cells underwent for each population and to specifically quantify the extent of aneuploidy. As previously discussed, using *in vitro* systems to investigate aneuploidy is challenging because of aneuploidy induced apoptosis. However, the use of matrigel and skin organ culture and CRE induced inactivation of *Mad2* combined with interphase FISH might be a way to capture the transition of cells from normal to aneuploidy in real time.

Although *Krt76* and *Mad2* disruption impact the skin via completely different mechanisms and affect distinct cellular processes, both alleles lend a few common lessons about the use of targeted skin specific mutagenesis. For both genes, *in vitro* systems are notoriously difficult to work with. As previously discussed, aneuploid cells do not survive in culture making the overexpression of *MAD2* and subsequent analysis difficult. *KRT76*

was impossible to overexpress in mammalian systems as attempts resulted in misfolded recombinant protein accumulation in the golgi. Indeed, the entire intermediate filament family are notoriously difficult to work with, *in vitro*, especially because of problems with insolubility. For these reasons, bacterial systems were used to overexpress the domains of KRT76 as these systems allow for protein production on a larger scale. Even with the bacterial system, there were still issues with insolubility of the KRT76 domains. However, for the head and tail domains, large enough cultures yielded sufficient amounts of recombinant protein. This protein product was then used to pull down CLDN1 from pawpad protein lysate successfully.

Another commonality between *Mad2* and *Krt76* is the lack of reliable antibodies to the proteins. For MAD2, one antibody worked for western blot, but was not specific enough for immunofluorescence. Instead, a CRE antibody, that was also unreliable (and has since been withdrawn from sale), was used to try to identify the specific cells that had recombined to lose *Mad2* function. Again, this method did not work well, and as discussed, posed a major problem with quantifying the number of cells that were *Mad2* deficient and aneuploid. Similar issues arose with antibody specificity to KRT76. Ideally, the pull downs using bacterial synthesized recombinant protein could have been replaced with co-IP's using skin lysate. However, the KRT76 antibody did not work on western blot and cross reacted with other keratins, making it unsuitable for this application.

## CHAPTER 6: CONCLUSIONS

Both *Krt76* and *Mad2* mouse based studies underscore the importance of generating mutant mice to study gene function. In both cases, *in vitro* methods fell short in providing the experimental opportunities required to identify the mechanisms by which these genes function. Knowledge of the consequences of gene inactivation in the mouse is important to our understanding of normal gene function and is lacking for many genes. Despite problems with targeting and validation, this thesis maintains that studying knockout mice is the best way to elucidate mammalian gene function *in vivo*. The novel genes described in this study as well as the known genes with novel roles in skin biology have been tested to different degrees, providing more information on roles of the genes themselves, the mechanisms of skin function, and the ability of the screen to produce reliable knockouts.

Gene driven screens provide a new way to solve the problem of systematic annotation of gene function. The benefits of functionally annotating the genome through such screens include the generation of targeted BAC libraries, extensive targeted ES cell banks and, as highlighted in this thesis, the association of gene disruption with a phenotype. Identifying the role of each gene in normal development and function through its disruption in the mouse has the secondary outcome of uncovering a diverse range of models for genetically linked human diseases. Through characterization of genetically mutant mice, the reverse genetic strategy discussed in this thesis allows for unbiased gene driven research on disease. Most of the knowledge we have about the genetic underpinnings of human disease comes from loss of function scenarios, be it directed or spontaneous genetic disruptions. With advances in gene targeting, specific genes can now be systematically targeted to maximize genome coverage and provide new models for disease. On a broad scale, this study retains the power of phenotype driven screens, and indeed, has recapitulated some of the results derived from earlier forward mutagenesis approaches. This study has also identified novel roles for known genes, especially those in skin biology, on which it was focused. Finally, my work highlights novel genetic contributors to skin function, showing deleterious developmental and functional consequences of gene inactivation in the skin, and providing, for the first time, characterization of cutaneous phenotypes derived from a high throughput gene driven screen.





## CHAPTER 7: EXTENDED MATERIALS AND METHODS

### 7.1 STANDARD MOLECULAR TECHNIQUES

#### 7.1.1 Bacterial cell maintenance

Bacterial cells were routinely maintained in Luria-Bertani (LB) broth and LB agar plates. LB broth was made of 10g/L tryptone, 5g/L yeast extract and 10g/L NaCl in H<sub>2</sub>O. LB agar plates contained the same as LB broth, with the addition of 1.5% agarose I (Amresco). For antibiotic selection, ampicillin was added to a final concentration of 100µg/ml.

#### 7.1.2 Cloning of DNA fragments

PCR amplified DNA fragments were cloned using the TOPO® cloning into a Gateway® Entry vectors, according to the manufacturer's instructions. L/R recombination into bacterial and mammalian expression vectors was performed as per manufacturer's instruction.

#### 7.1.3 Transformation of bacterial cells

DNA was transformed into chemically competent NEB 5-alpha Competent *E. coli* - High Efficiency (New England Biolabs C2987; NEB). Approximately 100ng of plasmid DNA was incubated in 50ul of competent cells and incubated on ice for 30 min. Cells were heat-shocked at 42°C for 45 seconds and then allowed to recover by adding 250µl LB medium and incubating for 30 min at 37°C, shaking. 100 µl of the mixture was plated on a LB agar plate with appropriate antibiotics and incubated overnight at 37°C.

#### 7.1.4 Purification of DNA from bacterial cells

Colonies were picked and incubated overnight at 37°C in the appropriate antibiotic. Plasmid DNA was isolated from bacteria with the Qiagen Plasmid Mini/Midi kits as per manufacturer's instructions.

#### 7.1.5 Preparation of genomic DNA for genotyping

Tail clips from mice were incubated in 75µl Lysis Buffer (25 mM NaOH, 0.2 mM EDTA) at 95°C, with occasional vortexing, until dissolved (approximately 30-60 min). The lysis mix was then incubated on ice for 5 min and neutralised with 75µl

Neutralisation buffer (40 mM Tris-HCl). The final DNA mixtures were stored at 4°C until analysed by PCR. 0.5-1.5µl of this mix was used in subsequent PCR reactions.

#### **7.1.6 RNA isolation**

Total RNA was isolated using the RNAqueous Kit (Ambion) or Trizol (Invitrogen), as per the manufacturer's instructions.

#### **7.1.7 DNA and RNA quantification**

Purified DNA and RNA samples were quantified in a NanoDrop 2000 spectrophotometer (Thermo Scientific). The NanoDrop 2000 measures the amount of ultraviolet light absorbed by a sample at 260 nm and uses the Beer-Lambert law to calculate the amount of nucleic acid present in that sample.

#### **7.1.8 DNase treatment of RNA**

Total RNA samples were DNase treated with the TURBO DNA-free kit (Ambion). Briefly, 5µg of RNA was mixed with 2.5µl 10x TURBO DNase buffer, 0.5µl TURBO DNase and RNase-free H<sub>2</sub>O to the final volume of 25µl. The reaction was propagated for 40 min at 37°C and stopped by incubation with 2.5µl DNase Inactivation reagent for 5 min at room temperature. The DNase Inactivation reagent was removed by centrifugation for 1.5 min at 10 000 x g.

#### **7.1.9 cDNA synthesis**

First-strand synthesis was performed by mixing 1µg of DNase treated RNA, 0.5µg random primers (Promega), 20U RNase inhibitor (SUPERase In; Ambion) and nuclease-free H<sub>2</sub>O to a final volume of 14µl. The reaction was incubated at 70°C for 5 min and ice for 5 min. The following was then added: 5µl 5x M-MLV reaction buffer (Promega), 5µl 10 mM dNTPs (Promega) and 200 U M-MLV H<sup>-</sup> Reverse Transcriptase (Promega). The reaction was incubated 10 min at 22°C, 90 min at 42°C and 15 min at 70°C. cDNA synthesis was verified by PCR amplification of β-actin (Fw: 5'-ACCAACTGGGACGATATGGAGAAG-3'; Rv: 5'-TACGACCAGAGGCATACAGGGACA-3') or Gapdh (Fw: 5'-GCCAAAAGGGTCATCATCTC, Rv: 5'-GCATGGACTGTGGTCATGAG-3')

#### **7.1.10 Amplification of DNA by Polymerase Chain Reaction (PCR)**

A typical PCR reaction contained 0.2mM dNTPs (Promega), 1x Thermopol buffer (NEB), 0.2  $\mu$ M forward and reverse primer and 0.4U Taq polymerase (NEB). Reactions were prepared in 20 or 50  $\mu$ l reaction volumes with 0.5-1.5 $\mu$ l genomic DNA mix, 1-50 ng plasmid or 0.5-1 $\mu$ l cDNA. Reactions were cycled in a MasterCycler EP Gradient S model 5345 (Eppendorf). This PCR mix was used for all PCR amplifications, unless stated otherwise.

#### **7.1.11 PCR screening of ligations**

Recombinant clones were identified by subjecting white colonies to PCR. Single white colonies were picked with a sterile pipette tip, streaked on a LB agar plate with appropriate antibiotics and dipped into a PCR tube containing PCR mix. The LB agar plate was incubated at 37°C overnight. PCR reactions were cycled as following:

1x	2 min at 94°C
2x	20s at 94°C, 30s at 63°C, 40s at 72°C
2x	20s at 94°C, 30s at 61°C, 40s at 72°C
2x	20s at 94°C, 30s at 59°C, 40s at 72°C
2x	20s at 94°C, 30s at 57°C, 40s at 72°C
30x	20s at 94°C, 30s at 55°C, 40s at 72°C
1x	5 min at 72°C

Denaturation temperatures result in bacterial cell lysis allowing PCR amplification from plasmid DNA. The PCR mix contained M13 forward and reverse primers which bind DNA sequences in the pGEM T-easy vector flanking the cloning region. Blue colonies were used as a negative control. Bacterial colonies with detectable products were prepped from the LB agar plate as previously described.

#### **7.1.12 DNA sequencing**

Sequencing of PCR products required removal of excess nucleotides and primers. This was done by addition of 5U Exonuclease I (NEB) and 5U Antarctic Phosphatase (NEB) to each PCR sample and incubation for 30 min at 37°C, followed by 20 min at 80°C. DNA sequencing reactions contained 0.75x Thermopol buffer (NEB), 0.5 $\mu$ M forward or reverse primer and 1 $\mu$ l BigDye Premix v3.1 (Applied Biosystems). The amount of DNA

added was 1-2µl for PCR templates or 400 ng for plasmids. Total reaction volume was 20µl. Sequencing reactions were cycled as following:

1x      3 min at 95°C  
30x     30s at 95°C, 15s at 50°C, 4 min at 60°C

After cycling, the reactions were precipitated by incubation in 80µl of 80% isopropanol for 20 min at room temperature. Samples were then centrifuged at top speed for 20 min in a table-top centrifuge. Supernatants were discarded and samples were washed in 80µl of 80% isopropanol and centrifuged at top speed for 10 min in a table-top centrifuge. Supernatants were discarded and samples were finally air-dried for 10 min. Sequencing samples were run by Micromon at Monash University Clayton campus or the Australian Genome Research Facility (AGRF) at the University of Melbourne, both suited in VIC Australia.

#### **7.1.13 Restriction endonuclease digestion of DNA**

Restriction digests were performed according to the manufacturer's specifications (usually NEB).

#### **7.1.14 Separation of nucleic acid by gel-electrophoresis**

RNA and DNA fragments acquired through PCR or restriction endonuclease digests were separated by gel electrophoresis in 1-2% w/v 1x TAE agarose gels submerged in 1xTAE (40mM Tris, 20mM acetic acid, 1mM EDTA, pH 8.4). To visualise RNA/DNA fragments, SYBR Safe DNA Gel Stain (Invitrogen) was added to agarose gels 1:25 000 v/v. Prior to loading, RNA/DNA fragments were mixed with Orange G loading dye (10mM Tris-HCl pH 7.5, 60mM EDTA, 60% glycerol, 0.15% (w/v) orange G, 0.03% xylene cyanol FF) 1:5. Gels were subsequently run at 60-100 volts until sufficient separation was achieved. RNA/DNA fragments were imaged in an Image Station 4000 MM Pro (Kodak) under the SYBR Safe filter and settings.

#### **7.1.15 Protein quantification**

Protein samples were quantified using Detergent Compatible Protein Assay (DC assay; Bio-Rad) as per the manufacturer's instructions. This is a colorimetric assay based on Lowry's protein assay, compatible with a wide range of different detergents. Protein

samples were mixed with the DC assay reagents in a 96-well microplate (BD Falcon) and absorbance was measured at 750 nm in a FLUOstar Omega (BMG Labtech) microplate reader. With the inclusion of standard protein samples (suspended in the same buffer as samples), a standard curve was generated in order to calculate the protein concentration in samples using the Beer-Lambert law.

#### **7.1.16 Sodium dodecyl sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)**

To separate proteins, samples were run in Sodium dodecyl sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE). Equal amounts of protein lysates were mixed with 5x SDS loading dye (Thermo Scientific) and boiled for 5 min at 100°C. When proteins were eluted directly from magnetic beads, beads were mixed with 2xSDS loading dye and boiled for 15 min at 100°C. Samples were then loaded on polyacrylamide gels (4-20% Mini-Protean TGX precast gels; Bio-Rad) submerged in 1x SDS buffer (25mM Tris, 0.192M Glycine, 0.1% SDS). Gels were routinely run at 100V at 4°C until the dye front ran off. Gels were allowed to equilibrate in 1x transfer buffer (25mM Tris, 0.192M Glycine, 20% Methanol) for 15-30 min at 4°C following electrophoresis. Proteins were then wet transferred to polyvinylidene fluoride (PVDF) membrane (Millipore) in 1x transfer buffer at 250 mA for 1h at 4°C.

#### **7.1.17 Western blotting**

Following transfer of protein samples, membranes were blocked in 5% Bovine Serum Albumin (BSA) in Tris buffered saline (TBS-T; 50 mM Tris, 150 mM NaCl, 0.05% Tween-20, pH7.4) or in 5% non-fat milk in TBS-T, for 1h at room temperature. Primary antibodies were diluted in 1% of respective block and were allowed to react with membranes overnight at 4°C. Secondary antibodies conjugated with horseradish peroxidase (HRP) were diluted in TBS-T and incubated with membranes for 1h at room temperature. To minimise background, excess antibody was washed away by rocking membrane in TBS-T 5x5 min at room temperature. Protein bands were visualised by subjecting membranes to enhanced chemiluminescence as per the manufacturer's instructions (ECL plus; GE Healthcare) and exposing membranes to X-Ray film (Kodak or GE Healthcare).

#### **7.1.18 Recombinant Protein Expression: Bacterial Expression**

Recombinant HIS-tagged proteins were produced by IPTG induction (0.4 mM) of T7 Express lysY/I<sup>q</sup> Competent *E.coli* (New England Biolabs C3013I) transformed with HIS-tag expressing control vector, pET-30a+ (Novagen) or HIS-tagged KRT76 domains in pDEST17 gateway backbone (Invitrogen), grown for 6-8 hours at 37°C in low salt LB, supplemented with 100 µg/ml ampicillin or 50 µg/mL kanamycin (as required). Recombinant protein was purified using 0.1 ml per 1 ml of culture of PopCulture lysis reagent (Novagen), 1 µl per mL of culture of 40U/ml of Lysonase bioprocessing reagent (Novagen), protease inhibitors (Sigma P8849), and His-Mag beads (Novagen) according to manufacturer's protocols. Bound recombinant HIS and HIS-KRT76 protein were washed and stored at 4°C as a 1:2 resin slurry in Tris-saline pH 7.4 containing protease inhibitors. Paw pad skin of adult was collected in RIPA lysis buffer and incubated with HIS or HIS-KRT76 overnight at 4°C. HisMag bead-bound HIS and HIS-KRT76 + lysates were then washed four times in Tris-saline pH 7.4 including 1% Triton X-100 and immunoblotted for CLDN1 (Santa Cruz Biotechnology, clone XX7 ) and the HIS tag (Sigma-Aldrich, clone HIS-1).

#### **7.1.19 Recombinant Protein Expression: Mammalian Expression**

A host of different mammalian cell lines were used to attempt over expression of Krt76 recombinant tagged protein including Human Embryonic Kidney 293 (HEK293), immortal human keratinocytes (HaCaT), CV-1 (simian) in Origin, and carrying the SV40 genetic material (COS), lung epithelial cell lines (HPL1), and an immortalized fibroblast (3T3) cells. Cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% FBS (Thermofisher Scientific), 100U/ml penicillin (Invitrogen) and 100µg/ml streptomycin (Invitrogen). The cells were grown at 37°C in a humidified incubator containing 5% CO<sub>2</sub>. To passage cells, 0.25% Trypsin (Invitrogen) was used. Cell stocks were kept in a mix of 10% DMSO (Merck) and 90% FBS.

For transfection, cells were grown in 15cm<sup>2</sup> tissue culture dishes and transiently transfected at approximately 70% confluency with. This was performed in the presence of Fugene HD (Roche) as per the manufacturer's instructions. Each transfection reaction contained 15µg of DNA and 45µl Fugene (DNA:Fugene ratio of 1:3). Successful

transfections were verified by RT-PCR or by performing immunocytochemistry for the vector tag (HA tag).

#### **7.1.20 Immunofluorescence (IF)**

Tissues were collected and washed in PBS and fixed in 4% PFA for 4 hrs at room temperature, washed 3xPBS and stored in 70% Ethanol until processing. Tissues were processed into paraffin using Leica Peloris II tissue processor. Paraffin processed tissues were embedded into paraffin blocks and histological sections were cut at 10-12  $\mu$ m with a Microtec x1/Microm x1/Reichert Jung x1 microtome and mounted onto Superfrost slides (Thermo Scientific). For antibody staining, slides were baked at 60°C for 10 minutes and then taken through 3 x xylene washes, 3x ethanol, and into water. Antibodies and concentrations are noted in manuscripts. All slides were mounted with Mowiol supplemented with 0.2% DABCO. Cells were imaged by confocal microscopy using Leica SP5 confocal microscope at Monash Micro Imaging.

#### **7.1.21 Statistical analysis**

Statistical analyses were performed using the two-tailed Student's *t*-test or two-way ANOVA. P-values of  $\leq 0.05$  were considered significant. Graphs display the Mean value  $\pm$  Standard Error of the Mean (SEM).

#### **7.1.22 Animal housing**

The offspring was maintained at the Animal Research Laboratories (ARL) at Monash University, Clayton campus, under Specific Pathogen Free (SPF) conditions. The animals were maintained according to standard animal housing practices under the provisions of Monash University ethical clearance. Animal husbandry skills and general care was provided by the staff of the ARL.

### **7.2 GENERATION AND ANALYSIS OF TRANSGENIC MICE**

#### **7.2.1 Expression analysis by Reverse-Transcriptase PCR (RT-PCR)**

RNA samples were used as template in cDNA synthesis (described above). Primers against the 3' end of the open reading frames of Krt76 were designed with the Primer3

software and are described in the below table.. Each PCR reaction contained 0.5µl cDNA and Krt76 specific primers. The PCR reactions were cycled as following:

1x	2 min at 94°C
2x	20s at 94°C, 30s at 63°C, 30s at 72°C
2x	20s at 94°C, 30s at 61°C, 30s at 72°C
2x	20s at 94°C, 30s at 59°C, 30s at 72°C
2x	20s at 94°C, 30s at 57°C, 30s at 72°C
35x	20s at 94°C, 30s at 55°C, 30s at 72°C
1x	5 min at 72°C

Equal amounts of PCR reactions were analysed by gel electrophoresis and visualised under SYBR safe settings as previously described.

### 7.2.2 Expression analysis by quantitative RT-PCR

Multiplex qPCRs using TaqMan probes (Invitrogen) were used to quantify Krt76 mRNA transcript. FAM labeled Krt76 ( probe: Mm01305200\_m1) designed to bind in exon2 and VIC labelled Gapdh (Mm99999915\_g1) as a reference gene, were used with Taqman Fast Advanced Mastermix (Invitrogen PN 4444556) were used in a single reaction containing 0.5µl cDNA as per manufacturer's instruction.

### 7.2.3 Genotyping Krt76 alleles

Using the Primer3 software, wildtype forward and reverse primers were designed to bind upstream and downstream of the cassette insertion site in intron 1-2 (K76\_Geno2\_F/R). The tm1a allele was detected using a forward primer was designed to bind the cassette (K76\_NeoF) using the WT reverse primer (K76\_Geno2\_R) for amplification. The tm1c allele was detected using primers on either side of the cassette in the 1-2 intron (K76FLP\_F/R) allowing the tm1c allele to be distinguished from the wildtype allele by a size difference resulting from the remaining frt site and LoxP sites. The tm1d allele could be detected by reusing forward primers from the WT or tm1c alleles in conjunction with a primer designed to bind in the 2-3 intron (downstream of the recombined exon) (K76\_3'arm). See below table for expected product sizes .



The offspring was genotyped with the following primers:

			expected product size			
	name	sequence	wt	tm1a	tm1c	tm1d
5'	K76_Geno2_F	agaaatctgcctcccagtg	458	np	531	np
3'	K76_Geno2_R	atagcagctgggtgtgtgctg				
5'	K76_Neo_F	tctcatgctggagtctctcg	np	602	np	np
3'	K76_Geno2_R	atagcagctgggtgtgtgctg				
5'	K76FLP_fwd	CCGTCCATCCTCTGACTGTT	250	np	361	np
3'	K76FLP_rev	TTGAGACTGCTGGTCCTCCT				
5'	K76_Geno2_F	agaaatctgcctcccagtg	1176	np	1333	428
3'	K76_3'arm	CCCAAACCCACTGCTTAAAA				
5'	K76FLP_fwd	CCGTCCATCCTCTGACTGTT	1054	np	1211	306
3'	K76_3'arm	CCCAAACCCACTGCTTAAAA				

The PCR reactions contained 0.2mM dNTPs (Promega), 1x Phire reaction buffer (Finnzymes), 0.5µM forward and reverse primer, 0.2µl Phire Hot Start DNA Polymerase (Finnzymes) and 1µl genomic DNA mix. The total reaction volume was 10µl and was cycled as following:

1x 30s at 98°C  
 36x 5s at 98°C, 5s at 60°C, 15s at 72°C  
 1x 1 min at 72°C

PCR reactions were analysed by gel electrophoresis and visualised under SYBR safe settings as previously described.

#### 7.2.4 LacZ staining of embryos and adult mice

E12.5 embryos mice were fixed in 4% PFA for 30 min at 4°C. Alternatively, adult mice were perfusion fixed in 4% PFA, followed by incubation in 4% PFA for 30 min at 4°C. Samples were placed on a rocker and washed 3x20 min in PBS at 4°C. Samples were placed in LacZ staining solution (0.01% Sodium Deoxycholate, 0.02% Igepal, 2mM MgCl<sub>2</sub>, 5mM Potassium Ferricyanide, 5mM Potassium Ferrocyanide, PBS) freshly supplemented with X-Gal (Promega) to a concentration of 1mg/ml. The reaction was allowed to proceed in the dark at room temperature/37°C until sufficient staining was detected. Samples were finally washed in PBS and fixed in 4% PFA.

### **7.2.5 Wound-healing experiment and macroscopic wound analysis**

Mice were wounded at P37 with 5 mm in diameter disposable biopsy punches (Livingstone). Wounds were imaged next to a ruler at day 0, 1, 3, 5, 7 and 10. Wounds were also collected at 0, 1, 3, 5, 7 and 10 using 8 mm disposable biopsy punches. The entry wounds were used for RNA and protein isolation and the exit wounds were used for histology. For histology, wounds were fixed overnight in 4% PFA at 4°C and washed in PBS 3x20 min at room temperature. Wounds were then divided in two by cutting across the widest part of the wound and placed in 70% ethanol. Wounds were finally dispatched to Monash Histology for processing into paraffin. Wound area measurements were performed in the image software Image J.

### **7.2.6 Re-epithelialisation analysis days post-wounding**

Paraffin sections (8µm) from the widest part of 2 day old wounds were collected and dispatched to Monash Histology for Hematoxylin and Eosin (H&E) staining. Sections were imaged in the Olympus DotSlide BX51 microscope system with the 20x objective. The degree of re-epithelialisation was measured in the DotSlide software. Re-epithelialisation was defined as the ratio of the [distance between the new epithelium and the wound edge] and the [total width of the wound].

### **7.2.7 Cornified Envelope isolation**

Analysis of the size of corneocytes in the cornified lipid envelope (CLE) assay was performed as previously published<sup>47</sup>.

### **7.2.8 Toluidine blue barrier function assays**

Dye exclusion assays were performed as previously published<sup>58</sup>. Briefly, ED18.5 embryos were harvested, killed by immersion in ice cold PBS and subjected to a methanol gradient. Embryos were then immersed in toluidine blue dye and destained with PBS.

### **7.2.9 Trans epidermal water loss (TEWL) assays**

Assays of epidermal barrier function were performed essentially as previously described<sup>58</sup>. TEWL assays were performed using dorsal skin samples excised from the lateral

thoracolumbar region of P3 pups. Skins were photographed with a Leica dissecting scope. camera mounted on a Zeiss Stemi microscope. Comparison of TEWL was made using logistic regression models. All other statistical analyses were performed using the statistical software package STATA Version 7 (Stata Corporation USA).

#### **7.2.10 Biotin tracer assays**

TJ permeability assays were undertaken as previously described<sup>64,160</sup>. Briefly, a solution of 10 mg/ml EZ-Link™ Sulfo-NHS-LC-Biotin (Pierce) in PBS containing 1 mM CaCl<sub>2</sub> was injected into the paw pads of P3 pups. Paw pads were incubated at room temperature for 30 minutes prior to frozen sectioning and IHC with conjugated Streptavidin antibody (Alexafluor).

#### **7.2.11 Protein Fractionation Assays**

All fractionation experiments were performed on epidermis of P3 animals. Pups were euthanized (Pentobarbital) and skin was removed as previously described<sup>159</sup>. Skins were floated on 2.3U/mL Dispase (Gibco® Life Technologies) in PBS overnight at 4°C. The epidermis was separated from the dermis and protein fractionated using a Qproteome Cell Compartment Kit (Qiagen). Western blots for E-cadherin (Invitrogen) and total Histone H3 (Cell Signaling) were performed on the nuclear and membrane fractions. Image Quant software was used to calculate densitometry and quantify protein levels. Claudin-1 levels in the membrane fraction were normalized to E-cadherin for each sample.



## References

- 1 Cuénot, L. Notes et revues. *Arch. Zool. Exp. Gen* (1902).
- 2 Paigen, K. One hundred years of mouse genetics: an intellectual history. I. The classical period (1902-1980). *Genetics* **163**, 1-7 (2003).
- 3 Nadeau, J. H. *et al.* Sequence interpretation. Functional annotation of mouse genome sequences. *Science (New York, N.Y.)* **291**, 1251-1255 (2001).
- 4 Waterston, R. H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-562, doi:10.1038/nature01262 (2002).
- 5 Dietrich, W. *et al.* A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**, 423-447 (1992).
- 6 Dietrich, W. F. *et al.* A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nature genetics* **7**, 220-245, doi:10.1038/ng0694supp-220 (1994).
- 7 Dietrich, W. F. *et al.* A comprehensive genetic map of the mouse genome. *Nature* **380**, 149-152, doi:10.1038/380149a0 (1996).
- 8 Gondo, Y. Trends in large-scale mouse mutagenesis: from genetics to functional genomics. *Nature reviews. Genetics* **9**, 803-810, doi:10.1038/nrg2431 (2008).
- 9 Hrabe de Angelis, M. & Balling, R. Large scale ENU screens in the mouse: genetics meets genomics. *Mutation research* **400**, 25-32 (1998).
- 10 Popp, R. A., Bailiff, E. G., Skow, L. C., Johnson, F. M. & Lewis, S. E. Analysis of a mouse alpha-globin gene mutation induced by ethylnitrosourea. *Genetics* **105**, 157-167 (1983).
- 11 Harbach, P. R., Filipunas, A. L., Wang, Y. & Aaron, C. S. DNA sequence analysis of spontaneous and N-ethyl-N-nitrosourea-induced hpert mutations arising in vivo in cynomolgus monkey T-lymphocytes. *Environmental and molecular mutagenesis* **20**, 96-105 (1992).
- 12 Noveroske, J. K., Weber, J. S. & Justice, M. J. The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mammalian genome : official journal of the International Mammalian Genome Society* **11**, 478-483 (2000).
- 13 Hitotsumachi, S., Carpenter, D. A. & Russell, W. L. Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 6619-6621 (1985).
- 14 Russell, W. L. *et al.* Dose--response curve for ethylnitrosourea-induced specific-locus mutations in mouse spermatogonia. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 3589-3591 (1982).
- 15 Russell, W. L. *et al.* Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 5818-5819 (1979).
- 16 van der Weyden, L., Adams, D. J. & Bradley, A. Tools for targeted manipulation of the mouse genome. *Physiological genomics* **11**, 133-164, doi:10.1152/physiolgenomics.00074.2002 (2002).
- 17 Skarnes, W. C. *et al.* A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337-342, doi:10.1038/nature10163 (2011).
- 18 Edwards, A. M. *et al.* Too many roads not taken. *Nature* **470**, 163-165, doi:10.1038/470163a (2011).

- 19 White, J. K. *et al.* Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* **154**, 452-464, doi:10.1016/j.cell.2013.06.022 (2013).
- 20 Testa, G. *et al.* A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *Genesis (New York, N.Y. : 2000)* **38**, 151-158, doi:10.1002/gene.20012 (2004).
- 21 Copeland, N. G., Jenkins, N. A. & Court, D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nature reviews. Genetics* **2**, 769-779, doi:10.1038/35093556 (2001).
- 22 Chan, W. *et al.* A recombineering based approach for high-throughput conditional knockout targeting vector construction. *Nucleic acids research* **35**, e64, doi:10.1093/nar/gkm163 (2007).
- 23 Murphy, K. C. Use of bacteriophage lambda recombination functions to promote gene replacement in Escherichia coli. *Journal of bacteriology* **180**, 2063-2071 (1998).
- 24 Court, D. L., Sawitzke, J. A. & Thomason, L. C. Genetic engineering using homologous recombination. *Annual review of genetics* **36**, 361-388, doi:10.1146/annurev.genet.36.061102.093104 (2002).
- 25 Lee, E. C. *et al.* A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56-65, doi:10.1006/geno.2000.6451 (2001).
- 26 Murphy, K. C. Lambda Gam protein inhibits the helicase and chi-stimulated recombination activities of Escherichia coli RecBCD enzyme. *Journal of bacteriology* **173**, 5808-5821 (1991).
- 27 Muyrers, J. P., Zhang, Y. & Stewart, A. F. Techniques: Recombinogenic engineering--new options for cloning and manipulating DNA. *Trends in biochemical sciences* **26**, 325-331 (2001).
- 28 Sternberg, N. & Hamilton, D. Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. *Journal of molecular biology* **150**, 467-486 (1981).
- 29 Sauer, B. Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae. *Molecular and cellular biology* **7**, 2087-2096 (1987).
- 30 Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 5166-5170 (1988).
- 31 Hoess, R. H., Wierzbicki, A. & Abremski, K. The role of the loxP spacer region in P1 site-specific recombination. *Nucleic acids research* **14**, 2287-2300 (1986).
- 32 Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156 (1981).
- 33 Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7634-7638 (1981).
- 34 Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255-256 (1984).
- 35 Pettitt, S. J. *et al.* Agouti C57BL/6N embryonic stem cells for mouse genetic resources. *Nature methods* **6**, 493-495, doi:10.1038/nmeth.1342 (2009).

- 36 Ayadi, A. *et al.* Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mammalian genome : official journal of the International Mammalian Genome Society* **23**, 600-610, doi:10.1007/s00335-012-9418-y (2012).
- 37 Wilson, S. I. *et al.* The status of Wnt signalling regulates neural and epidermal fates in the chick embryo. *Nature* **411**, 325-330, doi:10.1038/35077115 (2001).
- 38 Stern, C. D. Neural induction: old problem, new findings, yet more questions. *Development (Cambridge, England)* **132**, 2007-2021, doi:10.1242/dev.01794 (2005).
- 39 Hemmati-Brivanlou, A. & Melton, D. Vertebrate embryonic cells will become nerve cells unless told otherwise. *Cell* **88**, 13-17 (1997).
- 40 M'Boneko, V. & Merker, H. J. Development and morphology of the periderm of mouse embryos (days 9-12 of gestation). *Acta anatomica* **133**, 325-336 (1988).
- 41 Fuchs, E. & Green, H. Changes in keratin gene expression during terminal differentiation of the keratinocyte. *Cell* **19**, 1033-1042 (1980).
- 42 Cheng, X. & Koch, P. J. In vivo function of desmosomes. *The Journal of dermatology* **31**, 171-187 (2004).
- 43 Segre, J. A., Bauer, C. & Fuchs, E. Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nature genetics* **22**, 356-360, doi:10.1038/11926 (1999).
- 44 Ting, S. B. *et al.* The epidermis of grhl3-null mice displays altered lipid processing and cellular hyperproliferation. *Organogenesis* **2**, 33-35 (2005).
- 45 Fuchs, E. Scratching the surface of skin development. *Nature* **445**, 834-842, doi:10.1038/nature05659 (2007).
- 46 Koster, M. I. & Roop, D. R. Mechanisms regulating epithelial stratification. *Annual review of cell and developmental biology* **23**, 93-113, doi:10.1146/annurev.cellbio.23.090506.123357 (2007).
- 47 Colbert, M. C., McCoon, P. E., Day, K. H., Lane, A. T. & Goldsmith, L. A. Monoclonal antibodies to two different epitopes in a 30-kD CNBr peptide of the K1 and K2 keratins. *The Journal of investigative dermatology* **95**, 647-652 (1990).
- 48 Fine, J. D. *et al.* The classification of inherited epidermolysis bullosa (EB): Report of the Third International Consensus Meeting on Diagnosis and Classification of EB. *Journal of the American Academy of Dermatology* **58**, 931-950, doi:10.1016/j.jaad.2008.02.004 (2008).
- 49 Arin, M. J. The molecular basis of human keratin disorders. *Human genetics* **125**, 355-373, doi:10.1007/s00439-009-0646-5 (2009).
- 50 Steinert, P. M., Marekov, L. N., Fraser, R. D. & Parry, D. A. Keratin intermediate filament structure. Crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *Journal of molecular biology* **230**, 436-452, doi:10.1006/jmbi.1993.1161 (1993).
- 51 van Haelst, M. M. *et al.* Molecular study of 33 families with Fraser syndrome new data and mutation review. *American journal of medical genetics. Part A* **146A**, 2252-2257, doi:10.1002/ajmg.a.32440 (2008).
- 52 Slavotinek, A. M. & Tiffet, C. J. Fraser syndrome and cryptophthalmos: review of the diagnostic criteria and evidence for phenotypic modules in complex malformation syndromes. *Journal of medical genetics* **39**, 623-633 (2002).
- 53 Smyth, I. & Scambler, P. The genetics of Fraser syndrome and the blebs mouse mutants. *Human molecular genetics* **14 Spec No. 2**, R269-274, doi:10.1093/hmg/ddi262 (2005).

- 54 McGregor, L. *et al.* Fraser syndrome and mouse blebbed phenotype caused by mutations in FRAS1/Fras1 encoding a putative extracellular matrix protein. *Nature genetics* **34**, 203-208, doi:10.1038/ng1142 (2003).
- 55 Varnum, D. S. & Fox, S. C. Head blebs: a new mutation on chromosome 4 of the mouse. *The Journal of heredity* **72**, 293 (1981).
- 56 Vrontou, S. *et al.* Fras1 deficiency results in cryptophthalmos, renal agenesis and blebbed phenotype in mice. *Nature genetics* **34**, 209-214, doi:10.1038/ng1168 (2003).
- 57 Plagens, G. M. An embryological study of a special strain of deformed x-rayed mice, with special reference to the etiology and morphogenesis of the abnormalities. *Journal of Morphology* **55**, 151-183, doi:10.1002/jmor.1050550110 (1933).
- 58 Hardman, M. J., Sisi, P., Banbury, D. N. & Byrne, C. Patterned acquisition of skin barrier function during development. *Development (Cambridge, England)* **125**, 1541-1552 (1998).
- 59 Wilson, C. *et al.* Cells within the bulge region of mouse hair follicle transiently proliferate during early anagen: heterogeneity and functional differences of various hair cycles. *Differentiation; research in biological diversity* **55**, 127-136 (1994).
- 60 Kalinin, A. E., Kajava, A. V. & Steinert, P. M. Epithelial barrier function: assembly and structural features of the cornified cell envelope. *BioEssays : news and reviews in molecular, cellular and developmental biology* **24**, 789-800, doi:10.1002/bies.10144 (2002).
- 61 Candi, E., Schmidt, R. & Melino, G. The cornified envelope: a model of cell death in the skin. *Nature reviews. Molecular cell biology* **6**, 328-340, doi:10.1038/nrm1619 (2005).
- 62 Nemes, Z. & Steinert, P. M. Bricks and mortar of the epidermal barrier. *Experimental & molecular medicine* **31**, 5-19, doi:10.1038/emm.1999.2 (1999).
- 63 Oliveira, S. S. & Morgado-Diaz, J. A. Claudins: multifunctional players in epithelial tight junctions and their role in cancer. *Cell Mol Life Sci* **64**, 17-28 (2007).
- 64 Furuse, M. *et al.* Claudin-based tight junctions are crucial for the mammalian epidermal barrier: a lesson from claudin-1-deficient mice. *The Journal of cell biology* **156**, 1099-1111, doi:10.1083/jcb.200110122 (2002).
- 65 Fitch, K. R. *et al.* Genetics of dark skin in mice. *Genes & development* **17**, 214-228, doi:10.1101/gad.1023703 (2003).
- 66 Thomas, A. J. & Erickson, C. A. The making of a melanocyte: the specification of melanoblasts from the neural crest. *Pigment cell & melanoma research* **21**, 598-610, doi:10.1111/j.1755-148X.2008.00506.x (2008).
- 67 Mayer, T. C. The migratory pathway of neural crest cells into the skin of mouse embryos. *Developmental biology* **34**, 39-46 (1973).
- 68 Yoshida, H. *et al.* Neural and skin cell-specific expression pattern conferred by steel factor regulatory sequence in transgenic mice. *Developmental dynamics : an official publication of the American Association of Anatomists* **207**, 222-232, doi:10.1002/(sici)1097-0177(199610)207:2<222::aid-aja10>3.0.co;2-9 (1996).
- 69 Kelsh, R. N., Harris, M. L., Colanesi, S. & Erickson, C. A. Stripes and belly-spots -- a review of pigment cell morphogenesis in vertebrates. *Seminars in cell & developmental biology* **20**, 90-104, doi:10.1016/j.semcdb.2008.10.001 (2009).



- 70 Plikus, M. & Chuong, C. M. Making waves with hairs. *The Journal of investigative dermatology* **122**, vii-ix, doi:10.1111/j.0022-202X.2004.22436.x (2004).
- 71 Plikus, M. V. & Chuong, C. M. Complex hair cycle domain patterns and regenerative hair waves in living rodents. *The Journal of investigative dermatology* **128**, 1071-1080, doi:10.1038/sj.jid.5701180 (2008).
- 72 Stenn, K. S. & Paus, R. Controls of hair follicle cycling. *Physiological reviews* **81**, 449-494 (2001).
- 73 Cotsarelis, G., Sun, T. T. & Lavker, R. M. 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 (1990).
- 74 Braun, K. M. *et al.* Manipulation of stem cell proliferation and lineage commitment: visualisation of label-retaining cells in wholemounts of mouse epidermis. *Development (Cambridge, England)* **130**, 5241-5255, doi:10.1242/dev.00703 (2003).
- 75 Singer, A. J. & Clark, R. A. Cutaneous wound healing. *The New England journal of medicine* **341**, 738-746, doi:10.1056/NEJM199909023411006 (1999).
- 76 Reinke, J. M. & Sorg, H. Wound repair and regeneration. *European surgical research. Europäische chirurgische Forschung. Recherches chirurgicales europeennes* **49**, 35-43, doi:10.1159/000339613 (2012).
- 77 Hassold, T. *et al.* Human aneuploidy: incidence, origin, and etiology. *Environmental and molecular mutagenesis* **28**, 167-175, doi:10.1002/(sici)1098-2280(1996)28:3<167::aid-em2>3.0.co;2-b (1996).
- 78 Brown, S. Miscarriage and its associations. *Seminars in reproductive medicine* **26**, 391-400, doi:10.1055/s-0028-1087105 (2008).
- 79 Fuchs, E. & Weber, K. Intermediate filaments: structure, dynamics, function, and disease. *Annual review of biochemistry* **63**, 345-382, doi:10.1146/annurev.bi.63.070194.002021 (1994).
- 80 Omary, M. B., Ku, N. O., Tao, G. Z., Toivola, D. M. & Liao, J. "Heads and tails" of intermediate filament phosphorylation: multiple sites and functional insights. *Trends in biochemical sciences* **31**, 383-394, doi:10.1016/j.tibs.2006.05.008 (2006).
- 81 Parry, D. A. Hard alpha-keratin intermediate filaments: an alternative interpretation of the low-angle equatorial X-ray diffraction pattern, and the axial disposition of putative disulphide bonds in the intra- and inter-protofilamentous networks. *International journal of biological macromolecules* **19**, 45-50 (1996).
- 82 Roop, D. Defects in the barrier. *Science (New York, N.Y.)* **267**, 474-475 (1995).
- 83 Blanco, S. *et al.* The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. *PLoS genetics* **7**, e1002403, doi:10.1371/journal.pgen.1002403 (2011).
- 84 Smith, C. L., Goldsmith, C. A. & Eppig, J. T. The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome biology* **6**, R7, doi:10.1186/gb-2004-6-1-r7 (2005).
- 85 Bassett, J. H. *et al.* Rapid-throughput skeletal phenotyping of 100 knockout mice identifies 9 new genes that determine bone strength. *PLoS genetics* **8**, e1002858, doi:10.1371/journal.pgen.1002858 (2012).
- 86 Goldsmith, L. A. My organ is bigger than your organ. *Archives of dermatology* **126**, 301-302 (1990).
- 87 Chuong, C. M. *et al.* What is the 'true' function of skin? *Experimental dermatology* **11**, 159-187 (2002).

- 88 St Sauver, J. L. *et al.* Why patients visit their doctors: assessing the most prevalent conditions in a defined American population. *Mayo Clinic proceedings* **88**, 56-67, doi:10.1016/j.mayocp.2012.08.020 (2013).
- 89 Dyer, J. A. Practice gaps. Propranolol to treat hemangiomas of infancy: safety and side effect recognition. *JAMA dermatology* **149**, 485-486, doi:10.1001/jamadermatol.2013.3 (2013).
- 90 Awgulewitsch, A. Hox in hair growth and development. *Die Naturwissenschaften* **90**, 193-211, doi:10.1007/s00114-003-0417-4 (2003).
- 91 Brown, S. D., Chambon, P. & de Angelis, M. H. EMPReSS: standardized phenotype screens for functional annotation of the mouse genome. *Nature genetics* **37**, 1155, doi:10.1038/ng1105-1155 (2005).
- 92 Botchkarev, V. A. & Paus, R. Molecular biology of hair morphogenesis: development and cycling. *Journal of experimental zoology. Part B, Molecular and developmental evolution* **298**, 164-180, doi:10.1002/jez.b.33 (2003).
- 93 Hasson, T. *et al.* Unconventional myosins in inner-ear sensory epithelia. *The Journal of cell biology* **137**, 1287-1307 (1997).
- 94 Irvine, A. D. & McLean, W. H. Human keratin diseases: the increasing spectrum of disease and subtlety of the phenotype-genotype correlation. *The British journal of dermatology* **140**, 815-828 (1999).
- 95 Sundberg, J. P. & Silva, K. A. What color is the skin of a mouse? *Veterinary pathology* **49**, 142-145, doi:10.1177/0300985811417244 (2012).
- 96 Muller-Rover, S. *et al.* A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages. *The Journal of investigative dermatology* **117**, 3-15, doi:10.1046/j.0022-202x.2001.01377.x (2001).
- 97 McLean, W. H. & Moore, C. B. Keratin disorders: from gene to therapy. *Human molecular genetics* **20**, R189-197, doi:10.1093/hmg/ddr379 (2011).
- 98 Lewin, A. S., Glazer, P. M. & Milstone, L. M. Gene therapy for autosomal dominant disorders of keratin. *The journal of investigative dermatology. Symposium proceedings / the Society for Investigative Dermatology, Inc. [and] European Society for Dermatological Research* **10**, 47-61, doi:10.1111/j.1087-0024.2005.10207.x (2005).
- 99 Zhu, P. *et al.* A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Molecular cell* **27**, 609-621, doi:10.1016/j.molcel.2007.07.024 (2007).
- 100 Nijnik, A. *et al.* The critical role of histone H2A-deubiquitinase Mym1 in hematopoiesis and lymphocyte differentiation. *Blood* **119**, 1370-1379, doi:10.1182/blood-2011-05-352666 (2012).
- 101 Szeverenyi, I. *et al.* The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. *Human mutation* **29**, 351-360, doi:10.1002/humu.20652 (2008).
- 102 Chamcheu, J. C. *et al.* Keratin gene mutations in disorders of human skin and its appendages. *Archives of biochemistry and biophysics* **508**, 123-137, doi:10.1016/j.abb.2010.12.019 (2011).
- 103 Uitto, J., Richard, G. & McGrath, J. A. Diseases of epidermal keratins and their linker proteins. *Experimental cell research* **313**, 1995-2009, doi:10.1016/j.yexcr.2007.03.029 (2007).
- 104 Montcouquiol, M. *et al.* Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature* **423**, 173-177, doi:10.1038/nature01618 (2003).
- 105 Kibar, Z. *et al.* Novel mutations in VANGL1 in neural tube defects. *Human mutation* **30**, E706-715, doi:10.1002/humu.21026 (2009).

- 106 Yin, H., Copley, C. O., Goodrich, L. V. & Deans, M. R. Comparison of phenotypes between different vangl2 mutants demonstrates dominant effects of the Looptail mutation during hair cell development. *PloS one* **7**, e31988, doi:10.1371/journal.pone.0031988 (2012).
- 107 Higgins, C. A., Westgate, G. E. & Jahoda, C. A. From telogen to exogen: mechanisms underlying formation and subsequent loss of the hair club fiber. *The Journal of investigative dermatology* **129**, 2100-2108, doi:10.1038/jid.2009.66 (2009).
- 108 Jensen, K. B. & Watt, F. M. Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11958-11963, doi:10.1073/pnas.0601886103 (2006).
- 109 Schweizer, J., Langbein, L., Rogers, M. A. & Winter, H. Hair follicle-specific keratins and their diseases. *Experimental cell research* **313**, 2010-2020, doi:10.1016/j.yexcr.2007.02.032 (2007).
- 110 Jensen, K. B. *et al.* Lrig1 expression defines a distinct multipotent stem cell population in mammalian epidermis. *Cell stem cell* **4**, 427-439, doi:10.1016/j.stem.2009.04.014 (2009).
- 111 Festa, E. *et al.* Adipocyte lineage cells contribute to the skin stem cell niche to drive hair cycling. *Cell* **146**, 761-771, doi:10.1016/j.cell.2011.07.019 (2011).
- 112 Takegahara, N. *et al.* Integral roles of a guanine nucleotide exchange factor, FARP2, in osteoclast podosome rearrangements. *Faseb J* **24**, 4782-4792, doi:10.1096/fj.10-158212 (2010).
- 113 Toyofuku, T. *et al.* FARP2 triggers signals for Sema3A-mediated axonal repulsion. *Nature neuroscience* **8**, 1712-1719, doi:10.1038/nn1596 (2005).
- 114 Benitah, S. A., Frye, M., Glogauer, M. & Watt, F. M. Stem cell depletion through epidermal deletion of Rac1. *Science (New York, N.Y.)* **309**, 933-935, doi:10.1126/science.1113579 (2005).
- 115 Stenn, K. Exogen is an active, separately controlled phase of the hair growth cycle. *Journal of the American Academy of Dermatology* **52**, 374-375, doi:10.1016/j.jaad.2004.07.040 (2005).
- 116 Smyth, I. *et al.* A mouse model of harlequin ichthyosis delineates a key role for Abca12 in lipid homeostasis. *PLoS genetics* **4**, e1000192, doi:10.1371/journal.pgen.1000192 (2008).
- 117 Dobles, M., Liberal, V., Scott, M. L., Benezra, R. & Sorger, P. K. Chromosome missegregation and apoptosis in mice lacking the mitotic checkpoint protein Mad2. *Cell* **101**, 635-645 (2000).
- 118 Schweizer, J. *et al.* New consensus nomenclature for mammalian keratins. *The Journal of cell biology* **174**, 169-174, doi:10.1083/jcb.200603161 (2006).
- 119 Omary, M. B., Ku, N. O., Strnad, P. & Hanada, S. Toward unraveling the complexity of simple epithelial keratins in human disease. *The Journal of clinical investigation* **119**, 1794-1805, doi:10.1172/jci37762 (2009).
- 120 Simpson, C. L., Patel, D. M. & Green, K. J. Deconstructing the skin: cytoarchitectural determinants of epidermal morphogenesis. *Nature reviews. Molecular cell biology* **12**, 565-580, doi:10.1038/nrm3175 (2011).
- 121 Kim, S. & Coulombe, P. A. Emerging role for the cytoskeleton as an organizer and regulator of translation. *Nature reviews. Molecular cell biology* **11**, 75-81, doi:10.1038/nrm2818 (2010).

- 122 Ku, N. O., Zhou, X., Toivola, D. M. & Omary, M. B. The cytoskeleton of digestive epithelia in health and disease. *The American journal of physiology* **277**, G1108-1137 (1999).
- 123 Oriolo, A. S., Wald, F. A., Ramsauer, V. P. & Salas, P. J. Intermediate filaments: a role in epithelial polarity. *Experimental cell research* **313**, 2255-2264, doi:10.1016/j.yexcr.2007.02.030 (2007).
- 124 Toivola, D. M., Tao, G. Z., Habtezion, A., Liao, J. & Omary, M. B. Cellular integrity plus: organelle-related and protein-targeting functions of intermediate filaments. *Trends in cell biology* **15**, 608-617, doi:10.1016/j.tcb.2005.09.004 (2005).
- 125 Kim, S., Wong, P. & Coulombe, P. A. A keratin cytoskeletal protein regulates protein synthesis and epithelial cell growth. *Nature* **441**, 362-365, doi:10.1038/nature04659 (2006).
- 126 Ku, N. O., Toivola, D. M., Strnad, P. & Omary, M. B. Cytoskeletal keratin glycosylation protects epithelial tissue from injury. *Nature cell biology* **12**, 876-885, doi:10.1038/ncb2091 (2010).
- 127 Xu, J. *et al.* The emerging role of iron dyshomeostasis in the mitochondrial decay of aging. *Mechanisms of ageing and development* **131**, 487-493, doi:10.1016/j.mad.2010.04.007 (2010).
- 128 Gu, L. H. & Coulombe, P. A. Keratin function in skin epithelia: a broadening palette with surprising shades. *Current opinion in cell biology* **19**, 13-23, doi:10.1016/j.ceb.2006.12.007 (2007).
- 129 Geisler, N. & Weber, K. The amino acid sequence of chicken muscle desmin provides a common structural model for intermediate filament proteins. *The EMBO journal* **1**, 1649-1656 (1982).
- 130 Lane, E. B. & McLean, W. H. Keratins and skin disorders. *The Journal of pathology* **204**, 355-366, doi:10.1002/path.1643 (2004).
- 131 Parry, D. A., Strelkov, S. V., Burkhard, P., Aebi, U. & Herrmann, H. Towards a molecular description of intermediate filament structure and assembly. *Experimental cell research* **313**, 2204-2216, doi:10.1016/j.yexcr.2007.04.009 (2007).
- 132 Fuchs, E. & Cleveland, D. W. A structural scaffolding of intermediate filaments in health and disease. *Science (New York, N.Y.)* **279**, 514-519 (1998).
- 133 Osborn, M. Intermediate filaments as histologic markers: an overview. *The Journal of investigative dermatology* **81**, 104s-109s (1983).
- 134 Coulombe, P. A. & Omary, M. B. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. *Current opinion in cell biology* **14**, 110-122 (2002).
- 135 Niessen, C. M. Tight junctions/adherens junctions: basic structure and function. *The Journal of investigative dermatology* **127**, 2525-2532, doi:10.1038/sj.jid.5700865 (2007).
- 136 Kirschner, N., Bohner, C., Rachow, S. & Brandner, J. M. Tight junctions: is there a role in dermatology? *Archives of dermatological research* **302**, 483-493, doi:10.1007/s00403-010-1058-z (2010).
- 137 De Benedetto, A. *et al.* Tight junction defects in patients with atopic dermatitis. *The Journal of allergy and clinical immunology* **127**, 773-786 e771-777, doi:10.1016/j.jaci.2010.10.018 (2011).
- 138 Kirschner, N. *et al.* Alteration of tight junction proteins is an early event in psoriasis: putative involvement of proinflammatory cytokines. *The American journal of pathology* **175**, 1095-1106, doi:10.2353/ajpath.2009.080973 (2009).

- 139 Fanning, A. S., Jameson, B. J., Jesaitis, L. A. & Anderson, J. M. The tight junction protein ZO-1 establishes a link between the transmembrane protein occludin and the actin cytoskeleton. *The Journal of biological chemistry* **273**, 29745-29753 (1998).
- 140 Ambatipudi, S. *et al.* Downregulation of keratin 76 expression during oral carcinogenesis of human, hamster and mouse. *PloS one* **8**, e70688, doi:10.1371/journal.pone.0070688 (2013).
- 141 Collin, C., Ouhayoun, J. P., Grund, C. & Franke, W. W. Suprabasal marker proteins distinguishing keratinizing squamous epithelia: cytokeratin 2 polypeptides of oral masticatory epithelium and epidermis are different. *Differentiation; research in biological diversity* **51**, 137-148 (1992).
- 142 Baran, W., Szepietowski, J. C. & Szybejko-Machaj, G. Expression of p53 protein in psoriasis. *Acta dermatovenerologica Alpina, Panonica, et Adriatica* **14**, 79-83 (2005).
- 143 de Rie, M. A., Goedkoop, A. Y. & Bos, J. D. Overview of psoriasis. *Dermatologic therapy* **17**, 341-349, doi:10.1111/j.1396-0296.2004.04037.x (2004).
- 144 Thewes, M., Stadler, R., Korge, B. & Mischke, D. Normal psoriatic epidermis expression of hyperproliferation-associated keratins. *Archives of dermatological research* **283**, 465-471 (1991).
- 145 Lin, J. Y. & Fisher, D. E. Melanocyte biology and skin pigmentation. *Nature* **445**, 843-850, doi:10.1038/nature05660 (2007).
- 146 Berta, M. A., Baker, C. M., Cottle, D. L. & Watt, F. M. Dose and context dependent effects of Myc on epidermal stem cell proliferation and differentiation. *EMBO molecular medicine* **2**, 16-25, doi:10.1002/emmm.200900047 (2010).
- 147 Cottle, D. L. *et al.* c-MYC-induced sebaceous gland differentiation is controlled by an androgen receptor/p53 axis. *Cell reports* **3**, 427-441, doi:10.1016/j.celrep.2013.01.013 (2013).
- 148 Madsen, P., Rasmussen, H. H., Leffers, H., Honore, B. & Celis, J. E. Molecular cloning and expression of a novel keratinocyte protein (psoriasis-associated fatty acid-binding protein [PA-FABP]) that is highly up-regulated in psoriatic skin and that shares similarity to fatty acid-binding proteins. *The Journal of investigative dermatology* **99**, 299-305 (1992).
- 149 Ogawa, E. *et al.* Epidermal FABP (FABP5) regulates keratinocyte differentiation by 13(S)-HODE-mediated activation of the NF-kappaB signaling pathway. *The Journal of investigative dermatology* **131**, 604-612, doi:10.1038/jid.2010.342 (2011).
- 150 Paladini, R. D., Takahashi, K., Bravo, N. S. & Coulombe, P. A. Onset of re-epithelialization after skin injury correlates with a reorganization of keratin filaments in wound edge keratinocytes: defining a potential role for keratin 16. *The Journal of cell biology* **132**, 381-397 (1996).
- 151 Segre, J. A. Epidermal barrier formation and recovery in skin disorders. *The Journal of clinical investigation* **116**, 1150-1158, doi:10.1172/jci28521 (2006).
- 152 Sjo, A., Magnusson, K. E. & Peterson, K. H. Protein kinase C activation has distinct effects on the localization, phosphorylation and detergent solubility of the claudin protein family in tight and leaky epithelial cells. *The Journal of membrane biology* **236**, 181-189, doi:10.1007/s00232-010-9289-7 (2010).
- 153 Mrsny, R. J. *et al.* A key claudin extracellular loop domain is critical for epithelial barrier integrity. *The American journal of pathology* **172**, 905-915, doi:10.2353/ajpath.2008.070698 (2008).

- 154 Watson, R. E. *et al.* Altered claudin expression is a feature of chronic plaque psoriasis. *The Journal of pathology* **212**, 450-458, doi:10.1002/path.2200 (2007).
- 155 Dhawan, P. *et al.* Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer. *The Journal of clinical investigation* **115**, 1765-1776, doi:10.1172/JCI24543 (2005).
- 156 French, A. D. *et al.* PKC and PKA phosphorylation affect the subcellular localization of claudin-1 in melanoma cells. *International journal of medical sciences* **6**, 93-101 (2009).
- 157 Hough, C. D. *et al.* Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. *Cancer research* **60**, 6281-6287 (2000).
- 158 Adams, N. & Gale, N. in *Mammalian and Avian Transgenesis — New Approaches Principles and Practice* (eds Shirley Pease & Carlos Lois) Ch. 7, 131-172 (Springer Berlin Heidelberg, 2006).
- 159 Lichti, U., Anders, J. & Yuspa, S. H. Isolation and short-term culture of primary keratinocytes, hair follicle populations and dermal cells from newborn mice and keratinocytes from adult mice for in vitro analysis and for grafting to immunodeficient mice. *Nature protocols* **3**, 799-810, doi:10.1038/nprot.2008.50 (2008).
- 160 Chen, Y., Merzdorf, C., Paul, D. L. & Goodenough, D. A. COOH terminus of occludin is required for tight junction barrier function in early *Xenopus* embryos. *The Journal of cell biology* **138**, 891-899 (1997).
- 161 Hesse, M., Magin, T. M. & Weber, K. Genes for intermediate filament proteins and the draft sequence of the human genome: novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. *Journal of cell science* **114**, 2569-2575 (2001).
- 162 Simpson, E. M. *et al.* Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. *Nature genetics* **16**, 19-27, doi:10.1038/ng0597-19 (1997).
- 163 Bailey, D. W. Four approaches to estimating number of histocompatibility loci. *Transplantation proceedings* **2**, 32-38 (1970).

## Appendix I Supplementary Information for Chapter 2

### Identification of genes important for cutaneous function revealed by a large scale reverse genetic screen in the mouse

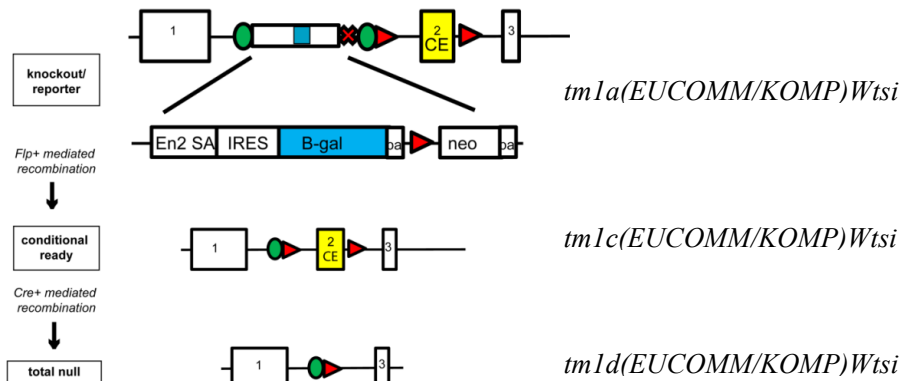
DiTommaso et al.

## Supplementary Information pages 81-118

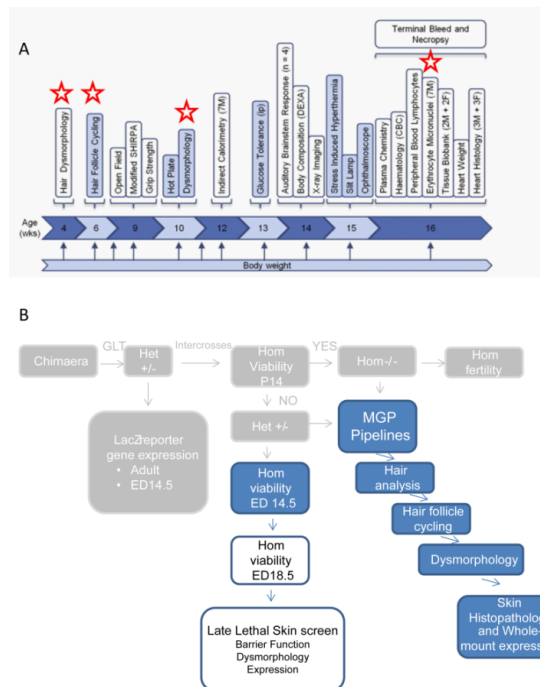
### CONTENTS

<b>Supplementary Figure 1. EUCOMM/KOMP knockout first conditional-ready targeting construct.</b> These lines are predicted to produce null alleles and prediction of gene expression patterns using the integrated B-gal reporter ( <i>tm1a</i> ( <i>EUCOMM</i> )/ <i>WTSl</i> )). Conditional ( <i>tm1c</i> ( <i>EUCOMM</i> )/ <i>WTSl</i> )) and total knockout( <i>tm1d</i> ( <i>EUCOMM</i> )/ <i>WTSl</i> )) alleles can be generated using cre and flp mediated recombination. ....	148
<b>Supplementary Figure 2. Pipelines and phenotype tests (adapted from White et al).</b> MGP phenotyping pipeline shows phenotyping tests and age at which test is undertaken. Red star = skin specific test. Generation of knockouts from the chimera stage to necropsy is shown. Grey = knockout generation strategy, blue = pre-existing/implemented skin screens; white = proposed skin screens.....	148
<b>Supplementary Table 1. Genes and phenotypes identified in gross clinical phenotype and hair follicle cycling screens.</b> 7 males and 7 female from 362 different transgenic lines were screened for defects in hair follicle cycling and grossly evident skin abnormalities.....	149
<b>Supplementary Table 2. Genes and phenotypes identified in skin histopathology screen.</b> Paraffin embedded, H&E stained skin from a cohort including 508 mutants and 46 wildtypes were assessed for histologic lesions.....	181
<b>Supplementary Table 3. Genes and MP terms discovered across all tests.</b> Each screen identified genes that were unique to the specific test, thus increasing the number of genes flagged. Unique MP terms were also described by adding different tests.....	182
<b>Supplementary Table 4. Screen parameters.</b> The Sanger-MGP pipeline comprises 263 parameters, with 43 directly relating to the integument. Parameters relating to skin, hair, teeth, and nails are indicated with a 1, whereas non skin parameters are indicated with a 0.....	190

## SUPPLEMENTARY FIGURES AND TABLES



**Supplementary figure 1. EUCOMM/KOMP knockout first conditional-ready targeting construct.** These lines are predicted to produce null alleles and prediction of gene expression patterns using the integrated B-gal reporter (*tm1a (EUCOMM)WTSI*). Conditional (*tm1c (EUCOMM)WTSI*) and total knockout (*tm1d (EUCOMM)WTSI*) alleles can be generated using cre and flp mediated recombination.



**Supplementary Figure 2. Pipelines and phenotype tests (adapted from White et al).**

MGP phenotyping pipeline shows phenotyping tests and age at which test is undertaken. Red star = skin specific test. Generation of knockouts from the chimera stage to necropsy is shown. Grey = knockout generation strategy, blue = pre-existing/implemented skin screens; white = proposed skin screens.



**Supplementary Table 1. Genes and phenotypes identified in gross clinical phenotype and hair follicle cycling screens. 7 males and 7 female from 362 different transgenic lines were screened for defects in hair follicle cycling and grossly evident skin abnormalities. (yellow indicates skin phenotype).**

Gene	Genotype	Dys- morpho logy	Follicle Cycling Phenot ype?	Phenotype	Strain
Cdk5rap2	Hom	Yes	no	MP:0000432 - abnormal head morphology MP:0000346 - abnormal head morphology, MP:0000432 - broad head, MP:0001148 - abnormal external male genitalia morphology, MP:0001270 - bulging abdomen, MP:0009199 - enlarged testis	C57BL/6N C57BL/6JTy;C57BL/6N
Cenpl	Het & Hom	Yes	no	MP:0002697 - abnormal eye size	C57BL/6JTy;C57BL/6N
Csrp2bp	Hom	Yes	no	MP:0000585 - abnormal tail morphology, MP:0002111 - kinked tail	C57BL/6N
Dnasel12	Hom	Yes	no	MP:0000443 - abnormal snout morphology	C57BL/6N
Exoc3l2	Hom	Yes	no	MP:0000443 - abnormal snout morphology, MP:0005358 - abnormal incisor morphology, MP:0005287 - narrow eye opening	C57BL/6JTy;C57BL/6N
Kdm4b	Het & Hom	Yes	no	MP:0001510 - abnormal coat appearance, MP:0010179 - rough coat, MP:0000416 - sparse hair	C57BL/6JTy;C57BL/6N C57BL/6JTy;C57BL/6N
Krt76	Het	Yes	no	MP:0001191 - abnormal skin condition, MP:0001192 - scaly skin	C57BL/6JTy;C57BL/6N
Lrig1	Hom	Yes	no	MP:0000443 - abnormal snout morphology, MP:0000120 - abnormal incisor morpholog	C57BL/6JTy;C57BL/6N
Ltbp1	Het & Hom	Yes	no	MP:0000371 - abnormal coat/hair pigmentation, MP:0002075 - diluted coat color, MP:0000367 - abnormal coat/ hair morphology, MP:0000418 - localized hair loss, MP:0001148 - abnormal external male genitalia morphology, MP:0009199 - enlarged testis	C57BL/6JTy;C57BL/6N C57BL/6Jlco;C57BL/10
Myo5a	Hom	Yes	yes	MP:0000372 - belly spot, MP:0000556 - abnormal hindlimb morphology, MP:0000556 - abnormal hindlimb morphology, MP:0003051 - curly tail, MP:0000592 - short tail	C57BL/6
Myo6	Het & Hom	Yes	no	abnormal coat appearance [MP:0001510], MP:0000443 - abnormal snout morphology	C57BL/6JTy;C57BL/6N
Mysm1	Het & Hom	Yes	no	abnormal nail morphology [MP:0000579]	C57BL/6N
Nsun2	Het & Hom	Yes	yes	MP:0000443 - abnormal snout morphology, MP:0005358 - abnormal incisor morphology	C57BL/6JTy;C57BL/6N
Ptkab1	Het	Yes	no	MP:0001324 - abnormal eye pigmentation, MP:0009859 - eye opacity, MP:0002697 - abnormal eye size, MP:0005287 - narrow eye opening	C57BL/6JTy;C57BL/6N
Slc25a21	Hom	Yes	no	MP:0010096 - abnormal incisor color	C57BL/6JTy;C57BL/6N
Sk4	Het & Hom	Yes	no	MP:0001324 - abnormal eye pigmentation, MP:0009859 - eye opacity [mp:0009859]	C57BL/6JTy;C57BL/6N
Sparc	Hom	Yes	no	MP:0003849 - greasy coat, MP:0000574 - abnormal foot pad morphology, MP:0000575 - dark foot pads	C57BL/6N
Spns2	Hom	Yes	no	MP:0000445 - short snout	C57BL/6N
Vangl1	Het & Hom	Yes	no	MP:0000443 - abnormal snout morphology	C57BL/6N
Ywhae	Hom	Yes	no		OTHER
Zc3hc1	Het & Hom	Yes	no		C57BL/6JTy;C57BL/6N
Nom1	Het	no	yes	MP:0008858 - abnormal hair cycle anagen phase	C57BL/6JTy;c;C57BL/6N
Nf1	Het	no	yes	MP:0008858 - abnormal hair cycle anagen phase	C57BL/6N
Trpcap	Het	no	yes	MP:0008858 - abnormal hair cycle anagen phase	C57BL/6N
1190002H23Rik	Hom	no	no	no phenotype detected	C57BL/6N
1500011B03Rik	Hom	no	no	no phenotype detected	C57BL/6N
1700058C13Rik	Het & Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
2010107G12Rik	Hom	no	no	no phenotype detected	C57BL/6N
2210016L21Rik	Hom	no	no	no phenotype detected	C57BL/6N

2310007B03RIK	Hom	no	no	no phenotype detected	C57BL/6N
2310057J18RIK	Hom	no	no	no phenotype detected	C57BL/6N
3010026O09RI k	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
3110001I22RIK	Hom	no	no	no phenotype detected	C57BL/6N
4930471M23RI k	Hom	no	no	no phenotype detected	C57BL/6N
4932414ND04RIK	Hom	no	no	no phenotype detected	C57BL/6N
4932438H23RIK	Hom	no	no	no phenotype detected	C57BL/6N
4933425L06RIK	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
5830415F09RIK	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
9030625A04RIK (laec1)	Hom	no	no	no phenotype detected	C57BL/6N
A430084P05RIK	Hom	no	no	no phenotype detected	C57BL/6N
Abca4	Hom	no	no	no phenotype detected	C57BL/6N
Abcd1	Hom/Hemi	no	no	no phenotype detected	C57BL/6N
Abhd5	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Acot6	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Acpl2	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Adam3	Hom	no	no	no phenotype detected	C57BL/6N
Aff3	Het & Hom	no	no	no phenotype detected	C57BL/6N
Agtr2	Hom/Hemi	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Akap9	Het & Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Akt2	Het & Hom	no	no	no phenotype detected	129/SvEv
Aldh16a1	Hom	no	no	no phenotype detected	C57BL/6N
Aldh2	Hom	no	no	no phenotype detected	C57BL/6
Alg13	Hom/Hemi	no	no	no phenotype detected	C57BL/6N
Amfr	Het & Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Anot11	Hom	no	no	no phenotype detected	C57BL/6N
Ampd3	Hom	no	no	no phenotype detected	C57BL/6N
Ankrd13a	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Anks4b	Hom	no	no	no phenotype detected	CBA/Ca:129P2
Anp32e	Hom	no	no	no phenotype detected	C57BL/6N
Anxa6	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Ap4e1	Hom	no	no	no phenotype detected	C57BL/6N
Apoe	Hom	no	no	no phenotype detected	C57BL/6
Apo17a	Hom	no	no	no phenotype detected	C57BL/6N
Apo1	Hom/Hemi	no	no	no phenotype detected	C57BL/6N
App12	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Arhgap18	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Arhgap25	Hom	no	no	no phenotype detected	C57BL/6N

Arhgef4	Hom	no	no	no phenotype detected	C57BL/6N
Arid4a	Het & Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Arl4d	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Arpc1b	Hom	no	no	no phenotype detected	C57BL/6N
Arsg	Hom	no	no	no phenotype detected	C57BL/6N
Arvrf	Hom	no	no	no phenotype detected	C57BL/6N
Astef1	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Atipif1	Hom	no	no	no phenotype detected	C57BL/6N
Bbx	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
BC013712	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
BC017643	Het & Hom	no	no	no phenotype detected	C57BL/6N
Bdh2	Hom	no	no	no phenotype detected	C57BL/6N
Blzf1	Hom	no	no	no phenotype detected	C57BL/6N
Brdt	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Btbd11	Hom	no	no	no phenotype detected	C57BL/6N
Cabp1	Hom	no	no	no phenotype detected	C57BL/6N
Cadml	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Calococ2	Hom	no	no	no phenotype detected	C57BL/6N
Cand2	Hom	no	no	no phenotype detected	C57BL/6N
Caprin2	Hom	no	no	no phenotype detected	C57BL/6N
Casp12	Hom	no	no	no phenotype detected	C57BL/6N
Cbx5	Hom	no	no	no phenotype detected	C57BL/6N
Cbx6	Hom	no	no	no phenotype detected	C57BL/6
Cbx7	Hom	no	no	no phenotype detected	C57BL/6
Cc2d1b	Het & Hom	no	no	no phenotype detected	C57BL/6N
Ccdc104	Hom	no	no	no phenotype detected	C57BL/6N
Ccdc106	Hom	no	no	no phenotype detected	C57BL/6N
Ccdc160	Hom/Hemi	no	no	no phenotype detected	C57BL/6N
Cdc57	Het & Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Cdc77	Hom	no	no	no phenotype detected	STOCK
Cdh23	Het & Hom	no	no	no phenotype detected	C57BL/6N
Cdkn2a	Hom	no	no	no phenotype detected	C57BL/6N
Cfh	Hom	no	no	no phenotype detected	C57BL/6N
Chia	Hom	no	no	no phenotype detected	C57BL/6N
Cish	Hom	no	no	no phenotype detected	C57BL/6N
Cldn16	Hom	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> :C57BL/6N
Clk1	Hom	no	no	no phenotype detected	C57BL/6N
Cntm6	Hom	no	no	no phenotype detected	C57BL/6N
Cnat6	Het & Hom	no	no	no phenotype detected	C57BL/6N
Cog9	Hom	no	no	no phenotype detected	C57BL/6N

Crtf3	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Cnnal1	Hom	no	no	no phenotype detected	C57BL/6
Cybb561	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
D11Wsub9e	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Dbn1	Het & Hom	no	no	no phenotype detected	C57BL/6N
Dcdc2c	Hom	no	no	no phenotype detected	C57BL/6N
Demnd1c	Hom	no	no	no phenotype detected	C57BL/6N
Dlg2	Hom	no	no	no phenotype detected	C57BL/6N
Dlg4	Hom	no	no	no phenotype detected	C57BL/6Jlco
Dopey2	Hom	no	no	no phenotype detected	C57BL/6Jlco.129P2
Dsccl	Het & Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Duspl	Hom	no	no	no phenotype detected	C57BL/6N
Dusp26	Hom	no	no	no phenotype detected	C57BL/6N
Dusp3	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Dusp4	Hom	no	no	no phenotype detected	C57BL/6N
Etna1	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Ehd1	Het & Hom	no	no	no phenotype detected	C57BL/6N
Eif2c1	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Eif4e3	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Elk4	Hom	no	no	no phenotype detected	C57BL/6
Entpd1	Hom	no	no	no phenotype detected	C57BL/6
Eps15	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Escr1	Hom	no	no	no phenotype detected	C57BL/6N
Espn	Hom	no	no	no phenotype detected	C57BL/6N
Expi	Hom	no	no	no phenotype detected	C57BL/6N
Fahd2a	Hom	no	no	no phenotype detected	C57BL/6N
Fam107b	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Fam134c	Hom	no	no	no phenotype detected	C57BL/6N
Fam63a	Hom	no	no	no phenotype detected	C57BL/6
Fam73b	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Fam96a	Hom	no	no	no phenotype detected	C57BL/6N
Farp2	Hom	no	no	no phenotype detected	C57BL/6N
Fbxl7	Hom	no	no	no phenotype detected	C57BL/6N
Fbxo47	Het & Hom	no	no	no phenotype detected	C57BL/6N
Fbxo7	Hom	no	no	no phenotype detected	C57BL/6N
Fgfy	Hom	no	no	no phenotype detected	C57BL/6N
Fkbp7	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Fto	Hom	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Fundct1	Hom/Hemi	no	no	no phenotype detected	C57BL/6JTyrc57BL/6N
Gba2	Hom	no	no	no phenotype detected	C57BL/6N

Gclc	Hom	no	no	no	no phenotype detected	C57BL/6N
Gif	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
Glt2	Hom	no	no	no	no phenotype detected	CBA/Ca.129P2
Glt8d2	Hom	no	no	no	no phenotype detected	C57BL/6N
Gm13547	Hom	no	no	no	no phenotype detected	C57BL/6N
Gm239	Hom	no	no	no	no phenotype detected	C57BL/6N
Gpc5	Hom	no	no	no	no phenotype detected	C57BL/6N
Gpc6	Hom	no	no	no	no phenotype detected	C57BL/6N
Gprc5b	Hom	no	no	no	no phenotype detected	C57BL/6N
Grii2a	Hom	no	no	no	no phenotype detected	C57BL/6N
Grcr1	Het & Hom	no	no	no	no phenotype detected	C57BL/6Jlco;.129P2
Gsk3a	Hom	no	no	no	no phenotype detected	STOCK
Gys2	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
H2-M5	Hom	no	no	no	no phenotype detected	C57BL/6N
Hac1	Hom	no	no	no	no phenotype detected	C57BL/6N
Hdac8	Hom/Hemi	no	no	no	no phenotype detected	C57BL/6N
Heat8	Hom	no	no	no	no phenotype detected	C57BL/6N
Herc3	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
Hmx3	Hom	no	no	no	no phenotype detected	STOCK
Hp1bp3	Het & Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
Hprt	Hom/Hemi	no	no	no	no phenotype detected	C57BL/6N
Ide	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
Ido1	Hom	no	no	no	no phenotype detected	C57BL/6N
Iftm3	Hom	no	no	no	no phenotype detected	C57BL/6Jlco;C57BL/6JTy <sup>r</sup>
Il10b	Hom	no	no	no	no phenotype detected	.129P2
Il22ra1	Hom	no	no	no	no phenotype detected	C57BL/6N
Inpp1	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
Irf1	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N
Jhdm1d	Hom	no	no	no	no phenotype detected	C57BL/6N
Jmjd5	Het & Hom	no	no	no	no phenotype detected	C57BL/6N
Kazn	Hom	no	no	no	no phenotype detected	C57BL/6N
Kcne2	Hom	no	no	no	no phenotype detected	C57BL/6N
Kcnv2	Hom	no	no	no	no phenotype detected	C57BL/6N
Kdm4c	Hom	no	no	no	no phenotype detected	C57BL/6N
Kdm4d	Hom	no	no	no	no phenotype detected	C57BL/6N
Kif21b	Hom	no	no	no	no phenotype detected	C57BL/6N
Kng2	Hom	no	no	no	no phenotype detected	C57BL/6N
Kpn	Hom	no	no	no	no phenotype detected	C57BL/6JTy <sup>r</sup> ;C57BL/6N

Lamc3	Hom	no	no	no	no phenotype detected	C57BL/6N
Laptm5	Hom	no	no	no	no phenotype detected	C57BL/6N
Lepr <sup>ot</sup>	Hom	no	no	no	no phenotype detected	C57BL/6N
Lnx2	Hom	no	no	no	no phenotype detected	C57BL/6N
Lpar5	Hom	no	no	no	no phenotype detected	C57BL/6N
Lrrc16a	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Lrrc67	Hom	no	no	no	no phenotype detected	C57BL/6N
Maneal	Hom	no	no	no	no phenotype detected	C57BL/6N
Map2k2	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Map3k1	Het & Hom	no	no	no	no phenotype detected	C57BL/6N
Mapk10	Hom	no	no	no	no phenotype detected	C57BL/6N
Matn1	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6.12 9S5
Matn3	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6.12 9S5
Mcph1	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Meik	Hom	no	no	no	no phenotype detected	C57BL/6N
Mfrp	Hom	no	no	no	no phenotype detected	C57BL/6N
Mfsd2b	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Mgsc3	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Mier1	Het & Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Mir96	Het & Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Mlec	Hom	no	no	no	no phenotype detected	C3HeB/FeJ
Mrap2	Hom	no	no	no	no phenotype detected	C57BL/6N
Mst1	Hom	no	no	no	no phenotype detected	C57BL/6N
Mtfa1	Het & Hom	no	no	no	no phenotype detected	129/SvEv
Mtfa3	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6Jlco
Mx1	Hom	no	no	no	no phenotype detected	C57BL/6N
Myd88	Hom	no	no	no	no phenotype detected	C57BL/6N
My12b	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Myo15	Hom	no	no	no	no phenotype detected	C57BL/6N
Myo7a	Hom	no	no	no	no phenotype detected	C57BL/6N
Naga	Hom	no	no	no	no phenotype detected	C57BL/6N
Necab2	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Nelf	Hom	no	no	no	no phenotype detected	C57BL/6N
Nkx1	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Nlrl	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Nlpa1	Hom	no	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Nkiras2	Hom	no	no	no	no phenotype detected	C57BL/6N
Oas1g	Hom	no	no	no	no phenotype detected	C57BL/6N

Oas12	Hom	no	no	no	no phenotype detected	C57BL/6N
Ocm	Hom	no	no	no	no phenotype detected	C57BL/6N
Ogfod1	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Opn	Hom	no	no	no	no phenotype detected	C57BL/6N
Osbp19	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Otub2	Hom	no	no	no	no phenotype detected	C57BL/6N
Otud6b	Hom	no	no	no	no phenotype detected	C57BL/6
Pabpc1l	Hom	no	no	no	no phenotype detected	C57BL/6N
Pabpc4	Hom	no	no	no	no phenotype detected	C57BL/6N
Pacsin3	Hom	no	no	no	no phenotype detected	C57BL/6N
Pex3	Hom	no	no	no	no phenotype detected	C57BL/6N
Phf20	Het & Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Phf12	Hom	no	no	no	no phenotype detected	C57BL/6N
Plk3cb	Het & Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Plk3cg	Hom	no	no	no	no phenotype detected	C57BL/6N
Plk3r2	Hom	no	no	no	no phenotype detected	C57BL/6N
Pitpm1	Hom	no	no	no	no phenotype detected	C57BL/6N
1100001H23R1						
K (Plbd1)	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Pld5	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Plekg1	Hom	No	No	No	no phenotype detected	C57BL/6N
Plekhm2	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Plin2	Hom	no	no	no	no phenotype detected	C57BL/6N
Pp2r2b	Hom	no	no	no	no phenotype detected	C57BL/6N
Ppp3ca	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Ppp5c	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Prdx6	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Prkcz	Het	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Prrnt2	Hom	no	no	no	no phenotype detected	C57BL/6N
Prrnt3	Het & Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Prrnt7	Het & Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Prrnt8	Hom	no	no	no	no phenotype detected	C57BL/6N
Prpsap2	Hom	no	no	no	no phenotype detected	C57BL/6N
Ptges2	Hom	no	no	no	no phenotype detected	C57BL/6N
Pus7l	Hom	no	no	no	no phenotype detected	C57BL/6N
Rab7l1	Hom	no	no	no	no phenotype detected	C57BL/6N
Rad18	Hom	no	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Rab1	Hom	no	no	no	no phenotype detected	C57BL/6N

Ragef11	Hom	no	no	no phenotype detected	C57BL/6N
Rassf1	Hom	no	no	no phenotype detected	C57BL/6JTy;1.29S5
Rbbp7	Hom/Hemi	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Rcor2	Het & Hom	no	no	no phenotype detected	C57BL/6N
Rdh16	Hom	no	no	no phenotype detected	C57BL/6N
Ren1	Hom	no	no	no phenotype detected	C57BL/6Jlco
Rftn2	Hom	no	no	no phenotype detected	C57BL/6N
Rhd	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Rhobtb3	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Rhot2	Hom	no	no	no phenotype detected	C57BL/6N
Rhou	Hom	no	no	no phenotype detected	C57BL/6N
Rhox13	Hom/Hemi	no	no	no phenotype detected	C57BL/6N
Ripply3	Hom	no	no	no phenotype detected	C57BL/6N
Rnf10	Hom	no	no	no phenotype detected	C57BL/6N
ROSA26	Hom	no	no	no phenotype detected	C57BL/6N
Rtbdn	Het & Hom	no	no	no phenotype detected	C57BL/6N
Rufy2	Het & Hom	no	no	no phenotype detected	C57BL/6N
S100b	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Sag	Hom	no	no	no phenotype detected	C57BL/6N
Sec16b	Hom	no	no	no phenotype detected	C57BL/6N
Sec24a	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Secisbp2	Hom	no	no	no phenotype detected	C57BL/6N
Sepr1	Hom	no	no	no phenotype detected	C57BL/6N
Sesn3	Hom	no	no	no phenotype detected	C57BL/6N
Setmar	Hom	no	no	no phenotype detected	C57BL/6N
Sgpl2	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Sirt2	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Sirt3	Hom	no	no	no phenotype detected	C57BL/6N
Slc16a2	Hom/Hemi	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Slc22a21	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Slc25a29	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Slc25a4	Hom	no	no	no phenotype detected	C57BL/6N
Slc25a43	Hom/Hemi	no	no	no phenotype detected	C57BL/6N
Slc30a8	Hom	no	no	no phenotype detected	C57BL/6N
Slc35f1	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Slc38a10	Hom	no	no	no phenotype detected	C57BL/6N
Slc41a3	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Slc44a5	Hom	no	no	no phenotype detected	C57BL/6N
Slc5a2	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N
Slc9a8	Hom	no	no	no phenotype detected	C57BL/6JTy;C57BL/6N



Slco2b1	Hom	no	no	no phenotype detected	C57BL/6N
Smarca11	Hom	no	no	no phenotype detected	C57BL/6N
Smyd2	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Smyd3	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Smyd4	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Smyd5	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Snap47	Hom	no	no	no phenotype detected	C57BL/6N
Snx5	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Socs7	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Sgstm1	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Srsf4	Het & Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
St18	Hom	no	no	no phenotype detected	C57BL/6N
Stard13	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Stard5	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Stard6	Hom	no	no	no phenotype detected	C57BL/6N
Stk32a	Hom	no	no	no phenotype detected	C57BL/6N
Stk39	Hom	no	no	no phenotype detected	C57BL/6N
Stx8	Hom	no	no	no phenotype detected	C57BL/6N
Suv420h1	Het & Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Synj2	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Syt11	Hom	no	no	no phenotype detected	C57BL/6N
Tatdn3	Hom	no	no	no phenotype detected	C57BL/6N
Tbc1d10a	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Tcf4	Het & Hom	no	no	no phenotype detected	C57BL/6N
Tff1	Hom	no	no	no phenotype detected	C57BL/6N
Thump3	Hom	no	no	no phenotype detected	C57BL/6N
Tnfrsf4	Hom	no	no	no phenotype detected	C57BL/6N
Tmcl	Het & Hom	no	no	no phenotype detected	STOCK
Tmc6	Hom	no	no	no phenotype detected	C57BL/6N
Tmem165	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Tmem189	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Tmem9	Hom	no	no	no phenotype detected	C57BL/6N
Tmprss15	Hom	no	no	no phenotype detected	C57BL/6N
Tnfrap1	Hom	no	no	no phenotype detected	C57BL/6N
Top3b	Hom	no	no	no phenotype detected	OTHER
Tp52l2	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Traf3ip3	Hom	no	no	no phenotype detected	C57BL/6N
Traf1	Hom	no	no	no phenotype detected	C57BL/6JYr;C57BL/6N
Trim17	Hom	no	no	no phenotype detected	C57BL/6N

Trim66	Hom	no	no	no phenotype detected	C57BL/6N
Trmt10a	Hom	no	no	no phenotype detected	C57BL/6N
Trpc2	Hom	no	no	no phenotype detected	C57BL/6N
Tsga14	Hom	no	no	no phenotype detected	C57BL/6N
Ttc39b	Het & Hom	no	no	no phenotype detected	C57BL/6
Ttli3	Hom	no	no	no phenotype detected	C57BL/6JTy; C57BL/6N
Ttli4	Hom	no	no	no phenotype detected	C57BL/6N
Twf1	Hom	no	no	no phenotype detected	CBA/Ca;129P2
Ube3b	Het & Hom	no	no	no phenotype detected	C57BL/6JTy; C57BL/6N
Uck1	Hom	no	no	no phenotype detected	CBA/Ca;129P2
Usp33	Hom	no	no	no phenotype detected	C57BL/6N
Usp4	Hom	no	no	no phenotype detected	C57BL/6N
Usp42	Hom	no	no	no phenotype detected	C57BL/6N
Vwa5b1	Hom	no	no	no phenotype detected	C57BL/6N
Wbp2	Hom	no	no	no phenotype detected	C57BL/6N
Wdr37	Hom	no	no	no phenotype detected	C57BL/6N
Wdr47	Hom	no	no	no phenotype detected	C57BL/6JTy; C57BL/6N
Whm	Hom	no	no	no phenotype detected	C57BL/6N
Yipf1	Hom	no	no	no phenotype detected	C57BL/6JTy; C57BL/6N
Zip182	Het &	no	no	no phenotype detected	C57BL/6N
Zip184	Hom/Hemi	no	no	no phenotype detected	C57BL/6
Zip367	Hom	no	no	no phenotype detected	C57BL/6JTy; C57BL/6N
Zp1	Hom	no	no	no phenotype detected	C57BL/6N

**Supplementary Table 2.** Genes and phenotypes identified in skin histopathology screen. Paraffin embedded, H&E stained skin from a cohort including 508 mutants and 46 wildtypes were assessed for histologic lesions (grey=Recessive lethal or subviable)

Gene	Genotype	Phenotype	Age weeks	Sex	Background	Diet
1700058C13Rik	1700058C13Rik/1700058C13R	<a href="#">dermatitis [MP:0001194]</a>	15.6	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Aldh18a1	Aldh18a1/+	<a href="#">abnormal hypodermis fat layer morphology [MP:0011156]</a>	16.1	Female	B6N	Diet
Anks4b	Anks4b/Anks4b	<a href="#">skin fibrosis [MP:0009932]</a>	16.3	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Pas5b	Aprin/+	<a href="#">dermatitis [MP:0001194]</a>	15.7	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Cadm1	Cadm1/Cadm1	<a href="#">skin inflammation [MP:0004947]</a>	15.9	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Capzb	Capzb/+	<a href="#">skin inflammation [MP:0004947]</a>	15.9	Female	B6	High-Fat Diet
Cbx7	Cbx7/Cbx7	<a href="#">skin inflammation [MP:0004947]</a>	15.6	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Cdcd57	Cdcd57/Cdcd57	<a href="#">abnormal hypodermis fat layer morphology [MP:0011156]</a>	16	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Cldn16	Cldn16/Cldn16	<a href="#">dermatitis [MP:0001194]</a>	16	Female	B6N	Crow
Cntm4	Cntm4/Cntm4	<a href="#">folliculitis [MP:0000376]</a>	16.1	Female	B6N	High-Fat Diet
Cnot6	Cnot6/Cnot6	<a href="#">folliculitis [MP:0000376]</a>	15.9	Female	B6N	Crow
Coq9	Coq9/Coq9	<a href="#">dermatitis [MP:0001194]</a>	15.9	Female	B6N	High-Fat Diet
Cyfp1	Cyfp1/+	<a href="#">dermatitis [MP:0001194]</a>	16	Female	B6N	High-Fat Diet
Dars2	Dars2/+	<a href="#">dermatitis [MP:0001194]</a>	16.1	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Dhcr24	Dhcr24/+	<a href="#">skin inflammation [MP:0004947]</a>	15.4	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Eif2c1	Eif2c1/Eif2c1	<a href="#">dermatitis [MP:0001194]</a>	15.9	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Farp2	Farp2/Farp2	<a href="#">abnormal hair shedding [MP:0008861]</a>	15.9	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Fkbp7	Fkbp7/Fkbp7	<a href="#">dermatitis [MP:0001194]</a>	15.7	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Fundc1	Fundc1/Fundc1	<a href="#">dermatitis [MP:0001194]</a>	16	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Gbel	Gbel/+	<a href="#">folliculitis [MP:0000376]</a>	16	Female	B6Br;dB6Dnk;B6N-Ty<-C-Brd>	High-Fat Diet
Hsp90aa1	Hsp90aa1/Hsp90aa1	<a href="#">granulomatous inflammation [MP:0002500]</a>	16	Female	B6N	Crow

Il10rb	Il10rb/Il10rb	<a href="#">folliculitis [MP:0000376]</a>	16	Female	B6N	Crow
Kdm4c	Kdm4c/Kdm4c	<a href="#">enlarged sebaceous gland [MP:0000652]</a>	16	Female	B6N	High-Fat Diet
Lrrg1	Lrrg1/Lrrg1	<a href="#">abnormal hair shedding [MP:0008861]</a>	16	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Mms22l	Mms22l/+	<a href="#">dermatitis [MP:0001194]</a>	16.6	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Myh9	Myh9/+	<a href="#">dermatitis [MP:0001194]</a>	15.4	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Nsun2	Nsun2/+	<a href="#">abnormal hair shedding [MP:0008861]</a>	16.3	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Orc3	Orc3/+	<a href="#">dermatitis [MP:0001194]</a>	16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Pfkl	Pfkl/+	<a href="#">hair follicle degeneration [MP:0003413], abnormal hair cycle catagen phase [MP:0008859]</a>	16	Female	B6N	Crow
Plk3cb	Plk3cb/Plk3cb	<a href="#">granulomatous inflammation [MP:0002500]</a>	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Plid5	Plid5/Plid5	<a href="#">dermatitis [MP:0001194]</a>	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Plekhm2	Plekhm2/Plekhm2	<a href="#">folliculitis [MP:0000376]</a>	16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Ppp3ca	Ppp3ca/Ppp3ca	<a href="#">folliculitis [MP:0000376]</a>	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Ptkcz	Ptkcz/Ptkcz	<a href="#">abnormal coat/hair pigmentation [MP:0002075]</a>	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Prmt3	Prmt3/Prmt3	<a href="#">abnormal hypodermis fat layer morphology [MP:0011156]</a>	16.4	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Rad18	Rad18/Rad18	<a href="#">abnormal arrector pili muscle morphology [MP:0010679]</a>	16	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Rars	Rars/+	<a href="#">dermatitis [MP:0001194]</a>	15.7	Female	B6N	High-Fat Diet
Rhou	Rhou/Rhou	<a href="#">dermatitis [MP:0001194]</a>	15.7	Female	B6N	High-Fat Diet
Ripk4	Ripk4/+	<a href="#">dermatitis [MP:0001194]</a>	15.9	Female	B6N	High-Fat Diet
Sag	Sag/Sag	<a href="#">dermatitis [MP:0001194]</a>	16	Female	B6N	Crow
Scn3b	Scn3b/Scn3b	<a href="#">myositis [MP:0004510]</a>	15.6	Female	B6N	High-Fat Diet
Sec16b	Sec16b/Sec16b	<a href="#">dermatitis [MP:0001194]</a>	16	Female	B6N	High-Fat Diet
Slc25a29	Slc25a29/Slc25a29	<a href="#">dermatitis [MP:0001194]</a>	15.6	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Spnb2	Spnb2/+	<a href="#">abnormal hair shaft morphology [MP:0003809]</a>	16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Suv420h1	Suv420h1/+	<a href="#">folliculitis [MP:0000376]</a>	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Ubp1	Ubp1/+	<a href="#">dermatitis [MP:0001194]</a>	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet

1100001H23Rik (Pbdt1)	1100001H23Rik/1100001H23 Rik	no	16	Female	Brd> B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	Diet High-Fat Diet
1190002H23Rik	1190002H23Rik/1190002H23 Rik	no	15.6	Female	B6N	High-Fat Diet
Fam134c	1300010M03Rik/1300010M03 Rik	no	15.6	Female	B6N	High-Fat Diet
1500011B03Rik	1500011B03Rik/1500011B03R ik	no	15.7	Female	B6N	High-Fat Diet
1500011B03Rik	1500011B03Rik/1500011B03R ik	no	16.3	Female	B6N	Chow
1700042B14Rik	1700042B14Rik/1700042B14R ik	no	16.3	Female	B6N	Chow
2010003O18Rik (Gatc)	2010003O18Rik/+ 2010107G12Rik/2010107G12 Rik	no	15.7	Female	B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	Chow High-Fat Diet
2010107G12Rik	2210016L21Rik/2210016L21Ri k	no	15.9	Female	B6N	High-Fat Diet
2210016L21Rik		no	16.1	Female	B6N	High-Fat Diet
2310046K01Rik	2310046K01Rik/+ 23100571J8Rik/23100571J8Ri	no	16.1	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	High-Fat Diet
23100571J8Rik	k	no	15.9	Female	B6N	High-Fat Diet
3010026O09Rik	3010026O09Rik/3010026O09 Rik	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	High-Fat Diet
Fam107b	3110001A13Rik/3110001A13R ik	no	16.1	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	High-Fat Diet
3110001I22Rik	3110001I22Rik/3110001I22Rik	no	16	Female	B6N	High-Fat Diet
3110035E14Rik	3110035E14Rik/3110035E14Ri	no	16	Female	B6N	Chow
3830417A13Rik	3830417A13Rik/3830417A13R ik	no	15.6	Female	B6N	Chow
4930471M23Rik	4930471M23Rik/4930471M23 Rik	no	15.9	Female	B6N	High-Fat Diet
4932414N04Rik	4932414N04Rik/4932414N04 Rik	no	15.9	Female	B6N	High-Fat Diet
4932438H23Rik	4932438H23Rik/4932438H23 Rik	no	15.9	Female	B6N	Chow
4933425L06Rik	4933425L06Rik/4933425L06Ri k	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	High-Fat Diet
5830415F09Rik	5830415F09Rik/5830415F09Ri k	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd>	High-Fat Diet
A430084P05Rik	A430084P05Rik/A430084P05R ik	no	16.3	Female	B6N	High-Fat Diet
A830019P07Rik	A830019P07Rik/A830019P07R no	no	15.9	Female	B6N	Chow

Abca4	ik	Abca4/Abca4	no	15.7	Female	B6N	Chow
Abcd1		Abcd1/Abcd1	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Abhd5		Abhd5/Abhd5	no	16.3	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Acot6		Acot6/Acot6	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Acpl2		Acpl2/Acpl2	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Acs14		Acs14/Acs14	no	16.4	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	Chow
Actn4		Actn4/+	no	16	Female	B6N	Chow
Actr6		Actr6/+	no	16.3	Female	B6N	Chow
Adam17		Adam17/+	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Adam3		Adam3/Adam3	no	16	Female	B6N	High-Fat Diet
Aft3		Aft3/Aft3	no	15.9	Female	B6N	Chow
Agpat3		Agpat3/+	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Agtr2		Agtr2/Agtr2	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Akap9		Akap9/Akap9	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Alch16a1		Alch16a1/Alch16a1	no	15.6	Female	B6N	Chow
Alch2		Alch2/Alch2	no	15.7	Female	B6	High-Fat Diet
Alg13		Alg13/Alg13	no	16.3	Female	B6N	Chow
Amfr		Amfr/Amfr	no	15.4	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Amotl1		Amotl1/Amotl1	no	15.1	Female	B6N	High-Fat Diet
Ampd3		Ampd3/Ampd3	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Anp32e		Anp32e/Anp32e	no	16.3	Female	B6N	High-Fat Diet
Anxa6		Anxa6/Anxa6	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Ap4e1		Ap4e1/Ap4e1	no	16.1	Female	B6N	Diet
Apol7a		Apol7a/Apol7a	no	15.9	Female	B6N	Chow
Apool		Apool/Apool	no	15.9	Female	B6N	Chow
App12		App12/App12	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet

Arhgap25	Arhgap25/Arhgap25	no	16.3	Female	B6N	High-Fat Diet
Arhgef4	Arhgef4/Arhgef4	no	16.1	Female	B6N	High-Fat Diet
Arhgef7	Arhgef7/+	no	16.1	Female	B6N	Crow
Artd4a	Artd4a/Artd4a	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Ar14d	Ar14d/Ar14d	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Arpc1b	Arpc1b/Arpc1b	no	15.9	Female	B6N	Crow
Arpc2	Arpc2/+	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Arpc3	Arpc3/+	no	16	Female	B6N	Crow
Arpc4	Arpc4/+	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Arsg	Arsg/Arsg	no	15.6	Female	B6N	Crow
Arvcf	Arvcf/Arvcf	no	15.7	Female	B6N	Crow
Aspm	Aspm/Aspm	no	16	Female	B6N	Crow
Aste1	Aste1/Aste1	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Asx1	Asx1/+	no	16	Female	B6N	High-Fat Diet
Atg16l1	Atg16l1/+	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Atp5a1	Atp5a1/+	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Atpaf2	Atpaf2/+	no	16	Female	B6N	High-Fat Diet
Atpif1	Atpif1/Atpif1	no	15.6	Female	B6N	Diet
B9d2	B9d2/+	no	15.6	Female	B6N	Crow
BC013712	BC013712/BC013712	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
BC017643	BC017643/BC017643	no	16.3	Female	B6N	High-Fat Diet
Bdh2	Bdh2/Bdh2	no	15.9	Female	B6N	High-Fat Diet
Blzf1	Blzf1/Blzf1	no	15.7	Female	B6N	Diet
Brdt	Brdt/Brdt	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Brpf3	Brpf3/+	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Btdb11	Btdb11/Btdb11	no	16	Female	B6N	Diet

Cabp1	Cabp1/Cabp1	no	16.1	Female	B6N	Chow
Calcoco2	Calcoco2/Calcoco2	no	15.3	Female	B6N	Chow
Cand2	Cand2/Cand2	no	15.9	Female	B6N	Chow
Caprin2	Caprin2/Caprin2	no	15.9	Female	B6N	Diet
Casp12	Casp12/Casp12	no	16	Female	B6N	Chow
Cbx5	Cbx5/Cbx5	no	15.7	Female	B6N	Chow
Cbx6	Cbx6/Cbx6	no	15.6	Female	B6N	Chow
CC2d2a	CC2d2a/+	no	15.7	Female	B6N	High-Fat Diet
Ccdc104	Ccdc104/Ccdc104	no	16.1	Female	B6N	High-Fat Diet
Ccdc106	Ccdc106/Ccdc106	no	15.9	Female	B6N	High-Fat Diet
Ccdc137	Ccdc137/+	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<-C-Brd>	Diet
Ccdc160	Ccdc160/Ccdc160	no	15.6	Female	B6N	Chow
Ccdc77	Ccdc77/Ccdc77	no	16.3	Female	B6N	Chow
Ccl22	Ccl22/+	no	16.1	Female	B6N	Chow
Cdc6	Cdc6/+	no	15.6	Female	B6N	High-Fat Diet
Cdca8	Cdca8/+	no	16	Female	B6N	High-Fat Diet
Cdk5rap2	Cdk5rap2/Cdk5rap2	no	15.4	Female	B6N	High-Fat Diet
Cdkn2a	Cdkn2a/Cdkn2a	no	16	Female	B6N	High-Fat Diet
Cds2	Cds2/+	no	16.3	Female	B6Br;d;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Cenpj	Cenpj/+	no	16.4	Female	B6Br;d;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Cenpj	Cenpj/Cenpj	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Cfh	Cfh/Cfh	no	15.7	Female	B6N	High-Fat Diet
Chd7	Chd7/+	no	15.7	Female	129S5;B6N	High-Fat Diet
Chia	Chia/Chia	no	16.6	Female	B6N	High-Fat Diet
Cish	Cish/Cish	no	15.6	Female	B6Br;d;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Clk1	Clk1/Clk1	no	15.9	Female	B6N	High-Fat Diet
Cluap1	Cluap1/+	no	15	Female	B6Br;d;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet



Cntnm6	Cntnm6/Cntnm6	no	16	Female	B6N	Chow
Hprt	CMV-Cre/CMV-Cre	no	16.1	Female	B6N	High-Fat Diet
Cntfr	Cntfr/+	no	16.4	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Cog2	Cog2/+	no	15.7	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Cog8	Cog8/+	no	16.1	Female	B6N	High-Fat Diet
Cog4	Cog4/+	no	15.7	Female	B6N	Chow
Cox6a1	Cox6a1/+	no	15.4	Female	B6N	High-Fat Diet
Creb3l1	Creb3l1(e)/Creb3l1(e)	no	15.6	Female	B6N	Chow
Thumpr3	CreER/CreER	no	16.3	Female	B6N	Chow
Crtf3(b)	Crtf3(b)/Crtf3(b)	no	16.3	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	Chow
Crtf3	Crtf3/Crtf3	no	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Csrp2bp	Csrp2bp/Csrp2bp	no	16	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Ctcf	Ctcf/+	no	16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Cyb561	Cyb561/Cyb561	no	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
D11Wsu99e	D11Wsu99e/D11Wsu99e	no	16.3	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Daam2	Daam2/Daam2	no	16.1	Female	B6N	Chow
Dbn1	Dbn1/Dbn1	no	16	Female	B6N	High-Fat Diet
Dcdc2c	Dcdc2c/Dcdc2c	no	15.7	Female	B6N	Chow
Dctn5	Dctn5/+	no	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Dcx	Dcx/Dcx	no	16	Female	B6N	Chow
Ddx27	Ddx27/+	no	16.3	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Dennd1c	Dennd1c/Dennd1c	no	15.9	Female	B6N	Chow
Dlg2	Dlg2/Dlg2	no	16	Female	B6N	Chow
Dlg4	Dlg4/Dlg4	no	15.9	Female	B6Jlco;129P2	High-Fat Diet
Mir96	Dmdo/Dmdo	no	15.3	Female	C3Fe	High-Fat Diet
Dmx12	Dmx12/+	no	15.7	Female	B6N	Chow
Dnase12	Dnase12/Dnase12	no	16.4	Female	B6N	High-Fat Diet

Donson	Donson/+	no	15.7	Female	B6N	High-Fat Diet
Dopey2	Dopey2/Dopey2	no	15.1	Female	B6N	Chow
Dph2	Dph2/+	no	15.9	Female	B6N	Chow
Dscc1	Dscc1/Dscc1	no	16.1	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Dsg1b	Dsg1b/Dsg1b	no	16	Female	B6N	Chow
Dusp1	Dusp1/Dusp1	no	16	Female	B6N	High-Fat Diet
Dusp26	Dusp26/Dusp26	no	15.9	Female	B6N	High-Fat Diet
Dusp3	Dusp3/Dusp3	no	15.9	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Dusp4	Dusp4/Dusp4	no	15.9	Female	B6N	High-Fat Diet
Dync12	Dync12/+	no	16	Female	B6N	Diet
Dynl1	Dynl1/+	no	15.9	Female	B6N	Chow
Dynrb1	Dynrb1/+	no	16	Female	B6N	Chow
Fars2	Fars2/+	no	16.1	Female	B6N	High-Fat Diet
Efn1	Efn1/Efn1	no				
Egfr	Egfr/+	no	16	Female	B6N	High-Fat Diet
Ehd1	Ehd1/+	no	15.7	Female	B6N	High-Fat Diet
Eif2b2	Eif2b2/+	no	16.4	Female	B6N	High-Fat Diet
Eif4e3	Eif4e3/Eif4e3	no	15.9	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Elk4	Elk4/Elk4	no	15.7	Female	B6	Diet
Eil2	Eil2/Eil2	no	16	Female	B6N	Chow
Elmo1	Elmo1/Elmo1	no	16.6	Female	B6N	Chow
Entpd1	Entpd1/Entpd1	no	16	Female	B6	High-Fat Diet
Esc1	Esc1/Esc1	no	15.7	Female	B6N	High-Fat Diet
Esc2	Esc2/+	no	15.4	Female	B6N	High-Fat Diet
Espn	Espn/Espn	no	15.3	Female	B6N	High-Fat Diet
Exoc3l2	Exoc3l2/Exoc3l2	no	15.7	Female	B6N	Diet
Expi	Expi/Expi	no	16.3	Female	B6N	Chow

Ezh2 Ezf	Ezh2/+ Ezf/+	no no	15.7 16	Female Female	B6N B6N	High-Fat Diet Chow High-Fat Diet Chow High-Fat Diet
Fahd2a Fam175b	Fahd2a/Fahd2a Fam175b/Fam175b	no no	16 15.9	Female Female	B6N B6N	High-Fat Diet Chow High-Fat Diet
Fam63a Fam96a	Fam63a/Fam63a Fam96a/Fam96a	no no	15.6 16.4	Female Female	B6 B6N	High-Fat Diet Chow High-Fat Diet
Farsa	Farsa/+	no	16.1	Female	B6N	Chow
Fbx17	Fbx17/Fbx17	no	16.1	Female	B6N	Chow
Fbxo47	Fbxo47/Fbxo47	no	15.4	Female	B6N	Chow
Fbxo7	Fbxo7/Fbxo7	no	15.6	Female	B6N	Chow
Fg8y	Fg8y/Fg8y	no	16.3	Female	B6N	Chow
ROSA26	FLPeR/FLPeR	no	15.9	Female	B6N	High-Fat Diet
Fto	Fto/Fto	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<C- Brd>	High-Fat Diet
Fyn	Fyn/+	no	15.6	Female	B6N	Chow
Gap43	Gap43/+	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C- Brd>	High-Fat Diet
Gba2	Gba2/Gba2	no	16.7	Female	B6N	High-Fat Diet
Gbf1	Gbf1/+	no	15.7	Female	B6N	Chow
Gclc	Gclc/Gclc	no	15.9	Female	B6N	High-Fat Diet
Gfm1	Gfm1/+	no	15.7	Female	B6N	High-Fat Diet
Gif	Gif/Gif	no	16.1	Female	B6N B6Br;d;B6Dnk;B6N-Tyr<C- Brd>	High-Fat Diet High-Fat Diet
Git2	Git2/Git2	no	16	Female	CBA/CA;129P2	High-Fat Diet
Glg1	Glg1/+	no	15.1	Female	B6N	Chow
Git8d2	Git8d2/Git8d2	no	15.9	Female	B6N	High-Fat Diet
Gm13547	Gm13547/Gm13547	no	15.7	Female	B6N	Chow
Gm239	Gm239/Gm239	no	16	Female	B6N	Chow
Gmn	Gmn/+	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<C- Brd>	High-Fat Diet
Gpc5	Gpc5/Gpc5	no	15.3	Female	B6N	Chow
Gpc6	Gpc6/Gpc6	no	16.1	Female	B6N	High-Fat Diet
Gprc5b	Gprc5b/Gprc5b	no	16.3	Female	B6N	Chow
Grb2	Grb2/+	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C- Brd>	High-Fat Diet

			Brd>			Diet
Grn1	Grn1/+	no	15.1	Female	B6N	High-Fat Diet
	Gsk3a/Gsk3a	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
	Gys2/Gys2	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
	H2-M5/H2-M5	no	16.4	Female	B6N	High-Fat Diet
Hac1	Hac1/Hac1	no	15.7	Female	B6N	Chow
Hdac1	Hdac1/+	no	15.6	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Hdac8	Hdac8/Hdac8	no	15.9	Female	B6N	High-Fat Diet
Heatr8	Heatr8/Heatr8	no	15.7	Female	B6N	High-Fat Diet
Hira	Hira/+	no	15.9	Female	B6N	High-Fat Diet
Hp1bp3	Hp1bp3/Hp1bp3	no	15.7	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Hmx3	hx/hx	no	14	Female	STOCK Hmx3<hx>	Chow
Ido1	Ido1/Ido1	no	16	Female	B6N	High-Fat Diet
Ido2	Ido2/Ido2	no	16.1	Female	B6N	Chow
Ifitm3	Ifitm3/Ifitm3	no	16.7	Female	B6lco;B6Brd;129P2-Tyr<C-Brd>	High-Fat Diet
Ifi80	Ifi80/Ifi80	no	16.1	Female	B6N	Chow
Ikbb	Ikbb/+	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Il22ra1	Il22ra1/Il22ra1	no	16.4	Female	B6N	High-Fat Diet
Imp1	Imp1/Imp1	no	16.3	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Ints12	Ints12/+	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Irf1	Irf1/Irf1	no	15.9	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Irf7	Irf7/Irf7	no	16	Female	B6N	Chow
Jarid2	Jarid2/+	no	16.4	Female	B6	High-Fat Diet
Jmjd5	Jmjd5/+	no	15.6	Female	B6N	Chow
Kazn	Kazn/Kazn	no	15.9	Female	B6N	High-Fat Diet
Kcnv2	Kcnv2/Kcnv2	no	15.9	Female	B6N	Chow
Kdm4b	Kdm4b/+	no	15.9	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet





						Brd> B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet
Ndufr33	Ndufr33/+	no	16	Female		
Neif	Neif/Neif	no	16.1	Female	B6N High-Fat Diet	
Nfkb1	Nfkb1/Nfkb1	no	16.3	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	
						B6 B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet
Nfya	Nfya/+	no	15.9	Female		
Ninj2	Ninj2/Ninj2	no	15.9	Female	B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	
Nin1	Nin1/Nin1	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	
Nipa1	Nipa1/Nipa1	no	16.1	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet
Nkiras2	Nkiras2/Nkiras2	no	16.1	Female		
Nploc4	Nploc4/+	no	15.7	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	
Nsun2	Nsun2/Nsun2	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet
Oas1g	Oas1g/Oas1g	no	16	Female	B6N Chow	
Oas12	Oas12/Oas12	no	15.7	Female	High-Fat Diet	
Ogfod1	Ogfod1/Ogfod1	no	15.9	Female	High-Fat Diet	
Optn	Optn/Optn	no	15.6	Female	B6N High-Fat Diet	B6N Diet
Otub2	Otub2/Otub2	no	15.9	Female	B6N High-Fat Diet	
Otub6b	Otub6b/Otub6b	no	15.9	Female	B6 High-Fat Diet	
Oxt1	Oxt1/+	no	15.6	Female	B6N Chow	
Pabpc1l	Pabpc1l/Pabpc1l	no	16	Female	B6N High-Fat Diet	B6N B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet
Pabpc4	Pabpc4/Pabpc4	no	16.3	Female	B6N High-Fat Diet	
Pabpn1l	Pabpn1l/Pabpn1l	no	15.6	Female	B6N Chow	
Pacsin3	Pacsin3/Pacsin3	no	15.7	Female	B6N Chow	
Pex3	Pex3/Pex3	no	16	Female	B6N High-Fat Diet	B6N B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet
Pfn1	Pfn1/+	no	15.9	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	
Pht20	Pht20/Pht20	no	16.1	Female	B6Brd;B6Dnk;B6N-Tyr<-C- Brd> High-Fat Diet	
Phtf2	Phtf2/Phtf2	no	16.1	Female	B6N Chow	

Plk3cb	Plk3cb/+	no	15.7	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Plk3cg	Plk3cg/Plk3cg	no	16.3	Female	B6N	High-Fat Diet	
Plk3r2	Plk3r2/Plk3r2	no	15.9	Female	B6N	High-Fat Diet	
Pip5k1c	Pip5k1c/+	no	15.9	Female	B6N	High-Fat Diet	
Pitpnm1	Pitpnm1/Pitpnm1	no	16.1	Female	B6N	High-Fat Diet	
Plekhhg1	Plekhhg1/Plekhhg1	no	16.1	Female	B6N	High-Fat Diet	
Plin2	Plin2/Plin2	no	16	Female	B6N	High-Fat Diet	
Plxnb2	Plxnb2/+	no	15.7	Female	B6N	High-Fat Diet	
Pnpo	Pnpo/+	no	16.7	Female	B6 B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Pnpt1	Pnpt1/+	no	16.4	Female	B6N	High-Fat Diet	
Polg2	Polg2/+	no	15.9	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Polr1e	Polr1e/+	folliculitis [MP:0000376], chronic inflammation [MP:0002499]			15.9	Female	High-Fat Diet
Ppp2r2b	Ppp2r2b/Ppp2r2b	no	16.1	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Ppp5c	Ppp5c/Ppp5c	no	16.1	Female	B6N	Chow	
Praf2	Praf2(e)/Praf2(e)	no	16.4	Female	B6N	High-Fat Diet	
Prkab1	Prkab1/Prkab1	no	16.1	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Prmt1	Prmt1/+	no	15.9	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Prmt2	Prmt2/Prmt2	no	15.9	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Prmt5	Prmt5/+	no	16	Female	B6N	Chow	
Prps1	Prps1/+	no	15.7	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Prpsap2	Prpsap2/Prpsap2	no	16	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Trnprs15	Prss7/Prss7	no	15.9	Female	B6N B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet	
Psat1	Psat1/+	no	15.6	Female	B6Jlco	High-Fat Diet	
Dlg2	PSD93/PSD93	no	16	Female	B6Jlco	High-Fat Diet	



Ptges2	Ptges2/Ptges2	no	15.7	Female	B6N	High-Fat Diet
Phn2r	Phn2r/Phn2r	no	16.1	Female	B6N	Chow
Pydn2	Pydn2/+	no	15.7	Female	B6N	High-Fat Diet
Pus7l	Pus7l/Pus7l	no	15.7	Female	B6N	Diet
Rab15	Rab15/Rab15	no	16.1	Female	B6N	High-Fat Diet
Rab5c	Rab5c/+	no	15.9	Female	B6N	Chow
Rab7l1	Rab7l1/Rab7l1	no	16	Female	B6N	High-Fat Diet
Ralb	Ralb/Ralb	no	15.7	Female	B6N	Diet
Ralgapb	Ralgapb/+	no	15.7	Female	B6N	Chow
Ragef1l	Ragef1l/Ragef1l	no	15.7	Female	B6N	High-Fat Diet
Rasal2	Rasal2/Rasal2	no	15.7	Female	B6N	Chow
Rbbp7	Rbbp7/Rbbp7	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
RbmX	RbmX/RbmX	no	16	Female	B6N	Chow
Rcor2	Rcor2/Rcor2	no	15.9	Female	B6N	High-Fat Diet
Rdh16	Rdh16/Rdh16	no	16.1	Female	B6N	Diet
Ren1	REKO/REKO	no	16.4	Female	B6Jlco	High-Fat Diet
Rftn2	Rftn2/Rftn2	no	15.9	Female	B6N	Diet
Rhd	Rhd/Rhd	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Rhobtb3	Rhobtb3/Rhobtb3	no	15.1	Female	B6Br;d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Rhot1	rhot1/+	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet
Rhot2	Rhot2/Rhot2	no	16	Female	B6N	Diet
Rhox13	Rhox13/Rhox13	no	15.7	Female	B6N	Chow
Ripply3	Ripply3/Ripply3	no	15.9	Female	B6N	High-Fat Diet
Rnaseh2b	Rnaseh2b/+	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<c-Brd>	Diet
Rnaseh2c	Rnaseh2c/+	no	15.7	Female	B6N	High-Fat Diet
Rnf10	Rnf10/Rnf10	no	16	Female	B6N	Chow
Rpap2	Rpap2/+	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<c-Brd>	High-Fat Diet

Rpgrip1l	Rpgrip1/+	no	16.3	Female	B6N	Chow
Rpn2	Rpn2/+	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Rqcd1	Rqcd1/+	no	16.4	Female	B6N	High-Fat Diet
Ribdn	Ribdn/Ribdn	no	15.6	Female	B6N	Chow
Rtf1	Rtf1/+	no	16.1	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Rufy2	Rufy2/Rufy2	no	15.7	Female	B6N	Chow
S100b	S100b/S100b	no	16.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Uba2	Sae2/+	no	15.6	Female	B6N	High-Fat Diet
Sar1b	Sar1b/+	no	15.7	Female	B6N	Diet
Sc4mol	Sc4mol/+	no	15.9	Female	B6N	High-Fat Diet
Scot1	Scot1/+	no	16.1	Female	B6N	High-Fat Diet
Sdhc	Sdhc/+	no	15.9	Female	B6N	High-Fat Diet
Sec1	Sec1/Sec1	no	15.6	Female	B6N	Chow
Sec23a	Sec23a/+	no	15.6	Female	B6N	High-Fat Diet
Sec24a	Sec24a/Sec24a	no				
Secisbp2	Secisbp2/+	no	16.1	Female	B6N	High-Fat Diet
Seh1l	Seh1l/+	no	15.6	Female	B6N	Diet
Sei1l	Sei1l/+	no	16.4	Female	B6N	Chow
Seik	Seik/Seik	no	15.9	Female	B6N	Chow
Sepn1	Sepn1/Sepn1	no	16.4	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Setdb1	Setdb1/+	no	16.3	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Setmar	Setmar/Setmar	no	16	Female	B6N	High-Fat Diet
Sgol2	Sgol2/Sgol2	no	16	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	Diet
Slc22a21	Slc22a21/Slc22a21	no	15.7	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Sirt2	Sirt2/Sirt2	no	15.9	Female	B6Br;d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Sirt3	Sirt3/Sirt3	no	15.9	Female	B6N	Diet
Skida1	Skida1/+	no	16.1	Female	B6N	Chow

Slc16a6	Slc16a6/Slc16a6	no	15.9	Female	B6N	Crow
Slc25a21	Slc25a21/Slc25a21	no	16.1	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Slc25a4	Slc25a4/Slc25a4	no	16.1	Female	B6N	Crow
Slc25a43	Slc25a43/Slc25a43	no	16	Female	B6N	High-Fat Diet
Slc30a8	Slc30a8/Slc30a8	no	15.9	Female	B6N	High-Fat Diet
Slc35f1	Slc35f1/Slc35f1	no	15.9	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	Diet
Slc38a10	Slc38a10/Slc38a10	no	15.6	Female	B6N	Crow
Slc41a3	Slc41a3/Slc41a3	no	16.4	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Slc44a5	Slc44a5/Slc44a5	no	15.7	Female	B6N	High-Fat Diet
Slc5a2	Slc5a2/Slc5a2	no	15.9	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Slc5a6	Slc5a6/+	no	15.1	Female	B6N	High-Fat Diet
Slc9a8	Slc9a8/Slc9a8	no	16.1	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Slco2b1	Slco2b1/Slco2b1	no	15.7	Female	B6N	High-Fat Diet
Slx4	Slx4/Slx4	no	16.3	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Smnrcal1	Smnrcal1/Smnrcal1	no	16.1	Female	B6N	High-Fat Diet
Smc3	Smc3/+	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Sms	Sms/+	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Smyd3	Smyd3/Smyd3	no	15.4	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Smyd4	smyd4/smyd4	no	15.9	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Smyd5	Smyd5/Smyd5	no	16.1	Female	B6N	Crow
Snap29	Snap29/+	no	15.7	Female	B6N	High-Fat Diet
Snap47	Snap47/Snap47	no	15.9	Female	B6N	Crow
Snapc4	Snapc4/+	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Snf8	Snf8/+	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Snx5	Snx5/Snx5	no	16	Female	B6Brd;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet

Sparc	Sparc/Sparc	no	15.6	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Spns2	Spns2/spns2	no	16.3	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	Crow
Spopl	Spopl/Spopl	no	16	Female	B6N	Crow
Sqstm1	Sqstm1/Sqstm1	no	15.6	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Src	Src/+	no	15.7	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Srsf4	Srsf4/Srsf4	no	15.7	Female	B6N	High-Fat Diet
St18	St18/St18	no	15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Stard13	Stard13/Stard13	no	15.6	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Stard5	Stard5/Stard5	no	16	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Stard6	Stard6/Stard6	no	15.7	Female	B6N	High-Fat Diet
Stk32a	Stk32a/Stk32a	no	15.9	Female	B6N	High-Fat Diet
Stk39	Stk39/Stk39	no	16	Female	B6N	High-Fat Diet
Supt7l	Supt7l/+	no	15.7	Female	B6N	Diet
Supv31l	Supv31l/+	no	14.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Suv420h1	Suv420h1/Suv420h1	no	15.6	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Sympk	Sympk/+	no	16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Synj2	Synj2/Synj2	no	15.9	Female	B6N	Diet
Syt1	Syt1/+	no	16.7	Female	B6N	Crow
Syt1	Syt1/Syt1	no	16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Tbcl10a	Tbcl10a/Tbcl10a	no	15.7	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Tcf4	Tcf4/Tcf4	no	16	Female	B6N	Crow
Tff1	Tff1/Tff1	no	16	Female	B6N	Crow
Tm9sf4	Tm9sf4/Tm9sf4	no	16	Female	B6N	High-Fat Diet
Tmc6	Tmc6/Tmc6	no	16	Female	B6N	Crow
Tmem165	Tmem165/Tmem165	no	16	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet

Trim189	Trim189/Trim189	no	no	15.9	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Trim9	Trim9/Trim9	no	no	15.9	Female	B6N	High-Fat Diet
Trim98	Trim98/+	no	no	16.6	Female	B6N	High-Fat Diet
Trimod1	Trimod1/+	no	no	16	Female	B6N	High-Fat Diet
Trfaip1 Trnk Tomm20l	Trfaip1/Trfaip1 Trnk/Trnk Tomm20l/Tomm20l	no no no	no	15.7 15.7 15.6	Female Female Female	B6N B6N B6N	Diet Chow High-Fat Diet
Trpd52l2	Trpd52l2/Trpd52l2	no	no	15.7	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Trpl1	Trpl1/+	no	no	16.1	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Traf3ip3	Traf3ip3/Traf3ip3	no	no	16	Female	B6N	High-Fat Diet
Trex1	Trex1/+	no	no	15.9	Female	B6N	High-Fat Diet
Trim17	Trim17/Trim17	no	no	15.9	Female	B6N	Chow
Trim29	Trim29/Trim29	no	no	16	Female	B6N	Chow
Trim45	Trim45/+	no	no	15.9	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Trim66	Trim66/Trim66	no	no	16.4	Female	B6N	High-Fat Diet
Trpc2	Trpc2/Trpc2	no	no	16.1	Female	B6N	Chow
Trpc4ap	Trpc4ap/+	no	no	16	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Tsfm	Tsfm/+	no	no	16.4	Female	B6N	Chow
Tsga14	Tsga14/Tsga14	no	no	16.9	Female	B6N	High-Fat Diet
Ttll3	Ttll3/Ttll3	no	no	16	Female	B6Br;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
Ttll4	Ttll4/Ttll4	no	no	16.1	Female	B6N	High-Fat Diet
Twf1	Twf1/Twf1	no	no	15.9	Female	B6N	High-Fat Diet
Uck1	Uck1/Uck1	no	no	15.7	Female	B6N	High-Fat Diet
Uhf1	Uhf1/+	no	no	15.9	Female	B6N	High-Fat Diet
Ush1c	Ush1c/+	no	no	16	Female	B6N	High-Fat Diet

Ush1c	Ush1c/Ush1c	no	15.7	Female	B6N	Chow
Usp22	Usp22/+	no	15.9	Female	B6N	Chow
Usp33	Usp33/Usp33	no	15.9	Female	B6N	High-Fat Diet
Usp4	Usp4/Usp4	no	16	Female	B6N	High-Fat Diet
Usp42	Usp42/Usp42	no	15.4	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Vangl1	Vangl1/Vangl1	no	16.3	Female	B6N	High-Fat Diet
Vps53	Vps53/+	no	15.9	Female	B6N	Chow
Vwa5b1	Vwa5b1/Vwa5b1	no	15.9	Female	B6N	High-Fat Diet
Wbp2	Wbp2/Wbp2	no	15.7	Female	B6N	High-Fat Diet
Wdr37	Wdr37/Wdr37	no	16.6	Female	B6N	Chow
Whrn	Whrn/Whrn	no	16	Female	B6N	High-Fat Diet
Wnt3	Wnt3/+	no	16.1	Female	B6N	Chow
Xbp1	Xbp1/+	no	15.9	Female	B6N	Chow
Ywhae	Ywhae/Ywhae	no	15.7	Female	OTHER	Chow
Zc3hc1	Zc3hc1/Zc3hc1	no	16	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Zfp106	Zfp106/Zfp106	no	12.1	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Zfp182	Zfp182/Zfp182	no	15.6	Female	B6N	Chow
zfp184	zfp184/zfp184	no	15.3	Female	B6	High-Fat Diet
Zfp367	Zfp367/Zfp367	no	16	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Zkscan14	Zkscan14/+	no	15.9	Female	B6N	High-Fat Diet
Zmynd8	Zmynd8/+	no	16.3	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
Zp1	Zp1/Zp1	no	15.9	Female	B6N	Diet
Zscan10	Zscan10/+	no	15.6	Female	B6N	Chow
Zz23	Zz23/+	no	16.1	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
WT	+/+	no	16	Female	B6N	High-Fat Diet
WT	+/+	no	15.6	Female	B6Br;B6Dnk;B6N-Tyr<-C-Brd>	High-Fat Diet
WT	+/+	no	15.9	Female	B6	High-Fat Diet

WT	+/+	no		16	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		16.7	Female	B6	High-Fat Diet
WT	+/+	no		16.1	Female	B6Jico	High-Fat Diet
WT	+/+	no		16	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.7	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		16	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		13.9	Female	STOCK Hmx3<hx>	Crow
WT	+/+	no		16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.7	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.5	Female	B6N	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		16.1	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.9	Female	B6N	High-Fat Diet
WT	+/+	no		15.7	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		15.4	Female	B6N	High-Fat Diet
WT	+/+	no		15.9	Female	B6Br.d;B6Dnk;B6N-Tyr<C-Brd>	High-Fat Diet
WT	+/+	no		16.3	Female	B6N	High-Fat Diet
WT	+/+	no		16.1	Female	B6N	High-Fat Diet
WT	+/+	no		15.9	Female	B6N	High-Fat Diet





**Supplementary Table 3. Genes and MP terms discovered across all tests.** Each screen identified genes that were unique to the specific test, thus increasing the number of genes flagged. Unique MP terms were also described by adding different tests.

Genes Identified by Skin Screens	Hair follicle cycling	Histopathology
Krt76	Myo5a	Aldh18a1
Lrig1	Nf1	Anks4b
Myo5a	Nom1	Arpc1b
Myo6	Nsun2	Ccdc57
Mysm1	Trpc4ap	Farp2
Nsun2		Kdm4c
Prkab1		Krt76
Sparc		Lrig1
Vangl1		Nsun2
		PKl
		PKcz
		Prmt3
		Rad18
		Scn3b
		unique genes: 23

MP terms Identified by Skin Screens	Phenotype Term	TEST
MP:0000367	abnormal coat/ hair morphology	Dysmorphology
MP:0000371	abnormal coat/hair pigmentation	Dysmorphology
MP:0000373	belly spot	Dysmorphology
MP:0000416	sparse hair	Dysmorphology
MP:0000418	localized hair loss	Dysmorphology
MP:0000575	dark foot pads	Dysmorphology
MP:0001191	abnormal skin condition	Dysmorphology
MP:0001192	scaly skin	Dysmorphology
MP:0001510	abnormal coat appearance	Dysmorphology
MP:0002075	diluted coat color	Dysmorphology
MP:0003849	greasy coat	Dysmorphology
MP:0010096	abnormal incisor color	Dysmorphology
MP:0010179	rough coat	Dysmorphology
(no MP term)	non-synchronous follicle cycling	Hair Follicle Cycling
MP:0008838	abnormal hair cycle anagen phase	Hair Follicle cycling
MP:0010137	delayed exit from anagen phase	Hair Follicle cycling
MP:0008861	abnormal hair shedding	Hair Follicle cycling/Histopathology
MP:0000652	enlarged sebaceous gland	Histopathology
MP:0003809	abnormal hair shaft morphology	Histopathology
MP:0009932	skin fibrosis	Histopathology
MP:0010206	pigment incontinence	Histopathology
MP:0010679	abnormal arrector pili muscle morphology	Histopathology
MP:0011156	abnormal hypodermis fat layer morphology	Histopathology
(no MP term)	premature exogen	Histopathology
MP:0004510	myositis	Histopathology

**Supplementary Table 4. Screen parameters.** The Sanger-MGP pipeline comprises 263 parameters, with 43 directly relating to the integument. Parameters relating to skin, hair, teeth, and nails are indicated with a 1, whereas non skin parameters are indicated with a 0.

Parameter	Skin Specific?
Homozygous Viability at P14	0
Fertility	0
Auditory Brainstem Response Threshold	0
Body Composition (DEXA) Body Weight	0
Body Composition (DEXA) Bone Mineral Content	0
Body Composition (DEXA) Bone Mineral Density	0
Body Composition (DEXA) Fat Mass	0
Body Composition (DEXA) Fat Percentage Estimate	0
Body Composition (DEXA) Lean Mass	0
Body Composition (DEXA) Nose to Tailbase Length	0
Citrobacter Challenge Bacterial Shedding	0
Citrobacter Challenge Cecal Contents	0
Citrobacter Challenge Cecal Patch	0
Citrobacter Challenge Cecum	0
Citrobacter Challenge Colon	0
Citrobacter Challenge Colon Weight	0
Citrobacter Challenge Liver	0
Citrobacter Challenge Spleen	0
Dysmorphology Anagen Phase	1
Dysmorphology Coat Colouration	1
Dysmorphology Coat Intact	1
Dysmorphology Coat Texture	1
Dysmorphology Dorsal Coat	1
Dysmorphology Dorsal Coat Coarse	1
Dysmorphology Dorsal Coat Fine	1
Dysmorphology Dorsal Coat Long	1
Dysmorphology Dorsal Coat Rough	1
Dysmorphology Dorsal Coat Short	1
Dysmorphology Dorsal Coat Skin Visible	1
Dysmorphology Dorsal Coat Sparse	1
Dysmorphology Ear Morphology	0
Dysmorphology Ear Size	0
Dysmorphology Eye Colouration	0
Dysmorphology Eye Size	0
Dysmorphology Eyes	0
Dysmorphology Forelimb Size	0
Dysmorphology Forepaw Foot Pads	1
Dysmorphology Forepaw Nails	1
Dysmorphology Forepaw Size	0
Dysmorphology Genitalia	0
Dysmorphology Genitalia Morphology	0
Dysmorphology Genitalia Size	0
Dysmorphology Guard Hairs	1
Dysmorphology Hair Behind Ears	1
Dysmorphology Hair On Tail	1
Dysmorphology Head Morphology	0
Dysmorphology Head Size	0
Dysmorphology Hind Paw Size	0
Dysmorphology Hindlimb Foot Pads	1
Dysmorphology Hindlimb Nails	1
Dysmorphology Hindlimb Size	0
Dysmorphology Incisor	1
Dysmorphology Incisor Colour	1

Dysmorphology Mouth Morphology	0
Dysmorphology Skin Condition	1
Dysmorphology Skin Pigmentation	1
Dysmorphology Snout Morphology	0
Dysmorphology Tail Length	0
Dysmorphology Tail Morphology	1
Dysmorphology Tail Tip Hairs	1
Dysmorphology Ventral Coat Coarse	1
Dysmorphology Ventral Coat Fine	1
Dysmorphology Ventral Coat Long	1
Dysmorphology Ventral Coat Rough	1
Dysmorphology Ventral Coat Short	1
Dysmorphology Ventral Coat Skin Visible	1
Dysmorphology Ventral Coat Sparse	1
Dysmorphology Ventral coat	1
Dysmorphology Vibrissae	1
Dysmorphology Vibrissae Absent	1
Dysmorphology Vibrissae Curly	1
Dysmorphology Vibrissae Disorientated	1
Dysmorphology Vibrissae Long	1
Dysmorphology Vibrissae Short	1
Dysmorphology Whiskers	1
Erythrocyte micronuclei	0
Eye Morphology Blood Presence	0
Eye Morphology Cornea	0
Eye Morphology Cornea Opacity	0
Eye Morphology Cornea Vascularization	0
Eye Morphology Eyelid Closure	0
Eye Morphology Eyes Bulging	0
Eye Morphology Iris Pigment	0
Eye Morphology Left Eye Blood Vessels	0
Eye Morphology Left Eye Blood Vessels Pattern	0
Eye Morphology Left Eye Blood Vessels Structure	0
Eye Morphology Left Optic Disc	0
Eye Morphology Left Retina	0
Eye Morphology Left Retina Pigmentation	0
Eye Morphology Left Retina Structure	0
Eye Morphology Left eye Bergmeister's Papilla	0
Eye Morphology Lens	0
Eye Morphology Lens Opacity (non suture or snowflake)	0
Eye Morphology Pupil	0
Eye Morphology Pupil Dilation	0
Eye Morphology Pupil Light Response	0
Eye Morphology Pupil Position	0
Eye Morphology Pupil Shape	0
Eye Morphology Right Eye Blood Vessels	0
Eye Morphology Right Eye Blood Vessels Pattern	0
Eye Morphology Right Eye Blood Vessels Structure	0
Eye Morphology Right Optic Disc	0
Eye Morphology Right Retina	0
Eye Morphology Right Retina Pigmentation	0
Eye Morphology Right Retina Structure	0
Eye Morphology Right eye Bergmeister's Papilla	0
Glucose Tolerance (ip) Fasted Plasma Glucose	0
Glucose Tolerance (ip) ip-GTT	0
Grip Strength All Paws Trial 1	0
Grip Strength All Paws Trial 2	0
Grip Strength All Paws Trial 3	0
Grip Strength Fore Paws Trial 1	0

Grip Strength Fore Paws Trial 2	0
Grip Strength Fore Paws Trial 3	0
Haematology (CBC) Haematocrit	0
Haematology (CBC) Haemoglobin	0
Haematology (CBC) Mean Corpuscular Haemoglobin	0
Haematology (CBC) Mean Corpuscular Haemoglobin Conc.	0
Haematology (CBC) Mean Corpuscular Volume	0
Haematology (CBC) Mean Platelet Volume	0
Haematology (CBC) Platelet Count	0
Haematology (CBC) Red Blood Cell Count	0
Haematology (CBC) Red Blood Cell Distribution Width	0
Haematology (CBC) White Blood Cell Count	0
Heart Weight	0
Hot Plate Latency to respond	0
Indirect Calorimetry Cumulative Food Intake	0
Indirect Calorimetry Energy Expenditure	0
Indirect Calorimetry Energy Expenditure/Animal	0
Indirect Calorimetry Respiratory Exchange Ratio	0
Indirect Calorimetry VCO2	0
Indirect Calorimetry VO2	0
Indirect Calorimetry Water Intake	0
Indirect Calorimetry X-Activity Ambulatory	0
Indirect Calorimetry X-Activity Total	0
Modified SHIRPA Body Position	0
Modified SHIRPA Contact Righting Reflex	0
Modified SHIRPA Convulsions	0
Modified SHIRPA Corneal Touch Reflex	0
Modified SHIRPA Evidence of Biting	0
Modified SHIRPA Gait (inc. Ataxia)	0
Modified SHIRPA Headbobbing/Circling	0
Modified SHIRPA Lacrimation	0
Modified SHIRPA Limb Grasping	0
Modified SHIRPA Locomotor Activity	0
Modified SHIRPA Palpebral Closure	0
Modified SHIRPA Pelvic Elevation	0
Modified SHIRPA Pinna Touch Reflex	0
Modified SHIRPA Positional Passivity	0
Modified SHIRPA Startle Response	0
Modified SHIRPA Tail Elevation	0
Modified SHIRPA Touch Escape	0
Modified SHIRPA Transfer Arousal	0
Modified SHIRPA Tremor	0
Modified SHIRPA Trunk Curl	0
Modified SHIRPA Vocalisation	0
Open Field % Distance Moved in Centre - Total	0
Open Field % Time Spent in Centre - Total	0
Open Field Average Speed at Centre	0
Open Field Average Speed at Periphery	0
Open Field Latency To Enter Centre	0
Open Field Number Entries to Centre	0
Open Field Total Distance Travelled - Centre	0
Open Field Total Distance Travelled - Periphery	0
Open Field Total Distance Travelled Whole Arena	0
Open Field Total Rears	0
Open Field Total Rest Time at Periphery	0
Open Field Total Resting Time in Centre	0
Open Field Total Time at Periphery	0
Open Field Total Time in Centre	0
Open Field Whole Arena Average Speed	0

Open Field Whole Arena Resting Time	0
Peripheral blood lymphocyte FACS Total T cell	0
Peripheral blood lymphocyte FACS T helper	0
Peripheral blood lymphocyte FACS T cytotoxic	0
Peripheral blood lymphocyte FACS T regulatory	0
Peripheral blood lymphocyte FACS Activated T helper	0
Peripheral blood lymphocyte FACS Activated T cytotoxic	0
Peripheral blood lymphocyte FACS NK	0
Peripheral blood lymphocyte FACS NKT	0
Peripheral blood lymphocyte FACS Total B cell	0
Peripheral blood lymphocyte FACS Mature B cell	0
Peripheral blood lymphocyte FACS Monocytes	0
Peripheral blood lymphocyte FACS Granulocytes	0
Plasma Chemistry Alanine Aminotransferase	0
Plasma Chemistry Albumin	0
Plasma Chemistry Alkaline Phosphatase	0
Plasma Chemistry Amylase	0
Plasma Chemistry Aspartate Aminotransferase	0
Plasma Chemistry Calcium	0
Plasma Chemistry Chloride	0
Plasma Chemistry Cholesterol	0
Plasma Chemistry Creatine Kinase	0
Plasma Chemistry Creatinine	0
Plasma Chemistry Fructosamine	0
Plasma Chemistry Glucose	0
Plasma Chemistry Glycerol	0
Plasma Chemistry High Density Lipoprotein	0
Plasma Chemistry Iron	0
Plasma Chemistry Lactate Dehydrogenase	0
Plasma Chemistry Low Density Lipoprotein	0
Plasma Chemistry Magnesium	0
Plasma Chemistry NEFAC	0
Plasma Chemistry Phosphate	0
Plasma Chemistry Potassium	0
Plasma Chemistry Sodium	0
Plasma Chemistry Total Bilirubin	0
Plasma Chemistry Total Protein	0
Plasma Chemistry Triglycerides	0
Plasma Chemistry Urea	0
Plasma Chemistry Uric Acid	0
Salmonella Challenge Cecal Contents	0
Salmonella Challenge Liver	0
Salmonella Challenge Spleen	0
Stress Induced Hyperthermia Basal Body Temperature	0
Stress Induced Hyperthermia Change in Temperature	0
Stress Induced Hyperthermia Post Stress Temperature	0
Weight Curve	0
X-ray Imaging Brachydactylism	0
X-ray Imaging Clavicle	0
X-ray Imaging Digit Integrity	0
X-ray Imaging Femur	0
X-ray Imaging Fibula	0
X-ray Imaging Fusion of Processes	0
X-ray Imaging Fusion of Ribs	0
X-ray Imaging Fusion of Vertebrae	0
X-ray Imaging Humerus	0
X-ray Imaging Joints	0
X-ray Imaging Kyphosis	0
X-ray Imaging Lordosis	0

X-ray Imaging Mandible	0
X-ray Imaging Maxilla	0
X-ray Imaging Number of Caudal Vertebrae	0
X-ray Imaging Number of Cervical Vertebrae	0
X-ray Imaging Number of Digits	0
X-ray Imaging Number of Lumbar Vertebrae	0
X-ray Imaging Number of Pelvic Vertebrae	0
X-ray Imaging Number of Ribs Left	0
X-ray Imaging Number of Ribs Right	0
X-ray Imaging Number of Thoracic Vertebrae	0
X-ray Imaging Pelvis	0
X-ray Imaging Polysyndactylism	0
X-ray Imaging Processes on Vertebrae	0
X-ray Imaging Radius	0
X-ray Imaging Scapula	0
X-ray Imaging Scoliosis	0
X-ray Imaging Shape of Ribcage	0
X-ray Imaging Shape of Ribs	0
X-ray Imaging Shape of Spine	0
X-ray Imaging Shape of Vertebrae	0
X-ray Imaging Skull Shape	0
X-ray Imaging Spinous Processes	0
X-ray Imaging Syndactylism	0
X-ray Imaging Teeth	0
X-ray Imaging Tibia	0
X-ray Imaging Transitional Vertebrae	0
X-ray Imaging Transverse Processes	0
X-ray Imaging Ulna	0
X-ray Imaging Zygomatic Bone	0
External collaboration - Eye histopathology	0
External collaboration - Brain histopathology	0
External collaboration - Skin histopathology	1
External collaboration - Tail epidermis wholemount	1
<b>Total Skin specific</b>	<b>42</b>

## Appendix II Other Published works

Liakath-Ali K, Vancollie VE, Heath E, Smedley DP, Estabel J, Sunter D, DiTommaso T, White JK, Ramirez-Solis R, Smyth I, Steel KP, Watt FM. **Novel skin phenotypes revealed by a genome-wide mouse reverse genetic screen.**

Nat Commun. 2014 Apr 11;5:3540. doi: 10.1038/ncomms4540.

Fu Y, Mukhamedova N, Ip S, D'Souza W, Henley KJ, DiTommaso T, Kesani R, Ditiatkovski M, Jones L, Lane RM, Jennings G, Smyth I, Kile BT, Sviridov D. **ABCA12 Regulates Macrophage Cholesterol Homeostasis and Atherosclerosis**

Cell Metab. 2013 Aug 6;18(2):225-38. doi: 10.1016/j.cmet.2013.07.003.

White JK, Gerdin AK, Karp N, Ryder E, Buljan M, Bussell JN, Salisbury J, Clare S, Ingham NJ, Podrini C, Houghton R, Estabel J, Bottomley JR, Melvin DG, Sunter D, Adams NC **The Sanger Institute Mouse Genetics Project**, Tannahill D, Logan DW, MacArthur DG, Flint J, Mahajan VB, Tsang SH, Smyth I, Watt FM, Skarnes WC, Dougan G, Adams DJ, Ramirez-Solis R, Bradley A, Steel KP. \*The Sanger Institute Mouse Genetics Project Phenotyping team: Baker L, Barnes C, Beveridge R, Cambridge E, Carragher D, Chana P, Clarke K, Hooks Y, Igosheva N, Ismail O, Jackson L, Kane L, Lacey R, Tino Lafont D, Lucas M, Maguire S, McGill K, McIntyre RE, Messenger S, Mottram L, Mulderrig L, Pearson S, Protheroe HJ, Roberson L, Salisbury G, Sanderson M, Sanger D, Shannon C, Thompson P, Tuck E, Vancollie VE. Genotyping and targeting team: Brackenbury L, Bushell W, Cook R, Dalvi P, Gleeson D, Habib B, Hardy M, Liakath KA, Miklejewska E, Price S, Sethi D, Trenchard E, von Schiller D, Vyas S, West AP, Woodward J, Wynn E. Mouse Informatics: Evans A, Gannon D, Griffiths M, Holroyd S, Iyer V, Kipp C, Lewis M, Li W, Oakley D, Richardson D, Smedley D. Tissue culture team: Agu C, Bryant J, Delaney L, Gueorguieva NI, Tharagonnet H, Townsend AJ. Mouse Production team: Biggs D, Brown E, Collinson A, Dumeau C, Grau E, Harrison S, Harrison J, Ingle CE, Kundi H, Madich A, Mayhew D, Metcalf T, Newman S, Pass J, Pearson L, Reynolds H, Sinclair C, Wardle-Jones H, Woods M. Mouse Facility team: Alexander L, Brown T, Flack F, Frost C, Griggs N, Hrnčiarova S, Kirton A, McDermott J, Rogerson C, White G, Zielesinsli P. Collaborators: DiTommaso T, Edwards A, Heath E, Mahajan MA, Yalcin B. **Genome-wide Generation and Systematic Phenotyping of Knockout Mice Reveals New Roles for Many Genes.** Cell. 2013 Jul 18;154(2):452-64. doi: 10.1016/j.cell.2013.06.022.

Wiradjaja F, DiTommaso T, Smyth I. **Basement membranes in development and disease.**

Birth Defects Res C Embryo Today. 2010 Mar;90(1):8-31.

Manuscripts to follow.

## ARTICLE

Received 9 Jan 2014 | Accepted 4 Mar 2014 | Published 11 Apr 2014

DOI: 10.1038/ncomms4540

OPEN

# Novel skin phenotypes revealed by a genome-wide mouse reverse genetic screen

Kifayathullah Liakath-Ali<sup>1,2,3</sup>, Valerie E. Vancollie<sup>4</sup>, Emma Heath<sup>1</sup>, Damian P. Smedley<sup>4</sup>, Jeanne Estabel<sup>4</sup>, David Sunter<sup>4</sup>, Tia DiTommaso<sup>5,†</sup>, Jacqueline K. White<sup>4</sup>, Ramiro Ramirez-Solis<sup>4</sup>, Ian Smyth<sup>5</sup>, Karen P. Steel<sup>4,6</sup> & Fiona M. Watt<sup>1</sup>

Permanent stop-and-shop large-scale mouse mutant resources provide an excellent platform to decipher tissue phenogenomics. Here we analyse skin from 538 knockout mouse mutants generated by the Sanger Institute Mouse Genetics Project. We optimize immunolabelling of tail epidermal wholemounts to allow systematic annotation of hair follicle, sebaceous gland and interfollicular epidermal abnormalities using ontology terms from the Mammalian Phenotype Ontology. Of the 50 mutants with an epidermal phenotype, 9 map to human genetic conditions with skin abnormalities. Some mutant genes are expressed in the skin, whereas others are not, indicating systemic effects. One phenotype is affected by diet and several are incompletely penetrant. In-depth analysis of three mutants, *Krt76*, *Myo5a* (a model of human Griscelli syndrome) and *Mysm1*, provides validation of the screen. Our study is the first large-scale genome-wide tissue phenotype screen from the International Knockout Mouse Consortium and provides an open access resource for the scientific community.

<sup>1</sup>Centre for Stem Cells and Regenerative Medicine, King's College London, Guy's Hospital, London SE1 9RT, UK. <sup>2</sup>Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK. <sup>3</sup>Wellcome Trust—Medical Research Council Stem Cell Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK. <sup>4</sup>Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridge CB10 1SA, UK. <sup>5</sup>Department of Anatomy and Developmental Biology, Monash University, Wellington Road, Clayton, Victoria 3800, Australia. <sup>6</sup>Wolfson Centre for Age-Related Diseases, King's College London, Guy's Campus, London SE1 1UL, UK. † Present address: Brigham Regenerative Medicine Center, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. Correspondence and requests for materials should be addressed to F.M.W. (email: fiona.watt@kcl.ac.uk).



The large amount of genetic information accumulated in the post-genomic era needs to be transformed into knowledge<sup>1</sup>. To study the functions of large numbers of genes *in vivo*<sup>2</sup> requires a shift from gene-specific to genome-wide approaches. Previous studies have catalogued libraries of mutant mice that lack specific classes of proteins<sup>3</sup> or exhibit behavioural phenotypes<sup>4</sup>. However, a large-scale screen (>100 mutants) for phenotypes in a specific tissue via targeted gene knockouts has not been reported previously.

For many years, researchers have been studying genes involved in epidermal homeostasis with an emphasis on stem cells and disease<sup>5,6</sup>. Since the first description of the hairless mouse in 1924, many spontaneous and random mutations with skin phenotypes have been studied by classical (forward) genetic approaches<sup>7</sup>. Owing to technical limitations, a large-scale high-throughput systematic functional screen for genes involved in skin homeostasis was not feasible until recently. Such a screen would facilitate the identification of novel genes that are involved in skin homeostasis, cancer, aging, infection, wound repair and sensation.

Genome-wide approaches to epidermal function include short interfering RNA-based genetic screens in cultured human epidermal cells<sup>8</sup> and RNA interference-mediated gene knockdown via *in utero* microinjection of lentiviral vectors<sup>9</sup>. Neither approach can identify effects of genes that are not expressed in the epidermis but influence epidermal homeostasis by local or systemic signalling. A further challenge is that it is extremely laborious to assess phenotypes in large areas of adult skin by relying on conventional histological sections. This can be overcome by preparing wholemounts of intact tail epidermis<sup>10</sup> in which the interfollicular epidermis (IFE), sebaceous glands (SGs) and hair follicles (HFs) are easily discerned. Virtually all epidermal stem cell populations express keratin (K)14 and one well-defined stem cell reservoir is in a region known as the HF bulge, which expresses K15 and is readily identified by its location<sup>11</sup>.

We have employed a high-throughput reverse genetics screen to identify skin phenotypes in mutant mice generated from the Sanger Institute Mouse Genetics Project (MGP)<sup>12,13</sup>. Identifying the phenotype correlations of genes involved in skin homeostasis paves the way to interrogate functional genetic interactions and find therapeutic targets for human skin diseases.

## Results

**Screen strategy.** The screening strategy was designed to analyse the tail epidermis of mutants available from the terminal phenotyping pipeline of the MGP. The screen was typically performed on viable adult homozygote mutants, or heterozygotes in the case of mutants with homozygote embryonic lethality. We analysed the skin without prior knowledge of the genotype. In principle, the screen could have been carried out with a single skin sample per genotype. However, we typically received two samples of each genotype, although we did not necessarily receive them at the same time. Pipeline design, generation of mice and primary phenotyping are described elsewhere<sup>12,13</sup>.

To facilitate analysis of large numbers of samples, the tail epidermal wholemount protocol was modified to reduce the time required for labelling and increase the number of samples that could be labelled simultaneously (Fig. 1a). Tissue was simultaneously labelled with fluorescently conjugated antibodies to K14 and K15 and subsequently counterstained with 4',6-diamidino-2-phenylindole (DAPI) nuclear counterstain. Epidermal sheets were labelled in individual wells of 24-well plates and up to 3 plates (72 specimens) were labelled simultaneously. In addition, pieces of intact tail skin were paraffin embedded for detailed histology

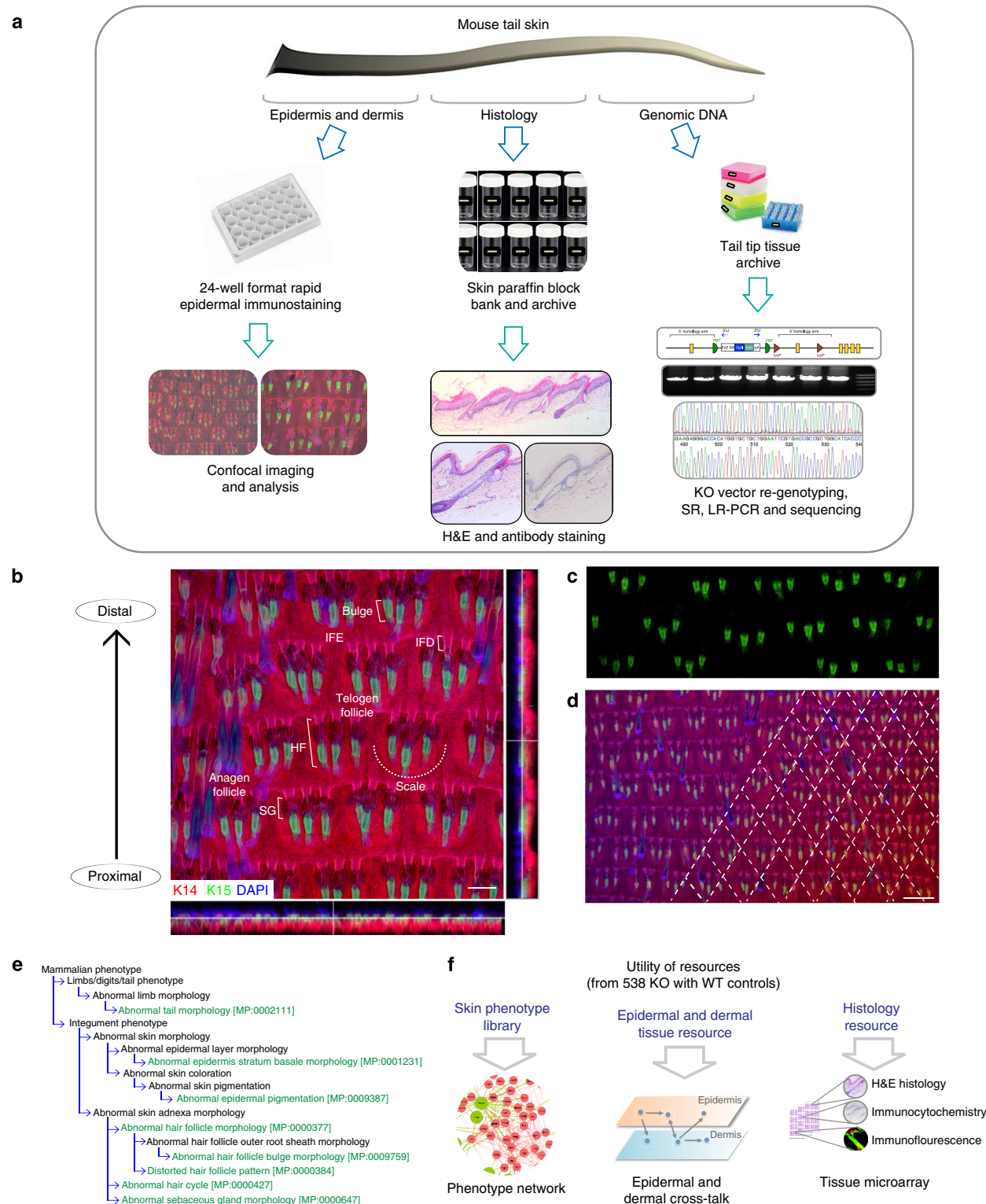
and DNA was extracted for verification of genotype if required (Fig. 1a).

Wholemounts were analysed by non-automated high-content confocal imaging (Fig. 1b–d; Supplementary Movie 1). In addition, any obvious macroscopic abnormalities, such as altered skin pigmentation or abnormal tail length, were recorded. For wild-type (WT) reference data, a panel of tail epidermal wholemount images from multiple WT genetic strains was generated for comparison with the respective mutant lines (Supplementary Fig. 1; see Methods). Many previous reports have highlighted the need to use specific vocabularies for consistent phenotypic annotation of mammals<sup>14,15</sup>. We therefore used existing mammalian phenotype (MP) terms to classify our phenotypes under the broad 'integument phenotype' and 'limbs/digits/tail phenotype' categories of the Mouse Genome Informatics database ([http://www.informatics.jax.org/searches/MP\\_form.shtml](http://www.informatics.jax.org/searches/MP_form.shtml); Fig. 1e).

Eight specific MP terms were used to annotate epidermal abnormalities (Fig. 1e). 'Abnormal tail morphology' was based on macroscopic evaluation (for example, thickened, kinked, scaly) but often manifested in abnormalities that were also detectable in epidermal wholemounts. In the tail, HFs are typically arranged in groups of three (triplets) and in staggered rows; when this pattern was abnormal, the phenotype was scored as 'distorted HF pattern'. Since the hair growth cycle is not synchronized in mice of the ages analysed (older than 13 weeks), the cycle was scored as 'abnormal hair cycle' if the majority of hair HFs were in anagen or telogen. When bulge size or K15 expression was altered, the skin was scored as 'abnormal HF bulge morphology' and when the shapes of individual follicles were distorted, we scored the phenotype as 'abnormal HF morphology'. We also scored abnormalities in the tail IFE, including scale versus interscale differentiation<sup>16</sup> ('abnormal epidermis stratum basale morphology'), and in SG morphology (abnormal size; irregular surface). In tail epidermis, melanocytes, identified by their dark colour, are normally confined to the HF or scale IFE and when this pattern was disrupted, the phenotype was scored as 'abnormal epidermal pigmentation'. The results were entered into a structured tick sheet before annotation and deposition at the Sanger Mouse Resources Portal (<http://www.sanger.ac.uk/mouseportal/>). This open access resource can be exploited to interrogate many aspects in skin biology, including information on the macroscopic appearance of living mice and histology of the skin from sites other than tail (Fig. 1f).

**Novel skin phenotypes.** Mutant mouse lines for a total of 557 alleles belonging to 538 unselected genes covering all mouse chromosomes except Y were analysed (Supplementary Data 1 and 2, and Supplementary Fig. 2a) along with matched WT control samples. The specific allele types are listed in Supplementary Data 2 and homozygotes are designated '–/–' in the text for simplicity. Gene ontology (GO) predicted that the majority of the encoded proteins are located in the cytoplasm and nucleus, with many involved in chromosome and chromatin organization. Large numbers of proteins were significantly enriched for GO terms corresponding to catalytic (28%; log<sub>10</sub> P value, –13.7356) and protein binding (30%; log<sub>10</sub> P value, –19.6965) functions (Supplementary Fig. 2b and Supplementary Data 2).

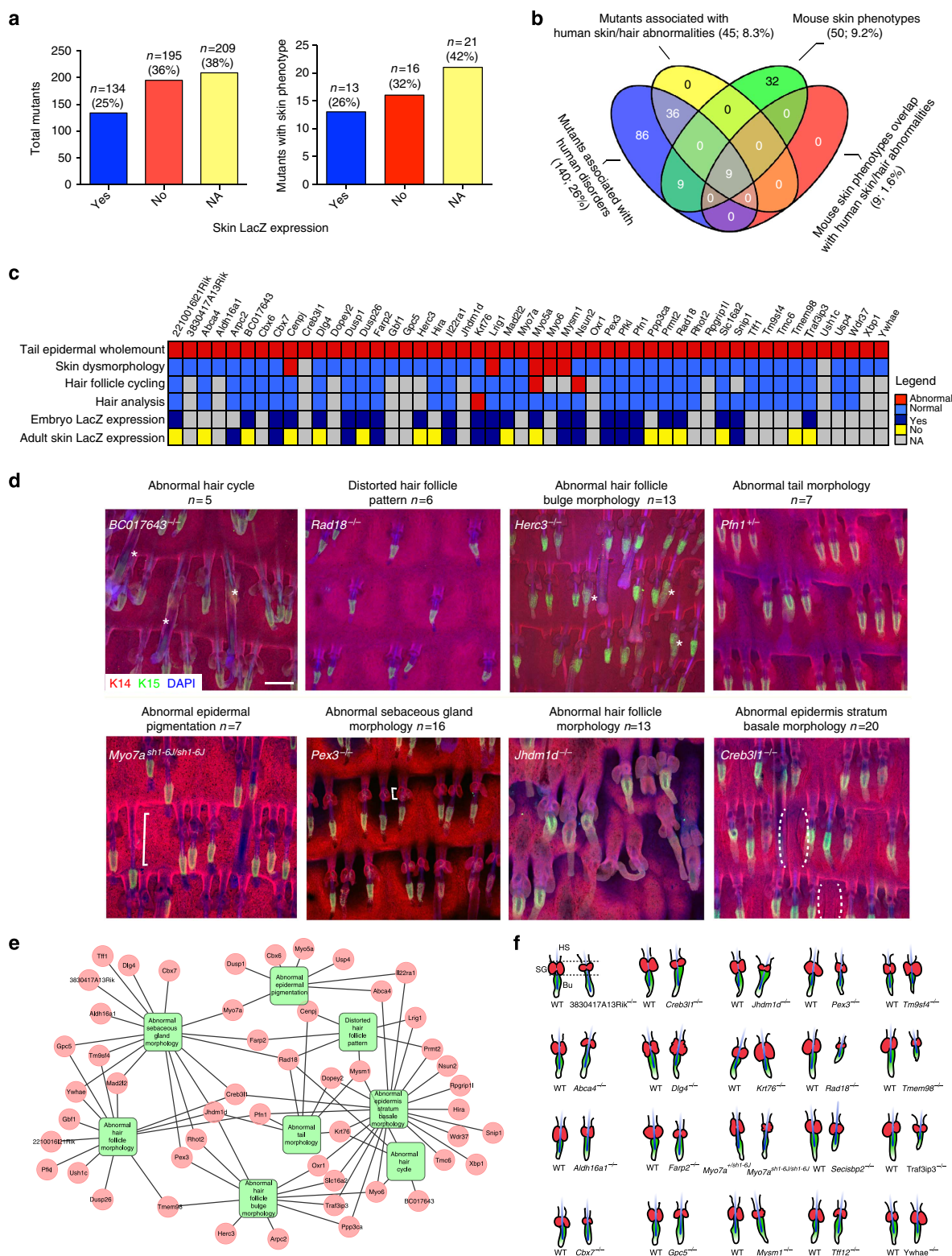
Most mice were engineered to express a *lacZ* reporter via the promoter of the deleted gene<sup>12</sup>. Skin *lacZ* expression was observed in 25% of mutant lines ( $n=134$ ) and was absent in 36% ( $n=195$ ). In 38% ( $n=209$ ) *lacZ* expression was either not evaluated or not determined, because the mice were deemed unsuitable for analysis or else a read-out of *lacZ* expression could not be obtained (Fig. 2a). We observed abnormalities in 9.2%



**Figure 1 | Skin screen strategy and exploitation of the resource.** (a) Screen strategy. On receipt, tail skin was divided into three parts: one was processed for epidermal wholemount immunostaining, one for paraffin conventional histology and one for genomic DNA extraction. (b) WT epidermal wholemount image labelled for K14 (red), K15 (green) and DAPI (blue) showing hair follicle (HF), interfollicular epidermis (IFE), infundibulum (IFD), sebaceous gland (SG), bulge, scale and anagen and telogen HFs. (c) Single channel (K15-488) image showing labelling of HF bulge. (d) Low magnification image of epidermis emphasizing patterned arrangement of HF triplets (one triplet per diamond marked with dashed lines). (e) MP ontology tree. MP terms in green were used to annotate skin phenotypes in this study. (f) Schema shows how resources generated from this study can be further utilized. Scale bars, 100  $\mu$ m (b,c); 200  $\mu$ m (d).

( $n=50$ ) of mutant lines. Skin *lacZ* expression was observed in 26% (13/50) and absent in 32% (16/50) of mutants with epidermal phenotypes (Fig. 2a, Supplementary Data 3 and 4, and Supplementary Fig. 3).

Fourteen percent of hits ( $n=7$ ) from the tail epidermal wholemount screen were also found independently in the MGP screen of live mice for macroscopically observable skin-related parameters (skin dysmorphology, HF cycling and





hair analysis; Fig. 2c; data available from <http://www.sanger.ac.uk/mouseportal>). Two of the hits in our screen, leucine-rich transmembrane protein *Lrig1* (ref. 17) and the RNA methyltransferase *Nsun2* (ref. 18), have previously been identified as playing a role in epidermal homeostasis and thus served as positive controls (Fig. 2c and Supplementary Fig. 4). Further information about skin phenotypes is available via the portal under the heading 'skin histopathology' or by searching for individual genes under the category 'integument'.

Representative images of each category of tail epidermal abnormality are shown in Fig. 2d. Several mutants exhibited abnormalities in more than one MP term. For example, the *Jhdm1d*<sup>-/-</sup> (also known as *Kdm7a*) phenotype was recorded as 'abnormal HF morphology' and also 'abnormal tail morphology', 'abnormal SG morphology' and 'abnormal HF bulge morphology' (Fig. 2d,e). Figure 2f shows outlines of individual HF with associated SG in the subset of mutants that showed HF/SG phenotypes, together with WT examples for comparison.

'Abnormal epidermal stratum basale morphology' and 'abnormal SG morphology' (40%; *n* = 20 and 34%; *n* = 17, respectively) represented the maximum number of 'hits'. The IFE abnormalities included loss of distinct scale/interscale boundaries and crowding of basal layer cells, indicative of increased proliferation. In some cases, such as *Creb3l1* (Fig. 2d), both abnormalities were present; since the scales form through more rapid basal cell proliferation than in interscales<sup>16</sup>, lack of scales could potentially be a consequence of IFE hyperproliferation. SGs are sac-like structures emerging from the junctional zone of HF with two to three lobes<sup>19</sup>. Most of the abnormalities related to SG size (for example, *Pex3*<sup>-/-</sup> had small SG and *Krt76*<sup>-/-</sup> had large SG), but *Dlg4*<sup>-/-</sup> and *Myo7a*<sup>-/-</sup> mutants displayed SG with irregular surfaces (Fig. 2f).

One of the most striking features of the HF phenotypes was that they were not obligatorily linked to one another. Thus, a mutant with abnormal bulge morphology (for example, *Herc3*<sup>-/-</sup>) could have a normal hair cycle. Conversely, the hair cycle could be perturbed in an otherwise normal epidermis (for example, *BC017643*<sup>-/-</sup>; Fig. 2d,e). In contrast, abnormal tail morphology was always linked to other epidermal phenotypes (Fig. 2e).

The mutant mice were generated as part of two different analysis pipelines, one of which included a high-fat diet while the other involved a standard chow diet. In most cases, it was not possible to look for an influence of diet on mutant epidermis, although we saw no effect of diet on WT control tail epidermis (see Methods). However, *Tm9sf4*<sup>-/-</sup> mutants went through both pipelines and an effect of diet was observed. Mutant mice fed on chow had smaller SGs than WT, whereas mutants fed on the high-fat diet did not (Supplementary Fig. 5).

We also observed incomplete penetrance of phenotype in some mutants, in particular *Rad18* and *Jhdm1d* (Fig. 2d and Supplementary Fig. 6), both of which exhibited strong phenotypes in four of eight categories examined but only for one mouse per genotype, the others being phenotypically normal

(Supplementary Data 4). This is of interest because some mutations associated with skin abnormalities have incomplete penetrance among human populations<sup>20</sup>.

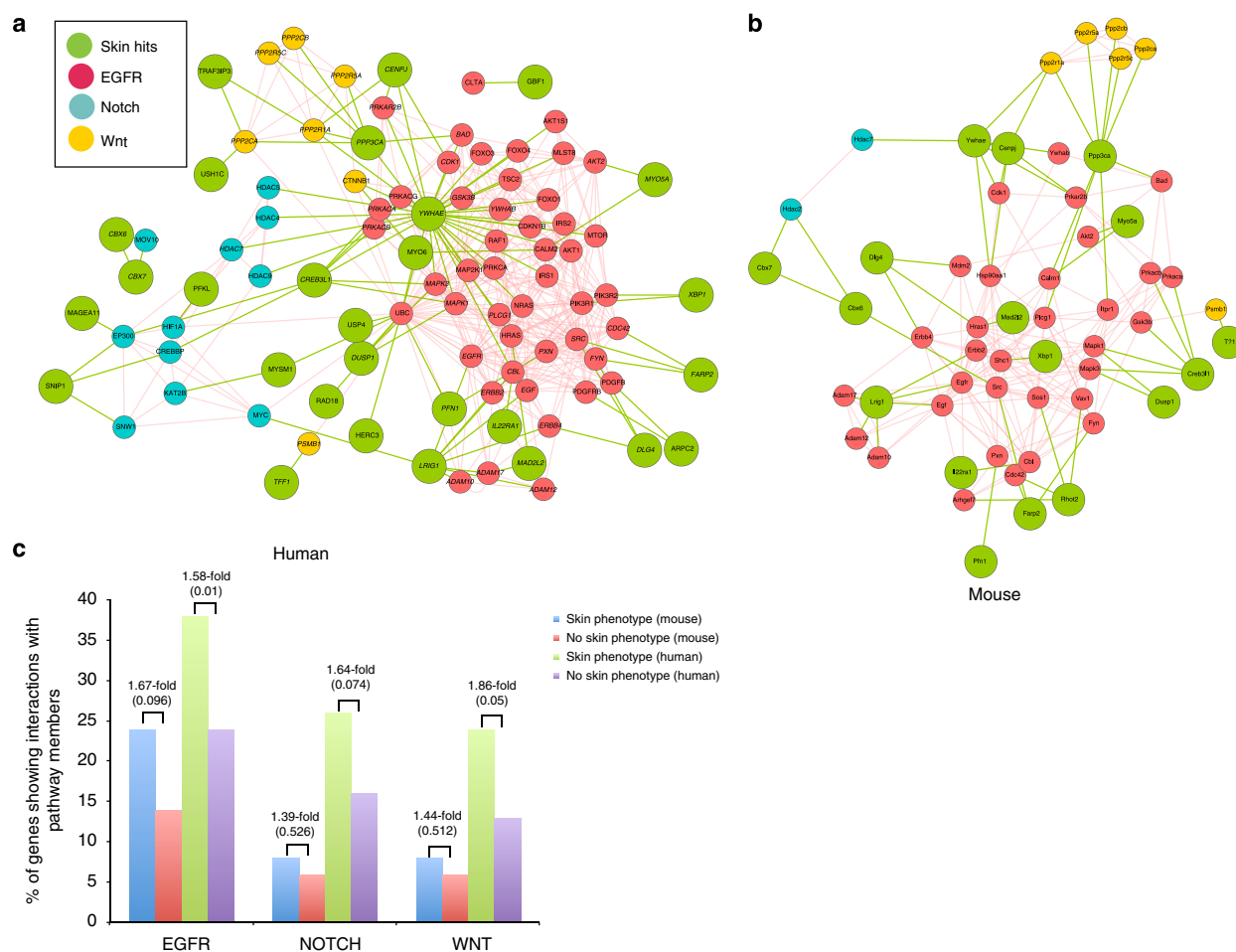
Most mice analysed were female, since they are less likely to fight and sustain skin wounds that could confound the analysis. To control for potential sexual dimorphism in skin structures, we analysed genes involved in oestrogen signalling. We extracted a list of 368 genes from the oestrogen responsive genes database<sup>21</sup> (<http://datam.i2r.a-star.edu.sg/ergdbV2/>). We found that 12 of the 538 genes in our screen are oestrogen responsive; however, none scored positive for any epidermal abnormalities (Supplementary Data 1).

#### Association with human disorders and signalling pathways.

We used the PhenExplorer application<sup>22</sup> (<http://www.compbio.charite.de/phenexplorer>) to examine whether any of the mutants in our screen were already linked to human diseases, in particular those with a skin phenotype. The key advantage of the PhenExplorer application is that users can search individual phenotypic features associated with any human disease by visualizing specific phenotypic terms with unique IDs. As an alternative strategy, to ensure we had not missed any known rare diseases or skin phenotype associations, we also interrogated the Human Phenotype Ontologies database, diseases of Online Mendelian Inheritance in Man and Orphanet (rare diseases and orphan drugs) databases. Twenty-six percent (*n* = 140) of the mouse-mutant lines we analysed (*n* = 538) were linked with human disorders, including 18 out of 50 mutants with an epidermal phenotype. Of those 18 mutants, 9 were associated with a human skin phenotype. These mutants provide useful experimental models to study the pathophysiology of human skin diseases or test potential therapeutics (Fig. 2b and Supplementary Data 3).

Egfr, Wnt and Notch signalling pathways regulate many aspects of epidermal homeostasis in humans and mice<sup>23</sup>. We determined whether the 50 genes with an epidermal phenotype were enriched for interactions with these pathways relative to the background population of 488 genes that did not (Fig. 3a–c; Supplementary Fig. 7). A binomial probability calculation using the background rate of interactions of the 487 genes was performed to test whether the observed enrichments for the 50 skin genes were statistically significant. There was statistically significant enrichment for EGFR (epidermal growth factor receptor; *P*-value < 0.05, by exact binomial probability) interactions, and borderline enrichment (*P* values from 0.05 to 0.1, by exact binomial probability) for the NOTCH and WNT pathways in humans, where more interaction data are available compared with mice. The main potential links to the WNT pathway were via a family of protein phosphatases that are of unknown function in skin, while the potential NOTCH interactors included histone deacetylases. In the case of EGFR interactions, we found associations with many genes known to affect epidermal differentiation and homeostasis, including cell

**Figure 2 | Phenotype summary.** (a) *lacZ* reporter expression in skin. NA, not applicable because not determined. Left hand panel shows data for all mutants. Right hand panel shows data for those mutants with a skin phenotype. (b) Venn diagram showing overlap between the 50 genes identified with skin phenotypes in the screen, genes in the screen that are associated with any human disease or human skin/hair abnormalities, and genes associated with both mouse and human skin/hair abnormalities. (c) Heatmap showing all 50 phenotypic hits in the tail epidermal wholemount screen (top row), *LacZ* expression data from the MGP (bottom two rows) and the results of three macroscopic skin-related tests available through the MGP (middle rows). (d) Representative tail epidermal wholemount images of phenotypic hits from eight categories annotated using MP terms. Images are 3D maximum projections, except for *Pex3*<sup>-/-</sup>, which shows a single stack. (e) Association of each hit with the eight phenotypic categories annotated using MP terms. (f) Outlines of individual mutant HF and SG (red) compared with WT. WT and mutants are matched for genetic background. Scale bars, 100  $\mu$ m.



**Figure 3 | Potential interactions of mutant genes with the Egfr, Notch and Wnt pathways.** (a) Human, (b) mouse. Genes common in human and mouse are italicized. (c) Number of mutants interacting with human and mouse Egfr, Notch and Wnt pathways, showing statistical analysis of enrichment by an exact binomial probability test.

cycle regulators, Foxo transcription factors and the MAPK signalling pathway<sup>24</sup>.

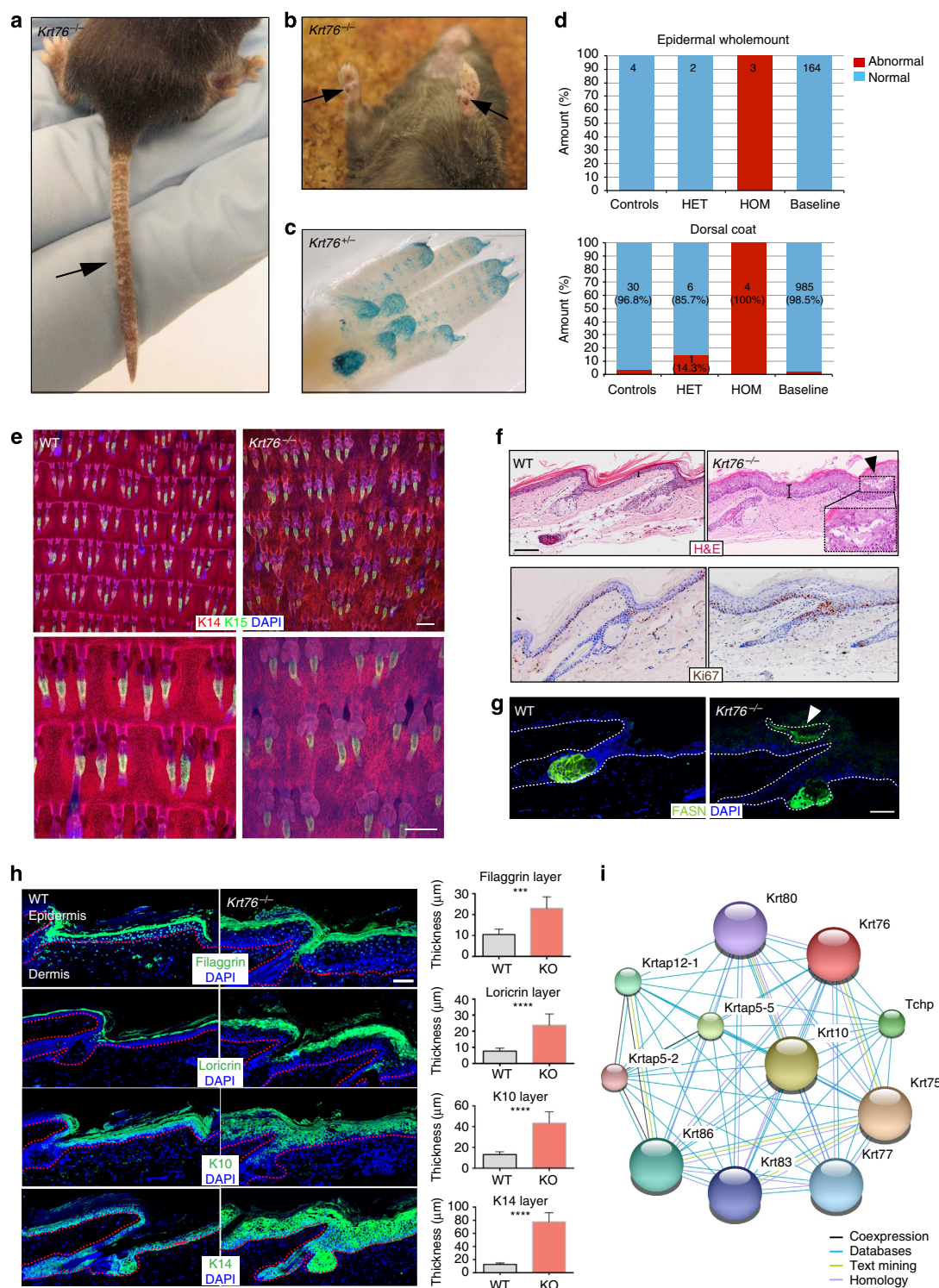
We also performed an unbiased search for cellular processes and phenotypes affected by the mutant genes. Enrichment analysis was performed using the ToppFun tool in the ToppGene suite (<http://www.toppgene.cchmc.org>) on all 50 skin phenotype hits annotated according to the specific MP terms (Supplementary Table 1). The default Bonferroni method for multiple testing was used to output results with a false discovery rate below 0.05. The background data sets used to calculate enrichment were all those supplied by ToppGene: for example, protein functions represented by all human gene GO annotations, disease–gene associations from Online Mendelian Inheritance in Man and phenotype annotations from the Human Phenotype Ontologies and Mouse Genome Informatics databases. The genes linked to abnormal skin pigmentation had known functions in pigmentation, such as granule transport ( $6.104E-3$ ), and were associated with abnormal eye pigmentation ( $2.460E-2$ ). The other seven MP terms yielded only vestibular hair cell degeneration ( $4.214E-2$ ) as a previously identified phenotype.

As further validation of the screen, we explored the phenotype of three hits in greater depth. *Krt76* was selected because it is a keratin and so any effects should be intrinsic to epidermal cells<sup>25</sup>. The type V myosin *Myo5a* was examined because it is mutated in

a human condition called Griscelli syndrome that has skin phenotypes<sup>26</sup>. The histone H2A deubiquitinase *Mysm1* was selected because although it had a striking skin phenotype its reported functions are in the haemopoietic system<sup>27,28</sup>.

***Krt76* mutants have aberrant epidermis and SG.** Keratin intermediate filament proteins protect epithelial cells from mechanical and non-mechanical stresses<sup>29</sup>. *Krt76* is a type II intermediate filament (<http://www.interfil.org>) that is downregulated in human oral carcinoma<sup>25</sup>. *Krt76*<sup>−/−</sup> mice exhibit preneoplastic changes in the gingivobuccal epithelium<sup>25</sup>, but *Krt76* function in the epidermis is unknown.

*Krt76*<sup>−/−</sup> mice had flaky tails and both the tail and footpads were darkly pigmented (Fig. 4a,b). *lacZ* reporter expression was observed in footpads (Fig. 4c). *Krt76* heterozygotes had no obvious skin abnormalities, although an abnormal dorsal hair coat was observed in a small percentage (Fig. 4d). Tail epidermal wholemounts were abnormal in four phenotypic categories, including abnormal stratum basale and SG morphology (Fig. 4e). Analysis of haematoxylin and eosin-stained tail skin revealed epidermal thickening and basal layer hyperproliferation (Fig. 4f). The scales were no longer clearly defined and instead the granular layer (which is normally lacking in the scale) was continuous.



**Figure 4 | *Krt76*<sup>-/-</sup> phenotype.** (a) Flaky tail skin (arrow). (b) Darkly pigmented forepaws (arrows). (c) *lacZ* reporter expression in footpad (downloaded from Sanger Mouse Portal). (d) Stacked bars showing proportion of *Krt76*<sup>-/-</sup> mice with abnormalities observed in epidermal wholemounts or dorsal coat colour. HET, heterozygous; HOM, homozygous; control, WT mice run each week through the Sanger MGP to ensure that all testing conditions are normal; baseline, all controls were run for a particular background, pipeline and diet combination. (e) Confocal epidermal wholemount images of WT and *Krt76*<sup>-/-</sup> mice. (f) Histological sections stained with haematoxylin and eosin (upper panel) and Ki67 (lower panel). Ectopic sebocytes in the IFE (arrowhead) are shown at higher magnification in the insert. Brackets indicate IFE thickness. (g) Immunostaining for the SG differentiation marker, FASN. Ectopic sebocytes are marked by arrowheads. (h) Immunostaining for the markers indicated. Thickness of each layer was quantified using Fiji (Image J) software. Stack bars show results of unpaired *t*-test. Error bars denote mean (*n* = 3) with s.d., \*\*\**P* < 0.0002, \*\*\*\**P* < 0.0001. (i) STRING network image shows possible interaction partners of *Krt76*. Scale bars, 50 μm (upper panel) and 100 μm (lower panel) (e); 200 μm (f,g); 100 μm (h).



The tail IFE also contained clusters of large cells with pale pink cytoplasm (Fig. 4f), resembling the ectopic sebocytes that form upon abnormal epidermal *Myc* activation<sup>10,30</sup>. These cells were confirmed as sebocytes on the basis of *FASN* and *SCD1* expression (Fig. 4g). Consistent with IFE hyperproliferation, there was suprabasal expression of *Krt14* and an increase in the number of suprabasal layers expressing the differentiation markers *Loricrin*, *Filaggrin* and *Krt10* (Fig. 4h). As predicted from the changes to the granular layer, *Filaggrin* labelling was continuous rather than confined to the interscale<sup>16</sup>.

We performed known and predicted protein–protein interaction analysis using the STRING database (<http://www.string-db.org/>)<sup>31</sup> to look for functional protein associations with *Krt76*. There are no experimentally recorded interactions of *Krt76*, but predicted interactions retrieved other keratins or intermediate filament-associated proteins such as *Tchp*, *Krt77*, *Krt86* and *Krt10*. Numerous genetic skin disorders are associated with keratin mutations, including *Krt10*, which is mutated in recessive epidermolytic ichthyosis with hyperkeratosis<sup>32</sup> (Fig. 4i).

***Myo5a* null mice phenocopy Griscelli syndrome.** Mutations affecting the actin-based motor protein MYO5A underlie a rare human genetic condition known as Griscelli syndrome<sup>26</sup>, which is characterized by hypopigmentation of the skin and hair. Several spontaneous mutations in the dilute locus were previously reported in mice with a lighter coat colour in a non-agouti background, leading researchers to map the *Myo5a* gene to this locus<sup>33</sup>. In *Myo5a*<sup>−/−</sup> mice, the coat colour abnormality was obvious (Fig. 5a,b). Tail epidermal wholemounts (both immunostained and unstained) revealed abnormal morphology of melanocytes and increased intracellular accumulation of melanin pigment (Fig. 5c,d), consistent with previous reports that the coat-colouration phenotype in dilute mice results from abnormal melanosome transport from melanocytes into epidermal keratinocytes<sup>34</sup>. Abnormal pigmentation was confirmed by conventional histology (Fig. 5d). STRING network analysis revealed experimentally validated interaction partners of *Myo5a*, of which *Rab27a* and *Mrp1* are known to have a role in melanosome transport in melanocytes<sup>35</sup> (Fig. 5e).

We also identified abnormal epidermal pigmentation in mice with a spontaneous mutation in *Myo7a*, which had unusually dark and large melanocytes in the IFE (Fig. 2d and <http://www.sanger.ac.uk/mouseportal/>). Surprisingly, these mutants did not show any coat-colour abnormalities, which may reflect different genetic backgrounds of the *Myo7a* and *Myo5a* mutants (Supplementary Data 2). *Myo7a* also belongs to the actin-based family of motor proteins and is involved in melanosome transportation by binding to *Rab27a* in retinal pigment epithelia, the same mechanism by which *Myo5a* contributes to epidermal pigmentation<sup>34,36,37</sup>.

**Deletion of *Mysm1* leads to abnormal HF patterning.** *Mysm1* is a histone H2A deubiquitinase<sup>38</sup>, previously identified as a positive regulator of the androgen receptor<sup>39</sup>. *Mysm1*<sup>−/−</sup> mutants all had white kinky tails (Fig. 6a), and displayed abnormalities in the IFE and HF (Fig. 6b). Abnormal HF patterning was observed, with HF no longer arranged as staggered rows of triplets. In addition, the HF cycle was disturbed, with most HFs being in telogen in the specimens examined (Fig. 6c). Within the IFE, the scale pattern was completely disrupted<sup>16</sup>, although, in contrast to the *Krt76*<sup>−/−</sup> mutant, regions of scale differentiation lacking the granular layer could still be observed (Fig. 6d,e). STRING network analysis identified the transcription factor *Tcf3* as an experimentally validated interactor, which is of interest, given the known role of *Tcf3* in epidermal homeostasis, where it functions

in both a Wnt-dependent and Wnt-independent manner<sup>40</sup> (Fig. 6f). In addition, Wnt signalling is known to regulate scale pattern<sup>16</sup>.

*Mysm1* mutants also had a white ‘belly spot’, which is a hallmark of melanoblast migration defects in BL6 background mice (<http://www.sanger.ac.uk/mouseportal/phenotyping/MAHN/mp-report/integument>). Many similar belly spot phenotypes have been reported, with different genes involved<sup>41,42</sup>. *Mysm1* is downregulated in mice with mutations influencing eye pigmentation<sup>43</sup>. Although the ubiquitin proteasome system is known to play a role in regulating skin pigmentation<sup>44</sup>, we did not observe a strong pigmentation defect in tail epidermal wholemounts.

## Discussion

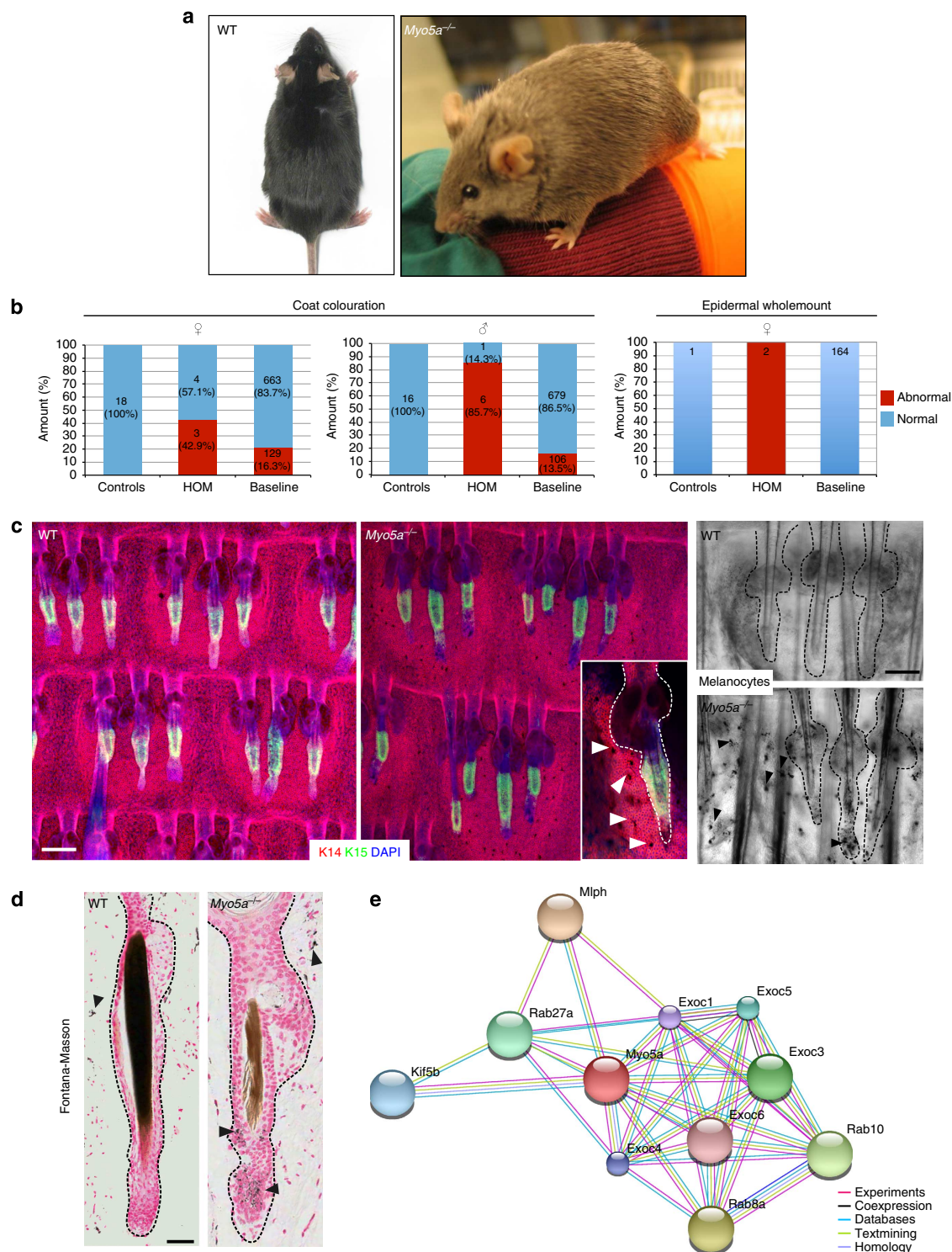
Now that large numbers of mouse mutants are available through the International Knockout Mouse Consortium<sup>45</sup>, there is a need to screen phenotypes via correspondingly high-throughput approaches. By making phenotype data available immediately via an open access resource, results can be rapidly and widely disseminated, and potential interactions between genes or between phenotypes can be revealed. We have used this approach to carry out the first large-scale tissue-specific screen of mutants from the MGP.

Seventy-eight percent of the 250 adult mouse mutants analysed in the initial MGP screen had a phenotype affecting at least one tissue or physiological process<sup>13</sup>. From our screen of 538 mutants, we identified 50 genes, mutations in which led to some form of tail epidermal phenotype. As part of the screen, we collected tail skin for paraffin embedding. This has created a valuable archive from which we can create tissue microarrays, to retrospectively explore other aspects of the skin, such as innervation, connective tissue and vasculature<sup>46</sup>. We examined the tail skin phenotypes of three mutants in depth, with larger numbers of mice (Supplementary Data 4), and in each case validated the conclusions from the initial screen.

For most mutants a functional role in skin was not anticipated. Of those genes with an epidermal phenotype and linked *lacZ* data, approximately equal numbers were expressed or not expressed in the skin, the latter indicating a non-cell autonomous effect. The majority of mutant genes that were not expressed in skin showed phenotypes in other body systems, particularly affecting the immune system and homeostasis/metabolism (<http://www.sanger.ac.uk/mouseportal/>). For example, *Mysm1* has so far been primarily implicated in regulating immune cell function, which can indirectly affect the HFs and IFE<sup>5</sup>. There is a precedent for this in other tissues: *Trim45* is not detected in the cartilage or bone, yet on deletion results in brittle bones and a high bone mineral content<sup>47</sup>.

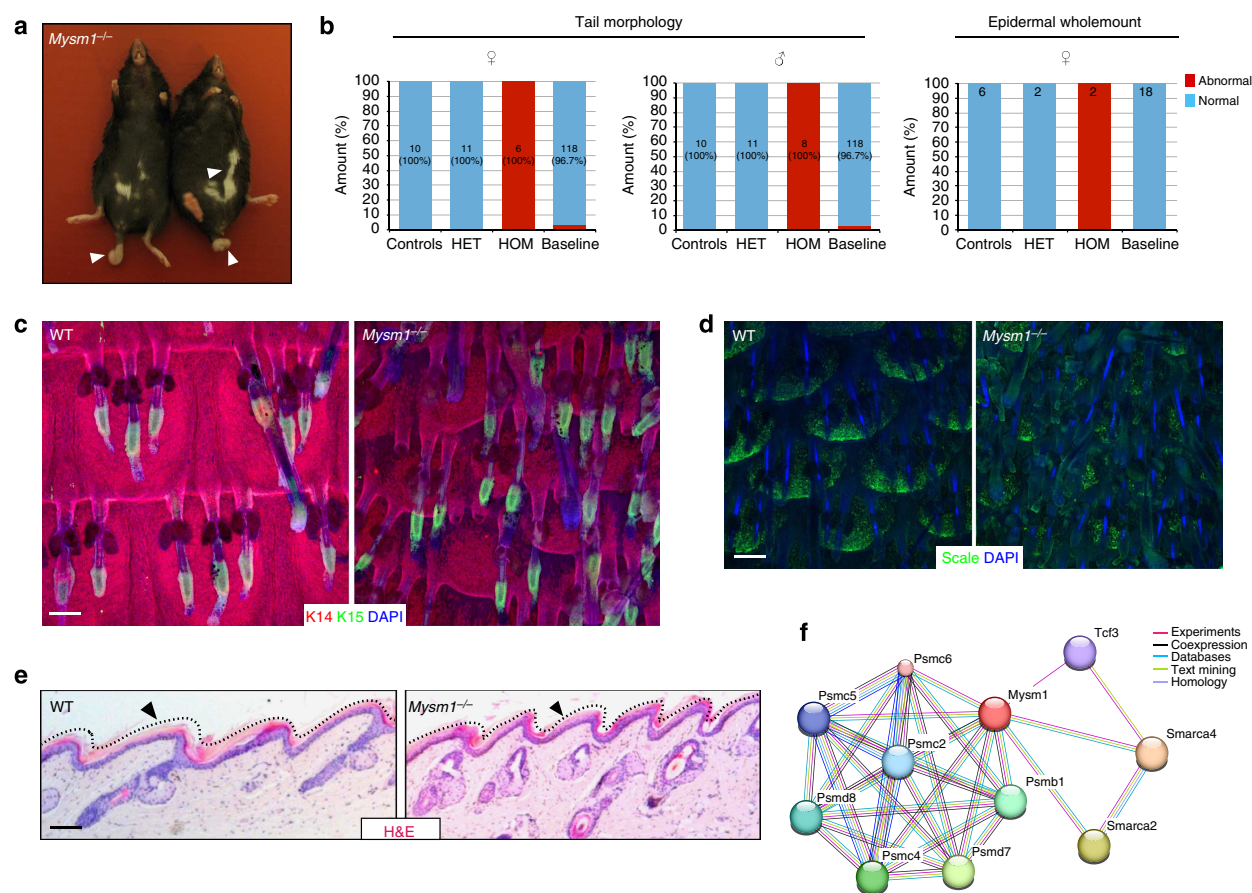
Some mutants had multiple skin phenotypes, while others were distinct. Many mice with defects in the hair growth cycle or HF morphology have been described previously<sup>5</sup>. However, we found a surprisingly high incidence of SG abnormalities. For example, *Pex3* plays a role in peroxisome biogenesis<sup>48</sup> and *Pex3*<sup>−/−</sup> mice had abnormally small SGs. In humans, peroxisome deficiency results in inflammatory hair disorders, such as primary cicatricial or scarring alopecia and lichen planopilaris, with SG atrophy as a key pathological feature. It is reported that *Pex3* and its regulator peroxisome proliferator-activated receptor- $\gamma$  are significantly downregulated in lichen planopilaris<sup>49</sup>. Another gene with a role in lipid metabolism was *Aldh16a1*, which belongs to the aldehyde dehydrogenase (ALDH) protein family. These mutants exhibited smaller SGs than WT. Previous reports suggest a role for other ALDHs in SG<sup>50,51</sup>, indicating that *Aldh16a1* may have a similar function.

Although we did not initially design the tail epidermal screen to identify pigmentation defects, these were readily observed,



**Figure 5 | *Myo5a*<sup>-/-</sup> phenotype.** (a) Coat colour of WT and *Myo5a*<sup>-/-</sup> mice on a BL6 genetic background. (b) Stack bar graphs shows coat colour and epidermal abnormalities. See the legend of Fig. 4d. (c) Epidermal wholemounts. Note abnormal melanocyte morphology in IFE interscale (arrowheads in magnified insert). Bright-field image showing melanosome clumping in *Myo5a*<sup>-/-</sup> epidermal melanocytes. (d) Melanocyte-specific Fontana-Masson staining of skin sections showing aggregation of melanosomes in melanocytes in *Myo5a*<sup>-/-</sup> HF. Note also differences in colour of WT and mutant hair shafts. Individual HF with SGs are demarcated with dashed lines. (e) STRING network image shows known (experimental) and predicted interaction partners of *Myo5a*. Scale bars, 100  $\mu$ m (c); 200  $\mu$ m (d).





**Figure 6 | *Mysm1*<sup>-/-</sup> phenotype.** (a) Adult *Mysm1*<sup>-/-</sup> homozygous mutant mice with short kinked white tails and white belly patches (arrows). (b) Stacked bars show incidence of abnormal tail morphology and epidermal wholemounts. See also the legend of Fig. 4d. (c,d) Epidermal wholemount images show disrupted HF pattern in *Mysm1*<sup>-/-</sup> mice when compared with WT. In d wholemounts were stained with the scale-specific marker ab1653 (ref. 16). (e) Haematoxylin and eosin (H&E)-stained sections showing parakeratotic interscale in WT and mutant tail epidermis. (f) STRING network image shows known (experimental) and predicted interaction partners of *Mysm1*. Scale bars, 100  $\mu$ m (c–e).

both macroscopically and in wholemounts. While *Myo5a* phenocopied human Griscelli syndrome, a role for *Myo7a* in epidermal pigmentation was not predicted. These two mouse models could potentially be used to investigate melanosome transfer into epidermal cells and the consequences of transfer deficiency for ultraviolet carcinogenesis. A further category of genes that are particularly interesting are those characterized by IFE hyperproliferation, since these were often associated with loss of scale patterning, supporting the concept that scale and interscale turn over at different rates and providing a potential tool for identifying the distinct pools of scale and interscale stem cells<sup>16</sup>. Six of nine mutants (*Abca4*, *Hira*, *Nsun2*, *Rpgrip11*, *Slc16a* and *Tmc6*) in the ‘abnormale stratum basale’ category have skin phenotypes in humans (Supplementary Data 3).

Since the mutants were unselected and covered a wide variety of biological functions, we asked whether mutants with a skin phenotype mapped to the Wnt, Notch and Egfr pathways, all of which play key roles in epidermal differentiation and homeostasis<sup>52,53</sup>. We identified four mutants, *Cenpj*, *Farp*, *Lrig1* and *Rad18*, with distorted HF patterning that are direct targets of the Egfr pathway. Two other mutants, *Mysm1* and *Prmt2*, mapped to the Wnt and NF- $\kappa$ B pathways, respectively.

Our study represents the first large-scale tissue-specific screen of mutant mice. The largest number of mutants screened previously was 100, in the case of bone abnormalities<sup>47</sup>.

We believe that the data and resources generated will accelerate research in many labs with an interest in skin function and pathology. In the past, generating knockout mice was time consuming and expensive, and frequently led to duplication of effort in different labs. Our advances include high-throughput analysis of the epidermis by combining an optimized wholemount labelling protocol with phenotyping based on standard ontologies. The bottleneck in our screen has been in evaluating individual wholemounts by confocal microscopy and collecting images of mutant phenotypes by three-dimensional (3D) Z-stacks. This approach is slow, requires a considerable amount of data storage and to avoid subjective calls requires a second person to examine a subset of wholemounts.

Future epidermal wholemount screens could be improved by replacing visual inspection with automated, high-throughput, high content imaging, which would also allow us to examine more mice per genotype. When we designed the screen, in 2006, many epidermal stem cell populations had not been identified<sup>11</sup>. Nevertheless, the combination of K14 and K15 labelling has provided a satisfactory and robust means of identifying the major regions of the epidermis. Given the large number of mutants with abnormal SG, the next iteration of the screen could include a lipophilic dye, such as LipidTox, to enhance visualization of the SG without appreciably lengthening the staining procedure. Finally, while the MGP allows integration of multiple data sets

for a given mouse mutant, the skin field would undoubtedly benefit from a platform that would facilitate integration of the skin screen with screens of *in vitro* and *in utero* epidermal knockdowns<sup>8,9</sup>, together with gene expression and genome-wide chromatin immunoprecipitation data sets. This holds great promise for skin research.

## Methods

**Mouse mutants.** Mutant mice were generated by high-throughput gene targeting in embryonic stem cells<sup>12</sup> by the Wellcome Trust Sanger Institute's MGP as part of the International Knockout Mouse Consortium. Mice were maintained at the Sanger Institute, the Wellcome Trust-MRC Cambridge Stem Cell Institute or the Cancer Research UK Cambridge Research Institute with local ethical approval from each institution and in compliance with the UK Animals (Scientific Procedures) Act, 1986.

**Tail skin samples.** The MGP is a high-throughput, large-scale primary screen designed to detect changes involving large biological effect sizes. For most tests in the pipeline (including skin dysmorphology in 10-week-old mice and HF cycling and hair analysis in 43-day-old mice, presented in Fig. 2c), 7 male and 7 female mice are assessed per mutant line. However, for the epidermal wholemount screen ( $n = 538$  unselected knockout strains assessed for 557 alleles), two female mutants (range 1–9 mutant mice) were typically assessed per line. Female mice were chosen because they are less aggressive than males, thereby reducing the incidence of skin wounding and inflammation<sup>54</sup>. There is also sexual dimorphism in SG size and other skin-associated structures in mice<sup>55</sup>. Where possible, when abnormalities were detected in the primary screen, additional mice were screened to confirm the finding.

To establish a reference range for each strain assessed, epidermal wholemounts from different strains of WT mice fed on different diets (mouse breeder diet—chow; 5021, Labdiet) or high-fat diet—(21.4% crude fat content, Western RD, 829100, Special Diets Services) were analysed ( $n = 168$ , B6Brd;B6Dnk;B6N-Tyrc-Brd high fat;  $n = 155$ , C57BL/6N high fat;  $n = 56$ , C57BL/6N chow;  $n = 18$ , B6Dnk;B6N chow;  $n = 15$ , 129Sv;  $n = 12$ , B6Brd;B6Dnk;B6N-Tyrc-Brd chow;  $n = 11$ , B6Dnk;B6N high fat;  $n = 10$ , C3HeB/F<sub>1</sub>;  $n = 8$ , CBACa;129P2;  $n = 8$ , B6Jlco;  $n = 4$ , B6Jlco;129P2;  $n = 4$ , B6Jlco;B6Brd;129P2-Tyrc-Brd;  $n = 4$ , B6Brd;B6Dnk;B6N;129S5-Tyrc-Brd;  $n = 2$ , 129S5;B6N;  $n = 2$ , AK BKS). Wholemounts were analysed for each of the eight MP ontologies to establish the range of normal phenotypes.

**Epidermal wholemount staining.** Intact tails were collected from  $\geq 14$ -week-old mutant and WT mice of the same genetic background. Tails were labelled with unique IDs and either processed immediately or stored at 4 °C for less than 3 days. The skin was manually detached from the caudal vertebrae, and then divided into three parts for epidermal wholemount labelling, histology and re-genotyping.

The epidermal wholemount labelling procedure<sup>10</sup> was modified to reduce the time required and allow large numbers of samples to be processed simultaneously. In brief, a scalpel was used to slit the tail lengthways. Pieces (0.5  $\times$  0.5 cm<sup>2</sup>) of skin were incubated in 5 mM EDTA in PBS at 37 °C for 4 h. Forceps were used to gently peel the epidermis from the dermis as an intact sheet in a proximal to distal direction, corresponding to the orientation of the hairs, and then the epidermis was fixed in 4% paraformaldehyde (PFA; Sigma) for 1 h at room temperature. Fixed epidermal sheets were washed in PBS and stored in PBS containing 0.2% sodium azide at 4 °C for up to 1 year before labelling. Remarkably, epidermal sheets were still suitable for immunostaining after 3 years in storage.

Mouse monoclonal K14 (LL002) and K15 (LHK-15) antibodies<sup>56</sup> were directly conjugated to Alexa Fluor 555 and 488 (A20174, A10235 Molecular Probes, Invitrogen), respectively, according to the manufacturer's instructions. In brief, 50  $\mu$ l 1 M sodium bicarbonate was added to 500  $\mu$ l of 2 mg ml<sup>-1</sup> purified protein. Alexa Fluor 555/488 reactive dye was added and incubated at room temperature for 1 h on a magnetic stirrer. Unconjugated dye was removed via column chromatography and conjugated antibody was eluted with 10  $\times$  elution buffer by gravity flow. Antibody dilutions (K14—1:200; K15—1:100) were optimized each time a new batch of directly conjugated antibodies was prepared. ab1653 (marker of scale epidermis) was used at 1:250 (Abcam, ab1653).

PFA-fixed epidermal sheets were placed in a 24-well tissue culture plate and blocked in PB (permeabilization & blocking) buffer (0.5% skimmed milk powder, 0.25% fish skin gelatin, 0.5% TritonX100, 1  $\times$  HEPES-buffered saline) for 30 min at 37 °C in a shaking incubator at 100 r.p.m. Directly conjugated K14 and K15 antibodies and DAPI diluted in PB were added to each well and incubated for 2 h in 37 °C in a shaking incubator at 100 r.p.m. Samples were washed three times with PBS/0.2% Tween solution at room temperature for 15 min per wash, and then gently rinsed with MilliQ water and mounted on glass slides in ProLong antifade medium (Invitrogen).

**Confocal microscopy.** Image acquisition of wholemounts was performed using a Nikon A1 or Leica SP5 TCS confocal microscope. Signals in the DAPI, Alexa 488 and Alexa555 channels were merged. For the Nikon microscope, lasers of 405, 488

and 561 nm wavelength were used with the following objectives: 4X dry (Plan Fluor 4X/0.13 numerical aperture) with Z-range:  $\sim 425$  nm; Z-step:  $\sim 55$ ; pixel dwell time (PDT): 5 mS and dry (Plan Fluor 10X/0.3 numerical aperture) with Z-range:  $\sim 125$  nm; Z-step: 9.4; PDT = 5 or 10.4 mS settings. 3D maximal projection images (1,024  $\times$  1,024 d.p.i.) and 3D reconstruction movies were generated using NIS Elements version 4.00.04 (Nikon Instruments Inc.).

When imaging with the Leica SP5 confocal microscope, lasers of 405, 488 and 561 nm wavelength were used and Z-stacks were acquired at 100 Hz with an optimal stack distance and 1,024  $\times$  1,024 d.p.i. resolution using 5X (0.07 dry), and (0.40 dry) HCX PL S-APO objectives. Leica Application Suite version 8.2.1 software (Leica Microsystems) was used for image acquisition and maximal projection.

**Skin phenotype calls.** Phenotype calls were based on macroscopic observation of tails and confocal microscopy of epidermal wholemounts. In some mutant mice, pigmentation in the IFE was altered and this was recorded under bright-field illumination.

Any change in overall tail structure was called with an MP term 'abnormal tail morphology (MP:0002111)'. Altered K14 expression (for example, reduced staining) was called as 'abnormal epidermis stratum basale morphology (MP:0001231)'. Other changes in the IFE, including hyperproliferation (detected as increased basal layer density and envelopment of the HF infundibulum), and hypocellularity were also recorded under this term. Abnormal K15 expression and/or abnormal bulge morphology was called as 'abnormal HF bulge morphology (MP:0009759)'. Any abnormalities in HF patterning (rows or triplets) were called with the term 'distorted HF pattern (MP:0000384)', while altered HF morphology, such as distorted follicles, was called as 'abnormal HF morphology (MP:0000377)'. The hair growth cycle in tail epidermis is less highly synchronized than in the back skin and the mice analysed were not age matched<sup>10</sup>. Nevertheless, obvious disturbance of the hair growth cycle, such as all HF being in anagen or telogen, was recorded with the term 'abnormal hair cycle (MP:0000427)'. Changes in SG size, number or morphology were recorded as 'abnormal SG morphology (MP:0000647)'. Altered pigmentation was called with the term 'abnormal epidermal pigmentation (MP:0009387)'.

Samples were examined and data were submitted without prior knowledge of genotype. When a mutant had no phenotype, the results were recorded on the tick sheet in the Sanger Mouse Genotype Portal (<http://www.sanger.ac.uk/mouseportal>). When a phenotype was abnormal, the information was summarized on the tick sheet and an illustrative 3D maximal projection image ( $\times 4$ ,  $\times 10$  or both) was uploaded to the portal. Images were optimized globally for brightness, contrast and colour balance using Photoshop CS5 (Adobe image suite) before uploading. If there was incomplete allelic penetrance of a gene, this was recorded in the portal.

**Skin-specific lacZ reporter gene expression.** lacZ reporter gene expression in skin was determined in wholemount tissue preparations from adult heterozygous mutant mice (6–12 weeks old)<sup>57</sup>. In brief, under terminal anaesthesia, mice were perfused with fresh cold 4% PFA, pH 8. Tissues were collected, fixed by immersion in 4% PFA for 30 min, and then incubated in lacZ staining solution containing 0.1% X-Gal (Invitrogen) at 4 °C for 48 h followed by post-fixation with 4% PFA at 4 °C. Tissues were cleared and stored in glycerol. Abdomen, pinna and tail skin were examined and staining was compared with a panel of standardized images. Images were taken using a Leica MZ16A microscope and processed by Imagic software. The pictures were annotated with MA terms before being released externally (<http://www.sanger.ac.uk/mouseportal>).

**Histochemistry.** Tissue was fixed overnight in 10% neutral buffered formalin and was paraffin embedded. Sections were labelled with haematoxylin and eosin, antibody to Ki67 or Fontana-Masson stain by the histology core services of the Wellcome Trust-MRC Cambridge Stem Cell Institute or the CRUK Cambridge Research Institute. Sections were examined and photographed on an Axiophot microscope with an AxioCam HRC camera (Zeiss, Germany). Other antibodies were used at the following dilutions for immunofluorescence: mouse anti-FASN (SCBT; sc48357)—1:100; rabbit anti-Filaggrin (Covance; PRB-417P)—1:100; rabbit anti-Loricrin (Covance; PRB-145P)—1:100; and mouse anti-K10 (Covance; MMS-159S)—1:100. Alexa fluor (Life Technologies) dye-conjugated secondary antibodies were used at 1:250 dilution.

**Long- and short-range PCR and DNA sequencing.** Tail DNA was prepared using a standard proteinase K and SDS lysis method<sup>58</sup>. Long- and short-range PCR were performed to re-confirm genotypes<sup>12</sup>. The following universal cassette and gene-specific primers were used for PCR amplification and DNA sequencing (see also Supplementary Fig. 4):

lacZ-genotype-For-5'-TGTATGAACGGTCTGGTCTTTG-3';  
lacZ-genotype-Rev-5'-CCAATCCACATCTGTGAAAGAA-3';  
L1-5'-CAACGGGTCTTCTGTGTAGTCC-3';  
LR-5'-ACTGATGGCGAGCTCAGACC-3';  
LRR-5'-GGTCTGAGCTCGCCATCAGT-3';

R1-5'-AAGGCGCATAACGATACCAC-3';  
 R2-5'-CCGCTACTGCGACTATAGA-3';  
 R2R-5'-TCTATAGTCGACGTAGGCGG-3';  
 R3-5'-GCGGATAACAATTTACACAGGA-3';  
 R4-5'-TGTAACACGACGGCCAGT-3';  
 PNF-5'-ATCCGGGGGTACCGCGTCGAG-3';  
 Rad18-5' HA-5'-TGTGTAATCCTGCCTATCC-3';  
 Rad18-3' HA-5'-CACCTTGGCTTTAAGTGG-3';  
 Rad18-out5' HA-5'-ACTTGGCTGGGAAGACATTG-3';  
 Rad18-out3' HA-5'-CCAGGATGTAACCTGCACT-3'.

**GO and pathway network analysis.** Enrichment of GO terms was performed using GoTermFinder (<http://www.go.princeton.edu/cgi-bin/GoTermFinder>)<sup>59</sup>. The REVIGO web server (<http://www.revigo.irb.hr/>) was used to generate GO scatter plots<sup>60</sup>. Localization of genes in the mouse karyotype was performed using Ensembl ([http://www.ensembl.org/Mus\\_musculus/Info/Index](http://www.ensembl.org/Mus_musculus/Info/Index)).

Direct (physical) and indirect (functional) protein associations were taken from the STRING database<sup>31</sup> and members of each pathway were extracted from Reactome<sup>61</sup>. Interactions were assessed in both mouse and human. Enrichment analysis was performed using ToppGene<sup>62</sup> on all 50 skin phenotype hits as well as the genes annotated to each specific MP term.

**Availability of biological resources and accessing the data.** Data generated by this study can be viewed via the Sanger Mouse Resources Portal at <http://www.sanger.ac.uk/mouseportal/> by entering the gene name of interest in the search menu. Users can then navigate through 'Phenotyping Test based Heatmap' to the 'Tail Epidermis Wholemount' section. Information, including ordering instructions, for all the mutant lines described in this study can be obtained either from the Sanger Mouse Resources Portal or the International Mouse Phenotyping Consortium URL: <http://www.mousephenotype.org/>.

## References

- Brenner, S. Sequences and consequences. *Phil. Trans. R. Soc. B* **365**, 207–212 (2010).
- Lander, E. S. Initial impact of the sequencing of the human genome. *Nature* **470**, 187–197 (2011).
- Tang, T. *et al.* A mouse knockout library for secreted and transmembrane proteins. *Nat. Biotechnol.* **28**, 749–755 (2010).
- Kumar, V. *et al.* Second-generation high-throughput forward genetic screen in mice to isolate subtle behavioral mutants. *Proc. Natl Acad. Sci. USA* **108**(Suppl 3): 15557–15564 (2011).
- Arwert, E. N., Hoste, E. & Watt, F. M. Epithelial stem cells, wound healing and cancer. *Nat. Rev. Cancer* **12**, 170–180 (2012).
- Fuchs, E. Scratching the surface of skin development. *Nature* **445**, 834–842 (2007).
- Benavides, F., Oberszyn, T. M., VanBuskirk, A. M., Reeve, V. E. & Kusewitt, D. F. The hairless mouse in skin research. *J. Dermatol. Sci.* **53**, 10–18 (2009).
- Mulder, K. W. *et al.* Diverse epigenetic strategies interact to control epidermal differentiation. *Nat. Cell Biol.* **14**, 753–763 (2012).
- Beronja, S. *et al.* RNAi screens in mice identify physiological regulators of oncogenic growth. *Nature* **501**, 185–190 (2013).
- Braun, K. M. *et al.* Manipulation of stem cell proliferation and lineage commitment: visualisation of label-retaining cells in wholemounts of mouse epidermis. *Development* **130**, 5241–5255 (2003).
- Kretschmar, K. W. & Watt, F. M. Markers of epidermal stem cell subpopulations in adult mammalian skin. *Cold Spring Harb Perspect Biol.* doi: 10.1101/cshperspect.a013631 (2014).
- Skarnes, W. C. *et al.* A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337–342 (2011).
- White, J. K. *et al.* Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* **154**, 452–464 (2013).
- Eppig, J. T. Algorithms for mutant sorting: the need for phenotype vocabularies. *Mamm Genome* **11**, 584–589 (2000).
- Smith, C. L., Goldsmith, C. A. & Eppig, J. T. The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome Biol.* **6**, R7 (2005).
- Gomez, C. *et al.* The interfollicular epidermis of adult mouse tail comprises two distinct cell lineages that are differentially regulated by Wnt, Edaradd, and Lrig1. *Stem Cell Rep.* **1**, 19–27 (2013).
- Jensen, K. B. *et al.* Lrig1 expression defines a distinct multipotent stem cell population in mammalian epidermis. *Cell Stem Cell* **4**, 427–439 (2009).
- Blanco, S. *et al.* The RNA-methyltransferase Mtsu (NSun2) poises epidermal stem cells to differentiate. *PLoS Genet.* **7**, e1002403 (2011).
- Frances, D. & Niemann, C. Stem cell dynamics in sebaceous gland morphogenesis in mouse skin. *Dev. Biol.* **363**, 138–146 (2012).
- Kellermayer, R. *et al.* Bigenic connexin mutations in a patient with hidrotic ectodermal dysplasia. *Eur. J. Dermatol.* **15**, 75–79 (2005).
- Tang, S., Han, H. & Bajic, V. B. ERGDB: estrogen responsive genes database. *Nucleic Acids Res.* **32**, D533–D536 (2004).
- Robinson, P. N. *et al.* The human phenotype ontology: a tool for annotating and analyzing human hereditary disease. *Am. J. Hum. Genet.* **83**, 610–615 (2008).
- Lopez-Pajares, V., Yan, K., Zarnegar, B. J., Jameson, K. L. & Khavari, P. A. Genetic pathways in disorders of epidermal differentiation. *Trends Genet.* **29**, 31–40 (2013).
- Doma, E., Rupp, C. & Baccarini, M. EGFR-ras-raf signaling in epidermal stem cells: roles in hair follicle development, regeneration, tissue remodeling and epidermal cancers. *Int. J. Mol. Sci.* **14**, 19361–19384 (2013).
- Ambatipudi, S. *et al.* Downregulation of keratin 76 expression during oral carcinogenesis of human, hamster and mouse. *PLoS One* **8**, e70688 (2013).
- Pastural, E. *et al.* Griscelli disease maps to chromosome 15q21 and is associated with mutations in the myosin-Va gene. *Nat. Genet.* **16**, 289–292 (1997).
- Alsultan, A., Shamseldin, H. E., Osman, M. E., Aljabri, M. & Alkuraya, F. S. MYSM1 is mutated in a family with transient transfusion-dependent anemia, mild thrombocytopenia, and low NK- and B-cell counts. *Blood* **122**, 3844–3845 (2013).
- Wang, T. *et al.* The control of hematopoietic stem cell maintenance, self-renewal, and differentiation by Mym1-mediated epigenetic regulation. *Blood* **122**, 2812–2822 (2013).
- Gu, L. H. & Coulombe, P. A. Keratin function in skin epithelia: a broadening palette with surprising shades. *Curr. Opin. Cell Biol.* **19**, 13–23 (2007).
- Watt, F. M. & Collins, C. A. Role of beta-catenin in epidermal stem cell expansion, lineage selection, and cancer. *Cold Spring Harb. Symp. Quant. Biol.* **73**, 503–512 (2008).
- Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–D815 (2013).
- Gutierrez, J. A. *et al.* A novel non-sense mutation in keratin 10 causes a familial case of recessive epidermolytic ichthyosis. *Mol. Genet. Genomic Med.* **1**, 108–112 (2013).
- Mercer, J. A., Seperack, P. K., Strobel, M. C., Copeland, N. G. & Jenkins, N. A. Novel myosin heavy chain encoded by murine dilute coat colour locus. *Nature* **349**, 709–713 (1991).
- Fukuda, M. & Kuroda, T. S. Missense mutations in the globular tail of myosin-Va in dilute mice partially impair binding of Slac2-a/melanophilin. *J. Cell Sci.* **117**, 583–591 (2004).
- Hume, A. N., Ushakov, D. S., Tarafder, A. K., Ferenczi, M. A. & Seabra, M. C. Rab27a and MyoVa are the primary Mplh interactors regulating melanosome transport in melanocytes. *J. Cell Sci.* **120**, 3111–3122 (2007).
- Klomp, A. E., Teofilo, K., Legacki, E. & Williams, D. S. Analysis of the linkage of MYRIP and MYO7A to melanosomes by RAB27A in retinal pigment epithelial cells. *Cell Motil. Cytoskeleton* **64**, 474–487 (2007).
- Gibbs, D. *et al.* Function of MYO7A in the human RPE and the validity of shaker1 mice as a model for Usher syndrome 1B. *Invest. Ophthalmol. Vis. Sci.* **51**, 1130–1135 (2010).
- Clague, M. J., Coulson, J. M. & Urbe, S. Deciphering histone 2A deubiquitination. *Genome Biol.* **9**, 202 (2008).
- Zhu, P. *et al.* A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol. Cell* **27**, 609–621 (2007).
- Nguyen, H. *et al.* Tcf3 and Tcf4 are essential for long-term homeostasis of skin epithelia. *Nat. Genet.* **41**, 1068–1075 (2009).
- Lindsay, C. R. *et al.* P-Rex1 is required for efficient melanoblast migration and melanoma metastasis. *Nat. Commun.* **2**, 555 (2011).
- Li, A. *et al.* Rac1 drives melanoblast organization during mouse development by orchestrating pseudopod-driven motility and cell-cycle progression. *Dev. Cell* **21**, 722–734 (2011).
- Trantow, C. M., Cuffy, T. L., Fingert, J. H., Kuehn, M. H. & Anderson, M. G. Microarray analysis of iris gene expression in mice with mutations influencing pigmentation. *Invest. Ophthalmol. Vis. Sci.* **52**, 237–248 (2011).
- Ando, H., Ichihashi, M. & Hearing, V. J. Role of the ubiquitin proteasome system in regulating skin pigmentation. *Int. J. Mol. Sci.* **10**, 4428–4434 (2009).
- Bucan, M., Eppig, J. T. & Brown, S. Mouse genomics programs and resources. *Mamm. Genome* **23**, 479–489 (2012).
- Driskell, R. R. *et al.* Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature* **504**, 277–281 (2013).
- Bassett, J. H. *et al.* Rapid-throughput skeletal phenotyping of 100 knockout mice identifies 9 new genes that determine bone strength. *PLoS Genet.* **8**, e1002858 (2012).
- Schmidt, F. *et al.* The role of conserved PEX3 regions in PEX19-binding and peroxisome biogenesis. *Traffic* **13**, 1244–1260 (2012).
- Karnik, P. *et al.* Hair follicle stem cell-specific PPARGgamma deletion causes scarring alopecia. *J. Invest. Dermatol.* **129**, 1243–1257 (2009).
- Cheung, C., Davies, N. G., Hoog, J. O., Hotchkiss, S. A. & Smith Pease, C. K. Species variations in cutaneous alcohol dehydrogenases and aldehyde



- dehydrogenases may impact on toxicological assessments of alcohols and aldehydes. *Toxicology* **184**, 97–112 (2003).
51. Muzio, G., Maggiora, M., Paiuzzi, E., Oraldi, M. & Canuto, R. A. Aldehyde dehydrogenases and cell proliferation. *Free Radic. Biol. Med.* **52**, 735–746 (2012).
  52. Petiot, A. *et al.* A crucial role for Fgfr2-IIIb signalling in epidermal development and hair follicle patterning. *Development* **130**, 5493–5501 (2003).
  53. Schlake, T. & Sick, S. Canonical WNT signalling controls hair follicle spacing. *Cell Adh. Migr.* **1**, 149–151 (2007).
  54. Van Loo, P. L., Van Zutphen, L. F. & Baumans, V. Male management: coping with aggression problems in male laboratory mice. *Lab. Anim.* **37**, 300–313 (2003).
  55. Azzi, L., El-Alfy, M., Martel, C. & Labrie, F. Gender differences in mouse skin morphology and specific effects of sex steroids and dehydroepiandrosterone. *J. Invest. Dermatol.* **124**, 22–27 (2005).
  56. Jensen, K. B. & Watt, F. M. Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence. *Proc. Natl. Acad. Sci. USA* **103**, 11958–11963 (2006).
  57. Valenzuela, D. M. *et al.* High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat. Biotechnol.* **21**, 652–659 (2003).
  58. Goldenberger, D., Perschil, I., Ritzler, M. & Altwegg, M. A simple 'universal' DNA extraction procedure using SDS and proteinase K is compatible with direct PCR amplification. *PCR Methods Appl.* **4**, 368–370 (1995).
  59. Boyle, E. I. *et al.* GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched gene ontology terms associated with a list of genes. *Bioinformatics* **20**, 3710–3715 (2004).
  60. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS One* **6**, e21800 (2011).
  61. Croft, D. *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691–D697 (2011).
  62. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009).

## Acknowledgements

This work was funded by the Wellcome Trust (including Grant number 098051 to K.P.S.), Medical Research Council and European Union (including EUMODIC contract number LSHG-CT-2006-037188 to K.P.S.). F.M.W. gratefully acknowledges the financial support of the Wellcome Trust, MRC, CR-UK and EU. I.S. acknowledges support from

a Discovery Project Grant from the Australian Research Council (DP1092723) and NH&MRC R. Douglas Wright and ARC Future Fellowships. We thank Celine Gomez for providing *Mysm1* scale staining, Elizabeth MacRae for technical support and Giacomo Donati for *Krt76 Scd1* staining. We thank Beate Lichtenberger, Gernot Walko and Ryan Driskell for suggestions on the manuscript, Wesley Chua for help with statistical analysis and all Wattlab members for helpful discussions. We thank Chris Lelliott for suggestions on the phenotype screen. We thank the imaging and histology core facilities at the Wellcome Trust-MRC Cambridge Stem Cell Institute and John Harris and all staff at the Nikon@King's imaging centre for their excellent support. We also thank the histopathology core facilities of the CR-UK Cambridge Research Institute for technical support. We thank the staff of the Sanger Institute's Mouse Genetics Project for generating, characterising and collecting samples from the mutant mice used in this screen.

## Author contributions

F.M.W. and I.S. conceived the project. K.L. performed experiments and had overall responsibility for the screen. E.H., T.D. and I.S. also performed experiments. V.V. contributed to the construction of data. D.P.S. performed computational network analysis. J.E., D.S., J.W., R.R. and K.P.S. managed the S.I.M.G.P. for production and primary phenotyping of mutants. K.L. and F.M.W. analysed the data and wrote the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Liakath-Ali, K. *et al.* Novel skin phenotypes revealed by a genome-wide mouse reverse genetic screen. *Nat. Commun.* 5:3540 doi: 10.1038/ncomms4540 (2014).



This work is licensed under a Creative Commons Attribution 3.0 Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>

# ABCA12 Regulates ABCA1-Dependent Cholesterol Efflux from Macrophages and the Development of Atherosclerosis

Ying Fu,<sup>1</sup> Nigora Mukhamedova,<sup>1</sup> Sally Ip,<sup>2</sup> Wilissa D'Souza,<sup>1</sup> Katya J. Henley,<sup>4,5</sup> Tia DiTommaso,<sup>2</sup> Rajitha Kesani,<sup>1</sup> Michael Ditiatkovski,<sup>1</sup> Lynelle Jones,<sup>2</sup> Rachael M. Lane,<sup>4,5</sup> Garry Jennings,<sup>1</sup> Ian M. Smyth,<sup>2,3</sup> Benjamin T. Kile,<sup>4,5</sup> and Dmitri Sviridov<sup>1,2,\*</sup>

<sup>1</sup>Baker IDI Heart and Diabetes Institute, Melbourne, Victoria 3004, Australia

<sup>2</sup>Department of Biochemistry and Molecular Biology

<sup>3</sup>Department of Anatomy and Developmental Biology  
Monash University, Clayton, Victoria 3800, Australia

<sup>4</sup>Cancer and Hematology Division, The Walter and Eliza Hall Institute for Medical Research, Parkville, Victoria 3052, Australia

<sup>5</sup>Department of Medical Biology, The University of Melbourne, Parkville, Victoria 3010, Australia

\*Correspondence:

<http://dx.doi.org/10.1016/j.cmet.2013.07.003>

## SUMMARY

ABCA12 is involved in the transport of ceramides in skin, but it may play a wider role in lipid metabolism. We show that, in *Abca12*-deficient macrophages, cholesterol efflux failed to respond to activation with LXR agonists. *Abca12* deficiency caused a reduction in the abundance of *Abca1*, *Abcg1*, and *Lxrβ*. Overexpression of *Lxrβ* reversed the effects. Mechanistically, *Abca12* deficiency did not affect expression of genes involved in cholesterol metabolism. Instead, a physical association between *Abca1*, *Abca12*, and *Lxrβ* proteins was established. *Abca12* deficiency enhanced interaction between *Abca1* and *Lxrβ* and the degradation of *Abca1*. Overexpression of ABCA12 in HeLa-ABCA1 cells increased the abundance and stability of ABCA1. *Abca12* deficiency caused an accumulation of cholesterol in macrophages and the formation of foam cells, impaired reverse cholesterol transport in vivo, and increased the development of atherosclerosis in irradiated *Apoe*<sup>-/-</sup> mice reconstituted with *Apoe*<sup>-/-</sup>*Abca12*<sup>-/-</sup> bone marrow. Thus, ABCA12 regulates the cellular cholesterol metabolism via an LXRβ-dependent posttranscriptional mechanism.

## INTRODUCTION

The dysregulation of cholesterol homeostasis in macrophages is a pivotal event in the pathogenesis of atherosclerosis, the major cause of cardiovascular disease. Even small fluctuations of cellular cholesterol content can cause apoptosis and necrosis. Macrophages, like all cells, possess a sophisticated regulatory network controlling cellular cholesterol homeostasis (Maxfield and Tabas, 2005). These pathways control, on the one hand, cholesterol biosynthesis and the uptake of cholesterol in low-

density lipoproteins (LDLs) and, on the other, the release of cellular cholesterol to the extracellular acceptors, such as apolipoprotein A-I (apoA-I) or high-density lipoprotein (HDL). Although mechanisms controlling cholesterol delivery have been investigated in great detail (Brown and Goldstein, 2009), the mechanisms responsible for regulating cholesterol efflux are less clear. They are particularly relevant for macrophages, given that cholesterol efflux in these cells is the major determinant of cholesterol homeostasis.

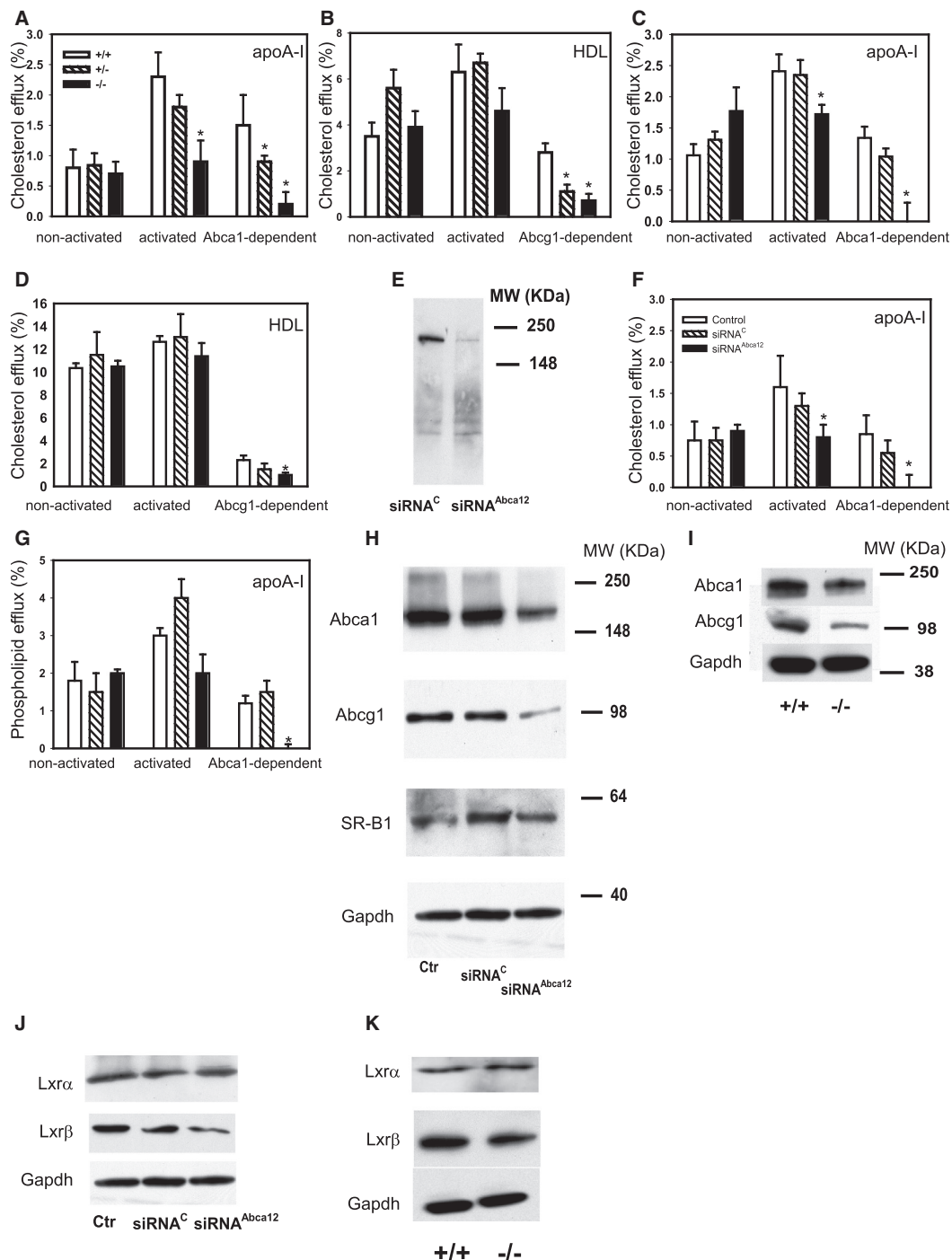
Three transporters responsible for regulated cholesterol efflux have been described: ABCA1, which is responsible for the efflux to lipid-free apoA-I, and ABCG1 and SR-B1, which regulate efflux to the mature HDL (Yvan-Charvet et al., 2010). However, numerous studies have indicated that our understanding of the regulation of cholesterol homeostasis is far from complete (for review, see Fitzgerald et al., 2010). Recently, we described a mouse carrying a mutation in another transporter, *Abca12* (Smyth et al., 2008), a gene best known for being mutated in a skin disease, Harlequin ichthyosis (HI). In the skin, ABCA12 plays a critical role in the transport of ceramides in keratinocytes (Kells et al., 2005); however, we found that, in fibroblasts lacking *Abca12*, cholesterol homeostasis was also perturbed (Smyth et al., 2008). This prompted us to investigate the broader role of ABCA12 in cholesterol metabolism and the development of atherosclerosis.

## RESULTS

### *Abca12* Deficiency Impairs Cholesterol Efflux from Macrophages

To examine the role of *Abca12* in cholesterol efflux, we derived macrophages from the fetal livers of mouse embryos carrying the *E112* allele of *Abca12* (Smyth et al., 2008). ABCA1 is the sole determinant of specific cholesterol efflux to lipid-free apoA-I; therefore, apoA-I was used as an acceptor in order to assess ABCA1-dependent efflux. In cells not activated with LXR agonist, cholesterol efflux to apoA-I was similar in cells of all genotypes (Figure 1A). After the activation of *Abca1* expression with the LXR agonist TO-901317, cholesterol efflux was elevated in *Abca12*<sup>+/+</sup>, but not mutant, cells. Consequently,





**Figure 1. Cholesterol and Phospholipid Efflux from Abca12-Deficient Cells**

(A and B) Cholesterol efflux to apoA-I (A) or HDL (B) from liver macrophages isolated from the *Abca12*<sup>E112</sup> strain. (C and D) Cholesterol efflux to apoA-I (C) or HDL (D) from liver macrophages isolated from the *Abca12*<sup>Lx12</sup> strain. (E) Abundance of Abca12 in RAW 264.7 cells after knockdown of Abca12 with siRNA. (F) Cholesterol efflux to apoA-I from RAW 264.7 cells after knockdown of Abca12 with siRNA. (G) Phospholipid efflux to apoA-I from liver macrophages isolated from *Abca12*<sup>E112</sup> strain. (H) Abundance of Abca1, Abcg1, and SR-B1 in RAW 264.7 cells after knockdown of Abca12 with siRNA. (I) Abca1 and Abcg1 abundance in liver macrophages isolated from the *Abca12*<sup>Lx12</sup> strain.

(legend continued on next page)

specific Abca1-dependent efflux was reduced in *Abca12*<sup>+/E112</sup> cells and almost undetectable in *Abca12*<sup>E112/E112</sup> cells (Figure 1A).

ABCG1 and SR-B1 are the principle mediators of specific cholesterol efflux to HDL. Expression of ABCG1, but not SR-B1, is also activated by LXR agonists. We used HDL as an acceptor and cells activated with LXR agonist to test the effect of *Abca12* deficiency on ABCG1-dependent efflux. In the absence of agonist, efflux was comparable in cells of all genotypes. Upon treatment with an LXR agonist, cholesterol efflux increased in *Abca12*<sup>+/+</sup>, but not *Abca12*<sup>E112/E112</sup>, cells (Figure 1B). Similar results were obtained with macrophages derived from mice carrying the independent *Lx12* knockout allele of *Abca12* (Yanagi et al., 2010). Fetal macrophages with this mutation exhibited similar defects in cholesterol efflux to apoA-I (Figure 1C) and HDL (Figure 1D) upon activation with LXR agonist. Nonspecific efflux was unaffected by a loss of *Abca12* (Table S1 available online). These observations were also confirmed in the RAW 264.7 mouse macrophage cells upon small interfering RNA (siRNA)-mediated *Abca12* knockdown (Figures 1E and 1F). The effects of *Abca12* deficiency on phospholipid efflux, a process also controlled by ABCA1, were similar to the effects on cholesterol efflux, except that phospholipid efflux was not impaired in heterozygous cells (Figure 1G).

Given that LXR activation triggers the upregulation of *Abca1* and *Abcg1* expression, we measured the effect of *Abca12* deficiency on levels of these proteins. *Abca12* depletion by siRNA<sup>*Abca12*</sup> resulted in a substantial reduction in the levels of *Abca1* and *Abcg1*, but not SR-B1, in cells activated with LXR agonist (Figure 1H). The level of *Abca1* after *Abca12* silencing was 66.1% ± 4.4% of control (mean ± SEM; n = 7). Similar results were observed in *Abca12*<sup>Lx12/Lx12</sup> cells (Figure 1I). Interestingly, in both siRNA-treated RAW 264.7 cells and *Abca12*<sup>Lx12/Lx12</sup>-derived macrophages, the abundance of *Lxrβ*, but not *Lxrα*, was also reduced (Figures 1J and 1K). The level of *Lxrβ* after *Abca12* silencing was 58.5% ± 5.1% of control (mean ± SEM; n = 6). Collectively, these data demonstrate that *Abca12* deficiency impairs the ability of macrophages to upregulate the levels of key ATP-binding cassette (ABC) transporters and cholesterol efflux in response to LXR activation. Both ABCA1 and ABCG1 are known to play a significant role in cellular cholesterol homeostasis. Whereas the precise role of ABCG1 is still contentious, the role of ABCA1 is well established. Therefore, we focused more detailed mechanistic studies on the relationship between ABCA12 and ABCA1.

### The Role of *Lxrβ* in Impairment of Cholesterol Efflux in *Abca12*-Deficient Cells

Given that LXR primarily mediates its effects via the regulation of gene transcription, we examined the impact of *Abca12* deficiency on the expression of genes involved in lipid metabolism. Using real time RT-PCR, we assessed messenger RNA (mRNA) levels of *Lxrα*, *Lxrβ*, two LXR target genes (*Abca1* and *Abcg1*), and several other genes involved in cellular lipid metabolism (*Ldlr*, *Scarb1*, *Msr1*, *Abca3*, and *Hmgcr*). As expected,

*Abca12* mRNA was undetectable in *Abca12*<sup>Lx12/Lx12</sup> macrophages (Figure 2A). *Abca12* deficiency did not have a statistically significant effect on *Abca1* and *Abcg1* expression with or without stimulation by LXR agonists, nor did it have an effect on the expression of any other genes tested (Figure 2A). Identical outcomes were observed when gene expression was analyzed in RAW 264.7 cells with siRNA knockdown of *Abca12* (Figure 2B). Thus, a loss of *Abca12* does not impair cholesterol efflux via transcriptional regulation of the key mediators of cholesterol homeostasis. This is to be expected, given the functional redundancy between *Lxrβ* and *Lxrα* and the only partial decrease of *Lxrβ* levels in *Abca12*-deficient cells.

To control for agonist-specific effects, we treated cells with a chemically distinct LXR agonist, GW3965 (Joseph et al., 2002), and observed similar results (Figures 2C and 2D). We also tested whether the effects of *Abca12* deficiency are mediated through another regulator of ABC transporters, PPAR (Ogata et al., 2009). Treatment with the PPARα/γ agonist RO4509851 did not reverse the effect of *Abca12* deficiency (Figure 2C).

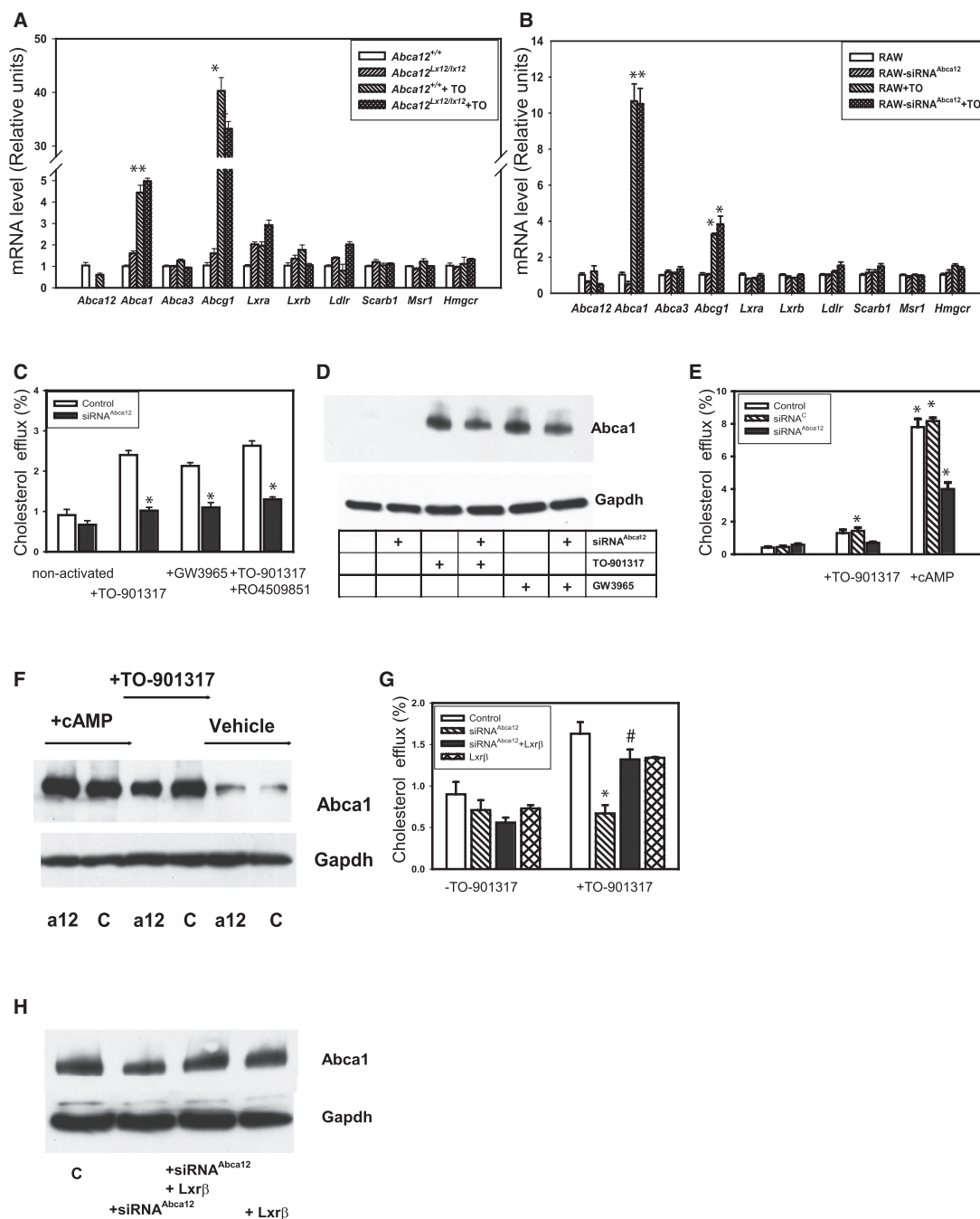
In addition to the LXR-dependent regulatory pathway, murine cells can activate *Abca1* expression and cholesterol efflux in response to treatment with cAMP (Haidar et al., 2002; Le Goff et al., 2006). When RAW 264.7 *Abca12* knockdown cells were activated with cAMP, cholesterol efflux was elevated by about half the level observed in cells transfected with siRNA<sup>c</sup> (Figure 2E). cAMP modulates cholesterol efflux via several transcriptional (Le Goff et al., 2006) and posttranscriptional (Haidar et al., 2002) mechanisms. Although the abundance of *Abca1* protein was reduced in siRNA<sup>*Abca12*</sup>-transfected cells activated with LXR agonist, the levels of *Abca1* mRNA (Figure S1) and protein (Figure 2F) were unaffected by *Abca12* deficiency after activation with cAMP. We did not observe any effect of *Abca12* deficiency on the cAMP-induced changes in the expression of nine genes involved in the regulation of lipid metabolism (Figure S1). These observations suggest that the effects of *Abca12* deficiency on *Abca1* abundance are specific for the *Lxr*-dependent pathway. However, partial inhibition of cholesterol efflux indicates that *Abca12* deficiency affected *Abca1* functionality independently of its abundance.

To investigate the specific role of *Lxrβ* in the impairment of cholesterol efflux in *Abca12*-deficient cells, we conducted “rescue” experiments in which RAW 264.7 cells were transfected with siRNA<sup>*Abca12*</sup>, *Lxrβ* cDNA, or a combination of the two. Knockdown of *Abca12* rendered cells unresponsive to treatment with LXR agonist, whereas the transfection of *Abca12*-deficient cells with heterologous *Lxrβ* cDNA largely restored cholesterol efflux and *Abca1* abundance (Figures 2G and 2H). Transfection with *Lxrβ* cDNA alone had no effect (Figure 2G). These findings confirmed that the effect of *Abca12* deficiency on cholesterol efflux and *Abca1* abundance was at least partially mediated by *Lxrβ*. Altogether, these data demonstrate that *Abca12* regulates *Lxrβ* abundance and *Abca1* protein level and functionality via a posttranscriptional mechanism.

(J) *Lxrα* and *Lxrβ* abundance in RAW 264.7 cells after knockdown of *Abca12* with siRNA.

(K) *Lxrα* and *Lxrβ* abundance in liver macrophages isolated from *Abca12*<sup>Lx12</sup> strain.

Activated cells are cells that have been preincubated with LXR agonist TO-901317 (final concentration, 4 μM). \*p < 0.01 (versus WT cells). Pooled data are represented as mean ± SEM. See also Table S1.



**Figure 2. Involvement of Lxrβ in the Impairment of Cholesterol Efflux in Abca12-Deficient Cells**

(A) Expression of genes involved in lipid metabolism in *Abca12*<sup>Lx12/Lx12</sup> and *Abca12*<sup>Lx12/Lx12</sup> macrophages with or without activation with LXR agonist TO-901317 as assessed by real-time RT-PCR. \*p < 0.01 (versus nonactivated cells).

(B) Expression of genes involved in lipid metabolism in RAW 264.7 macrophages after knockdown of *Abca12*. Pooled data are represented as mean ± SEM. \*p < 0.01 (versus nonactivated cells).

(C) Cholesterol efflux to apoA-I from RAW 264.7 cells after knockdown of *Abca12* by siRNA and treatment with LXR agonist TO-901317 (4 μM), GW3965 (1 μM), or a combination of TO-901317 (4 μM) and PPARα/γ agonist RO4509851 (5 μM).

(D) Abundance of *Abca1* in RAW 264.7 cells after knockdown of *Abca12* by siRNA and activation with LXR ligands.

(E) Cholesterol efflux to apoA-I from RAW 264.7 cells after knockdown of *Abca12* and activation by LXR agonist TO-901317 (4 μM) or cAMP (0.3 mM).

(F) Abundance of *Abca1* in RAW 264.7 cells after knockdown of *Abca12* and activation of *Abca1* expression by TO-901317 or cAMP.

(legend continued on next page)



### Mechanisms of Modulation of Cholesterol Efflux by Abca12

Previous studies have proposed that LXR $\beta$  can bind directly to ABCA1, reducing its activity, and LXR agonists interfere with this interaction, reversing the inhibitory effect of LXR $\beta$  (Hozoji-Inada et al., 2011). An additional posttranslational mechanism that regulates ABCA1 activity and abundance is its relocalization: when ABCA1 is displaced from the cell surface, it no longer supports cholesterol efflux and is rapidly degraded (Neufeld et al., 2001). Therefore, we investigated the effects of Abca12 deficiency on Abca1 localization and colocalization of Abca1 and Lxr $\beta$ . The abundance of Abca1 in nonactivated RAW 264.7 cells was low, and there was a degree of colocalization of Abca1 and Lxr $\beta$  (colocalization coefficient,  $0.56 \pm 0.09$ ) (Figures 3A–3C). In cells activated with LXR agonist, the abundance of Abca1 increased considerably, but the colocalization of Abca1 and Lxr $\beta$  did not change significantly in comparison to the nonactivated cells (colocalization coefficient,  $0.51 \pm 0.04$ ) (Figures 3D–3F). Remarkably, the colocalization of Abca1 with Lxr $\beta$  dramatically increased in cells activated with LXR agonist and treated with siRNA<sup>Abca12</sup> (colocalization coefficient,  $0.81 \pm 0.04$ ,  $p < 0.01$  for cells treated with siRNA<sup>Abca12</sup> versus cells treated with siRNA<sup>0</sup>) (Figures 3D–3I, lower magnification images are shown in Figure S2A). The distribution of Abca1 was investigated via the biotinylation of cell-surface Abca1. We observed a reduction in the abundance of cell-surface Abca1 in Abca12-deficient cells (Figure 3J). This was also confirmed with the use of cross-sections of individual cells (Figure 3K) and in confocal three-dimensional reconstitutions, both of which showed a reduction of plasma membrane Abca1 upon Abca12 knockdown (Movie S1, control, and Movie S2, Abca12 knockdown). The level of Abca12 in wild-type (WT) RAW 264.7 cells was insufficient for reliable detection; however, overexpression of heterologous Abca12 showed clear regions of overlap with Abca1 (Figure 3L).

To investigate whether Abca1 stability was affected by Abca12 deficiency, we radioactively labeled newly synthesized Abca1 and performed pulse-chase experiments. Analysis demonstrated that the degradation rate of Abca1 was increased when Abca12 was knocked down (Figures 3M and 3N). Furthermore, the effect was confirmed when cell-surface Abca1 was labeled with biotin and the abundance of biotinylated Abca1 was followed; i.e., the rate of Abca1 degradation increased in Abca12-deficient cells (Figure S2B).

Clearly, Abca12 is able to regulate the stability of Abca1. To investigate whether a physical interaction between Abca12, Abca1, and Lxr $\beta$  may form the basis for these effects, we employed a proximity ligation assay (PLA) (Baima et al., 2010). In this assay, the number of fluorescent dots in a given cell is proportional to the number of interaction events between a pair of heterotypic proteins. When the Abca12 and Abca1 pair was studied, an abundant interaction was found, and knockdown of Abca12 significantly reduced this interaction (Figure 3O). When the Abca1 and Lxr $\beta$  pair was analyzed, the activation of

cells with an LXR agonist increased their interaction, and this was increased further after Abca12 silencing (Figure 3P and quantitation in Figure S2C). However, as expected, the abundance of both Abca1 and Lxr $\beta$  was affected by the addition of LXR agonist and Abca12 silencing (Figure S2D shows the abundance of Abca12 and Lxr $\beta$  in this experiment), and, when the interaction between Abca1 and Lxr $\beta$  was normalized to the abundance of the interacting proteins, it was statistically significantly reduced after the addition of LXR agonist and elevated after Abca12 silencing (Figure 3P). We also observed an abundant interaction between Abca12 and Lxr $\beta$ , which was reduced when Abca12 was knocked down (Figure 3Q).

Next, we examined HeLa cells stably transfected with GFP-ABCA1 expressed under the control of a cytomegalovirus (CMV) promoter (parental HeLa cells do not express homologous ABCA1) (Mukhamedova et al., 2007). When HeLa-ABCA1 cells were transfected with tGFP-ABCA12, the abundance of ABCA1 significantly increased (Figure 4A). Overexpression of ABCA12 in HeLa cells caused a modest elevation of cholesterol efflux; however, in HeLa-ABCA1 cells, this overexpression did not increase cholesterol efflux in addition to the efflux caused by the presence of heterologous ABCA1 (Figure 4B). Then, we investigated the distribution of ABCA1 and ABCA12 by confocal microscopy. Consistent with previous reports (Mujawar et al., 2006; Mukhamedova et al., 2007), most of the ABCA1 in HeLa-ABCA1 cells localized to the perinuclear area, a smaller proportion being localized at the plasma membrane (Figure 4C). Transfection with ABCA12 caused the redistribution of ABCA1 to the outer cytoplasm and plasma membrane (Figure 4D), which was similar to the distribution of ABCA12 (Figure 4E); there was a significant colocalization of ABCA1 and ABCA12 (Figure 4F).

To investigate whether overexpression of ABCA12 affected the stability of ABCA1, we used biotin labeling of cell-surface ABCA1 and found that, although the abundance of total ABCA1 was considerably higher in cells transfected with ABCA12 (Figure 4G, lane 1), levels of ABCA1 at the cell surface were similar (Figure 4G, lane 2). This observation explained the lack of effect of ABCA12 on cholesterol efflux in HeLa-ABCA1 cells. The rate of degradation of biotinylated ABCA1 was significantly reduced by ABCA12 (Figures 4G and 4H).

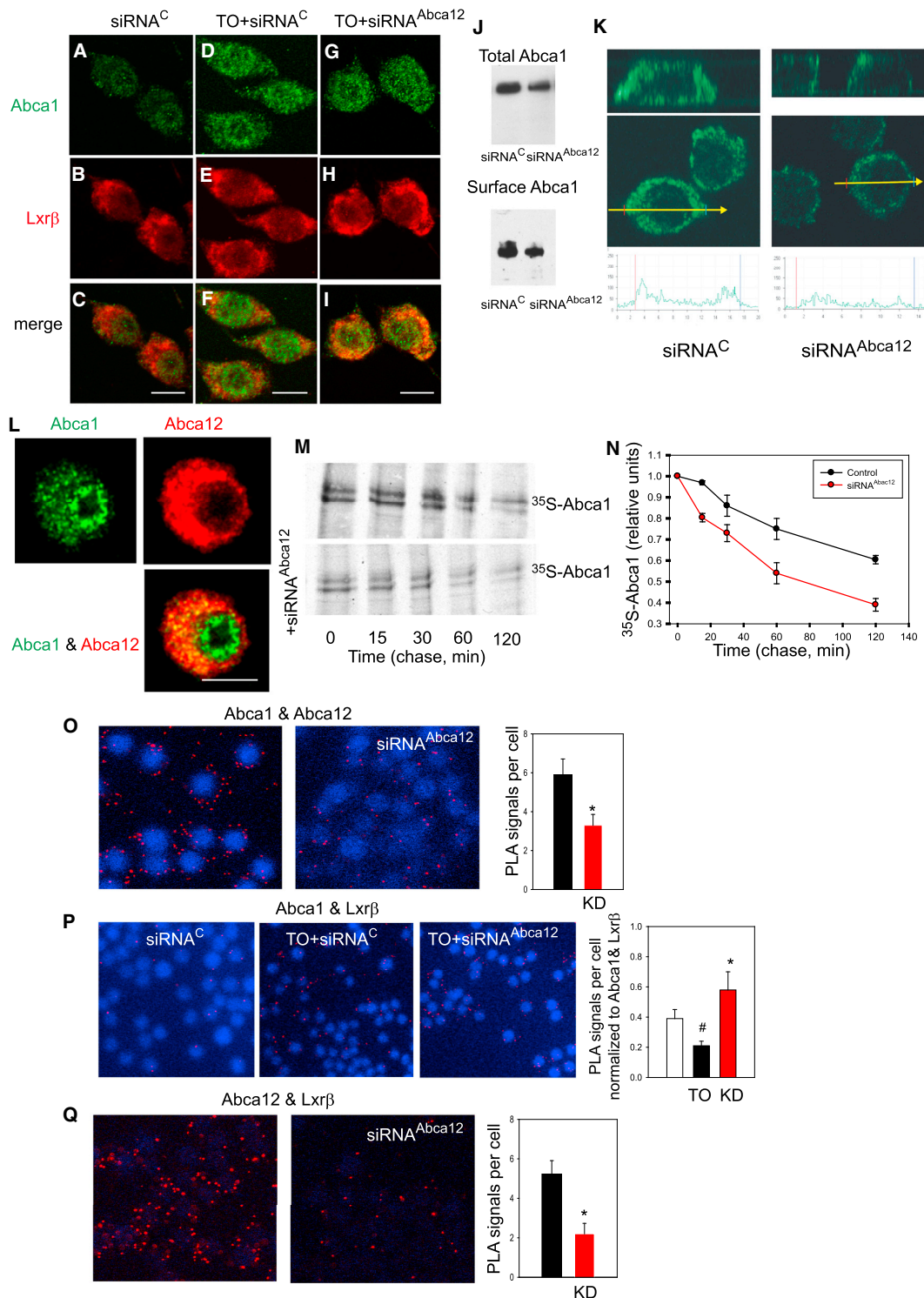
In pull-down experiments, an antibody against ABCA1 coprecipitated ABCA12 irrespective of whether or not proteins were crosslinked prior to immunoprecipitation (Figure 4I). In the reciprocal experiment, when ABCA12 was immunoprecipitated, ABCA1 coimmunoprecipitated with ABCA12 (Figure 4J). When LXR $\beta$  was immunoprecipitated from HeLa-ABCA1 cells transfected with ABCA12, both ABCA1 and ABCA12 coimmunoprecipitated with LXR $\beta$  (Figure 4K). Experiments with the PLA assay confirmed the interaction between ABCA1 and ABCA12 in this model (Figure 4L). They also revealed that the overexpression of ABCA12 increased the interaction of ABCA1 and LXR $\beta$  and caused a relocalization of this complex to a region surrounding the nucleus (Figure 4M). Note that cells in this experiment were

(G) Cholesterol efflux to apoA-I from RAW 264.7 cells after knockdown of Abca12 and cotransfection with Lxr $\beta$ . Pooled data are represented as mean  $\pm$  SEM.

\* $p < 0.01$  (versus control cells); # $p < 0.001$  (versus cells treated with siRNA).

(H) Abundance of Abca1 in RAW 264.7 cells after knockdown of Abca12 and cotransfection with Lxr $\beta$ .

See also Figure S1.



**Figure 3. Mechanism of Modulation of Cholesterol Efflux by Abca12: Studies on RAW 264.7 Cells**

(A–I) Confocal microscopy of Abca1 (A, D, and G; green), Lxrβ (B, E, and H; red), and a merge of the two (C, F, and I; yellow) in RAW 264.7 cells nonactivated (A–C) or activated with LXR agonist TO-901317 (D–I) with (G–I) or without (A–F) Abca12 knockdown. The scale bar represents 10  $\mu$ m.

(J) Abundance of total (upper panel) and cell-surface (bottom panel) Abca1 in RAW 264.7 cells with or without Abca12 knockdown.

(legend continued on next page)

not treated with LXR agonist and that the abundance of ABCA1 was elevated in cells overexpressing ABCA12.

Altogether, these findings indicate that ABCA12 interacts with ABCA1 and that a loss of ABCA12 reduces ABCA1 abundance on the cell surface and enhances its degradation. ABCA12 and ABCA1 are also likely to interact with LXR $\beta$ .

### Abca12-Deficient Macrophages Accumulate Cholesterol

In macrophages, deficiencies in cholesterol efflux lead to the accumulation of cholesterol and the transformation of macrophages into foam cells. To investigate the potential contribution of Abca12 to this process, we assessed the rate of cholesteryl ester biosynthesis. In *Abca12*<sup>E12/E12</sup> cells, the rate of cholesteryl ester biosynthesis was 3-fold higher compared to that of WT counterparts, whereas, under conditions of cholesterol loading, this difference was 2-fold (Figure 5A). Similar effects were observed in the Abca12-deficient RAW 264.7 cells (Figure 5B). Significantly more cells were stained with Oil Red O in Abca12 knockdown cells in comparison to control cells (Figures 5C and 5D).

Then, we performed a lipidomics analysis in order to investigate the effect of Abca12 knockdown on lipid accumulation in RAW 264.7 cells. Without cholesterol loading, Abca12 knockdown resulted in a modest elevation of the abundance of cholesteryl esters, ceramides, and sphingomyelins (Figure 5E). With cholesterol loading, Abca12 deficiency led to a significant accumulation of cholesterol, cholesteryl esters, ceramides, sphingomyelins, triglycerides, and phosphatidylcholine (Figure 5E). Full lipidomics data are presented in Table S2.

Altogether, these findings demonstrate that Abca12 deficiency results in the accumulation of lipids in macrophages.

### Loss of Abca12 Impairs Reverse Cholesterol Transport and Promotes Atherosclerosis In Vivo

Next, we assessed the effect of Abca12 deficiency on reverse cholesterol transport (RCT) in vivo. RAW 264.7 macrophages transfected with control or Abca12 siRNA were loaded with cholesterol, labeled, and transplanted into the peritoneal cavity of mice. The amount of labeled cholesterol exported from macrophages with Abca12 knockdown to plasma and feces was reduced 2-fold in comparison to control macrophages (Figures 6A and 6C). The amount of cholesterol in the liver was unaffected (Figure 6B), probably because of the transient nature of this pool. The ratio of free cholesterol to cholesteryl esters in plasma was

$0.56 \pm 0.27$  and  $0.57 \pm 0.49$  ( $p = 0.96$ ) for control and Abca12 knockdown groups, respectively.

Then, we investigated the effect of Abca12 deficiency on the development of atherosclerosis in the *Apoe*<sup>-/-</sup> mice. Given that *Abca12*<sup>E12/E12</sup> mice are not viable (Smyth et al., 2008), we used a bone marrow chimera model, transplanting fetal liver cells (FLCs) harvested from *Apoe*<sup>-/-</sup> *Abca12*<sup>+/+</sup> or *Apoe*<sup>-/-</sup> *Abca12*<sup>E12/E12</sup> embryos into *Apoe*<sup>-/-</sup> mice in which endogenous hematopoiesis had been ablated by irradiation. Bone marrow chimerism was confirmed by flow cytometric analysis of peripheral blood leukocytes 8 weeks posttransplantation, and mice exhibiting >90% donor engraftment commenced a high-fat diet. After 16 weeks of a high-fat diet, mice transplanted with *Apoe*<sup>-/-</sup> *Abca12*<sup>E12/E12</sup> FLCs exhibited double the number of the plaques in comparison to *Abca12*<sup>+/+</sup> counterparts (Figures 6D and 6H;  $p < 0.002$ ). The biggest effect was observed in the aortic arch, the area most susceptible to the development of atherosclerosis, (Figures 6E and 6H). There was a trend toward increased atherosclerosis in the thoracic aorta (Figure 6F), whereas no difference was found in the abdominal aorta (Figures 6G). Analysis of cross-sections of the aortic sinus revealed no difference between the two groups of mice in the abundance of macrophages in the plaque and expression of VCAM-1 on the surface of endothelium (Figures 6I and 6J). We found no difference in the abundances of collagen (Figure 6K), a marker of protein oxidation, nitrotyrosine (Figure 6L), or in the size of the necrotic core (Figure 6M).

Blood cell counts in the two groups of mice after 16 weeks on a high-fat diet revealed two statistically significant differences: in *Apoe*<sup>-/-</sup> *Abca12*<sup>E12/E12</sup> mice, the percentages of leukocytes were lower ( $6.3 \pm 0.5$  versus  $12.1\% \pm 1.0\%$ ;  $p < 0.002$ ), and the percent of mononuclear cells was higher ( $81.4\% \pm 1.9\%$  versus  $69.9\% \pm 1.4\%$ ;  $p < 0.01$ ). Plasma triglycerides, total, non-HDL, and HDL cholesterol levels in mice transplanted with *Apoe*<sup>-/-</sup> *Abca12*<sup>+/+</sup> or *Apoe*<sup>-/-</sup> *Abca12*<sup>E12/E12</sup> FLCs were similar (Table S3), indicating that the enhanced atherosclerosis was not the result of changes in plasma lipid profile.

These results indicate that Abca12 deficiency inhibits RCT in vivo and causes the accumulation of cholesterol in macrophages, promoting their transformation into foam cells and contributing to the development of atherosclerosis.

## DISCUSSION

ABCA12 is a full-length ubiquitously expressed ABC transporter. In the skin, ABCA12 is responsible for the transport of

(K) Intracellular distribution of Abca1 in RAW 264.7 cells or RAW 264.7 cells after Abca12 knockdown. The top panels show cross-sectional images derived from stacking the images from sequential scanning of cells shown in the middle panels at different depths. The bottom panel shows fluorescence intensity measured along the arrow in the middle panel.

(L) Intracellular distribution of Abca1 (green) and Abca12 (red) in RAW 264.7 cells overexpressing heterologous Abca12. The scale bar represents 10  $\mu$ m.

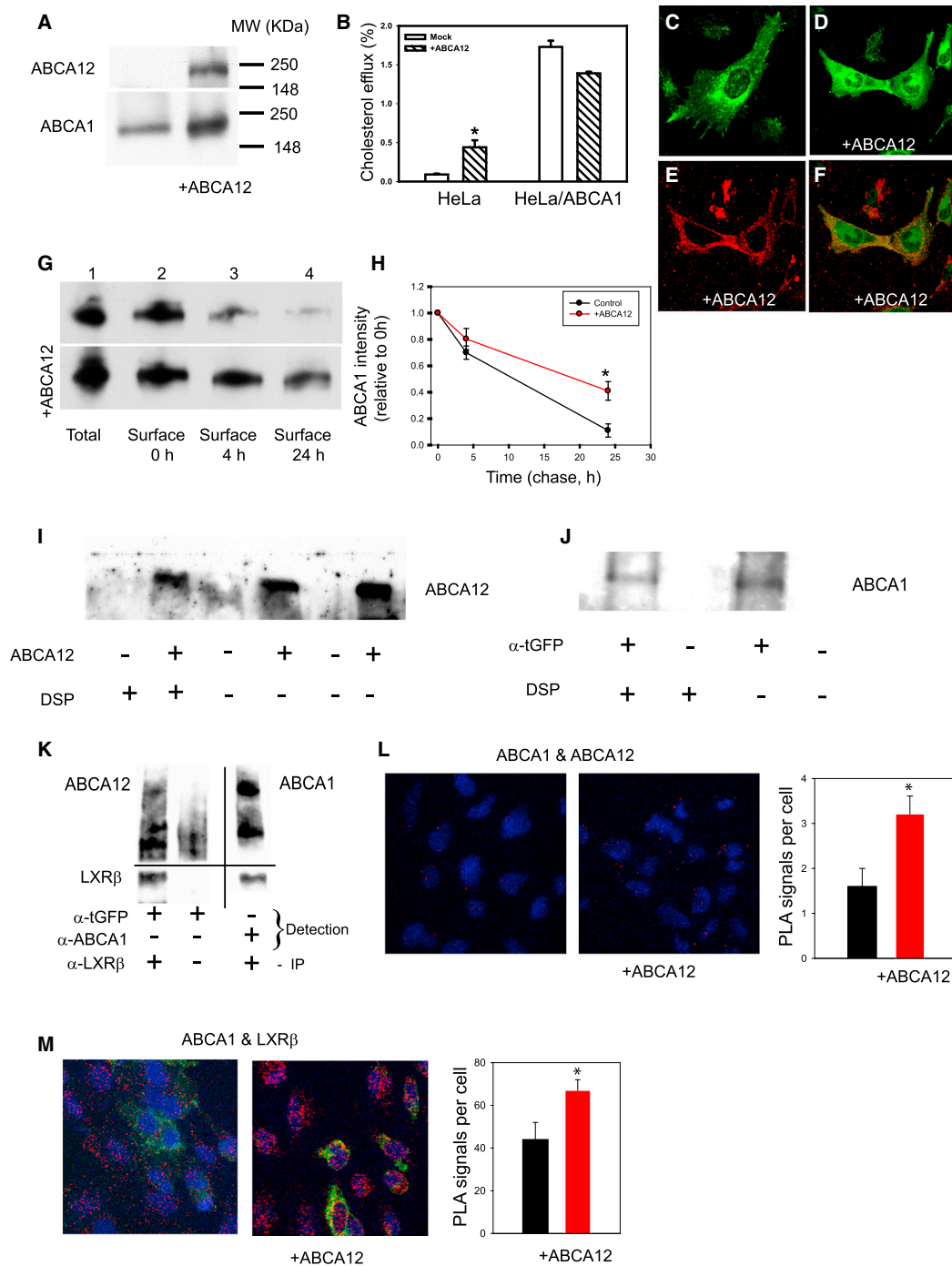
(M and N) A pulse-chase experiment of Abca1 degradation in RAW 264.7 cells with or without Abca12 knockdown. Proteins were pulse-labeled with <sup>35</sup>S-labeled methionine and cysteine, Abca1 immunoprecipitated, and isolated on SDS-PAGE. Autoradiography of the Abca1 bands (M) and quantitation of the radioactivity in these bands (N) are shown. Means  $\pm$  the range of two experiments are shown.

(O) Proximity ligation assay (PLA) of the Abca1 and Abca12 pair in RAW 264.7 cells with or without Abca12 knockdown.

(P) PLA of the Abca1 and Lxr $\beta$  pair in RAW 264.7 cells with or without activation with TO-901317 and with or without Abca12 knockdown. Graph shows PLA values normalized to the abundance of Abca1 and Lxr $\beta$ .

(Q) PLA of the Abca12 and Lxr $\beta$  pair in RAW 264.7 cells with or without Abca12 knockdown. Quantitation of PLA assay (right panels in O–Q) is described in the Experimental Procedures. Means  $\pm$  SEM of PLA signal per cell are shown. \* $p < 0.01$ (versus control); # $p < 0.01$ (versus nonactivated cells).

See also Figure S2 and Movies S1 and S2.



#### Figure 4. Mechanism of Modulation of Cholesterol Efflux by ABCA12: Studies on HeLa-ABCA1 Cells

(A) Abundance of ABCA12 and ABCA1 in HeLa-ABCA1 cells after transient transfection with ABCA12.

(B) Cholesterol efflux to apoA-I after the overexpression of ABCA12. \* $p < 0.05$  (versus untransfected cells).

(C and D) Intracellular distribution of ABCA1 in HeLa-ABCA1 cells transfected with or without ABCA12.

(E) Intracellular distribution of ABCA12 in HeLa-ABCA1 cells transfected with ABCA12.

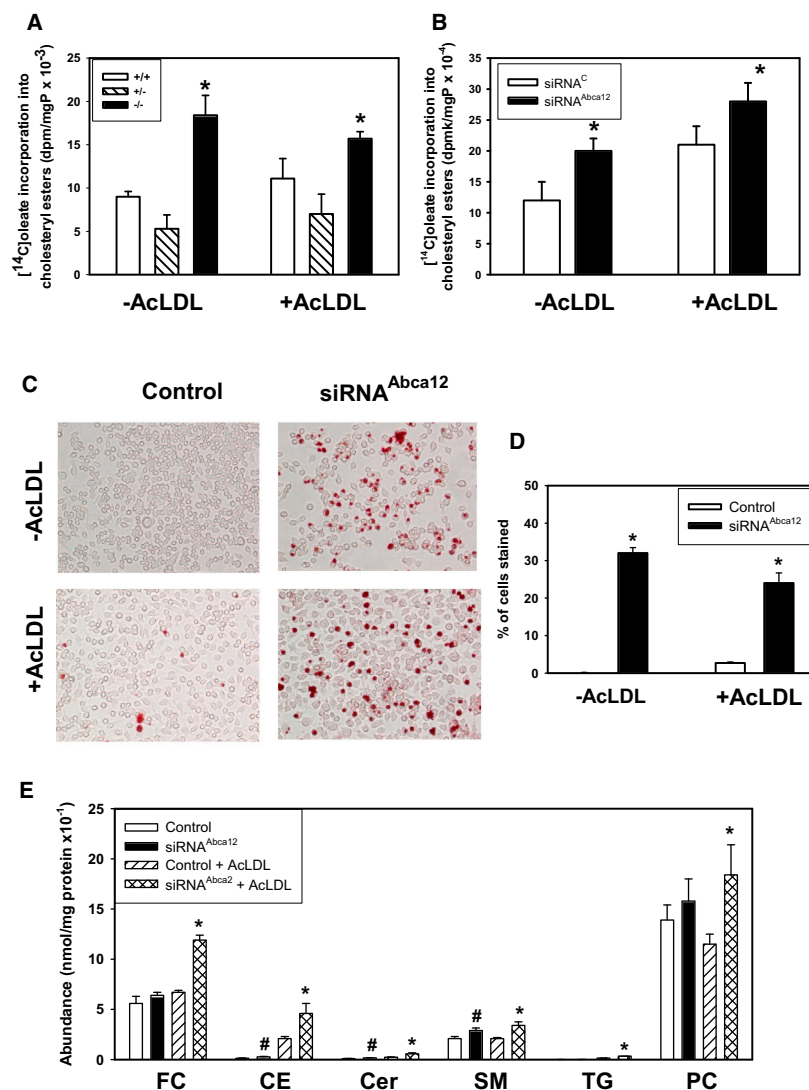
(F) A merger of (D) and (E).

(G) The time course of degradation of biotinylated surface Abca1 in HeLa/ABCA1 cells with or without the overexpression of ABCA12.

(H) Quantitation of the bands in three experiments similar to that shown in (G). Means  $\pm$  SEM are shown. \* $p < 0.01$  (versus time zero).

(legend continued on next page)





**Figure 5. Lipid Accumulation in Abca12-Deficient Cells**

(A) Cholesteryl ester biosynthesis in liver macrophages isolated from the *Abca12*<sup>E112/E112</sup> strain with or without loading the cells with acLDL. (B) Cholesteryl ester biosynthesis in RAW 264.7 cells after knockdown of Abca12 with siRNA with or without loading the cells with acLDL. (C) Oil Red O staining of RAW 264.7 cells after knockdown of Abca12 with siRNA. (D) Quantitation of Oil Red O staining of RAW 264.7 after knockdown of Abca12 with siRNA (percentage of stained cells, n = 5). Pooled data are represented as the mean ± SEM. \*p < 0.01 (versus WT or control cells). (E) Abundance of the individual lipid species in control versus Abca12 knockdown RAW 264.7 cells after activation with LXR agonist TO-901317 with or without loading with acLDL. \*p < 0.001 (versus acLDL-loaded control cells); #p < 0.01 versus non-acLDL-loaded control cells.

See also Table S2.

ABCA12 may not be limited to skin and may be an element of a pathway regulating lipid metabolism in various tissues, including macrophages, cells central to the development of atherosclerosis.

The main finding of this study is that ABCA12 is essential for the regulation of cellular cholesterol homeostasis. A loss of Abca12 renders macrophages unresponsive to stimulation by LXR agonists, reducing the abundance of cholesterol transporters and the rate of cholesterol efflux. These changes cause the accumulation of cholesterol in macrophages, their transformation into foam cells, and the development of atherosclerosis. In addition, Abca12 deficiency caused a reduction in the abundance of LXRβ; reconstitution experiments employing

heterologous Lxrβ suggest that LXRβ plays a key role in mediating the effects of a loss of ABCA12. LXRs are nuclear receptors which control the transcription of genes (Calkin and Tontonoz, 2010; Zelcer and Tontonoz, 2006); however, our findings support a posttranscriptional model of LXRβ action. It has been proposed that LXRβ binds to ABCA1, preventing ATP hydrolysis, and the interaction of ABCA1 with apoA-I; LXR agonists interfere with this interaction (Hozoji-Inada et al., 2011). Our studies have established an interaction between ABCA12, ABCA1, and

ceramides, and its mutation leads to defects in skin barrier function, leading to dehydration (Kelsell et al., 2005). Recently, our group (Smyth et al., 2008) and others (Yanagi et al., 2008; Zuo et al., 2008) have described mouse models of HI caused by mutations in *Abca12*. These mice are not viable and die shortly after birth. Using embryonic fibroblasts, we have demonstrated that *Abca12*<sup>-/-</sup> cells have severe defects in cholesterol efflux and exceptional susceptibility to lipid overloading (Smyth et al., 2008). These findings suggested that the physiological role of

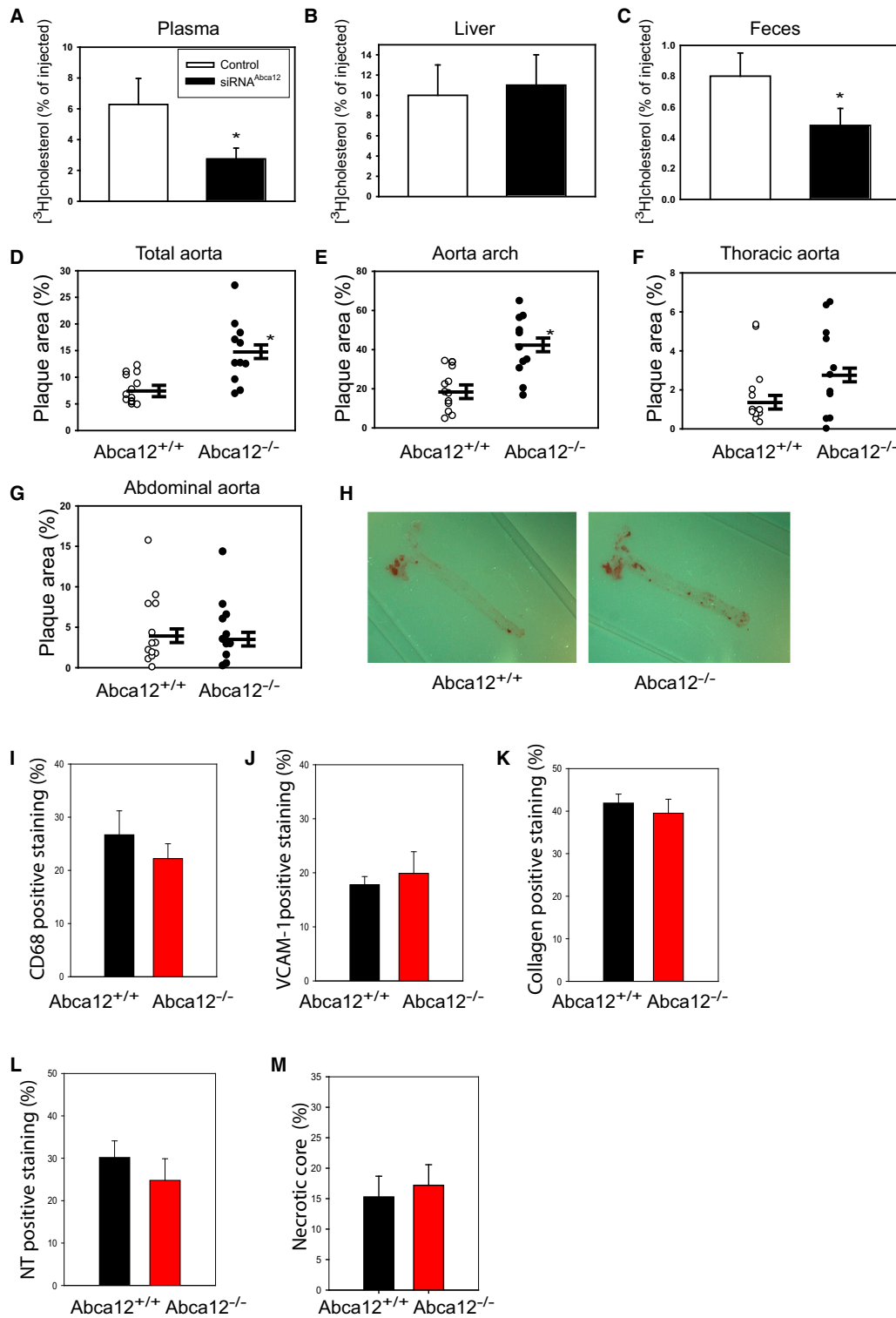
(I) Western blot after the immunoprecipitation of ABCA1 with an antibody against ABCA1 and developing with an antibody against ABCA12.

(J) Western blot after the immunoprecipitation of ABCA12 with an antibody against tGFP and developing with an antibody against ABCA1.

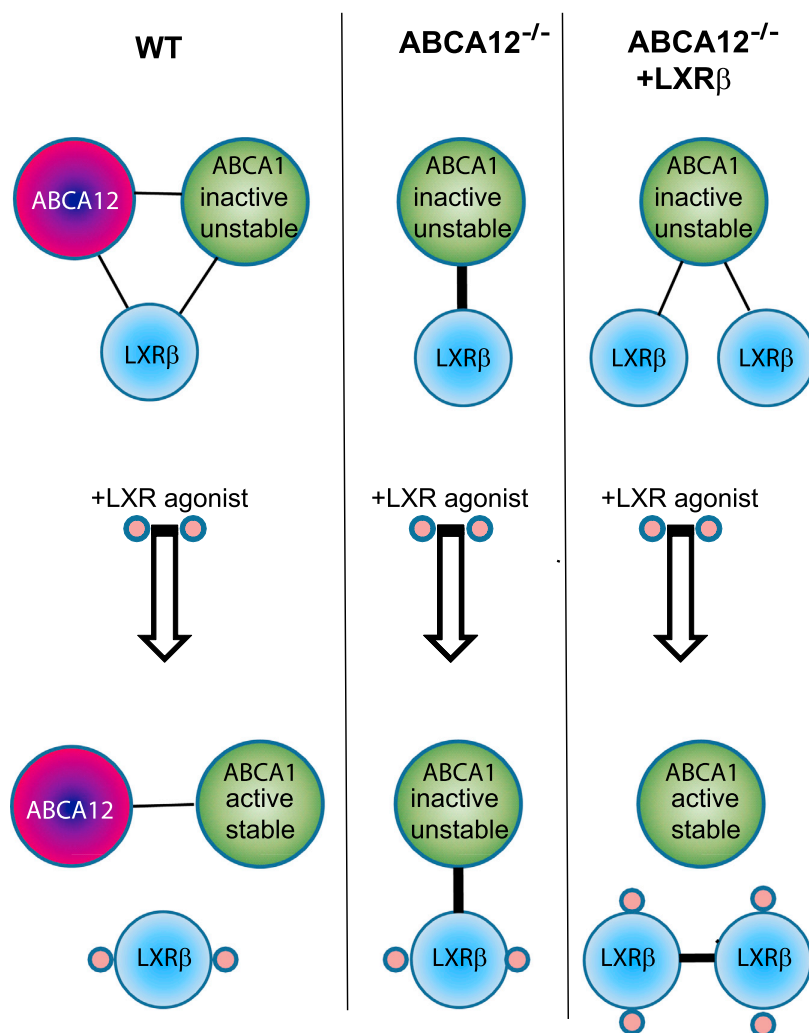
(K) Western blot after the immunoprecipitation of LXRβ and developing with antibodies against ABCA1 (top right quarter), tGFP (an ABCA12 tag; top left quarter), or LXRβ (bottom half). No primary antibody was added to the immunoprecipitation of the sample in the middle lane. HeLa-ABCA1 cells were transfected with ABCA12.

(L) PLA of the Abca1 and Abca12 pair in HeLa-ABCA1 cells transfected with or without ABCA12.

(M) PLA of the Abca1 and Lxrβ pair in HeLa-ABCA1 cells transfected with or without ABCA12. Means ± SEM of PLA signal per cell are shown in the right panels. \*p < 0.01.



(legend on next page)



**Figure 7. A Proposed Mechanism for Regulating ABCA1 via ABCA12 and LXRβ**

“Inactive” ABCA1 exists as a complex with ABCA12 and LXRβ. The addition of an LXR agonist leads to the dissociation of LXR from ABCA1-ABCA12 and the activation of ABCA1, enabling it to hydrolyse ATP, bind apoA-I, and remain on the cell surface, where it is stable. In the absence of ABCA12, LXRβ binds more firmly to ABCA1, and this association is not broken by LXR agonist; consequently, ABCA1 remains inactive and unstable. Overexpression of LXRβ leads to the excess of LXRβ and formation of LXRβ dimers in the presence of an agonist, and these dimers dissociate from ABCA1.

and the complex may be mislocalized; as a result, the LXR agonist is unable to induce dissociation. Consequently, ABCA1 remains dysfunctional and unable to bind apoA-I, leading to its internalization and degradation along with LXRβ. Overexpression of LXRβ favors the formation of LXRβ dimers, which bind more weakly to ABCA1, allowing agonists to disrupt the complex. Indirect evidence in support of this hypothesis include: (1) the binding of ABCA1 to ABCA12 and LXRβ, (2) partial colocalization of ABCA12 and ABCA1, (3) increased colocalization and interaction between LXRβ and ABCA1 after the silencing of ABCA12, (4) the ability of heterologous ABCA12 to cause a relocalization of ABCA1 and LXRβ similar to that observed by Hozoji-Inada et al. (2011), and (5) the complementary effects of ABCA12 deficiency and ABCA12 overexpression on ABCA1 stability and localization. On

LXRβ that is required for ABCA1 localization to the cell surface and is necessary for its normal activity and stability. We propose (Figure 7) that, in the presence of LXR agonist, the ABCA1-ABCA12-LXRβ complex dissociates, leading to a stimulation of cholesterol efflux, as suggested by Hozoji-Inada et al. (2011). In the absence of ABCA12, LXRβ binds more firmly to ABCA1,

the other hand, when cells were activated with cAMP, ABCA12 deficiency did not affect ABCA1 abundance but still partially reduced cholesterol efflux in the absence of LXR agonist. Furthermore, ABCA12 deficiency only partially reduced ABCA1 abundance but completely blocked LXR-dependent elevation of cholesterol efflux. Altogether, this suggests that there may

**Figure 6. The Effect of Abca12 Knockout on Reverse Cholesterol Transport and Development of Atherosclerosis in Vivo**

(A–C) The appearance of [ $^3$ H]cholesterol in plasma (A), liver (B), and feces (C) 24 hr after transplanting cholesterol-loaded and -labeled RAW 264.7 macrophages transfected with either siRNA<sup>C</sup> (n = 6) or siRNA<sup>Abca12</sup> (n = 5).

(D–G) Two groups of ApoE knockout mice were transplanted with either *Abca12*<sup>E112/E112</sup> (n = 10) or WT (n = 13) bone marrow. Animals were kept on a high-fat diet for 16 weeks. Lipid accumulation and distribution and the size (area) of atherosclerotic lesions in aorta were quantitatively assessed with en face analysis after staining with Sudan IV. Data are represented as individual values, and the horizontal dash represents the mean  $\pm$  SEM. \*p < 0.002.

(H) Representative photographs of aortas after staining with Sudan IV.

(I) Analysis of macrophage infiltration within the aortic sinus region after staining with anti-CD68.

(J) Analysis of the abundance of VCAM-1 in sections from the aortic sinus region.

(K) Analysis of the abundance of collagen in sections from the aortic sinus region.

(L) Analysis of the abundance of nitrotyrosine (NT) in sections from the aortic sinus region.

(M) Analysis of the size of the necrotic core in sections from the aortic sinus region.

(I–M) The percentage of staining in the plaque is shown. The mean  $\pm$  SEM of n = 6 in each group are shown.

See also Table S3.

be additional mechanisms underpinning the effects of ABCA12 deficiency on ABCA1 functionality. The only published mechanism that might explain the effects of cAMP on ABCA1 functionality is via the phosphorylation of ABCA1 (Haidar et al., 2002), which may be affected by the disruption of the ABCA1-ABCA12 complex. Although the lack of direct evidence supporting this hypothesis means that we cannot definitively exclude an unrelated mechanism, such as LXR-dependent ubiquitination (Zelcer et al., 2009), we consider this the most likely mode of action of ABCA12. It is important to recognize that the proposed mechanism of the effect of ABCA12 deficiency on ABCA1 cannot be a priori extended to ABCG1. Moreover, the effects of ABCA12 deficiency on ABCG1 may have contributed to the reduced reverse cholesterol transport and the development of atherosclerosis in animal experiments.

Impairment of cholesterol homeostasis plays a central role in cardiovascular, neurodegenerative, skin, and many other disorders. The current dogma suggests that the main players in regulating cholesterol homeostasis are the sterol regulatory element-binding protein (SREBP)-dependent pathways (which regulate the delivery of cholesterol to cells) and the LXR-dependent pathways (which regulate the removal of cholesterol from cells). In this study, we have identified ABCA12 as an important player in the “removal” arm of cholesterol homeostasis. We show that ABCA12 plays a key role in the posttranscriptional regulation of cholesterol transporter ABCA1 in vitro and demonstrate a pivotal role for ABCA12 deficiency in the development of atherosclerosis. Our findings also suggest that ABCA12 defines a part of a pathway involved in regulation of cholesterol homeostasis and a potential contributor to the development of atherosclerosis.

## EXPERIMENTAL PROCEDURES

### Cells

The *Abca12<sup>E112</sup>* and *Abca12<sup>Lx12</sup>* mouse strains were described previously (Smyth et al., 2008; Yanagi et al., 2010). *Abca12<sup>E112</sup>* mice were backcrossed for ten generations, and the mutation was mapped to a narrow 4.7 Mb interval. In this interval, only the *Abca12* gene was found to be altered. The uterus was removed on embryonic day (E) 12.5 of pregnancy, and embryos were isolated under aseptic condition. The embryonic liver was dissected and cells were suspended and cultured in a six-well tissue culture plate in Roswell Park Memorial Institute (RPMI)-1640 medium containing 10% fetal bovine serum and 20% L929-cell-conditioned medium as a source of colony-stimulating factors as described by Feng et al. (2005). After 24 hr depletion of adherent cells and debris, nonadherent macrophages were moved to fresh Petri dishes and grown to confluence. Embryos were genotyped, and cultured macrophages were characterized morphologically as described by Suzuki et al. (2000); 95% of cells were deemed to be macrophages.

RAW 264.7 cells were maintained as described previously (Mukhamedova et al., 2008). HeLa-ABCA1 cells stably expressing GFP-tagged ABCA1 were a kind gift of A. Remaley and were described previously (Mukhamedova et al., 2007).

### Transfections

RAW 264.7 mouse macrophages were transfected with *Abca12*-specific siRNA or scrambled siRNA (control) and treated with TO-901317 (4  $\mu$ M) for 18 hr. Transfection was performed with Lipofectamine RNAiMAX (Invitrogen) following the manufacturer's protocol. *Abca12* siRNA and scrambled siRNA were from Santa Cruz Biotechnology.

Lxr $\beta$  and murine *Abca12* genes were inserted into pCMV6-AC-GFP plasmid (Origene). Human ABCA12 plasmid was a kind gift of D. Kelsell. Transfections with Lxr $\beta$  plasmid were performed as described previously (Escher

et al., 2005). Transfection of HeLa cells with ABCA12 plasmid was conducted as described previously (Mukhamedova et al., 2007).

### HDL and Apolipoprotein A-I

HDL (1.083 < d < 1.21 g/L) was isolated by sequential centrifugation in KBr solutions and delipidated, and apoA-I was purified by gel filtration chromatography as described previously (Brace et al., 2010).

### Cholesterol Efflux

Cholesterol efflux was measured as described previously (Mukhamedova et al., 2008). In brief, cells were incubated in serum-containing medium supplemented with <sup>3</sup>H-labeled cholesterol (75 kBq/ml) for 48 hr. Then, cells were washed with PBS and incubated for 18 hr in serum-free medium in the presence or absence of LXR agonist TO-901317 (final concentration, 4  $\mu$ M). Then, ApoA-I or isolated HDL were added to the final concentration of 30  $\mu$ g/ml, and cells were incubated for 2 hr at 37°C. The efflux was calculated as a proportion of radioactivity moved from cells to medium (minus efflux to a medium without acceptors). ABC-dependent efflux was defined as the difference in the efflux to apoA-I or HDL from cells activated with or without LXR agonist.

### Western Blot Analysis

The antibodies against ABCA1, ABCG1, SR-B1, LXR $\alpha$ , LXR $\beta$ , and GAPDH were from Abcam and Aviva Systems Biology. Semiquantitative analysis of western blots was performed by densitometry and presented as a proportion of control after normalization to loading control.

For *Abca12* analysis, cellular proteins were biotinylated with 2 mM EZ-Link NHS-LC-Biotin (Thermo Scientific) at room temperature for 1 hr. Then, cells were lysed with NP-40 lysis buffer, and 200  $\mu$ g of each the lysates was incubated with 2.5  $\mu$ g anti-goat ABCA12 antibody (Novus). A complex of *Abca12* protein and an antibody against *Abca12* was immunoprecipitated with Sepharose-protein G beads. The protein equivalent of 100  $\mu$ g of original protein lysate was run on the SDS-PAGE, and western blot was performed and detected with horseradish-peroxidase-conjugated streptavidin.

### Immunoprecipitation

Equal amounts (150  $\mu$ g) of total cell lysates were used to immunoprecipitate ABCA1, LXR $\beta$ , and ABCA12-turboGFP proteins. In brief, radio-immunoprecipitation assay (RIPA) lysates were incubated with 2  $\mu$ g of rabbit polyclonal ABCA1 (Novus), turboGFP (Evrogen), or LXR $\beta$  (Abcam) antibody overnight at 4°C. Protein A-Sepharose beads (Thermo Scientific) were added to the antibody and protein mixture. After incubation for 3 hr with rotating at 4°C, beads were washed with PBS.

### Pulse-Chase Experiments

RAW 264.7 cells were transfected with *Abca12*-siRNA or scrambled siRNA as described above. After 48 hr posttransfection, the cells were activated with 4  $\mu$ M TO-901317 for 3 hr in culture medium and then incubated in methionine- and cysteine-free RPMI (Sigma-Aldrich) with TO-901317 for 1 hr. The cells were pulsed with 100  $\mu$ Ci/ml of <sup>35</sup>S-labeled methionine and cysteine (PerkinElmer, EasyTag EXPRESS Protein Labeling Mix) for 2 hr, washed in PBS, and chased in RPMI medium supplemented with 1.5 mg/ml L-methionine and 0.5 mg/ml L-cysteine (Sigma-Aldrich) for the indicated times. ABCA1 protein was determined by immunoprecipitation with an affinity-purified anti-rabbit ABCA1 antibody, separated by SDS-PAGE, transferred to polyvinylidene difluoride membrane, and exposed to autoradiography film.

### Biotinylation of Cell-Surface Proteins

Cells were washed three times with ice cold PBS and biotinylated with Sulfo-NHS-SS-Biotin as described previously (Cui et al., 2012). After biotinylation, cells were washed twice with 50 mM Tris and 150 mM NaCl to quench the reaction and once with PBS. After incubation in serum-free RPMI medium at 37°C for 0, 4, or 24 hr, cells were lysed with RIPA buffer. Biotinylated proteins were purified with UltraLink Plus immobilized streptavidin gel (Thermo Scientific).

### Cholesteryl Ester Biosynthesis

Cells were incubated for 18 hr in serum-containing medium in the presence or absence of 50  $\mu$ g/ml of acetylated LDL. Then, cholesteryl ester biosynthesis



was assessed by incorporating [ $^{14}\text{C}$ ] oleic acid into cholesteryl esters over 2 hr as described previously (Mujawar et al., 2006). The staining of cells with Oil Red O was performed as described previously (Smyth et al., 2008).

#### Real-Time RT-PCR

PCR primers were obtained from TaqMan Gene Expression Assays with the following gene assay IDs: Mm00442646\_m1, Mm00550501\_m1, Mm00613683\_m1, Mm00437390\_m1, Mm00443451\_m1, Mm00437262\_m1, Mm00450234\_m1, Mm00446214\_m1, Mm00440169\_m1, and Mm01282499\_m1. Real-time PCR reactions were performed on a 7500 Fast System from ABI Applied Biosystems. The relative amount of mRNA was calculated with the comparative  $C_T$  method. Gene expression was normalized to 18 s rRNA.

#### Lipidomics

RAW264.7 cells were transfected with siRNA<sup>Abca12</sup> or siRNA<sup>c</sup>, activated with 4  $\mu\text{M}$  TO-901317 for 24 hr, and incubated with or without 50  $\mu\text{g}/\text{ml}$  acetylated LDL overnight. Cells were washed with PBS, resuspended into 20 mM Tris and 500 mM NaCl solution containing 0.1 mM butylated hydroxytoluene, and sonicated. Ten microliters (20  $\mu\text{g}$  of protein) of each the supernatant were subjected to lipids extraction and analysis via mass spectrometry.

#### Confocal Microscopy

For imaging, RAW 264.7 or HeLa-ABCA1 cells were grown in a Lab-Tek II Chamber Slide System (Thermo Scientific). In HeLa-ABCA1 cells, ABCA1 was fused with GFP; in RAW 264.7 cells, it was stained with monoclonal antibody against ABCA1 (Abcam) and AlexaFluor 488 anti-mouse IgG secondary antibody (Invitrogen). Lxr $\beta$  was stained with polyclonal primary antibody (Abcam) and AlexaFluor 633 anti-rabbit secondary antibody (Invitrogen). Abca12 staining in RAW 264.7 cells was performed with in-house antibodies against Abca12 and AlexaFluor 488 anti-mouse IgG secondary antibody and in HeLa-ABCA1 cells with an antibody against ABCA12 (Novus) and Texas Red 633 anti-rabbit secondary antibody. Colocalization was quantitated with Zeiss LSM Image Browser as described previously (Zinchuk and Grossenbacher-Zinchuk, 2011; Zinchuk et al., 2007) and expressed as an unweighted colocalization coefficient of Lxr $\beta$  with Abca1 on full images (ten images, six to eight cells per image, and three experiments).

#### In Situ Proximity Ligation Assay

The general principle of the Duolink in situ coimmunoprecipitation assay has been described previously (Baima et al., 2010). Cells were seeded at  $1 \times 10^4$  cells per well on eight-well Lab-Tek II Chamber Slides and incubated to 80% confluence. After 48 hr posttransfection, cells were washed with PBS, fixed in 4% paraformaldehyde for 15 min, and permeabilized in 0.01% Triton X-100 for 10 min. Nonspecific binding was blocked by incubating cells for 30 min at 37°C with a blocking solution provided by the manufacturer. The cells were incubated overnight at 4°C with two antibodies raised in different species. The proximity ligation assay was conducted with a Duolink Orange Detection kit. Images were quantified by counting total number of PLA signals in each image and dividing by the number of nuclei in the field of view. Counting was conducted with ImageJ in automatic mode. A total of several hundred individual cells were counted for each group of samples.

#### In Vivo Reverse Cholesterol Transport Assay

Cholesterol efflux in vivo was measured with the use of a model described by Wang et al. (2007) with previously described modifications (Mukhamedova et al., 2008).

#### Bone Marrow Transplantation

After backcrossing ten generations to C57BL/6, the *Abca12*<sup>E112</sup> mutation was backcrossed onto a C57BL/6 ApoE-deficient background. Then, C57BL/6 CD45.2 *Abca12*<sup>-/-</sup> ApoE<sup>-/-</sup> mice were intercrossed, and E13.5 livers were harvested from embryos. Two million fetal liver cells were transplanted into adult C57BL/6 CD45.1 ApoE<sup>-/-</sup> recipients in whom endogenous hematopoiesis had been ablated by irradiation (11 Gy in 2 equal doses 2–3 hr apart). After 8 weeks posttransplantation, donor engraftment was determined by flow cytometric analysis of peripheral blood leukocyte in conjunction with CD45-specific monoclonal antibody. Then, 8 weeks postreconstitution, bone marrow

chimera mice were put on a high-fat diet (21% fat, 0.15% cholesterol) for 16 weeks. Lipid accumulation and distribution and size (area) of atherosclerotic lesions in aorta were quantitatively assessed with an en face analysis after being stained with Sudan IV.

Heart tissue containing the aortic sinus was embedded in optimum cutting temperature compound (Tissue-Tek) and frozen in order to assess plaques within the aortic sinus region. Consecutive 10  $\mu\text{m}$  sections spanning 240  $\mu\text{m}$  of the aortic sinus were cut and collected. Sections were stained with Masson's trichrome stain in order to detect collagen. Sections were also examined for macrophage content as well as inflammatory and oxidation markers by immunohistochemistry with the following primary antibodies: anti-CD68 (AbD, Serotec), anti-VCAM-1 (BD Pharmingen), and anti-Nitrotyrosine (Millipore). Images were quantified with Image-Pro Plus 6.0.

All animal experiments complied with the regulatory standards of and were approved by the Walter and Eliza Hall Institute Animal Ethics Committee and/or the Alfred Medical Research and Education Precinct Animal Ethics Committee.

#### Statistics

All in vitro experiments were performed in quadruplicates and repeated two to three times. Statistical analysis was performed with Student's *t* tests and one-way ANOVA modified by a step-down Bonferroni procedure, and both tests produced identical results. Means  $\pm$  SEM are shown.

#### SUPPLEMENTAL INFORMATION

Supplemental Information contains two figures, three tables, and two movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cmet.2013.07.003>.

#### ACKNOWLEDGMENTS

This study was supported by project grants (GNT545906 and GNT1020281), program grants (GNT1016647 and GNT1036352), fellowships (to D.S. and B.T.K.), and an Independent Medical Research Institutes Infrastructure Support Scheme grant (361646) from the Australian National Health and Medical Research Council and a Victorian State Government Operational Infrastructure Support grant. B.T.K. is supported by the Sylvia and Charles Viertel Foundation. I.M.S. is supported by a Monash Fellowship and a Future Fellowship from the Australian Research Council. Photomicrographs were collected with equipment provided by the Monash Micro Imaging facility.

Received: May 26, 2011

Revised: March 5, 2013

Accepted: July 2, 2013

Published: August 6, 2013

#### REFERENCES

- Baima, E.T., Guzova, J.A., Mathialagan, S., Nagiec, E.E., Hardy, M.M., Song, L.R., Bonar, S.L., Weinberg, R.A., Selness, S.R., Woodard, S.S., et al. (2010). Novel insights into the cellular mechanisms of the anti-inflammatory effects of NF- $\kappa$ B essential modulator binding domain peptides. *J. Biol. Chem.* 285, 13498–13506.
- Brace, R.J., Sorrenson, B., Sviridov, D., and McCormick, S.P.A. (2010). A gel-based method for purification of apolipoprotein A-I from small volumes of plasma. *J. Lipid Res.* 51, 3370–3376.
- Brown, M.S., and Goldstein, J.L. (2009). Cholesterol feedback: from Schoenheimer's bottle to Scap's MELADL. *J. Lipid Res. Suppl.* 50, S15–S27.
- Calkin, A.C., and Tontonoz, P. (2010). Liver x receptor signaling pathways and atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* 30, 1513–1518.
- Cui, H.L., Grant, A., Mukhamedova, N., Pushkarsky, T., Jennelle, L., Dubrovsky, L., Gaus, K., Fitzgerald, M.L., Sviridov, D., and Bukrinsky, M. (2012). HIV-1 Nef mobilizes lipid rafts in macrophages through a pathway that competes with ABCA1-dependent cholesterol efflux. *J. Lipid Res.* 53, 696–708.

- Escher, G., Hoang, A., Georges, S., Tchoua, U., El-Osta, A., Krozowski, Z., and Sviridov, D. (2005). Demethylation using the epigenetic modifier, 5-azacytidine, increases the efficiency of transient transfection of macrophages. *J. Lipid Res.* 46, 356–365.
- Feng, R.Q., Du, L.Y., and Guo, Z.Q. (2005). In vitro cultivation and differentiation of fetal liver stem cells from mice. *Cell Res.* 15, 401–405.
- Fitzgerald, M.L., Mujawar, Z., and Tamehiro, N. (2010). ABC transporters, atherosclerosis and inflammation. *Atherosclerosis* 211, 361–370.
- Haidar, B., Denis, M., Krimbou, L., Marcil, M., and Genest, J., Jr. (2002). cAMP induces ABCA1 phosphorylation activity and promotes cholesterol efflux from fibroblasts. *J. Lipid Res.* 43, 2087–2094.
- Hozoji-Inada, M., Munehira, Y., Nagao, K., Kioka, N., and Ueda, K. (2011). Liver X receptor beta (LXRbeta) interacts directly with ATP-binding cassette A1 (ABCA1) to promote high density lipoprotein formation during acute cholesterol accumulation. *J. Biol. Chem.* 286, 20117–20124.
- Joseph, S.B., McKilligan, E., Pei, L., Watson, M.A., Collins, A.R., Laffitte, B.A., Chen, M., Noh, G., Goodman, J., Hagger, G.N., et al. (2002). Synthetic LXR ligand inhibits the development of atherosclerosis in mice. *Proc. Natl. Acad. Sci. USA* 99, 7604–7609.
- Kelsell, D.P., Norgett, E.E., Unsworth, H., Teh, M.T., Cullup, T., Mein, C.A., Dopping-Hepenstal, P.J., Dale, B.A., Tadini, G., Fleckman, P., et al. (2005). Mutations in ABCA12 underlie the severe congenital skin disease harlequin ichthyosis. *Am. J. Hum. Genet.* 76, 794–803.
- Le Goff, W., Zheng, P., Brubaker, G., and Smith, J.D. (2006). Identification of the cAMP-responsive enhancer of the murine ABCA1 gene: requirement for CREB1 and STAT3/4 elements. *Arterioscler. Thromb. Vasc. Biol.* 26, 527–533.
- Maxfield, F.R., and Tabas, I. (2005). Role of cholesterol and lipid organization in disease. *Nature* 438, 612–621.
- Mujawar, Z., Rose, H., Morrow, M.P., Pushkarsky, T., Dubrovsky, L., Mukhamedova, N., Fu, Y., Dart, A., Orenstein, J.M., Bobryshev, Y.V., et al. (2006). Human immunodeficiency virus impairs reverse cholesterol transport from macrophages. *PLoS Biol.* 4, e365.
- Mukhamedova, N., Fu, Y., Bukrinsky, M., Remaley, A.T., and Sviridov, D. (2007). The role of different regions of ATP-binding cassette transporter A1 in cholesterol efflux. *Biochemistry* 46, 9388–9398.
- Mukhamedova, N., Escher, G., D'Souza, W., Tchoua, U., Grant, A., Krozowski, Z., Bukrinsky, M., and Sviridov, D. (2008). Enhancing apolipoprotein A-I-dependent cholesterol efflux elevates cholesterol export from macrophages in vivo. *J. Lipid Res.* 49, 2312–2322.
- Neufeld, E.B., Remaley, A.T., Demosky, S.J., Stonik, J.A., Cooney, A.M., Comly, M., Dwyer, N.K., Zhang, M., Blanchette-Mackie, J., Santamarina-Fojo, S., and Brewer, H.B., Jr. (2001). Cellular localization and trafficking of the human ABCA1 transporter. *J. Biol. Chem.* 276, 27584–27590.
- Ogata, M., Tsujita, M., Hossain, M.A., Akita, N., Gonzalez, F.J., Staels, B., Suzuki, S., Fukutomi, T., Kimura, G., and Yokoyama, S. (2009). On the mechanism for PPAR agonists to enhance ABCA1 gene expression. *Atherosclerosis* 205, 413–419.
- Smyth, I., Hacking, D.F., Hilton, A.A., Mukhamedova, N., Meikle, P.J., Ellis, S., Satterley, K., Collinge, J.E., de Graaf, C.A., Bahlo, M., et al. (2008). A mouse model of harlequin ichthyosis delineates a key role for Abca12 in lipid homeostasis. *PLoS Genet.* 4, e1000192.
- Suzuki, A., Zheng, Y., Kondo, R., Kusakabe, M., Takada, Y., Fukao, K., Nakauchi, H., and Taniguchi, H. (2000). Flow-cytometric separation and enrichment of hepatic progenitor cells in the developing mouse liver. *Hepatology* 32, 1230–1239.
- Wang, X., Collins, H.L., Ranalletta, M., Fuki, I.V., Billheimer, J.T., Rothblat, G.H., Tall, A.R., and Rader, D.J. (2007). Macrophage ABCA1 and ABCG1, but not SR-BI, promote macrophage reverse cholesterol transport in vivo. *J. Clin. Invest.* 117, 2216–2224.
- Yanagi, T., Akiyama, M., Nishihara, H., Sakai, K., Nishie, W., Tanaka, S., and Shimizu, H. (2008). Harlequin ichthyosis model mouse reveals alveolar collapse and severe fetal skin barrier defects. *Hum. Mol. Genet.* 17, 3075–3083.
- Yanagi, T., Akiyama, M., Nishihara, H., Ishikawa, J., Sakai, K., Miyamura, Y., Naoe, A., Kitahara, T., Tanaka, S., and Shimizu, H. (2010). Self-improvement of keratinocyte differentiation defects during skin maturation in ABCA12-deficient harlequin ichthyosis model mice. *Am. J. Pathol.* 177, 106–118.
- Yvan-Charvet, L., Wang, N., and Tall, A.R. (2010). Role of HDL, ABCA1, and ABCG1 transporters in cholesterol efflux and immune responses. *Arterioscler. Thromb. Vasc. Biol.* 30, 139–143.
- Zelcer, N., and Tontonoz, P. (2006). Liver X receptors as integrators of metabolic and inflammatory signaling. *J. Clin. Invest.* 116, 607–614.
- Zelcer, N., Hong, C., Boyadjian, R., and Tontonoz, P. (2009). LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. *Science* 325, 100–104.
- Zinchuk, V., and Grossenbacher-Zinchuk, O. (2011). Quantitative Colocalization Analysis of Confocal Fluorescence Microscopy Images. In *Current Protocols in Cell Biology*/Current Protocols in Cell Biology (New York: John Wiley & Sons), pp. 4.19.11–14.19.16.
- Zinchuk, V., Zinchuk, O., and Okada, T. (2007). Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: pushing pixels to explore biological phenomena. *Acta Histochem. Cytochem.* 40, 101–111.
- Zuo, Y., Zhuang, D.Z., Han, R., Isaac, G., Tobin, J.J., McKee, M., Welti, R., Brissette, J.L., Fitzgerald, M.L., and Freeman, M.W. (2008). ABCA12 maintains the epidermal lipid permeability barrier by facilitating formation of ceramide linoleic esters. *J. Biol. Chem.* 283, 36624–36635.



# Genome-wide Generation and Systematic Phenotyping of Knockout Mice Reveals New Roles for Many Genes

Jacqueline K. White,<sup>1</sup> Anna-Karin Gerdin,<sup>1</sup> Natasha A. Karp,<sup>1</sup> Ed Ryder,<sup>1</sup> Marija Buljan,<sup>1</sup> James N. Bussell,<sup>1</sup> Jennifer Salisbury,<sup>1</sup> Simon Clare,<sup>1</sup> Neil J. Ingham,<sup>1</sup> Christine Podrini,<sup>1</sup> Richard Houghton,<sup>1</sup> Jeanne Estabel,<sup>1</sup> Joanna R. Bottomley,<sup>1</sup> David G. Melvin,<sup>1</sup> David Sunter,<sup>1</sup> Niels C. Adams,<sup>1</sup> The Sanger Institute Mouse Genetics Project,<sup>1,2,3,5,6,8</sup> David Tannahill,<sup>1</sup> Darren W. Logan,<sup>1</sup> Daniel G. MacArthur,<sup>1</sup> Jonathan Flint,<sup>2</sup> Vinit B. Mahajan,<sup>3</sup> Stephen H. Tsang,<sup>4</sup> Ian Smyth,<sup>5</sup> Fiona M. Watt,<sup>6</sup> William C. Skarnes,<sup>1</sup> Gordon Dougan,<sup>1</sup> David J. Adams,<sup>1</sup> Ramiro Ramirez-Solis,<sup>1</sup> Allan Bradley,<sup>1</sup> and Karen P. Steel<sup>1,7,\*</sup>

<sup>1</sup>Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK

<sup>2</sup>Wellcome Trust Centre for Human Genetics, Oxford OX3 7BN, UK

<sup>3</sup>Omics Laboratory, University of Iowa, Iowa City, IA 52242, USA

<sup>4</sup>Harkness Eye Institute, Columbia University, New York, NY 10032, USA

<sup>5</sup>Monash University, Melbourne, Victoria 3800, Australia

<sup>6</sup>Wellcome Trust Centre for Stem Cell Research, University of Cambridge, Cambridge CB2 1QR, UK

<sup>7</sup>Wolfson Centre for Age-Related Diseases, King's College London, Guy's Campus, London SE1 1UL, UK

<sup>8</sup>A full list of The Sanger Institute Mouse Genetics Project contributors may be found in the [Supplemental Information](#)

\*Correspondence:

<http://dx.doi.org/10.1016/j.cell.2013.06.022>

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

## SUMMARY

Mutations in whole organisms are powerful ways of interrogating gene function in a realistic context. We describe a program, the Sanger Institute Mouse Genetics Project, that provides a step toward the aim of knocking out all genes and screening each line for a broad range of traits. We found that hitherto unpublished genes were as likely to reveal phenotypes as known genes, suggesting that novel genes represent a rich resource for investigating the molecular basis of disease. We found many unexpected phenotypes detected only because we screened for them, emphasizing the value of screening all mutants for a wide range of traits. Haploinsufficiency and pleiotropy were both surprisingly common. Forty-two percent of genes were essential for viability, and these were less likely to have a paralog and more likely to contribute to a protein complex than other genes. Phenotypic data and more than 900 mutants are openly available for further analysis.

## INTRODUCTION

The availability of well-annotated genome sequences for a variety of organisms has provided a strong foundation on which much biological knowledge has been assembled, including the generation of comprehensive genetic resources. This has been

achieved in several model organisms, including *E. coli*, *S. cerevisiae*, *S. pombe*, *A. thaliana*, *C. elegans*, and *D. melanogaster*, greatly facilitating studies focused on single genes and enabling genome-wide genetic screens.

Annotation of the human genome has identified over 20,000 protein-coding genes as well as many noncoding RNAs. Despite the dramatic increase in the knowledge of variation in human genomes, the normal function of many genes is still unknown or predicted from sequence analysis alone, and consequently, the disease significance of rare variants remains obscure. Furthermore, there remains a large bias toward research on a small number of the best-known genes (Edwards et al., 2011). Realizing the full value of the complete human genome sequence requires broadening this focus, and the availability of comprehensive biological resources will facilitate this process.

The mouse is a key model organism for assessing mammalian gene function, providing access to conserved processes such as development, metabolism, and physiology. Genetic studies in mice, mostly via targeted mutagenesis in ES cells, have described a function for 7,229 genes ([http://ftp.informatics.jax.org/pub/reports/MGI\\_PhenotypicAllele.rpt](http://ftp.informatics.jax.org/pub/reports/MGI_PhenotypicAllele.rpt), February 2013). The vast majority of these studies have been directed at previously studied (known) genes, driven by previous biological knowledge. Phenotype-driven screens have also identified genes associated with specific phenotypes, although to a smaller extent. Although targeted mutagenesis has been very successful, the global distribution of the effort has resulted in significant heterogeneity in allele design, genetic background of mice used, and their phenotypic analysis. Furthermore, the biological focus of most targeted knockout experiments is constrained by the expertise of the

specific research group. As a result, many phenotypes have not been detected, and consequently, the full biological function of many genes studied using knockout mice is significantly underreported.

Some efforts to generate and phenotype sizeable sets of new targeted alleles of genes of interest have been reported previously (e.g., Tang et al., 2010). These studies focused on specific categories of molecules such as secreted and transmembrane proteins or other “drugable” targets. Other research centers have established mouse clinics, with the aim of carrying out a comprehensive analysis of the phenotypes of mutant lines of specific interest (e.g., Fuchs et al., 2012; Wakana et al., 2009; Laughlin et al., 2012).

The genome-wide set of targeted mutations in ES cells established by the KOMP, EUCOMM, and MirKO programs (Skarnes et al., 2011; Prosser et al., 2011; Park et al., 2012) provides an opportunity to conduct systematic, large-scale gene function analysis in a mammalian system without the variables inherent in studies by individual groups. The Sanger Institute’s Mouse Genetics Project (MGP) was one of the first programs to pursue this objective, established in 2006 when the first targeted ES cells became available. The MGP later expanded to contribute to a European phenotyping effort, EUMODIC, and more recently has become a founding member of the International Mouse Phenotyping Consortium (IMPC). Summaries of the developing efforts and aspirations of the IMPC have been reported (e.g., Brown and Moore, 2012; Ayadi et al., 2012). As the first established large-scale project using the KOMP/EUCOMM ES cells, the MGP has provided pilot data to inform the design of the international effort, such as the advantages of a single pipeline design, optimum numbers of mice, and details of variance for specific phenotyping tests. To date, the MGP has generated more than 900 lines of mutants using KOMP/EUCOMM resources (<http://www.sanger.ac.uk/mouseportal/>), and here, we describe the analysis of 489 of these for viability and fertility and 250 lines that have passed through a systematic screen for adult phenotypes, providing a glimpse into the wealth of biological insight that will emerge from these programs. Publicly available data enable the construction of new hypotheses, and the mouse mutants provide an invaluable resource for follow-up studies.

## RESULTS

### Genes and Alleles

Mice carrying targeted knockout first conditional-ready alleles from the KOMP/EUCOMM ES cell resources (Figures S1A and S1B available online; Skarnes et al., 2011) were established on a C57BL/6 genetic background. The mutants generated are listed in Tables S1 and S2, and all are available through public repositories including EMMA (<http://www.emmanet.org/>) and KOMP (<http://www.komp.org/>). Two classes of alleles are represented: those targeted with a promoter-driven selectable marker, and those with promoterless targeting vectors. Most are expected to be null alleles based on previous experience with this design (Mitchell et al., 2001; Testa et al., 2004). Data from 25 alleles showed that most (15) had <0.5% of normal transcript level detected in liver with a minority (4) showing a “leaki-

ness” of ~20% (column X; Table S2). The structure of each allele was confirmed when established in mice (Figure S1C).

### Viability

Viability was assessed at postnatal day 14 (P14) by genotyping offspring of heterozygous crosses (Figure 1A). Data from 489 targeted alleles are summarized in Figure 2A. Overall, 58% were fully viable, whereas 29% produced no homozygotes at P14 and were classed as lethal, consistent with the proportion of homozygous embryonic/perinatal lethal mutants reported by MGI (2,183 of 7,229 lines of mice [30%]; [ftp://ftp.informatics.jax.org/pub/reports/MGI\\_PhenoGenoMP.rpt](ftp://ftp.informatics.jax.org/pub/reports/MGI_PhenoGenoMP.rpt), February 2013). A further 13% produced fewer than 13% homozygotes and were considered to be subviable. Genes required for survival included alleles generated with both promoter-driven and promoterless selection cassettes, but the latter were significantly more likely to be lethal (Figure 2B; Table S3) despite a greater level of persistent gene expression (11 of 14 promoter-driven compared with 4 of 11 promoterless alleles with <0.5% expression; column X; Table S2).

Alleles classed as lethal or subviable at P14 were further assessed at E14.5 (Figure 2C). From 143 alleles examined, 48% (68 genes) produced no homozygotes, indicating embryonic lethality and complete resorption by E14.5. One-third of alleles ( $n = 45$ ) produced the expected number of homozygous embryos, whereas 30 (21%) produced fewer than expected homozygotes. Of the 75 mutant lines that produced homozygous embryos, 34 exhibited one or more morphological defect (Figure 2D). Some mutants ( $n = 23$ ) presented with specific abnormalities including craniofacial defects and polydactyly, whereas 11 lines displayed only generalized indicators of developmental defects, edema, and/or growth retardation (Figure 2D). Examples are illustrated in Figures 2E–2G.

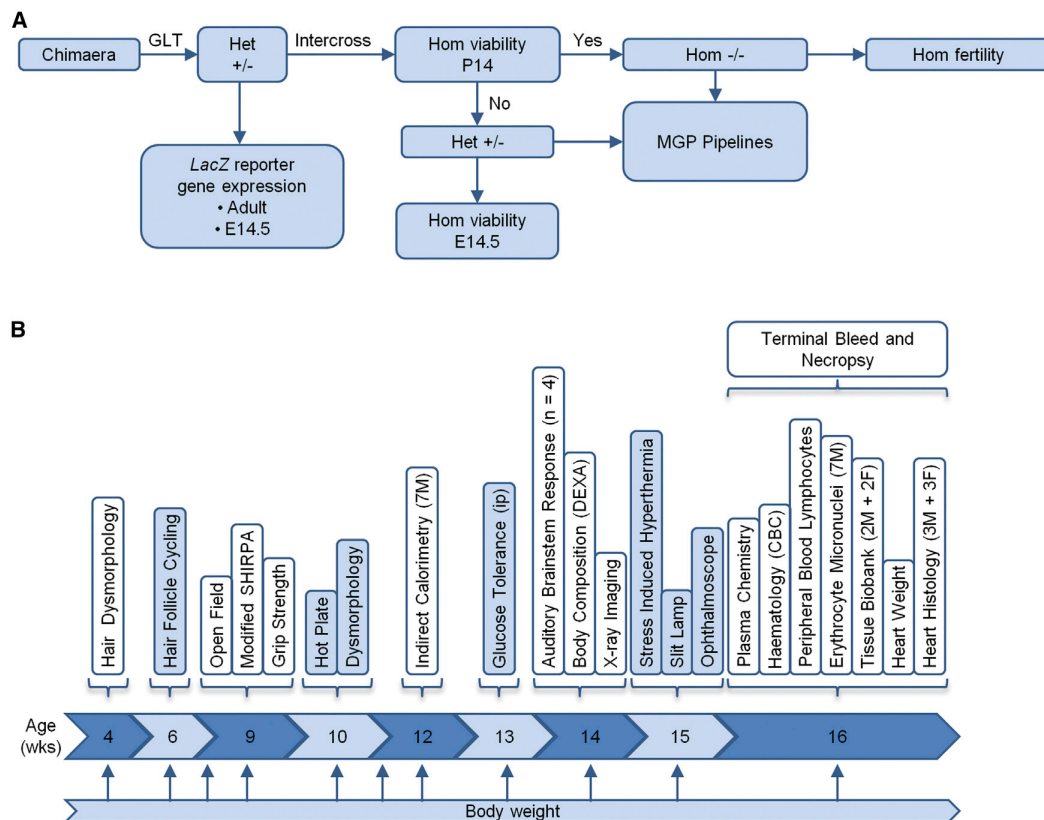
### Fertility

Fertility of heterozygotes was assessed from heterozygous intercrosses. Of 489 alleles assessed, all heterozygotes were able to produce offspring. Homozygous mutants for 307 of the viable lines were then assessed. A homozygous infertility rate of 5.2% ( $n = 16$ ) was observed (Figure 2H), strongly male biased with 15 of 16 genes exhibiting male infertility. A total of 11 genes affected only males, whereas just 1 was female specific (*Pabpc1l*). Of these 16 genes, 7 have not previously been associated with infertility. Although some were good candidates such as *Usp42*, expressed during mouse spermatogenesis (Kim et al., 2007), others are novel genes such as *3010026O09Rik* and may suggest new pathways or mechanisms influencing fertility.

### Adult Phenotypes

We report here the results of our screen of the first 250 lines to complete all primary phenotyping pipelines. In contrast to previous focused screens by Mitchell et al. (2001) and Tang et al. (2010), a broad range of gene products was included. The 250 genes reported span all chromosomes except Y (Figure S1D) and include eight control lines published previously and 87 genes proposed by the research community. For 34 of the 250 genes, no functional information has been published. A comparison of this gene set with all mouse genes indicates minimal





**Figure 1. Illustration of the Phenotyping Pipelines**

(A) An overview of the typical workflow from chimera to entry into phenotyping pipelines, encompassing homozygous (Hom) viability, fertility, and target gene expression profiling using the *lacZ* reporter. Het, heterozygous.

(B) The Sanger Institute MGP clinical phenotyping pipeline showing tests performed during each week. Seven male and seven female mutant mice are processed for each allele screened. In addition, seven male and seven female WT controls per genetic background are processed every week.

See also Figure S1 and Tables S1 and S2.

GO term enrichment spread over a variety of processes and underrepresentation only in sensory perception of smell, indicating that the gene set can be regarded as a reasonable sample of the genome.

A series of tests was used (Figure 1B), designed to detect robust variations in phenotypes that were key indicators of a broad spectrum of disease categories. Of the 250 reported lines, 104 were either lethal or subviable; most of these were screened as heterozygotes ( $n = 90$ ), and the remaining lines were screened as homozygotes and/or hemizygotes ( $n = 160$ ). All mutant lines generated passed through all primary phenotypic screens. For most tests in the pipeline, seven males and seven females were used, tested in small batches so that the data for each genotype were gathered on different days (Figure S2A). Assays culminated in the collection of samples at 16 weeks of age (Table S4). The primary screen included a high-fat diet challenge to exacerbate any latent phenotypes. Separate pipelines included challenges with two infectious agents: *Salmonella* Typhimurium and *Citrobacter rodentium* (Table S4).

Phenotypic data from the first 250 mutant alleles through the adult pipelines are summarized in Tables S1 and S2, with signif-

icant differences from the control baseline (hits) indicated by a red box. To make robust phenotypic calls, a reference range method was implemented that uses accumulated wild-type (WT) data to identify and refine the 95% reference range (Figure S2B). Mutant data were compared to the relevant reference range and variant phenotypes determined using a standardized set of rules (Figure S3). We aimed to highlight phenotypes with large effect sizes. This approach results in conservative calls and minimizes false positives. There was very little missing data (2.14% of all calls; Table S2). The maximum number of parameters collected per line was 263. Of these, 147 were categorical variables, for example normal or abnormal teeth, whereas 116, such as plasma magnesium levels, exhibited a continuous distribution from which outliers were identified. Examples of parameters with continuous variables (cholesterol, high-density lipoprotein [HDL], low-density lipoprotein [LDL], mean weights, and auditory brainstem responses [ABR]) are illustrated in Figure 3.

Gene expression was examined by whole-mount *lacZ* reporter gene expression in 41 tissues and organs of adults, typically using heterozygotes ( $\geq 6$  weeks old;  $n = 243$  lines; Table S1).

Ubiquitous expression was recorded for eight lines (3.3%) and complete absence of expression in nine lines (3.7%). Of the remaining lines, 168 (69.1%) showed expression in <20 of the tissues, suggesting a relatively specific expression pattern, whereas 58 (23.9%) were more broadly expressed ( $\geq 20$  tissues with detectable *lacZ* expression).

The data and images can be viewed on the Sanger Institute's mouse portal, accompanied by step-by-step examples of how to access the data (<http://www.sanger.ac.uk/mouseportal/>). Much of the raw data can be downloaded from the MGP Phenotyping Biomart (<http://www.sanger.ac.uk/htgt/biomart/martview/>) for further analysis. Summaries can be found by searching for each gene of interest in Wikipedia ([http://en.wikipedia.org/wiki/Category:Genes\\_mutated\\_in\\_mice](http://en.wikipedia.org/wiki/Category:Genes_mutated_in_mice)) and Mouse Genome Informatics (<http://www.informatics.jax.org/>).

### Many Unexpected Phenotypes Discovered

A few examples of the wide range of phenotypes we discovered are illustrated in Figure 4. Body weight and fat/lean composition were among the most common anomalies, with both overweight ( $n = 2$ ) and underweight ( $n = 21$ ) mutants discovered. The *Kptn* mutant is an example of an unexpected phenotype. *Kptn* is a putative actin binding protein proposed as a candidate for deafness because it is expressed in sensory hair cells (Bearer et al., 2000). Instead, the homozygous *Kptn* mutant has increased body weight on a high-fat diet (Figure 4A) and increased bacterial counts following *Salmonella* Typhimurium challenge but normal hearing (Table S2). Additional new phenotypes were detected in genes that had been published previously, such as reduced grip strength and ankylosis of the metacarpophalangeal joints in *Dnase1l2* mutants (Fischer et al., 2011), delayed response in the hot plate test in *Git2* mutants (Schmalzigaug et al., 2009), and small sebaceous glands in *Cbx7* mutants (Forzati et al., 2012) (Figures 4B–4E, 4G, and 4H). Phenotypes were also detected in genes that had not been published previously, such as impaired hearing in *Fam107b* mutants and elevated plasma magnesium concentration in *Rg9mtd2* mutants (Figures 4F and 4I, respectively). These examples demonstrate that many phenotypes will be missed unless they are specifically looked for and illustrate the value of carrying out a broad range of screens with all mutants going through all screens. They also reveal our collective inability to predict phenotypes based on sequence or expression pattern alone.

### Haploinsufficient and Nonessential Genes

Haploinsufficient phenotypes were detected in 38 of 90 (42%) of these lines. Thus, haploinsufficiency is relatively common, suggesting that screening heterozygotes of knockout lines can yield valuable insight into gene function and provide models for dominantly inherited human disorders. All 90 genes screened as heterozygotes had at least 1 hit (usually viability) and together gave a total of 181 hits (ranging from 1 to 14 per line), an average of 2.0 per line, or 1.0 per line if we consider that abnormal viability is a feature of the homozygote. The distributions of phenotypic hits are shown in Figures 5A and 5B. Two examples of haploinsufficiency are illustrated in Figures 4J–4N.

A total of 837 phenotypic variants were detected in the 250 mutant lines, 1.27% of the total calls (Tables S1 and S2). Of

the lines screened as homozygotes or hemizygotes, 35% (56 of 160) appeared completely normal in our screen. There are several possible explanations for the lack of a detected phenotype, such as incomplete inactivation of the gene, a subtle change in phenotype not detected by our screen, or the gene may be nonessential. So far, there is no overlap between the 56 mouse lines with no detected phenotype and genes homozygously inactivated in humans, but both data sets are limited in coverage to date (MacArthur et al., 2012). The remaining 104 homozygous/hemizygous lines gave a total of 656 hits (range 0–41 per line), an average of 6.3 hits per line.

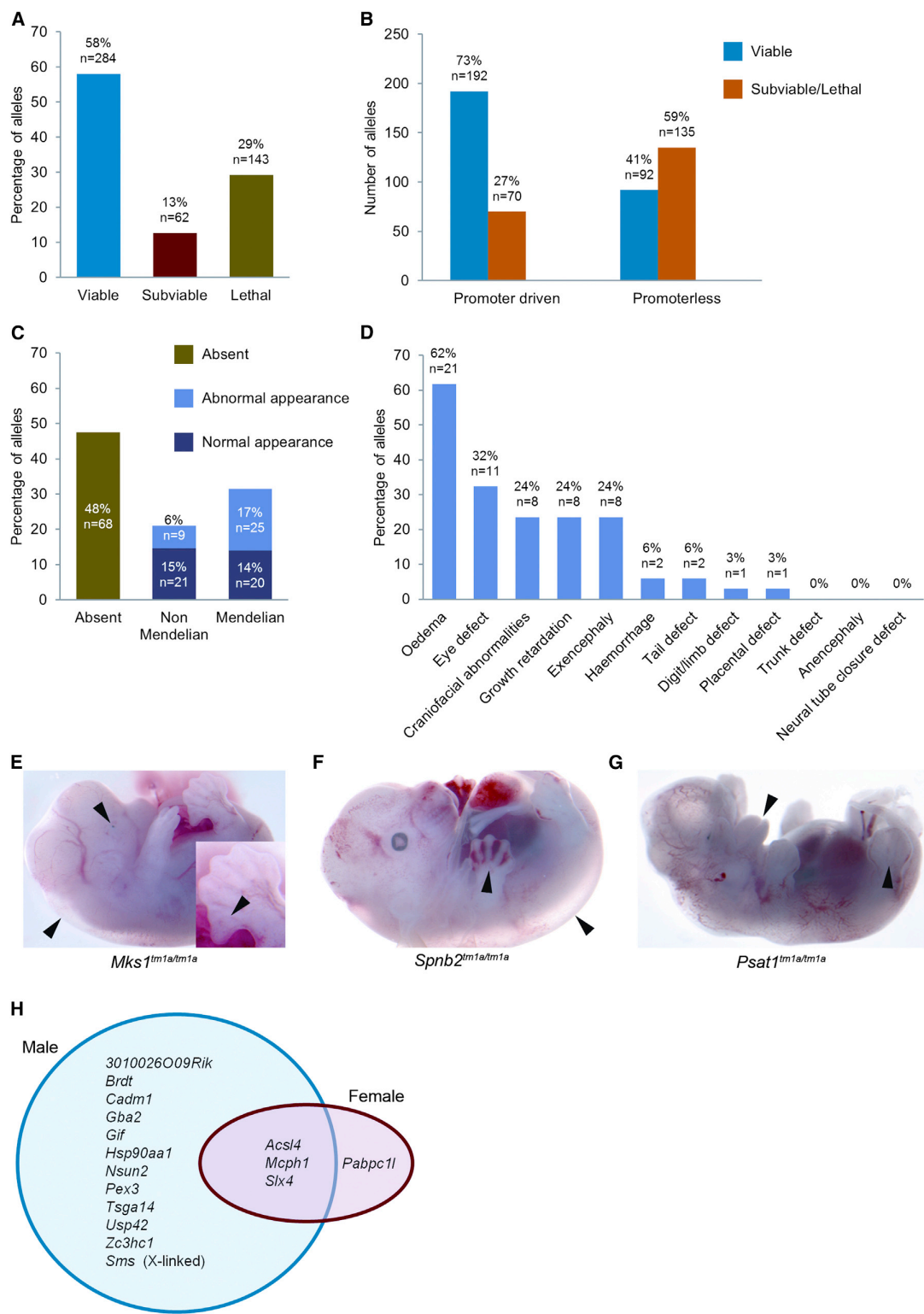
### Sensitivity of the MGP Screen

To assess the sensitivity of our screen, the phenotypes were compared with published data on alternative alleles where available. A total of 91 of 250 genes had published data reported in MGI (Table S5), and for 61 of these, our observations detected features of the published phenotypes. Importantly, for 56 genes, a new phenotype was detected by our screen (column K, Table S5). For 31 genes, features of the published phenotype were assessed but not detected by our pipeline. For example, *Asxl1<sup>tm1Bc/tm1Bc</sup>* mice are published as being viable (Fisher et al., 2010), but we found that *Asxl1<sup>tm1a(EUCOMM)Wtsi</sup>* homozygotes were lethal, with none detected among 276 progeny from heterozygous intercrosses ( $\chi^2 = 95.13$ ,  $df = 2$ ;  $p < 2.2 \times 10^{-16}$ ). These discrepant cases may reflect differences in the allele and/or genetic background. In other cases (77 genes), the reported characteristics required a specialized test not included in our screen, such as the calcium signaling defect in cardiomyocytes of *Anxa6<sup>tm1Moss/tm1Moss</sup>* mutants (Hawkins et al., 1999).

### New Mouse Models for Human Disease

The data set reported here includes 59 orthologs of known human disease genes. We compared our data with human disease features described in OMIM (Table S6). Approximately half (27) of these mutants exhibited phenotypes that were broadly consistent with the human phenotype. However, many additional phenotypes were detected in the mouse mutants suggesting additional features that might also occur in patients that have hitherto not been reported. Interestingly, a large proportion of genes underlying recessive disorders in humans are homozygous lethal in mice (17 of 37 genes), possibly because the human mutations are not as disruptive as the mouse alleles. Of the 59 genes, 26 represent the first mouse mutant with publicly available data. Three examples (*Sms*, *Ap4e1*, and *Smc3*) representing the first targeted mouse mutant for each gene are illustrated in Figure 6, and all show similar phenotypic features to their human counterparts.

*SMS* mutations in humans cause X-linked Snyder-Robinson syndrome involving hypotonia, unsteady gait, diminished muscle mass, kyphoscoliosis, osteoporosis, facial asymmetry, and intellectual disability (Cason et al., 2003). Hemizygous *Sms* mutant male mice showed reduced muscle strength, lean mass and bone mineral density, lumbar lordosis (Figures 6A–6E), and growth retardation, recapitulating features of the human disease. In addition, male infertility was detected, suggesting a feature



(legend on next page)

that may not have been recognized in humans with SMS mutations.

Spastic paraplegia 51, autosomal recessive, is caused by mutations in *AP4E1* and leads to spastic tetraplegia with hyper-reflexia and generalized hypertonia, microcephaly, intellectual disability and dilated ventricles, cerebellar atrophy, and/or abnormal white matter (Abou Jamra et al., 2011; Moreno-De-Luca et al., 2011). Homozygous *Ap4e1* mutant mice displayed several similarities such as increased lateral ventricle area, decreased corpus callosum span, and decreased rearing (Figures 6F–6J). In addition, hematological changes suggestive of anemia were detected in female *Ap4e1* mutants, which have not been reported in humans.

The third example is *SMC3*, associated with dominantly inherited Cornelia de Lange syndrome 3, featuring facial dysmorphism, hirsutism, growth retardation, neurodevelopmental delay, and upper-limb anomalies (Deardorff et al., 2007). *Smc3* mutant mice displayed homozygous lethality (prior to E14.5) and reduced heterozygote viability at P14 (45% instead of the expected 67%;  $p' = 0.0064$ ). Surviving *Smc3* heterozygotes showed reduced body weight, and a subset showed a distinct craniofacial morphology (Figures 6K–6M). Distinct *Smc3* expression in hair follicles and key brain substructures was revealed using *lacZ* (Figures 6N and 6O), noteworthy because of the hirsutism and neurodevelopmental delay aspects of Cornelia de Lange syndrome 3. In addition, an increase in the number of helper and cytotoxic T cells was observed in the mutant mice, again indicating an aspect that might contribute to the phenotype of patients but that has not yet been reported.

### Pleiotropic Effects of Mutations

The phenotypes detected in this study vary from discrete specific defects (e.g., decreased platelet cell number in *Crif3<sup>tm1a</sup>/tm1a* mutants) to complex phenotypes in which many organ systems are involved (e.g., *Spns2<sup>tm1a(KOMP)Wtsi</sup>* homozygotes show eye, hearing, and immune defects; Nijnik et al., 2012). The distribution of phenotypic hits is shown in Figures 5A and 5B for homozygous and heterozygous mutants, respectively. The peak for homozygotes was the category with no detected abnormalities, whereas the second biggest group consists of mutants with just one phenotypic call. The lines examined as heterozy-

gotes all have at least one hit (viability), but 20 lines have in addition one other abnormal phenotype, and a handful have several. Classifying parameters into five disease categories, we analyzed the distribution of disease areas represented across all 250 mouse lines. The most common phenotypic call was in the category reproduction, development, and musculoskeletal (Figure 5C).

Some abnormal phenotypes are clearly not primary effects; for example, reduced weight may be a secondary consequence of a number of different primary defects. Given that certain phenotypic features would be expected to co-occur frequently, reflecting physiological or developmental associations, a principal component analysis was conducted to look for correlated patterns in the data. Plotting principal component 1 against 2 revealed four main clusters of mouse lines (colored ovoids in Figure 5D). The separation along principal component 2 arises from viability. The remaining separation of clusters marked by red and green from clusters marked blue and yellow (Figure 5D) arises from body weight and associated variables, including DEXA measurements and energy use (Figure 5E). Body weight is a common covariable in disease (Reed et al., 2008), so it is not surprising that it dominates the principal component analysis.

### Features of Essential Genes

Genes are generally defined as essential if they are required for survival or fertility. Studies in yeast and worms suggest that genes with paralogs are much less likely to be essential, presumably because the paralog can compensate for the function of the inactivated gene (Gu et al., 2003; Conant and Wagner, 2004). Previous analyses of published data on mouse knockouts did not find a significant difference in essential genes between singleton and duplicated genes (Liang and Li, 2007; Liao and Zhang, 2007). However, the published gene set is biased toward genes involved in development (Makino et al., 2009). In contrast, we found that genes in our set without a paralog were more than twice as likely to be essential, a significant effect (Table S3; Figure 7A).

We next asked if the essential genes in our gene set are more likely to be involved in a protein complex, using an experimentally validated data set of human protein complexes from the

### Figure 2. Homozygous Viability and Fertility Overview

(A) Homozygous viability at P14 was assessed in 489 EUCOMM/KOMP targeted alleles. A minimum of 28 live progeny were required to assign viability status. Lines with 0% homozygotes were classed as lethal, >0% and ≤13% as subviable, and >13% as viable.

(B) Comparison of homozygous viability data from targeted alleles carrying either a promoter-driven or promoterless neomycin selection cassette.

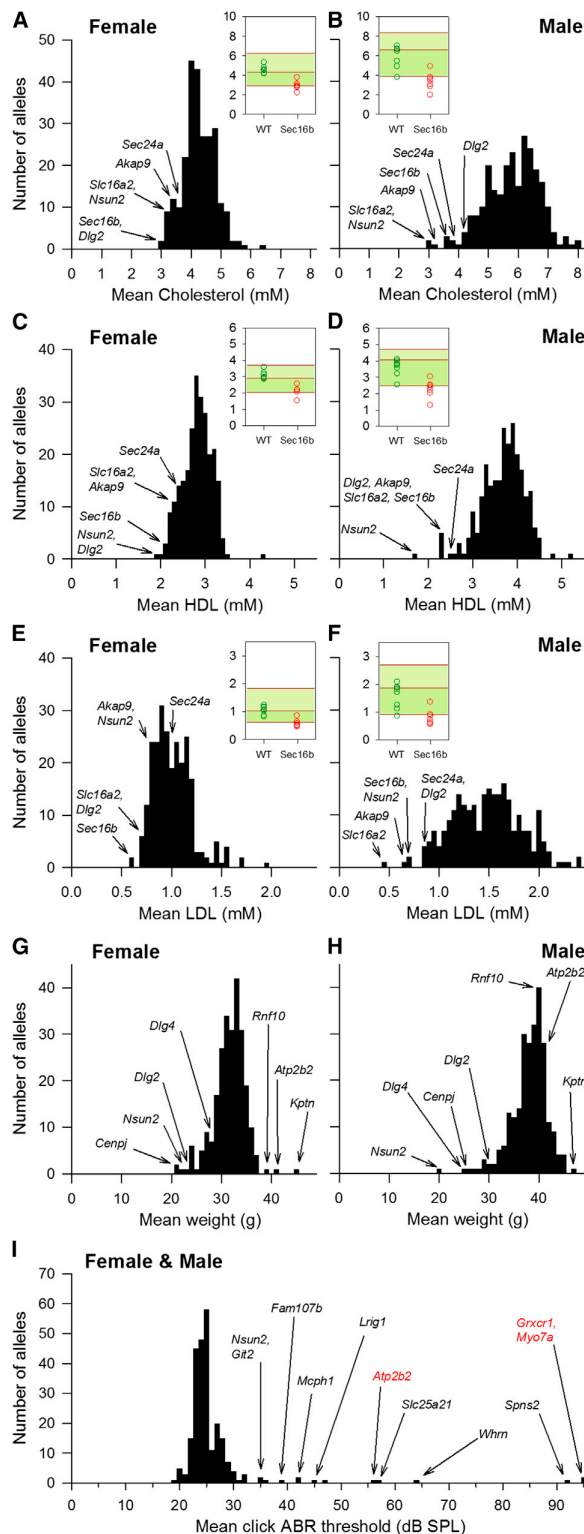
(C) Lines classed as lethal or subviable at P14 were further assessed for viability at E14.5. Of the 205 targeted alleles eligible for this recessive lethality screen, 143 are reported here. A total of 28 embryos were required to assign viability status, and outcomes were categorized by both the number and dysmorphology of homozygous offspring.

(D) A basic dysmorphology screen encompassing 12 parameters was performed on all embryos for the 75 targeted alleles classed as viable or subviable at E14.5. A total of 34 targeted alleles showed one or more abnormality, and the percentage incidence is presented.

(E–G) Examples of E14.5 dysmorphology (arrowheads indicate abnormalities) are presented. Homozygous progeny were detected at a Mendelian frequency in all three examples. Sixty-seven percent (six of nine) *Mks1<sup>tm1a/tm1a</sup>* embryos presented with edema, polydactyly, and eye defects (E). Sixty-two percent (five of eight) *Spnb2<sup>tm1a/tm1a</sup>* embryos presented with edema and hemorrhage (F). Eighty-six percent (six of seven) *Psat1<sup>tm1a/tm1a</sup>* embryos presented with growth retardation, exencephaly, and craniofacial abnormalities (G).

(H) Fertility was assessed in homozygous viable lines (307 mouse lines assessed from a total of 331 eligible lines). At least four independent 6-week-old mice of each sex were mated for a minimum of 6 weeks, and if progeny were born, the line was classed as fertile, regardless of if the progeny survived to weaning. Of note is the strong skew toward male (blue circle) fertility issues (15 of 16 genes) compared to 4 of 15 genes that displayed female (red circle) fertility issues. See also Table S3.





**Figure 3. Data Distributions for Selected Parameters**

(A–F) Distribution of mean total cholesterol (A and B), mean HDL cholesterol (C and D), and mean LDL cholesterol (E and F) at 16 weeks of age in both sexes for 250 unique alleles. Outliers are identified by gene name. The insets in (A)–(F) present the data for one outlier, *Sec16b<sup>tm1a/tm1a</sup>* (red circles represent individual mice), compared to the WT controls processed during the same week (green circles), and a cumulative baseline of all WT mice of that age, sex, and genetic background (>260 WT mice) is presented as the median and 95% confidence interval.

(G and H) Distribution of mean body weight at 16 weeks in (G) female and (H) male mutant lines of mice. Outliers are identified by gene name.

(I) Distribution of mean click ABR threshold at 14 weeks (typically  $n = 4$ , independent of sex). Outliers are identified by gene name including positive controls highlighted in red.

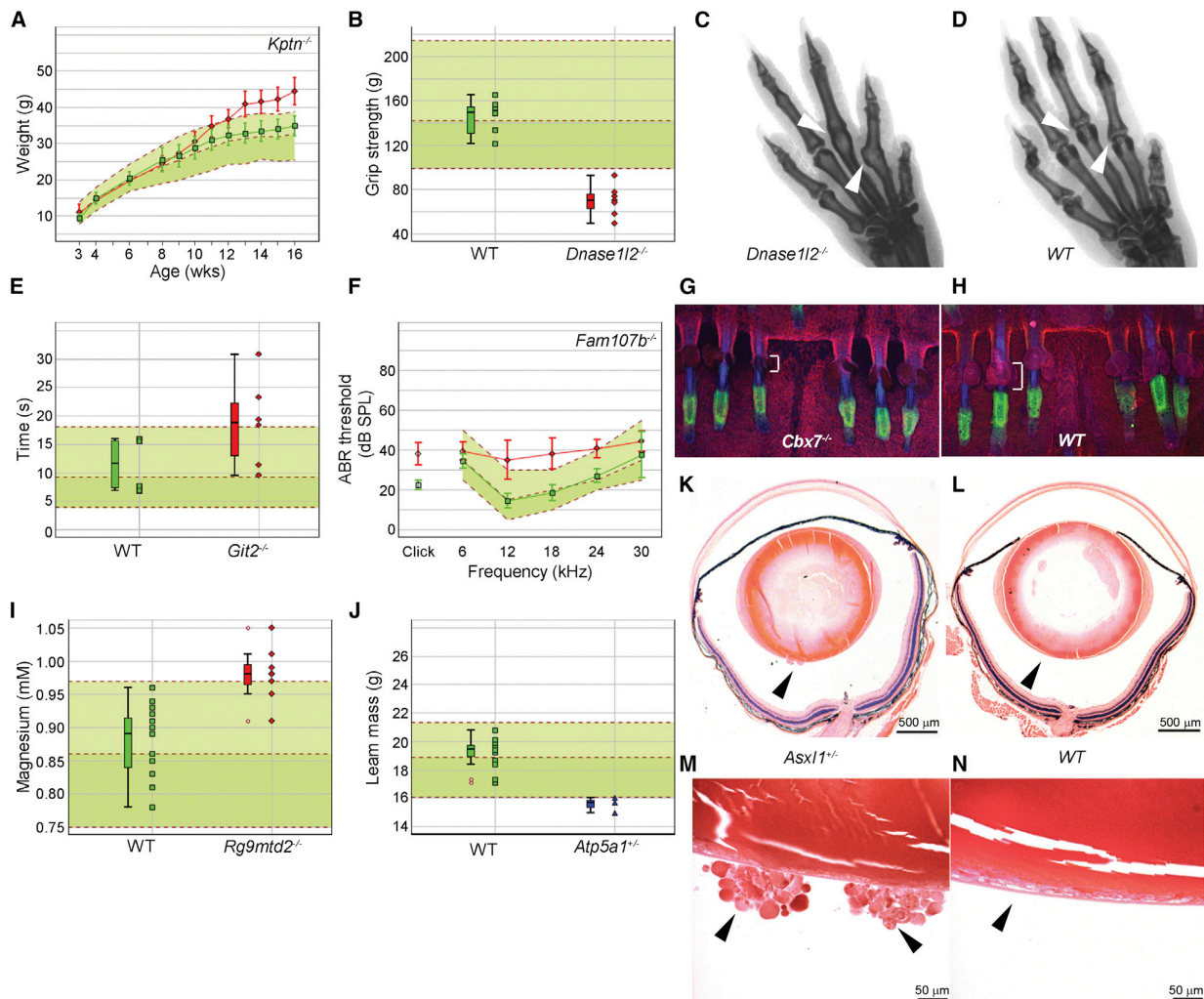
CORUM database (Ruepp et al., 2010). We found that genes with a human ortholog that is part of a complex were significantly more likely to be essential (Table S3; Figure 7B).

Finally, we asked if there were certain types of gene products that were more likely to be important for viability/fertility than others. In humans, transcription factor mutations appear enhanced in prenatal disease, and enzymes are overrepresented in diseases with onset in the first year after birth (Jimenez-Sanchez et al., 2001). We investigated four classes of protein identified by GO terms: transcription factors ( $n = 7$ ), transmembrane proteins ( $n = 50$ ), enzymes ( $n = 131$ ), and chromatin-associated proteins ( $n = 24$ ). Numbers of each were limited, but there was no significant enrichment for essential genes among any of the four groups (Tables S2 and S3).

In summary, we found that essential genes were less likely to have a paralog and more likely to be part of a protein complex, but no specific class of protein appeared more likely to be predictive of essentiality.

### Annotating the Function of Novel Genes

There is a large bias in the literature toward analysis of known genes (Edwards et al., 2011), but are genes that have yet to be examined experimentally less likely to underlie disease? Genes in our set that had no associated publications (other than high-throughput genome-wide reports) were compared with genes where some aspect of their function had been described. The proportion of essential genes among the novel set was not significantly different from the known genes (Figure 7C). Furthermore, there was no significant difference in the number of hits observed per line between known and novel genes (Tables S2 and S3). As a second test, we asked if genes with orthologs involved in human disease (having an OMIM disease ID) were enriched in essential genes or the number of phenotypic hits compared with genes not (yet) ascribed to human disease, but there was no significant difference (Tables S2 and S3; Figure 7D). Finally, we compared genes that had been proposed for inclusion by the community ( $n = 87$ ) with those with no specific request to ask if genes of interest to the community were more likely to be essential or to have detected phenotypes. There was no significant difference between the two groups (Table S3). Thus, known genes are no more likely to be involved in disease than novel genes, emphasizing that much new biology will be uncovered from the analysis of mutations in novel genes.



**Figure 4. Examples of Novel Phenotypes from a Wide Range of Assays with Particular Focus on Novel Genes**

(A) Elevated body weight gain of *Kptn*<sup>tm1a/tm1a</sup> females (n = 7) fed a high-fat diet from 4 weeks of age. Mean  $\pm$  SD body weight is plotted against age for *Kptn*<sup>tm1a/tm1a</sup> females (red line) and local WT controls run during the same weeks (n = 16; green line). The median and 95% reference range (2.5% and 97.5%; dotted lines) for all WT mice of the same genetic background and sex (n = 956 females) are displayed on the pale green background.

(B) Reduced grip strength in *Dnase1l2*<sup>tm1a/tm1a</sup> males (n = 7) (red symbols) compared with controls (n = 8) (green symbols) and the reference range (n = 289). Each mouse is represented as a single symbol on the graph. Median, 25<sup>th</sup> and 75<sup>th</sup> percentile (box), and the lowest and highest data point still within 1.5 $\times$  the interquartile range (IQR) (whiskers) are shown.

(C and D) Ankylosis of the metacarpophalangeal joints (arrowheads) shown by X-ray in *Dnase1l2*<sup>tm1a/tm1a</sup> mice (C) (six of seven males; five of seven females) compared with WT controls (D) correlates with reduced grip strength (B).

(E) Increased latency to respond to heat stimulus in *Git2*<sup>Git(XG510)Byg/ Git(XG510)Byg</sup> females (n = 6) (red symbols) compared with controls (n = 4) (green symbols) and the reference range (n = 115), with box and whisker plots on the left (see Figure 4B legend).

(F) Mild hearing impairment at the middle range of frequencies in *Fam107b*<sup>tm1a/tm1a</sup> mutants (n = 8) (red line shows mean  $\pm$  SD) compared with controls (n = 10) and the reference range (n = 440).

(G) Smaller sebaceous glands (indicated by bracket) in *Cbx7*<sup>tm1a/tm1a</sup> mutant tail skin hairs compared with WT (H).

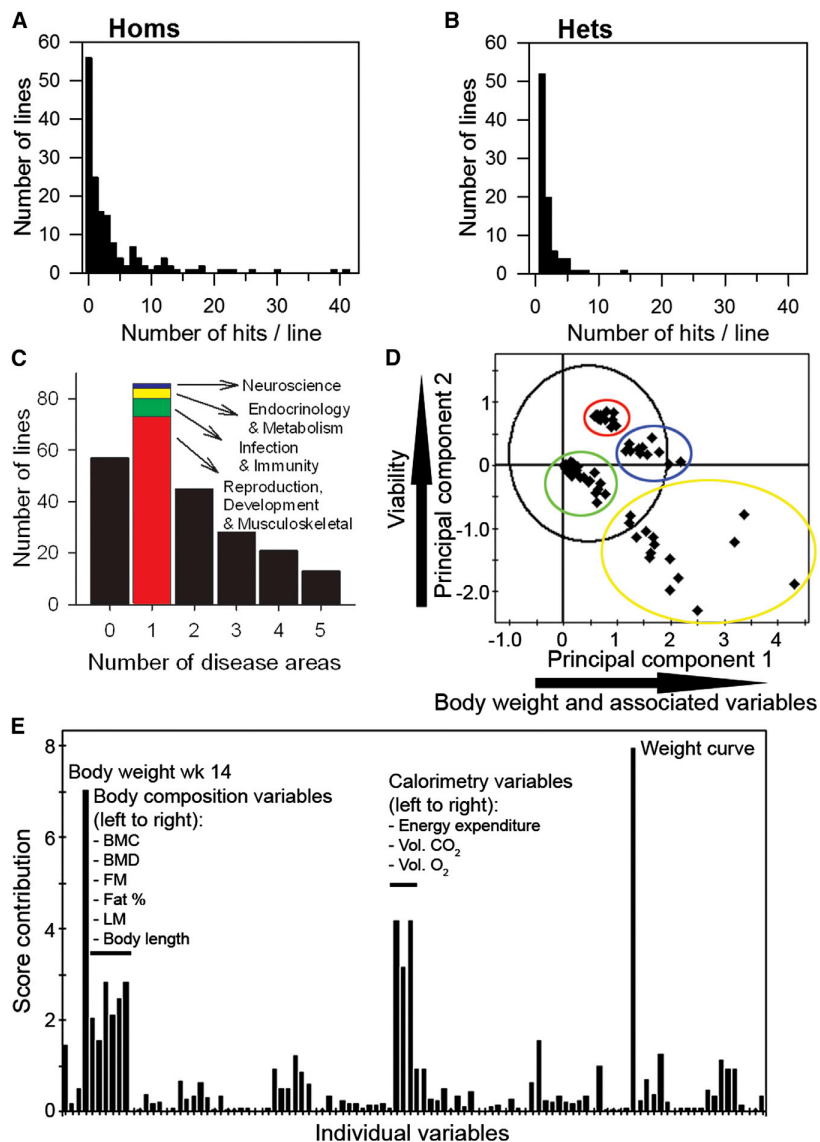
(I) Increased plasma magnesium levels in *Rg9mtd2*<sup>tm1a/tm1a</sup> males (n = 8) (red symbols) compared with local controls (n = 15) (green symbols) and the reference range (n = 241), with box and whisker plots on the left (see Figure 4B legend).

(J) Decreased lean mass in *Atp5a1*<sup>tm1a/+</sup> females (n = 3) (blue symbols) compared with local controls (n = 15) (green symbols) and the reference range (n = 757), with box and whisker plots on the left (see Figure 4B legend).

(K and L) Histopathology showed opacities in the vitreous of eyes from *Asx11*<sup>tm1a/+</sup> mice (K) (arrowheads; scale bar, 500  $\mu$ m) compared with empty vitreous in WT (L).

(M and N) Higher magnification revealed round opacities extending from the lens into the vitreous (arrowheads; scale bar, 50  $\mu$ m) in *Asx11*<sup>tm1a/+</sup> mice (M) compared with a normal lens contained within the lens capsule in WT mice (N).

See also Table S5.



**Figure 5. Characteristics of Phenotypic Hits Detected**

(A) Distribution of the number of phenotypic hits in each line screened as homozygotes showing the peak at no hits but a long tail of lines with multiple hits up to 41.

(B) Distribution of hits in lines screened as heterozygotes; all lines had at least one hit (for viability) with a spread up to 14 hits.

(C) Distribution of lines with hits in different disease areas showing a peak of lines with just one area affected (colors indicate which areas) but some lines with multiple disease areas involved, indicating a high degree of pleiotropy.

(D) Principal component analysis score scatterplot showing the deviation of each gene from the first two principal components to visualize the clustering in genes within the multidimensional space. The black ovoid represents the Hotelling's  $T^2$  95% confidence limits. Colored ovoids mark four different clusters of mutant lines. The two main principal components (or latent variables) in the model are significant in explaining 19.2% and 11.7% of the variation, respectively, and are predictive.

(E) Principal component analysis contribution plot indicating the contribution of the variables to the separation between the red and green clusters compared to the blue and yellow clusters in (D). Major phenotypic contributions are labeled. Key to variables is presented in Table S7.

standard allele design with the option of generating conditional mutations, and all are available from public repositories.

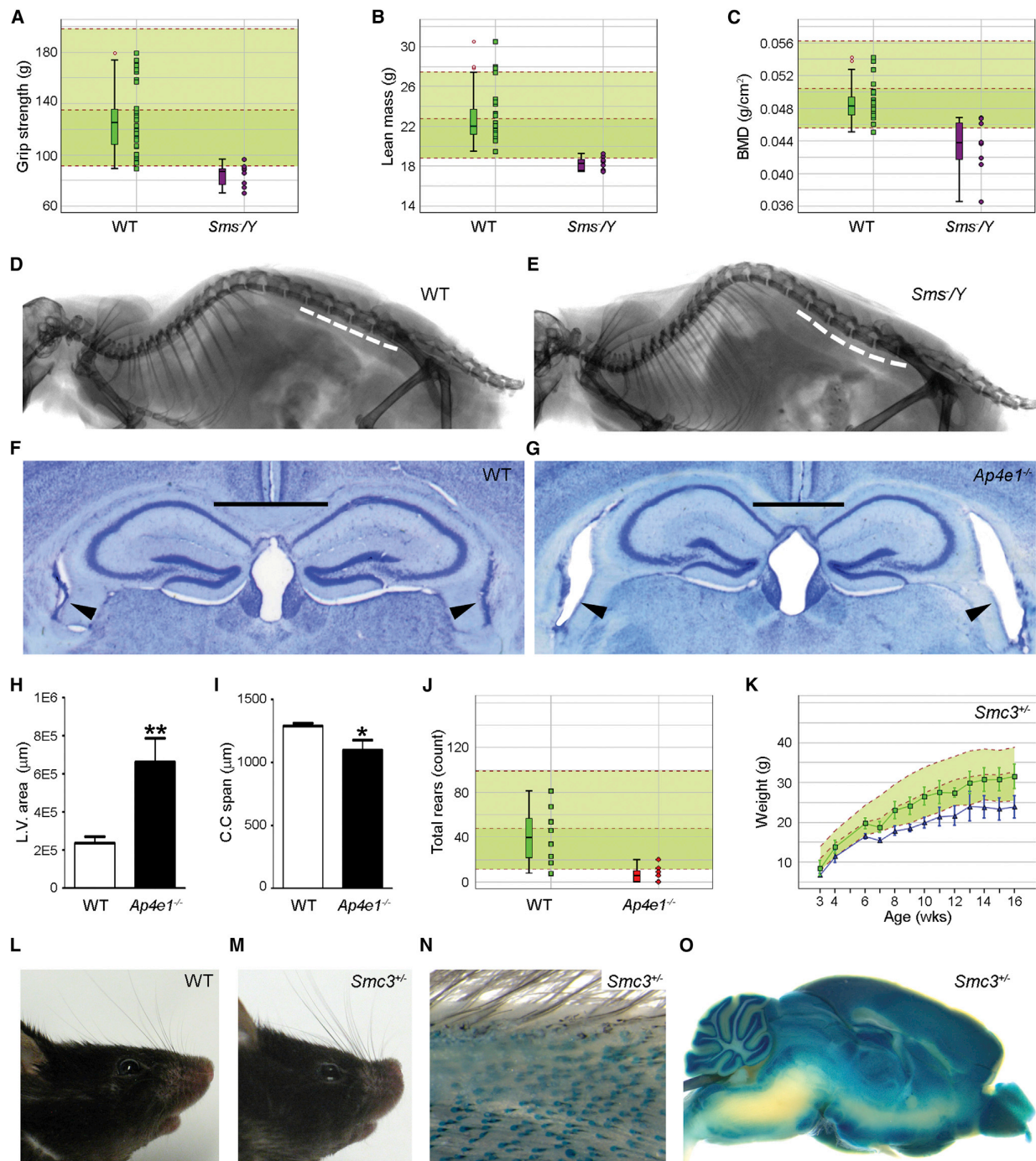
The phenotyping described here was not intended to provide an exhaustive characterization of the phenotype of the mutant lines but, rather, to place mutant alleles into broad categories by using screens, generating a pool of genetic resources from which individual mutants can be selected based on their phenotype for secondary follow-up studies. Several of the mutants have been

analyzed further following an initial phenotypic observation in the screen, and these add to the depth of our knowledge of biological mechanisms of disease (e.g., Nijnik et al., 2012; Crossan et al., 2011). As the assembled data expands, it will become possible to discern patterns between phenotypes and come to more holistic conclusions about categories of genes. Genes linked by common phenotypes can be grouped together to test for regulatory or other functional interactions and ultimately placed into pathways that in turn will implicate other genes in the disease process. For example, of the four genes associated with abnormal fasting glucose levels in our data set, *Slc16a2* can be linked to *Ldha* via regulation of L-triiodothyronine (Friesema et al., 2006; Miller et al., 2001), but the other two genes, *Nsun2* and *Cyb561*, have no reported regulatory links apart from in vitro protein-protein interactions, so these represent

## DISCUSSION

Genetic studies in mice via targeted mutagenesis of ES cells have been successful at illuminating selected aspects of the function of more than 7,000 mammalian genes. However, until recently, these studies have been conducted by individual laboratories and largely directed at previously studied genes. The focused collection of phenotypic information from these mutants has been very information rich, but many aspects remain undetected because they are outside the area of interest of the laboratory generating the mutant. Individual endeavors have led to wide variation in allele design and genetic backgrounds used, and all too often, the mutant is not available to other groups for further analysis. In contrast, the mutant mice described here have the advantage of a common genetic background and a





**Figure 6. Correlated Disease Characteristics in Knockouts of Three Known Human Disease Genes**

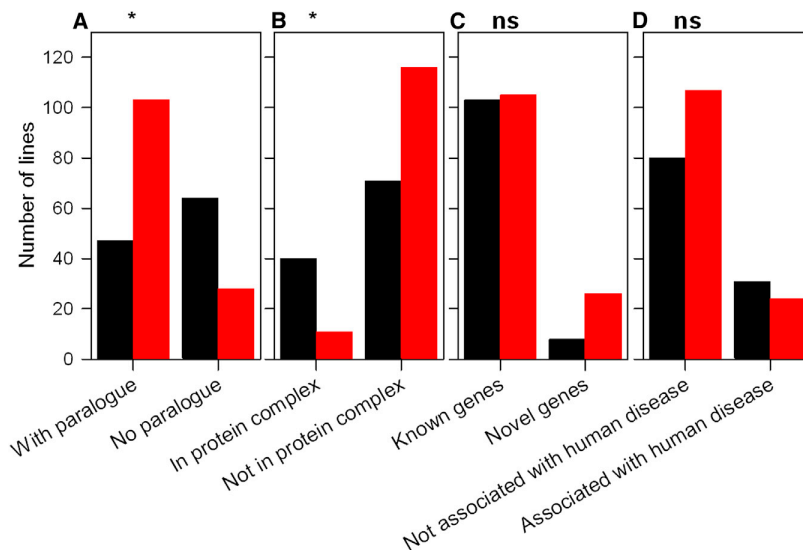
(A–E) Male hemizygotes for the *Sms* mutation showed similar features to X-linked Snyder-Robinson syndrome.

(A) Reduced grip strength in *Sms/Y* mice ( $n = 8$ ) (purple symbols) compared with WT controls ( $n = 30$ ) (green symbols) and the reference range ( $n = 793$ ). Each mouse is represented as a single symbol on the graph, with box and whisker plots on the left (see Figure 4B legend).

(B and C) Decreased lean mass (B) and bone mineral density (C) in *Sms/Y* mice ( $n = 8$ ) (purple symbols) compared with controls ( $n = 27$ ) (green symbols) and the reference range ( $n = 753$ ), with box and whisker plots on the left (see Figure 4B legend).

(D and E) Lumbar lordosis shown by X-ray (seven of eight males) in *Sms/Y* (E) compared with WT (D).

(legend continued on next page)



**Figure 7. Features Associated with Essential Genes**

Essential genes (black bars) are compared with genes that are not essential for viability (red bars). The asterisk (\*) indicates significant difference. ns, no significant difference in proportion of essential genes between the two categories. Statistics are presented in Table S3.

(A) Genes with no paralogue show a significantly larger proportion of essential lines than genes with at least one paralogue.

(B) Genes predicted to contribute to protein complexes showed a significantly larger proportion of essential lines than genes not predicted to contribute to a complex.

(C) Novel genes showed no significant difference in proportion of essential genes or number of hits than known genes.

(D) Genes known to underlie human disease were no more likely to be essential than genes not yet associated with human disease.

candidates to investigate further. Already some broad conclusions can be drawn from the data set, such as the value of analyzing novel genes, the increased incidence of essentiality in genes with no paralogue, and the increased number of genes required for male compared to female fertility. Many completely unexpected associations between genes and phenotypes have been discovered, illustrating the value of a broad-based screen.

Another aspect of our study was the examination of heterozygous mutants, a genotype that often is not studied by individual laboratories. Although this was restricted to mutants that displayed lethality or subviability of homozygotes, it revealed a number of genes with haploinsufficiency, a feature commonly associated with mutations in the human genome but rarely described in mouse knockouts.

The tests used in screening varied considerably in their complexity, cost, and suitability in a high-throughput scenario. The performance of these tests across 250 alleles provided insight into those that should be included or excluded in the efforts to examine 5,000 alleles through the activities of the IMPC. Key considerations are variance in the control group, specificity, sensitivity, effect size, and redundancy.

The major contribution of null alleles will be an improved understanding of biological processes and molecular mechanisms of disease. The null allele will give insight into the temporal and spatial requirements for the gene and will contribute to the establishment of gene networks involved in mammalian disease processes. Furthermore, our data set demonstrates that many features of human Mendelian diseases can be found in the corresponding mouse mutant. The mouse alleles studied here are expected to be null alleles or strong hypomorphs, which may not always reflect the consequence of the human mutation. However, null alleles should reveal haploinsufficiency and recessive effects due to deleterious mutations such as frameshift and nonsense mutations. Null alleles in the mouse are likely to make the largest impact upon understanding human diseases caused by rare variants of large effect size. Complex multifactorial diseases, which may depend on human-specific variants with small effect size or more specific molecular effects such as gain-of-function mutations, will require more customized approaches such as knockin of specific human mutations. Alternative approaches using the mouse for discovering loci underlying complex disease include the Hybrid Mouse Diversity panel and the Collaborative Cross (reviewed by Flint and Eskin, 2012).

(F–J) *Ap4e1<sup>tm1a/tm1a</sup>* mice displayed similarities to spastic quadriplegic cerebral palsy 4.

(F–I) Increased lateral ventricle area (arrowheads in F and G) and decreased corpus callosum span (solid lines in F and G) in *Ap4e1<sup>tm1a/tm1a</sup>* mice (G) compared with WT mice (F) with measurements plotted (mean ± SD) in (H) and (I), respectively (\*p < 0.05, \*\*p < 0.01; n = 3 mutant males and 34 WT males). Error bars in (H) and (I) are SD.

(J) Decreased rearing in *Ap4e1<sup>tm1a/tm1a</sup>* females (n = 7) (red symbols) compared with WT controls (n = 8) (green symbols) and the reference range (n = 180), with box and whisker plots on the left (see Figure 4B legend).

(K–O) Surviving *Smc3<sup>tm1a/+</sup>* mice showed similar features to Cornelia de Lange syndrome 3.

(K) Decreased body weight in *Smc3<sup>tm1a/+</sup>* females (n = 7) fed on high-fat diet. Mean ± SD body weight is plotted against age for *Smc3<sup>tm1a/+</sup>* females (blue line), WT mice (n = 24; green line), and the reference range (n = 948).

(L and M) Distinct craniofacial abnormalities in *Smc3<sup>tm1a/+</sup>* mice including upturned snout (M) (three of seven males, one of seven females), which was not observed in WT (L) (n = 850 male and 859 female).

(N and O) The *lacZ* reporter gene revealed a distinct *Smc3* expression pattern including (N) hair follicles and (O) key brain substructures, noteworthy because of the hirsutism and neurodevelopmental delay aspects of Cornelia de Lange syndrome 3.

See also Table S6.

These allow interrogation of many different loci simultaneously and study of epistatic interactions and can lead to identification of single gene variants causing disease (e.g., Orozco et al., 2012; Andreux et al., 2012), when variants affecting the trait of interest are present in the founders. ENU mutagenesis is another powerful technique that can be used to produce allelic series of mutations with differing effects upon function of single genes (e.g., Andrews et al., 2012). However, the null alleles that we describe here are a complement to these alternative approaches and will be invaluable for defining mechanisms of gene function on a standard genetic background.

The study described in this report builds on the large KOMP/EUCOMM resource of targeted mutations in mouse ES cells (Skarnes et al., 2011) and illustrates the breadth of phenotypic information that can be garnered from an organized effort. The Clinical Phenotyping Pipeline optimized here has been adopted by several other programs within the IMPC; multiple groups are now working together to extend what is described in this report for 250 genes to 5,000 genes over the next 4 years with the vision that this will eventually cover all protein-coding genes. The primary phenotypes and genetic resources emerging from these programs will make a significant contribution to our understanding of mammalian gene function.

## EXPERIMENTAL PROCEDURES

### Animals

Mice carrying knockout first conditional-ready alleles (Figures S1A and S1B) were generated from the KOMP/EUCOMM targeted ES cell resource using standard techniques. Eight in-house lines were included as known mutant controls. Details of the 250 lines can be found in Table S2. All lines are available from <http://www.knockoutmouse.org/>; or [mouseinterest@sanger.ac.uk](mailto:mouseinterest@sanger.ac.uk). Mice were maintained in a specific pathogen-free unit under a 12 hr light, 12 hr dark cycle with ad libitum access to water and food. The care and use of mice were in accordance with the UK Home Office regulations, UK Animals (Scientific Procedures) Act of 1986.

### Genotyping and Allele Quality Control

Short-range, long-range and quantitative PCR strategies (<http://www.knockoutmouse.org/kb/25/>) were used to evaluate the quality of each allele (Figure S1C). A subset of these assays was used to genotype offspring. The degree of knockdown in homozygotes was assessed by qRT-PCR of adult liver in a subset known to show expression in liver. Details are given in Extended Experimental Procedures.

### Phenotyping Pipeline and Tests

The typical workflow from chimera to primary phenotyping pipelines and an outline of the clinical phenotyping pipeline are presented in Figure 1. Details of batch size are given in Figure S2A. These pipelines include established tests used to characterize systematically every line of mice as described in Table S4.

### Histochemical Analysis of the *lacZ* Reporter

Adult whole-mount *lacZ* reporter gene expression was carried out essentially as described by Valenzuela et al. (2003).

### Statistical and Bioinformatic Analysis

For continuous data, including time course, a reference range approach was used to identify phenotypic variants as detailed in Figures S3A and S3B. Fisher's exact test was used to assess categorical data (Figure S3C). These automated calls were complemented by a manual assessment made by biological experts. An example of the establishment of the reference range is given in Figure S2B.

Downstream data analysis was performed using SPSS (version 17.0.2), R, and SIMCA-P (V-12.0, Umetrics). The data structure and biological question determined the statistical test used; details are in Table S3. Principal components analysis was performed in SIMCA-P (<http://www.umetrics.com>).

Further details of analyses and gene annotations are given in Extended Experimental Procedures.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, three figures, seven tables, and a complete list of contributors from the Sanger Institute Mouse Genetics Project and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2013.06.022>.

## ACKNOWLEDGMENTS

We thank Alex Bateman, Lars Barquist, Steve O'Rahilly, Keith Burling, Seth Grant, Pentao Liu, and Lorraine Everett for advice and the EUMODIC consortium for discussions. This work was supported by the Wellcome Trust (grant Nos. 098051 to Wellcome Trust Sanger Institute and RG45277 PCAG/116 to F.M.W.), Medical Research Council (to K.P.S. and F.M.W.), European Commission (EUMODIC contract No. LSHG-CT-2006-037188), NIH (EY08213 to S.H.T. and 5K08EY020530-02 to V.B.M.), Research to Prevent Blindness (to S.H.T. and V.B.M.), Australian Research Council (DP1092723 to I.S.), and Cancer Research UK (to D.J.A.). J.K.W. and K.P.S. conceived and devised the single phenotyping pipeline and principles of analysis and presentation of the data; R.R.S., J.K.W., R.H., J.R.B., and E.R. managed mouse production and genotyping; J.N.B. and J.S. managed mouse breeding; A.K.G., C.P., J.E., D.S., N.I., and J.K.W. managed mouse phenotyping and analyzed data; the Sanger Institute Mouse Genetics Project team contributed to all aspects of the work; S.C. and G.D. led the infectious challenge screen; V.B.M. and S.H.T. led the eye histopathology screen; I.S. and F.M.W. led the skin screen; J.F. led the brain histopathology screen; D.J.A. led the micronucleus screen; J.N.B. and J.R.B. managed mouse distribution; J.K.W., A.K.G., M.B., and K.P.S. compiled Tables S1, S2, S5, and S6; N.A.K., D.M., D.S., and M.B. carried out annotation and statistical analysis; N.C.A., D.G.M., and W.C.S. led development of informatics support; D.W.L. wrote Wiki pages for the mutants; J.K.W., D.J.A., R.R.S., and K.P.S. led the project; and K.P.S., A.B., and J.K.W. wrote the paper with contributions from all authors.

Received: March 14, 2013

Revised: May 10, 2013

Accepted: June 17, 2013

Published: July 18, 2013

## REFERENCES

- Abou Jamra, R., Philippe, O., Raas-Rothschild, A., Eck, S.H., Graf, E., Buchert, R., Borck, G., Ekici, A., Brockschmidt, F.F., Nöthen, M.M., et al. (2011). Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature. *Am. J. Hum. Genet.* 88, 788–795.
- Andreux, P.A., Williams, E.G., Koutnikova, H., Houtkooper, R.H., Champy, M.F., Henry, H., Schoonjans, K., Williams, R.W., and Auwerx, J. (2012). Systems genetics of metabolism: the use of the BXD murine reference panel for multiscalar integration of traits. *Cell* 150, 1287–1299.
- Andrews, T.D., Whittle, B., Field, M.A., Balakrishnan, B., Zhang, Y., Shao, Y., Cho, V., Kirk, M., Singh, M., Xia, Y., et al. (2012). Massively parallel sequencing of the mouse exome to accurately identify rare, induced mutations: an immediate source for thousands of new mouse models. *Open Biol.* 2, 120061.
- Ayadi, A., Birling, M.C., Bottomley, J., Bussell, J., Fuchs, H., Fray, M., Gailus-Durner, V., Greenaway, S., Houghton, R., Karp, N., et al. (2012). Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mamm. Genome* 23, 600–610.

- Bearer, E.L., Chen, A.F., Chen, A.H., Li, Z., Mark, H.F., Smith, R.J.H., and Jackson, C.L. (2000). 2E4/Kaptn (KPTN)— a candidate gene for the hearing loss locus, DFNA4. *Ann. Hum. Genet.* 64, 189–196.
- Brown, S.D., and Moore, M.W. (2012). Towards an encyclopaedia of mammalian gene function: the International Mouse Phenotyping Consortium. *Dis. Model. Mech.* 5, 289–292.
- Cason, A.L., Ikeguchi, Y., Skinner, C., Wood, T.C., Holden, K.R., Lubbs, H.A., Martinez, F., Simensen, R.J., Stevenson, R.E., Pegg, A.E., and Schwartz, C.E. (2003). X-linked spermine synthase gene (SMS) defect: the first polyamine deficiency syndrome. *Eur. J. Hum. Genet.* 11, 937–944.
- Conant, G.C., and Wagner, A. (2004). Duplicate genes and robustness to transient gene knock-downs in *Caenorhabditis elegans*. *Proc. Biol. Sci.* 271, 89–96.
- Crossan, G.P., van der Weyden, L., Rosado, I.V., Langevin, F., Gaillard, P.H., McIntyre, R.E., Gallagher, F., Kettunen, M.I., Lewis, D.Y., Brindle, K., et al.; Sanger Mouse Genetics Project. (2011). Disruption of mouse *Slx4*, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nat. Genet.* 43, 147–152.
- Deardorff, M.A., Kaur, M., Yaeger, D., Rampuria, A., Korolev, S., Pie, J., Gil-Rodríguez, C., Arnedo, M., Loeys, B., Kline, A.D., et al. (2007). Mutations in cohesin complex members *SMC3* and *SMC1A* cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. *Am. J. Hum. Genet.* 80, 485–494.
- Edwards, A.M., Isserlin, R., Bader, G.D., Frye, S.V., Willson, T.M., and Yu, F.H. (2011). Too many roads not taken. *Nature* 470, 163–165.
- Fischer, H., Szabo, S., Scherz, J., Jaeger, K., Rossiter, H., Buchberger, M., Ghannadan, M., Hermann, M., Theussl, H.C., Tobin, D.J., et al. (2011). Essential role of the keratinocyte-specific endonuclease *DNase1L2* in the removal of nuclear DNA from hair and nails. *J. Invest. Dermatol.* 131, 1208–1215.
- Fisher, C.L., Pineault, N., Brookes, C., Helgason, C.D., Ohta, H., Bodner, C., Hess, J.L., Humphries, R.K., and Brock, H.W. (2010). Loss-of-function Additional sex combs like 1 mutations disrupt hematopoiesis but do not cause severe myelodysplasia or leukemia. *Blood* 115, 38–46.
- Flint, J., and Eskin, E. (2012). Genome-wide association studies in mice. *Nat. Rev. Genet.* 13, 807–817.
- Forzati, F., Federico, A., Pallante, P., Abbate, A., Esposito, F., Malapelle, U., Sepe, R., Palma, G., Troncone, G., Scarfò, M., et al. (2012). *CBX7* is a tumor suppressor in mice and humans. *J. Clin. Invest.* 122, 612–623.
- Friesema, E.C., Kuiper, G.G., Jansen, J., Visser, T.J., and Kester, M.H. (2006). Thyroid hormone transport by the human monocarboxylate transporter 8 and its rate-limiting role in intracellular metabolism. *Mol. Endocrinol.* 20, 2761–2772.
- Fuchs, H., Gailus-Durner, V., Neschen, S., Adler, T., Afonso, L.C., Aguilar-Pimentel, J.A., Becker, L., Bohla, A., Calzada-Wack, J., Cohrs, C., et al. (2012). Innovations in phenotyping of mouse models in the German Mouse Clinic. *Mamm. Genome* 23, 611–622.
- Gu, Z., Steinmetz, L.M., Gu, X., Scharfe, C., Davis, R.W., and Li, W.H. (2003). Role of duplicate genes in genetic robustness against null mutations. *Nature* 421, 63–66.
- Hawkins, T.E., Roes, J., Rees, D., Monkhouse, J., and Moss, S.E. (1999). Immunological development and cardiovascular function are normal in annexin VI null mutant mice. *Mol. Cell. Biol.* 19, 8028–8032.
- Jimenez-Sanchez, G., Childs, B., and Valle, D. (2001). Human disease genes. *Nature* 409, 853–855.
- Kim, Y.K., Kim, Y.S., Yoo, K.J., Lee, H.J., Lee, D.R., Yeo, C.Y., and Baek, K.H. (2007). The expression of *Usp42* during embryogenesis and spermatogenesis in mouse. *Gene Expr. Patterns* 7, 143–148.
- Laughlin, M.R., Lloyd, K.C., Cline, G.W., and Wasserman, D.H.; Mouse Metabolic Phenotyping Centers Consortium. (2012). NIH Mouse Metabolic Phenotyping Centers: the power of centralized phenotyping. *Mamm. Genome* 23, 623–631.
- Liang, H., and Li, W.H. (2007). Gene essentiality, gene duplicability and protein connectivity in human and mouse. *Trends Genet.* 23, 375–378.
- Liao, B.Y., and Zhang, J. (2007). Mouse duplicate genes are as essential as singletons. *Trends Genet.* 23, 378–381.
- MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B., et al.; 1000 Genomes Project Consortium. (2012). A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335, 823–828.
- Makino, T., Hokamp, K., and McLysaght, A. (2009). The complex relationship of gene duplication and essentiality. *Trends Genet.* 25, 152–155.
- Miller, L.D., Park, K.S., Guo, Q.M., Alkharouf, N.W., Malek, R.L., Lee, N.H., Liu, E.T., and Cheng, S.Y. (2001). Silencing of Wnt signaling and activation of multiple metabolic pathways in response to thyroid hormone-stimulated cell proliferation. *Mol. Cell. Biol.* 21, 6626–6639.
- Mitchell, K.J., Pinson, K.I., Kelly, O.G., Brennan, J., Zupicich, J., Scherz, P., Leighton, P.A., Goodrich, L.V., Lu, X., Avery, B.J., et al. (2001). Functional analysis of secreted and transmembrane proteins critical to mouse development. *Nat. Genet.* 28, 241–249.
- Moreno-De-Luca, A., Helmers, S.L., Mao, H., Burns, T.G., Melton, A.M., Schmidt, K.R., Fernhoff, P.M., Ledbetter, D.H., and Martin, C.L. (2011). Adaptor protein complex-4 (AP-4) deficiency causes a novel autosomal recessive cerebral palsy syndrome with microcephaly and intellectual disability. *J. Med. Genet.* 48, 141–144.
- Nijnik, A., Clare, S., Hale, C., Chen, J., Raisen, C., Mottram, L., Lucas, M., Estabel, J., Ryder, E., Adissu, H., et al.; Sanger Mouse Genetics Project. (2012). The role of sphingosine-1-phosphate transporter *Spns2* in immune system function. *J. Immunol.* 189, 102–111.
- Orozco, L.D., Bennett, B.J., Farber, C.R., Ghazalpour, A., Pan, C., Che, N., Wen, P., Qi, H.X., Mutukulu, A., Siemers, N., et al. (2012). Unraveling inflammatory responses using systems genetics and gene-environment interactions in macrophages. *Cell* 151, 658–670.
- Park, C.Y., Jeker, L.T., Carver-Moore, K., Oh, A., Liu, H.J., Cameron, R., Richards, H., Li, Z., Adler, D., Yoshinaga, Y., et al. (2012). A resource for the conditional ablation of microRNAs in the mouse. *Cell Rep.* 1, 385–391.
- Prosser, H.M., Koike-Yusa, H., Cooper, J.D., Law, F.C., and Bradley, A. (2011). A resource of vectors and ES cells for targeted deletion of microRNAs in mice. *Nat. Biotechnol.* 29, 840–845.
- Reed, D.R., Lawler, M.P., and Tordoff, M.G. (2008). Reduced body weight is a common effect of gene knockout in mice. *BMC Genet.* 9, 4.
- Ruepp, A., Waagele, B., Lechner, M., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Mewes, H.W. (2010). CORUM: the comprehensive resource of mammalian protein complexes—2009. *Nucleic Acids Res.* 38(Database issue), D497–D501.
- Schmalzigaug, R., Rodriguez, R.M., Phillips, L.E., Davidson, C.E., Wetsel, W.C., and Premont, R.T. (2009). Anxiety-like behaviors in mice lacking *GIT2*. *Neurosci. Lett.* 451, 156–161.
- Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T., et al. (2011). A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* 474, 337–342.
- Tang, T., Li, L., Tang, J., Li, Y., Lin, W.Y., Martin, F., Grant, D., Solloway, M., Parker, L., Ye, W., et al. (2010). A mouse knockout library for secreted and transmembrane proteins. *Nat. Biotechnol.* 28, 749–755.
- Testa, G., Schaft, J., van der Hoeven, F., Glaser, S., Anastassiadis, K., Zhang, Y., Hermann, T., Stremmel, W., and Stewart, A.F. (2004). A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *Genesis* 38, 151–158.
- Valenzuela, D.M., Murphy, A.J., Frendewey, D., Gale, N.W., Economides, A.N., Auerbach, W., Poueymirou, W.T., Adams, N.C., Rojas, J., Yasenchak, J., et al. (2003). High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat. Biotechnol.* 21, 652–659.
- Wakana, S., Suzuki, T., Furuse, T., Kobayashi, K., Miura, I., Kaneda, H., Yamada, I., Motegi, H., Toki, H., Inoue, M., et al. (2009). Introduction to the Japan Mouse Clinic at the RIKEN BioResource Center. *Exp. Anim.* 58, 443–450.



## EXTENDED EXPERIMENTAL PROCEDURES

### Online Database

Results can be accessed at <http://www.sanger.ac.uk/mouseportal/>, accompanied by step-by-step examples of how to navigate the data. Alternatively, much of the raw data can be downloaded from the MGP Phenotyping Biomart at <http://www.sanger.ac.uk/htgt/biomart/martview/>. Advice on navigating this Biomart is provided at [ftp://ftp.sanger.ac.uk/pub/mgp/extracting\\_mouse\\_genetic\\_program\\_raw\\_phenotyping\\_data.docx](ftp://ftp.sanger.ac.uk/pub/mgp/extracting_mouse_genetic_program_raw_phenotyping_data.docx). Results have also been summarized in Wikipedia ([http://en.wikipedia.org/wiki/Category:Genes\\_mutated\\_in\\_mice](http://en.wikipedia.org/wiki/Category:Genes_mutated_in_mice)) and Mouse Genome Informatics (<http://www.informatics.jax.org/>).

### Animals

The mutant lines screened are listed in Table S2. We generated most of the mutant lines (242/250) reported in detail here using the EUCOMM/KOMP knockout first conditional-ready targeted ES cell resource on a C57BL/6N background (Skarnes et al., 2011) (Figures S1A and S1B). We maintained the mice on a consistent inbred C57BL/6N background (n = 47 lines), or for early lines on mixed C57BL/6 backgrounds (e.g., 190 lines were maintained on a C57BL/6N;C57BL/6Brd-*Tyr<sup>c-Brd</sup>* background), to minimize variation in screening results due to strain variation and to facilitate comparison across mutant lines. Eight mutant lines with known phenotypes from other sources were included in early screening as positive controls. All lines are available from <http://www.knockoutmouse.org/>; or [mouseinterest@sanger.ac.uk](mailto:mouseinterest@sanger.ac.uk). Mice were maintained ad libitum on Mouse Breeders Diet (LabDiets 5021-3, IPS, Richmond, USA) unless otherwise stated.

### Genotyping and Allele Quality Control

ES cell QC was performed as described (Skarnes et al., 2011). Furthermore, extensive quality control of each allele was performed in mice using a panel of short-range PCR assays (specific for the mutant or wild-type allele, the *lacZ* reporter gene, 5' FRT site or 3' loxP site), quantitative (q) PCR assays (*neo* cassette and loss of wild-type allele counting systems), and long-range (LR) PCR assays (5' and 3' using one primer in the cassette and another outside of the homology arms of the allele design) as summarized in Figure S1C. Further details of the QC protocols are given at <http://www.knockoutmouse.org/kb/25/>.

Typically, mice were genotyped at postnatal day (P)14 using a combination of the three short-range PCR assays or the qPCR *neo* cassette allele-counting assay. Upon completion of phenotyping, genotyping was repeated and data were only accepted from mice for which the second genotype was concordant with the P14 genotype.

A subset of 25 lines, selected because of previously-reported gene expression in liver, was used to assess the degree of knock-down resulting from the targeting event. A TaqMan assay was devised to detect wild-type splicing between the exons on either side of the mutagenic cassette. Samples showing wild-type expression by qRT-PCR were confirmed by end-point RT-PCR and sequencing. ~350ng of total RNA extracted from liver was used in each reaction, performed in triplicate as duplex reactions using the RNA-to-Ct one step kit (Applied Biosystems) with *Gapdh* or *B2m* as endogenous controls and analyzed using a 7900HT qPCR machine with RQ manager software v1.2 (Applied Biosystems).

### Phenotyping Pipeline and Tests

Viability was assessed at P14 and, for those lines classed as lethal or sub-viable at P14, again at E14.5 (Figure 1A). A minimum of 28 genotype-confirmed, live progeny from heterozygous intercrosses were required to assess viability. Based on exact binomial probability calculations, zero homozygotes from 28 progeny gave 95% confidence that the probability of homozygote survival was  $\leq 40\%$ . Lines were classed as homozygous fertile if offspring were born from homozygous parents, regardless of whether the offspring survived to weaning.

At 4 weeks of age, mice undergoing the Clinical Phenotyping Pipeline (Figure 1B) were transferred from Mouse Breeders Diet to a high fat (21.4% fat by crude content; 42% calories provided by fat) dietary challenge (Special Diet Services Western RD 829100, SDS, Witham, UK) for the remainder of the pipeline. This pipeline included the established phenotyping tests described in Table S4. For most tests in this pipeline 7 male and 7 female mutants were used, with the exceptions of the Auditory Brainstem Response (n = 4, independent of gender) and erythrocyte micronuclei (7 males), both deemed to be sufficient with the reduced numbers, and indirect calorimetry (7 males) and biobanking of 41 tissues and organs in paraffin blocks (2 males, 2 females), both limited by operational constraints. If homozygotes were lethal, difficult to produce due to sub-viability or were unsuitable for screening due to welfare concerns, we used heterozygotes for screening. For each genetic background, control cohorts (7M + 7F) were run each week.

A second primary pipeline included challenges with two infectious agents, *Citrobacter rodentium* (8 females) and *Salmonella* Typhimurium (8 males), with matched controls run simultaneously (Table S4). In both challenges we looked at colonization of target tissues at 14 and 28 days post infection. Tissue was biobanked in paraffin and serum taken from the *Salmonella* challenged animals to measure antigen specific IgG (and subclass) antibodies.

### Statistical and Bioinformatic Analysis

For GO term enrichment using TermFinder (Boyle et al., 2004) only high quality experimental evidence codes were included (EXP, IDA, IMP, IPI, IGI, IEP and IC). Positive control lines were excluded from the gene set for this analysis. For GO term enrichment using



FuncAssociate v2.0 (Berriz et al., 2009) (<http://lama.mshri.on.ca/funcassociate/>), the evidence code IEA (Inferred from Electronic Annotation) was excluded. This software used a gene association file downloaded from <ftp.geneontology.org> on 26<sup>th</sup> September 2011. Of the 14139 GO terms queried, 83 (0.59%) were classed as being over-represented and 6 (0.04%) under-represented in our gene set. Revigo was used to reduce the GO term redundancy to give a representative subset of terms (Supek et al., 2011). The gene set was found to be under-represented in only one area: “Sensory perception of smell.” The gene set was found to be over-represented in 20 spread over a variety of processes with no one area dominating.

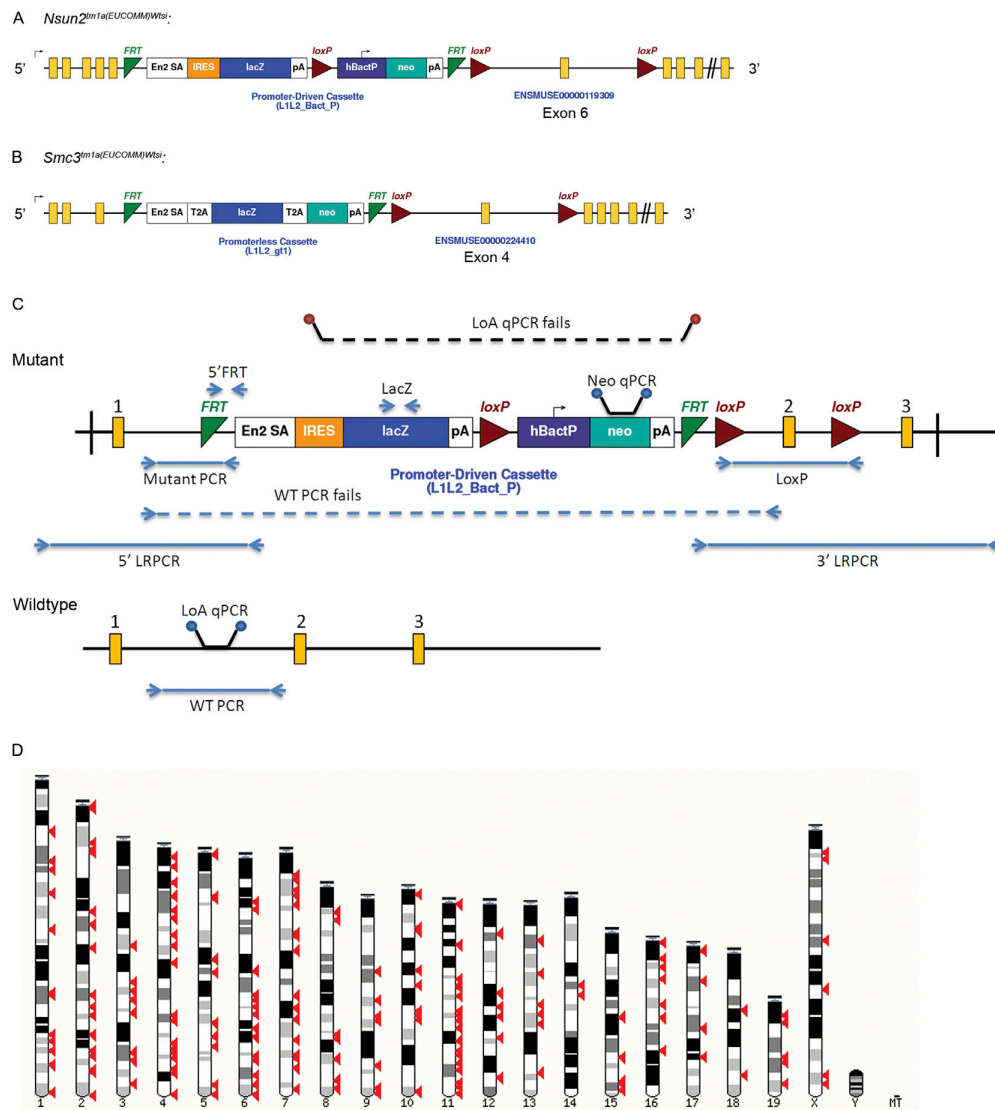
Principal components analysis, performed in SIMCA-P (<http://www.umetrics.com>), was used to assess gene clustering based on variation across the phenotypic hits. The methodology was as default, except that data rescaling was turned off as it was not required for this binary data set. Model predictivity was confirmed by a cross validation procedure (Wold, 1978). The differences between the clusters were investigated using a contribution plot, which showed the differences in scaled units between the highlighted groups for all variables in the model.

Using G\*Power (Version 3.1.3) sensitivity approach (Erdfelder et al., 1996) we calculated the power of our gene set to ask if specific types of protein were more likely to be important for viability and fertility. For a target power of 80%, with the numbers of genes we had in each category, the transcription factor data set needed an effect size (ES) of 45%, the transmembrane proteins needed an ES of 19.3%, the enzyme data set needed an ES of 19% and the chromatin-associated proteins needed an ES of 47% to be detected at the 0.05 threshold.

For protein characterization, amino acid sequences were obtained from Ensembl, release 59, and the longest protein product was chosen as a representative for each gene. Paralog assignments were also obtained from Ensembl (Flicek et al., 2011). Transcription factor annotations were obtained from the DBD database (Wilson et al., 2008) and transmembrane regions were predicted with Phobius (Käll et al., 2004). Involvement in a protein complex was inferred from CORUM, an experimentally-validated data set of human protein complexes (Ruepp et al., 2010).

#### SUPPLEMENTAL REFERENCES

- Berriz, G.F., Beaver, J.E., Cenik, C., Tasan, M., and Roth, F.P. (2009). Next generation software for functional trend analysis. *Bioinformatics* 25, 3043–3044.
- Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20, 3710–3715.
- Erdfelder, E., Faul, F., and Buchner, A. (1996). GPOWER: a general power analysis program. *Behav. Res. Methods Instrum. Comput.* 28, 1–11.
- Flicek, P., Amode, M.R., Barrell, D., Beal, K., Brent, S., Chen, Y., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., et al. (2011). Ensembl 2011. *Nucleic Acids Res.* 39(Database issue), D800–D806.
- Käll, L., Krogh, A., and Sonnhammer, E.L. (2004). A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* 338, 1027–1036.
- Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6, e21800.
- Wilson, D., Charoensawan, V., Kummerfeld, S.K., and Teichmann, S.A. (2008). DBD—taxonomically broad transcription factor predictions: new content and functionality. *Nucleic Acids Res.* 36(Database issue), D88–D92.
- Wold, S. (1978). Cross-validated estimation of the number of components in factor and principal components models. *Technometrics* 20, 397–405.

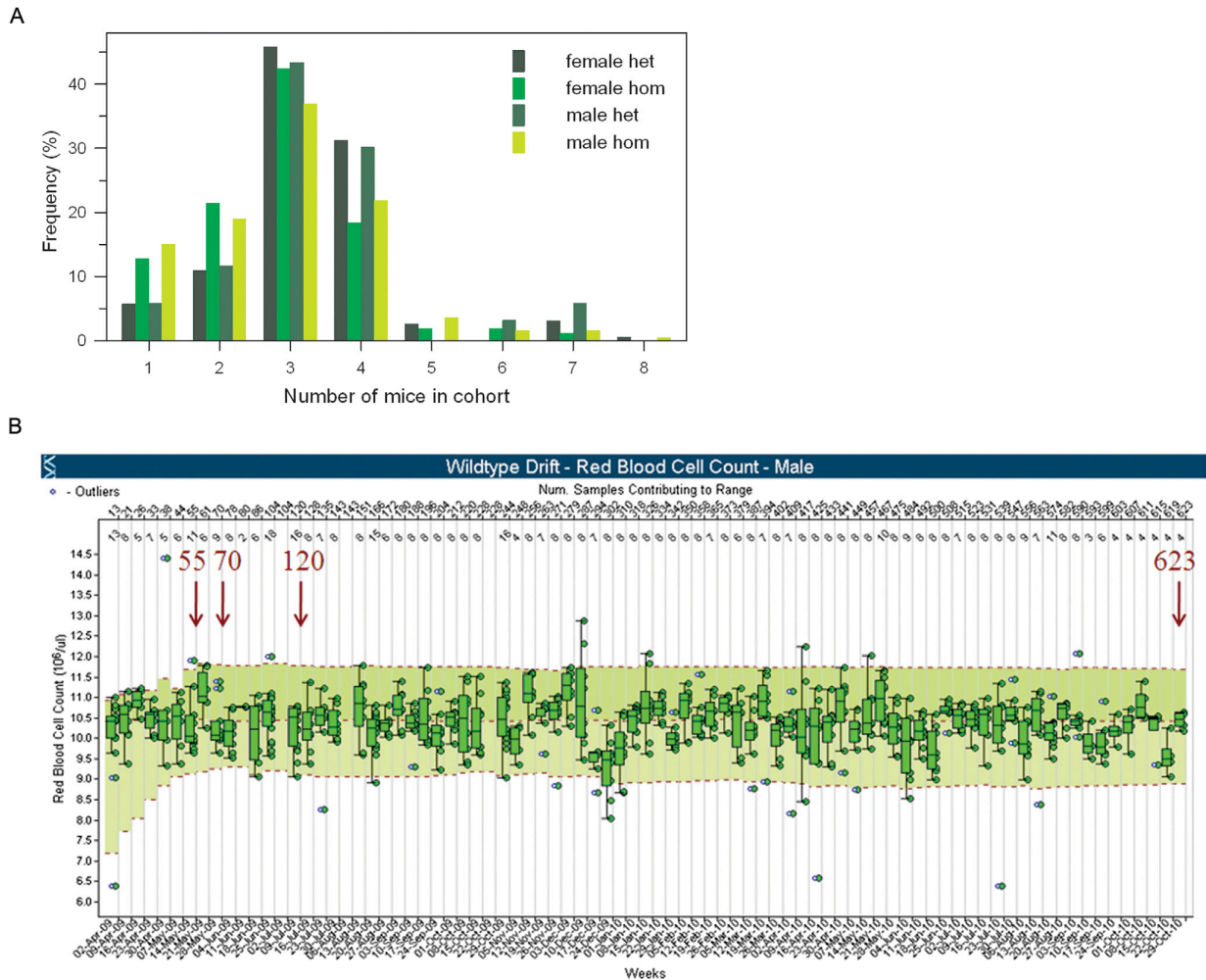


**Figure S1. Allele Design, Genotyping, and Chromosomal Distribution of Genes Selected, Related to Figure 1**

(A and B) Examples of the allele designs used. Illustration of the two main alleles used, A, *Nsun2*<sup>tm1a</sup> contains a promoter-driven targeting vector, and B, *Smc3*<sup>tm1a</sup> contains a promoterless targeting vector [gene build Mouse NCBIM37, (Ensembl 66: Feb 2012)]. The promoterless allele design is biased toward genes that are expressed in ES cells. The alleles are expected to be null alleles, but assessment of the degree of knockdown and the extent of off-target effects on nearby genes has not been carried out systematically.

(C) Genotyping and quality control of mice. *ES cells*: Long-range (LR) PCR, using one primer in the cassette and another outside of the homology arms of the allele design, was used to confirm the targeting on either the 3' or 5' side of the vector prior to micro-injection. *Mice*: To determine the genotype and confirm gene identity, three short-range PCR assays were used: mutant allele-specific, wild-type allele-specific and to detect the *lacZ* gene. Targeting was confirmed by either LRPCR, loss of the wild-type specific short-range PCR product in homozygotes or a qPCR assay confirming loss of the wild-type allele. Presence of the 3' LoxP site was detected by either qPCR or short-range PCR assays. Further details of the QC protocols are available from: <http://www.knockoutmouse.org/kb/25/>. Initially mice were genotyped using a combination of the three short-range PCR assays, but to facilitate high-throughput, we later switched to a qPCR neo cassette counting-based system. Initial genotyping was carried out using ear punches from ~14 day old mice, so that mice of the desired genotypes for screening could be identified and weaned together. Genotyping was repeated at the far-end of the pipeline after culling, and data were only accepted from mice for which the second genotype was concordant with the 14 day genotype.

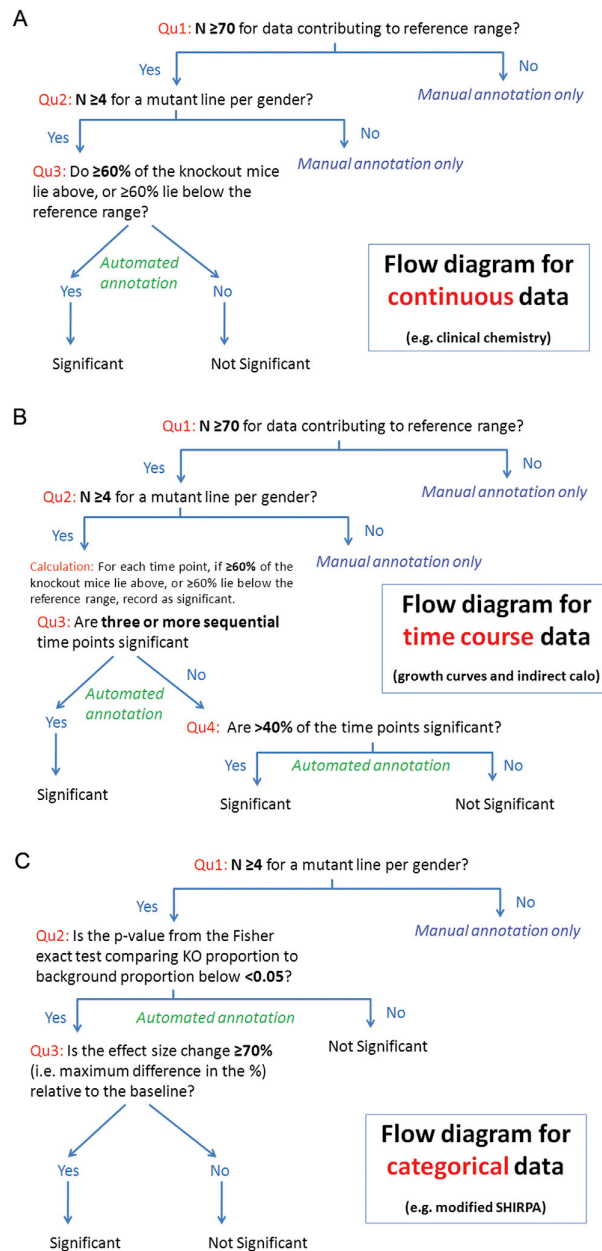
(D) Genomic distribution of genes studied. An illustration of the mouse karyotype showing the location of genes targeted (red arrowheads) across all chromosomes except Y.



**Figure S2. Batch Size and Baseline Variation over Time, Related to Experimental Procedures**

(A) Batch size of mutant mice. Frequency distribution of cohort size of mice of the same genotype issued to the phenotyping pipeline at a time. For each mutant allele, typically 3 mice of a defined sex and zygosity were issued to the Clinical Phenotyping Pipeline at one time. However, the number ranged from 1-8 mice issued in a single batch or cohort.

(B) Baseline variation over time. Example of baseline week to week variation seen in the control data. Example shown is red blood cell count presented weekly from 02/04/09 to 29/10/10 for male mice for the strain group B6Brd;B6Dnk;B6N-Tyr<sup>C-Brd</sup>. Each boxplot represents data collected from control mice in one week. The size of this effect is significant as shown by some of the box plots not overlapping each other, indicating a Cohen's  $d > 3$ . The pale green area indicates the 95% reference range calculated from the 2.5 and 97.5 percentile values as the data accumulate. Red arrows show the cumulative total of animals contributing to the reference range from 55 mice in May 2009 up to 623 mice in October 2010. The reference range becomes stable after about 70 control mice.



**Figure S3. Decision-Making Process for Calling Hits, Related to Experimental Procedures**

The figures show the process we used to call significant hits for three different types of data: (A) continuous, (B) time course and (C) categorical.

## FULL AUTHOR LIST

Jacqueline K White<sup>1</sup>, Anna-Karin Gerdin<sup>1</sup>, Natasha A Karp<sup>1</sup>, Ed Ryder<sup>1</sup>, Marija Buljan<sup>1</sup>, James N Bussell<sup>1</sup>, Jennifer Salisbury<sup>1</sup>, Simon Clare<sup>1</sup>, Neil J Ingham<sup>1</sup>, Christine Podrini<sup>1</sup>, Richard Houghton<sup>1</sup>, Jeanne Estabel<sup>1</sup>, Joanna R Bottomley<sup>1</sup>, David G Melvin<sup>1</sup>, David Sunter<sup>1</sup>, Niels C Adams<sup>1</sup>, The Sanger Institute Mouse Genetics Project<sup>1, 2, 3, 5, 6, \*</sup>, David Tannahill<sup>1</sup>, Darren W Logan<sup>1</sup>, Daniel G MacArthur<sup>1</sup>, Jonathan Flint<sup>2</sup>, Vinit B Mahajan<sup>3</sup>, Stephen H Tsang<sup>4</sup>, Ian Smyth<sup>5</sup>, Fiona M Watt<sup>6</sup>, William C Skarnes<sup>1</sup>, Gordon Dougan<sup>1</sup>, David J Adams<sup>1</sup>, Ramiro Ramirez-Solis<sup>1</sup>, Allan Bradley<sup>1</sup>, and Karen P Steel<sup>1, 7#</sup>

### \*The Sanger Institute Mouse Genetics Project

#### Phenotyping team:

Lauren Baker<sup>1</sup>, Caroline Barnes<sup>1</sup>, Ryan Beveridge<sup>1</sup>, Emma Cambridge<sup>1</sup>, Damian Carragher<sup>1</sup>, Prabhjoat Chana<sup>1</sup>, Kay Clarke<sup>1</sup>, Yvette Hooks<sup>1</sup>, Natalia Igosheva<sup>1</sup>, Ozama Ismail<sup>1</sup>, Hannah Jackson<sup>1</sup>, Leanne Kane<sup>1</sup>, Rosalind Lacey<sup>1</sup>, David Tino Lafont<sup>1</sup>, Mark Lucas<sup>1</sup>, Simon Maguire<sup>1</sup>, Katherine McGill<sup>1</sup>, Rebecca E McIntyre<sup>1</sup>, Sophie Messenger<sup>1</sup>, Lynda Mottram<sup>1</sup>, Lee Mulderrig<sup>1</sup>, Selina Pearson<sup>1</sup>, Hayley J Protheroe<sup>1</sup>, Laura-Anne Roberson<sup>1</sup>, Grace Salisbury<sup>1</sup>, Mark Sanderson<sup>1</sup>, Daniel Sanger<sup>1</sup>, Carl Shannon<sup>1</sup>, Paul C Thompson<sup>1</sup>, Elizabeth Tuck<sup>1</sup>, Valerie E Vancollie<sup>1</sup>

#### Genotyping and targeting team:

Lisa Brackenbury<sup>1</sup>, Wendy Bushell<sup>1</sup>, Ross Cook<sup>1</sup>, Priya Dalvi<sup>1</sup>, Diane Gleeson<sup>1</sup>, Bishoy Habib<sup>1</sup>, Matt Hardy<sup>1</sup>, Kifayathullah Liakath Ali<sup>1</sup>, Evelina Miklejewska<sup>1</sup>, Stacey Price<sup>1</sup>, Debarati Sethi<sup>1</sup>, Elizabeth Trenchard<sup>1</sup>, Dominique von Schiller<sup>1</sup>, Sapna Vyas<sup>1</sup>, Anthony P West<sup>1</sup>, John Woodward<sup>1</sup>, Elizabeth Wynn<sup>1</sup>

#### Mouse Informatics:

Arthur Evans<sup>1</sup>, David Gannon<sup>1</sup>, Mark Griffiths<sup>1</sup>, Simon Holroyd<sup>1</sup>, Vivek Iyer<sup>1</sup>, Christian Kipp<sup>1</sup>, Morag Lewis<sup>1</sup>, Wei Li<sup>1</sup>, Darren Oakley<sup>1</sup>, David Richardson<sup>1</sup>, Damian Smedley<sup>1</sup>

#### Tissue culture team:

Chukwuma Agu<sup>1</sup>, Jackie Bryant<sup>1</sup>, Liz Delaney<sup>1</sup>, Nadia I Gueorguieva<sup>1</sup>, Helen Tharagounet<sup>1</sup>, Anne J Townsend<sup>1</sup>

#### Mouse Production team:

Daniel Biggs<sup>1</sup>, Ellen Brown<sup>1</sup>, Adam Collinson<sup>1</sup>, Charles-Etienne Dumeau<sup>1</sup>, Evelyn Grau<sup>1</sup>, Sarah Harrison<sup>1</sup>, James Harrison<sup>1</sup>, Catherine E Ingle<sup>1</sup>, Helen Kundi<sup>1</sup>, Alla Madich<sup>1</sup>, Danielle Mayhew<sup>1</sup>, Tom Metcalf<sup>1</sup>, Stuart Newman<sup>1</sup>, Johanna Pass<sup>1</sup>, Laila Pearson<sup>1</sup>, Helen Reynolds<sup>1</sup>, Caroline Sinclair<sup>1</sup>, Hannah Wardle-Jones<sup>1</sup>, Michael Woods<sup>1</sup>

#### Mouse Facility team:

Liam Alexander<sup>1</sup>, Terry Brown<sup>1</sup>, Francesca Flack<sup>1</sup>, Carole Frost<sup>1</sup>, Nicola Griggs<sup>1</sup>, Silvia Hrnčiarova<sup>1</sup>, Andrea Kirton<sup>1</sup>, Jordan McDermott<sup>1</sup>, Claire Rogerson<sup>1</sup>, Gemma White<sup>1</sup>, Pawel Zielezinski<sup>1</sup>

#### Collaborators:

Tia DiTommaso<sup>5</sup>, Andrew Edwards<sup>2</sup>, Emma Heath<sup>6</sup>, Mary Ann Mahajan<sup>3</sup>, Binnaz Yalcin<sup>2</sup>

1. Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
2. Wellcome Trust Centre for Human Genetics, Oxford, OX3 7BN, UK
3. Omics Laboratory, University of Iowa, Iowa City, IA 52242, USA

4. Harkness Eye Institute, Columbia University, New York, NY 10032, USA
  5. Monash University, Melbourne, Victoria 3800, Australia
  6. Wellcome Trust Centre for Stem Cell Research, University of Cambridge, Cambridge, CB2 1QR UK
  7. Wolfson Centre for Age-Related Diseases, King's College London, Guy's Campus, London SE1 1UL, UK
- # Corresponding and senior author

# Basement Membranes in Development and Disease

Fenny Wiradjaja, Tia DiTommaso and Ian Smyth\*

Basement membranes (BMs) are specializations of the extracellular matrix that act as key mediators of development and disease. Their sheet like protein matrices typically serve to separate epithelial or endothelial cell layers from underlying mesenchymal tissues, providing both a biophysical support to overlying tissue as well as a hub to promote and regulate cell–cell and cell–protein interactions. In the latter context, the BM is increasingly being recognized as a mediator of growth factor interactions during development. In this review, we discuss recent findings regarding the structure of the BM and its roles in mediating the normal development of the embryo, and we examine congenital diseases affecting the BM which impact embryonic development and health in later life. **Birth Defects Research (Part C) 90:8–31, 2010.** © 2010 Wiley-Liss, Inc.

**Key words:** basement membrane; extracellular matrix; development; disease

a layer of support for overlying tissue, and this is achieved by the ability of components such as laminin and collagen IV to assemble to form supra-molecular structures which to a large extent form the sheet like matrices apparent in EM. However, the interactions of these components are far more complex than a simple polymerization, and so to understand the role of the BM in development and disease it is necessary to understand the building blocks of which it is composed.

## BASEMENT MEMBRANES—COMPONENTS AND FUNCTIONS

Every organ in the body is made up of epithelial or endothelial cell layers which line their cavities and surfaces, and an underlying mesenchymal cell layer, which feeds proliferation and differentiation signals to the epithelium. The tightly packed connected cells of the epithelium are separated from loosely packed mesenchymal cells, by a sheet-like fibrous material known as the basement membrane (BM). As well as providing a structural support for the surrounding tissues, the BM relays functional messages between the two cell layers and acts as a transit for various growth factors. When viewed by electron micro-

copy (EM), typical BMs, such as those of the epidermis, are composed of an electron dense, basal, lamina densa and an overlying lamina lucida; although, in other organs variations exist, as in the kidney, where the lamina densa is sandwiched between the lamina rara interna and externa (Fig. 1A, B). Such elaborations to the basic EM structure are found in almost every organ system, and the composition of the BM itself is also highly divergent depending on its site within the body and on the tissue which it supports. In general terms, BMs are assembled from four primary components; collagen, laminin, heparan sulfate proteoglycans (HSPGs), and nidogen (Timpl, 1996; Timpl and Brown, 1996; LeBleu et al., 2007). One of the primary roles of BMs is to form

## Basic Constituents of the BM Collagens.

The principal component by weight of the BM is collagen IV. In humans, there are six collagen IV genes, *COL4A1–COL4A6*, encoding the  $\alpha 1$ – $\alpha 6$  chains, which are arranged head-to-head in pairs in their chromosomal loci; *COL4A2* and *COL4A1* on chromosome 13, *COL4A4* and *COL4A3* on chromosome 2, and *COL4A6* and *COL4A5* on chromosome X (Soininen et al., 1988; Mariyama et al., 1992; Leinonen et al., 1994; Sugimoto et al., 1994). Each collagen IV  $\alpha$ -chain (~180 kDa) consists of an N-terminal 7S domain (26 kDa), a core collagenous domain (120 kDa) containing the characteristic Gly-X-Y repeats, and a C-terminal noncollagenous NC1 domain. Col-

Fenny Wiradjaja and Tia DiTommaso are from the Department of Biochemistry and Molecular Biology, Monash University, Clayton, Melbourne, Australia  
Ian Smyth is from the Department of Biochemistry and Molecular Biology and Department of Anatomy and Developmental Biology, Monash University, Clayton, Melbourne, Australia

\*Correspondence to: Ian M. Smyth, Building 76/3 STRIP Complex, Monash University, Clayton Campus, Wellington Rd, Clayton, Melbourne 3800, Australia. E-mail: [REDACTED]

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/bdrc.20172







effect between the two molecules (Ylikarppa et al., 2003).

Other collagenous elements are present in different tissue and developmental contexts. The epithelial collagens XVII and VII are found mostly in the BM of the dermal-epidermal junction (DEJ) and BM lining the mucous membrane of the orifices. Collagen XVII is a homotrimer of  $\alpha 1$  chains consisting of a short N-terminal cytoplasmic domain, ADAM (a disintegrin and metalloproteinase) protease-cleavage site, a helical collagenous ectodomain interspersed with non-collagenous motifs (NC1-16) and a C-terminal tail that inserts into the lamina densa and binds to laminin-332 (Fig. 1C). The N-terminal domain interacts with integrin  $\beta 4$ , BP230, and plectin (Hopkinson et al., 1998; Hopkinson and Jones, 2000; Koster et al., 2003; Litjens et al., 2003) in the cytoplasmic plaque of the epidermal hemidesmosomes, multiprotein adhesion complexes which attaches epithelial cells to the BM (see later discussion about skin disease). Collagen VII is the longest collagen and a key epidermal BM constituent, comprising three  $\alpha 1$  chains which assemble to form an N-terminal noncollagenous domain, NC1, a triple collagenous helix, and a C-terminal NC2 domain (Fig. 1C; Burgeson, 1993). The Gly-X-Y motif of the helical domain is often interrupted at evolutionarily conserved intervals (Li et al., 1993). ColVII undergoes proteolytic cleavage by bone morphogenetic protein (BMP)-1 to generate its mature form (Rattenholl et al., 2002), which binds collagen IV and the  $\beta 3$  and  $\gamma 2$  chains of laminin-332 in the skin, and associates with dermal collagen networks to anchor BM.

### Laminins

The most abundant noncollagenous protein in the BM is laminin (see Miner and Yurchenco, 2004; Durbeej, 2009 for excellent reviews). In humans, there are 11 laminin genes (*LAMA1-5*, *LAMB1-3*, and *LAMC1,2*), which encode  $\alpha$ ,  $\beta$ , and  $\gamma$  laminin chains. Each iso-

form of laminin exists in a predominantly cruciform heterotrimeric  $\alpha$ - $\beta$ - $\gamma$  form,  $\sim 900$  kDa in size (Fig. 1E), of which there are 16 trimeric combinations (Tzu and Marinkovich, 2008). By convention, these trimers are expressed as Laminin "XYZ," where X, Y, and Z describe the constituent  $\alpha$ ,  $\beta$ , and  $\gamma$  chains, respectively. All 16 laminin heterotrimer combinations have been detected in BM, and are often expressed in a tissue-specific manner. All three laminin chains contain C-terminal domains I and II, and an N-terminal short arm composed of rod-like domains III (LEb) and V (LEa), and globular domains IV (LF, L4) and VI (Fig. 1E). The C-terminus of the  $\alpha$ -chain includes an extra motif used for cell binding, consisting of five homologous globular motifs, LG1-5, and collectively referred to as the G domain (Tzu and Marinkovich, 2008). Laminin self assembles via the N-terminal globular domains by noncovalent interactions (Odenthal et al., 2004). The C-terminal domains I and II of each chain are coiled around one another in a helical conformation similar to that of collagen, forming the long arm of the molecule, with the N-termini emerging from the coiled-coil domain like a three-pronged fork (cruciform, see Fig. 1E) (Yurchenco and Cheng, 1994). Like collagen IV, many of the laminins are capable of assembling into higher order structures, which form a key structural element of BMs. This is mediated to a large extent by the N-terminal domain VI, also called the LN module, which is required for  $\text{Ca}^{2+}$ -dependent interaction between the three chains to initiate self-assembly of laminins (Yurchenco et al., 1992). These LN modules serve as a focal interaction for the assembly of higher order laminin networks essential for normal BM function.

As well as self-self interactions, the laminins are responsible for interactions with other BM components as well as with cells on either face of the BM. In the former context, the  $\alpha 1$  long arm has been shown to bind the N-terminus of

the HSPG agrin, a minor BM component required for synaptic transmission at the neuromuscular junction (Kammerer et al., 1999). Domains III and V consist of EGF-like repeats called the LE modules, which appear to function as spacer elements between the two domains, rather than having EGF activity (Tzu and Marinkovich, 2008). One of the LE repeats in domain III of laminin  $\gamma 1$  (LEb) binds nidogen-1 (Poschl et al., 1996), which in turn binds to collagen IV, connecting the laminin and collagen IV supramolecular networks. In the latter context, the crystal structure of the LG5 domain of the  $\alpha 2$  chain reveals a sandwich of 14  $\beta$ -sheet strands with a calcium ion bound to one side of the sandwich surrounded by conserved acidic residues (Tisi et al., 2000). This domain is responsible for  $\text{Ca}^{2+}$ -dependent binding to heparin and to  $\alpha$ -dystroglycan, a transmembrane protein that links the BM and cytoskeletal components. In the  $\alpha 2$  chain, this binding is triggered by proteolytic cleavage of the N-terminal propeptide in the LG3 motif (Talts et al., 1999). Interaction of the G domain with the cell surface proteins  $\alpha$ -dystroglycan and  $\beta 1$  integrin facilitates laminin polymerization and disruption of the actin cytoskeleton associated with these receptors also abolished laminin assembly (Colognato et al., 1999). Domain IV of the human  $\gamma 2$  chain is a potential binding site of fibulin-2 (Utani et al., 1997), whereas mouse  $\alpha 1$  IVa binds to fibulin-1. In addition, domain IVa of the  $\alpha 1$  serves as a regulator of cell adhesion as it can mask an RGD domain in the immediately adjacent LE module in domain V, which is the binding site for the cell surface receptor integrin  $\alpha v \beta 3$  (Schulze et al., 1996).

### Nidogens

Nidogen-1 and -2 (also known as entactins) are sulfated glycoproteins composed of three globular-like domains, G1, G2, and G3, composed of nidogen, EGF-like, thyroglobulin type 1, and LDLR

subdomains (Fig. 1F). The proteins are extensively modified by equal proportions of N- and O-linked oligosaccharide chains (Ho et al., 2008). The two nidogen genes are expressed ubiquitously, but not equally, in BMs throughout the body and nidogen-2 is also found in other protein matrices in the skeleton and vasculature (Kohfeldt et al., 1998). One of their key functions is to act as bridging molecules between different components of the BM. Nidogen-1 binds to laminin  $\gamma$ 1 via its G3 domain (Mayer et al., 1993; Ries et al., 2001) and collagen IV via both G2 and G3 domains, and is thought to be involved in bridging these two molecules and their supramolecular networks, as well as for ternary complex formation between laminin and the proteoglycan components of BM via the G3 and G2 domain, respectively (Fox et al., 1991; Aumailley et al., 1993). Null *Nid1* mice display normal BMs (Murshed et al., 2000) and only subtle defects in the eye and vasculature (Dong et al., 2002) and wound healing (Baranowsky et al., 2009), as along with evidence that the usually scant nidogen-2 can compensate for its loss. More profound defects are apparent in animals lacking both genes (see below).

#### *Heparan sulfate proteoglycans*

Perlecan is the most prevalent BM HSPG, and like many other BM components, is a complex multi-domain protein with number of discrete binding partners. It is synthesized and secreted as a single molecule composed of five distinct domains, I-V (Fig. 1G). Perlecan is involved in regulating many signaling pathways and is perhaps best recognized for its HS mediated and core fibroblast growth factor (FGF) binding ability, and consequent roles in angiogenesis and tumor induction (Aviezer et al., 1994; Mongiat et al., 2000). Domain I contains binding sites for heparan sulfate (HS) chains and an SEA (sperm, protein- $\alpha$ -enterokinase-agrin) motif for glycosylation. The HS chain attachment is

important for cell adhesion and BM anchoring via interaction with laminin, collagen IV, and fibronectin; and for growth factor and angiogenesis regulation via interaction with FGF2, PDGF, VEGF, and angiopoietin-3 (Iozzo, 2005). Domain II consists of four LDL (low density lipoprotein) receptor class A repeats and an immunoglobulin (Ig)-like motif. Domain III contains globular modules similar to the laminin domain IV, separated by EGF-like LE (laminin type EGF) repeats, which associate with various growth factors to modulate mitogenic and angiogenic activity (Sasaki et al., 2000). Domain IV is made up entirely of Ig-like repeats and is involved in cell adhesion and motility (Hopf et al., 1999). Domain V consists of three LG motifs similar to laminin  $\alpha$  chains, which interact with surrounding cells via  $\alpha$ 2 $\beta$ 1 integrins on the cell surface, and bind the ECM components heparin, nidogen-1, laminin-nidogen complex, fibulin-2,  $\alpha$ -dystroglycan, ECM-1, and collagen XVIII (Brown et al., 1997; Iozzo, 2005). Domain V binds  $\alpha$ -dystroglycan in vitro with higher affinity than laminin  $\alpha$ 1 and  $\alpha$ 2 chains (Talts et al., 1999). Perlecan colocalizes with  $\alpha$ -dystroglycan in neuromuscular junctions and recruits acetylcholinesterases to terminate synaptic transmission (Peng et al., 1999). Like the other major BM constituent nidogen, a key role for perlecan is to bind to collagen IV and laminin, thus bringing the two supramolecular networks together (Battaglia et al., 1992; Yurchenco and Cheng, 1994).

Various minor HSPG constituents are expressed in the BM of specific tissues, such as agrin, syndecan, and the collagen-HSPG hybrid collagens XVIII and XV previously discussed (LeBleu et al., 2007). Agrin contains an N-terminal NtA domain, a central SEA domain flanked by two Ser/Thr-rich regions (which serve as attachment sites for heparan and chondroitin sulfate chains), and three LG motifs at the C-terminus (Winzen et al., 2003). It forms a network with  $\alpha$ -dystroglycan, laminin, and integrin, linking the BM to

the synaptic cells at neuromuscular junctions to cluster acetylcholine receptors necessary for synaptic transmission (Patthy and Nikolics, 1993). Syndecan is a transmembrane HSPG which connects heparin-binding growth-factor receptors at the cell surface to the actin cytoskeleton. There are four types of syndecans, SDC1-4, and all are type I transmembrane proteins with a large GAG-bearing ectodomain, a transmembrane domain, and a short cytoplasmic tail (Kirn-Safran et al., 2009). Because of their ability to bind growth factor receptors, syndecans can regulate signal transduction. For example, syndecan-1 has been implicated in the phosphorylation of the MAP kinase p44/p42 which is important for cell adhesion and chemotaxis (Pakula et al., 2007).

### THE ROLE OF THE BASEMENT MEMBRANE IN DEVELOPMENT

The role of BMs in regulating differentiation, development, tissue homeostasis and response to infection, injury, and disease are exceedingly complex. In this review, we will focus on the biology of those BM components which play a role in embryonic development and which contribute to inherited diseases, with particular attention paid to genes and proteins which play specific roles in the BM rather than generalized roles in the ECM. A summary of the genes and proteins which form the major constituents of the BM, their expression, developmental roles, mouse knockout phenotypes, and associated human diseases are presented in Table 1. The diversity of these developmental roles and the broad spectrum of diseases associated with dysfunction of the BM stand testament to the importance and specialization of this structure in different tissues and organs.

#### Early Embryonic Development

Cell specification and movement during pre-implantation and peri-implantation development are key

**TABLE 1. Basement Membrane Components in Development and Disease**

Gene	Protein	BM expression	Null phenotype	Human disease
<i>Lama1</i>	Laminin $\alpha 1$	Adult epithelia and embryonic basement membranes	Lethal E6.5-E7.5 Embryos lack Reichardt's Membrane (Miner et al., 2004)	
<i>Lama2</i>	Laminin $\alpha 2$	Muscle, nervous system	Lethal after weaning Muscular dystrophy, neuropathy, defective sperm formation, defective tooth differentiation (Miyagoe et al., 1997; Kuang et al., 1998)	Merosin congenital muscular dystrophy (Heibling-Leclerc et al., 1995)
<i>Lama3</i>	Laminin $\alpha 3$	Stratified epithelia	Lethal shortly after birth Skin blistering (Ryan et al., 1999)	Junctional epidermolysis bullosa (Kivirikko et al., 1995)
<i>Lama4</i>	Laminin $\alpha 4$	Vasculature	Variably penetrant death around birth Defects in vessels, nervous system, and heart (Patton et al., 2001)	Cardiomyopathy (Knoll et al., 2007)
<i>Lama5</i>	Laminin $\alpha 5$	Many embryonic and adult tissues	Lethal before E16.5 Exencephaly, syndactyly, defective lung lobation, defective hair follicle differentiation, defective tooth formation (Miner et al., 1998; Miner and Li, 2000; Nguyen et al., 2002; Li et al., 2003; Fukumoto et al., 2006; Rebustini et al., 2007)	
<i>Lamb1</i>	Laminin $\beta 1$	Ubiquitous	Lethal at E5.5 Absence of Reichardt's and embryonic BM (Miner et al., 2004)	
<i>Lamb2</i>	Laminin $\beta 2$	Ubiquitous	Lethal at P21 Retinal, renal, and neuromuscular defects (Noakes et al., 1995a,b; Libby et al., 1999)	Pierson Syndrome (Zenker et al., 2004a)
<i>Lamb3</i>	Laminin $\beta 3$	Stratified epithelia	Lethal at birth Skin blistering (Kuster et al., 1997)	Junctional epidermolysis bullosa (Pulkkinen et al., 1994b)
<i>Lamc1</i>	Laminin $\gamma 1$	Ubiquitous	Lethal at E5.5 Absence of Reichardt's and embryonic BM (Smyth et al., 1999)	
<i>Lamc2</i>	Laminin $\gamma 2$	Stratified epithelia	Lethal shortly after birth Skin blistering (Meng et al., 2003)	Junctional epidermolysis bullosa (Aberdam et al., 1994; Pulkkinen et al., 1994a)
<i>Lamc3</i> <i>Fras1</i>	Laminin $\gamma$ <i>Fras1</i>	Epithelia Embryonic basement membranes—multiple organs	No phenotype (Denes et al., 2007) Variably penetrant lethality from E14.5	Fraser Syndrome (McGregor et al., 2003)

TABLE 1. (Continued)

Gene	Protein	BM expression	Null phenotype	Human disease
<i>Frem1</i>	Frem1	Embryonic basement membranes—multiple organs	Epidermal blistering, renal a/dysgenesis, lung defects (Phillips, 1970; Petrou et al., 2005) Variably penetrant lethality from E14.5 Epidermal blistering, renal a/dysgenesis, neural tube closure defects (Varnum and Fox, 1981) Variably penetrant lethality from E14.5 Epidermal blistering, renal a/dysgenesis, neural tube closure defects, lung defects (Little and Bagg, 1923; Timmer et al., 2005)	BNAR Syndrome (Alazami et al., 2009)
<i>Frem2</i>	Frem2	Embryonic basement membranes—multiple organs		Fraser Syndrome (Jadeja et al., 2005)
<i>Ecm1</i>	Ecm1	Broad	—	Lipoid proteinosis (Hamada et al., 2002)
<i>Nid-1</i> <i>Nid-2</i>	Nidogen-1 Nidogen-2	Ubiquitous Ubiquitous	No phenotype (Murshed et al., 2000) No phenotype (Schymeinsky et al., 2002). In combination with <i>Nid1</i> —neuromuscular defects, syndactyly, delayed lung development, cardiovascular defects (Bader et al., 2005)	
<i>Hspg2</i>	Perlecan	Ubiquitous	Lethal at E11.5 Cardiac/outflow tract defects, skeletal dysgenesis, exencephaly, microcephaly (Arikawa-Hirasawa et al., 1999; Costell et al., 1999, 2002)	Schwartz-Jampel Syndrome (Nicole et al., 2000), Dystrophic Dysplasia Silverman Handmaker (Arikawa-Hirasawa et al., 2001a) HANAC Syndrome (Plaisier et al., 2007), Porencephaly (Gould et al., 2005)
<i>Col4a1</i>	Collagen IV $\alpha 1$	Ubiquitous	Eye defects, renal disease, stroke, porencephaly (Poschl et al., 2004; Van Agtmael et al., 2005)	
<i>Col4a2</i>	Collagen IV $\alpha 2$	Ubiquitous	Eye defects, hemorrhage, porencephaly (Favor et al., 2007)	
<i>Col4a3</i>	Collagen IV $\alpha 3$	All BMs, late acquisition during development, multiple organs	Nephropathy (Beirowski et al., 2006; Abrahamson et al., 2007)	Alport's Syndrome (Lemmink et al., 1994), Thin Basement Membrane Nephropathy (Hudson et al., 2003)
<i>Col4a4</i>	Collagen IV $\alpha 4$	All BMs, late acquisition during development, multiple organs		Thin Basement Membrane Nephropathy (Gubler, 2008)

TABLE 1. (Continued)

Gene	Protein	BM expression	Null phenotype	Human disease
<i>Lama1</i>	Laminin $\alpha 1$	Adult epithelia and embryonic basement membranes	Lethal E6.5-E7.5	
<i>Col4a5</i>	Collagen IV $\alpha 5$	All BMs, late acquisition during development, multiple organs	Nephropathy (Rheault et al., 2004)	X-linked Alport's Syndrome (Barker et al., 1990)
<i>Col7a1</i>	Collagen VII	Skin, kidney, choroid plexus, cornea larynx, breast, prostate, cervix	Neonatal lethality Skin blistering (Heinonen et al., 1999)	Dystrophic epidermolysis bullosa (Christiano et al., 1993; Hilal et al., 1993)
<i>Col15a1</i>	Collagen XV	Capillaries, peripheral nervous system, developmental expression in heart, lung, and kidney	Cardiac and muscle defects (Eklund et al., 2001)	
<i>Col17a1</i>	Collagen XVII	Skin, placenta, bronchus	Neonatal lethality Skin blistering (Nishie et al., 2007)	Junctional epidermolysis bullosa (McGrath et al., 1995a)
<i>Col18</i>	Collagen XVIII	Choroid plexus, skin, liver, kidney, villi	Hydrocephalus, eye defects (Fukai et al., 2002; Utriainen et al., 2004)	Knobloch Syndrome (Sertie et al., 2000)
<i>Agrin</i>	Agrin	Broad, with high expression in nervous system, kidney, lung	Defective synaptogenesis (Gautam et al., 1996)	

BM components are broadly expressed during development and in adult tissues, and their loss leads to a range of mouse phenotypes and human diseases.

requirements for the establishment of cell lineages and during processes such as gastrulation. It is, therefore, not surprising that BMs play a key role in regulating these processes both structurally and in the context of regulating growth factor presentation and activity during cell differentiation. The earliest BMs are established around implantation and components of the BM are expressed very early in embryogenesis. Laminin-1 ( $\alpha 1\beta 1\gamma 1$ ) is the earliest laminin expressed during mouse embryonic development, with  $\beta 1$  and  $\gamma 1$  mRNA detected at 2–4 cell stage, and  $\alpha 1$  chain appearing in the morula (Cooper and MacQueen, 1983; Dziadek and Timpl, 1985). Collagen IV chains are also expressed early in embryogenesis (Adamson and Ayers, 1979). BM structures are associated with two primary cellular interfaces of the early implanted embryo; the notably thick membrane associated with Reichardt's membrane and the embryonic BM which separates the epiblast from the visceral endoderm. Studies in genetically engineered mice suggest that the functional role of many BM proteins at early developmental stages is modest, and that as a consequence relatively simple BMs are permissive for early development. For example, the most abundant collagen IV protomer,  $\alpha 1\alpha 1\alpha 2$ , appears to be dispensable for BM synthesis and assembly during early embryogenesis, as other BM components are deposited at their correct sites up to embryonic day (E) 10.5 in targeted double null  $\alpha 1$  and  $\alpha 2$  mutants (Poschl et al., 2004). Likewise, embryos lacking either nidogen-1 or -2 separately survive to birth (Murshed et al., 2000; Schymeinsky et al., 2002), and those lacking both nidogen-1 and -2 survive early embryogenesis (although they play a role in later development, see below). Similarly, mice carrying mutations in perlecan also survive through implantation and gastrulation (Costell et al., 1999).

The laminins, however, are known to be key regulators of



early cell specification and fetal growth. In mice lacking *Lama5* (which consequently have no laminin-511), both the embryonic BM and Reichardt's membrane form normally, probably because laminin-111 can compensate (Miner et al., 1998). However, mice lacking *Lama1* (which produce laminin-511 but not laminin-111) develop a normal embryonic BM but lack Reichardt's membrane (Miner et al., 2004). As a consequence, the parietal endoderm fails to migrate, leading to embryonic lethality at E6.5. Some phenotypic rescue of this phenotype was reported by overexpression of *Lama5* on a *Lama1* background, suggesting that laminin-511 is able to compensate for the absence of laminin-111. Finally, when *Lamb1* or *Lamc1* are deleted, removing both laminin-511 and -111, embryonic development terminates at E5.5 and neither BM is observed (Smyth et al., 1999). The expression of the laminins and collagen IV during these early developmental stages is thought to be regulated by several key signaling pathways. Significant in vivo and in vitro evidence from embryoid body culture experiments highlight a role for the fibroblast growth factors (Fgfs), which act upstream of *Gata6*, to induce endoderm differentiation by regulating laminin expression (Li et al., 2004). The induction of BM gene expression and the formation of the embryonic BMs also play a key role in the specification and polarization of the epiblast. Loss of *Smad4* results in an expansion of the BM which is associated with endodermal thickening and which potentially interrupts signaling between these cells and the epiblast (Costello et al., 2009).

Studies of the laminins indicate that a simple BM composed of laminins, and in particular laminin-111, is all that are required for early developmental events. It is interesting to note that those laminin isoforms required for early development are also those which are most closely related to the ancestral laminin genes in invertebrates,

where mammalian *Lama1* and 2 are orthologs to *Drosophila*  $\alpha 1,2$  and *C. elegans*  $\alpha A$ , whereas *Lama5* derives from  $\alpha 3,5$  and  $\alpha B$ , respectively (Miner et al., 2004). Loss of these proteins in invertebrates also has profound, but not early, effects on embryogenesis. Mutation of *Drosophila* *Lan A* (which encodes  $\alpha 3,5$ ) results in embryonic lethality as a consequence of abnormal development of the muscles, heart, and endoderm (Henchcliffe et al., 1993; Yarnitzky and Volk, 1995), although early patterning of the mesoderm is not affected. Antero-posterior (A-P) patterning of the oocyte and neuronal migration are also defective in these flies (Garcia-Alonso et al., 1996; Deng and Ruohola-Baker, 2000). Loss of *wing blister* (*wb*), which encodes  $\alpha 1,2$ , also plays late roles in development, as loss leads to wing blistering and defects in vascular, tracheal, and muscle development (Martin et al., 1999). Finally, flies lacking *LanB1*, which encodes the sole  $\gamma$  laminin subunit and whose loss should abrogate the formation of any trimers, develop until the end of embryogenesis. BMs in these flies are absent, and loss is accompanied by a failure to recruit other BM components such as collagen IV and perlecan, and the formation of many organs and tissues is inhibited (Urbano et al., 2009). Mutations in the paralogous worm genes display semi penetrant growth arrest at embryonic or larval stages but relatively little in the way of profound early embryonic defects (Kramer, 2005; Kao et al., 2006). Together, these studies indicate that the early developmental requirement for laminin observed during mammalian development is absent in invertebrates, and that these functions have been co-opted to permit the structural and cell signaling mechanisms necessary at these developmental stages in mammals.

### Organogenesis and Tissue Differentiation

As the organs of the body develop, the embryonic BMs begin

to specialize to acquire their often highly refined adult roles. As a consequence, a number of different programs of organogenesis and embryonic tissue differentiation are affected by loss of certain genes encoding BM proteins. Over the last 10 years, the application of gene targeting and transgenic approaches in the mouse has uncovered a number of different developmental roles for these primary BM proteins.

### Collagens

As the major structural component of the extracellular matrix, collagens play diverse roles in tissue architecture and cellular organization. Despite the fact that collagen IV is the most abundant structural BM component and that its scaffold integrates with a range of other BM proteins, its ablation results in surprisingly late defects during embryogenesis. Homozygous null deletion of both of its  $\alpha$  chains still permits development of the embryo to E9.5, indicating that it is dispensable for the formation of early embryonic BMs. Homozygotes die around E10.5–E11.5 due to general BM deficiencies and a failure of Reichardt's membrane (Poschl et al., 2004). Other BM components in underlying epithelia are present in mutant embryos, albeit at reduced levels. Vascular development in these mice is normal, although subtle differences in the organization of capillary networks suggest that this process might rely on the presence of collagen IV containing BMs. In the epidermis, collagen VII, which makes up anchoring fibrils which link the BM to the papillary dermis, is required for stabilization of this interface, and ablation of *Col7a1* leads to extensive blistering in utero at its site of action below the BM lamina densa (Heinonen et al., 1999) (see following section). Likewise collagen XVII, a transmembrane protein, plays a key role in the stabilization of the dermal-epidermal junction (DEJ) during late gestation, and deletion in the mouse leads to subepidermal blistering (Nishie

et al., 2007). These mice also demonstrate a role for this protein in epithelial–mesenchymal interactions which lead to normal enamel formation in the developing tooth (Asaka et al., 2009). As with other BM components, extraneous factors in the ECM can significantly impact on collagen function in the BM and elsewhere in the body. A prime example is the glycoprotein SPARC, whose expression during embryogenesis is necessary for laminar assembly of BM collagen components (Martinek et al., 2008), but which also plays a much broader role in adult tissue homeostasis (Clark and Sage, 2008).

### Laminins

As well as their key roles in early embryogenesis discussed earlier, the laminins are also important regulators of later development and organogenesis. Perhaps, the best studied in this respect is *Lama5*. Inactivation of the gene results in defects in a number of different developing organs, including the lung, kidney, neural tube, and placenta. *Lama5* null mice die at around E16.5, most probably due to abnormalities in vascular and trophoblast cell differentiation in the placenta (Miner et al., 1998). In the developing limb, *Lama5* plays a role in stabilizing the BM, such that loss results in structural defects allowing extrusion of the underlying limb mesenchyme. These mislocalized cells are then thought to contribute to fusion of neighboring digits and syndactyly. A similar process also appears to be the cause of lung lobulation defects in these mice, highlighting a specific role for laminins in the formation and maintenance of the pleural BM (Nguyen et al., 2002). Later developmental processes in the lung, such as alveolarization, were also found to be affected, with an absence of alveolar cells types and a reduction in the number of alveolar capillaries. Strikingly, these effects were apparent in the absence of ultrastructural deficiencies in the BM, suggesting that cell

signaling mediated by engagement of laminin-511 with cell surface receptors may arbitrate these effects. Unlike the lungs, epithelial morphogenesis in the submandibular glands is altered, suggesting a role for the laminins in integrin signaling via regulation of Fgf receptor expression (Rebustini et al., 2007). Renal agenesis is also a feature of these animals, apparently derived from a failure of the ureteric bud to grow into or engage the metanephric mesenchyme. In *Lama5* mice, the glomeruli, future filtrative component of the kidney, lack BMs, thus affecting the attachment of endothelial cells and podocytes (Miner and Li, 2000). The observation that hypomorphic *Lama5* mutants develop polycystic kidney disease supports an ongoing role of the protein in kidney function (Shannon et al., 2006). BM specialization in the kidney is a key requirement for later organ function, and will be discussed at length later in this review in the context of inherited disease. Neural tube closure is defective in null embryos (Miner et al., 1998) and epidermal BMs are discontinuous in *Lama5* mutant embryos, leading to defects in adhesion and hair follicle induction and growth (Li et al., 2003). Although not completely analogous, loss of the gene leads to defects in fin formation in zebrafish (Webb et al., 2007).

Other laminins also appear to play specific developmental roles. Laminins-211 and -221 are associated with the structural integrity of muscle cells and their attachment to tendons. Loss of *Lama2* leads to muscular dystrophy in mice (Miyagoe et al., 1997), and loss of *Lamb2* contributes to defects in the myotendinous junction as well as further defects in the kidney (Noakes et al., 1995b), neural system (Noakes et al., 1995a; Knight et al., 2003), and the retina (Libby et al., 1999). In the kidney, a switch from *Lamb1* to *Lamb2* occurs in the GBM at the capillary loop stage during glomerular development. Ablation of *Lamb2* consequently results in persistence of *Lamb1* in the GBM,

which is accompanied by development of nephrotic disease (Noakes et al., 1995b). Laminins are also essential for normal stability and development of the epidermis, primarily through the action of laminin-332. Deletion of *lamb3* leads to adhesion defects affecting the DEJ (Kuster et al., 1997), and similar defects are observed in mice lacking *Lama3* (Ryan et al., 1999). The latter animals also display defects in the formation of tooth enamel. Interactions of the laminins with other BM components are also likely to contribute to developmental defects and serve to highlight the cooperative function of many of the BM proteins. For example, branching morphogenesis of the kidneys and lungs are driven in vitro by *Lamb2*–nidogen-1 and *Lamc1*–nidogen-1 interactions (Ekblom et al., 1994; Kadoya et al., 1997), whereas ablation of the nidogen binding site in *Lamc1* results in renal agenesis in the majority of embryos, as a consequence of defective Wolffian duct development and failure of ureteric bud formation (Willem et al., 2002). Defects in lung development affecting the air–blood interface were also observed in these mice.

### Nidogens

Although nidogen-1 and -2 have divergent expression in adult tissues, they display extensive coexpression during embryonic development. Loss of nidogen-1 (Murshed et al., 2000) or nidogen-2 (Schymeinsky et al., 2002) alone does not impact on normal development, possibly as a consequence of redistribution or upregulation of the other gene (Murshed et al., 2000; Miosge et al., 2002); however, combinatorial deletion leads to perinatal lethality in mice (Bader et al., 2005). In combination, the genes are not required for organ specification; however, a number of different fetal malformations ultimately contribute to lethality. In the lung, tissue compaction during late embryogenesis is abnormal in double mutants, indicating restricted differentiation

compared with wild-type littermates, with relatively little saccular differentiation observed. At birth, double null lungs consequently exhibited poor expansion of sacculi, thickened septae and mesenchyme, small alveolar spaces, and evidence of defects in late stage cell differentiation. A proportion (~10%) of double mutant embryos displayed renal aplasia, suggesting a role for the nidogens in ureteric bud outgrowth. In those kidneys which formed a range of variably penetrant phenotypes including cyst formation, cortical and medullary size differences and hydronephrosis were observed. These mice also display abnormalities of the cardiac BM unrelated to the protein interactions with *Lamc1* (Willem et al., 2002). As with *Lama5* mutants, combinatorial loss of nidogen-1 and -2 leads to defects in limb development (Bose et al., 2006), where both proteins are required to maintain the BM underlying the limb apical ectodermal ridge. Their loss leads to subsequent mesenchymal extrusion through the BM, resulting in syndactyly. Although the dermal-epidermal junction of the skin is normal, stabilization of capillary BMs in the dermis is dramatically affected (Mokkapati et al., 2008).

### *Perlecans*

Perlecan is a scaffold protein capable of interacting with a broad range of partner proteins in a developmental context, both through its core protein, and through its heparan sulfate modifications. In particular, the molecule can act to regulate the bioavailability of Fgfs in the extracellular environment (reviewed in Ornitz, 2000). Inactivation of perlecan in the mouse predominantly leads to embryonic death at E11.5 and mutant embryos are characterized by a broad range of abnormalities in the heart, brain, kidney, and skeleton (Arikawa-Hirasawa et al., 1999; Costell et al., 1999). Perlecan null embryos exhibit significant levels of pericardial hemorrhage (Costell et al., 1999), as

well as later developmental defects in the cardiac outflow tract (Costell et al., 2002; Gonzalez-Iriarte et al., 2003). Null hearts have a generalized reduction in BM components, lack BMs in cardiomyocytes, and display mechanical instability of the heart wall (Sasse et al., 2008). Those mice that survive these cardiac defects die shortly after birth, probably as a result of a range of skeletal abnormalities affecting the ribs (Arikawa-Hirasawa et al., 1999). As with *Lama5* mutants, exencephaly is a feature of roughly half of null embryos. The remaining mice exhibit microcephaly as a consequence of thin cerebral walls and abnormal ganglionic eminences. In brains with intact BMs, but which lack perlecan, cell division in the forebrain is reduced and neurogenesis is abnormal, suggesting that the protein regulated growth factor activity in the neuroepithelium (Giros et al., 2007). Knockdown of the gene in zebrafish leads to severe defects in the development of the cardiovascular system and an unexpected myopathic phenotype (Zoeller et al., 2008), whereas inactivation of the *Drosophila* perlecan homolog impacts on Fgf and sonic hedgehog (Shh) signaling in the developing CNS (Park et al., 2003).

### *Summary*

It remains unclear as to the extent to which many of these BM components contribute to postnatal disease. Although many are known to play a role in organogenesis (*Lama5*, collagen IV, and others), the degree to which overt postnatal phenotypes in surviving mice are caused by developmental, rather than homeostatic functions remains to be elucidated. In the face of increasing evidence that the BM plays an active and dynamic role in cell signaling, it is possible that some of these defects may be a consequence of altered cell differentiation during embryonic development. Similarly, the severity of mutations in some BM components, noticeably the laminins, implies that other devel-

opmental requirements may be obscured. The limited double knockout experiments so far undertaken to examine interactions of BM proteins *in vivo* have illuminated a number of developmental roles for the BM. In future, conditional and combinatorial knockout approaches to analyzing BM protein function will likely uncover novel roles for these proteins during organogenesis and development.

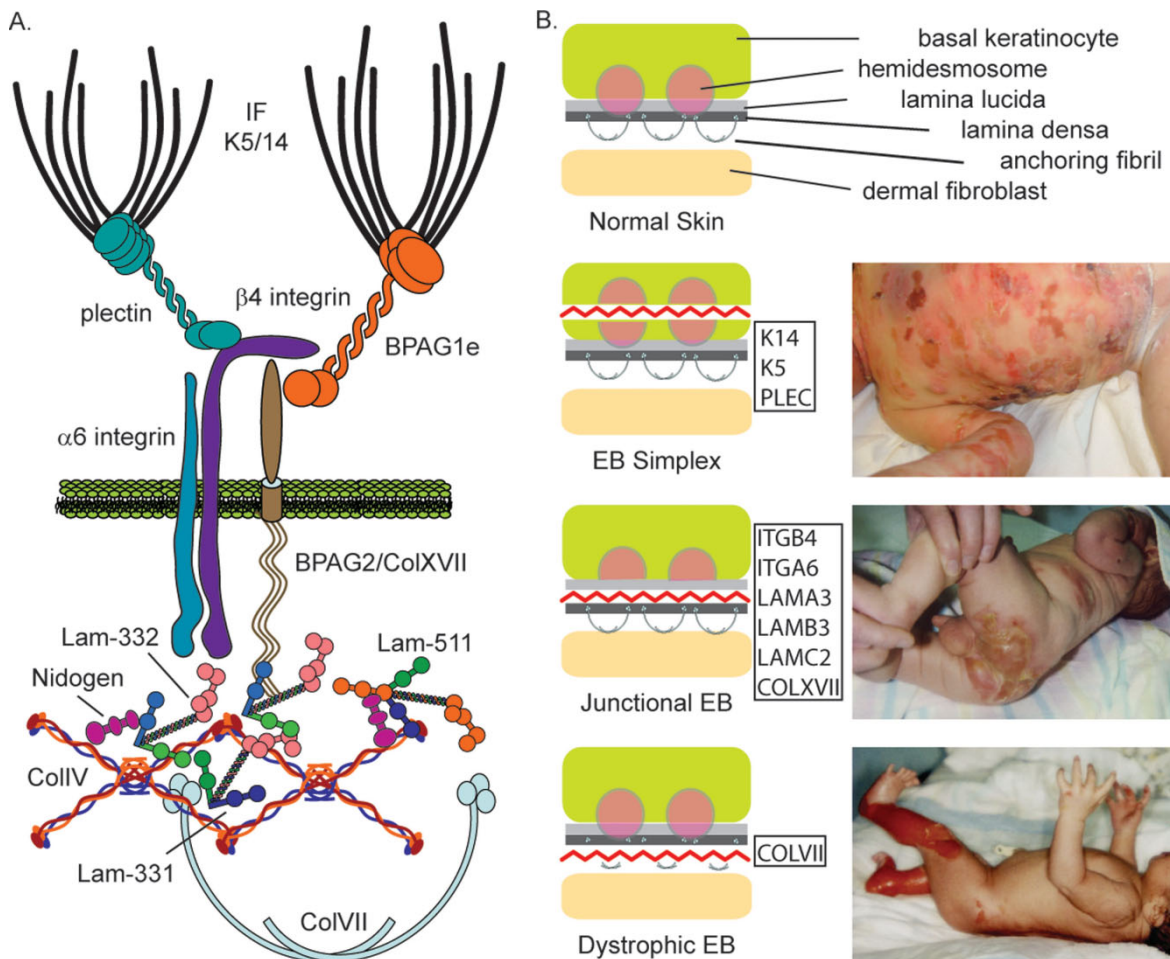
## THE ROLE OF THE BASEMENT MEMBRANE IN DISEASE

The importance of the BM in the development of disease has been highlighted by the description of numerous inherited conditions in which mutations affect the composition or function of the BM. For a summary of these conditions refer to Table 1.

### *The Skin*

In the skin, the epidermis is separated from the underlying dermis by a BM consisting of, among other proteins, laminin-511, laminin-332, collagen IV  $\alpha1\alpha2-\alpha5\alpha6$ , nidogen-1 and -2, and perlecan (McMillan et al., 2003; Marionnet et al., 2006; Sugawara et al., 2008) (Fig. 2A). The skin is anchored to the underlying dermis by a specialized protein complex called the hemidesmosome, which serves to link the intermediate filament (IF) network of the basal keratinocyte through the BM itself. Keratin 5/14 IF networks link to this complex via plectin and BPAG1 (Ruhrberg and Watt, 1997), which in turn associate with transmembrane proteins integrin  $\alpha6\beta4$  and collagen XVII (Breitkreutz et al., 2009). The anchoring filament component, laminin-332, interacts with other laminins, the transmembrane collagen XVII (Tasanen et al., 2004), and integrin  $\alpha6\beta4$  (Sonnenberg et al., 1993), acting as a key linker between the keratinocyte and the BM. The protein also associates with collagen VII, nidogen, and fibulin (Sasaki et al., 2001).





**Figure 2.** The dermal/epidermal junction and epidermolysis bullosa. The dermal epidermal junction links a complex of intermediate filament, linker, and transmembrane proteins with the underlying basement membrane (**A**). The keratin 5/keratin 14 rich intermediate filament (IF) network of the basal keratinocyte is linked via plectin/BPAG1e molecules to the transmembrane integrin  $\alpha 6$ , integrin  $\beta 4$ , and collagen XVII/BPAG2 proteins. The contact between the underlying basement membrane is supported through laminin-332 which in turn associates with itself and other laminins to form a multimeric protein scaffold. Laminin-332 cannot self polymerize and, therefore, associates with laminin-331 to further integrate into laminin networks. Laminin-332 further associates with the collagen IV based protein network in the lamina densa by crosslinking through the  $\gamma 2$  chain to nidogen molecules. Collagen VII, which makes up the sub lamina densa anchoring fibrils, associates with the underlying dermal fibroblasts and with laminin-332 via its NC domain. Mutations in different proteins of this interface result in different forms of the skin blistering disease Epidermolysis Bullosa (EB) (**B**). Blistering in EB in different layers of the DEJ junction depends on the site of action of the given protein. Schematics of the DEJ interface, the plane of blistering (red line), and the genes mutated in different types of EB are noted. The panel of patient phenotypes is, from top; EB simplex, junctional EB, and a newborn with aplasia cutis on the legs with a localized form of recessive dystrophic EB. Patient images were provided by Prof. Dedee Murrell.

In particular, the association with nidogen links the laminin network with the collagen IV network, and the association with collagen VII (Rousselle et al., 1997) is also important in establishing links with the dermis (Litjens et al., 2006).

#### Epidermolysis bullosa

The importance of BM protein-protein interactions to epidermal

integrity is highlighted by the defects apparent in a group of genetic disorders known as epidermolysis bullosa (EB). EB is characterized by bullous lesions (blisters) which form spontaneously or in response to minor mechanical surface trauma, and which are usually present at birth. Mutations in genes expressed in the cutaneous BM which encode for structural proteins in the skin are responsible

for EB disease phenotypes. To date, over 1000 different mutations across 13 different genes have been associated with EB (Fine et al., 2008; Arin, 2009). The EB disorders are classified into four major subtypes: EB simplex, junctional EB, recessive dystrophic EB, and dominant dystrophic EB (Fine et al., 2008). The classification of these major EB subtypes depends to a large extent on the

plane of tissue rupture. Intraepidermal loss of tissue integrity characterizes epidermolysis bullosa simplex (EBS), which is further separated into suprabasal and basal subtypes. Junctional and dystrophic forms of EB are characterized by skin cleavage at the dermal-epidermal interface (intralamina lucida) and in the upper dermis (sublamina densa), respectively. The degree of skin fragility varies depending on the type of EB, as well as the underlying molecular defects.

**Epidermolysis bullosa simplex.** EBS was the first skin fragility disorder shown to be caused by mutations in a gene encoding an intermediate filament (IF) protein (Coulombe et al., 1991; Lane et al., 1992). Accounting for more than half of all EB cases, EBS is the most common EB subtype and is clinically characterized by non-scarring blisters caused by minor frictional trauma. Although not strictly a lesion in the BM, we include this condition to highlight the broader impact that BM specialization in different tissues can have on disease development in those tissues. In EBS, basal keratinocytes that are subjected to minor mechanical stress or trauma rupture, and subsequent intraepidermal blistering is seen. EBS presents as the least severe of the EB subtypes, though severity depends on the nature and position of the inherited mutation (Letai et al., 1993; Pfendner et al., 2005; Has and Bruckner-Tuderman, 2006b). Most cases of EBS are caused by mutations in keratin 14 (K14) or keratin 5 (K5), the type I and type II IF proteins (Fuchs et al., 1981; Fuchs, 1995), which provide mechanical stability and structural support to basal keratinocytes. They comprise a central  $\alpha$  helical rod domain, consisting of four  $\alpha$  helical coils, 1A, 1B, 2A, and 2B, separated by nonhelical linker domains, L1, L12, and L2. Highly conserved helix initiation and helix termination peptides flank the  $\alpha$  helical rod and are involved in keratin intermediate filament interactions/assembly and

filament elongation (Steinert et al., 1993). By insertion at desmosomes and hemidesmosomes, keratin filaments play a role in cytoskeletal scaffolding, cell-cell stability, and attachment to connective tissue (Roop, 1995; Arin, 2009) (Fig. 2A). Not all EBS cases are linked to mutations in K14/5 genes. EBS with muscular dystrophy (EBS-MD) is caused by mutations in plectin (PLEC1), which contribute to epidermal stability by regulating keratin IF insertion into the inner hemidesmosomal plaque (McLean et al., 1996; Has and Bruckner-Tuderman, 2006b). Mutations in plectin have been associated with lethal EBS, characterized by in utero blistering and other developmental abnormalities (Charlesworth et al., 2003).

**Dystrophic epidermolysis bullosa.** Dystrophic epidermolysis bullosa (DEB) is characterized by loss of skin integrity and blistering below the level of the BM. All forms of DEB are caused by a defect in collagen VII (Christiano et al., 1993; Hilal et al., 1993), which is synthesized by fibroblasts and keratinocytes (Ryynanen et al., 1992) and which interacts with dermal collagen bundles as well as with laminin-332 in the hemidesmosome (Litjens et al., 2006). Collagen VII is the longest collagen and is made up of three  $\alpha$ 1 VII chains, each containing a central, triple-helical collagenous domain (Burgeson, 1993), consisting of repeating and often interrupted Gly-X-Y sequences (Fig. 1C). The position of the interruptions in the collagenous domains is evolutionarily conserved (Li et al., 1993). Collagen VII is the main component of anchoring fibrils and its interaction with ECM proteins is critical to establish stable epidermal-dermal cohesion. Anchoring fibrils extend from the lamina densa of the BM into the underlying dermis. In DEB, mutations in the collagen VII gene, *COL7A1*, lead to loss of collagen VII function. Dominant negative dystrophic EB is elicited by mutations resulting in glycine substitutions within the triple helical domain.

The position of the glycine substitution determines the subsequent effect, with some but not all glycine substitutions resulting in severe phenotypes (Bruckner-Tuderman et al., 1999). Although mutations resulting in DEB are difficult to detect and assess (Has and Bruckner-Tuderman, 2006b), making prenatal diagnosis difficult, there is evidence that normal fibroblasts can act as a source of therapeutic collagen VII (Fritsch et al., 2008). Recessive dystrophic epidermolysis bullosa (RDEB) is caused by mutations in *COL71A* which often encode premature termination codons, resulting in truncated collagen VII (Varki et al., 2007). Patients with RDEB have an increased risk of squamous cell carcinoma, with 70% of RDEB cases leading to squamous cell carcinomas by age 45 (Fine et al., 2009). Polymorphisms in the *MMP1* gene are also thought to act as a modifier in regulating disease severity (Titeux et al., 2008). The connection between collagen type VII deficiency, increased collagenase, and squamous cell carcinomas (SCC) is not fully understood, and debate remains open regarding the effect of the loss of BM collagen VII in cutaneous SCC. Recently, a study using siRNAs to knockdown collagen VII expression in cutaneous SCC has shown that loss of collagen VII in the BM promotes invasive behavior in SCC, indicating that *COL7A1* may play a tumor suppressing role in skin (Martins et al., 2009).

**Junctional epidermolysis bullosa.** Junctional epidermolysis bullosa (JEB) is characterized by mechanically induced blistering and cleavage at the dermo-epidermal junction. Various levels of blistering severity have been described in JEB. The JEB phenotype can be attributed to mutations in six genes: *COL17A1*, *LAMA3*, *LAMB3*, *LAMC2*, *ITGA6*, and *ITGB4* (McGrath et al., 1995c; Ashton et al., 2001b; Nakano et al., 2002c; Bauer and Lanschuetzer, 2003b; Muhle et al., 2005a; Varki et al., 2006c), which are all con-

stituent components of the hemidesmosome or anchoring fibrils.

There are three subtypes of JEB: JEB-Herlitz, JEB non-Herlitz, and JEB with pyloric atresia. Herlitz JEB (JEB-H) is caused by mutations in genes encoding the three chains of laminin 332. As we have previously discussed in this review, laminins play a crucial role in tissue integrity and cell adhesion to the BM. In particular, mutations in *LAMA3*, *LAMB3*, or *LAMC2* can lead to dramatically reduced or absent laminin 332, causing subsequent fragility, blistering, and extensive detachment of the epidermis seen in JEB (Aberdam et al., 1994; Pulkkinen et al., 1994a; Kivirikko et al., 1995). Recessive JEB-H is the most severe subtype of the EB and is lethal in infancy, causing death in the first year of life (Muhle et al., 2005b). The mean life span for individuals suffering from JEB-H is ~5 months. Morbidity is due to both cutaneous and extracutaneous complications of this disease, which can include corneal and conjunctival disorders, upper respiratory tract complications, gastrointestinal symptoms, urethral meatus stenosis, and various intraoral diseases. The non-Herlitz subtype of JEB (JEB-nH) has a far more favorable prognosis than that of JEB-H (Yancey and Hintner, 2010). Like JEB-H, JEB-nH can be caused by mutations in *LAMA3*, *LAMB3*, and *LAMC2* (Nakano et al., 2002a; Varki et al., 2006b); however, JEB-nH can also result from mutations in *COL17A1* (McGrath et al., 1995b). In the latter case, these mutations predominantly lead to premature termination codons and subsequent absence of collagen XVII in the BM (Nakano et al., 2002b; Bauer and Lanschuetzer, 2003a; Varki et al., 2006a). There is a broad range in severity and onset of clinical phenotypes associated with JEB-nH but generally JEB-nH is a life-long disease characterized by skin blistering, skin atrophy, nail and dental dysmorphology, and alopecia; with onset ranging from infancy to young adulthood. Finally, mutations *ITGA6* and

*ITGB4* encoding for the  $\alpha 6\beta 4$  integrin subunits are responsible for the subtype of JEB with pyloric atresia (JEB-PA). The relationship between the nature of the mutation and the clinical phenotype is more difficult to ascertain in JEB-PA. However, a majority of mutations in  $\alpha 6\beta 4$  integrin involve the  $\beta 4$  subunit (Ashton et al., 2001a). Forms of JEB-PA with mild skin involvement have favorable prognoses after surgery to correct the pyloric atresia (Has and Bruckner-Tuderman, 2006a).

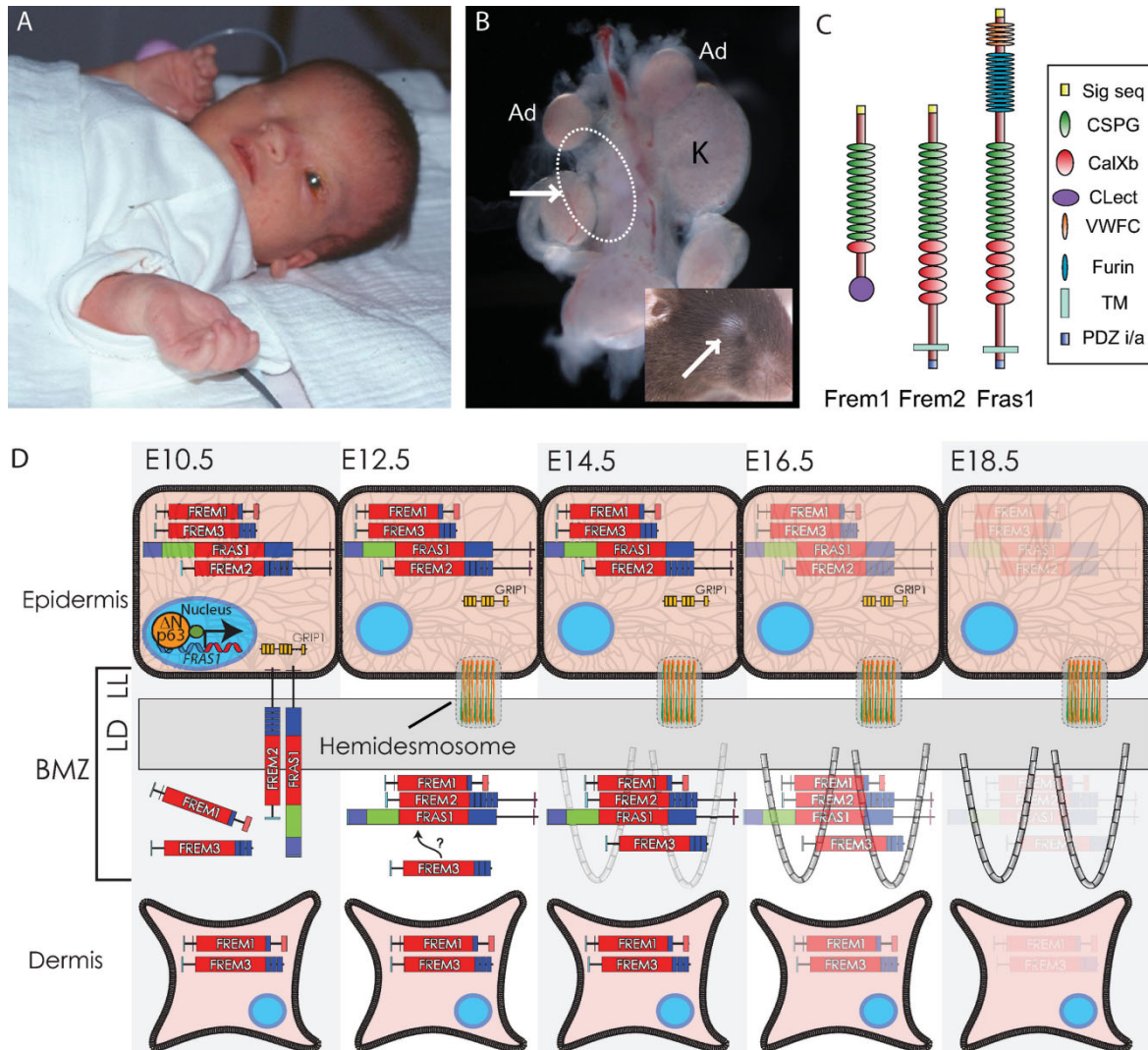
### Fraser syndrome

Fraser Syndrome (FS) is a rare autosomal recessive disorder characterized by loss of embryonic epidermal adhesion specifically during midgestation, defects in renal development, and a range of other developmental phenotypes affecting many organ systems in the body (Slavotinek and Tiffet, 2002; van Haelst et al., 2008). Cryptophthalmia, in which skin covers the globe of the eye, is a key diagnostic feature (Fig. 3A). The defects in adhesion of the epidermis in utero are similar to those observed in DEB, with blister cavities forming below the BM lamina densa. A key to understanding Fraser Syndrome has been the family of blebs mouse mutants which phenocopy the disease in many different organ systems (Fig. 3B) (Smyth and Scambler, 2005). These mice called, *blebbed* (Phillips, 1970), *head blebs* (Varnum and Fox, 1981), *eye blebs* (Chapman and Hummel, 1963), and *myelencephalic blebs* (Little and Bagg, 1923), have provided key insights into the causes and biology of the disease. These mice develop extensive blisters on the limbs and sides of the head around E11, which resolve before birth. Unlike EB patients and their mouse models, blebs mice have no overt postnatal defects in epidermal stability.

The blebs mutants were first proposed as models of FS in the early 1990s (Winter, 1990; Darling and Gossler, 1994), and positional cloning in both the mouse and

human identified mutations in the *Fras1* gene as the cause of the *blebbed* mouse and a significant proportion of Fraser Syndrome cases (McGregor et al., 2003; Vrontou et al., 2003). *Fras1* encodes a large epidermally expressed transmembrane domain protein with core chondroitin sulfate proteoglycan (CSPG), calcium exchange ( $\text{Calx}\beta$ ), furin, and von Willebrand factor C (vWFC) domains (Fig. 3C). Despite its transmembrane domain, *Fras1* localizes to the sub lamina densa and in particular to anchoring fibrils in the dermis, suggesting that the ectodomain of the protein is cleaved (Dalezios et al., 2007). Further insights into the role of *Fras1* have been provided by the cloning of the other blebs mutants. The myelencephalic blebs mutation affects *Frem2*, an epidermally expressed transmembrane protein containing CSPG and  $\text{Calx}\beta$  domains, whose ectodomain is also cleaved and which localizes to the sub lamina densa (Petrou et al., 2007b). A small proportion of FS mutations are found in *FREM2*, but they are missense as opposed to the primarily truncating mutations in *FRAS1* (Jadeja et al., 2005). The mouse head blebs phenotype is caused by mutations in *Frem1* (Smyth et al., 2004), a related mesenchymally expressed, secreted, BM localized protein lacking the vWF and Furin domains and associated with an N-terminal lectin motif. *Frem1* engages integrins  $\alpha v$  and  $\alpha 8$  through an N-terminal RGD domain (Kiyozumi et al., 2005), and localizes to the fibrillar structures in the sub lamina densa but is also detected in the lamina lucida (Petrou et al., 2007b). Interestingly, mutations in humans *FREM1* do not cause a "classic" Fraser Syndrome-like phenotype, despite the overt similarity with the other blebs mice. Instead mutations in the gene have recently been reported in kindreds presenting with bifid nose, renal agenesis, and ano-rectal malformations (Alazami et al., 2009). The final blebs mutant, eye blebs, is not a BM gene, but instead encodes a multi-PDZ domain protein which plays a key role in localizing *Fras1*, and probably *Frem2*, to the cell surface (Takamiya et al., 2004).





**Figure 3.** The Fras/Frem proteins in epithelial adhesion. Mutations in either FREM2 or FRAS1 cause Fraser Syndrome, in which embryonic epidermal blistering gives rise to prominent defects in the epidermis [cryophthalmia, (A)] and a range of other developmental defects. The blebs family of mouse mutants also carry mutations in these proteins and phenocopy the human disorder, including renal agenesis (B, arrow) and cryptophthalmia (B, inset). The Fras and Frem genes encode large, secreted multidomain proteins of which Frem2 and Fras1 are transmembrane polypeptides (C). A model for the action of these proteins in the BM (D) suggests that the earliest signs of Frem expression coincide with Fras1 at E10.5 in epithelial cells (in the mouse), whereas significant basement membrane zone (BMZ) deposition is not observed until later at E12.5. Grip1 encodes an intracellular PDZ domain protein necessary for secretion of Fras1 and Frem2. Expression of Frem1 and Frem3 is also reported in the dermis. The Fras and Frem proteins form a mutually stabilizing complex and are localized to fibrils in the sublamina densa. As development progresses (from left to right), expression of the Fras and Frem proteins decreases in parallel with an increase in the expression of collagen VII. LL, lamina lucida; LD, lamina densa. Patient image was kindly provided by Prof. Peter Scambler.

Fras1, Frem2, and Frem1 all localize to the BM, where the proteins are thought to form an oligomeric complex required for their reciprocal stabilization at this location (Kiyozumi et al., 2006), although differences in tissue specific expression between the genes

suggest that they may have independent functions (Petrou et al., 2007a,b; Pavlakis et al., 2008). This means to an extent that the phenotypes derived from mutation of one of these components might actually reflect a loss or partial loss of all three species. Although

it is clear that the Fras1 and Frem proteins contribute to the BM in the developing epidermis during a specific temporal window (Fig. 3D) and their stabilizing function acts at the level of the dermal-lamina densa interface, how they achieve this and how they interact with

other BM components remains to be determined. A common phenotype of both the blebs mice and FS patients is a high incidence of renal agenesis or dysgenesis (Short et al., 2007; van Haelst et al., 2008). This has been best studied in *Fras1* mutant mice in which ureteric bud formation is relatively normal, and induction and invasion of the metanephric mesenchyme occurs, but kidney development fails to progress (Vrontou et al., 2003). In part, this seems to be a consequence of defects in glial cell-derived neurotrophic factor (Gdnf) and growth differentiation factor (Gdf) signaling, as both molecules can affect phenotypic rescue in culture (Pitera et al., 2008). In surviving adult mice, the GBM in both blebbed (*Fras1*) and myelencephalic blebs (*Frem2*) mice is abnormal, suggesting that the proteins play an ongoing role in maintenance of the glomerulus (Pitera et al., 2008). The development of other organs in FS patients and the blebs mice are also affected. As with *Lama5* mutants, *Fras1* mutant mice exhibit lobar fusion of the lungs and this is also characterized by extravasation of blood cells from the alveolar capillaries into the respiratory lumen (Petrout et al., 2005).

#### *Lipoid proteinosis*

Lipoid proteinosis is a recessive skin disorder characterized by generalized thickening of the skin, mucosal infiltration, and vascular defects of BM components like collagen IV (Fleischmajer et al., 1984; Mirancea et al., 2007). The disorder is caused by mutations in an extracellular glycoprotein called ECM1 (Hamada et al., 2002), and auto-antibodies to this protein also characterize an inflammatory mucocutaneous fragility disorder known as lichen sclerosis (Oyama et al., 2003). ECM1 colocalizes with collagen IV in the skin BM (Sercu et al., 2008b) and is thought to interact with a range of different BM binding partners, including perlecan (Mongiati et al., 2003), laminin-332, and collagen IV (Sercu et al., 2007). Mutations in the gene induce

defects in the epidermal and mucosal BMs, which include irregular collagen and laminin deposition, but normal spatial localization of the hemidesmosomes and basal integrins (Mirancea et al., 2007). As well as its role in the skin, the protein is a known susceptibility locus for ulcerative colitis (Fisher et al., 2008), and is increasingly recognized as a factor regulating metastatic tumor progression (Sercu et al., 2008a).

#### **Diseases Affecting the Kidney**

The kidney is unique in that it contains three different BMs, the glomerular BM (GBM) (Fig. 1B), the BM lining the tubular epithelium, and that of Bowman's capsule. During early kidney development, the BM collagen IV network consists of  $\alpha1\alpha1\alpha2$  protomers, as the kidney develops further the collagen IV species in the renal BM changes. The GBM separates the glomerular endothelium from the epithelial cells of the podocytes (Fig. 1B). In the mature state, it is made up of laminin-521, collagen IV  $\alpha3\alpha4\alpha5$ , perlecan, and nidogen (Miner, 1999). Collagen IV  $\alpha3$  and  $\alpha4$  chains are synthesized and secreted by podocytes, which themselves form slit diaphragms in the GBM for size-selective filtration barriers (Miner, 1999). Furthermore, the GBM has charge properties that facilitate its function as a semipermeable barrier through which blood plasma is filtered. Collagen IV  $\alpha3\alpha4\alpha5$  protomers connect in a lateral fashion to create a tighter network than that of  $\alpha1\alpha1\alpha2$ , thus improving the stability and strength of the BM, to prepare the GBM for higher levels of stress to the adult kidney (Yurchenco and Ruben, 1987). The tubular epithelium which regulates water and electrolyte balance has collagen IV  $\alpha3\alpha4\alpha5$  network and laminin-511 in its BM, whereas Bowman's capsule BM displays  $\alpha1\alpha1\alpha2$ - $\alpha5\alpha5\alpha6$  collagen IV, laminin-111, and laminin-511 network (Miner, 1999). It is interesting to note that only laminin-521 (in the GBM) has been shown to be important for renal function, since deletion of the  $\beta2$  chain in mouse results in proteinuria induced by a defec-

tive glomerular filtration barrier and death within 5 weeks of age (Noakes et al., 1995b). The importance of the renal BM is highlighted by the diverse range of inherited diseases caused by mutations in BM components.

#### *Alport syndrome*

There are several kidney diseases which result from defective GBM that arise from abnormalities in collagen IV, collagen III, and fibronectin deposition (Gubler, 2008). The most common collagen IV-associated renal disease is Alport syndrome, of which there are three types: X-linked, autosomal recessive, and autosomal dominant (Thorner, 2007; Gubler, 2008). Symptoms can vary but usually present as hematuria, hearing loss, and ocular lesions. The latter symptoms occur because collagen IV  $\alpha3\alpha4\alpha5$  is also an essential part of the BM lining the cochlea and the eye (Cheong et al., 1994; Cosgrove et al., 1996a). The X-linked Alport syndrome (~85% of all cases) is caused by a mutation in the *COL4A5* gene (Barker et al., 1990) and analysis of the GBM shows the absence of collagen IV  $\alpha3\alpha4\alpha5$  heterotrimers, but the presence of collagen  $\alpha1\alpha1\alpha2$ , which are normally restricted to the mesangial and endothelial side of the GBM (Gubler, 2008). It is interesting to note that even though only the  $\alpha5$  chain is mutated,  $\alpha3$  and  $\alpha4$  chains are also absent from the GBM, indicating that all three chains are necessary for proper assembly of the structural collagen network. Mutations of the *Col4a5* gene in dog have been shown to decrease mRNA transcript levels of  $\alpha3$  and  $\alpha4$  chains (Thorner et al., 1996); however, this is not reflected in human and mouse, where a post-transcriptional mechanism is the more likely explanation for the absence of these chains (Miner and Sanes, 1996; Nakanishi et al., 1996; Cosgrove et al., 1996b). Because of the mutated *COL4A5* gene, the skin of Alport Syndrome males also lacks collagen IV  $\alpha5\alpha5\alpha6$ , thus immunostaining for collagen IV  $\alpha5$  chain from skin biopsies, which are much less invasive

than renal biopsies, is a useful diagnosis tool for X-linked Alport syndrome in males (Haas, 2009). The disease is progressive in affected males, leading to end-stage renal disease, whereas heterozygous females only experience hematuria, with or without mild proteinuria (Meleg-Smith et al., 1998). At the ultrastructural level, the GBM in Alport's exhibits alternating zones of thinning and thickening accompanied by lamellation, splitting, and fragmentation of the lamina densa into strands, giving the GBM a "basket-weave" appearance (Haas, 2009).

Autosomal recessive Alport Syndrome (10–15% of cases) is severe in affected females, arising from homozygous or compound heterozygous mutations in the *COL4A3* or *COL4A4* genes in chromosome 2 (Mochizuki et al., 1994). Collagen IV  $\alpha 5$  is distributed normally in the skin BM and can be observed in the Bowman's capsule and collecting ducts (as part of the  $\alpha 5\alpha 5\alpha 6$  network), but collagen IV  $\alpha 3\alpha 4\alpha 5$  heterotrimers are absent from the GBM. The morphology of the GBM appears less severe than the X-linked form, with thinning of the GBM similar to Thin Basement Membrane Nephropathy (see below) (Haas, 2009). Autosomal dominant Alport Syndrome (>5% of cases) has been associated with allelic mutations or compound heterozygous *COL4A3* or *COL4A4* mutations (Jefferson et al., 1997; van der Loop et al., 2000; Ciccarese et al., 2001). Various types of mutations have been reported including missense, nonsense, splice mutations, deletions, and insertions (Thorner, 2007; Gubler, 2008). Symptoms are similar to the other two types of Alport's syndrome, but tend to be milder and rarely progress to end-stage renal disease without ocular involvement.

#### *Thin basement membrane nephropathy*

Another renal disease caused by defective collagen IV $\alpha 3$  or  $\alpha 4$  which can be confused with Alport syndrome is thin glomerular BM nephropathy (TBMN). Around 40% of TBMN sufferers are also heterozygous carriers of the autosomal

recessive Alport syndrome affecting *COL4A3* or *COL4A4* genes. The disease is inherited in an autosomal dominant manner, with persistent hematuria a general symptom, but no proteinuria. The disease is quite benign and does not usually progress to end stage renal disease. TBMN is sometimes referred to as benign familial hematuria, although this is not very accurate since some TBMN cases do not have a family history, and not all familial hematuria is benign (Thorner, 2007). The GBM structure is uniformly thin, but splitting or lamellation is rare. A useful diagnostic tool is immunostaining of kidney biopsies with antibodies to collagen IV  $\alpha 5$  or  $\alpha 3$  chains, which are normal in TBMN but abnormal or completely absent in Alport syndrome patients (Haas, 2009). It is not known why some mutations in *COL4A3* or *COL4A4* genes can give rise to autosomal Alport syndrome whereas others cause the more benign TBMN.

#### *Hereditary angiopathy with nephropathy, aneurysms, and cramps*

Mutations in the *COL4A1* gene which affects the  $\alpha 1\alpha 1\alpha 2$  chain expressed in the kidney mesangial matrix, the BM of the Bowman's capsule, the loop of Henle, collecting duct, distal tubule, and proximal tubule have been identified in humans as the cause of a rare autosomal dominant syndrome, hereditary angiopathy with nephropathy, aneurysms, and cramps (HANAC) (Plaisier et al., 2007). The nephropathy manifests as hematuria or cysts. At the ultrastructural level, the BM of the Bowman's capsule and tubules show thickening, lamellation, and splitting of the lamina densa. Interestingly, the tubular BM morphology is reminiscent of the Alport syndrome "basket-weave" appearance of the GBM. The BM of the interstitial capillary also shows lamellations, and in some parts detachment from the endothelial cell lining. In addition, since the  $\alpha 1$  chain is a part of the collagen IV network maintaining the BM of the skin, it is not surpris-

ing that the dermal-epidermal junction BM displays lamina densa duplication to some extent. Another common feature of HANAC is retinal arteriolar tortuosity, thus diagnosis can be made on the basis of hematuria and other external features, confirmed by sequencing of the *COL4A1* gene (Plaisier et al., 2007; Gubler, 2008). Interestingly, certain rare dominant or semidominant mutations in *COL4A1* also cause a type 1 porencephaly in which degenerative lesions affect the cerebellum, and are associated with risk of stroke (Gould et al., 2005).

#### *Pierson syndrome*

Pierson syndrome is caused by mutations in the *LAMB2* gene which codes for the laminin  $\beta 2$  chain, found in the GBM of the mature kidney, and BMs of synapses and the eye (Zenker et al., 2004b). Symptoms include congenital nephritis (hematuria and proteinuria), glomerular lesions, and microcoria (contraction of the pupil of the eye), and may vary widely, possibly due to the nature of the mutations, of which frameshifts (truncating) and missense changes have been reported, with the latter demonstrating a milder phenotype (Choi et al., 2008; Kagan et al., 2008; Zhao et al., 2009). The severity of congenital nephritis does not always correlate to the ocular lesions. Ultrastructures of the GBM present thinning of the BM with irregular contour, lamellation of the lamina densa, and the lack of podocyte foot processes (Choi et al., 2008). *Lamb2* null mice develop severe proteinuria shortly after birth and die within 5 weeks of age, and although the GBM appears normal, the podocyte foot processes is fused. In addition, laminin  $\beta 1$  is detected in the GBM of these mice, indicating that this chain cannot compensate for the loss of  $\beta 2$  chain in the GBM (Noakes et al., 1995b).

#### *Nail-Patella syndrome*

Nail-Patella syndrome (NPS) is an autosomal dominant disease

caused by the abnormal accumulation of the usually absent collagen III in the renal GBM and the mesangial matrix. Characteristic symptoms are nail hypoplasia/dysplasia, bone abnormalities, proteinuria, hematuria, and bilateral microcoria (Dai et al., 2009). The mutation that gives rise to NPS has been mapped to the *LMX1B* locus, which codes for an LIM homeodomain transcription factor required during vertebrate dorso-ventral patterning. Consequently, *Lmx1b*<sup>-/-</sup> mice display developmental defects in dorso-ventral patterning resulting in abnormal limbs, brain defects, and ocular defects, with kidney defects observed in 30% of cases (Dai et al., 2009). *Lmx1b* is expressed in the podocytes of the kidney, and as such is needed for normal formation and maintenance of the podocyte foot processes (Chen et al., 1998). Studies in *Lmx1b* null mice revealed dramatically reduced number of foot processes and lack of slit diaphragms typical of normal podocytes. Furthermore, expression of genes necessary for normal glomerular function, including collagen IV  $\alpha 3$  and  $\alpha 4$ , podocin, and the nephrin-binding CD2AP, is downregulated in these null mice, implicating *Lmx1b* in the regulation of podocyte formation and functions (Morello et al., 2001; Miner et al., 2002; Rohr et al., 2002). The functional role of *Lmx1b* in the kidney may extend to the GBM, since its loss resulted in diminished levels of collagen IV  $\alpha 3$  and  $\alpha 4$  chains, with lesions in the GBM (Morello et al., 2001). However, in contrast to the mouse model, immunohistological analysis of seven human NPS patients with severe glomerular disease revealed normal distribution of collagen IV  $\alpha 3$ ,  $\alpha 4$ , podocin, and CD2AP (Heidet et al., 2003).

#### *Collagen III glomerulopathy and fibronectin glomerulopathy*

Another renal disease caused by collagen III abnormalities that is not caused by *LMX1B* mutations is collagen III glomerulopathy, which is inherited in an autosomal recessive manner and character-

ized by persistent proteinuria, and sometimes hypertension. The glomerulus is dramatically enlarged due to the accumulation of collagen III in the GBM. It can be distinguished from NPS by the expansion of the mesangial matrix and thickened capillary walls, and the lack of GBM lesion (Gubler, 2008). Fibronectin glomerulopathy is an autosomal dominant disease caused by parietal and mesangial fibrillar deposits of fibronectin, and symptoms include proteinuria, hematuria, and hypertension. The causative genes have not been identified in either disease, although it has been confirmed that fibronectin glomerulopathy is not caused by mutations in the fibronectin gene or nearby genes encoding villin and desmin (Hildebrandt et al., 1996)

#### **BM Defects Affecting the Muscle and Cartilage**

##### *Congenital muscular dystrophy 1A (merosin-deficient CMD)*

Congenital muscular dystrophy 1A (CMD1A) is inherited in an autosomal recessive manner, and is clinically characterized by early onset of muscle fiber degeneration, peripheral neuropathy, and white matter lesions in the CNS (Miyagoe-Suzuki et al., 2000). It was originally called merosin-deficient CMD, due to the lack of merosin (which was later renamed laminin-2/laminin-211) on the basal lamina of skeletal muscles and nerves (Burgeson et al., 1994; Tome et al., 1994). Merosin/laminin-211 is predominantly found in the BM of skeletal muscle fibers. Here, it binds dystrophin underlying the sarcolemma and connects it to the surrounding ECM via  $\alpha$  and  $\beta$ -dystroglycans forming the dystrophin glycoprotein complex (DGC), which in turn links the actin cytoskeleton of the muscle cell to the ECM to maintain mechanical stability of the muscle fiber (Leivo and Engvall, 1988; Ervasti and Campbell, 1991; Ibraghimov-Beskrovnaya et al., 1992). Missense deletions, insertions, or nonsense mutations in the human *LAMA2* have been found in CMD1A

patients, resulting in the loss of functional laminin-211 (Helbling-Leclerc et al., 1995). The dy/dy mouse model of merosin-deficient CMD exhibits irregular interruptions of the basal lamina in the skeletal muscle fibers (Arahata et al., 1993). Another mutant, the dy2J dystrophic mouse, lacks parts of domain VI of the laminin  $\alpha 2$  chain resulting from mutation in the splicing site (Xu et al., 1994). The lack of LN motif in this mutant blocks laminin self-assembly during BM formation resulting in the net effect of loss of functional laminin-211.

##### *Schwartz-Jampel syndrome and dyssegmental dysplasia of the Silverman-Handmaker type*

Defects in HSPG2/perlecan underlie both Schwartz-Jampel syndrome (SJS) and dyssegmental dysplasia of the Silverman-Handmaker type (DDSH) (Nicole et al., 1995; Arikawa-Hirasawa et al., 2001a) although the phenotypes are quite distinct, possibly due to the different levels of functional perlecan present in the BM (Stum et al., 2005). SJS is an autosomal recessive disorder characterized by permanent muscle stiffness and skeletal dysplasia. Patients survive into adulthood with relatively mild symptoms which include short stature, osteochondrodysplasia, myotonia, and a "mask-like" face (blepharospasm and pursed lips) (Nicole et al., 2000). Deletions, insertions, missense, nonsense, and splice site mutations in *HSPG2* have been identified in SJS, which result in perlecan mRNA instability. A mouse model containing a homozygous missense mutation based on the same lesion in a human SJS patient appeared normal at 1 month of age, but later developed locomotor defects (Stum et al., 2005). DDSH is a severe autosomal recessive dwarfism that arises through skeletal dysplasia characterized by vertebral segmentation defects, micromelia (short limbs), and decreased joint mobility. The segmentation defects are caused by multiple ossifications in irregular patches in the vertebral bodies. Affected individuals are either still



born or die a few days after birth. Several *HSPG2* mutations have been detected in DDSH patients including exon 34 duplication, splice site mutations, and frame-shifts, which ultimately result in premature truncation of the protein and loss of function (Arikawa-Hirasawa et al., 2001a). The defects in cartilage development may arise from BM destabilization and impaired Fgf signaling (Arikawa-Hirasawa et al., 2001b). This has been demonstrated in an *Hspg2*<sup>-/-</sup> mouse model showing abnormal cartilage matrix with reduced and disorganized collagen fibrils and GAGs, accompanied by reduced chondrocyte proliferation and abnormal Fgf1 accumulation on the chondrocyte cell surface in the growth plate (Arikawa-Hirasawa et al., 1999). The muscle stiffness of SJS and DDSH has been postulated to arise from reduced acetylcholinesterases (AChE) at the neuromuscular junction (NMJ) caused by the absence of functional perlecan (Arikawa-Hirasawa et al., 2002). Under normal conditions, perlecans interact with  $\alpha$ -dystroglycans and recruit AChE to the NMJ to degrade acetylcholine (ACh), the primary neurotransmitter at the NMJ, to terminate synaptic transmission (Peng et al., 1999). Consequently, ACh accumulates at the NMJ of perlecan mutants, leading to hyperexcitability of muscle fibers and the subsequent muscle stiffness.

### BM Defects Affecting the Eye

As previously discussed, retinopathies are a common observation in patients with defects in primary BM proteins such as Alport syndrome. Optic involvement is also a feature of mutations in *COL18A1*, which underlie Knobloch Syndrome, an autosomal recessive disorder characterized by retinal degeneration, myopia, macular defects, and impaired occipital neural tube closure (Sertie et al., 2000).

### CONCLUSIONS AND OUTSTANDING QUESTIONS

It is clear from studies of mutant model organisms that the BM plays diverse roles in regulating

early embryonic development, organogenesis, tissue differentiation, and adult homeostasis. This is reflected in the broad range of inherited diseases caused by perturbation of genes which contribute to, or are involved in, the regulation of the BM. In assessing this range of developmental and adult disease it is clear that beyond the requirement for BMs in the embryo at implantation, they are developmentally important in several organ systems, including the skin, kidney, lung, and cardiovascular systems. Each of these tissues possesses BMs uniquely composed and specialized for particular functional roles. In the context of the skin, an organ subject to significant physical stress before and after birth, the hemidesmosomes and their interacting BM components are key regulators of stability in the tissue. Likewise, in the glomerulus, in the vasculature, and in the heart, which are subject to significant hydrostatic pressure, the BMs play a key role in regulating tissue integrity. Although a great deal is known about the composition and specializations of BMs from these different organs, the recent description of the *Fras1* and *Frem* proteins suggests that perhaps this picture is not yet complete, and that other components and regulators of the BM remain to be elucidated, especially in a developmental context. Similarly, while significant advances in understanding the direct interaction of different BM components has been made using a variety of elegant in vitro approaches, the significance of these many different interactions in vivo has yet to be fully explored. Perlecan is a classic example in this respect. The protein has at least 25 described binding partners (White-lock et al., 2008), but the extent to which these binding partners engender function in vivo temporally (development versus adult homeostasis) or specifically (organs and tissues) is yet to be determined. It will only be through the generation of combinatorial mutations in these genes that such

effects will be revealed. Similarly, the broader roles of those BM components which give rise to embryonic lethal phenotypes will only be revealed through carefully restricted temporal inactivation using Cre-Lox or similar approaches, or through the study of permissive hypomorphic mutants. It is also clear that the BM is capable of acting as much more than a simple physical support, both through direct interaction of BM components with cell surface receptors and by acting as either a sink or focally localized source of growth factors for presentation to neighboring cells. Our understanding of the biology of this proteinaceous layer will continue to be informed and refined by analysis of disease causes and phenotype in humans, where the BM can be rightly considered a critical modulator of disease severity and progression not only in inherited conditions but also during infection, damage, and oncogenesis.

### ACKNOWLEDGMENTS

The authors apologize their colleagues whose publications were unable to cite due to space restraints. The authors thank Lorine Wilkinson, Melissa Little, Pete Scambler, and Dedee Murrell for providing images for this review, and Kieran Short for his help with figure preparation. This review was supported by a project grant to I.S. from the NH&MRC, and I.S. acknowledges support through the NH&MRC R. Douglas Wright and Monash Fellowship schemes.

### REFERENCES

- Aberdam D, Galliano MF, Vailly J, et al. 1994. Herlitz's junctional epidermolysis bullosa is linked to mutations in the gene (LAMC2) for the gamma 2 subunit of nicein/kalinin (LAMININ-5). *Nat Genet* 6:299-304.
- Abrahamson DR, Isom K, Roach E, et al. 2007. Laminin compensation in collagen alpha3(IV) knockout (Alport) glomeruli contributes to permeability defects. *J Am Soc Nephrol* 18:2465-2472.
- Adamson ED, Ayers SE. 1979. The localization and synthesis of some collagen types in developing mouse embryos. *Cell* 16:953-965.



- Alazami AM, Shaheen R, Alzahrani F, et al. 2009. FREM1 mutations cause bifid nose, renal agenesis, and anorectal malformations syndrome. *Am J Hum Genet* 85:414–418.
- Arahata K, Hayashi YK, Mizuno Y, et al. 1993. Dystrophin-associated glycoprotein and dystrophin co-localization at sarcolemma in Fukuyama congenital muscular dystrophy. *Lancet* 342:623–624.
- Arikawa-Hirasawa E, Watanabe H, Takami H, et al. 1999. Perlecan is essential for cartilage and cephalic development. *Nat Genet* 23:354–358.
- Arikawa-Hirasawa E, Wilcox WR, Le AH, et al. 2001a. Dyssegmental dysplasia, Silverman-Handmaker type, is caused by functional null mutations of the perlecan gene. *Nat Genet* 27:431–434.
- Arikawa-Hirasawa E, Wilcox WR, Yamada Y. 2001b. Dyssegmental dysplasia, Silverman-Handmaker type: unexpected role of perlecan in cartilage development. *Am J Med Genet* 106:254–257.
- Arikawa-Hirasawa E, Rossi SG, Rotundo RL, Yamada Y. 2002. Absence of acetylcholinesterase at the neuromuscular junctions of perlecan-null mice. *Nat Neurosci* 5:119–123.
- Arin MJ. 2009. The molecular basis of human keratin disorders. *Hum Genet* 125:355–373.
- Asaka T, Akiyama M, Domon T, et al. 2009. Type XVII collagen is a key player in tooth enamel formation. *Am J Pathol* 174:91–100.
- Ashton GH, Sorelli P, Mellerio JE, et al. 2001a. Alpha 6 beta 4 integrin abnormalities in junctional epidermolysis bullosa with pyloric atresia. *Br J Dermatol* 144:408–414.
- Ashton GH, Sorelli P, Mellerio JE, et al. 2001b. Alpha 6 beta 4 integrin abnormalities in junctional epidermolysis bullosa with pyloric atresia. *Br J Dermatol* 144:408–414.
- Aumailley M, Battaglia C, Mayer U, et al. 1993. Nidogen mediates the formation of ternary complexes of basement membrane components. *Kidney Int* 43:7–12.
- Aviezer D, Hecht D, Safran M, et al. 1994. Perlecan, basal lamina proteoglycan, promotes basic fibroblast growth factor-receptor binding, mitogenesis, and angiogenesis. *Cell* 79:1005–1013.
- Bader BL, Smyth N, Nedbal S, et al. 2005. Compound genetic ablation of nidogen 1 and 2 causes basement membrane defects and perinatal lethality in mice. *Mol Cell Biol* 25:6846–6856.
- Baranowsky A, Mokkapat S, Bechtel M, et al. 2010. Impaired wound healing in mice lacking the basement membrane protein nidogen 1. *Matrix Biol* 29:15–21.
- Barker DF, Hostikka SL, Zhou J, et al. 1990. Identification of mutations in the COL4A5 collagen gene in Alport syndrome. *Science* 248:1224–1227.
- Battaglia C, Mayer U, Aumailley M, Timpl R. 1992. Basement-membrane heparan sulfate proteoglycan binds to laminin by its heparan sulfate chains and to nidogen by sites in the protein core. *Eur J Biochem* 208:359–366.
- Bauer JW, Lanschuetzer C. 2003a. Type XVII collagen gene mutations in junctional epidermolysis bullosa and prospects for gene therapy. *Clin Exp Dermatol* 28:53–60.
- Bauer JW, Lanschuetzer C. 2003b. Type XVII collagen gene mutations in junctional epidermolysis bullosa and prospects for gene therapy. *Clin Exp Dermatol* 28:53–60.
- Beirowski B, Weber M, Gross O. 2006. Chronic renal failure and shortened lifespan in COL4A3+/- mice: an animal model for thin basement membrane nephropathy. *J Am Soc Nephrol* 17:1986–1994.
- Bose K, Nischt R, Page A, et al. 2006. Loss of nidogen-1 and -2 results in syndactyly and changes in limb development. *J Biol Chem* 281:39620–39629.
- Breitkreutz D, Mirancea N, Nischt R. 2009. Basement membranes in skin: unique matrix structures with diverse functions? *Histochem Cell Biol* 132:1–10.
- Brown JC, Sasaki T, Gohring W, et al. 1997. The C-terminal domain V of perlecan promotes beta1 integrin-mediated cell adhesion, binds heparin, nidogen and fibulin-2 and can be modified by glycosaminoglycans. *Eur J Biochem* 250:39–46.
- Bruckner-Tuderman L, Hopfner B, Hammami-Hauasli N. 1999. Biology of anchoring fibrils: lessons from dystrophic epidermolysis bullosa. *Matrix Biol* 18:43–54.
- Burgeson RE. 1993. Type VII collagen, anchoring fibrils, and epidermolysis bullosa. *J Invest Dermatol* 101:252–255.
- Burgeson RE, Chiquet M, Deutzmann R, et al. 1994. A new nomenclature for the laminins. *Matrix Biol* 14:209–211.
- Chapman DB, Hummel KP. 1963. Eye blebs (eb). *Mouse News Lett* 28:32.
- Charlesworth A, Gagnoux-Palacios L, Bonduelle M, et al. 2003. Identification of a lethal form of epidermolysis bullosa simplex associated with a homozygous genetic mutation in plectin. *J Invest Dermatol* 121:1344–1348.
- Chen H, Lun Y, Ovchinnikov D, et al. 1998. Limb and kidney defects in Lmx1b mutant mice suggest an involvement of LMX1B in human nail patella syndrome. *Nat Genet* 19:51–55.
- Cheong HI, Kashtan CE, Kim Y, et al. 1994. Immunohistologic studies of type IV collagen in anterior lens capsules of patients with Alport syndrome. *Lab Invest* 70:553–557.
- Choi HJ, Lee BH, Kang JH, et al. 2008. Variable phenotype of Pierson syndrome. *Pediatr Nephrol* 23:995–1000.
- Christiano AM, Greenspan DS, Hoffman GG, et al. 1993. A missense mutation in type VII collagen in two affected siblings with recessive dystrophic epidermolysis bullosa. *Nat Genet* 4:62–66.
- Ciccarese M, Casu D, Ki Wong F, et al. 2001. Identification of a new mutation in the alpha4(IV) collagen gene in a family with autosomal dominant Alport syndrome and hypercholesterolemia. *Nephrol Dial Transplant* 16:2008–2012.
- Clark CJ, Sage EH. 2008. A prototypic extracellular matrix protein in the tumor microenvironment—where there's SPARC, there's fire. *J Cell Biochem* 104:721–732.
- Colognato H, Winkelman DA, Yurchenco PD. 1999. Laminin polymerization induces a receptor-cytoskeleton network. *J Cell Biol* 145:619–631.
- Cooper AR, MacQueen HA. 1983. Subunits of laminin are differentially synthesized in mouse eggs and early embryos. *Dev Biol* 96:467–471.
- Cosgrove D, Kornak JM, Samuelson G. 1996a. Expression of basement membrane type IV collagen chains during postnatal development in the murine cochlea. *Heart Res* 100:21–32.
- Cosgrove D, Meehan DT, Grunkemeyer JA, et al. 1996b. Collagen COL4A3 knockout: a mouse model for autosomal Alport syndrome. *Genes Dev* 10:2981–2992.
- Costell M, Gustafsson E, Aszodi A, et al. 1999. Perlecan maintains the integrity of cartilage and some basement membranes. *J Cell Biol* 147:1109–1122.
- Costell M, Carmona R, Gustafsson E, et al. 2002. Hyperplastic conotruncal endocardial cushions and transposition of great arteries in perlecan-null mice. *Circ Res* 91:158–164.
- Costello I, Biondi CA, Taylor JM, et al. 2009. Smad4-dependent pathways control basement membrane deposition and endodermal cell migration at early stages of mouse development. *BMC Dev Biol* 9:54.
- Coulombe PA, Hutton ME, Letai A, et al. 1991. Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: genetic and functional analyses. *Cell* 66:1301–1311.
- Dai JX, Johnson RL, Ding YQ. 2009. Manifold functions of the Nail-Patella syndrome gene Lmx1b in vertebrate development. *Dev Growth Differ* 51:241–250.
- Dalezios Y, Papasozomenos B, Petrou P, Chalepakis G. 2007. Ultrastructural localization of Fras1 in the sublamina densa of embryonic epithelial basement membranes. *Arch Dermatol Res* 299:337–343.
- Darling S, Gossler A. 1994. A mouse model for Fraser syndrome? *Clin Dysmorphol* 3:91–95.

- Denes V, Witkovsky P, Koch M, et al. 2007. Laminin deficits induce alterations in the development of dopaminergic neurons in the mouse retina. *Vis Neurosci* 24:549–562.
- Deng WM, Ruohola-Baker H. 2000. Laminin A is required for follicle cell-oocyte signaling that leads to establishment of the anterior-posterior axis in *Drosophila*. *Curr Biol* 10:683–686.
- Dong L, Chen Y, Lewis M, et al. 2002. Neurologic defects and selective disruption of basement membranes in mice lacking entactin-1/nidogen-1. *Lab Invest* 82:1617–1630.
- Durbecq M. 2010. Laminins. *Cell Tissue Res* 339:259–268.
- Dziadek M, Timpl R. 1985. Expression of nidogen and laminin in basement membranes during mouse embryogenesis and in teratocarcinoma cells. *Dev Biol* 111:372–382.
- Eklund P, Eklund M, Fecker L, et al. 1994. Role of mesenchymal nidogen for epithelial morphogenesis in vitro. *Development* 120:2003–2014.
- Eklund L, Piuhola J, Komulainen J, et al. 2001. Lack of type XV collagen causes a skeletal myopathy and cardiovascular defects in mice. *Proc Natl Acad Sci USA* 98:1194–1199.
- Elamaa H, Sormunen R, Rehn M, et al. 2005. Endostatin overexpression specifically in the lens and skin leads to cataract and ultrastructural alterations in basement membranes. *Am J Pathol* 166:221–229.
- Ervasti JM, Campbell KP. 1991. Membrane organization of the dystrophin-glycoprotein complex. *Cell* 66:1121–1131.
- Favor J, Gloeckner CJ, Janik D, et al. 2007. Type IV procollagen missense mutations associated with defects of the eye, vascular stability, the brain, kidney function and embryonic or postnatal viability in the mouse, *Mus musculus*: an extension of the Col4a1 allelic series and the identification of the first two Col4a2 mutant alleles. *Genetics* 175:725–736.
- Fine JD, Eady RA, Bauer EA, et al. 2008. The classification of inherited epidermolysis bullosa (EB). Report of the Third International Consensus Meeting on Diagnosis and Classification of EB. *J Am Acad Dermatol* 58:931–950.
- Fine JD, Johnson LB, Weiner M, et al. 2009. Epidermolysis bullosa and the risk of life-threatening cancers: the National EB Registry experience, 1986–2006. *J Am Acad Dermatol* 60:203–211.
- Fisher SA, Tremelling M, Anderson CA, et al. 2008. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 40:710–712.
- Fleischmajer R, Krieg T, Dziadek M, et al. 1984. Ultrastructure and composition of connective tissue in hyalinos cutis et mucosae skin. *J Invest Dermatol* 82:252–258.
- Fox JW, Mayer U, Nischt R, et al. 1991. Recombinant nidogen consists of three globular domains and mediates binding of laminin to collagen type IV. *EMBO J* 10:3137–3146.
- Fritsch A, Loeckermann S, Kern JS, et al. 2008. A hypomorphic mouse model of dystrophic epidermolysis bullosa reveals mechanisms of disease and response to fibroblast therapy. *J Clin Invest* 118:1669–1679.
- Fuchs E. 1995. Keratins and the skin. *Annu Rev Cell Dev Biol* 11:123–153.
- Fuchs EV, Coppock SM, Green H, Cleveland DW. 1981. Two distinct classes of keratin genes and their evolutionary significance. *Cell* 27:75–84.
- Fukai N, Eklund L, Marneros AG, et al. 2002. Lack of collagen XVIII/endostatin results in eye abnormalities. *EMBO J* 21:1535–1544.
- Fukumoto S, Miner JH, Ida H, et al. 2006. Laminin alpha5 is required for dental epithelium growth and polarity and the development of tooth bud and shape. *J Biol Chem* 281:5008–5016.
- Gaetzner S, Deckers MM, Stahl S, et al. 2005. Endostatin's heparan sulfate-binding site is essential for inhibition of angiogenesis and enhances in situ binding to capillary-like structures in bone explants. *Matrix Biol* 23:557–561.
- Garcia-Alonso L, Fetter RD, Goodman CS. 1996. Genetic analysis of Laminin A in *Drosophila*: extracellular matrix containing laminin A is required for ocellar axon pathfinding. *Development* 122:2611–2621.
- Gautam M, Noakes PG, Moscoso L, et al. 1996. Defective neuromuscular synaptogenesis in agrin-deficient mutant mice. *Cell* 85:525–535.
- Giros A, Morante J, Gil-Sanz C, et al. 2007. Perlecan controls neurogenesis in the developing telencephalon. *BMC Dev Biol* 7:29.
- Gonzalez-Iriarte M, Carmona R, Perez-Pomares JM, et al. 2003. Development of the coronary arteries in a murine model of transposition of great arteries. *J Mol Cell Cardiol* 35:795–802.
- Gould DB, Phalan FC, Breedveld GJ, et al. 2005. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. *Science* 308:1167–1171.
- Gubler MC. 2008. Inherited diseases of the glomerular basement membrane. *Nat Clin Pract Nephrol* 4:24–37.
- Haas M. 2009. Alport syndrome and thin glomerular basement membrane nephropathy: a practical approach to diagnosis. *Arch Pathol Lab Med* 133:224–232.
- Hamada T, McLean WH, Ramsay M, et al. 2002. Lipoid proteinosis maps to 1q21 and is caused by mutations in the extracellular matrix protein 1 gene (ECM1). *Hum Mol Genet* 11:833–840.
- Has C, Bruckner-Tuderman L. 2006a. Molecular and diagnostic aspects of genetic skin fragility. *J Dermatol Sci* 44:129–144.
- Has C, Bruckner-Tuderman L. 2006b. Molecular and diagnostic aspects of genetic skin fragility. *J Dermatol Sci* 44:129–144.
- Heidet L, Bongers EM, Sich M, et al. 2003. In vivo expression of putative LMX1B targets in nail-patella syndrome kidneys. *Am J Pathol* 163:145–155.
- Heinonen S, Mannikko M, Klement JF, et al. 1999. Targeted inactivation of the type VII collagen gene (Col7a1) in mice results in severe blistering phenotype: a model for recessive dystrophic epidermolysis bullosa. *J Cell Sci* 112(Part 21):3641–3648.
- Helbling-Leclerc A, Zhang X, Topaloglu H, et al. 1995. Mutations in the laminin alpha 2-chain gene (LAMA2) cause merosin-deficient congenital muscular dystrophy. *Nat Genet* 11:216–218.
- Henchcliffe C, Garcia-Alonso L, Tang J, Goodman CS. 1993. Genetic analysis of laminin A reveals diverse functions during morphogenesis in *Drosophila*. *Development* 118:325–337.
- Hilal L, Rochat A, Duquesnoy P, et al. 1993. A homozygous insertion-deletion in the type VII collagen gene (COL7A1) in Hallopeau-Siemens dystrophic epidermolysis bullosa. *Nat Genet* 5:287–293.
- Hildebrandt F, Strahm B, Prochoreff A, et al. 1996. Glomerulopathy associated with predominant fibronectin deposits: exclusion of the genes for fibronectin, villin and desmin as causative genes. *Am J Med Genet* 63:323–327.
- Ho MS, Bose K, Mokkapat S, et al. 2008. Nidogens-extracellular matrix linker molecules. *Microsc Res Technol* 71:387–395.
- Hopf M, Gohring W, Kohfeldt E, et al. 1999. Recombinant domain IV of perlecan binds to nidogens, laminin-nidogen complex, fibronectin, fibulin-2 and heparin. *Eur J Biochem* 259:917–925.
- Hopkinson SB, Findlay K, deHart GW, Jones JC. 1998. Interaction of BP180 (type XVII collagen) and alpha6 integrin is necessary for stabilization of hemidesmosome structure. *J Invest Dermatol* 111:1015–1022.
- Hopkinson SB, Jones JC. 2000. The N terminus of the transmembrane protein BP180 interacts with the N-terminal domain of BP230, thereby mediating keratin cytoskeleton anchorage to the cell surface at the site of the hemidesmosome. *Mol Biol Cell* 11:277–286.
- Hudson BG, Tryggvason K, Sundaramoorthy M, Neilson EG. 2003. Alport's syndrome, Goodpasture's syndrome, and type IV collagen. *N Engl J Med* 348:2543–2556.

- Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, et al. 1992. Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* 355:696-702.
- Iozzo RV. 2005. Basement membrane proteoglycans: from cellar to ceiling. *Nat Rev Mol Cell Biol* 6:646-656.
- Jadeja S, Smyth I, Pitera JE, et al. 2005. Identification of a new gene mutated in Fraser syndrome and mouse myelencephalic blebs. *Nat Genet* 37:520-525.
- Jefferson JA, Lemmink HH, Hughes AE, et al. 1997. Autosomal dominant Alport syndrome linked to the type IV collagen alpha 3 and alpha 4 genes (COL4A3 and COL4A4). *Nephrol Dial Transplant* 12:1595-1599.
- Kadota Y, Salmivirta K, Talts JF, et al. 1997. Importance of nidogen binding to laminin gamma1 for branching epithelial morphogenesis of the submandibular gland. *Development* 124:683-691.
- Kagan M, Cohen AH, Matejas V, et al. 2008. A milder variant of Pierson syndrome. *Pediatr Nephrol* 23:323-327.
- Kammerer RA, Schulthess T, Landwehr R, et al. 1999. Interaction of agrin with laminin requires a coiled-coil conformation of the agrin-binding site within the laminin gamma1 chain. *EMBO J* 18:6762-6770.
- Kao G, Huang CC, Hedgecock EM, et al. 2006. The role of the laminin beta subunit in laminin heterotrimer assembly and basement membrane function and development in *C. elegans*. *Dev Biol* 290:211-219.
- Khoshnoodi J, Pedchenko V, Hudson BG. 2008. Mammalian collagen IV. *Microsc Res Technol* 71:357-370.
- Kirn-Safran C, Farach-Carson MC, Carson DD. 2009. Multifunctionality of extracellular and cell surface heparan sulfate proteoglycans. *Cell Mol Life Sci* 66:3421-3434.
- Kivirikko S, McGrath JA, Baudoin C, et al. 1995. A homozygous nonsense mutation in the alpha 3 chain gene of laminin 5 (LAMA3) in lethal (Herlitz) junctional epidermolysis bullosa. *Hum Mol Genet* 4:959-962.
- Kiyozumi D, Osada A, Sugimoto N, et al. 2005. Identification of a novel cell-adhesive protein spatiotemporally expressed in the basement membrane of mouse developing hair follicle. *Exp Cell Res* 306:9-23.
- Kiyozumi D, Sugimoto N, Sekiguchi K. 2006. Breakdown of the reciprocal stabilization of QBRICK/Frem1, Fras1, and Frem2 at the basement membrane provokes Fraser syndrome-like defects. *Proc Natl Acad Sci USA* 103:11981-11986.
- Knight D, Tolley LK, Kim DK, et al. 2003. Functional analysis of neurotransmission at beta2-laminin deficient terminals. *J Physiol* 546:789-800.
- Knoll R, Postel R, Wang J, et al. 2007. Laminin-alpha4 and integrin-linked kinase mutations cause human cardiomyopathy via simultaneous defects in cardiomyocytes and endothelial cells. *Circulation* 116:515-525.
- Kohfeldt E, Sasaki T, Gohring W, Timpl R. 1998. Nidogen-2: a new basement membrane protein with diverse binding properties. *J Mol Biol* 282:99-109.
- Koster J, Geerts D, Favre B, et al. 2003. Analysis of the interactions between BP180, BP230, plectin and the integrin alpha6beta4 important for hemidesmosome assembly. *J Cell Sci* 116:387-399.
- Kramer JM. 2005. Basement membranes. *WormBook*, 1-15.
- Kuang W, Xu H, Vachon PH, Engvall E. 1998. Disruption of the lama2 gene in embryonic stem cells: laminin alpha 2 is necessary for sustenance of mature muscle cells. *Exp Cell Res* 241:117-125.
- Kuster JE, Guarnieri MH, Ault JG, et al. 1997. IAP insertion in the murine Lamb3 gene results in junctional epidermolysis bullosa. *Mamm Genome* 8:673-681.
- Lane EB, Rugg EL, Navsaria H, et al. 1992. A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering. *Nature* 356:244-246.
- LeBleu VS, Macdonald B, Kalluri R. 2007. Structure and function of basement membranes. *Exp Biol Med* (Maywood) 232:1121-1129.
- Leinonen A, Mariyama M, Mochizuki T, et al. 1994. Complete primary structure of the human type IV collagen alpha 4(IV) chain. Comparison with structure and expression of the other alpha (IV) chains. *J Biol Chem* 269:26172-26177.
- Leivo I, Engvall E. 1988. Merosin, a protein specific for basement membranes of Schwann cells, striated muscle, and trophoblast, is expressed late in nerve and muscle development. *Proc Natl Acad Sci USA* 85:1544-1548.
- Lemmink HH, Mochizuki T, van den Heuvel LP, et al. 1994. Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal recessive Alport syndrome. *Hum Mol Genet* 3:1269-1273.
- Letai A, Coulombe PA, McCormick MB, et al. 1993. Disease severity correlates with position of keratin point mutations in patients with epidermolysis bullosa simplex. *Proc Natl Acad Sci USA* 90:3197-3201.
- Li J, Tzu J, Chen Y, et al. 2003. Laminin-10 is crucial for hair morphogenesis. *EMBO J* 22:2400-2410.
- Li L, Arman E, Ekblom P, et al. 2004. Distinct GATA6- and laminin-dependent mechanisms regulate endodermal and ectodermal embryonic stem cell fates. *Development* 131:5277-5286.
- Li MH, Fan P, Brodsky B, Baum J. 1993. Two-dimensional NMR assignments and conformation of (Pro-Hyp-Gly)10 and a designed collagen triple-helical peptide. *Biochemistry* 32:7377-7387.
- Libby RT, Lavalley CR, Balkema GW, et al. 1999. Disruption of laminin beta2 chain production causes alterations in morphology and function in the CNS. *J Neurosci* 19:9399-9411.
- Litjens SH, Koster J, Kuikman I, et al. 2003. Specificity of binding of the plectin actin-binding domain to beta4 integrin. *Mol Biol Cell* 14:4039-4050.
- Litjens SH, de Pereda JM, Sonnenberg A. 2006. Current insights into the formation and breakdown of hemidesmosomes. *Trends Cell Biol* 16:376-383.
- Little CC, Bagg HJ. 1923. The occurrence of two heritable types of abnormality among descendants of X-rayed mice. *Am J Roentgenol* 10:975-989.
- Marionnet C, Pierrard C, Vioux-Chagnoleau C, et al. 2006. Interactions between fibroblasts and keratinocytes in morphogenesis of dermal epidermal junction in a model of reconstructed skin. *J Invest Dermatol* 126:971-979.
- Mariyama M, Zheng K, Yang-Feng TL, Reeders ST. 1992. Colocalization of the genes for the alpha 3(IV) and alpha 4(IV) chains of type IV collagen to chromosome 2 bands q35-q37. *Genomics* 13:809-813.
- Martin D, Zusman S, Li X, et al. 1999. Wing blister, a new *Drosophila* laminin alpha chain required for cell adhesion and migration during embryonic and imaginal development. *J Cell Biol* 145:191-201.
- Martinek N, Shahab J, Saathoff M, Ringuette M. 2008. Hemocyte-derived SPARC is required for collagen-IV-dependent stability of basal laminae in *Drosophila* embryos. *J Cell Sci* 121:1671-1680.
- Martins VL, Vyas JJ, Chen M, et al. 2009. Increased invasive behavior in cutaneous squamous cell carcinoma with loss of basement-membrane type VII collagen. *J Cell Sci* 122:1788-1799.
- Mayer U, Nischt R, Poschl E, et al. 1993. A single EGF-like motif of laminin is responsible for high affinity nidogen binding. *EMBO J* 12:1879-1885.
- McGrath JA, Gatalica B, Christiano AM, et al. 1995a. Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. *Nat Genet* 11:83-86.
- McGrath JA, Gatalica B, Christiano AM, et al. 1995b. Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. *Nat Genet* 11:83-86.

- McGrath JA, Gatalica B, Christiano AM, et al. 1995c. Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. *Nat Genet* 11:83–86.
- McGregor L, Makela V, Darling SM, et al. 2003. Fraser syndrome and mouse blebbed phenotype caused by mutations in FRAS1/Fras1 encoding a putative extracellular matrix protein. *Nat Genet* 34:203–208.
- McLean WH, Pulkkinen L, Smith FJ, et al. 1996. Loss of plectin causes epidermolysis bullosa with muscular dystrophy: cDNA cloning and genomic organization. *Genes Dev* 10:1724–1735.
- McMillan JR, Akiyama M, Shimizu H. 2003. Epidermal basement membrane zone components: ultrastructural distribution and molecular interactions. *J Dermatol Sci* 31:169–177.
- Meleg-Smith S, Magliato S, Cheles M, et al. 1998. X-linked Alport syndrome in females. *Hum Pathol* 29:404–408.
- Meng X, Klement JF, Leperi DA, et al. 2003. Targeted inactivation of murine laminin gamma2-chain gene recapitulates human junctional epidermolysis bullosa. *J Invest Dermatol* 121:720–731.
- Miner JH. 1999. Renal basement membrane components. *Kidney Int* 56:2016–2024.
- Miner JH, Cunningham J, Sanes JR. 1998. Roles for laminin in embryogenesis: exencephaly, syndactyly, and placental pathology in mice lacking the laminin alpha5 chain. *J Cell Biol* 143:1713–1723.
- Miner JH, Li C. 2000. Defective glomerulogenesis in the absence of laminin alpha5 demonstrates a developmental role for the kidney glomerular basement membrane. *Dev Biol* 217:278–289.
- Miner JH, Morello R, Andrews KL, et al. 2002. Transcriptional induction of slit diaphragm genes by Lmx1b is required in podocyte differentiation. *J Clin Invest* 109:1065–1072.
- Miner JH, Li C, Mudd JL, et al. 2004. Compositional and structural requirements for laminin and basement membranes during mouse embryo implantation and gastrulation. *Development* 131:2247–2256.
- Miner JH, Sanes JR. 1996. Molecular and functional defects in kidneys of mice lacking collagen alpha 3(IV): implications for Alport syndrome. *J Cell Biol* 135:1403–1413.
- Miner JH, Yurchenco PD. 2004. Laminin functions in tissue morphogenesis. *Annu Rev Cell Dev Biol* 20:255–284.
- Miosge N, Sasaki T, Timpl R. 2002. Evidence of nidogen-2 compensation for nidogen-1 deficiency in transgenic mice. *Matrix Biol* 21:611–621.
- Mirancea N, Hausser I, Metze D, et al. 2007. Junctional basement membrane anomalies of skin and mucosa in lipoid proteinosis (hyalinosis cutis et mucosae). *J Dermatol Sci* 45:175–185.
- Miyagoe Y, Hanaoka K, Nonaka I, et al. 1997. Laminin alpha2 chain-null mutant mice by targeted disruption of the Lama2 gene: a new model of merosin (laminin 2)-deficient congenital muscular dystrophy. *FEBS Lett* 415:33–39.
- Miyagoe-Suzuki Y, Nakagawa M, Takeda S. 2000. Merosin and congenital muscular dystrophy. *Microsc Res Tech* 48:181–191.
- Mochizuki T, Lemmink HH, Mariyama M, et al. 1994. Identification of mutations in the alpha 3(IV) and alpha 4(IV) collagen genes in autosomal recessive Alport syndrome. *Nat Genet* 8:77–81.
- Mokkapat S, Baranowsky A, Mirancea N, et al. 2008. Basement membranes in skin are differently affected by lack of nidogen 1 and 2. *J Invest Dermatol* 128:2259–2267.
- Mongiat M, Fu J, Oldershaw R, et al. 2003. Perlecan protein core interacts with extracellular matrix protein 1 (ECM1), a glycoprotein involved in bone formation and angiogenesis. *J Biol Chem* 278:17491–17499.
- Mongiat M, Taylor K, Otto J, et al. 2000. The protein core of the proteoglycan perlecan binds specifically to fibroblast growth factor-7. *J Biol Chem* 275:7095–7100.
- Morello R, Zhou G, Dreyer SD, et al. 2001. Regulation of glomerular basement membrane collagen expression by LMX1B contributes to renal disease in nail patella syndrome. *Nat Genet* 27:205–208.
- Muhle C, Jiang QJ, Charlesworth A, et al. 2005a. Novel and recurrent mutations in the laminin-5 genes causing lethal junctional epidermolysis bullosa: molecular basis and clinical course of Herlitz disease. *Hum Genet* 116:33–42.
- Muhle C, Jiang QJ, Charlesworth A, et al. 2005b. Novel and recurrent mutations in the laminin-5 genes causing lethal junctional epidermolysis bullosa: molecular basis and clinical course of Herlitz disease. *Hum Genet* 116:33–42.
- Murshed M, Smyth N, Miosge N, et al. 2000. The absence of nidogen 1 does not affect murine basement membrane formation. *Mol Cell Biol* 20:7007–7012.
- Nakanishi K, Yosikawa N, Iijima K, Nakamura H. 1996. Expression of type IV collagen alpha 3 and alpha 4 chain mRNA in X-linked Alport syndrome. *J Am Soc Nephrol* 7:938–945.
- Nakano A, Chao SC, Pulkkinen L, et al. 2002a. Laminin 5 mutations in junctional epidermolysis bullosa: molecular basis of Herlitz vs. non-Herlitz phenotypes. *Hum Genet* 110:41–51.
- Nakano A, Chao SC, Pulkkinen L, et al. 2002b. Laminin 5 mutations in junctional epidermolysis bullosa: molecular basis of Herlitz vs. non-Herlitz phenotypes. *Hum Genet* 110:41–51.
- Nakano A, Chao SC, Pulkkinen L, et al. 2002c. Laminin 5 mutations in junctional epidermolysis bullosa: molecular basis of Herlitz vs. non-Herlitz phenotypes. *Hum Genet* 110:41–51.
- Nguyen NM, Miner JH, Pierce RA, Senior RM. 2002. Laminin alpha 5 is required for lobar septation and visceral pleural basement membrane formation in the developing mouse lung. *Dev Biol* 246:231–244.
- Nicole S, Ben Hamida C, Beighton P, et al. 1995. Localization of the Schwartz-Jampel syndrome (SJS) locus to chromosome 1p34-p36.1 by homozygosity mapping. *Hum Mol Genet* 4:1633–1636.
- Nicole S, Davoine CS, Topaloglu H, et al. 2000. Perlecan, the major proteoglycan of basement membranes, is altered in patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia). *Nat Genet* 26:480–483.
- Nishie W, Sawamura D, Goto M, et al. 2007. Humanization of autoantigen. *Nat Med* 13:378–383.
- Noakes PG, Gautam M, Mudd J, et al. 1995a. Aberrant differentiation of neuromuscular junctions in mice lacking s-laminin/laminin beta 2. *Nature* 374:258–262.
- Noakes PG, Miner JH, Gautam M, et al. 1995b. The renal glomerulus of mice lacking s-laminin/laminin beta 2: nephrosis despite molecular compensation by laminin beta 1. *Nat Genet* 10:400–406.
- Odenthal U, Haehn S, Tunggal P, et al. 2004. Molecular analysis of laminin N-terminal domains mediating self-interactions. *J Biol Chem* 279:44504–44512.
- Ornitz DM. 2000. FGFs, heparan sulfate and FGFRs: complex interactions essential for development. *Bioessays* 22:108–112.
- Oyama N, Chan I, Neill SM, et al. 2003. Autoantibodies to extracellular matrix protein 1 in lichen sclerosis. *Lancet* 362,118–123.
- Pakula R, Melchior A, Denys A, et al. 2007. Syndecan-1/CD147 association is essential for cyclophilin B-induced activation of p44/42 mitogen-activated protein kinases and promotion of cell adhesion and chemotaxis. *Glycobiology* 17:492–503.
- Park Y, Rangel C, Reynolds MM, et al. 2003. Drosophila perlecan modulates FGF and hedgehog signals to activate neural stem cell division. *Dev Biol* 253:247–257.
- Patthy L, Nikolics K. 1993. Functions of agrin and agrin-related proteins. *Trends Neurosci* 16:76–81.
- Patton BL, Cunningham JM, Thyboll J, et al. 2001. Properly formed but improperly localized synaptic specializations in the absence of laminin alpha4. *Nat Neurosci* 4:597–604.



- Pavlikakis E, Makrygiannis AK, Chiotaki R, Chalepakis G. 2008. Differential localization profile of Fras1/Frem proteins in epithelial basement membranes of newborn and adult mice. *Histochem Cell Biol* 130:785–793.
- Peng HB, Xie H, Rossi SG, Rotundo RL. 1999. Acetylcholinesterase clustering at the neuromuscular junction involves perlecan and dystroglycan. *J Cell Biol* 145:911–921.
- Petrou P, Pavlikakis E, Dalezios Y, et al. 2005. Basement membrane distortions impair lung lobation and capillary organization in the mouse model for fraser syndrome. *J Biol Chem* 280:10350–10356.
- Petrou P, Chiotaki R, Dalezios Y, Chalepakis G. 2007a. Overlapping and divergent localization of Frem1 and Fras1 and its functional implications during mouse embryonic development. *Exp Cell Res* 313:910–920.
- Petrou P, Pavlikakis E, Dalezios Y, Chalepakis G. 2007b. Basement membrane localization of Frem3 is independent of the Fras1/Frem1/Frem2 protein complex within the sublamina densa. *Matrix Biol* 26:652–658.
- Pfendner EG, Sadowski SG, Uitto J. 2005. Epidermolysis bullosa simplex: recurrent and de novo mutations in the KRT5 and KRT14 genes, phenotype/genotype correlations, and implications for genetic counseling and prenatal diagnosis. *J Invest Dermatol* 125:239–243.
- Phillips RJS. 1970. Blebbed, bl. *Mouse News Lett* 42:26.
- Pitera JE, Scambler PJ, Woolf AS. 2008. Fras1, a basement membrane-associated protein mutated in Fraser syndrome, mediates both the initiation of the mammalian kidney and the integrity of renal glomeruli. *Hum Mol Genet* 17:3953–3964.
- Plaisier E, Gribouval O, Alamowitch S, et al. 2007. COL4A1 mutations and hereditary angiopathy, nephropathy, aneurysms, and muscle cramps. *N Engl J Med* 357:2687–2695.
- Poschl E, Mayer U, Stetefeld J, et al. 1996. Site-directed mutagenesis and structural interpretation of the nidogen binding site of the laminin gamma1 chain. *EMBO J* 15:5154–5159.
- Poschl E, Schlotzer-Schrehardt U, Brachvogel B, et al. 2004. Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development* 131:1619–1628.
- Pulkkinen L, Christiano AM, Airenne T, et al. 1994a. Mutations in the gamma 2 chain gene (LAMC2) of kalinin/laminin 5 in the junctional forms of epidermolysis bullosa. *Nat Genet* 6:293–297.
- Pulkkinen L, Christiano AM, Gerecke D, et al. 1994b. A homozygous nonsense mutation in the beta 3 chain gene of laminin 5 (LAMB3) in Herlitz junctional epidermolysis bullosa. *Genomics* 24:357–360.
- Rattenholl A, Pappano WN, Koch M, et al. 2002. Proteinases of the bone morphogenetic protein-1 family convert procollagen VII to mature anchoring fibril collagen. *J Biol Chem* 277:26372–26378.
- Rebustini IT, Patel VN, Stewart JS, et al. 2007. Laminin alpha5 is necessary for submandibular gland epithelial morphogenesis and influences FGFR expression through beta1 integrin signaling. *Dev Biol* 308:15–29.
- Rheault MN, Kren SM, Thielen BK, et al. 2004. Mouse model of X-linked Alport syndrome. *J Am Soc Nephrol* 15:1466–1474.
- Ries A, Gohring W, Fox JW, et al. 2001. Recombinant domains of mouse nidogen-1 and their binding to basement membrane proteins and monoclonal antibodies. *Eur J Biochem* 268:5119–5128.
- Rohr C, Prestel J, Heidet L, et al. 2002. The LIM-homeodomain transcription factor Lmx1b plays a crucial role in podocytes. *J Clin Invest* 109:1073–1082.
- Roop D. 1995. Defects in the barrier. *Science* 267:474–475.
- Rousselle P, Keene DR, Ruggiero F, et al. 1997. Laminin 5 binds the NC-1 domain of type VII collagen. *J Cell Biol* 138:719–728.
- Ruhrberg C, Watt FM. 1997. The plakins family: versatile organizers of cytoskeletal architecture. *Curr Opin Genet Dev* 7:392–397.
- Ryan MC, Lee K, Miyashita Y, Carter WG. 1999. Targeted disruption of the LAMA3 gene in mice reveals abnormalities in survival and late stage differentiation of epithelial cells. *J Cell Biol* 145:1309–1323.
- Ryynanen J, Sollberg S, Parente MG, et al. 1992. Type VII collagen gene expression by cultured human cells and in fetal skin. Abundant mRNA and protein levels in epidermal keratinocytes. *J Clin Invest* 89:163–168.
- Sasaki T, Larsson H, Tisi D, et al. 2000. Endostatin derived from collagens XV and XVIII differ in structural and binding properties, tissue distribution and anti-angiogenic activity. *J Mol Biol* 301:1179–1190.
- Sasaki T, Gohring W, Mann K, et al. 2001. Short arm region of laminin-5 gamma2 chain: structure, mechanism of processing and binding to heparin and proteins. *J Mol Biol* 314:751–763.
- Sasse P, Malan D, Fleischmann M, et al. 2008. Perlecan is critical for heart stability. *Cardiovasc Res* 80:435–444.
- Schulze B, Mann K, Poschl E, et al. 1996. Structural and functional analysis of the globular domain IVa of the laminin alpha 1 chain and its impact on an adjacent RGD site. *Biochem J* 314 (Part 3):847–851.
- Schymeinsky J, Nedbal S, Miosge N, et al. 2002. Gene structure and functional analysis of the mouse nidogen-2 gene: nidogen-2 is not essential for basement membrane formation in mice. *Mol Cell Biol* 22:6820–6830.
- Sercu S, Poumay Y, Herphelin F, et al. 2007. Functional redundancy of extracellular matrix protein 1 in epidermal differentiation. *Br J Dermatol* 157:771–775.
- Sercu S, Zhang L, Merregaert J. 2008a. The extracellular matrix protein 1: its molecular interaction and implication in tumor progression. *Cancer Invest* 26:375–384.
- Sercu S, Zhang M, Oyama N, et al. 2008b. Interaction of extracellular matrix protein 1 with extracellular matrix components: ECM1 is a basement membrane protein of the skin. *J Invest Dermatol* 128:1397–1408.
- Sertie AL, Sossi V, Camargo AA, et al. 2000. Collagen XVIII, containing an endogenous inhibitor of angiogenesis and tumor growth, plays a critical role in the maintenance of retinal structure and in neural tube closure (Knobloch syndrome). *Hum Mol Genet* 9:2051–2058.
- Shannon MB, Patton BL, Harvey SJ, Miner JH. 2006. A hypomorphic mutation in the mouse laminin alpha5 gene causes polycystic kidney disease. *J Am Soc Nephrol* 17:1913–1922.
- Short K, Wiradjaja F, Smyth I. 2007. Let's stick together: the role of the Fras1 and Frem proteins in epidermal adhesion. *IUBMB Life* 59:427–435.
- Slavotinek AM, Tiffet CJ. 2002. Fraser syndrome and cryptophthalmos: review of the diagnostic criteria and evidence for phenotypic modules in complex malformation syndromes. *J Med Genet* 39:623–633.
- Smyth I, Du X, Taylor MS, et al. 2004. The extracellular matrix gene Frem1 is essential for the normal adhesion of the embryonic epidermis. *Proc Natl Acad Sci USA* 101:13560–13565.
- Smyth I, Scambler P. 2005. The genetics of Fraser syndrome and the blebs mouse mutants. *Hum Mol Genet* 14(Spec No. 2):R269–R274.
- Smyth N, Vatansever HS, Murray P, et al. 1999. Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation. *J Cell Biol* 144:151–160.
- Soininen R, Huotari M, Hostikka SL, et al. 1988. The structural genes for alpha 1 and alpha 2 chains of human type IV collagen are divergently encoded on opposite DNA strands and have an overlapping promoter region. *J Biol Chem* 263:17217–17220.
- Sonnenberg A, de Melker AA, Martinez de Velasco AM, et al. 1993. Formation of hemidesmosomes in cells of a transformed murine mammary tumor cell line and mechanisms involved in adherence of these cells to laminin and kalinin. *J Cell Sci* 106 (Part 4):1083–1102.
- Steinert PM, Marekov LN, Fraser RD, Parry DA. 1993. Keratin intermediate

- filament structure. Crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *J Mol Biol* 230:436-452.
- Stum M, Davoine CS, Fontaine B, Nicole S. 2005. Schwartz-Jampel syndrome and perlecan deficiency. *Acta Myol* 24:89-92.
- Sugawara K, Tsuruta D, Ishii M, et al. 2008. Laminin-332 and -511 in skin. *Exp Dermatol* 17:473-480.
- Sugimoto M, Ohashi T, Ninomiya Y. 1994. The genes COL4A5 and COL4A6, coding for basement membrane collagen chains alpha 5(IV) and alpha 6(IV), are located head-to-head in close proximity on human chromosome Xq22 and COL4A6 is transcribed from two alternative promoters. *Proc Natl Acad Sci USA* 91:11679-11683.
- Takamiya K, Kostourou V, Adams S, et al. 2004. A direct functional link between the multi-PDZ domain protein GRIP1 and the Fraser syndrome protein Fras1. *Nat Genet* 36:172-177.
- Talts JF, Andac Z, Gohring W, et al. 1999. Binding of the G domains of laminin alpha1 and alpha2 chains and perlecan to heparin, sulfatides, alpha-dystroglycan and several extracellular matrix proteins. *EMBO J* 18:863-870.
- Tasanen K, Tunggal L, Chometon G, et al. 2004. Keratinocytes from patients lacking collagen XVII display a migratory phenotype. *Am J Pathol* 164:2027-2038.
- Thorner PS. 2007. Alport syndrome and thin basement membrane nephropathy. *Nephron Clin Pract* 106:c82-c88.
- Thorner PS, Zheng K, Kalluri R, et al. 1996. Coordinate gene expression of the alpha3, alpha4, and alpha5 chains of collagen type IV. Evidence from a canine model of X-linked nephritis with a COL4A5 gene mutation. *J Biol Chem* 271:13821-13828.
- Timmer JR, Mak TW, Manova K, et al. 2005. Tissue morphogenesis and vascular stability require the Frem2 protein, product of the mouse myelencephalic blebs gene. *Proc Natl Acad Sci USA* 102:11746-11750.
- Timpl R. 1996. Macromolecular organization of basement membranes. *Curr Opin Cell Biol* 8:618-624.
- Timpl R, Brown JC. 1996. Supramolecular assembly of basement membranes. *Bioessays* 18:123-132.
- Tisi D, Talts JF, Timpl R, Hohenester E. 2000. Structure of the C-terminal laminin G-like domain pair of the laminin alpha2 chain harboring binding sites for alpha-dystroglycan and heparin. *EMBO J* 19:1432-1440.
- Titeux M, Pendaries V, Tonasso L, et al. 2008. A frequent functional SNP in the MMP1 promoter is associated with higher disease severity in recessive dystrophic epidermolysis bullosa. *Hum Mutat* 29:267-276.
- Tome FM, Evangelista T, Leclerc A, et al. 1994. Congenital muscular dystrophy with merosin deficiency. *C R Acad Sci III* 317:351-357.
- Tzu J, Marinkovich MP. 2008. Bridging structure with function: structural, regulatory, and developmental role of laminins. *Int J Biochem Cell Biol* 40:199-214.
- Urbano JM, Torgler CN, Molnar C, et al. 2009. Drosophila laminins act as key regulators of basement membrane assembly and morphogenesis. *Development* 136:4165-4176.
- Utani A, Nomizu M, Yamada Y. 1997. Fibulin-2 binds to the short arms of laminin-5 and laminin-1 via conserved amino acid sequences. *J Biol Chem* 272:2814-2820.
- Utriainen A, Sormunen R, Kettunen M, et al. 2004. Structurally altered basement membranes and hydrocephalus in a type XVIII collagen deficient mouse line. *Hum Mol Genet* 13:2089-2099.
- Van Agtmael T, Schlotzer-Schrehardt U, McKie L, et al. 2005. Dominant mutations of Col4a1 result in basement membrane defects which lead to anterior segment dysgenesis and glomerulopathy. *Hum Mol Genet* 14:3161-3168.
- van der Loop FT, Heidet L, Timmer ED, et al. 2000. Autosomal dominant Alport syndrome caused by a COL4A3 splice site mutation. *Kidney Int* 58:1870-1875.
- van Haelst MM, Maiburg M, Baujat G, et al. 2008. Molecular study of 33 families with Fraser syndrome new data and mutation review. *Am J Med Genet A* 146:2252-2257.
- Varki R, Sadowski S, Pfendner E, Uitto J. 2006a. Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants. *J Med Genet* 43:641-652.
- Varki R, Sadowski S, Pfendner E, Uitto J. 2006b. Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants. *J Med Genet* 43:641-652.
- Varki R, Sadowski S, Pfendner E, Uitto J. 2006c. Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants. *J Med Genet* 43:641-652.
- Varki R, Sadowski S, Uitto J, Pfendner E. 2007. Epidermolysis bullosa. II. Type VII collagen mutations and phenotype-genotype correlations in the dystrophic subtypes. *J Med Genet* 44:181-192.
- Varnum DS, Fox SC. 1981. Head blebs: a new mutation on chromosome 4 of the mouse. *J Hered* 72:293.
- Vrontou S, Petrou P, Meyer BI, et al. 2003. Fras1 deficiency results in cryptophthalmos, renal agenesis and blebbed phenotype in mice. *Nat Genet* 34:209-214.
- Webb AE, Sanderford J, Frank D, et al. 2007. Laminin alpha5 is essential for the formation of the zebrafish fins. *Dev Biol* 311:369-382.
- Whitelock JM, Melrose J, Iozzo RV. 2008. Diverse cell signaling events modulated by perlecan. *Biochemistry* 47:11174-11183.
- Willem M, Miosge N, Halfter W, et al. 2002. Specific ablation of the nidogen-binding site in the laminin gamma1 chain interferes with kidney and lung development. *Development* 129:2711-2722.
- Winter RM. 1990. Fraser syndrome and mouse "bleb" mutants. *Clin Genet* 37:494-495.
- Winzen U, Cole GJ, Halfter W. 2003. Agrin is a chimeric proteoglycan with the attachment sites for heparan sulfate/chondroitin sulfate located in two multiple serine-glycine clusters. *J Biol Chem* 278:30106-30114.
- Xu H, Christmas P, Wu XR, et al. 1994. Defective muscle basement membrane and lack of M-laminin in the dystrophic dy/dy mouse. *Proc Natl Acad Sci USA* 91:5572-5576.
- Yancey KB, Hintner H. 2010. Non-herlitz junctional epidermolysis bullosa. *Dermatol Clin* 28:67-77.
- Yarnitzky T, Volk T. 1995. Laminin is required for heart, somatic muscles, and gut development in the *Drosophila* embryo. *Dev Biol* 169:609-618.
- Ylikarppa R, Eklund L, Sormunen R, et al. 2003. Double knockout mice reveal a lack of major functional compensation between collagens XV and XVIII. *Matrix Biol* 22:443-448.
- Yurchenco PD, Cheng YS. 1994. Laminin self-assembly: a three-arm interaction hypothesis for the formation of a network in basement membranes. *Contrib Nephrol* 107:47-56.
- Yurchenco PD, Cheng YS, Colognato H. 1992. Laminin forms an independent network in basement membranes. *J Cell Biol* 117:1119-1133.
- Yurchenco PD, Ruben GC. 1987. Basement membrane structure in situ: evidence for lateral associations in the type IV collagen network. *J Cell Biol* 105:2559-2568.
- Zenker M, Aigner T, Wendler O, et al. 2004a. Human laminin beta2 deficiency causes congenital nephrosis with mesangial sclerosis and distinct eye abnormalities. *Hum Mol Genet* 13:2625-2632.
- Zenker M, Tralau T, Lennert T, et al. 2004b. Congenital nephrosis, mesangial sclerosis, and distinct eye abnormalities with microcoria: an autosomal recessive syndrome. *Am J Med Genet A* 130:138-145.
- Zhao D, Ding J, Wang F, et al. 2009. The first Chinese Pierson syndrome with novel mutations in LAMB2. *Nephrol Dial Transplant* 25:776-778.
- Zoeller JJ, McQuillan A, Whitelock J, et al. 2008. A central function for perlecan in skeletal muscle and cardiovascular development. *J Cell Biol* 181:381-394.