Influence of Genetic Background on Genetically Engineered Mouse Phenotypes

Thomas Doetschman

Abstract

The history of mouse genetics, which involves the study of strain-dependent phenotype variability, makes it clear that the genetic background onto which a gene-targeted allele is placed can cause considerable variation in genetically engineered mouse (GEM) phenotype. This variation can present itself as completely different phenotypes, as variations in penetrance of phenotype, or as variable expressivity of phenotype. In this chapter we provide examples from gene-targeting literature showing each of these types of phenotype variation. We discuss ways in which modifier genes can affect the phenotype of a mouse with a mutant gene, and we give examples of modifier locus identification. We also review approaches to minimize gene polymorphism and flanking gene differences between experimental animals, and between them and their controls. In addition, we discuss the advantages and disadvantages of performing the first analysis of a knockout mouse on a mixed genetic background. We conclude that a mixed background provides the quickest preview of possible strain-dependent phenotypes (1, 2). Finally, we review recent approaches to improving genetic diversity by generating new inbred strains that encompass a broader range of alleles within the mouse species.

Key words: Knockout, mouse, genetic background, genetic engineering, penetrance, expressivity, modifier gene.

1. Introduction

The theoretical basis for an understanding of Mendelian inheritance of complex traits and the importance of genetic background in mouse studies was presented in a note to Science by C. C. Little in 1914 (3), an argument he later used to explain the Mendelian nature of the apparent non-Mendelian inheritance of tumor transplantation susceptibility (4). It was this understanding that led him
to believe in the importance of developing inbred mouse strains as tools for investigating complex human disease. A demonstration of the power of utilizing the genetic diversity of mouse to generate models for a complex trait was provided by Gunther Schlager (5) who used an eight-way cross of common inbred strains to develop advanced intercross lines of mice with differential blood pressures. More recently, striking phenotypic differences have been found in closely related strains, for example, in response to proteoglycan-induced arthritis among ten different C3H substrains (6). These and other studies make it clear that genetically engineered mouse (GEM) phenotype can be background dependent.

The influence of genetic background on GEM phenotype became apparent in some of the early knockout mice. Hynes reported that fibronectin knockouts had considerable variation in phenotype which he attributed to the analyses being done on embryos of a 129 and C57BL/6 background (Hynes George 1993). Baribault and colleagues (7) showed that a keratin-8 deficiency on different backgrounds leads to quite different phenotypes of midgestational lethality or adult colorectal hyperplasia (8). Other notable background-dependent differential GEM phenotypes have been found in EGFR-deficient mice with phenotypes ranging from a peri-implantation lethality to a weaning-age lethality due to abnormalities in multiple organs (9), TGFβ1-deficient mice with phenotypes ranging from preimplantation to weaning-age lethals (reviewed in (2, 10)), and GEM models for cystic fibrosis with the presence (11) or absence (12) of the lung disease. Another cystic fibrosis model was shown to have background-dependent differences in the severity of intestinal obstruction (13).

2. Methods

2.1. Background-Dependent Variability in Penetrance and Expressivity

The presence of phenotypic variability in penetrance and/or expressivity is nearly always due to the knockout allele being present on a mixed genetic background. This was the case for the fibronectin, keratin-8, cystic fibrosis, and TGFβ1 examples mentioned above. Generation of congenic strains for the mutant allele usually leads to a more consistent phenotype. Incomplete penetrance and variable expressivity in GEMs can also result from environmental influences. Our experience with TGFβ- and SMAD3-deficient mice will primarily be used here to illustrate these points.

We have maintained our Tgfb1 knockout strain on a mixed genetic background of 129/SvJ and CF-1. On this mixed background only half of the homozygous mutant animals are born, and they subsequently die of a weaning-age autoimmune disease (14). Maintenance on a mixed background is required to prevent loss of
the autoimmune phenotype because on several inbred backgrounds nearly all $Tgfb1^{-/-}$ animals die of preimplantation lethalities. Similarly, on a different mixed genetic background of 129/Sv, C57BL/6, and NIH/Olac, Akhurst's group found that half of $Tgfb1^{-/-}$ animals die of a yolk sac developmental defect (15), and that survival to birth ranges from nearly 0% on a C57BL/6J/Ola background to about 80% on an NIH/Ola background (16). In both $Tgfb1$ knockout strains, there is clearly an incomplete penetrance of phenotype.

When rescued by rendering the mice immunocompetent ($Tgfb1^{-/-}Rag2^{-/-}$ mice) we have found that a TGFβ1 deficiency leads to a colitis-associated colon cancer at 100% penetrance when the mice are of mixed but primarily 129/Sv background (approximately 85% 129 and 15% CF-1) (17); whereas, on a primarily C3H background (approximately 85% C3H and 15% 129 plus CF-1) no colon cancer is detectable in $Tgfb1^{-/-}Prkdc^{scid/scid}$ immunodeficient mice, even though they have colitis (18). Consequently, there is considerable genetic background influence on TGFβ1-deficient GEM phenotypes, most but not all of which express variable penetrance and expressivity.

The $Tgfb2$ knockout strain has been maintained on a mixed genetic background of 129 and Black Swiss, and these mice die from midgestation stage to birth and have severe heart, skeletal, ear, and eye defects (19–22). In general, nearly all of the many congenital defects in these mice present with incomplete penetrance and variable expressivity. Although at first glance it might seem that this variability would be problematic for determining the mechanisms underlying specific defects, it has been useful for correlating the extent of changes in pathways upstream and downstream of TGFβ2 with the variation in penetrance and/or expressivity of the defect. This type of analysis can therefore provide a degree of built-in experimental control.

The $Tgfb3$ knockout strain is also maintained on a mixed genetic background of 129 and Black Swiss. These mice die within the first 18 h after birth due to a completely penetrant cleft palate with widely varying expressivity; whereas, in C57BL/6 congenics the expressivity of the cleft palate is very high with low variability. Hence, the developmental progression of the defect could more fully be characterized on the mixed background (23).

Environmental influences can also contribute significantly to incomplete penetrance and variable expressivity of GEM phenotype. Although $Tgfb1^{-/-}Rag2^{-/-}$ mice can develop colon cancer at 100% penetrance (17), a drop to 0% penetrance occurs in the absence of all enteric flora (germ-free mice), and this drop in penetrance is maintained if the gut flora of the mice are reconstituted with *Helicobacter hepaticus*-free flora (18). Similarly, $Smad3^{-/-}$ mice on a 129 background have recently been found to develop colon cancer if *Helicobacter hepaticus* is present (24).
2.2. Modifier Genes

Genetic background-dependent differences in the expressivity of the intestinal obstruction of a GEM model for cystic fibrosis has been used to screen for modifier loci that support increased longevity (13). About 30% of F2 progeny from 129/Sv × CD1 F1 intercrosses were found to live at least 6 weeks while the rest died of bowel obstruction by 2 weeks of age. Polymorphic markers for the two backgrounds were screened for association with increased survival age. A locus on proximal chromosome 7 was identified. Complementary studies on chloride conductance indicated that the increased longevity correlated with upregulation of a compensating chloride current. Guilbault et al. (25) has reviewed cystic fibrosis GEMs.

Several studies by Akhurst have identified loci for genetic modifiers of the embryonic yolk sac lethality in TGFβ1-deficient mice. As mentioned above, the penetrance of this lethality varies considerably between the C57BL/6 J/Ola and NIH/Ola backgrounds (16). Their first genetic screen was based upon survival of F2 Tgfb1−/− animals to birth and yielded a locus on NIH/Ola chromosome 5 (Tgfbkm1NIH) that accounts for three-fourths of the survival effect. In a follow-up study it was determined that a second locus on chromosome 12 (Tgfbkm3NIH) modified the first locus to increase its ability to support survival to birth (26). A third study was based upon data that the TGFβ1 deficiency on a 129S2/SvHsd background had a higher incidence of survival to birth (30%) than when on a C57BL/6NTac (0%) (27). Analysis of survival to birth in F1 progeny of reciprocal crosses revealed differential maternal imprinting effects in which F1s of C57 mothers have a much higher survival-to-birth rate. In addition, an F1 intercross genome scan revealed a chromosome 1 modifier (Tgfbkm129) which accounts for 90% of the survival, independent from the maternal effect.

2.3. Elimination of Gene Polymorphism and Flanking Gene Problems

The flanking gene allele problem was originally addressed by Smithies and Maeda (28). They were concerned about phenotyping GEMs for complex genetic diseases (atherosclerosis and essential hypertension in their case) when there could be allelism in trait-modifying genes flanking the targeted gene. This is especially problematic when one is using GEMs to prove causality for a candidate gene drawn from human linkage studies that show disease association with a particular chromosomal region. In this case one does not know whether it is the targeted candidate gene, a particular allele of another gene in that chromosomal region, or a complex interaction between the two that phenocopies the human disease trait. They suggested that if heterozygous F1 offspring of germline chimeras were crossed with their wild-type F1 littermates, rather than being selfed, then at the F2 generation, wild-type control animals could be screened for those that are non-allelic (129) at the flanking regions (see Fig. 23.1). This scheme solves the flanking gene problem, but not potential polymorphic differences in unlinked regions of the genome.
Wolfer et al. (29) reviewed breeding schemes designed to control for the widely recognized problem of genetic background differences in general. They discussed the Banbury Conference on Genetic Background (30) recommendation of producing co-isogenic strains (e.g., co-isogenic 129 and congenic B6) followed by phenotype analysis on F1 hybrids of the two strains (see Fig. 23.2). This basically solves all polymorphism problems except for the flanking gene problem, but is quite expensive as it requires maintenance of two strains for each targeted gene. However, they discuss breeding schemes to screen for flanking gene effects that may contribute significantly to a knockout phenotype (see Fig. 23.2). Finally, they point out that another solution to these

![Breeding scheme diagram](image)

* indicates targeted modification

Fig. 23.1 Breeding scheme to eliminate flanking allele differences between experimental and control animals. F1 offspring from both 129 gametes are genetically identical except for the genetically modified allele. Selfing the wild-type F1 animals (second row, right) yields F2 animals, some of which have only 129 alleles in the flanking region (bottom row, far right animal). Selfing the F1 animals that carry the genetic modification will yield homozygous mutant animals which will always be non-allelic in the flanking region. Finally, crossing the wild-type and heterozygous F1 animals will yield heterozygous animals nearly half of which will be non-allelic in the flanking region. Arrows (third row) indicate experimental, heterozygous, and control animals that do not have the flanking gene problem, though each animal will have a different allelic mixture in non-flanking genome regions.
Fig. 23.2. Breeding scheme to eliminate all genetic background differences except in the flanking region. Congenic and co-isogenic strains are generated either by crossing germline chimera to C57 or 129 mates, respectively. Congenic strain can be generated through multiple backcrosses or by speed congenics procedures. Crossing the co-isogenic and congenic strains will yield animals of genetic backgrounds similar to that of F1 hybrids except for the region flanking the modified gene. Hence, experimental, control, and heterozygous animals can be compared with identical (except for flanking region) genetic backgrounds. Flanking allele effects can be tested by phenotypic comparison of the wild-type and heterozygous animal with the modified gene on the C57 congenic chromosome.
problems is to use conditional knockout alleles which allow for comparison between the “on” and “off” states in animals in which the genetic backgrounds are completely identical.

In general, keeping in mind the recommendations of these two papers (28, 30) during phenotype analysis of genetically modified genes should allow the investigator to identify any major roles played by differences in genetic background.

2.4. Value of Initially Analyzing Null Phenotypes on a Mixed Genetic Background

Knockout mice are usually generated by crossing a germline chimera, in which the knockout allele is on a 129 background, with an animal of any desired background. With speed congenic techniques congenic animals can be generated within 1–2 years (31, 32). This is the case even though there are breeding schemes to test for or eliminate flanking gene effects. The resulting offspring are then intercrossed to generate homozygous mutant animals that will either be inbred 129 strain or F2 generation mice with a 50/50 mixture of 129 and the other desired background. With the exception of doing the gene targeting in an ES cell of another background, putting the targeted allele on a background other than 129 requires a standard backcrossing scheme. For convenience, the first homozygous mutant animals can be produced both on a 129 inbred background and on a mixed background. If one produces experimental and control mice from Fn generations derived from the mixed background strain, each experimental and control animal will have a different mixture of the two original backgrounds, assuming that the mixed strain is maintained as an advanced intercross line (33). The question arises as to which background is better for phenotype analysis. Obviously, the more backgrounds the better; however, with limited resources, choices must be made.

We suggest that the background most likely to provide the widest range of phenotypes is the mixed background. This is due to the considerable background dependence of knockout phenotypes discussed previously. On a mixed background this phenotype variation could often play itself out as incomplete penetrance and variable expressivity. These, in turn, would likely decrease as the targeted allele were moved to a more inbred state. Consequently, the mixed-background knockouts potentially display a wide range of phenotypes, and those phenotypes with incomplete penetrance and variable expressivity would be candidate phenotypes upon which a modifier gene search could be based. If resources allow phenotype analysis of a second knockout strain, the 129 strain knockout would be most appropriate because it would reveal which of those mixed-strain phenotypes may have 129 strain modifiers, and, by elimination, which phenotypes may have modifiers on the other strain of the original mixture. It is for these reasons that we always make our first phenotype screen on a mixed strain. The TGFβ1 knockout mouse provides an important case in point. Had the decision been made to put the knockout
allele on the C57BL/6 or 129 background, embryonic lethalities would have precluded discovery of the important roles played by TGFβ1 in autoimmunity (34), platelet aggregation (35), colon cancer (17), and cardiac hypertrophy (36).

2.5. Inbred Strain-Specific GEMs: Present and Future

As we have seen, there is significant background-dependent phenotypic variability found in GEM strains. Nonetheless, the power of combining the inherent phenotypic differences between inbred strains with GEMs is underutilized. Ideally one should be able to choose ES cell lines from the inbred strain(s) that are most appropriate for the phenotype/disease/defect to be investigated. Given advances in the understanding of embryonic “stemness” (37–39) procedures will likely be developed that improve the ability to generate strain-specific ES cell lines. Such cell lines could then be used to put “disease” genes, for example, into inbred strains with differences in susceptibility to that disease. This would be akin to studies that have used congenic lines for the ApoE KO gene to model and investigate aspects of hyperlipidemia (40–42). This direct ES cell approach would save time by avoiding backcrossing and it would eliminate the “flanking gene” problem.

With completion of the mouse genome sequencing project (43) it is now known that the nearly 500 traditional strains of inbred mice (http://www.informatics.jax.org/external/festing/search_form.cgi) represent from only one-third (44) to one-half (45) of the genetic diversity present in the Mus musculus species. Of the ancestral subspecies groups, domesticus, castaneus, and musculus, there is a disproportionately low representation of castaneus in the traditional strains. Hence, the value of mouse genetics for investigating complex disease would be greatly improved were a new set of inbred strains developed that more fully represented the full potential of mouse genetic diversity. To this end the Complex Trait Consortium (http://www.complextrait.org/) has initiated the Collaborative Cross (46) in which 1000 new recombinant inbred strains will be derived from a randomized cross of eight inbred strains that more fully represent the genetic diversity of Mus musculus (http://lsd.ornl.gov/mgg/projects/collabcross.html) than that presently available. The addition of these new RI strains will expand the availability of genetic background choices upon which to make GEMs.

3. Conclusion

The study of mouse genetics has taken an exciting step forward with the advent of gene targeting via homologous recombination in ES cells. The knockout mice are wonderfully informative because of the unexpected and wide-ranging phenotypes that can
result. Cognizance of the importance of genetic background on differences in knockout phenotype and of approaches for analyzing the genetics of those differences will broaden our understanding of the complexities of gene function.

References


