Effect of the Genetic Background on the Phenotype of Mouse Mutations

XAVIER MONTAGUTELLI
Unité de Génétique des Mammifères, Institut Pasteur, Paris, France.

Abstract. An increasing number of scientific articles report that the phenotype of a given single gene mutation in mice is modulated by the genetic background of the inbred strain in which the mutation is maintained. This effect is attributable to so-called modifier genes, which act in combination with the causative gene. The modulation of the phenotype can be major, as exemplified in the case of several mouse models of polycystic kidney disease. Because of the existence of inbred strains and the possibility of developing congenic strains, the effect of the genetic background can be analyzed in mice, including the identification of major modifier genes. Furthermore, by transferring a given mutation into different genetic backgrounds, mouse models can be manipulated with the aim of more accurately mimicking specific features of human diseases.

Medical genetics deals with two types of inherited conditions. Single-gene (so-called mendelian) disorders, such as cystic fibrosis or Duchenne muscular dystrophy, account for the largest number of genetic diseases, although the incidence of each disease is low or very low. In contrast, multifactorial traits, such as hypertension or insulin-dependent diabetes, are attributable to various associations of susceptibility genes, the effects of which are modulated by environmental factors. These definitions no longer reflect the actual pathogenesis of genetic diseases. Increasing numbers of publications report on the variability of phenotypes among patients who share the same molecular defect in a gene responsible for a monogenic disease (1).

The same observations have long been made for mice. There are a limited number of cases in which the phenotype of mice is entirely controlled by a mutation at the causative locus. One such case involves the albino mutation, which affects the tyrosinase gene. In the absence of tyrosinase, the coat color of mice is invariably white. At the other extreme, traits such as systemic lupus erythematosis or insulin-dependent diabetes are under the control of a large number of genes (2,3). In the majority of cases, the phenotype is associated with a single gene mutation but is modulated by the genotype at other loci. For example, the ob and db mutations (which affect the leptin and leptin receptor genes, respectively) are responsible for hyperglycemia and obesity in the C57BL/6 genetic background but induce overt diabetes in the related C57BL/Ks inbred strain (4,5). C57BL/6 mice that are heterozygous for the Apc\textsuperscript{Min} (multiple intestinal neoplasia) mutation, which is a germline mutation in the Apc (adenomatous polyposis coli) gene, develop an average of approximately 30 tumors throughout their intestines, whereas their heterozygous F1 progeny produced from crosses with AKR/J mice exhibit only an average of approximately 6 polyps (6). There are also several examples of kidney disease mutations for which the severity of macroscopic and histologic lesions is strongly influenced by genes from the genetic background (7–10). Two examples are presented and discussed below. Similar reports on mutations in which a gene was inactivated by homologous recombination have been published. The phenotypes induced by targeted disruption of the epidermal growth factor receptor gene ranged from peri-implantation death of embryos with a CF-1 background to survival for up to 3 wk after birth for animals with a CD-1 background (11). A modifier locus that influences the severity of the phenotype induced by inactivation of the Cftr (cystic fibrosis transmembrane conductance regulator) locus was mapped to mouse chromosome 7 (12).

The effects of the genetic background should be distinguished from two other sources of phenotypic variability. The phenotypes associated with a mutation can vary among genetically identical individuals. This variation can be attributable either to incomplete penetrance, when a fraction of the individuals carrying the disease-associated genotype fail to develop the abnormal phenotype, or to variable expressivity, when individuals carrying the disease-associated genotype exhibit phenotypes with various degrees of severity. In human populations, phenotypic variations are often attributable to genetic heterogeneity, with patients carrying different mutations in the same major causative gene or mutations in different genes and with the remainder of the genome also being variable. The effect of the genetic background can be invoked in the intermediate situation where phenotypic variations are observed among individuals who carry the same mutation in the same causative gene but differ in a number of other genes. A few of these genes control part of the phenotypic variation and hence are called “modifier loci.”
The identification of such modifier loci is very difficult in human populations for at least three reasons, as follows: (1) these loci control quantitative variations and not a qualitative trait, and their mapping uses quantitative trait locus (QTL) mapping statistical methods; (2) phenotypic variation may be controlled by more than a few genes, each of which has a modest or weak effect; and (3) the effects are likely to be influenced by environmental factors. One attractive option for linkage studies is the analysis of discordant sibling pairs who share the same genotype at the disease locus and experience similar environmental influences but are dissimilar at the modifier loci (1).

A variety of powerful genetic tools are available in mice, facilitating the genetic mapping of modifier loci. Inbred strains have been developed in large numbers, including strains derived from mice of different species (Mus spretus) or subspecies (such as Mus musculus castaneus) that offer a higher degree of genetic polymorphism. The genetic map of mice is almost as dense as that of human subjects, with more than 10,000 loci. Experimental crosses can be performed as desired, with a short generation time. Different strategies are available to manipulate the genome by transgenesis or homologous recombination, with a variety of refinements. Finally, the environment in which the animals are maintained can be very strictly standardized.

With most mouse mutations, the effects of the genetic background have been detected among the backcross or intercross progeny generated to map the causative gene. Considerable variations were observed in some cases, especially when an intersubspecific cross was used. I illustrate this point with the kat mutation, the phenotype for which is associated with polycystic kidney disease, progressive normocytic normochromic anemia, small testes with male sterility, and other features, including dwarfism and facial dysmorphism (13). To map the kat mutation, Janaswami et al. (13) crossed a homozygous female mouse with a Mus musculus castaneus male mouse to produce an F2 population. Homozygous mutant F2 mice, which were unambiguously identified by their facial dysmorphism, exhibited polycystic kidney disease with a highly variable degree of severity, as estimated from gross examinations, kidney/body weight ratios, and histologic features. Anemia, as measured by hematocrit values, was also quite variable. F2 mice were then genotyped for microsatellite markers spread along the mouse chromosomes, and appropriate statistical analyses (ANOVA and multiple regression analysis) detected three chromosomal regions associated with more or less severe kidney and/or anemia phenotypes (10).

A very efficient strategy to genetically map modifier loci is to select, among the segregating progeny, the individuals who exhibit the most extreme phenotypes. In fact, for QTL mapping, individuals with phenotypes >1 SD from the mean represent 33% of the total population but contribute approximately 81% of the total linkage information (14). For example, this strategy was used by Woo et al. (15) in the mapping of loci modifying the phenotype of the polycystic kidney disease (pcy) mutation. Two major modifier loci were identified by genotyping only 46 mice, of a total of 3105 F2 progeny. The analysis of more mice increased the logarithm of odds score at the two QTL but did not reveal any other modifier loci.

A very powerful genetic tool available in mice is the development of congenic strains, with the aim of transferring a mutation or a chromosomal segment from one inbred genetic background to another. The crosses involved are presented in Figure 1. They consist of backcrossing the original (donor) inbred strain with the recipient inbred strain at least 10 times. The locus (or the chromosomal region) to be transferred is selected for at every generation. After 10 backcrosses, the only region of the genome that originates from the donor strain is a single chromosomal segment containing the locus selected for. At that stage, heterozygous mice are intercrossed to produce mice homozygous for the introgressed region. The same strategy can be used to transfer each QTL identified in the mapping cross from one parental strain into the genetic background of the other strain. Although this process is long (approximately 3 yr), it yields closely related inbred strains that differ in only

![Figure 1. Development of a congenic strain. Mice heterozygous (or homozygous) for the mutation with inbred background A are crossed with mice of strain B. Among the F1 animals, mice heterozygous for the mutation are selected either by progeny testing or by genotyping with linked molecular markers (unnecessary if the parent from strain A was homozygous). These F1 animals are crossed with strain B to produce N2 progeny, among which only mice heterozygous for the mutation are used for the next generation. The same cycle is repeated until N10 mice are obtained. Heterozygous N10 mice are crossed to produce mice homozygous for the mutation, with a genetic background very close to strain B.](image-url)
one small region of their genomes. A more sophisticated protocol, called “speed congenics,” has been proposed to shorten the generation time by counterselecting the alleles from the donor strain (16).

Congenic strains might sometimes produce misleading results because of the chromosomal segment that is introgressed together with the gene of interest. After 10 backcrosses, the fragment from the donor strain is still approximately 20 cM in length and contains approximately 1.3% of the genome (300 to 1000 genes). If the trait analyzed is controlled by a large number of genes, as is the case, for example, for neural or behavioral processes, it can be difficult to distinguish between the effect of the gene of interest and the effects of genes located in the flanking chromosomal segment (17). One way to minimize this effect is to derive a series of partially overlapping congenic strains. If these strains are derived independently, the residual segments from the donor strain are different and the genes that are not very tightly linked to the mutation are not present in all strains. By comparing the phenotypes of the different congenic strains, it is possible to distinguish between the effect of the mutation itself and the effects of a number of the flanking genes.

Despite this limitation, congenic strains are invaluable genetic tools in several respects. First, the phenotype of the same mutation can be studied in different genetic backgrounds. Second, these strains allow confirmation of the existence of a QTL in a candidate region identified by linkage analysis in a two-generation cross. They are also helpful for study of the pathophysiologic features associated with a given genomic region. Because the congenic strain differs from the recipient strain by a small chromosomal segment selected for during the backcross generations, any phenotypic differences observed between these two strains can be attributed to this segment. Congenic strains are powerful tools to refine the location of a QTL by generation of several “subcongenic” strains carrying smaller, partially overlapping, chromosomal segments (such subcongenic strains are produced from a two-generation cross from the original congenic mice). This strategy can narrow the candidate interval to <1 cM, a resolution that cannot be approached by linkage studies. An additional advantage of congenic strains is their usefulness for the detection of loci with minor phenotypic effects. After a major modifier locus has been mapped and introgressed in a congenic strain, a two-generation cross between this congenic strain and the donor strain can be generated. Because the major modifier locus is not polymorphic in the cross, its major effect no longer hides the more modest effects of other loci. Finally, congenic strains are ideal for testing epistatic interactions between modifier loci, because homogeneous populations can be produced for each genotypic combination and their phenotypes can be compared. Congenic strains are therefore ideal tools for the analysis of interactions between genes responsible for mendelian traits and the genetic background, although a single congenic strain might not always be sufficient to yield clear conclusions.

The fact that background genes can influence the phenotype induced by a single gene mutation (a spontaneous mutation, a transgene, or a gene inactivated by homologous recombination) should not be surprising. It is likely that few genes work in such a way that the effects of their alteration cannot be modulated by other genes. Metabolic pathways, developmental processes, and the levels of gene expression result from protein-protein and protein-DNA interactions. The effects of the modification of one component of these networks are influenced by the other components. In other words, a mutation induces a phenotype in a particular genomic context.

The importance of these interactions has been particularly emphasized by researchers working with transgenic and knockout mice. Transgenic mice are often obtained from F1 hybrid female mice mated with F1 hybrid male mice. The transgenic founders (usually called F0) are in fact genetically F2 mice from crosses between two inbred strains. If brother-sister mating is used to maintain the transgenic line, starting with these F2 mice (as has often been performed), then a new inbred strain (called a recombinant inbred strain), which is a hybrid between the two parental strains, is produced. This strategy has a number of disadvantages, particularly the fact that the transgene is maintained in an original genetic background that cannot be reproduced and for which no reference data are available. The same is true for knockout mice that have been maintained on a mixed 129×B6 genetic background. Several recommendations regarding the breeding of genetically engineered mice, to avoid these problems, have been published (18). Most importantly, the mice should always be produced on a pure inbred background or repeatedly backcrossed to a standard inbred strain to create a congenic strain. For knockout mice, the availability of embryonic stem cells derived from various inbred backgrounds would be very valuable.

Several general considerations should be kept in mind when the effects of modifier genes are being considered. First, there is no ideal genetic background that would be appropriate for every class of phenotypes. This is particularly true for neurologic phenotypes; 129/Sv mice show an absence of the corpus callosum and perform poorly on memory tasks, whereas C57BL/6J mice exhibit poor performance in passive avoidance tests (19). The second point is that taking advantage of the phenotypic differences associated with different genetic backgrounds provides a way to make mouse models more closely resemble particular characteristics of human diseases, not only in the degree of severity but also in the type of lesions. Because of the genetic heterogeneity that is a feature of most human genetic diseases, human patients present a range of phenotypes that can be quite wide. In contrast, a mouse mutation maintained in an inbred background produces much more homogeneous phenotypes, which can be compared with those of a subset of patients (Figure 2). By transferring the mutation into a different genetic background, the narrow window of phenotypes can be moved toward more or less severity.

The identification of modifier genes is a powerful means to dissect the pathophysiologic processes underlying diseases and to discover new biologic mechanisms. In their study of the kat mutation, Upadhya et al. (10) identified three chromosomal regions that influence the severity of polycystic kidney disease (as measured by the kidney/body weight ratio) and anemia (as
measured by hematocrit values). It is generally assumed that anemia is a consequence of renal failure resulting from decreased production of erythropoietin by the affected kidney. One of the QTL was found to have effects on both traits. The second QTL affected only the kidney/body weight ratio, whereas the third was associated only with anemia. This finding indicates that there are genes that can independently modulate the severity of each class of symptoms. With the development of congenic strains for each of these QTL, it will become possible to identify the genes themselves and to better understand their biologic effects.

We have clearly demonstrated that mice (and rats) offer all necessary genetic tools (including genetic, physical, and expression maps) for dissection of the genetic control underlying the effects of the genetic background. This medical genetics field will surely be very active in the coming years, now that all of these tools are available. In a few cases, it has been possible to identify the actual modifier gene, using a candidate gene approach. For example, the effect of the genetic background on the number of polyps that develop in the intestines of Apc

\[\text{Min}^1\] mice could be attributed primarily to a single locus called \textit{Mom1} (modifier of \textit{Min}), localized on distal chromosome 4 (20). The \textit{Pla2g2a} gene, which encodes secretory phospholipase A2, was mapped to the same chromosomal region. On the basis of similar map positions, the expression of \textit{Pla2g2a} in the intestines, and the correlation between the level of \textit{Pla2g2a} expression and the susceptibility to intestinal neoplasia, this gene was proposed as a strong candidate for being the \textit{Mom1} locus (21). This was later confirmed by Cormier \textit{et al.} (22) with mice transgenic for a cosmid containing a functional copy of the AKR allele of the \textit{Pla2g2a} gene in a C57BL/6 background [the \textit{Apc

\[\text{Min}^1\] mutation induces 3 to 4 times more polyps in the C57BL/6 background than in the (C57BL/6×AKR)F1 background]. Mice overexpressing \textit{Pla2g2a} \textit{AKR} exhibited a twofold reduction in tumor load and a significant reduction in tumor size, compared with those carrying a single copy of the resistance allele \textit{Pla2g2a} \textit{AKR} (22).

What will be the effect of mouse studies on the field of human genetics? It is hoped that the identification of modifier loci in mice will indicate good candidates for genes that control the severity of diseases in human subjects (23). However, it is likely that modifier genes are not always conserved between distant species such as mice and humans. In that case, the gene identified in mice might indicate a metabolic pathway or a pathophysiologic mechanism that is more likely to be conserved between species. Finally, when chromosomal segments are shown to affect the phenotype, although the gene itself is not cloned, chromosomal homologies between mice and human subjects (which currently include approximately 75% of the genome) should suggest candidate regions in human subjects to examine for modifier loci.

**References**


