Implications of Mouse Genotype for Phenotype

Judy M. Hickman-Davis

Genetic engineering of inbred mice offers the capability to separate interactions of multiple genes that control host resistance to pathogens. An understanding of the negative impact that genetic variability of mouse substrains can have on data is necessary for the design of experiments to dissect out complicated gene interactions.

The study of the host response to infectious diseases allows for the identification of components of the immune system that are essential for host defense. Characterization of these components then allows the development of therapies that may augment natural host resistance. Virulence factors of the invading organism and genetic factors that influence the ability of the host to mount an appropriate immune response to the pathogen together determine the outcome of infection. Human resistance to infectious disease is a multifactorial trait involving multiple gene interactions and strong environmental influences. Unfortunately, studies of human populations are extremely difficult and inherently biased by population heterogeneity, linkage disequilibrium, and the lack of suitable controls. Because of this, the inheritance of genetic resistance and susceptibility to pathogens has been studied using laboratory mice. As early as 1933, Leslie Webster used quantitative measurements such as survival times and pathogen tissue loads in inbred mouse strains to identify dominant resistance factors to Bacillus enteritidis (now Salmonella enteritidis) (13).

Historically, the concept that alterations within a single gene would lead to a single altered biochemical process led early biologists and physiologists to expand the role of the mouse in research. The inbred mouse provided a genetically well-defined model with the advantage of high reproducibility not available with standard outbred mouse strains. The primary advantages to the laboratory mouse are 1) the existence of commercially available inbred strains allowing for the limitation of genetic variability between animals; 2) extensive genetic homology between mice and humans (181 conserved linkage groups, ~90% of the mouse genome) (9); 3) ease of controlled breeding; 4) availability of reagents; and, more recently, 5) the ability to genetically engineer mice.

Much of the initial characterization of the immune system was a result of the identification and use of spontaneous mutations for the study of immune regulation of autoimmunity (Foxn1nu, commonly referred to as “nude”), immunodeficiency (Prkdcscid, severe combined immunodeficiency), and lymphoproliferative diseases (Fas(lpr), lymphoproliferation). Originally, researchers working with inbred mouse strains relied on these spontaneously occurring mutations or those

J. M. Hickman-Davis is Research Assistant Professor in the Department of Anesthesiology, University of Alabama at Birmingham, Birmingham, AL 35249-0006.
induced by exposure to chemicals or radiation to identify specific genes and their functions. As recently as 1998, a missense mutation in the Toll-like receptor-4 gene (Tlr4) was identified in C3H/HeJ mice, and this has contributed significantly to the understanding of lipopolysaccharide-dependent signal transduction (8). The identification of coisogenic mice, or mice that differ from the original inbred strain of mouse by a single mutation, has allowed for direct comparison of these mutants using wild-type littermates as controls. Fabricated congenic mutant inbred strains were also created to approximate the coisogenic strains. Congenic mice are generated using traditional breeding methods to backcross a mouse of strain 1 to carry an allele of interest from a mouse of strain 2. After 10 backcross generations, the genome will be >99% strain 1 with the allele from strain 2. Some passenger genes, particularly genes closely related to the selected marker, will remain with the allele of interest from strain 2; therefore, congenic mice differ from coisogenic mice in that they contain the gene of interest as well as closely linked genes. Although single gene effects are of importance, the overall resistance of an animal to any particular infectious agent is generally the result of the effect of genes at multiple loci.

“The development of spontaneous mutations with biological significance is rare.”

Studies to separate these multiple gene effects generally use recombinant inbred mouse strains. Recombinant inbred mouse lines represent inbred lines derived from the F2 progeny of parental inbred strains. These lines are inbred for 20 or more generations, and as a result each recombinant inbred line is considered to be homozygous at every genetic locus. These mice represent a special subset of inbred mice used for the precise chromosomal localization of alleles linked to host resistance phenotypes or “mapping.” Analysis of an infection phenotype between multiple recombinant inbred strains should reveal a pattern of continuous variation, with the identification of resistant, intermediate, and susceptible strains. Once resistant and susceptible strains of mouse are identified, these strains can then be crossed to identify the mode of inheritance involved. Mapping of a particular phenotypic trait to a precise chromosomal location is achieved by comparing the distribution of alleles within a recombinant inbred strain. If the phenotypic difference is associated with a single chromosomal region, it provides a map position for the gene of effect and strong evidence that the phenotype being studied is controlled by a particular genetic locus.

Experimental approaches using recombinant inbred mice have significantly impacted our understanding of complicated protective immune mechanisms. Attempts to separate the contributions of multiple genes or quantitative gene effects have been made with respiratory mycoplasmosis. Murine respiratory mycoplasmosis due to Mycoplasma pulmonis is the animal model of choice for human respiratory mycoplasmal disease. Studies involving 17 inbred mouse strains of differing Bcg and H-2 haplotypes have demonstrated that disease resistance to mycoplasmas is a complex trait and that immune responses and lesion development may be independently controlled by different genetic mechanisms (1).

Although resistance to disease is usually the cumulative result of multiple gene effects, recombinant inbred mice have also been used successfully to identify some large single gene or qualitative gene effects. For example, resistant A/J and susceptible C57BL/6J mouse strains have been crossed to identify a locus variously named Bcg, Ity, or Lsh. This locus generates Nramp1 (natural resistance-associated macrophage protein), which is located on chromosome 1, and is responsible for resistance to the unrelated intracellular parasites Mycobacterium bovis, Salmonella typhimurium, and Leshmania donovani. Likewise, the Lr gene on chromosome 2 was identified using susceptible BALB/c and resistant C57BL/6 strains and is responsible for resistance to Listeria monocytogenes (7).

The development of spontaneous mutations with biological significance is rare. Likewise, chemical- and radiation-induced mutations are not predictable events. Since the 1980s, gene targeting has allowed for the generation of mice carrying specific gene mutations. Techniques for gene addition or disruption (gene knockout) include homologous recombination, pronuclear injection, and viral transgenesis. Targeted mutation (homologous recombination) allows for the introduction of a defined genetic mutation at a specific site within the genome. In this approach, a target vector is designed with a selectable marker flanked by sequences homologous to the targeted gene. The vector is then introduced into pluripotent embryonic stem (ES) cells by transfection. The homologous flanking sequences permit insertion of the vector into the gene target, and successfully targeted cells are identified by their selectable marker. Gene-targeted ES cells are introduced into developing mouse embryos, where they contribute to both somatic and germ tissues of the newborn. The results of this will vary from the complete loss of gene function to subtle changes in gene expression.

The majority of targeting experiments use ES cells derived from the 129 inbred mouse substrains because these cells can be manipulated in culture and still remain competent to repopulate the mouse germ line. Unfortunately, there is a wide range of genetic variability within 129 substrains and the ES cells derived from them. Three lineages of 129 substrains have been identified according to deliberate outcrossings to introduce new alleles: 1) the parental substrains, 2) the Steel substrains, and 3) the Ter substrains (Fig. 1) (12). To maximize the frequency of homologous recombination, the substrain genomic library from which a targeting construct is prepared should match the substrain of the ES cells into which it is introduced. Moreover, if the mutation is to be maintained on the 129 background, the resultant chimeras should be backcrossed to their genetically matched substrain. Because gene segregation can occur at multiple alleles during the homologous recombination event, initial crossings of these mice will produce progeny that are only approximately coisogenic to the parent strain. To produce a true coisogenic 129 knockout strain, these mice must be backcrossed onto the parental strain in accordance with the percent match of...
the ES cell line and the parental substrain.

The use of these techniques in the development of mouse models has led to the characterization of many genetic traits affecting host susceptibility to bacterial and viral infections. A prime example of this is the host resistance Nramp1 gene. Nramp1 was mapped as a dominant trait to chromosome 1 using recombinant inbred mouse strains. A causal relationship between Nramp1 and resistance to intracellular pathogens was then confirmed using both transgenic (gain of function) and gene-targeting (loss of function) technologies. The locus of this protein was also cloned, and the protein was functionally characterized. Nramp1 is a highly conserved transmembrane protein thought to play a role in the transport of cations into the phagolysosome (4). Currently, a “mouse-to-man” approach is being used to identify important gene regions in humans that correspond to the murine Nramp1 alleles and susceptibility or resistance to mycobacterial infections (4). In these studies, gene cloning and congenic mouse models have proven invaluable for studies with human populations in endemic disease areas. In addition, gene-targeting technologies are now being applied to bacteria in an effort to characterize virulence factors attributable to the pathogen itself. Homologous recombination was used to knock out the sodC gene encoding sequence for Cu,Zn superoxide dismutase in the intracellular pathogen S. typhimurium. These studies identified Cu,Zn superoxide dismutase as an important part of the bacterial protective mechanism against host defense. sodC bacterial mutants were used in combination with Nramp1-positive and -negative congenic mice to fully characterize host pathogen interactions (2). The identification of an Nramp1 homologue in Mycobacterium tuberculosis suggests an exciting new area for research into the role of Nramp1 in bacteria.

It is generally accepted that some initial studies of a newly created transgenic mouse may be carried out with mice of a mixed genetic background. However, if these mice are not backcrossed but are instead maintained (continuously inbred) at this initial stage, the mixed genetic background will become fixed in the strain, making the provision of genetically matched controls an impossibility. Likewise, new combinations of alleles may alter phenotype or interfere with expression of the targeted gene. If preliminary studies are to be performed using mice of mixed genetic background, larger numbers of mice should be examined to allow for variations in phenotype due to differences in genetic background (11). There is little argument that strain background and genetics impact the outcome of infectious disease (7, 9). Several reviews have recently outlined specific recommendations concerning the use of transgenic mice and the difficulty in identifying suitable controls (11). Some of these guidelines should be directly applied to the development of experimental models of infectious processes: 1) inbred mice should be used whenever possible; 2) publications must contain accurate, detailed information regarding the genetic background of the mice studied and the gene manipulated; and 3) mutations should be maintained in congenic lines to protect against phenotypic changes that result in the loss of targeted loci. It should be remembered that a congenic mouse strain generated from mice with unrelated genetic backgrounds could be expected to be 99.9% derived from the host only after 10 generations of backcrossing.

Although controlling for subtle variation within the gene pool of closely related mouse substrains may seem time consuming and pedantic, these minor alterations within the genome can have major impacts on mutant gene expression. A classic example for this concept is the diabetic (Lepr<sup>db</sup>) mouse mutation. C57BL/KsJ-Lepr<sup>db</sup> mice develop obesity and are severely diabetic, making them a useful model for

FIGURE 1. 129 substrains are easily identifiable by coat and eye color; however, strain designations are very similar and technically complicated, contributing to the confusion surrounding embryonic stem cell culture lineages. The nomenclature for 129 substrains listed above was recently revised: 129/ReJ is now 129P1/ReJ; 129/J is now 129P3/J; 129/Sv<sup>4+<br> Tyr-c Mgf SI</sup>/+ is now 129S1/Sv<sup>+/+</sup>Tyr-c Mgf SI/+; and 129/SvEms-<sup>+</sup>Ter /J is now 129T2/SvEmsJ. Adapted from Simpson et al. Nat Genet 16: 19–27, 1997.
human type II diabetes, whereas the closely related substrain C57BL/6-Lepr<sup>db</sup> develops obesity and insulin resistance but does not develop permanent diabetes. Similarly, the lymphoproliferation (Fas<sup>sp</sup>) mutation in the Fas-encoding gene has been used as a model for connective tissue diseases and was developed using the MRL/Mp mouse strain. MRL/Mp-Fas<sup>sp</sup> mice develop early-onset lymphadenopathy and glomerulonephritis and have shortened life spans, whereas DBA/2-Fas<sup>sp</sup> mice do not even develop lymphadenopathy. The importance of this becomes obvious when one considers that the Fas<sup>sp</sup> mutation occurred spontaneously and was identified primarily on the basis of the lymphadenopathy phenotype, i.e., this mutation might never have been discovered if it had originally occurred on the DBA/2 background. These examples demonstrate the dramatic effect that background genes can have on the phenotypic expression of a single mutant gene (10).

Inbred mouse strains have been reported to evolve “quickly,” with a large degree of genetic divergence being recorded between different but closely related strains. Population genetics predict that the average rate of genetic diversity within mammalian populations is ~9%, meaning that 9 out of every 100 gene loci are occupied by different alleles. The calculated diversity for inbred mice is 33%, with the observation that the distribution of alleles among inbred strains is nonsimple and nonrandom (6). The generation of transgenic mice through targeted mutations may create a selective pressure for secondary compensatory mutations. Transgenic mice with mixed genetic backgrounds are especially susceptible to selective pressures whereby deleterious mutations might be lost. Likewise, the possibility that genetic contamination may contribute to the genetic diversity of inbred lines must always be considered.

Genetic contamination of mouse strains can occur in any commercial or private environment and can have serious implications, both personal, from the standpoint of professional advancement, and financial, in terms of wasted grant money and time. In the early 1980s, Kahan and colleagues, cancer researchers at the Universities of Wisconsin and Minnesota, published a letter in Science (5) reporting that there were major differences in the expression patterns of histocompatibility antigens and an isozyme of glucose phosphate isomerase within groups of inbred BALB/c mice from the same commercial supplier. These markers are commonly used to identify that inbred mouse strains are indeed genetically identical at every locus. Interpretations of data obtained from difficult transplant procedures, such as those conducted by Kahan et al., are dependent on the fact that mice are genetically compatible. The end result of this particular case of “genetic impurity” was the disruption of research, loss of time and money by at least three research institutions, and a lawsuit against the commercial animal vendor (3). Currently, genetic monitoring is considered a standard procedure for most of the larger commercial animal vendors.

The ability of the host to resist infection is highly complicated and controlled by multiple genes. The inbred mouse stands as a reproducible model for identification of genes involved in important mechanisms of host defense. The ability to control for genetic background has opened new and unexpected avenues for identifying single gene effects, and the high degree of homology with humans offers a strong possibility for the identification of genes that are important for the control of human disease. This identification is crucial for the understanding of basic immune functions and the development of new therapeutic approaches. Although the ability to manipulate the mouse genome has become more accessible to the average researcher, the nomenclature for the cataloging of mouse strains, the plasticity of the mouse genome, and the complicated vernacular surrounding the use of inbred mouse strains has become more difficult. To avoid results in which the normal strain phenotype conflicts with that of the target gene, researchers need to have a clear understanding of classic and molecular mouse genetics and know the particular characteristics of a strain before they choose it as a model.

I would like to thank Drs. I. Davis, J. R. Lindsey, and S. Matalon for their support and helpful discussions.

I am currently supported by Grant RR-00149 from the National Institutes of Health.

References