The Problem of Passenger Genes in Transgenic Mice
Aldons J. Lusis, Janet Yu, Susanna S. Wang

Abstract—Gene targeted or transgenic mice are frequently created on one genetic background and the mutation then transferred to a second genetic background by a series of backcrosses, with selection for the mutation at each generation. This is the same experimental design as that used in the production of classical congenic strains in which a genomic region from a “donor” strain is transferred to the background of a “recipient” strain by backcrossing and selection for ten or more generations. While the strategy largely results in the replacement of the donor background with the recipient background, the region flanking the selected gene remains of donor origin. Thus, the genes in this flanking region are “passengers” that travel with the selected gene, and when comparing transgenic (or knockout) mice with non-transgenic littermates, allelic differences in the passenger genes can conceivably influence the traits of interest. It now appears that this passenger gene effect may be an important confounder in studies of cardiovascular and metabolic traits that are influenced by dozens of common variations present among inbred mouse strains. Here, we discuss the problem and some approaches to avoid it.

Keywords: Gene targeting ■ genetic background ■ congenic strains

A recent issue of Arteriosclerosis, Thrombosis and Vascular Biology carried reports from 3 different laboratories which studied atherosclerosis in mice carrying a null mutation of the cholesterol transporter ABCG1. In 2 studies, lesions decreased, and in 1 study lesions increased.1–4 Several subsequent studies of atherosclerosis in ABCG1 transgenic or knockout mice have not eliminated the confusion.5,6 There are many similar contrasting findings with transgenic mice in the literature. For example, a recent paper in the Journal of Clinical Investigation reported that mice carrying null mutations of the macrophage scavenger receptors CD36 or SRA did not have reduced atherosclerosis,7,8 in sharp contrast to previous studies.9–11 With a trait as complex as atherosclerosis of the macrophage scavenger receptors CD36 or SRA

minimizing the impact of a contaminating strain background

The importance of genetic background in the phenotype exhibited by a naturally occurring mutant, knockout, or transgenic mice is well established. One of the first instances in which this was noted was for the diabetes and obese mutations, which cause diabetes on a C57BLKS/J background but not a C57BL/6 background.12 Since then, many similar examples have been noted (eg,13–15).

When studying knockout or transgenic mice, therefore, investigators usually try to minimize background effects by breeding these mutations onto common genetic backgrounds for comparative studies (Figure 1). Gene targeted mice are usually initially created on a strain 129 background because embryonic stem cells from 129 mice are, for unknown reasons, much more efficient for the creation of targeted mice than stem cells from other strains. Because strain 129 mice breed poorly, the null mutation is generally transferred to another background, usually to a C57BL/6 background, by backcrossing for 5 to 10 generations. Similarly, for reasons of efficiency, transgenic mice are frequently made on a heterozygous background and then backcrossed to strain C57BL/6. The C57BL/6 background is particularly widely used, allowing comparisons of results with other engineered mice on a common background (reviewed in16).

When a transgenic or gene targeted mouse is crossed to a new, inbred genetic background, there is a 50% chance that a locus from the original background will be removed at each generation. For example, if strain 129 is crossed to strain C57BL/6, the first generation mice (N1) will consist of 50% 129 and 50% C57BL/6. If these mice are then crossed to C57BL/6, the progeny will consist on average of 75% C57BL/6 and 25% 129. The chance that a particular region from strain 129 (donor) is retained will decrease with additional backcrossing. After 10 generations (N10), the chance is only about 0.1%, and so the mouse is considered to be “congenic”, consisting of the selected locus from the donor 129 strain on the recipient C57BL/6 genetic background. If the selected locus is to be maintained in a homozygous state, 2 such N10 mice are intercrossed and the homozygous mice are used for further studies. In most cases, the congenic strain
is compared with the inbred background strain and any differences ascribed to the targeted gene. Similar logic applies to transgenes bred to a common background.

One potential problem with such comparisons is that donor chromosomal regions unlinked to the selected locus may remain by chance. Although the likelihood that a particular region is retained is small, there is a possibility that some donor DNA remains, particularly if fewer than 10 backcrosses are used. This problem can be addressed in part by intercrossing the congenic strain with the background strain and comparing homozygous, heterozygous, and wild-type mice at the locus of interest. With this design, any contaminating DNA should be spread at random among the 3 genotypic groups. (It is possible that, in some cases, there may be selection as well as chance that plays a role in the retention of unlinked donor chromosomal regions.)

A larger problem has to do with the region flanking the selected locus. Thus, unless separated by recombination, the flanking region from the donor strain will travel with the selected locus. For example, if a “passenger” gene is 10 cM from the selected locus, the chance that it will be retained after N generations of backcrossing is about 39% (Table 1). A 10-cM region of the mouse corresponds to roughly 20 million bp or about 200 genes.

As an example, Figure 2 shows the 129-derived regions for a mouse in which a locus of interest (a paraoxonase 3 null mutation, Pon3<sup>−/−</sup>) has been transferred onto an apolipoprotein E–null mutation by 6 generations of backcrossing (J. Yu and A. Lusis 2007, unpublished). The 129 regions were recognized using polymorphic genetic markers that distinguish the 129 and C57BL/6 genomes (Figure 2). The Apoe<sup>−/−</sup> mouse that serves as the background has been backcrossed to C57BL/6J for over 10 generations, but as can be seen, it retains about 22 megabases (Mb) derived from strain 129 at the Chr. 7 locus. The passenger region around the Pon3 gene, on chromosome 6, consists of about 42 Mb derived from strain 129. In addition, there is an unlinked region from strain 129 distal on chromosome 17, about 20 Mb in size.

Such passenger regions from strain 129, or any other donor strain, can lead to incorrect conclusions if allelic differences between the donor and recipient strains affect the phenotypes under study. Two inbred strains will likely differ at numerous loci, perhaps dozens, that affect complex traits such as atherosclerosis, adiposity, and plasma lipid levels. For example, over 30 loci for atherosclerosis or obesity or HDL levels have been mapped in various crosses in mice. If such an allelic variation occurs near the targeted gene, and is carried with the gene during backcrossing, the phenotype of the targeted mice will be attributable to both the effects of the targeted gene (if any) and the passenger gene.

Recent gene expression studies have revealed the remarkable level of quantitative genetic variation in transcript levels between inbred strains of mice. For example, a recent study of the genetics of gene expression in a cross between the common inbred strains C3H and C57BL/6 revealed that about one quarter of the genes expressed in liver or adipose tissue exhibit genetic variations that directly affect their transcript levels. This level of variation will be similar between other pairs of common inbred strains, such as 129, DBA, BALB/c, or AKR.

To illustrate the problem, we list in Table 2 the genes whose expression levels show common variations resulting from the passenger region flanking the Pon3 gene knockout shown in Figure 2. Altogether, about 40 genes exhibited variations in transcript levels, and a number of these have been found to affect atherosclerosis in transgenic studies, including paraoxonase, 2 lipoprotein lipase, and Cxcr6.

**Table 1. Probability (%) of a Passenger Gene Being Retained Along With the Differential Gene After n Backcross Generations**

<table>
<thead>
<tr>
<th>n</th>
<th>0.5</th>
<th>0.25</th>
<th>0.10</th>
<th>0.05</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>50.00</td>
<td>75.00</td>
<td>90.00</td>
<td>95.00</td>
<td>99.00</td>
</tr>
<tr>
<td>4</td>
<td>12.50</td>
<td>42.19</td>
<td>72.90</td>
<td>85.74</td>
<td>97.02</td>
</tr>
<tr>
<td>6</td>
<td>3.13</td>
<td>23.73</td>
<td>59.05</td>
<td>77.38</td>
<td>95.02</td>
</tr>
<tr>
<td>8</td>
<td>0.78</td>
<td>13.35</td>
<td>47.83</td>
<td>69.83</td>
<td>93.21</td>
</tr>
<tr>
<td>10</td>
<td>0.20</td>
<td>7.50</td>
<td>38.74</td>
<td>63.02</td>
<td>91.35</td>
</tr>
<tr>
<td>12</td>
<td>0.05</td>
<td>4.22</td>
<td>31.38</td>
<td>56.88</td>
<td>89.53</td>
</tr>
<tr>
<td>14</td>
<td>&lt;0.01</td>
<td>2.38</td>
<td>25.41</td>
<td>51.33</td>
<td>87.75</td>
</tr>
<tr>
<td>16</td>
<td>&lt;0.01</td>
<td>1.34</td>
<td>20.59</td>
<td>46.33</td>
<td>86.01</td>
</tr>
<tr>
<td>18</td>
<td>&lt;0.01</td>
<td>0.76</td>
<td>16.68</td>
<td>41.81</td>
<td>84.29</td>
</tr>
<tr>
<td>20</td>
<td>&lt;0.01</td>
<td>0.42</td>
<td>13.51</td>
<td>37.74</td>
<td>82.62</td>
</tr>
</tbody>
</table>

From Flaherty. The “differential gene” refers to a targeted gene or a transgene that is selected for retention at each generation. The frequency of recombination between 2 loci depends on their genetic distance (1 centiMorgan = 1% recombination).
Clearly, then, any conclusion about the effect of \textit{Pon3} on atherosclerosis is confused by the passenger problem. The various studies of ABCG1, Cd36, and SRA1 cited at the beginning of this review are likely to have different passenger gene effects because the congenics were generated separately and the amount of donor material remaining in a particular congenic is a matter of chance, and so will differ from laboratory to laboratory. Thus, these and most other studies of knockout or transgenic mice are confounded by passenger genetic variations of unknown effect. (When studying knockouts or transgenics with dramatic phenotypes not present among common strains, the risk of complications is small. A prior hypothesis concerning the effects of the knockout also diminishes the chance of drawing erroneous conclusions.)

What can be done to avoid this problem? It would, of course, be best to construct knockouts or transgenics on a pure genetic background, so that the resulting mice are “isogenic” rather than congenic. Transgenic mice can be generated rather efficiently on an FVB background, although this makes comparisons with other studies, usually on a C57BL/6 background, difficult. Significant progress in making targeted mice in C57BL/6 or other inbred backgrounds has been made, but obtaining chimeras and germline transmission is still not nearly as efficient as with the use of 129 or mixed embryonic stem cell backgrounds. For transgenic production, the gold standard remains analyzing multiple independent transgenic lines. Because transgenes insert at random locations in the genome, the passenger problem is then eliminated or reduced. This is, however, not a solution to the problem of gene targeting. A possible, although cumbersome, solution would be to complement the knockout with a transgenic; thus, if the transgene alters the effect of the knockout, one can conclude with confidence that the results are attributable to the targeted gene.

If a congenic approach must be used, we suggest that a genome-wide scan be performed to test for passenger genes, both linked and unlinked to the gene of interest. The above list consists of genes (indicated as gene symbols) whose expression was significantly affected (lod score $>4.3$) by the 0–40 Mb locus of mouse chromosome 6 in livers of a cross between C3H and C57BL/6 mice.21

Table 2. Genes Whose Expression Is Significantly Perturbed by Common Genetic Variation in the Region 0 to 40 Mb of Chromosome 6

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Gene Symbol</th>
<th>Gene Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>2310036D22Rik</td>
<td>Arf5</td>
<td>Dkk4</td>
</tr>
<tr>
<td>4921524J17Rik</td>
<td>Cald1</td>
<td>MMT00064476</td>
</tr>
<tr>
<td>5830426I05Rik</td>
<td>Chchd3</td>
<td>Ear6</td>
</tr>
<tr>
<td>6330514A18Rik</td>
<td>Lpl</td>
<td></td>
</tr>
<tr>
<td>A930016D02Rik</td>
<td>Gcc1</td>
<td>Co1a2</td>
</tr>
<tr>
<td>B630005N14Rik</td>
<td>Ppp3cc</td>
<td>Smc1l2</td>
</tr>
<tr>
<td>2410075B13Rik</td>
<td>Cnot4</td>
<td>Ml2</td>
</tr>
<tr>
<td>2102 Arterioscler Thromb Vasc Biol. October 2007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. The genome constitution of a paraoxonase 3 (\textit{Pon3}) gene targeted mouse after 6 generations of backcrossing to C57BL/6 strain on an ApoE-null background. The regions derived from B6 are shown as gray bars and the regions from 129 as black bars. The positions along each chromosome, from 1 to X, are indicated in bp, with centromeres at the left. The parental origin of the genomic material was determined by genotyping with the Affymetrix MegAllele Genotyping Mouse 5K SNP panel (Affymetrix).25

The above list consists of genes (indicated as gene symbols) whose expression was significantly affected (lod score $>4.3$) by the 0–40 Mb locus of mouse chromosome 6 in livers of a cross between C3H and C57BL/6 mice.21
While not a solution to the problem, this will likely identify any contaminating regions not linked to the gene of interest and also define the size of the region flanking the gene of interest.

Acknowledgments
We thank Rosa Chen for help in the preparation of this manuscript.

Sources of Funding
Our research is supported by National Institutes of Health grants HL30568, HL28481, and DK071673.

Disclosures
None.

References
The Problem of Passenger Genes in Transgenic Mice
Aldons J. Lusis, Janet Yu and Susanna S. Wang

Arterioscler Thromb Vasc Biol. 2007;27:2100-2103; originally published online August 9, 2007; doi: 10.1161/ATVBAHA.107.147918

Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2007 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/27/10/2100

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at:
http://atvb.ahajournals.org//subscriptions/