Utilization of mouse models in the discovery of human disease genes

Karen J. Moore

The mouse is a powerful model organism for studying human disease. By the use of various breeding strategies in combination with positional cloning and positional candidates, many genes underlying human disease have been and will continue to be discovered through the use of mouse genetics. The author illustrates how the mouse can be used to discover human disease genes underlying simple mendelian and complex polygenic diseases and genes involved in the modification of disease.

The past few years have seen a change in the strategies taken to identify therapeutic agents. Pharmaceutical companies are focusing more on targeting molecules implicated in the genetic component of diseases than they did a decade ago. Two important factors have influenced this change. The first is the recognition that most common human diseases, such as obesity, cardiovascular disease, type 2 diabetes, osteoporosis, autoimmune diseases and diseases of the CNS, have a component of heritability or genetic predisposition. The second results from the Human Genome Project, which has provided a plethora of tools to facilitate an ever-accelerating rate of disease gene identification through positional cloning, positional candidacy and association studies. Concomitant biological studies of disease genes have led to the elucidation of pertinent biochemical and physiological pathways, which in turn identify additional disease targets for therapeutic interaction. The starting points of genetics and genomics for the identification of pathways implicated in common human diseases have, in parallel, brought about the enhanced use of model organisms for the genetic dissection and simplification of disease pathways (Fig. 1).

Advantages of the mouse as a model organism

The mouse is a model organism that offers many cumulative advantages over other model systems such as yeast, *Drosophilia*, *Caenorhabditis elegans* and the rat. The mouse has been used in genetic studies as a model organism since the turn of the century, when it was used to show that Mendel's laws applied to mammalian inheritance[1,2] and hence, by inference, to humans. Indeed, the mouse was subsequently used to investigate the heritability of cancer. What are the current advantages of the mouse as a model organism? First and foremost it is a mammal and therefore many of its biochemical pathways will be similar to those in humans. Mice have a relatively short generation time, they are prolific breeders and can be maintained in a laboratory situation, thus allowing control of the environment, which is a major contributory factor that is extremely difficult to control for in studies of complex human diseases.

An additional benefit of the mouse as a model organism is the ability to control the genetics. There is a wide range of inbred strains of mice available that allows rigorous genetic control within a strain yet genetic variability between strains. Also, the ease with which mouse embryos can be manipulated allows both overexpression and underexpression of genes through production of transgenic and knockout mice. Lastly, a century of research

Karen J. Moore, Millennium Pharmaceutical Inc., 640 Memorial Drive, Cambridge, MA 02139, USA. tel: +1 617 679 7111, fax: +1 617 374 9379, e-mail: moore@mpi.com

 1359-6446/99/$ – see front matter PII: S1359-6446(99)01304-5

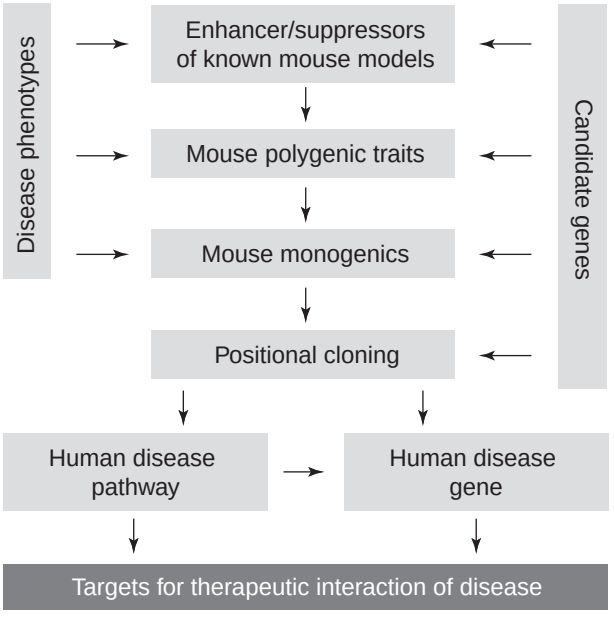

Figure 1. *The mouse can be used to identify the genes involved in simple mendelian disease or in complex polygenic diseases. Through the use of various breeding regimes, complex polygenic disease models can be reduced to more simple models (as described in the text), which can ultimately be used in disease gene identification by positional cloning. Alternatively, the positional candidate gene approach might lead to gene identification, as indicated on the right. Once the mouse gene for a model of a human disease is discovered, it is easy to identify the human ortholog. The human gene might or might not be a gene that is directly involved in the human disease. However, it is likely that the biological pathway so identified will lead to other human genes, variants of which are directly causative in human disease. Regardless of whether the causative human disease genes are identified or not, targets for drug discovery can be found in the relevant biological pathways.*

employing the mouse has resulted in abundant documentation about phenotypes of mouse mutants and strains that resemble both simple, mendelian human disease and complex multigenic diseases. Hence, the stage is set to use the mouse as a surrogate, or 'furry little human', for human disease gene discovery.

Mouse models of simple, mendelian inherited diseases

There is a multitude of examples of mouse mutations whose phenotypes resemble rare human diseases. Subsequent identification of the mutated mouse genes has indeed lead to the identification of human orthologs that are responsible for rare human diseases such as piebaldism[3], Waardenburg syndrome type II (Ref. 4), Usher's syndrome[5], Chediak–Higashi syndrome[6,7], Hermansky–Pudlak syndrome[8] and dwarfism[9,10]. Additionally, single gene trait mouse mutants have been used to identify genes that elucidate biochemical pathways pertinent to common human diseases. The best examples of this are the cloning of the genes underlying the five single-gene mouse models of obesity: *obese*[11], *diabetes*[12], *fat*[13], *tubby*[14,15] and *agouti lethal-yellow*[16,17] (Fig. 2).

The cloning of *obese* and *diabetes* confirmed earlier studies that indicated that the genes encode a ligand (leptin) and its concomitant receptor (leptin receptor) whose signal is vital for regulating body energy balance[18]. Downstream of leptin signaling is the agouti pathway. The agouti lethal-yellow (A^y) mouse has intense yellow pigmentation of the hair, late-onset obesity, hyperinsulinemia, hyperglycemia and hyperphagia[19], which is a consequence of the ectopic expression of a 131 amino acid peptide that is normally expressed in the skin[16,17]. Pharmacological studies indicate that the overabundance of the agouti protein antagonizes the binding of melanocyte-stimulating hormone (α-MSH) (or similar peptides) to melanocortin receptors[20]. The yellow hair pigmentation is the result of antagonizing the melanocortin receptor, MCR1, and the obese phenotype is driven by antagonizing the MCR4 receptor[21]. The natural ligands of the melanocortin receptors, α-MSH and similar peptides, are derived from processing the polypeptide pro-opiomelanocortin (POMC) into smaller peptides. The transcription of POMC is thought to be downstream of the leptin signal transduction pathway. It has also been hypothesized that the product of the *fat* mouse, carboxypeptidase E, is involved in trimming the two C-terminal amino acids from the cleavage products of POMC to activate small peptides such as α-MSH. In a relatively short period of time, since the cloning of these five monogenic obesity mouse mutants, it has become possible to build and connect multiple pathways that are involved in energy balance in the mouse, providing more therapeutic targets than those provided by the original five mouse genes.

What is the relevance to human obesity? To date only rare, extremely obese individuals have been identified who harbor mutations in leptin[22,23] and the leptin receptor[24], rendering these molecules inactive. No human variants of the *agouti* gene have yet been found in obese individuals; however, the agouti pathway has been implicated in human obesity through several observations.

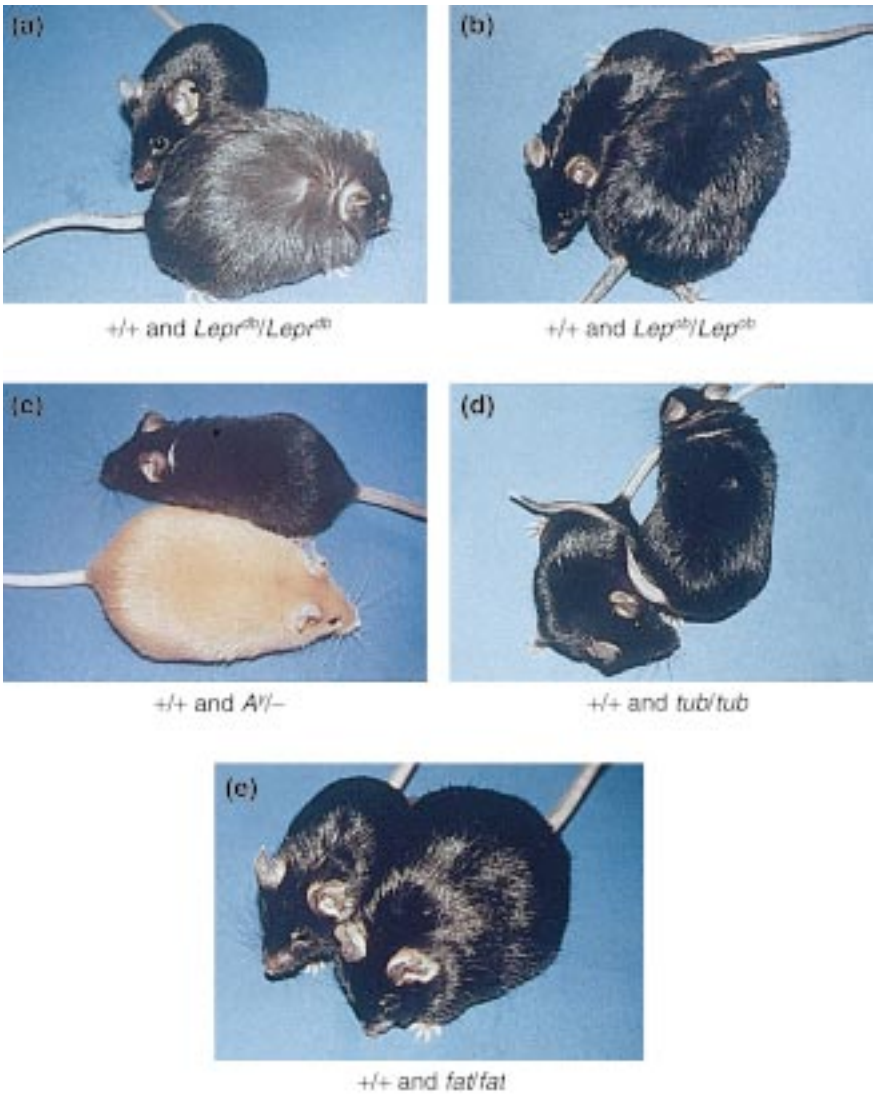

Figure 2. *Each of the monogenic mouse obese mutants is shown in a separate panel along with its normal littermate, indicated +/+. The mutant mouse is to the right in (a), (b), (d) and (e) and is the lower mouse in (c). (a) Shows the leptin receptor mutant (*Lepr^{db}*), previously known as* diabetes*; (b) shows the leptin mutant (*Lep^{ob}*), previously known as* obese*; (c) shows the* agouti lethal-yellow *mutant (*A^{y}*); (d) shows the* tubby (tub) *mutation; and (e) shows the* fat *mutation.*

Mutations of POMC have been found in extremely obese individuals[25] and association studies have implicated *MC4R* in obesity[26]. Also, mutations inactivating the processing enzyme proconvertase 1, which, like carboxypeptidase E, is implicated in processing POMC to smaller peptide hormones, have been found in obese individuals[27]. Clearly, the pathways identified by cloning the mouse obese mutants are validated pathways for extreme obesity in humans, but are they relevant to the more common types of obesity? More work needs to be done to answer this, but it is very likely that subtle changes in gene expression, perhaps due to promoter variants, by contrast to the ablation mutations described above, will result in later-onset, less extreme obesity. There are also some potential hints to this effect from mouse polygenic studies (see below).

Mouse models of complex multigenic diseases

The advent of high-density genetic maps not only provided the tools for the positional cloning of monogenic traits but has also facilitated the identification of genetic loci involved in multigenic traits in humans, mice, rats and cows. Several linkages of quantitative trait loci (QTLs) have already been detected in mouse for complex disorders such as alcohol sensitivity, atopy, diabetes, autoimmunity, CNS disorders and behavior. There is also copious research identifying QTLs of mouse obesity. These studies used a variety of inbred mouse strain pair combinations and measured obesity in multiple ways, including high fat diet-induced obesity, simple whole body weight determination, weight of specific fat pads and total adiposity. In total, these investigations have identified 48 QTLs (Refs 28,29). Because of the overlapping genetic map location of some of the QTLs, the 48 may represent anywhere from 29 to 48 independent loci responsible for some aspect of body energy balance and/or body weight. None of the genes responsible for these QTLs have yet been identified, partly because these studies are relatively recent. It is at this stage of identifying loci involved in complex traits that the mouse reveals many of its advantages over similar investigations in human populations. The populations available for human and bovine complex disease analysis often segregate multiple loci, each of which may drive a subcomponent of the phenotype. This often renders the ability to

clone disease genes using positional cloning difficult, but not impossible. Such complexities can be circumvented when using the mouse.

Through various breeding strategies each QTL can be dissected from the others to produce a series of monogenic mouse strains that each harbor one QTL (congenic strains)[30,31]. These monogenic strains can then be used in more simple positional cloning strategies. Examples of positional cloning of monogenic mouse mutations are described above and have become commonplace today. Consequently, it will be possible to clone some QTLs by this reductionism approach. As promising as this approach is, it is not the panacea and caveats to this approach have already been encountered. Some investigators have discovered that when a complex trait is dissected into its individual components there is no residual disease phenotype, or subphenotype, in the newly created mouse strains. There are several potential reasons for this:

- The QTLs interact epistatically with (i.e. are dependent upon) one or more of the other QTLs such that in isolation each has no affect upon phenotype
- The percentage of the phenotypic variance of the original phenotype that each of the individual loci drives is too low to be detected when in isolation
- The breeding strategies did not transfer the pertinent genomic fragment into the congenic strains
- Once individualized, the QTLs cannot be detected using the original phenotypic assay, but rather a subphenotype is now needed to detect the locus
- Not all detected QTLs are true genetic loci, some may be false positives resulting from QTL mapping

Only a few mouse QTL studies have progressed to the individual congenic stage, the most elegant of which is a mouse model of systemic lupus erythematosus (SLE)[32]. What is particularly exciting about this study is that each of the individual congenic strains now has a different subphenotype of the disease, each of which is a known component of the fully fledged SLE disease in humans. One of the congenic strains has a predisposition to break immune tolerance to chromatin, one has B-cell hyperreactivity and the third has high antinuclear antibody production and predisposition to end-organ disease (and may represent two genes)[33] (Fig. 3). In addition to the incredible parallel between the human and mouse phenotypes of this mouse model of human disease, it is intriguing that the genomic region of one of the mouse *Sle* QTLs is evolutionarily equivalent to a region in the human genome within which a genetic linkage for human SLE has been identified.

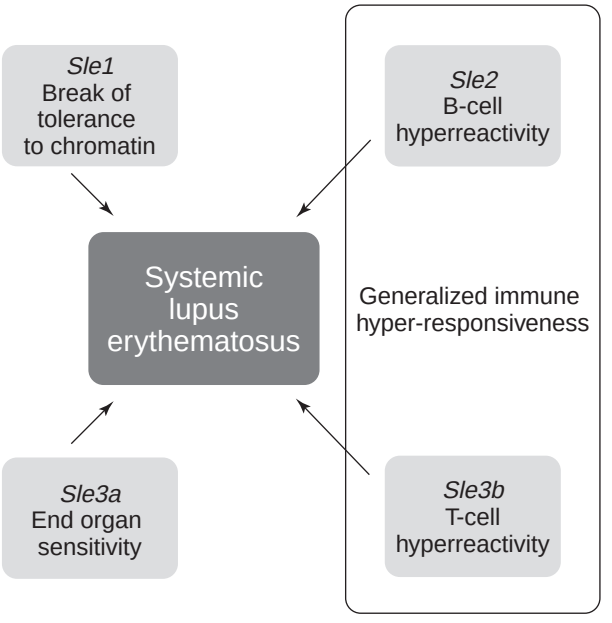

Figure 3. *Complex diseases such as systemic lupus erythematosus (SLE) are often the cumulative effects of many genes, each of which when variant can cause a single component of the total disease. Genetic dissection of a mouse model of SLE has resulted in the production of strains of mice that each carry a single gene:* Sle1, Sle2 *or* Sle3. *Each of these strains has a different pathogenesis that constituted the original disease.*

When the mouse gene is identified it will be an excellent candidate for the human SLE gene, once again exemplifying the relevance of using mouse genetics to identify human disease genes.

The five genes that have been cloned from mouse monogenic obese models have been described and they indicate that numerous loci have also been identified in obesity QTL studies. Are the QTLs and the monogenic obesity loci mutually exclusive? Perhaps not, there are QTLs that map coincident with the *leptin* gene and with the *agouti* gene and an obesity QTL in rat that maps coincident with the *tubby* gene[28]. No one has yet proven that these are the causative genes for these QTLs, but they offer tantalizing candidates. This highlights an important part of the strategy for cloning mouse disease genes identified through QTL analysis. It is not always necessary to undertake genetic dissection of a complex trait if within the genetic confidence limit of the QTL there is an appealing candidate gene that subsequently can be shown to be the gene causing the QTL variation. This strategy, known as

the positional candidate approach to gene cloning, has in the past been applied very successfully in identifying monogenic disease genes in both humans and mouse. It has also been applied to one complex trait locus in the mouse to identify a gene affecting colon polyp number in a mouse model of adenomatous polyposis coli[34]. In the past, the candidate gene approach has relied upon compelling biological annotation of the gene that maps within a disease linkage confidence region. With the advent of transcriptional profiling of very large numbers of genes simultaneously it will be easier to look at the expression of all genes known to map within a disease locus region. This approach is again particularly compatible with the mouse, where there is controlled genetics and hence less noise in the system, where tissue source and quantity are not limiting and biological annotation to mapped genes will not be as critical. Using a combination of genetic linkage and genomic region-specific transcriptional profiling, it is anticipated that the positional candidate approach will become even more successful, that positional cloning will become less necessary and that congenic strains for QTLs will need to be made for fewer loci.

Enhancer/suppressor screens

Enhancer/suppressor screens are a powerful method for identifying additional genes in a given pathway and have been used extensively to identify new mutations in yeast, *Drosophilia* and *C. elegans*. There are currently only a few examples of mouse suppressor genes, including the spontaneous mutation *dilute suppressor*[35] and three suppressors of *agouti* phenotypes[36–38] (Fig. 4). There are also examples of modifiers of disease induced by *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis[39]. As mouse genetics enters the last year of the twentieth century there will be an explosion of such mouse variants. Mutagenesis of the mouse is entering a renaissance, with the emphasis on discovering new mouse models of human disease. Sophisticated phenotypic screens are being developed for the mice derived from the mutagenesis regimes in order to close the 'phenotype gap' that currently exists. Funding agencies are enthusiastically supporting this.

An additional source of enhancer/suppressor genes comes from studies that exploit the genetic variance of the inbred strains. As described above, inbred strains can differ phenotypically, allowing QTL analysis to be performed. An additional twist to this approach is to drive a particular phenotype using over- or underexpression alleles of a known gene (i.e. transgenic or knockout mice) and then place these alleles on multiple genetic backgrounds and assay for modification of phenotype. The modification loci can then be mapped using QTL analysis. Such an approach has been used very effectively in an atherosclerosis study with the Apoe null allele and has indicated the presence of strong enhancer and suppressor genes[35,40].

Figure 4. *Suppression of the yellow pigmentation of* agouti lethal-yellow *(A^y) by two independent genes:* mahogany *(*mg*) and* mahoganoid *(*md*). The five mice are: far left,* $A^y/-$,md/md*; next left* $A^y/-$,md/+*; top right,* $A^y/-$,+/+*; middle right,* $A^y/-$,mg/+*; bottom right,* $A^y/-$,mg/mg*. It can be seen that a single copy of either* mg *or* md *acts to partially suppress the production of yellow pigment while two copies result in much greater suppression. What is not clear from this photograph, because of the different ages of these mice, is that both* md *and* mg *suppress the obesity of (A^y/–) mice also*[37].

Summary

This is an exciting time in mouse genetics. The mouse has been used as a surrogate to study human diseases for decades. However, there is currently a heightened interest that is concomitant with an ever-growing technical ability to use mouse genetics to identify genes underlying both rare, monogenic human diseases as well as polygenic human diseases. In time, mouse genetics will have proved to be a critical early step in the development of therapeutic drugs for many of our common diseases.

ACKNOWLEDGEMENTS

I thank Robert Tepper, Deborah Nagle and John Smutko for their helpful comments.

REFERENCES

1 Cuenot, L. (1902) *Arch. Zool. Exp. Gen.* 3, 27–30
2 Cuenot, L. (1905) *Arch. Zool. Exp. Gen.* 4, 123–132
3 Spritz, R.A. *et al.* (1993) *Hum. Mol. Genet.* 2, 1499–1500
4 Moore, K.J. (1995) *Trends Genet.* 11, 442–448
5 Weil, D. *et al.* (1995) *Nature* 374, 60–61
6 Nagle, D.L. *et al.* (1996) *Nat. Genet.* 14, 307–311
7 Barbosa, M.D. *et al.* (1996) *Nature* 382, 262–265
8 Feng, G.H. *et al.* (1997) *Hum. Mol. Genet.* 6, 793–797
9 Sornson, M.W. *et al.* (1996) *Nature* 384, 327–333
10 Wu, W. *et al.* (1998) *Nat. Genet.* 18, 147–149
11 Zhang, Y. *et al.* (1994) *Nature* 372, 425–432
12 Tartaglia, L.A. *et al.* (1995) *Cell* 83, 1263–1271
13 Naggert, J.K. *et al.* (1995) *Nat. Genet.* 10, 135–142
14 Kleyn, P.W. *et al.* (1996) *Cell* 85, 281–290
15 Noben-Trauth, K. *et al.* (1996) *Nature* 380, 534–538
16 Miller, M.W. *et al.* (1993) *Genes Dev.* 7, 454–467
17 Bultman, S.J., Michaud, E.J. and Woychik, R.P. (1992) *Cell* 71, 1195–1204
18 Houseknecht, K.L. *et al.* (1998) *J. Anim. Sci.* 76, 1405–1420
19 Manne, J., Argeson, A.C. and Siracusa, L.D. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 4721–4724
20 Lu, D. *et al.* (1994) *Nature* 371, 799–802
21 Huszar, D. *et al.* (1997) *Cell* 88, 131–141
22 Montague, C.T. *et al.* (1997) *Nature* 387, 903–908
23 Strobel, A. *et al.* (1998) *Nat. Genet.* 18, 213–215
24 Clement, K. *et al.* (1998) *Nature* 392, 398–401
25 Comuzzie, A.G. and Allison, D.B. (1998) *Science* 280, 1374–1377
26 Chagnon, Y.C. *et al.* (1997) *Mol. Med.* 3, 663–673
27 Jackson, R.S. *et al.* (1997) *Nat. Genet.* 16, 303–306
28 Chagnon, Y.C., Perusse, L. and Bouchard, C. (1998) *Obes. Res.* 6, 76–92
29 Suto, J. *et al.* (1998) *Mamm. Genome* 9, 506–510
30 Darvosi, A. (1998) *Nat. Genet.* 18, 19–24
31 Markel, P. *et al.* (1997) *Nat. Genet.* 17, 280–284
32 Morel, L. *et al.* (1994) *Immunity* 1, 219–229
33 Morel, L. *et al.* (1997) *J. Immunol.* 158, 6019–6028
34 Dietrich, W.F. *et al.* (1993) *Cell* 75, 631–639
35 Moore, K.J. *et al.* (1988) *Genetics* 119, 933–941
36 Mather, K. and North, S.B. (1940) *J. Genet.* 40, 229–241
37 Miller, K.A. *et al.* (1997) *Genetics* 146, 1407–1415
38 Dinulescu, D.M. *et al.* (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 12707–12712
39 Bode, V.C. *et al.* (1988) *Genetics* 118, 299–305
40 Hayes, D.M. *et al. Arterioscler. Thromb. Vasc. Biol.* (in press)