Insertional mutation of 'classical' and novel genes in transgenic mice

MIRIAM H. MEISLER

Approximately 5% of established transgenic lines carry insertional mutations. The mutated genes may be directly isolated using the transgene DNA as a molecular probe. These mutants provide useful models of human inherited disorders and developmental abnormalities.

Approximately 500 loci responsible for genetic disease in the mouse have been described and mapped during the past 80 years – the so-called 'classical mutations' of the mouse[1]. As a result of progress in genetic mapping and molecular technology, these interesting genes have become accessible to molecular analysis[2]. The relevance of these loci to human disease was recently underscored by the demonstration that the molecular defects in the mouse mutants *Small eye*[3], *Splotch*[4], and *retinal degeneration slow*[5] are homologous to those in the human disorders aniridia[3], Waardenburg syndrome[6,7] and one form of retinitis pigmentosa[8], respectively. The most successful method to date for identification of the mutated genes in the classical mutants has been genetic mapping of candidate genes near the mutant loci. More than 15 mutations have been identified in this way during the past three years.

Another useful approach to identifying the molecular basis of interesting mouse mutants is the analysis of insertional mutations. These mutations are generated when foreign DNA becomes integrated within a functional gene and prevents or alters its expression. The foreign DNA can subsequently be used as a molecular tag to clone the disrupted gene. Spontaneous insertion of retroviral-like elements can produce mutations, as observed for alleles of the *dilute*[9] and *hairless*[10] loci. Genes may also be interrupted experimentally as a result of the random insertion of foreign DNA during the production of transgenic mice by microinjection of fertilized eggs or retroviral infection of embryos (see R. Beddington, centerfold of this issue). The overall frequency of insertional mutation in transgenic lines, including visible mutations and prenatal lethals, has been estimated[11–13] to be between 5% and 10%. The visible mutations are of particular interest as models for human disorders. Genomic libraries can be prepared from insertional mutants and screened with the transgene as probe, in order to recover sequences from the insertion site which should include the disrupted gene. Insertional mutations thus offer a direct route from a complex genetic abnormality to the underlying molecular defect.

Insertional mutation of 'classical' loci

A number of visible transgenic mutations have been found to be allelic with classical mutations (Table 1).

TABLE 1. Insertional mutation of 13 'classical' mouse loci in transgenic mice

Chrom.	Locus symbol	Locus name	Phenotype	Method[a]	Ref.
1	*dt*	*dystonia*	Severe muscle weakness at four days of age	M	30
2	*ld*	*limb deformity*	Fused radius and ulna, missing kidneys	M, C	21
3	*rcm*	*rostral cerebellar malformation*	Disrupted cerebellar striatum	M	–[b]
6	*ho*	*hotfoot*	Behavior (constant running), sterility	M	31
10	*pg*	*pygmy*	25% of normal size, defective response to growth hormone	M	32
10	*dl*	*downless*	Thick coat, bald patches, missing teeth	M	17
10	*Sl*	*Steel*	Defective growth factor; sterile, anemic	M	14
12	*twi*	*twitcher*	Tremor, seizures, demyelination	M	33
12	*iv*	*situs inversus viscerum*	Right–left inversion of viscera, leg deformity	M, C	20
13	*Xt*	*extra toes*	Abnormal morphogenesis, homozygous lethal	M	34
13	*pcd*	*Purkinje cell degeneration*	Ataxia, photoreceptor cell degeneration, male sterility, abnormal sperm	M	35
14	*hr*	*hairless*	Disorganization of hair follicle	M	–[c]
18	*Tw*	*Twirler*	Circling, abnormalities of the inner ear	M	–[d]

[a] M, microinjection; C, transcript cloned.
[b] L. Kozak, pers. commun.
[c] Meisler *et al.*, unpublished; based on noncomplementation of the *hr* allele.
[d] Meisler *et al.*, unpublished; based on linkage data and phenotypic similarity, without complementation test.

©1992 Elsevier Science Publishers Ltd (UK)

TABLE 2. Insertional mutation of 16 previously unknown loci in transgenic mice

Chrom.	Locus symbol	Locus name	Phenotype	Insert[a]	Ref.
Visible, viable phenotypes					
14	*sys*	*symplastic spermatids*	Male sterility, multinucleated syncytia of round spermatids	M	36
15	*ple*	*perinatal lethality*	Newborn death	M	37
16	*ckr*	*chakragati*	Recessive circling of CNS origin	M	18,38
?	*Mvp17*	–	Kidney failure; glomerular sclerosis	R, C	13
?	*gcd*	*germ cell deficient*	Defective germ cells at day 11.5 of gestation	M	39
6	*mnd-2*	*motor neuron disease 2*	Degeneration of spinal motor neurons, muscle wasting, thymic involution	M	–[b]
?	*Wo*	*Wocko*	Abnormal development of vestibular system of inner ear; homozygous lethal	M	40
?	AE24	–	Sperm mobility, flagellar disruption	M	41
?	–	*microtia*	1st and 2nd branchial arch syndrome	M	42
?	*adp*	*acrodysplasia*	Deformity of fore and hind limbs	M	43
Prenatal lethal					
6	*RB3*	–	Developmental arrest at day 5 of gestation	M	44
8	*Mov-34*	–	Lethal at time of implantation	R, C	23
10	*Hb58*	–	Prenatal lethal (day 10)	M	45
11	*Cola-1*	procollagen 1 α-1	Day 12 embryonic lethal	R, C	22
?	413.d	–	Embryonic disorganization at day 7	R	25
?	R197	–	Not viable beyond egg cylinder (day 6.5)	M	46

[a] M, microinjection; R, retroviral infection; C, transcript cloned.
[b] Meisler *et al.*, unpublished.

In one series analysed in my laboratory, animals from transgenic lines generated by microinjection were examined at four to six weeks of age for visible abnormalities in external appearance, behavior and internal organs. We detected two dominant phenotypes among 200 lines tested (coat color and circling behavior) and three recessives among 120 lines (motor neuron disease, kidney agenesis and hair loss). Genetic mapping and complementation analysis indicate that three of the five represent re-mutation of the classical loci *Steel*[14], *Twirler* and *hairless* (Table 1). The high proportion of classical loci mutated in our lines is not unique. Among 23 visible, viable transgenic mutations that have been described, there are 13 mutations at previously known loci (Table 1) and 10 mutations at new loci (Table 2).

One factor contributing to the re-isolation of classical mutations is the method of ascertainment. Mutations are detected in the subset of genes for which inactivation can result in viable (i.e. liveborn) animals with visible abnormalities. Most of the classical loci were originally identified as spontaneous mutations by a similar analysis of large numbers of mice at facilities such as the Jackson Laboratory, Harwell and Oak Ridge. The fact that roughly half of the visible insertional mutations are allelic with classical loci suggests that half of this subset of genes have already been detected. Since the number of classical mutations described as of 1987 was just under 500 (Ref. 1), the total number of such genes in the mouse genome appears to be approximately 1000. This is similar in magnitude to the estimated 3000 human disease loci (Ref. 15 and V. McKusick, pers. commun.), which are the human counterparts of the visible, viable mutations of the mouse.

Identification of new loci

Insertional mutagenesis is also a powerful method for identification of previously unknown mouse loci. Some recent examples are listed in Table 2. The new mutants include models for kidney failure, motor neuron disease and infertility. Genetic mapping and phenotypic description of these mutants will expand the list of mouse models. Compared with other types of mutations, the chromosomal mapping of insertional mutants is greatly facilitated by the ability to use the transgenic sequences as linkage markers and as probes for fluorescence *in situ* hybridization (FISH) of metaphase chromosomes.

A partial list of recessive prenatal lethal insertional mutations is included in Table 2. These mutations can be recognized by the absence of homozygotes among offspring of heterozygous transgenic mice. A more extensive list is available in a recent review[16].

Isolation of the interupted genes

A general approach to the cloning of interrupted genes is outlined in Box 1. Although the procedure is straightforward, practical problems can complicate the process. The presence of repetitive sequences around the insertion site can make it difficult to isolate single-copy flanking sequences. While some microinjected fragments produce simple insertions, deletions around the insertion site[11,13,17] may remove some or all of the interrupted gene. Rearrangements at the site of insertion, including inversions[18] and translocations[19], can also separate the transgene from exons of the mutated gene. Nonetheless, transcripts have been successfully cloned from two microinjected insertions (*legless*[20] and *limb deformity*[21]), three retroviral insertions [*Cola-1* (Ref. 22), *Mov-34* (Ref. 23) and *Mvp17* (Ref. 13)] and the spontaneous retroviral insert at the *dilute* locus[9]. Retroviral inserts may be advantageous in this context, since they usually generate single-copy inserts without loss of adjacent sequences.

The cloned transcripts that have been isolated are of considerable interest. For example, mutations at the *situs inversus viscerum* locus disrupt the normal development of asymmetry of the viscera. Transgene sequences from the *legless* mutant were used to isolate transcripts that appear to be allelic with *situs inversus*[20] and will contribute to understanding an important developmental process. Cloning of the insertion site from the *limb deformity* mutant led to the discovery of a new class of structural proteins designated formins[21], which are present in a wide variety of cells. The product of the *dilute* locus was found to be a novel myosin heavy chain which is expressed at high levels in neurons. Attempts to isolate other loci in Table 1 are in progress. The success rate of this cloning strategy should become clear within the next few years.

Is the frequency of visible mutations consistent with random insertion of microinjected transgenes?

At first glance, the frequency of visible mutations in transgenic lines, and the frequency of re-mutation of classical loci, may appear surprisingly high. It is possible to make a very rough estimate of the expected rate of such mutations. It is assumed that most insertions into genes will result in null alleles. The target for generation of visible mouse mutations, as discussed above, appears to be approximately 1000 genes, or 1% of the estimated 10^5 genes in the mammalian genome. If we also assume that the average gene length is 10^4 bp, sufficient to encode a protein of 30 kDa with a 10:1 ratio of noncoding to coding sequence, then 30% of genomic DNA will be intragenic. Under these assumptions, 30% of random insertions will disrupt a gene, and the frequency of visible mutations in transgenic lines would be 1% of 30%, or 0.3%. Observed rates of visible mutations, 3% in our lines and 1.5% (2/139) as reported in Ref. 11, exceed the estimate by five- to tenfold. Several factors may contribute to this excess, including preferential insertion into genes, underestimation of the target for visible mutations, and disruption of more than one gene by a significant proportion of insertions. In view of the uncertainties associated with these estimates, it would be premature to propose that insertion is not random.

Box 1. From phenotype to gene: identification and analysis of insertional mutations in transgenic mice

(1) Examine transgenic heterozygotes from established transgenic lines for internal and external abnormalities indicative of dominant mutation.

(2) Cross heterozygous transgenic mice and screen their offspring to detect recessive mutations.

(3) Determine the chromosomal localization of mutations by FISH or linkage analysis of transgene DNA.

(4) Test complementation of any classical mutant with a similar phenotype that was previously mapped to the same chromosome region.

(5) Screen a library of genomic DNA from the insertional mutant using a transgene probe, to isolate mouse DNA flanking the insertion site.

(6) Isolate the wild-type allele by screening a mouse genomic DNA library with a probe from the flanking sequence.

(7) Identify exons in the flanking DNA and/or wild-type allele by available methods including sequence analysis, zoo blot, northern blot, screening of cDNA libraries and exon trapping; analyse protein sequence.

(8) Identify the molecular defect by comparison of wild-type and mutant alleles.

(9) If the insertional mutation appears to be allelic with a classical mutant, identify the molecular lesion in the spontaneous mutant allele, or correct the defect with a wild-type transgene.

Prenatal lethal mutations appear to account for more than 75% of insertional mutations[11,16]. As many as 10% of transgenic insertions may result in lethal mutations, a frequency similar to that observed for insertional mutation in the yeast genome. The issues of mutation rates, target sizes and random versus preferential insertion should be re-examined as more data become available for visible and lethal insertional mutations, and for retroviral inserts compared with microinjected transgenes.

Gene trapping and homologous recombination

When mutagenesis is the primary purpose for generating transgenic animals, mutation rates can be increased by microinjection of embryonic stem (ES) cells with 'gene trap' constructs and preselection for expression of the constructs resulting from insertion into genes[24,25]. Among 24 transgenic lines recently generated by this method, the mutation rate was 38%, with nine prenatal lethal mutations and no visible mutations[26]. In another recent study, three gene trap inserts were bred to homozygosity and two were associated with perinatal lethality[27].

Targeted inactivation of genes by homologous recombination is another important method for analysing the molecular basis of complex phenotypes[16]. The experimental control of the molecular lesion is ideal for interpreting mutant phenotypes. However, this method is limited to genes that have already been cloned, and cannot be used to discover

new loci. Targeted inactivation may reveal the basis for classical mutants, as occurred with *Wnt-1* and the cerebellar mutant, *swaying*[28].

Another useful innovation is the co-injection of tyrosinase with the transgene of interest, which makes it possible to recognize transgenic heterozygotes and homozygotes by visual inspection of coat color rather than analysis of DNA[29].

Future prospects

Thousands of transgenic mice are currently being generated for research on a wide variety of biological topics. These animals represent a valuable resource, with their high rate of molecularly accessible mutations. Realization of this potential will require systematic analysis of many transgenic lines, including crosses to generate homozygotes. At institutions where large numbers of transgenics are generated for multiple investigators, organization of a central screening facility would facilitate the detection of mutants. Screening transgenic mice with sophisticated diagnostic techniques to detect specific classes of defects, such as hearing or vision disorders, is also likely to be very productive. One large-scale mutagenesis project is currently in progress at the Oak Ridge National Laboratory, under the direction of Richard Woychik. We can anticipate that the list of transgenic models in Tables 1 and 2 will be greatly expanded in coming years.

Acknowledgements

I thank Julie Jones, Scott Keller, Chao-Nan Ting and Linda Samuelson for enthusiastic collaboration in the identification of insertional mutants, and the reviewers for valuable comments on the manuscript. Supported by NIH GM24872.

References

1 Green, M.C. (1989) in *Genetic Variants and Strains of the Laboratory Mouse* (2nd edn) (Lyon, M.F. and Searle, A.G., eds), pp. 12–403, Oxford University Press
2 Reith, A.D. and Bernstein, A. (1991) *Genes Dev.* 5, 1115–1123
3 Hill, R.E. *et al.* (1991) *Nature* 354, 522–525
4 Epstein, D.J., Vekemans, M. and Gros, P. (1991) *Cell* 67, 767–774
5 Travis, G.H. *et al.* (1989) *Nature* 338, 70–73
6 Tassabehji, M. *et al.* (1992) *Nature* 355, 635–637
7 Baldwin, C.T. *et al.* (1992) *Nature* 355, 637–638
8 Farrar, G.J. *et al.* (1991) *Nature* 354, 478–480
9 Mercer, J.A. *et al.* (1991) *Nature* 349, 709–713
10 Stoye, J.P. *et al.* (1988) *Cell* 54, 383–391
11 Palmiter, R.D. and Brinster, R.L. (1986) *Annu. Rev. Genet.* 20, 465–500
12 Gridley, T., Soriano, P., and Jaenisch, R. (1987) *Trends Genet.* 3, 162–166
13 Weiher, H. *et al.* (1990) *Cell* 62, 425–434
14 Keller, S.A., Liptay, S., Hajra, A. and Meisler, M.H. (1990) *Proc. Natl Acad. Sci. USA* 87, 10019–10022
15 McKusick, V. (1992) *On-Line Mendelian Inheritance in Man (OMIM)*, Welsh Library, Johns Hopkins University
16 Grindley, T. (1991) *New Biol.* 3, 1025–1034
17 Shawlot, W., Siciliano, M.J., Stallings, R.L. and Overbeek, P.A. (1989) *Mol. Biol. Med.* 6, 299–307
18 Ratty, A.K. *et al.* (1992) *Mamm. Genome* 3, 5–10
19 Franke, U. *et al.* (1992) *Mamm. Genome* 3, 209–216
20 Singh, G. *et al.* (1991) *Genes Dev.* 5, 2245–2255
21 Woychik, R.P. *et al.* (1991) *Nature* 346, 850–853
22 Schnieke, A., Harbers, K. and Jaenisch, R. (1983) *Nature* 304, 315–320
23 Gridley, T. *et al.* (1990) *Development* 109, 235–242
24 Gossler, A., Joyner, A.L., Rossant, J. and Skarnes, W.C. (1989) *Science* 244, 463–465
25 Robertson, E.J. (1991) *Biol. Reprod.* 44, 238–245
26 Friedrich, G. and Soriano, P. (1991) *Genes Dev.* 5, 1513–1523
27 Skarnes, W.C., Auerbach, B.A. and Joyner, A.L. (1992) *Genes Dev.* 6, 903–918
28 Thomas, K.R., Musci, T.S., Neumann, P.E. and Capecchi, M.R. (1991) *Cell* 67, 969–976
29 Beermann, F. *et al.* (1990) *EMBO J.* 9, 2819–2826
30 Kothary, R. *et al.* (1988) *Nature* 335, 435–437
31 Gordon, J.W. *et al.* (1990) *Dev. Biol.* 137, 349–358
32 Xiang, X., Benson, K.F. and Chada, K. (1990) *Science* 247, 967–969
33 Jensen, N.A. *et al.* (1989) *Mol. Biol. Med.* 6, 493–500
34 Pohl, T.M., Mattei, M.G. and Rüther, U. (1990) *Development* 110, 1153–1157
35 Krulewski, T.F., Neumann, P.E. and Gordon, J.W. (1989) *Proc. Natl Acad. Sci. USA* 86, 3709–3712
36 MacGregor, G.R. *et al.* (1990) *Proc. Natl Acad. Sci. USA* 87, 5016–5020
37 Beier, D.R. *et al.* (1989) *Genomics* 4, 498–504
38 Ratty, A.K. *et al.* (1990) *Brain Res. Mol. Brain Res.* 8, 355–358
39 Pellas, T.C. *et al.* (1991) *Proc. Natl Acad. Sci. USA* 88, 8787–8791
40 Crenshaw, E.B., III *et al.* (1991) *J. Neurosci.* 11, 1524–1530
41 Merlino, G.T. *et al.* (1991) *Genes Dev.* 5, 1395–1406
42 Otani, H. *et al.* (1991) *Anat. Anz.* 172, 1–9
43 DeLoia, J.A. and Solter, D. (1990) *Development Suppl.*, 73–79
44 Mark, W.H., Signorelli, K. and Lacy, E. (1985) *Cold Spring Harbor Symp. Quant. Biol.* 50, 453–464
45 Radice, G., Lee, J.J. and Costantini, F. (1991) *Development* 111, 801–811
46 Tan, S-S. (1991) *Dev. Biol.* 146, 24–37

M.H. Meisler is in the Department of Human Genetics, The University of Michigan Medical School, Medical Science II M4708, Ann Arbor, MI 48109-0618, USA.

Rice Genome Newsletter

The Rice Genome Research Program has begun to produce a newsletter, *Rice Genome*, which aims to inform all researchers interested in mapping and analysing plant genomes. The main objective is to enhance international cooperative research efforts for rice genome analysis and for the isolation and utilization of useful rice genes in plant breeding and biotechnology. The newsletter is available free of charge.

Write to: Editorial Office, *Rice Genome*,
Rice Genome Research Program, National Institute of Agrobiological Resources,
2-1-2, Kannondai, Tsukuba, Ibaraki 305, Japan.
FAX: +81-298-38-7468