

Genetic Manipulation in the Mouse Genome: An Overview on Reverse Genetics Approaches

The functions of any gene/gene product can be analyzed at multiple levels: 1) cellular, 2) biochemical, or 3) within a whole organism. Studies at the cellular and biochemical levels are performed in a cell culture or cell-free system; this is known as *in vitro* analysis. To understand gene function at the organism level, geneticists take an *in vivo* approach. Mouse molecular genetic analysis—either by forward or reverse genetics—is an effective way to study gene function in the context of the whole organism. The tools and approaches used for forward and reverse mouse genetics are, however, distinctively different.

The scope of this overview focuses on **mouse reverse genetics**, which employs a specific set of genetic manipulations in the mouse genome and is subdivided into two phases: 1) creating mutations, and 2) analyzing the phenotype(s) of the mutant mice. Here, we present various options that are available to investigators interested in using the mouse model to achieve their desired research goals. Emphasis is placed on how to select among the various strategies prior to making a commitment to a specific project.

1) Creating Mutations

There are two general strategies for introducing mutations into the germline of the mouse: 1) via homologous recombination in embryonic stem (ES) cells or 2) via direct microinjection of DNA into the fertilized egg. Both types of manipulations are technically feasible and are based upon the standard procedure of microinjection of pronucleus or blastocyst stage of mouse embryo. Basically, all mutations can be broadly categorized as either **gain-of-function** or **loss-of-function**. Both categories of mutations improve our understanding of gene function by altering the *in vivo* expression of the gene of interest. The types of mutations to be created are question-oriented. **Gain-of-function mutations** involve the *overexpression* or *ectopic expression* of the transgene while at the same time keeping the expression of the endogenous gene intact. The transgenic approach accomplishes this by pronuclei injection into the fertilized egg, and can be used for:

1. Tissue-specific or ectopic expression of the gene of interest.
2. Identification of promoter/enhancer elements for tissue-specific gene expression.
3. Creation of an inducible or tissue-specific cre/flip recombinase expression vector for spatial- or temporal-specific gene expression, used for deletion of the gene of interest (i.e., a “conditional” knockout). For temporal control of transgene expression, an inducible regulatory element is needed, such as a tetracycline- or ecdysone-based, hormone-modulated system.

Problems:

1. Variability of copy numbers = variability in the level of transgene expression.
2. Variability of the chromosomal integration site = variability in the level of transgene expression.
3. Few well-characterized tissue-specific regulatory elements exist; these elements are required in order for a conditional knockout to drive tissue-specific cre recombinase expression.

Solutions:

1. To reduce the copy number by using the loxP/cre system.
2. To achieve locus-specific integration of the transgene by using the homologous recombination/embryonic stem (ES) approach, i.e., the “knockin” or “replacement” approach.
3. To insert locus control elements into the constructs to achieve position-independent and copy number-dependent transgene expression
4. To avoid the lack of well-characterized, tissue-specific regulatory elements to direct tissue specific deletion or tissue specific expression, cDNA encoding cre/flip recombinase, or other genes of interest, can be ‘knocked’ into a specific endogenous locus whose expression matches or overlaps with the desired tissue and temporal expression pattern. Strategies such as in-frame

fusion, splicing acceptor/donor, internal ribosomal entry site (IRES) or other inducible systems can be incorporated seamlessly to subvert the endogenous tissue specific regulatory system to your specific needs (see section **2a**).

Loss-of-function mutations are manipulations in the mouse genome that result in either a complete loss of activity (null mutation), or a partial loss of activity (hypomorphic, haploinsufficient, or dominant-negative mutations). Loss-of-function mutations can be created by:

1. Deleting an endogenous locus through a homologous recombination approach in an embryonic stem cell line. The ES clone bearing the desired mutation can be selected, trimmed, or expanded, then injected into blastocytes. The deletion/replacement approach is the most common form of the loss-of-function mutation and is now considered the classic, conventional “knockout” approach.
2. Overexpressing a dominant-negative gene using the transgenic approach (also frequently used in the *Xenopus* or *Zebrafish* systems because no ES cells are available). This approach has been successfully used in the mouse to study the functions of genes that belong to a large gene family with a redundant and/or overlapping expression pattern.
3. Microinjecting a gene-specific antisense oligonucleotide or double-strand RNA (e.g., RNAi) to reduce endogenous gene expression temporally (used mainly in *Xenopus* and *Caenorhabditis elegans*). This approach can be extended to mouse embryos to study early developmental processes.
4. Creating more sophisticated mutations—such as large chromosomal deletions or subtle point mutations—by introducing loxP (or FRT) sites into the knockout construct. The removal of a selection marker flanked by loxP sites is necessary and can be achieved by the transient transfection of a cre recombinase expression vector at the ES cell stage. Such a trimming step can also be achieved by crossing to a cre recombinase-bearing strain with desired expression time courses.
5. Selectively deleting specific cell types during development by introducing alpha-toxin at the desired locus via homologous recombination.

Potential problems and possible solutions (more detail in section 2)

Problem: Lethality; animals fail to survive to the stage of interest for analysis

Solutions: Perform a conditional knockout (cre/loxP, flp/FRT system), or an inducible system (tet, or other post-translational inducible systems for fast kinetic gene regulation), a conditional expression (tet-inducible system), or use a combination of the cre/tet system approaches.

1. Problem: Multiple phenotypes involving many organ systems/tissue types; sometimes, unexpected phenotypes
Solution: Perform a conditional knockout with a tissue-specific deletion.
2. Problem: No phenotypes or subtle phenotypes observed due to possible redundancy/compensation problems
Solution: Use of knockin (i.e., replacement) approach, double knockout approaches, or the FRT/loxP combination, or a dominant negative approach.

2) Characterizing Mutant Phenotypes

The success of a knockout/transgenic project is judged by the quality of downstream characterization, hence the quality of the investigator’s publications. The major goal of generating knockout/transgenic mice is to elucidate the mechanism of *in vivo* biological/pathological processes at the cellular, molecular, and organism levels. The variables encountered in a petri dish or in acutely lesioned animal models (axotomy/toxin) are totally different from the variables introduced by the life history/developmental processes of an entire organism. The following difficulties often complicate and interfere with a smooth characterization of mutant phenotypes. It will be prudent to take these complications into consideration at the stage of knockout construct design.

Potential problems:

1. **Early versus late phenotypes:** Early-onset lethality will prohibit characterization of later phenotypes (e.g., kidney agenesis and enteric aganglionosis in glial-derived neurotrophic factor (GDNF)^{-/-} mutants prohibited the analysis of GDNF function in dopaminergic neurons or motoneuron function after birth).
2. **Multiple phenotypes involving many organ systems/tissue types:** The investigator may or may not care about phenotypes in other systems.
3. **Unexpected phenotypes:** What should the researcher do when the desired phenotypes are not found in ^{-/-} mutants while other severe phenotypes are plentiful in systems in which he/she does not have the appropriate analysis expertise? For example, the 'suspected' function, based on many studies of GDNF on dopaminergic neurons, was not readily found in homozygous mutants, but instead, was replaced by unexpected phenotypes such as kidney agenesis and enteric aganglionosis.
4. **Subtle phenotypes:** Many knockouts fall into this category of mice, for example, haploinsufficiency with variable penetrance as found in GDNF^{+/-} mice. However, the subtle phenotypes are difficult to characterize but may be more relevant to human diseases.
5. **No obvious phenotypes found in the areas of interest:** This result is due to redundancy and the overlapping expression of members of the same gene family (e.g., *Hox* genes, *En-1* and *En-2*, etc.).

Possible solutions:

A. Most of these problems can be effectively resolved by using the knock-in approach, by introducing a functional cDNA into a specific locus. Thus, expression of the knock-in gene will follow the temporal, spatial, and quantitative level of the endogenous gene expression being targeted at (but will be functionally unimpaired). Depending on the specific application, the cDNA being introduced by homologous recombination (the knockin) can be used:

1. **To mark the cells of interest (lineage tracing):** This can be accomplished by introducing various types of markers (i.e., LacZ; alkaline phosphatase (AP); green fluorescent protein, or other autofluorescent proteins), in combination with nuclear localization signal (NLS), Tau, or other subcellular targeting motifs, through in-frame fusion, splicing acceptor or IRES element, to mark the population of cells of interest. The ontogeny (fate), morphology, and density (numbers) of these cells can be monitored qualitatively and quantitatively by either *in situ* staining or other *in vitro* manipulations. For example, GFP-marked live cells can also be purified by fluorescence-activated cell sorter (FACS) analysis, or by time-lapse video recording in explant or dissociated culture. The FACS-sorted cells, purified in a stage-specific manner followed by either *in vivo* or *in vitro* manipulation, can be used as a source of complex mRNA for probing the cDNA microarray. Thus, the investigator can dissect genetic pathways in a specific cell lineage during development, or at various stages of tumor formation/progression, or study the effects of specific treatments on the whole organism or on the cell type of interest (e.g., axotomy, ischemia, excise, smoking, irradiation, or carcinogen).
2. **To produce a conditional knockout:** A conditional knockout can be achieved by the knock-in of a cre or flp recombinase (or a combination of these) into a specific gene locus (or two loci). The expression of cre (and flp) is under the control of the endogenous locus in a tissue-specific, time-dependent manner. No prior knowledge of the regulatory element is needed, except for its expression pattern, which is derived from *in situ* hybridization studies. The temporal/spatial-restricted cre/flp expression line will lead to a selective (or conditional) deletion of the gene of interest when crossed to a line of mice in which Lox P or FRT recognition sites are made to flank the gene (or genes) of interest. In addition, a combination of the cre/LoxP and flp/FRT systems will allow selective and simultaneous deletion of two loci of interest.
3. **In a gene replacement study:** The knockin of cDNA into the locus being targeted will usually lead to replacement/substitution of the endogenous gene with the cDNA of a different member of the same gene family or a downstream gene. This approach can be used to assess the redundancy of the family member or genetic pathway, and such a substitution may rescue the deficiency that is caused by deletion of the endogenous gene.

4. In the conditional expression approach: This approach was first reported in December 1999, and may be one of the best approaches developed to date that enables the investigator to gain total temporal control of gene expression *in vivo*. By manipulating both alleles of the same gene, this approach incorporates tet-inducible expression with the knockin of a tet transactivator into a specific genetic locus. Thus, gene expression is controlled in a temporal-specific manner. Because the transactivator (rtTA or tTA) is knocked into a specific genetic locus, the temporal expression of the cDNA under the tet-inducible promoter will recapitulate the spatial expression pattern of the endogenous gene. Thus, the level of tet-inducible gene expression can be controlled in a linear fashion. Using the recently improved rtTA transactivator, this approach may offer an investigator ultimate control over the level of expression of a specific gene, either below or above its normal expression level. The approach may be adapted to other post-translational inducible system for applications requiring fast induction kinetics.

B. Other Tools of Developmental Biology

The developmental biology tools listed below can also be used to characterize mutant mouse phenotypes:

1. Chimera study: The failure of a mutant ES-derived contribution to a particular lineage/organ system would suggest an important role for a gene during the development of that lineage/system.
2. Tissue recombination assay and explant culture system: This technique is best used to study the interaction of two tissue compartments during many developmental processes and to distinguish the extrinsic versus intrinsic effect(s) of the genetic manipulation on the specific process.
3. A completely ES-derived embryo: This approach may be used to speed up the production of homozygous mice with homozygous mutant ES cells that lead to an effective reduction of genetic variability in the population.

Summary

Diverse strategies for manipulating the mouse genome are developed in the past decade to study the genetic basis of complex biological processes. This overview is just a highlight and not meant to be comprehensive or exhaustive. Investigators who are interested in using this genetic tool are encouraged to read following papers (references therein) to get started. Obviously there is no single solution to all problems. However, there are many strategies that we can explore based on specific questions we intend to address, even when the system is complex.

Suggested readings for creating mutant strains:

1. A. Porter (Review) Controlling your losses: conditional gene silencing in mammals *Trends in Genetics* 14, 73-79, 1998
2. M. K. Shin, J. M. LeVorse, R. S. Ingram and S. M. Tilghman The temporal requirement for endothelin receptor-B signalling during neural crest development *Nature* 402, 496-501, 1999
3. S. Soukharev, J. L. Miller and B. Sauer Segmental genomic replacement in embryonic stem cells by double *lox* targeting *Nucleic Acids Research* 27 e21, 1999
4. S. Urlinger, U. Baron, M. Thellmann, M. T. Hasan, H. Bujard, and W. Hillen Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity *Proc. Natl. Acad. Sci. USA*, Vol. 97, 7963-7968, 2000
5. A. Nagy (Review) Cre recombinase: The universal reagent for genome tailing *Genesis* 26: 99-109, 2000
6. N. Hardouin and A. Nagy (Review) Gene-trap-based target site for cre-mediated transgenic insertion *Genesis* 26:245-252, 2000
7. U. Muller (Review) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis *Mech. Development* 82: 3-21, 1999