Demystified . . .

Gene knockouts

J P Iredale

Our knowledge of the identity, structure, and (in some cases) function of proteins that mediate growth and development, and physiological and pathophysiological processes, is being driven at an unprecedented rate by the application of modern molecular and cell biology techniques. The enormous power that molecular biology has brought to the study of the function of molecules at a cellular and subcellular level is perhaps most dramatically illustrated when a protein perceived as a key player in a biological process is removed and the resulting phenotype analysed. This process, in which a gene is functionally deleted, is termed “knocking out a gene”. For practical reasons, relating to the ease with which reproductive biology and embryonic stem cell culture can be manipulated, gene knockouts are established in mice. In this review I will describe the methods used to generate a mouse in which a gene has been targeted to prevent expression of the cognate protein in all cells of that animal’s body (conventional gene targeting). In addition, I will describe the potential problems and pitfalls that can occur when conventionally targeted mice are used to study gene function in vivo. In the final section of the review I will describe some of the more recent developments in the field of gene targeting and some of the strategies that have been used to overcome the problems of lethal mutations and the sometimes undesired effects of systemic gene inactivation. In addition, I will touch on some of the techniques that result in gene or chromosomal modification in vivo, which have developed directly from knockout technology. It is beyond the scope of this review to cover these latter areas in detail, and the interested reader is referred to the references at the end of the review.

Why inactivate a gene?

The pace of development of molecular biology and the speed with which data on the genome are being generated challenge our current knowledge of the functions of many recently identified genes. Often a given gene’s function can be predicted, in a limited way, from sequence homology with other related family members. Alternatively, the presence of specific elements in the predicted amino acid sequence of a gene (such as a membrane spanning domain) might provide important clues to the location and function of the encoded protein by a novel sequence. However, in many cases, function cannot be assigned as a result of sequence analysis and the presence of recognisable sequence motifs cannot provide proof of specific function. An important way of determining the role of a given gene product in vivo is to inactivate the gene and then observe the resulting phenotype in cell culture or in vivo. Creating a gene knockout provides the most direct way to achieve this goal.

Molecular and cell biology have also greatly facilitated our understanding of pathological processes in vivo in human tissues, as well as in animal models and cell culture systems. In general, these studies are descriptive and, with the exception of cell culture systems, the opportunity to complement descriptive data with mechanistic studies is limited by technical considerations. This distinction between descriptive and mechanistic data is critical to the development of our understanding of pathology. Descriptive studies may show clear chronological evidence of the association of a protein with a particular process; however, this remains simply an association until the target protein is manipulated and the observed process critically altered. This is difficult to achieve in many systems in vivo, usually requiring the development of delivery systems or vectors that can prevent the synthesis of a particular protein or inactivate it (for example, by the use of antisense delivering viral vectors or the injection of inactivating antibody or ligand). These methods are costly and some tissues and organs remain non-targetable. The development of a gene knockout mouse facilitates the study of a protein’s function in vivo.

A final and important stimulus to the development of gene knockout animals is the establishment of models of inherited human disease. These models provide valuable data on the detailed pathological processes that result from the loss of gene function (for example, those events that might occur in a tissue before a disease becomes clinically manifest and therefore not available for study as human biopsy material). In addition, these models may also prove to be pivotally important in the development of new treatments. This is particularly likely to be true for gene therapy, in which the functional effectiveness of each individual strategy to replace absent or defective gene function will have to be investigated and demonstrated in model systems.

Targeting a gene

Two processes are crucial to the generation of a gene knockout animal. The first is the observa-
Design and construction of the targeting vector

Genomic DNA is a continuous molecule in which the genes are divided up into exons (those parts of the DNA represented in the final mRNA) interspersed with introns (which contain non-coding and regulatory regions) (fig 1). Before the inactivation vector can be designed and constructed, the target gene (exons and introns) will have to be identified, cloned, mapped, and, at least in part, sequenced. This is accomplished by screening a genomic library derived from the ES cell line strain, using the relevant cDNA for a probe. Clones that hybridise with the probe (and so demonstrate homology) are amplified, restriction mapped, and sequenced. Ultimately, a clone is derived that represents a significant part of the gene, or the whole gene, and which contains the critical exon to be targeted (see below). The isolated clone is then restriction mapped and sequenced to identify convenient sites for subcloning of fragments and the exon–intron boundaries, and to define and map crucial exon sequences.

The specific inactivation strategy will vary from gene to gene, but will always involve the targeting of an exon that is crucial for the biological function of the protein (such as the sequence that encodes the active site of an enzyme). Gene inactivation is achieved by replacing this exon with one that has been critically mutated. Because homologous recombination will occur infrequently, even in successfully transfected cells, the mutation that disrupts the exon is usually induced by introducing a rogue sequence that can also act as a marker for recombination. This is achieved using a cassette of DNA that contains the sequence for a selectable marker (often neomycin resistance) with appropriate regulatory elements. By using the cassette to disrupt the target exon, gene expression can be prevented and cellular selection can be undertaken, on the basis of neomycin resistance, to enhance the efficiency of screening for successful recombinants.5-6

Once the exon to be targeted has been determined, the inactivation vector can be designed and constructed. This review will focus on the use of a simple “classic” targeting strategy using positive and negative selection for targeted recombinants. A detailed discussion of other methods of targeting and a detailed description of more advanced methods of conditional targeting have been published elsewhere.10-15

The inactivation vector is constructed by the sequential subcloning of fragments, generated by polymerase chain reaction (PCR) of the cloned genomic sequence, into a bacterial plasmid vector, which has the neomycin resistance cassette strategically placed in the plasmid polylinker (multiple cloning site) (fig 1A and B). An advantage of using PCR to generate the arms of homology is that primer design can be used to induce further mutations in the
Targeting ES cells: the induction of a mutation by homologous recombination

ES cells can be grown in tissue culture and will retain their pluripotency if certain crucial culture conditions are met. ES cells are grown in culture and harvested when they are rapidly dividing. Dividing ES cells are transfected with the inactivation vector, typically by electroporation, where the passage of a brief electric current renders the ES cell membrane permeable to large molecules such as DNA. In successful transfectants, the inactivation vector has three possible fates: (1) it might not be incorporated into the host DNA (fig 1D); (2) it might be incorporated into the host DNA in a random manner (fig 1F); and (3) in a small number of successful transfectants there will be homologous recombination with replacement of the targeted exon by the mutated homologue containing the neomycin resistance cassette (fig 1E). In this event, and in the presence of a random incorporation of the vector, the cell will demonstrate G418 resistance and the transfected population can be enriched for clones that have incorporated the vector by incubation in the presence of G418 (G418 inhibits eukaryotic ribosomes and will therefore select for cells expressing neomycin resistance). However, a random recombinant will be more likely to have incorporated the HSVTK cassette also and would therefore be susceptible to gancyclovir. By contrast, homologous recombination, which is driven by the specific sequence of the long and short arms of homology, will be less likely to have resulted in recombination of the HSVTK cassette (fig 1E and F). Thus, by incubating the transfected cells in the presence of both neomycin and gancyclovir, the ES cells will be enriched predominantly for successfully targeted recombinants (fig 1E).

Over a period of seven to 10 days, the ES cells that survive in G418 and gancyclovir grow to form colonies, which are visible to the naked eye and can be physically lifted from the plate. Lifted colonies are then divided, with half being replated to provide a living clonal line and half lysed for DNA extraction and analysis. The extracted DNA is used to screen for successful homologous recombination events by PCR using mutually exclusive primer pairs. When putative targeted clones have been identified, successful targeting is confirmed by Southern analysis and clones with the mutated allele are expanded to provide ES cells for implantation.

In some circumstances, the targeted ES cells may themselves be useful for experimental work. Indeed, ES cells can be used in vitro to study the effects of gene inactivation on embryogenesis. Usually, the selection conditions are altered to favour the survival of clones with two targeted alleles, or a further round of targeting may be required to inactivate the second allele, although X-linked alleles can be targeted and functionally inactivated after one round of transfection of male ES cells.

Development of chimaeric mice: implantation of the targeted ES cells

The targeted ES cells are microinjected into blastocysts derived from pregnant mice (fig 2). In turn, these blastocysts are implanted into pseudopregnant females (obtained by mating normal female mice with vasectomised males) and allowed to proceed to term and through normal postnatal development. The presence of a second unmutated allele in each cell normally prevents the loss of the inactivated allele from resulting in premature fetal or premature postnatal wastage.

If the ES cells retain their pluripotency, the mutated cell line will be represented in the germ line of the chimaeric animals. Backcrossing of the chimaeras with normal mice is then undertaken to generate heterozygotes for the null allele (fig 2). Ultimately, backcrossed heterozygous mice can be interbred to yield homozygous mice that contain two mutated alleles: a gene knockout mouse.

Analysis of the gene knockout mouse

It is incumbent on the researcher who creates a gene knockout mouse to demonstrate functional inactivation of the targeted allele. This usually involves confirming the knockout genotype in the adult animal by Southern
analysis and undertaking mRNA studies and protein assays to detect and quantify the activity of the protein encoded by the targeted gene. When effective targeting of both alleles has been confirmed in the mouse in vivo, analysis of the resulting phenotype can be undertaken.

Often the phenotype of the gene knockout mouse is apparent. Indeed many mutations are lethal in utero or neonatally. Nevertheless, even a lethal phenotype can be informative about the function of the targeted gene product. For other targeted alleles, the researcher may not expect the phenotype to be apparent in the normal mouse. This may occur—for example, in the case of a mediator expressed during inflammation. Clearly, the analysis of such a knockout might involve the initiation of a specific model inflammatory process to observe the phenotypic response. For others, the phenotype might be quite unexpected, and may lead directly to the identification of a new role for the inactivated gene product or other proteins with which it interacts.

Thus, the development of the technology to inactivate gene function provides an exciting and extremely powerful research tool. However, the resulting phenotype may not be apparent even after extensive investigation. Moreover, the lethality of certain mutations and the difficulty of interpreting certain phenotypes might mean that the investigator obtains only very limited information after a considerable amount of experimental effort. These problems and the possible strategies that can be applied to resolve them are considered below.

Problems and pitfalls in using gene knockout mice

One major potential problem with the use of gene knockout mice is the fact that because the mouse is a different species it might use a different protein to fulfill the function of the target protein in humans. This might be a related protein if the target is a member of a family, or it might be a protein with a less direct structural relationship, which subserves a similar function. Thus, gene knockout technology is not universally successful in determining the roles of proteins in a manner that is directly applicable to human disease. However, these potential problems are offset, at least in part, by the relatively large volume of data accrued about murine immunology, physiology, and pathophysiology, which in many cases will allow the investigator to either plan alternative experimental strategies, or to anticipate phenotypic variations in the mouse.

As mentioned above, a lethal gene knockout may or may not supply useful information. It is possible to develop a lethal knockout that ultimately yields little valuable information except that the targeted gene is a vital one. It might also be difficult to generate a gene knockout line if the targeted gene is important for fertility or the processes of reproduction, implantation, and gestation. In this eventuality, the experiment may proceed to the stage of breeding homozygous knockout animals but might then be associated with a high fetal loss that precludes the generation of adult mice for study.

An additional but perhaps less apparent problem is that the gene knockout mouse has a normal phenotype. In this circumstance, the investigator is left with a potentially endless search for a phenotype in which he or she has to determine just how far they will go to document any effects of gene deletion. The possible explanations for a normal phenotype are many. The most obvious one is that the target protein is not critical for the functional role that earlier studies suggested. This is a concept that may be rather galling to grasp if one has built a career around the importance of a particular protein. Alternatively, the mutation may have resulted in a phenotype that is too subtle to be apparent. For example, some knockout phenotypes have been demonstrated to result in unexpected behavioral changes, and it will
systemic gene deletion can be obtained by administration of interferon. (IFN) inducible MX promoter) results in a mouse in which Cre gene expression and Cre from a transgene under the control of an inducible promoter (in this case the interferon inducible MX promoter) results in a mouse in which Cre gene expression and systemic gene deletion can be obtained by administration of interferon.

be apparent to the reader that subtle behavioural changes may well go unnoticed. A further problem is that of biological redundancy. Many key proteins are members of families or super families. As such, the deletion of a single member may be compensated for by sibling proteins, either by a distinct compensatory upregulation, or by virtue of the sheer size of a family of proteins with overlapping functions. A final possible explanation for a normal phenotype is that the targeting, although technically successful, has not prevented the expression of a truncated or mutated protein with bioactivity.

Lastly, an important drawback to the development of gene knockout animals may be cost. At the present time, the cost of creating a gene knockout mouse lies between £15 000–20 000 ($25 000–33 000). This may at first seem prohibitively expensive; however, if it is necessary to inactivate a protein functionally in vivo for research purposes, then alternative strategies based on inactivating antibodies or antisense technology may prove equally expensive and are likely to involve considerable time and costs in establishing appropriate delivery systems.

Conditional gene targeting and the further development of gene targeting technology

Because “classic” gene knockouts may result in a lethal phenotype or be associated with phenotypes that result from compensatory gene expression, a number of laboratories have been active in developing alternative “conditional” gene targeting strategies in which the targeted gene remains intact and is expressed normally until inactivation is required. Foremost among the systems developed to obtain conditional gene deletion is the Cre/lox system. Cre (causes recombination events) recombinase is an enzyme derived from the P1 phage which mediates excision (and in effect deletion) of genomic DNA flanked by 34 bp loxP sites in the same orientation (an allele or DNA segment thus flanked is termed “floxed”) to leave a single loxP motif. Cre recombinase has been shown to be active in vivo in murine cells. Moreover, loxP sites can be introduced into the non-regulatory introns of genomic DNA in a manner that does not prevent normal expression of the encoded protein from the exons. Thus, by targeting a gene in ES cells in a manner identical to that described above, but using a vector which introduces a neomycin resistance cassette and strategically placed loxP sites, an allele in ES cells can be appropriately mutated (fig 3A and B). As before, targeted clones in which homologous recombination has taken place can be selected for by G418 resistance. Successful targeting in such clones is then confirmed by PCR.

Targeted ES cells containing the loxP sites are then transfected with a Cre expressing plasmid and cultured in gancyclovir. Unsuccessful recombinase activity or a recombination event that deletes genomic DNA but preserves the selection cassettes will be lethal for the cell in the presence of gancyclovir. A further possible result of Cre gene expression is the excision of the major part of the targeted area as a result of recombination involving the loxP sites closest to the 5' and 3' ends. This is, in effect, a systemic mutation and an alternative to a classically targeted systemic mutation as described above, a so called type 1 deletion (fig 3C). Alternatively, Cre mediated recombination compatible with survival in the presence of gancyclovir is the result of a type 2 deletion (fig 3C). A type 2 deletion results in loss of the selection cassette but with the essential gene exon(s) intact but floxed and therefore vulnerable to Cre mediated excision in the future. The cells containing a type 2 deletion are cultured and introduced into blastocysts and chimaeras and, ultimately, homozygous mice are derived as described above. By these means an animal can be created in which the target gene is expressed as normal, but because of the presence of strategic loxP sites the gene remains vulnerable to Cre mediated deletion in adulthood.

The second limb to the development of conditional gene deletion is the development of Cre expression systems for adult mice in vivo. These fall into two groups. It is now well established that the Cre gene can be expressed from a transgene under the control of a cell specific promoter (cell specific Cre gene expression). The alternative strategy is to have the Cre gene regulated by an inducible promoter, which is in turn responsive to an external stimulus (temporal or inducible Cre expression). Both of these strategies have now been used successfully to mediate gene deletion in adult mice in vivo by crossing an animal with a floxed target gene with a mouse transgenic for either cell specific promoter driven Cre expression (for example, using a T cell lineage promoter) or inducible Cre expression (for example, using the interferon inducible MX promoter).
the former example, the result is that gene deletion is conditional to those cells that express the gene encoding Cre (fig 4), and the more generalised results of gene loss are not observed. In the latter example, gene deletion occurs in many (in fact potentially all) cells after Cre gene expression is induced, but this occurs in the adult mouse at a time determined by the investigator (fig 5).

### The future of gene targeting

The evolving experience with Cre/lox technology and the continuing development of murine strains with new cell specific Cre driven transgenes means that the techniques will become more widely applicable. In addition, alternative inducible Cre promoter systems are being developed. The range of cells in which Cre is after inducible expression is by no means exhaustively studied and alternative inducible promoters might ultimately be shown to be more or less efficient in individual cell types. The use of loxP sites flanking introduced stop codons might provide a mechanism for keeping a transgene suppressed until it is “knocked in” by Cre gene expression. Cre is not the only experimental recombina

### Postscript

My aim in this explanatory review was to attempt to explain the science and technology of gene knockouts to the general reader. For reasons of brevity, this review has only really skirted over the science and applications of gene targeting. For this reason and to facilitate further study by those readers whose interest has been stimulated by this introduction, I have attempted to limit the reference list to a series of excellent review articles that illustrate the different facets of gene targeting described above, so that it may function more as a bibliography than an exhaustive reference list. The interested reader is also referred to the excellent review series: “Molecular medicine in genetically engineered animals” published in the *Journal of Clinical Investigation* between January and August 1996.

The author is grateful to Dr Jurgen Roes for his valuable advice on the preparation of this manuscript. At the time of writing Dr JP Iredale was in receipt of an MRC Clinician Scientist Fellowship. Dr JP Iredale also gratefully acknowledges the support of the Wessex Medical Trust and the British Liver Trust.

Demystified ... gene knockouts.

J P Iredale

*Mol Path* 1999 52: 111-116
doi: 10.1136/mp.52.3.111

Updated information and services can be found at:
http://mp.bmj.com/content/52/3/111.citation

These include:

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/