Current and future approaches using genetically modified mice in endocrine research

Rachel A. Davey and Helen E. MacLean

Department of Medicine, University of Melbourne, Austin Health, Heidelberg, Victoria, Australia

Submitted 15 March 2006; accepted in final form 4 May 2006

Davey, Rachel A., and Helen E. MacLean. Current and future approaches using genetically modified mice in endocrine research. Am J Physiol Endocrinol Metab 291: E429–E438, 2006. First published May 9, 2006; doi:10.1152/ajpendo.00124.2006.—Genetically modified mouse models have been used widely to advance our knowledge in the field of endocrinology and metabolism. A number of different approaches to generate genetically modified mice are now available, which provide the power to analyze the role of individual proteins in vivo. However, there are a number of points to be considered in the use and interpretation of these models. This review discusses the advantages and disadvantages involved in the generation and use of different genetically modified mouse models in endocrine research, including conventional techniques (e.g., overexpression, knockout, and knock-in models), tissue- and/or time-specific deletion of target genes [e.g., Cre-loxP and short interfering (si)RNA transgenic approaches], and gene-trap approaches to undertake functional genomics. This review also highlights the many factors that should be considered when assessing the phenotype of these mouse models, many of which are relevant to all murine physiological studies. These approaches are a powerful means by which to dissect the function of genes and are revolutionizing our understanding of endocrine physiology and metabolism.

OVEREXPRESSION OF TARGET GENES

The first widely used approach to study gene function in vivo was to produce transgenic mice that overexpress target genes. This requires the full-length coding sequence (cDNA) of a gene to be cloned downstream of a promoter to provide ubiquitous or tissue-specific expression (Fig. 1A). The choice of promoter can provide universal expression, tissue specificity, time specificity, or the ability to be regulated.

Advantages/Disadvantages

The major advantages of the transgenic approach are that it is relatively straightforward and inexpensive, high levels of target gene expression can be achieved, and transgenic overexpressing mice often demonstrate an obvious phenotype. One disadvantage is that the site of integration of the transgene into the genome can seriously affect tissue specificity and levels of transgene expression (41), so a number of founder lines may need to be screened. The use of insulator sequences to protect the transgene expression from effects of integration site have so far provided variable success (51). Many promoters have now been used to drive transgene expression in endocrine tissues. However, even well-characterized promoters are often expressed at lower levels in nontarget tissues, so rigorous analysis should include examination of transgene expression in a range of tissues. Although overexpression models by their nature produce nonphysiological levels of gene and protein expression, they can still provide valuable insights into normal gene function in vivo.

Examples of Overexpression Models

Numerous overexpression models have been reported in the literature over a number of years that have increased substantially our understanding of the physiological role of endocrine genes (3). For example, the paradigm of regulation of pancreatic insulin secretion is that glucose metabolism increases the intracellular ATP-to-ADP ratio (ATP/ADP), causing inhibition...
A mutant variant of the Kir6.2 subunit of the KATP channel that elegantly in vivo by overexpression in pancreatic ATP/ADP sensitivity in insulin secretion was demonstrated plasma membrane, Ca²⁺ entry, and insulin release. This is transferred into a pseudopregnant female, resulting in transgenic offspring overexpressing the target gene specifically in the tissue of interest. B: alternatively, targeting construct consisting of short interfering (si)RNA for the target gene under the control of a tissue- and/or time-specific promoter can be used to generate knockdowns. The sense/loop/antisense sequence forms a short hairpin structure, which when expressed is cleaved to form siRNAs.

of K⁺-sensitive ATP (K_{ATP}) channels, depolarization of the plasma membrane, Ca²⁺ entry, and insulin release. The role of ATP/ADP sensitivity in insulin secretion was demonstrated elegantly in vivo by overexpression in pancreatic β-cells of a mutant variant of the Kir6.2 subunit of the K_{ATP} channel that has reduced ATP sensitivity (22). β-Cell-specific expression was achieved using the widely used rat insulin promoter (RIP) to create transgenic mice with hyperactive K_{ATP} channels in the pancreatic β-cells. Transgenic mice had reduced ATP sensitivity, hyperglycemia, and insulinopenia in peripheral tissues, and were hypoinsulinemic and severely hyperglycemic, demonstrating the critical role of normal K_{ATP} channel activity in physiological control of insulin secretion.

More recent variations using transgenic overexpression illustrate the ongoing relevance of this approach. Using a combination of global knockouts and tissue-specific transgenetics can provide information about tissue-specific function. For example, knockout mice lacking the gene encoding proopiomelanocortin (POMC) have early-onset obesity, demonstrating that POMC pathways are essential for normal energy homeostasis (54). To dissect the role of POMC peptides in the central nervous system vs. peripheral actions, transgenic mice were created that overexpress Pomc in pituitary cells but not neurons, allowing secretion of POMC-derived peptides into the circulation (45). Compound Pomc knockout/transgenetics were generated that effectively had neuron-specific deficiency of POMC. These mice exhibited an increased severity of the metabolic phenotype compared with Pomc knockouts, most likely due to the actions of glucocorticoids to increase appetite and metabolic rate. Thus the use of this combined global knockout/transgenic overexpression strategy demonstrated a critical role of central actions of POMC peptides in regulation of energy balance that could not be rescued by restoring peripheral actions. This illustrates how precisely developed overexpression models can be used to localize tissue-specific actions of factors in vivo.

A modification of the transgenic approach is to block the action of a target protein through the use of transgenes encoding dominant negative protein variants (24) or, more recently, using transgenic RNA interference (RNAi) approaches. These provide simplified versions of tissue-specific target knockdown compared with the more complex Cre-loxP approach described earlier (Tissue-Specific Deletion of Target Genes) and can provide useful information in cases where gene knockouts are lethal or uninformative. RNAi has been used predominantly in vitro to suppress gene expression through the use of target-specific small interfering RNAs (siRNAs). More recently, in vivo approaches have used viral or other forms of delivery to silence gene expression in particular organs or target cells. For example, intravenous injection of an adenoviral siRNA construct against the ATP-binding cassette transporter A1 (ABCA1) gene targeted the liver of infected mice and decreased endogenous ABCA1 protein levels by ~50% (36). This caused a decrease in plasma HDL cholesterol levels, and at 7 days postinfection mice had an alteration in postprandial lipoprotein metabolism. Thus studies indicate that gene silencing using acute delivery of siRNAs in vivo can produce demonstrable physiological responses. However, the duration of gene knockdown is limited and may not be sustained sufficiently to allow the development of a phenotype.

An extension of the in vivo siRNA approach is to generate transgenic mice with tissue-specific expression of transgenes encoding siRNAs (Fig. 1B), which allows prolonged gene silencing. A limited number of reports have already described the success of this approach. For example, in a transgenic mouse model of amyotrophic lateral sclerosis (ALS), expression of a mutant variant of the SOD1 gene produces a toxic protein that induces the disease. Transgenic mice were developed that expressed anti-SOD1 siRNA in the nervous system (38). These mice showed an 80% reduction in mutant copper/zinc superoxide dismutase-1 protein levels, and the siRNA silencing effect was maintained up to the age of 11 mo and was inherited stably over three generations. More importantly, ALS/anti-SOD1 double transgenic mice had a complete suppression of the ALS phenotype when examined up to the age of 10 mo. Thus this study demonstrates that it is possible to generate transgenic siRNA mice in which gene expression is sustained and in which a physiologically relevant outcome is achieved.
The advantage of using the transgenic siRNA knockdown approach is that sustained and targeted gene knockdown can be achieved in vivo by using the transgenic methodology, which is relatively rapid and straightforward. However, in addition to the disadvantages associated with generating transgenics discussed previously, a number of other potential problems in this relatively new approach require careful consideration. The specificity of gene knockdown can be influenced by the tissue specificity of siRNA transgene expression. Because RNA polymerase III promoters are often used to drive siRNA expression, the choice of promoters and tissue specificity of expression may be limited. In addition, target gene silencing may not be specific, as siRNAs can have off-target effects. The degree of knockdown can be affected by the efficacy of different siRNAs, and the level of knockdown in different cell types may vary, which could be dependent on the levels of the endogenous mRNA or other as yet unidentified factors. In addition, target gene silencing may not be specific, as siRNAs can have off-target effects. Thus, further studies are likely to be required before the utility of this approach can be fully evaluated.

Overall, the use of transgenic overexpression models has provided a wealth of information regarding the normal actions of endocrine and metabolic factors. The relatively low cost and ease of generating these models, particularly combined with more recent modifications to the approach, will ensure its continued use in studying gene function in vivo.

**GLOBAL KNOCKOUT MOUSE MODELS**

The development of global knockout mouse models is also becoming comparatively routine as an increasing number of laboratories and commercial services are set up to facilitate the generation of knockout mice. The generation of knockouts utilizes cultured embryonic stem (ES) cells to target gene deletion in vitro by homologous recombination (Fig. 2).

**Advantages/Disadvantages**

Global knockout mouse models are more predictable and reproducible than overexpression models, because the target gene is ablated in all tissues. Phenotypes caused by loss of target gene function are likely to provide insight into the physiological roles of these gene products. In addition, novel or unexpected actions of target genes often emerge because, unlike transgenic overexpression models, global knockout models are not limited to a particular tissue or system. However, although generation of knockouts is comparatively routine, it is nevertheless still time consuming and expensive compared with the conventional transgenic approach. In addition, global deletion of factors essential for embryonic development will result in embryonic or neonatal lethality, prevent-
invited review

1. However, deletion of osteoblasts. The Runx (runt domain) family of transcription factors at later stages. For example, Men1 gene knockouts were generated to investigate the role of the Men1 gene product (menin), which is mutated in multiple endocrine neoplasia type 1. However, deletion of Men1 is embryonic lethal, and knockout embryos have widespread tissue abnormalities associated with defects in cell differentiation and senescence (6). Thus, although this study demonstrated a critical role for menin in embryonic development, this global knockout model did not allow the investigation of its role as a tumor suppressor in endocrine tissues.

The power of global knockout models is that they can uncover unexpected actions of target genes. A classic example of this is the identification of Runx2 (also known as CBFA1, PEBP2αA, or AML3) as a master transcriptional regulator in endocrine tissues. The Runx (runt domain) family of transcription factors was initially characterized from the study of acute myeloid leukemia, where it was found that the Runx1 (CBFA2, PEB2αB, AML1) gene frequency undergoes translocations (13). The lack of fetal liver hematopoiesis in Runx1 knockout embryos confirmed the critical regulatory role of Runx1 in hematopoietic development (39). In contrast, the expression pattern of the related family member Runx2 had been shown to be limited to the thymus and testis, and many T cell-specific gene promoters were known to contain Cbf response elements; therefore, it was hypothesized that Runx2 may be involved in T lymphocyte-specific transcriptional regulation. Surprisingly, the striking phenotype of the Runx2 knockout mice was their lack of formation of the entire ossified skeleton, with development of only the cartilaginous skeleton (21, 31). Runx2-null mice failed to form differentiated osteoblasts, uncovering this novel and critical role for Runx2 in bone. Thus, by not focusing on gene function in specific tissues, phenotypic analysis of global knockout models can potentially reveal actions in novel or multiple organs or systems.

Tissue-Specific Deletion of Target Genes

The emergence of the Cre-loxP system has overcome the difficulties often encountered when traditional global knockout models such as embryonic or neonatal lethality are used. The Cre-loxP system achieves this, as the deletion of the target gene can be controlled in a tissue- and/or time-specific manner. As such, there is a wide range of Cre-loxP mouse models described in the literature in all scientific disciplines.

To use the Cre-loxP system in mice, two mouse lines are required: the floxed line, in which the region of the gene targeted for deletion is flanked by two 34-base pair loxP sites, and a second mouse line in which the expression of Cre recombinase (Cre, a site-specific recombinase) is under the control of a time- and/or tissue-specific promoter. The floxed mouse line is produced by homologous recombination in ES cells, whereas the Cre transgenic mouse line is produced by standard transgenic approaches (overexpression of target genes). In mice inheriting both the floxed target gene and the Cre transgene, the Cre recognizes pairs of loxP sites and deletes the intervening sequence (Fig. 3). The loxP sites are placed within the intron of the gene and therefore theoretically do not interfere with the normal function of the gene. As such, the floxed target gene functions normally in all tissues and is deleted only in tissues where Cre is expressed.

Advantages/Disadvantages

The most significant advantage of the Cre-loxP system is its flexibility. Unlike the standard global knockout approaches, a number of options are available with the Cre-loxP system. It can be readily used to generate global knockout mouse lines by breeding the floxed target mouse line with a line that expresses Cre ubiquitously. Using the Cre-loxP mouse system to generate global knockout models is particularly advantageous if deletion of the target gene results in infertility (30, 40). The Cre-loxP system can also be used to investigate the role of a target gene at specific stages of development by breeding the floxed target mouse line with an inducible Cre mouse line, thus allowing the deletion of the target gene to be controlled in a time-specific manner. A significant feature of the Cre-loxP system is its potential for expansion. Once developed, the target floxed mouse line can be bred with a number of different tissue- and/or time-specific Cre mouse lines, enabling the investigation into the role of the target gene at a number of sites. Similarly, the Cre-expressing mice can be crossed with a number of different floxed mouse lines to study a number of genes in a tissue of interest.

Researchers should also be aware, however, of the potential disadvantages that may be encountered when using the Cre-loxP system. For example, when the Cre mouse line is being designed, it is often difficult to find a promoter to drive the expression of Cre specifically to the tissue of interest and with sufficient activity to result in complete excision of the target gene, and, as for all transgenics, expression is largely dependent on the site of integration. In the floxed line, the placement of the loxP sites and the presence of the selection cassette (such as the neomycin resistance gene neo) can also have an impact on the expression levels of the target gene in the control mice. With the inducible Cre models, the inducer may have a significant impact on the phenotype. For example, a widely used inducible Cre model is one in which the expression of Cre is activated by the anti-estrogen tamoxifen. Tamoxifen, however, can also have profound effects on estrogen-responsive organs such as the bone, uterus, breast, and liver, independently of target gene deletion.
Examples of Tissue-Specific Deletion of Target Genes

An example of the power of tissue-specific gene deletion comes from the study of the role of the enzyme phosphoenolpyruvate carboxykinase (PEPCK) in gluconeogenesis and hepatic energy metabolism. The Cre-loxP system was used to specifically delete the PEPCK enzyme in liver, as global PEPCK knockouts die within 3 days (43). Results from this study suggest that PEPCK exerts much less control of gluconeogenesis than previously thought. However, observations of marked steatosis during fasting in these mice indicated that PEPCK also regulates hepatic lipid metabolism in addition to having marked alterations in the expression of a variety of hepatic genes involved in energy metabolism. Of significant interest from this study was the observation that the presence of the neo cassette in the PEPCK-floxed mouse line resulted in a functionally impaired allele. Liver and kidney PEPCK protein and enzyme activity was reduced by 10–20% in floxed mice with the neo cassette compared with wild-type and floxed mice that had the neo cassette removed. This study clearly highlights the potential interference by the presence of a selection marker within an intron on target gene expression.

Deletion in a time-specific manner is also a very useful approach to study gene function. For example, although it is well established that IGF-I and its receptor are expressed in the heart and act in an autocrine/paracrine manner to regulate blood pressure, the role of IGF secreted from the liver and acting in an endocrine manner was poorly understood. To address the potential endocrine action of IGF-I to regulate blood pressure, Tivesten et al. (50) generated a liver-specific inducible IGF-I knockout using the Cre-loxP system. Liver-derived IGF-I was deleted in mice at 4 wk of age by use of the Mx1 promoter, the expression of which can be induced by treatment with interferon (IFN)-α or -β or polynosinic-polycytidylic acid (44). Liver-specific IGF-I knockouts had increased blood pressure accompanied by secondary decreases in markers of cardiac performance, thereby identifying that circulating liver-derived IGF-I is involved in the physiological regulation of blood pressure and peripheral resistance. This study is an excellent example of how the Cre-loxP system can be applied to ascertain the physiological role of a growth factor that can act both locally and systemically.

The Cre-loxP system can be used to identify the role of endocrine factors at specific times of pre- and postnatal development. For example, it is well established that bone mineral density (BMD) is correlated with circulating IGF-I levels (1, 26). To address the role of circulating IGF-I during the attainment of peak cortical BMD, liver-derived IGF-I was deleted in mice between 4 and 8 wk of age by use of a tamoxifen-inducible Cre-loxP system (53). Although the reduction in circulating hepatic IGF-I levels resulted in a reduction in circulating IGF-I levels. Although the reduction in control floxed IGF-I mice was lower in magnitude than that observed in liver IGF-I
Cre-loxP mice, this study clearly demonstrated not only the power of the inducible approach but also the requirement for control mice treated with the inducer for correct interpretation.

**KNOCK-IN MODELS**

In addition to creating global and tissue-specific knockouts, ES cell gene targeting can also be used to generate knock-in models. In these models, the endogenous gene is replaced with a mutant variant (Fig. 4) in order to address the role of specific functional domains, amino acid residues, or signaling pathways in vivo.

**Advantages/Disadvantages**

The power of knock-in models is that they can be used to determine the effects of subtle changes in protein structure or function. Knock-in mice can be used to model human diseases (14) or to determine the functional significance of particular receptor-signaling pathways (18). When knock-in models are being generated, it is critical to demonstrate that the only change in the targeted locus is the introduced knock-in mutation and that other sequences, such as an intronic neo cassette, do not decrease endogenous gene expression.

**Examples of Knock-In Animal Models**

An example of using knock-ins to study different functional activities of a protein comes from study of the glucocorticoid receptor (GR). Members of the steroid receptor superfamily act classically as ligand-dependent transcription factors. The GR is essential for survival, as GR-null mice die shortly after birth with widespread abnormalities (10). To investigate the importance of DNA binding in GR function, a knock-in model was generated with a point mutation that abolishes receptor dimerization and transactivation (GR<sup>dim</sup>) (37). The targeting vector used in this study also had the neo cassette flanked by loxP sites so that it could be removed in targeted ES cells prior to the generation of mice. In contrast to the GR knockout mice, GR<sup>dim/dim</sup> mice were viable and fertile, although they had abnormalities in the hypothalamic-pituitary axis and thymus, among other tissues. This model is now being used to investigate the different roles of GR-dependent transactivation and transrepression; more generally, it illustrates the precision and power of functional studies that can now be undertaken using knock-in models.

**PHENOTYPIC ANALYSIS**

Many factors need to be taken into consideration when one is assessing the phenotype of a genetically modified mouse strain. A number of these examples discussed below are relevant not only to genetically modified mouse models but to all murine physiological studies.

Caution must be taken when extrapolating the findings in mouse models to the human endocrine system, as a number of important differences exist between the two species. For example, the steroidogenic enzymes differ in their tissue and developmental expression between mice and humans. Unlike in humans, the rat adrenal does not produce androgens, as they do not express the 17α-hydroxylase/17, 20-lyase cytochrome P-450 (P450c17) gene, also known as Cyp17. P450c17 catalyzes the 17α hydroxylation of progesterone and pregnenolone (34), which are required for the production of androgens and glucocorticoids. As a result, the adrenals of mice produce corticosterone rather than cortisol, which is produced by the adrenals in humans. Differences also exist in substrate preference of the steroidogenic enzymes between species (33). One such example is 17β-hydroxysteroid dehydrogenase (17HSD1) which uses estrogens as a substrate in humans, whereas in rodents the enzyme catalyzes the conversion of both estrogens and androgens (33).

Environmental factors, for example animal handling and housing procedures, have also been identified to play an important role in modulating the phenotype of genetically modified mouse models. Factors such as cage density, diet, and light-dark cycles can have an impact on the phenotype of a mouse line (9), and as a result standard operating procedures have been developed to allow direct comparison between phenotypic analyses performed in different research laboratories (http://www.emprex.har.mrc.ac.uk) (16). Other experimental protocols that should be controlled throughout the duration of the experiment include the length of fasting time and the site and anesthesia procedure for blood collection (32). Alterations in these procedures may have a significant impact on the interpretation of the data. For example, we have shown here that serum parathyroid hormone (PTH) levels are significantly elevated in blood samples taken by cardiac puncture (16.9 ± 5.1 pmol/l) compared with levels measured in blood collected from the tail vein (2.0 ± 0.5 pmol/l, P < 0.0001), suggesting that PTH drains from its site of production in the parathyroids directly into the left ventricle of the heart.

When mouse models for the study of endocrine and metabolic bone diseases are used, one of the most obvious differences between mice and humans is posture and, therefore, the biomechanical loading patterns of the bone (27). Also, linear growth in humans ceases after epiphyseal closure, whereas in mice the growth plate does not fuse (27). Despite the fact that the epiphyses remain open in mice, linear growth is minimal in mice by 4 mo of age (5).

Despite the differences that exist between the endocrine systems of rodents and humans, significant knowledge can be gained from mouse models if they are used and interpreted.
appropriately. Of particular importance when the phenotype of genetically modified mouse models is being interpreted is their genetic background, because a number of endocrine parameters differ markedly between inbred strains of mice. The process of generating genetically modified mouse models, in particular by homologous recombination techniques or by mating two mouse lines when Cre-loxP models are generated, results in offspring of mixed genetic background. A large number of differences exist between the mouse strains commonly used for transgenic technology (e.g., C57BL/6, Sv129/J, FVB, and ICR). Reproductive traits such as litter size, testis weight, uterine weight, sperm production, and hormone-induced ovulation rate differ between mouse strains (8, 47). The capacity of the Leydig cells to produce testosterone is also strain related and determined by the amount of Cyp11a1 (P450scc) per Leydig cell (34). Of particular importance to the study of diabetes and other metabolic diseases is that different inbred strains have different inherent metabolic characteristics. For example, DBA/2 mice gain more weight and fat than diet- and age-matched C57BL/6 or 129T2 mice, whereas resting energy expenditure is greater in C57BL/6 mice than in DBA or 129T2 mice (15). In addition, the number and mass of pancreatic islet cells differ significantly between mouse strains, with C57BL/6 mice having the lowest mass per body weight of islet and β-cells (7). Other parameters that differ between different inbred strains of mice include bone mineral density (2, 5), mechanical strength (2), thermoregulation (28), and kidney and adrenal weight indexes (12). An extensive database containing physiological, anatomic, and behavioral data for different inbred strains of mice has been established by the Jackson Laboratories (http://www.jax.org/phenome) (17).

A mix of genetic backgrounds can, therefore, interfere significantly with the interpretation of the data. To avoid this problem, the recombinant mouse lines should be backcrossed to a single genetic background for a number of generations. A minimum of five generations of backcrossing will achieve >96% homogeneity.

The animals used as controls should also be carefully considered, particularly when the Cre-loxP system is used. As discussed in *Tissue-Specific Deletion of Target Genes*, the presence of a selection cassette such as neo and/or the position of the floxed sites within an intron of the target gene may result in variable levels of target gene expression. For this reason, it is essential to include the floxed mouse line as a control in addition to wild-type and Cre littermates. Also when inducible model systems are used, a control group (such as the floxed mouse) that has been treated with the inducer should always be included for comparison (*Tissue-Specific Deletion of Target Genes*).

In situations where genetically modified mice have a mild, or lack of, discernible phenotype, applying a physiological stress may uncover target gene functions. This could include dietary manipulation, for example, by placing mice on a high-fat diet (25); altering hormonal status, such as by performing ovariectomy or orchidectomy (19); inducing metabolic stress through ischemia or other methods (49); or altering environmental factors, such as varying light-dark cycle length (46) or instituting an exercise regime (20). When one is assessing the phenotype, it may also be necessary to study the genetically modified mouse line at different ages to fully ascertain the effects of the genetic modification (52). This is particularly important if the effects are present in early life and are compensated for by other regulatory systems or, alternatively, are of late onset.

It is obvious from this review that a large number of issues regarding phenotypic interpretation exist. However, taking all these factors into consideration when analyzing mouse studies will ensure valid and significant conclusions result. This rigorous approach will increase our understanding of the actions of target genes in endocrinology.

**FUNCTIONAL GENOMICS**

Currently, functional knockout data are available from only ~10% of the estimated 25,000 mouse genes (11). With the exponential increase in data from genomics and proteomics, one of the greatest challenges facing researchers is to analyze the functional significance of this information in a timely and cost-effective manner. The power of using genetically modified mouse models to provide physiologically significant data is clear. However, the time and resources required to generate conventional transgenic or knockout mice as described previously may preclude this as a widespread analytic tool. Fortunately, high-throughput mutagenesis strategies such as gene trapping have now been developed, which combine the power of random mutagenesis with the precision of molecularly defined lesions to generate a large pool of targeted ES cells for generation of mutant mouse strains and functional genomics.

Gene trapping involves the production of random insertional mutations in ES cells (48), using “trap” vectors that allow selection of cells in which coding regions have been targeted for mutagenesis. Conventional gene-trap vectors comprise a promoterless selection/reporter gene such as β-geo (the β-galactosidase and neomycin resistance fusion gene), flanked upstream by a splice acceptor site and downstream by a poly(A) tail (Fig. 5A). Vectors are introduced into ES cells by electroporation or retroviral insertion, where they integrate into the genome randomly. Selection of gene-targeted ES cells occurs because insertion of the gene trap cassette downstream of a promoter interrupts the transcription of the trapped gene and leads to the generation of a fusion transcript expressing β-geo, allowing selection of ES cells through neomycin resistance. The targeted gene locus is determined using primers in the selection gene to perform 5′ rapid amplification of cDNA ends (RACE) and sequencing, identifying the upstream endogenous exon(s). Gene trap ES cell lines are then used to generate knockout mice as described for standard knockouts (GLOBAL KNOCKOUT MOUSE MODELS and Fig. 2). Conventional promotorless gene trap mutations generally lead to a null mutation, although, depending on the site of insertion, fusion proteins may retain partial function. Use of a reporter gene in the conventional gene trap cassette also allows the expression pattern of the targeted gene to be determined in vivo, because expression of the reporter gene is driven by the endogenous promoter. Because conventional gene trap vectors rely on expression of the selectable marker in ES cells, genes that are highly expressed in ES cells are most easily trapped.

The alternate poly(A) vector approach uses a gene trap vector containing a strong promoter and transcription start site upstream of the gene trap cassette with a splice donor sequence downstream (Fig. 5B). Insertion of this gene trap cassette into
a gene causes downstream exons to be trapped through the splice donor sequence and leads to constitutive expression of a fusion protein that confirms neomycin resistance. In more refined versions of the poly(A) trap vectors, the selection gene (e.g., neo) is driven by a strong promoter and has a downstream splice donor site, as in the previous example. When the vector inserts into a gene, expression from the endogenous gene promoter leads to upstream exons being trapped through the SA site. This results in production of a fusion protein consisting of the upstream trapped exons β-gal and neomycin, thereby allowing selection of targeted ES cells. The reporter gene will only be expressed when the endogenous gene promoter is active, which may not occur in ES cells, but still allows the expression pattern of the gene to be determined in vivo. The advantage of this approach is that it does not require the targeted gene to be expressed in ES cells, and therefore all genes should theoretically be trapped equally.

Newer-generation gene trap vectors contain additional refinements, such as having the gene trap cassette flanked by loxP and/or FRT sites, allowing conditional restoration of normal gene expression.

Advantages/Disadvantages

The major advantage of these gene trap approaches is that they offer the opportunity for large-scale functional genomics. The International Gene Trap Consortium (IGTC) is a worldwide collaborative group formed to make gene trap cell lines publicly available to the scientific community on a noncollaborative basis, and for a nominal handling fee (29). The IGTC website (http://www.genetrap.org) provides centralized access to all publicly available annotated gene trap cell lines and provides the ability to search by gene name, chromosomal location or sequence, or biological pathway. At last report, over 49,000 cell lines were available, representing ~9,000 mouse genes, many of which have multiple gene-trapped alleles. This provides a significant resource for researchers worldwide to generate knockout mice rapidly, using well-characterized and annotated gene trap ES cell lines. One potential disadvantage of this approach is that, depending on the site of insertion and gene splicing, partial function may be retained by the fusion protein or reduced levels of wild-type protein expression.

AJP-Endocrinol Metab • VOL 291 • SEPTEMBER 2006 • www.ajpendo.org

Fig. 5. Steps involved in gene-trap mutagenesis. A: conventional trap vector consisting of the promoterless selection/reporter gene β-geo [encoding β-galactosidase (β-gal) and neomycin resistance (neo)], flanked upstream by a splice acceptor site (SA) and downstream by a poly(A) region for mRNA stability (pA), is introduced into ES cells, where it is randomly integrated into the genome. When the vector inserts into a gene, expression from the endogenous gene promoter leads to upstream exons being trapped through the SA site. This results in production of a fusion protein consisting of the upstream trapped exons β-gal and neomycin, thereby allowing selection of targeted ES cells. B: poly(A) vector consists of the promoterless gene encoding β-gal (LacZ) and the neo gene driven by a strong promoter [e.g., the phosphoglycerate kinase (PGK) promoter], flanked by an SA site upstream and a splice donor (SD) site downstream. When the poly(A) vector inserts into a gene in ES cells, downstream exons are trapped through the SD site. Activity of the endogenous gene promoter results in production of fusion protein-1, consisting of the first exons of the trapped gene and β-gal. Activity of the strong PGK promoter results in production of fusion protein-2, consisting of neomycin and the downstream exons of the trapped gene. Expression of fusion protein-2 will occur independently of the target gene expression in ES cells.
transcript may still be generated. However, these hypomorphic alleles may result in subtle phenotypes that can still provide insights into normal gene function, although careful characterization is required.

The development of new generation vectors that offer the ability to create conditional knockout alleles will only enhance the power of the gene trap approach for large scale mutagenesis of the mammalian genome, providing a way forward to allow comprehensive investigation of gene function in vivo.

Conclusion

The approaches described in this review can be used to address a wide range of biological questions, provided that the mouse models are appropriately designed and analyzed. The advantages and disadvantages of current and future approaches have been assessed. With the number of approaches continuing to expand, these provide powerful tools to dissect the function of genes, which will continue to revolutionize our understanding of endocrine physiology and metabolism.

With the vast number of studies using genetically modified mouse models currently in the literature, it is often difficult to fully interpret the significance of the findings reported. Therefore, this review has highlighted the many factors to be taken into account when one is designing, generating, and evaluating these mouse models. Addressing these raised issues will substantially improve the level of rigor within the field of endocrine research, which in turn will lead to a significant advancement in our understanding of endocrine systems.

ACKNOWLEDGMENTS

We thank Howard Morris for scientific contributions and Jeffrey Zajac for helpful discussions.

GRANTS

R. A. Davey is supported by National Health and Medical Research Council (NHMRC; Australia) Grant 350346, H. E. MacLean by NHMRC Grant 350334, and R. D. Wright by Career Development Award 359226.

REFERENCES


