Published in final edited form as:

Methods Enzymol. 2010 ; 477: 243–269. doi:10.1016/S0076-6879(10)77013-0.

Gene Trap Mutagenesis in the Mouse

Roland H. Friedel^{1,2} and Philippe Soriano^{2,*}

¹Department of Neurosurgery, Mount Sinai School of Medicine, 1468 Madison Avenue, New York, NY 10029

²Department of Developmental & Regenerative Biology, Mount Sinai School of Medicine, 1468 Madison Avenue, New York, NY 10029

Abstract

Gene trapping in mouse embryonic stem (ES) cells is an efficient method for the mutagenesis of the mammalian genome. Insertion of a gene trap vector disrupts gene function, reports gene expression, and provides a convenient tag for the identification of the insertion site. The trap vector can be delivered to ES cells by electroporation of a plasmid, by retroviral infection, or by transposon-mediated insertion. Recent developments in trapping technology involve the utilization of site-specific recombination sites, which allow the induced modification of trap alleles in vitro and in vivo. Gene trapping strategies have also been successfully developed to screen for genes that are acting in specific biological pathways. In this chapter, we review different applications of gene trapping, and we provide detailed experimental protocols for gene trapping in ES cells by retroviral and transposon gene trap vectors.

1. Introduction

The generation of mutations in the mouse is a prerequisite for understanding processes as diverse as development, disease, and cancer, and it is therefore the subject of concerted efforts worldwide to disrupt the function of all mouse genes. Three main strategies have been used to generate mutations: first, gene targeting by homologous recombination is used to produce defined mutations in specific genes; second, point mutations or small deletions can be generated by chemical mutagenesis using agents such as ENU or EMS; and third, random gene trap mutagenesis can be applied genome-wide using plasmid, virus, or transposon DNA vectors. These three approaches create different types of alleles, and they complement each other in the creation of allelic series for all mouse genes. The recent advances in sequencing of mammalian genomes have facilitated the expansion of these approaches to large scale mutagenesis programs, which aim at generating public resources of mouse mutant alleles (for review, see (Gondo, 2008)). It will be this collection of allelic variants, with mutations ranging from complete deletion to subtle modifications, which will eventually allow us to dissect gene function at a detailed level.

In this chapter, we review different approaches of gene trapping in mouse embryonic stem (ES) cells, and we provide detailed experimental protocols for trapping by retroviral and

*corresponding author: philippe.soriano@mssm.edu.

transposon vectors. The last section contains a practical guide for the ordering and handling of gene trap clones from consortia. Readers are also referred to the accompanying chapter (Stanford, 2010) that also discusses gene trap mutagenesis.

2. Gene trapping strategies

The concept of using insertional mutagenesis to disrupt gene function takes its source from the classic work of Barbara McClintock with transposons in maize. In the mouse, endogenous or exogenous proviruses as well as microinjected DNA can disrupt gene function by insertional mutagenesis (Copeland et al., 1983; Schnieke et al., 1983; Wagner et al., 1983). It was shown that transgene expression can be subject to gene regulatory elements in the vicinity of the insertion site (Jaenisch et al., 1981), a finding subsequently extended to transgenes encoding an easily followed reporter gene, such as lacZ (Korn et al., 1992; Kothary et al., 1988).

Based on these observations, the first generation of trapping vectors contained a minimal internal promoter that drives gene expression if the trap insertion has occurred in the vicinity of an endogenous enhancer element, hence the term “enhancer trap” (Fig. 1A). The enhancer trapping strategy is mainly attractive for screens that focus on generating lines with specific expression patterns (Gossler et al., 1989; Korn et al., 1992). However, enhancer trap insertions do not always create functional mutations, and this strategy has not been exploited on a wider scale.

With the advent of methods for culturing germline competent ES cells, the concept of gene trap mutagenesis was expanded to a variety of approaches (see below; Friedrich and Soriano, 1991; Gossler et al., 1989; von Melchner et al., 1992). The insertion of a gene trap vector in ES cells simultaneously disrupts gene function, reports expression pattern of the trapped gene, and provides a convenient tag for the molecular identification of the insertion site. Gene trapping has since been proven to be a highly efficient method to generate large number of mutations, and it has been successfully applied by several large-scale projects to generate libraries of trapped ES cell clones (Table 1).

In the first years of gene trapping, one of the salient features was its effectiveness for the discovery of novel genes. This aspect of trapping seems less important now that the mouse genome has been sequenced. However, it has been recognized that the mammalian genome may also contain large numbers of yet unidentified non-protein coding genes, and thus gene trapping may be a very valuable approach for the discovery of these novel genes (Roma et al., 2007).

2.1 Promoter and polyA trapping strategies

Two main strategies were developed for the mutagenic trapping of mouse genes: “promoter trapping” and “polyA trapping”. The currently more widely applied strategy is promoter trapping, in which the trap vector contains a 5′ splice acceptor that splices to the upstream exon of the trapped gene (Friedrich and Soriano, 1991; Gossler et al., 1989; Skarnes et al., 1992; Tate et al., 1998). In this way, the endogenous promoter of the trapped gene is exploited to drive the expression of the reporter gene (Fig. 1B).

A promoter trap design that has been proven efficient to enrich for insertions into genes is the translational fusion of the reporter β -galactosidase to the selection gene neomycin phosphotransferase (neo; the fusion is termed β geo; Friedrich and Soriano, 1991). Since this design involves expression of the selection gene from an endogenous promoter, insertions into transcriptionally silent regions are eliminated during selection. However, it should be noted that promoter trapping can also be performed with a vector containing a separate selection marker that is expressed by an exogenous promoter (e.g., β actin-neo-polyA or PGK-neo-polyA; (Friedrich and Soriano, 1991; Gossler et al., 1989)). This alternative strategy is useful for screens designed to achieve a more unbiased distribution of insertions throughout the genome. About 90–95% of insertions of a promoter trap with a PGK-neo selection gene are estimated to be in loci that are not expressed in ES cells (Friedrich and Soriano, 1991). A recent large-scale gene trap effort with a PGK-neo-polyA selection gene has demonstrated the feasibility of this approach by generating a trap library that has a mutational coverage of more than 90% of all mouse genes (Gragerov et al., 2007). As this strategy does not enrich for gene insertions, traps were not characterized by insertion site cloning of isolated clones, instead, traps in genes of interest can be identified by PCR screening of ES cell pools, and individual ES cell clones are subsequently obtained by successive dilutions of the pools (Gragerov et al., 2007).

The reporter of a promoter trap can be translated independently from the trapped gene if it includes its own ATG translation start codon. This is a useful property for the detection of cells that express a trapped gene, since the reporter protein can diffuse throughout the cell body. An alternative strategy involves a reporter without ATG, which results in chimeric fusion proteins containing N-terminal parts of the endogenous protein and the reporter protein. The fusion proteins can be directed to a variety of different subcellular compartments, and this feature can be exploited to screen for genes encoding compartment-specific proteins, as has been successfully shown in a promoter trap screen for nuclear proteins (Tate et al., 1998).

A promoter trap approach that is specifically designed to enrich for genes encoding secreted and transmembrane proteins is the “secretory trap” (Fig. 1C; (De-Zolt et al., 2006; Leighton et al., 2001; Mitchell et al., 2001; Skarnes et al., 1995)). This class of genes is underrepresented in standard promoter traps, because such trap fusions with secretory proteins contain a signal sequence and are secreted from cells. A secretory trap vector encodes a splice acceptor that is followed by a transmembrane domain and reporter/selection gene(s). This design enriches for secretory genes, since only trap vector insertion downstream of signal sequence (or a type II transmembrane domain) create a topology that allows for efficient drug selection: the vector-borne transmembrane domain anchors the fusion protein in the membrane and places the reporter/selection marker in the cytosol, where it has normal enzymatic activity. In contrast, trap insertions in non-secretory genes are mostly non-productive, since the absence of a signal sequence leads to an insertion of the transmembrane domain in type II orientation, which places the reporter/selection element in the extracellular space or the ER/lysosomal lumen.

A further variant of promoter traps is the “exon trap”, which lacks a splice acceptor and needs to insert directly into an exon to form a functional trap (Fig. 1D; (Hicks et al., 1997;

von Melchner et al., 1992)). The mutational target of exon trap vectors is much less than that of promoter traps, since exons comprise a much smaller fraction of the genome than introns. For this reason, the exon trap vector is less efficient in high-throughput screens, and it has not been used as widely. However, only this type of vector can trap genes that do not possess introns.

As promoter traps that utilize the splice acceptor- β geo (SA- β geo) design depend on expression by the endogenous promoter to build up drug resistance, some genes are not accessible to this strategy simply because they are not expressed at high enough levels in mouse ES cells. Based on the results of large scale promoter trap screens with SA- β geo vectors, it is currently estimated that about 60% of mouse genes can be mutated with this vector type in ES cells (Skarnes et al., 2004). A design that aims at increasing the pool of genes accessible to promoter traps is to incorporate binding sites for transcription factors that are specifically active in mouse ES cells. Promising results have been obtained with a promoter trap vector that contains binding sites for the transcription factor Oct4, which is thought to enhance expression of trapped genes by recruiting Oct4 to the trapped locus (Schnütgen et al., 2008). The binding sites can be removed at a later stage from the trapped allele by excision through the FLP recombinase (Schnütgen et al., 2008).

The “polyA trap” strategy was devised for the mutation of genes independent of their expression level in mouse ES cells. The method relies on using a vector with a strong internal promoter that drives expression of a resistance marker lacking its own polyA signal, followed by a splice donor (Fig. 1E; (Niwa et al., 1993; Salminen et al., 1998; Yoshida et al., 1995; Zambrowicz et al., 1998)). The splice donor design serves to enrich for insertions into genes, since polyA trap insertions outside of genes cannot form polyadenylated transcripts, leading to transcript degradation and elimination of ES cells during drug selection. Gene insertions of the polyA trap vector form a stable transcript by splicing the resistance gene to an endogenous polyA signal. A 5′ splice acceptor with reporter gene is usually included in polyA trap vectors to monitor the expression of the trapped gene.

Although the polyA trap design should allow in theory to trap all genes regardless of expression levels, the results of polyA trapping screens have often been below expectations, since insertions have been mainly found in 3′ introns (Osipovich et al., 2005; Zambrowicz et al., 2003). One reason for this shortcoming could be the mechanism of nonsense-mediated decay (NMD), which degrades transcripts that contain a premature stop-codon upstream of a splice junction (Shigeoka et al., 2005). However, the NMD mechanism may not affect every type of polyA trap vector. In a gene trap screen for growth factor-induced genes, a polyA trap component of the vector was used to generate 3′ RACE products for more than 2,000 clones, indicating stable transcript expression from this vector (Chen et al., 2004a). A recent study also suggests that an optimized combination of promoter and splice donor elements may overcome the problems of polyA trapping (Tsakiridis et al., 2009). The development of new polyA trap vectors is an active field of research, and these attempts include the insertion of mRNA destabilizing elements after the splice donor to more efficiently eliminate unspliced transcripts (Ishida and Leder, 1999; Tsakiridis et al., 2009), or of an IRES sequence to protect transcripts against NMD (Shigeoka et al., 2005). However, these strategies have yet to be validated in a large-scale gene trap setting.

2.2 Trap modifications with site-specific recombinases

The Cre, Flp, and ϕ C31 recombinases mediate site specific recombination at lox, FRT, and attB/P recognition sites, respectively, and allow the directed modification of the initial trap allele (see (Schnütgen et al., 2006) for review). Codon optimized versions of these enzymes have been engineered for efficient translation in mammalian cells (Buchholz et al., 1998; Raymond and Soriano, 2007). Recently, the Dre recombinase, which recognizes rox sites, has also become available for mammalian genome engineering (Anastassiadis et al., 2009).

A trap vector can be flanked with loxP recombination sites in tandem orientation, which enables Cre mediated excision of the trap vector in rescue experiments (Ishida and Leder, 1999). A reversion of a phenotype after trap vector excision indicates that the mutation was specifically caused by the insertion of the trap vector. An important application of recombinases in gene trapping is the conditional trap approach, which allows the time and tissue specific ablation of gene function (Schnütgen et al., 2005; Xin et al., 2005). The conditional designs are based on trap cassettes that are flanked by recombination sites in opposing orientations. This type of site arrangement causes inversion of the cassette upon recombination, rather than excision.

The “Flex” system utilizes two arrays of loxP/lox5171 and FRT/F3 sites that flank the trap cassette in opposing orientation (Fig. 1F; (Schnütgen et al., 2005)). The initial trap allele of a Flex vector is a mutagenic promoter trap insertion. By applying Flp recombinase, e.g. by breeding to a Flp germline deleter mouse, the trap cassette can be inverted -thereby creating a neutral conditional allele, since the splice acceptor has been removed from the direction of transcription. The conditional trap can be selectively reactivated in a time- and tissue-specific manner by applying Cre recombinase, which re-inverts the trap cassette back to its original mutagenic orientation.

The “double switch” trap vector consists of an upstream 5′ splice acceptor-GFP reporter component, which is in inverse orientation to gene transcription and flanked by half mutant lox66 and lox71 sites in opposing orientation, and a downstream 3′ polyA trap component flanked by FRT sites (Fig. 1G; (Xin et al., 2005)). The polyA trap component can be excised by Flp recombinase, which creates a conditional allele. This allele can be rendered mutagenic in a second step by Cre-mediated inversion of the splice acceptor-GFP component.

The conditional version of a trap insertion may in some instances not be completely innocuous. For instance, it is possible that the insertion interferes with a cis-acting regulatory element, or the inverted cassette may disrupt another gene in antisense orientation to the initially trapped gene. These scenarios should be ascertained by phenotypic comparison of conditional trap mice with wildtype and null mutant mice.

Another useful application of site-specific recombination sites is the exchange of cassettes in existing trap lines (“recombinase mediated cassette exchange”, RMCE). In this way, the initial trap allele can be modified in any conceivable way, e.g. by inserting cDNAs encoding a fluorescent protein reporter, a Cre recombinase, a mutant version of the trapped gene, or a shRNA construct. The RMCE design is based on a cassette flanked by tandem

recombination sites, which mediate the exchange of the residing cassette with an incoming vector. Vectors suitable for RMCE have been reported with flanking loxP sites (Hardouin and Nagy, 2000), half-mutant lox66/71 sites (Araki et al., 1999; Singla et al., 2010; Taniwaki et al., 2005), or heterotypic lox sites (Osipovich et al., 2005). Similar RMCE designs are also feasible with vectors that harbor heterotypic FRT sites or attB/P recombination sites (Baer and Bode, 2001).

2.3 Induction trapping and phenotypic screens in ES cells

A mutagenesis project typically aims at identifying genes that are involved in a specific biological process. Induction trapping is a strategy for the enrichment of trap events for genes that are involved in such processes.

For example, trap clones can be subjected to treatment with the differentiation factor retinoic acid (RA) *in vitro*, and the readout of the lacZ reporter can be used to select for genes that are sensitive to regulation by RA (Forrester et al., 1996). Conceptually similar approaches have involved induction of clones with growth factors (Bonaldo et al., 1998), a secreted Engrailed-2 homeodomain (Mainguy et al., 2000), or the differentiation of trapped clones to endothelial cells (Hirashima et al., 2004). ES cells can also be differentiated via embryoid bodies to different cell lineages of all three germ layers (Stuhlmann, 2003). Bradley and colleagues devised an induction trapping strategy that allows trapping of genes only transiently expressed upon induction (Chen et al., 2004b). This strategy utilizes a Cre recombinase as intermediate readout to activate a permanent resistance marker from a responder allele. The use of a reporter allowing both positive and negative selection coupled with flow sorting has also been used successfully to identify hepatocyte growth factor (HGF)-responsive genes (Medico et al., 2001).

A potential limitation of applying induction trapping strategies in ES cells may be that these cells are not appropriate responders for most exogenous stimuli. It may be biologically more appropriate to induce target genes in specialized cell lines or primary tissue, and to identify genes regulated by a specific stimulus by DNA microarray hybridization or massive parallel sequencing. Mouse mutants can then be conveniently generated by resorting to available gene trap resources. In two studies that performed gene trap screens in combination with DNA arrays, trap vectors contained promoter and splice donor to generate a trap transcript that can be used for the amplification of 3' RACE fragments. DNA arrays were generated from the 3' RACE fragments, thus representing the collection of gene trap clones. Arrays were then hybridized with cDNA probes that were specific for certain organs (Matsuda et al., 2004), or with probes derived from cells that were stimulated with PDGF (Chen et al., 2004a).

Another approach to direct gene trapping towards genes in a specific pathway is to perform a phenotypic screen in ES cells. However, most gene trap insertions will cause heterozygous mutations (except for insertions in sex chromosomes), which will rarely generate detectable phenotypes. A strategy to obtain homozygous mutations is the application of ES cells that show high levels of somatic recombination due to a deficiency in the DNA helicase gene Bloom, which is mutated in human Bloom syndrome (Guo et al., 2004). Somatic recombination is a process of crossing-over of sister chromosomes during mitosis, which

leads to daughter cells that are homozygote for genes distal to the cross-over site. Bloom^{-/-} ES cells with a heterozygous trap insertion will contain cells that are homozygous for the trapped allele after about 13 rounds of cell division. Phenotypic screening in Bloom^{-/-} ES cells has been applied to identify genes in the DNA mismatch repair machinery (Guo et al., 2004; Wang et al., 2009), host factors for retroviral infection (Wang and Bradley, 2007), and genes of the RNAi machinery (Trombly et al., 2009). Another strategy for homozygosing trap mutations in ES cells is the use of a gene trap vector with a hypomorphic version of the neo resistance gene (Lin et al., 2006). By simply increasing the drug concentration during neo selection, cells are selected that have doubled the gene dosage of the resistance marker by spontaneous loss-of-heterozygosity at the trapped locus.

3. Design of the gene trap vector

3.1 Splicing elements

A critical point in design of the gene trap cassette is the choice of splice acceptors and donors. For 5' splicing in promoter trapping, we recommend the adenoviral major late gene exon 2 splice acceptor (Friedrich and Soriano, 1991). We have not observed bypassing of this splice acceptor, which could potentially result in non-mutagenic or hypomorphic trap insertions, in any of twenty-four mutant mouse lines that we have analyzed at a molecular level. The adenoviral major late gene exon 2 splice donor is suitable as 3' splice donor for polyA trapping approaches, as this splice donor was successfully used for the cloning of 3' RACE transcripts from the ROSAFARY vector (Chen et al., 2004a). The splicing elements discussed above can be obtained from parts of pROSAFARY plasmid from the Addgene plasmid library (www.addgene.org).

3.2 Reporter and selection genes

The most commonly used reporter is the β -galactosidase enzyme (β gal), which can be used to visualize gene expression by a color reaction of fixed embryos or tissue sections. For promoter trapping, the β gal enzyme is often used as a direct fusion to the neomycin phosphotransferase gene (β geo; (Friedrich and Soriano, 1991)). A marker that allows labeling of cell membranes (e.g. axon tracts) is the placental alkaline phosphatase (PLAP; (Leighton et al., 2001)). For live imaging of gene expression, fluorescent proteins previously used in gene trap vectors include green fluorescent protein (GFP) (Ishida and Leder, 1999), Venus (Tanaka et al., 2008), and fusions of GFP to neo (Chen and Chen, 2004) or β -geo (Tsakiridis et al., 2007).

An essential component of a gene trap vector is the resistance marker, typically a neomycin phosphotransferase (neo) gene, which allows for selection of clones that harbor the gene trap insertion. It should be noted that some vectors contain a hypomorphic variant of the neo that is less sensitive to drug selection (Skarnes et al., 1995). Alternative resistance markers for traps in ES cells are puromycin N-acetyl transferase and hygromycin phosphotransferase.

A strategy that allows co-expression of multiple reporters (or resistance genes) under the control of the same promoter is to place an internal ribosomal entry site (IRES) between the open reading frames. Drawbacks of this strategy are that the gene following the IRES

sequence is often translated at reduced levels and that the efficiency of translation may vary from cell type to cell type (Merrick, 2004). Another approach to ensure equal amounts of translation products is the use of viral 2A self cleaving peptides (Szymczak et al., 2004). Genes that have been connected by 2A linker sequences are transcribed as one continuous open reading frame, and during translation one peptide bond in the 2A peptide is omitted by a ribosomal skip mechanism. The 2A peptides are 18 aa long, and leave a 17 aa sequence at the C-terminus of the upstream protein, and one proline residue at the N-terminus of the downstream protein. This strategy is therefore only effective as long as the final protein products remain functional with these additional amino acids.

3.3 Plasmid, retroviral, and transposon vectors

Gene trap vectors can be delivered into the ES cell genome as plasmids, retroviruses, or transposons. Each of these methods has unique advantages and limitations. Plasmids are easy to prepare, have no significant size limitations, and can be introduced into ES cells simply by electroporation. The main drawback of plasmid traps is that undefined events can occur at the insertion site, such as deletions, and concatemeric insertions. Such complex insertions complicate the analysis of the mutation, and they hamper the efficient identification of genomic integration sites by linker-mediated PCR.

Retroviruses require little experimental work, since they are harvested as supernatants from a producer cell line, before application to ES cells. Moreover they offer the advantage of clean, single insertions, which facilitate the high-throughput cloning of the genomic junctions. Retroviruses preferentially insert into the 5' region of genes, which is also advantageous for disrupting gene function with promoter traps.

Gene trapping with transposons can be performed with class II transposons, which move in the genome by a cut-and-paste mechanism. Transposon traps can be easily introduced into ES cells by co-electroporation of a transposon trap plasmid together with a plasmid encoding the appropriate transposase. The main advantage of transposon traps, however, resides in their ability to transpose in somatic tissues at later stages, and they have thus been used very successfully for the identification of genes implicated in cancer (Collier et al., 2005; Dupuy et al., 2005); see also chapters by Bradley, Dupuy, and Takeda in this volume). Two transposon systems are currently used for mutagenesis in the mouse: Sleeping Beauty, which has been synthetically reconstructed from salmonid fish (Ivics et al., 1997), and PiggyBac, which has been originally identified as transposon in the insect cabbage looper (Cary et al., 1989). Transposon insertions occur at specific target sites, which are characterized by the nucleotide sequences "TA" for Sleeping Beauty, and "TTAA" for PiggyBac. The PiggyBac transposon system appears to transpose with higher activity, and it is therefore currently the preferred system for transposon mutagenesis in ES cells (Wang et al., 2009; Wang et al., 2008). Moreover, the lower transposition rate of the Sleeping Beauty transposon leads frequently to insertions in the vicinity of the donor site, a characteristic known as "local hopping". Further advantages of the PiggyBac system are large cargo sizes (Ding et al., 2005), and a clean excision (Sleeping Beauty leaves a "footprint" mutation at the excised locus) (Luo et al., 1998; Wang et al., 2008). In contrast to retroviruses however, transposons seem to have less preference for the insertion into the 5' end of genes.

However, a bias for insertions of transposons into transcribed regions of the genome has been observed (Liang et al., 2009).

4. Gene trapping protocol for retroviral vectors

4.1 Generation of retroviral vectors

A retroviral gene trap vector is generated by cloning of a plasmid that contains a gene trap cassette into two flanking 5' and 3' long terminal repeat (LTR) elements of the Moloney Murine Leukemia Virus (MMLV). The total size of the retroviral vector, including LTRs and trap cassette, should not exceed 11kb. The direction of the trap cassette should be in reverse to the 5'-3' orientation of the viral LTRs (Friedrich and Soriano, 1991), and the LTRs should be engineered to lack viral enhancers (Soriano et al., 1991) or enhancers and promoters (Chen et al, 2004a).

The plasmid for the production of the ROSAFARY gene trap virus (Chen et al, 2004a), which incorporates the design criteria described above, can be ordered from the Addgene plasmid library (www.addgene.org). The plasmid for the conditional rsFLP-ROSA β geo gene trap vector (Schnütgen et al., 2005) is available through the EUCOMM program (www.eucomm.org). The latter plasmid contains a "PINCO" backbone with the EBNA1 gene of the Epstein-Barr virus (Grignani et al., 1998), which leads to increased virus titers upon transfection of packaging cell lines.

4.2 Production of virus supernatant

Retrovirus particles for gene trapping can be easily obtained by transfecting a retroviral gene trap plasmid into a packaging cell line that carries stable integrations of the gag, pol, and env genes of the MMLV. We have been routinely using the "GP+E86" (Markowitz et al., 1988; Soriano et al., 1991) and the "Phoenix-Eco" cells line ((Swift et al., 2001); <http://www.stanford.edu/group/nolan>) with stable and transient transfection protocols, respectively.

For the generation of stably transfected GP+E86 cells, ten microgram linearized vector plasmid are electroporated in 1×10^6 cells in 0.8 ml phosphate-buffered saline (PBS). After selecting in presence of the appropriate antibiotic for 10–12 days, resistant clones are pooled, expanded and frozen in batch. To collect infectious viral particles, virus-producing cells are grown to near confluence and overlaid with fresh medium. After 16 hr, virus-containing medium supernatant is filtered through a 0.45 μ m filter to remove cells and debris before being used for infection. The supernatant is either applied directly on ES cells (the half-life of retroviruses at 37°C is approximately 6 hours), or stored at –80°C. However, freezing and thawing of virus supernatants reduces titers by about half. To determine the effective colony generating viral titer, serial 10-fold dilutions of the viral supernatant are prepared to infect log phase 3T3 cells or ES cells in 6-well dishes. Media should contain 4 μ g/ml polybrene (Hexadimethrin bromide, Sigma) to increase infection efficiency by enhancing virus adsorption on target cell membranes. Cells are selected in antibiotic-containing medium until colonies can be counted by staining with 0.5% methylene blue in 50% methanol.

For virus production by transient transfection, Phoenix cells are expanded on 10 cm tissue culture dishes until they reach subconfluency. Ten microgram circular plasmid DNA are transfected per dish by the Calcium-phosphate method or with a commercial lipid-based transfection reagent. On the day following transfection, overlay the transfected cells with 7 ml of ES cell media, incubate cells over night and collect the supernatant on the following day. The transfected cells can be conditioned again overnight for a second harvest of virus supernatant.

4.3 Infection of ES cells and clone picking

Culture ES cells according to the protocols for the particular cell line. Each ES cell line has its own specific requirements for feeders, media, trypsinization, and split ratios (see chapters by Skarnes and Stewart in this volume). Grow ES cells to subconfluency in a 10 cm tissue culture dish and infect with virus supernatant in a total volume of 4 ml media in the presence of 4 µg/ml polybrene (see Fig. 2A for an overview of the workflow). After six hours, change to normal media without virus. On the following day, start selection with media containing 150µg/ml active concentration of G418 (Invitrogen 11811-031). Change selection media every day for 8–10 days. ES clones should be visible after 7 days, and colonies can be usually picked after 10 days of selection.

For picking of clones, wash cells on selection plate once with PBS and cover them again with PBS. Pick clones with a glass capillary or a micropipette in a volume of 10µl by gently scratching them from the plate. Transfer picked clones into a 96-well U-bottom plate containing 30 µl of Trypsin in each well and incubate for 10 min at 37°C. Add 150µl of medium (+G418) to each well and triturate cells thoroughly to generate a suspension of single cells. Transfer cell suspension to a 96-well flat-bottom plate. Change medium (+G418) on 96-well plates every day. After 2–4 days, when the majority of wells have reached subconfluence, split cells to four 96-well plates (copies A and B for freezing cells, copies C and D for DNA analysis). Copies A and B are frozen for archiving when the majority of wells have reached subconfluence. For freezing of cells, wash each well with 100 µl PBS, trypsinize with 25 µl Trypsin for 10 min at 37°C, add 75 µl media and triturate cells, add 100 µl 2x freezing media (20% FCS, 10% DMSO in DMEM) and mix, and cover with 50 µl mineral oil. For long term storage, is it advisable to freeze cells in racks that hold 96 microtubes (e.g., Neptune Minitube system, or Thermo Scientific Matrix 2D Barcoded Storage Tubes). Cultivate copies C and D for 2–3 extra days until they appear overgrown in order to get sufficient amounts of cells for DNA analysis.

5. Gene trapping protocol for transposon vectors

5.1 Generation of transposon plasmids

This protocol is described for PiggyBac vectors, but can be easily adapted for the application of the Sleeping Beauty transposon. For gene trapping with a PiggyBac transposon, generate a plasmid that contains a gene trap cassette that is flanked by the 5' and 3' PiggyBac terminal repeats (TR). For efficient transposition of PiggyBac, the minimal sizes for the TRs have been determined to be 310bp for the 5'TR and 235bp for the 3'TR (these TRs contain also non-repeating internal sequences that enhance transposition) (Li et al., 2005). A plasmid

containing the minimal PiggyBac repeats can be obtained from Dr. Fraser at the University of Notre Dame (piggybac.bio.nd.edu). A PiggyBac transposon can carry up to 9.1 kb of foreign DNA without a significant drop in transposition efficiency (Ding et al., 2005). As transposase, either the original insect transposase (iPBase), or a PiggyBac transposase that has been optimized for mammalian codon usage (mPBBase) can be used (Cadinanos and Bradley, 2007). In our hands, the activity of original iPBase expressed under the synthetic CAG promoter (Wang et al., 2008) was adequate for gene trapping experiments in ES cells.

5.2 Electroporation of ES cells with PiggyBac transposon

The PiggyBac transposon system is introduced into ES cells by co-electroporation of two circular plasmids: a plasmid encoding the transposase, and a plasmid containing the PiggyBac transposon gene trap vector. The transposase enzyme will cut out the transposon and paste it randomly at a TTAA target site in the ES cell genome. Depending on the amounts of DNA that are electroporated, multiple independent insertions of PiggyBac transposons can occur in a single clone. To reduce the frequency of clones with multiple insertions, it is advisable to use low amounts of plasmid DNA for electroporation. We typically electroporate 0.4 µg PiggyBac transposon plasmid and 2 µg iPBase plasmid for 10^7 cells in a total volume of 800 µl PBS. Parameter settings are 250 V, 500 µF, and “exponential decay” mode for a Biorad Gene Pulser with a 4 mm cuvette, which should result in a τ discharge time of about 0.7 ms. Applying these parameters for a PiggyBac promoter-trap vector with drug-selectable marker, we obtain more than 100 clones per electroporation, of which about 95% harbor single insertions (see also (Wang et al., 2009; Wang et al., 2008)).

For one round of electroporation, expand ES cells in one 10 cm culture dish until they reach subconfluency (see Fig. 2B for an overview of the workflow). On the day of electroporation, change media on cells 3 hours before electroporation. Trypsinize and count cells in a Neubauer chamber. Spin down 10^7 cells each and resuspend with PBS to a total volume of 800 µl. Add desired amount of plasmid DNA (dissolved in Tris-EDTA, in a volume of no more than 20 µl), transfer cell/DNA mix into an electroporation cuvette of 4 mm gap width, and electroporate as described above. The electroporated cells are plated on three 10 cm dishes. Antibiotic selection is applied about 20 hr after electroporation by changing media to selective media with 150 µg/ml active G418 (Invitrogen 11811-031). ES clones are cultivated in selective media for the following 7–10 days. Perform picking, expansion, and freezing of clones as described in chapter 4.3.

6. Identification of trap insertion sites by Splinkerette PCR

The insertion sites of gene traps have been previously mainly determined by methods that generate cDNA fragments containing parts of the vector cassette and exons flanking the trap insertion. For promoter traps, intronic locations can be determined by sequencing exons upstream of the trap insertion with the 5' RACE method, and for polyA traps, insertions can be identified by 3' RACE to sequence exons downstream of the insertion. Detailed protocols for 5' and 3' RACE analysis of gene trap clones can be found in (Chen and Soriano, 2003). However, these methods rely on the analysis of mRNA, which can be a limiting factor for genes that are expressed at low levels in ES cells. In addition, the results

give only information about the intron in which the trap vector inserted, but not about the precise genomic location of the insertion, which is a prerequisite for PCR genotyping of gene trap mouse lines from genomic tail DNA.

With the advent of the complete sequence of the mouse genome, it became feasible to identify gene trap insertion loci by analysis of the flanking genomic sequences at the insertion site. Different methods for the cloning of junction fragments that contain vector and flanking genomic sequences are available, such as plasmid rescue, inverse PCR, and linker-mediated PCR. A method that has proven to be particularly efficient for this purpose is the “splinkerette”, an optimized protocol of linker-mediated PCR (Devon et al., 1995). It has been successfully applied for the high throughput identification of gene trap insertions (Horn et al., 2007), and it can be applied for both retroviral or transposon insertions. The main improvement of splinkerette in respect to other linker-mediated strategies is the introduction of a hairpin loop in the linker oligos, which suppresses amplification of unspecific side products (Fig. 3).

Splinkerette is performed in two separate reactions for the 5' and 3' junctions of the vector. We provide here a protocol for the use of the restriction enzyme *Bst*YI, which recognizes a R'GATCY site, and was successfully used for a large scale annotation of a gene trap library (Horn et al., 2007). The reactions can also be repeated with other restriction enzymes, such as *Sau*3AI, which may increase the recovery rate of junction fragments. The following protocol serves for the cloning of transposon insertion sites. The protocol can also be applied for the cloning of retroviral insertion sites, only the vector specific primers need to be adjusted (see (Horn et al., 2007; Uren et al., 2009)).

For DNA preparation, ES cells are grown in 96 well plates until they reach confluency. Plates are washed once with PBS and can then be stored dry at -20°C until further processing. Incubate each well with 50 µl lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM EDTA, 10 mM NaCl, 0.5% (w/v) sodium dodecyl sulfate (SDS); add proteinase K to 1mg/mL fresh before use) for 2 hours at 60°C. Add 150 µl ice cold Ethanol-Salt Solution (150 µl 5M NaCl to 10 ml of Ethanol) per well, let sit at room temperature for one hour. The precipitated DNA threads stick to the bottom of the well. Wash wells three times with 150 µl 70% EtOH. Drain plates by slowly inverting on paper towels. Air dry precipitated DNA and re-dissolve with 40 µl of 5 mM Tris, pH8.0 overnight at room temperature.

For digest, mix 9 µl of the genomic DNA with 5 units of *Bst*YI in a total volume of 20 µl with buffers supplied by the manufacturer. Incubate reactions in a 96-well plate in a PCR thermal cycler at 60°C for one hour. Stop digest by heat inactivation at 80°C for 20 min. For alternative restriction enzymes, adjust conditions accordingly.

Prepare the double-stranded splinkerette adapter by combining 400 pmol each of the splinkerette primers HMSpAA and HMSpBB (see Table 2) in a total volume of 40 µl of 5 mM Tris-HCl, pH 7.5, 25 mM NaCl, 0.5mM DTT. Denature for 3 min at 97°C and allow strands to anneal by slowly cooling to room temperature. Set up ligation in a final volume of 30 µl using 20 µl of heat inactivated restriction digest, 1 µl of the annealed splinkerette

adapter mix and 400 units T4-DNA ligase (New England Biolabs) in the provided ligation buffer. Incubate ligation overnight at 16°C.

Purify ligation products using a PCR purification kit (QIAGEN) and elute the DNA into 45 µl of 5mM Tris-HCl, pH 8.0. Processing of 96 well plates can be performed with multiwell purification plates. Set up PCR reactions in 0.2-ml thin-walled PCR tubes in a volume of 15 µl containing 5 µl of the adapter ligated DNA, 1x ThermoPol PCR buffer (NEB), 0.2 mM dNTPs, 0.2 mM of each primer (HMSp1 and PB5'-1 for 5' junction, and HMSp1 and PB3'-1 for 3' junction; Table 2), and 0.4 unit Taq Polymerase (NEB). The primary PCR is performed with the following cycling conditions: 2 min 30 sec at 95°C for initial denaturation, then 30 cycles of 15s at 95°C, 30s at 65°C, and 2 min at 72°C, and a final elongation step at 72°C for 7 min. Use 1 µl of the primary amplification reaction without further purification in a nested PCR with the primers HMSp2 and PB5'-2, and HMSp2 and PB3'-2, respectively (Table 2), using the same PCR conditions as for the primary PCR. Use 1 µl of the final amplification product for direct sequencing with primers PB5'-seq or PB3'-seq (Table 2).

7. Ordering and handling of gene trap clones from consortia

Collections of gene trap clones have been generated by several laboratories worldwide (see Table 1 for weblinks to gene trap consortia and databases). The annotation data of these clones has been combined in a central database, accessible at the website of the International Gene Trap Consortium (IGTC, www.igtc.org). Other useful databases for accessing gene trap clones can be found at the UniTrap website (unitrap.cbm.fvg.it), and at the Mouse Genome Informatics website (www.informatics.jax.org), which lists available gene trap clones on its gene information pages. To find gene trap clones for your gene of interest, enter the gene name in the search query, or blast your sequence against the gene trap sequence tag database. The search result will provide an annotation for each clone, including ordering information, the specific ES cell line that was used, and the insertion site of the clone. The insertion site will be mapped to an intronic location, if 5'/3' RACE was used to identify the clone, or to the exact genomic coordinates, if the clone was identified by genomic PCR. The databases typically list for each gene multiple ES cell clones representing traps in different introns of the gene. For most phenotypic studies, trap clones with insertions in the 5' region of the gene should be chosen, since these are more likely to be functionally mutagenic. The germline transmission capacity of gene trap ES cell clones is in practice about 50–70%. That is, not every gene trap clone will result in germline transmission, and it may be necessary in some cases to inject an alternative gene trap clone.

Gene trap ES cell clones are shipped in cryo-vials on dry ice, and should be placed immediately in liquid nitrogen storage upon receipt. The sending institution will provide protocols for the handling of the ES cell line and its specific media requirements. Before injection of a gene trap clone into blastocysts for the generation of chimeric mice, the identity of the obtained cell clone should be re-confirmed by RT-PCR. For this purpose, split a fraction of the gene trap cell line during expansion onto one 35mm dish, let cells grow to confluency, and use cells to prepare RNA and subsequently oligo-dT primed cDNA. Perform the diagnostic RT-PCR reaction with a primer is located in an exon upstream of the

annotated trap insertion (or downstream, in the case of poly-A traps), and a matching reverse primer that is specific for the gene trap vector cassette.

In few sporadic cases, partial by-passing of the gene trap splice acceptor has been observed, resulting in hypomorphic or non-mutagenic alleles. The full mutagenicity of a gene trap allele should therefore be verified in homozygous mutant animals. The amount of wild type transcript or protein should be assessed in homozygous mutant and wild type littermates by Northern or Western blot analysis, respectively.

Genotyping of gene trap mouse lines can be easily accomplished by using genomic PCR with a crude DNA preparation of the tail tip. We recommend a three-primer PCR strategy, with one primer specific for the vector, and two primers flanking the insertion site (see (Chen and Soriano, 2003) for details).

The β -galactosidase (β gal) enzyme of *E. coli* is frequently included as a reporter in gene trapping vectors to monitor the expression of the trapped gene. Detailed protocols for the application of this reporter in wholemount embryo and tissue section staining can be found in (Chen and Soriano, 2003). For gene trap lines with an appropriate expression pattern of the trapped gene, it is also possible to take advantage of the β gal reporter for genotyping of mice by X-gal staining of the tail tip.

8. Outlook

Gene trapping has been one of the most successful strategies for the mutagenesis of the mammalian genome, and it is a cornerstone in the world-wide efforts to generate a public library of ES cells with mutations in all mouse genes. Although several consortia have shifted their efforts now on gene targeting, which allows the tailored modification of all genes, gene trapping remains by far the most efficient mutagenesis strategy for the generation of large collections of mutant alleles. Powerful new applications of gene trapping are also emerging in methodologies that aim at modifying gene trap alleles *in vivo*. The most prominent examples are conditional gene traps that can be activated by recombinases in a time and tissue-specific manner, and the utilization of gene trapping together with somatic transposon mutagenesis *in vivo*, which has been shown to be a very effective tool for identifying cancer-causing genes. These methods are likely to shed insight into physiological processes well beyond embryonic development, where gene trap mutagenesis has been mainly applied so far. Much progress has been made in the development of gene trap technology over the past decade, and it will be exciting to see its future applications.

Acknowledgments

We thank our colleagues for critical reading of the manuscript. Work in the authors laboratory is supported by grants from the National Institute of Child Health and Human Development to P.S.

References

- Anastassiadis K, Fu J, Patsch C, Hu S, Weidlich S, Duerschke K, Buchholz F, Edenhofer F, Stewart AF. Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice. *Dis Model Mech*. 2009; 2:508–15. [PubMed: 19692579]

- Araki K, Imaizumi T, Sekimoto T, Yoshinobu K, Yoshimuta J, Akizuki M, Miura K, Araki M, Yamamura K. Exchangeable gene trap using the Cre/mutated lox system. *Cell Mol Biol (Noisy-le-grand)*. 1999; 45:737–50. [PubMed: 10512203]
- Baer A, Bode J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr Opin Biotechnol*. 2001; 12:473–80. [PubMed: 11604323]
- Bonaldo P, Chowdhury K, Stoykova A, Torres M, Gruss P. Efficient gene trap screening for novel developmental genes using IRES beta geo vector and in vitro preselection. *Exp Cell Res*. 1998; 244:125–36. [PubMed: 9770356]
- Buchholz F, Angrand PO, Stewart AF. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotechnol*. 1998; 16:657–62. [PubMed: 9661200]
- Bult CJ, Kadin JA, Richardson JE, Blake JA, Eppig JT. The Mouse Genome Database: enhancements and updates. *Nucleic Acids Res*. 2009
- Cadinanos J, Bradley A. Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res*. 2007; 35:e87. [PubMed: 17576687]
- Cary LC, Goebel M, Corsaro BG, Wang HG, Rosen E, Fraser MJ. Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology*. 1989; 172:156–69. [PubMed: 2549707]
- Chen WV, Chen Z. Differentiation trapping screen in live culture for genes expressed in cardiovascular lineages. *Dev Dyn*. 2004; 229:319–27. [PubMed: 14745956]
- Chen WV, Delrow J, Corrin PD, Frazier JP, Soriano P. Identification and validation of PDGF transcriptional targets by microarray-coupled gene-trap mutagenesis. *Nat Genet*. 2004a; 36:304–12. [PubMed: 14981515]
- Chen WV, Soriano P. Gene trap mutagenesis in embryonic stem cells. *Methods Enzymol*. 2003; 365:367–86. [PubMed: 14696359]
- Chen YT, Liu P, Bradley A. Inducible gene trapping with drug-selectable markers and Cre/loxP to identify developmentally regulated genes. *Mol Cell Biol*. 2004b; 24:9930–41. [PubMed: 15509795]
- Collier LS, Carlson CM, Ravimohan S, Dupuy AJ, Largaespada DA. Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature*. 2005; 436:272–6. [PubMed: 16015333]
- Collins FS, Rossant J, Wurst W. A mouse for all reasons. *Cell*. 2007; 128:9–13. [PubMed: 17218247]
- Copeland NG, Hutchison KW, Jenkins NA. Excision of the DBA ecotropic provirus in dilute coat-color revertants of mice occurs by homologous recombination involving the viral LTRs. *Cell*. 1983; 33:379–87. [PubMed: 6305507]
- De-Zolt S, Schnutgen F, Seisenberger C, Hansen J, Hollatz M, Floss T, Ruiz P, Wurst W, von Melchner H. High-throughput trapping of secretory pathway genes in mouse embryonic stem cells. *Nucleic Acids Res*. 2006; 34:e25. [PubMed: 16478711]
- Devon RS, Porteous DJ, Brookes AJ. Splinkerettes—improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res*. 1995; 23:1644–5. [PubMed: 7784225]
- Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell*. 2005; 122:473–83. [PubMed: 16096065]
- Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature*. 2005; 436:221–6. [PubMed: 16015321]
- Forrester LM, Nagy A, Sam M, Watt A, Stevenson L, Bernstein A, Joyner AL, Wurst W. An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid in vitro. *Proc Natl Acad Sci U S A*. 1996; 93:1677–82. [PubMed: 8643689]
- Friedel RH, Seisenberger C, Kaloff C, Wurst W. EUCOMM—the European conditional mouse mutagenesis program. *Brief Funct Genomic Proteomic*. 2007; 6:180–5. [PubMed: 17967808]
- Friedrich G, Soriano P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev*. 1991; 5:1513–23. [PubMed: 1653172]
- Gondo Y. Trends in large-scale mouse mutagenesis: from genetics to functional genomics. *Nat Rev Genet*. 2008; 9:803–10. [PubMed: 18781157]

- Gossler A, Joyner AL, Rossant J, Skarnes WC. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science*. 1989; 244:463–5. [PubMed: 2497519]
- Gragerov A, Horie K, Pavlova M, Madisen L, Zeng H, Gragerova G, Rhode A, Dolka I, Roth P, Ebbert A, Moe S, Navas C, et al. Large-scale, saturating insertional mutagenesis of the mouse genome. *Proc Natl Acad Sci U S A*. 2007; 104:14406–11. [PubMed: 17720809]
- Grignani F, Kinsella T, Mencarelli A, Valtieri M, Riganelli D, Grignani F, Lanfranccone L, Peschle C, Nolan GP, Pelicci PG. High-efficiency gene transfer and selection of human hematopoietic progenitor cells with a hybrid EBV/retroviral vector expressing the green fluorescence protein. *Cancer Res*. 1998; 58:14–9. [PubMed: 9426049]
- Guo G, Wang W, Bradley A. Mismatch repair genes identified using genetic screens in Blm-deficient embryonic stem cells. *Nature*. 2004; 429:891–5. [PubMed: 15215866]
- Hansen GM, Markesich DC, Burnett MB, Zhu Q, Dionne KM, Richter LJ, Finnell RH, Sands AT, Zambrowicz BP, Abuin A. Large-scale gene trapping in C57BL/6N mouse embryonic stem cells. *Genome Res*. 2008; 18:1670–9. [PubMed: 18799693]
- Hansen J, Floss T, Van Sloun P, Fuchtbauer EM, Vauti F, Arnold HH, Schnutgen F, Wurst W, von Melchner H, Ruiz P. A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *Proc Natl Acad Sci U S A*. 2003; 100:9918–22. [PubMed: 12904583]
- Hardouin N, Nagy A. Gene-trap-based target site for cre-mediated transgenic insertion. *Genesis*. 2000; 26:245–52. [PubMed: 10748462]
- Hicks GG, Shi EG, Li XM, Li CH, Pawlak M, Ruley HE. Functional genomics in mice by tagged sequence mutagenesis. *Nat Genet*. 1997; 16:338–44. [PubMed: 9241269]
- Hirashima M, Bernstein A, Stanford WL, Rossant J. Gene-trap expression screening to identify endothelial-specific genes. *Blood*. 2004; 104:711–8. [PubMed: 15090446]
- Horn C, Hansen J, Schnutgen F, Seisenberger C, Floss T, Irgang M, De-Zolt S, Wurst W, von Melchner H, Noppinger PR. Splinkerette PCR for more efficient characterization of gene trap events. *Nat Genet*. 2007; 39:933–4. [PubMed: 17660805]
- Ishida Y, Leder P. RET: a poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells. *Nucleic Acids Res*. 1999; 27:e35. [PubMed: 10572187]
- Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*. 1997; 91:501–10. [PubMed: 9390559]
- Jaenisch R, Jahner D, Nobis P, Simon I, Lohler J, Harbers K, Grotkopp D. Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. *Cell*. 1981; 24:519–29. [PubMed: 7237558]
- Korn R, Schoor M, Neuhaus H, Henseling U, Soininen R, Zachgo J, Gossler A. Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimaeric embryos with a high frequency and detect endogenous genes. *Mech Dev*. 1992; 39:95–109. [PubMed: 1283314]
- Kothary R, Clapoff S, Brown A, Campbell R, Peterson A, Rossant J. A transgene containing lacZ inserted into the dystonia locus is expressed in neural tube. *Nature*. 1988; 335:435–7. [PubMed: 3138544]
- Leighton PA, Mitchell KJ, Goodrich LV, Lu X, Pinson K, Scherz P, Skarnes WC, Tessier-Lavigne M. Defining brain wiring patterns and mechanisms through gene trapping in mice. *Nature*. 2001; 410:174–9. [PubMed: 11242070]
- Li X, Harrell RA, Handler AM, Beam T, Hennessy K, Fraser MJ Jr. piggyBac internal sequences are necessary for efficient transformation of target genomes. *Insect Mol Biol*. 2005; 14:17–30. [PubMed: 15663772]
- Liang Q, Kong J, Stalker J, Bradley A. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *Genesis*. 2009; 47:404–8. [PubMed: 19391106]
- Lin Q, Donahue SL, Moore-Jarrett T, Cao S, Osipovich AB, Ruley HE. Mutagenesis of diploid mammalian genes by gene entrapment. *Nucleic Acids Res*. 2006; 34:e139. [PubMed: 17062627]
- Luo G, Ivics Z, Izsvak Z, Bradley A. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*. 1998; 95:10769–73. [PubMed: 9724779]

- Mainguy G, Montesinos ML, Lesaffre B, Zevnik B, Karasawa M, Kothary R, Wurst W, Prochiantz A, Volovitch M. An induction gene trap for identifying a homeoprotein-regulated locus. *Nat Biotechnol.* 2000; 18:746–9. [PubMed: 10888842]
- Markowitz D, Goff S, Bank A. A safe packaging line for gene transfer: separating viral genes on two different plasmids. *J Virol.* 1988; 62:1120–4. [PubMed: 2831375]
- Matsuda E, Shigeoka T, Iida R, Yamanaka S, Kawaichi M, Ishida Y. Expression profiling with arrays of randomly disrupted genes in mouse embryonic stem cells leads to in vivo functional analysis. *Proc Natl Acad Sci U S A.* 2004; 101:4170–4. [PubMed: 15010531]
- Medico E, Gambarotta G, Gentile A, Comoglio PM, Soriano P. A gene trap vector system for identifying transcriptionally responsive genes. *Nat Biotechnol.* 2001; 19:579–82. [PubMed: 11385465]
- Merrick WC. Cap-dependent and cap-independent translation in eukaryotic systems. *Gene.* 2004; 332:1–11. [PubMed: 15145049]
- Mikkers H, Allen J, Knipscheer P, Romeijn L, Hart A, Vink E, Berns A. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nat Genet.* 2002; 32:153–9. [PubMed: 12185366]
- Mitchell KJ, Pinson KI, Kelly OG, Brennan J, Zupicich J, Scherz P, Leighton PA, Goodrich LV, Lu X, Avery BJ, Tate P, Dill K, et al. Functional analysis of secreted and transmembrane proteins critical to mouse development. *Nat Genet.* 2001; 28:241–9. [PubMed: 11431694]
- Niwa H, Araki K, Kimura S, Taniguchi S, Wakasugi S, Yamamura K. An efficient gene-trap method using poly A trap vectors and characterization of gene-trap events. *J Biochem.* 1993; 113:343–9. [PubMed: 8387481]
- Nord AS, Chang PJ, Conklin BR, Cox AV, Harper CA, Hicks GG, Huang CC, Johns SJ, Kawamoto M, Liu S, Meng EC, Morris JH, et al. The International Gene Trap Consortium Website: a portal to all publicly available gene trap cell lines in mouse. *Nucleic Acids Res.* 2006; 34:D642–8. [PubMed: 16381950]
- Osipovich AB, Singh A, Ruley HE. Post-entrapment genome engineering: first exon size does not affect the expression of fusion transcripts generated by gene entrapment. *Genome Res.* 2005; 15:428–35. [PubMed: 15741512]
- Raymond CS, Soriano P. High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS One.* 2007; 2:e162. [PubMed: 17225864]
- Roma G, Cobellis G, Claudiani P, Maione F, Cruz P, Tripoli G, Sardiello M, Peluso I, Stupka E. A novel view of the transcriptome revealed from gene trapping in mouse embryonic stem cells. *Genome Res.* 2007; 17:1051–60. [PubMed: 17540781]
- Roma G, Sardiello M, Cobellis G, Cruz P, Lago G, Sanges R, Stupka E. The UniTrap resource: tools for the biologist enabling optimized use of gene trap clones. *Nucleic Acids Res.* 2008; 36:D741–6. [PubMed: 17942430]
- Salminen M, Meyer BI, Gruss P. Efficient poly A trap approach allows the capture of genes specifically active in differentiated embryonic stem cells and in mouse embryos. *Dev Dyn.* 1998; 212:326–33. [PubMed: 9626507]
- Schnieke A, Harbers K, Jaenisch R. Embryonic lethal mutation in mice induced by retrovirus insertion into the alpha 1(I) collagen gene. *Nature.* 1983; 304:315–20. [PubMed: 6308457]
- Schnütgen F, De-Zolt S, Van Sloun P, Hollatz M, Floss T, Hansen J, Altschmied J, Seisenberger C, Ghyselinck NB, Ruiz P, Chambon P, Wurst W, et al. Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. *Proc Natl Acad Sci U S A.* 2005; 102:7221–6. [PubMed: 15870191]
- Schnütgen F, Hansen J, De-Zolt S, Horn C, Lutz M, Floss T, Wurst W, Noppinger PR, von Melchner H. Enhanced gene trapping in mouse embryonic stem cells. *Nucleic Acids Res.* 2008; 36:e133. [PubMed: 18812397]
- Schnütgen F, Stewart AF, von Melchner H, Anastassiadis K. Engineering embryonic stem cells with recombinase systems. *Methods Enzymol.* 2006; 420:100–36. [PubMed: 17161696]
- Shigeoka T, Kawaichi M, Ishida Y. Suppression of nonsense-mediated mRNA decay permits unbiased gene trapping in mouse embryonic stem cells. *Nucleic Acids Res.* 2005; 33:e20. [PubMed: 15687378]

- Singla V, Hunkapiller J, Santos N, Seol AD, Norman AR, Wakenight P, Skarnes WC, Reiter JF. Floxin, a resource for genetically engineering mouse ESCs. *Nat Methods*. 2010; 7:50–2. [PubMed: 19966808]
- Skarnes WC, Auerbach BA, Joyner AL. A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev*. 1992; 6:903–18. [PubMed: 1592261]
- Skarnes WC, Moss JE, Hurtley SM, Beddington RS. Capturing genes encoding membrane and secreted proteins important for mouse development. *Proc Natl Acad Sci U S A*. 1995; 92:6592–6. [PubMed: 7604039]
- Skarnes WC, von Melchner H, Wurst W, Hicks G, Nord AS, Cox T, Young SG, Ruiz P, Soriano P, Tessier-Lavigne M, Conklin BR, Stanford WL, et al. A public gene trap resource for mouse functional genomics. *Nat Genet*. 2004; 36:543–4. [PubMed: 15167922]
- Soriano P, Friedrich G, Lawinger P. Promoter interactions in retrovirus vectors introduced into fibroblasts and embryonic stem cells. *J Virol*. 1991; 65:2314–9. [PubMed: 2016762]
- Stryke D, Kawamoto M, Huang CC, Johns SJ, King LA, Harper CA, Meng EC, Lee RE, Yee A, L'Italien L, Chuang PT, Young SG, et al. BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. *Nucleic Acids Res*. 2003; 31:278–81. [PubMed: 12520002]
- Stuhlmann H. Gene trap vector screen for developmental genes in differentiating ES cells. *Methods Enzymol*. 2003; 365:386–406. [PubMed: 14696360]
- Swift S, Lorens J, Achacoso P, Nolan GP. Rapid production of retroviruses for efficient gene delivery to mammalian cells using 293T cell-based systems. *Curr Protoc Immunol*. 2001:17C. Chapter 10, Unit 10.
- Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF, Vignali DA. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat Biotechnol*. 2004; 22:589–94. [PubMed: 15064769]
- Tanaka TS, Davey RE, Lan Q, Zandstra PW, Stanford WL. Development of a gene-trap vector with a highly sensitive fluorescent protein reporter system for expression profiling. *Genesis*. 2008; 46:347–56. [PubMed: 18615730]
- Taniwaki T, Haruna K, Nakamura H, Sekimoto T, Oike Y, Imaizumi T, Saito F, Muta M, Soejima Y, Utoh A, Nakagata N, Araki M, et al. Characterization of an exchangeable gene trap using pU-17 carrying a stop codon-beta geo cassette. *Dev Growth Differ*. 2005; 47:163–72. [PubMed: 15840001]
- Tate P, Lee M, Tweedie S, Skarnes WC, Bickmore WA. Capturing novel mouse genes encoding chromosomal and other nuclear proteins. *J Cell Sci*. 1998; 111(Pt 17):2575–85. [PubMed: 9701556]
- To C, Epp T, Reid T, Lan Q, Yu M, Li CY, Ohishi M, Hant P, Tsao N, Casallo G, Rossant J, Osborne LR, et al. The Centre for Modeling Human Disease Gene Trap resource. *Nucleic Acids Res*. 2004; 32:D557–9. [PubMed: 14681480]
- Trombly MI, Su H, Wang X. A genetic screen for components of the mammalian RNA interference pathway in Bloom-deficient mouse embryonic stem cells. *Nucleic Acids Res*. 2009; 37:e34. [PubMed: 19223321]
- Tsakiridis A, Tzouanacou E, Larralde O, Watts TM, Wilson V, Forrester L, Brickman JM. A novel triple fusion reporter system for use in gene trap mutagenesis. *Genesis*. 2007; 45:353–60. [PubMed: 17492751]
- Tsakiridis A, Tzouanacou E, Rahman A, Colby D, Axton R, Chambers I, Wilson V, Forrester L, Brickman JM. Expression-independent gene trap vectors for random and targeted mutagenesis in embryonic stem cells. *Nucleic Acids Res*. 2009; 37:e129. [PubMed: 19692586]
- Uren AG, Mikkers H, Kool J, van der Weyden L, Lund AH, Wilson CH, Rance R, Jonkers J, van Lohuizen M, Berns A, Adams DJ. A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. *Nat Protoc*. 2009; 4:789–98. [PubMed: 19528954]
- von Melchner H, DeGregori JV, Rayburn H, Reddy S, Friedel C, Ruley HE. Selective disruption of genes expressed in totipotent embryonal stem cells. *Genes Dev*. 1992; 6:919–27. [PubMed: 1317320]

- Wagner EF, Covarrubias L, Stewart TA, Mintz B. Prenatal lethalties in mice homozygous for human growth hormone gene sequences integrated in the germ line. *Cell*. 1983; 35:647–55. [PubMed: 6580963]
- Wang W, Bradley A. A recessive genetic screen for host factors required for retroviral infection in a library of insertionally mutated Blm-deficient embryonic stem cells. *Genome Biol*. 2007; 8:R48. [PubMed: 17407599]
- Wang W, Bradley A, Huang Y. A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. *Genome Res*. 2009; 19:667–73. [PubMed: 19233961]
- Wang W, Lin C, Lu D, Ning Z, Cox T, Melvin D, Wang X, Bradley A, Liu P. Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*. 2008; 105:9290–5. [PubMed: 18579772]
- Xin HB, Deng KY, Shui B, Qu S, Sun Q, Lee J, Greene KS, Wilson J, Yu Y, Feldman M, Kotlikoff MI. Gene trap and gene inversion methods for conditional gene inactivation in the mouse. *Nucleic Acids Res*. 2005; 33:e14. [PubMed: 15659575]
- Yoshida M, Yagi T, Furuta Y, Takayanagi K, Kominami R, Takeda N, Tokunaga T, Chiba J, Ikawa Y, Aizawa S. A new strategy of gene trapping in ES cells using 3'RACE. *Transgenic Res*. 1995; 4:277–87. [PubMed: 7655516]
- Zambrowicz BP, Abuin A, Ramirez-Solis R, Richter LJ, Piggott J, BeltrandelRio H, Buxton EC, Edwards J, Finch RA, Friddle CJ, Gupta A, Hansen G, et al. Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention. *Proc Natl Acad Sci U S A*. 2003; 100:14109–14. [PubMed: 14610273]
- Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C, Sands AT. Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature*. 1998; 392:608–11. [PubMed: 9560157]

Fig. 1.

Schematic diagrams of different types of gene trap alleles

A) Insertion of an enhancer trap vector with minimal promoter (mP) nearby a transcriptionally active gene activates expression of the β -galactosidase reporter (β gal).
 B) A promoter trap insertion, in which a splice acceptor (SA) fuses an endogenous transcript to a β -galactosidase/neomycin phosphotransferase (β geo) reporter/selection cassette. The fusion transcript of the promoter trap confers drug resistance, reports the expression pattern, and mutates the function of the endogenous trapped gene.

Note that promoter trapping can also be performed with vectors that express drug resistance by a strong heterologous promoter (see text for details).

C) The secretory trap vector is designed to trap genes that encode secretory signal peptide (SS). Secretion of the fusion protein from the endogenous gene and the trap vector is prevented by a transmembrane domain (TM) encoded in the trap vector. The TM anchors the fusion protein at the intracellular side of the cell membrane, thereby conferring drug resistance to trapped cells.

D) The insertion of an exon trap vector into an exon of a transcriptionally active gene leads to a direct fusion of a selection gene (neomycin phosphotransferase, neo) to the endogenous gene.

E) A polyA trap vector has two components: an upstream element with 5' splice acceptor (SA β gal) to report gene expression, and a downstream element with promoter, selection gene, and splice donor (P-neo-SD), to select for drug-resistant intergenic insertions.

F) The conditional "Flex" design contains a promoter trap cassette that is flanked by upstream and downstream arrays of heterologous recombination sites for the FLP (semicircles) and Cre recombinases (triangles); for details see (Schnütgen et al., 2005). Successive application of these recombinases leads to inversion (conditional allele) and re-inversion (re-instated mutagenic allele) of the promoter trap cassette.

G) The conditional double switch vector contains an upstream 5' splice acceptor-GFP reporter component (SA-IRES-eGFP-pA), which is in inverse orientation to gene transcription and flanked by half mutant lox66 and lox71 sites (triangles). A downstream polyA trap component (P-puro-SD) is flanked by FRT recombination sites (semi-circles). Excision of the polyA trap component by FLP recombinase creates the conditional allele, and subsequent inversion of the reporter element generates the mutated allele (see (Xin et al., 2005) for details).

Fig. 2.
Workflow of gene trapping by retroviral infection (A) or by electroporation with transposon vectors (B).

Fig. 3.

Identification of gene trap insertion sites by Splinkerette PCR. (1) A DNA preparation of an ES cell clone is digested with *BstYI* restriction endonuclease. (2) ligation to the Splinkerette linker. (3) In the first round of PCR, only the vector-specific primer V1 can anneal and initiate synthesis of a strand that complements the Splinkerette linker. (4) The linker-specific primer L1 and the vector-specific primer V1 PCR amplify a junction fragment of the vector and the adjacent genomic DNA. (5) A PCR with nested primers L2 and V2 serves to improve specificity and yield of the reaction. (6) Generation of unspecific products is minimized, since the primer L1 cannot anneal to linkers that contains the self-complementary "hook" region.

Table 1**Gene Trap Resources**

Program	Website	Reference
International Gene Trap Consortium (IGTC)	www.genetrap.org	(Nord et al., 2006; Skarnes et al., 2004)
German Gene Trap Consortium (GGTC)	www.genetrap.de	(Hansen et al., 2003)
Sanger Institute Gene Trap Resource (SIGTR)	www.sanger.ac.uk/genetrap/	
Centre for Modelling Human Disease (CMHD)	www.cmhd.ca/genetrap/	(To et al., 2004)
TIGEM-IRBM Gene TRAP Project	www.genetrap.tigem.it/	
Texas A&M Institute of Genomic Medicine (TIGM)	www.tigm.org	(Hansen et al., 2008)
Database for exchangeable genetrap clones	www.egtc.ip	(Taniwaki et al., 2005)
Soriano lab	www.drbsinai.org/pc/soriano.html	(Chen et al., 2004a)
Centre for Mammalian Functional Genomics	escells.ca	(Hicks et al., 1997)
Baygenomics	baygenomics.ucsf.edu	(Stryke et al., 2003)
European Conditional Mouse Mutagenesis (EUCOMM)	www.eucomm.org	(Friedel et al., 2007)
Unitrap gene trap browser	unitrap.cbm.fvg.it	(Roma et al., 2008)
Knockout Mouse Project (KOMP)	www.knockoutmouse.org	(Collins et al., 2007)
Mouse Genome Informatics	www.informatics.iax.org	(Bult et al., 2009)
Mutant Mouse Regional Resource Centers (MMRRC)	www.mmrrc.org	
European Mouse Mutant Cell Repository (EuMMCR)	www.eummc.org	

Table 2

Splinkerette primers

Adapter	
HMSpAA	CGAAGAGTAACCGTTGCTAGGAGAGACCGTGGCTGAATGAGACTGGTGTGCGACTAGTGG
HMSpBB	<i>GATC</i> ccactagtgtcgacaccagtctctaa <i>TTTTTTTTTCAAAAAA</i>
1 st round PCR	
HMSp1	CGAAGAGTAACCGTTGCTAGGAGAGACC
PB5'-1	CAAAATCAGTGACACTTACCGCATTGACAA
PB3'-1	TAAATAAACCTCGATATACAGACCGATAAA
2 nd round PCR	
HMSp2	GTGGCTGAATGAGACTGGTGTGCGAC
PB5'-2	CTTACCGCATTGACAAGCACGCCTCACGGG
PB3'-2	ATATACAGACCGATAAAACACATGCGTCAA
Direct sequencing	
PB5'-seq	TTAGAAAGAGAGAGCAATATTCAAGAATG
PB3'-seq	TTTTACGCATGATTATCTTTAACGTACGTC

HM primers from (Mikkers et al., 2002)

PB primers from (Wang et al., 2009)