Review

Insertional mutagenesis strategies in zebrafish
Sridhar Sivasubbu*, Darius Balciunas†, Adam Amsterdam‡ and Stephen C Ekker†§

*Institute of Genomics and Integrative Biology, Council for Scientific and Industrial Research, Mall Road, Delhi 110007, India. †University of Minnesota, Department of Genetics, Cell Biology and Development, Arnold and Mabel Beckman Center for Transposon Research, Church St SE, Minneapolis, Minnesota 55455, USA. ‡Center for Cancer Research, Department of Biology, Massachusetts Institute of Technology, Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. §Current address: Mayo Clinic College of Medicine, Department of Biochemistry and Molecular Biology, Rochester, Minnesota 55905, USA.

Correspondence: Stephen C Ekker. E-mail: ekker.stephen@mayo.edu

Abstract

We review here some recent developments in the field of insertional mutagenesis in zebrafish. We highlight the advantages and limitations of the rich body of retroviral methodologies, and we focus on the mechanisms and concepts of new transposon-based mutagenesis approaches under development, including prospects for conditional 'gene trapping' and 'gene breaking' approaches.

Introduction

The human genome contains over 20,000 predicted genes as well as an array of conserved noncoding genetic information [1], which will have to be functionally validated employing a variety of model systems. Traditional vertebrate model systems such as mouse have now been complemented by an array of genetically tractable organisms, including rat, the diploid frog *Xenopus tropicalis*, and zebrafish (*Danio rerio*).

The mouse has been the traditional 'Cadillac' of vertebrate models because of its mammalian nature, its rapid breeding time, and the extensive molecular genetic tools that have been developed for use in this animal. In addition, the unique development of embryonic stem (ES) cells for targeted genetic modification using homologous recombination, along with new recombineering techniques, now allow one to manipulate fully the genome of this organism [2]. Insertional methods for genetic modification and gene disruption have been developed extensively for this organism, largely focused on the advantages of working with ES cells *in vitro* prior to the necessary and laborious whole organism studies. The latter work makes the mouse a relatively expensive and complex model organism to use, and this limitation makes *in vivo* genome-wide genetic approaches impractical in many important scientific scenarios. To address this need, several vertebrate systems complementary to those of the mouse have been developed, as illustrated by the zebrafish *Danio rerio*. It is noteworthy that the insertional approaches now being deployed for the latter system would be especially appropriate methodology for genome modification of other important model vertebrates that lack ES cells, such as rat and pig.

Advantages of the zebrafish include its external fertilization, high fecundity, rapid development, production of optically clear embryos, and relatively short generation time for a vertebrate [3]. These qualities, in addition to the high degree of genetic conservation [4-6] reflected in the developmental gene pathways and regulatory mechanism, contribute to its emergence as a model for obtaining insights into fundamental human physiology. Current established forward genetic tools for the zebrafish include chemical (*N*-ethyl-*N*-nitrosourea [ENU]; for review [7]) and insertional (retroviral [8-11]) mutagens. Reverse genetic methods include morpholino antisense oligonucleotides [12] and targeted lesion detection, called TILLING (targeting-induced local lesions in genomes) [13].

Chemical mutagenesis using ENU has been key to establishing the zebrafish as a forward genetic model system. ENU produces random point mutations in the germline, and

these single base pair changes result in a high frequency of mutant phenotypes (for review [7]). Multiple large-scale chemical mutagenesis screens using ENU [14,15] have successfully produced and characterized an impressive collection of zebrafish mutants that affect various biological processes. In spite of the high efficiency in generation of point mutations, the major limitation in this approach is the identification of genes whose mutations are responsible for the particular phenotype [16]; the identities of the mutated genes have been reported for about 154 chemically induced mutations out of approximately 1,740 mutants recovered from the two large-scale chemical mutagenesis screens [17].

An alternative approach is insertional mutagenesis [18], in which an exogenous DNA serves as a mutagen and also functions as a molecular tag for identifying the gene whose disruption causes the phenotype. One effective mutagen for zebrafish is the pseudotyped retrovirus, which is composed of a genome based on the Moloney murine leukemia virus and the envelope glycoprotein of the vesicular somatitis virus [8,19]. Injection of this retrovirus into 1,000-cell to 2,000-cell stage zebrafish embryos [19] results in chimeric embryos in which different cells have integrations of the viral sequences in different random sites in the genome. By passing these insertions through the germline and inbreeding them, one finds that about 1 in 80 of such insertions result in mutant phenotypes. This method has been used successfully in zebrafish in a large-scale forward genetic screen, identifying more than 500 mutations and about 350 loci; the insertional nature of the mutagen facilitated the rapid molecular characterization of the genetic loci, with 335 cloned to date [8-11]. One major challenge to the field has been the inability to develop a similarly mutagenic, high-titer retrovirus with robust expression [17,20,21]; this limitation has hindered the generation of expression based mutagenic retroviral vectors for the zebrafish.

Recent new initiatives in zebrafish include enhancer [21-24] and gene trapping [25] approaches. Enhancer trapping has been an integral part of the *Drosophila* genome project [26], suggesting that this and related methodologies should provide additional means for deciphering gene function in zebrafish. However, enhancer trapping, unlike gene trapping, is not expected to enrich for mutations because inserts need not be in genes nor disrupt them. Indeed, in zebrafish no enhancer or gene trap vectors have yet yielded any phenotypic mutants [22,23,25,27,28].

To address the limitations of mutagenicity found in enhancer trapping and related approaches, in one recent study [29] the investigators used a modified vector dubbed the 'gene break', named after a human insertional mutation mechanism [30]. This approach confers new capabilities upon insertional vectors in zebrafish; when integrated into an intron this cassette can direct all or nearly all splicing into the reporter construct and thereby quantitatively terminate

transcription of the endogenous gene. This vector could be used in zebrafish *in vivo* for genome-wide forward and reverse genetic applications. In this review we compare published zebrafish insertional mutagenesis strategies. In addition, we discuss how new tools can be developed to complement those developed in mouse for conditional mutagenesis and chromosomal engineering applications to untap the genome-wide screening potential found in the zebrafish model system.

Insertional mutagenesis: history and perspective

Insertional mutagenesis in vertebrates has a rich past using many vectors and approaches. The most prominent tool for delivering DNA as a mutagen is based on an engineered viral vector. Newer methods include the use of transposons (see below) and other methodologies restricted to mouse ES cells that harness those cells' uniquely high (in vertebrates) activity homologous recombination machinery (such as MICER and other tools; for review [2]). This review focuses on methods of general interest to the broader vertebrate genetics community and is based on the extensive literature with viruses, while also providing a trajectory for the next generation of transposon-based vectors, especially those in active use in zebrafish.

One key issue for all insertional methods is the mechanism of mutagenicity. Insertion of DNA into most locations on a vertebrate genome has little or no effect on any gene or gene product. However, and regardless of vector deployed, random insertion of DNA into a genome will, at a modest frequency, result in the disruption of an exon (Figure 1a). The vector independent potential consequences on the genetic locus of such an insertion event are many. First, the insertion may prematurely end the coding region by introducing a premature stop codon, truncating the resulting protein product (Figure 1a). Other consequences can manifest at the RNA level, resulting in a destabilization of the transcript such as through nonsense mediated mRNA decay.

The ability of a vector to cause exon disruption is a function of many different variables, including the frequency of exons encoded by the genome, the insertion site bias (or lack thereof) inherent in the vector, and the availability of such sequences to insertion by the mutagen. As exons tend to encode only 1% to 2% of most vertebrate genomes [1], the basal rate of inducing genetic modification by exon disruption is typically very low. Strategies for further increasing the rate of gene inhibition over this basal mechanism, often using vectors in which mutagenic insertions can be enriched by selection, as well as conditional allele generation strategies, are described below.

Retroviral insertional mutagen in zebrafish

The most extensively studied insertional mutagen to date in zebrafish is the pseudotyped retrovirus [8-11,31,32]. These
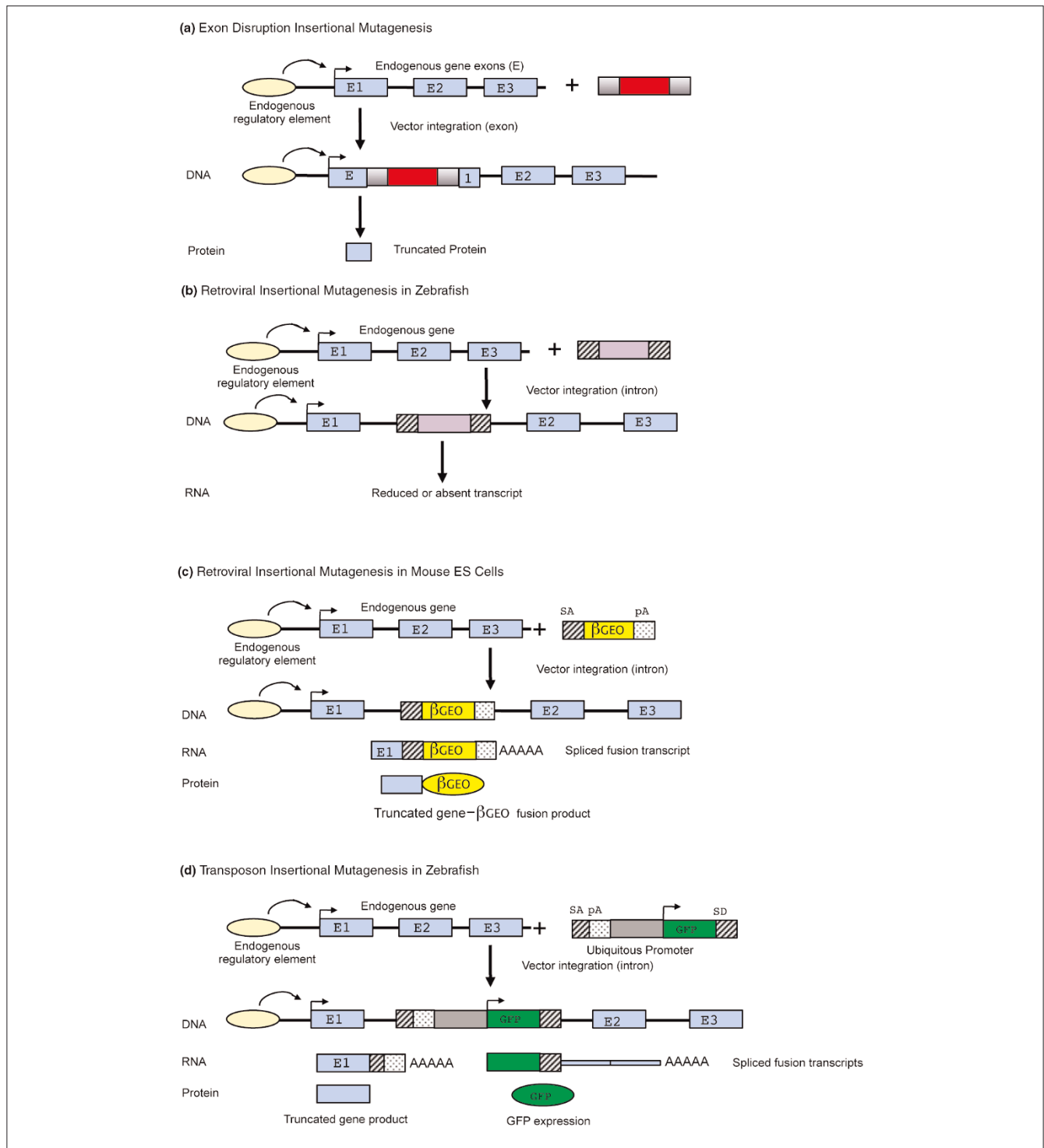
Figure 1

Insertional mutagenesis strategies used in vertebrates. In each case, a schematized endogenous locus is represented by exons (E) and an endogenous regulatory element. A nonintegrated vector is also shown above, with integrated vector below. Transcriptional start sites are shown as an arrow above each diagram. **(a)** Integration of DNA into a coding exon can mutate the locus, resulting in a truncated gene product. **(b)** Retroviral insertional mutagenesis alters the tagged locus using multiple methods, including the loss of the encoding transcript. **(c)** 5' gene trapping in mouse embryonic stem cells. Shown is one approach whereby the resulting fusion transcript encodes a truncated gene product fused to the selectable marker protein. **(d)** Insertional mutagenesis in zebrafish using transposons. Based on 3' exon or poly(A) trapping methods, this approach uses two components: a transcriptional termination cassette to truncate the integrated locus and a separate 3' exon trap gene finding cassette. See text for details. pA, polyadenlyation signal; SA, splice acceptor; SD, splice donor.

retroviral vectors have been used to molecularly characterize more mutations to date in zebrafish than all other methods combined [17]. This tool is being further deployed in a reverse genetic approach by Znomics, Inc. (Portland, OR, USA) [33], a project in which tens of thousands of insertions are being mapped to the genome and are individually recoverable through cryopreserved sperm.

Retroviruses appear to cause mutations in zebrafish by several major mechanisms, including exon disruption (Figure 1a) and gene silencing caused by insertion into an intron (Figure 1b). Nearly 30% of mutagenic insertions recovered from a large-scale mutagenesis screen were in exons, although about half of these are in 5' untranslated regions [32]. In all cases examined, such insertions lead to a complete loss of wild-type gene product [11] (Amsterdam A, unpublished observations). Most of the other 70% of mutagenic insertions are in introns (with the last few in promoters). In part because of a preference of Maloney-based viruses to insert near to the 5' end of genes [34,35], most of these insertions are in the first intron. For reasons that are not clear, these insertions usually result in the reduction or complete abrogation of endogenous RNA expression [11] (Amsterdam A, unpublished observations). Intronic insertions can also lead to aberrant splicing, resulting in skipped exons and either frameshift mutations or internally truncated gene products [36-38]. Another way in which intronic insertions can be mutagenic is illustrated by one of the viruses used in this screen employing a 'splice-in, splice-out' gene trap [39], which causes a frameshift in the subsequent exon. Although mRNAs containing this trapped exon are found only in a modest subset of the cases in which the virus has inserted into an intron in the correct orientation to be utilized, this may be due either to nonsense-mediated mRNA decay or to general loss of gene expression, as is seen in other intronic insertions. For this gene trap virus, but not for a virus that lacks the gene trap, there is a bias toward the trap orientation among mutagenic insertions that have landed in introns past the ATG (Amsterdam A, unpublished observations), which is consistent with the trap contributing to mutagenicity of this virus, even in cases in which trap containing messages are not observed.

Gene trapping: tapping the rich history of insertional mutagens in mouse embryonic stem cells

The genetics of the murine system has been revolutionized by *in vitro* ES cell culture, which allows one to modify and confirm the resulting genetic changes induced in these totipotent cells prior to the laborious and slow downstream *in vivo* studies. Many approaches to the genetic modification of ES cells have been deployed, including the addition of DNA by electroporation, viruses, and transposons. For example, a recent method for insertional mutagenesis in

mice using gene trap transposons has recently been reported [40]. For the sake of simplicity, here we highlight the body of work done largely using retroviral vectors (extensive reviews of this topic are provided elsewhere [2,41]).

One major distinguishing feature of insertional mutagenesis in ES cells described here from that of the retroviral *in vivo* work in zebrafish is the fundamental requirement for expression-compliant modifications; in each case, there is a selection process employed to distinguish modified from unmodified cells *in vitro*. ES cell technology offers the ability to conduct genome-wide analyses through clonal selection, rapid amplification of cDNA ends (RACE) analyses, and ES cell cryopreservation. The major bottleneck in mouse genomics is the significant investment in resources required to study even a single gene *in vivo* in this model system.

One key method pioneered in ES cells uses an expression-based approach called promoter or 5' gene trapping (Figure 1c) [2,41]. Five prime gene traps combine the potential of highly efficient mutagenesis with the ease with which the mutant loci can be cloned. Five prime gene trap vectors typically contain a splice acceptor immediately upstream of a promoterless reporter used for ES cell selection (such as βgeo). Integration of the gene trap vector in a promoter, exon, or intron of transcriptionally active loci can generate a fusion transcript between the upstream coding sequence and the reporter (Figure 1c). With a high efficiency splice acceptor and poly(A) signal serving as an artificial 3' terminal exon, this trap can disrupt the expression of the trapped locus by inducing truncation of the 'hijacked' transcript [2,41-44]. The fusion transcripts generated by gene trap integration also serves as templates for identification and cloning of the disrupted gene, using a polymerase chain reaction based technique called 5' RACE. One key mutagenicity mechanism in nearly all gene trapping methods is the ability to truncate transcription at the point of insertion, which is achieved by the inclusion of a high quality transcriptional termination cassette. Without such a module, splicing around the trap can readily occur, thus resulting in an insertion without effectively knocking out function at the insertion locus.

Limitations of this basic method include the requirement for endogenous expression of the locus in ES cells. This vector system can be used to help identify endogenous, tissue-specific expression, but only after differentiation of the ES cells *in vitro* or after the generation of the transgenic animal from these modified ES cells. Furthermore, because the trap vector insertion can occur in any one of the three reading frames, only one-third of the loci that are in the correct reading frame to yield a functional reporter fusion protein will be identified using this method.

A rich array of modified gene traps have now been generated to help address some of these and other limitations of the

original gene trap design for use in ES cells [2]. An example approach known as poly(A) trapping [2,41,45-49] is the panel of insertions generated by Lexicon Pharmaceuticals [50]. In this example, a 5'-style gene trap as a mutagenicity cassette and a 3' gene trap are used to enrich for intragenic insertions, regardless of the expression status of the locus in ES cells (Figure 1d; see below for a detailed description of a related approach deployed in zebrafish). The 3' gene trap consists of a constitutive promoter/enhancer that drives the expression of a reporter gene that contains a downstream splice donor instead of a polyadenylation signal (poly(A) signal). Integration of the 3' gene trap vectors in the proper *cis* orientation results in a spliced poly(A) signal from the endogenous gene. This generates a stable fusion transcript between the vector derived reporter and the downstream exons of the trapped gene, resulting in functional reporter expression. Therefore, insertions with viable reporter expression are enriched for intragenic events. The fusion transcript also serves as a template for identification and cloning of the trapped genes by 3'-RACE. Because trapping does not depend on the expression status or relative abundance of the endogenous transcript, nearly all genes in an organism should be available for screening using this approach, which is a key advantage of 3' gene traps over 5' gene traps. Despite the inclusion of the 3' gene trap component in this vector, the primary mutagenicity mechanism for the Lexicon ES cell panel is thought to be due to the inclusion of the 5' transcriptional termination cassette in *cis* to the 3' gene trap [50].

Transposon-based insertional mutagens in zebrafish

Two groups have employed different 5' gene trap vectors in zebrafish. In the first case, the trap was not necessarily used for selection but possibly to increase the mutagenicity of the retroviral vector described in the previous section for insertional mutagenesis [39]. This vector utilized a splice-in, splice-out vector, similar to the basic gene trap design used for mouse ES cells (Figure 1c), with intronic insertion in the correct orientation inducing a frameshift and probably causing either a truncated protein or a loss of gene product due to nonsense-mediated mRNA decay. The inclusion of this gene trap cassette in the retroviral forward genetic screen did successfully generate some additional mutations in this screening work, but the overall mutagenicity rate of the trap containing vector was not dramatically different from that of a non-trap-containing virus [9-11]. Because this trap lacked an easily detectable reporter, it was not useful for screening strategies that specifically select trapped insertions for further breeding.

A 5' gene trap containing a splice acceptor and the green fluorescent protein (GFP) gene was used in a *Tol2* based transposon insertional study in zebrafish [25]. Integration of the gene trap in the proper orientation and reading frame resulted in GFP expression in temporally and spatially restricted patterns. Using this approach it was shown that endogenous transcripts could be successfully trapped. However, in the only case examined, expression of the endogenous locus was only reduced fourfold, implying that the trap could be spliced around at levels that may allow phenotypic consequences to be avoided. In this study, thirty-six trapped lines were homozygosed with no visible phenotypes, but because only about 5% of zebrafish genes may show embryonic phenotypes when mutated, this is not enough lines to conclude whether this type of gene trap can reliably mutate genes. Paradoxically, the experience of mutagenic retroviral insertions suggests one potential hypothesis for why this sort of gene trap may not be a very effective mutagen. Because many insertions in introns (the type of insertion required to activate this trap) can abrogate or severely reduce gene expression, it may be that many such insertions are not detected as trap events because the GFP reporter cannot be visualized in the absence of expression of the endogenous gene. It may therefore be a subset of intronic insertions that are selected, namely those that do not have an appreciable impact on expression of the locus; in these cases, if sufficient message is able to splice around the trap, then phenotypes might not be observed. A careful analysis directly comparing the effects on gene expression of trapping cassettes that have inserted into a given locus that either do or do not express the trap reporter will be required to determine whether this explanation might account for the modest mutagenic success noted for conventional 5' gene traps in zebrafish.

Recently, a combined 5'-3' gene trap ('gene breaking') vector was developed to trap genes in zebrafish [29], in a method similar to the Lexicon traps described for ES cells in the previous section. We use the newer term 'gene breaking' [30] to emphasize the dual module nature of this approach, which includes a 5' transcriptional terminator cassette to mutate the gene in concert with 3' gene trapping as an alternative strategy to select for intragenic vector integrations (Figure 1d). It is worth noting that the ability of this kind of vector to mutate genes upon intronic insertion is almost exclusively due to the 5' transcriptional termination cassette, a function that is independent of the 3' gene trapping mechanism. Although the employment of a transcriptional terminator in the gene breaking trap vector allows suppression of splicing around the trapping vector, the trapped gene expression domain cannot readily be identified using this basic approach. Alternative approaches to add this feature to gene breaking transposons are underway (Balciunas D, Ekker SC, unpublished data).

Practical limitations on the mutagenicity of gene trapping and related approaches

Integration of an insertional mutagenesis vector in an intron of a gene is generally expected to interfere with the synthesis

of the normal spliced transcript. However, the ability to reduce endogenous gene expression depends on the efficiency of the splicing, polyadenylation, and any other transcriptional termination signals in the insertional mutagenesis vector. Employment of a weak splicing signal in the gene trap will allow splicing of the endogenous transcript around the trap insertion and cause restoration of undisrupted wild-type transcript, and this has shown to be one of the major hurdles in creating null alleles using gene traps in mouse [51-53]. In zebrafish, gene trapping can occur, but the effective level of gene disruption can be variable [25,39]. The inability of a particular gene trap to disrupt endogenous gene expression could be due to choosing vector components that might not work efficiently in the employed model system. In this regard, it is important to consider the efficiency and limitations of each individual component before assembling the final vector. Not only do the vector components need to work, but they must also work efficiently for many aspects of trapping to be successful. One way to test the vector components is to conduct artificial test trapping studies. In this regard, before assembling the gene breaking trap vector, we conducted artificial test trapping studies and vector components were selected based on the performance of the components in these studies [29]. We have also noticed that vector components originating from nonpiscine sources tend to perform poorly in trapping contexts in zebrafish. Examples include the SV40 poly(A) signal and triple poly(A) signal sequence originating from murine vectors [54].

Many 3' gene traps have an inherent 3' bias when integrating in genes caused by the reduction in expression from the resulting fusion transcript with long, untranslated 3' sequences that are now subject to nonsense-mediated mRNA decay; thus, insertions might truncate but still leave a viable protein from the 5' transcript. This problem has been addressed in part by the inclusion of an internal ribosomal entry sequence (IRES) immediately ahead of the splice donor [55]; the inclusion of the IRES separately induces translation of the otherwise untranslated 3' exons, alleviating nonsense-mediated mRNA decay. Potential precocious *trans* activation at the endogenous locus by the strong enhancer/promoter traditionally used in 3' gene traps is also a possible concern for this gene identification strategy.

Perspectives and future insertional mutagenesis applications for the zebrafish

Analysis of the human genome has revealed that only a small fraction of the genome is spanned by exons (about 1%); this is in contrast to introns, which account for almost one-quarter (approximately 24%) of the genome [1]. In addition, exons are on average small (encoding an average of only 50 codons) and are separated by long introns (some exceeding 10 kilobases) [56]. This disproportionate size of exons relative to introns is probably true for other vertebrates as

well [57,58]. This poses a challenge for conducting insertional mutagenesis because any sufficiently random insertional mutagen is more likely to integrate into the intron than an exon. Depending upon the insertional vector used, intronic insertions may have an insufficient impact on the amount of wild-type transcript, and thus they often result in hypomorphic alleles [41]. Therefore, an ideal insertional mutagenesis vector would be one that can suppress splicing around the insertional mutagenesis vector and ensure the complete or near-complete disruption of endogenous gene expression.

We have presented the current status of insertional mutagenesis in zebrafish and have discussed in detail the recent advances in vectors that have made it possible to develop a viable gene tagging, identification, and mutagenicity method (gene breaking) for zebrafish. This technique differs from the existing zebrafish mutagenesis approaches in several ways. The gene finding and the gene mutating functions that are usually coupled together in a trapping vector have been separated into independent modules, so that one can use different 'gene finding cassettes' such as 3' gene trap or an enhancer trap in conjunction with the 'mutagenesis cassette' for the effective disruption of the trapped transcripts. Thus, the gene breaking vector is modular in design, providing flexibility to the user to pick and choose the individual modules to trap genes with diverse functions. The gene breaking trap is also designed to be independent of delivery mechanism. This was deliberately done because insertional mutagenesis studies in *Drosophila* have shown that, to achieve genome-wide saturation mutagenesis, multiple delivery tools that have distinct global and local gene tagging behavior must be employed [26,59]. Operating independently of delivery mechanism offers the flexibility to use alternative delivery tools in zebrafish, such as *Tol2* transposon [23,25,60-62], *Ac/Ds* transposon [63], and retrovirus [19], or to employ *in vivo* delivery methods such as transposase-expressing animals [64,65].

The gene breaking strategy offers several advantages in screening and molecular characterization of tagged loci over existing mutagenesis systems. Because the 3' gene trap employed for gene finding is not dependent on expression of the endogenous trapped gene, this trap can identify and mutate genes that cannot be isolated in a classical gene trap screen [25] or a phenotype-driven insertional mutagenesis screen [9]. By recovering the trap lines and identifying all trapped transcripts, one potentially could use the gene breaking traps for a reverse genetics screen by establishing a sequence-based library of insertional alleles in genes that might have adult phenotype. Such an approach would permit application of a sequence-based prioritization for identifying subtle or adult phenotypes caused by mutations in pre-selected genes. Similar insertional libraries have been created in other model organisms such as fly and
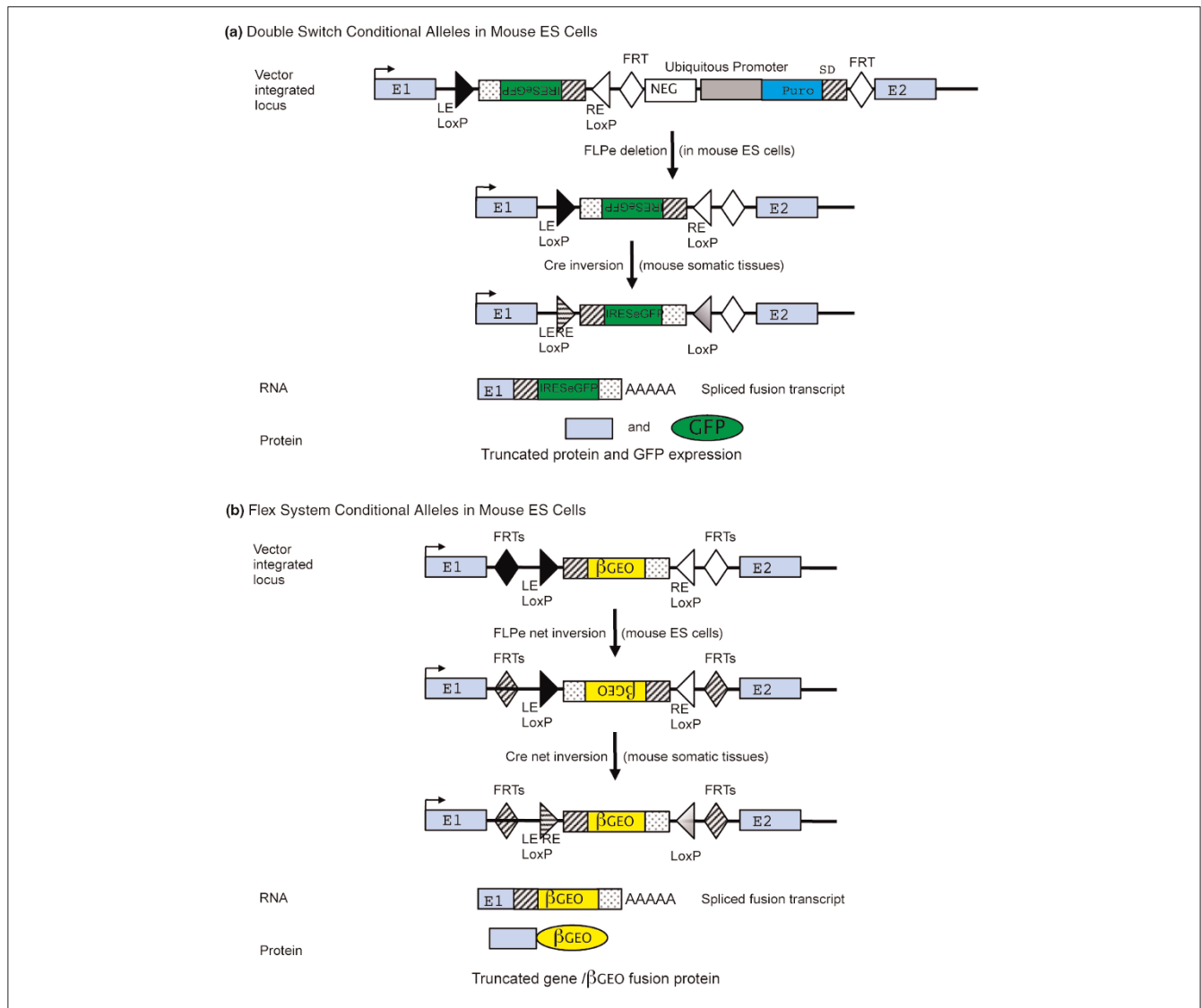
Figure 2
Conditional insertional mutagenesis strategies. In each case, a vector-integrated locus is shown. Sequences required for Cre-mediated recombination are designated 'LoxP' and shown as a triangle; similarly, sequences required for FLP-mediated recombination are designated 'FRT' and shown as a diamond. These recombinase-requiring sequences are complex components and have been greatly simplified in this diagram. Recombinase sequences that have been converted to an inactive form after inversion are shown hatched. Selection and modification after integration is conducted in mouse embryonic stem (ES) cells *in vitro*, resulting in a nonmutagenic insertional allele. The locus is shut off in mice *in vivo* after Cre-mediated inversion and resolution. See text for details. E, exons.

mouse. Thousands of fly lines have been created using the P-element transposon [26], and efforts are underway in several laboratories to establish a collection of mouse ES cells with gene trap insertions (for review [2]). The creation of an insertional library in fish might also be achieved by the use of retroviral vectors, and is in fact largely underway [33], although it is not yet clear what proportion of the insertions will have an impact on gene expression because they have been selected only by insertion site, and not by phenotype.

Prospects for conditional insertional mutations: illustrated examples of conditional gene traps in mouse

Many vertebrate genes are used in diverse tissues and in distinct temporal windows of development. The ability to control gene function in time and space critical for the elucidation of gene function *in vivo*. Consequently, gene trapping in ES cells (Figure 1c) has been modified by several groups to include unidirectional recombination technology (Figure 2). In both cases, the final ES construct is an

inverted gene trap cassette designed to serve only as a transcriptional terminator after Cre-mediated inversion in the mouse *in vivo*.

In the 'double switch system' [66], a 3' exon trap is used to enrich for intragenic insertions in ES cells; this vector also contains a negative selection cassette to allow subsequent deletion of this part of the vector through FLP-mediated recombinational deletion (Figure 2a). In this configuration the insertion is not mutagenic, but the targeted gene can be shut off by unidirectional Cre-mediated inversion and resolution of the trap *in vivo* in the mouse. Thus, inactivation of the gene can be spatiotemporally regulated by appropriate expression of the Cre recombinase. One advantage of the 'double switch' generated alleles is the use of GFP to follow endogenous gene expression noninvasively *in vivo*; reporter expression is independent of reading frame through the use of an IRES upstream of the reporter (GFP). The 'flex system' [56] provides an alternative strategy (Figure 2b) that uses two different unidirectional recombination methods. The vector initially integrates as a standard 5' gene trap (compare with Figure 1c); after ES cell selection for intragenic insertions through expression of the reporter, FRT-mediated recombination and resolution causes inversion of the trap, resulting in a nonmutagenic orientation of the 5' trap vector. This cassette also blocks transcription at the targeted locus after a unidirectional Cre-mediated inversion and resolution of the trap *in vivo* in the mouse. The pioneering work using these kinds of approaches for the generation of conditional insertional alleles suggests that similar methods will be very powerful in other vertebrate systems such as zebrafish, rat, and pig.

Conclusion

Insertional vectors have been developed to allow one to use functional, expression-based or sequence-based criteria to distinguish and prioritize mutagenic integrations for subsequent analyses. In addition, a number of these vectors have the potential for downstream chromosomal engineering approaches for both somatic and germline applications. Insertional mutagenesis is poised to become an integral component of the molecular genetic toolbox for the zebrafish.

Note added in proof

A paper describing an academic project similar to the Znomics insertional library [33] was recently published [67].

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

References

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, *et al.*: **The sequence of the human genome.** *Science* 2001, **291:**1304-1351.
2. Raymond CS, Soriano P: **Engineering mutations: deconstructing the mouse gene by gene.** *Dev Dyn* 2006, **235:**2424-2436.
3. Kimmel CB: **Genetics and early development of zebrafish.** *Trends Genet* 1989, **5:**283-288.
4. Barbazuk WB, Korf I, Kadavi C, Heyen J, Tate S, Wun E, Bedell JA, McPherson JD, Johnson SL: **The syntenic relationship of the zebrafish and human genomes.** *Genome Res* 2000, **10:**1351-1358.
5. Postlethwait JH, Woods IG, Ngo-Hazelett P, Yan YL, Kelly PD, Chu F, Huang H, Hill-Force A, Talbot WS: **Zebrafish comparative genomics and the origins of vertebrate chromosomes.** *Genome Res* 2000, **10:**1890-1902.
6. Liu TX, Zhou Y, Kanki JP, Deng M, Rhodes J, Yang HW, Sheng XM, Zon LI, Look AT: **Evolutionary conservation of zebrafish linkage group 14 with frequently deleted regions of human chromosome 5 in myeloid malignancies.** *Proc Natl Acad Sci USA* 2002, **99:**6136-6141.
7. Patton EE, Zon LI: **The art and design of genetic screens: zebrafish.** *Nat Rev Genet* 2001, **2:**956-966.
8. Gaiano N, Amsterdam A, Kawakami K, Allende M, Becker T, Hopkins N: **Insertional mutagenesis and rapid cloning of essential genes in zebrafish.** *Nature* 1996, **383:**829-832.
9. Amsterdam A, Burgess S, Golling G, Chen W, Sun Z, Townsend K, Farrington S, Haldi M, Hopkins N: **A large-scale insertional mutagenesis screen in zebrafish.** *Genes Dev* 1999, **13:**2713-2724.
10. Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, Hopkins N: **Identification of 315 genes essential for early zebrafish development.** *Proc Natl Acad Sci USA* 2004, **101:**12792-12797.
11. Golling G, Amsterdam A, Sun Z, Antonelli M, Maldonado E, Chen W, Burgess S, Haldi M, Artzt K, Farrington S, *et al.*: **Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development.** *Nat Genet* 2002, **31:**135-140.
12. Nasevicius A, Ekker SC: **Effective targeted gene 'knockdown' in zebrafish.** *Nat Genet* 2000, **26:**216-220.
13. Wienholds E, Plasterk RH: **Target-selected gene inactivation in zebrafish.** *Methods Cell Biol* 2004, **77:**69-90.
14. Driever W, Solnica-Krezel L, Schier AF, Neuhauss SC, Malicki J, Stemple DL, Stainier DY, Zwartkruis F, Abdelilah S, Rangini Z, *et al.*: **A genetic screen for mutations affecting embryogenesis in zebrafish.** *Development* 1996, **123:**37-46.
15. Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden FJ, Jiang YJ, Heisenberg CP, *et al.*: **The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*.** *Development* 1996, **123:**1-36.
16. Talbot WS, Hopkins N: **Zebrafish mutations and functional analysis of the vertebrate genome.** *Genes Dev* 2000, **14:**755-762.
17. Amsterdam A, Hopkins N: **Mutagenesis strategies in zebrafish for identifying genes involved in development and disease.** *Trends Genet* 2006, **22:**473-478.
18. Jaenisch R: **Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus.** *Proc Natl Acad Sci USA* 1976, **73:**1260-1264.
19. Lin S, Gaiano N, Culp P, Burns JC, Friedmann T, Yee JK, Hopkins N: **Integration and germ-line transmission of a pseudotyped retroviral vector in zebrafish.** *Science* 1994, **265:**666-669.
20. Linney E, Udvadia AJ: **Construction and detection of fluorescent, germline transgenic zebrafish.** *Methods Mol Biol* 2004, **254:**271-288.
21. Ellingsen S, Laplante MA, Konig M, Kikuta H, Furmanek T, Hoivik EA, Becker TS: **Large-scale enhancer detection in the zebrafish genome.** *Development* 2005, **132:**3799-3811.
22. Balciunas D, Davidson AE, Sivasubbu S, Hermanson SB, Welle Z, Ekker SC: **Enhancer trapping in zebrafish using the Sleeping Beauty transposon.** *BMC Genomics* 2004, **5:**62.

23. Parinov S, Kondrichin I, Korzh V, Emelyanov A: **Tol2 transposon-mediated enhancer trap to identify developmentally regulated zebrafish genes in vivo.** *Dev Dyn* 2004, **231:**449-459.
24. Wadman SA, Clark KJ, Hackett PB: **Fishing for answers with transposons.** *Mar Biotech (NY)* 2005, **7:**135-141.
25. Kawakami K, Takeda H, Kawakami N, Kobayashi M, Matsuda N, Mishina M: **A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish.** *Dev Cell* 2004, **7:**133-144.
26. Bellen HJ, Levis RW, Liao G, He Y, Carlson JW, Tsang G, Evans-Holm M, Hiesinger PR, Schulze KL, Rubin GM, *et al.*: **The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes.** *Genetics* 2004, **167:**761-781.
27. Balciunas D, Ekker SC: **Trapping fish genes with transposons.** *Zebrafish* 2005, **1:**335-341.
28. Kotani T, Nagayoshi S, Urasaki A, Kawakami K: **Transposon-mediated gene trapping in zebrafish.** *Methods* 2006, **39:**199-206.
29. Sivasubbu S, Balciunas D, Davidson AE, Pickart MA, Hermanson SB, Wangensteen KJ, Wolbrink DC, Ekker SC: **Gene-breaking transposon mutagenesis reveals an essential role for histone H2afza in zebrafish larval development.** *Mech Dev* 2006, **123:** 513-529.
30. Wheelan SJ, Aizawa Y, Han JS, Boeke JD: **Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution.** *Genome Res* 2005, **15:**1073-1078.
31. Amsterdam A: **Insertional mutagenesis in zebrafish.** *Dev Dyn* 2003, **228:**523-534.
32. Amsterdam A, Becker TS: **Transgenes as screening tools to probe and manipulate the zebrafish genome.** *Dev Dyn* 2005, **234:**255-268.
33. **Znomics, Inc. homepage** [http://www.znomics.com/]
34. Wu X, Li Y, Crise B, Burgess SM: **Transcription start regions in the human genome are favored targets for MLV integration.** *Science* 2003, **300:**1749-1751.
35. Wu X, Burgess SM: **Integration target site selection for retroviruses and transposable elements.** *Cell Mol Life Sci* 2004, **61:** 2588-2596.
36. Sun Z, Hopkins N: **vhnf1, the MODY5 and familial GCKD-associated gene, regulates regional specification of the zebrafish gut, pronephros, and hindbrain.** *Genes Dev* 2001, **15:** 3217-3229.
37. Yan YL, Miller CT, Nissen RM, Singer A, Liu D, Kirn A, Draper B, Willoughby J, Morcos PA, Amsterdam A, *et al.*: **A zebrafish sox9 gene required for cartilage morphogenesis.** *Development* 2002, **129:**5065-5079.
38. Maldonado E, Hernandez F, Lozano C, Castro ME, Navarro RE: **The zebrafish mutant vps18 as a model for vesicle-traffic related hypopigmentation diseases.** *Pigment Cell Res* 2006, **19:**315-326.
39. Chen W, Burgess S, Golling G, Amsterdam A, Hopkins N: **High-throughput selection of retrovirus producer cell lines leads to markedly improved efficiency of germ line-transmissible insertions in zebra fish.** *J Virol* 2002, **76:**2192-2198.
40. Keng VW, Yae K, Hayakawa T, Mizuno S, Uno Y, Yusa K, Kokubu C, Kinoshita T, Akagi K, Jenkins NA, *et al.*: **Region-specific saturation germline mutagenesis in mice using the Sleeping Beauty transposon system.** *Nat Methods* 2005, **2:**763-769.
41. Stanford WL, Cohn JB, Cordes SP: **Gene-trap mutagenesis: past, present and beyond.** *Nat Rev Genet* 2001, **2:**756-768.
42. Gossler A, Joyner AL, Rossant J, Skarnes WC: **Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes.** *Science* 1989, **244:**463-465.
43. Wurst W, Rossant J, Prideaux V, Kownacka M, Joyner A, Hill DP, Guillemot F, Gasca S, Cado D, Auerbach A, *et al.*: **A large-scale gene-trap screen for insertional mutations in developmentally regulated genes in mice.** *Genetics* 1995, **139:**889-899.
44. 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 USA* 1996, **93:**1677-1682.
45. 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 (Tokyo)* 1993, **113:**343-349.
46. 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:**277-287.

47. 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:**326-333.
48. Zambrowicz BP, Friedrich GA: **Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse.** *Int J Dev Biol* 1998, **42:**1025-1036.
49. 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.
50. Zambrowicz BP, Abuin A, Ramirez-Solis R, Richter LJ, Piggott J, BeltrandelRio H, Buxton EC, Edwards J, Finch RA, Friddle CJ, *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 USA* 2003, **100:**14109-14114.
51. McClive P, Pall G, Newton K, Lee M, Mullins J, Forrester L: **Gene trap integrations expressed in the developing heart: insertion site affects splicing of the PT1-ATG vector.** *Dev Dyn* 1998, **212:**267-276.
52. Sam M, Wurst W, Kluppel M, Jin O, Heng H, Bernstein A: **Aquarius, a novel gene isolated by gene trapping with an RNA-dependent RNA polymerase motif.** *Dev Dyn* 1998, **212:** 304-317.
53. Voss AK, Thomas T, Gruss P: **Efficiency assessment of the gene trap approach.** *Dev Dyn* 1998, **212:**171-180.
54. Soriano P: **Generalized lacZ expression with the ROSA26 Cre reporter strain.** *Nat Genet* 1999, **21:**70-71.
55. 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.
56. Schnutgen F, De-Zolt S, Van Sloun P, Hollatz M, Floss T, Hansen J, Altschmied J, Seisenberger C, Ghyselinck NB, Ruiz P, *et al.*: **Genomewide production of multipurpose alleles for the functional analysis of the mouse genome.** *Proc Natl Acad Sci USA* 2005, **102:**7221-7226.
57. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, *et al.*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409:**860-921.
58. Sterner DA, Carlo T, Berget SM: **Architectural limits on split genes.** *Proc Natl Acad Sci USA* 1996, **93:**15081-15085.
59. Thibault ST, Singer MA, Miyazaki WY, Milash B, Dompe NA, Singh CM, Buchholz R, Demsky M, Fawcett R, Francis-Lang HL, *et al.*: **A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac.** *Nat Genet* 2004, **36:**283-287.
60. Koga A, Suzuki M, Inagaki H, Bessho Y, Hori H: **Transposable element in fish.** *Nature* 1996, **383:**30.
61. Balciunas D, Wangensteen KJ, Wilber A, Bell JB, Sivasubbu S, Wang X, Hackett PB, Largaespada DA, McIvor RS, Ekker SC: **Harnessing a high cargo-capacity transposon for genetic applications in vertebrates.** *PLoS Genet* 2006, **2:**e169 doi:10.1371/journal.pgen.0020169.
62. Urasaki A, Morvan G, Kawakami K: **Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition.** *Genetics* 2006, **174:** 639-649.
63. Emelyanov A, Gao Y, Naqvi NI, Parinov S: **Trans-kingdom transposition of the maize dissociation element.** *Genetics* 2006, **174:**1095-1104.
64. Dupuy AJ, Fritz S, Largaespada DA: **Transposition and gene disruption in the male germline of the mouse.** *Genesis* 2001, **30:** 82-88.
65. Horie K, Kuroiwa A, Ikawa M, Okabe M, Kondoh G, Matsuda Y, Takeda J: **Efficient chromosomal transposition of a Tc1/mariner-like transposon Sleeping Beauty in mice.** *Proc Natl Acad Sci USA* 2001, **98:**9191-9196.
66. Xin HB, Deng KY, Shui B, Qu S, Sun Q, Lee J, Greene KS, Wilson J, Yu Y, Feldman M, *et al.*: **Gene trap and gene inversion methods for conditional gene inactivation in the mouse.** *Nucleic Acids Res* 2005, **33:**e14.
67. Wang D, Jao LE, Zheng N, Dolan K, Ivey J, Zonies S, Wu X, Wu K, Yang H, Meng Q, Zhu Z, Zhang B, Lin S, Burgess SM: **Efficient genome-wide mutagenesis of zebrafish genes by retroviral insertions.** *Proc Natl Acad Sci USA* 2007, **104:**12428-12433.