nature genetics

Toward simpler and faster genome-wide mutagenesis in mice

Sen Wu[1], Guoxin Ying[2], Qiang Wu[2] & Mario R Capecchi[1,2]

Here we describe a practical Cre-*lox*P and *piggyBac* transposon–based mutagenesis strategy to systematically mutate coding sequences and/or the vast noncoding regions of the mouse genome for large-scale functional genomic analysis. To illustrate this approach, we first created *lox*P-containing loss-of-function alleles in the protocadherin α, β and γ gene clusters (*Pcdha*, *Pcdhb* and *Pcdhg*). Using these alleles, we show that, under proper guidance, Cre-*lox*P site-specific recombination can mediate efficient *trans*-allelic recombination *in vivo*, facilitating the generation of large germline deletions and duplications including deletions of *Pcdha*, and *Pcdha* to *Pcdhb*, simply by breeding (that is, at frequencies of 5.5%–21.6%). The same breeding method can also generate designed germline translocations between nonhomologous chromosomes at unexpected frequencies of greater than 1%. By incorporating a *piggyBac* transposon to insert and to distribute *lox*P sites randomly throughout the mouse genome, we present a simple but comprehensive method for generating genome-wide deletions and duplications, in addition to insertional loss-of-function and conditional rescue alleles, again simply by breeding.

An international effort to generate knockouts of all mouse genes has been initiated[1–3]; however, this effort will concentrate on coding regions, which represent about 2.5% of the genome. The remaining 97.5% noncoding region is often referred to as 'junk DNA'. On the basis of comparisons among the newly sequenced mammalian genomes and partial sequencing of other vertebrate genomes, more than 300,000 conserved noncoding elements (CNEs; also referred to as conserved nongenic sequences or CNGs) have been identified in this presumed junk DNA. Many of these CNEs show greater sequence conservation among disparate vertebrate species than does the average protein-coding sequence[4–10]. Increasing evidence suggests that such noncoding regions have important regulatory roles, particularly in the modulation of genes controlling development. Mutations in these regions often cause significant disease phenotypes. To assess their functions directly, however, it is unrealistic to generate specific deletions of all CNEs in the mouse. A more practical approach to dissect the functional roles of such noncoding regions would be to systematically generate relatively large deletions of up to several hundred kilobase pairs that encompass several CNEs.

In *Drosophila melanogaster*, the transposon-based gene-trap has been used to generate a large collection of *FRT*-bearing alleles, enabling investigators to use Flp-*FRT* site-specific recombination to mediate *trans*-allelic recombinations *in vivo* between homologous chromosomes to generate large deletions and duplications covering the whole genome[11,12]. In the mouse, an *in vitro* Cre-*lox*P–based method in embryonic stem (ES) cells has been used to generate megabase deletions and duplications[13,14]. This *in vitro* protocol is, however, very labor-intensive and requires many rounds of ES cell genomic manipulations. An *in vivo* Cre-*lox*P method (designated TAMERE), which uses the *Sycp1*-Cre driver and is based on homologous chromosome pairing during meiosis, has been used to generate *trans*-allelic recombination in mice[15]. Although this method has been successful in generating deletions and duplications in the closely linked *Hoxd* genes[16], it has produced relatively small deletions[17,18]. By taking advantages of naturally occurring crossover to bring two *lox*P-containing alleles onto the same chromosome, an alternative method (designated STRING) has been described that uses Cre-mediated recombination to generate deletions and inversions of several megabases[19]. Because STRING is dependent on using naturally occurring homologous recombination–induced crossovers to bring the *lox*P site to the same chromosome, this methodology is limited to *lox*P sites that are separated on homologous chromosomes by several megabase pairs to ensure a reasonable crossover frequency. What would be useful is a single methodology that would permit efficient generation of Cre-*lox*P–based deletions and/or duplications, that would be independent of whether the *lox*P sites are on the same or separate chromosome homologs, and that would be applicable continuously for the generation of deletions and/or duplications ranging from a few thousand base pairs to several megabase pairs. This methodology is described herein.

We start with five distinct *lox*P-containing knockout alleles in the *Pcdh* gene clusters and a powerful Cre driver, and we show that with this system highly efficient *trans*-allelic recombination between homologous chromosomes occurs in somatic cells and in the germ line. We use this methodology to generate germline alleles containing large deletions and duplications, including deletion of the α cluster (228 kb)

[1]Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah 84112, USA. [2]Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112, USA. Correspondence should be addressed to M.R.C. (mario.capecchi@genetics.utah.edu).

Received 26 February; accepted 4 May; published online 17 June 2007; doi:10.1038/ng2060

and deletion of the α to β cluster (730 kb), with very high efficiencies. We further show that germline translocations between nonhomologous chromosomes bearing *lox*P sites can be similarly generated by simple breeding, again at very high frequencies.

The *piggyBac* transposon derived from the cabbage looper moth, *Trichoplusia ni*, has been shown to transpose very efficiently in mammalian cells and in the mouse[20]. To extend the Cre-*lox*P–mediated methodology for generating deletions and/or duplications to the whole mouse genome, we have also designed a *piggyBac* transposon–based gene-trap strategy that not only provides the means for very efficient introduction of *lox*P sites throughout the mouse genome by breeding, but also allows *in vivo* generation of genome-wide, sophisticated gene-trap alleles that complement the more conventional gene knockout strategy.

RESULTS

Generation of *Pcdh* mutant alleles

The clustered *Pcdh* genes provide an ideal locus to test methods for generating genomic manipulations such as large deletions and duplications. This unusual locus (**Fig. 1a**) in the mouse contains 58 very similar genes, encoding *trans*-plasma membrane adhesion molecules, that are arranged into three sequentially linked clusters (*Pcdha*, *Pcdhb* and *Pcdhg*, also known as *Pcdhα*, *Pcdhβ* and *Pcdhγ*, respectively), spanning a region of about 1 Mb of DNA[21–23]. The protein products from this locus are generally localized to synaptic junctions in the central nervous system[24–26], and have been proposed to be important in the establishment and maintenance of neuronal connections. The α and γ proteins are thought to have distinct intracellular signaling pathways, owing to their highly divergent intracellular domains[21]. A γ cluster deletion[25], generated by *in vitro* Cre-*lox*P–mediated recombination[13,14], and a γ constant region deletion[27,28] have been shown to cause neonatal lethality in the mouse. Initial analysis of these mice suggests that the γ cluster is required for synapse development, and for survival of interneurons in the spinal cord and specific neurons in the brain[25,27]. However, the function of the α and β clusters, or that of any individual *Pcdh* gene, remains undefined.

To dissect systematically the function of the α and β clusters of the *Pcdh* genes, we created five knockout mouse lines (**Fig. 1a** and **Supplementary Fig. 1** online) in the *Pcdh* clusters. In the *Pcdha1*^{eGFP} allele, the *Pcdha1* variable region was replaced in frame by an enhanced green fluorescent protein (eGFP) reporter gene. In the *Pcdhac1*^{AP} allele, an alkaline phosphatase (*AP*) knock-in replaced in frame the *Pcdhac1* variable exon. In the *Pcdhac2*^{lacZ} allele, a *lacZ* knock-in replaced the *Pcdhac2* variable exon. Each of these alleles permitted determination and internal comparisons of individual *Pcdha* gene expression patterns in the embryonic or adult mouse. The *Pcdha* conditional allele was created by flanking the first two constant exons with *lox*P sites. Cre-mediated recombination of this conditional allele generated the *Pcdha*^{Con} allele, which contained a deletion of α constant exons 1 and 2. In the *Pcdhga1*^{hrGFP} allele, an hrGFP knock-in replaced the *Pcdhga1* variable exon. Each allele was confirmed to carry the prescribed genetic alterations by DNA blot analysis (**Fig. 1b**). Genotyping of these alleles was performed by PCR (**Supplementary Table 1** online). Because all of these alleles contained

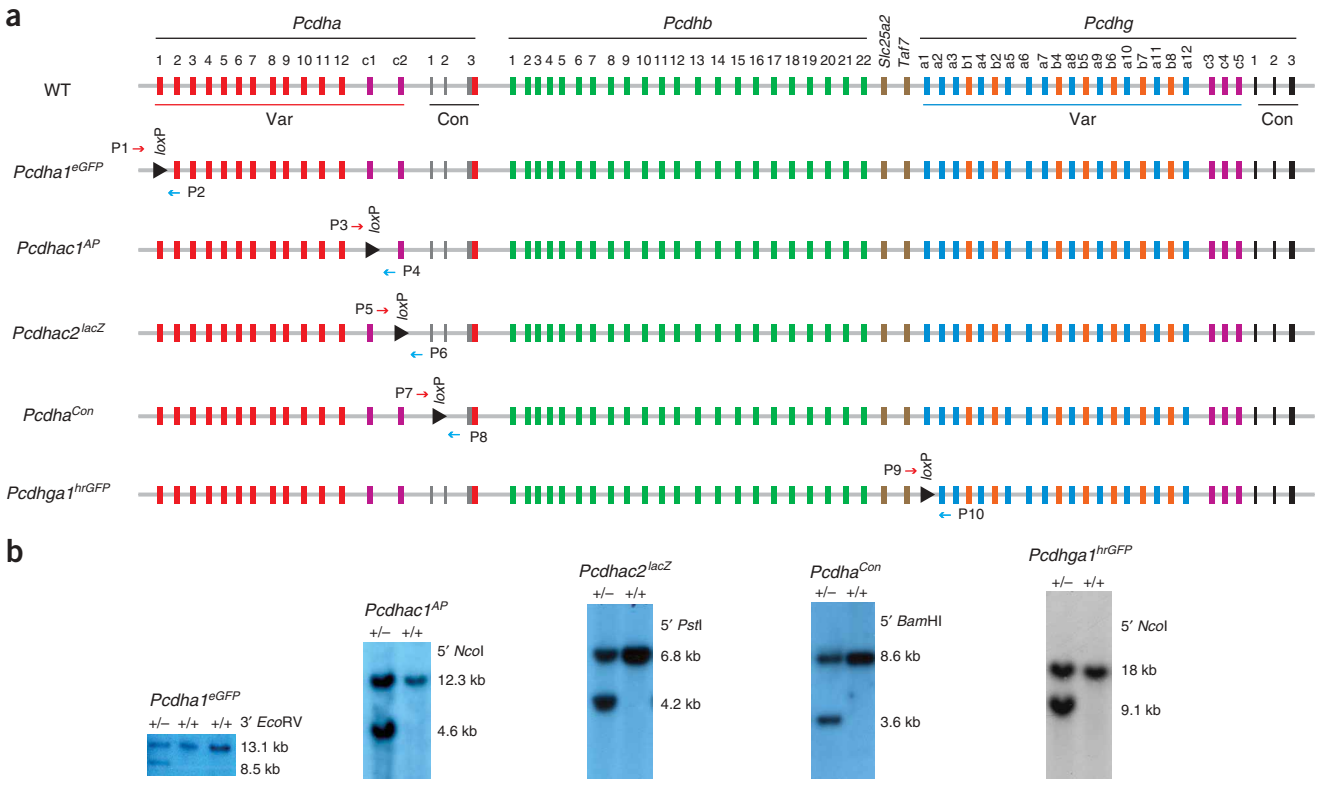


Figure 1 Generation of mouse *Pcdh* alleles. (**a**) The wild-type mouse *Pcdh* clusters and the five gene-targeted mutant alleles. In the wild-type locus, the α and γ clusters share similar genomic structures: each of the 14 α and 22 γ variable exons are separately spliced, respectively, to the three α and γ constant exons; by contrast, the β cluster has no constant exons. All of the five mutant alleles, designated *Pcdha1*^{eGFP}, *Pcdhac1*^{AP}, *Pcdhac2*^{lacZ}, *Pcdha*^{Con} and *Pcdhga1*^{hrGFP} (see text and **Supplementary Fig. 1** for details), carry a *lox*P site, whose orientation is indicated by a black triangle. Small arrows indicate the positions of the PCR primers used to characterize the mutant alleles. (**b**) DNA blot transfer analysis confirmed the gene structure of each of the five mutant alleles.

a *lox*P site, they were used to identify an efficient Cre driver for *in vivo* generation of large deletions and duplications.

Each of the above described *Pcdh* alleles was bred to homozygosity in mice. Each genotype was obtained at the expected mendelian ratio, and homozygotes were viable and fertile. The phenotype of these mice is being analyzed in detail for changes in neural circuitry and behavior. These results will be reported separately.

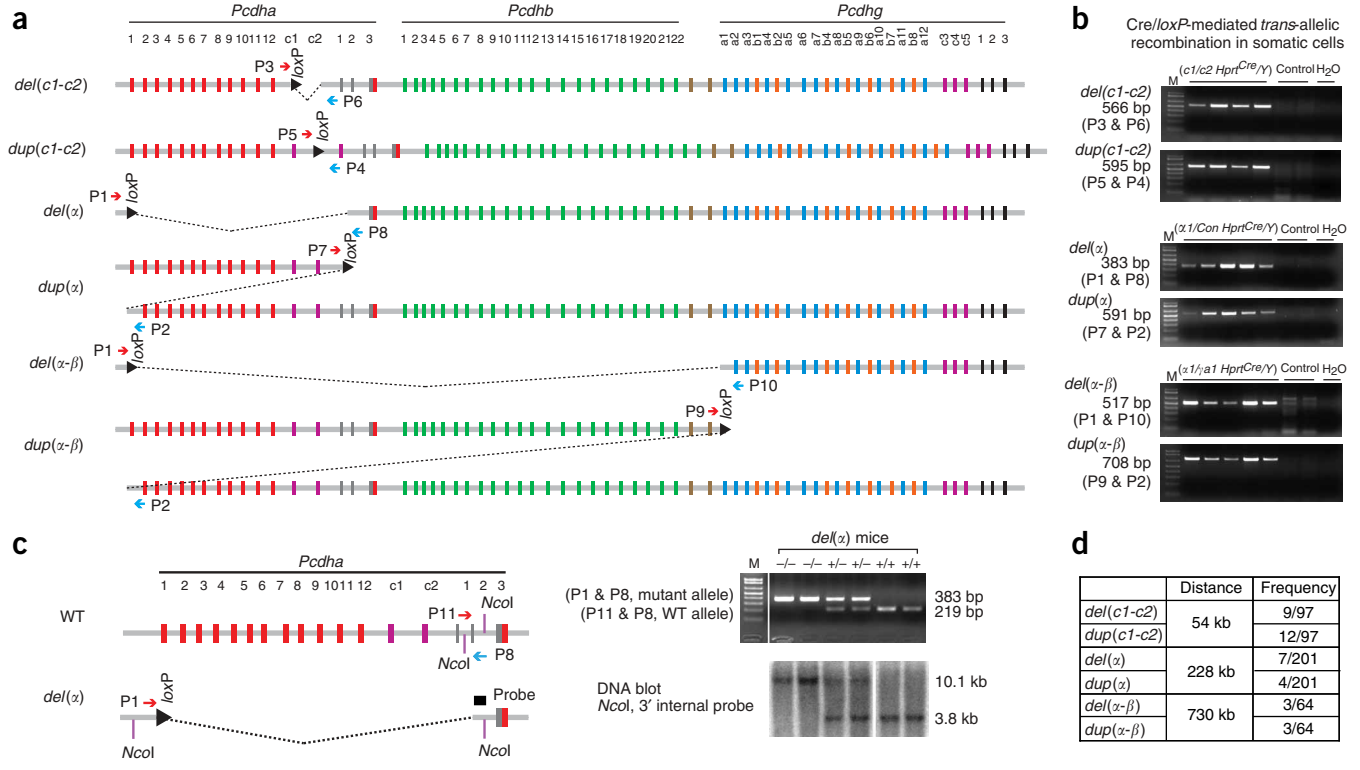
Somatic *trans*-allelic Cre-*lox*P recombination *in vivo*

Because Cre recombinase activity is dose dependent, we reasoned that a strong, constitutively expressed *Cre* transgene might be able to drive *trans*-allelic recombination efficiently during both mitotic and meiotic cell divisions. The *Hprt*-Cre driver[29], in which a strong CAG promoter–driven Cre expression cassette is inserted into the X-linked *Hprt* locus, seemed to be a good candidate for this purpose. To test this hypothesis, we first crossed mice carrying the five *Pcdh* alleles, which all carried *lox*P sites with the same chromosomal orientation, with mice carrying *Hprt*-Cre to generate compound heterozygous males that contained the following sets of alleles: $Pcdhac1^{AP/+}Pcdhac2^{lacZ/+}Hprt^{Cre}$/Y, $Pcdha1^{eGFP/+}Pcdha^{Con/+}Hprt^{Cre}$/Y and $Pcdha1^{eGFP/+}Pcdhga1^{hrGFP/+}Hprt^{Cre}$/Y. We then analyzed tail DNA from these mice by PCR to detect the corresponding Cre-mediated deletion and duplication alleles. We demonstrated that each mouse had undergone the predicted Cre-*lox*P–mediated *trans*-allelic recombination event in somatic cells, and that both the deletion and duplication alleles were present (**Fig. 2a**,**b**). We sequenced all of the PCR products and confirmed the presence of the predicted junction sequences generated by the Cre-mediated site-specific recombination events. Although these experiments did not yield the frequencies of the long-range *trans*-allelic recombination events, they encouraged us to test for germline transmission of these events.

Large deletions and duplications in germ line

Heterozygous males containing the above combinations of *Pcdh* alleles and *Hprt*-Cre were mated to wild-type C57BL/6J females. We detected germline transmission of all six recombined alleles, *del(c1–c2)*, *dup(c1–c2)*, *del(α)*, *dup(α)*, *del(α–β)* and *dup(α–β)*, in the offspring at high frequencies (**Fig. 2a**,**c**,**d**). The Cre-mediated *trans*-chromosomal recombination frequency between *lox*P sites separated by the shortest distance (54 kb, between $Pcdhac1^{AP}$ and $Pcdhac2^{lacZ}$) yielded the highest frequency (~21.6%; **Fig. 2d**). Even when the *lox*P sites were separated by over 700 kb, as in the $Pcdha1^{eGFP}$ and $Pcdhga1^{hrGFP}$ alleles, the *trans*-chromosomal recombination frequency was still high (~9.3%). These results suggest that large genome-wide deletions and duplications ranging from tens to hundreds of kilobase pairs can be similarly generated by this simple breeding procedure.

Unexpectedly, mice homozygous for the *del(α)* allele were viable and fertile with no apparent gross phenotype. Histochemical analysis



	Distance	Frequency
del(c1-c2)	54 kb	9/97
dup(c1-c2)		12/97
del(α)	228 kb	7/201
dup(α)		4/201
del(α-β)	730 kb	3/64
dup(α-β)		3/64

Figure 2 Generation of large deletions and duplications by Cre-*lox*P–mediated *trans*-allelic recombination. The $Pcdha1^{eGFP}$, $Pcdhac1^{AP}$, $Pcdhac2^{lacZ}$, $Pcdha^{Con}$ and $Pcdhga1^{hrGFP}$ alleles were used for *in vivo* generation of large deletions and duplications. (**a**) Genomic structures of deletions and duplications that can be generated by Cre-*lox*P–mediated *trans*-allelic recombination. (**b**) Cre-*lox*P mediates efficient *trans*-allelic recombination between chosen *Pcdh* mutant alleles in somatic cells. Tail DNA from 4-week-old mice with different combinations of alleles were subjected to PCR analysis to detect the corresponding deletion and duplication alleles. *c1/c2* $Hprt^{Cre}$/Y corresponds to $Pcdhac1^{AP/+}Pcdhac2^{lacZ/+}Hprt^{Cre}$/Y (top); *α1/Con* $Hprt^{Cre}$/Y corresponds to $Pcdha1^{eGFP/+}Pcdha^{Con/+}Hprt^{Cre}$/Y (center) and *α1/γa1* $Hprt^{Cre}$/Y corresponds to $Pcdha1^{eGFP/+}Pcdhga1^{hrGFP/+}Hprt^{Cre}$/Y (bottom). The predicted deletion and duplication alleles were detected in tail DNA of each male containing the appropriate *Pcdh* *lox*P alleles and *Hprt*-Cre. (**c**) As an example, a germline transmission of the recombined deletion allele *del(α)* was confirmed by both PCR and DNA blot in 3-week-old F_2 mice generated by intercrossing heterozygous *del(α)* mice. (**d**) Frequency of germline transmission of the corresponding duplication and deletion alleles.

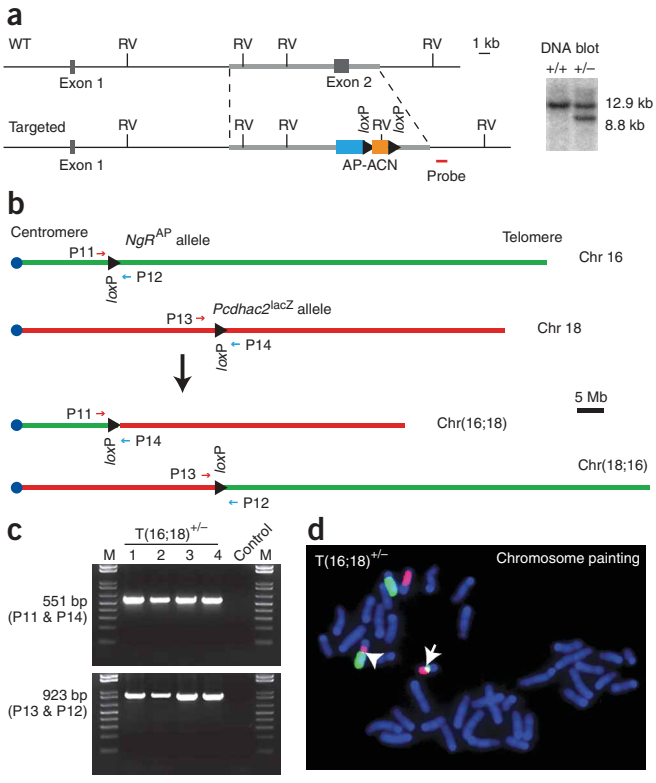


Figure 3 Generation of a Cre-*lox*P–mediated germline translocation between nonhomologous chromosomes. (**a**) Gene targeting in the *NgR* locus. To generate a loss-of-function allele of the *NgR* gene, an *AP* gene reporter self-excising neo selection cassette (ACN)[47] replaced more than 95% of the amino acids of NgR. DNA blot transfer analysis, using *Eco*RV (RV) digest and a 3′ flanking probe, identified a 12.9-kb band for wild type and a predicted 8.8-kb band for the targeted allele. (**b**) Both *lox*P sites in the NgR^{AP} allele and the $Pcdhac2^{lacZ}$ alleles have the same orientation relative to their centromeres. Cre-*lox*P–mediated recombination between these loci generated a translocation between these chromosomes. (**c**) Germline transmission of reciprocal T(16;18) translocations was confirmed by PCR. Lanes 1 and 2, two F_1 adult mice; lanes 3 and 4, two 2-week-old F_2 offspring from an F_1 mouse crossed to wild type. These four mice are heterozygous for the reciprocal translocation T(16;18). (**d**) Chromosome painting of metaphase chromosomes prepared from heterozygous embryonic day 16 (E16) F_2 fibroblast culture. Green, chromosome 16; red, chromosome 18; blue, 4′,6′-diamidino-2-phenylindole (DAPI). Arrow indicates Chr(16;18); arrowhead indicates Chr(18;16).

with Nissl staining revealed apparently normal gross anatomy of the mutant mouse brain (**Supplementary Fig. 2** online). For example, the thickness and layering of the cerebral cortex appeared normal. Detailed analysis of the long-range deletions and duplications *del(α)*, *dup(α)*, *del(α–β)* and *dup(α–β)* will be reported separately.

Germline generation of chromosomal translocations

To examine whether the same breeding strategy can also generate germline translocations between nonhomologous chromosomes, we bred mice carrying the $Pcdhac2^{lacZ}$ allele on chromosome 18 to mice carrying an NgR^{AP} allele on chromosome 16 (**Fig. 3a**), in which the gene encoding the Nogo receptor (*NgR*; official symbol *Rtn4r*) was replaced in frame by an *AP* reporter gene, and to mice carrying *Hprt*-Cre to produce compound heterozygous males ($Pcdhac2^{lacZ/+}NgR^{AP/+}Hprt^{Cre}/Y$). The $Pcdhac2^{lacZ}$ and NgR^{AP} alleles each carried *lox*P sites in the same chromosomal orientation relative to the centromeres. We then crossed these males to C57BL/6J wild-type females. Out of 182 offspring, 2 were shown to inherit a germline transmission of the balanced reciprocal T(16;18) translocations (**Fig. 3b**,**c**). The two mice, in which 81.6% of chromosome 16 was swapped with 58.9% of chromosome 18 and vice versa, appeared normal, healthy and fertile. In living mice, it seems that the reciprocal translocations always co-segregate together, because nonreciprocal transmission of these translocation products was not detected in the next generation (**Fig. 3c**,**d**). Among 27 F_2 offspring generated from crosses of F_1 to wild-type mice, 14 reciprocal translocations were detected, but mice containing individual translocations were not detected. This result suggests that being either haploid or trisomic for these portions of chromosome 16 and 18 causes embryonic lethality.

To ensure that the high frequency of the above chromosomal translocation was not an attribute of the particular chromosomes or loci chosen for these experiments, we arbitrarily chose to repeat the experiment with two other chromosomes (6 and 11) and two other loci, for which we had *lox*P site insertions in the appropriate orientation relative to the centromere (namely, the *Hoxa* and *Hoxb* complexes). Compound heterozygous males containing $Hoxa5^{loxP/+}Hoxb5^{loxP/+}Hprt^{Cre}/Y$ alleles were crossed to C57BL/6J wild-type females, and 1.1% of the newborns (F_1) were detected to contain germline transmission of the predicted reciprocal translocations, which again seemed to co-segregate even in the F_2 generation (data not shown).

Gene-trap using the *piggyBac* transposon

The generation of genome-wide chromosome rearrangements requires *lox*P sites distributed throughout the mouse genome. We considered that the efficient transposition of the *piggyBac* transposon in the mouse germ line[20] offered an ideal method for achieving this goal, as has been shown in *D. melanogaster*[30,31]. We therefore designed a *piggyBac* gene-trap strategy that would produce multipurpose loss-of-function and conditional rescue alleles, and that would provide a broad distribution of *lox*P sites in the mouse genome (**Fig. 4**).

To maximize the utility of the *piggyBac* gene-trap vector, we first defined which transposon sequences would be required for efficient transposition, and which sequences could be deleted to increase cargo size and to abolish, simultaneously, cryptic RNA splicing through transposon sequences. On the basis of this consideration, three gene-trap vectors were constructed: ZG-l, ZG-m and ZG-s (**Fig. 4a**). ZG-l contained the 63 bp of 3′ terminal repeats with the adjacent 932 bp of internal sequences, and 35 bp of 5′ terminal repeats with the adjacent 645 bp of internal sequences. ZG-m differed from ZG-l in the exclusion of 542 bp of 3′ internal sequence, which deleted all but one of the potential cryptic splice donor sites. ZG-s varied from ZG-m in that it lacked all potential splice donor sites and had a deletion of an additional 450 bp of internal sequences.

These gene-trap vectors were designed to generate multipurpose alleles (**Fig. 4b**). First, successful *piggyBac* transposition into an intron of an endogenous gene should produce a loss-of-function gene-trap allele. Second, it allows easy removal of the sequences flanked by the two *FRT* sites by Flp recombinase to generate the Flp allele. In contrast to the gene-trap allele, this Flp allele can be bred to any Cre driver without concern for unwanted germline Cre-mediated recombination involving these alleles, which is an important requirement in some biological experiments. Third, because the splice acceptor in the gene-trap allele is flanked by two *lox*P sites, Cre-mediated recombination of this allele should generate a wild-type revertant allele (Cre allele),

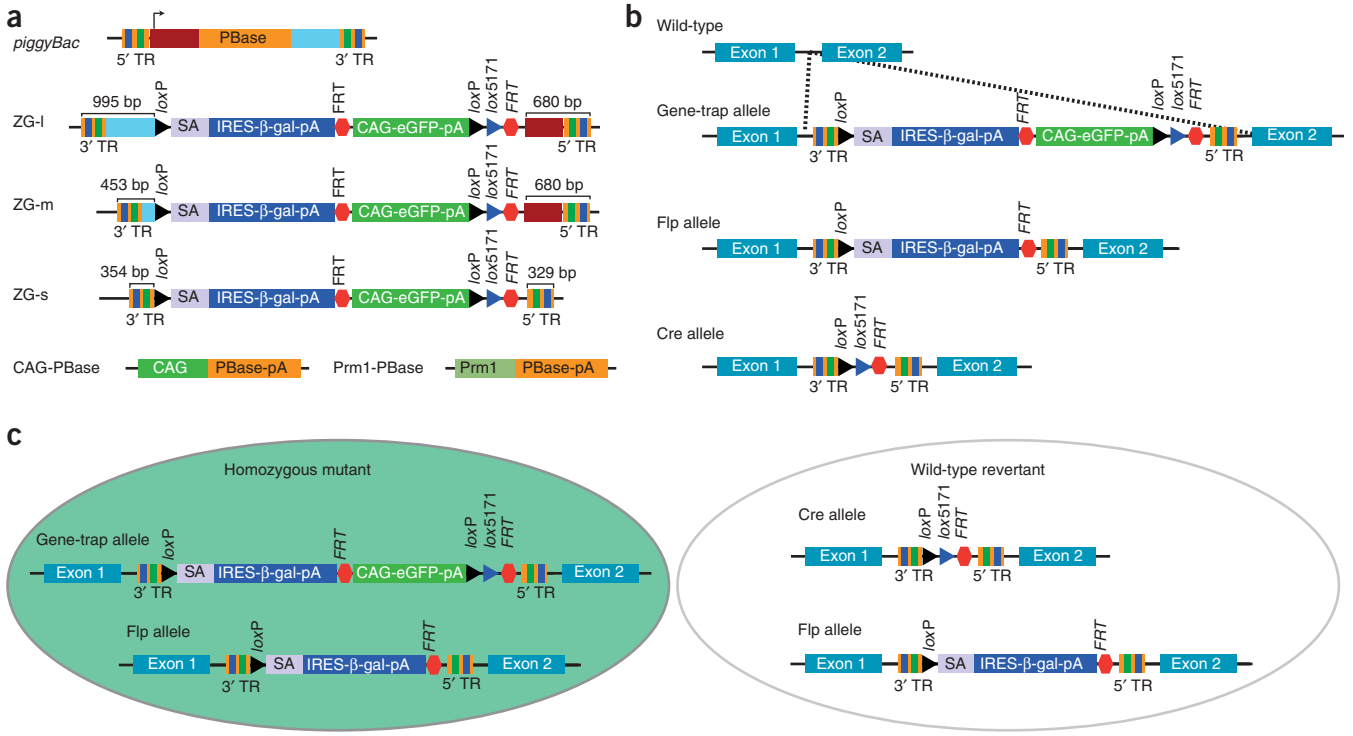


Figure 4 Optimization of *piggyBac* for multipurpose gene-trapping. (**a**) Three gene-trap vectors (ZG-l, ZG-m and ZG-s) and two transposase constructs (CAG-PBase and Prm1-PBase). (**b**) Successful *piggyBac* transposition into an intron of an endogenous gene produces a gene-trap allele. The presence of *lox*P and *FRT* sites facilitates further *in vivo* manipulation of this gene-trap allele. (**c**) Gene-trap alleles can be used as conditional rescue alleles.

because the remaining sequence at the locus, without a splice acceptor site, is very likely to be innocuous. Thus, the gene-trap allele can be interrogated conditionally by reverting the loss-of-function allele to wild-type function in chosen tissues by the presence of Cre. We designated these alleles as 'conditional rescue alleles' (**Fig. 4b**,**c**).

To compare the three gene-trap constructs in the mouse, we created transgenic mice through pronuclear injection of circular plasmid containing a strong CAG promoter–driven *piggyBac* transposase (CAG-PBase) plus circular plasmid containing the ZG-l, ZG-m or ZG-s sequence. Transgenic founder mice containing only ZG-l (two lines), ZG-m (three lines) or ZG-s (three lines) without CAG-PBase were obtained. All of these founders were fluorescent, as expected, because all three constructs carried a constitutive CAG promoter–driven *eGFP* gene. Sequencing of the inverse PCR products, generated from genomic DNA from each founder, confirmed that most of the integrations were precise *piggyBac*-mediated transpositions (**Supplementary Table 2** online). Because lack of internal sequences in ZG-s did not seem to affect the transposition efficiency, ZG-s, which has no cryptic splice donor sites, was chosen for all subsequent germline transposition studies. We also designed variations of ZG-s vector to expand its versatility and/or to limit its potential problem (**Supplementary Fig. 3** online).

To produce large numbers of *lox*P-bearing gene-trap alleles by germline transposition, we next created double transgenic mouse lines through pronuclear co-injection of ZG-s, and protamine (Prm1) promoter-driven *piggyBac* transposase (Prm1-PBase), which expresses *piggyBac* transposase in the male germline only[20,32]. Eighteen founder mice (12 males and 6 females) were obtained. After confirming that the Prm1 promoter was active only in male founders, we used female founders to maintain a line and produce more male founders. Male founders were used for crossing to wild-type C57BL/6J females

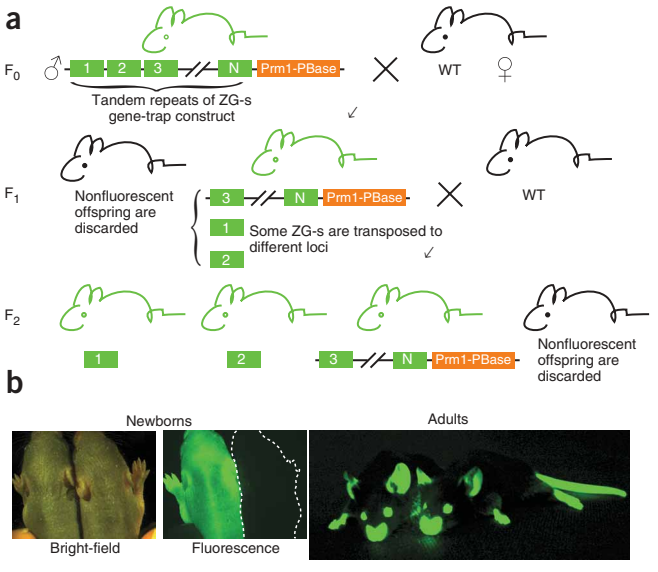


Figure 5 Germline transposition. (**a**) Breeding scheme to generate new *piggyBac* gene-trap alleles by germline transposition. Male founders (F_0) containing tandem ZG-s and Prm1-PBase (in the same locus) are bred to wild-type C57BL/6J females. New transposition events (shown as boxes 1 and 2) in the F_1 generation can be separated as F_1 gametes, and stable alleles are obtained in the F_2 generation. (**b**) Newborns and adults carrying a ZG-s gene-trap construct can be easily detected under fluorescence.

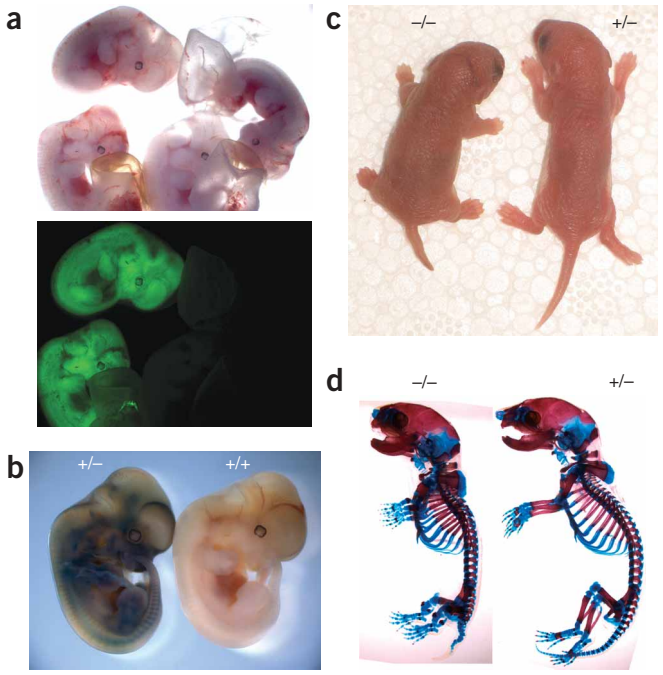


Figure 6 Gene-trap alleles recapitulate a conventional knockout phenotype. To verify that *piggyBac* gene-trap can generate null alleles, we used a *Ror2* gene-trap allele as an example. (**a**) E12 embryos containing *Ror2* gene-trap alleles are fluorescent, owing to the presence of the CAG-eGFP reporter in the ZG-s gene-trap construct. (**b**) A heterozygous embryo from **a** is stained for β-galactosidase and shows strong expression in developing bones and forebrain. (**c**) Postnatal day 0 mice homozygous for the *Ror2* gene-trap allele have shortened body, limbs and tail. (**d**) Skeleton preparation of the two mice in **c** further confirms extensive abnormalities in the bones.

to generate new transposition events. Because the Prm1 promoter is active only after meiosis, only those offspring (about half) that inherit the transgenes (Prm1-PBase and ZG-s) can contain new transpositions (**Fig. 5a**). The ZG-s gene-trap construct includes a constitutively expressed *eGFP* gene (**Fig. 4a**), which is a very convenient marker to identify which offspring may contain new transposition events (**Fig. 5b**). Results from inverse PCR confirmed and identified the new transposition events that occurred in germ line through simple breeding (**Supplementary Table 3** online). On average, we obtained one new transposition event per offspring per generation, which is consistent with previous results[20].

We next determined whether individual gene-trap alleles generated by breeding indeed functioned as multipurpose alleles. First, we found that the CAG-eGFP reporter was ubiquitously expressed during development in embryos containing gene-trap alleles (**Fig. 6a** and data not shown). Second, we carried out X-Gal staining on mice that were heterozygous for *piggyBac* vector insertions at the *Ror2*, *Runx2* (also known as *Cbfa1*) or *Slc14a2* loci (**Fig. 6b** and **Supplementary Fig. 4** online). The β-galactosidase expression patterns of these gene-trap alleles recapitulated the published expression profiles for these genes[33–35]. We further performed RT-PCR on mRNA isolated from mice carrying these alleles. Sequencing of these RT-PCR products confirmed that the endogenous splicing is hijacked by the splice acceptor in the ZG-s gene-trap construct. Third, to demonstrate that *piggyBac* gene-trapping does produce null alleles, we bred, as an example, a *Ror2* gene-trap allele to homozygosity in mice. In this allele, the insertion occurs in the second intron of the *Ror2* gene. The phenotype of homozygous *Ror2* gene-trap alleles perfectly reproduced the null phenotype reported by conventional gene knockout[33] (**Fig. 6c**,**d**). We also crossed this allele to a Flp deleter mouse and the CAG-eGFP reporter was efficiently recombined out, as expected, resulting in nonfluorescent mice and demonstrating that we could obtain a stable Flp allele (**Fig. 4b** and data not shown).

Fourth, to test whether our gene-trap alleles could be used as conditional rescue alleles, we crossed the *Ror2* gene-trap allele to the *Hprt*-Cre driver to obtain a Cre allele (**Fig. 4b**,**c**). Again, the Cre allele of *Ror2* functioned as a wild-type revertant allele, and mice homozygous for the Cre allele were nonfluorescent and phenotypically wild type (data not shown), showing that our gene-trap alleles can be used as conditional rescue alleles. Last, to test the possibility of using the *lox*P sites present in *piggyBac* gene-trap alleles containing the *Mdc1* and *Runx2 piggyBac* insertion mutations to generate large deletions and duplications, we bred mice carrying these gene-trap alleles to mice carrying *Hprt*-Cre. Among all of the founder mice that were positive for the two *trans*-allelic gene-trap alleles and *Hprt*-Cre, the predicted Cre-*lox*P–mediated *trans*-allelic recombination events (that is, both the deletion and the duplication alleles) were present in somatic cells, similar to what was observed for the *Pcdh* locus (**Supplementary Fig. 4**). The distance between these two loci on chromosome 17 is roughly 8.6 Mb.

DISCUSSION

Through breeding of mice possessing *lox*P-containing mutant alleles with the *Hprt*-Cre mouse, we have generated large germline deletions and duplications of the *Pcdh* locus. Deletions of >700 kb were obtained at a high frequency of ~9.3%. Translocations between nonhomologous chromosomes were also generated with efficiencies of ~1%. This frequency is interesting because we would anticipate that Cre-mediated recombination between *lox*P sites located on nonhomologous chromosomes would be lower, as we observed, than between *lox*P sites located on homologous chromosomes. This reasoning sets the lower limit of the frequency of generating any chosen chromosomal deletion or duplication by our methodology as >1%. This high frequency of recombination does not seem to be locus-specific, because *Hprt*-Cre generated translocations between other nonhomologous chromosomes at similar frequencies. A merit of this method is that we have not observed Cre-related cellular toxicity with the *Hprt*-Cre driver, as has been observed with other strong Cre drivers[32]. The mice carrying this Cre driver are stable, appear normal, live normal life spans and propagate this allele to their progeny with the expected mendelian frequency.

Many cancers, leukemias, lymphomas and sarcomas seem to be initiated by the presence of hybrid gene products produced by specific chromosomal translocations. Apart from one case[36], it has not been possible to model these cancers in the mouse directly through the generation of the appropriate chromosomal translocation by controlling, for example, when and in which cells the translocations are generated. The reason for this failure is that previously reported Cre-mediated chromosomal translocation occurs at a frequency of $\sim 10^{-6}$. Under these conditions, the pool of cells in the mouse that carry the chromosomal translocation is not sufficiently large to allow for spontaneous occurrence of the secondary genetic events that are normally required for tumor progression. With an improvement of four orders of magnitude in ability to generate such chromosomal translocations, direct modeling of these cancers in the mouse by generating the appropriate translocation becomes feasible.

We have described a gene-trap system based on the *piggyBac* transposon for mutagenesis in mice. This gene-trap system has the advantages that it has higher transposition frequency in mammals than previously described transposon systems and that, in the process of transposition to a new site, it does not leave behind a DNA 'scar', which could itself be mutagenic. Consistent with previous observations[20], the *piggyBac* transpositions seem to be randomly distributed throughout the mouse genome (**Supplementary Tables 2** and **3**). Many transpositions occur in intronic sequences, generating gene-trap alleles. Another advantage of the *piggyBac* gene-trap is that it uses breeding to generate new multipurpose alleles without *in vitro* ES cell manipulations or *in vivo* blastocyst injections. The number of insertional alleles can be rapidly expanded by breeding (**Fig. 5**) to generate a collection of mouse strains that could approach containment of null alleles in most genes in a relatively short period of time.

For example, how many mice would need to be generated to produce a collection of mice that carry at least one insertion in 15,000 different mouse genes? An estimate of this number can be readily calculated by using the Poisson distribution. From the known mouse genomic sequence, we can calculate the fraction of the genome occupied by genes (that is, the fraction occupied by gene 'footprints', defined as the intron and exon sequences covered by a gene plus 50 bp of sequence 5′ and 3′ to the first and last exons, respectively). On the basis of the annotated mouse genomic sequence, this number is 38.1%. Because many predominantly centromeric repetitive sequences remain to be annotated, however, 36% is presumed to be a better estimate of this fraction. The number of genes present in the mouse genome is also needed for the above calculation, and at present GenBank lists this number as 31,805. On the basis of annotation of the human genomic sequence and the fact that the criteria for gene composition of the mouse genome are not very stringent, this number is likely to be inflated. To compensate for the absence of a stringent number, we made two calculations based on the mouse genome containing 30,000 or 22,000 genes. Last, to permit use of the Poisson distribution to make this calculation, we need to assume that the insertions in the mouse genome are random and that the size of the gene footprint is the same for all genes, both of which are oversimplifications. Nonetheless, with these numbers the Poisson distribution suggests that the generation of mice containing at least one *piggyBac* insertion in 15,000 different genes would require breeding of 57,760 (if 30,000 genes) or 69,970 (if 22,000 genes) mice. If we started with ten breeding pairs, saved all of the progeny, and assumed a generation time of 12 weeks and an average litter size of seven (the average litter size for C57BL/6), then these numbers of mice could be generated in ~1.5 years.

Equally important, about half of the transpositions would be located in the intergenic regions. Therefore, the more or less randomly distributed *lox*P sites contained in the above large collection of mice could be further used to obtain genome-wide deletions and duplications within the vast noncoding regions of the genome. On the basis of these pilot experiments, we provide a framework for implementing a simple, fast and highly cost-effective large-scale mouse mutagenesis protocol.

We used the CAG promoter–driven *eGFP* reporter as a convenient marker in our *piggyBac* gene-trap vectors. In theory, this strong CAG promoter could interfere with local gene expression. This effect is not likely to be critical to the generation of loss-of-function insertion alleles, however, because by definition the locus is not functional and therefore alteration of its expression would not matter. Indeed, we have directly compared a mouse possessing the *piggyBac* insertion mutation with a mouse containing the corresponding gene-targeted mutation, and found that their mutant phenotypes were indistinguishable. Problems might arise if the CAG promoter interferes with the expression of other genes located near the insertion locus. There are a number of simple solutions to remedy this potential problem (**Supplementary Fig. 3**). One of the simplest is to invert the order of the *FRT* and *lox*P sites present on the right arms of the *piggyBac* insertion vector. With the inverted order, the eGFP cassette could be excised by the Flp recombinase without excising the *lox*P site, which therefore could be subsequently used to generate the conditional revertant allele.

As described, the *piggyBac* gene-trap alleles provide the capacity to conditionally interrogate the function of the disrupted allele as a 'rescuable' allele. Conventional conditional knockout alleles are 'on→off', whereby the genes are always on until Cre is provided in preselected tissues or during chosen temporal periods of the mouse, to excise the gene of interest in those restricted tissues and/or temporal periods. By contrast, the revertable gene-trap alleles generated by *piggyBac* transposition are 'off→on', whereby the genes are always off until Cre is selectively provided, thereby conferring the potential to rescue the mutant phenotype by activation of the gene in chosen tissues and/or during chosen temporal periods. This design should provide a useful and informative complement to the conventional on→off Cre-*lox*P–mediated conditional mutant approach, particularly for studies in development where timing is a crucial parameter. The ability to approach the same problem using both on→off and off→on conditional mutagenesis should provide informative insight. The CAG-eGFP cassette in the *piggyBac* insertion vector is an effective reporter for revealing the precise pattern and extent of excision achieved by the Cre driver chosen to revert the insertion mutant phenotype, because only wild-type cells will not be fluorescent. Gene-trap alleles may also offer a simple approach for mosaic analysis[37] because mutant cells are labeled by eGFP, whereas heterozygous cells are not (**Fig. 4c**), facilitating observation of mutant cells in an otherwise wild-type environment.

The *piggyBac* gene-trap could be used in large-scale screens for specific phenotypes, such as the generation of tumors in a sensitized mouse genetic background. The *piggyBac* gene-trap system should also facilitate loss-of-function studies for model organisms such as rats, chickens and zebrafish, where transgenic technology is available for the introduction of foreign DNA by DNA injection at single-cell stage or by retrovirus infections, but where ES cells capable of contributing to the germ line are lacking such that the organisms are inaccessible to gene targeting to modify their genome. Combined use of the Cre-*lox*P and *piggyBac* gene-trap systems should facilitate the generation of genome-wide deletions and duplications in these vertebrate models, as it does in the mouse.

METHODS

Creation of mutant *Pcdh* and *NgR* alleles. Targeting vectors were constructed from BAC clones containing the respective genomic regions. A streamlined modular cloning method involving a combination of λ phage-mediated recombination ('recombineering') and Gateway cloning was used (S.W., G.Y., Q.W. & M.R.C., unpublished). Standard ES cell manipulation and blastocyst injections were used to obtain germline knockout alleles. Mouse experiments were approved by the University of Utah Institutional Animal Care and Use Committee and were conducted strictly in accordance with the relevant protocol.

Chromosome painting. A detailed protocol is available from the authors on request. We derived fibroblasts from an embryonic day 16 embryo containing the reciprocal T(16;18) translocations. We prepared metaphase chromosomes according to a standard protocol[38]. We performed chromosome hybridization by using green chromosome-16 and red chromosome-18 paints in accordance with instructions from the manufacturer (Applied Spectral Imaging).

Generation of *piggyBac* transgenic mice. The full-length *piggyBac* element is a 2,472-bp autonomous transposon that is flanked by inverted repeats and encodes a functional transposase (PBase)[39,40]. It specifically inserts into a TTAA target site[40]. In addition to the 5′ and 3′ inverted terminal repeats, the PBase requires internal sequences for efficient transposition[41]; however, the minimal requirements for these internal sequences have not been determined in the mouse. Because the 5′ end of *piggyBac* has a residual promoter, we chose to put the 3′ end before a reporter gene in our gene-trap design. The 3′ end of *piggyBac* has a few cryptic splice donor sites, which could also compromise gene-trapping, as has been observed in *D. melanogaster*[30].

To construct the gene-trap vector, the *piggyBac* transposon 5′ and 3′ sequences were derived from the plasmid C4-PBss (see Acknowledgments). The internal ribosomal entry site (IRES) sequence was derived from pIRES2-eGFP (Clontech). The adenovirus splice acceptor was derived from pBigT (see Acknowledgments). A codon-optimized *lacZ* gene was derived from the nls-lacZ (nuclear localization signal β-galactosidase) in pBroad2-LacZnls (InvivoGen), and the sequence for nuclear localization was removed by PCR-based mutagenesis. All three gene-trap constructs were designed as follows. The splice acceptor (SA) was a widely used adenovirus SA[42]. The *lacZ* reporter with the bovine growth hormone polyadenylation signal (BGH-pA) was preceded by an IRES sequence from the encephalomyocarditis virus, which enabled translation of *lacZ* independent of the trapped endogenous genes. Three stop codons in three different reading frames were included between SA and IRES, preventing unwanted translation of truncated endogenous protein into the IRES. The IRES-*lacZ* sequence was followed by a CAG promoter–driven *eGFP* gene with the SV40 polyadenylation signal. The whole gene-trap cassette was flanked by two *lox*P sites. An independent *lox5171* site was also included for a potential increase in *trans*-allelic recombination efficiency, as has been suggested by *in vitro* data[43]. Standard cloning and recombineering were used to assemble all of the components into the full-length gene-trap vector ZG-l. This vector was further modified by recombineering and PCR-based mutagenesis to obtain the medium-sized ZG-m and the small-sized ZG-s gene-trap vectors (described in text). PBase was derived from plasmid 286 (ref. 44; see Acknowledgments) for the construction of CAG-PBase and Prm1-PBase vectors. The sequences of the final vectors were confirmed by sequencing. Transgenic mouse lines containing the above vectors were created by standard pronuclear injection methodology.

Inverse PCR to identify *piggyBac* insertion sites. To isolate genomic DNA for either inverse PCR or DNA blot analysis, about half-centimeter pieces of tail were put into 480 μl of lysis buffer (50 mM Tris (pH 8), 100 mM EDTA, 1% SDS and 100 mM NaCl). Twenty-five microliters of proteinase K (20 mg/ml stock) were added and the mixture was incubated at 55 °C overnight. A 0.25-ml aliquot of 6 M NaCl was added. The reaction mixture was kept on ice for 10 min, shaken vigorously for 2 min, and then spun at 14,000 r.p.m. (20,000*g*) for 10 min at 4 °C. The supernatant was transferred to an Eppendorf tube, and ~1 ml of 100% ethanol was added. A capillary tube was used to spool out the DNA, which was washed in 70% ethanol and dissolved in 200 μl of TE buffer. Twenty microliters (about 5 μg) of the DNA was digested with the appropriate restriction enzyme in a 50-μl reaction volume for 2 h. The enzyme digestion was purified on a Qiagen column and eluted with 50 μl of water, all of which (~44 μl) was mixed with 5 μl of 10× T4 DNA ligase buffer and 1 μl (5 units) of T4 DNA ligase (Fermentas, EL0011) in a total volume of 50 μl for ligation reaction at room temperature (22 °C) for 2 h. The ligation reaction was purified on a Qiagen column, and eluted with 50 μl of water. Three microliters was used in a 25-μl PCR reaction (**Supplementary Table 1** contains primer sequences). For identifying the ZG-1 3′ junction, the DNA was digested with *Msp*I and the PCR primers were PB37inv3R and PB38inv3F. For the ZG-m junction, the DNA was digested with *Msp*I, and the primers were PB39inv3F and PB37inv3R. For ZG-s, *Msp*I digestion was used for the 3′ junction and the primers were PB40inv3R and PB41inv3F. The Roche Expand long template PCR kit was used for all inverse PCR reactions. The PCR conditions were as follows: 92 °C for 2 min; 30 cycles of 92 °C for 10 s, 55 °C for 30 s and 68 °C for 3 min; 68 °C for 10 min. Starting from tenth cycle, extension was increased by 20 s for each successive cycle.

Skeleton preparation and X-gal staining. For analysis of *piggyBac* mutants, newborn skeletons were prepared as described[45]. X-gal staining of embryos, tissue in whole mount, or sections was done as described[46].

Accession codes. Sequences of *piggyBac* gene-trap plasmids have been submitted to GenBank (EF591488, EF591489, EF591490, EF591491 and EF591492). Sequences of all other plasmids used in this study are available from the authors on request.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank J.R. Mann (City of Hope's Beckman Research Institute) for the *Hprt*-Cre mouse; J.-F. Cheng (Lawrence Berkeley National Laboratory) for the *Pcdh* BAC clones; C. Zou and H. Peng (laboratory of Q.W.) for constructing *Pcdha1* and *Pcdhga1* targeting vectors and performing DNA blot experiments; M. Hockin for help with chromosome painting; G. Karan for fluorescence photography of adult mice; members of M.R.C.'s laboratory for comments on the manuscript and the University of Utah Transgenic/Gene Targeting Facility for pronuclear injections. pBigT was a gift of F. Constantini (Columbia University), plasmid C4-PBss was gift of R.S. Mann (Columbia University) and plasmid 286 was a gift of A. Handler (University of Florida). We are grateful for technical support from the ES cell culture, mouse surgery and husbandry staff in M.R.C.'s laboratory, in particular S. Barnett, L. Byers, C. Lenz, K. Lustig, J. Tomlin and J. Shuhua. G.Y. and Q.W. were supported by a grant from the American Cancer Society. Q.W. is a Basil O'Connor Scholar.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturegenetics.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Austin, C.P. *et al.* The knockout mouse project. *Nat. Genet.* **36**, 921–924 (2004).
2. Auwerx, J. *et al.* The European dimension for the mouse genome mutagenesis program. *Nat. Genet.* **36**, 925–927 (2004).
3. The International Mouse Knockout Consortium. A mouse for all reasons. *Cell* **128**, 9–13 (2007).
4. Dermitzakis, E.T., Reymond, A. & Antonarakis, S.E. Conserved non-genic sequences — an unexpected feature of mammalian genomes. *Nat. Rev. Genet.* **6**, 151–157 (2005).
5. Bejerano, G. *et al.* Ultraconserved elements in the human genome. *Science* **304**, 1321–1325 (2004).
6. Boffelli, D., Nobrega, M.A. & Rubin, E.M. Comparative genomics at the vertebrate extremes. *Nat. Rev. Genet.* **5**, 456–465 (2004).
7. Sandelin, A. *et al.* Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. *BMC Genomics* **5**, 99 (2004).
8. Margulies, E.H. *et al.* Comparative sequencing provides insights about the structure and conservation of marsupial and monotreme genomes. *Proc. Natl. Acad. Sci. USA* **102**, 3354–3359 (2005).
9. Vavouri, T., McEwen, G.K., Woolfe, A., Gilks, W.R. & Elgar, G. Defining a genomic radius for long-range enhancer action: duplicated conserved non-coding elements hold the key. *Trends Genet.* **22**, 5–10 (2006).
10. Bejerano, G., Haussler, D. & Blanchette, M. Into the heart of darkness: large-scale clustering of human non-coding DNA. *Bioinformatics* **20**, I40–I48 (2004).
11. Golic, K.G. & Golic, M.M. Engineering the *Drosophila* genome: chromosome rearrangements by design. *Genetics* **144**, 1693–1711 (1996).
12. Ryder, E. *et al.* The DrosDel collection: a set of P-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. *Genetics* **167**, 797–813 (2004).
13. Zheng, B., Sage, M., Sheppeard, E.A., Jurecic, V. & Bradley, A. Engineering mouse chromosomes with Cre-*loxP*: range, efficiency, and somatic applications. *Mol. Cell. Biol.* **20**, 648–655 (2000).
14. Mills, A.A. & Bradley, A. From mouse to man: generating megabase chromosome rearrangements. *Trends Genet.* **17**, 331–339 (2001).
15. Herault, Y., Rassoulzadegan, M., Cuzin, F. & Duboule, D. Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat. Genet.* **20**, 381–384 (1998).
16. Kmita, M., Fraudeau, N., Herault, Y. & Duboule, D. Serial deletions and duplications suggest a mechanism for the collinearity of Hoxd genes in limbs. *Nature* **420**, 145–150 (2002).
17. Zakany, J. & Duboule, D. Hox genes and the making of sphincters. *Nature* **401**, 761–762 (1999).
18. Genoud, N. *et al.* Disruption of Doppel prevents neurodegeneration in mice with extensive Prnp deletions. *Proc. Natl. Acad. Sci. USA* **101**, 4198–4203 (2004).
19. Spitz, F., Herkenne, C., Morris, M.A. & Duboule, D. Inversion-induced disruption of the Hoxd cluster leads to the partition of regulatory landscapes. *Nat. Genet.* **37**, 889–893 (2005).
20. Ding, S. *et al.* Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**, 473–483 (2005).

21. Wu, Q. & Maniatis, T. A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* **97**, 779–790 (1999).
22. Wu, Q. *et al.* Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. *Genome Res.* **11**, 389–404 (2001).
23. Wu, Q. Comparative genomics and diversifying selection of the clustered vertebrate protocadherin genes. *Genetics* **169**, 2179–2188 (2005).
24. Kohmura, N. *et al.* Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron* **20**, 1137–1151 (1998).
25. Wang, X. *et al.* Gamma protocadherins are required for survival of spinal interneurons. *Neuron* **36**, 843–854 (2002).
26. Phillips, G.R. *et al.* Gamma-protocadherins are targeted to subsets of synapses and intracellular organelles in neurons. *J. Neurosci.* **23**, 5096–5104 (2003).
27. Weiner, J.A., Wang, X., Tapia, J.C. & Sanes, J.R. Gamma protocadherins are required for synaptic development in the spinal cord. *Proc. Natl. Acad. Sci. USA* **102**, 8–14 (2005).
28. Hambsch, B., Grinevich, V., Seeburg, P.H. & Schwarz, M.K. γ-Protocadherins, presenilin-mediated release of C-terminal fragment promotes locus expression. *J. Biol. Chem.* **280**, 15888–15897 (2005).
29. Tang, S.H., Silva, F.J., Tsark, W.M. & Mann, J.R. A Cre/*lox*P-deleter transgenic line in mouse strain 129S1/SvImJ. *Genesis* **32**, 199–202 (2002).
30. Bonin, C.P. & Mann, R.S. A piggyBac transposon gene trap for the analysis of gene expression and function in *Drosophila*. *Genetics* **167**, 1801–1811 (2004).
31. Thibault, S.T. *et al.* A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. *Nat. Genet.* **36**, 283–287 (2004).
32. Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S. & Capecchi, M.R. Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc. Natl. Acad. Sci. USA* **97**, 13702–13707 (2000).
33. DeChiara, T.M. *et al.* Ror2, encoding a receptor-like tyrosine kinase, is required for cartilage and growth plate development. *Nat. Genet.* **24**, 271–274 (2000).
34. Komori, T. *et al.* Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* **89**, 755–764 (1997).
35. Jung, J.Y. *et al.* Expression of urea transporters in potassium-depleted mouse kidney. *Am. J. Physiol. Renal Physiol.* **285**, F1210–F1224 (2003).
36. Forster, A. *et al.* Engineering de novo reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. *Cancer Cell* **3**, 449–458 (2003).
37. Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D. & Luo, L. Mosaic analysis with double markers in mice. *Cell* **121**, 479–492 (2005).
38. Dracopoli, N.C. *et al.* in *Current Protocols in Human Genetics* 4.1.1–4.1.4 (John Wiley & Sons, Hoboken, New Jersey, 2006).
39. Cary, L.C. *et al.* Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* **172**, 156–169 (1989).
40. Fraser, M.J., Ciszczon, T., Elick, T. & Bauser, C. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Mol. Biol.* **5**, 141–151 (1996).
41. Li, X. *et al.* piggyBac internal sequences are necessary for efficient transformation of target genomes. *Insect Mol. Biol.* **14**, 17–30 (2005).
42. Friedrich, G. & Soriano, P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5**, 1513–1523 (1991).
43. Liu, P., Jenkins, N.A. & Copeland, N.G. Efficient Cre-loxP-induced mitotic recombination in mouse embryonic stem cells. *Nat. Genet.* **30**, 66–72 (2002).
44. Handler, A.M. & Harrell, R.A. II Transformation of the Caribbean fruit fly, *Anastrepha suspensa*, with a piggyBac vector marked with polyubiquitin-regulated GFP. *Insect Biochem. Mol. Biol.* **31**, 199–205 (2001).
45. Boulet, A.M. & Capecchi, M.R. Multiple roles of Hoxa11 and Hoxd11 in the formation of the mammalian forelimb zeugopod. *Development* **131**, 299–309 (2004).
46. Economides, K.D., Zeltser, L. & Capecchi, M.R. Hoxb13 mutations cause overgrowth of caudal spinal cord and tail vertebrae. *Dev. Biol.* **256**, 317–330 (2003).
47. Bunting, M., Bernstein, K.E., Greer, J.M., Capecchi, M.R. & Thomas, K.R. Targeting genes for self-excision in the germ line. *Genes Dev.* **13**, 1524–1528 (1999).