A protocol for constructing gene targeting vectors: generating knockout mice for the cadherin family and beyond

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

[1]Howard Hughes Medical Institute and [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).

Published online 29 May 2008; doi:10.1038/nprot.2008.70

We describe here a streamlined procedure for targeting vector construction, which often is a limiting factor for gene targeting (knockout) technology. This procedure combines various highly efficient recombination-based cloning methods in bacteria, consisting of three steps. First step is the use of Red-pathway-mediated recombination (recombineering) to capture a genomic fragment into a Gateway-compatible vector. Second, the vector is modified by recombineering to include a positive selection gene *neo*, from a variety of modular reagents. Finally, through a simple *in vitro* Gateway recombination, the modified genomic fragment is switched into a vector that contains negative selection cassettes, as well as unique sites for linearization. To demonstrate the usefulness of this protocol, we report targeted disruptions of members of the cadherin gene family, focusing on those that have not been previously studied at the molecular genetic level. This protocol needs ~2 weeks to construct a targeting vector, and several vectors can be easily handled simultaneously using common laboratory setup.

INTRODUCTION

Gene targeting, the use of homologous recombination in mouse embryonic stem (ES) cells to modify mouse genes precisely[1–5], allows researchers to create virtually any desired modification in the genome of a living mouse (for an introductory review, see ref. 6). The predominant use of gene targeting is to generate mice with loss-of-function mutations, so-called 'knockout mice'. To date, thousands of mouse genes have been disrupted by gene targeting[7,8]. Several major international programs have been initiated with the goal to disrupt every gene in mouse ES cells[7–9]. Yet, mice with custom-designed mutant alleles that address investigators' specific experimental requirements are still in increasing demand.

A limiting factor of conventional gene targeting is the time-consuming and technically challenging process of targeting vector construction. Because most targeting vectors are plasmids of well over 20 kb in size, they can be difficult and time-consuming to construct using conventional restriction enzyme-based cloning methods. PCR-based methods[10] using presently available high-fidelity polymerases, though straightforward, are generally avoided for this purpose because they often introduce unwanted mutations when amplifying the large DNA templates required for generating the homology arms. Much sequencing effort is needed to ensure mutation-free PCR product, particularly for DNA amplified from strains (e.g., most commonly used 129Sv DNA) with no published genome sequences—published genome sequences are from the mouse strain C57BL/6J[11]. It is costly and time-consuming to determine whether sequence differences with C57BL/6J are polymorphisms specific to 129Sv DNA versus true mutations in long PCR products. Thus, the overall effort can be more than that for conventional enzyme-based cloning methods.

A number of recombination-based methods in bacteria and yeast have been used for molecular cloning[12–16]. A greatly improved method which uses the Rac prophage recombinase pair (RecE and RecT)[17,18] requires only ~50 bp of homology arms to mediate the desired recombination events in bacteria. However, the efficiency of this procedure is reduced when used for generating large DNA constructs, required for gene targeting vector construction[19,20]. Later, another improvement was introduced by expressing λ phage Red-recombination proteins Redα, Redβ and Redγ from a defective λ-prophage in bacterial chromosome[19,21,22], with a 10–100-fold increase in efficiency. Both improved recombination methods (also known as recombineering) have been used extensively for targeting vector construction[19,23,24].

Recombineering has prompted several attempts to develop high-throughput methods for targeting vector construction. An example is the REC method[25], which couples library screening with targeting vector construction. However, the REC method requires relatively complex phage manipulation. Furthermore, the homology arms screened out from the library may not be compatible with subsequent confirmation analysis (e.g., Southern blot strategy). Therefore, the adoption of REC has been limited. In another high-throughput method, bacterial artificial chromosome (BAC) clones are directly modified to create targeting vectors[26] through Red-recombination. However, manipulation of BAC clones is technically more challenging than handling plasmids[26–29], and Southern blot or PCR screening of ES cells modified with these vectors is very difficult[26]. BAC vectors are thereby not routinely adopted as plasmid vectors; however, they have certain advantages in situations such as creating 'humanized mice'[30].

High-throughput approaches[26,31] often emphasize production speed, rather than individual researchers' needs for modifying their genes of interest[32]. The international knockout programs are focusing on disrupting genes in ES cells, and because of budget only a very small number of genes from these disrupted ES cells will be chosen to make mice[7]. Previous large-scale efforts have produced mutant ES cell libraries that cover a significant portion of mouse genes[33–35], yet researchers have been reluctant to use these lines to produce mice; apart from the cost, an important factor is that investigators are often more interested in research-oriented designs that best meet their experimental requirements.

Many excellent protocols exist for targeting vector construction, and for knockout production in general[19,23,26–30,36,37]. As a complement to these protocols, we describe here a fully tested, simple and efficient protocol that allows for rapid production of sophisticated targeting vectors. This procedure breaks down the complexity of targeting vector construction into a few simple modular steps. For beginners in the field, to understand gene targeting technology and this protocol quickly, a shortcut is to read the following references (refs. 6,17–19,38–40).

We use an improvement of the Red-recombination method, where Red recombinases are expressed from a very low-copy plasmid pKD46 (ref. 40). As mentioned earlier, several other similar plasmids and the chromosome-based recombineering[19] (e.g., DY380) have been used successfully for molecular cloning in bacteria[23,31,41–43]. Chromosome-based recombineering[19] (e.g., DY380) has been used more commonly due to its much higher recombinogenic efficiency; however, it does often require the use of relatively long homology arms (200–500 bp each) to mediate the recombination events, which entails several additional cloning steps thereby reducing many of the advantages associated with the simplicity of recombineering protocols[19,31]. Our own comparison suggests that pKD46 has about the same efficiency as the chromosome-based system, while still requiring only a short homology of ~50 bp to mediate efficient recombination.

Furthermore, we created a series of modules to streamline the construction procedure. Together, these components generate a flexible new protocol for targeting vector construction that incorporates the use of both Red-recombination and Gateway recombination[44], different self-excision *neo* cassettes and many small yet important technical details. We have paid special attention to allow for targeting vector construction from DNAs that are difficult to maintain in bacteria and to help avoid many trivial yet frustrating troubles commonly encountered by beginners with limited cloning experience. This protocol has been tested extensively over a 7-year period, and has been used to generate a broad spectrum of knockout mouse lines in previous publications[45–47].

To illustrate the efficacy of this modular cloning protocol, we show, as an example, our effort to disrupt members of the cadherin family (**Table 1**) systematically. The cadherin family of cell adhesion genes is among the largest gene families in the mouse genome, containing >100 clustered and dispersed members[48,49]. These genes are important players in constructing the body plan during development, playing disparate roles in processes such as the epithelial-mesenchymal transitions, synaptic formation, axon guidance and neural circuit establishment[50], planar cell polarity and organ shape formation and cell sorting and tissue morphogenesis[51]. Mutations in members of the cadherin family can lead to dramatic phenotypes including neuronal diseases and tumor metastasis. Despite two decades of extensive work in this area, functional studies in the mouse for many of these genes are still lacking.

For simplicity, examples provided in this protocol are mostly loss-of-function targeting vectors. Once principles in this protocol are established, targeting vectors for other more sophisticated mutations can be readily created by this protocol with very minor adaptations—which are mentioned accordingly in the procedure. In addition, we have been able to use the same modular cassettes and cloning methods described here to generate targeting vectors for use in other mammalian cell types (e.g., human embryonic stem cells; Y. Liu, S.W. & M.R.C., unpublished data) with similar targeting frequencies to those for mouse ES cells. Gene targeting is experiencing an ever greater demand in the post-genomic era. Even with the completion of the large-scale knockout mouse projects[7], the need for more custom-designed loss- or gain-of-function alleles, such as point mutations, gene-swaps, Cre-drivers and many others, will inevitably continue to rise.

TABLE 1 | Targeted alleles of cadherin and protocadherin genes.

Gene name	Allele name	Targeting vector	Germline transmission
Pcdha (Mid)	*del(Mid)*	Yes	Yes
Pcdha (CIE)	*del(CIE)*	Yes	Yes
Pcdha Type A	A^{CFP}	Yes	Yes
Pcdha Type B	B^{YFP}	Yes	Yes
Pcdha Type A	*delA*	Yes	Yes
Pcdha Type B	*delB*	Yes	Yes
Pcdha (Down)	*del(Down)*	Yes	Yes
Pcdhb1	$Pcdhb1^{EGFP}$	Yes	In progress
Pcdhb22	$Pcdhb22^{EGFP}$	Yes	In progress
Celsr2	$Celsr2^{EGFP}$	Yes	Yes
Celsr3	$Celsr3^{EGFP}$	Yes	Yes
Fat2	$Fat2^{EGFP}$	Yes	Yes
Fat3	$Fat3^{nlacZ}$	Yes	Yes
Fat4	$Fat4^{EGFP}$	Yes	Yes
Dscam	$Dscam^{EGFP5'}$	Yes	In progress
Dscam	$Dscam^{EGFP3'}$	Yes	In progress
DscamL1	$DscamL1^{EGFP5'}$	Yes	In progress
DscamL1	$DscamL1^{EGFP3'}$	Yes	In progress
Dchs1	$Dchs1^{EGFP}$	Yes	In progress
Dchs2	$Dchs2^{EGFP}$	Yes	In progress
Cdh8	$Cdh8^{EGFP}$	Yes	In progress
Cdh13	$Cdh13^{EGFP}$	Yes	In progress
Cdh18	$Cdh18^{EGFP}$	Yes	In progress
Cdh19	$Cdh19^{EGFP}$	Yes	In progress
Cdh20	$Cdh20^{EGFP}$	Yes	In progress
Cdh22	$Cdh22^{EGFP}$	Yes	In progress
Cdh24	$Cdh24^{EGFP}$	Yes	In progress
Pcdh1	$Pcdh1^{EGFP}$	Yes	In progress
Pcdh7	$Pcdh7^{EGFP}$	Yes	In progress
Pcdh10	$Pcdh10^{EGFP}$	Yes	In progress
Pcdh17	$Pcdh17^{EGFP}$	Yes	In progress
Pcdh18	$Pcdh18^{EGFP}$	Yes	In progress
Pcdh19	$Pcdh19^{EGFP}$	Yes	In progress
Pcdh20	$Pcdh20^{EGFP}$	Yes	In progress

Knockout mouse production; an overview

The procedure for a knockout mouse production has been reviewed many times[38,39,52–54]. First, principally in ES cells derived from agouti brown mice (strain 129Sv), designed modifications engineered in genomic DNA of a targeting vector are transferred into the endogenous locus by homologous recombination. ES cell clones containing the modified DNA are enriched by positive and negative selection[2] and identified by Southern blot analysis. Second, targeted ES cells are injected into blastocysts derived from black mice (C57BL/6J) to generate chimeric (brown/black) mice. When targeted ES cells contribute to formation of the germ line of chimeric mice, progeny with the desired mutations are obtained. The flow chart in **Figure 1** explains the different stages of the whole gene targeting process. Although this protocol is focused on

targeting vector construction, a better understanding of the overall gene targeting procedure from start to finish can help design a good targeting strategy. Therefore, in addition to describing a detailed method for vector construction, we are also summarizing briefly technical details that are essential for a successful gene targeting experience.

Targeting vector design

An efficient targeting vector design uses both positive and negative selection[2,38,52]. Most current targeting vectors are replacement vectors[1,39], which are a modified genomic fragment of an endogenous gene, and are used to replace the latter. The genomic fragments present on each side of the modification are generally referred to as homology arms (**Fig. 2a**). After a targeting vector is introduced into ES cells, they can result in rare events of either homologous recombination or random integration. Therefore, a positive selection gene (e.g., *neo*) present on the targeting vector is needed to select for these rare events and a negative selection gene (e.g., herpes simplex virus thymidine kinase (*HSV-tk*) gene) to further select against random integration event. Although there is no single design that fits all needs, there are some general principles that can help design a good targeting strategy and avoid mistakes. To design an efficient targeting strategy, the sequences of the mouse gene of interest and its homologs and functional studies of the gene should be incorporated into the planning. Many genes have complex genomic structures, with multiple introns and exons spanning hundreds of kilobases. However, gene targeting, in its current most-used form with a plasmid-based targeting vector, can only efficiently delete up to ~15 kb in ES cells (S.W. and M.R.C., unpublished data). (For larger deletions, an efficient *in vivo* breeding technique can be used[46].) Therefore, prior knowledge about the genomic organization of a gene is often useful to decide which part of the gene to modify. The known genomic sequence can also be used to position the homology arms to reduce the repetitive DNA sequences contained within the targeting vector. The presence of excessive repetitive DNA can significantly reduce the targeting frequency (S.W. and M.R.C., unpublished data).

Simple loss-of-function knockout. In the simple loss-of-function approach (**Fig. 2a**), a reporter gene, for example, GFP or β-galactosidase, is often used to replace a part or all of the coding sequence of a gene. By placing the reporter gene in frame with the endogenous start codon AUG, the expression pattern of the endogenous gene can be easily followed *in vivo*. For most genes, replacing the coding sequence a few amino acids after the start codon AUG with a reporter cassette ensures a null allele and recapitulates the endogenous gene expression pattern. But for more complex genes with multiple promoters and/or alternative splicing, a reporter with a strong transcription stop might be needed to ensure a true null allele. An alternative strategy to obtain a loss-of-function allele for such complicated genes is to replace their most important domains with a reporter, or delete the entire gene through *in vivo* breeding[46]. We have compared different reporters *in vivo*: enhanced GFP (EGFP), enhanced yellow fluorescent protein (EYFP), enhanced cyan fluorescent protein (ECFP), Cre, lacZ and alkaline phosphatase (AP). It appears that EGFP, lacZ and Cre are good choices. If higher resolution is required, the inclusion of a nuclear localization signal for the reporter gene is beneficial. For the purpose of visually labeling neurons and their projections, a τ-EGFP reporter has been described previously[55].

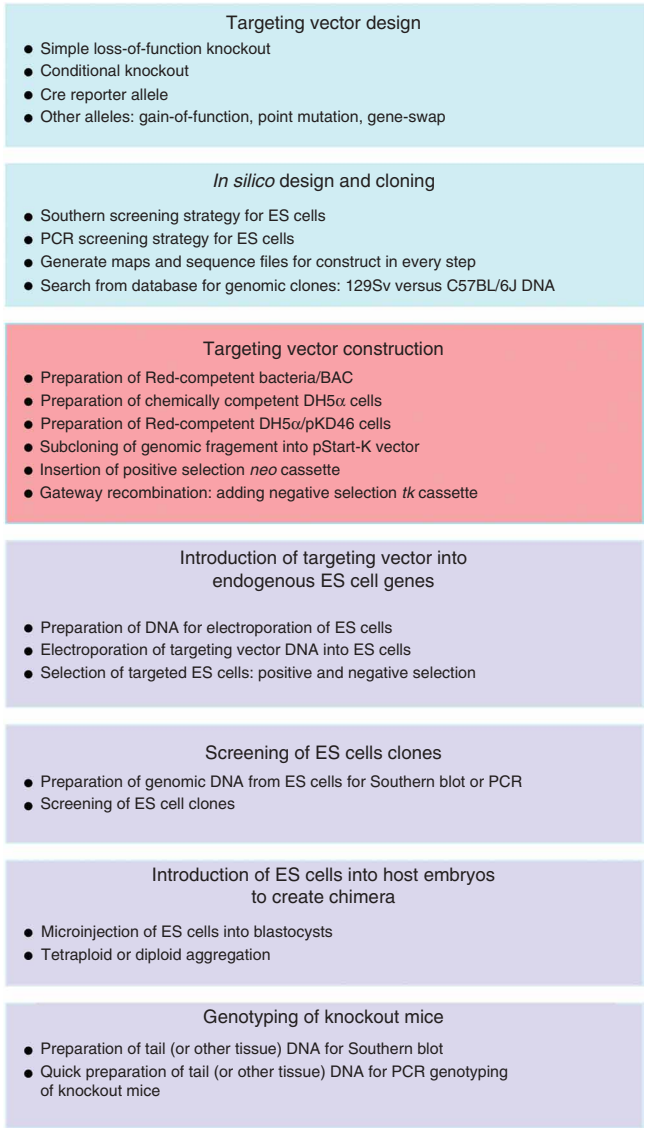

Figure 1 | Flow chart showing step by step how to design and construct a targeting vector: considerations before and after targeting vector cloning. BAC, bacterial artificial chromosome; ES, embryonic stem.

Conditional knockout. Cre/*lox*P and Flp/*FRT* site-specific recombination systems have become an essential part of mouse knockout technology[56–59] and have made a conditional knockout approach possible[60–63]. Cre recombinase (derived originally from bacteriophage P1) and Flp recombinase (derived originally from yeast) recognize their target DNA sequences named *lox*P and *FRT*, respectively, each having a 34 bp consensus core sequence[64–67]. Depending on the orientation and structure of *lox*P or *FRT*, Cre or Flp-mediated recombination can result in deletions, insertions, inversions and exchanges[46,56]. Because a significant portion of the ~25,000 mouse genes are essential for survival, loss-of-function alleles often result in embryonic or postembryonic lethality, precluding analysis of that gene's function at stages past lethality. A conditional allele overcomes this limitation by flanking the gene

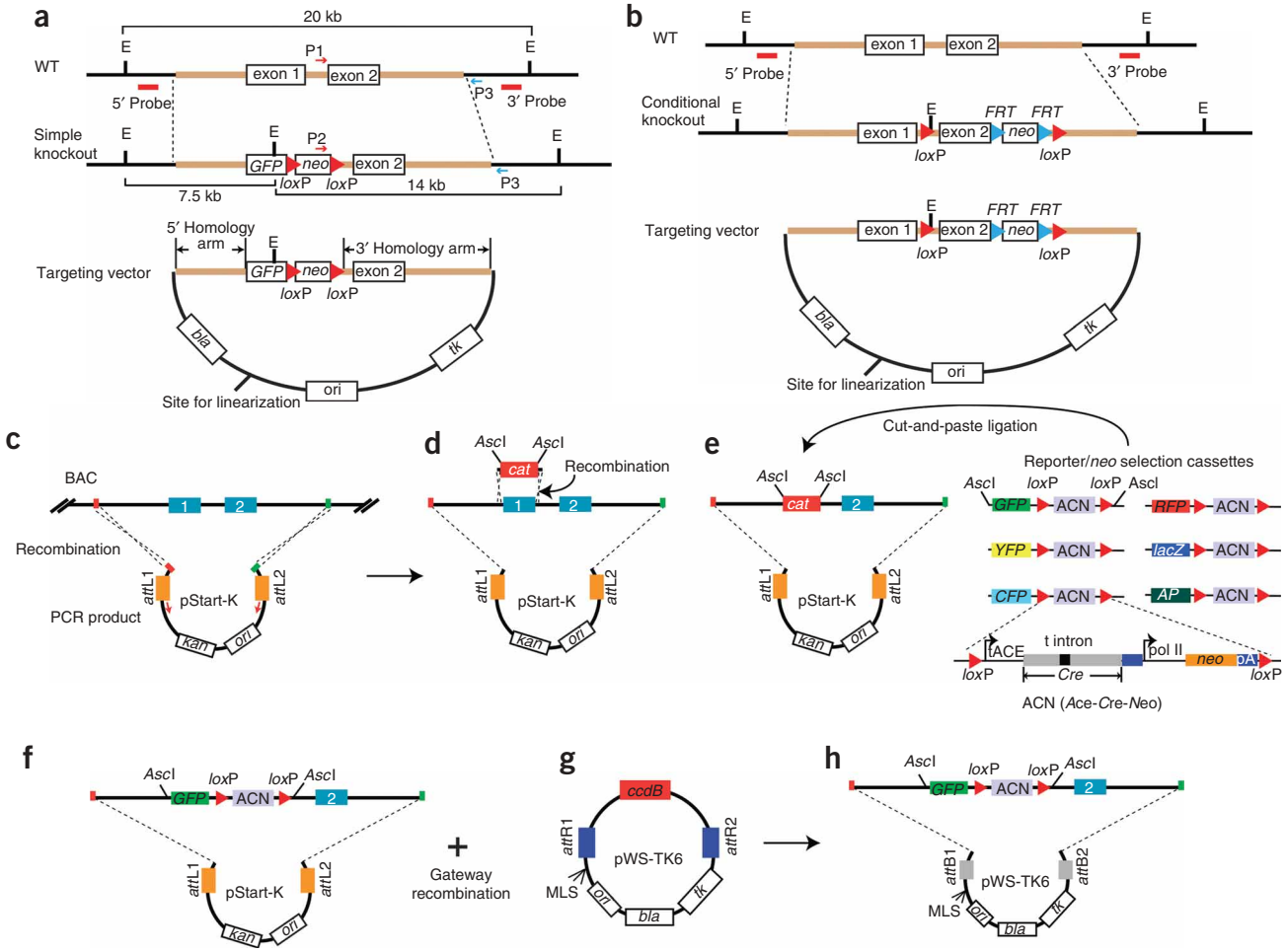

Figure 2 | Two basic targeting vector designs and construction of a simple loss-of-function targeting vector. Schematics showing general design of a loss-of-function allele (**a**) and a conditional knockout allele (**b**). A Southern blotting strategy for embryonic stem (ES) cell screening should be designed before construction of targeting vector. A good Southern strategy should easily distinguish between correctly targeted allele and random integrations. In general, a downshifted band for the targeted allele is easier to identify than an upshifted band, as the latter could often be hidden in the noise from a suboptimal Southern blot. Although 5′ and 3′ flanking probes are used to identify desired recombination event, internal probes (*neo*) are used to identify unwanted random integrations. The principle of Southern blot is illustrated by a simplified example in panel **a**, where a restriction enzyme (E) is used for digestion of DNA isolated from ES cells, to allow detection of fragments using 5′, 3′ and internal *neo* probes. Because the targeting vector contains an additional restriction enzyme site compared with the wild-type (WT) gene sequence, using the 5′ flanking probe, a 7.5-kb band (downshift) is expected for the targeted allele, compared with the 20-kb WT band (size limit is ~25–30 kb for efficient gel separation and transfer). For the 3′ flanking probe, a 14-kb band (downshift) is expected for the targeted allele, compared with the 20-kb WT band. For the internal *neo* probe, a 14-kb band is expected for the targeted allele, while random integrations can show variations. (**c**) To construct the targeting vector, Red-recombination is first used to pull out a genomic fragment from a chosen BAC clone that can be used for the homology arms in the targeting vector. On the other hand, genomic fragment can also be subcloned into pStart-K using restriction enzyme-based method. pStart-K is created as a Gateway-compatible vector, with a low-copy origin for DNA replication. (**d**) The region of interest (e.g., exon 1) is replaced with an *Asc*I-flanked chloramphenicol resistance gene (*cat*) by another round of Red-recombination. Other desired changes, such as insertion, point mutation and so on, can also be introduced in this step through λ phage Red-mediated recombination, or restriction enzyme-based cloning. (**e**) A set of *Asc*I-flanked reporter cassettes. They all contain a self-excision *neo* cassette (designated ACN) used for selection in mouse ES cells, which is subsequently automatically deleted in the male germline[85]. If endogenous polyadenylation signal is preferred, we have also designed a series of reporter/*neo* cassettes that do not carry a polyadenylation signal. (**f**) Subcloning of any reporter cassette from panel **e** into the *Asc*I site of the vector shown in panel **e** will result in the vector shown in panel **f**. (**g**) TK vectors were created as Gateway-compatible vectors and also contain multiple restriction sites for linearization (MLS) of the targeting vector. (**h**) The final targeting vector is obtained by a Gateway recombination of the vectors shown in panels **f** and **g**. *AP*, human placental alkaline phosphatase gene; *att*B1, *att*B2, *att*L1, *att*L2, *att*R1 and *att*R2, target DNA sequences for site-specific Gateway recombinases; *bla*, ampicillin resistance gene; *cat*, chloramphenicol resistance gene; *ccdB*, the F plasmid-encoded gene that inhibits growth of *Escherichia coli*; *CFP*, cyan fluorescent protein; *Cre*, Cre recombinase gene; *FRT*, target DNA sequence for site-specific Flp recombinase; *kan*, kanamycin resistance gene; *lacZ*, β-galactosidase gene; *lox*P, target DNA sequence for site-specific Cre recombinase; *neo*, neomycin resistance gene; *ori*, origin of replication; pA, polyadenylation site; pol II, RNA polymerase II promoter; *RFP*, red fluorescent protein; tACE, testic-specific promoter for angiotensin converting enzyme; t-intron, the small t intron from SV40 virus; *tk*, herpes simplex virus thymidine kinase gene; *YFP*, yellow fluorescent protein.

of interest with *lox*P sites (**Fig. 2b**). In this way, the gene of interest is only disrupted when Cre recombinase is provided. Because many Cre drivers, each with specific expression in different tissues and developmental stages, have been generated[68], the *lox*P-flanked gene can be excised in desired tissues and developmental periods. Conditional alleles can also be designed to offer more creative

uses in many unique ways. Conditional alleles can be engineered with a reporter gene in a way that the reporter is not expressed before Cre-mediated removal of the endogenous gene, but expressed from the endogenous locus after Cre-mediated recombination[69]. Conditional rescue alleles can be equally informative if designed well[70].

Cre reporter allele. Cre recombinase, as a reporter, has become increasingly popular for many investigators because they can be used for lineage analysis, conditional mutagenesis and conditional cell ablation[71]. Previously, these so-called Cre drivers were mostly created through pronuclear injection-based transgenesis[56]. Recently, growing numbers of Cre drivers are created through gene targeting as a knockin/null allele, or as an internal ribosomal entry site (IRES) version without disrupting the endogenous gene function. In the knock-in design, *Cre* gene replaces in frame the endogenous gene coding sequence, concurrently generating a null allele. In the IRES version, *Cre* gene is inserted after the endogenous gene coding sequence, and is transcribed contiguously with the endogenous transcript. Encephalomyocarditis virus (EMCV) IRES is the most commonly used IRES. As the IRES sequence allows translation initiation in the middle of the mRNA, the IRES-Cre created through gene targeting, like the knockin Cre allele, faithfully recapitulates the endogenous gene expression. For many genes, Cre expression from either knockin Cre or IRES-Cre is sufficient to effect efficient recombination, resulting in the same expression pattern and lineage pattern (S.W., Y. Wu and M.R.C., unpublished data). When IRES-Cre or other IRES-reporter is inserted right after the stop codon (with ~30 bp residual cloning sequences between the stop codon and the EMCV IRES in our hands), ideal bicistronic expression is consistently obtained *in vivo* (S.W. and M.R.C., unpublished data). Owing to its usefulness, several modifications of Cre have been generated. These include GFP-Cre fusion protein, and tamoxifen-inducible CreER and CreERT2 (Cre fused with estrogen receptor)[72–74]. However, CreER or CreERT2 are generally not 100% efficient in induction and can have unwanted leaky expression before tamoxifen induction in certain situations.

Other alleles. In addition to the loss-of-function and conditional alleles, researchers are also using gene targeting to generate precise point mutations, gain-of-function alleles, tetracycline inducible alleles and other sophisticated alterations[55,75–77]. For example, conditional gain-of-function studies are increasingly popular. In a common design, a ubiquitous promoter, Rosa26, is followed by a *lox*P-flanked transcriptional stop, then the gene of interest[71,78–80]. The gene of interest can be expressed in a tissue specific and temporal manner by providing Cre recombinase in a controlled manner.

In silico design and cloning

To design an efficient targeting vector for the above purposes, softwares such as Gene Construction Kit or Vector NTI are essential. They can help speed up the cloning procedure by taking advantage of the fully sequenced mouse genome information. Simple mistakes (e.g., designed fusion protein has an unwanted frameshift) unfortunately happen more often than you think but can be easily avoided using a software. The mouse genome informatics website is a very useful database resource for the laboratory mouse, providing integrated genetic, genomic and biological data for all mouse genes. Sequence information on mouse genes can also be obtained from the NCBI website (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090) or the UCSC Genome Bioinformatics Site (http://genome.ucsc.edu). BAC clones containing genes of interest can be readily searched from the UCSC Genome Bioinformatics Site by displaying BAC end pairs. Individual C57BL/6J BAC clones containing the desired genomic regions can be ordered directly from the BACPAC Resource Center (http://bacpac.chori.org/). Individual 129 BAC clones containing a gene of interest may need to be screened from RPCI-22 (129S6/SvEvTac) mouse BAC library, which has not been end sequenced. Other sources for 129 BAC clones include an end sequenced 129Sv BAC library[81].

A total of 8–15 kb homology to the target locus is normally used. Each homology arm should be >1 kb[82,83] (**Fig. 2a**). If both homology arms contain an excess of repetitive DNA sequences, then the targeting frequency will be low (S.W. and M.R.C., unpublished data), and longer arms or a shift in the position of the targeting vector may be required to obtain a successful targeting. Whether homology arms contain repetitive DNA sequences can be examined with BLAT (http://genome.ucsc.edu/cgi-bin/hgBlat) or RepeatMasker (http://www.repeatmasker.org/). For ES cells screening, a strategy for Southern blot analysis (**Fig. 2a**) should be considered during this designing phase of targeting vector construction. Not all restriction enzymes are equally efficient in digesting genomic DNA from ES cells. Some good sites that have been tested in our lab are *Acc*65I, *Bam*HI, *Bgl*I, *Bgl*II, *Bsr*GI, *Cla*I, *Eco*RI, *Eco*RV, *Hind*III, *Asp*718, *Nco*I, *Pst*I, *Rsr*II, *Spe*I, *Sph*I, *Sca*I (Roche), *Sst*I and *Xba*I. Template used for Southern probe may be 200–2,000 bp. When designing probe templates for Southerns, we check the template sequence with BLAT or RepeatMasker to avoid regions with repeats, because repeats in probes usually result in very high background noise. PCR product amplified from genomic DNA (e.g., BAC, ES cells DNA, tail DNA, etc.) can often be directly used as a probe template, but it is better to perform a TOPO-TA cloning (Invitrogen) of the PCR product because using a template cutout from a plasmid generates cleaner Southerns. Ideally, an alternate Southern screening strategy should also be designed. Probes for Southern transfer analysis should be tested before starting construction of the targeting vector. Alternatively, a PCR screening strategy can also be devised at this stage. If PCR is used as primary screening method, ideally, it is better to keep one homology arm <3 kb for easy PCR amplification later on. Primers for PCR, particularly the outside primer, should be tested at this stage. It is a good practice to generate maps and sequence files for every cloning step designed for a targeting vector before starting bench work.

Targeting vector construction

Using our targeting vector construction protocol, the first cloning step is to subclone a genomic fragment of the chosen gene from a BAC clone to generate the homology arms. To facilitate this and subsequent steps (**Fig. 2c–h**), we created a series of vectors, pStart-C2 (chloramphenicol resistant, Cam^r), pStart-K (kanamycin resistant, Kan^r), pStart-T2 (tetracycline resistant, Tet^r) and others (**Fig. 3a**, and data not shown). The different resistance genes in these Gateway-compatible, low-copy-number replicating plasmids provide choices for subcloning DNA from different sources.

To capture a fragment from the chosen BAC clone (usually Cam^r) into the pStart-K vector, two oligonucleotides of ~75 bases, including 50 bases homology to the 5′ or 3′ end points of the BAC region of interest, are used to amplify pStart-K to obtain a linear PCR product, which is, in turn, used to recombine out the genomic region of interest from the BAC (**Fig. 2c**). All subsequent modifications of the target locus are performed on this plasmid, thereby avoiding more difficult manipulation of intact BAC clones[26,31].

To insert a *neo*-positive selection cassette (for positive selection in ES cells) into the BAC subcloned fragment, first two oligos with homology (~50 bases) to the 5′ and 3′ ends of the insertion site of the genomic DNA in the BAC subcloned fragment are used to PCR amplify a different resistance gene (e.g., *cat*). This resistance gene, flanked by two *Asc*I (rare cutter) sites, is introduced into the genomic DNA captured by pStart-K through the use of Red-mediated recombination (**Fig. 2d**). Next, through a standard restriction enzyme-based cloning, a reporter/*neo* cassette can be readily inserted into the *Asc*I sites in the genomic DNA of pStart-K to replace the gene of interest. The *neo* gene in the reporter/*neo* cassette will be used as the positive selection marker in ES cells, because they can render cells resistant to the drug G418. We constructed a series of convenient reporter/*neo* cassettes that are all flanked by *Asc*I sites (**Fig. 2e**).

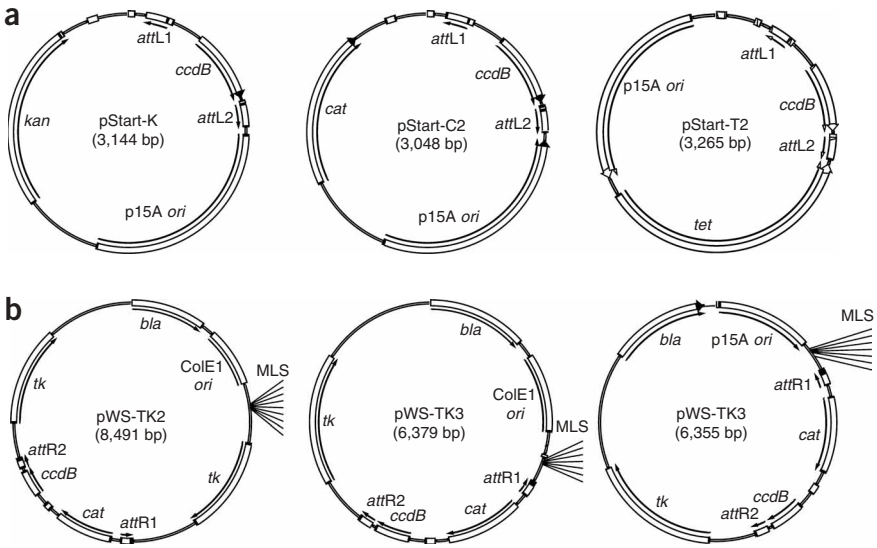

Figure 3 | Gateway-compatible vectors. (**a**) Although it is possible to use the commercially available high-copy Gateway plasmid (e.g., pENTR1a, Invitrogen) to subclone a genomic fragment by recombineering, we have found that it is often difficult to grow high-copy plasmid with some mammalian genomic DNA in bacteria (e.g., *Hoxc8* and *Hoxc10*). One simple solution we have found is to use low-copy plasmids such as those with p15A origin of replication. Therefore, we designed a series of Gateway-compatible, low-copy vectors from the plasmids pACYC177 and pACYC184 (New England Biolabs). Panel **a** shows three of these vectors, pStart-K, pStart-C2 and pStart-T2, which can all be used to pull out a genomic fragment from an existing plasmid or BAC clone. (**b**) We have also developed a series of Gateway destination vectors that can be used to add negative selection cassettes, as well as linearization sites, for targeting vectors. pWS-TK2 is a high-copy plasmid with two *tk* genes. pWS-TK3 is a high-copy plasmid with one *tk* gene. pWS-TK6 is a low-copy plasmid with one *tk* gene. All three TK vectors have multiple restriction sites for linearization (MLS). *att*L1, *att*L2, *att*R1 and *att*R2 target DNA sequences for site-specific Gateway recombinases; *bla*, ampicillin resistance gene; *cat*, chloramphenicol resistance gene; *ccdB*, the F plasmid-encoded toxin gene; *kan*, kanamycin resistance gene; *ori*, origin of replication; *tet*, tetracycline resistance gene; *tk*, herpes simplex virus thymidine kinase gene.

Because the presence of *neo* may affect neighboring gene expression, most current protocols use a *lox*P-flanked *neo* cassette for selection in ES cells that can subsequently be removed by crossing founders to a Cre deleter mouse[84]. To shorten this time-consuming breeding process (>5 months), a very effective *neo* cassette (designated ACN) using Cre-*lox*P system was developed for automatic self-excision of the *neo* gene in the male germ line[85]. We took advantage of this useful tool and built a series of reporter/*neo* self-excision cassettes (**Fig. 2e** and data not shown). We believe the use of reporter/*neo* cassettes, and Flp-*FRT*-based cassettes (S.W. and M.R.C., unpublished data), on a large scale should save years of mouse husbandry work.

In the final step, one or two negative selection cassettes, *HSV-tk* gene, are added to the targeting vector (**Fig. 2f**) for negative selection in ES cells[2]. *HSV-tk* in ES cells can be selected against using a drug FIAU (1-(2-deoxy-2-fluoro-β-D-arabinofuranosyl)-5-iodouracil). We constructed a series of Gateway-compatible TK vectors containing the *HSV-tk* gene (**Figs. 2g** and **3b**), with either high-copy or low-copy origins of replication. For stable genomic DNA, high-copy TK vectors can be used to facilitate DNA preparation, while low-copy TK vectors may be used to reduce potential problems with growth of vectors carrying unstable genomic DNA. Although including two *tk* genes, one at each end of the two homology arms in a targeting vector, increases enrichment of homologous recombinants[38], this configuration may cause instability during bacterial culture for some genomic DNAs. All TK vectors contain several engineered restriction sites for linearization of the targeting vector, before ES cell electroporation[86]. Through a simple Gateway recombination reaction, the modified genomic DNA with its reporter/*neo* cassette is transferred from pStart-K to a TK vector (**Fig. 2h**). The same cloning strategy described above for generating null alleles can be used for constructing conditional allele vectors, with only minor modifications (**Fig. 4**).

We have found that the use of a low-copy replicating plasmid as the carrier of the BAC subcloned fragment has many advantages. First, in bacteria some mammalian DNA sequences (e.g., *Hoxc8* and *Hoxc10*) are difficult to maintain in high-copy replicating plasmids, whereas in low-copy replicating plasmids many of these difficult DNAs are tolerated. Second, modifications of high-copy plasmids (containing ColE1, pUC or similar origins of DNA replication) by Red-mediated recombination almost always generates concatemers[19,21], containing both the original and the modified plasmids that are very difficult to separate. Again, use of a low-copy plasmid (containing an origin of replication such as p15A) usually eliminates such problems. In summary, the combined use of low-copy-plasmid (15–20 copies

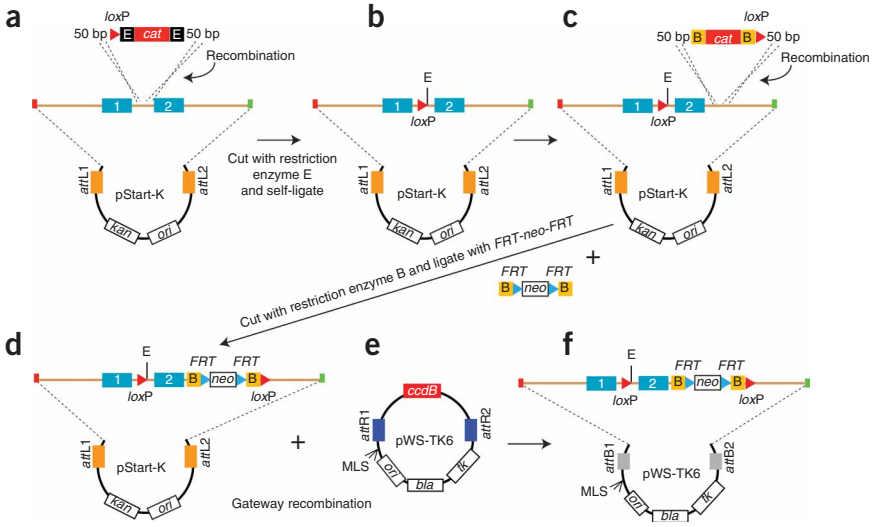

Figure 4 | Schematic for the construction of conditional targeting vectors. Most common conditional vectors use a design that includes two *lox*P sites to flank the region of interest. When choosing positions for inserting *lox*P sites and the *neo* cassette, care should be taken not to disrupt the endogenous transcription, translation and splicing. If possible, the first *lox*P site should be placed within an intron. If the first *lox*P has to be inserted before the start codon, we usually choose to insert it ~10 bp before the ATG (insertion to further upstream might affect transcription), minimizing the potential effect on the Kozak sequence and translation. Because *lox*P sequence in one orientation has two ATGs that could potentially be used as start codons, it is better to use the other orientation that has no ATG in the reading frame. When possible, the second *lox*P site and *neo* are also inserted within an intron. Another possible position for the second *lox*P is right after the stop codon. (**a**) To insert the first *lox*P site, two oligos are designed for PCR amplifying a resistance gene (*cat*). The forward oligo has this formula: 50 bases for homologous recombination, *lox*P site, a unique restriction enzyme site (E) and 20–25 bases as PCR primer. The reverse oligo has similar formula but without the *lox*P site. It should be noted that site E is also designed for use in Southern screening of ES cells, ensuring that 5′ *lox*P site is targeted. The PCR and Red-recombination conditions are the same as those described for generations of simple loss-of-function alleles. (**b**) The resulting plasmid from **a** is cut with the unique restriction enzyme E and self-ligates to obtain the plasmid in **b**. (**c**) To insert the second *lox*P site and *neo* cassette, the forward oligo has this formula: 50 bases for homologous recombination, a unique restriction enzyme site (B), and 20–25 bases as PCR primer. The reverse oligo has similar formula: 50 bases for homologous recombination, *lox*P site, a unique restriction enzyme site (B) and 20–25 bases as PCR primer. The PCR and Red-recombination conditions are the same as above. (**d**) The resulting plasmid in **c** is cut with the unique restriction enzyme B and ligated to a site B-flanked *FRT-neo-FRT* cassette to obtain the plasmid in **d**. The negative selection cassette TK (**e**) is added in a similar manner as for the simple loss-of-function targeting vectors by Gateway recombination (**f**). *att*B1, *att*B2, *att*L1, *att*L2, *att*R1 and *att*R2, target DNA sequences for site-specific Gateway recombinases; *bla*, ampicillin resistance gene; *cat*, chloramphenicol resistance gene; *ccdB*, the F plasmid-encoded toxin gene; *FRT*, target DNA sequence for site-specific Flp recombinase; *kan*, kanamycin resistance gene; *ori*, origin of replication; *tk*, herpes simplex virus thymidine kinase gene.

per bacterial cell) for cloning and low temperature (30–32 °C instead of 37 °C) for bacterial growth can solve many problems associated with large construct cloning.

Introduction of targeting vector into endogenous ES cell genes

During early days of gene targeting, microinjection was first used for DNA introduction into ES cells[86,87]. However, electroporation has since replaced this tedious method and is nowadays the commonly used method for introducing finished targeting vectors into ES cells[1]. Attempts to use chemical transfection reagents (e.g., calcium phosphate coprecipitation, Fugene-6, lipofectamine) in ES cells have not been successful[88,89].

In the experiments of early days, when DNA was microinjected into the nuclei of mammalian cells, it was found that multiple copies of exogenous DNA molecules can be inserted into one site of the recipients' chromosomes, and all of them are in the same orientation[90]. This and other studies (reviewed in ref. 52) demonstrate that mammalian cells contain intrinsic ability to mediate homologous recombination between newly introduced DNA and cognate DNA sequence on the chromosome. Targeting vectors were soon designed to take advantage of this cellular machinery for homologous recombination to modify endogenous genes[1,2,5,91].

How to choose a certain ES cell line is based on multiple considerations. Different ES cell lines may result in radically different targeting frequencies for the same targeting vector. Most current knockout experiments are performed in ES cell lines derived from strain 129Sv[81]. The finished mouse genome sequencing, however, was performed on C57BL/6J DNA, although other strains are being sequenced[11]. As these two strains have sequence variations from gene to gene that can affect gene targeting efficiency, the use of isogenic DNA for targeting vector construction is often very beneficial[81,82,92]. However, in a 129Sv and C57BL/6J hybrid ES cell line (e.g., G4), genomic DNA from either 129Sv or C57BL/6J can be used to generate homology arms for the targeting vectors[93]. An additional advantage of this G4 line is that complete ES cell-derived animals can be efficiently obtained through tetraploid aggregation.

When 129Sv-derived ES cells (e.g., R1 cells[94] or G4 cells[93]) are used, the F1 heterozygotes produced from breeding of chimera with C57BL/6J mice will have a mixed 129Sv/C57BL/6J genetic background. Because genetic background can affect phenotype analysis[95,96], backcrossing of knockout lines with C57BL/6J is often needed in fields, such as neuroscience and pharmaceutical discovery. ES cells derived from C57BL/6J and other inbred mouse lines have also been used successfully in many laboratories for chimera generation and germline transmission, allowing direct production of mutant mouse lines on these chosen backgrounds[97–100].

Selection of targeted ES cell clones: positive and negative selection

In a typical experiment of electroporation, 10 million ES cells are transfected with linearized targeting vector. Roughly half of the cells can survive electroporation, yet only a very small fraction (~0.1%) of those survived cells are transformed[1]—they are results of homologous recombination or random integration. A positive selection (*neo*) cassette on the targeting vector can render the ES cells resistant to the aminoglycoside G418 (geneticin, a derivative

of gentamycin), and allow direct selection of these rare events of transformation. Although the *neo* gene is the most widely used positive selection gene[1], other positive selection marker genes have also been used in ES cells, such as the hygromycin resistance gene, the puromycin resistance gene, the zeocin resistance gene, the blasticidin resistance gene, or *Hprt* gene, for example[23,101,102].

Another problem is, on average only ~1 in 1,000 G418 resistant clones are products of a homologous recombination event, while the others are random integration events[1]. To enrich for targeting events, inclusion of a negative selection cassette adjacent to one or two ends of the homology arms on the targeting vector can be very useful. During homologous recombination, the negative selection cassette is lost. If random integration occurs, the negative selection cassette is incorporated into the genome. By selecting against cells containing the negative selection cassette, cells with homologous recombination events are enriched. The most widely used negative selection cassette in ES cells is the *HSV-tk* gene[2]. Positive and negative selections are achieved simply by adding drugs in culture medium of ES cells, for example, G418 for *neo* and FIAU for *tk*. Diphtheria toxin gene has also been used as a negative selection agent[103,104]. Despite the enrichment, many of the selected ES clones still contain nonhomologous recombination products. To identify clones that result from homologous recombination, screening using Southern blot analysis or PCR is needed.

Introduction of ES cells into host embryos to create chimera

Introduction of targeted ES cells into host embryos is an important step in the knockout technology[36,37]. This is most commonly performed through microinjection of ES cells into blastocysts. When targeted ES cells contribute to formation of the germ line of chimeric mice, progeny with the desired mutations are obtained. Alternatively, aggregation of ES cells with diploid or tetraploid embryos can be used.

Tetraploid aggregation was developed more than a decade ago[37,94,105,106]. By aggregation of ES cells and developmentally compromised tetraploid embryos, completely ES cells-derived embryos (essentially heterozygotes) can be obtained in a single step—the tetraploid component usually contributes only to the extraembryonic membranes. This aggregation procedure is one generation faster than blastocyst microinjection in terms of obtaining heterozygotes. Although embryos generated in this way using most of the previously derived ES cell lines would not survive to term or maturity, the recently established G4 ES cell line has shown great success in tetraploid aggregation[93].

MATERIALS

REAGENTS

- Plasmids: pKD3 and pKD46 are provided by Dr. Barry Wanner. The other plasmids are created in the Capecchi laboratory. Information on plasmids not listed below (e.g., Cre, IRES-Cre) is available from the authors
- pAP-ACN (GenBank accession number EU530628)
- pECFPpA-ACN (GenBank accession number EU530627)
- pEGFPpA-ACN (GenBank accession number EU530626)
- pEYFPpA-ACN (GenBank accession number EU530630)
- pKD3 (GenBank accession number AY048742)
- pKD46 (GenBank accession number AY048746)
- pnlacZ-ACN (GenBank accession number EU530629)
- pStart-C2 (GenBank accession number EU530622)
- pStart-K (GenBank accession number EU530620)
- pStart-T2 (GenBank accession number EU530621)
- pWS-TK2 (GenBank accession number EU530623)
- pWS-TK3 (GenBank accession number EU530624)
- pWS-TK6 (GenBank accession number EU530625)
- Agarose, GenePure ME (ISC BioExpress, cat. no. E-3121-500)
- Ampicillin (American Pharmaceutical Partners, Inc.)
- Bromophenol Blue (Sigma, cat. no. B7021)
- BSA (Sigma, cat. no. A3912)
- $CaCl_2$ (Sigma, cat. no. C3881)
- Chloramphenicol (Sigma, cat. no. C0378)
- Chloroform (Fisher, cat. no. C298-500)
- DH5α (Invitrogen)
- DMSO (Sigma, cat. no. D8418)
- dNTP mix (Fermentas, cat. no. R0192)
- FIAU (Moravek, cat. no. M251)
- Formamide (Fisher, cat. no. BP227-500)
- G50 columns (ProbeQuant G-50 micro columns; Amersham)
- Gateway LR clonase enzyme mix (Invitrogen, cat. no. 11791-019)
- Glycerol (Fisher, cat. no. BP229-4)
- Geneticin (G418, Invitrogen, cat. no. 11811)
- Herring sperm DNA (Sigma, cat. no. D3159)
- Kanamycin (Sigma, cat. no. K137725G)
- KCl (Fisher, cat. no. BP366-500)
- L-Arabinose (Difco, cat. no. 0159-15)
- LIF (Chemicon, cat. no. ESG1107)
- $MnCl_2$ (Sigma, cat. no. M3634)
- Na_2HPO_4 (Fisher, cat. no. BP393-3)
- NaCl (Fisher, cat. no. S640-10)
- NaH_2PO_4 (Fisher, cat. no. BP329-1)
- NaOH (Fisher, cat. no. BP359-500)
- Phenol (Sigma, cat. no. P4557)
- PIPES (Sigma, P1851-500G)
- Proteinase K (Invitrogen, cat. no. 25530031)
- QIAprep Spin Miniprep Kit (Qiagen, cat. no. 27106)
- QIAquick PCR Purification Kit (Qiagen, cat. no. 28104)
- QIAGEN Plasmid Maxi Kit (Qiagen, cat. no. 12163)
- Random Primer Labeling Kit (Stratagene, cat. no. 300385)
- Restriction enzymes (NEB)
- SDS (Roche, cat. no. 11667262001)
- Shrimp alkaline phosphatase (Roche, cat. no. 1758250)
- Spermidine 3-HCl (Sigma, cat. no. S2501)
- T4 DNA ligase (Fermentas, cat. no. EL0011)
- *Taq* DNA polymerase (Fermentas, cat. no. EL0402)
- TOPO-TA cloning kit (Invitrogen, cat. no. 45-0640)
- Tris–base (Roche, cat. no. 11814273001)
- Tris–HCl (Roche, cat. no. 10812846001)
- Trisodium citrate (Sigma, cat. no. S4641-1KG)
- Xylene Cyanol (Kodak, cat. no. T1579)

EQUIPMENT

- Centrifuge (J2-21M; Beckman)
- Centrifuge (J-6M; Beckman)
- Electroporation device (Gene Pulser Xcell; BIO-RAD)
- Gel documentation system (Alpha Innotech)
- GeneAmp PCR system 9700 (Applied Biosystems)
- Glass capillary (KIMAX-51; Kimble)
- Hybond-N+ nylon membrane (Amersham)
- Hybridization oven (Hybridiser HB-1D; Techne, cat. no.)
- Refrigerated benchtop centrifuge (5417R; Eppendorf)
- UV crosslinker (UV Stratalinker 1800; Stratagene)
- UV transilluminator

REAGENT SETUP

10× Electroporation buffer (for ES cells) 20 mM HEPES, 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 6 mM dextrose. Store at room temperature (RT, 21–24 °C). Good for at least 1 year.

ES cell lysis buffer 100 mM NaCl; 20 mM Tris, pH 7.6; 10 mM EDTA; 0.5% SDS. Store at RT (21–24 °C). Good for at least 1 year.

HotShot lysis solution 25 mM NaOH; 0.2 mM EDTA. Store at RT (21–24 °C) in a plastic bottle. Good for up to 6 months.

Hybridization solution 50% Formamide; 5× SSC; 5× Denhardt; 0.05 M sodium phosphate, pH 6.5; 0.5% SDS; 100 μg ml^{-1} herring sperm DNA. Do not autoclave. Store at RT (21–24 °C). Good for at least 2 months.
10× loading dye 95% Formamide; 10 mM EDTA; 0.025% SDS; 0.17% Xylene Cyanol; 0.17% Bromophenol Blue.
Transformation buffer (TB) 10 mM PIPES, from 0.5 M pH 6.7 stock; 55 mM $MnCl_2$; 15 mM CaCl2; 250 mM KCl. Filter through 0.45-μm filter. Should be made fresh.
Spermidine stock (100 mM, pH 7.0) 127 mg of Spermidine 3-HCl, 5 ml of ddH_2O and 1 drop of 1 M NaOH. Can be aliquoted into 1 ml each in Eppendoff tubes and stored at −20 °C. Good for at least 5 years.
SOB medium For 1 l, Bacto tryptone 20 g, Bacto yeast extract 5 g, NaCl 0.5 g, 1 M KCl 2.5 ml. Autoclave and cool to RT (21–24 °C). Store at RT (21–24 °C). Good for at least 1 year. Just before use, add 10 ml of sterile 1 M $MgCl_2$ for 1 l.
SOC medium Add 2 ml of sterile 1 M glucose to 100 ml of SOB medium. Store at RT (21–24 °C). Good for at least 1 year.
Sodium phosphate buffer 1 M, pH 6.5 Mix 1 M Na_2HPO_4 and 1 M NaH_2PO_4 to obtain pH 6.5. Store at RT (21–24 °C). Good for >3 years.
20× SSC 300 mM Trisodium citrate, 3 M NaCl; adjust pH to 7.0 with a few drops of 10 M NaOH. Store at RT (21–24 °C). Good for >3 years.
Tail lysis buffer 50 mM Tris, pH 8; 100 mM EDTA; 1% SDS; 100 mM NaCl. Store at RT (21–24 °C). Good for >3 years.
50× Tris–acetate–EDTA (TAE) buffer 2 M Tris base, 1 M acetate, 100 mM EDTA. Store at RT (21–24 °C). Good for >3 years.
Tris–EDTA (TE) (pH 7.5) buffer 10 mM Tris, 1 mM EDTA. Store at RT (21–24 °C). Good for at least a year.

PROCEDURE

Preparation of Red-competent bacteria/BAC ● TIMING 1 d

1| Inoculate bacteria containing the BAC clone of interest into 5 ml of SOB, 20 μg ml^{-1} chloramphenicol, and grow at 37 °C for 3–5 h or at 30–32 °C overnight.
? TROUBLESHOOTING

2| Centrifuge the cells at 2,000*g* in J-6M (Beckman) for 5 min at 4 °C.

3| Discard the supernatant. Resuspend the cells (by pipetting up and down) in 1 ml of 10% (vol/vol) ice cold glycerol (prechilled to 0–4 °C in ice-water bath or at 4 °C), transfer into a 1.7-ml microtube and centrifuge at 8,000*g* and 4 °C in a benchtop centrifuge for 10 s.
▲ CRITICAL STEP 10% (vol/vol) ice cold glycerol is prepared ahead of time in milliQ water without autoclave or filtration (precooled to 0–4 °C in ice-water bath or at 4 °C).

4| Discard the supernatant, resuspend the cells in 1 ml of 10% (vol/vol) ice cold glycerol and centrifuge at 8,000*g* and 4 °C in benchtop centrifuge for 10 s.

5| Discard the supernatant and resuspend the cells in 100 μl of 10% (vol/vol) ice cold glycerol; divide into two 50-μl aliquots.
■ PAUSE POINT One aliquot is stored directly in −80 °C as a backup, and the other is kept on ice for immediate use.

6| Transform 10 ng of the Red recombinase-expressing plasmid pKD46 (ref. 40) to one tube of the above freshly made electrocompetent cells. The electroporation conditions are 0.1-cm cuvette, 1.8 kV, 25 μF capacitance and 200 Ω (BIO-RAD, Gene Pulser Xcell).

7| Immediately after the pulse (within seconds), add 300 μl of SOC medium and transfer the cells into a microtube. Without further incubation, directly spread 50 μl and the remaining of the cells on two LB-agar plates (100 μg ml^{-1} ampicillin; 20 μg ml^{-1} chloramphenicol), respectively.

8| Incubate the plates at 30–32 °C for 24–30 h.
▲ CRITICAL STEP pKD46 is temperature sensitive. Do not grow at 37 °C.
? TROUBLESHOOTING

9| Inoculate a single colony from the above plate into a 15-ml tube containing 5 ml of SOB medium (100 μg ml^{-1} ampicillin; 20 μg ml^{-1} chloramphenicol), and incubate the culture at 30–32 °C with shaking (200–300 r.p.m.).

10| After determining the OD_{600} (e.g., 4), inoculate an appropriate amount into a 250-ml flask containing 50 ml of SOB medium (100 μg ml^{-1} ampicillin; 20 μg ml^{-1} chloramphenicol) to reach a final OD_{600} of ~0.1–0.2.

11| To induce the Red recombination system, add L-arabinose (powder or freshly made liquid stock) to a final concentration of 0.1–0.2%, and incubate bacteria at 30–32 °C with shaking (200–300 r.p.m.). Doubling time is ~1.5–2 h.
▲ CRITICAL STEP pKD46 is temperature sensitive. Do not grow at 37 °C.

12| When OD_{600} is 0.4–0.8, transfer the culture into a 50-ml conical tube and leave on ice for 10 min with occasional swirling.

13| Centrifuge the cells for 15 min at 2,000*g* and 4 °C in a precooled J-6M (Beckman) or similar centrifuge.

14| Discard the supernatant, resuspend the cells gently in 50 ml of ice cold 10% glycerol and centrifuge at 2,000*g* and 4 °C for 15 min.

15| Discard the supernatant, resuspend the cells gently in 25 ml of ice cold 10% glycerol and centrifuge at 2,000*g* and 4 °C for 15 min.

16| Repeat Step 15 once.

17| Finally, discard the supernatant and remove any remaining liquid by pipetting. Add 100 µl of fresh, ice cold 10% (vol/vol) glycerol to resuspend the pellet. Divide the cells into 50–100 µl aliquots.
■ **PAUSE POINT** Keep one tube on ice for immediate use. The remaining tubes are frozen directly in −80 °C (no dry ice or liquid nitrogen is required).

Subcloning of genomic fragment into pStart-K vector ● TIMING 2 d

18| Design two oligonucleotides for PCR amplification of the pStart-K vector. For simplicity, throughout this protocol, we use the targeting vector construction for a protocadherin gene, *Pcdh1*, as an example (see **Supplementary Table 1** online for other examples). The upstream oligo sequence is: CCCTATCTCCCAGAACCGGCTATTAGCCTCTGCAGGCTTCCATGCACCTGcgactgaattggttcctttaaagc; and the downstream oligo sequence is: GCGCTGTCTGTATGTCCGGTAGCAAGCACCAGACTTTAAGATATATGTCTgccgcactcgagatatctagaccca. The uppercase sequence within each oligo matches 50 bp at the two junctions of the desired BAC fragment. The lowercase oligo sequence matches the backbone of pStart-K.
▲ **CRITICAL STEP** All of the oligos up to 130 bases in this protocol are synthesized at a normal 40-nmol scale, and no purification such as HPLC or PAGE is necessary. Therefore, sequences such as *lox*P and *FRT* can be conveniently included (**Fig. 4**). (Although mutations may occur during synthesis of long oligos, the correct ones appear to have some advantage to be selected during recombination.)

19| Use these two oligos to amplify the plasmid template pStart-K by PCR (5× 25-µl reactions: pStart-K, 50 ng; 10× buffer, 12.5 µl; 25 mM $MgCl_2$, 10 µl; primers, 2.5 µl each at 10 µM; 10 mM dNTP, 2.5 µl; *Taq*, 1.25 µl; H_2O is added to a total volume of 125 µl). PCR conditions are 94 °C 2 min; 30 cycles of 94 °C 30 s, 59 °C 30 s and 72 °C 90 s; 72 °C 7 min. Examine 2 µl of PCR products on a 1% agarose gel by electrophoresis.
▲ **CRITICAL STEP** For PCR involving long oligos, it is better to use regular *Taq* polymerase, because proofreading polymerases (e.g., Pfu, Herculase, Expand and so on) tend to be much less efficient. Although *Taq* is more likely to introduce mutations during amplification, as long as the junctions are sequenced when the clones are obtained, it does not affect the genomic fragment captured.
? TROUBLESHOOTING

20| Combine the amplified PCR products; purify them using a single QIAprep spin column or QIAquick PCR purification column following the manufacturer's instruction. In brief, mix PCR products with five times (volume) PB buffer and load onto a column and centrifuge. Wash once with PE buffer and elute with 100 µl H_2O.

21| Digest the entire eluted DNA with 2 µl (20 U) of *Dpn*I in appropriate buffer for 1–2 h at 37 °C. Re-purify the reaction on a Qiagen column as in Step 20 and elute it with 40 µl H_2O. DNA concentration is normally ~200 ng μl^{-1} or higher.

22| Electroporate the purified PCR product (5–10 µl) into 50 µl of the Red-competent bacteria containing the desired BAC (from Step 17). The parameters for electroporation are 0.1-cm cuvette, 1.8 kV, 25 µF capacitance and 200 Ω.
? TROUBLESHOOTING

23| Immediately after the pulse, add 1 ml of SOC medium, transfer the cells into a 1.7-ml microtube, incubate at 37 °C with shaking (200–300 r.p.m.) for 1 h and spread the cells onto LB-agar plates (50 µg ml^{-1} kanamycin).
? TROUBLESHOOTING

24| Pick eight small Kan^r colonies and grow in 5 ml of SOB medium at 37 °C overnight for preparation of DNA minipreps using QIAprep spin columns following the manufacturer's instruction.
▲ **CRITICAL STEP** Usually several hundred to thousand colonies appear on the plate with variations in size. Try to avoid picking large colonies because they are usually self-recombined pStart-K. Use QIAprep spin columns for DNA minipreps, as plasmid DNA prepared without columns can often have contaminations of BAC DNA (Cam^r) that can affect the use of chloramphenicol resistance gene for later steps in the procedure.

25| Analyze these DNA preps by restriction enzyme digest to identify the ones that have successfully captured the genomic fragment (in this case *Pcdh1*) into pStart-K: for a 20 µl reaction in a 1.7-ml microtube, add 5 µl (~100–200 ng) of the DNA, 1–2 U (usually 0.1–0.2 µl for most restriction enzymes) of restriction enzyme, appropriate buffer and add H_2O to 20 µl. Incubate at 37 °C or other appropriate temperature for the enzyme for 1–2 h. Separate the reaction on 1% agarose gel. Clones with expected pattern are designated pStart-K-Pcdh1.
? TROUBLESHOOTING

26| Sequence two pStart-K-Pcdh1 clones from above to further confirm the junction regions of captured fragment, that is, the *att*L1 and *att*L2 sites and the first and last 50 bp of the genomic fragment. Universal primers for the sequencing reactions are WS275 and WS276 (present on the backbone of pStart-K; see **Supplementary Table 1** for oligonucleotide sequences).

Preparation of chemically competent DH5α cells ● TIMING 2 d

27| Even with recombineering, standard restriction enzyme-based cloning is still needed. Chemically super-competent DH5α cells ($>1 \times 10^8$ transformants per μg DNA) are essential for everyday cloning of targeting vectors. We follow the Inoue method[107] with slight modifications for preparation of chemically competent DH5α cells.

28| Inoculate DH5α bacteria (from a stock or a tube of competent cells) using an inoculating loop onto an LB-agar plate and incubate at 37 °C overnight.

29| Pick a single colony and grow in 5 ml SOB medium at 37 °C overnight.

30| After determining the OD_{600} (e.g., 4), transfer ~1 ml into 250 ml of SOB medium in a 2-l flask to obtain OD_{600} of ~0.01–0.02. Incubate the culture at RT (21–24 °C) with shaking (200–300 r.p.m.). Doubling time is ~3–4 h.

31| When the OD_{600} of the cells reaches ~0.5–0.6, transfer the culture into a large centrifuge tube and leave on ice for 10–20 min with occasional shaking.

32| Centrifuge the bacteria at 2,500*g* and 4 °C for 10 min. Gently resuspend the pellet in 80 ml of ice cold TB, leave on ice for 10 min.

33| Centrifuge the bacteria at 2,500*g* and 4 °C for 10 min. Gently resuspend the pellet in 20 ml of ice cold TB.

34| Add 1.4 ml of DMSO to the tube and mix well. Keep the tube on ice for 10 min.

35| Divide the bacteria into 0.5–1 ml aliquots, freeze in liquid nitrogen and store at −80 °C.

■ **PAUSE POINT** If stored at −80 °C, competent cells are good for at least 3 years. When thawed for use, the 0.5–1 ml aliquot can be further aliquoted and refrozen without noticeable reduction of transformation efficiency. This procedure produces tenfold more competent cells than many other commonly used protocols, yet the efficiency of the cells is still as high as 1×10^8 to 5×10^8 transformants per μg plasmid DNA.

Preparation of Red-competent DH5α/pKD46 cells ● TIMING 2 d

36| Transform 50 μl of chemically competent DH5α cells (from Step 35) with pKD46 (~10 ng) by 42 °C heat shock for 1 min. Right after heat shock, directly spread the transformed cells onto an LB-agar plate (100 μg ml^{-1} ampicillin) and grow at 30–32 °C for ~24–30 h.

▲ **CRITICAL STEP** pKD46 is temperature sensitive. Do not grow at 37 °C.

37| Pick several big colonies and grow in 20 ml of SOB medium overnight at 30–32 °C.

38| After determining the OD_{600} (e.g., 4), inoculate an appropriate amount (~10 ml) into a 1-l flask containing 250 ml of SOB medium to reach a final OD_{600} of ~0.1–0.2 and add 0.25 g L-arabinose powder to the culture. Incubate the culture at 30–32 °C with shaking (200–300 r.p.m.) for another 2–4 h.

39| When the OD_{600} of the cells reaches ~0.4–0.8, transfer the culture into a large centrifuge tube and leave on ice for 10–20 min with occasional shaking.

40| Centrifuge the cells for 5 min at 4,000*g* and 4 °C. Discard the supernatant, and resuspend the pellet gently in 200 ml of ice cold 10% (vol/vol) glycerol.

41| Centrifuge the cells for 10 min at 4,000*g* and 4 °C. Discard the supernatant, and resuspend the pellet gently in 200 ml of ice cold 10% (vol/vol) glycerol.

42| Centrifuge the cells for 10 min at 4,000*g* and 4 °C. Discard the supernatant, and resuspend the pellet gently in 100 ml of ice cold 10% (vol/vol) glycerol.

43| Finally centrifuge the cells at 4,000*g* for 10 min at 4 °C. Discard the supernatant, and remove the residual liquid by pipetting. Gently resuspend the pellet in a fresh 0.5 ml of ice cold 10% (vol/vol) glycerol. Divide the cells into 50 μl aliquots; snap freeze in liquid nitrogen.

▲ **CRITICAL STEP** If possible, all the steps in this section are performed in a 4 °C cold room to ensure a very high transformation efficiency. If these cells are needed for many targeting vectors, the procedure above can be easily scaled up.

■ **PAUSE POINT** Cells are stored at −80 °C (good for at least 3 years).

Insertion of positive selection *neo* cassette ● **TIMING 3 d**

44| For the simple loss-of-function allele of *Pcdh1*, for example, design oligos to delete the *Pcdh1* exon 1 and concomitantly introduce an *Asc*I restriction enzyme site at this position. The upstream oligo sequence is GTCTTCTTGTAGTTCTCCTGATTCTGGAGCCTGCCAGGATGGGGCCTCTGA GGCGCGCCagcattacacgtcttgagcgattgt and the downstream oligo sequence is CTCCCTCATGATCTAGTCGATCATGGCGGGTAAGACACACCTGCTCTATCAGGCGCGCCcacttaacggctgacatgggaatta. Uppercase sequences are 50 bases homology to flanking genomic DNA of *Pcdh1* exon 1—homology arms for Red-recombination. The GGCGCGCC is the *Asc*I consensus sequence. Lowercase sequences are primers for the chloramphenicol resistance gene (*cat*) in the plasmid pKD3 (ref. 40). Although other *cat*-containing plasmids can also be used as a PCR template, the defective origin of replication (oriRγ) present on pKD3 reduces potential carrier-over background for Steps 48–50 below.
▲ **CRITICAL STEP** Other enzyme sites can also be used. We choose to use an *Asc*I site because all the other reporter/*neo* cassettes (**Fig. 2e**) are similarly flanked by *Asc*I. Because *Asc*I is a rare cutter and its overhang is compatible with a few other restriction enzymes such as *Mlu*I, it is possible to use *Asc*I-flanked reporter/*neo* for virtually any gene.
? TROUBLESHOOTING

45| Perform PCR using these two oligos to amplify the chloramphenicol resistance gene (*cat*) in the plasmid pKD3 (5× 25-µl reactions: pKD3, 50 ng; 10× buffer, 12.5 µl; 25 mM $MgCl_2$, 10 µl; primers, 2.5 µl each at 10 µM; 10 mM dNTP, 2.5 µl; *Taq*, 1.25 µl; H_2O is added to a total volume of 125 µl). PCR conditions are 94 °C 2 min; 30 cycles of 94 °C 30 s, 59 °C 30 s and 72 °C 60 s; 72 °C 7 min.

46| Combine and purify PCR products through a QIAprep spin column as described in Step 20. Elute with 100 µl H_2O.

47| Digest the entire eluted DNA with 2 µl (20 U) of *Dpn*I in appropriate buffer for 1–2 h at 37 °C. Repurify the reaction on a Qiagen column as in Step 20 and elute it with 40 µl H_2O. DNA concentration is normally ~100 ng μl^{-1} or higher.

48| Electroporate the purified PCR product (2–5 µl, ~200–500 ng) plus pStart-K-Pcdh1 (2–5 µl, ~100–500 ng; from Step 26) into 50 µl of Red-competent DH5α/pKD46 cells (from Step 43). Electroporation conditions are 0.1-cm cuvette, 1.8 kV, 25 µF capacitance and 200 Ω.

49| Transfer the electroporated cells into 1 ml of SOC medium and incubate the mixture at 37 °C with shaking (200–300 r.p.m.) for 1 h. Centrifuge briefly at 20,000*g* for 3–5 s at RT. Discard supernatant and resuspend the pellet in the remaining liquid of ~100 µl. Spread the bacteria onto LB-agar plates (30 µg ml^{-1} chloramphenicol) and incubate at 37 °C overnight.
▲ **CRITICAL STEP** For some DNA that is difficult to maintain at 37 °C, incubation can also be carried out at 30–32 °C.

50| Pick four medium-large colonies (>90% are recombinants) and grow in 5 ml of SOB medium at 37 °C overnight for DNA minipreps (as described in Step 24). Analyze these DNA preps by restriction enzyme digest as in Step 25 to identify correct clones.
? TROUBLESHOOTING

51| To confirm the presence of the predicted junction regions, sequence two minipreps with the correct restriction patterns using primers WS187 and WS188 (see **Supplementary Table 1** for oligonucleotide sequences). Designate the resulting plasmid pStart-K-Pcdh1Asc.

52| Cut 2–5 µg of pStart-K-Pcdh1Asc in a 50 µl reaction with 1 µl (10 U) of *Asc*I restriction enzyme. Incubate at 37 °C for 1–2 h.

53| Purify the reaction using a Qiagen column as in Step 20 and elute with 48 µl H_2O.

54| Mix the eluted DNA (~43 µl) with 5 µl of 10× Shrimp AP buffer and 2 µl of Shrimp AP (Roche) in a total volume of 50 µl for dephosphorylation at 37 °C for 10 min. Inactivate the reaction at 65 °C for 15 min. This inactivated DNA can be directly used for ligation.

55| Set up a standard ligation to insert a pre-cut *Asc*I-EGFP-ACN-*Asc*I cassette or other reporter cassettes (**Fig. 2e**) as follows: 12 µl (~200–500 ng) of above purified recipient DNA (Step 54), 5 µl (~200 ng) of pre-cut *Asc*I-EGFP-ACN-*Asc*I cassette, 2 µl of T4 DNA ligase buffer and 1 µl of T4 DNA ligase in a total volume of 20 µl for ligation reaction at RT (21–24 °C) for 2 h.

56| Transform chemically competent DH5α cells (100 µl, $>10^8$ transformants per µg; from Step 35) with 5–10 µl of the above ligation reaction mixture by heat shock at 42 °C for 1 min. Add 1 ml of SOC medium, and incubate at 37 °C with shaking (200–300 r.p.m.) for 1 h. Centrifuge briefly at 20,000*g* for 3–5 s at RT. Discard supernatant and resuspend the pellet in the remaining liquid of ~100 µl. Spread on LB-agar plates (50 µg ml^{-1} kanamycin). Pick four to eight colonies for culture and preparation of DNA minipreps. Perform restriction enzyme digest as in Step 25 to identify correct clones. Designate the clones with expected restriction pattern pStart-K-Pcdh1-EGFP (**Fig. 2f**).
? TROUBLESHOOTING

Gateway recombination: adding negative selection *tk* cassette ● TIMING 2 d

57| Finally, introduction of the *HSV-tk* gene into the targeting vector requires a Gateway recombination reaction. Set up a reaction mixture that contains: LR reaction buffer (5×), 1 μl; pStart-K-Pcdh1-EGFP (e.g., 2 μl (from Step 56)); pWS-TK6/ linearized with *Sal*I, 1 μl; LR clonase enzyme mix, 1 μl.

58| Incubate at 25 °C for 1 h, add 0.5 μl of proteinase K solution and incubate the reaction for 10 min at 37 °C.

59| Transform chemically competent DH5α cells (100 μl, $>10^8$ transformants per μg; from Step 35) with 2 μl of the above reaction mixture (from Step 58) by heat shock at 42 °C for 1 min. Without adding SOC medium or incubation, spread 10 and 90 μl of the transformed bacteria on LB-agar plates (100 μg ml^{-1} ampicillin).

60| Incubate the plates at 30–32 °C for 20–30 h. Pick two colonies and grow in 5 ml of SOB medium at 30–32 °C for ~20 h for preparation of DNA minipreps using QIAprep spin columns following the manufacturer's instruction.
▲ **CRITICAL STEP** Do not incubate at 37 °C or higher. Usually >90% colonies are correct.
? TROUBLESHOOTING

61| Analyze these DNA preps by restriction enzyme digest as in Step 25 to identify correct clones, which are designated pTV-Pcdh1-EGFP.

62| Further prepare large quantity of pTV-Pcdh1-EGFP DNA (~500 μg) using QIAGEN Plasmid Maxi Kit following the manufacturer's instruction.

63| Perform extensive restriction enzyme digest of the final targeting vector, and sequence important regions (e.g., check reporter sequence to see whether it is in frame as designed) as a quality control step.

Preparation of DNA for electroporation of ES cells ● TIMING 6 h

64| To linearize targeting vectors for electroporation of ES cells, prepare ~100–150 μg of clean DNA by digesting 200 μg of targeting vector DNA (from Step 62; assuming a 70% recovery) with appropriate restriction enzyme at 1–2 U per μg DNA in a total volume of 500-μl reaction for ~4 h.

65| To determine the completeness of the digest by agarose gel electrophoresis, run ~50 ng of digested DNA along with 50 ng of uncut targeting vector and ladder in 0.8% agarose gel in 1× TAE at 100 V for ~1 h. If the gel is run long enough, uncut DNA should be found to migrate at a slower speed.
▲ **CRITICAL STEP** Do not load >50 ng of DNA on the agarose gel to check completeness of digest; otherwise the band is not sharp and looks like smear.

66| To purify the 500 μl DNA digest (from Step 64), add 1 volume of phenol and 1 volume of chloroform. Hand-shake the tube vigorously, and centrifuge at 20,000*g* for 3–5 min at RT in a benchtop centrifuge.

67| Transfer the supernatant into a new tube, and add an equal volume of chloroform. Hand-shake the tube vigorously, and centrifuge at 20,000*g* for 3–5 min at RT.

68| Transfer the supernatant into a new microtube. Add 1.6–2 volumes of ethanol (without adding salt). Mix the tube by gentle inversions, and DNA becomes web-like.

69| Centrifuge the tube at 3,000*g* for 2 min, and discard the supernatant.
▲ **CRITICAL STEP** Higher centrifugation speed makes the pellet hard to dissolve later.

70| Rinse the DNA pellet with 0.5 ml of 70% ethanol by shaking. Centrifuge the tube at 3,000*g* for 2 min, and discard the supernatant.

71| Centrifuge the tube briefly at 3,000*g* for 10 s, and remove residual liquid by pipetting. The linearized DNA pellet needs to be fully dissolved (by gentle pipetting up and down) in 100 μl of TE (pH 7.5, filtered).
▲ **CRITICAL STEP** Do not vortex DNA. If DNA is difficult to dissolve, it can be incubated at 55 °C for 15 min. If DNA is not fully dissolved in TE, it will affect targeting efficiency in ES cells.

72| To determine the concentration, use 1 μl of resuspended DNA to check $OD_{260/280}$. Adjust the concentration to 1 μg μl^{-1}.
■ **PAUSE POINT** Store the linearized DNA at 4 °C until use (long-term storage should be at −20 °C).

Electroporation of targeting vector DNA into ES cells ● TIMING 3 weeks

73| Electroporate linearized targeting vector into ES cells (either R1 or G4 cells)[93,94] according to a previously published method[1,82] with modifications as follows. In brief, grow ES cells on mouse embryonic fibroblast feeder plates with leukemia

inhibitory factor[108]. Perform electroporation at RT (21–24 °C) by pulsing two times at 650 V on a capacitor discharge instrument. Immediately transfer cells into regular DMEM. After 24–28 h, change medium to DMEM containing G418 (380 μg ml^{-1}) and FIAU (1 × 10^{-7} M) selection drugs for *neo* and *tk*, respectively. After ~6–9 d, pick individual cell clones into DMEM containing only G418 (330 μg ml^{-1}) (no FIAU).

▲ **CRITICAL STEP** Find out ideal electroporation conditions for your machine. We use Promega Biotec X-cell 2000 in the Capecchi laboratory. Also, different ES cells demonstrate markedly different sensitivities for each of the commonly used selection drugs, such as G418, hygromycin, puromycin, zeocin, blasticidin, FIAU and so on. Selection is achieved by adding drugs into culture medium. For each drug, selection parameters (dose and timing) for different ES cells need to be established before starting experiments with targeting vectors. In our hands, the F1 hybrid G4 cells[93] usually are superior than the R1 cells[94] in terms of targeting frequency.

Preparation of genomic DNA from ES cells for Southern blot or PCR ● **TIMING 8 h**

74| To prepare DNA from ES cells for Southern blot analysis or PCR, lyse cells in 1.7-ml microtubes in ~500 μl of ES cell lysis buffer. (Alternatively, prepare DNA in a 96-well format[37].)

75| Add 5 μl of proteinase K (from 20 mg ml^{-1} stock) to a final concentration 0.2 mg ml^{-1}, and incubate the tubes at 37 °C for 2–4 h without shaking.

76| Add 250 μl of saturated NaCl (~6 M) to each tube, and handshake vigorously 100–200 times.

77| Leave the tubes on ice for 10 min, and centrifuge on a benchtop centrifuge at 20,000*g* for 10 min at 4 °C.

78| Collect supernatant (~700 μl) into a new 2-ml microtube. Add ~1.2 ml 100% ethanol.

79| Use a glass capillary for each tube to spool out the DNA. The DNA normally sticks to the tip of glass capillary. Alternatively, DNA can be spun down at 20,000*g* for 2 min at RT (21–24 °C).

80| Wash the DNA briefly by dipping into another tube containing 70% ethanol. If DNA was spun down in Step 79, wash the DNA with 0.5–1 ml of 70% ethanol.

81| Dissolve each DNA in 150 μl TE (pH 7.5). If difficult to dissolve, incubate at 55 °C for 30 min and then hand-shake well.

■ **PAUSE POINT** Quality of DNA prepared by this procedure is sufficient for most Southen or PCR analyses. For higher quality, standard phenol:chloroform extraction can be used to further purify DNA.

Screening of ES cell clones

82| Screening of ES cells can be carried out by Southern blotting (option A) or PCR (option B).

(A) Southern blot screening of ES cells ● **TIMING 3 d**

(i) Prepare radiolabeled probe using Stratagene's Prime-It II Random Primer Labeling Kit or the Ready-To-Go DNA Labelling Beads (Amersham) following the manufacturer's instructions.

(ii) Purify probes with G50 columns (e.g., ProbeQuant G-50 Micro Columns from Amersham).

(iii) Set up the DNA digest as follows: 5–10 μg of genomic DNA (usually 10–15 μl) are digested with restriction enzyme (10–20 U) in 25-μl reactions in the presence of 4 mM spermidine for 12–20 h at an appropriate temperature for the enzyme.

(iv) To separate the digested DNA, add 3 μl of 10× loading dye, heat the samples at 65 °C for 10 min and load them on 0.8–1% agarose gel in 1× TAE. DNA standard (e.g., 1 kb plus DNA from Invitrogen) should also be included.

(v) Electrophorese the samples overnight at 20–50 V. Photograph the gel under UV light along with a fluorescent ruler.

(vi) For more efficient transfer of DNA fragment >20–25 kb, treat the gel with 0.2 M HCl for 10 min, or UV irradiate the gel for 5 min.

(vii) Transfer DNA into Hybond-N+ nylon membrane (Amersham) by downward capillary transfer using a protocol described previously[55,109]. In brief, denature the gel in 3 M NaCl, 0.4 M NaOH for 1 h and blot DNA onto Hybond N+ membrane in 3 M NaCl, 8 mM NaOH for 2 h to overnight.

(viii) Neutralize the nylon membrane in 0.2 M sodium phosphate, pH 6.5 and UV-crosslink with UV Stratalinker 1800 for 100 s (set at AUTO CROSSLINK).

(ix) Use the wet membrane directly for hybridization.

■ **PAUSE POINT** Alternatively, dry the membrane at RT (21–24 °C) for later use. Store the membrane at RT (21–24 °C) in a plastic bag. Good for at least a year.

(x) To hybridize, place the membrane into a large hybridization tube and prehybridize it in an oven at 42 °C in ~25 ml of hybridization solution.

(xi) After 1 h of prehybridization, discard the solution.

(xii) Hybridize the membrane using a radioactive probe (see INTRODUCTION for information on design and preparation of probes. The probe is boiled for 10 min and snap cooled on ice before use) in ~15 ml hybridization solution at 42 °C overnight.

(xiii) Wash the hybridized membrane twice in 2× SSC, 0.1% SDS at RT (21–24 °C) for 15 min and once in 0.2× SSC, 0.1% SDS at 42 °C for 10–15 min.

(xiv) Place the membrane into a Kodak film cassette with film at −80 °C for overnight to several days. On the other hand, place the membrane into a phosphor imaging cassette at RT (21–24 °C) for a few hours to overnight and visualize using Typhoon PhosphorImager (e.g., see **Fig. 5**).

(v) The hybridized membrane may also be stripped for hybridization with another probe. To strip the membrane, place it in 0.5% SDS, 100 °C until the signal disappears (from a few seconds to several minutes).

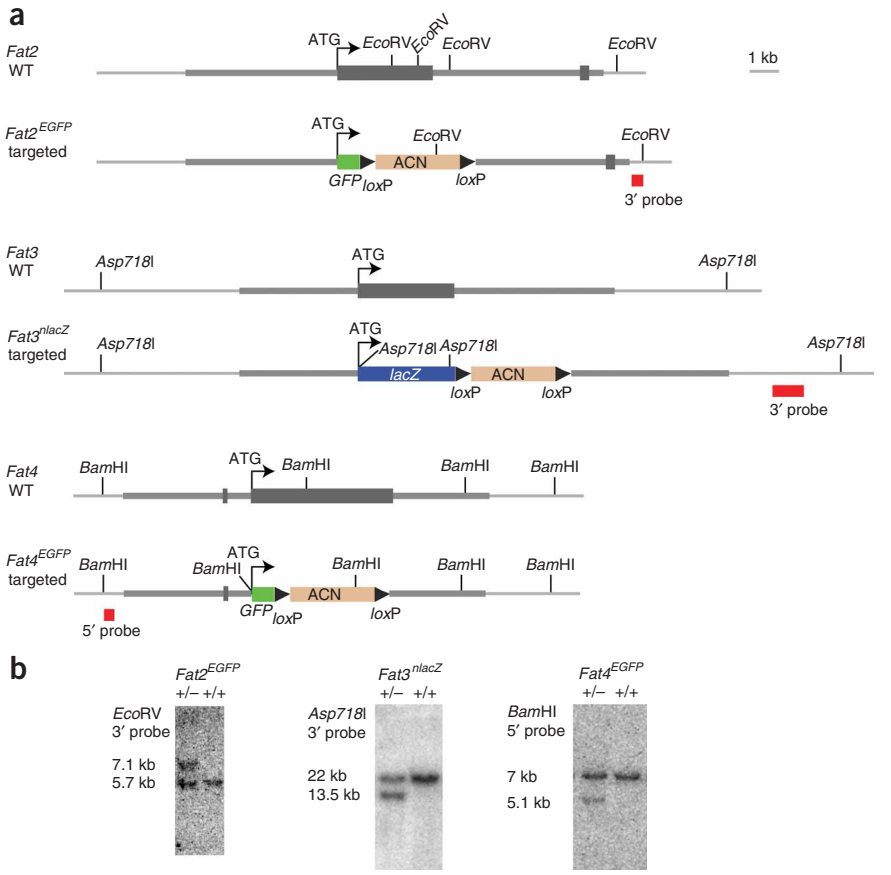

Figure 5 | Three examples of the knockout mice generated: *Fat2*^{EGFP}, *Fat3*^{nlacZ} and *Fat4*^{EGFP} knockout mice. (**a**) Targeting strategies for *Fat2*^{EGFP}, *Fat3*^{nlacZ} and *Fat4*^{EGFP} alleles. (**b**) Southern blot analysis of embryonic stem (ES) cells for *Fat2*^{EGFP}, *Fat3*^{nlacZ} and *Fat4*^{EGFP} alleles. *Fat2*, *Fat3* and *Fat4* are three giant protocadherin genes, each encoding a protein with >4,000 amino acids[114–116].

(B) PCR screening of ES cells

● TIMING **1 d**

(i) The Roche Expand long template PCR kit is used for all PCRs. Dilute 2 µl of genomic DNA prepared from ES cells (Step 81) in 20 µl of HotShot lysis solution.

(ii) Boil for 15 min, and neutralize with 20 µl of 40 mM Tris–Cl.

(iii) Use 2 µl of the neutralized DNA in a 12.5–µl PCR. PCR is performed according to the manufacturer's instruction.

▲ CRITICAL STEP If primers are not designed well, PCR can be difficult. We usually design several pairs of primers for test. It is important to find a good flanking primer that is outside of the homology arms (e.g., common primer P3 in **Fig. 2a**). Control primers P1 and P3 are designed to amplify a wild-type fragment of about the same size of the mutant fragment. If P1/P3 can readily amplify the expected fragment, P3 is then likely to be a good flanking primer. We try to use the same good internal primers (e.g., primers from *neo* or *GFP*).

Microinjection of ES cells into blastocysts

83| Microinject ES cells into host blastocysts as described previously[36].

Genotyping of knockout mice

84| Genotyping of knockout mice can be carried out by Southern blotting (option A) or PCR (option B). To confirm germline transmission of a gene-targeted allele, Southern blot analysis is usually needed. Once confirmed by Southern blot, PCR is used for routine genotyping of mice from subsequent breeding.

(A) Preparation of tail (or other tissue) DNA for Southern blot ● TIMING **1 d**

(i) To isolate genomic DNA from tails or other tissues for Southern blot analysis and PCR, put tails (~0.5 cm) into 1.7-ml microtubes containing 480 µl of tail lysis buffer.

(ii) Add 25 µl of proteinase K (from 20 mg ml^{-1} stock) and incubate the mixture at 55 °C overnight.

(iii) Add 0.25 ml of 6 M NaCl to each tube. Leave on ice for 10 min.

(iv) Shake the tubes vigorously for 2 min and centrifuge at 20,000*g* for 10 min at 4 °C.

(v) Transfer the supernatant into a new microtube, and add ~1 ml 100% ethanol.

(vi) Use a capillary tube to spool out the DNA, and wash by dipping in a tube containing 70% ethanol. Alternatively, DNA can be spun down at 20,000*g* for 2 min at RT (21–24 °C) and washed with 0.5–1 ml of 70% ethanol.

(vii) Dissolve DNA in 200 µl TE in a 1.5-ml microtube.
(viii) Perform Southern blot as described in Step 82.

(B) Quick preparation of tail (or other tissue) DNA for PCR genotyping of knockout mice ● TIMING **2 h**

(i) Cut 1–1.5-mm long tails (or other tissues, such as yolk sac for embryos) and boil in 100 µl Hotshot lysis solution for 1 h; neutralize with 100 µl of 40 mM Tris–Cl[110].

(ii) Use an aliquot of 2 µl for PCR genotyping in a 12.5-µl reaction.

● TIMING

Steps 1–17, preparation of Red-competent bacteria/BAC: 1 d
Steps 18–26, subcloning of genomic fragment into pStart-K vector: 2 d
Steps 27–35, preparation of chemically competent DH5α cells: 2 d
Steps 36–43, preparation of Red-competent DH5α/pKD46 cells: 2 d
Steps 44–56, insertion of positive selection *neo* cassette: 3 d
Steps 57–63, Gateway recombination: adding negative selection *tk* cassette: 2 d
Steps 64–72, preparation of DNA for electroporation of ES cells: 6 h
Step 73, electroporation of targeting vector DNA into ES cells: ~3 weeks
Steps 74–81, preparation of genomic DNA from ES cells for Southern blot or PCR: 8 h
Step 82, screening of ES cell clones: 1–3 d
Step 83, microinjection of ES cells into blastocysts: 1 d
Step 84, Genotyping of knockout mice: 1–3 d

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2**.

TABLE 2 | Troubleshooting table.

Steps	Problem	Possible reason	Solution
1	Bacteria did not grow	High chloramphenicol concentration	Ensure correct chloramphenicol concentration is used
8	No colony on plate	High chloramphenicol concentration	Ensure correct chloramphenicol concentration is used
		High incubation temperature	Check whether incubator temperature is correctly set at 30–32 °C
		Not enough pKD46	Try to use more pKD46 (up to 50 ng) in Step 6
19	Little or no PCR product.	High-fidelity polymerase used	Ensure *Taq* polymerase is used
		Not enough elongation time	Ensure enough elongation time for PCR amplification is used
		Annealing temperature too high	Try to lower annealing temperature (e.g., 57.5 °C)
22	Electric arcing during electroporation	Salt in PCR products	Ensure purified PCR product is dissolved in H_2O instead of Tris–EDTA
			Try to lower the volume of PCR product used
			Try to purify PCR with standard Phenol:chloroform extraction
			Try to use slightly more competent bacterial cells
23	Few colonies on plate	LB-agar plate with wrong antibiotics used	Ensure only kanamycin is used for selection (no ampicillin or chloramphenicol)
		Bacteria not competent enough	Ensure bacteria made in Step 17 are competent ($>10^9$ transformants per µg)

TABLE 2 | Troubleshooting table (continued).

Steps	Problem	Possible reason	Solution
	Bacteria grew as lawn on plate, no single colonies could be picked	Low kanamycin concentration	Ensure kanamycin concentration is correct (50 μg ml^{-1}) Try to increase kanamycin concentration to 75 μg ml^{-1}
25	Low yield of DNA	Not enough growth time	Try to increase culture time up to 20 h in SOB at 37 °C
	Only self-recombined pStart-K vector, no expected recombinants	Self-recombination between the *att*L1 and *att*L2 of pStart-K	To capture difficult genomic regions, it is better to leave out the almost identical *att*L1 and *att*L2 sites (identities = 93/95) in pStart-K in the PCR product (that is, if oligonucleotides are designed to match red arrows in **Fig. 2c**), and the background of self-recombined pStart-K is greatly reduced
		Wrong orientation of homology arms in oligos	Ensure 5′ oligo uses reverse sequence and 3′ oligo uses forward sequence
44	Little or no PCR product	See troubleshooting for Step 19	See troubleshooting for Step 19
50	Unexpected restriction pattern	Concatemer formation	Although concatemer formation is a common problem when high-copy plasmids (containing ColE1, pUC or similar origins of DNA replication) are modified by Red-mediated recombination[19,21], the use of a low-copy plasmid such as pStart-K usually eliminates this problem Pick more colonies to find one without concatemer Concatemer can often be separated by retransformation A more efficient solution is to digest the concatemer DNA with a restriction enzyme that cuts each vector backbone once and only once, re-ligate and perform another transformation of DH5α bacteria
56	Few or no colonies	Not enough DNA in ligation reaction	Ensure the efficiency of DH5α cells is $>10^8$. For ligation of a large construct, it is better to use a larger amount of DNA, for example, a few hundred nanograms of DNA for the backbone vector and the insert
60	Few colonies or correct clones	High incubation temperature	Ensure 30–32 °C is used for incubation of the plate and culture

ANTICIPATED RESULTS

Targeting cadherin gene family

We have recently reported the generation of mouse loss-of-function alleles for individual members of the protocadherin gene clusters, as well as long-range deletions that cover 14 *Pcdha* genes and 22 *Pcdhb* genes[46]. We designed targeting strategies for disrupting 34 additional classic cadherin and protocadherin genes[49,51,111–113] (**Table 1**, **Figs. 5** and **6**), following the above procedure. To create these targeting vectors, we started in each case with BAC clones containing the desired genomic regions.

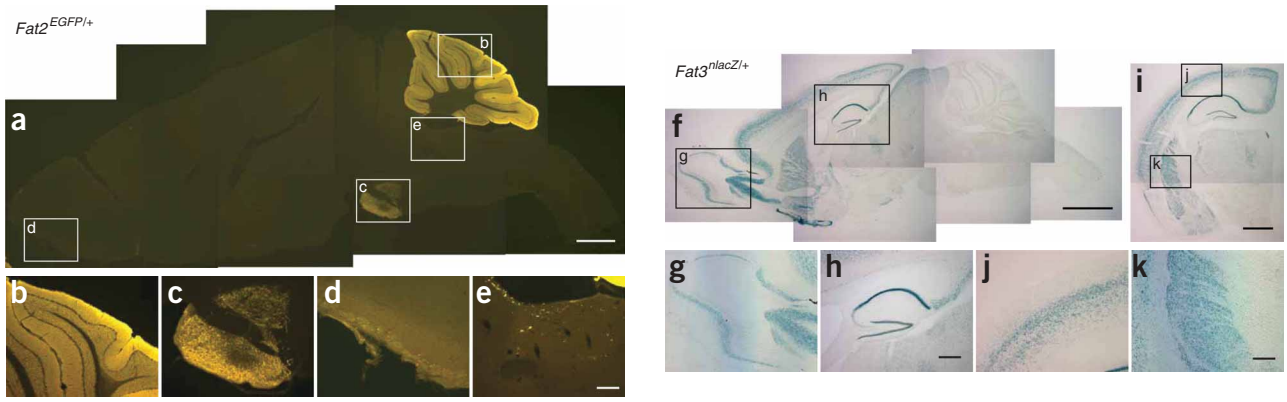

Figure 6 | Expression patterns of *Fat2* and *Fat3* protocadherin genes revealed by knockin reporter alleles. (**a**–**e**) Fat2 expression pattern. (**a**) Fat2 is mainly expressed in the cerebellum (**b**) and pontine nuclei (**c**) as shown in a sagittal section. (**d**,**e**) Weak expression is also observed in distinct neurons in the glomerular layer of the olfactory bulb (**d**) and in the pontine central gray and the vestibular nucleus (**e**). (**f**–**k**) Fat3 expression pattern. Fat3 expression is strong in the forebrain, but weak in the midbrain and hindbrain in a sagittal section (**f**). (**g**,**h**) Many regions in the forebrain, including the mitral cell layer of the olfactory bulb and anterior olfactory nuclei (**g**) and the hippocampus (**h**), have strong Fat3 expression. In the hippocampus, Fat3 is expressed strongly in the CA1 and the dentate gyrus but at a very low level in the CA3 and dentate hilus (**h**). (**i**–**k**) Fat3 shows a layer-specific expression pattern in the cortex in a coronal section (**i**). Strong labeling is mainly localized to layers 4 and 5 of the neocortex (**j**). Strong signals are also observed in striatum (**k**). Scale bars: (**a**,**i**), 1 mm; (**b**–**e**,**j**,**k**), 200 μm; (**f**), 2 mm; (**g**,**h**), 400 μm.

The modular nature of our cloning method made it possible to handle dozens of targeting vectors at a time. We were able to subclone genomic regions for these genes into the pStart-K vector with an average efficiency of >40% (four of ten colonies picked contained the genomic region of interest) (**Fig. 2c**). For the next step (**Fig. 2d**), we used Red-competent DH5α/pKD46 cells (see Procedure Steps 36–43). With only 50 bp homology for recombination, this step (**Fig. 2d**) had >90% efficiency. We next used restriction enzyme-based cloning to insert the reporter/*neo* cassette into the *Asc*I site (**Fig. 2e**). Although this was nondirectional cloning (single *Asc*I site), almost all transformants in this ligation step contained an insert, with ~50% having the desired orientation. For the final step (**Fig. 2f**,**g**), we could choose from a series of TK vectors (**Fig. 3b**). Because Gateway recombination (Invitrogen) is very efficient, virtually every clone contained the desired insert (that is, homology arms with reporter/*neo* cassette). When these plasmid-based vectors were electroporated into mouse ES cells, the overall targeting efficiency was 7.0%, higher than the reported targeting frequency obtained with targeting vectors generated from intact BACs (3.8%)[26]. So far, germline transmission has been obtained for 12 of the targeted cell lines (**Table 1**). Overall, G4 ES cells have a >90% success rate in germline transmission using blastocyst microinjection method, for each positive ES cell clone injected.

To further compare the different reporter/systems *in vivo*, we analyzed the expression patterns of these knockout alleles. In general, all of the reporter/*neo* cassettes tested could be used to examine the endogenous gene expression pattern. However, different reporters showed different properties *in vivo* (**Figs. 6**–**8**). Our results suggest that EGFP and nlacZ (lacZ containing a nuclear localization signal) reporters are superior to others tested. Although EGFP can usually be observed directly using

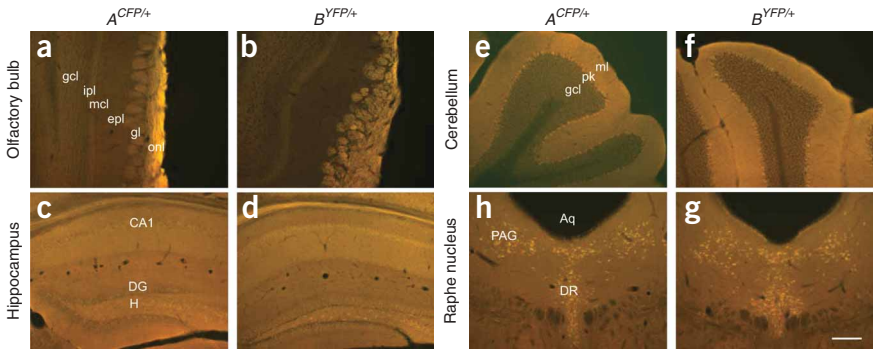

Figure 7 | Pcdha type A and B expression patterns in the adult mouse brain revealed by staining of enhanced cyan fluorescent protein (ECFP) and enhanced yellow fluorescent protein (EYFP) reporters. Expression of Pcdha types A and B in the olfactory bulb (**a**,**b**), hippocampus (**c**,**d**), cerebellum (**e**,**f**) and raphe nucleus (**g**,**h**) are shown. Pcdha type A (**a**,**c**,**e** and **h**) and type B (**b**,**d**,**f** and **g**) expression patterns. Pcdha types A and B have similar expression pattern in the brain. In the olfactory bulb, Pcdha types A and B are strongly expressed in the olfactory nerve layer and glomerular layer but very weak in other layers (**a**,**b**). In the hippocampus, the signal is observed in the pyramidal cell layer and granule cell layer and hilus (**c**,**d**). In the cerebellum, both types A and B are strongly expressed in the Purkinje cell layer and the molecular layer, but weakly in the granular cell layer (**e**,**f**). Both are also strongly expressed in the midbrain raphe nucleus (**g**,**h**). In summary, we wanted to test the possibility of using two different fluorescent reporter pairs ECFP and EYFP for distinguishing proteins *in vivo*. However, we were not able to observe the direct fluorescence due to the low expression levels of ECFP and EYFP. In the future, EGFP and mCherry might be a good combination for tracing two proteins *in vivo*. Aq, aqueduct; CA1, cornu ammonis region 1; DG, dentate gyrus; DR, dorsal raphe; epl, external plexiform layer; gcl, granule cell layer; gl, glomerular layer; H, hilus; ipl, internal plexiform layer; mcl, mitral cell layer; ml, molecular layer; onl, olfactory nerve layer; PAG, periaqueductal gray; pk, Purkinje cell layer. Scale bar, 200 μm.

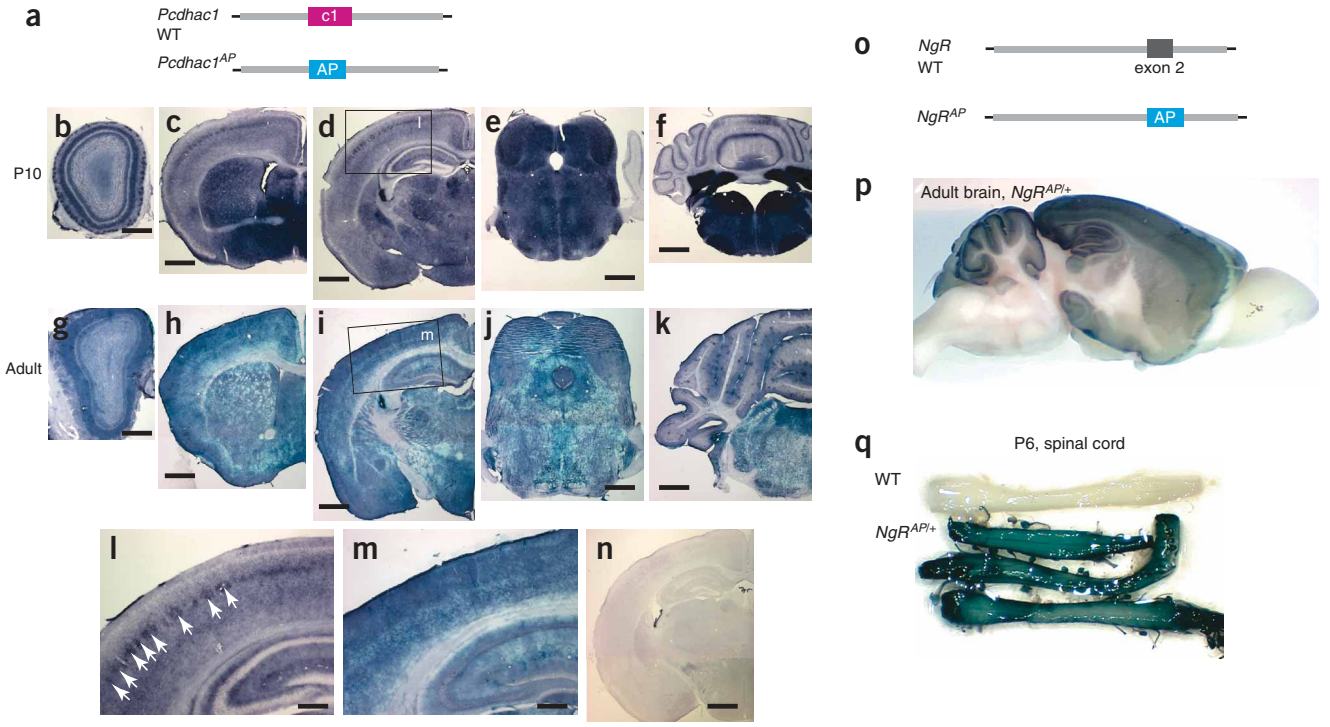

Figure 8 | Alkaline phosphatase (AP) as a reporter to probe gene expression patterns. (**a**–**n**) Pcdhac1 gene expression patterns and (**o**–**q**) NgR gene expression patterns. Schematic design of Pcdhac1-AP (**a**) and its expression patterns in the brain at P10 (**b**–**f**) and adult (**g**–**k**). Pcdhac1 is widely expressed at various brain regions throughout rostra-caudal levels, including olfactory bulb (**b**,**g**), cortex, striatum and septum (**c**,**h**), cortex, hippocampus and thalamus (**d**,**i**), midbrain (**e**,**j**) and cerebellum and pons (**f**,**k**). At P10, the strongest expression is in the olfactory bulb (**b**), basal ganglia (**c**), midbrain (**e**) and cerebellar nucleus and Pons (**f**); while at adult, the expression seems evenly distributed throughout the brain (**g**–**k**). In the P10 cortex, Pcdhac1 shows an interesting patch-like expression pattern (arrows in **l**, high magnification of boxed area in **d**) in the sensory cortical region (**c**, **d** and **l**), possibly in layer 2/3. This pattern is no longer seen in the adult cortex, where Pcdhac1 is localized in all layers ((**h**), (**i**) and (**m**), high magnification of boxed area in **i**). Note that no expression was detected in the hippocampal pyramidal cells, dentate granule cells and cerebellar granule cells in either P10 or adult brain. AP staining using P10 WT brain is negative (n). As a separate example, we examined AP expression in an NgR^{AP} allele[71] (**o**). NgR is expressed in the nervous system (**p**,**q**). Scale bar 500 μm for **b**,**g**,**l** and **m**, 1 mm for **c**–**f**,**h**–**k** and **n**.

fluorescence microscopy without immunostaining, the nlacZ reporter yields very high resolution at the cellular level. For example, in the $Fat2^{EGFP}$ allele (**Fig. 6**), EGFP is very strongly expressed in the cerebellum and pons (**Fig. 6a**–**e**), in a pattern very similar to the previously reported Fat2 antibody staining pattern[114]. In the $Fat3^{nlacZ}$ allele, nlacZ is widely expressed in the nervous system of adults. Owing to the robustness of lacZ (**Fig. 6f**–**k**), we were able to obtain a clearer expression pattern for Fat3 than the previously reported antibody staining patterns[115]. In contrast to EGFP and nlacZ, the AP reporter yielded lower resolution patterns, particularly in the nervous system (**Fig. 8**).

ACKNOWLEDGMENTS We thank Drs. A. Francis Stewart, Neal G. Copeland and Barry L. Wanner for providing their phage recombination systems, Dr. Jan-Fang Cheng (Lawrence Berkeley National Laboratory) for providing BAC clones for the clustered *Pcdh* locus, Drs. Changjiang Zou and Can Li (the Wu laboratory) for help with DNA blot experiments, Haiyan Peng (the Wu laboratory) for help on targeting vector construction and members of the Capecchi laboratory for comments on the article. We are grateful for the excellent technical support from ES cell culture and mouse surgery and husbandry staff in M.R.C. lab, in particular Sheila Barnett, Lois Byers, Carol Lenz, Arunth Lingam, Karl Lustig, Julie Tomlin and Joan Shuhua. G.Y. and Q.W. are supported by an American Cancer Society grant.

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

1. Thomas, K.R. & Capecchi, M.R. Site-directed, mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512 (1987).
2. Mansour, S.L., Thomas, K.R. & Capecchi, M.R. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348–352 (1988).
3. Evans, M.J. & Kaufman, M.H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
4. Martin, G.R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**, 7634–7638 (1981).
5. Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. & Kucherlapati, R.S. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317**, 230–234 (1985).
6. Capecchi, M.R. Targeted gene replacement. *Sci. Am.* **270**, 52–59 (1994).
7. International Mouse Knockout Consortium, Collins, F.S., Rossant, J. & Wurst, W. A mouse for all reasons. *Cell* **128**, 9–13 (2007).
8. Austin, C.P. *et al.* The knockout mouse project. *Nat. Genet.* **36**, 921–924 (2004).
9. Auwerx, J. *et al.* The European dimension for the mouse genome mutagenesis program. *Nat. Genet.* **36**, 925–927 (2004).

10. Akiyama, K., Watanabe, H., Tsukada, S. & Sasai, H. A novel method for constructing gene-targeting vectors. *Nucleic Acids Res.* **28**, E77 (2000).
11. Waterston, R.H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
12. Hamilton, C.M., Aldea, M., Washburn, B.K., Babitzke, P. & Kushner, S.R. New method for generating deletions and gene replacements in *Escherichia coli*. *J. Bacteriol.* **171**, 4617–4622 (1989).
13. Yang, X.W., Model, P. & Heintz, N. Homologous recombination based modification in *Escherichia coli* and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nat. Biotechnol.* **15**, 859–865 (1997).
14. Bradshaw, M.S., Bollekens, J.A. & Ruddle, F.H. A new vector for recombination-based cloning of large DNA fragments from yeast artificial chromosomes. *Nucleic Acids Res.* **23**, 4850–4856 (1995).
15. Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. & Cullin, C. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **21**, 3329–3330 (1993).
16. Oliner, J.D., Kinzler, K.W. & Vogelstein, B. *In vivo* cloning of PCR products in *E. coli*. *Nucleic Acids Res.* **21**, 5192–5197 (1993).
17. Zhang, Y., Buchholz, F., Muyrers, J.P. & Stewart, A.F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.* **20**, 123–128 (1998).
18. Zhang, Y., Muyrers, J.P., Testa, G. & Stewart, A.F. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* **18**, 1314–1317 (2000).
19. Liu, P., Jenkins, N.A. & Copeland, N.G. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* **13**, 476–484 (2003).
20. Wang, J. *et al.* An improved recombineering approach by adding RecA to lambda Red recombination. *Mol. Biotechnol.* **32**, 43–53 (2006).
21. Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97**, 5978–5983 (2000).
22. Lee, E.C. *et al.* A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56–65 (2001).
23. Angrand, P.O., Daigle, N., van der Hoeven, F., Scholer, H.R. & Stewart, A.F. Simplified generation of targeting constructs using ET recombination. *Nucleic Acids Res.* **27**, e16 (1999).
24. Copeland, N.G., Jenkins, N.A. & Court, D.L. Recombineering: a powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* **2**, 769–779 (2001).
25. Zhang, P., Li, M.Z. & Elledge, S.J. Towards genetic genome projects: genomic library screening and gene-targeting vector construction in a single step. *Nat. Genet.* **30**, 31–39 (2002).
26. Valenzuela, D.M. *et al.* High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat. Biotechnol.* **21**, 652–659 (2003).
27. Testa, G. *et al.* Engineering the mouse genome with bacterial artificial chromosomes to create multipurpose alleles. *Nat. Biotechnol.* **21**, 443–447 (2003).
28. Yang, Y. & Seed, B. Site-specific gene targeting in mouse embryonic stem cells with intact bacterial artificial chromosomes. *Nat. Biotechnol.* **21**, 447–451 (2003).
29. Cotta-de-Almeida, V., Schonhoff, S., Shibata, T., Leiter, A. & Snapper, S.B. A new method for rapidly generating gene-targeting vectors by engineering BACs through homologous recombination in bacteria. *Genome Res.* **13**, 2190–2194 (2003).
30. Wallace, H.A. *et al.* Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell* **128**, 197–209 (2007).
31. Chan, W. *et al.* A recombineering based approach for high-throughput conditional knockout targeting vector construction. *Nucleic Acids Res.* **35**, e64 (2007).
32. Accili, D. A note of caution on the Knockout Mouse Project. *Nat. Genet.* **36**, 1132 (2004).
33. Adams, D.J. *et al.* Mutagenic insertion and chromosome engineering resource (MICER). *Nat. Genet.* **36**, 867–871 (2004).
34. Skarnes, W.C. *et al.* A public gene trap resource for mouse functional genomics. *Nat. Genet.* **36**, 543–544 (2004).
35. Hansen, J. *et al.* A large-scale, gene-driven mutagenesis approach for the functional analysis of the mouse genome. *Proc. Natl. Acad. Sci. USA* **100**, 9918–9922 (2003).
36. Hogan, B., Beddington, R., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1994).
37. Joyner, A.L. (ed.) *Gene Targeting: A Practical Approach* (Oxford University Press, New York, 1999).
38. Capecchi, M.R. Altering the genome by homologous recombination. *Science* **244**, 1288–1292 (1989).
39. Koller, B.H. & Smithies, O. Altering genes in animals by gene targeting. *Annu. Rev. Immunol.* **10**, 705–730 (1992).
40. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000).
41. Datta, S., Costantino, N. & Court, D.L. A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**, 109–115 (2006).
42. Court, D.L. *et al.* Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene* **315**, 63–69 (2003).
43. Court, D.L., Sawitzke, J.A. & Thomason, L.C. Genetic engineering using homologous recombination. *Annu. Rev. Genet.* **36**, 361–388 (2002).
44. Hartley, J.L., Temple, G.F. & Brasch, M.A. DNA cloning using *in vitro* site-specific recombination. *Genome Res.* **10**, 1788–1795 (2000).
45. Gaufo, G.O., Wu, S. & Capecchi, M.R. Contribution of Hox genes to the diversity of the hindbrain sensory system. *Development* **131**, 1259–1266 (2004).
46. Wu, S., Ying, G., Wu, Q. & Capecchi, M.R. Toward simpler and faster genome-wide mutagenesis in mice. *Nat. Genet.* **39**, 922–930 (2007).
47. Wu, Y., Wang, G., Scott, S.A. & Capecchi, M.R. Hoxc10 and Hoxd10 regulate mouse columnar, divisional and motor pool identity of lumbar motoneurons. *Development* **135**, 171–182 (2008).
48. Wu, Q. & Maniatis, T. A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* **97**, 779–790 (1999).
49. Wu, Q. & Maniatis, T. Large exons encoding multiple ectodomains are a characteristic feature of protocadherin genes. *Proc. Natl. Acad. Sci. USA* **97**, 3124–3129 (2000).
50. Price, S.R., De Marco Garcia, N.V., Ranscht, B. & Jessell, T.M. Regulation of motor neuron pool sorting by differential expression of type II cadherins. *Cell* **109**, 205–216 (2002).
51. Takeichi, M. The cadherin superfamily in neuronal connections and interactions. *Nat. Rev. Neurosci.* **8**, 11–20 (2007).
52. Capecchi, M.R. The new mouse genetics: altering the genome by gene targeting. *Trends Genet.* **5**, 70–76 (1989).
53. Capecchi, M.R. Generating mice with targeted mutations. *Nat. Med.* **7**, 1086–1090 (2001).
54. Soriano, P. Gene targeting in ES cells. *Annu. Rev. Neurosci.* **18**, 1–18 (1995).
55. Tvrdik, P. & Capecchi, M.R. Reversal of Hox1 gene subfunctionalization in the mouse. *Dev. Cell* **11**, 239–250 (2006).
56. Branda, C.S. & Dymecki, S.M. Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell* **6**, 7–28 (2004).
57. Glaser, S., Anastassiadis, K. & Stewart, A.F. Current issues in mouse genome engineering. *Nat. Genet.* **37**, 1187–1193 (2005).
58. Yu, Y. & Bradley, A. Engineering chromosomal rearrangements in mice. *Nat. Rev. Genet.* **2**, 780–790 (2001).
59. Rajewsky, K. *et al.* Conditional gene targeting. *J. Clin. Invest.* **98**, 600–603 (1996).
60. Gu, H., Marth, J.D., Orban, P.C., Mossmann, H. & Rajewsky, K. Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103–106 (1994).
61. Gu, H., Zou, Y.R. & Rajewsky, K. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell* **73**, 1155–1164 (1993).
62. Dymecki, S.M. Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc. Natl. Acad. Sci. USA* **93**, 6191–6196 (1996).
63. Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc. Natl. Acad. Sci. USA* **85**, 5166–5170 (1988).
64. Hoess, R.H., Ziese, M. & Sternberg, N. P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc. Natl. Acad. Sci. USA* **79**, 3398–3402 (1982).
65. McLeod, M., Craft, S. & Broach, J.R. Identification of the crossover site during FLP-mediated recombination in the *Saccharomyces cerevisiae* plasmid 2 microns circle. *Mol. Cell. Biol.* **6**, 3357–3367 (1986).
66. Sternberg, N. & Hamilton, D. Bacteriophage P1 site-specific recombination. I. Recombination between *lox*P sites. *J. Mol. Biol.* **150**, 467–486 (1981).
67. Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M. & Hoess, R. Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harb. Symp. Quant. Biol.* **45**, 297–309 (1981).
68. Nagy, A. Cre recombinase: the universal reagent for genome tailoring. *Genesis* **26**, 99–109 (2000).
69. Moon, A.M. & Capecchi, M.R. Fgf8 is required for outgrowth and patterning of the limbs. *Nat. Genet.* **26**, 455–459 (2000).
70. Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* **445**, 661–665 (2007).

71. Wu, S., Wu, Y. & Capecchi, M.R. Motoneurons and oligodendrocytes are sequentially generated from neural stem cells but do not appear to share common lineage-restricted progenitors *in vivo*. *Development* **133**, 581–590 (2006).
72. Harfe, B.D. *et al.* Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* **118**, 517–528 (2004).
73. Hayashi, S. & McMahon, A.P. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev. Biol.* **244**, 305–318 (2002).
74. Indra, A.K. *et al.* Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases. *Nucleic Acids Res.* **27**, 4324–4327 (1999).
75. Greer, J.M., Puetz, J., Thomas, K.R. & Capecchi, M.R. Maintenance of functional equivalence during paralogous Hox gene evolution. *Nature* **403**, 661–665 (2000).
76. Gossen, M. *et al.* Transcriptional activation by tetracyclines in mammalian cells. *Science* **268**, 1766–1769 (1995).
77. Lamartina, S. *et al.* Stringent control of gene expression *in vivo* by using novel doxycycline-dependent trans-activators. *Hum. Gene Ther.* **13**, 199–210 (2002).
78. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70–71 (1999).
79. Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4 (2001).
80. Haldar, M., Hancock, J.D., Coffin, C.M., Lessnick, S.L. & Capecchi, M.R. A conditional mouse model of synovial sarcoma: insights into a myogenic origin. *Cancer Cell* **11**, 375–388 (2007).
81. Adams, D.J. *et al.* A genome-wide, end-sequenced 129Sv BAC library resource for targeting vector construction. *Genomics* **86**, 753–758 (2005).
82. Deng, C. & Capecchi, M.R. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol. Cell. Biol.* **12**, 3365–3371 (1992).
83. Hasty, P., Rivera-Perez, J., Chang, C. & Bradley, A. Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. *Mol. Cell. Biol.* **11**, 4509–4517 (1991).
84. Tang, S.H., Silva, F.J., Tsark, W.M. & Mann, J.R. A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvImJ. *Genesis* **32**, 199–202 (2002).
85. 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).
86. Thomas, K.R., Folger, K.R. & Capecchi, M.R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419–428 (1986).
87. Zimmer, A. & Gruss, P. Production of chimaeric mice containing embryonic stem (ES) cells carrying a homoeobox Hox 1.1 allele mutated by homologous recombination. *Nature* **338**, 150–153 (1989).
88. Zwaka, T.P. & Thomson, J.A. Homologous recombination in human embryonic stem cells. *Nat. Biotechnol.* **21**, 319–321 (2003).
89. Vasquez, K.M., Marburger, K., Intody, Z. & Wilson, J.H. Manipulating the mammalian genome by homologous recombination. *Proc. Natl. Acad. Sci. USA* **98**, 8403–8410 (2001).
90. Folger, K.R., Wong, E.A., Wahl, G. & Capecchi, M.R. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Mol. Cell. Biol.* **2**, 1372–1387 (1982).
91. Doetschman, T. *et al.* Targeted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* **330**, 576–578 (1987).
92. te Riele, H., Maandag, E.R. & Berns, A. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Natl. Acad. Sci. USA* **89**, 5128–5132 (1992).
93. George, S.H. *et al.* Developmental and adult phenotyping directly from mutant embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **104**, 4455–4460 (2007).
94. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J.C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **90**, 8424–8428 (1993).
95. Gerlai, R. Gene-targeting studies of mammalian behavior: is it the mutation or the background genotype? *Trends Neurosci* **19**, 177–181 (1996).
96. Wolfer, D.P., Crusio, W.E. & Lipp, H.P. Knockout mice: simple solutions to the problems of genetic background and flanking genes. *Trends Neurosci.* **25**, 336–340 (2002).
97. Seong, E., Saunders, T.L., Stewart, C.L. & Burmeister, M. To knockout in 129 or in C57BL/6: that is the question. *Trends Genet.* **20**, 59–62 (2004).
98. Schoonjans, L. *et al.* Improved generation of germline-competent embryonic stem cell lines from inbred mouse strains. *Stem Cells* **21**, 90–97 (2003).
99. Auerbach, W. *et al.* Establishment and chimera analysis of 129/SvEv- and C57BL/6-derived mouse embryonic stem cell lines. *Biotechniques* **29**, 1024–1028, 1030, 1032 (2000).
100. Schuster-Gossler, K. *et al.* Use of coisogenic host blastocysts for efficient establishment of germline chimeras with C57BL/6J ES cell lines. *Biotechniques* **31**, 1022–1024 1026 (2001).
101. 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).
102. van Deursen, J., Lovell-Badge, R., Oerlemans, F., Schepens, J. & Wieringa, B. Modulation of gene activity by consecutive gene targeting of one creatine kinase M allele in mouse embryonic stem cells. *Nucleic Acids Res.* **19**, 2637–2643 (1991).
103. Yanagawa, Y. *et al.* Enrichment and efficient screening of ES cells containing a targeted mutation: the use of DT-A gene with the polyadenylation signal as a negative selection maker. *Transgenic Res.* **8**, 215–221 (1999).
104. Yagi, T. *et al.* Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diphtheria toxin A-fragment gene in negative selection. *Proc. Natl. Acad. Sci. USA* **87**, 9918–9922 (1990).
105. Nagy, A. *et al.* Embryonic stem cells alone are able to support fetal development in the mouse. *Development* **110**, 815–821 (1990).
106. Wood, S.A., Allen, N.D., Rossant, J., Auerbach, A. & Nagy, A. Non-injection methods for the production of embryonic stem cell-embryo chimaeras. *Nature* **365**, 87–89 (1993).
107. Inoue, H., Nojima, H. & Okayama, H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**, 23–28 (1990).
108. Nichols, J., Evans, E.P. & Smith, A.G. Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development* **110**, 1341–1348 (1990).
109. Chomczynski, P. One-hour downward alkaline capillary transfer for blotting of DNA and RNA. *Anal. Biochem.* **201**, 134–139 (1992).
110. Truett, G.E. *et al.* Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* **29**, 52 54 (2000).
111. Wu, Q. *et al.* Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. *Genome Res.* **11**, 389–404 (2001).
112. Wu, Q. Comparative genomics and diversifying selection of the clustered vertebrate protocadherin genes. *Genetics* **169**, 2179–2188 (2005).
113. Sugino, H. *et al.* Genomic organization of the family of CNR cadherin genes in mice and humans. *Genomics* **63**, 75–87 (2000).
114. Nakayama, M., Nakajima, D., Yoshimura, R., Endo, Y. & Ohara, O. MEGF1/fat2 proteins containing extraordinarily large extracellular domains are localized to thin parallel fibers of cerebellar granule cells. *Mol. Cell. Neurosci.* **20**, 563–578 (2002).
115. Nagae, S., Tanoue, T. & Takeichi, M. Temporal and spatial expression profiles of the Fat3 protein, a giant cadherin molecule, during mouse development. *Dev. Dyn.* **236**, 534–543 (2007).
116. Rock, R., Schrauth, S. & Gessler, M. Expression of mouse dchs1, fjx1, and fat-j suggests conservation of the planar cell polarity pathway identified in Drosophila. *Dev. Dyn.* **234**, 747–755 (2005).