nature genetics

Current issues in mouse genome engineering

Stefan Glaser, Konstantinos Anastassiadis & A Francis Stewart

The mouse is the foremost vertebrate experimental model because its genome can be precisely and variously engineered. Now that the mouse genome has been sequenced and annotated, the task of mutating each gene is feasible, and an international cooperation is providing mutated embryonic stem cells and mice as readily available resources. Because these resources will change biomedical research, decisions about their nature will have far-reaching effects. It is therefore timely to consider topical issues for mouse genome engineering, such as the background genotype; homologous, site-specific and transpositional recombination; conditional mutagenesis; RNA-mediated interference; and functional genomics with embryonic stem cells.

Methods for engineering the mouse genome have been developed for all fundamental mechanisms of mutagenesis, accessing the germline genome through gametes, zygotes and embryonic stem (ES) cells (**Table 1**). Some of these genome engineering methods are also applicable to other model systems such as rat or zebrafish, but the mouse is unique because of ES cells. The ability to reconstitute fertile mice from cells in culture has led to the development of a range of sophisticated genome engineering technologies, most notably gene targeting, gene trapping and conditional mutagenesis. The unique properties of ES cells also underpin two international efforts, the Knock-Out Mouse Project (KOMP) and European Conditional Mouse Mutagenesis (EUCOMM), which aim to mutate every gene in the mouse[1,2].

Here we discuss issues relevant to the decisions behind KOMP and EUCOMM, among other topical matters. Several reviews on mouse genome engineering[3–6] provide greater detail than is possible here.

The configuration of an ideal ES cell line

For reasons that may be related to the ability of mice to undergo diapause, so far the germline properties of ES cells have been found solely in mouse and in only a few mouse strains[7,8]. ES cells from the 129 strain were the first to be derived[9,10] and have become the most widely used; however, these strains have some genetic weaknesses and are generally disfavored by immunologists and neurologists, among other scientists[11,12]. Consequently, rather than remain within a congenic background for genetic clarity, a suboptimal practice has arisen. Mutations are made in 129 cells and then repeatedly crossed to the desired background (usually C57BL/6), even though the congenic purity of C57BL/6 cannot be restored. In addition, the choice of background for transgenesis by pronuclear injection has been determined mainly by superovulation yield. Subsequent crossing of a transgenic line (*e.g.*, a Cre transgenic line built on one particular background) with a targeted line originally built on the 129 background but often crossed to another is common practice. This messy state of genetic affairs is an outcome of the ways in which the technologies were developed. Now that there are choices, it is worth considering whether the merits of congenic genetic practice can be resumed. The choice of the genetic background for the resources generated by KOMP and EUCOMM will have a lasting impact on practice in mouse experimental biology.

At present, the only alternative to the 129 strain is C57BL/6 (ref. 13). Ideally, a few C57BL/6 ES cell lines with sufficient reliability (in terms of chimera and germline transmission efficiencies) will emerge to replace 129. If so, then C57BL/6 will become the main genetic background for ES cell work and pronuclear injections. Until then, mutagenesis in the 129 ES cell background is sure to continue because these cells are robust.

F_1 ES cell lines, usually from a 129 × C57BL/6 cross, offer an alternative to 129 cell lines. Owing to hybrid vigor, some F_1 cell lines have good germline properties and can generate mice derived completely from ES cells by tetraploid complementation[14,15]. Thus, functional analyses (*e.g.*, using RNA-mediated interference (RNAi) knockdown), can be rapidly done with 'transient' mice without the need to establish mouse lines and breed[16]. Because this approach is speedy, it is certain to find many applications. The genetic differences between 129 and C57BL/6 mean, however, that the merits of maintaining a congenic background are discarded.

Pre-engineered ES cells present some possibilities. For example, ES cells carrying integrated transgenes encoding site-specific recombinases offer advantages, such as simplified removal of selectable markers or built-in switches for inducible RNAi or conditional mutagenesis[17,18]. Alternatively, ES cells engineered to carry strategically placed, site-specific recombinase target sites can facilitate functional studies[19,20]. Regardless of whether pre-engineering can arrive at a single ES cell format, or whether the genotype issue will be resolved soon, the idea that one, optimally configured ES cell line will emerge to anchor mouse research in the near future is probably just wishful thinking, although it is not without appeal.

Homologous recombination

Mouse ES cells seem to have higher frequencies of homologous recombination than do other cultured cells. This higher frequency is probably due, in part, to the very rapid growth and accompanying high activity of DNA replication of these cells, in which S phase accounts for a relatively greater proportion of the cell cycle. Whatever the exact explanation, translating the potential for homologous recombination into routine

Genomics, BioInnovationsZentrum, Dresden University of Technology, Am Tatzberg 47, 01307, Dresden, Germany. Correspondence should be addressed to A.F.S. (stewart@biotec.tu-dresden.de).

Published online 27 October 2005; doi:10.1038/ng1668

success with gene targeting in ES cells has not been without complications, mainly related to the ways in which the targeting constructs have been built. It is important to use both isogenic DNA (*i.e.*, homologous DNA for homologous recombination[21]) and homology arms longer than those that can be easily made by PCR. Both of these aspects have become simple and convenient in recent years. It is now very easy to obtain (nearly) isogenic DNA from genome repositories because BACs containing either C57BL/6 or 129 DNA are annotated onto the silicon genome (Ensembl Mouse, http://www.ensembl.org/Mus_musculus/). Larger targeting constructs are also easy to make by using Red/ET recombineering methodology[22,23] and subcloning from BACs[24] (**Fig. 1**).

Table 1 Options for manipulating the mouse genome

Mutagenesis	Path 1 (using zygotes)	Path 2 (using ES cells)
Point	ENU (in testis)	ENU
Illegitimate recombination	Transgenesis by pronuclear injection	Transgenesis by transfection; gene traps
Homologous recombination		Targeting; targeting to introduce *lox*P sites (conditional mutagenesis)
Site-specific recombination	Through mating: parent 1, recombinase; parent 2, target sites	Recombination achieved by transfection or viral delivery, Cre protein translocation, or ligand activation of endogenous Cre
Transpositional	Through mating: parent 1, transposase; parent 2, inverted repeats	

Although longer targeting constructs increase targeting frequency[25–28], we still do not know what lengths are optimal. Such optima probably vary from locus to locus, and some loci might be recalcitrant to targeting regardless of homology length. The reasons behind locus-specific variations in targeting frequencies remain unknown but could involve putative negative or positive elements, chromatin status, or spatial aspects of the nucleus. A better understanding of these issues would be not only fundamentally interesting but also relevant to the practicalities of genome engineering. As it stands, the design of efficient targeting constructs is guided by three parameters: isogenic is better; bigger is better; and promoterless selection, if possible, is better. The more that these parameters can be incorporated in a targeting construct, the greater the chance of success. Targeting remains unpredictably variable, however, for reasons that are not understood.

How reliant is targeting on isogenicity? A recent study using large BAC-derived, nonisogenic targeting constructs suggests that dependence on isogenicity diminishes with increasing size[27]. But a second BAC-based targeting study using isogenic DNA has reported a much higher average frequency[28] (albeit from only five examples). These data indicate that both size and isogenicity contribute to targeting efficiency and that optimal results can be achieved by their combination.

Because large BAC-based targeting constructs are easy to make and provide both higher targeting frequencies and a partial solution to the isogenicity issue, should BAC-based targeting become the primary method? There are reasons to think twice. BAC-based targeting constructs are too large to permit a thorough characterization of the targeting event. This problem of size has required the development of new methods of screening for correct recombinants[27,28]. Although both methods are creative and excellent, the main problem remains; namely, the molecular detail of the recombination event is not verifiable. For example, it is likely that long homology arms will contain repeats; at what frequency will these repeats promote deletion in the genome? In addition, imperfect recombination events can occur, such as duplications at the ends of the targeting construct[29]. The use of very large targeting constructs makes routine inspection for these events difficult. If the BAC is not isogenic, it is difficult to determine where along the homology arms the crossover points lie.

BAC-based targeting constructs also preclude the use of promoterless targeting constructs to enhance targeting frequencies. For expression of a selectable marker, promoterless targeting constructs must integrate into an active gene in the correct orientation. Therefore, most random integrations will be unsuccessful and targeting frequencies will be improved, sometimes to 100%. This strategy also enables targeting to be screened by 5′ rapid amplification of cDNA ends, which is fast and easy. Of course, the targeted gene must be expressed sufficiently for these benefits; however, about two-thirds of all genes seem to be expressed in ES cells[30]. Therefore, promoterless targeting can substantially reduce the effort spent in screening for homologous recombinants, which is the most laborious step.

For these reasons, we think that the first step in targeting a new gene should be to determine whether it is expressed sufficiently to permit promoterless targeting and, if so, to subclone by 'recombineering' a backbone that is optimized for both promoter trapping and Southern-blot

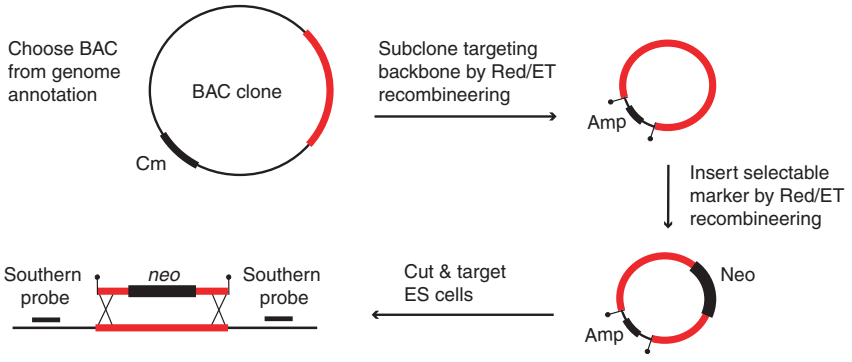

Figure 1 Scheme for rapidly constructing optimized targeting vectors. After planning the targeting strategy according to whether a promoterless strategy can be used and what Southern-blotting strategy will be used, an isogenic BAC encompassing the region of interest is obtained from the genome resource. The region is subcloned by, for example, Red/ET 'recombineering' into a moderate copy plasmid, which includes restriction sites to cutoff the vector. A selectable cassette, such as the neomycin resistance (*neo*) cassette, is inserted into the targeting construct, which is then subjected to restriction digestion and targeting in ES cells. Amp, ampicillin; Cm, chloramphenicol; Neo, neomycin.

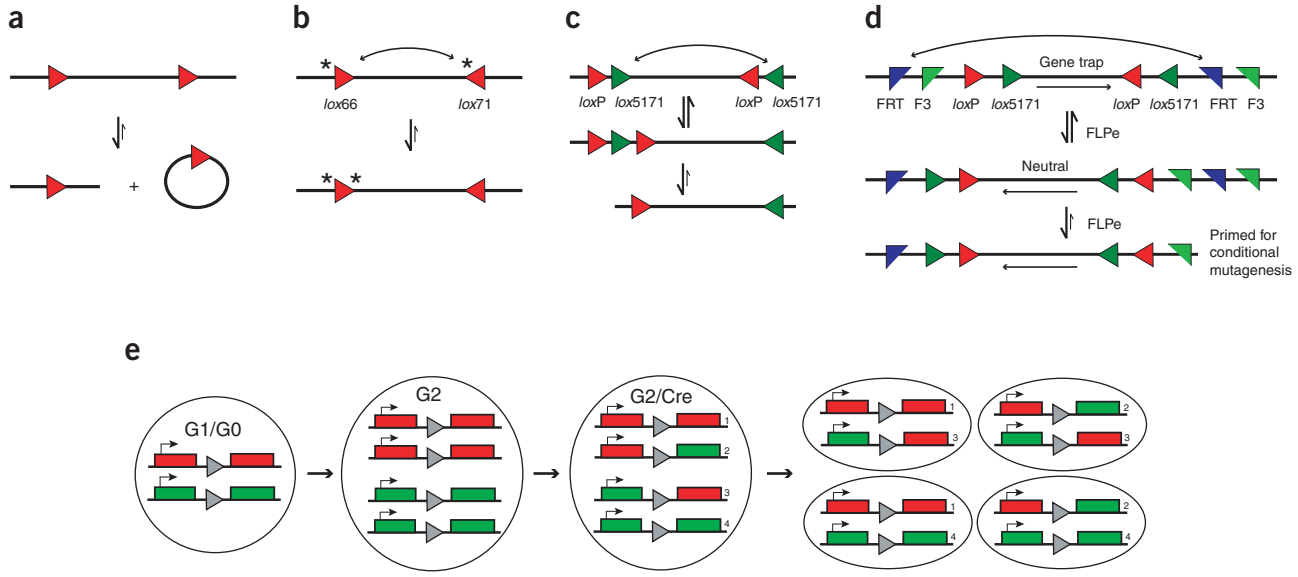

Figure 2 Strategies to impose asymmetry onto Cre recombination. (**a**) Deletion of a DNA interval between two directly repeated *lox*P sites is asymmetrical because of the difference between intra- and intermolecular recombination. (**b**) Inversion between two inverted *lox*P sites can be made directional by using the mutated *lox*66 and *lox*71sites. Both can sustain recombination, but the product includes a doubly mutated *lox* site that is not well recognized by Cre such that the back reaction is disfavored. (**c**) In FIEx, two pairs of heterotypic *lox* sites are placed in an interwoven, inverted configuration. Recombination is possible only between the homotypic sites, resulting in an inversion. This places the other pair of *lox* sites in a direct orientation, facilitating deletion between them and causing asymmetry, as in **a**. (**d**) In gene trapping, the cassette must integrate in an orientation to capture the transcript. After characterization, the cassette is then inverted to the neutral orientation by FLP action on sites placed in a FIEx configuration. The cassette is then ready for Cre-induced conditional mutagenesis. (**e**) Cre-mediated chromosomal translocations produce asymmetries after sister chromatid exchange. One of the four possible product cells will contain a chromosomal arm that is homozygous with respect to a mutation (red boxes).

analysis of the recombination event. For genes that cannot be targeted with a promoterless construct, we recommend using a subclone that is designed to maximize the length of homology, while retaining a convenient way to verify correct recombination by Southern blotting. These gene-specific targeting constructs, Southern-blot probes and results can facilitate future targeting to create different alleles. Ideally, the next few years will bring a steady accumulation of these reagents toward a readily available, nearly complete set for genome targeting.

Site-specific recombination and conditional mutagenesis

The site-specific recombinases Cre and FLP have been applied to various strategies of genome manipulation in many living systems. The most precise is conditional mutagenesis in the mouse, which arises from a combination of homologous recombination and the properties of Cre recombinase[3–6]. To reserve Cre for this task, the thermostable FLP recombinase FLPe is used for working tasks, such as the removal of selectable genes[31,32]. Conditional mutagenesis relies on placing two *lox*P sites in the otherwise unaltered gene of interest[33]. On exposure to Cre, recombination between the *lox*P sites deletes the sequence between the two sites. Therefore, the *lox*P sites must be placed either side of an essential section of the gene. Preferably, all of the gene should be flanked by *lox*P sites, but this is impractical unless the gene is very small.

Site-specific recombination depends on random collision between two recombination target sites[34]. For genes embedded in chromatin in the nucleus or during M phase events, the rate of random collision will be affected by at least two parameters. The first, which is measurable, is the distance between the two sites along a chromosome: the further the two sites are apart, the less often they are likely to collide, leading to lower rates of recombination. The second is the 'position effect', a source of variability that is not understood. Consequently, designing conditional alleles involves an unpredictable factor. Pragmatically, it is a balance between minimizing the distance between *lox*P sites and excising exons that either encode a vital segment of the protein or result in a frame-shifted open reading frame when omitted from the mRNA. This balance means that the design of each conditional allele requires a careful consideration of the gene structure. Work is underway to develop a generic strategy for conditional mutagenesis that will be applicable to most genes without the need to consider the individual gene architecture. Similar to earlier progress in allele design, gene trapping is leading the way[35].

The basic approach in generic conditional mutagenesis is to place a cassette flanked by inverted *lox*P sites into an early intron. In one orientation, the cassette is neutral because it is unrecognized by the splicing machinery. On exposure to Cre recombinase, the cassette is inverted to a dominant orientation, which captures splicing of the transcript and then terminates it, thereby preventing the production of downstream mRNA. Although this approach is simple, its implementation is not straightforward because Cre-mediated recombination is inherently reversible. Without a source of asymmetry in the DNA, Cre-*lox*P recombination equilibrates at a substrate to product ratio of 50:50. For deletion strategies, the asymmetry comes from the difference between forward intramolecular recombination and reverse intermolecular recombination (**Fig. 2a**).

For inversion strategies, asymmetries can be imposed in two ways: either by using mutated *lox* sites[36], called *lox*66 and *lox*71 (**Fig. 2b**), or by Flip excision[37] (FlEx; **Fig. 2c**). Inversion strategies in generic conditional mutagenesis depend on an intronic cassette that is neutral in one orientation but dominant in the other with respect to the splicing

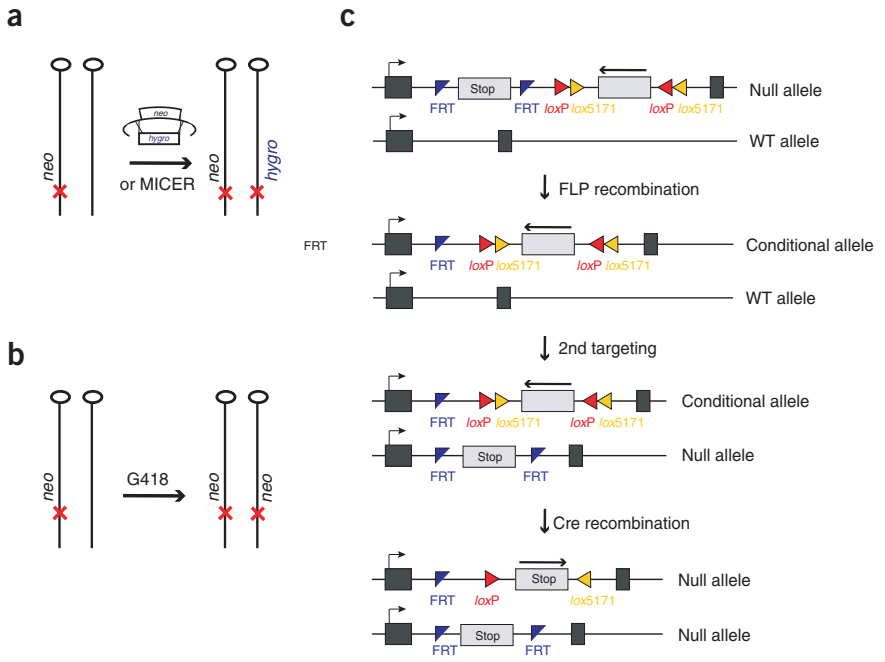

Figure 3 Strategies to create homozygosity in ES cells. (**a**) Double targeting of the second allele is possible by using a different selectable marker either by exchanging the selection marker from *neo*, for example, to the hygromycin resistance (*hygro*) cassette in the first targeting construct or by obtaining a clone from the MICER resource. (**b**) Selection for rare chromosomal dysjunction can be achieved by increasing the selection pressure for the first targeted allele (*e.g.*, by selecting with G418). (**c**) When the gene function cannot be eliminated from ES cells without adverse effects, a conditional strategy is required. Here, for example, the starting allele is converted to a neutral FlEx conditional allele by FLP recombination to remove the FRT-flanked stop cassette. The second allele is targeted with the original targeting construct without the FlEx cassette to permit conditional mutagenesis by Cre. This starting allele is also an alternative to the double FlEx conditional gene trapping strategy (**Fig. 2d**). The strategy shown in **Figure 2e** will also create homozygosity. WT, wild-type.

machinery. Work with gene trapping has shown that splice acceptor and polyadenylation stop cassettes, when integrated into early exons, are dominant in most cases[38,39]. The challenge now lies in developing cassettes that are dominant in the sense orientation but completely neutral in the opposite orientation.

Gene trapping

Gene trapping is the most cost-effective way to create gene-specific mutations. Various gene trapping laboratories have coordinated their activities in the International Gene Trap Consortium[30] and mutations in more than one-third of genes are now available (http://www.igtc.org.uk/). Through their efforts and those of a commercial gene trapping activity[39], some inherent characteristics of gene trapping have been identified. Not surprisingly, genes that are strongly expressed in ES cells are the most easily trapped. Large genes are also favored, as is integration into introns at the 5′ end. Although characteristics differ according to the gene trapping cassette and the gene, the most useful cassettes produce a null allele when they are integrated near the 5′ end of a gene. It is possible that gene traps that are integrated into the first or early exons capture splicing and terminate the transcript more reliably than those that are integrated at more 3′ regions. Verification of this observation requires more data.

Through the development of a double FlEx strategy (**Fig. 2d**), conditional gene traps are now being generated and more than 1,000 alleles are available from the International Gene Trap Consortium. Although preliminary results are encouraging, the utility of these alleles will depend on whether the gene trapping cassettes are indeed dominant in one orientation and neutral in the other, and whether they function as well *in vivo* as they do in ES cells. Further cassette development will probably be an important aspect of conditional mutagenesis in the future.

Functional genomics with ES cells

In addition to producing mice, ES cells form the basis of various functional models. Undifferentiated, they are a model for pluripotency and stem cell issues[40]. They can be differentiated into numerous cell lineages, thereby facilitating mechanistic analyses of lineage commitment and differentiation[41]. Thus, it is logical to consider how the libraries of mutated ES cells can be applied to functional genomics. These libraries provide cells with a mutation in one allele; autosomal genes usually need to be mutated on the other allele for functional analysis. If the first allele has been mutated by targeting, then mutagenesis of the second allele can be greatly facilitated by reusing the targeting construct after the selectable marker is exchanged[42] (**Fig. 3a**). Thus, targeted ES cell lines available in the public resource should be accompanied with the targeting reagents (targeting construct, generic cassette to exchange the selectable marker, subclones for Southern-blot analysis) to facilitate targeting of the second allele or the generation of further alleles.

If the first allele has been mutated by trapping, it may be possible to use a clone from the Mutagenic Insertion and Chromosome Engineering Resource (MICER)[43] to mutate the second allele. MICER clones are annotated, random subclones of strain 129 that are readily obtainable (http://www.ensembl.org/Mus_musculus). Because these clones are insertion vectors (rather than the conventional replacement targeting vectors), they result in high rates of homologous recombination, albeit with the complication that the targeting site becomes duplicated so that the allele generated may not be null. Nevertheless, MICER can provide a convenient option.

A third general strategy to mutate both alleles involves creating homozygosity by selecting for the duplication of whole, or nearly whole, chromosomes. One mechanism involves selecting for rare defects in chromosomal segregation[44] (**Fig. 3b**). Another involves the use of *lox*P sites placed near centromeres, coupled with Cre recombinase and a suitable selection strategy[18] (**Fig. 2e**). These chromosomal conversion strategies for establishing homozygosity, however, cannot be easily applied to the seven chromosomes that carry imprinted regions. In addition, the Cre-based strategy requires the prior creation of a specific ES cell line with integrated *lox*P sites for each chromosome. Thus, the use of chromosomal conversion to achieve homozygosity will probably be confined to exercises that concentrate on a few selected chromosomal arms.

Whether starting from a gene trap, targeted allele or *lox*P chromosome, strategies to mutate both alleles benefit from a conditional mutagenesis approach. If a gene is essential to ES cells, both of its alleles cannot be

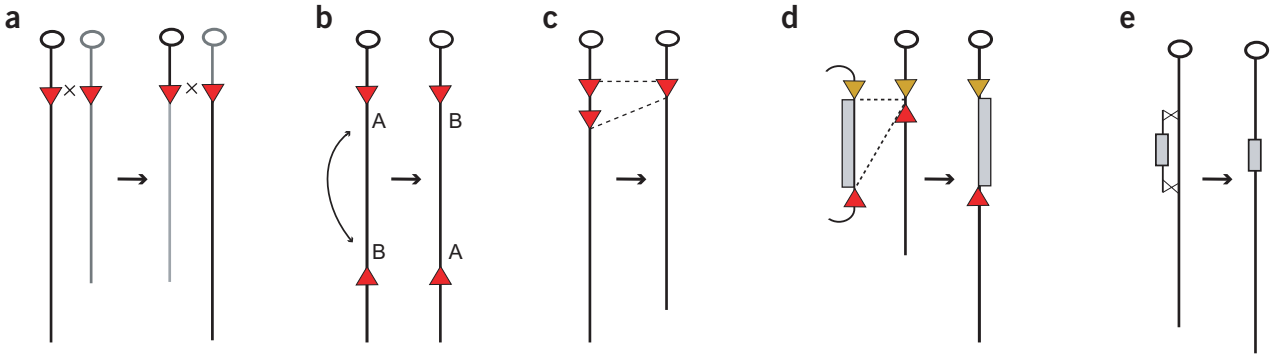

Figure 4 Engineering above 10 kb. (**a**) Chromosomal translocations can be mediated by Cre through *lox*P sites (triangles) on heterologous chromosomes with selection, or on homologous chromosomes with or without selection. (**b**) Inversions between inverted *lox*P sites (designated A and B) mediated by Cre create balancer chromosomes. (**c**) Large deletions mediated by Cre are possible unless the deleted interval contains a haploinsufficiency gene. (**d**) Regional exchanges of up to 200 kb can be achieved by RMCE, which requires a heterospecific pair of recombination target sites. (**e**) Smaller regional exchanges may be achievable by using a large targeting construct containing the exchange region flanked by homology arms.

mutated to null alleles. In this situation, conversion of one allele to a conditional one permits mutagenesis of the other (**Fig. 3c**). Even when the gene is not essential, a conditional allele is preferable because the other allele can be mutated without provoking compensatory adaptations.

RNAi

For loss-of-function engineering, RNAi provides an alternative to DNA mutagenesis[45]. RNAi can be used for rapid functional genomics with ES cells without the complications of the second allele discussed above. This flexibility has a lot of appeal and RNAi will undoubtedly have a key role in functional genomics with ES cells. Used in combination with F_1 cells, it has already proved very effective at generating embryos derived completely from ES cells for functional tests during development[15]. Unlike DNA mutagenesis, RNAi also facilitates large-scale screens, in particular when short interfering RNAs are used for transient knock-down of expression[46] or as pools of expression constructs for identifying a selectable outcome[47].

Given these advantages, will RNAi replace DNA mutagenesis? As currently applied, RNAi provokes degradation of specific mRNA and is therefore a balance between ongoing production and degradation. Although it is likely that this balance can be maintained stably in cultured cells, little is known about what happens when RNA levels fluctuate during development or owing to changes in environmental conditions such as stress. For these and other reasons, much more experience is needed before the strengths and limitations of RNAi applications in the mouse or in differentiating ES cells are appreciated. It is too early to decide. For example, recent data from adult mice show a currently inexplicable variation in the degree of expression knock-down among cell types[48]. These preliminary data suggest that cell type variability could make the interpretation of RNAi phenotypes exceedingly complicated.

Superficially, DNA mutagenesis seems to be at a disadvantage to RNAi because both alleles need to be mutated at the DNA level, whereas a single RNAi transgene knocks-down both alleles. For functional work with ES cells, the availability of gene trap or targeted ES cells diminishes this disadvantage because one allele has been already mutated. For establishing mouse lines, retaining the function of one allele is very useful. To establish mouse lines with RNAi knock-down of essential genes, a strategy for conditional expression of the RNAi is required. Such strategies are being developed[49,50], but it is not known whether several conditional RNAi lines will need to be established to obtain a useful line, or how this may be achieved reproducibly. For example, Seibler *et al.*[48] knocked a short hairpin RNAi transgene into the *ROSA26* locus to achieve reproducible, ubiquitous expression, but they did not achieve uniform knock-down.

A further point to consider is that RNAi can produce only a knock-down in expression. Mutagenesis at the DNA level is required for point mutagenesis, for creating humanized alleles, for generating fusion proteins and for various other tasks, all of which will benefit from expanding the reagents for DNA mutagenesis.

Cre is the sharpest tool in the box

Though a useful tool, Cre recombinase is dangerous. Its expression can provoke mutagenesis through strand breakage or recombination at cryptic *lox* sites in the genome[51,52]. It is essential to be aware of the potential for undesired Cre-mediated mutagenesis to influence the experimental outcome. This is particularly relevant in experiments involving Cre-provoked tumorigenesis and other situations where the product cell is clonally amplified after Cre has acted. For experiments using conditional mutagenesis, a comparison of the Cre transgene in the heterozygous *lox* allele background with the Cre transgene in the homozygous *lox* (or heterozygous *lox* and null) allele background is an essential control.

Though not wishing to diminish the importance of this warning, we would suggest that so far the collective experience with Cre gives grounds for reasonable confidence. There is a workable window between intended *lox*P recombination and cryptic *lox* activity. In addition, the increasingly popular strategy of fusing a site-specific recombinase with a steroid receptor[53] presents a way to limit Cre activity and thereby to lessen the chances for cryptic damage in the genome.

Genome engineering above 10 kb

Cre-*lox*P recombination has been applied to the full size range of engineering tasks, from the short-range, highly efficient events that underline conditional mutagenesis to the most difficult that involve translocations between nonhomologous chromosomes. For example, Cre-mediated chromosomal translocations have been used to recapitulate leukemiogenic translocations[54] (**Fig. 4a**). Because this rare event is amplified by

the leukemia, it need not be particularly efficient. Cre-mediated translocations between homologous chromosomes or chromatids, however, occur at workable efficiencies without amplification. Notable applications include chromosomal engineering through the germ line[55,56] and somatic cell marking[57]. Cre-mediated large inversions also have been used to develop balancer mouse lines, which have advantages in genotyping and stock maintenance[58] (**Fig. 4b**). Similarly, large deletions mediated by Cre are also feasible without selection[56] (**Fig. 4c**).

For engineering on a more moderate scale, both Cre and FLP have been used for recombination-mediated cassette exchange (RMCE)[59]. This genome engineering strategy requires antibiotic selection in cultured cells and is therefore unique to the mouse because of ES cells. It involves the exchange of a genomic region flanked by heterotypic recombinase target sites (*e.g.*, *lox*P and *lox*5171 or FRT and F3) with an incoming region that is also flanked by these sites (**Fig. 4d**). RMCE seems to be practical up to, and perhaps beyond, BAC-sized DNA exchanges (up to 200 kb) and is the method of choice for replacing a mouse gene with its human counterpart (called 'humanizing') and other permutations of regional replacements. During RMCE, ES cells often need to be transfected and cloned under selection several times, which can result in impaired germline transmission. Thus, RMCE applications frequently include the generation of mice and derivation of new ES cell lines that carry intermediate steps of the engineering plan. A less laborious approach for regional replacements based on only one round of targeting with a complex targeting construct may be a preferable method[60,61] (**Fig. 4e**).

Space for new tools

As evident from the discussion so far, there are more applications for site-specific recombination than there are site-specific recombinases. Ideally, each application would have its own recombinase (*e.g.*, Cre for conditional mutagenesis, FLPe for removing selection cassettes, a third for RMCE, a fourth for chromosomal engineering). Consequently, there is still a need for new useful recombinases.

Recently, it seemed that ϕC31, a recombinase from *Streptomyces*, could be added to the toolkit; however, its performance has not matched its promise, possibly owing to a fundamental problem. ϕC31 is a member of the serine site-specific recombinases. Unlike the tyrosine site-specific recombinases, such as Cre and FLP, the serine recombinases require a topologically wrapped substrate, which is usually established by the supercoiling energy available in prokaryotic genomes[62]. The observation that ϕC31 could recombine linear DNA *in vitro* was therefore unexpected and led to the definition of a subclass called the 'large serine' recombinases and prompted tests in eukaryotes[63,64]. Although ϕC31 has been applied usefully to RMCE in ES cells, an unusual asymmetry between intermolecular and intramolecular recombination was noted[20]. This asymmetry suggests that the large serine recombinases also need to wrap the substrate. If so, then ϕC31 will be useful only for intermolecular exercises such as RMCE.

Recently, an explicit search for a Cre-like recombinase has uncovered a very similar enzyme, termed Dre, which recombines *rox* sites but not *lox*P sites[65]. It is to be hoped that Dre will be as usefulas Cre and become the long-sought new member of the toolkit.

The most complete set of mutagenic options

N-Ethyl-*N*-nitrosourea (ENU)-based mutagenesis of ES cells adds a new dimension to point mutagenesis in the mouse genome, which was previously accessible only by ENU treatment of testis[66]. Because ES cells can tolerate a high mutagenic burden and still retain germline transmission, it is possible to screen a library of ENU-mutated ES cells by tilling to find mutations in a chosen gene. Because a large proportion of point mutations affect splicing, gene-specific splicing mutations also can be efficiently sought by screening cDNA pools from such libraries[67]. ENU mutagenesis may not be particularly suited to functional genomics with ES cells because of the multiplicity of mutations that it generates. In mice, these mutations are bred away from the object of study; however, they will complicate functional analyses based on ES cells, particularly those aimed at creating homozygosity through 'homozygosing' chromosomes.

In contrast to the central role of transposons in the genetics of most model systems, applications of transposition in the mouse are relatively undeveloped[68]. The recent addition of transposition to the mouse repertoire, through Sleeping Beauty, Minos or piggyBac[69–72], completes the full set of genetic tools available. Because the existing resources for and capabilities of gene trapping and targeting are very advanced, it is unlikely that transposition will have a strong role in genotype-driven research. Rather, it will probably join ENU mutagenesis as a tool for phenotype-driven research. Compared to ENU mutagenesis, transposition has the distinct advantage that the site of mutagenesis can be identified rapidly. Transgenic lines that support efficient transposition in the germ line have been developed to permit phenotypic screens based on feasible breeding program sizes[71,72]. In addition, transposons can be applied as somatic mutagens to screen for tumor promoting mutations on sensitized backgrounds[73].

Mouse genome engineering is replete with options. One of the main challenges involves finding the optimal ways in which to combine the options. This is a welcome and tractable problem. Another important issue involves determining the best way to use RNAi. Beyond one-by-one analysis lies the challenge of genetic interactions. The vast scale of combinatorial mutagenesis means that the definition of interactions and pathways will require maximal input from all relevant sources, including proteomics, expression genetics and genetical genomics[74,75].

ACKNOWLEDGMENTS

We thank H. von Melchner, B. Skarnes, A. Smith, G. Testa and G. Yancopoulos for discussions. This work was supported by funding from the Sixth Research Framework Programme of the European Union, Project FunGenES and the VW Program on Conditional Mutagenesis.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests (see the *Nature Genetics* website for details).

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

1. Austin, C.P. *et al.* The knockout mouse project. *Nat. Genet.* **36**, 921–924 (2004).
2. Auwerx, J. *et al.* The European dimension for the mouse genome mutagenesis program. *Nat. Genet.* **36**, 925–927 (2004).
3. 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).
4. van der Weyden, L., Adams, D.J. & Bradley, A. Tools for targeted manipulation of the mouse genome. *Physiol. Genomics* **11**, 133–164 (2002).
5. Metzger, D. & Chambon, P. Site-and time-specific gene targeting in the mouse. *Methods* **24**, 71–80 (2001).
6. von Melchner, H. & Stewart, A.F. Engineering of ES cell genomes with recombinase systems. in *Handbook of Stem Cells,* 609–623 (Elsevier, Burlington, Massachusetts, 2004).
7. Gardner, R.L. & Brook, F.A. Reflections on the biology of embryonic stem cells. *Int. J. Dev. Biol.* **41**, 235–243 (1997).
8. Buehr, M. & Smith, A. Genesis of embryonic stem cells. *Phil. Trans. R. Soc. Lond. B* **358**, 1397–1402 (2003).
9. Evans, M.J. & Kaufman, M.H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
10. 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).
11. Corcoran, L.M. & Metcalf, D. IL-5 and Rp105 signaling defects in B cells from commonly used 129 mouse substrains. *J. Immunol.* **163**, 5836–5842 (1999).

12. Crawley, J.N. Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacology (Berl.)* **132**, 107–124 (1997).
13. 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).
14. Tanaka, M., Hadjantonakis, A.K. & Nagy, A. Aggregation chimeras. Combining ES cells, diploid and tetraploid embryos. *Methods Mol. Biol.* **158**, 135–154 (2001).
15. Eggan, K. & Jaenisch, R. Differentiation of F1 embryonic stem cells into viable male and female mice by tetraploid embryo complementation. *Methods Enzymol.* **365**, 25–39 (2003).
16. Kunath, T. *et al.* Transgenic RNA interference in ES cell-derived embryos recapitulates a genetic null phenotype. *Nat. Biotechnol.* **21**, 559–561 (2003).
17. Seibler, J. *et al.* Rapid generation of inducible mouse mutants. *Nucleic Acids Res.* **31**, e12 (2003).
18. O'Gorman, S., Dagenais, N.A., Qian, M. & Marchuk, Y. Protamine–Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **94**, 14602–14607 (1997).
19. Liu, P., Jenkins, N.A. & Copeland, N.G. Efficient Cre-*lox*P-induced mitotic recombination in mouse embryonic stem cells. *Nat. Genet.* **30**, 66–72 (2002).
20. Belteki, G., Gertsenstein, M., Ow, D.W. & Nagy, A. Site-specific cassette exchange and germline transmission with mouse ES cells expressing ϕC31 integrase. *Nat. Biotechnol.* **21**, 321–324 (2003).
21. 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).
22. Zhang, Y., Buchholz, F., Muyrers, J.P. & Stewart, A.F. A new logic for DNA engineering using recombination in *E. coli*. *Nat. Genet.* **20**, 123–128 (1998).
23. 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).
24. Zhang, Y., Muyrers, J.P., Testa, G. & Stewart, A.F. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* **18**, 1314–1317 (2000).
25. Thomas, K.R., Deng, C. & Capecchi, M.R. High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. *Mol. Cell. Biol.* **12**, 2919–2923 (1992).
26. Hasty, P., Rivera-Perez, J. & Bradley, A. The length of homology required for gene targeting in embryonic stem cells. *Mol. Cell. Biol.* **11**, 5586–5591 (1991).
27. Valenzuela, D.M. *et al.* High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat. Biotechnol.* **21**, 652–659 (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. Moens, C.B., Auerbach, A.B., Conlon, R.A., Joyner, A.L. & Rossant, J.A. Targeted mutation reveals a role for N-myc in branching morphogenesis in the embryonic mouse lung. *Genes Dev.* **6**, 691–704 (1992).
30. Skarnes, W.C. *et al.* The International Gene Trap Consortium. *Nat. Genet.* **36**, 543–544 (2004).
31. Rodriguez, C.I. *et al.* High efficiency FLPe deleter mice provide a complement to Cre-*lox*P for *in vivo* genetic engineering. *Nat. Genet.* **25**, 139–140 (2000).
32. Schaft, J., Ashery-Padan, R., van der Hoeven, F., Gruss, P. & Stewart, A.F. Efficient FLP recombination in mouse ES cells and oocytes. *Genesis* **31**, 6–10 (2001).
33. Gu, H., Zou, Y.R. & Rajewsky, K. Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-*lox*P-mediated gene targeting. *Cell* **73**, 1155–1164 (1993).
34. Ringrose, L., Chabanis, S., Angrand, P.-O., Woodroofe, C. & Stewart, A.F. Quantitative comparison of DNA looping *in vitro* and *in vivo*: Chromatin increases effective DNA flexibility at short distances. *EMBO J.* **18**, 6630–6641 (1999).
35. Schnutgen, F. *et al.* Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. *Proc. Natl. Acad. Sci. USA* **102**, 7221–7226 (2005).
36. Albert, H., Dale, E.C., Lee, E. & Ow, D.W. Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J.* **7**, 649–659 (1995).
37. Schnutgen, F. *et al.* A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse. *Nat. Biotechnol.* **21**, 562–565 (2003).
38. 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).
39. Zambrowicz B.P., *et al.* Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention. *Proc. Natl. Acad. Sci. USA* **100**, 14019–14114 (2003).
40. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* **19**, 1129–1155 (2005).
41. Wobus, A.M. & Boheler, K.R. Embryonic stem cells: prospects for developmental biology and cell therapy. *Physiol. Rev.* **85**, 635–678 (2005).
42. Testa, G. *et al.* A reliable expression reporter cassette for multipurpose, knock-out/conditional mouse alleles. *Genesis* **38**, 151–158 (2004).
43. Adams, D.J. *et al.* Mutagenic insertion and chromosome engineering resource (MICER). *Nat. Genet.* **36**, 867–871 (2004).
44. Lefebvre, L., Dionne, N., Karaskova, J., Squire, J.A. & Nagy, A. Selection for transgene homozygosity in embryonic stem cells results in extensive loss of heterozygosity. *Nat. Genet.* **27**, 257–258 (2001).
45. Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349 (2004).
46. Kittler, R. *et al.* An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* **432**, 1036–1040 (2004).
47. Berns, K. *et al.* A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* **428**, 431–437 (2004).
48. Seibler, J. *et al.* Single copy shRNA configuration for ubiquitous gene knockdown in mice. *Nucleic Acids Res.* **33**, e67 (2005).
49. Ventura, A. *et al.* Cre-lox-regulated conditional RNA interference from transgenes. *Proc. Natl. Acad. Sci. USA* **101**, 10380–10385 (2004).
50. Wiznerowicz, M. & Trono, D. Conditional suppression of cellular genes: lentivirus vector–mediated drug-inducible RNA interference. *J. Virol.* **77**, 8957–8961 (2003).
51. Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S. & Capecchi, M.R. Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proc. Natl. Acad. Sci. USA* **97**, 13702–13707 (2000).
52. Loonstra, A. *et al.* Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc. Natl. Acad. Sci. USA* **98**, 9209–9214 (2001).
53. Logie, C. & Stewart, A.F. Ligand-regulated site-specific recombination. *Proc. Natl. Acad. Sci. USA* **92**, 5940–5944 (1995).
54. Forster, A. *et al.* Engineering *de novo* reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. *Cancer Cell* **3**, 449–458 (2003).
55. Herault, Y., Rassoulzadegan, M., Cuzin, F. & Duboule, D. Engineering chromosomes in mice through targeted meiotic recombination (TAMERE). *Nat. Genet.* **20**, 381–384 (1998).
56. Spitz, F., Herkenne, C., Morris, M.A. & Duboule, D. Inversion-induced disruption of the *Hoxd* cluster leads to the partition of regulatory landscapes. *Nat. Genet.* **37**, 889–893 (2005).
57. Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D. & Luo, L. Mosaic analysis with double markers in mice. *Cell* **121**, 479–492 (2005).
58. Zheng, B. *et al.* Engineering a mouse balancer chromosome. *Nat. Genet.* **22**, 375–378 (1999).
59. Baer, A. & Bode, J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr. Opin. Biotechnol.* **12**, 473–478 (2001).
60. Luo, J.L. *et al.* Knock-in mice with a chimeric human/murine p53 gene develop normally and show wild-type p53 responses to DNA damaging agents: a new biomedical research tool. *Oncogene* **20**, 320–328 (2001).
61. Testa, G. *et al.* Engineering the mouse genome with bacterial artificial chromosomes to create multi-purpose alleles. *Nat. Biotechnol.* **21**, 443–447 (2003).
62. Oram, M., Szczelkun, M.D. & Halford, S.E. Recombination. Pieces of the site-specific recombination puzzle. *Curr. Biol.* **5**, 1106–1109 (1995).
63. Smith, M.C. & Thorpe, H.M. Diversity in the serine recombinases. *Mol. Microbiol.* **44**, 299–307 (2002).
64. Olivares, E.C. *et al.* Site-specific genomic integration produces therapeutic Factor IX levels in mice. *Nat. Biotechnol.* **20**, 1124–1128 (2002).
65. Sauer, B. & McDermott, J. DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages. *Nucleic Acids Res.* **32**, 6086–6095 (2004).
66. Chen, Y. *et al.* Genotype-based screen for ENU-induced mutations in mouse embryonic stem cells. *Nat. Genet.* **24**, 314–317 (2000).
67. Greber, B., Lehrach, H. & Himmelbauer, H. Mouse splice mutant generation from ENU-treated ES cells—a gene-driven approach. *Genomics* **85**, 557–562 (2005).
68. Miskey, C., Izsvak, Z., Kawakami, K. & Ivics, Z. DNA transposons in vertebrate functional genomics. *Cell. Mol. Life Sci.* **62**, 629–641 (2005).
69. Largaespada, D.A. Generating and manipulating transgenic animals using transposable elements. *Reprod. Biol. Endocrinol.* **1**, 80 (2003).
70. Drabek, D. *et al.* Transposition of the *Drosophila hydei* Minos transposon in the mouse germ line. *Genomics* **81**, 108–111 (2003).
71. Ding, S. et al. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**, 473–483 (2005).
72. Bestor, T.H. Transposons reanimated in mice. *Cell* **122**, 322–325 (2005).
73. Dupuy, A.J., Akagi, K., Largaespada, D.A., Copeland, N.G. & Jenkins, N.A. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature* **436**, 221–226 (2005).
74. Churchill, G.A. *et al.* The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* **36**, 1133–1137 (2004).
75. Fraser, A.G. & Marcotte, E.M. A probabilistic view of gene function. *Nat. Genet.* **36**, 559–564 (2004).