© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

Engineering the mouse genome with bacterial artificial chromosomes to create multipurpose alleles

Giuseppe Testa[1], Youming Zhang[2], Kristina Vintersten[3], Vladimir Benes[3], W.W.M. Pim Pijnappel[1], Ian Chambers[4], Andrew J.H. Smith[4], Austin G. Smith[4], and A. Francis Stewart[1]*

Published online 10 March 2003; doi:10.1038/nbt804

The mouse is the leading vertebrate model because its genome can be altered by both random transgenesis and homologous recombination with targeting constructs. Both approaches have been hindered by the size and site limitations implicit in conventional *Escherichia coli* DNA-engineering methods. Homologous recombination in *E. coli*, or 'recombineering', has overcome these limitations for bacterial artificial chromosome (BAC) transgenesis[1–3]. Here we applied Red/ET recombineering (using the lambda Redα/Redβ recombinase pair)[4–6] to generate a 64 kilobase targeting construct that carried two selectable cassettes permitting the simultaneous mutation of the target gene, *Mll*, at sites 43 kb apart in one round of mouse embryonic stem (ES) cell targeting. The targeting frequency after dual selection was 6%. Because the two selectable cassettes were flanked by FRT or *lox*P sites, three more alleles can be generated by site-specific recombination. Our approach represents a simple way to introduce changes at two or more sites in a genetic locus, and thereafter generate allele combinations. The size of BAC templates offers new freedom for the design of targeting constructs. Combined with the use of two selectable cassettes placed far apart, BAC-based targeting constructs may be applicable to tasks such as regional exchanges, deletions, and insertions.

We made a targeting construct from a mouse *Mll* BAC using Red/ET recombination[4–6]. Our purpose was to create a *lox*P translocation model for human mixed-lineage leukemia (*MLL*)[7], to introduce a protein tag at the N terminus, and to establish a versatile allele from which several mouse lines could be generated to dissect *Mll* function. *MLL* translocations, prevalent in childhood leukemias[8], map to a short breakpoint cluster region of 8.3 kilobases that includes intron 11[9], which in mouse is located ~43 kilobases from the initiating methionine.

End sequencing from the vector established that the BAC contained 83 kb 5′ to the *Mll*-initiating methionine, and ended on the 3′ side at the 25th *Mll* intron. In the first steps (Fig. 1), the BAC was shaved on both sides by replacement with ampicillin and gentamicin selection cassettes, to leave homology arms for ES cell recombination of 5 kilobases on the 5′ side of the initiating methionine and 6.6 kilobases on the 3′ side of intron 11. The BAC was shaved for three reasons: first, to avoid hampering conventional Southern analysis that distinguishes homologous from random integrants; second, to choose the boundaries of the targeting construct according to prior design of a convenient Southern analysis strategy; third, to place sites for the rare cutting restriction enzyme, PI-*Sce*I, such that the vector could be severed from the targeting construct before ES cell electroporation. This was accomplished by inclusion of PI-*Sce*I sites in the synthetic oligonucleotides, between the 5′ *Mll* homology arms and 3′ PCR primers, used for the BAC shaving steps. After 5′ and 3′ shavings, the BAC was 64 kilobases long.

The next steps were simplified by the use of selectable cassettes that included dual prokaryotic and eukaryotic promoters to permit consecutive selection for correct recombinants in *E. coli* and then in ES cells[10]. The first cassette was designed to place the tandem affinity purification (TAP) protein tag[11] between the first and second codons of *Mll* using the hygromycin phosphotransferase gene (hygro). The hygro cassette was flanked by two FRT sites so that, upon FLPe recombination[12], the cassette would be removed to restore the open reading frame (ORF) as ATG-TAP-FRT-*Mll*. The second cassette, flanked by *lox*P sites, was placed in *Mll* intron 11. It includes a splice acceptor from engrailed-2 (en-2), the encephalomyocarditis virus (EMCV) internal ribosomal entry site (IRES), followed by a β-galactosidase-neomycin fusion (βgeok) and the simian virus-40 (SV40) late polyadenylation signal. βgeok is a variation of the βgeo cassette[13], in which a bacterial promoter and a ribosome-binding site were cloned between β-galactosidase and the *neo* gene, while keeping the fusion protein reading frame. This allows kanamycin selection in *E. coli* and G418 selection in ES cells from the βgeo fusion protein. Apart from ligating the oligonucleotides to the selectable cassettes (to avoid PCR mutagenesis), we followed the same procedure as for BAC shaving, adding at each step the appropriate antibiotic to identify colonies that had integrated the newly added selectable marker. In both rounds, >95% of resistant colonies were the intended homologous recombinants.

Feeder-independent E14TG2α ES cells were electroporated with 90 μg of the targeting construct that had been digested with PI-*Sce*I (Fig. 1C). Three regimens of selection—G418 alone, hygromycin alone, or both—were applied to evaluate whether the targeting construct retained its integrity or had broken (Table 1). Breakage or degradation would dissociate the two selectable marker cassettes. Alternatively, coresistance for the unselected marker would indicate that the electroporated DNA fragment remained intact during genomic integration. Overall ES cell colony numbers were similar to those produced with smaller targeting constructs under equivalent electroporation and selection conditions. As expected, fewer colonies were obtained by double selection, reflecting the additional challenge posed by co-selection. In both singly selected cases, most colonies were coresistant for the unselected cassette. This implies that either the construct remained intact upon integration in most cases, or the frequency of unselected second-site integration after construct breakage is high, or both.

Analysis of 47 doubly selected colonies revealed that 3 were homologous recombinants at both 5′ and 3′ sites (Fig. 2A,B, and data not shown). This frequency is similar to, if not higher than, frequencies seen in targeting experiments with conventionally sized constructs. Interestingly, a fourth colony (DB2; Fig. 2A,B) was positive for homologous recombination at the 3′, but not the 5′, end. Because this colony was resistant to both antibiotics, the construct either broke, with a hygromycin cassette integrating elsewhere in the genome, or provoked a local rearrangement similar to one noted during the targeting of N-*myc*[14]. Hence breakage and independent integration are possible. We generated chimeric mice from ES colony AC4. The genetic modification was transmitted through the germline to the fourth generation (Fig. 2C) with complete co-segregation of both selection cassettes (data not shown). Thus, we can exclude the unlikely scenario of simultaneous homologous recombination on different alleles.

[1]*Biotec, Technische Universität Dresden, and* [2]*Gene Bridges, GmbH, c/o Max Planck Institute for Molecular Cell Biology and Genetics, Pfotenhauerstr. 108, Dresden, Germany 01307.* [3]*European Molecular Biology Laboratory, Meyerhofstr. 1, Heidelberg, Germany 69117.* [4]*Centre for Genome Research, University of Edinburgh, Roger Land Building, Edinburgh, UK EH9 3JQ.* **Corresponding author (stewart@mpi-cbg.de).*

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

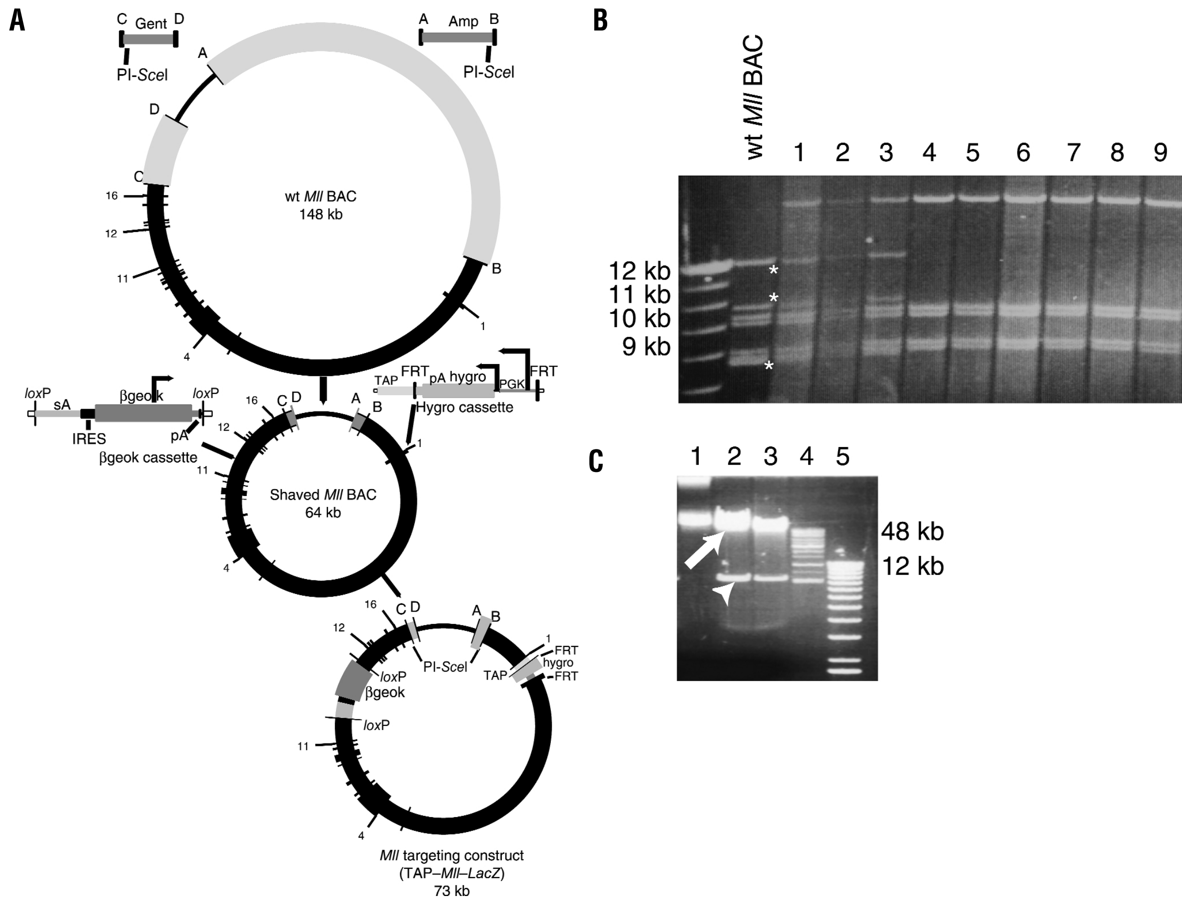

Figure 1. Red/ET engineering of a BAC targeting construct. (A) Schematic representation of the BAC modification strategy. The wild-type (wt) and the shaved *Mll* BACs, as well as the final *Mll* targeting construct, are drawn to scale with exons represented by numbered boxes. Amp and Gent represent the two PCR products used in the two BAC shaving rounds, conferring resistance to ampicillin and gentamicin, respectively. The oligonucleotide homology arms A, B, C, and D were taken from positions as indicated on the wt *Mll* BAC. A and D flank the BAC vector. The regions replaced with the Amp and Gent PCR products (respectively in the first and second deletion round) are depicted on the wt BAC in light gray; the remaining *Mll* region after shaving is in black. The oligonucleotides for the shaving steps each had 63 nucleotides of homology to the BAC. The two cassettes TAP-hygro and βgeok were targeted to the shaved *Mll* BAC in two successive rounds of Red/ET recombination to produce the final *Mll* targeting construct, TAP-*Mll-LacZ*. The BAC was obtained from Research Genetics, and all engineering was done in their host, HS996. (B) Restriction-fingerprinting analysis of BAC DNA after the first deletion round. A representative set of *Bam*HI digests shows that the intended deletion was obtained in 100% of colonies examined. Colonies 1, 2, and 3 represent mixed colonies, in which the intended deletion occurred only on one of two copies of the BAC within one cell. White asterisks indicate the wt bands missing from the deleted BACs. (C) 0.4% agarose gel electrophoresis of PI-*Sce*I digests of the TAP-*Mll-LacZ* construct before ES cell electroporation. Lane 1, uncut TAP-*Mll-LacZ*; lanes 2 and 3, TAP-*Mll-LacZ*, digested with PI-*Sce*I; lane 4, high-molecular-weight marker (Gibco BRL), whose highest band corresponds to 48 kb; 1 kb ladder (Gibco, BRL), whose highest band corresponds to 12 kb. The white arrow indicates the excised TAP-*Mll-LacZ* construct, whereas the white arrowhead indicates the BAC vector. A total of 5 liters were grown to give a yield of 270 μg BAC using the Large Construct Maxiprep kit (Qiagen).

To confirm that the *Mll* allele had been correctly engineered, we deleted the FRT and *lox*P cassettes in ES cells by co-electroporation of FLPe and Cre expression plasmids to restore the *Mll* mRNA and ORF with attached protein tag. Western blot analysis of extracts from ES cells (Fig. 2D) and brain (Fig. 2E) showed that the expected, TAP-tagged protein was expressed. We have also deleted both cassettes in mice by crossing to FLPe, Cre, and FLPe:Cre deleter lines. All progeny carrying the recombinase transgene(s) showed complete corresponding recombination (data not shown). Hence, all data show that the *Mll* gene was engineered correctly and simultaneously at two sites 43 kilobases apart.

As developed here by a first example, BAC-based targeting constructs permit the use of co-selection for two selectable cassettes separated by substantial distances. We now discuss four of the many implications for genomic engineering (Fig. 3).

A genomic region can be simultaneously engineered at two sites separated by distances accessible with BACs. Previously, this could be achieved only by two consecutive rounds of targeting in ES cells. Our approach replaces laborious consecutive targeting rounds in ES cells with rapid consecutive targeting rounds in *E. coli*. Not only is this easier, but it also reduces the handling of ES cells and the risks of impaired germline transmission or other second-site mutagenesis.

By selecting for both cassettes at the two ends of the construct, the region becomes a third area for introducing genetic change(s). This has particular relevance to the introduction of small mutations in coding regions or the placement of *lox*P sites for conditional strategies in chosen neutral intronic positions. Combined with recent methods for introducing seamless point mutations and small insertions in BACs[15,16], this represents a convenient means to discretely mutate given site(s) distant from selectable cassettes.

Recombination-mediated cassette exchange (RMCE) is the major current strategy for exchanging or deleting more than a few kilobases[17]. In RMCE, two different Site Specific Recombination Target sites (SSRTs) are integrated by homologous recombination in two rounds of targeting; in a third round, a site-specific recombinase then mediates exchange of an incoming SSRT-flanked interval with the endogenous one. Within the size range of BACs, our results present (i) a simultane-

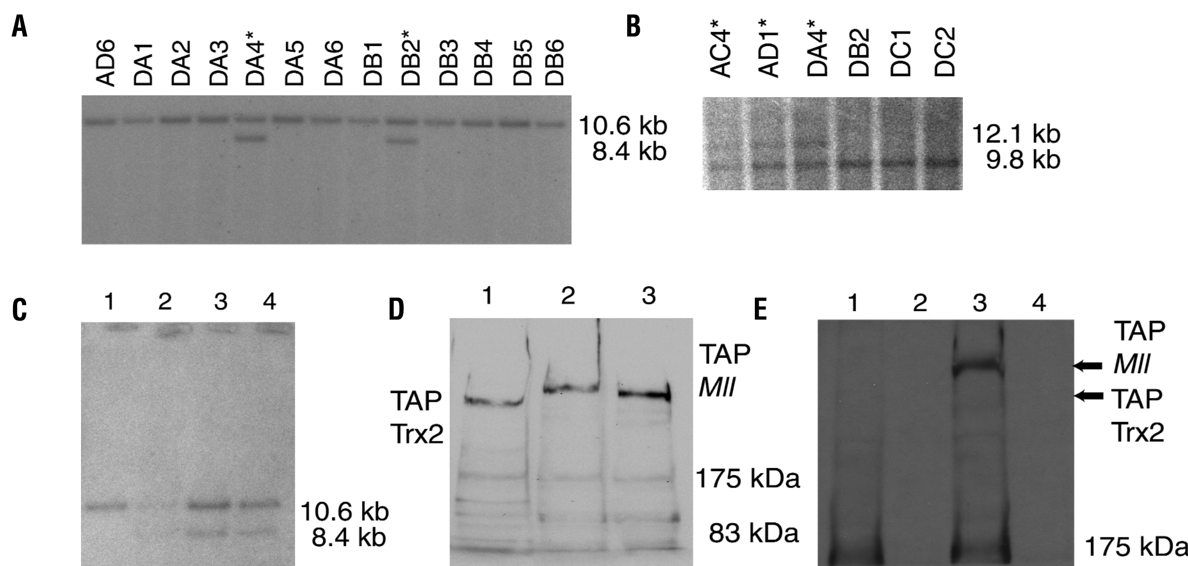

Figure 2. Analysis of the allele in ES cells and mice. (A) Southern analysis strategy to detect homologous recombination at the 3′ end of the TAP-*Mll-LacZ* construct with an external probe (3′ probe indicated in Fig. 3). The wild-type (10.6 kb) and the recombined (8.4 kb) bands are indicated. Colonies indicated by an asterisk are correct recombinants. (B) Southern analysis strategy to detect homologous recombination at the 5′ end of the TAP-*Mll-LacZ* construct with an external probe (5′ probe indicated in Fig. 3). The wild-type (9.8 kb) and the correctly recombined (12.1 kb) bands are indicated. (C) Germline transmission assessed by Southern blot hybridization with the 3′ external probe on a litter of four embryos harvested at embryonic day 10.5. Embryo 1 is wild type; embryos 2–4 are heterozygous for the targeted allele. The wild-type (10.6 kb) and the recombined (8.4 kb) bands are indicated. (D) Western blots of protein extracts from a recombined derivative of clone AC4. ES cell clone AC4 was co-electroporated with expression vectors for FLPe and Cre. Recombination at either FRT or *lox*P sites was 5%, and 1.4% of colonies were doubly recombined. Nuclear extract was prepared from ES cells (lane 2) and embryoid bodies (lane 3), loaded on a 5% polyacrylamide gel, and probed with a PAP (peroxidase-anti-peroxidase)-coupled IgG antibody (Sigma) to detect the protein A moiety of the TAP tag. In lane 1, nuclear extract from a mouse organ expressing a TAP-tagged *Mll* homolog (Trx2), which is also large (310 kDa) was loaded as a size control. Other protein size markers are indicated. (E) Western blot of brain protein extract. Nuclear (lanes 1 and 3) and cytoplasmic (lanes 2 and 4) extracts were prepared from the brain of a wild-type mouse (lanes 1 and 2, respectively) and a high-grade chimera (>80%), generated by blastocyst injection of the doubly recombined ES cell (lanes 3 and 4, respectively). Extracts were loaded on a 5% polyacrylamide gel and probed with a PAP-coupled IgG antibody to detect the protein A moiety of the TAP tag. The sizes of the highest protein marker band (175 kda) and TAP-Trx2 are indicated.

ous, not consecutive, way to insert the two SSRTs for later RMCE, and (ii) the possibility that the exchange of an interval can be directly accomplished in one step of homologous recombination in ES cells. The second possibility may not be feasible if lengthy regions of nonhomology between the two selectable cassettes substantially impair the efficiency of homologous recombination in ES cells. Further work is required to explore this possibility. It may also be possible to use BAC-based dual selectable cassettes as recombination 'scissors' to delete BAC-sized regions of the mouse genome.

Multipurpose alleles make use of site-specific recombination to change an allele from one state to another[18,19]. The most popular application involves the creation of conditional alleles. Usually *lox*P sites are placed on either side of an exon(s) of the gene to be conditionally mutated by Cre recombinase in the mouse, whereas FRT sites surround the selectable cassette and FLPe recombinase is used to delete the selectable cassette. Before FLPe recombination, the presence of the FRT cassette may also create a useful hypomorphic allele. Here we took advantage of the freedom offered by BAC-based targeting constructs to design a complex multipurpose allele that could serve at least four goals. Other variations of designed multipurpose alleles are now possible. For example, we suggest the application of BAC targeting constructs to insert *lox*P sites at positions between the selectable cassettes, rather than flanking one of the selectable cassettes. In the absence of flanking *lox*P sites, the selectable cassette could be placed well outside of the gene. Ideally, however, it would be flanked by SSRTs from a third site-specific recombinase-SSRT system. We note that the application in mice of a third system, based on phage phiC31 from *Streptomyces*, is underway[20]. Molecular evolution of Cre or FLP to recognize different target sites also provides options for an additional recombinase system[21]. An allele built with three site-specific recombinase options could theoretically generate at least seven alleles.

BAC-based targeting constructs permit the design of multipurpose alleles that can combine various options, such as knockout, hypomorphic, and conditional alleles, with expression reporters, protein tags, knockins, and gain-of-function cassettes. Our *Mll* allele is an example of such new combinatorial alleles built in BACs. Notably, *Mll* has been targeted three times, with each mutation showing strikingly different phenotypes[22]. Our *Mll* multipurpose allele series represents a more systematic approach to the dissection of *Mll* function in development and illustrates a general approach for other studies.

We have shown that a reasonably large targeting construct was efficiently electroporated into ES cells and predominantly integrated intact. Homologous recombination was recorded at frequencies similar to, if not higher than, those reported for standard vectors. We speculate than an advantage of large targeting constructs will emerge as further

Table 1. Results from the three selection regimens.

Selection scheme	G418 200 µg/ml	hygromycin 160 µg/ml	G418 200 µg/ml plus hygromycin 160 µg/ml
Colonies/dish[a]	900	680	450
Coresistance[b]	83% ± 5	76%	

[a]The absolute number of colonies/plate.
[b]Percentage of colonies initially selected with one drug that were also resistant to the other drug.

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

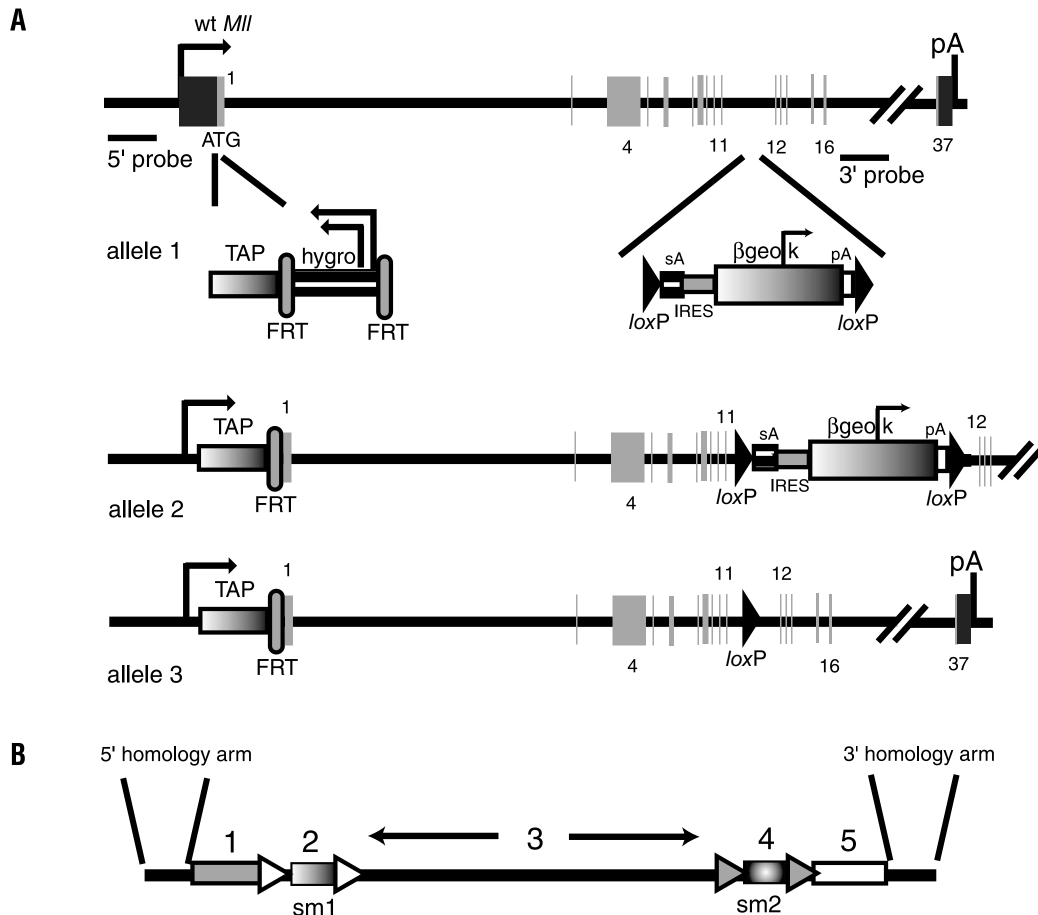

Figure 3. Schematic representation of multi-purpose alleles. (A) Diagram of the wt mouse *Mll* gene with the two targeting cassettes depicted below representing allele 1. Exons are shown as grey boxes. The two black arrows in the hygro cassette in opposite orientation to the *Mll* gene represent the prokaryotic (Tn903) and eukaryotic (PGK) promoters used for selection in *E. coli* and ES cells, and the small arrow in the middle of the βgeok represents the prokaryotic Tn5 promoter. Allele 1 is intended to be a complete knock-out with the TAP-FRT-hygro cassette disturbing the protein reading frame. Allele 2 is generated after FLPe recombination to restore the reading frame and position the TAP tag onto the N-terminus of Mll. Allele 2 is designed to be a tagged hypomorphic (or dominant negative) allele due to transcript truncation after exon 11. It is also amenable to conditional gene repair strategies upon Cre deletion of the βgeok cassette in intron 11. Allele 3 is generated after FLPe and Cre recombination and is designed to be a tagged wt allele for biochemical characterization. Furthermore, Cre mediated interchromosomal translocation[7] between the loxP site in intron 11 and a loxP site in a relevant intron of an *Mll* translocation partner represents a conditional model for *Mll* leukemiogenesis. (B) Schematic representation of the general architecture of large, dual selection targeting constructs. The homology arms lie at the 5′ and 3′ ends and are separated by five positions that can be occupied by a range of different cassettes or modules. The external positions (1 and 5) present options for inclusion of knock-in elements including fusion tags, expression reporters, and coding or *cis*-element mutations. In positions 2 and 4 the selectable marker cassettes (sm1 and sm2) are flanked by SSRTs. Selective use of site specific recombination removes the cassettes and generates the various alleles. Position 3 is the intervening region between the selectable marker cassettes. It could be engineered to contain point mutations and/or additional SSRTs to add greater flexibility to conditional strategies. It may also be possible to use position 3 for the replacement or insertion of large stretches of foreign DNA so that, by double selection, these foreign DNA sections can be placed in chosen mouse genomic sites. Abbreviations: ATG, start codon of *Mll*; βgeok, fusion of βgalactosidase and neomycin genes, including a prokaryotic promoter for kanamycin selection; hygro, hygromycin phosphotransferase gene; IRES, internal ribosomal entry site; pA, polyadenylation signal; sA, splice acceptor; TAP, Tandem Affinity Purification protein tag.

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

experience is gathered. Since a larger region of homologous DNA is included, it is possible that targeting frequencies will typically be reasonable to good, and that there will be fewer examples of the extremely poor frequencies sometimes experienced with conventionally sized targeting constructs.

Experimental protocol

Animals. All animal experiments were conducted in a licensed animal facility in accordance with the German Animal Welfare Act (Tierschutzgesetz), following the guidelines of the European Convention for the Protection of Vertebrate Animals Used for Experimental and Other Scientific Purposes (1986).

BAC engineering. BACs were obtained from Research Genetics in *E. coli* HS996. All engineering was conducted in *E. coli* HS996, and the BACs were directly sequenced as described[23].

For the first round of shaving (Fig. 1A), a PCR product containing the β-lactamase gene (Amp) replaced the 5′ end of the *Mll* BAC from the end of the BAC vector to ~5 kb upstream of the initiating ATG. The PCR primer oligos included at their 5′ ends 63 nucleotides of homology, and one included the 38 bp restriction site for the intein-encoded endonuclease PI-*Sce*I. The Amp PCR product was electroporated into *E. coli* HS996 cells harboring the *Mll* BAC and the R6K-αβγ plasmid, which expresses the lambda phage proteins Redα, Redβ, and Redγ[6]. Recombinant colonies were identified by double selection for chloramphenicol and ampicillin, and then restriction fingerprinting (Fig. 2A) and end sequencing of the insertion site. The approach for the second round of shaving was the same as above except the PCR product included the gentamicin-resistance gene and the deletion left 6.6 kb downstream of the βgeok insertion site in intron 11. Recombinant colonies were identified by triple selection for chloramphenicol, ampicillin, and gentamicin. These integration cassettes were first built as plasmids by insertion of *Mll* PCR products or oligos at either side of

the TAP-Hygro or βgeok cassettes. For each engineering step, cells harboring the BAC were transformed with the Red/ET expression plasmid pR6K-αβγ by standard procedures. Single colonies were picked and grown in 5 ml Luria (L) broth overnight. Then, 0.7 ml was transferred into 70 ml of L-broth (without glucose) and grown at 37 °C. At OD_{600} = 0.1–0.15, 0.7 ml of 10% L-arabinose (wt/vol) was added to induce Red protein expression. At OD_{600} 0.25–0.4, cells were centrifuged (10 min, 3,800*g*, –5 °C, Sorvall SS34 rotor; plastic tubes) and resuspended in 30 ml ice-cold 10% glycerol (vol/vol) (repeated three times). After the final centrifugation, decanting, and wiping the tube with a tissue, cells were resuspended in a minimum residual volume (<500 µl final resuspended volume) and 50 µl aliquots were immediately electroporated with 0.3–1 µg of linear DNA fragment (PCR product or a fragment excised from a plasmid). Colonies were identified on selection plates containing appropriate combinations of the following antibiotics (tetracycline 25 µg/ml for pR6K-αβγ, chloramphenicol 12.5 µg/ml, ampicillin 50 µg/ml, gentamicin 3 µg/ml, kanamycin 20 µg/ml, hygromycin 50 µg/ml).

BAC DNA was prepared using the Large Construct Maxiprep Kit (Qiagen, Düsseldorf, Germany). Five liters yielded 270 µg. After PI-*Sce*I digestion of the targeting construct (10 µg, 20 U enzyme, 200 µl final volume, 37 °C, overnight incubation), the DNA was extracted with phenol-chloroform, precipitated with ethanol, and resuspended in PBS.

ES cells. Mouse E14Tg2a (ref. 24) ES cells were cultured without feeders in medium supplemented with recombinant leukemia inhibitory factor (LIF) as described[24]. Electroporation was carried out with a BioRad Gene Pulser using standard conditions (3 µF, 800 V), adding 90 µg DNA in 90 µl PBS (without calcium and magnesium ions) to 7.4×10^{7} cells, resuspended in 710 µl PBS. Cells were seeded at a density of 2×10^6 cells/10 cm plate. At 24 h after electroporation, drug selection was started at the following concentrations: G418 (200 µg/ml); hygromycin (160 µg/ml).

ES cell injections were done using standard procedures[25].

Protein nuclear extracts from ES cells, embryoid bodies, and mouse organs were done with the CellLytic NuClear extraction kit from Sigma (Deisenhofen, Germany). Western blotting was done following standard procedures. The protein A part of the TAP tag[11] was visualized with a peroxidase-coupled anti-peroxidase rabbit antibody (Sigma) only (no first antibody), and the ECL system (Amersham).

Acknowledgments
We wish to thank Konstantinos Anastassiadis, William Brown, Frank van der Hoeven, Robin Lovell-Badge, and Daniela Nebenius-Oosthuizen for discussions and help. This work was partly funded by a grant from the Volkswagen Foundation, Program on Conditional Mutagenesis. The work was initiated at the European Molecular Biology Laboratory (EMBL), Heidelberg.

Competing interests statement
The authors declare competing financial interests: see the Nature Biotechnology website (http://www.nature.com/naturebiotechnology) for details.

Received 2 October 2002; accepted 15 January 2003

1. Muyrers, J.P., Zhang, Y. & Stewart, A.F. Techniques: recombinogenic engineering—new options for cloning and manipulating DNA. *Trends Biochem. Sci.* **26**, 325–331 (2001).
2. 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).
3. Heintz, N. BAC to the future: the use of bac transgenic mice for neuroscience research. *Nat. Rev. Neurosci.* **2**, 861–870 (2001).
4. 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).
5. Muyrers, J.P., Zhang, Y., Testa, G. & Stewart, A.F. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.* **27**, 1555–1557 (1999).
6. Zhang, Y., Muyrers, J.P., Testa, G. & Stewart, A.F. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* **18**, 1314–1317 (2000).
7. Testa, G. & Stewart, A.F. Creating a translocation. Engineering interchromosomal translocations in the mouse. *EMBO Rep.* **1**, 120–121 (2000).
8. Rowley, J.D. The critical role of chromosome translocations in human leukemias. *Annu. Rev. Genet.* **32**, 495–519 (1998).
9. Nilson, I. *et al.* Exon/intron structure of the human ALL-1 (*MLL*) gene involved in translocations to chromosomal region 11q23 and acute leukaemias. *Br. J. Haematol.* **93**, 966–972 (1996).
10. 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).
11. Rigaut, G. *et al.* A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032 (1999).
12. Buchholz, F., Angrand, P.O. & Stewart, A.F. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat. Biotechnol.* **16**, 657–662 (1998).
13. Zambrowicz, B.P. *et al.* Disruption of overlapping transcripts in the ROSA βgeo26 gene trap strain leads to widespread expression of β-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl. Acad. Sci. USA* **94**, 3789–3794 (1997).
14. 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).
15. Muyrers, J.P. *et al.* Point mutation of bacterial artificial chromosomes by ET recombination. *EMBO Rep.* **1**, 239–243 (2000).
16. Ellis, H.M., Yu, D., DiTizio, T. & Court, D.L. High-efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. USA* **98**, 6742–6746 (2001).
17. Baer, A. & Bode, J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr. Opin. Biotechnol.* **12**, 473–480 (2001).
18. Meyers, E.N., Lewandoski, M. & Martin, G.R. An Fgf8 mutant allelic series generated by Cre- and Flp-mediated recombination. *Nat. Genet.* **18**, 136–141 (1998).
19. Nagy, A. *et al.* Dissecting the role of N-*myc* in development using a single targeting vector to generate a series of alleles. *Curr. Biol.* **8**, 661-664 (1998).
20. Andreas, S., Schwenk, F., Kuter-Luks, B., Faust, N. & Kuhn, R. Enhanced efficiency through nuclear localization signal fusion on phage PhiC31-integrase: activity comparison with Cre and FLPe recombinase in mammalian cells. *Nucleic Acids Res.* **30**, 2299-2306 (2002).
21. Buchholz, F. & Stewart, A.F. Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nat. Biotechnol.* **19**, 1047–1052 (2001).
22. Ayton, P. *et al.* Truncation of the *Mll* gene in exon 5 by gene targeting leads to early preimplantation lethality of homozygous embryos. *Genesis* **30**, 201–212 (2001).
23. Benes, V., Kilger, C., Voss, H., Paabo, S. & Ansorge, W. Direct primer walking on P1 plasmid DNA. *Biotechniques* **23**, 98–100 (1997).
24. 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).
25. Joyner, A.L. (ed.). *Gene Targeting. A Practical Approach* (Oxford University Press, New York, 2000).

Site-specific gene targeting in mouse embryonic stem cells with intact bacterial artificial chromosomes

Yi Yang and Brian Seed*

Published online 10 March 2003; doi:10.1038/nbt803

Homologous recombination in *Escherichia coli* simplifies the generation of gene targeting constructs for transduction into mouse embryonic stem (ES) cells[1–7]. Taking advantage of the extensive homology provided by intact bacterial artificial chromosomes (BACs), we have developed an efficient method for preparing targeted gene disruptions in ES cells. Correctly integrated clones were identified by a simple screening procedure based on chromosomal fluorescence *in situ* hybridization (FISH). To date, five mutant lines have been generated and bred to homozygosity by this approach.

Site-specific gene disruption in mice has been an important tool for the analysis of gene function *in vivo*[8,9]. However, the process is resource intensive and the targeting frequencies typically are low. As currently practiced, creation of a targeting vector requires the formation in bacteria of a replica of the desired disrupted gene, followed by introduction of the replica into ES cells and identification of the properly targeted DNA. Orthotopic integration can be

Department of Molecular Biology, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston, MA 02114.
**Corresponding author (seed@molbio.mgh.harvard.edu).*

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology