Invited Review: Functional genomics in the mouse: powerful techniques for unraveling the basis of human development and disease

Clifford W. Bogue

J Appl Physiol 94:2502-2509, 2003. doi:10.1152/japplphysiol.00209.2003

You might find this additional info useful...

This article cites 52 articles, 19 of which can be accessed free at:
http://jap.physiology.org/content/94/6/2502.full.html#ref-list-1

This article has been cited by 3 other HighWire hosted articles

An interview study of phenotypic characterization of genetically-modified mice
R Thon, H Vondeling, J Lassen, A K Hansen and M Ritskes-Hoitinga
Lab Anim, July , 2009; 43 (3): 278-283.
[Abstract] [Full Text] [PDF]

Sleep disruption is related to allelic variation in the *ob* gene
Ralph Lydic
Am J Physiol Regul Integr Comp Physiol 2006; 290 (4): R892-R893.
[Full Text] [PDF]

C57BL/6J and B6.V-LEP^{OB} mice differ in the cholinergic modulation of sleep and breathing
Christopher L. Douglas, Grant N. Bowman, Helen A. Baghdoyan and Ralph Lydic
J Appl Physiol, March , 2005; 98 (3): 918-929.
[Abstract] [Full Text] [PDF]

Updated information and services including high resolution figures, can be found at:
http://jap.physiology.org/content/94/6/2502.full.html

Additional material and information about ***Journal of Applied Physiology*** can be found at:
http://www.the-aps.org/publications/jappl

This information is current as of June 5, 2012.

Journal of Applied Physiology publishes original papers that deal with diverse areas of research in applied physiology, especially those papers emphasizing adaptive and integrative mechanisms. It is published 12 times a year (monthly) by the American Physiological Society, 9650 Rockville Pike, Bethesda MD 20814-3991. Copyright © 2003 by the American Physiological Society. ISSN: 0363-6143, ESSN: 1522-1563. Visit our website at http://www.the-aps.org/.

Downloaded from jap.physiology.org on June 5, 2012

J Appl Physiol 94: 2502–2509, 2003;
10.1152/japplphysiol.00209.2003.

highlighted topics

Genetic Models in Applied Physiology

Invited Review: Functional genomics in the mouse: powerful techniques for unraveling the basis of human development and disease

Clifford W. Bogue
Yale Child Health Research Center, Section of Critical Care and Applied Physiology, Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut 06519

Bogue, Clifford W. Invited Review: Functional genomics in the mouse: powerful techniques for unraveling the basis of human development and disease. *J Appl Physiol* 94: 2502–2509, 2003; 10.1152/japplphysiol.00209.2003.—Now that near-complete DNA sequences of both the mouse and human genomes are available, the next major challenge will be to determine how each of these genes functions, both alone and in combination with other genes in the genome. The mouse has a long and rich history in biological research, and many consider it a model organism for the study of human development and disease. Over the past few years, exciting progress has been made in developing techniques for chromosome engineering, mutagenesis, mapping and maintenance of mutations, and identification of mutant genes in the mouse. In this mini-review, many of these powerful techniques will be presented along with their application to the study of development, physiology, and disease.

phenotype; genome

THE MOUSE: A KEY MODEL ORGANISM FOR STUDYING HUMAN DEVELOPMENT AND DISEASE

The mouse has a long and profitable history in biomedical research, and its origin as the leading model system for biomedical research can be traced back to the beginning of human civilization (23, 34). From early on in history, humans noticed spontaneously arising coat-color mutants and recorded their observations for thousands of years; by the 1700s, many varieties of mice had been domesticated as pets by mouse fanciers in Japan and China. This practice was adopted by the Europeans, who subsequently began breeding these strains with local mice, creating the progenitors of modern laboratory mice (5, 23, 45). Genetic mapping in mice began in 1915 when Haldane and colleagues (16) reported linkages between the pink-eye dilution and albino loci on a linkage group that was eventually assigned to mouse chromosome 7.

Although the publication of the human draft genome sequence has heralded a new era of research in genetics (26, 51), the obvious logistical obstacles of carrying out experimental studies in humans necessitate the use of a model organism. The mouse is quite similar to humans physiologically and is considered by many to be the ideal model organism. Some important features that make the mouse such an ideal organism in which to study biological processes relevant to humans include its relatively short generation time (~10 wk), its small size, the history of over a century of genetic studies, and the existence of many inbred strains and hundreds of spontaneous mutations. In addition, practical techniques are now available for random mutagenesis and directed engineering of the genome through knockout, knockin, and transgenic techniques. It is thought that mice and humans diverged from a common ancestor ~65–75 million years ago, yet most salient aspects of mammalian physiology have not diverged significantly in these lineages during this time. Both organisms have the same organ systems, similar reproductive cycles, similar skeletons, and quite similar biochemistries, physiologies, and pathologies. We now know that the similarity of humans and mice extends to their genomes as well.

Mus musculus, a species of mouse, is one of the five key model organisms sequenced as part of the Human Genome Project. In 1998–1999, the National Institutes of Health published an action plan for mouse genomics that, among other things, called for a working draft sequence of the mouse genome by 2003. The Mouse Genome Sequencing Consortium, an international consortium dedicated to producing such a working draft sequence, recently achieved that goal with its publica-

Address for reprint requests and other correspondence: C. W. Bogue, Yale Child Health Research Center, Section of Critical Care and Applied Physiology, Dept. of Pediatrics, Yale Univ. School of Medicine, 464 Congress Ave., New Haven, CT 06519 (E-mail: clifford.bogue@yale.edu).

 8750-7587/03 $5.00 Copyright © 2003 the American Physiological Society

Downloaded from jap.physiology.org on June 5, 2012

tion of the genomic sequence of the mouse strain C57BL/6J (52). Comparison of the mouse and human genome sequences reveals extensive similarities and some intriguing differences. Humans and mice have ~30,000 genes each, although the mouse genome is ~14% smaller than the human genome (2.5 vs. 2.9 gigabytes). There is a remarkable degree of synteny between the two genomes: over 90% of the mouse and human genomes can be partitioned into regions of conserved synteny in which the gene order on the chromosome is conserved. Of the ~30,000 protein-coding genes in the mouse genome, 99% of these have a sequence match in the human genome; when considering the entire mouse and human genomes at the nucleotide level, there is ~40% identity. The opportunities to use the information contained in the mouse and human genomes to study human disease and to devise new therapies to treat human disease with the use of the mouse as a model system are extraordinary. This new information will radically change the way that experimental genetics can be done.

FUNCTIONAL GENOMIC APPROACHES TO STUDY GENE FUNCTION AND REGULATION

Understanding the function of genes and other parts of the genome is known as functional genomics. The Human Genome Project is just the first step in understanding humans at the molecular level. Now that the sequencing phase of the human and mouse genomes is complete, many questions remain unanswered, including the function of most of the estimated 30,000–35,000 mouse and human genes. In the following sections, I will present a few of the techniques that are being applied to the daunting yet exciting task of functional genomic analysis in the mouse.

Targeted mutagenesis. Presently, very few of the ~30,000 genes present in the mouse genome have been mutated; those that have been mutated were mostly by homologous recombination in embryonic stem cells (gene knockout). Although this technique is both time and resource intensive, it remains the best way to determine the function of a specific gene in vivo. Many of the genes that have been mutated via targeted mutagenesis are cataloged at the Jackson Laboratory (http://tbase.jax.org/). These targeted mutations have provided a significant amount of information to biomedical research so far. There have been, and continue to be, a number of ongoing refinements to the technique of targeted mutagenesis that serve to strengthen the power of this technique for studying gene function. For instance, rather than just disrupting the function of an entire gene by homologous recombination (knockout), it is also possible to introduce subtle missense and/or gain of function mutations (knockin) in a specific gene or gene regulatory element (41, 44). Thus the function of specific domains of a gene can be elucidated in vivo, as can the role of gene regulatory elements such as tissue-specific enhancers. In addition, one major advance in the field of targeted mutagenesis is the development of techniques for conditional control of gene expression in vivo. The most common techniques used for conditional gene expression in the mouse make use of binary transgenic systems in which the conditional expression of a gene is controlled by the interaction of two integral components: an “effector” transgene, whose product interacts, in turn, with a

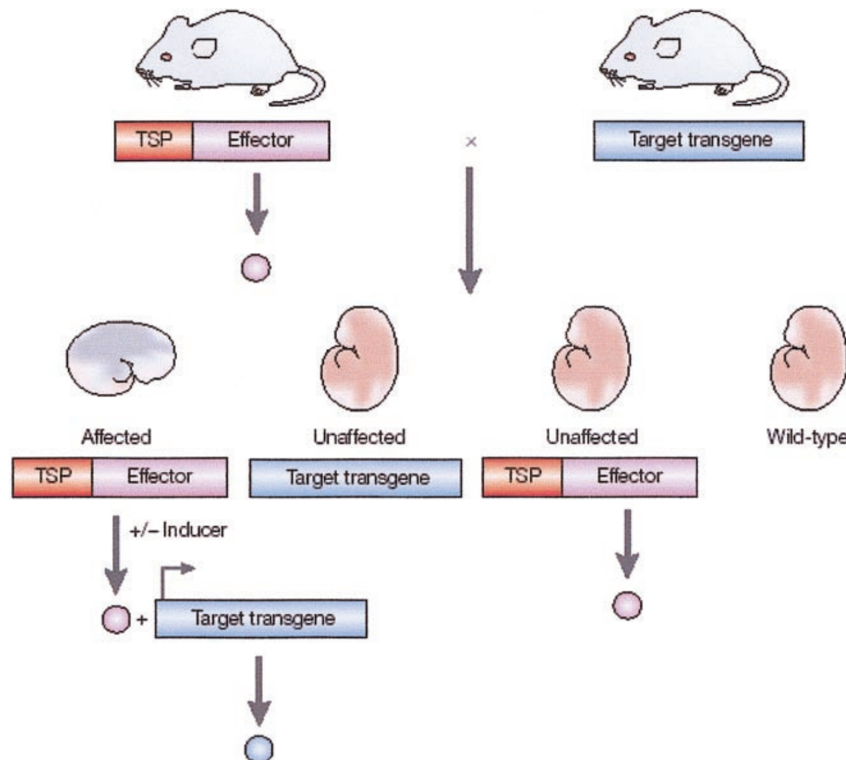

Fig. 1. Binary transgenics in the mouse. The binary transgenic system is based on the product of an “effector” transgene acting on a “target” transgene to activate or silence the transgene of choice. The effector transgene is expressed in specific temporospatial patterns under the control of a tissue-specific promoter (TSP). Expression of the target transgene should only occur in doubly transgenic F1 progeny derived from mouse lines that carry either the target transgene or the effector transgene or in the presence of an exogenous inducer that allows the appropriate interaction of the effector on the target. The effector gene product-target gene interaction results in target activation. In some cases, (such as when using a site-specific DNA recombinase such as Cre), the effector can also inactivate the target transgene. In most inducible binary systems, the effector gene product-target transgene interaction depends on the presence of an exogenously added inducer. However, withdrawal of an inducer can also result in effector action. (Reproduced from Ref. 28 with permission from Nature Publishing Group.)

Downloaded from jap.physiology.org on June 5, 2012

target transgene (Fig. 1) (Ref. 10 and reviewed in Ref. 28). There are generally two categories of binary systems used for conditional gene expression. In one, the effector transgene transactivates the transcription of the target transgene; this technique is commonly used for targeting expression of a gene to certain tissues or developmental time points. In the other, the effector is a site-specific DNA recombinase that rearranges the target gene, thereby activating or silencing it.

The most widely used binary systems for gene transactivation are the tetracycline-regulated systems originally developed by Gossen and Bujard (13). In these systems, the effector is a chimeric construct that includes the potent transactivator VP16 fused to the *Escherichia coli* tetracycline repressor protein (TetR). This effector can bind both tetracycline and a target transgene constructed of the 19-bp operator sequences (*tetO*) of the *tet* operon. This then results in gene transcription. In the original version of this system, the tetracycline-controlled transactivator is not able to bind DNA when the inducer, tetracycline, is present. This system is referred to as tTa or the "tet-off" system. In a modified version developed in 1995, referred to as "reverse tTa" (rtTa), the effector binds to the target transgene only when the inducer is present (14). The principal difference between the two systems is the kinetics of transgene induction. In the tTa system, suppression of transgene expression depends on continuous exposure to the inducer. On withdrawal of the inducer, the timing of transgene expression depends on the rate of clearance of the inducer. In the rtTa system, the addition of the inducer rapidly induces transgene expression (perhaps even within hours), but subsequent repression and transgene inactivation depend on the kinetics of clearance of the inducer. Presently, doxycycline is the inducer of choice due to its low cost, easy availability, and its ability to efficiently activate rtTa and inactivate tTa at doses that are well below those that cause cytotoxicity (2, 14). A second binary system that has been generated for inducible transactivation of a gene product is the *Gal4*-based system, which is based on the transcriptional activator *Gal4* from *Saccharomyces cerevisae*. *Gal4* transactivates target genes by binding to upstream activator sequences and is the first nonviral binary system to be used in mice (2, 14). This system shows some promise but presently is not as widely used as the tetracycline-based systems.

The development of site-specific DNA recombinases for use in mice to conditionally control gene expression is a significant advance in the field of functional genomics. The biggest disadvantage of traditional gene knockouts is the fact that irreversible gene deletion takes place in the germline and often results in embryonic lethality. Because many genes function at many different stages of development and in adult tissues, one can only assess the earliest, nonredundant function of a gene, and analysis of its function later in development is precluded. Additionally, in a global gene knockout, it can be difficult to decipher cell autonomous from noncell autonomous or systemic effects of the gene mutation. Therefore, the use of site-specific DNA recombination allows the investigator to circumvent these limitations by silencing a gene at the specific time and place of their choosing. There are two types of site-specific recombinases that are commonly used to conditionally control gene expression in the mouse: Cre recombinase from the bacteriophage P1 (42) and Flp recombinase from *Saccharomyces cerevisae* (38). Both of these recombinases catalyze DNA recombination between two 34-bp recognition sequences that are similar in their secondary structure but differ in their primary sequence (Fig. 2). The recognition sequences for Cre and Flp are *loxP* and *FRT*, respectively. To generate a tissue-specific knockout,

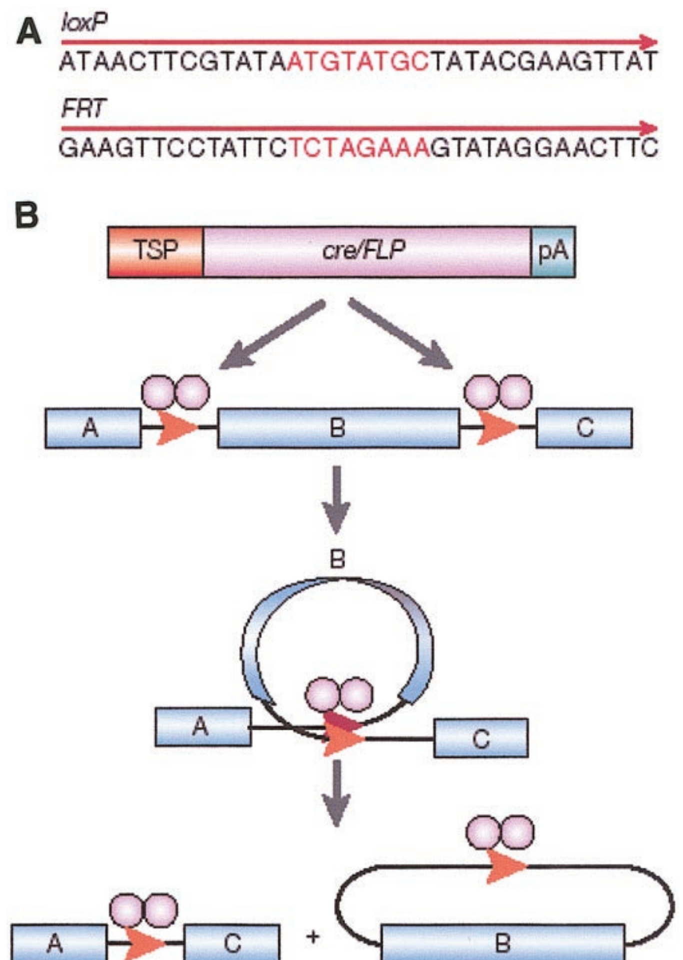

Fig. 2. Controlling gene expression by DNA recombination. *A*: the 34-bp *loxP* and *FRT* sites each consist of two 13-bp inverse repeats (black) that flank an 8-bp core sequence (red). This core sequence confers directionality to these sites (red arrows). *B*: dimers of Cre or Flp (pink) catalyze in *cis* the conservative recombination between two directly repeated *loxP* or *FRT* sites (red arrowheads), resulting in the formation of a synaptic structure, the excision of *region B* and the juxtaposition of *regions A* and *C*. If *region B* is an essential region of a gene, then the recombination event results in gene inactivation. Recombination can also activate gene expression. For example, transcription from a promoter in *region A* could fail to reach protein-coding sequences in *region C* if polyadenylation sites exist in *region B*; excising *region B* would therefore activate transcription of *region C*. pA, polyadenylation site; TSP, tissue-specific promoter. (Reproduced from Ref. 28 with permission from Nature Publishing Group.)

Downloaded from jap.physiology.org on June 5, 2012

homologous recombination in embryonic stem cells is used to generate an allele in which the gene targeted for inactivation has *loxP* or *FRT* recognition sequences inserted so that recombination between the recognition sites leads to gene deletion or inactivation ("floxed" or "FRTed"). In general, the *loxP* or *FRT* sites are placed in intronic sequence so that, in the absence of Cre or Flp recombinase, the function of the targeted gene is not disrupted. Then, to effect tissue-specific gene deletion, these mice are mated to mice that harbor a transgene composed of the Cre or Flp recombinase expressed under the control of a tissue-specific promoter/enhancer. Since the first published report of the use of Cre recombinase to generate a tissue-specific knockout, the generation of additional tissue-specific Cre-expressing mice has advanced rapidly, and now there are many lines available for this purpose (reviewed in Ref. 28).

YACs, BACs, PACs and chromosome engineering. Over the past 5 yr, significant new advances in the use of phage-based *E. coli* homologous recombination systems have been made, enabling genomic DNA in yeast, phage, and bacterial artificial chromosomes to be modified and subcloned without depending on restriction endonucleases or DNA ligases. This relatively new technique for genome and chromosome engineering is referred to as recombinogenic engineering or recombineering (36). As noted previously, targeted genetic mutation is a crucial part of the functional genomics in the mouse. However, until recently, the process of generating the targeting and selection constructs for homologous recombination in embryonic stem (ES) cells involved extensive, and often time-consuming, DNA cloning techniques that relied heavily on the presence of appropriate and unique restriction endonuclease cleavage sites. At times, this can be a formidable obstacle to constructing appropriate targeting vectors. In addition, recapitulating human disease-causing genetic mutations by deleting or rearranging megabase-sized regions of a chromosome is difficult to achieve with standard recombinant DNA techniques. Thus the recent development of yeast- and phage-based homologous recombination systems has greatly accelerated the pace of developing transgenic and knockout constructs and has made it possible to engineer large segments of genomic DNA, such as those carried on yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), or P1 artificial chromosomes (PACs).

Since at least 1993, it has been known that it is possible to mutate the yeast genome by deleting a yeast gene and replacing it with a selectable marker with the use of homologous recombination (4). This is possible because in *Saccharomyces cerevisae* there exists a very efficient DNA double-strand break and repair recombination pathway that recombines transformed, linear, double-stranded DNA with homologous sites in the yeast genome (50). Bradshaw and colleagues (7, 8) utilized this technique to generate, via homologous recombination, targeting vectors for mouse knockouts. They developed a shuttle vector named pClasper that is a plasmid that can be moved from yeast to bacteria. This enabled them to subclone DNA by gap repair in yeast, referred to as in vivo cloning, and then to transform pClasper into bacteria for production of sufficient quantities for use in generating transgenic mice (7, 8). More recently, systems have been developed for performing in vivo cloning in bacteria to modify DNA that is cloned into BAC and PAC vectors (reviewed in Ref. 11). Mutagenesis of BACs and PACs in *E. coli* enables an investigator to rapidly alter the structure of genes in their native contexts without the limitation of using restriction endonucleases and DNA ligases (Fig. 3). This allows the use of large segments of cloned DNA (200–300 kb). The systems that have been developed recently are based on the *E. coli* recA system (53) and on bacteriophage systems (27, 36). In fact, the technology now exists that enables one to mutagenize BAC DNA without the necessity for selection (49) and without leaving exogenous DNA at the mutated site (35, 37). These recombination methods greatly simplify the generation of transgenic and knockout constructs for use in gene function and regulation studies (1, 27). In addition, the in vivo identification of gene regulatory elements will be much easier with these recombineering techniques. This is particularly true for regulatory

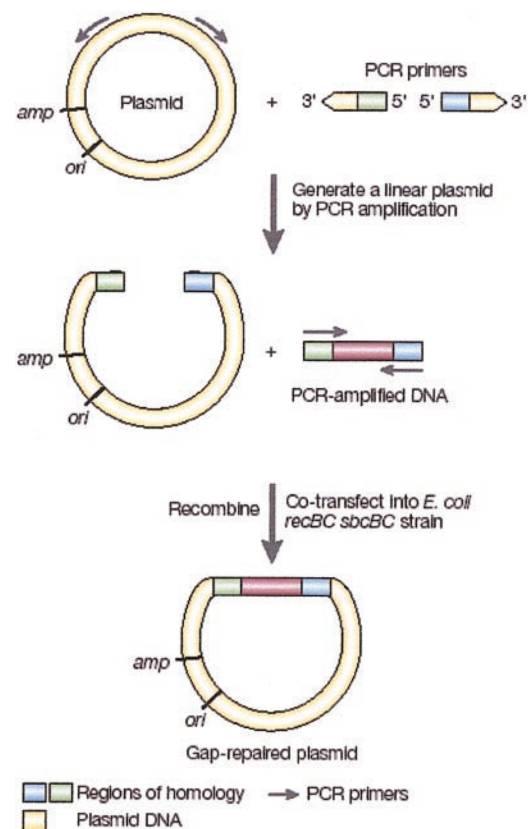

Fig. 3. In vivo cloning. This method of in vivo cloning uses two linear DNAs, a vector and a target DNA, which carry stretches of homology to each other at their ends. Both linear DNAs are electroporated into competent cells to allow homologous recombination between them, thereby repairing the plasmid DNA by closing the circle. *amp*, ampicillin resistance gene; *ori*, origin of replication. (Reproduced from Ref. 11 with permission from Nature Publishing Group.)

Downloaded from jap.physiology.org on June 5, 2012

elements that are far from the genes that they regulate and for genes that are regulated by multiple, different regulators or combinations of regulators.

Finally, the Cre-*loxP* system can be used for chromosome engineering to produce specific inversions or deletions to study the function of a specific region of a chromosome. This technique was first pioneered by the Bradley laboratory (40) and combines the power of gene targeting with Cre-*loxP* technology to generate mouse strains that harbor chromosome rearrangements (46). Generation of defined chromosomal rearrangements first involves selection of the endpoints for the region to be rearranged and modification of those endpoints in two separate steps of gene targeting in ES cells (Fig. 4). *LoxP* sites are placed so that they flank the region to undergo rearrangement. The type of chromosome rearrangement that is generated depends on the orientation of the *loxP* sites relative to each other. If the *loxP* sites are in the same orientation, then the region between them is deleted or duplicated (depending on whether the *loxP* sites are on the same or different chromosomes). If they are in the opposite orientation, then the region between them is inverted. The use of this technique has been extensively tested in the Bradley laboratory (29, 30, 40) with mouse chromosomes 11, 4, and 16. They have shown that, by using the positive selection strategy, Cre-mediated recombination is efficient enough to catalyze site-specific recombination between *loxP* sites that are separated by genomic intervals up to 60 cM in size. One additional feature that can be added to facilitate mapping, stock maintenance, and genetic screening is to "tag" these mutations with a dominant coat-color marker, as described by Zheng et al. (54). In this modification, a dominant coat-color marker, such as K14 agouti, is included in the 5′ and 3′ cassettes so that mice in which the rearrangement has occurred can easily be identified by coat color. There are several valuable uses for chromosome engineering. Two of the many uses are mutagenesis screens that use deletion and inversion strains to define the genetic basis of specific chromosomal syndromes in humans, such as DiGeorge syndrome. Large numbers of mouse strains with gene deletions or inversions generated by chromosome engineering can be used in mutagenesis screens to determine the function of genes within a given genomic interval (regional mutagenesis screen) (22). These screens can be used to correlate mutant phenotypic information with specific chromosomal rearrangements, which is extremely useful for defining gene function and for generating models of human disease. Many human birth defects have been attributed to both chromosome deletions and duplications. DiGeorge syndrome is the most common human chromosomal microdeletion syndrome and is associated with a heterozygous deletion on chromosome 21 [del(22)(q11)] (6). Lindsay and coworkers (29) recently generated a mouse model of DiGeorge syndrome by using chromosome engineering and showed that mice heterozygous for a 1.2-Mb deletion on mouse chromosome 16 have a cardiovascular phenotype similar to humans with DiGeorge syndrome; also, using the same techniques, they subdivided the 1.2-Mb region to identify the causative gene (30).

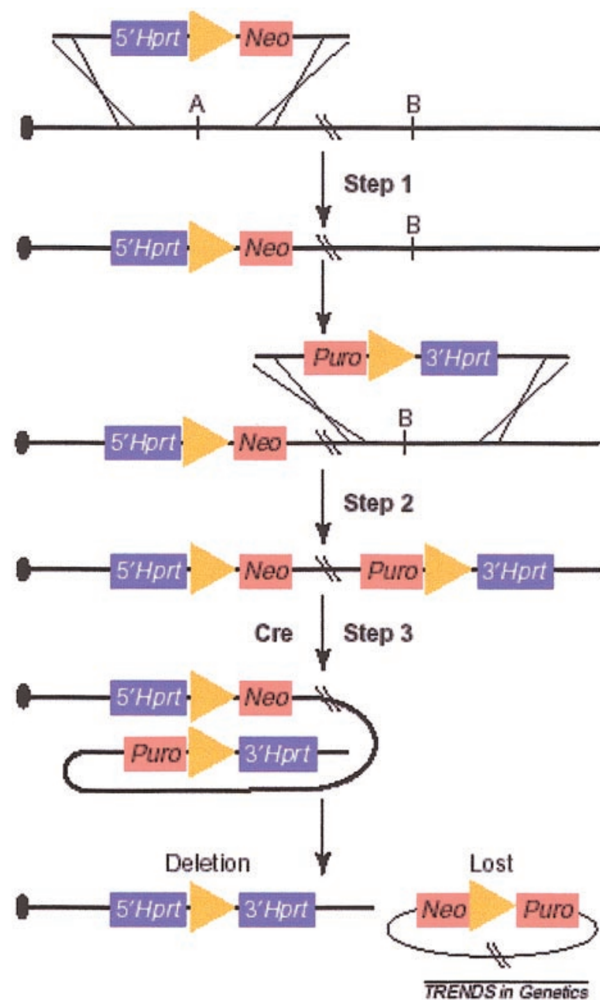

Fig. 4. Generation of defined chromosome rearrangements. The endpoints selected for the rearranged interval (*A* and *B*) are modified using two consecutive steps of gene targeting in the same embryonic stem (ES) cell. A positive selection scheme with an *Hprt* minigene is used to identify clones containing rearrangements induced by Cre. The *Hprt* minigene was divided into two nonfunctional halves to generate the 5′ and 3′ halves of the *Hprt* cassette (blue), both of which share an intron containing a *loxP* site (orange triangles). Two positive selectable markers, *PGKneobpA* and *PGKpurobpA* (red), use the PGK promoter to drive expression of the neomycin and puromycin resistance genes, respectively. *Step 1* results in the integration of a *loxP* site, the 5′ *Hprt* cassette, and the *PGKneobpA* cassette (*Neo*) at *endpoint A*. *Step 2* results in the integration of a second *loxP* site, the 3′ *Hprt* cassette, and *PGKpurobpA* (*Puro*) at *endpoint B*. In *step 3*, Cre is used to catalyze site-specific recombination between *loxP* sites. If the *loxP* sites are in the same orientation, as shown, the intervening region is deleted. However, if the *loxP* sites are in opposite orientation (not shown), a chromosome inversion occurs. (Reproduced from Ref. 33 with permission from Elsevier Science.)

Chemical mutagenesis. Another type of genetic screen used in mice is a whole genome chemical mutagenesis screen. Chemical mutagenesis is a potent approach to generate a large mutant mouse resource and is a screen that focuses on phenotypes rather than on a specific gene, like targeted mutagenesis, or a

Downloaded from jap.physiology.org on June 5, 2012

specific chromosome or region of a chromosome, like chromosome engineering. In fact, chemical mutagenesis has the potential to generate a large number of genetic mutations with a similar phenotype. Additionally, because chemical mutagenesis has the ability to induce single-base changes in DNA, subtle mutations can be introduced in a gene that may be more informative about the gene's function than a null mutation. The most commonly used mutagen is *N*-ethyl-*N*-nitrosourea (ENU), a compound that causes single-point mutations by ethylation. ENU predominately modifies A/T base pairs and results in missense mutations 64% of the time, nonsense mutations 10% of the time, and splicing errors 26% of the time (21, 24). Male mice treated with ENU undergo effective mutagenesis in the early spermatogonial cells. Thus, after an initial sterile period, the males will continue to generate mutated sperm for the remainder of their reproductive life. In contrast to other mutagenesis techniques, such as targeted mutations, ENU produces point mutations that are randomly distributed in the genome, thereby providing a greater opportunity for uncovering phenotypic diversity at a particular genetic locus. Despite these advantages, one significant challenge in performing ENU mutagenesis is detecting the underlying mutation. This requires extensive backcrossing to follow the phenotype in parallel with simple sequence length polymorphism markers.

There are two types of whole genome genetic screens that utilize ENU mutagenesis (reviewed in Ref. 21). One screens for viable and fertile mutants that represent allelic series, modifiers, or dominant mutations. This screen is the easiest because it involves treating males with ENU, mating them with wild-type females, and analyzing the phenotype of the offspring. The second type screens for recessive mutations, which requires a three-generation cross. There are a number of genome-wide mutagenesis screens underway, and because of the recent completion of the mouse genome, interest in performing these types of phenotype-based screens is increasing.

Identification of mammalian regulatory sequences. One of the important findings of the human genome project is that there is a significant amount of noncoding DNA. Although this noncoding DNA serves a number of different functions, a small fraction of it is felt to contain gene regulatory sequences. Until recently, despite their importance, these gene regulatory sequences have been quite difficult to identify. Classical techniques for identifying *cis*-regulatory sequences have involved a trial-and-error approach, including generation of deletion constructs for determining the minimal sequences needed for transcription in cell culture, DNase I hypersensitivity assays, DNA footprinting and gel shift assays, and the use of transgenic mice harboring reported gene constructs linked to various-sized genomic DNA fragments. However, these techniques have consisted largely of unguided searches of genomic sequence. However, with the completion of the human and mouse genomes and the genomes of other organisms that are well under way to completion, the use of computational methods provides new ways to screen the genome and accurately predict which sequences serve as gene regulators (reviewed in Ref. 39).

A relatively new approach to identifying mammalian gene regulatory sequences uses interspecies sequence comparison to identify highly conserved noncoding sequences, which are likely to be gene regulators. With the availability of sequence data from a number of different organisms, it has become apparent that noncoding sequences conserved across species often function experimentally as gene regulatory elements. Most often, species comparisons are performed between the mouse and human genomes, and studies based on these comparisons have been quite fruitful in identifying conserved DNA sequences as true regulatory sequences (15, 31, 47). These comparisons are made possible by the development of global alignment algorithms that align large genomic intervals and identify areas of conservation (3, 12). Two software programs developed for visualizing sequence alignment outputs are VISTA (32) and PIPmaker (43). These two programs, which are available on the Web, allow an investigator to analyze sequence data from two or more species and to visually identify noncoding regions that are conserved and that lie in the vicinity of genes of interest. One important issue to consider in using interspecies genomic comparisons is what species should be used in the comparison. Clearly, one factor that plays an important role is the availability of sequence from different species distant from human. In addition, from sequence comparisons performed already, it is becoming clear that different regions of the genome evolve independently from each other, some faster than others. For instance, available evidence suggests that the β-globin locus control region has evolved quickly (15, 31), whereas the T-cell receptor loci have evolved much more slowly (19, 25). In regions that have evolved quickly, comparisons of closely related mammals, such as human and mouse or whale, should allow easy identification of conserved noncoding sequence. However, for regions that have evolved much more slowly, it may be necessary to perform comparisons between more distantly related mammals or nonmammalian vertebrates (such as birds, reptiles, or fish) (48).

FUTURE CHALLENGES

Although these are exciting times in mouse genetics, there remain challenges for the future. One of the major challenges, which has already been alluded to, is closing what has been referred to the "phenotype gap" (9). This term refers to the large disparity between the number of genes to be mutated and the number of known mutant phenotypes. Clearly, large-scale mutatagenesis screens are a step in the right direction but will require improved methods for analyzing phenotypes and detecting physiological abnormalities in mice. High-throughput phenotype screens that utilize quantitative analytic methods are needed, which is a

Downloaded from jap.physiology.org on June 5, 2012

challenge in the mouse. Despite its superiority as a genetic model, the small size of the mouse is a challenge to performing quantitative physiological analyses. One area where some progress has been made is in the development of methods to detect abnormalities of the cardiovascular system. Miniaturization of analytic techniques that have been used in humans and larger animals have led to important advances in assaying cardiovascular traits in small animals such as the mouse (17, 18). Some of the techniques that have been adapted to mice include echocardiography, electrocardiography, telemetry, metabolic and hemodynamic exercise studies, and electrophysiological studies. Development of additional imaging technologies, such as mini-magnetic resonance and mini-computed tomography are needed.

One ongoing project that is of particular interest to physiologists is the Rat Genome Project. Early on, the mouse became the mammalian model of choice for geneticists, whereas the rat became the model of choice for physiologists. Rat strains have been selected and bred to have traits of biomedical interest, and, since the late 1800s, investigators have created more than 240 inbred rat strains that have a number of phenotypes, such as hypertension, immunologic defects, and cancer (20). An international effort at sequencing the rat genome is now underway and will be an extremely valuable resource of comparative genomics. Recently, a systems biology approach was used to study cardiovascular and renal phenotypes in the rat. Stoll et al. (44) analyzed 239 phenotypes and mapped 81 of those traits onto the genome, and aggregates of traits ("quantitative trait loci) were identified on four chromosomes (chromosomes 1, 2, 7, and 18). Interestingly, these investigators used a new analytical approach, which they term physiological profiling, to assess changes in the system biology of the cardiovascular system in response to allelic substitutions. These types of approaches hold great promise for closing the phenotype gap.

REFERENCES

1. **Angrand PO, Daigle N, van der Hoeven F, Scholer HR, and Stewart AF.** Simplified generation of targeting constructs using ET recombination. *Nucleic Acids Res* 27: e16, 1999.
2. **Baron U and Bujard H.** Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol* 327: 401–421, 2000.
3. **Batzoglou S, Pachter L, Mesirov JP, Berger B, and Lander ES.** Human and mouse gene structure: comparative analysis and application to exon prediction. *Genome Res* 10: 950–958, 2000.
4. **Baudin A, Ozier-Kalogeropoulos O, Denouel A, Lacroute F, and Cullin C.** A simple and efficient method for direct gene deletion in *Saccharomyces cerevisae*. *Nucleic Acids Res* 21: 3329–3330, 1993.
5. **Beck J, Lloyd S, Hafenparast M, Lennon-Pierce M, Eppig J, Festing M, and Fisher E.** Genealogies of mouse inbred strains. *Nat Genet* 24: 23–25, 2000.
6. **Botta A, Lindsay EA, Jurecic V, and Baldini A.** Comparative mapping of the DiGeorge syndrome region in mouse shows inconsistent gene order and differential degree of gene conservation. *Mamm Genome* 8: 890–895, 1997.
7. **Bradshaw MS, Bollekens JA, and Ruddle FH.** A new vector for recombination-based cloning of large DNA fragments from yeast artificial chromosomes. *Nucleic Acids Res* 23: 4850–4856, 1995.
8. **Bradshaw MS, Shashikant CS, Belting HG, Bollekens JA, and Ruddle FH.** A long-range regulatory element of Hoxc8 identified by using the pClasper vector. *Proc Natl Acad Sci USA* 93: 2426–2430, 1996.
9. **Brown SD and Peters J.** Combining mutagenesis and genomics in the mouse—closing the phenotype gap. *Trends Genet* 12: 433–435, 1996.
10. **Byrne GW and Ruddle FH.** Multiplex gene regulation: a two-tiered approach to transgene regulation in transgenic mice. *Proc Natl Acad Sci USA* 86: 5473–5477, 1989.
11. **Copeland NG, Jenkins NA, and Court DL.** Recombineering: a powerful new tool for mouse functional genomics. *Nat Rev Genet* 2: 769–779, 2001.
12. **Delcher AL, Kasif S, Fleischmann RD, Peterson J, White O, and Salzberg SL.** Alignment of whole genomes. *Nucleic Acids Res* 27: 2369–2376, 1999.
13. **Gossen M and Bujard H.** Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci USA* 89: 5547–5551, 1992.
14. **Gossen M, Freundlieb S, Bender G, Muller G, Hillen W, and Bujard H.** Transcriptional activation by tetracyclines in mammalian cells. *Science* 268: 1766–1769, 1995.
15. **Gottgens B, Barton LM, Gilbert JG, Bench AJ, Sanchez MJ, Bahn S, Mistry S, Grafham D, McMurray A, Vaudin M, Amaya E, Bentley DR, Green AR, and Sinclair AM.** Analysis of vertebrate SCL loci identifies conserved enhancers. *Nat Biotechnol* 18: 181–186, 2000.
16. **Haldane JBS, Sprunt AD, and Haldan NM.** Reduplication in mice. *J Genet* 5: 133–135, 1915.
17. **Hoit BD, Kiatchoosakun S, Restivo J, Kirkpatrick D, Olszens K, Shao H, Pao YH, and Nadeau JH.** Naturally occurring variation in cardiovascular traits among inbred mouse strains. *Genomics* 79: 679–685, 2002.
18. **Hoit BD and Nadeau JH.** Phenotype-driven genetic approaches in mice: high-throughput phenotyping for discovering new models of cardiovascular disease. *Trends Cardiovasc Med* 11: 82–89, 2001.
19. **Hood L, Koop BF, Rowen L, and Wang K.** Human and mouse T-cell-receptor loci: the importance of comparative large-scale DNA sequence analyses. *Cold Spring Harb Symp Quant Biol* 58: 339–348, 1993.
20. **Jacob HJ and Kwitek AE.** Rat genetics: attaching physiology and pharmacology to the genome. *Nat Rev Genet* 3: 33–42, 2002.
21. **Justice MJ, Carpenter DA, Favor J, Neuhauser-Klaus A, Hrabe de Angelis M, Soewarto D, Moser A, Cordes S, Miller D, Chapman V, Weber JS, Rinchik EM, Hunsicker PR, Russell WL, and Bode VC.** Effects of ENU dosage on mouse strains. *Mamm Genome* 11: 484–488, 2000.
22. **Justice MJ, Zheng B, Woychik RP, and Bradley A.** Using targeted large deletions and high-efficiency *N*-ethyl-*N*-nitrosourea mutagenesis for functional analyses of the mammalian genome. *Methods* 13: 423–436, 1997.
23. **Keeler CE.** *The Laboratory Mouse: Its Origin, Heredity and Culture*. Cambridge, MA: Harvard Univ. Press, 1931.
24. **Klysik J.** Mice and humans: chromosome engineering and its application to functional genomics. *Acta Biochim Pol* 49: 553–569, 2002.
25. **Koop BF and Hood L.** Striking sequence similarity over almost 100 kilobases of human and mouse T-cell receptor DNA. *Nat Genet* 7: 48–53, 1994.
26. **Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L,

Downloaded from jap.physiology.org on June 5, 2012

Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, et al. Initial sequencing and analysis of the human genome. *Nature* 409: 860–921, 2001.

27. **Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing DA, Court DL, Jenkins NA, and Copeland NG.** A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* 73: 56–65, 2001.
28. **Lewandoski M.** Conditional control of gene expression in the mouse. *Nat Rev Genet* 2: 743–755, 2001.
29. **Lindsay EA, Botta A, Jurecic V, Carattini-Rivera S, Cheah YC, Rosenblatt HM, Bradley A, and Baldini A.** Congenital heart disease in mice deficient for the DiGeorge syndrome region. *Nature* 401: 379–383, 1999.
30. **Lindsay EA, Vitelli F, Su H, Morishima M, Huynh T, Pramparo T, Jurecic V, Ogunrinu G, Sutherland HF, Scambler PJ, Bradley A, and Baldini A.** Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* 410: 97–101, 2001.
31. **Loots GG, Locksley RM, Blankespoor CM, Wang ZE, Miller W, Rubin EM, and Frazer KA.** Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science* 288: 136–140, 2000.
32. **Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, and Dubchak I.** VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* 16: 1046–1047, 2000.
33. **Mills AA and Bradley A.** From mouse to man: generating megabase chromosome rearrangements. *Trends Genet* 17: 331–339, 2001.
34. **Morse H.** *The Mouse in Biomedical Research*. New York: Academic, 1981.
35. **Muyrers JP, Zhang Y, Benes V, Testa G, Ansorge W, and Stewart AF.** Point mutation of bacterial artificial chromosomes by ET recombination. *EMBO J* 1: 239–243, 2000.
36. **Muyrers JP, Zhang Y, and Stewart AF.** Techniques: recombinogenic engineering—new options for cloning and manipulating DNA. *Trends Biochem Sci* 26: 325–331, 2001.
37. **Nefedov M, Williamson R, and Ioannou PA.** Insertion of disease-causing mutations in BACs by homologous recombination in *Escherichia coli*. *Nucleic Acids Res* 28: E79, 2000.
38. **O'Gorman S, Fox DT, and Wahl GM.** Recombinase-mediated gene activation and site-specific integration in mammalian cells. *Science* 251: 1351–1355, 1991.
39. **Pennacchio LA and Rubin EM.** Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet* 2: 100–109, 2001.
40. **Ramirez-Solis R, Liu P, and Bradley A.** Chromosome engineering in mice. *Nature* 378: 720–724, 1995.
41. **Roebroek AJ, Wu X, and Bram RJ.** Knockin approaches. *Methods Mol Biol* 209: 187–200, 2003.
42. **Sauer B and Henderson N.** Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome. *Nucleic Acids Res* 17: 147–161, 1989.
43. **Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, and Miller W.** PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res* 10: 577–586, 2000.
44. **Shastry BS.** Gene disruption in mice: models of development and disease. *Mol Cell Biochem* 181: 163–179, 1998.
45. **Silver L.** *Mouse Genetics.* New York: Oxford Univ. *Press*, 1995.
46. **Smith AJ, De Sousa MA, Kwabi-Addo B, Heppell-Parton A, Impey H, and Rabbitts P.** A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. *Nat Genet* 9: 376–385, 1995.

46a. **Stoll M, Cowley AW Jr, Tonellato PJ, Greene AS, Kaldunski ML, Roman RJ, Dumas P, Schork NJ, Wang Z, and Jacob HJ.** A genomic-systems biology map for cardiovascular function. *Science* 294: 1723–1726, 2001.

47. **Sumiyama K, Irvine SQ, Stock DW, Weiss KM, Kawasaki K, Shimizu N, Shashikant CS, Miller W, and Ruddle FH.** Genomic structure and functional control of the Dlx3–7 bigene cluster. *Proc Natl Acad Sci USA* 99: 780–785, 2002.
48. **Sumiyama K, Kim CB, and Ruddle FH.** An efficient *cis*-element discovery method using multiple sequence comparisons based on evolutionary relationships. *Genomics* 71: 260–262, 2001.
49. **Swaminathan S, Ellis HM, Waters LS, Yu D, Lee EC, Court DL, and Sharan SK.** Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis* 29: 14–21, 2001.
50. **Szostak JW, Orr-Weaver TL, Rothstein RJ, and Stahl FW.** The double-strand-break repair model for recombination. *Cell* 33: 25–35, 1983.
51. **Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, et al.** The sequence of the human genome. *Science* 291: 1304–1351, 2001.
52. **Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, et al.** Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520–562, 2002.
53. **Yang XW, Model P, and 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.
54. **Zheng B, Mills AA, and Bradley A.** A system for rapid generation of coat color-tagged knockouts and defined chromosomal rearrangements in mice. *Nucleic Acids Res* 27: 2354–2360, 1999.

Downloaded from jap.physiology.org on June 5, 2012