Immunity, Vol. 15, 409–418, September, 2001, Copyright ©2001 by Cell Press

Genome-Wide ENU Mutagenesis to Reveal Immune Regulators

Review

Keats A. Nelms and Christopher C. Goodnow[1]
Australian Cancer Research Foundation
Genetics Laboratory
Medical Genome Centre
John Curtin School of Medical Research
Canberra ACT 2601
Australia

A complete list of molecular components for immune system function is now available with the completion of the human and mouse genome sequences. However, identification and functional annotation of genes involved in immunological processes require a discovery methodology that can efficiently and broadly analyze the complex interplay of these components in vivo. Our recent experience indicates that genome-wide chemical mutagenesis in the mouse is an extremely powerful methodology for the identification of genes required for complex immunological processes.

Mining the Phenome: Opportunity and Challenge

The assembled sequences of the human and mouse genomes provide a remarkable opportunity in the form of a complete list of genes controlling the immune system. Solutions to long standing immunological questions such as the cause and treatment of autoimmunity, the prevention of allergy and transplant rejection, and the enhancement of suitable immune responses to infection, parasitism, and cancer lie in using this list to identify the molecular circuitry underpinning immune cell responses. For example, the limitations of the current gold standard immunosuppressive drugs cyclosporin and tacrolimus probably stem from their interference with tolerogenic as well as immunogenic signaling in lymphocytes (Glynne et al., 2000b). Revealing more effective targets for long-term graft acceptance or treatment of autoimmunity will depend upon resolving biochemical pathways that are exclusively immunogenic from pathways that serve both immunogenic and tolerogenic responses to antigen.

With the genome sequence in hand as a surface map, how do we mine the "phenome," the many layers of molecular and cellular behavior that extend deep beneath the sequence? How do we find and recognize key regulatory and effector molecules? Sequence homology and patterns of mRNA or protein expression provide valuable clues for where to start digging. Functional studies in cell culture, by overexpression, for example, provide an early indication of function, but one that is often incomplete or misleading. The in vivo phenotype of genetic variants in humans and animals provides the firmest definition of function. In this regard, the mouse is the Rosetta stone for immunology, by allowing the cellular and biochemical roles of individual gene products to be clearly revealed through allelic variants such as *H-2* congenic strains, spontaneous mutants such as *nude*, and the numerous transgenic and knockout strains that are a modern workhorse.

Established sources for genetic variants in the mouse nevertheless limit investigation to isolated patches of the phenome. A limited range of alleles exists among inbred and outbred mouse strains, since strongly deleterious alleles are rapidly selected against. Relatively few spontaneous mouse mutants exist for immunology, because it is hard to recognize an immunological mutant by looking in a cage during the course of normal inbreeding. Transgenic mice, especially where the transgene is targeted by homologous recombination, allow limitless possibilities in terms of the genes and types of alleles that can be investigated, but they require a great deal of background knowledge and investment in construction and analysis which must be performed one gene at a time. Large scale insertional knockout programs are accelerating the construction of mouse mutants (Zambrowicz et al., 1998), but because the approach still examines one gene at a time, these approaches also limit the field to analyzing a small fraction of the genome and superficial digging into the phenome. To illuminate immune regulators, it is often necessary to perform complex antigen challenges or crosses to susceptible strains or transgenic models, efforts that are prohibitive when scaling up a "one gene at a time" approach.

Genome-wide, random mutagenesis with the ethylating chemical N-ethyl-N-nitrosourea (ENU) provides a way to produce and screen genetic variants for immunological functions in the mouse in a highly parallel manner, thousands of genes at a time, allowing much wider and deeper mining of the phenome. Here, we summarize our recent experience and the current status of this strategy, and discuss its strengths and limitations.

Overall Concept

The overall aim of ENU mutagenesis is to produce a high rate of point mutations genome wide, at random, in the mouse germline (Figure 1). As a complement to gene knockout strategies, genome-wide mutagenesis in mice requires no foreknowledge about the target sequence or kind of mutation that should be induced. Introducing point mutations has the advantage that it most often inactivates or alters the function of individual protein domains rather than eliminating the protein altogether, making it harder for related proteins to compensate, better mimicking the effects of drugs or natural genetic variants, illuminating active sites in proteins, and revealing specific functions of alternatively spliced forms. Breeding of the mutagenized stocks can be performed such that libraries of mice are created where potentially every gene product is mutated to inactivate discrete domains in individual mice. The mutagenized mouse libraries can then be screened for key immunological functions, to identify genetic variants in many of the key regulatory components and provide animal models so as to understand their role.

The concept of genome-wide mutagenesis is well proven in fruit flies and nematode worms. It was used

[1] E-mail: chris.goodnow@anu.edu.au

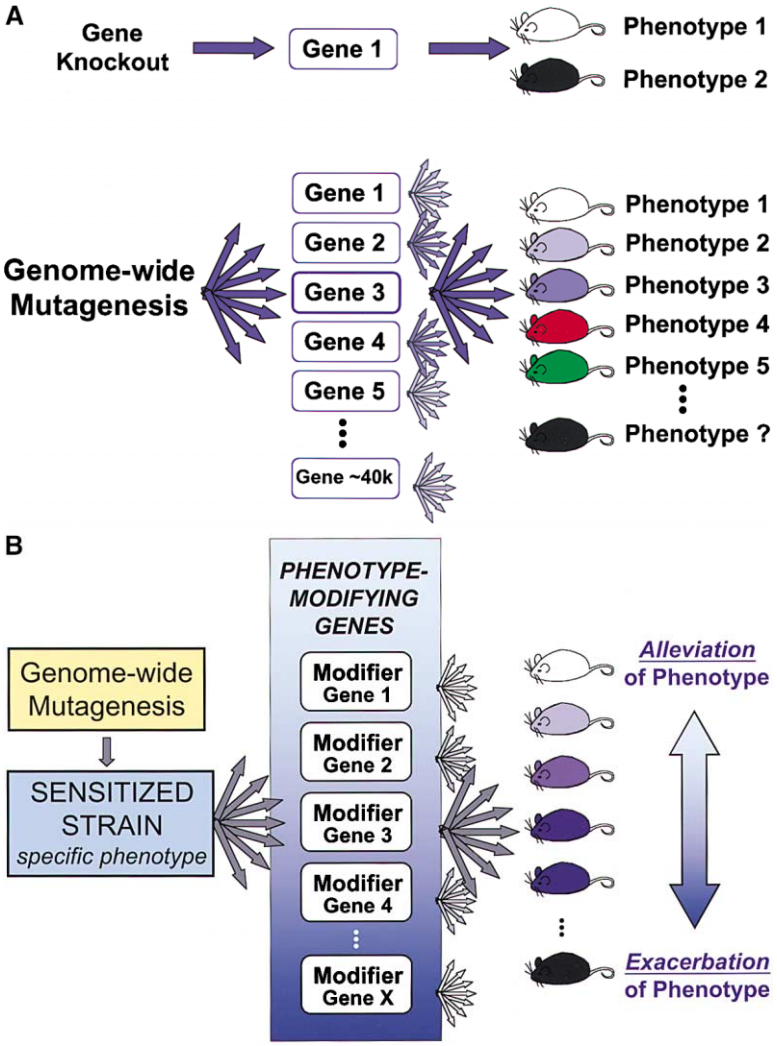

Figure 1. Genome-Wide Random Mutagenesis Is a Powerful Approach to Identifying Genes Involved in a Wide Range of Phenotypes in Mice

(A) Comparison of gene knockout analysis of gene function to genome-wide mutagenesis. Gene knockouts result in the inactivation of one gene at a time and a limited range of phenotypes. Genome-wide ENU mutagenesis results in the simultaneous, random mutation of many genes and can result in a wide range of phenotypes due to the varied effect of point mutations within a gene ranging from complete inactivation to subtle alteration of protein function.

(B) Performing ENU mutagenesis on a specific sensitized strain whose genetic background predisposes it to a specific phenotype can lead to the identification of a set of modifier genes that may not be detected when mutagenesis is performed on a wild-type genetic background. Random mutations in such modifier genes may lead to the alleviation or exacerbation of the phenotype in the sensitized strain.

by T.H. Morgan to create the first genetic maps of Mendelian inheritance in *Drosophila*, and revealed many of the key genes controlling early embryonic patterning and development (Nusslein-Volhard and Wischaus, 1980), to name but two examples. Tens of thousands of mutagenized gametes and hundreds of thousands of progeny need to be produced and screened to survey most of the genes in invertebrates. What makes it possible to carry out comparable efforts in mice is the extraordinary mutagenic efficiency of ENU in mouse spermatogonial stem cells. ENU has been shown to create a new loss-of-function allele in any given locus, typically once in every 700 gametes (Hitotsumachi et al., 1985). This means that it is possible to survey the majority of genes in the genome for immunological or other functions in a library of 14,000 mice derived from 700 gametes. Because each mouse in such a library is potentially homozygous for a different set of perhaps 12 loss-of-function mutations, the immunological function of tens of thousands of genes can be tested in a highly parallel manner rather than one gene at a time. Multiplex screening of genetic variants enables deeper exploration through intensive immunological tests and incorporation of special transgenic models (discussed below).

Identifying ENU Mouse Mutants in One Generation

The supermutagenic effect of ENU for mouse spermatogonial stem cells was discovered by Russell and colleagues at Oak Ridge Laboratories in Tennessee, using an assay called the "specific locus test" that was developed in the post-World War II period to measure the heritable mutagenic effects of radiation (Russell et al., 1979). This one-generation breeding assay is performed by treating normal male mice (generation 0; G0 mice) with ENU by intraperitoneal injection, creating a high burden of point mutations in the spermatogonial stem cells. After waiting 7 weeks for mature sperm to be formed from the mutagenized spermatogonia, the G0 males are mated with untreated females from a tester strain. The tester strain females are homozygous for seven recessive visible mutations affecting pigment and ear size, and the treated G0 males start with wild-type alleles at each of these loci. Consequently, most progeny (generation 1; G1) appear wild-type because they inherit a wild-type copy of each gene from the G0 father and a defective copy from the mother. Each time a G1 mouse is conceived from a sperm carrying a new loss-of-function mutation in one of the seven specific loci, the new mutation is immediately visible by a difference

in color or ear size. By testing thousands of such G1 progeny, Russell and colleagues showed that a new loss-of-function mutation in a specific locus was produced on average once in every 700 gametes using an optimized dose and treatment regime (Hitotsumachi et al., 1985). Adaptations of this one-generation breeding strategy to other gene loci have established the generality of this mutation rate, although some loci appear particularly easy to mutate to loss of function while others are more difficult (Bode, 1984; Cordes and Barsh, 1994; Lewis et al., 1991; Shedlovsky et al., 1993). The predominant mutation is in an A:T base pair, either as a transition or transversion; 64% of the mutations are missense, 10% nonsense, and 26% cause errors in mRNA splicing (Justice et al., 1999).

Similar one-generation breeding strategies have been used to discover new ENU-induced mutations at previously unknown loci by breeding ENU-treated males with wild-type females rather than a tester strain. Since only the paternal genome is mutated in this case, only dominant mutations can be detected in the resulting G1 animals. Dominant mutations are of three chief types: gain-of-function, dominant interfering, and haploinsufficiency. Rare gain-of-function mutants can be extremely informative, as in the Lurcher mutant strain where a point mutation in the δ2 glutamate receptor results in apoptosis of Purkinje cells and the characteristic gait of these mice (Zuo et al., 1997). Many dominant interfering mutations causing cataracts have been induced by ENU using a one-generation breeding strategy (Favor et al., 1991). Most of these appear to reflect misfolded proteins that interfere with normal lens structure (Klopp et al., 1998). Misfolded proteins can be toxic to cells and cause phenotypes that have little relevance to the normal function of the protein. A small fraction of genes are overtly haploinsufficient, such that loss of function in a single copy reduces activity so drastically that a role for the gene can be readily detected in the heterozygous state. A number of the homeobox-containing developmental regulators exhibit this characteristic, and this may partly reflect monoallelic expression in individual cells. In the immune system, monoallelic expression of the Pax5 transcription factor occurs in early B cell progenitors and in mature B cells (Nutt et al., 1999). When mutations occur in one Pax5 allele, the monoallelic expression pattern leads to haploinsufficiency and results in the deletion of B cells expressing the mutant allele.

ENU-induced dominant visible mutations have been produced in PAX6 (small eye, Neuherberg), the brachyury gene (T^{kt1}), and the c-kit tyrosine kinase (white spotting). In the latter two examples, the dominant defects are only detected at the very end of the neural tube or at the extremities of melanocyte migration, where the dosage of the normal gene product appears to be at its most limiting. Thus, the ability to detect the loss of one copy of these genes may primarily reflect fortuitous use of a highly sensitive assay. A sensitive assay for diurnal rhythm involving activity measurements repeated over 30 days was used to identify a single dominant mutation, Clock, among over a thousand G1 animals tested (Vitaterna et al., 1994). In reality, many gene products are limiting in signaling pathways in development (Simon et al., 1991) and in the immune system (Cornall et al., 1998; Cyster et al., 1996; Rathmell et al., 1996), such that the loss of one copy can register as a clear phenotype, provided one assays in a sufficiently sensitive way. Strategies to tease these out with sensitized/enhancer screens are discussed further below.

Loss-of-function mutations in tumor suppressor genes represent a variation on haploinsufficiency. Inheritance of a single wild-type copy of a tumor suppressor gene is often insufficient to prevent cells from losing the remaining copy and becoming neoplastic. A dominant ENU mutation causing multiple intestinal adenomas, Min, was originally detected in a one-generation screen because of anemia secondary to bleeding colonic cancer polyps (Moser et al., 1990). This strain carries a loss-of-function mutation in the mouse homolog of the human tumor suppressor gene *APC*, and has been a valuable preclinical model for developing drugs against human colon cancer. In our own program, we have also identified two dominant mutations that represent loss of function in tumor suppressor genes (P. Papathanasiou, A. Loy, K.A.N., and C.C.G., unpublished data).

The one-generation breeding strategy has been used on a large scale recently in two extensive mouse mutagenesis programs conducted at Harwell in the United Kingdom (Nolan et al., 2000) and in Munich, Germany (Hrabe de Angelis et al., 2000). These programs screened tens of thousands of G1 animals for behavioral and blood defects, yielding over 300 confirmed mutant mouse strains. Of these new mutant strains, 39 were identified to have mutations affecting the immune system with alterations in basal immunoglobulin levels (including IgE) or altered expression of surface proteins on peripheral blood lymphocytes (Hrabe de Angelis et al., 2000).

Three-Generation Strategies to Detect Recessive Mutations Genome Wide

The majority of functional gene defects induced by ENU will be loss-of-function variants that behave as recessive traits, and will only be detectable after breeding to homozygosity or to an existing mutant. A shortcut to reveal new recessive mutations in two generations is by breeding individual G1 males with females carrying a chromosomal deletion (Rinchik et al., 1990; Shedlovsky et al., 1988). Only the small fraction of the genome encompassed by the deletion can be screened for recessive mutations in this way. To reveal recessive mutations on a genome-wide basis requires a three-generation pedigree structure (Figure 2). To screen a large fraction of genes, hundreds of such pedigrees must be created in parallel.

A three-generation pedigree structure which we have used successfully to reveal many new immunological mutants is shown schematically in Figure 2. In each pedigree, all of the mutagenized chromosomes that segregate among the offspring come from a single founder G1 male, who can be estimated to carry ~100 loss-of-function mutations. Each G1 male is mated to a wild-type female, and then to his G2 daughters. The resulting G3 offspring have a 1/8 chance of being homozygous for any individual mutation carried by the founding G1 male. If four G3 offspring from four G2 mothers are screened in each pedigree (16 G3 mice in total), it can be shown that at least one G3 animal would be expected

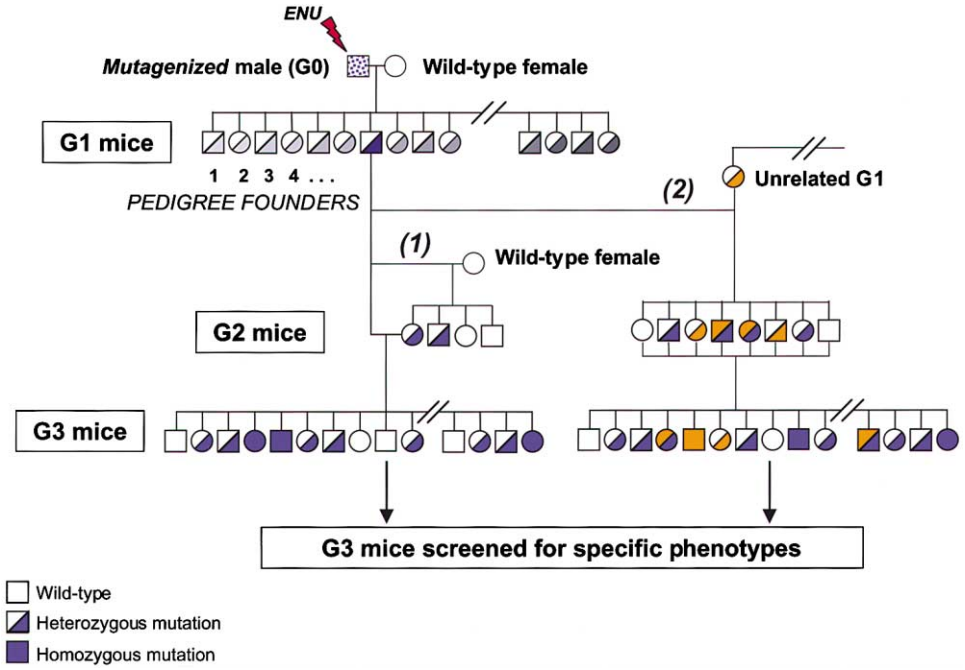

Figure 2. Breeding of ENU-Treated Mice to Reveal Recessive Mutations on Wild-Type Genetic Background

G1 mice carrying approximately 100 independent loss-of-function mutations become the founders of independent pedigrees by outcrossing G1 males to wild-type females (1) or intercrossing unrelated mutagenized G1 mice (2). Recessive mutations become homozygous in G3 mice, and these mice are screened for specific immunological phenotypes.

to be homozygous for 78 of the estimated 100 founding mutations in the pedigree. Assuming the G1 male carries 100 loss-of-function mutations, the G3 mice will be homozygous for a mode of 12 different loss-of-function mutations, providing a highly parallel method to survey gene function genome wide.

An alternative three-generation breeding strategy that we now prefer initiates each pedigree with a G1 male paired to an unrelated G1 female (Figure 2). Each of these founding G1 grandparents contributes ~100 loss-of-function mutations, and the total pool of ~200 mutations are brought to homozygosity by intercrossing their G2 progeny as brother-sister pairs. If four G3 progeny are tested from each of four G2 pairs (16 G3 mice in total), it can be shown that 100 of the 200 founding mutations will be represented by at least one homozygous animal. This two-founder strategy has four advantages over the daughter-father strategy: (1) it surveys a larger number of mutations per G3 animal; (2) it screens mutations on the X chromosome; (3) it is logistically more straightforward because no rotation of breeders is needed; and (4) the breeders are younger, and a higher proportion of unaffected G3 siblings are carriers to help propagate a recessive mutation of interest.

While three-generation strategies conceptually provide a highly parallel method to survey most of the genes in the genome for immunological functions, doubts about their feasibility have centered on four key issues. First, the large number of homozygous loss-of-function mutations in each G3 animal may prevent animals from surviving to maturity because of the burden of lethal mutations. Mature animals are essential to screen for immunological disorders, cancer, and many other processes relevant to human health. Second, the burden of mutations could mean that most phenotypic abnormalities observed represent complex interactions between multiple mutations that might prove difficult to resolve or fail to breed true. Third, it is logistically difficult to track hundreds of three-generation pedigrees and to screen the animals for nonvisible traits such as immunological abnormalities. Fourth, the application of ENU mutagenesis to immunologically well-characterized inbred strains such as C57BL/6 to facilitate detection and analysis of immune system mutants has not been extensively examined.

Genome-wide three-generation screens for recessive mutations affecting embryonic and neonatal mice have been successfully performed using the strategy above by several groups, attesting to the general feasibility of this approach. Bode and colleagues (Bode et al., 1988) mutagenized (C57BL/6xCBA) F1 animals to produce approximately 25 G3 progeny in each of 105 daughter-father pedigrees. The G3 mice were screened to discover mouse models of the human metabolic disorder phenylketonuria at 1 week of age while the G3 pups were still caged with their G2 mothers. Screening at this age minimizes the number of cages and tracking logistics enormously. Blood levels of phenylalanine were tested using the standard Guthrie test performed on human neonates, and one pedigree with several G3 animals affected by hyperphenylalaninemia was identified. Propagation and mapping of the mutation led to its identification in phenylalanine hydroxylase, providing a valuable model for the corresponding human disease. Harding et al. (1992) produced 135 pedigrees of ENU-mutagenized BTBR strain mice containing 10–15 G3 off-

spring in each, and screened weaned mice for amino acid metabolism defects by placing them on a high protein diet for 3 days and analyzing urine spots from the whole cage by ultraviolet illumination. They identified one recessive mutant with sarcosinemia, providing an animal model for this human genetic disorder. Anderson and colleagues (Kasarskis et al., 1998) mutagenized C57BL/6 mice and established 130 three-generation daughter-father pedigrees. The daughters were sacrificed at 10.5 days of gestation to screen for recessive mutations disrupting neural tube closure and morphology in the embryos. Four recessive mutant strains were identified that mapped to single genes on chromosomes 1, 6, 12, and 16.

To establish the feasibility of this strategy for revealing genes regulating the immune system and cancer in mature inbred mice, we mutagenized C57BL/6 mice and established 185 daughter-father three-generation pedigrees (K.A.N., A. Loy, P. Papathanasiou, L. Miosge, L. Wilson, S. Chaudhry, and C.C.G., unpublished data). At least 25 G3 offspring in each pedigree were aged to 6–9 months in order to detect mature onset disorders. Blood from the mice was screened for B and T cell subsets by flow cytometry, yielding over 20 different strains with recessive disorders ranging from absence of T cells or T cell subsets and deficiencies of T and/or B cells to hyperactivity of T cells and hypergammaglobulinemia. Four strains with spontaneous cancer predisposition were established, one with a dominant T cell leukemia, one with a range of tumors (primarily B or T lymphomas or osteosarcoma), one with primitive hemopoietic cell tumors, and one with hepatocellular carcinoma. In parallel, a large number of strains were established with recessive defects affecting many other body systems, notably the nervous system, obesity, skin and hair, skin mast cells, mucosal goblet cells, skeleton, and kidney. Almost all of the mutations breed true as simple Mendelian traits, and over half of the mutations appear to be novel, in that the chromosomal location and the phenotype do not correspond to any known mutant or knockout mouse strain.

Several key points are underscored by our experience with genome-wide screening for recessive immunological defects. Foremost is the sensitivity and specificity of different thresholds and parameters for recognizing putative mutants. Because rare mutant animals of interest must be readily detected against a background of hundreds of phenotypically normal individuals, it is important to use quantitative screening parameters that do not have too great a coefficient of variation in the parental strain. Multiparameter flow cytometry is ideally suited to this task. Because of the need to minimize other sources of variability, it is far preferable to mutagenize and screen a well-characterized inbred strain like C57BL/6 and to ensure pathogen-free status. Ideally, a threshold should be set for a screened trait such that less than 1% of normal animals would exceed the threshold. Specificity of the screen is greatly improved by repeat testing, hence a blood screen that can be repeated on apparently abnormal individuals is far better than a one-off assay such as a terminal test. Similarly, the accuracy of detecting rare events is enhanced significantly by redundant screens where a specific parameter is assayed by independent detection reagents or methods (Townsend et al., 2001). The presence of multiple affected siblings in a pedigree provides another repeat measure that greatly increases specificity to differentiate true mutants from chance animals, and this is a major advantage of recessive traits. Likewise, the ability to propagate from healthy sibling carriers is an additional advantage of recessive traits, compared to dominant traits which are very difficult to propagate if animals are infertile or if identifying carrier individuals precludes breeding.

Mapping and Identifying ENU-Induced Mutations

Two parallel paths of biological and genetic analyses can be pursued to pinpoint the cellular and molecular defect in each new immunological mutant generated (Figure 3A). Because each new strain is isogenic on the C57BL/6 strain background from the outset, this facilitates detailed immunohistochemistry, flow cytometry, cell transfer, mixing experiments, and constructing bone marrow radiation chimeras reconstituted with a mixture of mutant C57BL/6 and wild-type C57BL/6-Ly5a congenic bone marrow. For most of the immunological mutants we have isolated, these approaches have resolved the cell lineage and differentiation stage expressing the primary defect and often suggested the biochemical process that is disrupted. Once the primary cell type and differentiation stage is known, a molecular fingerprint of changes in gene expression caused by the mutation can be obtained by profiling mRNA on DNA microarrays. We have previously established the ability of microarrays to identify selective defects in lymphocyte signaling pathways and link these to downstream targets (Glynne et al., 2000a), and emphasized the need to perform at least two replicates (Glynne et al., 2000c) and the importance of starting with highly purified FACS-sorted cells that are as closely matched between test and control groups (Glynne et al., 2000b). Amplification of mRNA by in vitro transcription enables clear mRNA profiles to be readily measured from one or two million sorted resting lymphocytes (Glynne et al., 2000b).

In parallel to obtaining a precise cellular and molecular phenotype for each mutant, a generic mapping approach enables rapid localization of each mutation to a particular chromosomal region. In this approach, the mutation is mapped by cosegregation of the mutation with genetic markers interspersed throughout the genome. The initial step of this approach requires that mutants on the C57BL/6 background be outcrossed with a second inbred mapping strain resulting in F1 progeny that acquire one set of chromosomes from the C57BL/6 strain and the other set from the mapping strain (Figure 3A). Following the initial outcross, recessive mutations are mapped most efficiently by intercrossing F1 carrier mice in brother-sister pairings that yield F2 progeny of which 25% are homozygous for the recessive mutation (Figure 3A). The mapping efficiency of such F1 intercrosses is high, because every meiosis is informative (i.e., crossing over can occur between every pair of F1 chromosomes) and the fertility and fecundity of F1 mice is very high. An alternative to the intercross is the backcross strategy that is required when the mapped mutation is dominant or when a recessive mutation is suppressed by the genetic background of the mapping

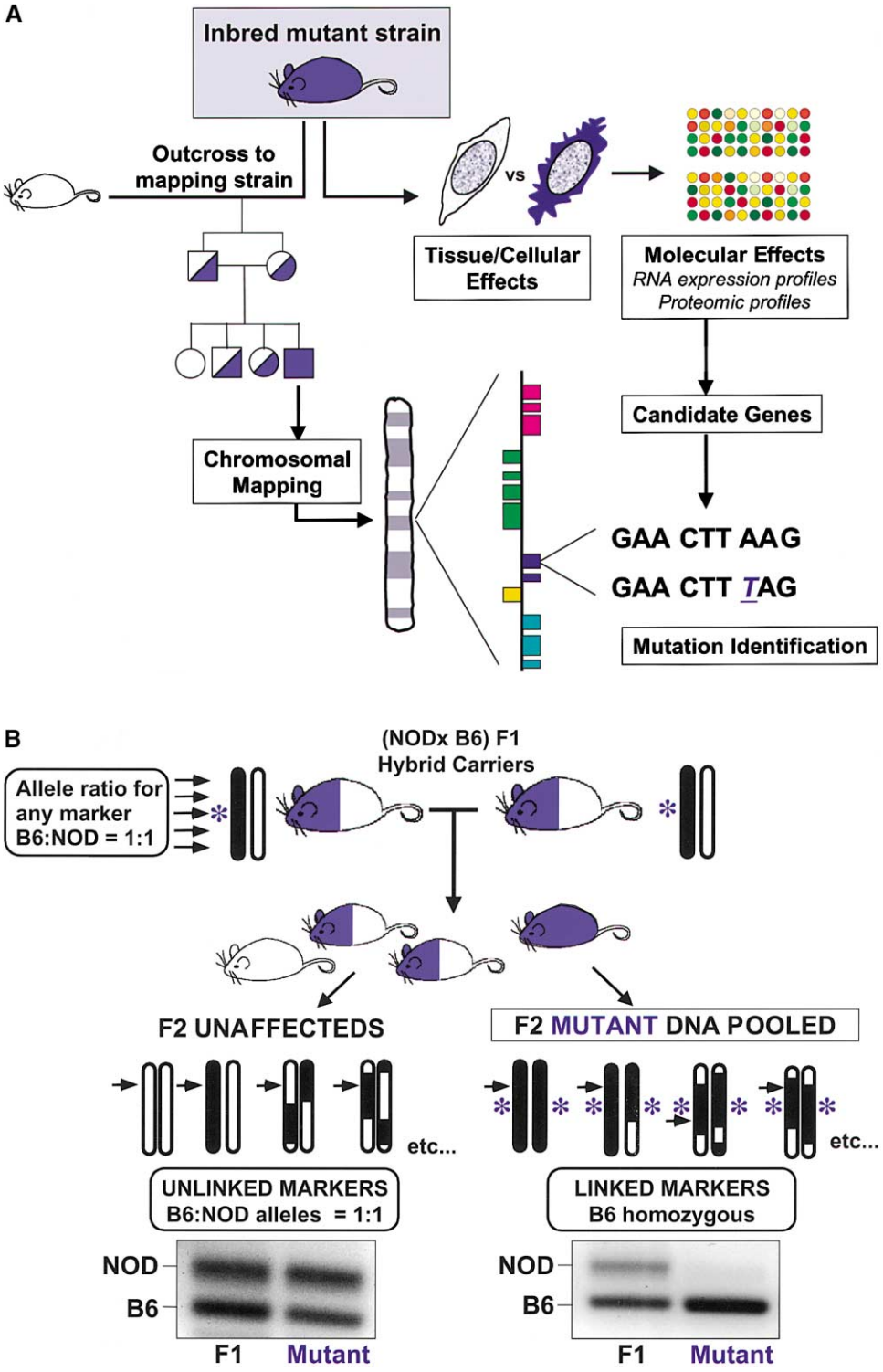

Figure 3. Gene Identification in Genome-Wide Mutagenesis

(A) Parallel biological and genetic analyses in novel strains of mutant mice lead to the identification of candidate genes whose mutation results in the phenotypes observed in the mice.

(B) Intercross breeding strategy for the mapping of recessive mutations by genetic linkage analysis. Mutant mice (B6) are outcrossed to a mapping strain (NOD) to produce F1 hybrid carrier mice that are then intercrossed to produce F2 progeny, 25% of which will be homozygous for the recessive mutation. DNA from F2 mutant mice is pooled and analyzed by microsatellite markers. B6 and NOD alleles are distinguished by the size of PCR fragments in gel electrophoresis. Markers closely linked to the mutation will yield only the B6 allelic fragment in F2 mutant mice. By contrast, unlinked markers will yield both B6 and NOD alleles in F2 mutant mice. DNA from F1 mice will give B6 and NOD alleles for all markers.

strain. In either the intercross or backcross strategy, two generations and 4–6 months are required, depending on the age that the mutant phenotype can be detected.

Meiotic recombination in F1 mice acts as a genetic shuffle that can be used to distinguish regions of the genome that are closely linked to the mutation from regions that are unlinked. These regions are determined by linkage analysis in the last step of the mapping process by analyzing the recombinant chromosomes inherited by F2 mice with a panel of genetic markers that span the genome (Figure 3B). In the case of recessive intercrosses, this is done by analysis of genomic DNA from F2 mice that exhibit the mutant phenotype and are thus homozygous for the mutation. Analysis of only mutant F2 mice ensures that the mutated gene and the genomic regions closely linked to it are homozygous for the C57BL/6 allele. In contrast, genomic regions that are unlinked and therefore inherited randomly will exhibit the 1:1 ratio of C57BL/6 and mapping strain alleles also observed in the F1 parents. Thus, analyses of pooled genomic DNA from mutant F2 mice with closely linked genetic markers will only yield the C57BL/6 allele, whereas analyses with unlinked markers will yield an equal mix of C57BL/6 and mapping strain alleles (Figure 3B). Using the pool of ~25 mutant mice derived from a total of 100 F2 intercross mice, linkage can be detected at a 95% confidence level between the mutation and genetic markers spaced at 28 cM intervals across the genome (Silver, 1995). This coverage is accomplished very efficiently with a panel of 60 independent genetic markers run on one F1 DNA control and one pooled DNA sample per marker.

Once a preliminary chromosomal map location is determined, the genetic location of the mutation is more accurately defined by fine mapping analyses. Such analyses require that individual mutant mice be analyzed with genetic markers in a defined chromosomal region. Ultimately, the location of each mutation is delimited by specific genetic markers that flank the site of the mutation and detect recombinant chromosomes in individual mutant F2 mice. Additional markers very close to the mutation will only very rarely have points of recombination between the marker and the mutation. This approach of pooling and performing linkage analysis on ~25 F2 intercross mutant mice is predicted to lead to a 1–2 cM map position for a mutation. Indeed, we have achieved this accuracy with many mutants. However, in some cases, the limited availability of usable genetic markers between C57BL/6 and our current mapping strain, NOD, have limited the mapped interval to 5–15 cM. Currently, simple sequence length polymorphism (SSLP) or "microsatellite" markers are the primary tool used for such genome-wide mapping and fine mapping analyses. At present, such markers can be restricted in their genomic density and utility between specific mouse strains. Thus, obtaining an accurate map position to less than 1–2 cM can be marker limited. With the compilation of genomic sequence information from multiple strains of mice, it is anticipated that a very high-density microsatellite and single-nucleotide polymorphism (SNP) marker panel will be readily assembled. Such a panel of high-density genetic markers would greatly enhance the accuracy and efficiency of mutation mapping procedures.

Obtaining an accurate map location for the mutated gene is the key step prior to initiating studies to establish the identity of the mutated gene. On average, 60–100 genes must be considered as candidates for the mutated gene when mutations are mapped to an interval of 1–2 cM corresponding to approximately 2–4 Mb of the genome. The ensuing process of identifying the mutated gene is less defined in comparison to the genetic mapping of the mutation, and can be greatly influenced by the results of biological analyses obtained on the mutant strain. This is particularly true when aspects of the mutant phenotype have been observed in previously described knockout or spontaneous mutant mice. In such cases, resequencing of one candidate gene often reveals the novel mutation. In other cases, no knockout or mutant strains exist that suggest a candidate gene in the mapped interval. In our experience, a majority of ENU-induced recessive mutations fall into this category and the identification of the mutated gene in these novel mutants requires a multifaceted approach.

After obtaining a chromosomal map location, addressing several key questions can aid in the identification of uncharacterized genes containing ENU-induced mutations. The first and most important question is: what are the genes that exist in the mapped interval? Genomic sequence databases and the annotation of these databases are the critical tools used to address this question. If the genomic sequence data is complete, sequencing and expression analysis strategies may be readily implemented to identify the mutated gene. A second key question addressed by the biological analysis of each mutant is whether a specific cell type or tissue is intrinsically affected in the mutant mice. In the case of immunological mutants, this may often be addressed using bone marrow chimeras as discussed above. When a specific cell or tissue is affected, identifying candidates is greatly facilitated by delineating the genes in the interval that are expressed specifically in the affected cells or tissue. Such data can be rapidly obtained using expression array assembled with gene sequences from the mapped interval. This expression data may also be used to determine whether the induced mutation occurs in a *cis*-acting element or RNA stability sequence that is required for expression in the affected cell or tissue. In contrast, mutations in splice sites are readily detected even from limited numbers of target cells using RT-PCR and resequencing of candidate genes while coding sequence mutations are detected by cDNA or exon resequencing.

Although identification of the genetic lesion in novel mutant mice is a potential "bottleneck" in the gene discovery process, this bottleneck can be readily addressed with existing technology. In particular, high-throughput resequencing of all the cDNAs or exons in the mapped interval is the most straightforward option and could potentially be used to identify all mutations in a relatively large mapped interval. Such efforts can be streamlined significantly by further delimiting the mapped interval with fine mapping. However, from a practical standpoint, large scale resequencing efforts may be cost prohibitive. Thus, emerging technologies focused on the rapid and economical identification of point mutations may play a key role in mutation detection in the near future.

Mining Deeper into the Phenome

A key issue that remains to be fully explored is how deeply we will be able to mine the phenome using Mendelian mutations in individual genes in mice, whether these be induced by targeted transgenesis, insertion trap, or with ENU. Only a subset of genes and mutations has a sufficiently extreme phenotype that they will be readily recognized by a broad, superficial survey of physiological systems. For example, the mouse small eye mutation was readily detected by its externally visible effect on eye development, but it was only decades later that striking abnormalities in pancreatic islet cell development were fortuitously revealed by specific immunohistochemical staining (Sander et al., 1997; St-Onge et al., 1997). The majority of illuminating mutations will likely be buried for one of several reasons outlined below, requiring special efforts to unearth them. The ability to screen hundreds of mutations at a time by ENU mutagenesis, and to do this on a transgenically sensitized background, makes it feasible to apply these extra efforts on a genome-wide scale.

Surrogate Phenotypes

Mutations in genes critical to specific immunological processes will frequently not be immediately apparent by a spontaneous clinical disorder or a straightforward immunological screen, and may in fact affect rare cell types or processes that are not obvious without specific histological or biochemical analysis. The example of subclinical defects in pancreatic islet cell development in small eye mutant mice was mentioned above (Sander et al., 1997; St-Onge et al., 1997). Identification of such phenotypes will require specialized knowledge and highly focused assays such as high-throughput immunohistological tests. Genes that alter such surrogate traits are much more likely to be revealed if the histological screen can examine mutations in hundreds of genes at a time rather than one gene at a time. Similarly, mutations that alter antigen-specific selection and responses will often be hidden by the variability and diversity of clones in a polyclonal repertoire of T and B cells. Many such defects will be made visible by performing ENU mutagenesis on the background of an antigen-specific TCR or Ig receptor transgenic strain (Figure 1B).

Phenotypes that Require a Challenge Test

An important class of immunological genes and mutations that require in depth screening procedures are those affecting the dynamics and type of response to immunological challenge. Included in this class would be gene mutations that suppress inflammatory, allergic, or autoimmune responses. A key issue in the identification of such mutations is the normal variability of immune responses even to highly purified antigens in an inbred strain. Even a false positive rate of 5% can lead to an unacceptable level of misdirected effort. Thus, identifying mutants that affect the response to immunological challenge on the background of an ENU-mutagenized strain will require highly reproducible immunization protocols. Again, such variability problems may be addressed by screening for ENU-induced mutations in mice that are genetically hardwired with a TCR or Ig transgene or a mutation, to give a more stereotypic response. The logistical issue of mouse fertility must be solved if the genetically challenged mice cannot breed.

Phenotypes that Require Time and/or Multiple Steps

Genes contributing to spontaneous disorders such as lymphoid cancer or autoimmune diabetes may normally be hidden by the need for multiple genetic abnormalities to be present before the defect becomes penetrant as a clinical phenotype, and an inordinate amount of time. One solution to this problem is to screen for surrogate phenotypes, as discussed above, but in many cases these will not exist. Another strategy is to test for these genes on a strain background where some of the steps have been taken to start with; for example, a p53 knockout strain that is already one step toward cancer, a TCR transgenic strain that already has a high level of subclinical autoimmune insulitis, or a drug treatment that mimics one of the steps.

Lethal Mutations

Perhaps 10% of mammalian genes serve essential functions during fetal life, such that homozygous mutant mice will not survive to an age that an important immunological function could be revealed. Targeted construction of conditional knockout alleles provides a powerful strategy to pursue immunological functions of essential genes on a small scale. One solution to this problem that has been used in *Drosophila* is to combine random ENU mutagenesis with an engineered chromosome that allows mutations to be made homozygous in subsets of somatic cells by recombinase-mediated exchange between chromosome homologs (Golic and Lindquist, 1989; Theodosiou and Xu, 1998). Given the spectacular advances in chromosome engineering that have been achieved (reviewed in Mills and Bradley, 2001), it is conceivable that the same strategy could be applied to screen for immunological genes chromosome by chromosome in mice.

Sensitized enhancer/suppressor screens provide another solution to revealing later roles for genes that serve essential embryonic functions, borrowing another leaf from the *Drosophila* geneticists' textbook. An elegant example of this strategy was used to reveal the intermediate genes between the sevenless receptor tyrosine kinase and its nuclear targets in *Drosophila* eye development (Simon et al., 1991). The intermediate genes include essential embryonic proteins such as Grb-2, Sos, and Ras, which were never revealed in recessive screens genome wide because of lethality. By partially crippling the receptor kinase, correct eye development became sensitive to the loss of a single copy of the intermediate genes in heterozygous mutant animals (allowing the mutant genes to be found), while the remaining wild-type copy of these genes remained sufficient to support embryonic development. Evidence that similar strategies could be employed in mice comes from the ability to reveal heterozygous loss-of-function deficits in numerous lymphocyte signaling molecules when placed onto a homogeneous lymphocyte repertoire of an Ig or TCR transgenic mouse (Cornall et al., 1998; Cyster et al., 1996; Rathmell et al., 1995). The key issue for the success of these approaches is the challenge of detecting a relatively subtle change in one mouse against a background of normal variability. The specificity of the enhancer of the sevenless screen benefits from measuring hundreds of replicate developmental events in the many ommatidia of each individual fly. The detection of heterozygous lymphocyte defects depends similarly on

comparing the distribution of surface Ig densities on many lymphocytes in individual mice.

Future Directions

The rapid expansion and increased accessibility of genomic technologies and information has resulted in the successful application of genome-wide ENU mutagenesis in the mouse for the discovery of genes in a wide variety of immunological, physiological, and developmental processes. These successes will undoubtedly lead to the increased use and inevitable refinements of the current methodology. Complete assembly and full annotation of the human and mouse genomic sequence databases as well as the availability of sequences from different mouse strains will likely have the greatest impact on gene discovery in the near term. In addition, application of the sequencing technologies and capabilities resulting from genome sequencing efforts will have a great impact on the rate of mutation identification in novel mutant mouse strains. The advent of new technologies such as high-throughput SNP mapping will also greatly expedite gene discovery. Ultimately, the fusion of all of these technologies will facilitate the understanding of many of the most complex questions currently facing immunologists.

References

Bode, V.C. (1984). Ethylnitrosourea mutagenesis and the isolation of mutant alleles for specific genes located in the T region of mouse chromosome 17. Genetics *108*, 457–470.

Bode, V.C., McDonald, J.D., Guenet, J.L., and Simon, D. (1988). hph-1: a mouse mutant with hereditary hyperphenylalaninemia induced by ethylnitrosourea mutagenesis. Genetics *118*, 299–305.

Cordes, S.P., and Barsh, G.S. (1994). The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor. Cell *79*, 1025–1034.

Cornall, R.J., Cyster, J.G., Hibbs, M.L., Dunn, A.R., Otipoby, K.L., Clark, E.A., and Goodnow, C.C. (1998). Polygenic autoimmune traits: Lyn, CD22, and SHP-1 are limiting elements of a biochemical pathway regulating BCR signaling and selection. Immunity *8*, 497–508.

Cyster, J.G., Healy, J.I., Kishihara, K., Mak, T.W., Thomas, M.L., and Goodnow, C.C. (1996). Regulation of B-lymphocyte negative and positive selection by tyrosine phosphatase CD45. Nature *381*, 325–328.

Favor, J., Neuhauser-Klaus, A., and Ehling, U.H. (1991). The induction of forward and reverse specific-locus mutations and dominant cataract mutations in spermatogonia of treated strain DBA/2 mice by ethylnitrosourea. Mutat. Res. *249*, 293–300.

Glynne, R., Akkaraju, S., Healy, J.I., Rayner, J., Goodnow, C.C., and Mack, D.H. (2000a). How self-tolerance and the immunosuppressive drug FK506 prevent B-cell mitogenesis. Nature *403*, 672–676.

Glynne, R., Ghandour, G., Rayner, J., Mack, D.H., and Goodnow, C.C. (2000b). B-lymphocyte quiescence, tolerance and activation as viewed by global gene expression profiling on microarrays. Immunol. Rev. *176*, 216–246.

Glynne, R.J., Ghandour, G., and Goodnow, C.C. (2000c). Genomic-scale gene expression analysis of lymphocyte growth, tolerance and malignancy. Curr. Opin. Immunol. *12*, 210–214.

Golic, K.G., and Lindquist, S. (1989). The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. Cell *59*, 499–509.

Harding, C.O., Williams, P., Pflanzer, D.M., Colwell, R.E., Lyne, P.W., and Wolff, J.A. (1992). sar: a genetic mouse model for human sarcosinemia generated by ethylnitrosourea mutagenesis. Proc. Natl. Acad. Sci. USA *89*, 2644–2648.

Hitotsumachi, S., Carpenter, D.A., and Russell, W.L. (1985). Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. Proc. Natl. Acad. Sci. USA *82*, 6619–6621.

Hrabe de Angelis, M.H., Flaswinkel, H., Fuchs, H., Rathkolb, B., Soewarto, D., Marschall, S., Heffner, S., Pargent, W., Wuensch, K., Jung, M., et al. (2000). Genome-wide, large-scale production of mutant mice by ENU mutagenesis. Nat. Genet. *25*, 444–447.

Justice, M.J., Noveroske, J.K., Weber, J.S., Zheng, B., and Bradley, A. (1999). Mouse ENU mutagenesis. Hum. Mol. Genet. *8*, 1955–1963.

Kasarskis, A., Manova, K., and Anderson, K.V. (1998). A phenotype-based screen for embryonic lethal mutations in the mouse. Proc. Natl. Acad. Sci. USA *95*, 7485–7490.

Klopp, N., Favor, J., Loster, J., Lutz, R.B., Neuhauser-Klaus, A., Prescott, A., Pretsch, W., Quinlan, R.A., Sandilands, A., Vrensen, G.F., and Graw, J. (1998). Three murine cataract mutants (Cat2) are defective in different γ-crystallin genes. Genomics *52*, 152–158.

Lewis, S.E., Barnett, L.B., Sadler, B.M., and Shelby, M.D. (1991). ENU mutagenesis in the mouse electrophoretic specific-locus test, 1. Dose-response relationship of electrophoretically-detected mutations arising from mouse spermatogonia treated with ethylnitrosourea. Mutat. Res. *249*, 311–315.

Mills, A.A., and Bradley, A. (2001). From mouse to man: generating megabase chromosome rearrangements. Trends Genet. *17*, 331–339.

Moser, A.R., Pitot, H.C., and Dove, W.F. (1990). A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. Science *247*, 322–324.

Nolan, P.M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I.C., Vizor, L., Brooker, D., Whitehill, E., et al. (2000). A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. Nat. Genet. *25*, 440–443.

Nusslein-Volhard, C., and Wischaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. Nature *287*, 795–801.

Nutt, S.L., Vambrie, S., Steinlein, P., Kozmik, Z., Rolink, A., Weith, A., and Busslinger, M. (1999). Independent regulation of the two Pax5 alleles during B-cell development. Nat. Genet. *21*, 390–395.

Rathmell, J.C., Cooke, M.P., Ho, W.Y., Grein, J., Townsend, S.E., Davis, M.M., and Goodnow, C.C. (1995). CD95 (Fas)-dependent elimination of self-reactive B cells upon interaction with $CD4^+$ T cells. Nature *376*, 181–184.

Rathmell, J.C., Townsend, S.E., Xu, J.C., Flavell, R.A., and Goodnow, C.C. (1996). Expansion or elimination of B cells in vivo: dual roles for CD40- and Fas (CD95)-ligands modulated by the B cell antigen receptor. Cell *87*, 319–329.

Rinchik, E.M., Carpenter, D.A., and Selby, P.B. (1990). A strategy for fine-structure functional analysis of a 6- to 11-centimorgan region of mouse chromosome 7 by high-efficiency mutagenesis. Proc. Natl. Acad. Sci. USA *87*, 896–900.

Russell, W.L., Kelly, E.M., Hunsicker, P.R., Bangham, J.W., Maddux, S.C., and Phipps, E.L. (1979). Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. Proc. Natl. Acad. Sci. USA *76*, 5818–5819.

Sander, M., Neubuser, A., Kalamaras, J., Ee, H.C., Martin, G.R., and German, M.S. (1997). Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. Genes Dev. *11*, 1662–1673.

Shedlovsky, A., King, T.R., and Dove, W.F. (1988). Saturation germ line mutagenesis of the murine t region including a lethal allele at the quaking locus. Proc. Natl. Acad. Sci. USA *85*, 180–184.

Shedlovsky, A., McDonald, J.D., Symula, D., and Dove, W.F. (1993). Mouse models of human phenylketonuria. Genetics *134*, 1205–1210.

Silver, L.M. (1995). Mouse Genetics: Concepts and Applications (New York: Oxford University Press).

Simon, M.A., Bowtell, D.D., Dodson, G.S., Laverty, T.R., and Rubin, G.M. (1991). Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase. Cell *67*, 701–716.

St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A., and

Gruss, P. (1997). Pax6 is required for differentiation of glucagon-producing α-cells in mouse pancreas. Nature *387*, 406–409.

Theodosiou, N.A., and Xu, T. (1998). Use of FLP/FRT system to study *Drosophila* development. Methods *14*, 355–365.

Townsend, S.E., Goodnow, C.C., and Cornall, R.J. (2001). Single epitope multiple staining to detect ultralow frequency B cells. J. Immunol. Methods *249*, 137–146.

Vitaterna, M.H., King, D.P., Chang, A.M., Kornhauser, J.M., Lowrey, P.L., McDonald, J.D., Dove, W.F., Pinto, L.H., Turek, F.W., and Takahashi, J.S. (1994). Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. Science *264*, 719–725.

Zambrowicz, B.P., Friedrich, G.A., Buxton, E.C., Lilleberg, S.L., Person, C., and Sands, A.T. (1998). Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature *392*, 608–611.

Zuo, J., De Jager, P.L., Takahashi, K.A., Jiang, W., Linden, D.J., and Heintz, N. (1997). Neurodegeneration in Lurcher mice caused by mutation in δ2 glutamate receptor gene. Nature *388*, 769–773.