TIMELINE

Trends in large-scale mouse mutagenesis: from genetics to functional genomics

Yoichi Gondo

Abstract | The primary goal of mouse mutagenesis programmes is to develop a fundamental research infrastructure for mammalian functional genomics and to produce human disease models. Many large-scale programmes have been ongoing since 1997; these culminated in the International Knockout Mouse Consortium (IKMC) in 2007 with the aim to establish knockout and conditional mouse strains for all mouse genes. This article traces the origins and rationale of these large-scale mouse mutagenesis programmes.

Genetics directly associates changes in DNA (that is, mutations) to biological function at the organism level (that is, phenotypes). In nature, however, mutants arise at extremely low frequency. To make genetic approaches to functional genomics feasible and practical, effective mutagenesis technologies are indispensable, accompanied by comprehensive methods to detect and score phenotypes. By harnessing a high-throughput, genome-wide mutagenesis system on a large scale, mouse genetics has become a fundamental means of carrying out functional genomics research. Indeed, the mouse is the organism with the oldest history in mutagenesis. The fact that X-rays could induce mutations was first shown in the mouse in 1923 (REF. 1), before the report of a similar phenomenon in fruitflies by Muller in 1927 (REF. 2). Later, in the 1980s, Russell and his colleagues[3–5] conducted extensive chemical mutagenesis studies in mice that included the use of the alkylating agent N-ethyl-N-nitrosourea (ENU).

The idea of starting a large-scale mouse mutagenesis project was raised in the 1990s. The human genome sequencing project was underway and was expected to reveal many candidate genes for human diseases and their putative biological functions. By the time this was achieved, technologies for generating transgenic[6] as well as gene-targeted or knockout mice[7] had become available (TIMELINE). Thus, experimental manipulation of human-homologous genes in the mouse was deemed indispensable for characterizing such candidates. Large-scale mouse mutagenesis projects began in 1997 (REFS 8,9), starting with ENU mutagenesis, and quickly spread over the world. The first phase of mouse mutagenesis — consisting of forward genetics or phenotype-driven mutagenesis[10] — corresponded to a classical mutagenesis approach: a mutant mouse is first recognized by its phenotype after random mutagenesis, and the causative mutation is subsequently mapped by positional cloning.

In 2006, international mouse knockout projects were initiated[11] for two main reasons. The first was to supplement the mutagenesis programme with reverse genetics (that is, gene-driven) mutagenesis. In this approach, a target gene is first disrupted by site-specific homologous recombination in embryonic stem (ES) cells, followed by the functional analysis of the mutation at the organismal level. The second reason was that the number of genes deduced by the completion of the human genome projects[12,13] was smaller than expected, which made it plausible to construct knockout mice for every gene in the genome.

This Timeline article describes the progress that has been made in large-scale mutagenesis in the mouse model, focusing principally on the past decade. I describe the key innovations that have made this progress possible, the benefits that the current generation of screens can offer and the challenges that they face. This article also highlights the importance of an interdisciplinary integration of expertise to further advance functional genomics.

Advantages of mouse as a model system

The mouse is chosen not only because it is closely related to humans but also because it has more than 100 years of history in genetic analysis[14]. Over this period many mutants were identified, a number of inbred lines were established and gene mapping had been conducted more extensively than in any other mammalian species — that is, until the completion of the Human Genome Sequencing Project made the human genome the most extensively mapped mammalian species. The well described syntenic relationships between human and mouse genomes imply that whenever candidate genes for human diseases are found, their homologues usually exist and are quickly identified in the mouse genome.

In addition, the mouse is currently the only species for which embryos can be manipulated using available ES-cell technologies. By merging these technologies with genetic engineering, the transgenic mouse[6] and then the knockout mouse[7] were developed in the 1980s. Chromosomal engineering technologies were also developed[15], which allowed balancer chromosomes (chromosomes with recessive lethal mutations and inverted segments that suppress recombination) to be constructed[16].

Technologies for freezing embryos[17] and gametes[18] are well established in the mouse, allowing *in vitro* fertilization to be combined with embryo transfer methods. Thus, valuable mouse lines can be easily and stably maintained in liquid nitrogen for many years while requiring minimal space and manpower.

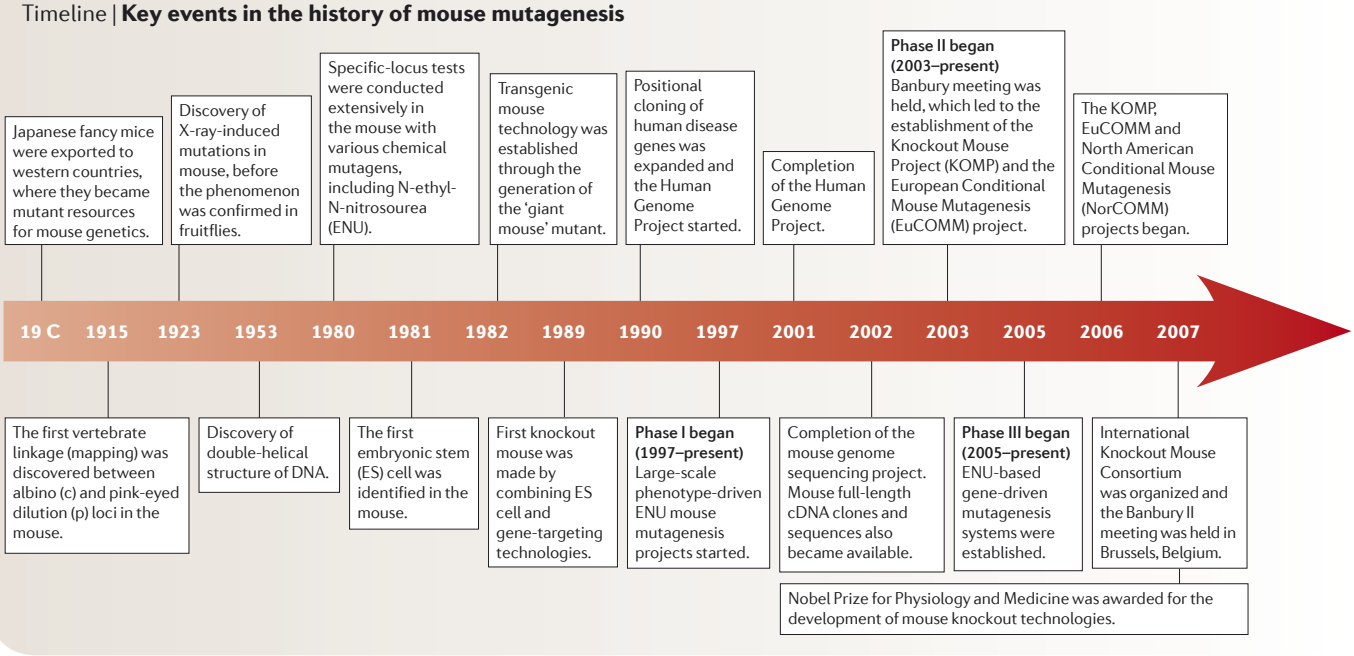
The elucidation of the human genome sequence aided the genotyping and positional cloning of mouse genes, for instance, by comparative genomics. To facilitate the identification of human homologues in the mouse genome, mouse geneticists proposed the immediate establishment of the mouse genome sequencing project. It was finished in 2002 (REF. 19), soon after the human genome sequencing project[12,13].

As shown in the TIMELINE, large-scale mouse mutagenesis is divided into three phases. In 1997, ENU mutagenesis began with the use of forward genetics, primarily focusing on dominant phenotypes. In 2003, a knockout mouse mutagenesis programme was proposed, principally to establish recessive mutant lines by reverse genetics. Now, a new reverse genetics that uses ENU as well as additional genetic tools has also become available.

Mouse mutagenesis phase I

ENU mouse mutagenesis projects. In 1997, two large-scale ENU mouse mutagenesis programmes were launched, at the Helmholtz Zentrum München German

Timeline | **Key events in the history of mouse mutagenesis**



Research Center for Environment and Health (GmbH, formally the GSF) in Germany[8] and at the Medical Research Council (MRC) in the UK[9]. Two more projects followed at RIKEN in Japan and at the Australian National University (ANU) in Australia in 1999. In 2000, many more ENU mouse mutagenesis projects were funded. Approximately twenty such projects (see the further information box) have been organized, each of which set itself the primary goal to produce and analyse the phenotypes of~10,000 G1 mice (see below and FIG. 1a).

All the projects used ENU, which was chosen because it is the most potent chemical mutagen[3–5] and because it induces single base-pair substitutions[20]. ENU is therefore more likely to induce mutations with partial function rather than complete null alleles. Candidate mutant mice are screened by a series of phenotyping tests, and the inheritance of each identified mutant phenotype is confirmed through genetic crosses. The crucial gene with the base-pair substitution responsible for the mutant phenotype is then pinpointed by positional cloning or candidate gene approaches.

Before the mouse mutagenesis projects began, there were some concerns about whether the efficiency of ENU mutagenesis was great enough to produce useful mutant lines efficiently and cost-effectively. As summarized in FIG. 1, at first it was practical to carry out genome-wide dominant-mutation screens on a large scale; however, the question remained of how many dominant mutations would be induced and what proportion of such mutations would be useful. Two independent pioneering ENU-mutagenesis studies that identified causative dominant mutations for tumorigenesis and circadian rhythm by Dove *et al.*[21,22] and Takahashi *et al.*[23,24], respectively, provided the rationale and incentive to initiate large-scale ENU mouse mutagenesis projects. In particular, the ApcMin mice generated by the Dove group were a good model for human familial adenomatous polyposis (FAP)[22]. The identification of the *Clock* gene by the Takahashi group[24] also proved that the ENU-based phenotype-driven approach allows the identification of biologically significant yet unknown genes in the mammalian genome.

Phenotype-driven approach. The most crucial factor for ENU mutagenesis is a reliable phenotyping platform, which must ensure sensitivity and reproducibility. In most of the projects, first generation (G1) mice are subjected to a battery of basic phenotypic assays, which include a series of quick morphological and behavioural tests (often referred to as SHIRPA[25]), and assays to ascertain haematological and biochemical values. Each project supplemented these with its own unique phenotyping assays. For instance, in-depth analyses of morphological and/or developmental anomalies were added by GSF, tests for sensory deficits or behavioural anomalies by the MRC, assays for common diseases/late-onset phenotypes by RIKEN, and assessments for immunological aberrations by ANU.

The mutagenesis flow is depicted in FIG. 1a. The choice of generation 0 (G0) parental strains was occasionally different among the projects. Each project selected the strain(s) that empirically was best suited to identifying a particular phenotype(s). One advantage of using G1 mice with an F1 hybrid background (an F1 hybrid is the offspring of two different inbred lines) is that the G2 progenies obtained from the backcrosses between a mutant-candidate G1 mouse and mouse of the maternal inbred strain immediately give rise to the material for genetic mapping (FIG. 1a). Genetic variations between the G0 parental strains, however, segregate in G2 and in later generations, potentially complicating phenotypic assessments.

Dominant phenotype screens. As depicted in FIG. 1a, in ENU dominant-mutation screens the mutagen is administered to G0 males; each mutagenized sperm transmits many ENU-induced mutations to the G1 progeny. To obtain as many candidate mutants as possible, the maximum number of G1 mice are produced and the primary phenotype screen is conducted in the G1 population as thoroughly as possible. The detection rate of mutants depends upon the number, type and accuracy of the phenotypic parameters, the observation period, and also on the environmental

and genetic background. In both the pioneering projects at GSF and MRC, approximately 2–3% of the G1 progeny exhibited some dominant phenotypes during basic screening within 3 months of birth. RIKEN extended the screening period to 18 months of age, and as a result the overall mutant detection rate in the G1 mice was slightly higher (about 3–4%). Even early in the mutagenesis projects quite a few dominant mutants were reported, including those affecting eye or whole-body morphology[26,27], immunological function[28], erythrocyte enzymatic activity in blood cells[29], and metabolism[30].

Recessive phenotype screens. ENU induces roughly 1,000-fold more recessive mutations than dominant ones. Only 2–4% of G1 mice carry dominant mutations; however, 20–40 recessive mutations per G1 mouse are expected to be induced by ENU, based on the specific-locus test[3–5] and on the estimated number of mouse genes[19]. The specific-locus test is used to easily detect newly arisen recessive mutations in specific loci by looking for mutations in genes that lead to recessive visible phenotypes, such as coat colour, in homozygotes. The early success of the dominant phenotype screens encouraged many projects to carry out recessive phenotypic screens as well, because many human genetic diseases are recessive. To assess recessive phenotypes, it is necessary to have a G3 generation (FIG. 1b). In this scheme, it is difficult to handle a large number of G1 strains owing to space limitations in the mouse facility; thus, some elaborate schemes have also been developed. One is a chromosomal region-specific screen, in which recessive hemizygous mutations are recovered as a result of loss of heterozygosity against a deletion-carrying chromosome[31,32] or against an inversion-carrying balancer chromosome that contains a dominant visible marker[16]. However, the usefulness of recessive mutants encouraged most of the ENU mouse projects to conduct the laborious but straightforward genome-wide recessive screens by randomly generating G3 mice as illustrated in FIG. 1b, at least as a feasibility study.

In the course of carrying out phenotype-driven mutagenesis, an ENU-based gene-driven mutagenesis system has also been developed specifically for recessive screening. This paradoxical approach will be discussed later in this article as a complementary system to the knockout mouse mutagenesis project.

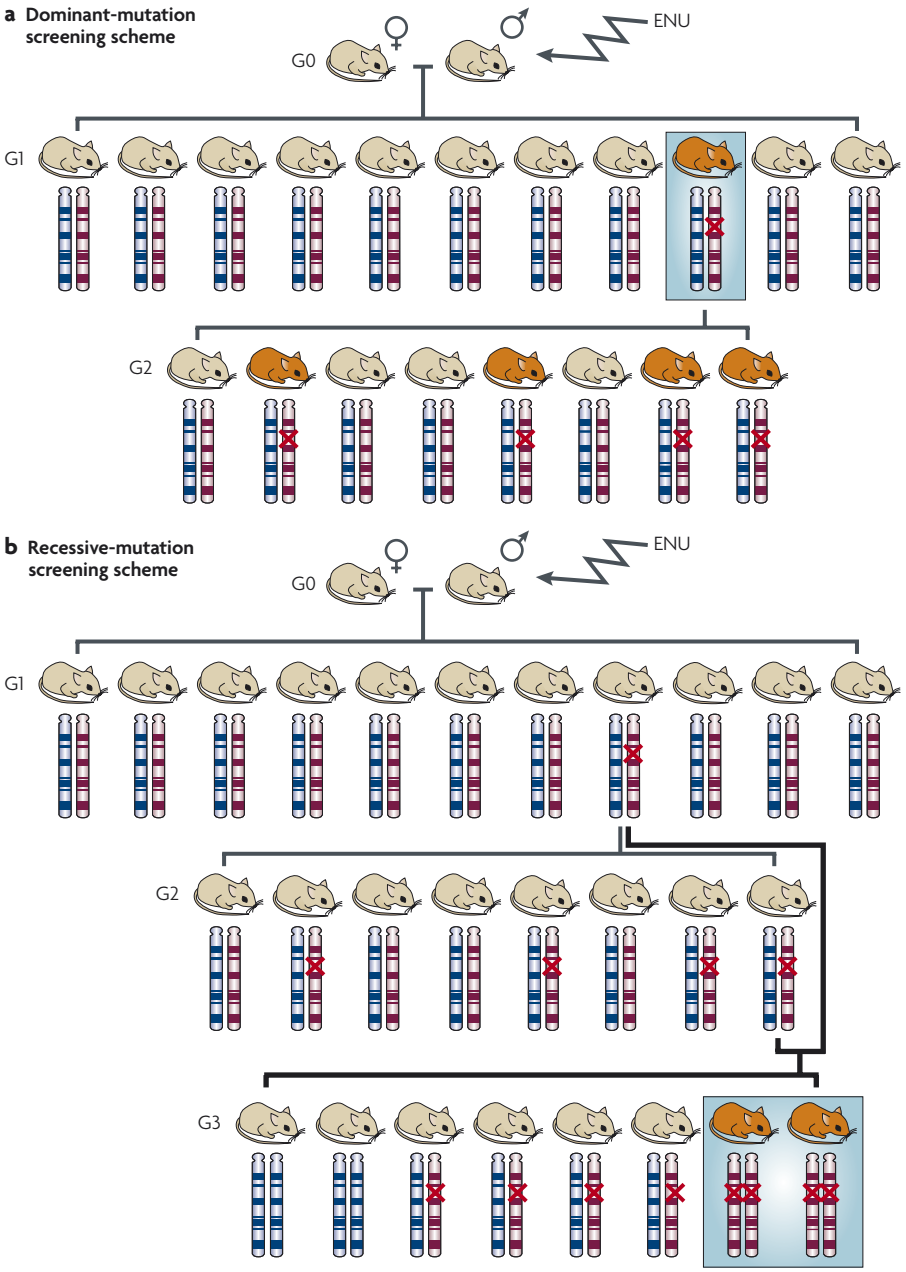


Figure 1 | **Mating scheme of ENU-mutagenesis protocols in the mouse.** Genome-wide dominant and recessive protocols for phenotype-driven mutagenesis screens. In both screening strategies, N-ethyl-N-nitrosourea (ENU) is administered to generation 0 (G0) males; the crucial step is to detect a mutant phenotype in the mice, indicated by blue rectangles. **a** | In the dominant-mutation screening scheme, G1 mice are subjected to a series of phenotypic tests. Any G1 animal that has a phenotypic anomaly is a mutant candidate, and is mated to produce G2 mice. When the candidate G1 has a heritable mutation, it is expected to segregate in a 1:1 ratio in the G2 progeny. Mutant animals are indicated by an orange coat and a red cross on the chromosome. When the G0 parental strains are different to each other, a G1 candidate can be mated to the G0 maternal strain to produce G2 backcross progeny. In this case, segregation of the mutant phenotype and of genetic markers, such as SNPs, gives direct genetic linkage information for mapping the identified mutation. **b** | In recessive-mutation screening, mutant phenotypes are only recognized in the G3 population. Each G1 mouse is expected to carry ~30 recessive mutations (Supplementary information S1 (box)); in this mating scheme, therefore, half of the 30 mutations in any G1 animal are transmitted to the G2. These mutations then become homozygous in a Mendelian 3:1 ratio in G3 mice; alternatively, any G3 mouse might carry ~4 recessive homozygous mutations (Supplementary information 1 (box)). In the recessive-mutation screening scheme the parents usually come from the same strain: if different parental G0 strains were used, the recessive traits in the parental strains would also co-segregate in G3, making the G3 phenotype data complicated to interpret.

Mouse mutagenesis phase II

Knockout mouse mutagenesis. Gene targeting, also called knockout mouse technology, made it possible to conduct site-directed mutagenesis in a mammalian genome for the first time. Before 1990, gene targeting required a high degree of expertise and so was practiced by only a handful of laboratories. Knockout mice had been contributing to the functional analysis of genes on a one-by-one basis since the early 1990s; however, generating knockout mice on a large-scale was thought to be impossible because of the required throughput and consequent high costs.

The crucial factor that drove knockout projects to be performed on a genome-wide scale was the fall in the estimated number of genes in the mammalian genome. After the completion of the human genome sequencing project in 2002 (REFS 12,13) the predicted number of human genes dropped suddenly from between 100,000 and 150,000 (REF. 33) to less than 30,000. The reduction in the technical hurdles associated with knockout-mouse construction and the sudden decrease in the total gene number in the human[12,13] and mouse genomes[19] synergistically promoted the genome-wide knockout mouse project, particularly with the aim of generating recessive human disease models.

Lessons from ENU mutagenesis. The ENU mouse mutagenesis programme provided several lessons and resources to the knockout mouse project. The phenotype-driven dominant ENU screens proved that mutant mice are indeed indispensable for the functional analysis of the gene. In addition, the basic infrastructure required for phenotyping, archiving mutant strains, managing databases and handling mice were well developed and were directly transferable to the knockout mouse mutagenesis projects. The ENU mouse mutagenesis community also established core centre systems for routine mutant production and phenotype screening as well as many specialized satellite researchers and laboratories to functionally analyse the mutants.

Conversely, some of the limitations that were identified during the early phase of the ENU-mutagenesis programme were resolved by the knockout mouse project. For instance, many recessive mutations segregate in each G1 animal in subsequent crosses, making it necessary to conduct complicated phenotype analyses. A second limitation to ENU mutagenesis is that G1 mice are expected to carry quite a few recessive lethal mutations, which make the litter size of G3 colonies much smaller than expected and necessitate more space and time for phenotyping and mapping. Finally, even when a recessive mutation is stably and reproducibly inherited in offspring, the phenotype occasionally disappears during the outcross, making mapping and positional cloning impossible.

Overall, knockout mouse mutagenesis seemed to be better for genome-wide recessive screening, and so in 2003 mouse geneticists working on projects such as ENU mutagenesis, gene targeting and trapping, and genetic and embryonic engineering gathered at the Banbury Center at the Cold Spring Harbor Laboratory, USA, to discuss the feasibility of this project.

> “The crucial factor that drove knockout projects to be performed on a genome-wide scale was the fall in the estimated number of genes in the mammalian genome.”

Banbury Meeting. After the Banbury Meeting[34,35], an ES-cell-based mutagenesis project in the mouse was officially proposed by two consortia — the Knockout Mouse Project (KOMP)[34] in the United States and the European Conditional Mouse Mutagenesis (EuCOMM)[35]. Both groups emphasized that the technological advancements and needs of recessive mutants by research communities justified a knockout mouse mutagenesis programme. Both consortia proposed a standardized knockout procedure by using the same ES-cell line (C57BL/6N) in addition to targeting or trapping vector(s)[36,37] that carried an expression marker gene such as β-galactosidase or GFP.

KOMP initially proposed a pyramidal approach because of the limited funds, facilities and manpower. At the bottom of the pyramid were all the available knockout mouse lines. If knockout mice did not exist, then the corresponding genes would be knocked out one by one in a high-throughput pipeline. All the knockout ES-cell lines would be archived and freely distributed to the research community. A portion of the ES-cell lines would be used to create live mice through germline transmission; these mice would be subjected to a series of basic phenotypic tests (tier 1 phenotyping). Transcriptome analysis would be conducted on a subset of tier 1 lines. A further round of basic phenotypic tests (tier 2 phenotyping) and specialized phenotyping would be applied to a further subset. Therefore, at the bottom of the pyramid the full set of genome-wide knockout ES-cell lines are archived and, at the top, some pioneering functional analyses are conducted. Decisions are often peer reviewed; for example, to establish which gene to target first or which knockout ES-cell line to subject to tier 2 phenotyping.

In addition to this strategy, EuCOMM emphasized that the efficiency of trapping mutagenesis in ES cells with a standardized conditional vector[36,37] could be maximized by using the Cre–*loxP* or the Flp–*FRT* recombination system, so that the targeted or trapped gene can be turned off in a specific cell type at a certain developmental stage[38]. Such temporal and spatial control of the gene knockout is particularly useful when the knockout allele is embryonic lethal.

Banbury II Meeting in Belgium. Another contributor to the large-scale knockout mouse project — the North American Conditional Mouse Mutagenesis (NorCOMM) in Canada — joined the field just after the Banbury Meeting. Eventually, three knockout and conditional mouse mutagenesis projects were officially started in 2006 and funded as 5-year projects; KOMP (funded by the US National Institutes of Health (NIH)), EuCOMM (funded by the European Commission) and NorCOMM (funded by Genome Prairie/Genome Canada). The overall number of genes to be knocked out, the budgets and the main participants are summarized in TABLE 1. The core of the large-scale knockout mouse project is to archive knockout ES-cell lines for all the 22,000 mouse genes by collecting existing knockout mice and by conducting targeting and trapping over 5 years[11]. For the functional analysis, the three consortia proposed to revive 920 ES-cell lines to live mice[11].

Just after the launch of the three knockout mouse projects, the International Knockout Mouse Consortium (IKMC) was organized[11] and the first IKMC meeting (also called the Banbury II Meeting) was held in 2007 in Brussels, Belgium[39]. One of the key issues was how to avoid duplication between the mutagenesis projects.

Table 1 | **Key features of the International Knockout Mouse Consortium**

Feature	KOMP	EUCOMM	NorCOMM
Number of ES-cell lines created by gene targeting (2006–2010)[11]	8,500	8,000	2,000
Number of ES-cell lines created by gene trapping (2006–2010)[11]	–	12,000	10,000
Number of mouse lines revived from ES cell lines (2006–2010)[11]	500	320	100
Funding agency*	US National Institutes of Health (US$68.8 million)	Sixth Framework Programme (FP6), European Commission (EUR13 million)	Genome Prairie (part of Genome Canada) (CAD$33.8 million)
Main participants*	• Canada: Samuel Lunenfeld Research Institute • UK: Wellcome Trust Sanger Institute • USA: Children's Hospital Oakland Research Institute, University of California (Davis), Regeneron Pharmaceuticals, The Jackson Laboratory, University of Pennsylvania, National Institute on Drug Abuse, 19 National Institues of Health institutions, Deltagen Inc., Lexicon Genetics, Mutant Mouse Regional Resource Centers	• France: Institut Clinique de la Souris (Strasbourg) • Germany: Helmholtz Zentruem Muenchen German Research Center for Environment and Health (GmbH; formerly GSF), University of Frankfurt, Charité — University Medical Berlin, Center for Cardiovascular Research, University of Technology (Dresden), • Italy: European Molecular Biology Laboratory (Monterotondo), Consiglio Nazionale delle Ricerche • UK: Wellcome Trust Sanger Instititute, Medical Research Council (Harwell)	• Canada: Univeristy of Manitoba, University of Toronto, Samuel Lunenfeld Research Institute, Hospital for Sick Children (Toronto), University of Calgary, University of Alberta, University of British Columbia, Toronto Centre for Phenogenomics

*Information is taken from the web sites of the Knockout Mouse Project (KOMP), European Conditional Mouse Mutagenesis (EUCOMM) and North American Conditional Mouse Mutagenesis (NorCOMM). ES, embryonic stem.

The IKMC decided to apply the 'knockout first and conditional ready' strategy[11,39]. In this approach, each gene is disrupted by a vector containing *loxP* elements, giving rise to a null allele. Then, variously engineered patterns of Cre expression allow the gene to be turned on or off, or even to replace the knockout allele with a designed sequence.

Another issue discussed at the IKMC meeting was how best to integrate the various mutagenesis schemes — those carried out by the members of knockout and conditional projects and ENU-mutagenesis projects as well as those by researchers conducting insertional mutagenesis with transposons. Recently, transposon-mediated insertional mutagenesis has become available and large-scale insertional mutagenesis projects with piggyBac[40] or sleeping beauty[41] transposons have been proposed.

Consortia for the infrastructures that are affiliated to the IKMC have also been organized. These consortia manage phenotyping[42] (EuroPhenome; European Mouse Phenotyping Resource of Standardized Screen, EMPReSS; The European Mouse Disease Clinic, EUMODIC), archiving and resourcing[43] (Federation of International Mouse Resources, FIMRe), and database management and informatics[44] (Mouse Genome Informatics, MGI; Coordination and Sustainability of International Mouse Informatics Resources, CASIMIR; and the Mouse Phenome Database). The framework of the project is depicted in FIG. 2.

Mouse mutagenesis phase III

Concurrently with the genome-wide knockout mouse projects, many dominant mutations (and some recessive ones) that were established by phenotype-driven mutagenesis with ENU are being subjected to gene identification by positional cloning. The phenotyping platform — which includes a standardized protocol, a database system to describe phenotypes using ontological labels, and an archiving and distribution system for mutant lines — is also being developed, mainly using already established ENU mutant mice. All of these infrastructures will be synergistically orchestrated for functional genome research, together with the recessive alleles that will be produced by the knockout and conditional mutagenesis programme. However, another resource for the functional annotation of the mouse genome is becoming available from the ENU mouse mutagenesis projects, this is known as mouse mutagenesis phase III.

ENU mutagenesis for a new reverse genetics. ENU randomly induces mutations in the whole genome; thus, in principle it cannot be used for site-specific mutagenesis. However, some feasibility studies[45,46] indicate that ENU-induced mutations could be identified in a target gene by sequence-based screens. Such new reverse genetics systems have already been expanded and implemented by several centres to provide ENU-based gene-driven mutagenesis to the research community[47–50]. The overall scheme of ENU-based gene-driven mutagenesis is depicted in FIG. 3.

ENU mutant mouse library for reverse genetics. A total of ~40,000 G1 mouse lines have been cryopreserved by ENU mutagenesis in frozen sperm archives[47–50]. The archives encompass ~1.2×10^8 base substitutions, because each G1 line carries ~3,000 ENU-induced mutations[46–50]. Summing up the published data[47–50], the 272 ENU-induced base-substitution mutations that were identified in protein-coding and splicing donor or acceptor sequences revealed that 182 (66.9%) were missense mutations, 26 (9.6%) were knockout-equivalent mutations and 64 (23.5%) were synonymous mutations. In this context a knockout-equivalent mutation is a nonsense mutation or a mutation in a splice donor or acceptor sequence that is likely to produce a truncated protein or no mRNA owing to nonsense-mediated decay. If we assume that a target gene consists of 2,000 base pairs of protein-coding sequences then, on

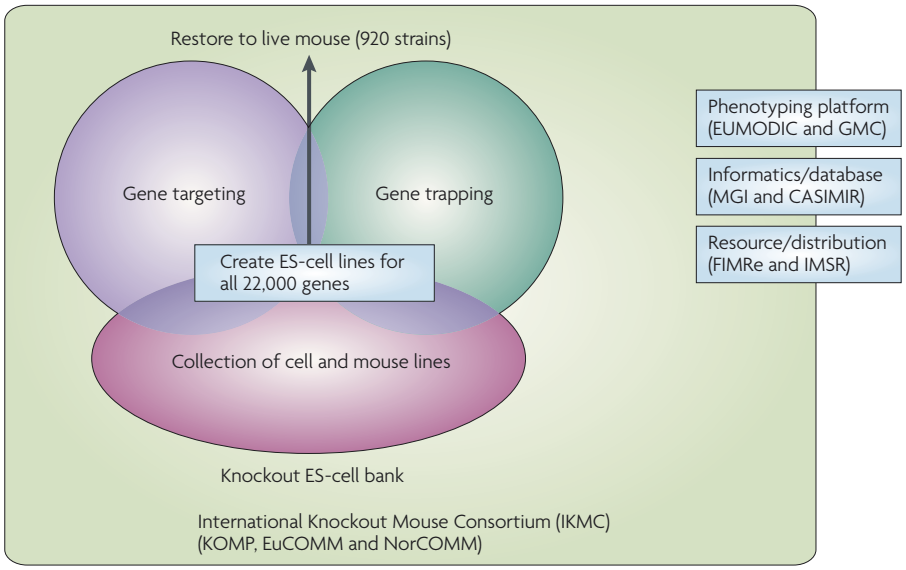


Figure 2 | **Knockout and trapping mutagenesis and related infrastructures.** Through collaborative efforts, knockout mouse strains for all 22,000 mouse genes will be archived as embryonic stem (ES)-cell lines. The International Knockout Mouse Consortium (IKMC)[11], the main members of which are the Knockout Mouse Project (KOMP) in the United States, European Conditional Mouse Mutagenesis (EuCOMM) in the European Union and North American Conditional Mouse Mutagenesis (NorCOMM) in Canada, plans to accomplish this in its 5-year programme. The first step was to collect existing knockout and trapped mouse lines. The research community can nominate genes to be prioritized for targeting, subject to review by the working groups. New mutants will be constructed using gene trapping with a standardized conditional vector. Remaining genes will then be targeted by a standardized conditional-ready targeting vector. In this 5-year period, a part of the collected and constructed knockout ES-cell lines will be restored to live mice, which will be subjected to some pioneering phenotyping. To succeed, the knockout project requires phenotyping platforms, the construction of informatics resources and databases and of resource and distribution infrastructures; examples of groups involved in setting up these requirements are shown in the figure. Some of the consortia and projects that are affiliated with mutagenesis platforms are summarized in TABLE 1. CASIMIR, Coordination and Sustainability of International Mouse Informatics Resources; EUMODIC, European Mouse Disease Clinic; FIMRe, Federation of International Mouse Resources; GMC, German Mouse Clinic; IMSR, International Mouse Strain Resource; MGI, Mouse Genome Informatics.

average, 80 point mutations exist per gene in the frozen sperm archives. This is large enough to be considered the 'mutant mouse library'[49]. Assuming that 1% of the mouse genome is protein-coding, then the 40,000 frozen G1 mouse lines already encompass 800,000 missense mutations and 115,000 knockout-equivalent alleles.

Cataloguing the ENU mutant mouse library. Useful mutations in the ENU mutant mouse libraries are currently screened by constructing PCR primer pairs for one target gene at a time. In the near future it will be possible to catalogue all the ~1.2×10^8 base substitutions in the archives of 40,000 G1 mouse lines by using next-generation sequencing or US$1000 resequencing technology. The estimated cost is simply $1000 multiplied by 40,000 G1 genomes; however, at the initial stage it is sufficient to resequence only the whole protein-coding complement (1–2%)[19] of the genome, at a cost of $400,000–$800,000. This catalogue will provide 115,000 knockout-equivalent mutations in addition to 800,000 missense mutations genome-wide. The sequence-based screening of the worldwide collection of ENU mutant mouse libraries might identify an average of one base substitution every 25 base pairs.

It is noteworthy that the large-scale detection of ENU-induced mutations in the mouse genome revealed no mutation hot spots or cold spots[51]. ENU mutations therefore occur randomly in the genome. By contrast, the efficiency of gene targeting varies from gene to gene, and insertional trapping or transposon mutagenesis is also significantly biased towards some genomic regions. So far, ENU-induced mutation bias has been found only between transcribed and non-transcribed strands of genomic sequences[51].

ENU and knockout come together. The two mutagenesis systems complement each other in many ways in phase III. For instance, the lethality problem of knocking out essential genes might well be evaded by analysing compound heterozygotes between a knockout allele and an ENU-induced allele derived from sequence-based screening. This is equivalent to a classical complementation test between recessive alleles at a locus.

The analysis of knockout alleles using conditional alleles and/or compound heterozygotes after phase II mutagenesis helps to elucidate the function of a single gene, and therefore contributes to our understanding of human monogenic traits and diseases. An allelic series of ENU-induced mutations in G1 mouse lines might provide a good tool to dissect gene function to the base-pair level at a locus. Modifier genes, suppressor genes and epistatic interactions might well be reflected in the ENU-induced mutant lines, as suggested by the disappearance of recessive phenotypes when they are outcrossed to other inbred strains, as described above. Sensitized screens that combine knockout and ENU-mutagenized mice have already provided a series of modifier mutations[52,53]. Strain-specific SNPs, 8.27 million of which have been identified[54], are yet another resource for the genome-wide discovery of modifier mutations. Such modifier mutations have also been identified in inbred and outbred strains by crossing them to ENU-induced ApcMin mice[55].

Conclusions

Based on the current progress of the worldwide knockout mouse project, monogenic traits and related diseases will initially be elucidated as a function of individual genes. Such an approach is supported by the site-directed nature of the gene knockouts created in the uniform C57BL/6N genetic background. In the near future, multigenic effects among several mutations will be better understood by combining an identified mutation, generated either by knockout or ENU reverse genetics, with other background ENU mutations or even strain-specific SNPs. Several other research resources are available for the mouse; for instance, chromosome substitution strains[56,57] (strains that contain an entire chromosome of a donor parent placed in the genetic background of the recipient parent; often called consomic strains) and collaborative crosses[58] (a series of recombinant inbred lines each derived from randomized crosses between strains) take

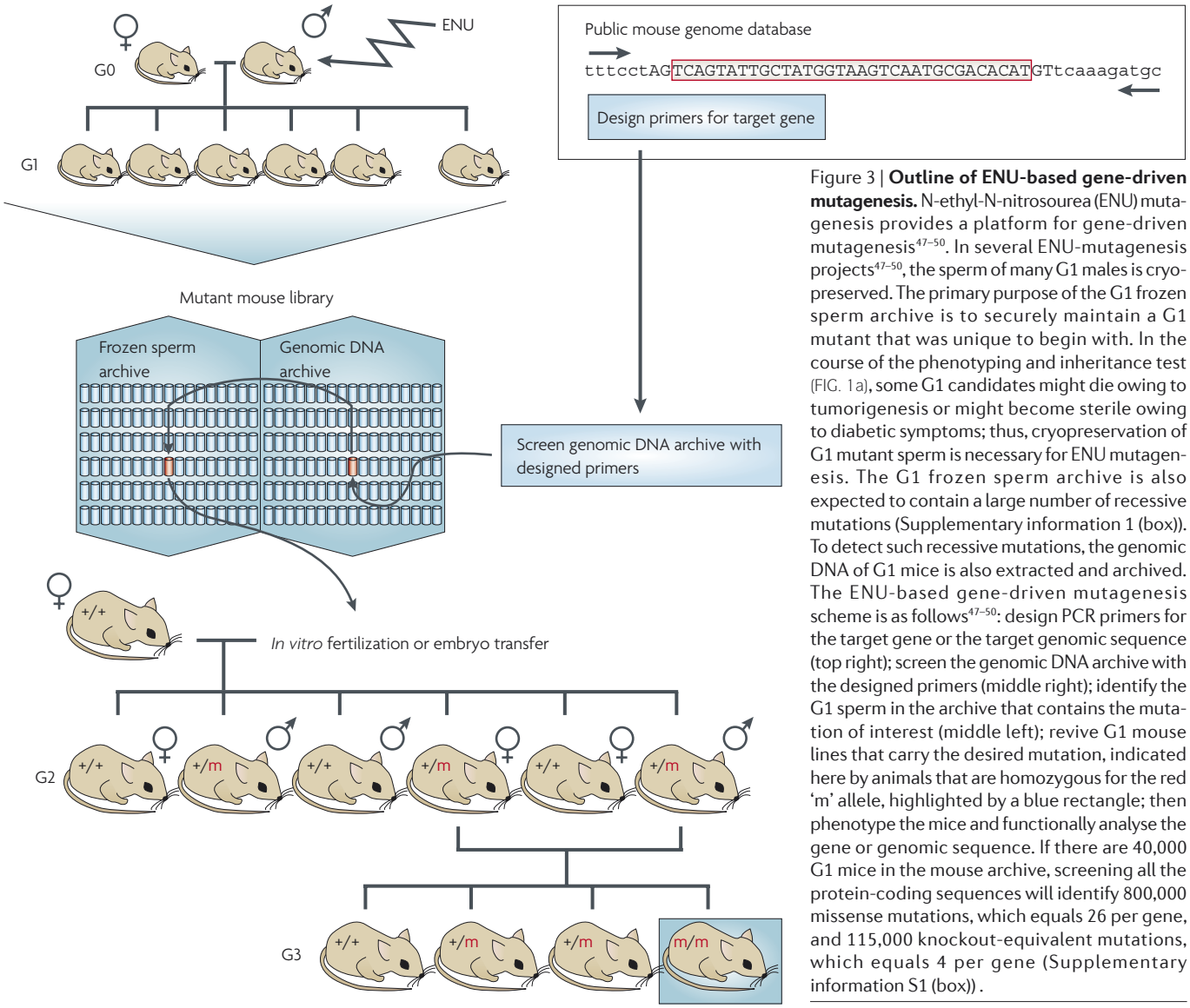


Figure 3 | **Outline of ENU-based gene-driven mutagenesis.** N-ethyl-N-nitrosourea (ENU) mutagenesis provides a platform for gene-driven mutagenesis[47–50]. In several ENU-mutagenesis projects[47–50], the sperm of many G1 males is cryopreserved. The primary purpose of the G1 frozen sperm archive is to securely maintain a G1 mutant that was unique to begin with. In the course of the phenotyping and inheritance test (FIG. 1a), some G1 candidates might die owing to tumorigenesis or might become sterile owing to diabetic symptoms; thus, cryopreservation of G1 mutant sperm is necessary for ENU mutagenesis. The G1 frozen sperm archive is also expected to contain a large number of recessive mutations (Supplementary information 1 (box)). To detect such recessive mutations, the genomic DNA of G1 mice is also extracted and archived. The ENU-based gene-driven mutagenesis scheme is as follows[47–50]: design PCR primers for the target gene or the target genomic sequence (top right); screen the genomic DNA archive with the designed primers (middle right); identify the G1 sperm in the archive that contains the mutation of interest (middle left); revive G1 mouse lines that carry the desired mutation, indicated here by animals that are homozygous for the red 'm' allele, highlighted by a blue rectangle; then phenotype the mice and functionally analyse the gene or genomic sequence. If there are 40,000 G1 mice in the mouse archive, screening all the protein-coding sequences will identify 800,000 missense mutations, which equals 26 per gene, and 115,000 knockout-equivalent mutations, which equals 4 per gene (Supplementary information S1 (box)).

advantage of strain-specific SNP variations. The elucidation of epistatic interactions is the ultimate challenge in genetics and functional genomics.

Such interdisciplinary studies should lead to a better understanding of many human diseases, because most of them have a complex genetic basis. One important challenge will be to produce mouse models of human psychiatric diseases. As reviewed recently[59], many human psychiatric diseases are affected by the subtle effects of alleles at many genes in combination with various environmental factors. Some human psychiatric diseases, however, exhibit high heritability; for instance, the heritability for schizophrenia, bipolar disorder and autism is 80% or more[59]. Such diseases might be well modelled by mutant mice. Indeed, mice carrying an ENU-induced missense mutation in the *Disc1* gene showed schizophrenia-like behaviour[60]. In addition, another missense mutation in the *Disc1* gene resulted in mice with depression-like symptoms[60]. The compound heterozygotes that are created by placing the knockout allele opposite a series of ENU-induced missense alleles should allow us to dissect the molecular functions of *Disc1* in depth. Outcrossing *Disc1* mutant mice to other inbred strains as well as changing environmental factors (including drug administrations) should give us many clues about the mechanisms of schizophrenia and depression and inform the development of preventive or therapeutic interventions.

As extrapolated from the above example, knockout and ENU mutant alleles with chromosome substitution strains and collaborative crosses, which have been developed only in the mouse, will together decode the 3 billion base pairs in the mammalian genome. This is the reason that the mouse has been chosen as the model system for functional genomics and human diseases. To realize this goal it is necessary to promote interdisciplinary collaborative efforts among researches in many disciplines, including genetics, physiology, pathology, molecular biology and informatics.

Yoichi Gondo is a member of the Mutagenesis and Genomics Team, RIKEN BioResource Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan.

e-mail: gondo@brc.riken.jp

doi:10.1038/nrg2431
Published online 10 September 2008

1. Little, C. C. & Bagg, H. J. The occurrence of two heritable types of abnormality among the descendants of X-rayed mice. *Amer. J. Roentgenol. Radiat. Therap.* **10**, 975–989 (1923).
2. Muller, H. J. Artificial transmutation of the gene. *Science* **66**, 84–87 (1927).
3. Russell, W. L. *et al.* Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. *Proc. Natl Acad. Sci. USA* **76**, 5818–5819 (1979).
4. Russell, W. L. *et al.* Dose-response curve for ethylnitrosourea-induced specific-locus mutations in mouse spermatogonia. *Proc. Natl Acad. Sci. USA* **79**, 3589–3591 (1982).
5. Hitotsumachi, S., Carpenter, D. A. & Russell, W. L. Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. *Proc. Natl Acad. Sci. USA* **82**, 6619–6621 (1985).
6. Palmiter, R. D. *et al.* Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* **300**, 611–615 (1982).
7. Capecchi, M. R. Altering the genome by homologous recombination. *Science* **244**, 1288–1292 (1989).
8. Hrabě de Angelis, M. *et al.* Genome-wide, large scale production of mutant mice by ENU mutagenesis. *Nature Genet.* **25**, 444–447 (2000).
9. Nolan, P. M. *et al.* A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nature Genet.* **25**, 440–443 (2000).
10. Takahashi, J. S., Pinto, L. H. & Vitaterna, M. H. Forward and reverse genetic approaches to behavior in the mouse. *Science* **264**, 1724–1733 (1994).
11. International Mouse Knockout Consortium. A mouse for all reasons. *Cell* **128**, 9–13 (2007).
12. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
13. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
14. Lyon, M. F., Rastan, S. & Brown, S. D. M. (eds) *Genetic Variants and Strains of the Laboratory Mouse* 3rd edn (Oxford University Press, Oxford, 1996)
15. Ramirez-Solis, R., Liu, P. & Bradley, A. Chromosome engineering in mice. *Nature* **378**, 720–724 (1995).
16. Kile, B. T. *et al.* Functional genetic analysis of mouse chromosome 11. *Nature* **425**, 81–86 (2003).
17. Whittingham, D. G., Leibo, S. P. & Mazur, P. Survival of mouse embryos frozen to −196°C and −269°C. *Science* **178**, 411–414 (1972).
18. Nakagata, N. Cryopreservation of mouse spermatozoa. *Mamm. Genome* **11**, 572–576 (2000)
19. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
20. Noveroske, J. K., Weber, J. S. & Justice, M. J. The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mamm. Genome* **11**, 478–483 (2000).
21. Moser, A. R., Pitot, H. C. & Dove, W. F. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* **247**, 322–324 (1990).
22. Su, L.-K. *et al.* Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science* **256**, 668–670 (1992).
23. Vitaterna, M. H. *et al.* Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* **264**, 719–725 (1994).
24. King, D. P. *et al.* Positional cloning of the mouse circadian *Clock* gene. *Cell* **89**, 641–653 (1997).
25. Rogers, D. C. *et al.* Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment. *Mamm. Genome* **8**, 711–713 (1997).
26. Favor, J. & Neuhäuser-Klaus, A. Saturation mutagenesis for dominant eye morphological defects in the mouse *Mus musculus*. *Mamm. Genome* **11**, 520–525 (2000).
27. Fuchs, H. *et al.* Screening of dysmorphological abnormalities — a powerful tool to isolate new mouse mutants. *Mamm. Genome* **11**, 528–530 (2000).
28. Flaswinkel, H. *et al.* Identification of immunological relevant phenotypes in ENU mutagenized mice. *Mamm. Genome* **11**, 526–527 (2000).
29. Pretsch, W. Enzyme-activity mutants in *Mus musculus*. I. Phenotypic description and genetic characterization of ethylnitrosourea-induced mutations. *Mamm. Genome* **11**, 537–542 (2000).
30. Inoue, M. *et al.* A series of maturity onset diabetes of the young, type 2 (MODY2) mouse models generated by a large-scale ENU mutagenesis program. *Hum. Mol. Genet.* **13**, 1147–1157 (2004).
31. Rinchik, E. M. & Carpenter, D. A. N-ethyl-N-nitrosourea mutagenesis of a 6- to 11-cM subregion of the Fah-Hbb interval of mouse chromosome 7: completed testing of 4557 gametes and deletion mapping and complementation analysis of 31 mutations. *Genetics* **152**, 373–383 (1999).
32. Hagge-Greenberg, A., Snow, P. & O'Brien, T. P. Establishing an ENU mutagenesis screen for the piebald region of mouse chromosome 14. *Mamm. Genome* **12**, 938–941 (2001).
33. Chikaraishi, D. M., Deeb, S. S. & Sueoka, N. Sequence complexity of nuclear RNAs in adult rat tissues. *Cell* **13**, 111–120 (1978).
34. Austin, C. P. *et al.* The knockout mouse project. *Nature Genet.* **36**, 921–924 (2004).
35. Auwerx, J. *et al.* The European dimension for the mouse genome mutagenesis program. *Nature Genet.* **36**, 925–927 (2004).
36. Testa, G. *et al.* A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. *Genesis* **38**, 151–158 (2004).
37. Schnütgen, F. *et al.* Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. *Proc. Natl Acad. Sci. USA* **102**, 7271–7226 (2005).
38. Metzger, D. & Chambon, P. Site- and time-specific gene targeting in the mouse. *Methods* **24**, 71–80 (2001).
39. Editorial. Mutant mice galore. *Nature* **446**, 469–470 (2007).
40. Sun, L. V. *et al.* PBmice: an integrated database system of piggyBac (PB) insertional mutations and their characterization in mice. *Nucleic Acid Res.* **36**, D729–D734 (2007).
41. Dupuy, A. J., Jenkins, N. A. & Copeland, N. G. Sleeping beauty: a novel cancer gene discovery tool. *Human Mol. Genet.* **15**, R75–R79 (2008).
42. The Eumorphia Consortium. EMPReSS: standardized phenotype screens for functional annotation of the mouse genome. *Nature Genet.* **37**, 1155 (2005).
43. FIMRe Board of Directors. FIMRe: Federation of International Mouse Resources: global networking of resource centers. *Mamm. Genome* **17**, 363–364 (2006).
44. The Mouse Phenotype Database Integration Consortium. Integration of mouse phenome data resources. *Mamm. Genome* **18**, 157–163 (2007).
45. Beier, D. R. Sequence-based analysis of mutagenized mice. *Mamm. Genome* **11**, 594–597 (2000).
46. Coghill, E. L. *et al.* A gene-driven approach to the identification of ENU mutants in the mouse. *Nature Genet.* **30**, 255–256 (2002).
47. Quwailid, M. M. *et al.* A gene-driven ENU-based approach to generating an allelic series in any gene. *Mamm. Genome* **15**, 585–591 (2004).
48. Augustin, M. *et al.* Efficient and fast targeted production of murine models based on ENU mutagenesis. *Mamm. Genome* **16**, 405–413 (2005).
49. Sakuraba, Y. *et al.* Molecular characterization of ENU mouse mutagenesis and archives. *Biochem. Biophys. Res. Commun.* **336**, 609–616 (2005).
50. Michaud, E. J. *et al.* Efficient gene-driven germ-line point mutagenesis of C57BL/6J mice. *BMC Genomics* **6**, 164 (2005).
51. Takahasi, K. R., Sakuraba, Y. & Gondo, Y. Mutational pattern and frequency of induced nucleotide changes in mouse ENU mutagenesis. *BMC Mol. Biol.* **8**, 52 (2007).
52. Rubio-Aliaga, I. *et al.* A genetic screen for modifiers of the Delta1-dependent Notch signaling function in the mouse. *Genetics* **175**, 1451–1463 (2007).
53. Matera, I. *et al.* A sensitized mutagenesis screen identifies *Gli3* as a modifier of *Sox10* neurocristopathy. *Human Molec. Genet.* **17**, 2118–2131 (2008).
54. Frazer, K. A. *et al.* A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. *Nature* **448**, 1050–1053 (2007).
55. Kwong, L. N. *et al.* Identification of *Mom7*, a novel modifier of $Apc^{Min/+}$ on mouse chromosome 18. *Genetics* **176**, 1237–1244 (2007).
56. Nadeau, J. H., Singer, J. B., Matin, A. & Lander, E. S. Analysing complex genetic traits with chromosome substitution strains. *Nature Genet.* **24**, 221–225 (2000).
57. Oka, A. *et al.* Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics* **166**, 913–924 (2004).
58. The Complex Trait Consortium. The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genet.* **36**, 1133–1137 (2004).
59. Burmeister, M., McInnis, M. G. & Zöllner, S. Psychiatric genetics: progress amid controversy. *Nature Rev. Genet.* **9**, 527–540 (2008).
60. Clapcote, S. J. *et al.* Behavioral phenotypes of *Disc1* missense mutations in mice. *Neuron* **54**, 387–402(2007).

FURTHER INFORMATION

AMMRA, Asian Mouse Mutagenesis and Resource Association: http://www.ammra.info
APF, Australian Phenomics Facility: http://www.apf.edu.au
Baylor College of Medicine ENU mouse mutagenesis: http://www.mouse-genome.bcm.tmc.edu/ENU/MutagenesisProj.asp
CARD, Center for Animal Resources and Development: http://cardb.cc.kumamoto-u.ac.jp/transgenic/index.jsp
CASIMIR, Coordination and Sustainability of International Mouse Informatics Resources: http://www.casimir.org.uk
CMHD, Center for Modeling Human Disease: http://www.cmhd.ca
DBCLS, Database Center for Life Sciences: http://dbcls.rois.ac.jp/en
EMMA, European Mouse Mutant Archive: http://www.emmanet.org
EMPReSS, European Mouse Phenotyping Resource of Standardized Screen (SOP database): http://empress.har.mrc.ac.uk
EuCOMM, European Conditional Mouse Mutagenesis: http://www.eucomm.org
EUMODIC, The European Mouse Disease Clinic: http://www.eumodic.org
Eumorphia, European Phenotyping Consortium 2002–2006: http://www.eumorphia.org
EURExpress II, gene expression atlas by RNA *in situ*: http://www.eurexpress.org/ee
EuroPhenome, Europhenome Mouse Phenotyping Resource: http://www.europhenome.org
Experimental Animal Division, RIKEN BioResource Center: http://www.brc.riken.jp/lab/animal/en
FIMRe, Federation of International Mouse Resources: http://www.fimre.org
GMC, German Mouse Clinic: http://www.mouseclinic.de
Helmholtz Zentrum Munchen German Research Center ENU mouse mutagenesis: http://www.helmholtz-muenchen.de/en/ieg/group-functional-genetics/enu-screen/index.html
IGTC, International Gene Trap Consortium: http://www.genetrap.org
IMSR, International Mouse Strain Resource: http://www.informatics.jax.org/imsr/index.jsp
INGENOtyping for ENU gene-driven mutagenesis: http://www.ingenium-ag.com
JAX Mice, The Jackson Laboratory: http://jaxmice.jax.org/index.html
JAX Mouse Heart, Lung, Blood and Sleep Disorders Center: http://pga.jax.org/index.html
JAX Mouse Phenome Database: http://aretha.jax.org/pub-cgi/phenome/mpdcgi?rtn=docs/home
JAX Neuroscience Mutagenesis Facility: http://nmf.jax.org
JMSR, Japan Mouse/Rat Strain Resource Database: http://shigen.lab.nig.ac.jp/mouse/jmsr/strainsList.jsp
KOMP, Knockout Mouse Project: http://www.komp.org
McLaughlin Research Institue ENU mouse mutagenesis: http://www.montana.edu/wwwmri/enump.html
Medical Research Council, archive for ENU gene-driven mutagenesis: http://www.har.mrc.ac.uk/services/dna_archive
Medical Research Council ENU mouse mutagenesis: http://www.mut.har.mrc.ac.uk/research/mutagenesis
MGI, Mouse Genome Informatics: http://www.informatics.jax.org
MMRRC, Mutant Mouse Regional Resource Centers: http://www.mmrrc.org
NorCOMM, North American Conditional Mouse Mutagenesis: http://norcomm.phenogenomics.ca/index.htm
Northwestern University ENU mouse mutagenesis: http://genomics.northwestern.edu/neuro
Oak Ridge National Laboratory Mutant Mouse Database: http://mouse.ornl.gov/mmdb
PBmice, PiggyBac Mutagenesis Information Center: http://www.idmshanghai.cn/PBmice
RIKEN ENU mouse mutagenesis: http://www.brc.riken.go.jp/lab/gsc/mouse/index.html
RIKEN Mutant Mouse Library, ENU-based gene-driven mutagenesis: http://www.brc.riken.go.jp/lab/mutants/genedriven.htm
TCP, Toronto Centre for Phenogenomics: http://www.phenogenomics.ca/services/phenotyping.html
Tennessee Mouse Genome Consortium ENU mouse mutagenesis: http://www2.tnmouse.org

SUPPLEMENTARY INFORMATION

See online article: S1 (box)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF