BMB
reports

Mini Review

Next-generation gene targeting in the mouse for functional genomics

Yoichi Gondo, Ryutaro Fukumura, Takuya Murata & Shigeru Makino*

Mutagenesis and Genomics Team, RIKEN BioResource Center, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

In order to elucidate ultimate biological function of the genome, the model animal system carrying mutations is indispensable. Recently, large-scale mutagenesis projects have been launched in various species. Especially, the mouse is considered to be an ideal model to human because it is a mammalian species accompanied with well-established genetic as well as embryonic technologies. In 1990's, large-scale mouse mutagenesis projects firstly initiated with a potent chemical mutagen, *N*-ethyl-*N*-nitrosourea (ENU) by the phenotype-driven approach or forward genetics. The knockout mouse mutagenesis projects with trapping/conditional mutagenesis have then followed as Phase II since 2006 by the gene-driven approach or reverse genetics. Recently, the next-generation gene targeting system has also become available to the research community, which allows us to establish and analyze mutant mice carrying an allelic series of base substitutions in target genes as another reverse genetics. Overall trends in the large-scale mouse mutagenesis will be reviewed in this article particularly focusing on the new advancement of the next-generation gene targeting system. The drastic expansion of the mutant mouse resources altogether will enhance the systematic understanding of the life. The construction of the mutant mouse resources developed by the forward and reverse genetic mutagenesis is just the beginning of the annotation of mammalian genome. They provide basic infrastructure to understand the molecular mechanism of the gene and genome and will contribute to not only basic researches but also applied sciences such as human disease modelling, genomic medicine and personalized medicine. [BMB reports 2009; 42(6): 315-323]

INTRODUCTION

Mutants are useful tools to directly understand the biological function of the gene and genome. For instance, the mechanism of tumorigenesis is now evident to be multistage progress with complicated interactions (1, 2). The mice carrying a disrupted p53 gene (p53 mutants) were proven to develop various tumors (p53 phenotypes) (3, 4). Based on this knowledge, the mechanism of tumorigenesis with the p53 gene function has been extensively studied and revealed at the molecular and cellular levels (5). At the same time, the established p53 mutant mice serve a good model for human tumorigenesis.

Mutants are, therefore, useful tools to directly understand the biological function of the gene and genome without knowing any molecular or cellular mechanisms. Rather, mutant mice and genetics lead to the understanding of the complicated molecular and cellular network of the life. Mutant mice are also indispensable for genomic medicine and personalized medicine as animal models to various human diseases.

Forward genetics and positional cloning

Classically, mutants have been identified and established by the recognition of phenotype first as shown by the pioneering work by Mendel. In old days, there are no technologies available to change the genomic DNA sequences except random mutagenesis. It was for the first time suggested in the mouse by Little and Bagg (6) that x-ray induces heritable mutations. Immediately after their finding, it was proved in fruit flies by Muller (7). Various chemical mutagens have also been identified but all of them randomly induce mutations in the genome. The identification of the genomic sequence change to reveal the responsible gene to the mutant phenotype had been impossible particularly in mammals. In 1980's, it became feasible to identify the site of mutation by positional cloning. Combining the genetic mapping method to narrow down the site of the mutation in the chromosome with the DNA cloning and sequencing technologies at and around the candidate genomic region, the site of the mutation is now able to be identified at the DNA sequence level. This classical approach to connect the site of the mutation to the phenotype is called "phenotype-driven mutagenesis" or "forward genetics" (Fig. 1).

By the positional cloning, various causative genes and specific genomic DNA changes have been identified for the human genetic diseases. As a few of many examples, the DMD (8), RB1 (9) and HTT (10) genes were positionally cloned from Xp21.2, 13q14.2 and 4p16.3 as a causative genes for Duchenne muscular dys-

*Corresponding author. Tel: 81-29-836-9232; Fax: 81-29-836-9098; E-mail: gondo@brc.riken.jp

Received 14 May 2009

Keywords: Functional genomics, Model animal, Mouse, Mutagenesis, Reverse genetics

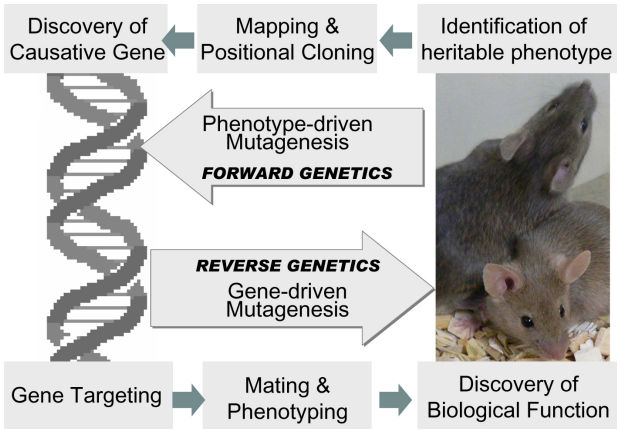

Fig. 1. Direct annotation of genome function by genetics. Genetics directly associates the biological trait (phenotype) to genomic DNA sequence (genotype). Mutant strains may be established by either phenotype or genotype first. Upper panel: when a phenotype is identified at first and it's heritability is then confirmed (like Mendel's classical experiment), the causative gene and change of genomic DNA sequence will be discovered by the genetic mapping and positional cloning. This genetic approach is called phenotype-driven mutagenesis or forward genetics. Lower panel: alternatively, mutant strains may be constructed by site-directed mutagenesis that targets a specific gene of the interest first. The mutant phenotype, namely, biological function will be elucidated by mating and analyzing the mutant animals. This approach is called gene-driven mutagenesis or reverse genetics.

trophy, retinoblastoma and Huntington's disease, respectively. To reveal the predisposition and mechanism of the disease progression in terms of the molecular function of the causative gene, it has become necessary to develop mice that carry a mutation(s) in the causative gene. Furthermore, when it is not clear which candidate gene is the true causative gene by the positional cloning, mice carrying mutations in the candidate gene will also provide a very useful tool to verify. Based on these needs, another genetic approach, namely, the site-directed mutagenesis in the mouse, was newly developed in the 1980's.

Reverse genetics and gene targeting

The developed site-directed mutagenesis in the mouse is called "gene targeting" or "knockout mouse", since this technology specifically disrupts the gene of interest. The disruption of the target gene was conducted by homologous recombination with the targeting vector carrying a positively selectable marker like the neomycin resistant gene, which was independently developed by Capecchi (11) and Smithies (12). This targeting manipulation was conducted in the mouse embryonic stem (ES) cells, developed by Evans (13, 14), so that the disrupted gene is transmitted to the live mice through the germline of the chimeric mice with the manipulated ES cells. Although the gene targeting requires high-precision skills and well-controlled tissue culture and mouse facilities, it soon became a standard technology to study the mammalian gene in early 1990's all over the world because it reveals the gene function directly at the organism level. The Nobel Prize in physiology or medicine in 2007 was awarded to Drs. Capecchi, Smithies and Evans because of the tremendous contribution and impact to the basic as well as applied sciences. The gene targeting or knockout mouse made it possible to reverse the genetic approach to elucidate gene function compared to the classical forward genetics. Thus, it is "reverse genetics" that allows us to conduct the "gene-driven mutagenesis" in the mosue (Fig. 1).

Phase I: ENU mouse mutagenesis

In 1990, many fundamental technologies to elucidate the gene and genome function were set; genomic DNA sequencing, positional cloning and forward and reverse genetics. To facilitate the basic infrastructure for the functional genomics, various "big projects" has been funded and organized in worldwide. The Human Genome Project firstly primed this trend (15, 16). In 1997, foreseeing the success of the Human Genome Project, two large-scale mouse mutagenesis projects with phenotype-driven approach were launched in Germany (17, 18) and UK (19, 20) as Phase I. Japan and Australia independently started a similar phenotype-driven mutagenesis project in 1999 and US and Canada joined in 2000, altogether approximately twenty of such large-scale phenotype-driven mutagenesis have been conducted in the world.

The phenotype-driven approach with a potent alkylating mutagen ENU was taken in Phase I mouse mutagenesis because 1) ENU had been well known to effectively induce point mutations in a genomewide manner (21-26), which resemble most of the cause of human genetic diseases, and 2) the phenotype-driven approach enable to identify yet-unknown genes responsible for human diseases. On the other hand, the ENU mutagenesis has been conducted mostly for dominant phenotype screenings but not for recessives (27). The causative gene must be identified by the positional cloning that has been expected to become much quicker and easier to conduct by the completion of the entire mouse genome sequencing but is still time-consuming and often difficult process even now (27). After all, many ENU-induced mutant strains were established and most of them are currently still under positional cloning. The large-scale ENU mouse mutagenesis is also expanded to the sensitized screening to reveal the modifiers for gene-gene interactions and effectors for gene-environmental interactions (27, 28).

Phase II: knockout/conditional mouse mutagenesis

In the 2003, large-scale gene-driven mutagenesis project in the mouse was proposed (29, 30). KOMP, EUCOMM and NorCOMM have indeed begun since 2006 as a 5-year project at US, EU and Canada, respectively (27, 31, 32). The primary goal of the knockout/conditional mutagenesis projects is to construct

mouse ES cell lines that cover at least one "knockout-first conditional-ready" mutant for every gene. The target of the random ENU mutagenesis is 3 billion base pairs of whole mouse genome; thus, the chance of having exactly the same base substitution by ENU among different projects is practically none. On the other hand, the knockout/conditional mutagenesis aims to disrupt merely about 30,000 mouse genes; thus, some coordination among the KOMP, EUCOMM and NorCOMM has been necessary to avoid the redundant construction of mutant ES cell lines. In 2007, the International Mouse Knockout Consortium (IMKC) was organized to arrange the effective international coordination and exchange the information among the participating groups (31, 32).

Nevertheless, the trapping rather than targeting has been the major method to construct knockout ES cell lines (27, 31). Therefore, duplicated disruption of the same gene is more or less unavoidable. Each project designed own trapping/conditional vectors in order to make even duplicated mutant ES cell lines of the same gene be able to provide some different utilities. Users must take account for the design and utility of the targeting/trapping vectors when users find several knockout ES cell lines are available for the gene of their interest. Another yet significant concern is that the knockout mutation often ends up with embryonic lethality or even no obvious phenotypes at all. The phenotypes observed from knockout mice were discussed at the Panel Discussion in the International Mammalian Genome Conference (33). Overall very rough estimation of the knockout phenotypes was that one third may be recessive embryonic lethal, the other one third may give rise to some phenotypes after born, and the remaining one third may have no phenotypes. It is one of the major reasons why the knockout/conditional mutagenesis projects designed the "knockout-first conditional-ready" targeting/trapping vectors (27, 31, 32). In order to fully utilize the knockout/conditional resource of the mutant ES cell lines, therefore, it is necessary to set various "CRE-zoo" transgenic mouse lines.

Users of the knockout/conditional mutant ES cell lines also concern about how they are able to obtain the mutant mouse stains. To make live mice from ES cells, it is necessary to have the technologies and facilities to construct germline chimeras with in vitro fertilization (IVF) and embryo transfer (ET). Nowadays, the IVF/ET with freezing technology (34-36) may become a routine work for many experimental mouse facilities in the specific-pathogen free (SPF) facilities and for the storage of precious mouse strains in liquid nitrogen tanks. However, not all of the mouse laboratories prevalently conduct the germline chimera construction, yet. Thus the effective distribution of established knockout resources is another issue to be concerned by the research community.

Phase III: next-generation gene-targeting

The development of the site-directed mutagenesis to introduce a single base substitution in the mouse genome has been a big dream in the mammalian genetics. The knock-in mouse is one of such technologies. More elaborated methods have also been developed. For instance, the double replacement targeting system has been designed and successfully developed (37-40). This double replacement targeting is an ideal universal gene replacement method, since it replaces the target gene to any designed sequences; a single base substitutions, insertions, deletions, and whole replacement to totally different sequences in the ES cells (40). In addition, it does not leave any artificial base pairs like loxP or selectable markers at all. The double replacement targeting system, however, was not adopted for the international knockout/conditional mutagenesis projects because of its extremely low efficiency of the second replacement step in the ES cells.

On the other hand, the next-generation gene targeting that enables to establish mutant mice carrying a single base pair substitution in the target gene was paradoxically developed by the random mutagenesis with ENU. The overview of this ENU-based gene-driven mutagenesis is summarized in Fig. 2. The construction of dual mutant archives and the development of the novel mutation discovery system were necessary to make the next-generation gene-targeting system feasible.

Dual archive construction

When the large-scale ENU mouse mutagenesis by the phenotype-driven approach was started at RIKEN in 1999, we decided to keep the original G1 mice as frozen sperm. G1 stands for Generation-1 that is the F1 offspring directly derived from the mating of ENU-treated F0 males (27). Thus, each G1 is the founder mouse of the ENU-induced mutations that will be inherited to the subsequent generations according to the Mendel's law. Primarily, the G1 mice are the subjects of the phenotype screening for the large-scale ENU mutagenesis (27). The original rationale to preserve all the G1 sperm in liquid nitrogen tank at RIKEN was to conduct various late-onset phenotype screenings and to find genetic traits similar to human common diseases from many G1 mice. For instance, high blood pressure, hearing and vision loss, diabetes, and/or tumorigenesis have been the candidate for late-onset phenotypes. Such traits associated to the senescence often accompany the sterility and/or lethality. Even if we identify a diabetic G1 mouse, it often becomes sterile concomitantly, in which case, we should lose this strain permanetly. Thus, we decided to preserve all the G1 sperm at the 3 months after the birth irrespective of any phenotype appearance at this time; thereby, even if we lose the G1 mouse due to its sterility or lethality, we can restore the strain from the frozen sperm.

The G1 sperm archive also preserves all the ENU-induced mutations in each G1 genome. If a large enough number of G1 sperm samples are accumulated, then, any gene may well be mutagenized by ENU. It is a matter of chance; *i.e.*, given more the number of G1 mice and higher the ENU-induced mutation rate, better chance to have more number of ENU-induced mutations per gene. At this time of period, this ENU-

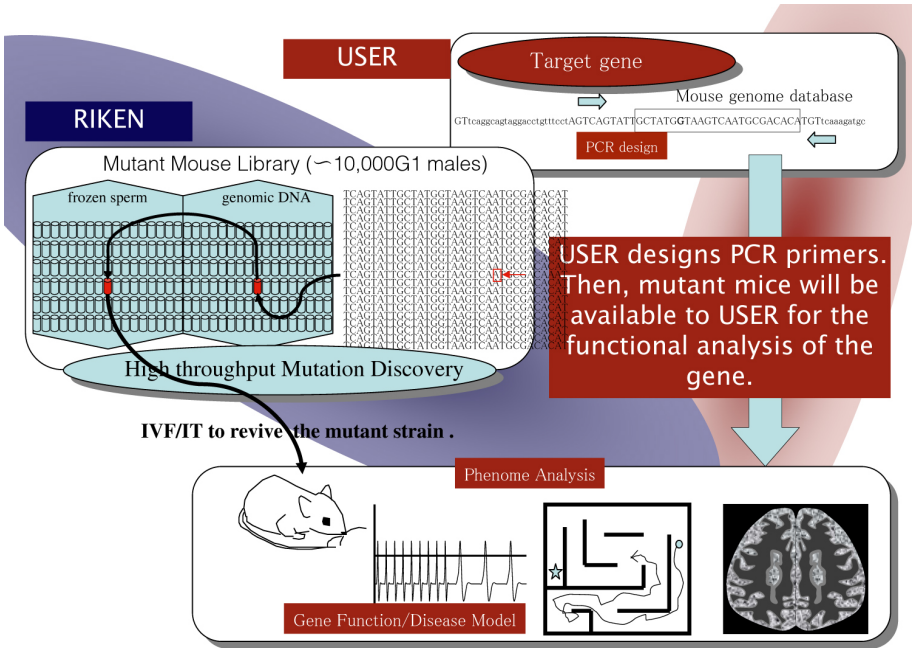

Fig. 2. Scheme of RIKEN gene-driven mutagenesis system (RGDMS). User who has own gene(s) of interest may design and send the PCR primer pair(s) to RIKEN. RIKEN then screens the ENU-mutagenized G1 genomic DNA archive and report all the discovered mutations to User. User decides which mutation to analyze and RIKEN revives the mutant strain as live mice from the fozen sperm by IVF/ET and send them to User for functional analysis of the gene.

based gene-driven strategy had been considered that it would have been merely a theoretical possibility but practically impossible. Firstly, there were no large enough G1 sperm archive at all in the world yet and nobody knows how many G1 sperm samples were necessary to make this ENU-based gene-driven strategy feasible at all. Moreover, even if we had had enough number of G1 sperm archive, no practical technologies were yet available to identify ENU-induced new mutations in a particular target gene in the 3 billion base pairs of the mouse genome.

The DNA sequencing technologies during 1990's had been drastically advanced and the sequencing cost had been inversely reduced. Thus, we decided to purify the genomic DNA from each G1 male in order to construct the ENU-mutagenized genomic DNA archive. At the same time, we have started a feasibility study to find ENU-induced mutations from the ENU-mutagenized genomic DNA archive. Once we find a mutation in a particular target gene, it now becomes possible to revive the discovered mutation as live mice by IVF/ET as depicted in Fig. 2. The dual archive of frozen sperm and genomic DNA from G1 mice are the key resources for the next-generation gene targeting.

Mutation discovery methods

We have started to experimentally detect newly induced mutations and estimate the ENU-induced mutation rate in the mouse genome. The first trial was to use the direct sequencing method of the PCR-amplified DNA fragments from the ENU-mutagenized genomic DNA archive. This method was found to be too expensive to conduct in a large scale at that time. In addition, it was extremely difficult to identify the site of ENU-induced mutations from sequencing chromatograms produced by the automated capillary sequencer like ABI3700.

We then tried several mutation discovery systems based on the heteroduplex detection. The G1 mice only inherited the ENU-induced mutations from F0 male; therefore, all the ENU-induced mutations in the G1 genome are heterozygous. After PCR amplification of a target sequence, the PCR fragments are denatured and annealed; then, paternally- and maternally-derived products randomly form double-stranded DNA again. If no ENU-induced mutations locate in the PCR products, all annealed fragments forms perfectly matched double-stranded DNA (homoduplex). When an ENU-induced mutation(s) exists in the target PCR sequence, then half of the annealed DNA fragments contain one mismatch (heteroduplex) in addition to the other half of homoduplex. Thus, we are able to quickly identify the G1 genomic DNA samples that carry an ENU-induced mutation(s) by distinguishing the heteroduplex-containing samples from solely homoduplex ones.

We firstly compared two heteroduplex detection methods; the dHPLC (Transgenomic WAVE system) (41, 42) and the temperature gradient capillary electrophoresis (TGCE) method (43-45) by using SCE9610 (SpectruMedix). We chose the TGCE method because of the throughput, cost performance and reproducibility to screen the ENU-induced mutations in the genomic DNA archive. Recently, we have added several other mutation discovery systems; for instance, the TILLING/Cel1 method (46, 47), the high-resolution melting method (48, 49) and the direct re-sequencing method using next-generation sequencers such as 454GS-FLX (Roche), GA-II (Illumina) and/or

SOLiD (AB).

ENU-induced mutation rate

In the early phase of the development of the ENU-based gene-driven mutagenesis of the mouse, there were several reports about the ENU-induced mutation rate. In 2000, Beier (50) reported a preliminary mutation rate of 16/Mb by directly sequencing a total of 370 kb in a small cohort of 192 G1 mice. In 2001, Gondo *et al*. (51) expected the ENU-induced mutation rate to be roughly 2/Mb based on the specific locus tests (21-24) with several assumptions and also presented a preliminary experimental finding of 1.4/Mb by screening a total of ~30 Mb from 2,000 G1 genomic DNA. In 2002, there were two short communications. Coghill *et al*. (52) found 0.5/Mb ENU-induced mutations by screening a total of 9.5 Mb from 2,230 G1 mice. Gondo *et al*. (53) reported 0.91/Mb in a total of 35 Mb screening. Conception *et al*. (54) found 1/Mb by screening 9.6 Mb in a rather small cohort of 500 G1. Three groups reported a large-scale experimental evaluation of ENU-induced mutation in the mouse. Quwailid *et al*. (55) found 0.99/Mb mutations in 27.4 Mb screening of 6,000 G1 archive. Augustin *et al*. (56) identified 0.35/Mb mutations by screening 612 Mb of 9367 G1 archive. Sakuraba *et al*. (57) detected 0.75/Mb mutations in a total of 197 Mb screening from the RIKEN archive of 9,224 G1 mice. The ENU-induced mutation rates seem to be quite different among the reported data even in the large-scale analyses. The differences of the ENU doses, mutation discovery systems and mouse strains may have affected the mutation rate. Nevertheless, the ENU-induced mutation rate in the mouse genome was roughly 0.5-1.0/Mb.

ENU-induced mutation: distribution and spectrum

Mutations often arise in limited regions more and less than other regions in the genome, which is called mutation "hot spots" and "cold spots", respectively. If such mutation hot and cold spots would be prevalent in the mouse genome, the expected distribution of ENU-induced mutations in the entire genome would be quite different from the experimentally examined patterns as described above. First of all, there were no hot or cold spots in the ENU-induced mutations so far detected at all (57). The discovered mutations were found according to the Poisson distribution that implies the random distribution. Namely, the numbers of discovered mutations are proportional to the length of the screening. For instance, we detected 65 and 80 ENU-induced mutation by screening 86.8 Mb and 110.7 Mb of coding and noncoding sequences, respectively (57). We screened a total of 63 loci covering all the autosomes. They are chosen not by the chromosomal locations but due to the interest of the gene function; thus, the 63 tested loci are regarded as random sampling in terms of chromosomal locations. The numbers of discovered mutations were randomly distributed in proportional to the size of the PCR amplicons (57). It is noteworthy that we also found many long conserved noncoding sequences (LCNS) that are highly conserved in vertebrates and studied the rate of ENU mutations. ENU equally induced the mutations in the LCNS as well (57, 58).

We also analyzed ENU-induced mutation spectrum in details (57, 59). The A/T to G/C transitions and A/T to T/A transversions were found more often than the random expectation and were consistent with the molecular knowledge of the ENU adducts in the DNA bases (57). Conversely, the G/C to C/G transversions were hardly detected (1 in 142 mutations or 0.7%) compared to the random expectation of 15.8% (57). Other than these preferences of base substitutions induced by ENU, all the discovered mutations were distributed randomly in entire tested genomic regions except the preference of non-transcribed sequences against transcribed sequences (59).

Naturally, the mutation spectra of ENU-induced mutations discovered by the gene-driven screening were significantly different from those by phenotype-driven screening (57, 59). As discussed above, ENU-induced mutations were equally discovered in coding exons, noncoding exons, introns, flanking sequences and even LCNS by gene-driven screening; on the other hand, by phenotype screening, no mutations were reported in noncoding exons and flanking sequences. When mutations were found in intron by phenotype-driven screening, most, if not all, of the mutations caused splicing anomalies (59). The A/T to T/A and G/C to C/G transversions were detected 1.8-folds and 4.7-folds more often, respectively, by phenotype-driven screening than by gene-driven screening (57), indicating that such base substitutions may cause more drastic functional changes than other substitutions.

Mutant mouse library

After obtaining the ENU-induced mutation rate and confirming the genomewide random distribution in the mouse genome, it became possible to estimate the nature of the dual archives constructed by ENU (27). The overall estimate of the ENU-induced mutations was given as 1/Mb; therefore, the average number of ENU-induced mutation per G1 mouse was 3,000. The RIKEN archive encompassed about 10,000 G1; therefore, a total of 30,000,000 base substitutions were preserved in liquid nitrogen tanks (Fig. 3). The protein coding sequences in the mouse genome is at least 1%; thus, the 300,000 mutations locate in the protein coding sequences. Assuming the number of mouse gene is 30,000, we are able to provide 10 different mutations per gene on average; thus, we consider the dual archive to be 10-fold coverage of mutant moue library (57). According to the peptide alteration pattern deduced from our data (57), 60% and 10% of the ENU-induced mutations in coding plus splicing donor/acceptor signal sequences were missense and knockout-type mutations, respectively (27). Therefore, 180,000 missense mutations and 30,000 knockout-type mutations are available in the RIKEN archive (Fig. 3).

Current progress and future directions

The next-generation gene-targeting system based on RIKEN mutant mouse library has been available to the research com-

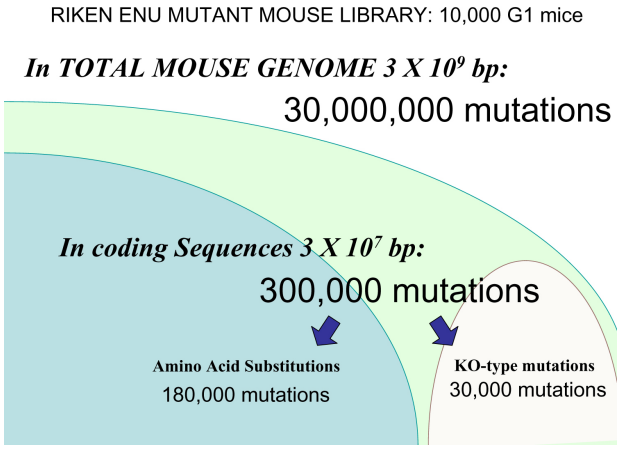

Fig. 3. Preserved ENU-induced mutations in the RIKEN mutant mouse library. Each G1 carries 3,000 ENU-induced mutations on average. Thus, a total of 30,000,000 mutations are preserved in the RIKEN library. Coding exons consist of at least 1% of the mouse genome; thus, 300,000 mutations locate in the coding sequences, 60% and 30% of which are missense and knockout-type mutations, respectively. Non sense and mutations in dinucleotide sequences of splicing donor/acceptor sites are regarded as KO-type.

munity since 2002 (www.brc.riken.go.jp/lab/mutants/genedriven.htm). In more than 300 requested target genes, we have so far been successfully discovered multiple mutant alleles for many genes. For instance, we found eleven ENU-induced mutations by screening the entire Nat1 gene; 5, 1 and 5 of which were missense, nonsense and synonymous mutations, respectively (60). One of the missense mutations, Ile95Thr, of the Nat1 gene was revived as mice and indeed exhibited a drastic reduction of its enzymatic activity (60). The remaining 10 alleles discovered in Nat1 have not been analyzed yet.

The discovery of multiple mutations is very common. For instance, we have so far partially identified: 6 missense, 1 KO-type and 2 synonymous mutations in both the β-catenin and Sufu genes; 5 missense, 1 KO-type and 2 synonymous mutations in the Smo gene; and 2 missense, 1 KO-type and 1 synonymous mutations in the Slc39a6 gene. We have so far revived 11 mutations out of the total of 23 missense and KO-type alleles found in the four genes. The phenotype analyses are currently undertaken for the 11 revived mutant strains and 7 strains have already exhibited some phenotypic anomalies.

Other examples are depression and schizophrenia models established in the Disc1 missense mutations of Gln31Leu and Leu100Pro, respectively, by Clapcote *et al*. (61). Some functional mutations in noncoding sequences have also been established. ENU-induced mutations in a highly conserved noncoding sequence, called MFCS1, affected the expression of the Shh gene that located ~1 Mb apart from the site of the mutations, exhibiting morphological anomalies in the limb/digit development (62, 63).

The RIKEN mutant mouse library is, therefore, a good open resource to the research community as a next-generation gene-targeting system. Currently, total of 40,000 G1 mice including RIKEN library have already been archived in the world; thus, the entire resource is potentially four-fold larger (27). In near future, when the $1,000 resequencing technology become available, it should be able to determine all the ENU induced mutations, or at least all the mutations in the coding sequences in the RIKEN and the other libraries. Then, users will be able to obtain multiple mutant alleles in any mouse genes for the functional annotation of the gene and genome without any screening or PCR primer designing at all.

Acknowledgements

This work is partly supported by a Grant-in-Aid for Scientific Research (A) to Y. G. (KAKENHI 21240043), for Grants-in-Aid for Young Scientists (B) to R. F. (KAKENHI 20790196) and S. M. (KAKENHI 21700454) from Japan Ministry of Education, Science, Sports and Culture (MEXT) and Japan Society for the Promotion of Science (JSPS). Y. G. is also a recipient of Research Grants of the Princess Takamatsu Cancer Research Fund, the Kato Memorial Bioresearch Foundation and the Fukuoka Cancer Society.

REFERENCES

1. Fearon, E. R. (1991) A genetic basis for the multi-step pathway of colorectal tumorigenesis. *Princess Takamatsu Symp.* **22**, 37-48.
2. Wu, X. and Pandolfi, P. P. (2001) Mouse models for muti-step tumorigenesis. *Trends Cell Biol.* **11**, S2-9.
3. Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C. A. Jr., Butel, J. S. and Bradley, A. (1992) Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215-221.
4. Gondo, Y., Nakamura, K., Nakao, K., Sasaoka, T., Ito, K., Kimura, M. and Katsuki, M. (1994) Gene replacement of the p53 gene with the lacZ gene in mouse embryonic stem cells and mice by using two steps of homologous recombination. *Biochem. Biophys. Res. Commun.* **202**, 830-837.
5. Toleco, F. and Wahl, G. M. (2006) Regulating the p53 pathway: *in vitro* hypothesis, *in vivo* veritas. *Nat. Rev. Cancer.* **6**. 909-923.
6. Little, C. C. and Bagg, H. J. (1923) The occurrence of two heritable types of abnormality among the descendants of X-rayed mice. *Amer. J. Roentgenol. Radiat. Therap.* **10**, 975-989.
7. Muller, H. J. (1927) Artificial transmutation of the gene. *Science* **66**, 84-87.
8. Monaco, A. P., Bertelson, C. J., Middlesworth, W., Colletti, C. A., Aldridge, J., Fischbeck, K. H., Bartlett, R., Pericak-Vance, M. A., Roses, A. D. and Kunkel, L. M. (1985) Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. *Nature* **316**, 842-845.
9. Friend, S. H., Bernards, R., Rogelj, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M. and Dryja, T. P. (1986) A human DNA segment with properties of the gene that predis-

poses to retinoblastoma and osteosarcoma. *Nature* **323**, 643-646.
10. The Hungtington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosome. *Cell* **72**, 971-983.
11. Thomas, K. R., Folger, K. R. and Capecchi, M. R. (1986) High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419-428.
12. Doetschman, T., Maeda, N. and Smithies, O. (1988) Targeted mutation of the Hprt gene in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8583-8587.
13. Evans, M. J. and Kaufman, M. H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
14. Robertson, E., Bradley, A., Kuehn, M. and Evans, M. (1986) Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* **323**, 445-448.
15. Watson, J. D. (1990) The Human Genome Project: past, present and future. *Science* **248**, 44-49.
16. Canter, C. R. (1990) Orchestrating the human genome project. *Science* **248**, 49-51.
17. Hrabé de Angelis, M. and Balling, R. (1998) Large scale ENU screens in the mouse: genetics meets genomics. *Mutat. Res.* **400**, 25-32.
18. Hrabé de Angelis M. H., Flaswinkel, H., Fuchs, H., Rathkolb, B., Soewarto, D., Marschall, S., Heffner, S., Pargent, W., Wuensch, K., Jung, M., Reis, A., Richter, T., Alessandrini, F., Jakob, T., Fuchs, E., Kolbm, H., Kremmer, E., Schaeble, K., Rollinski, B., Roscher, A., Peters, C., Meitinger, T., Strom, T., Steckler, T., Holsboer, F., Klopstock, T., Gekeler, F., Schindewolf, C., Jung, T., Avraham, K., Behrendt, H., Ring, J., Zimmer, A., Schughart, K., Pfeffer, K., Wolf, E. and Balling, R. (2000) Genome-wide, large scale production of mutant mice by ENU mutagenesis. *Nature Genetics* **25**, 444-447.
19. Brown, S. D. M. and Nolan, P. M. (1998) Mouse mutagenesis-systematic studies of mammalian gene function. *Human Molec. Genet.* **7**, 1627-1633.
20. Nolan, P. M., Peters, J., Strivens, M., Rogers, D., Hagan, J., Spurr, N., Gray, I. C., Vizor, L., Brooker, D., Whitehill, E., Washbourne, R., Hough, T., Greenaway, S., Hewitt, M., Liu, X., McCormack, S., Pickford, K., Selley, R., Wells, C., Tymowska-Lalanne, Z., Roby, P., Glenister, P., Thornton, C., Thaung, C., Stevenson, J. A., Arkell, R., Mburu, P., Hardisty, R., Kiernan, A., Erven, A., Steel, K. P., Voegeling, S., Guenet, J. L., Nickols, C., Sadri, R., Nasse, M., Isaacs, A., Davies, K., Browne, M., Fisher, E. M., Martin, J., Rastan, S., Brown, S. D. and Hunter, J. (2000) A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nature Genetics* **25**, 440-443.
21. 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. U.S.A.* **76**, 5818-5819.
22. Russell, W. L., Hunsicker, P. R., Raymer, G. D., Steele, M. H., Stelzner, K. F. and Thompson, H. M. (1982) Dose-response curve for ethylnitrosourea-induced specific-locus mutations in mouse spermatogonia. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3589-3591.
23. Russell, W. L., Hunsicker, P. R., Carpenter, D. A., Cornett, C. V. and Guinn, G. M. (1982) Effect of dose fractionation on the ethylnitrosourea induction of specific-locus mutations in mouse spermatogonia. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3592-3593.
24. 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. U.S.A.* **82**, 6619-6621.
25. Justice, M. J., Noveroske, J. K., Weber, J. S., Zheng, B. and Bradley, A. (1999) Mouse ENU mutagenesis. *Human Mol. Genet.* **8**, 1955-1963
26. Noveroske, J. K., Weber, J. S. and Justice, M. J. (2000) The mutagenic action of N-ethyl-N-nitrosourea in the mouse. *Mamm. Genome* **11**, 478-483.
27. Gondo, Y. (2008) Trends in large-scale mouse mutagenesis: from genetics to functional genomics. *Nature Reviews Genetics* **9**: 803-810
28. Beckers, J., Wurst, W. and Hrabé de Angelis, M. (2009) Towards better mouse models: enhanced genotypes, systemic phenotyping and envirotype modeling. *Nature Reviews Genetics* **10**, 371-380.
29. Austin, C. P., Battey, J. F., Bradley, A., Bucan, M., Capecchi, M., Collins, F. S., Dove, W. F., Duyk, G., Dymecki, S., Eppig, J. T., Grieder, F. B., Heintz, N., Hicks, G., Insel, T. R., Joyner, A., Koller, B. H., Lloyd, K. C., Magnuson, T., Moore, M. W., Nagy, A., Pollock, J. D., Roses, A. D., Sands, A. T., Seed, B., Skarnes, W. C., Snoddy, J., Soriano, P., Stewart, D. J., Stewart, F., Stillman, B., Varmus, H., Varticovski, L., Verma, I. M., Vogt, T. F., von Melchner, H., Witkowski, J., Woychik, R. P., Wurst, W., Yancopoulos, G. D., Young, S. G. and Zambrowicz, B. (2004) The knockout mouse project. *Nat. Genet.* **36**, 921-924.
30. Auwerx, J., Avner, P., Baldock, R., Ballabio, A., Balling, R., Barbacid, M., Berns, A., Bradley, A., Brown, S., Carmeliet, P., Chambon, P., Cox, R., Davidson, D., Davies, K., Duboule, D., Forejt, J., Granucci, F., Hastie, N., de Angelis, M. H., Jackson, I., Kioussis, D., Kollias, G., Lathrop, M., Lendahl, U., Malumbres, M., von Melchner, H., Müller, W., Partanen, J., Ricciardi-Castagnoli, P., Rigby, P., Rosen, B., Rosenthal, N., Skarnes, B., Stewart, A. F., Thornton, J., Tocchini-Valentini, G., Wagner, E., Wahli, W. and Wurst, W. (2004) The European dimension for the mouse genome mutagenesis program. *Nat. Genet.* **36**, 925-927.
31. International Mouse Knockout Consortium. (2007) A mouse for all reasons. *Cell* **128**, 9-13.
32. Editorial. (2007) Mutant mice galore. *Nature* **446**, 469-470.
33. Panel Discussion. (2005) In the 19th International Mammalian Genome Conference (IMGC), Strasbourg, France.
34. Whittingham, D. G., Leibo, S. P. and Mazur, P. (1972) Survival of mouse embryos frozen to -196° and -269°C. *Science* **178**, 411-414.
35. Nakagata, N. (1993) Production of normal young following transfer of mouse embryos obtained by *in vitro* fertilization between cryopreserved gametes. *J. Reprod. Fertil.* **99**, 77-80.
36. Nakagata, N. (2000) Cryopreservation of mouse spermatozoa. *Mamm. Genome* **11**, 572-576.

37. Askew, G. R., Doetschman, T. and Lingrel, J. B. (1993) Site-directed point mutations in embryonic stem cells: a gene-targeting tag-and-exchange strategy. *Mol. Cell Biol.* **13**, 4115-4124.
38. Stacey, A., Schnieke, A., McWhir, J., Cooper, J., Colman, A. and Melton, D. W. (1994) Use of double-replacement gene targeting to replace the murine alpha-lactalbumin gene with its human counterpart in embryonic stem cells and mice. *Mol. Cell Biol.* **14**, 1009-1016.
39. Wu, H., Liu, X. and Jaenisch, R. (1994) Double replacement: strategy for efficient introduction of subtle mutations into the murine Col1a-1 gene by homologous recombination in embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **91**, 2819-2823.
40. Gondo, Y., Nakamura, K., Nakao, K., Sasaoka, T., Ito, K., Kimura, M. and Katsuki, M. (1994) Gene replacement of the p53 gene with the lacZ gene in mouse embryonic stem cells and mice by using two steps of homologous recombination. *Biochem. Biophys. Res. Commun.* **202**, 830-837.
41. Gross, E., Arnold, N., Goette, J., Schwarz-Boeger, U. and Kiechle, M. (1999) A comparison of BRCA1 mutation analysis by direct sequencing, SSCP and DHPLC. *Hum. Genet.* **105**, 72-78.
42. Xiao, W. and Oefner, P. J. (2001) Denaturing high-performance liquid chromatography: a review. *Hum. Mutat.* **17**, 439-474.
43. Gao, Q. and Yeung, E. S. (2000) High-throughput detection of unknown mutations by using multiplexed capillary electrophoresis with poly (vinylpyrrolidone) solution. *Anal. Chem.* **72**, 2499-2506.
44. Li, Q., Liu, Z., Monroe, H. and Culiat, C. T. (2002) Integrated platform for detection of DNA sequence variants using capillary array electrophoresis. *Electrophoresis* **23**, 1499-1511.
45. Murphy, K., Hafez, M., Philips, J., Yarnell, K. Gutshall, K. and Berg, K. (2003) Evaluation of temperature gradient capillary electrophoresis for detection of the factor v leiden mutation: coincident identification of a novel polymorphism in factor v. *Mol. Diagn.* **7**, 35-40.
46. Oleykowski C. A., Bronson Mullins, C. R., Godwin, A. K. and Yeung, A. T. (1998) Mutation detection using a novel plant endonuclease. *Nucleic. Acids Res.* **26**, 4597-4602.
47. Till, B. J. , Reynolds, S. H., Greene, E. A., Codomo, C. A., Enns, L. C., Johnson, J. E., Burtner, C., Odden, A. R., Young, K., Taylor, N. E., Henikoff, J. G., Comai, L. and Henikoff, S. (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res.* **13**, 524-530.
48. Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G. and Pryor, R. J. (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin. Chem.* **49**, 853-860.
49. Bennett, C. D., Campbell, M. N., Cook, C. J., Eyre, D. J., Nay, L. M., Nielsen, D. R., Rasmussen, R. P. and Bernard, P. S. (2003) The LightTyper: high-throughput genotyping using fluorescent melting curve analysis. *Biotechniques* **34**, 1288-1295.
50. Beier, D. R. (2000) Sequence-based analysis of mutagenized mice. *Mamm. Genome* **11**, 594-597.
51. Gondo, Y., Minowa, O., Sezutsu, H., Sakuraba, Y., Toki, H., Wakana, S., Noda, T. and Shiroishi, T. (2001) *ENU mouse mutagenesis project in RIKEN GSC: archives for versatile screens,* The 15th International Mouse Genome Conference, Edinburgh, UK.
52. Coghill, E. L., Hugill, A., Parkinson, N., Davison, C., Glenister, P., Clements, S., Hunter, J., Cox, R. D. and Brown, S. D. (2002) A gene-driven approach to the identification of ENU mutants in the mouse. *Nat. Genet.* **30**, 255-256.
53. Gondo, Y., Wakana, S., Noda, T. and Shiroishi, T. (2002) *Approaches for the gene indentification: gene-driven mutagenesis and recessive mutations,* Workshop on ENU Mutagenesis: Planning for Saturation. pp.25 Chantilly, USA.
54. Concepcion, D., Seburn, K. L., Wen, G., Frankel, W. N. and Hamilton, B. A. (2004) Mutation rate and predicted phenotypic target sizes in ethylnitrosourea-treated mice. *Genetics* **168**, 953-959.
55. Quwailid, M. M., Hugill, A., Dear, N., Vizor, L., Wells, S., Horner, E., Fuller, S., Weedon, J., McMath, H., Woodman, P., Edwards, D. Campbell, D., Rodger, S., Carey, J., Roberts, A., Glenister, P., Lalanne, Z., Parkinson, N., Coghill, E. L., McKeone, R., Cox, S., Willan, J., Greenfield, A., Keays, D., Brady, S., Spurr, N., Gray, I., Hunter, J., Brown, S. D. and Cox, R. D. (2004) A gene-driven ENU-based approach to generating an allelic series in any gene. *Mamm. Genome* **15**, 585-591.
56. Augustin, M., Sedlmeier, R., Peters, T., Huffstadt, U., Kochmann, E., Simon, D., Schöniger, M., Garke-Mayerthaler, S., Laufs, J., Mayhaus, M., Franke, S., Klose, M., Graupner, A., Kurzmann, M., Zinser, C., Wolf, A., Voelkel, M., Kellner, M., Kilian, M., Seelig, S., Koppius, A., Teubner, A., Korthaus, D., Nehls, M. and Wattler, S. (2004) Efficient and fast targeted production of murine models based on ENU mutagenesis. *Mamm. Genome* **16**, 405-413.
57. Sakuraba, Y., Sezutsu, H., Takahasi, K. R., Tsuchihashi, K., Ichikawa, R., Fujimoto, N., Kaneko, S., Nakai, Y., Uchiyama, M., Goda, N., Motoi, R., Ikeda, A., Karashima, Y., Inoue, M., Kaneda, H., Masuya, H., Minowa, O., Noguchi, H., Toyoda, A., Sakaki, Y., Wakana, S., Noda, T., Shiroishi, T. and Gondo Y. (2005) Molecular characterization of ENU mouse mutagenesis and archives. *Biochem. Biophys. Res. Commun.* **336**, 609-616.
58. Sakuraba, Y., Kimura, T., Masuya, H., Noguchi, H., Sezutsu, H., Takahasi, K. R., Toyoda, A., Fukumura, R., Murata, T., Sakaki, Y., Yamamura, M., Wakana, S., Noda, T., Shiroishi, T. and Gondo, Y. (2008) Identification and characterization of new long conserved noncoding sequences in vertebrates. *Mamm. Genome* **19**, 703-712.
59. Takahasi, K. R., Sakuraba, Y. and Gondo, Y. (2007) Mutational pattern and frequency of induced nucleotide changes in mouse ENU mutagenesis. *BMC Molec. Biol.* **8**, 52.
60. Erickson, R. P., McQueen, C. A., Chau, B., Gokhale, V., Uchiyama, M., Toyoda, A., Ejima, F., Maho, N., Sakaki, Y. and Gondo Y. (2008) An N-ethyl-N-nitrosourea-induced mutation in N-acetyltransferase 1 in mice. *Biochem. Biophys. Res. Commun.* **370**, 285-288.
61. Clapcote, S. J., Lipina, T. V., Millar, J. K., Mackie, S., Christie, S., Ogawa, F., Lerch, J. P., Trimble, K., Uchiyama, M., Sakuraba, Y., Kaneda, H., Shiroishi, T., Houslay, M. D., Henkelman, R. M., Sled, J. G., Gondo, Y., Porteous, D. J. and Roder, J. C. (2007) Behavioral phenotypes of Disc1 mis-

sense mutations in mice. *Neuron* **54**, 387-402.
62. Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T. and Shiroishi, T. (2004) Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mamm. Genome* **15**, 23-34.
63. Masuya, H., Sezutsu, H., Sakuraba, Y., Sagai, T., Hosoya, M., Kaneda, H., Miura, I., Kobayashi, K., Sumiyama, K., Shimizu, A., Nagano, J., Yokoyama, H., Kaneko, S., Sakurai, N., Okagaki, Y., Noda, T., Wakana, S., Gondo, Y. and Shiroishi, T. (2007) A series of ENU-induced single-base substitutions in a long-range cis-element altering Sonic hedgehog expression in the developing mouse limb bud. *Genomics* **89**, 207-214.