Published in final edited form as:
Curr Protoc Mouse Biol. ; 2: 103–118. doi:10.1002/9780470942390.mo110200.

Discovery Genetics – The History and Future of Spontaneous Mutation Research

Muriel T. Davisson, **David E. Bergstrom**, **Laura G. Reinholdt**, and **Leah Rae Donahue**
The Jackson Laboratory, Bar Harbor, Maine, U.S.A

Muriel T. Davisson: dave.bergstrom@jax.org; Laura G. Reinholdt: laura.reinholdt@jax.org; Leah Rae Donahue: leahrae.donahue@jax.org

Abstract

Historically, spontaneous mutations in mice have served as valuable models of heritable human diseases, contributing substantially to our understanding of both disease mechanisms and basic biological pathways. While advances in molecular technologies have improved our ability to create mouse models of human disease through targeted mutagenesis and transgenesis, spontaneous mutations continue to provide valuable research tools for discovery of novel genes and functions. In addition, the genetic defects caused by spontaneous mutations are molecularly similar to mutations in the human genome and, therefore often produce phenotypes that more closely resemble those characteristic of human disease than do genetically engineered mutations. Due to the rarity with which spontaneous mutations arise and the animal intensive nature of their genetic analysis, large-scale spontaneous mutation analysis has traditionally been limited to large mammalian genetics institutes. More recently, ENU mutagenesis and new screening methods have increased the rate of mutant strain discovery, and high-throughput DNA sequencing has enabled rapid identification of the underlying genes and their causative mutations. Here, we discuss the continued value of spontaneous mutations for biomedical research.

Keywords

spontaneous mutations; mouse models; forward genetics; biomedical research

Introduction

Spontaneous mutations are detectable as random genetic events in any vivarium housing large colonies of mice. Historically, mice bearing spontaneous mutations have provided a rich source of animal models for human genetic diseases and basic biomedical research. Mutant mice are first recognized by observation of a biomedically relevant phenotype and are genetically analyzed using forward genetic approaches (*i.e.*, with analysis advancing from phenotype to gene). During the latter half of the 20th century, the evolution of traditional approaches and technological advances, most notably the complete sequencing of the mouse genome (Waterston et al. 2002), significantly shortened the time from mutant strain discovery to gene identification. Despite these advances, the concurrent development

Contributing Author: Muriel T. Davisson, Telephone: 207-288-6223, Fax: 207-288-6149, muriel.davisson@jax.org

and rapid advancement of gene targeting approaches contributed to a perception that spontaneous mutations had been superseded in their utility by engineered alleles, a perception that pervaded some segments of the scientific community, including funding agencies. Yet despite the advent of targeted mutagenesis and the fact that we can genetically engineer any mutation in any gene, reverse genetics (*i.e.*, with analysis advancing from gene to phenotype) still relies on *a priori* knowledge of gene location, structure and putative function. While this information exists for many genes, complete structural (exons, introns, UTRs) and functional annotation of the mammalian genome remains a work in progress. In this context, forward genetic approaches continue to reveal unsuspected gene functions and provide valuable models for further investigation. Moreover, with the recent advent of high throughput sequencing, it is now possible to identify the causative mutant gene in spontaneous mutant models at an unprecedented rate. Here we provide a rationale for the on-going utility of spontaneous mutations as valuable resources for gene discovery, briefly describe the history of developing spontaneous mutation mouse models, and provide an overview of current protocols for rapid forward genetic analysis and mutation identification.

Advantages of Spontaneous Mutation Analysis

Since the early 1900s, mice bearing spontaneous mutations have been a rich source of animal models of human genetic diseases for use in biomedical research. In the 21st century, they continue to provide important research tools that complement genetically engineered mutations. Their analysis can reveal novel functions for known genes, can identify genes underlying orthologous human diseases (where the causative gene has not yet been identified) and can immediately provide investigative insights into biomedically relevant phenotypes.

Phenotype-driven mutation detection—Spontaneous mutations manifest as, and are thus already known to cause, a biomedically relevant phenotype. The phenotype immediately provides information about the physiological function of the mutant gene and its biomedical relevance. Thus, the analysis of spontaneous mutations can reveal genes not yet suspected to underlie human diseases. For example, Seymour et al. showed that the mouse chronic proliferative dermatitis (*cpdm*) mutation causes a disease that closely mimics human hypereosinophilic syndromes, a rare, heterogeneous group of hematological and systemic diseases characterized by unexplained blood and tissue eosinophilia and dermatologic lesions. Identification of a causative mutation in the mouse gene *Sharpin* (SHANK-associated RH domain interacting protein) in *cpdm/cpdm* mutants predicts that mutations in the *Sharpin* gene may cause one of the human disease syndromes (Seymour et al. 2007). Similarly, cloning of the hurry-scurry *(hscy*) mutation led to discovery of the causative gene for human autosomal recessive deafness disorder DFNB67 (OMIM#610265) (Longo-Guess et al. 2005) and, more recently, identification of the mouse pirouette (*pi*) mutation led to discovery of the causative gene for human autosomal recessive DFNB25 (OMIM#613285) (Odeh et al. 2010).

Moreover, spontaneous mutants with phenotypes that don't directly mimic clinical symptoms of human disease can provide models in unexpected ways. For example, Donahue et al. showed that the disheveled hair and [small] ears mutation is a mutation in lamin A

(*Lmna*^{Dhe}), an intermediate filament protein of the nuclear membrane. Detailed analysis of the mouse phenotype revealed that it is not only a good model for craniofacial development, but also a useful model of human laminopathies and for progeria, specifically Hutchinson-Gilford Progeria Syndrome, OMIM#176670 (Odgren et al. 2010).

In addition to facilitating gene discovery and functional gene annotation, mouse strains carrying spontaneous mutations also can provide platforms for development of drug therapies. Alleviation of the mouse condition caused by mutations in the *Sharpin* gene (described above) by treatment with a proteasome inhibitor (Liang et al. 2011) suggests a potential therapeutic approach for related human diseases. Similarly, spontaneous models of human disease have been used for the development of gene therapy approaches. For example at The Jackson Laboratory, Chang and colleagues have developed a vision screening program using electroretinography that has led to the identification of the causative genes for many ocular diseases (one of which, *Rd3*, is a novel gene) and generated mouse models for three forms of human Leber congenital amaurosis (LCA2, LCA10, LCA12), as well as several models for various forms of human retinitis pigmentosa. These mouse models have been used successfully to develop gene therapy approaches for treating ocular disease (Pang et al. 2010, Li et al. 2011, Cideciyan et al. 2011). Table 1 gives details for these and a few other mouse models with biomedically relevant phenotypes.

Discovery of novel gene functions—The analysis of spontaneous mutations can lead to the discovery of new roles for known genes, often implicating biological processes not previously known or suspected to be involved in disease. For example, the sticky mutant, first identified by its abnormal fur, was later shown to display cerebellar Purkinje cell loss and ataxia, resulting from a missense mutation in the editing domain of the alanyl-tRNA synthetase gene (*Aars*^{sti}). This missense mutation compromises the proofreading activity of the enzyme during aminoacylation of tRNAs. Analysis of sticky mutant mice provided the first demonstration that low levels of mischarged transfer RNAs (tRNAs) can lead to an intracellular accumulation of misfolded proteins in specific neurons and cell death, identifying a novel mechanism underlying neurodegeneration (Lee et al. 2006). Analysis of woozy (*wz*) mutant mice also showed that protein accumulation can lead to neuronal death, in this case due to mutation in the gene for endoplasmic reticulum chaperone SIL1 homolog (S. cerevisiae) (*Sil1*) encoding a crucial endoplasmic reticulum (ER) chaperone. Studies of this mutant provided evidence that perturbation of ER chaperone function in terminally differentiated neurons leads to protein accumulation, ER stress and subsequent neurodegeneration (Zhao et al. 2005).

In another example, analysis of the iris in mice with several coat color variants revealed an unsuspected role for pigmentation genes in the etiology of a poorly understood set of ocular disorders –pigment dispersion syndromes (PDS). PDS are a major cause of angle-closure glaucoma, in which pigment released from the iris blocks drainage from the eye leading to increased intraocular pressure. Analysis of spontaneous coat color mutations has shown that mutant mice with alleles that affect melanosomes provide a rich source of models for PDS and this class of glaucomas (Anderson et al. 2008).

Pigment mutations also have identified over 15 mouse models for Hermansky-Pudlak Syndrome, an autosomal recessive disease characterized by pigment dilution and prolonged bleeding time. Research on these mouse models showed that HPS is caused by inhibition of intracellular organelle biogenesis or transfer (Nguyen et al. 2002) and continuing research is revealing the interacting genes involved in organelle biogenesis (Ciciotte et al. 2003, Gwynn et al. 2000, Gwynn et al. 2004). Table 2 lists selected examples of spontaneous mouse mutations that have revealed unsuspected functions for known genes.

Unbiased diversity of mutations and similarities to human disorders— Spontaneous mutations, because they are naturally occurring, produce a full and unbiased array of mutation types – single nucleotide variants (SNVs), small insertions or deletions (<100 bp, INDELs) and structural variants (including large insertions or deletions, repeat sequence expansions and chromosomal rearrangements) – similar to the variety of mutations underlying human genetic disorders. Thus, spontaneous mutations can produce a range of effects on protein function, more closely resembling those found in human genetic disease than do genetically engineered knockout mutations for the same gene. Spontaneous mutations can cause over-expression (hypermorphic), under-expression (hypomorphic) or alteration of function (dominant negative or gain-of-function) as well as the complete loss of function (null). For example, two spontaneous mouse mutations of the homeobox D13 gene (synpolydactyly homolog, *Hoxd13*^{spdh} and digit in Y-shaped finger and carpe, *Hoxd13*^{Dyc}) are in-frame duplications that cause expansion of a polyalanine repeat as in human synpolydactyly (SPD) (Bruneau et al. 2001, Cocquempot et al. 2009, Johnson et al. 1998). The resulting phenotypes closely resemble those of the human condition. Table 3 gives some additional examples of varying types of spontaneous mutations and the phenotypes they confer, compared to targeted mutations of the same genes.

Allelic Series: Multiple novel mutations in previously characterized genes create allelic series, and analysis of their associated phenotypic variation can provide important insights into domain-specific protein function, alternative splicing and modifier genes. The *Kit* oncogene provides a classic example of how an allelic series can enable gene function analysis by genotype-to-phenotype correlation. Mouse Genome Informatics (MGI) documents 136 spontaneous, induced and targeted mutations in the *Kit* oncogene. Phenotypes in heterozygotes with different alleles range from a white belly spot to variable degrees of whole body spotting, male sterility and anemia. Phenotypes in homozygotes range from preimplantation lethality to postnatal survival into adulthood as black-eyed white mice. Severity of the phenotype is correlated with intragenic location or type of mutation and the corresponding protein domain.

Similarly, multiple mutations in low-density lipoprotein receptor-related protein 4 (*Lrp4*) (see Table 2) provide another example of an allelic series that enables gene function analysis. The severely hypomorphic alleles *dan* (digitation anormale), on 129S2/SvPas, and *mdig* (malformed digits), on DBA/2J, cause variable brachydactyly, syndactyly and polysyndactyly, abnormally shaped nails and failure to nurse (Simon-Chazottes et al. 2006). Two chemically induced null mutations, mitaine (*mte*) and mitten (*mitt*), both on C57BL/6, cause embryonic lethality and, in addition to missing, abnormal and fused digits and

abnormal mammary buds, mutants exhibit small, nonfunctional lungs, kidney agenesis and abnormal embryonic neuromuscular synapses (Weatherbee et al. 2006). While some or all of these differences might be attributed to differences in genetic background, two targeted mutations on a 129S background similar to *dan* show only supernumerary and grooved incisors while *dan* produces digit defects but no tooth abnormalities on 129 (Weatherbee et al. 2006).

Genetic heterogeneity: Genetically heterogeneous disorders, where mutations in different genes give similar phenotypes, can reveal pathways or interacting genes. An example is the important role spontaneous mutation analysis has played in uncovering gene function and points of interaction in the mammalian vestibular system. Paffenholz et al. have shown that in head tilt (*het*) mice, a vestibular mutant lacking functional balance end organs, the causative gene underlying the mutant phenotype is *Nox3*, an NADPH oxidase with a previously unsuspected role in the inner ear (Paffenholz et al. 2004). Insight from the *Nox3* discovery led to the rapid identification and characterization of a mutation in *Noxo1*, a gene encoding an NADPH oxidase organizer, as a non-allelic gene causing a nearly identical phenotype in a mutant known as head slant (*hslt*) (Bergstrom et al. 2004, Kiss PJ 2006). Sanger-based sequencing and transgenic rescue confirmed *Noxo1* as the causative gene. A third, ENU-induced mutation known as neuromutagenesis facility 333 (*nmf333*) also confers an identical imbalance phenotype. Once mapped to a subchromosomal region, *nmf333* was quickly identified as a mutation in a unique NADPH oxidase component (cytochrome b-245, alpha polypeptide (*Cyba*) ($p22^{phox}$)) that is shared with NADPH oxidases throughout the body, including the phagocytic oxidase required for proper immune function (Nakano Y 2008). Armed with this fact, *nmf333* mutants were tested for primary immunodeficiency and, as suspected, exhibited a deficiency with striking similarities to the orthologous human condition, cytochrome b-negative, autosomal recessive, chronic granulomatous disease (OMIM #233690). Thus, *nmf333* mutants provide a unique and faithful mouse model for this life-threatening human ailment. Moreover, human *Cyba*-deficient patients can now be assessed for subclinical vestibular deficits. Collectively, the analysis of the head tilt, head slant and *nmf333* mutants has identified a new role for NADPH oxidase in the inner ear, identified three components of the complex, and provided an animal model and translational opportunities for the corresponding human condition.

Multiple strain backgrounds and discovery of modifying genes and pathways—Spontaneous mutations occur on a wide variety of genetically distinct strain backgrounds, whereas induced and genetically engineered mutations are confined to only a few strain backgrounds. Analysis of the phenotypic differences compared to genetic background can lead to a better understanding of heterogeneous human disorders. For example, two novel alleles of the *Phex* (phosphate regulating gene with homologies to endopeptidases on the X chromosome) gene result in phenotypic variability from each other and the original $Phex^{Hyp}$ mutation. The two alleles ($Phex^{Hyp\text{-}2J}$ and $Phex^{Hyp\text{-}Duk}$) are both null mutations, resulting from intragenic deletions of different parts of the gene, but they occurred on different genetic backgrounds, C57BL/6J and BALB/c, respectively. Both cause similar phenotypes of shortened hind legs and tail, shortened square trunk, hypophosphatemia, hypocalcemia and rachitic bone disease. Mice carrying the $Phex^{Hyp\text{-}Duk}$ mutation, however, also exhibit

background-dependent, variably expressed deafness, circling behavior and cranial dysmorphology, demonstrating the influence of modifying genes on *Phex*-related phenotypes. Evidence of degeneration of the organ of Corti and spiral ganglion occurs in the hearing impaired $Phex^{Hyp\text{-}Duk}$/Y mice, but not in the normal hearing $Phex^{Hyp\text{-}2J}$/Y mice (Lorenz-Depiereux et al. 2004). In humans, mutations in the *PHEX* gene cause X-linked hypophosphatemic rickets (XLH). Identification of the genetic modifiers of hearing and craniofacial dysmorphology in $Phex^{Hyp\text{-}Duk}$/Y mice may offer insight into the phenotypic variation of XLH seen in humans.

Phenotypic variation on different genetic backgrounds also enables discovery of modifier genes that suppress or enhance the phenotype resulting from another mutant gene. For example, the dactylaplasia mutation in the F-box and WD-40 domain protein 4 ($Fbxw4^{Dac}$) gene is semi-dominant on the SM/Ckc background on which it arose (Johnson et al. 1995). Homozygotes die prenatally or around the time of birth. Outcrosses to mice of other strain backgrounds revealed a modifier gene (*mdac*) that controls the severity of expression of the phenotype and is polymorphic among inbred strains. Homozygosity for the recessive allele *mdac* conveys the semi-dominant *Dac* phenotype seen in SM/Ckc mice. Presence of the dominant *Mdac* allele ameliorates the severity of the dactylaplasia and allows homozygous *Dac/Dac* mice to survive (Johnson et al. 1995). Likewise the curly bare (*cub*) mutation can cause two different phenotypes depending on the allelic composition of a modifying gene (*mcub*). When homozygous for the recessive allele of *mcub*, *cub/cub* mice are hairless. Presence of at least one copy of the dominant *Mcub* allele changes the phenotype of *cub/cub* mice to a full, curly coat (Johnson et al. 2003).

The analysis of modifier genes or genetic background genes that increase susceptibility to specific traits often reveals digenic or multigenic systems that provide a means for identifying interacting genes and molecular pathways. For example, mice of the A/- strains are more susceptible to cleft lip and cleft palate than mice of other strains. Homozygosity for mutant alleles of two unlinked genes, cleft lip 1 (*clf1*), which has been identified as a mutation in the wingless-type MMTV integration site 9B (*Wnt9b*) gene (Juriloff et al. 2006), and cleft lip 2 (*clf2*) confer this increased susceptibility (Juriloff et al. 2004). The two genes interact epistatically with each other and with environmental factors such as maternal effects (Plamondon et al. 2011).

Some modifier genes can totally suppress the phenotypic effect of a major mutant gene. There are many examples of important mutant phenotypes that would not have been discovered on the standard C57BL/6 or 129 strain backgrounds typically used for mutagenesis and genetic engineering in mice for this reason. For example, the severe cardiomyopathy caused by the thrombocytopenia and cardiomyopathy mutation ($Abcg5^{trac}$) mutation on the A/J background, where the mutation was discovered, disappears when the mutation is placed on the C57BL/6J background (Chase et al. 2010). Identification of suppressive modifier alleles can provide insight into therapeutic approaches for protecting individuals against disease.

Thus, genetic disorders originally identified as monogenic may turn out to be digenic or multigenic on further genetic analysis. Analysis of such systems in mouse models has great

potential to reveal genes modulating the severity of clinical symptoms in orthologous human disorders (Nadeau 2003, Romeo and McKusick 1994). Table 4 lists a few specific examples of phenotypic variability depending on modifier genes and genetic background.

Cost-effective source of new mutations—Last but not least, because spontaneous mutations are the byproducts of normal colony maintenance, their discovery is free. Astute technicians should watch for deviants in their colonies, not only to discover new spontaneous mutation models but also for genetic quality control to “clear” their colonies of such mutations. Although naturally occurring mutations cannot be planned, the alert animal care or research technician can discover them by careful observation during routine operations and the astute scientist can rationally exploit such chance observations to create new research opportunities from something not specifically sought. Now that analysis and gene identification can be done so rapidly, the cost from discovery to gene identification is often less than the cost of genetically engineering a mutation.

Challenges of Spontaneous Mutation Analysis

Spontaneous mutations are apparently random and occur at a low mutation rate in most genes. Estimates of mutation rates in mammals vary widely depending on the ascertainment method, the types of genes assessed and the calculation method. To our knowledge, the only estimate for spontaneous mutations causing visible phenotypic deviants (such as small size, neuromuscular disorders, coat color, coat texture, skeletal anomalies, behavioral abnormalities, etc.) was based on observation of over 7 million mice in Jackson Laboratory production mouse colonies in the 1960s (Schlager and Dickie 1971). The average mutation rate from wild type alleles for specific, mostly coat color loci was estimated at 11×10^{-6} per locus per gamete per generation, with different rates for different genes. The rate from recessive alleles to normal (reversions) was almost 6-fold lower (2.5×10^{-6} per locus per gamete per generation). In large production breeding colonies as at The Jackson Laboratory and other mouse resource centers, however, even a low mutation rate produces a large number of mutations. Only those that produce a visible phenotype will be detected. This is both a challenge and an advantage. Subtle phenotypes or those whose detection requires specialized screening will not be found by standard, day to day colony observation. Yet, as noted earlier, the presence of an obvious phenotype frequently identifies mutants that will be valuable for biomedical research. Discovery of disorders with subtle phenotypes that require screening for detection can be made economically feasible by increasing the mutation rate using chemical mutagenesis. Ethylnitrosourea (ENU) is the most effective (Russell et al. 1979, Probst and Justice 2010) chemical mutagen used in vivo. Chemically induced mutations, which are virtually always single base pair changes, provide many of the same advantages as spontaneous mutations. During the 1990s, NIH funded several ENU mutagenesis programs. Unfortunately, many of these programs lost funding prematurely, perhaps because of the large number of deviant mice produced and the time required for genetic analysis. Since then many new ENU-induced mouse models have been published that originated in such programs (Justice 2000, Boles et al. 2009). Moreover, rapid advancement of high throughput sequencing technologies over the last decade has resulted in a paradigm shift in the process of mutation identification. The result has been an upsurge

in interest, not only in ENU mutagenesis, but all forward genetic approaches as practical means of functional genome annotation.

History of Spontaneous Mutation Discovery and Development

The three largest sources of spontaneous (and induced) mutations in the 20th century were the Medical Research Council group at Harwell (U.K.), the Oak Ridge National Laboratory mouse biology program in Oak Ridge, Tennessee (USA) and The Jackson Laboratory in Bar Harbor, Maine (USA). At each of these institutions the juxtaposition of large mouse breeding colonies maintained by genetically defined breeding protocols and faculties of mammalian geneticists provided a unique combination of resources for the discovery and analysis of spontaneous mutations of biomedical significance. Both Harwell and Oak Ridge had large radiation and chemical risk assessment programs after World War II. In an effort to assess the mutagenicity of these agents, these programs involved the production of large cohorts of progeny from mice exposed to these agents that were examined carefully for anomalies. The Jackson Laboratory has historically maintained large production colonies of mice for the purpose of distribution throughout the scientific community. The identification of phenotypic deviants is an essential component of a genetic quality assurance program that functions to remove “deviant” mice with their parents and siblings from the breeding colonies. Such “phenotypic deviant” breeding units are transferred to Jackson Laboratory researchers and to external scientists through the Mouse Mutant Resource (MMR), which has an established program for analyzing and developing spontaneous mutation mouse models. The MMR at The Jackson Laboratory is an excellent example of the success of developing mouse disease model strains from single, rare phenotypic deviants bearing spontaneous mutations. The MMR was first funded by the National Science Foundation in 1959. Since the beginning of NIH funding in 1978, the MMR has developed approximately 700 new spontaneous mutation strains that have been used throughout the research community for disease modeling and basic biomedical research. A few of the most widely used strains include Snell's dwarf (*Pou1f1*^{dw}), obese (*Lep*^{ob}), diabetic (*Lepr*^{db}) and the mouse model of Duchenne muscular dystrophy (*Dmd*^{mdx}).

Protocols for Genetic Analysis – then and Now

The basic approach to analyzing spontaneous mutations remained the same over the last half of the 20th century with incremental advances leading to enhanced efficiency in the analysis of new phenotypic deviants and a reduction in the time required for causative gene identification. For example, chromosomal markers for genetically mapping new mutations have evolved from mutant genes detected by visible phenotypes, where only a few chromosomes could be tested in a single linkage cross, to polymorphic co-dominant markers - first biochemical, then molecular and finally single nucleotide polymorphisms (SNPs) - where virtually all chromosomes can be tested in a single cross. The complete sequencing and improved annotation of the mouse genome dramatically increased the number of potential mapping markers, and also significantly reduced the time required for mutant gene identification by providing regional candidate genes. Mutation detection has now transitioned from the positional cloning method, requiring high-resolution mapping, laborious physical mapping efforts (to construct contigs of BAC or YAC clones) and Sanger sequencing of PCR products, to approaches that use low-resolution mapping and high

throughput sequencing technologies (Arnold et al. 2011) (Boles et al. 2009) (D'Ascenzo et al. 2009). It now is possible to go from new mutation discovery to gene identification in under 6 months. In the 10 years prior to 2010, the MMR identified an average of 4-5 causative genes each year. Recently, with the implementation of high throughput sequencing approaches, that rate has tripled. As an example, the current method for genetic analysis of new spontaneous mutations in The Jackson Laboratory Mouse Mutant Resource can be distilled into three areas of effort: heritability analysis, linkage analysis and mutation identification.

Heritability analysis—As the name implies, heritability analysis of new phenotypic deviants is focused on ascertainment of phenotype transmission and mode of inheritance. To accomplish this new deviants are outcrossed to genetically unrelated mice and F1 and F2 progeny are observed for recapitulation of the deviant phenotype. Semi-dominant and dominant phenotypes will recur in F1 mice and recessive traits will recur among F2 progeny. If available, parents and siblings of deviant mice also are bred with the goal of propagating the phenotype (if it proves heritable). Recurrence of the phenotype in progeny from these crosses can also indicate heritability. If the phenotype closely resembles that of a previously identified mutation, it is tested for allelism by breeding mice that each carry one of the mutations (known as an allele or complementation test). With the increased efficiency of molecular linkage mapping strategies, however, it is sometimes more cost effective to establish linkage prior to testing for allelism. In some cases, allelism can be more efficiently assessed by direct sequencing of candidate genes rather than by testing for functional complementation, which may require importation of mutant mice or their reconstitution from cryopreservation. Some so-called “phenodeviant” traits occur at a low frequency in some inbred strains and are not heritable as single gene mutations. For example, white head blazes or belly spots occur in a small proportion of C3H and C57BL/6 mice and microphthalmia occurs in some C57BL/6 mice.

Linkage analysis—If the mode of inheritance is dominant or semi-dominant, linkage is established by a backcross, where affected F1 mice are backcrossed to wild type mice of the parent strain and the resulting backcross mice are used for mapping. If the mode of inheritance is recessive, linkage is established by an intercross, in which F1 animals are intercrossed and the resulting F2 animals are used for mapping. Generally, 12-15 affected and 9-10 unaffected individuals are sufficient to establish chromosomal linkage. Traditionally, wild-derived inbred strains, such as CAST/EiJ from *Mus musculus castaneus*, were used in crosses for establishing linkage and for fine mapping of causative mutations to small chromosomal regions (<5Mb). This approach maximizes the number of informative polymorphic mapping markers. If the strain background selected for the heritability outcross is an inbred strain that is genetically divergent from the strain on which the spontaneous mutation arose, then this cross also can provide linkage data. The use of divergent strain backgrounds in mapping crosses can be problematic, however, when the phenotype of interest is influenced by strain background, as is frequently the case. With the advent of high throughput sequencing and the complete genome sequencing of a large number of inbred strain backgrounds (Keane et al. 2011), the use of such divergent strain backgrounds for the linkage cross is no longer necessary. With the large number of single nucleotide

polymorphisms (SNPs) generated by high throughput discovery efforts (Frazer et al. 2007, Kirby et al. 2010,), more closely related strains can be used to establish linkage in the same cross with a single SNP panel. For example, to take advantage of large litter sizes, the Mouse Mutant Resource at The Jackson Laboratory now uses FVB/NJ or 129S1/SvImJ (depending on the coat color of the deviant) for virtually all linkage crosses. Also, over the last few years, it has become clear that fine mapping is no longer necessary due to the availability and applicability of high throughput sequencing technologies (Arnold et al. 2011, Boles et al. 2009, Zhang et al. 2009). Nevertheless, knowing the approximate chromosomal location still makes sequencing more efficient because one can focus sequence analysis on that chromosome. Also, ENU-induced mutations may require outcrosses because the original mutant may harbor more than one mutation. Once data analysis technology becomes as robust as sequencing technology, whole genome sequencing may obviate the need for linkage mapping (see next section).

Gene identification—The positional cloning approach, which combines knowledge of the map position of the phenotype with increasingly dense genetic maps and genomic sequence annotation, was the predominant method for gene identification until early in the 21st century. Traditionally, fine mapping crosses requiring 100s of progeny were used to map causative mutations to small intervals (< 5 Mb) containing candidate genes that could then be considered for molecular analysis. Of course, the candidate gene search changed dramatically after the mouse genome was sequenced (Waterston et al. 2002) and high quality gene annotation became available (Blake et al. 2011, Wilming et al. 2008, Pruitt et al. 2009, Pruitt et al. 2007). Currently, the gene models that result from these annotation efforts are available through public genome browsers (e.g. UCSC Genome Browser, Ensembl Genome Browser) and provide the basis for PCR amplification and DNA sequencing of individual exons of candidate genes, primarily via the Sanger method. In some cases, especially where causative mutations are not found in coding sequence, candidate genes also are evaluated by assaying gene expression and cDNA sequences using RT-PCR and, if necessary, by Southern and northern blot analyses. In all cases, DNA and RNA from mutant mice are compared with coisogenic littermate controls.

Today, massively parallel sequencing technologies (a.k.a. high throughput sequencing technologies) have become widely available and are rapidly changing the process by which causative mutations are discovered (Turner et al. 2009, Arnold et al. 2011, Boles et al. 2009, D'Ascenzo et al. 2009, Zhang et al. 2009). Indeed, the identification of causative mutations in the mouse genome in the absence of any linkage data is entirely possible through whole genome or whole exome re-sequencing; this approach is routinely used to identify potential causative mutations in human genomes where linkage data are rarely available. In considering the cost benefit of linkage analysis, the availability of resources for establishing linkage (i.e. vivarium space and cost) vs. validation burden (i.e. labor for amplification and sequencing of many candidate mutations across multiple samples) must be weighed carefully. Establishment of linkage significantly reduces the validation burden at minimal cost, assuming vivarium space and mice are available (Arnold et al. 2011). In our experience, linkage to a chromosome reduces validation by at least 20 times in an exome (38 candidate mutations to just one candidate mutation) (Fairfield et al. 2011). If sequencing

capacity is more readily available than vivarium space, sequencing of an unaffected sibling or parent can similarly reduce the validation burden.

One limitation to high throughput sequencing for mutation detection lies in the tools available for data analysis. Currently, the tools available for aligning high throughput sequencing data back to the reference genome and for subsequent SNV (single nucleotide variant) and INDEL (insertion/deletions) calling are robust for the detection of single nucleotide substitutions and small insertions and deletions, where, for example, the size of a detectable deletion depends heavily on the length of the sequencing reads and the capacity of the alignment algorithm to tolerate gaps. The tools for detection of larger insertions or deletions (a.k.a structural variants), molecular events that frequently underlie spontaneous mutants, are not as well developed and are not yet broadly available to wet bench scientists. Moreover, the success of these tools depends on the sequencing approach employed (for example, paired end library construction and the generation of longer reads). If high throughput sequencing doesn't reveal a likely candidate mutation and a large insertion or deletion is suspected, comparative genome hybridization is a robust alternative approach for the detection of structural variants. Regardless, while the rate at which sequence data can be generated at low cost still outpaces our ability to analyze the data, high throughput sequence analysis is a rapidly growing field and new algorithms for detection of structural variants are increasingly available in the scientific literature (Wong et al. 2010). Therefore, full implementation of high throughput sequencing for the detection of spontaneous mutations of all types is imminent and with it, significant diminution of the gap between phenotype and genotype.

Summary and Conclusions

In this paper, we have shown that spontaneous mutations in mice have provided, and continue to provide, an invaluable resource for analyzing both disease mechanisms and basic biological systems, and have provided important animal models of human genetic disease. Furthermore, in the 21st century, improved methods for the forward genetic analysis of spontaneous mutations will greatly enhance functional genome analysis by increasing the discovery rate of novel genes and unexpected gene functions. Spontaneous mutations also continue to provide valuable models of human disorders. The examples we have given in this paper are only a few of the many spontaneous mutations that have contributed to important biomedical discoveries since the first decades of the 20th century. Until the role of every gene or chromosomal sequence in the human and mouse genomes is known, spontaneous mutations will continue to contribute to understanding gene function.

Acknowledgments

We acknowledge funding by NIH/NCRR (P40RR001183), which supported research on many of the mutants described here. Additional funding came from EY015073 (Donahue), DC007431 (Bergstrom) and DK064674 (Davisson). We are grateful for the more than 50 years of federal funding by NSF and NIH/NCRR for the discovery and characterization of spontaneous mutants. We thank past and current members of the Mouse Mutant Resource (MMR) who have worked diligently over the last 50 years to characterize new spontaneous mutants, identify causative mutations and provide new models to the scientific community. Finally, we thank Susan Ackerman and Kenneth Johnson for valuable comments on the manuscript.

Literature Cited

The Mouse Genomes Project.

Anderson MG, Hawes NL, Trantow CM, Chang B, John SW. Iris phenotypes and pigment dispersion caused by genes influencing pigmentation. Pigment cell & melanoma research. 2008; 21(5):565–78. [PubMed: 18715234]

Arnold CN, Xia Y, Lin P, Ross C, Schwander M, Smart NG, Muller U, Beutler B. Rapid identification of a disease allele in mouse through whole genome sequencing and bulk segregation analysis. Genetics. 2011; 187(3):633–41. [PubMed: 21196518]

Bergstrom, DE.; Bergstrom, RA.; Munroe, RJ.; Schimenti, JC.; Gagnon, LH. The Nox3 and Noxo1 genes, encoding presumptive members of an NADPH oxidase complex, are required for normal vestibular function and development; 18th international mammalian genome conference; Seattle, WA. 17-22 Oct 2004; 2004.

Blake JA, Bult CJ, Kadin JA, Richardson JE, Eppig JT. The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. Nucleic Acids Res. 2011; 39(Database issue):D842–8. [PubMed: 21051359]

Boles MK, Wilkinson BM, Wilming LG, Liu B, Probst FJ, Harrow J, Grafham D, Hentges KE, Woodward LP, Maxwell A, Mitchell K, Risley MD, Johnson R, Hirschi K, Lupski JR, Funato Y, Miki H, Marin-Garcia P, Matthews L, Coffey AJ, Parker A, Hubbard TJ, Rogers J, Bradley A, Adams DJ, Justice MJ. Discovery of candidate disease genes in ENU-induced mouse mutants by large-scale sequencing, including a splice-site mutation in nucleoredoxin. PLoS Genet. 2009; 5(12):e1000759. [PubMed: 20011118]

Bruneau S, Johnson KR, Yamamoto M, Kuroiwa A, Duboule D. The mouse Hoxd13(spdh) mutation, a polyalanine expansion similar to human type II synpolydactyly (SPD), disrupts the function but not the expression of other Hoxd genes. Dev Biol. 2001; 237(2):345–53. [PubMed: 11543619]

Chang B, Khanna H, Hawes N, Jimeno D, He S, Lillo C, Parapuram SK, Cheng H, Scott A, Hurd RE, Sayer JA, Otto EA, Attanasio M, O'Toole JF, Jin G, Shou C, Hildebrandt F, Williams DS, Heckenlively JR, Swaroop A. In-frame deletion in a novel centrosomal/ciliary protein CEP290/NPHP6 perturbs its interaction with RPGR and results in early-onset retinal degeneration in the rd16 mouse. Human molecular genetics. 2006; 15(11):1847–57. [PubMed: 16632484]

Chase TH, Lyons BL, Bronson RT, Foreman O, Donahue LR, Burzenski LM, Gott B, Lane P, Harris B, Ceglarek U, Thiery J, Wittenburg H, Thon JN, Italiano JE Jr, Johnson KR, Shultz LD. The mouse mutation “thrombocytopenia and cardiomyopathy” (trac) disrupts Abcg5: a spontaneous single gene model for human hereditary phytosterolemia/sitosterolemia. Blood. 2010; 115(6):1267–76. [PubMed: 19846887]

Ciciotte SL, Gwynn B, Moriyama K, Huizing M, Gahl WA, Bonifacino JS, Peters LL. Cappuccino, a mouse model of Hermansky-Pudlak syndrome, encodes a novel protein that is part of the pallidin-muted complex (BLOC-1). Blood. 2003; 101(11):4402–7. [PubMed: 12576321]

Cideciyan AV, Rachel RA, Aleman TS, Swider M, Schwartz SB, Sumaroka A, Roman AJ, Stone EM, Jacobson SG, Swaroop A. Cone photoreceptors are the main targets for gene therapy of NPHP5 (IQCB1) or NPHP6 (CEP290) blindness: generation of an all-cone Nphp6 hypomorph mouse that mimics the human retinal ciliopathy. Human molecular genetics. 2011; 20(7):1411–23. [PubMed: 21245082]

Cocquempot O, Brault V, Babinet C, Herault Y. Fork stalling and template switching as a mechanism for polyalanine tract expansion affecting the DYC mutant of HOXD13, a new murine model of synpolydactyly. Genetics. 2009; 183(1):23–30. [PubMed: 19546318]

Cook SA, Collin GB, Bronson RT, Naggert JK, Liu DP, Akeson EC, Davisson MT. A mouse model for Meckel syndrome type 3. Journal of the American Society of Nephrology: JASN. 2009; 20(4): 753–64. [PubMed: 19211713]

Curtain MM, Donahue LR. A short face mutation on Chromosome 11. MGI Direct Data Submission. 2007

D'Ascenzo M, Meacham C, Kitzman J, Middle C, Knight J, Winer R, Kukricar M, Richmond T, Albert TJ, Czechanski A, Donahue LR, Affourtit J, Jeddeloh JA, Reinholdt L. Mutation discovery

in the mouse using genetically guided array capture and resequencing. Mamm Genome. 2009; 20(7):424–36. [PubMed: 19629596]

Fairfield H, Gilbert GJ, Barter M, Corrigan RR, Curtain M, Ding Y, D'Ascenzo M, Gerhardt DJ, He C, Huang W, Richmond T, Rowe L, Probst FJ, Bergstrom DE, Murray SA, Bult C, Richardson J, Kile B, Gut I, Hager J, Sigurdsson S, Mauceli E, Di Palma F, Lindblad-Toh K, Cunningham ML, Cox TC, Justice MJ, Spector MS, Lowe SW, Albert T, Donahue LR, Jeddeloh J, Shendure J, Reinholdt LG. Mutation discovery in mice by whole exome sequencing. Genome biology. 2011; 12(9):R86. [PubMed: 21917142]

Fan Y, Rudert WA, Grupillo M, He J, Sisino G, Trucco M. Thymus-specific deletion of insulin induces autoimmune diabetes. The EMBO journal. 2009; 28(18):2812–24. [PubMed: 19680229]

Frazer KA, Eskin E, Kang HM, Bogue MA, Hinds DA, Beilharz EJ, Gupta RV, Montgomery J, Morenzoni MM, Nilsen GB, Pethiyagoda CL, Stuve LL, Johnson FM, Daly MJ, Wade CM, Cox DR. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature. 2007; 448(7157):1050–3. [PubMed: 17660834]

Friedman JS, Chang B, Kannabiran C, Chakarova C, Singh HP, Jalali S, Hawes NL, Branham K, Othman M, Filippova E, Thompson DA, Webster AR, Andreasson S, Jacobson SG, Bhattacharya SS, Heckenlively JR, Swaroop A. Premature truncation of a novel protein, RD3, exhibiting subnuclear localization is associated with retinal degeneration. American journal of human genetics. 2006; 79(6):1059–70. [PubMed: 17186464]

Garvey SM, Rajan C, Lerner AP, Frankel WN, Cox GA. The muscular dystrophy with myositis (mdm) mouse mutation disrupts a skeletal muscle-specific domain of titin. Genomics. 2002; 79(2):146–9. [PubMed: 11829483]

Gwynn B, Ciciotte SL, Hunter SJ, Washburn LL, Smith RS, Andersen SG, Swank RT, Dell'Angelica EC, Bonifacino JS, Eicher EM, Peters LL. Defects in the cappuccino (cno) gene on mouse chromosome 5 and human 4p cause Hermansky-Pudlak syndrome by an AP-3-independent mechanism. Blood. 2000; 96(13):4227–35. [PubMed: 11110696]

Gwynn B, Martina JA, Bonifacino JS, Sviderskaya EV, Lamoreux ML, Bennett DC, Moriyama K, Huizing M, Helip-Wooley A, Gahl WA, Webb LS, Lambert AJ, Peters LL. Reduced pigmentation (rp), a mouse model of Hermansky-Pudlak syndrome, encodes a novel component of the BLOC-1 complex. Blood. 2004; 104(10):3181–9. [PubMed: 15265785]

Harris BS, Franz T, Ullrich S, Cook S, Bronson RT, Davisson MT. Forebrain overgrowth (fog): a new mutation in the mouse affecting neural tube development. Teratology. 1997; 55(4):231–40. [PubMed: 9216040]

Honarpour N, Du C, Richardson JA, Hammer RE, Wang X, Herz J. Adult Apaf-1-deficient mice exhibit male infertility. Dev Biol. 2000; 218(2):248–58. [PubMed: 10656767]

Honarpour N, Gilbert SL, Lahn BT, Wang X, Herz J. Apaf-1 deficiency and neural tube closure defects are found in fog mice. Proceedings of the National Academy of Sciences of the United States of America. 2001; 98(17):9683–7. [PubMed: 11504943]

Johnson KR, Lane PW, Cook SA, Harris BS, Ward-Bailey PF, Bronson RT, Lyons BL, Shultz LD, Davisson MT. Curly bare (cub), a new mouse mutation on chromosome 11 causing skin and hair abnormalities, and a modifier gene (mcub) on chromosome 5. Genomics. 2003; 81(1):6–14. [PubMed: 12573256]

Johnson KR, Lane PW, Ward-Bailey P, Davisson MT. Mapping the mouse dactylaplasia mutation, Dac, and a gene that controls its expression, mdac. Genomics. 1995; 29(2):457–64. [PubMed: 8666395]

Johnson KR, Sweet HO, Donahue LR, Ward-Bailey P, Bronson RT, Davisson MT. A new spontaneous mouse mutation of Hoxd13 with a polyalanine expansion and phenotype similar to human synpolydactyly. Hum Mol Genet. 1998; 7(6):1033–8. [PubMed: 9580668]

Johnson KR, Zheng QY, Noben-Trauth K. Strain background effects and genetic modifiers of hearing in mice. Brain research. 2006; 1091(1):79–88. [PubMed: 16579977]

Juriloff DM, Harris MJ, Dewell SL. A digenic cause of cleft lip in A-strain mice and definition of candidate genes for the two loci. Birth defects research Part A, Clinical and molecular teratology. 2004; 70(8):509–18.

Juriloff DM, Harris MJ, McMahon AP, Carroll TJ, Lidral AC. Wnt9b is the mutated gene involved in multifactorial nonsyndromic cleft lip with or without cleft palate in A/WySn mice, as confirmed by a genetic complementation test. Birth defects research Part A, Clinical and molecular teratology. 2006; 76(8):574–9.

Justice MJ. Capitalizing on large-scale mouse mutagenesis screens. Nature reviews Genetics. 2000; 1(2):109–15.

Karst SY, Ward-Bailey PF, Washburn LL, Bergstrom DE, Donahue LR, Davisson-Fahey MT. Vaginal imperforation; a new mutation on Chromosome 13 causing a reproductive phenotype. MGI Direct Data Submission. 2011

Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, Heger A, Agam A, Slater G, Goodson M, Furlotte NA, Eskin E, Nellaker C, Whitley H, Cleak J, Janowitz D, Hernandez-Pliego P, Edwards A, Belgard TG, Oliver PL, McIntyre RE, Bhomra A, Nicod J, Gan X, Yuan W, van der Weyden L, Steward CA, Bala S, Stalker J, Mott R, Durbin R, Jackson IJ, Czechanski A, Guerra-Assuncao JA, Donahue LR, Reinholdt LG, Payseur BA, Ponting CP, Birney E, Flint J, Adams DJ. Mouse genomic variation and its effect on phenotypes and gene regulation. Nature. 2011; 477(7364):289–94. [PubMed: 21921910]

Kirby A, Kang HM, Wade CM, Cotsapas C, Kostem E, Han B, Furlotte N, Kang EY, Rivas M, Bogue MA, Frazer KA, Johnson FM, Beilharz EJ, Cox DR, Eskin E, Daly MJ. Fine mapping in 94 inbred mouse strains using a high-density haplotype resource. Genetics. 2010; 185(3):1081–95. [PubMed: 20439770]

Kiss PJ, K J, Zhang Y, Baltrusaitis J, Sigmund CD, Thalmann R, Smith RJ, Verpy E, Bánfi B. Inactivation of NADPH oxidase organizer 1 results in severe imbalance. Curr Biol. 2006; 16(2): 208–13. Curr Biol. 2006 Jan 24; 16(2): 208-13. [PubMed: 16431374]

Lee CH, Reifsnyder PC, Naggert JK, Wasserfall C, Atkinson MA, Chen J, Leiter EH. Novel leptin receptor mutation in NOD/LtJ mice suppresses type 1 diabetes progression: I. Pathophysiological analysis. Diabetes. 2005; 54(9):2525–32. [PubMed: 16123339]

Lee JW, Beebe K, Nangle LA, Jang J, Longo-Guess CM, Cook SA, Davisson MT, Sundberg JP, Schimmel P, Ackerman SL. Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. Nature. 2006; 443(7107):50–5. [PubMed: 16906134]

Li X, Li W, Dai X, Kong F, Zheng Q, Zhou X, Lu F, Chang B, Rohrer B, Hauswirth WW, Qu J, Pang JJ. Gene therapy rescues cone structure and function in the 3-month-old rd12 mouse: a model for midcourse RPE65 leber congenital amaurosis. Investigative ophthalmology & visual science. 2011; 52(1):7–15. [PubMed: 21169527]

Liang Y, Seymour RE, Sundberg JP. Inhibition of NF-kappaB signaling retards eosinophilic dermatitis in SHARPIN-deficient mice. The Journal of investigative dermatology. 2011; 131(1):141–9. [PubMed: 20811394]

Longo-Guess CM, Gagnon LH, Cook SA, Wu J, Zheng QY, Johnson KR. A missense mutation in the previously undescribed gene Tmhs underlies deafness in hurry-scurry (hscy) mice. Proc Natl Acad Sci USA. 2005; 102:7894–99. [PubMed: 15905332]

Lorenz-Depiereux B, Guido VE, Johnson KR, Zheng QY, Gagnon LH, Bauschatz JD, Davisson MT, Washburn LL, Donahue LR, Strom TM, Eicher EM. New intragenic deletions in the *Phex* gene clarify XLH-related abnormalities in mice. Mamm Genome. 2004; 15:151–61. [PubMed: 15029877]

Merath KM, Chang B, Dubietzig R, Jeannotte R, Sidjanin DJ. A spontaneous mutation in Srebf2 leads to cataracts and persistent skin wounds in the lens opacity 13 (lop13) mouse. Mamm Genome. 2011

Nadeau JH. Modifier genes and protective alleles in humans and mice. Current opinion in genetics & development. 2003; 13(3):290–5. [PubMed: 12787792]

Nair KS, Hmani-Aifa M, Ali Z, Kearney AL, Salem SB, Macalinao DG, Cosma IM, Bouassida W, Hakim B, Benzina Z, Soto I, Soderkvist P, Howell GR, Smith RS, Ayadi H, John SW. Alteration of the serine protease PRSS56 causes angle-closure glaucoma in mice and posterior microphthalmia in humans and mice. Nature Genetics. 2011; 43(6):579–84. [PubMed: 21532570]

Nakano Y, LG C, Bergstrom DE, Nauseef WM, Jones SM, Bánfi B. Mutation of the Cyba gene encoding p22phox causes vestibular and immune defects in mice. J Clin Invest. 2008; 118(3): 1176–85. [PubMed: 18292807]

Nguyen T, Novak EK, Kermani M, Fluhr J, Peters LL, Swank RT, Wei ML. Melanosome morphologies in murine models of hermansky-pudlak syndrome reflect blocks in organelle development. The Journal of investigative dermatology. 2002; 119(5):1156–64. [PubMed: 12445206]

Noben-Trauth K, Zheng QY, Johnson KR. Association of cadherin 23 with polygenic inheritance and genetic modification of sensorineural hearing loss. Nature Genetics. 2003; 35(1):21–3. [PubMed: 12910270]

Odeh H, Hunker KL, Belyantseva IA, Azaiez H, Avenarius MR, Zheng L, Peters LM, Gagnon LH, Hagiwara N, Skynner MJ, Brilliant MH, Allen ND, Riazuddin S, Johnson KR, Raphael Y, Najmabadi H, Friedman TB, Bartles JR, Smith RJ, Kohrman DC. Mutations in Grxcr1 are the basis for inner ear dysfunction in the pirouette mouse. American journal of human genetics. 2010; 86(2):148–60. [PubMed: 20137774]

Odgren PR, Pratt CH, Mackay CA, Mason-Savas A, Curtain M, Shopland L, Ichicki T, Sundberg JP, Donahue LR. Disheveled hair and ear (Dhe), a spontaneous mouse Lmna mutation modeling human laminopathies. PloS one. 2010; 5(4):e9959. [PubMed: 20376364]

Paffenholz R, Bergstrom RA, Pasutto F, Wabnitz P, Munroe RJ, Jagla W, Heinzmann U, Marquardt A, Bareiss A, Laufs J, Russ A, Stumm G, Schimenti JC, Bergstrom DE. Vestibular defects in head-tilt mice result from mutations in Nox3, encoding an NADPH oxidase. Genes Dev. 2004; 18(5):486–91. [PubMed: 15014044]

Pang J, Boye SE, Lei B, Boye SL, Everhart D, Ryals R, Umino Y, Rohrer B, Alexander J, Li J, Dai X, Li Q, Chang B, Barlow R, Hauswirth WW. Self-complementary AAV-mediated gene therapy restores cone function and prevents cone degeneration in two models of Rpe65 deficiency. Gene therapy. 2010; 17(7):815–26. [PubMed: 20237510]

Pang JJ, Chang B, Hawes NL, Hurd RE, Davisson MT, Li J, Noorwez SM, Malhotra R, McDowell JH, Kaushal S, Hauswirth WW, Nusinowitz S, Thompson DA, Heckenlively JR. Retinal degeneration 12 (rd12): a new, spontaneously arising mouse model for human Leber congenital amaurosis (LCA). Molecular vision. 2005; 11:152–62. [PubMed: 15765048]

Plamondon JA, Harris MJ, Mager DL, Gagnier L, Juriloff DM. The clf2 gene has an epigenetic role in the multifactorial etiology of cleft lip and palate in the A/WySn mouse strain. Birth defects research Part A, Clinical and molecular teratology. 2011; 91(8):716–27.

Plosch T, Bloks VW, Terasawa Y, Berdy S, Siegler K, Van Der Sluijs F, Kema IP, Groen AK, Shan B, Kuipers F, Schwarz M. Sitosterolemia in ABC-transporter G5-deficient mice is aggravated on activation of the liver-X receptor. Gastroenterology. 2004; 126(1):290–300. [PubMed: 14699507]

Probst FJ, Justice MJ. Mouse mutagenesis with the chemical supermutagen ENU. Methods in enzymology. 2010; 477:297–312. [PubMed: 20699147]

Pruitt KD, Tatusova T, Klimke W, Maglott DR. NCBI Reference Sequences: current status, policy and new initiatives. Nucleic Acids Res. 2009; 37(Database issue):D32–6. [PubMed: 18927115]

Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 2007; 35(Database issue):D61–5. [PubMed: 17130148]

Radke MH, Peng J, Wu Y, McNabb M, Nelson OL, Granzier H, Gotthardt M. Targeted deletion of titin N2B region leads to diastolic dysfunction and cardiacatrophy. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(9):3444–9. [PubMed: 17360664]

Romeo G, McKusick VA. Phenotypic diversity, allelic series and modifier genes. Nature Genetics. 1994; 7(4):451–3. [PubMed: 7951309]

Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, Phipps EL. Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. Proc Natl Acad Sci U S A. 1979; 76(11):5818–9. [PubMed: 293686]

Sands MS, Wolfe JH, Birkenmeier EH, Barker JE, Vogler C, Sly WS, Okuyama T, Freeman B, Nicholes A, Muzyczka N, Chang PL, Axelrod HR. Gene therapy for murine mucopolysaccharidosis type VII. Neuromuscular disorders: NMD. 1997; 7(5):352–60. [PubMed: 9267850]

Schlager G, Dickie MM. Natural mutation rates in the house mouse. Estimates for five specific loci and dominant mutations. Mutat Res. 1971; 11(1):89–96. [PubMed: 5104863]

Seymour RE, Hasham MG, Cox GA, Shultz LD, Hogenesch H, Roopenian DC, Sundberg JP. Spontaneous mutations in the mouse Sharpin gene result in multiorgan inflammation, immune system dysregulation and dermatitis. Genes and immunity. 2007; 8(5):416–21. [PubMed: 17538631]

Signorini S, Liao YJ, Duncan SA, Jan LY, Stoffel M. Normal cerebellar development but susceptibility to seizures in mice lacking G protein-coupled, inwardly rectifying K+ channel GIRK2. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94(3):923–7. [PubMed: 9023358]

Simon-Chazottes D, Tutois S, Kuehn M, Evans M, Bourgade F, Cook S, Davisson MT, Guenet JL. Mutations in the gene encoding the low-density lipoprotein receptor LRP4 cause abnormal limb development in the mouse. Genomics. 2006; 87(5):673–7. [PubMed: 16517118]

Surmeier DJ, Mermelstein PG, Goldowitz D. The weaver mutation of GIRK2 results in a loss of inwardly rectifying K+ current in cerebellar granule cells. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93(20):11191–5. [PubMed: 8855331]

Turner DJ, Keane TM, Sudbery I, Adams DJ. Next-generation sequencing of vertebrate experimental organisms. Mamm Genome. 2009; 20(6):327–38. [PubMed: 19452216]

Wang J, Takeuchi T, Tanaka S, Kubo SK, Kayo T, Lu D, Takata K, Koizumi A, Izumi T. A mutation in the insulin 2 gene induces diabetes with severe pancreatic beta-cell dysfunction in the Mody mouse. J Clin Invest. 1999; 103(1):27–37. [PubMed: 9884331]

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, Antonarakis SE, Attwood J, Baertsch R, Bailey J, Barlow K, Beck S, Berry E, Birren B, Bloom T, Bork P, Botcherby M, Bray N, Brent MR, Brown DG, Brown SD, Bult C, Burton J, Butler J, Campbell RD, Carninci P, Cawley S, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Clee C, Collins FS, Cook LL, Copley RR, Coulson A, Couronne O, Cuff J, Curwen V, Cutts T, Daly M, David R, Davies J, Delehaunty KD, Deri J, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dodge S, Dubchak I, Dunn DM, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Fewell GA, Flicek P, Foley K, Frankel WN, Fulton LA, Fulton RS, Furey TS, Gage D, Gibbs RA, Glusman G, Gnerre S, Goldman N, Goodstadt L, Grafham D, Graves TA, Green ED, Gregory S, Guigo R, Guyer M, Hardison RC, Haussler D, Hayashizaki Y, Hillier LW, Hinrichs A, Hlavina W, Holzer T, Hsu F, Hua A, Hubbard T, Hunt A, Jackson I, Jaffe DB, Johnson LS, Jones M, Jones TA, Joy A, Kamal M, Karlsson EK, et al. Initial sequencing and comparative analysis of the mouse genome. Nature. 2002; 420(6915):520–62. [PubMed: 12466850]

Weatherbee SD, Anderson KV, Niswander LA. LDL-receptor-related protein 4 is crucial for formation of the neuromuscular junction. Development. 2006; 133(24):4993–5000. [PubMed: 17119023]

Weinert S, Bergmann N, Luo X, Erdmann B, Gotthardt M. M line-deficient titin causes cardiac lethality through impaired maturation of the sarcomere. The Journal of cell biology. 2006; 173(4): 559–70. [PubMed: 16702235]

Wilming LG, Gilbert JG, Howe K, Trevanion S, Hubbard T, Harrow JL. The vertebrate genome annotation (Vega) database. Nucleic Acids Res. 2008; 36(Database issue):D753–60. [PubMed: 18003653]

Wong K, Keane TM, Stalker J, Adams DJ. Enhanced structural variant and breakpoint detection using SVMerge by integration of multiple detection methods and local assembly. Genome Biol. 2010; 11(12):R128. [PubMed: 21194472]

Yoshida H, Kong YY, Yoshida R, Elia AJ, Hakem A, Hakem R, Penninger JM, Mak TW. Apaf1 is required for mitochondrial pathways of apoptosis and brain development. Cell. 1998; 94(6):739–50. [PubMed: 9753321]

Zhang Z, Alpert D, Francis R, Chatterjee B, Yu Q, Tansey T, Sabol SL, Cui C, Bai Y, Koriabine M, Yoshinaga Y, Cheng JF, Chen F, Martin J, Schackwitz W, Gunn TM, Kramer KL, De Jong PJ, Pennacchio LA, Lo CW. Massively parallel sequencing identifies the gene Megf8 with ENU-induced mutation causing heterotaxy. Proc Natl Acad Sci U S A. 2009; 106(9):3219–24. [PubMed: 19218456]

Zhao L, Longo-Guess C, Harris BS, Lee JW, Ackerman SL. Protein accumulation and neurodegeneration in the woozy mutant mouse is caused by disruption of SIL1, a cochaperone of BiP. Nat Genet. 2005; 37(9):974–9. [PubMed: 16116427]

Table 1

Selected examples of mouse mutations that led to discovery of causative genes in human diseases and provided models for drug testing.

Mouse				**Human**		
Mutation	**Gene**	**Phenotype**	**Citation**	**Disease**	**Gene**	**Clinical features**
pirouette, *pi*	glutaredoxin cysteine-rich 1, *Grxcr1*	hearing loss and vestibular dysfunction due to neuroepithelial defects in the inner ear	(Odeh et al. 2010)	autosomal recessive deafness type 25, DFNB25	*GRXCR1*	Profound early onset deafness
hurry-scurry, *hscy*	lipoma HMGIC fusion partner-like 5, *Lhfpl5*	Deafness, vestibular dysfunction	(Longo-Guess et al. 2005)	autosomal recessive nonsyndromic deafness type 67, DFNB67	*LHFPL5*	Profound early onset deafness
chronic proliferative dermatitis, *cpdm*	SHANK-associated RH domain interacting protein, *Sharpin*	blood and tissue eosinophilia, dermatologic lesions	(Seymour et al. 2007)	hypereosinophilic syndromes		blood and tissue eosinophilia, dermatologic lesions
retinal degeneration 3, *Rd3*	retinal degeneration 3, *Rd3*	Early onset retinal degeneration	(Friedman et al. 2006)	Leber congenital amaurosis type 12, LCA12 (OMIM#610612)	*RD3*	Early childhood partial or total blindness
retinal degeneration 12, *rd12*	retinal pigment epithelium 65, *Rpe65*	Early onset retinal degeneration	(Pang et al. 2005)	LCA2 (OMIM#204100,	*RPE65*	Early childhood partial or total blindness
retinal degeneration 16, *rd16*	centrosomal protein 290, *Cep290*	Early onset retinal degeneration	(Chang et al. 2006)	LCA10 (OMIM#611755; Joubert Syndrome 5 (OMIM#610188)	*CEP290*	Early childhood partial or total blindness
(elevated intraocular pressure)	RIKEN cDNA 1700027L20 gene, *1700027L20Rik*	elevated intraocular pressure, eventual blindness, posterior microphthalmia	(Nair et al. 2011)	Angle-closure glaucoma, several types		Glaucoma
muscular dystrophy with myositis, *mdm*	Titin, *Ttn*	causes severe limb muscular dystrophy with inflammation	(Garvey et al. 2002)	tibial and limb-girdle muscular dystrophies (OMIM#600334 & OMIM#608807, respectively) & cardiomyopathy (OMIM#604145).	*TTN*	distal myopathies, such as those listed at the left
mucopoly-saccharidosis VII, *mps*	Glucuronidase, beta, *Gusb*	Lysosomal storage disease, Disproportionate dwarfism, skeletal abnormalities	(Sands et al. 1997)	Mucopolysaccharidosis VII, Sly Syndrome (OMIM#253220)	*GUSB*	Lysosomal storage disease, various skeletal abnormalities

Table 2

Selected examples of mouse mutations that led to discovery of novel gene functions.

Mouse					
Mutation	**Gene**	**Phenotype**	**Citation**	**Mechanism**	**Human Disease or biomedical importance**
sticky, *sti*	alanyl-tRNA synthetase gene, *Aars*	cerebellar Purkinje cell loss, ataxia	(Lee et al. 2006)	Disruption of translation, intracellular accumulation of misfolded proteins	Disrupted translation fidelity causes neuro-degeneration
woozy, *wz*	endoplasmic reticulum chaperone SIL1 homolog (*S. cerevisiae*), *Sil1*	cerebellar Purkinje cell loss, adult onset ataxia	(Zhao et al. 2005)	ER chaperone function in terminally differentiated neurons leads to protein accumulation, ER stress and subsequent neuro-degeneration	Marinesco-Sjogren syndrome (OMIM#248800) cerebellar ataxia, progressive myopathy, and cataracts
digitation anormale, *dan* malformed digits, *mdig*	low density lipoprotein receptor-related protein 4. *Lrp4*	brachy-, poly- and syndactyly, abnormal nails	(Simon-Chazottes et al. 2006)	Identified an unsuspected function in skeletal development for a lipid metabolism protein	Cenani-Lenz syndactyly syndrome (OMIM#212780) fusion of radius & ulna, metacarpal & phalangeal abnormalities
Thrombocyto-penia and cardiomyopathy, *trac*	ATP-binding cassette, family g, member 5, *Abcg5*	cardiomyopathy	(Chase et al. 2010)	*Abcg5* not previously associated with cardiomyopathy	Sitosterolemia (OMIM#210250), high cholesterol absorption, atherosclerosis and coronary disease
lens opacity 13, *lop13*	Sterol regulatory element binding-transcription factor 2, *Srebf2*	nuclear cataracts, chronic skin wounds	(Merath et al. 2011)	SREBF2 deficiency leads to reduced cellular cholesterol which seems to be required for normal lens fiber cell maintenance	Study of the contribution of cholesterol and fatty acids to homeostasis in the lens and skin
short face, *Sofa*	phosphoribosylformylglycinamidine synthase (FGAR amidotransferase), *Pfas*	short nose, domed skull and wide set eyes	(Curtain and Donahue 2007) (Fairfield et al. 2011)	Affects aminoimidazole ribonucleotide biosynthesis pathway	Study of PFAS in craniofacial development
vaginal imperforation, *vgim*	lipoma HMGIC fusion partner-like 2, *Lhfpl2*	Complete closure of the vagina, swollen perineum, fluid-filled uterus	(Karst et al. 2011) (Fairfield et al. 2011)	Unknown, PFAS is integral to cellular membrane	Study of LHFPL2 in cell to cell interactions, adhesions

Table 3

Examples of the variety of mutation types that occur as spontaneous mutations and comparison with targeted mutations of the same genes.

Mutant gene/allele	Spontaneous mutation type	Spontaneous mutation phenotype	Citation	Targeted mutation phenotype	Citation
Titin; Muscular dystrophy with myositis, *Ttn*^{mdm}	Deletion of 4 small exons of very large gene, hypomorph	Severe muscle degeneration, myositis, premature death; model for tibial mucular dystrophy	(Garvey et al. 2002)	embryonic-perinatal lethality	(Weinert et al. 2006)
				Cardiovascular defects, not closely resembling cardiomyopathy in human patients	(Radke et al. 2007)
potassium inwardly-rectifying channel, subfamily J, member 6; weaver, *Kcnj6*^{wv}	gain-of-function, a single nucleotide transition affecting the highly conserved H5 domain of the channel	affects cerebellar granule cell migration	(Surmeier et al. 1996)	normal cerebellar development	(Signorini et al. 1997)
apoptotic peptidase activating factor 1; forebrain overgrowth, *Apaf1*^{fog}	partial loss of function, reduced mRNA protein levels & activity	abnormal cranium, spina bifida occulta, lowered survivor rate, but many survive to adulthood, some reduced fertility	(Honarpour et al. 2001, Harris et al. 1997)	Similar to *fog* but most pre- or perinatal lethality, survivors sterile	(Honarpour et al. 2000)
				Similar to *fog* but complete postnatal lethal	(Yoshida et al. 1998)
Insulin II; Akita, *Ins2*^{Akita}	dominant negative, single nucleotide transition that alters protein conformation	Affects many systems including adipose tissue and organs affected by diabetes, premature death	(Wang et al. 1999)	Complete postnatal lethality	(Fan et al. 2009)
ATP-binding cassette, sub-family G (WHITE), member 1; thrombocytopenia and cardiomyopathy, *Abcg5*^{trac}	G-to-A mutation in exon 10 of the gene, creating a premature stop codon predicted to produce a mutant protein that lacks the last 4 transmembrane domains	Cardiomyopathy on A/J; none on other backgrounds	(Chase et al. 2010)	Homeostasis, liver/billary	(Plosch et al. 2004)

Table 4

Selected examples of genetic background effect on phenotype.

Gene	Mutation	Genetic Background	Phenotype	Citation
ATP-binding cassette, sub-family G (WHITE), member 1, *Abcg5*	thrombocytopenia and cardiomyopathy, *trac*	A/J	cardiomyopathy	(Chase et al. 2010)
		Others, including C57BL/6J	no noticeable heart-related phenotype	
F-box and WD-40 domain protein 4, *Fbxw4*	Dactylaplasia, *Dac;* modifier of *Dac, mdac*	SM/Ckc, LG/Ckc, BALB/cJ, 129/J	Heterozygotes: missing and fused digits, Homozygotes: pre- or perinatal lethal, no feet single digit	(Johnson et al. 1995)
		Crosses to NZB/BINJ, DBA/2J, CBA/J. C3H/J,/. SWR/J	Homozygotes: viable and fertile, digits less severely affected	
		Others, including C57BL/6J	No phenotype	
Genes not yet identified	Curly bare, *cub;* modifier of *cub, mcub*	Multiple backgrounds	curly coat or complete hairlessness depending on allele of *mcub* in each background	(Johnson et al. 2003)
ATPase, Ca++ transporting, plasma membrane 2, *Atp2b2*	deaf waddler, *dfw*	Multiple backgrounds	Allelic variants of *Cdh23* (*mdfw, Ahl*) determine severity of hearing loss due to *dfw* mutation and age of age-related hearing loss when *dfw* mutations are not present	(Noben-Trauth et al. 2003) (Johnson et al. 2006)
G protein-coupled receptor 98, *Gpr98*cadherin 23, *Cdh23*	frings, *frings*modifier of *dfw,* age-related hearing loss, *ahl,*			
leptin receptor, *Lepr*	Diabetic 5J, $Lepr^{db\text{-}5J}$	NOD/LtJ	This allele of *Lepr* suppresses Type 1 autoimmune diabetes in NOD/LtJ mice	(Lee et al. 2005)
transmembrane 67, *Tmem67*	bilateral polycystic kidney, *bpck*	C57BL/6J x C3H/HeJ mixed	Polycystic kidney disease, lethality by 20 days of age	(Cook et al. 2009)
		C57BL/6J, C3H/HeJ	Prenatal lethality	