charles river | The Jackson Laboratory

RESEARCH MODELS

Strategies to Minimize Genetic Drift and Maximize Experimental Reproducibility in Mouse Research

Author:
Janine Low-Marchelli, PhD, Senior Technical Information Scientist, The Jackson Laboratory

Overview/Abstract:

Genetic drift occurs in any independent mouse breeding colony and has the potential to negatively affect experimental reproducibility and scientific conclusions. Spontaneous mutations caused by genetic drift may go unrecognized for years, until the specific research questions that depend on such mutations happen to be addressed. While it cannot be stopped completely, genetic drift and the impact on scientific discovery can be minimized through careful and thoughtful colony management practices. Because individual mouse breeding colonies may differ in size and management strategies, use of complete and accurate mouse strain nomenclature, including substrain designation, benefits the scientific community as a whole.

The Importance of Genetic Stability in Mouse Research

To the average life-science researcher, a mouse's genetic background may be an afterthought, if even a thought at all. A researcher's top priorities may be to understand disease, to publish, and to obtain funding. However, to successfully reach these goals, maintaining genetic stability, or *preventing genetic drift*, in a mouse colony should be of high importance.

Laboratory mice are unique, live elements in scientific research that change over the course of their lifetime, and importantly, from one generation to the next. After all, heritable changes in DNA sequence are the basis for species diversity and evolution in the wild. Even in the absence of evolutionary pressure, changes in DNA sequence occur. At first glance, these mutations seem to be silent, unimportant fluctuations in the genetic makeup of an individual. However, these seemingly insignificant mutations can become the source of unexplainable experimental irreproducibility.

Mouse researchers, then, are met with a conundrum. Generating mice for research requires breeding, but with breeding comes the inherent risk to propagate genetic diversity, and thus, to propagate experimental diversity. From one experiment to the next and from one publication to the next, data diversity is unconducive to scientific progress.

The purpose of this paper is to educate mouse researchers on the potential for genetic drift to impact research progress, to highlight best practices to minimize drift, and provide solutions to reverse drift if it arises in a mouse colony. Use of full official mouse strain nomenclature and careful reporting of breeding generation information in publications and grant proposals are some simple practices researchers can take that promote reproducibility and responsible animal use.

How Genetic Drift Arises and Its Prevalence in Mouse Colonies

Inbreeding, or sibling mating, is a powerful method to reduce heterozygosity at every genetic locus in the mouse genome, allowing for uniformity in phenotype and forming the basis for experimental reproducibility. Genetic homozygosity allows comparison of a single variable between a control and an experimental group, and thus, to be able to attribute any differences in readout to that variable.

Much like species in the wild, two populations of inbred laboratory mouse strains maintained in isolation from each other will change over time. Spontaneous mutations may occur in the form of single nucleotide polymorphisms (SNPs), deletions, inversions, duplications, and other such errors during DNA replication and meiosis. This process of spontaneous mutations appearing, disappearing, or becoming fixed in a population at random is called genetic drift (Lee Silver, 1995).

The amount of genetic drift occurring in any actively breeding colony varies, but is predicted to be rather frequent. The average breeding generation is 3-4 months long, with mice becoming sexually mature around 5-8 weeks of age. Offspring are typically born about 3 weeks after mating. Based on spontaneous mutation rates calculated from coat color mutations measured in over 1 million mice, 1 phenotypic mutation may arise every 1.8 breeding generations (Drake et al., 1998; Russell and Russell, 1996).

The risk of breeding a mouse that carries a spontaneous mutation in the germ line, and thus of propagating this mutation, is higher in small colonies than in large colonies (Figure 1A). For any given germ line mutation in a mouse, roughly half of its offspring will be heterozygous for this mutation (Figure 1B). In inbred breeding colonies, there is a 25% chance these mutations will become fixed (homozygous) in the population (Chamary and Hurst, 2004; Drake et al., 1998).

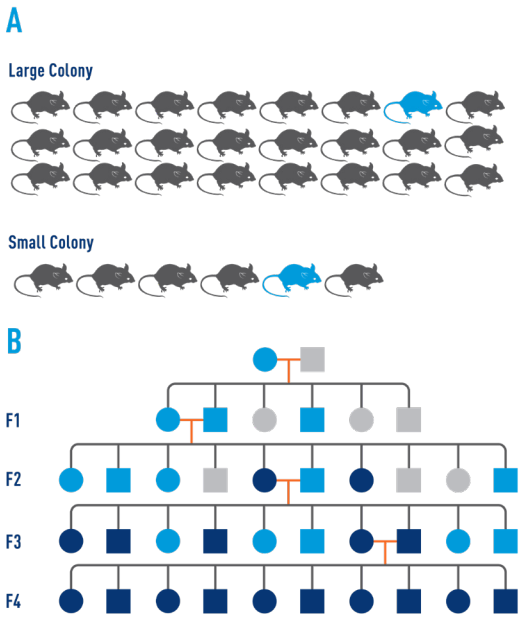


Figure 1. The risk of propagating a spontaneous mutation is higher in small colonies versus large colonies. A) The probability of using a mouse that carries any given mutation (light blue) for breeding is higher in a small colony than a large colony. B) In each round of breeding, there is a 25% chance that a new mutation will become more established in the population. For example, Mendelian inheritance predicts that the F1 generation will be composed of 50% wildtype (grey) and 50% heterozygous for the mutation (light blue). If by chance, two heterozygotes are used as breeders, the F2 generation will be composed of 25% wildtype, 50% heterozygotes, and 25% homozygotes (dark blue). This can continue until the entire colony is fixed homozygous for the mutation (F3, F4). However, the genome can drift in either direction depending on the genotypes of the mice used for breeding – the probability that the mutation becomes fixed is equivalent to the probability it will be lost entirely from the colony.

Indications that Genetic Drift Has Occurred: Substrain Designations

A substrain is a branch of an inbred strain that is *suspected* or *known* to be genetically different from the parent colony (http://www.informatics.jax.org/mgihome/nomen/strains.shtml#substrains).

Because genetic drift may differentially occur in two populations of any given inbred strain, substrain designation is a crucial component of nomenclature. Substrains are designated by adding a unique lab code assigned by the Institute for Laboratory Animal Research (ILAR) (http://dels.nas.edu/global/ilar/Lab-Codes). A lab code identifies the institute, lab, or investigator that produced or maintains a particular animal strain (Table 1). Because lab codes accumulate in the nomenclature, the strain's genealogy is understood from the name alone. For example, strain C57BL/6NJ was maintained for many years at the National Institutes of Health (N) and is now distributed by The Jackson Laboratory (J) (Figure 3). By extension, the substrain nomenclature gives a general indication that genetic variation between two strains exists.

Lab Code	Organization
Crl	Charles River Laboratories
Hsd	Envigo (formerly Harlan Laboratories)
J	The Jackson Laboratory
N	National Institutes of Health
Rj	Centre D'Elevage R. Janvier
Tac	Taconic Farms, Inc.

Table 1. Common laboratory codes found in mouse substrain nomenclature. The Institute for Laboratory Animal Research (ILAR) assigns and maintains unique identifiers for institutes, laboratories, or individual investigators who create and maintain mouse colonies.

Suspected genetic differences: generation number

Any strain that has been maintained separately from the parental strain for 20 consecutive inbred generations (~5-6 years) is suspected to carry genetic differences, and is therefore considered a substrain. Additionally, breeding generations are cumulative, such that if two labs obtain mice from the same common ancestor and breed for 10 generations, each lab has a different substrain from one another because the two strains are considered 20 generations apart (Figure 2).

The very first inbred mouse strains (including C57BL/6, DBA, C3H, BALB, CBA, and others) used for biological research were established almost 100 years ago and continue to be heavily published today. Because these strains exceed 200 inbred generations and because multiple institutions worldwide breed them, a considerable amount of genetic drift has occurred over time in all of these strains. Because of genetic drift, it is possible that observations made in existing substrains differ from observations made in the parental inbred strains from which they were derived.

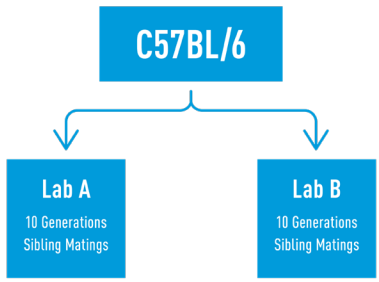


Figure 2. Substrain development. Substrains develop after 20 consecutive generations of inbreeding. While these labs have not surpassed 20 breeding generations individually, Lab A and Lab B are separated from each other by 20 generations. Appending laboratory codes to strain names can give a general indication of whether genetic drift has occurred in one substrain versus another.

Known genetic differences: substrain designation by observed phenotypic differences

Additionally, substrains are designated when a difference in phenotype is observed between two groups of inbred mice. However, unless these spontaneous mutations manifest obvious phenotypes, frequently after they become fixed homozygous in the colony, and attentive colony managers or researchers recognize something "off" about the mice, the mutations may be carried in a strain unnoticed for years. Thus, identifying drift may depend on individual labs asking questions whose answers happen to rely on such mutations to recognize that unexpected results are more than just "failed experiments," and later to identify the mutation that is responsible for the aberrant phenotype.

For example, the parental inbred strain C3H gave rise to two substrains from two Jackson Laboratory researchers, which for many years, did not seem to differ. Dr. Walter Heston bred the strain in the 1930s (now C3H/HeJ). In 1952, Heston transferred some of his mice to another Jackson Laboratory researcher, Dr. Henry Outzen (now C3H/HeOuJ). In the late 1960s, Heston's strain was found to be resistant to lipopolysaccharide (LPS), whereas Outzen's strain remained sensitive. Later, the mutation was mapped to *Tlr4*, a gene involved in pathogen recognition and innate immune system activation (Poltorak et al., 1998a; Watson et al., 1978). By the time the C to A substitution at nucleotide 2342 was identified in *Tlr4*, it had already become fixed in the Heston substrain, likely between 1958 and 1965 (Poltorak et al., 1998b). Had Heston's C3H substrain never been treated with LPS, it is possible that the *Tlr4* mutation would not have been identified and conclusions involving basic immunology in these strains may have become highly controversial.

Known genomic sequences are substrain-specific

Aside from chance discoveries, the only way to definitively identify whether genetic drift has occurred is to sequence the strain and compare to reference genomes. A C57BL/6J female was the first mouse to be completely sequenced by the Mouse Genome Sequencing Consortium (Chinwalla et al., 2002), www.ensembl.org/Mus_musculus). To date, 15 other major inbred mouse strains have been fully sequenced, all of which are "J" substrains, the official ILAR laboratory code for The Jackson Laboratory (Adams et al., 2015), www.ensembl.org/Mus_musculus/Info/Strains) (Table 2).

Mouse Strain, Full Nomenclature	JAX® Strain Number	Complete Sequence in Ensembl	Number of Datasets in MPD	Protected by GSP
C57BL/6J	000664	Y	237	Y
129S1/SvImJ	002448	Y	133	Y
A/J	000646	Y	177	
AKR/J	000648	Y	114	
B6.129P2-Apoe^{tm1Unc}/J	002052	Y	7	Y
BALB/cJ	000651	Y	93	
BALB/cByJ	001026		118	Y
C3H/HeJ	000659	Y	158	Y
C57BL/6NJ	005304	Y	2	Y
CAST/EiJ	000928	Y	97	
CBA/J	000656	Y	110	Y
DBA/1J	000670		36	Y
DBA/2J	000671	Y	166	Y
FVB/NJ	001800	Y	133	Y
LP/J	000676	Y	84	
NOD/ShiLtJ	001976	Y	106	Y
NOD.CB17-Prkdc^{scid}/J	001303	Y	8	Y
NZO/HlLtJ	002105	Y	49	
PWK/PhJ	003715	Y	43	
SPRET/EiJ	001146	Y	34	
WSB/EiJ	001145	Y	67	

Table 2. Sequence and phenotypic characterization data available for the most commonly used inbred strains is substrain specific. Only "J" mouse substrains have been fully sequenced and deposited in Ensembl (where Y = yes). Many of these substrains have SNP annotations, where the C57BL/6J substrain is the reference genome. Along with SNP information, thousands of substrain-specific phenotypic characteristics have been independently quantified and can be analyzed in the Mouse Phenome Database (MPD). Many of these strains are protected by The Jackson Laboratory's patented Genetic Stability Program (GSP) and are distributed by Charles River Laboratories in Europe and Japan.

An additional 20+ inbred strains have been sequenced using short-read approaches to identify SNPs, indels, and structural variations relative to the C57BL/6J mouse reference genome (Frazer et al., 2007 and www.sanger.ac.uk/science/data/mouse-genomes-project).

Furthermore, known SNP data for specific substrains can be found and compared in the Mouse Phenome Database (MPD), a collaborative standardized collection of genotypic and phenotypic data on the most commonly published mouse strains (http://phenome.jax.org).

Genetic Background Impacts Research Conclusions

As described earlier with the C3H example, substrains may acquire spontaneous mutations that have the potential to influence research conclusions. If these experiments are not properly controlled for, such as through use of appropriate substrains, disastrous consequences on experimental reproducibility may ensue. Whether these spontaneous mutations arise in a repository, from a vendor, or in individual laboratories, how can researchers know which is the "best" substrain to use for their experiments?

Unfortunately, there isn't an easy answer. The best way to determine whether genetic background matters is to perform controlled, side-by-side experiments and compare. Since it is impossible to test every substrain that exists for a particular experimental readout, the next best way to understand the potential impact of genetic background on research conclusions is to rely on what other researchers have observed, in the form of peer-reviewed, published literature, and to continue experiments that build on such knowledge using identical substrains.

C57BL/6 Substrains

Certainly, substrain differences exist in all inbred mouse strains. By far, however, the C57BL/6 strains are the most commonly published in the world, with over 37,000 entries in PubMed (Table 3). For this reason, this paper will focus on published differences only in the C57BL/6 family.

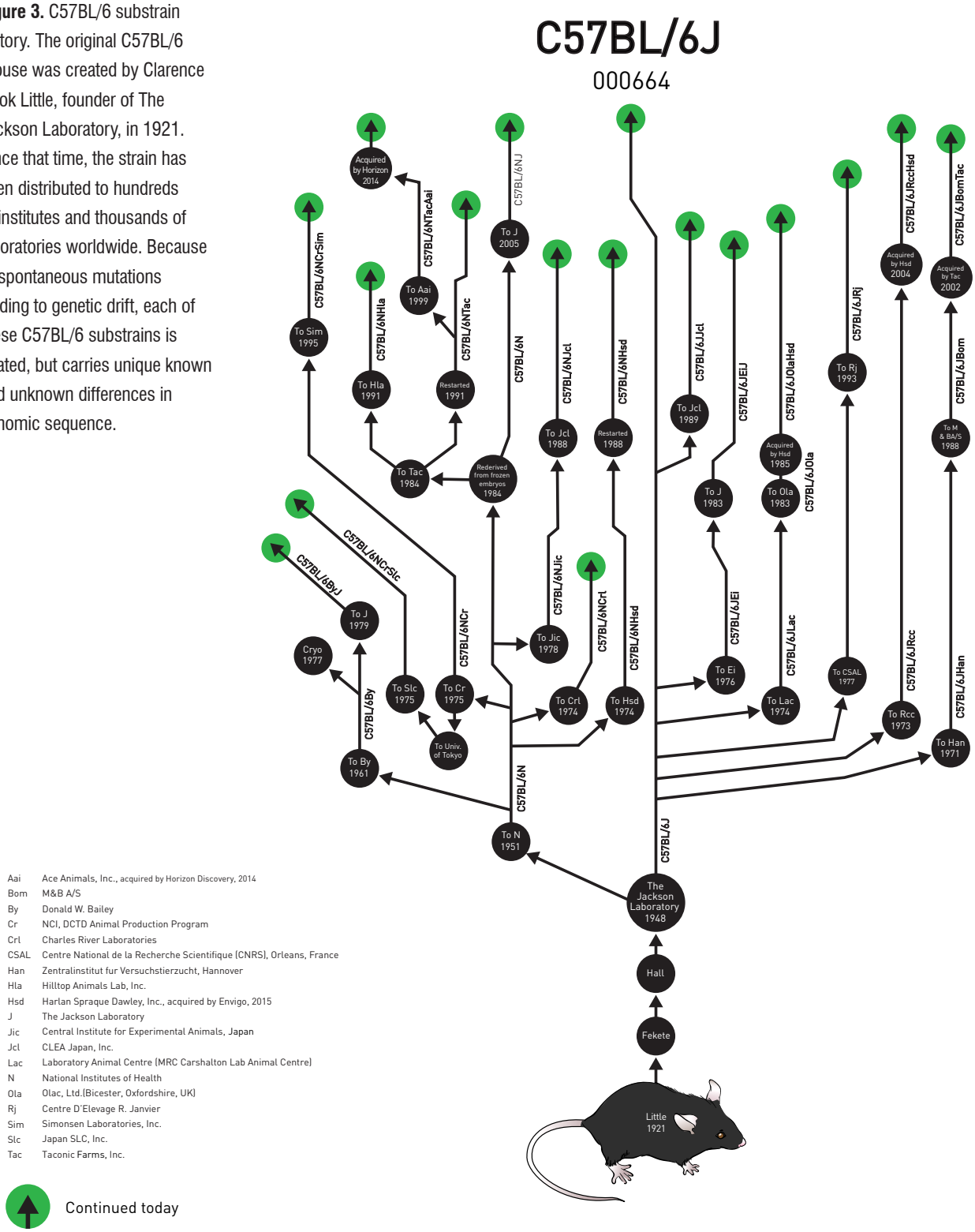
Currently there are over 16,000 entries using the original Jackson Laboratory C57BL/6J substrain. A few other entries exist for substrains derived from the original C57BL/6J. Roughly 1,200 entries use C57BL/6N-derived substrains. In the coming years, the use of C57BL/6N substrains is expected to grow significantly, as all 20,000 genes in the mouse genome will eventually be targeted in C57BL/6N ES cells through the International Knockout Mouse Consortium (IKMC) project (http://www.mousephenotype.org/).

Search Term	PubMed Entries
C57BL/6	37122
C57BL/6ByJ	112
C57BL/6J	16390
C57BL/6JOlaHsd	53
C57BL/6JBomTac	11
C57BL/6JRj	7
C57BL/6N	1182
C57BL/6NCrl	71
C57BL/6NJ	11
C57BL/6NHsd	41
C57BL/6NTac	78

Table 3. Prevalence of common C57BL/6 substrains in published literature. At the time of publication, the following search terms for individual C57BL/6 substrains were entered into the PubMed database and number of references were recorded.

The original C57BL/6J substrain from The Jackson Laboratory was sent to the National Institutes of Health (NIH) in 1951. The NIH substrain, C57BL/6N, was later distributed to several institutes, including Charles River Laboratories in 1974 (C57BL/6NCrl), to Harlan (now Envigo, C57BL/6NHsd) in 1974 and 1988, and to Taconic (C57BL/6NTac) in 1991. In 2005, the N substrain came back to The Jackson Laboratory, and is known as the C57BL/6NJ substrain. Currently, at least 100 generations of breeding separate C57BL/6J substrains from all C57BL/6N substrains (Figure 3).

Figure 3. C57BL/6 substrain history. The original C57BL/6 mouse was created by Clarence Cook Little, founder of The Jackson Laboratory, in 1921. Since that time, the strain has been distributed to hundreds of institutes and thousands of laboratories worldwide. Because of spontaneous mutations leading to genetic drift, each of these C57BL/6 substrains is related, but carries unique known and unknown differences in genomic sequence.



Aai	Ace Animals, Inc., acquired by Horizon Discovery, 2014
Bom	M&B A/S
By	Donald W. Bailey
Cr	NCI, DCTD Animal Production Program
Crl	Charles River Laboratories
CSAL	Centre National de la Recherche Scientifique (CNRS), Orleans, France
Han	Zentralinstitut fur Versuchstierzucht, Hannover
Hla	Hilltop Animals Lab, Inc.
Hsd	Harlan Spraque Dawley, Inc., acquired by Envigo, 2015
J	The Jackson Laboratory
Jic	Central Institute for Experimental Animals, Japan
Jcl	CLEA Japan, Inc.
Lac	Laboratory Animal Centre (MRC Carshalton Lab Animal Centre)
N	National Institutes of Health
Ola	Olac, Ltd.(Bicester, Oxfordshire, UK)
Rj	Centre D'Elevage R. Janvier
Sim	Simonsen Laboratories, Inc.
Slc	Japan SLC, Inc.
Tac	Taconic Farms, Inc.

Continued today

Several publications demonstrate heritable phenotypic differences between J and N substrains that have arisen due to genetic drift. Depending on the specific research question, some substrains may be preferred over others (Bryant, 2011). Some classic and recent examples are listed here:

- C57BL/6J mice express a mutant *Nnt* gene, which is involved in glucose-mediated insulin secretion, compared to C57BL/6N substrains (Freeman et al., 2006).
- C57BL/6J mice have strong preferences for alcohol while C57BL/6NCrl mice do not (Mulligan et al., 2008). Quantitative Trait Loci mapping studies comparing these substrains may lead to a better understanding of the genes involved in addiction.
- C57BL/6N substrains harbor the retinal degeneration allele *Crb*^{rd8} while the C57BL/6J substrain carries a wildtype allele (Mattapallil et al., 2012).
- C57BL/6JOlaHsd mice are homozygous for a spontaneous deletion in the genes encoding alpha synuclein and multimerin-1 (Specht and Schoepfer, 2001, 2004). While alpha synuclein aggregates in the nervous system in Parkinson's Disease, the deletion in the C57BL/6JOlaHsd substrain does not appear to contribute to prion disease-mediated synaptotoxicity (Asuni et al., 2010) but may have effects on motor neuron degeneration in general (Pelkonen and Yavich, 2011; Peña-Oliver et al., 2012). C57BL/6JOlaHsd mice also have reduced bone density compared to C57BL/6J and C57BL/6JRccHsd substrains (Liron et al., 2017).
- C57BL/6NHsd mice carry a *Dock2* mutation affecting B-cell signaling and immune tolerance that is not found in other major C57BL/6 substrains (Mahajan et al., 2016).

In this last, recent example, an approximately 10-year step back in research progress occured for one laboratory as a result of conclusions drawn from using two distinct C57BL/6 substrains over the course of its studies (www.jax.org/news-and-insights/jax-blog/2016/may/why-it-took-2-years-for-a-harvard-research-lab-to-get-back-to-research). The original studies were published using an undefined "C57BL/6" substrain as the genetic background for creating a *Siae* gene knockout (Cariappa et al., 2009). *Siae* was thought to contribute to B-cell development and signaling when initially published in 2009. The *Siae* mutation was later backcrossed to the specific C57BL/6J substrain from The Jackson Laboratory. Surprisingly, the experiments on the C57BL/6J background failed to reproduce the laboratory's previous publication (Mahajan et al., 2016). After several years of additional analysis of several commercially available C57BL/6 substrains, it was discovered that a copy number mutation in a different gene, *Dock2*, had spontaneously arisen in a strain of C57BL/6NHsd mice. *Dock2* was the actual causal mutation for these B-cell functions. This example should serve as a cautionary tale to closely monitor and understand the origins of any mice used in research. Because of genetic drift, inbred mouse substrains should not be used interchangeably.

It should be noted that in addition to the specific research question, the phenotypic effects of spontaneous mutations that have arisen due to genetic drift may depend on several contributing experimental factors. For instance, the *Nnt* mutation in C57BL/6J strains has been shown to have reduced insulin secretion *in vitro* compared to C57BL/6J mice rescued with transgenic wildtype *Nnt* (Freeman et al., 2006). In another study, no significant differences in insulin secretion were measured *in vitro* or *in vivo* in C57BL/6J and C57BL/6NTac substrains (Wong et al., 2010). Furthermore, the *Nnt* mutational status and relationship with diet-induced obesity and insulin responsivity is not straightforward, as it may depend on the fat content of the diet (Nicholson et al., 2010). Similarly, two J substrains (J, JWehi) and four N susbstrains (NTac, NHsd, NCrl, NJ) fed a low-fat diet were found to have similar insulin secretion profiles in response to glucose challenge. However, when fed a high fat diet, the C57BL/6NJ substrain demonstrated a reduced insulin response to glucose challenge that could not be explained by differences in *Nnt* status, weight gain, fat mass, food intake, or beta cell area (Hull et al., 2017).

Several other published differences exist between C57BL/6 substrains. Differences in behavior such as fear, anxiety,

pain, and response to amphetamines have been noted in the literature (summarized in Bryant et al., 2008). More broadly, differences exist across many other baseline measures. In particular, C57BL/6J and C57BL/6NTac substrains were compared in a comprehensive, standardized phenotyping pipeline of 413 parameters (EMPReSS) completed by four individual mouse centers of the European Mouse Disease Clinic (EUMODIC) consortium (Simon et al., 2013). Across the four phenotyping centers, the J and NTac mice differed in several areas including startle response, locomotor activity, grip strength, cardiovascular characteristics, metabolic parameters, and clinical chemistries.

Taken together, genetic background is one component of experimental design which may affect reproducibility and the ability to make generalizations about biological processes. Troublingly, of the nearly 37,000 entries in PubMed for "C57BL/6," the majority of these publications do not indicate a substrain.

Colony Management Practices That Limit Genetic Drift

All breeding colonies are subject to genetic drift. However, there are a number of colony management strategies that can limit drift, and therefore limit the effects of drift on experimental reproducibility. These strategies include use of proper nomenclature, thoughtful breeding practices, and cryopreservation. The following are some best practices to employ while maintaining a mouse colony.

Nomenclature and proper reporting

Use full and proper mouse strain nomenclature to remove uncertainty and allow identification of the exact substrain that is being investigated.

- **For daily colony management, use colored, preprinted labels with full nomenclature, including substrain designation(s) on cage cards and in lab notebooks.** Preprinting the labels reduces penmanship errors and improves nomenclature compliance. Use of differently colored labels or cage cards is especially important in busy, shared spaces where strains of similar nomenclature and appearance may be housed nearby.
- **Practice proper nomenclature in lab meeting presentation slides.** These casual, "unofficial" communications will eventually become "formal" communications when these figures and data are ultimately formatted for posters, talks, publications, and grant proposals.
- **In publications and grant proposals, use full nomenclature, including substrain designation(s), the first time the strain is mentioned.** Define how the strain will then be abbreviated in the text and figures ("hereafter referred to as…"). In the methods section, use full nomenclature and substrain designation(s). Identify the source of the strain such as lab name, institute, or vendor and catalog number of the strain. Include generation numbers and breeding schemes employed (see below). For further suggestions, consult the ARRIVE guidelines (www.nc3rs.org.uk/arrive-guidelines).

Inbreeding, pedigrees, and generation numbers

Inbreeding allows for faster identification of deviant phenotypes in a colony. Pedigrees allow affected and potentially affected inbred mice to be removed easily (Figure 4). Generation numbers allow quick identification of potential risk of genetic drift in the colony.

- **Inbreeding** - Only mate brothers and sisters.
- **Pedigrees** - Record dam (female) and sire (male) used in each breeding. Keep two or more pedigrees within a colony, never mixing breeders from one pedigree with another.
- **Generation numbers**
  - N = Number of backcross generation(s)
  - F = Filial, inbred (sister x brother) generation(s)
  - p = Cryopreserved
  - + = Separates generation information prior to importation
  - ? = Unknown generation number

For example, "N6F12+F8" refers to a strain that was backcrossed 6 times, filial mated 12 times, and transferred to another laboratory where it was filial mated 8 more times. Given that substrains arise in 20 generations of consecutive inbreeding and genetic drift has likely occurred, it would be wise to refresh the genetic background of this example colony.

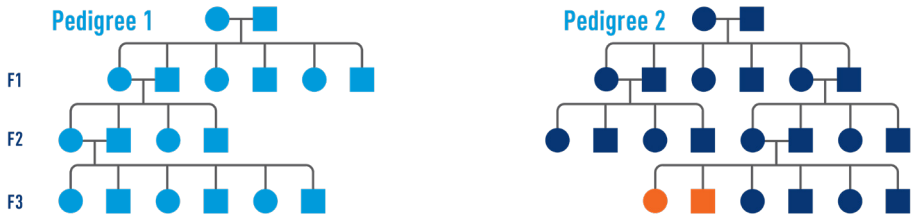


Figure 4. Maintaining a pedigreed colony. Only mate sister-brother (circle-square) pairs or trios in two or more pedigrees (light blue vs. dark blue), never mixing breeders from one pedigree with the other. If aberrant phenotypes (orange) arise in one pedigree, affected individuals are more easily identified and removed. The unaffected pedigree (Pedigree 1, light blue) can then be divided into two new pedigrees, without losing time rebuilding the colony *de novo*.

Data collection and regular evaluation

In addition to breeding practices, strains should regularly be observed for any changes in phenotype. Where genetic drift is concerned, changes in phenotype can mean anything observable or measurable: appearance, behavior, breeding performance, or experimental readout, to name a few examples. Identifying genetic drift relies on colonists and researchers to first notice phenotypic change, and second, to do something about it.

For some strains, comparison to baseline characteristics may help. The Mouse Phenome Database may have such information, which is searchable by strain or phenotype and includes all data collection protocols (http://phenome.jax.org) (Table 2).

If a phenotype has changed in a colony, genetic drift is one of many potential sources of variability to investigate. Some questions to consider:

- How many mice are affected and can they be traced to any particular cage or pedigree?
- How many years or generations has the colony been in the facility?
- When was the last refresh (see the following section) and what was the source of mice used?

Without careful colony notes or reference data, it may be impossible to determine whether a phenotype has changed.

Refreshing the genetic background

After 5-10 generations of inbreeding, mouse colonies should be refreshed, to remove or prevent genetic drift accumulation in the colony. Methods to refresh the genetic background may include the following:

- **Backcross.** Genetically Engineered Mutant Mouse (GEMM) strains may be backcrossed to the appropriate inbred or hybrid mouse strain purchased from a reputable mouse repository or vendor who practices methods to limit drift in their colonies. Backcrossing should be done through both the male and female germ lines to ensure both sex chromosomes are refreshed. If the strain is already being crossed as heterozygous or hemizygous to wildtype, using an inbred mouse directly from the vendor as the wildtype breeder serves to refresh the genetic background. When notating generation number, each backcross or refresh serves as an additional "N" (see the previous "Inbreeding, pedigrees, and generation numbers" section).
- **Purchase new breeders.** For inbred strains, the colony should be restarted with new breeder mice purchased directly from a trusted mouse repository or vendor who practices appropriate methods to limit genetic drift in their colonies.
- **Cryorecover from frozen stock.** The only method to stop genetic drift is to stop breeding mice. Low use and unique mouse strains should always be cryopreserved as either sperm or embryos to protect against genetic drift, ensure against loss of a strain, and to reduce animal use and maintenance costs. This cryopreserved material can be used to recover a colony that has experienced drift or breeding errors, or was lost to disease or natural disasters.

Verifying Genetic Background

- **Perform a genome scan to determine risk of contamination.** Genome scans or SNP arrays may allow differentiation between closely related substrains such as C57BL/6J vs C57BL/6N.
- **Sequence the genome.** SNP arrays will not identify genetic drift within a colony. The only way to know if a strain has undergone drift is to fully sequence its genome and compare to a reference sequence.

Advanced Methods to Limit Genetic Drift

If genetic drift occurs in any actively breeding colony, why would individual laboratories be able to refresh their colonies by repurchasing mice from a mouse repository or a vendor who would also experience drift?

Mouse repositories and vendors, for one, maintain much larger colonies that are less subject to genetic population bottlenecks than small colonies (Figure 1A) . Additionally, many repositories and vendors professionally practice the above-mentioned colony management strategies such as complete nomenclature, pedigrees/limited breeding, and cryopreservation, as well as more advanced methods.

To estimate genetic drift in the large production colonies at The Jackson Laboratory, C57BL/6J mice separated by 69 inbred generations and 19 years of continuous breeding were sequenced. Between these two snapshots in time, 669 unique SNPs were identified. Of these SNPs, seven changed the amino acid sequence or altered an RNA splice site. Thus, an estimated one mutation with potential impact on protein function occurs every 10 generations (7 SNPs/69 generations), not including larger perturbations such as deletions, inversions, and duplications, which may have profound phenotypic consequences. Considering the average graduate student or postdoctoral fellow's tenure in a lab may last five years, a Principal Investigator's individual research career may stretch 20 or more years, and scientific research as a whole will continue indefinitely, an individual mouse colony may experience significant genetic drift. As a repository that distributes mice worldwide over several cumulative years, The Jackson Laboratory is uniquely challenged to limit genetic drift as much as possible, so that researchers using these strains may continue to rely on a stable, reproducible genome.

The Jackson Laboratory strains are protected from accumulated genetic drift through a combination of several practices. All strains that carry The Jackson Laboratory "J" laboratory code are maintained through one or both of two programs designed to limit and detect genetic drift: the Genetic Stability Program (GSP) and the Genetic Quality Control Program. Strains that are distributed as "J" strains include all strains propagated and distributed directly from The Jackson Laboratory facilities in The United States as well as "J" strains propagated and distributed by Charles River Laboratories in Europe and Japan. To maintain continuity across these sites, colony management practices are reviewed regularly and approved by The Jackson Laboratory. This includes the use of identical cryopreserved material to either regularly infuse into existing live colonies (GSP described in the following section) or to recreate living colonies entirely. In sum, these actions effectively prevent substrain divergence.

Genetic Stability Program (GSP) for the most common inbred "J" strains

The most commonly used inbred "J" strains are maintained using a unique strategy that actively prevents the accumulation of genetic drift. The US Patent and Trademark Office awarded patents for The Jackson Laboratory's Genetic Stability Program (GSP) in 2009 and 2012 (Wiles et al., 2009, 2012, https://www.jax.org/jax-mice-and-services/find-and-order-jax-mice/why-jax-mice/patented-genetic-stability-program). GSP strains have been cryopreserved as two-cell embryos and are regularly re-infused into pedigreed foundation colonies to avoid cumulative genetic drift.

Without the GSP practice, a foundation breeding colony would be derived from a single brother-sister mating. Two to four times a year, a new brother-sister pair would be selected from the foundation colony as the new founder pair to re-establish the colony. Using this approach, a foundation colony today would be genetically different from the foundation colony years from now, because of genetic drift.

Under the GSP practice, pedigree-tracked stocks of cryopreserved embryos are derived from a single

foundation colony. The stock of embryos is used to re-establish the live breeding foundation colony every five generations. Periodically re-establishing the foundation colony with mice recovered from the cryopreserved embryos reduces the numbers of generations passed for a given time period. Therefore, "J" strains under the GSP practice are protected against genetic drift across space (at different geographical facilities), but also, importantly, across time (Table 3).

Genetic Quality Control Program

In addition to the GSP practices, all "J" strains are maintained through a Genetic Quality Control (GQC) Program (https://www.jax.org/jax-mice-and-services/find-and-order-jax-mice/why-jax-mice/genetic-quality-control-program). This program integrates many of the typical colony management practices that individual labs may use (as described earlier), but includes a very high degree of accountability.

Animal care professionals undergo a rigorous training program to identify phenotypic variants such as coat color, unusual body size, weight, skeletal structure, behavior, reproductive performance, tumor susceptibility, and life span. Any mice that deviate from the characteristics expected for a particular strain are further investigated and pedigrees can be traced and removed as needed.

Additionally, pedigreed lines are maintained in foundation colonies that are separate from expansion and distribution colonies. The pedigreed lines are regularly screened for genetic anomalies or evidence of genetic contamination using an SNP panel based on Petkov et al., 2004.

JAX™ Mice Bred by Charles River in Europe and Japan

The Jackson Laboratory and Charles River have a cooperative agreement to provide local supplies of JAX™ Mice to biomedical researchers in Europe, Japan, Korea, and Taiwan. Following strict adherence to The Jackson Laboratory's breeding protocols and genetic quality control guidelines, Charles River breeds JAX™ Mice in Europe and Japan that are equivalent in genetic quality to those bred by The Jackson Laboratory. For further information please see www.criver.com/jaxmice.

Conclusion

Genetic drift is an inevitable reality in actively breeding mouse colonies and may deeply impact research conclusions and reproducibility. While genetic drift cannot be eliminated completely, colony management strategies can be implemented both in individual laboratories and in large mouse repositories and vendors to maintain genetic stability. Reproducibility and scientific discovery rely on careful reporting of complete mouse substrain nomenclature and breeding information.

References

Adams, D.J., Doran, A.G., Lilue, J., and Keane, T.M. (2015). The Mouse Genomes Project: a repository of inbred laboratory mouse strain genomes. Mamm. Genome Off. J. Int. Mamm. Genome Soc. *26*, 403–412.

Asuni, A.A., Hilton, K., Siskova, Z., Lunnon, K., Reynolds, R., Perry, V.H., and O'Connor, V. (2010). Alpha-synuclein deficiency in the C57BL/6JOlaHsd strain does not modify disease progression in the ME7-model of prion disease. Neuroscience *165*, 662–674.

Boleij, H., Salomons, A.R., van Sprundel, M., Arndt, S.S., and Ohl, F. (2012). Not all mice are equal: welfare implications of behavioural habituation profiles in four 129 mouse substrains. PloS One 7, e42544.

Bryant, C.D. (2011). The blessings and curses of C57BL r 129 mouse substrains. PloS One 7, e42544. /6 substrains in mouse genetic studies. Ann. N. Y. Acad. Sci. *1245*, 31.

Bryant, C.D., Zhang, N.N., Sokoloff, G., Fanselow, M.S., Ennes, H.S., Palmer, A.A., and McRoberts, J.A. (2008). Behavioral differences among C57BL/6 substrains: implications for transgenic and knockout studies. J. Neurogenet. *22*, 315–331.

Cariappa A., Takematsu H., Liu H., Diaz S., Haider K., Boboila C., Kalloo G., Connole M., Shi H.N., Varki N., Varki A., Pillai S. (2008). B cell antigen receptor signal strength and peripheral B cell development are regulated by a 9-O-acetyl sialic acid esterase. J Exp Med. 2009 Jan 16;206(1):125-38.

Chamary, J.-V., and Hurst, L.D. (2004). Similar Rates but Different Modes of Sequence Evolution in Introns and at Exonic Silent Sites in Rodents: Evidence for Selectively Driven Codon Usage. Mol. Biol. Evol. *21*, 1014–1023.

Chinwalla, A.T., Cook, L.L., Delehaunty, K.D., Fewell, G.A., Fulton, L.A., Fulton, R.S., Graves, T.A., Hillier, L.W., Mardis, E.R., McPherson, J.D., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–562.

Drake, J.W., Charlesworth, B., Charlesworth, D., and Crow, J.F. (1998). Rates of Spontaneous Mutation. Genetics *148*, 1667–1686.

Frazer, K.A., Eskin, E., Kang, H .M., Bogue, M.A., Hinds, D.A., Beilharz, E.J., Gupta, R.V., Montgomery, J., Morenzoni, M.M., Nilsen, G.B., et al. (2007). A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature *448*, 1050–1053.

Freeman, H.C., Hugill, A., Dear, N.T., Ashcroft, F.M., and Cox, R.D. (2006). Deletion of nicotinamide nucleotide transhydrogenase: a new quantitive trait locus accounting for glucose intolerance in C57BL/6J mice. Diabetes *55*, 2153–2156.

Hull, R.L., Willard, J.R., Struck, M.D., Barrow, B.M., Brar, G.S., Andrikopoulos, S., and Zraika, S. (2017). High fat feeding unmasks variable insulin responses in male C57BL/6 mouse substrains. J. Endocrinol. *233*, 53–64.

Liron, T., Raphael, B., Hiram-Bab, S., Bab, I.A., and Gabet, Y. (2017). Bone Loss in C57BL/6J-OlaHsd Mice, a Substrain of C57BL/6J Carrying Mutated Alpha-Synuclein and Multimerin-1 Genes. J. Cell. Physiol.

Mahajan, V.S., Demissie, E., Mattoo, H., Viswanadham, V., Varki, A., Morris, R., and Pillai, S. (2016). Striking Immune Phenotypes in Gene-Targeted Mice Are Driven by a Copy-Number Variant Originating from a Commercially Available C57BL/6 Strain. Cell Rep. *15*, 1901–1909.

Mattapallil, M.J., Wawrousek, E.F., Chan, C.-C., Zhao, H., Roychoudhury, J., Ferguson, T.A., and Caspi, R.R. (2012). The Rd8 mutation of the Crb1 gene is present in vendor lines of C57BL/6N mice and embryonic stem cells, and confounds ocular induced mutant phenotypes. Invest. Ophthalmol. Vis. Sci. *53*, 2921–2927.

Mulligan, M.K., Ponomarev, I., Boehm, S.L., Owen, J.A., Levin, P.S., Berman, A.E., Blednov, Y.A., Crabbe, J.C., Williams, R.W., Miles, M.F., et al. (2008). Alcohol trait and transcriptional genomic analysis of C57BL/6 substrains. Genes Brain Behav. *7*, 677–689.

Nicholson, A., Reifsnyder, P.C., Malcolm, R.D., Lucas, C.A., MacGregor, G.R., Zhang, W., and Leiter, E.H. (2010). Diet-induced obesity in two C57BL/6 substrains with intact or mutant nicotinamide nucleotide transhydrogenase (Nnt) gene. Obes. Silver Spring Md *18*, 1902–1905.

Pelkonen, A., and Yavich, L. (2011). Neuromuscular pathology in mice lacking alpha-synuclein. Neurosci. Lett. *487*, 350–353.

Peña-Oliver, Y., Buchman, V.L., Dalley, J.W., Robbins, T.W., Schumann, G., Ripley, T.L., King, S.L., and Stephens, D.N. (2012). Deletion of alpha-synuclein decreases impulsivity in mice. Genes Brain Behav. *11*, 137–146.

Petkov, P.M., Cassell, M.A., Sargent, E.E., Donnelly, C.J., Robinson, P., Crew, V., Asquith, S., Haar, R.V., and Wiles, M.V. (2004). Development of a SNP genotyping panel for genetic monitoring of the laboratory mouse. Genomics *83*, 902–911.

Poltorak, A., Smirnova, I., He, X., Liu, M.-Y., Van Huffel, C., Birdwell, D., Alejos, E., Silva, M., Du, X., Thompson, P., et al. (1998a). Genetic and Physical Mapping of the Lps Locus: Identification of the Toll-4 Receptor as a Candidate Gene in the Critical Region. Blood Cells. Mol. Dis. *24*, 340–355.

Poltorak, A., He, X., Smirnova, I., Liu, M.-Y., Huffel, C.V., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., et al. (1998b). Defective LPS Signaling in C3H/HeJ and C57BL/10ScCr Mice: Mutations in Tlr4 Gene. Science *282*, 2085–2088.

Russell, L.B., and Russell, W.L. (1996). Spontaneous mutations recovered as mosaics in the mouse specific-locus test. Proc. Natl. Acad. Sci. *93*, 13072–13077.

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

Simon, M.M., Greenaway, S., White, J.K., Fuchs, H., Gailus-Durner, V., Wells, S., Sorg, T., Wong, K., Bedu, E., Cartwright, E.J., et al. (2013). A comparative phenotypic and genomic analysis of C57BL/6J and C57BL/6N mouse strains. Genome Biol. *14*, R82.

Specht, C.G., and Schoepfer, R. (2001). Deletion of the alpha-synuclein locus in a subpopulation of C57BL/6J inbred mice. BMC Neurosci. *2*, 11.

Specht, C.G., and Schoepfer, R. (2004). Deletion of multimerin-1 in alpha-synuclein-deficient mice. Genomics *83*, 1176–1178.

Watson, J., Kelly, K., Largen, M., and Taylor, B.A. (1978). The Genetic Mapping of a Defective Lps Response Gene in C3H/HeJ Mice. J. Immunol. *120*, 422–424.

Wiles, M.V., Taft, R., and Eicher, E.M. (2009). Methods for maintaining genetic stability of inbred animal strains.

Wiles, M.V., Taft, R., and Eicher, E.M. (2012). Methods for maintaining genetic stability of inbred animal strains.

Wong, N., Blair, A.R., Morahan, G., and Andrikopoulos, S. (2010). The deletion variant of nicotinamide nucleotide transhydrogenase (Nnt) does not affect insulin secretion or glucose tolerance. Endocrinology *151*, 96–102.