Published in final edited form as:

Nat Methods. 2018 December ; 15(12): 994–996. doi:10.1038/s41592-018-0224-7.

Comparing phenotypic variation between inbred and outbred mice

Alexander H. Tuttle¹, Vivek M. Philip², Elissa J. Chesler², and Jeffrey S. Mogil^{3,*}

¹UNC Neuroscience Centre, University of North Carolina at Chapel Hill, NC U.S.A.

²The Jackson Laboratory, Bar Harbor, ME USA

³Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC CANADA

Abstract

Inbred mice are preferred over outbred mice with the assumption that they display less trait variability. We compared coefficients of variation and failed to find evidence of greater trait stability in inbred mice. We conclude that contrary to conventional wisdom, outbred mice may be better subjects for most biomedical research.

The laboratory mouse is the most commonly used experimental non-human subject in biomedical research¹. For many decades, inbred mouse strains have been preferred over outbred stocks, with particular strains such as C57BL/6 and BALB/c used in wide-ranging biomedical applications². Inbred mice are preferentially chosen for immunological studies (to prevent alloimmune responses), population genetic mapping (to allow diallele crosses for known genetic markers to be used), and molecular genetic studies (to avoid background effects in mutagenesis and transgenics). The more general preference for inbred strains across biomedicine stems from the conventional wisdom that within-strain phenotypic variability observed in these animals should be lower than that associated with a sample of outbred animals, because in any given inbred strain phenotypic variability (V_p) equals environmental variability (V_e), whereas in outbred animals genetic variability (V_g) is present in addition to V_e and gene x environment interaction (V_{ge}). Based on this assumption, fewer inbred mice would be required to maintain similar statistical power, presenting practical and ethical advantages. However, the evidence for lower phenotypic variability amongst inbred mice is mixed, with some early (e.g., ³) and modern (e.g., ⁴) studies explicitly suggesting otherwise. Nonetheless, the idea that genetic heterogeneity leads to larger phenotypic variability is compelling, with proponents arguing against the use of outbred stocks in biomedical research (e.g., ^{5,6,7}).

The differences between inbred and outbred mice that might affect their relative utility as experimental subjects are well-documented. These characteristics were often selected in the

*Address for correspondence: Jeffrey S. Mogil, Ph.D. jeffrey.mogil@mcgill.ca.

Competing Financial Interests

None of the authors has any competing financial interests.

Data Availability

Data used in this manuscript are provided as Supplementary Information.

development of the inbred strains as models for particular diseases. Inbreeding depression leads to poor fecundity; for example, inbred breeders produce between 3 and 9 pups per litter on average⁸, whereas CD-1® (ICR) breeders produce 12 pups per litter⁹. Inbred mice are also extremely small compared to outbred mice (average 70 day-old male weights: inbred, 25.4 g; outbred, 34.7 g)⁷. Other aberrations compared to outbred (and wild-derived) mice are less well known and include abnormal stress/anxiety and aggression responses^{10,11}, activity levels¹², and social behavior^{13,14}. By strongly selecting for nonlethal alleles in the homozygous state, the development of inbred mice may have lead to a relaxation of selection pressure on otherwise fitness-related genes, resulting in strains that are “not merely idiosyncratic, but idiosyncratically debilitated”¹⁵. It is of interest that researchers who use rats as subjects prefer outbred stocks, such as Sprague Dawley, Long-Evans and Wistar, even though inbred rat strains are readily available².

To evaluate whether inbred lines or outbred stocks exhibit more phenotypic variability, we performed a systematic review of the primary literature and compared means and measures of variability from extracted data in which inbred and outbred mice were tested for the same trait contemporaneously. Our approach was similar to a previous effort comparing variability in male versus female mice¹⁶. We began by searching the primary English-language biomedical literature available on MEDLINE® using the MeSH terms “Mice”, “Animals, Inbred Strains” and “Animals, Outbred Strains”, with an AND logical operator. To broaden our search, we also used text strings corresponding to popular inbred strains (e.g., “C57 OR B6”, “BALB”, “C3H”, “DBA”) and outbred stocks (“Swiss”, “ICR”, “SW”, “CD-1 OR CD1”). Finally, we supplemented our search results with papers from the senior author’s database that fit our criteria. Our search was completed in April 2017, consisting of 107 published articles, with 741 distinct inbred/outbred CV comparisons. Raw data are provided as Supplementary Table 1. Inclusion criteria (Supplementary Fig. 1) were based on the presence of quantitative data—including means, either standard deviation (SD) or standard error of the mean (SEM), and sample sizes (*n*)—from intact, adult mice of at least one commercially available inbred strain and at least one outbred mouse stock representing the results of a single experiment. In papers with more than one eligible data set, we limited CV reporting to the first three reported measures to avoid data oversampling, excluding some available data in 48 papers. Data were also excluded if either means or variances equalled zero. The lead author (A.H.T.) extracted available quantitative data from the text (if available) or from graphs using xyscan software (New Haven, CT), as well as information pertaining to animal genotype, sex, housing, and specific experimental phenotype.

To test the hypothesis that the variability within strain differs by strain type, we categorized strains into one of two “strain types”—Inbred (I) or Outbred (O). For each published data set we assigned one of seven trait categorical labels—Anatomy, Behavior/CNS, Behavior/Other, Immune Function, Molecules, Organ Function—yielding approximately equal partitioning of the available data sets (*n*=31–50/category). We then determined the coefficient of variation ($CV = SD/mean$) for each strain reported in all studies. Using general linear mixed models we assessed the main effect of strain type (I or O), with strain and/or study as random effects. We found that strain type did not have an effect on variability within strain regardless of trait category (Fig. 1), the particular selection of mouse strains included in the

study, or the study itself. In each of these models, there was no significant effect of strain type on within-strain CV (all $p > 0.05$; Supplementary Table 2).

In addition to assessing phenotypic data in the primary literature, we tested our hypothesis directly using a large phenotyping dataset collected as part of a study on the new Diversity Outbred (DO) population at The Jackson Laboratory¹⁷. We directly compared CVs from (inbred) founder strains and the outbred DO populations derived from that founder stock on contemporaneously collected measures, using data publicly available in the Mouse Phenome Database (<https://phenome.jax.org/>). We compared the inbred CVs, computed per strain and per measure, to the average DO CV obtained from 1000 bootstrap subsamples. Each DO subsample was constrained to 8 males and 8 females, thereby allowing us to compare CVs computed with equal sample sizes. For each measure we performed two one-sided t -tests, testing if the inbred CV was greater than or less than that of the DO CV, adjusting p -values for multiple testing with a significance threshold of $FDR < 0.05$.

As shown in Figure 2 (also see Supplementary Table 3), in only 6/26 measures was the inbred CV lower than the outbred CV. For 3/26 measures, the inbred CV was greater than the DO CV. For all other measures there was no statistically significant difference. Thus, in most instances the inbred population is not less noisy than outbreds. This also demonstrates that DO mice are phenotypically stable for experiments using sample sizes comparable to inbred strain studies.

The expectation that inbred strains, lacking genetic heterogeneity, would display less phenotypic variation than outbred stocks is certainly a reasonable one. However, our review of the literature shows that outbred stocks are not, in fact, necessarily more phenotypically variable than inbred strains in the same experiments. Furthermore, our analysis of the DO data set demonstrates that this is true also when precisely the same pool of known genetic variation is being compared in the inbred and outbred populations. The phenotypic variability demonstrated by inbred mice may reflect the unusual (and unnatural) condition of fixed allelic states and allelic interactions. In response to environmental perturbation (e.g.,¹⁸), the inbred organism has a single state response, whereas the outbred organism has multiple allelic variants throughout each biological pathway, which may be tuned to respond to manipulations¹⁹. Background genetic variability may have a stabilizing effect on phenotypic endpoints, buffering the organism from the idiosyncratic influences of environmental variability, and other challenges inherent in experimental manipulations, that can affect outcomes within an isogenic stock²⁰. Therefore, the apparent paradox that the phenotypic variance for inbred strains, attributable only to V_e , is equal to outbred phenotypic variance, attributable to genetic and environmental sources ($V_g + V_e + V_{ge}$), may be due to an expansion of the organismal response to the environment that inflates V_e in inbred stocks.

If genetic variability is a stabilizing force in the determination of phenotypic outcomes, observable variability demonstrated by outbred populations could be defined as an escape from these buffering mechanisms to stabilize the phenotypic endpoints in the face of experimental or environmental perturbation. This provides a tractable framework for further genetic analyses to investigate the basis of genotype x environment interactions and other sources of individual variability in response to experimental treatment. We therefore

conclude that defined outbred stocks from heterogeneous backgrounds (even considering the fact that commercially available outbred stocks are far less genetically diverse than wild mice) are more appropriate and much more cost-effective research subjects in many biomedical research applications, except in cases where precise genotypic regulation or standardization is required (i.e., immunological, genetic, or molecular genetic applications, or where the idiosyncratic inbred strain is a recognized disease model). For most applications, the use of robust and diverse subjects is preferred, so that conclusions obtained will be maximally generalizable across conditions and populations. Such lack of generalizability is almost certainly a major contributor to the current replication crisis; contrary to expectation, the adoption of outbred mice as research subjects may improve future experimental replicability.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements and Contributions

This work was supported by funding from the Canadian Institutes for Health Research (FRN154281 to J.S.M.), the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03873 to J.S.M.), the Louise and Alan Edwards Foundation (J.S.M.), and the NIH National Institute on Drug Abuse (5 P50 DA039841 to E.J.C.).

References

1. Taylor K, Gordon N, Langley G & Higgins W Altern. Lab Anim 36, 327–342 (2008). [PubMed: 18662096]
2. Festing MFW ILAR J. 55, 399–404 (2014). [PubMed: 25541542]
3. Biggers JD & Claringbold PJ Nature 174, 596–597 (1954). [PubMed: 13203582]
4. Jensen VS, Porsgaard T, Lykkesfeldt J & Hvid H Am. J. Transl. Res 8, 3574–3584 (2016). [PubMed: 27648148]
5. Festing MFW Toxicol. Pathol 38, 681–690 (2010). [PubMed: 20562325]
6. Festing MFW Neurobiol. Aging 20, 237–244 (1999). [PubMed: 10537033]
7. Chia R, Achilli F, Festing MFW & Fisher EMC Nat. Genet 37, 1181–1186 (2005). [PubMed: 16254564]
8. Murray SA et al. PLoS One 5, e12418 (2010). [PubMed: 20811634]
9. Tanaka T Reprod. Toxicol 12, 613–617 (1998). [PubMed: 9875696]
10. Chalfin L et al. Nat. Commun 5, 4569 (2014). [PubMed: 25090970]
11. Fonio E, Golani I & Benjamini Y Nat. Meth 9, 1167–1170 (2012).
12. Dohm MR, Richardson CS & Garland T Jr. Am. J. Physiol 267, R1098–R1108 (1994). [PubMed: 7943421]
13. Nevison CM, Barnard CJ & Hurst JL Appl. Anim. Behav. Sci 81, 387–398 (2003).
14. Tuttle AH et al. Proc. Natl. Acad. Sci. U. S. A 114, 5515–5520 (2017). [PubMed: 28484016]
15. Miller RA et al. Neurobiol. Aging 20, 217–231 (1999). [PubMed: 10537031]
16. Prendergast BJ, Onishi KG & Zucker I Neurosci. Biobehav. Rev 40, 1–5 (2014). [PubMed: 24456941]
17. Logan RW et al. Genes Brain Behav. 12, 424–437 (2013). [PubMed: 23433259]
18. Mogil JS Lab Anim. 46, 136–141 (2017).
19. Carter GW, Hays M, Sherman A & Galitski T PLoS Genet. 8, e1003010 (2012). [PubMed: 23071457]
20. Phelan JP & Austad SN J. Gerontol 49, B1–B11 (1994). [PubMed: 8282970]

Fig. 1.

Coefficients of variation of all available studies in which inbred and outbred mice are directly compared. Data are presented overall (241 data sets) and divided into behavioral and non-behavioral measures (**a**) and broken down by trait category (**b**; see Supplementary Table 1). Each symbol represents a single inbred or outbred strain within a single experiment. P values are from models m1.a, m2.a and m3 in Supplementary Table 2.

Fig. 2.

Coefficients of variation (CV) within each inbred strain and the average CV across 1000 subsamples of the DO population. Each subsample consists of 16 mice (8 male; 8 female) to resemble the composition of each inbred strain studied. The error bars correspond to the standard deviation of DO subsample CVs. Asterisks or plus signs indicate significant differences where $FDR < 0.05$ for the comparison of average inbred CV to the DO CV using a one sided t -test; *, average inbred CV $<$ DO CV, +, average inbred CV $>$ DO CV, -, average inbred CV not significantly different from DO CV. Trait descriptions can be found in Supplementary Table 3.