

A Diallel of the Mouse Collaborative Cross Founders Reveals Strong Strain-Specific Maternal Effects on Litter Size

John R. Shorter,^{*,1} Paul L. Maurizio,^{*,†,1} Timothy A. Bell,^{*,‡} Ginger D. Shaw,^{*} Darla R. Miller,^{*} Terry J. Gooch,^{*} Jason S. Spence,^{*} Leonard McMillan,[§] William Valdar,^{*,‡,2} and Fernando Pardo-Manuel de Villena^{*,‡,2}

^{*}Department of Genetics, [†]Curriculum in Bioinformatics and Computational Biology, [‡]Lineberger Comprehensive Cancer Center, and [§]Department of Computer Science, University of North Carolina, Chapel Hill, North Carolina, 27599-7265

ORCID IDs: 0000-0003-4732-5526 (J.R.S.); 0000-0002-5859-6260 (P.L.M.); 0000-0002-2419-0430 (W.V.); 0000-0002-5738-5795 (F.P.-M.d.V.)

ABSTRACT Reproductive success in the eight founder strains of the Collaborative Cross (CC) was measured using a diallel-mating scheme. Over a 48-month period we generated 4,448 litters, and provided 24,782 weaned pups for use in 16 different published experiments. We identified factors that affect the average litter size in a cross by estimating the overall contribution of parent-of-origin, heterosis, inbred, and epistatic effects using a Bayesian zero-truncated overdispersed Poisson mixed model. The phenotypic variance of litter size has a substantial contribution (82%) from unexplained and environmental sources, but no detectable effect of seasonality. Most of the explained variance was due to additive effects (9.2%) and parental sex (maternal vs. paternal strain; 5.8%), with epistasis accounting for 3.4%. Within the parental effects, the effect of the dam's strain explained more than the sire's strain (13.2% vs. 1.8%), and the dam's strain effects account for 74.2% of total variation explained. Dams from strains C57BL/6J and NOD/ShiLtJ increased the expected litter size by a mean of 1.66 and 1.79 pups, whereas dams from strains WSB/EiJ, PWK/PhJ, and CAST/EiJ reduced expected litter size by a mean of 1.51, 0.81, and 0.90 pups. Finally, there was no strong evidence for strain-specific effects on sex ratio distortion. Overall, these results demonstrate that strains vary substantially in their reproductive ability depending on their genetic background, and that litter size is largely determined by dam's strain rather than sire's strain effects, as expected. This analysis adds to our understanding of factors that influence litter size in mammals, and also helps to explain breeding successes and failures in the extinct lines and surviving CC strains.

KEYWORDS

F1 cross
additive heritability
fertility
BayesDiallel
zero-truncated
Poisson
variance projection
Multiparent
Advanced
Generation Inter-
Cross (MAGIC)
multiparental
populations
MPP

A critical part of fitness is optimizing the balance between numbers of offspring, offspring body size, and reproductive timing, leading to strong selection for these traits (Smith and Fretwell 1974; Beauchamp and

Kacelnik 1990). Fundamental fitness traits, like litter size (*i.e.*, the number of pups born per litter), should have low contributions of genetic variance due to rapid fixation of alleles that increase fitness (Mousseau and Roff 1987; Price and Schluter 1991; Merilä and Sheldon 1999; Stirling *et al.* 2002; Peripato *et al.* 2004). Estimates of litter size heritability vary from 5 to 20% in mouse, rabbit, and porcine populations (Falconer 1960; Avalos and Smith 1987; Falconer and Mackay 1996; Johnson *et al.* 1999; García and Baselga 2002; Peripato *et al.* 2004; Gutiérrez *et al.* 2006) leading to the majority of variation in litter size being non-heritable. Non-heritable factors include seasonality, food availability, paternal and maternal effects, dominance and epistasis, as well as unknown factors. Additionally, sexual conflict over parental investment may exist between females and males, and this in turn could affect litter size (Trivers 1974; Hager and Johnstone 2003). These factors make it difficult to obtain clear measurements of litter

Copyright © 2019 Shorter *et al.*

doi: <https://doi.org/10.1534/g3.118.200847>

Manuscript received October 31, 2018; accepted for publication March 14, 2019; published Early Online March 15, 2019.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at Figshare: <https://figshare.com/s/fefa6216bf29b2baeb65>.

¹These authors contributed equally to this work.

²Corresponding authors: University of North Carolina, CB# 7264, Chapel Hill, NC 27599. E-mail: william.valdar@unc.edu, fernando@med.unc.edu

size in wild populations. A controlled laboratory environment is therefore necessary to measure genetic effects on litter size (Jinks and Broadhurst 1963; Roberts 1960; Bandy and Eisen 1984; Hoornbeek 1968; Peripato *et al.* 2004; Gutiérrez *et al.* 2006; Varona and Sorensen 2014).

The Collaborative Cross (CC) and its eight founder strains are an important resource for studying complex traits, for establishing mouse models for human disease, and for understanding the mouse Diversity Outbred (DO), which originated from the CC (Ferris *et al.* 2013; Chesler 2014; Rogala *et al.* 2014; Rasmussen *et al.* 2014; Gralinski *et al.* 2015, 2017; Schoenrock *et al.* 2017; Maurizio *et al.* 2018). The CC and its founder strains are also useful for studying reproductive ability due to their established record of breeding successes and failures. For example, litter size in the early CC lines shows a steady decline over the first six generations of inbreeding (Philip *et al.* 2011). In addition, nearly 95% of all CC lines have become extinct, primarily due to subspecific genomic incompatibilities (Shorter *et al.* 2017). Although standard reproductive phenotypes have been measured, mostly in male CC founders (Odet *et al.* 2015), the genetic and non-genetic control over CC breeding success has never been thoroughly characterized, and is likely to contain vital information that may be used to improve CC breeding in the future.

Given the growing popularity of the CC as a community resource, we investigated reproductive ability across the eight founders of the CC to better understand the genetic and non-genetic factors that affect their fertility and breeding. Using an 8×8 diallel design, we measured weaned litter sizes from 4,448 litters arising from 62 crosses across four years of breeding. Adapting a recently developed statistical model of diallel effects (Lenarcic *et al.* 2012), we quantified both the genetic contributions that shape litter size and the contributions of several environmental factors. Our results provide a detailed account of breeding patterns across and between the eight founders. This experiment also informs us about genetic combinations that are highly or marginally productive in the CC, helping us to better understand CC fertility problems and line extinction.

MATERIALS AND METHODS

The mouse inbred strains used in these experiments are the eight founder strains of the Collaborative Cross (CC) (Collaborative Cross Consortium 2012). The founders of the CC include five classical strains, A/J (A), C57BL/6J (B6), 129S1/SvImJ (129S1), NOD/ShiLtJ (NOD), NZO/HlLtJ (NZO), and three wild-derived strains, CAST/EiJ (CAST), PWK/PhJ (PWK), and WSB/EiJ (WSB). Mice originated from a colony maintained by Gary Churchill at the Jackson Laboratory, and were transferred to the FPMV lab at the University of North Carolina (UNC) in 2008. The original colony also produced most of the G1 breeders that populated the inbred funnels at ORNL, TAU and Geniad (Srivastava *et al.* 2017). All mice described here were reared by the FPMV lab at UNC. Mice were bred at the UNC Hillsborough vivarium from 2008–2010 and bred at the UNC Genetic Medicine Building (GMB) vivarium from 2010–2012. A total of 4,448 litters resulting from crosses between 1,478 individual dams and 1,238 individual sires were born from all eight inbred crosses and 54 of 56 reciprocal F1 hybrid crosses, excluding hybrids between NZO×CAST and NZO×PWK, which are known unproductive crosses (Chesler *et al.* 2008). The directions of all crosses are described as female by male (*i.e.*, dam.strain × sire.strain), unless otherwise noted. Litter size and sex were determined at weaning by visual inspection. Animals were kept on a 14-hour, 10-hour light/dark schedule with lights turned on at 7:00 AM; temperature was maintained at 20°–24° with relative humidity between 40–50%. Mice were housed in standard 20×30-cm ventilated polysulfone cages with standard laboratory grade Bed-O-Cob bedding. Water and Purina Prolab RMH3000

were available *ad libitum*. Mouse chow was supplemented with Fenbendazole (Feb 2010) two weeks before and two weeks after transportation to the GMB facility to eliminate possible pinworms. Selamectin treatment was dropped onto the coats of mice before transfer to remove mites from the cages. These treated cages were not opened until after their arrival at UNC GMB. All animal rearing and breeding was conducted in strict compliance with the *Guide for the Care and Use of Laboratory Animals* (Institute of Laboratory Animals Resources, National Research Council 1996, <https://www.ncbi.nlm.nih.gov/books/NBK232589/>). The Institutional Animal Care and Use Committee of the University of North Carolina approved all animal use and research described here.

STATISTICAL ANALYSIS

Testing environmental interactions

Significant environmental interactions were determined using ANOVA, using JMP 12 software (JMP, Version 12. SAS Institute Inc., Cary, NC, 1989–2007). Tested effects included a season effect (average weaned litter size in each month over every year), non-seasonal factors using a year-by-month effect (average weaned litter size for all months across all years), and a litter order effect (average weaned litter size in each subsequent dam litter). We performed the ANOVA on litter size counts from the eight inbred matings only, because of their robust sample sizes throughout the four years of breeding.

Diallel analysis of litter size

To estimate the overall contributions of heritable factors affecting litter size in our population, we adapted a previously published linear mixed model, BayesDiallel (Lenarcic *et al.* 2012), which performs this estimation for continuous phenotypes, to the setting of a discrete count-based phenotype. To do this we reimplemented the BayesDiallel Gibbs sampler as a generalized linear mixed model (GLMM) using the R software package MCMCglmm (Hadfield 2010).

Let y_i be the number of pups born to litter i , where y_i is a positive integer (zeros are not observed). For any categorical variable x , let the notation $x[i]$ indicate the value of x relevant to i : in particular, let $h[i]$ denote litter i 's batch $h = 1, \dots, 48$; let $r[i]$ denote its parity order $r = 1, \dots, 12$; and let $(j, k)[i]$ denote its parentage, defined by maternal strain j and paternal strain k , where $(j, k) \in \{1, \dots, 8\}^2$. The effect of parental strains on y_i was modeled using an overdispersed zero-truncated Poisson (ZTP) regression (data scales of the model in brackets):

$$\begin{aligned} y_i &\sim \text{ZTPois}(\lambda_i) && \text{[observed]} \\ \lambda_i &= g^{-1}(\ell_i) && \text{[expected]} \\ \ell_i &= \eta_i + \varepsilon_i && \text{[latent + overdis.]} \\ \eta_i &= \mu + \text{order}_{r[i]} + \text{batch}_{h[i]} + \mathbf{d}_{(j,k)[i]}^T \boldsymbol{\beta} && \text{[latent]} \end{aligned} \quad (1)$$

where $\text{ZTPois}(\lambda_i)$ denotes a Poisson distribution with an expected value $E[y_i] = \frac{\lambda_i}{1 - e^{-\lambda_i}}$ but is conditional on having observed that $y_i \neq 0$ (Appendix A), g is the link function $g(x) = \log(x)$, with inverse $g^{-1}(x) = e^x$, that relates λ_i to a latent scale ℓ_i , and η_i is a linear predictor on that latent scale with an error term $\varepsilon_i \sim N(0, \sigma^2)$ providing overdispersion. The linear predictor η_i is composed of the following: an intercept μ ; a litter (parity) order effect, modeled as the combination of a fixed effect slope, α , and a random effect with an independent level for each litter order, *i.e.*, $\text{order}_r \sim N(\alpha, \tau_{\text{order}}^2)$, where α is a fixed effect and τ_{order}^2 is the variance of the random deviations around α ; a batch effect, $\text{batch}_h \sim N(0, \tau_{\text{batch}}^2)$; and a linear predictor, $\mathbf{d}_{(j,k)}^T \boldsymbol{\beta}$, modeling the effect of the parental strain combination (j, k) .

The contribution of parental strains, $\mathbf{d}_{(j,k)}^T \boldsymbol{\beta}$, is equivalent to the ‘full’ (full, unsexed, ‘Babmwv’) model described in (Lenarcic *et al.* 2012), namely,

$$\mathbf{d}_{(j,k)}^T \boldsymbol{\beta} = \underbrace{a_j + a_k}_{\text{additive}} + \underbrace{m_j - m_k}_{\text{parental sex}} + \underbrace{I_{\{j=k\}} (\beta_{\text{inbred}} + b_j)}_{\text{inbred penalty}} + \underbrace{I_{\{j \neq k\}} (v_{jk} + S_{\{j < k\}} \cdot w_{jk})}_{\text{epistasis}} \quad (2)$$

where subscripted variables a, m, b, v, w model the effects of specific strains or strain-pairs, β_{inbred} models an overall effect, $I_{\{A\}}$ is an indicator and equal to 1 if A is true and 0 otherwise, $S_{\{A\}}$ is a sign variable equal to $\frac{1}{2}$ if A is true and $-\frac{1}{2}$ otherwise. In more detail, the a (“additive”) class represents strain-specific dosage effects, with a_1, \dots, a_8 modeled as $a_j \sim N_{\text{stz}}(0, \tau_a^2)$, where N_{stz} is a normal distribution subject to the sum-to-zero constraint $\sum a_j = 0$ [using the approach of Crowley *et al.* (2014), Appendix A]. The m (“parental sex”) class represents parent-of-origin effects, modeled as $m_j \sim N_{\text{stz}}(0, \tau_m^2)$, where a positive value of m_j implies that strain j increases litter size more when inherited through the maternal line, with the difference between maternal and paternal being $2m_j$ (since the design matrix entry for m_j is coded +1 for dam and -1 for sire). The effect of being inbred is composed of an overall (fixed) effect β_{inbred} and strain-specific (random) effects $b_j \sim N_{\text{stz}}(0, \tau_b^2)$. Epistatic effects are modeled as strain-pair specific deviations, with “symmetric” effects $v_{jk} \sim N_{\text{stz}}(0, \tau_v^2)$ representing the overall deviation from the rest of the model induced by the (unordered) strain combination j with k , and “asymmetric” effects $w_{jk} \sim N_{\text{stz}}(0, \tau_w^2)$ modeling a further deviation induced by differences in parent-of-origin.

To provide directional parent-of-origin effects for maternal and paternal strain we defined, by reparameterization of the additive and parental sex effect, the following two additional types of effect:

$$\text{dam}_j = a_j + m_j$$

$$\text{sire}_j = a_j - m_j$$

For example, $\text{dam}_{B6} = a_{B6} + m_{B6}$ is the B6-specific dam (maternal strain) effect and $\text{sire}_{B6} = a_{B6} - m_{B6}$ is the B6-specific sire (paternal strain) effect. Posterior samples of these quantities were obtained as a post-processing step by reparameterizing posterior samples of a_j and m_j . Obtaining these as a post-processing step, rather than explicit modification of Equation 2, preserves the original BayesDiallel model and therefore has no effect on the effect estimates (or variance projection estimates) for the other diallel categories. (Parameter definitions summarized in Table S1.)

The decomposition of diallel effects in Equation 2 and its dam vs. sire reformulation have parallels in earlier diallel literature. A reduced version of the decomposition composed of additive (a) and symmetric epistasis (v) estimate, respectively, the Generalized Combining Ability (GCA) and Specific Combining Ability (SCA) described by Sprague and Tatum (1942); including also the (reciprocal) asymmetric epistasis (w) term recapitulates Griffing’s Method 1, Model 1 (Griffing 1956). The inbred penalty (β_{inbred}, b) is comparable to dominance measures defined in, for example, Hayman (1954). The parental sex effects (m) are akin to the maternal effects of Topham (1966) and the “extranuclear” effects in the “bio” model of Cockerham and Weir (1977) and Zhu and Weir (1996), and the bio model’s reparameterization to indirectly estimate dam and sire effects has been described in, for example, Lynch and Walsh (1998). The estimation of dam and sire effects

directly, as explicit model parameters for the diallel, was proposed by Comstock and Robinson (1948, 1952). The main differences between these earlier analyses and ours are the simultaneous inclusion of all parameter groups, the use of a Bayesian random effects framework to allow all these parameters to be fitted, and the extension to modeling a non-Gaussian response.

Priors were chosen to be minimally informative. For fixed effects (μ, α , and β_{inbred}) we used $N(0, 1 \times 10^3)$. For variances of random effects ($\sigma^2, \tau_a^2, \tau_m^2, \tau_b^2, \tau_v^2, \tau_w^2, \tau_{\text{order}}^2, \tau_{\text{batch}}^2$) we used an inverse gamma distribution with scale and shape both equal to 0.001. Posterior effect estimates are presented as posterior mean (and median), and the 95% highest posterior density (HPD) interval.

In order to stably estimate the contribution of each variance class to the overall phenotype, we used the diallel variance projection [VarP; Crowley *et al.* (2014)]. This is a heritability-like measure that partitions the overall phenotypic variance of an idealized future diallel experiment into additive, parent-of-origin (parental sex), inbred (dominance), epistatic, and other random/fixed effects categories in the diallel. Rather than being based on estimated variance parameters (e.g., $\tau_a^2, \tau_b^2, \dots$), which are typically ill-informed by the data and thus both uncertain and sensitive to priors, the VarP uses the estimated effects themselves, both fixed (e.g., β_{inbred}) and random (i.e., the best linear unbiased predictors, or BLUPs, such as a_1, a_2, \dots), since these are well-informed, precise and regularized by shrinkage. The VarP calculation involves ratios of sums-of-squares in similar fashion to R^2 but for an idealized diallel that is both complete and balanced. As an R^2 -like measure, the VarP is reportable for both fixed and random effects, and includes confidence intervals arising from posterior uncertainty in those effects’ values.

VarPs were calculated both from the ZTP model described above, and, for comparison, from the standard Gaussian-outcome BayesDiallel model (using the BayesDiallel software) (Appendix A). The standard BayesDiallel model was applied to our data after litter size was subject to a variance-stabilizing transformation, the square root, this corresponding to the linear mixed model approximation to the ZTP,

$$\text{sqrt}(y_i) \sim \mu + \text{order}_{r[i]} + \text{batch}_{h[i]} + \mathbf{d}_{(j,k)[i]}^T \boldsymbol{\beta} + \varepsilon_i \quad (3)$$

Diallel analysis of sex ratio

To model diallel effects on sex ratio we recast the BayesDiallel model as a binomial GLMM. Letting y_i be the number of males out of a total of n_i pups for litter i , we model

$$\begin{aligned} y_i &\sim \text{Binom}(n_i, \pi_i) && \begin{bmatrix} \text{observed} \\ \text{expected} \end{bmatrix} \\ \pi_i &= g^{-1}(\ell_i) && \end{aligned} \quad (4)$$

where π_i is the expected proportion of males predicted for litter i , the link function is $g(x) = \text{logit}(x) = \log(x)/(1-x)$, with inverse link $g^{-1}(x) = \text{logit}^{-1}(x) = \text{expit}(x) = e^x/(1+e^x)$, and ℓ_i , which represents π_i on the latent scale, is modeled using the BayesDiallel hierarchy as in Equation 1.

Additional details about our statistical modeling approaches are provided in Appendix A (litter size) and Appendix B (sex ratio).

Data Availability

File S1 contains all breeding data used for the analysis in this study. File S2 contains the scripts and software used for the analysis in this study.

File S3 contains the Supplemental Figures and Tables. The data and code used to analyze the data are available at <https://github.com/mauriziopaul/litterDiallel>, with a static version available at <https://doi.org/10.5281/zenodo.2580307>. Phenotype data from this study is available on the Mouse Phenome Database (Bogue *et al.* 2018) at <https://phenome.jax.org/projects/Shorter3> with persistent identifier RRID:SCR003212, Accession number MPD:623, project Shorter3. Supplemental material available at Figshare: <https://figshare.com/s/fe6a6216bf29b2baeb65>.

RESULTS

Litter size is affected by housing facility but not season

We bred litters in an 8×8 inbred diallel of the CC founder strains, and generated all eight inbreds and 54 of 56 possible reciprocal F1 hybrids (Figure 1, Figure S1–S3). The eight inbred strains, displayed across the diagonal, were mated at higher frequencies both for maintenance of inbred strains and propagation of the diallel. For all 62 genetic inbred and hybrid crosses, we recorded the following information: mated pairs, wean dates, litter size at weaning, including total and sex-specific counts (File S1). This diallel cross was originally designed and maintained for the generation of F1 mice for several experimental projects (Koturbash *et al.* 2011; Aylor *et al.* 2011; Mathes *et al.* 2011; Kelada *et al.* 2012; Didion *et al.* 2012; Collaborative Cross Consortium 2012; Calaway *et al.* 2013; Crowley *et al.* 2014; Phillippi *et al.* 2014; Odet *et al.* 2015; Crowley *et al.* 2015; Morgan *et al.* 2016; Percival *et al.* 2016; Shorter *et al.* 2017; Oreper *et al.* 2017; Maurizio *et al.* 2018). As a result, certain reproductive measurements such as time between litters and maximum number of offspring per cross were necessarily biased by experimental breeding requirements. Average weaned litter size, however, is a reproductive trait that should be well-estimated independently of these factors.

We measured litter size and report the mean number of weaned pups per litter for the 62 viable crosses in the diallel (Figure 1). A wide distribution of litter sizes was observed, ranging from an average of 3.6 weaned pups for WSB×WSB crosses, to an average of 9.1 weaned pups for NOD×PWK crosses, with an overall mean of 5.5 weaned pups per litter. Examining average litter sizes of the inbred strains across all 4 years, we found negligible evidence of consistent seasonal effects ($F_{11,2283} = 1.272$ $P = 0.23$) (Figure S4) but positive evidence of non-seasonal patterns: a modeled ‘year.month’ covariate significantly affected litter size ($F_{47,2247} = 2.44$ $P < 0.0001$; see also Figure S5). The non-seasonal effect could be driven by the relocation of mice between vivariums, which occurred in February 2010. The housing facility may have an impact on litter size as well. To test this, we measured average litter size differences between the two housing facilities and found that two founder strains, 129S1 and WSB, had significantly larger average litter sizes at the Hillsborough facility than at UNC GMB with a difference of 5.87 to 4.42 weaned pups per litter ($P < 0.0001$) for 129S1, and 3.84 to 3.42 weaned pups per litter ($P = 0.029$) for WSB. The six other founder strains did not significantly differ in their average litter sizes between the two facilities, but tended to have smaller weaned litters at UNC GMB (see Discussion). For average litter size across the diallel, we tested for an effect of litter number (birth parity number, for a given mating pair) on the number of pups weaned in each litter. Previous research suggests that the first litter can be significantly smaller than subsequent litters due to various biological factors (De la Fuente and San Primitivo 1985). This effect was consistent in the diallel: we observed that there was significant reduction in litter size in the first litter ($P < 0.0001$) compared to the overall linear effect that parity has on reducing litter size (Figure S5).

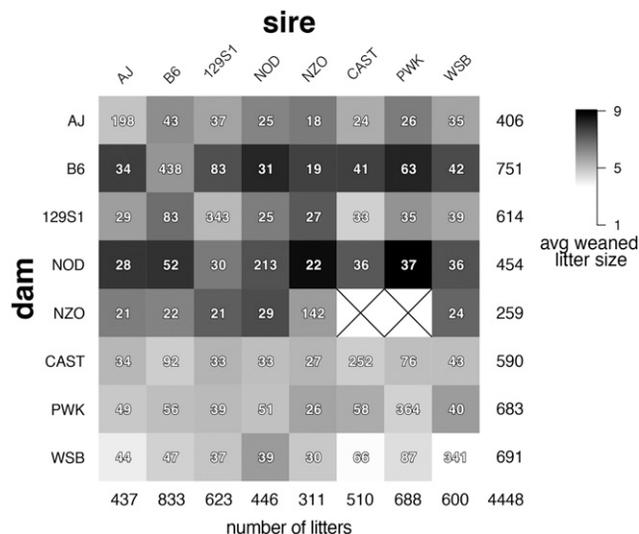


Figure 1 Diallel crossing scheme and weaned pup distribution. The number of litters observed per cross is given by the integers, with the largest sample sizes, along the diagonal, corresponding to the production of inbred parental strains. Column and row sums are given along the bottom row and rightmost column, respectively. A total of 4,448 litters were evaluated for this analysis, resulting in a total of 24,782 weaned pups. The shading within each box corresponds to the average number of weaned pups per litter in each cross, with averages ranging from 3.6 to 9.1 pups per litter. Litters for which no pups survived until weaning were not included in our analysis. The symbol “x” is used to indicate incompatible crosses that do not produce any litters.

Litter size is moderately heritable, and maternal effects account for the majority of explained variation

To estimate founder strain effects on litter size, we used a Bayesian regression model that decomposes the phenotypic variation in the diallel into genetic and parent-of-origin contributions (Lenarcic *et al.* 2012). Using this model, the percentage of the variance in litter size explained by diallel effects was 17.73%, with additive effects explaining 9.18% (VarP[additive]; this GCA-like measure being related to narrow sense heritability), parent-of-origin effects (VarP[parental.sex]) accounting for 5.77%, the fact of being inbred (VarP[inbred.overall]) at 1.43%, and strain-by-strain interactions (VarP[epistatic.symmetric] + VarP[asymmetric.epistatic]) at 3.40% (Figure 2A).

In more detail, we present estimates for all modeled diallel effects as posterior means and highest posterior density (HPD) intervals (Figure 2B). Parameters are divided into two groups: general effects and strain pair-specific effects. General effects comprise strain-specific additive effects (additive), strain-specific and overall inbred (inbred), and strain-specific parent-of-origin (parental sex) effects. Strain pair-specific (epistatic) effects are the effects that arise specifically in crosses of two heterologous strains, with ‘v’ referring to symmetric epistatic and ‘w’ referring to asymmetric epistatic effects (pairwise parent-of-origin effects).

Under the general effects, we see significant positive additive effects on average litter size from B6, NOD, and NZO and significant negative additive effects on litter size from CAST, PWK, and WSB strain dosages. A similar pattern is seen in the parental sex effects, where B6 and NOD dosages have a significant positive effect on average litter size, whereas CAST, PWK, and WSB have significant negative effects. The overall “inbred” effect is negative, indicating that inbred status decreases

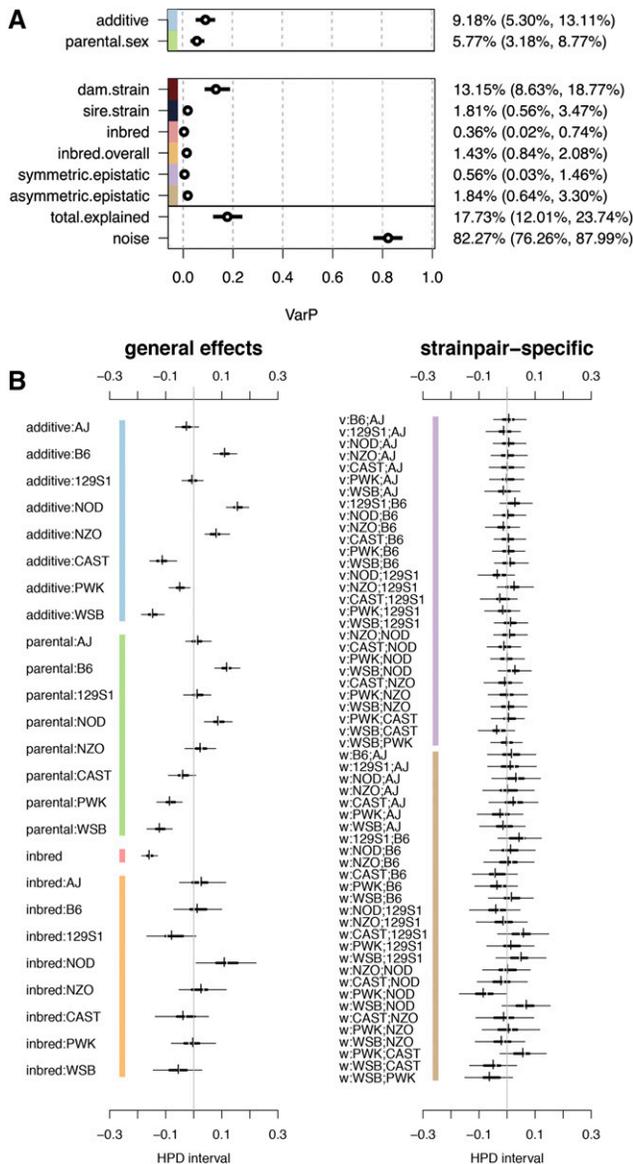


Figure 2 Diallel effects and variance contributions for weaned litter size. (A) Variance contributions of distinct effect classes, reported as posterior means and 95% HPDs of variance projections (VarPs) on weaned litter size. (B) Diallel effects, including (left) strain-specific additive, parental sex, and inbred effects, and (right) epistatic effects between each pairwise cross. For each parameter, thin and thick horizontal lines represent 95% and 50% highest posterior density (HPD) intervals of effects, respectively, and vertical break and dash give posterior median and mean, respectively. The effects are in relation to an overall mean litter size of 5.46 (95% HPD: 5.00-6.10). The gray vertical lines indicate zero. Effects are shown as the log, or latent, scale effects on the mean litter size attributable to each strain or strainpair and inheritance group, where values are centered at 0 for each random effect class. Intervals that exclude zero have non-negligible effects on the mean litter size. Labels with “v” or “w” refer to symmetric or asymmetric epistatic effects, respectively. Colored bars indicate corresponding variance classes in (A) and (B).

average litter size, regardless of parental strain. Each strain also has an individual inbred effect in addition to the overall inbred effect. When the individual inbred effects are taken into account with the overall “inbred” effect, inbred litters are on average slightly smaller

than their heterozygous counterparts. For strain pair-specific epistatic effects, there are a few marginally noteworthy effects, with the most prominent being a negative asymmetric epistatic effect for PWK×NOD.

To more clearly differentiate the contributions of the mother vs. the father strain, we reparameterized the BayesDiallel model to capture effects specific to dam.strain and sire.strain (Figure 3). B6 and NOD dams increase litter sizes by more than 1.26 fold, by an average of 1.66 and 1.79 pups, respectively, regardless of sire. CAST, PWK, and WSB dams tend to decrease average litter size by 0.90, 0.81, and 1.51 pups. The sire effect is similar, with NOD and NZO sires having larger litters and CAST sires producing smaller litters. As expected, we see that the dam.strain has a much larger influence on the variation of litter size compared with the sire.strain (13.15% vs. 1.81%).

No significant strain effects on sex ratio

We examined genetic and non-genetic effects on the average sex ratio per litter and found no evidence that sex ratio was skewed (Figure S6; Figure S7). The overall mean for sex bias, quantified as number of male pups weaned divided by the total number of weaned pups, was 0.4979, and did not significantly differ from our expectation of a 50:50 sex ratio (binominal test, two-tailed $P = 0.219$). For the eight inbred founders, 129S1 is the only strain that departed significantly from expectation, with slight reduction of males 0.475:0.525 (binominal test, two-tailed $P = 0.041$). However, correction for multiple testing shows no significant sex ratio bias of any inbred strain. The outbred crosses have substantially fewer litters and offspring than the inbred matings, leading to less balanced sex ratios; however, when multiple testing is accounted for, they also show no significant deviations from expected sex ratios.

DISCUSSION

We have investigated factors influencing litter size in the eight CC founders and their F1 hybrids using a new extension of the BayesDiallel model. We note that litter size is a component of reproductive performance, but distinct from total strain productivity; a study on total strain productivity would need to take into account litter size, numbers of litters, maximum reproductive age, and pup survival until breeding. Our results illustrate how mammalian litter size in a full diallel design is influenced by genotypic and environmental variation. These results present new information on CC founder strains’ reproductive performance, show that maternal effects and the environment play a large role in litter size variation, and provide no evidence for seasonality effects on litter size in a controlled animal facility. The results also address some of the factors that contributed to line extinction and breeding problems in generating the CC (Chesler *et al.* 2008; Philip *et al.* 2011; Collaborative Cross Consortium 2012; Shorter *et al.* 2017).

We estimated that genetic (additive, inbred, and epistatic) effects on average litter size explained 17.73% of variation, suggesting that most of the phenotypic variation arises from unexplained environmental effects. Compared with the overall average litter, we observed substantial positive effects of B6 and NOD strains, from both additive genetic and parent-of-origin parameters, and substantial negative effects of PWK and WSB (Figure 2). We also observed, as expected, that lower litter size was associated with being inbred (Figure 2A). These estimates are likely driven by the unique selection history of these inbred lines, and comparisons should be limited to these eight founder strains, and the CC and DO populations.

During the G1 and G2 out-crossing generations of the CC, mean litter size was lower for crosses involving wild-derived strains, CAST, PWK and WSB (Philip *et al.* 2011). A similar pattern is observed here, but these effects are determined to be specifically through the maternal strain, with CAST, PWK and WSB having negative “dam” effects on

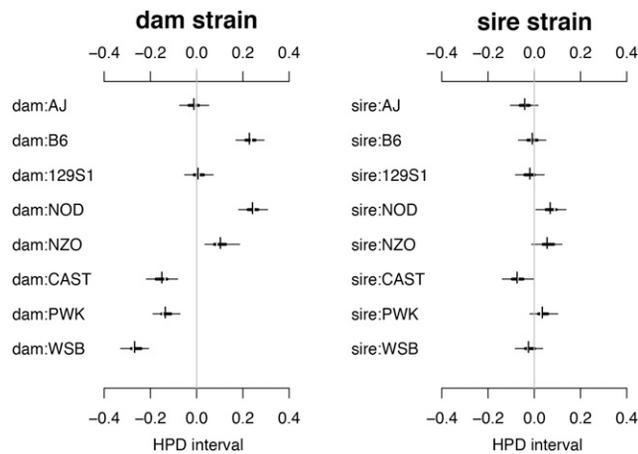


Figure 3 Dam.strain and sire.strain variance contributions and estimates of effects on weaned litter size. These effects are a reparameterization of additive and parental.sex effects from the previous analysis. Estimates for the maternal (“dam.strain”) and paternal (“sire.strain”) effects on litter size, as calculated from the additive and parental sex parameters in Figure 2, with HPD intervals defined correspondingly.

litter size (Figure 3). It is likely that selection pressure in classical lab strains is associated with larger litters compared with the wild-derived strains. Additionally, two of the wild-derived strains, CAST and PWK, are from a different subspecific origin than the other six CC founders. This likely contributes to decreased productivity through subspecific incompatibilities (Shorter *et al.* 2017).

We identify and report environmental factors that may influence litter size. The breeding of this diallel was performed across two different vivariums over the course of 4 years, and we see a significant effect from housing facility. The Hillsborough facility was associated with larger litters for all strains, especially 129S1 and WSB. The two facilities have many different factors that could explain these differences. The Hillsborough facility housed multiple species, including dogs and mice, had smaller rooms that held approximately 300 mouse cages, was remotely located in a rural area, had different laboratory personnel, had cage changes once a week, and was supplied with filtered well water. The diallel breeding at the GMB facility took place in a large central room, contained only laboratory mice, is located in a basement of a large seven story research building on campus, has cage changes every other week, and is supplied with filtered city water. It is possible that one or more of these factors, independently or in combination, affected productivity. Another finding was that seasonality, which has previously been shown to influence litter size and frequency of litters in mammals (Drickamer 1990), did not seem to significantly impact litter size in this study. This is likely due to consistent light-dark cycles and temperatures as well as a steady diet. We did observe a significant effect on litter size after the transfer of the mice to the UNC GMB facility (February 2010), which reduced overall litter sizes from March to June 2010. This may have been due to the use of Fenbendazole during the time of the transfer. Other factors, such as the sex of the laboratory personnel interacting with the animals, are generally known to influence rodent behavior and could contribute to some of the environmentally-induced variation we observed (Sorge *et al.* 2014).

Last, we measured sex ratio across all inbred and outbred crosses. Despite some departures from equality at the nominal significance threshold ($\alpha = 0.05$), no associations with founder strain dosage remain significant after correction for multiple testing. Recent work

has suggested a potential for bias in sex ratio driven by the male germline in *Mus musculus* (Conway *et al.* 1994; Macholán *et al.* 2008; Cocquet *et al.* 2009; Ellis *et al.* 2011; Cocquet *et al.* 2012; Turner *et al.* 2012; Larson *et al.* 2016), particularly in inter-subspecific hybrids that are mismatched for copy number of X- and Y-linked genes expressed in postmeiotic spermatids. Although there are no previous observations that suggest a bias in sex ratio in the CC or its founders, it remains an important characteristic to measure in a study on reproductive productivity.

To estimate heritable effects on litter size and sex ratio we extended the original BayesDiallel model of Lenarcic *et al.* (2012) in two new ways. First, to better understand and distinguish the effects arising from female and male parents, we reparameterized our strain-specific additive and parental-sex effects such that we could provide estimates of maternal strain and paternal strain effects separately. In the original BayesDiallel model, maternal and paternal strain effects are split into “additive” effects, which consolidates the effects they have in common, and “parental sex”, which models any remaining deviation between the two. Recognizing that additive effects from the sire are essentially wiped out by the additive effects, we instead collapsed the additive and parental-sex effects into “dam.strain” and “sire.strain” effects in a post-processing step on the posterior output. This allowed us to run the original BayesDiallel model while also viewing our data from the perspective of dam strain and sire strain contributions.

Second, we reimplemented the original MCMC sampler, designed for modeling a continuous outcome variable, in a general package MCMCglmm (Hadfield 2010) in order to model count and binary responses. Litter size, as measured, is most naturally distributed as Poisson, with zero-truncation owing to the fact that only successful litters were recorded. Sex ratio is most naturally modeled as a binomial, with an underlying (male) proportion between 0 and 1. Although it would be possible to obtain an approximate analysis by transformations to normality using the original BayesDiallel (and we do this for litter size for some otherwise hard-to-obtain quantities), we found such approximations to be inadequate for reliable estimation of higher order effects in the case of litter size and deeply flawed in the case of sex ratio. Although there is a computational cost, and added complexity to determining variance contributions, this new implementation achieves several objectives: 1) we no longer break the assumptions in the original model regarding normally distributed errors; 2) we easily accommodate overdispersion in our data; and 3) we can select from a large number of GLMMs models that more closely resemble the forms of our data observations. In addition, we believe this flexibility will be appealing to many other researchers who would like to model non-Gaussian distributed phenotypes using diallel designs, and we have provided the code in an R package litterDiallel (<https://doi.org/10.5281/zenodo.2580307>).

Overall, these results have implications for other avenues of future research. Future multiparental research populations should test for strain incompatibilities, reproductive phenotyping, and other health traits in a full diallel before the recombinant inbreeding begins (Odet *et al.* 2015). These future research populations should also use non-related wild-derived individuals from the same subspecific origin in order to increase genetic diversity without introducing hybrid incompatibilities.

ACKNOWLEDGMENTS

This work was supported in part by the following grants from the National Institutes of Health: P50GM076468, P50HG006582/P50MH090338, R01HD065024 (FP-MV), T32HD040127 (JRS), T32AI007419 (PLM), R01GM104125 (PLM, WV), and R35GM127000

(WV). The Collaborative Cross project is also supported by the University Cancer Research Funds granted to Lineberger Comprehensive Cancer Center (MCR012CCRI).

LITERATURE CITED

- Avalos, E., and C. Smith, 1987 Genetic improvement of litter size in pigs. *Anim. Sci.* 44: 153–163.
- Aylor, D. L., W. Valdar, W. Foulds-Mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011 Genetic analysis of complex traits in the emerging collaborative cross. *Genome Res.* 21: 1213–1222. <https://doi.org/10.1101/gr.111310.110>
- Bandy, T., and E. Eisen, 1984 Prenatal and postnatal effects in mouse lines selected for body weight and litter size: performance of postnatal dams and growth of progeny. *J. Anim. Sci.* 59: 896–907. <https://doi.org/10.2527/jas1984.594896x>
- Beauchamp, G., and A. Kacelnik, 1990 On the fitness functions relating parental care to reproductive value. *J. Theor. Biol.* 146: 513–522. [https://doi.org/10.1016/S0022-5193\(05\)80376-5](https://doi.org/10.1016/S0022-5193(05)80376-5)
- Bogue, M. A., S. C. Grubb, D. O. Walton, V. M. Philip, G. Kolishovski *et al.*, 2018 Mouse phenome database: an integrative database and analysis suite for curated empirical phenotype data from laboratory mice. *Nucleic Acids Res.* 46: D843–D850. <https://doi.org/10.1093/nar/gkx1082>
- Calaway, J. D., A. B. Lenarcic, J. P. Didion, J. R. Wang, J. B. Searle *et al.*, 2013 Genetic architecture of skewed X inactivation in the laboratory mouse. *PLoS Genet.* 9: e1003853. <https://doi.org/10.1371/journal.pgen.1003853>
- Chesler, E. J., 2014 Out of the bottleneck: the diversity outcross and collaborative cross mouse populations in behavioral genetics research. *Mamm. Genome* 25: 3–11. <https://doi.org/10.1007/s00335-013-9492-9>
- Chesler, E. J., D. R. Miller, L. R. Branstetter, L. D. Galloway, B. L. Jackson *et al.*, 2008 The collaborative cross at oak ridge national laboratory: developing a powerful resource for systems genetics. *Mamm. Genome* 19: 382–389. <https://doi.org/10.1007/s00335-008-9135-8>
- Cockerham, C., and B. Weir, 1977 Quadratic analyses of reciprocal crosses. *Biometrics* 33: 187–203. <https://doi.org/10.2307/2529312>
- Cocquet, J., P. J. Ellis, S. K. Mahadevaiah, N. A. Affara, D. Vaiman *et al.*, 2012 A genetic basis for a postmeiotic X vs. Y chromosome endogenous conflict in the mouse. *PLoS Genet.* 8: e1002900. <https://doi.org/10.1371/journal.pgen.1002900>
- Cocquet, J., P. J. Ellis, Y. Yamauchi, S. K. Mahadevaiah, N. A. Affara *et al.*, 2009 The multicopy gene sly represses the sex chromosomes in the male mouse germline after meiosis. *PLoS Biol.* 7: e1000244. <https://doi.org/10.1371/journal.pbio.1000244>
- Collaborative Cross Consortium, 2012 The genome architecture of the collaborative cross mouse genetic reference population. *Genetics* 190: 389–401. <https://doi.org/10.1534/genetics.111.132639>
- Comstock, R., and H. Robinson, 1948 The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. *Biometrics* 4: 254–266. <https://doi.org/10.2307/3001412>
- Comstock, R., and H. Robinson, 1952 Estimation of average dominance of genes, pp. 494–516 in *Heterosis*, chap 30., edited by Gowen, J. W., Iowa State College Press, Ames.
- Conway, S., S. Mahadevaiah, S. Darling, B. Capel, A. Rattigan *et al.*, 1994 Y353/b: a candidate multiple-copy spermiogenesis gene on the mouse Y chromosome. *Mamm. Genome* 5: 203–210. <https://doi.org/10.1007/BF00360546>
- Crowley, J. J., Y. Kim, A. B. Lenarcic, C. R. Quackenbush, C. J. Barrick *et al.*, 2014 Genetics of adverse reactions to haloperidol in a mouse diallel: a drug-placebo experiment and Bayesian causal analysis. *Genetics* 196: 321–347. <https://doi.org/10.1534/genetics.113.156901>
- Crowley, J. J., V. Zhabotynsky, W. Sun, S. Huang, I. K. Pakatci *et al.*, 2015 Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nat. Genet.* 47: 353–360 (erratum: *Nat. Genet.* 47: 690). <https://doi.org/10.1038/ng.3222>
- De la Fuente, L., and F. San Primitivo, 1985 Selection for large and small litter size of the first three litters in mice. *Genetique, Selection, Evolution* 17: 251.
- Didion, J. P., H. Yang, K. Sheppard, C.-P. Fu, L. McMillan *et al.*, 2012 Discovery of novel variants in genotyping arrays improves genotype retention and reduces ascertainment bias. *BMC Genomics* 13: 34. <https://doi.org/10.1186/1471-2164-13-34>
- Drickamer, L., 1990 Seasonal variation in fertility, fecundity and litter sex ratio in laboratory and wild stocks of house mice (*Mus domesticus*). *Lab. Anim. Sci.* 40: 284–288.
- Ellis, P. J., J. Bacon, and N. A. Affara, 2011 Association of sly with sex-linked gene amplification during mouse evolution: a side effect of genomic conflict in spermatids? *Hum. Mol. Genet.* 20: 3010–3021. <https://doi.org/10.1093/hmg/ddr204>
- Falconer, D., 1960 The genetics of litter size in mice. *J. Cell. Physiol.* 56: 153–167. <https://doi.org/10.1002/jcp.1030560414>
- Falconer, D., and T. Mackay, 1996 *Introduction to Quantitative Genetics*, Longmans Green, Harlow, Essex.
- Ferris, M. T., D. L. Aylor, D. Bottomly, A. C. Whitmore, L. D. Aicher *et al.*, 2013 Modeling host genetic regulation of influenza pathogenesis in the Collaborative Cross. *PLoS Pathog.* 9: e1003196. <https://doi.org/10.1371/journal.ppat.1003196>
- Garcia, M., and M. Baselga, 2002 Estimation of genetic response to selection in litter size of rabbits using a cryopreserved control population. *Livest. Prod. Sci.* 74: 45–53. [https://doi.org/10.1016/S0301-6226\(01\)00280-9](https://doi.org/10.1016/S0301-6226(01)00280-9)
- Gralinski, L. E., M. T. Ferris, D. L. Aylor, A. C. Whitmore, R. Green *et al.*, 2015 Genome Wide Identification of SARS-CoV Susceptibility Loci Using the Collaborative Cross. *PLoS Genet.* 11: e1005504. <https://doi.org/10.1371/journal.pgen.1005504>
- Gralinski, L. E., V. D. Menachery, A. P. Morgan, A. L. Totura, A. Beall *et al.*, 2017 Allelic variation in the toll-like receptor adaptor protein *Ticam2* contributes to SARS-Coronavirus pathogenesis in mice. *G3: Genes, Genomes, Genetics* 7: 1653–1663.
- Griffing, B., 1956 A generalized treatment of the use of diallel crosses in quantitative inheritance. *Heredity* 10: 31–50. <https://doi.org/10.1038/hdy.1956.2>
- Gutiérrez, J. P., B. Nieto, P. Piqueras, N. Ibáñez, and C. Salgado, 2006 Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. *Genet. Sel. Evol.* 38: 445. <https://doi.org/10.1186/1297-9686-38-5-445>
- Hadfield, J. D., 2010 MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* 33: 1–22. <https://doi.org/10.18637/jss.v033.i02>
- Hager, R., and R. A. Johnstone, 2003 The genetic basis of family conflict resolution in mice. *Nature* 421: 533–535. <https://doi.org/10.1038/nature01239>
- Hayman, B. I., 1954 The analysis of variance of diallel tables. *Biometrics* 10: 235–244. <https://doi.org/10.2307/3001877>
- Hoornbeek, F., 1968 Mating success and litter size variation within and between inbred and hybrid generations of rats. *J. Anim. Sci.* 27: 1378–1382. <https://doi.org/10.2527/jas1968.2751378x>
- Jinks, J., and P. Broadhurst, 1963 Diallel analysis of litter size and body weight in rats. *Heredity* 18: 319–336. <https://doi.org/10.1038/hdy.1963.33>
- Johnson, R. K., M. K. Nielsen, and D. S. Casey, 1999 Responses in ovulation rate, embryonal survival, and litter traits in swine to 14 generations of selection to increase litter size. *J. Anim. Sci.* 77: 541–557. <https://doi.org/10.2527/1999.773541x>
- Kelada, S. N., D. L. Aylor, B. C. Peck, J. F. Ryan, U. Tavarez *et al.*, 2012 Genetic analysis of hematological parameters in incipient lines of the collaborative cross. *G3: Genes, Genomes, Genetics* 2: 157–165.
- Koturbash, I., A. Scherhag, J. Sorrentino, K. Sexton, W. Bodnar *et al.*, 2011 Epigenetic mechanisms of mouse interstrain variability in genotoxicity of the environmental toxicant 1, 3-butadiene. *Toxicol. Sci.* 122: 448–456. <https://doi.org/10.1093/toxsci/kfr133>
- Larson, E. L., S. Keeble, D. Vanderpool, M. D. Dean, and J. M. Good, 2016 The composite regulatory basis of the large X-effect in mouse speciation. *Mol. Biol. Evol.* 34: 282–295.
- Lenarcic, A. B., K. L. Svenson, G. A. Churchill, and W. Valdar, 2012 A general Bayesian approach to analyzing diallel crosses of inbred

- strains. *Genetics* 190: 413–435. <https://doi.org/10.1534/genetics.111.132563>
- Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*, Sinauer Associates, Inc., Sunderland, MA.
- Macholán, M., S. J. Baird, P. Munclinger, P. Dufková, B. Bímová *et al.*, 2008 Genetic conflict outweighs heterogametic incompatibility in the mouse hybrid zone? *BMC Evol. Biol.* 8: 271. <https://doi.org/10.1186/1471-2148-8-271>
- Mathes, W. F., D. L. Aylor, D. R. Miller, G. A. Churchill, E. J. Chesler *et al.*, 2011 Architecture of energy balance traits in emerging lines of the collaborative cross. *Am. J. Physiol. Endocrinol. Metab.* 300: E1124–E1134. <https://doi.org/10.1152/ajpendo.00707.2010>
- Maurizio, P. L., M. T. Ferris, G. R. Keele, D. R. Miller, G. D. Shaw *et al.*, 2018 Bayesian diallel analysis reveals Mx1-dependent and Mx1-independent effects on response to influenza A virus in mice. *G3: Genes, Genomes. Genetics* 8: 427–445.
- Merilä, J., and B. C. Sheldon, 1999 Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity* 83: 103. <https://doi.org/10.1046/j.1365-2540.1999.00585.x>
- Morgan, A. P., C.-P. Fu, C.-Y. Kao, C. E. Welsh, J. P. Didion *et al.*, 2016 The mouse universal genotyping array: from substrains to subspecies. *G3: Genes, Genomes. Genetics* 6: 263–279.
- Mousseau, T. A., and D. A. Roff, 1987 Natural selection and the heritability of fitness components. *Heredity* 59: 181–197. <https://doi.org/10.1038/hdy.1987.113>
- Odet, F., W. Pan, T. A. Bell, S. G. Goodson, A. M. Stevans *et al.*, 2015 The founder strains of the Collaborative Cross express a complex combination of advantageous and deleterious traits for male reproduction. *G3: Genes, Genomes. Genetics* 5: 2671–2683.
- Oreper, D., Y. Cai, L. M. Tarantino, F. P.-M. de Villena, and W. Valdar, 2017 Inbred strain variant database (isvdb): a repository for probabilistically informed sequence differences among the collaborative cross strains and their founders. *G3: Genes, Genomes. Genetics* 7: 1623–1630.
- Percival, C. J., D. K. Liberton, F. Pardo-Manuel de Villena, R. Spritz, R. Marcucio *et al.*, 2016 Genetics of murine craniofacial morphology: diallel analysis of the eight founders of the collaborative cross. *J. Anat.* 228: 96–112. <https://doi.org/10.1111/joa.12382>
- Peripato, A., R. De Brito, S. Matioli, L. Pletscher, T. Vaughn *et al.*, 2004 Epistasis affecting litter size in mice. *J. Evol. Biol.* 17: 593–602. <https://doi.org/10.1111/j.1420-9101.2004.00702.x>
- Philip, V. M., G. Sokoloff, C. L. Ackert-Bicknell, M. Striz, L. Branstetter *et al.*, 2011 Genetic analysis in the collaborative cross breeding population. *Genome Res.* 21: 1223–1238. <https://doi.org/10.1101/gr.113886.110>
- Phillippi, J., Y. Xie, D. R. Miller, T. A. Bell, Z. Zhang *et al.*, 2014 Using the emerging collaborative cross to probe the immune system. *Genes Immun.* 15: 38–46. <https://doi.org/10.1038/gene.2013.59>
- Price, T., and D. Schluter, 1991 On the low heritability of life-history traits. *Evolution* 45: 853–861. <https://doi.org/10.1111/j.1558-5646.1991.tb04354.x>
- Rasmussen, A. L., A. Okumura, M. T. Ferris, R. Green, F. Feldmann *et al.*, 2014 Host genetic diversity enables Ebola hemorrhagic fever pathogenesis and resistance. *Science* 346: 987–991. <https://doi.org/10.1126/science.1259595>
- Roberts, R., 1960 The effects on litter size of crossing lines of mice inbred without selection. *Genet. Res.* 1: 239–252. <https://doi.org/10.1017/S0016672300000227>
- Rogala, A. R., A. P. Morgan, A. M. Christensen, T. J. Gooch, T. Bell *et al.*, 2014 The Collaborative Cross as a Resource for Modeling Human Disease: CC011/Unc, a New Mouse Model for Spontaneous Colitis. *Mamm. Genome* 25: 95–108. <https://doi.org/10.1007/s00335-013-9499-2>
- Schoenrock, S. A., D. Oreper, J. Farrington, R. C. McMullan, R. Ervin *et al.*, 2018 Perinatal nutrition interacts with genetic background to alter behavior in a parent-of-origin-dependent manner in adult Collaborative Cross mice. *Genes Brain Behav.* 17: e12438.
- Shorter, J. R., F. Odet, D. L. Aylor, W. Pan, C.-Y. Kao *et al.*, 2017 Male infertility is responsible for nearly half of the extinction observed in the mouse Collaborative Cross. *Genetics* 206: 557–572. <https://doi.org/10.1534/genetics.116.199596>
- Smith, C. C., and S. D. Fretwell, 1974 The optimal balance between size and number of offspring. *Am. Nat.* 108: 499–506. <https://doi.org/10.1086/282929>
- Sorge, R. E., L. J. Martin, K. A. Isbester, S. G. Sotocinal, S. Rosen *et al.*, 2014 Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat. Methods* 11: 629–632. <https://doi.org/10.1038/nmeth.2935>
- Sprague, G., and L. Tatum, 1942 General vs. specific combining ability in single crosses of corn. *J. Am. Soc. Agron.* 34: 923–932. <https://doi.org/10.2134/agronj1942.00021962003400100008x>
- Srivastava, A., A. P. Morgan, M. L. Najarian, V. K. Sarsani, J. S. Sigmon *et al.*, 2017 Genomes of the mouse Collaborative Cross. *Genetics* 206: 537–556. <https://doi.org/10.1534/genetics.116.198838>
- Stirling, D., D. Réale, and D. Roff, 2002 Selection, structure and the heritability of behaviour. *J. Evol. Biol.* 15: 277–289. <https://doi.org/10.1046/j.1420-9101.2002.00389.x>
- Topham, P. B., 1966 Diallel analysis involving maternal and maternal interaction effects. *Heredity* 21: 665–674. <https://doi.org/10.1038/hdy.1966.65>
- Trivers, R. L., 1974 Parent-offspring conflict. *Integr. Comp. Biol.* 14: 249–264.
- Turner, L. M., D. J. Schwahn, and B. Harr, 2012 Reduced male fertility is common but highly variable in form and severity in a natural house mouse hybrid zone. *Evolution* 66: 443–458. <https://doi.org/10.1111/j.1558-5646.2011.01445.x>
- Varona, L., and D. Sorensen, 2014 Joint analysis of binomial and continuous traits with a recursive model: a case study using mortality and litter size of pigs. *Genetics* 196: 643–651. <https://doi.org/10.1534/genetics.113.159475>
- Yu, G., 2009 Variance stabilizing transformations of Poisson, binomial and negative binomial distributions. *Stat. Probab. Lett.* 79: 1621–1629. <https://doi.org/10.1016/j.spl.2009.04.010>
- Zeileis, A. and C. Kleiber, 2018 *countreg: Count Data Regression*. R package version 0.2–1.
- Zhu, J., and B. Weir, 1996 Mixed model approaches for diallel analysis based on a bio-model. *Genet. Res.* 68: 233–240. <https://doi.org/10.1017/S0016672300034200>

Communicating editor: L. McIntyre

APPENDIX A

DIALLEL MODEL FOR LITTER SIZE

We collected data on litter size at weaning for 62 genetic crosses of inbred lines, across four years of breeding.

We use zero-truncated poisson (ZTP) regression for modeling our data. This type of regression is explicit in its framework accounting for discrete observations, flexible in its ability to use linear mixed models on the latent scale, and allows for parameterization of excess variance observed, in a way that standard Poisson regression does not. We account for the zero depletion in our data by using ZTP regression instead of standard Poisson regression, since we exclude observations of birth cohorts where no pups survived to weaning.

In Figure 4, the distribution of the observed data (litter size) is displayed for the WSB×WSB inbred mating, along with simulated data from a zero-truncated Poisson distribution based on the data mean.

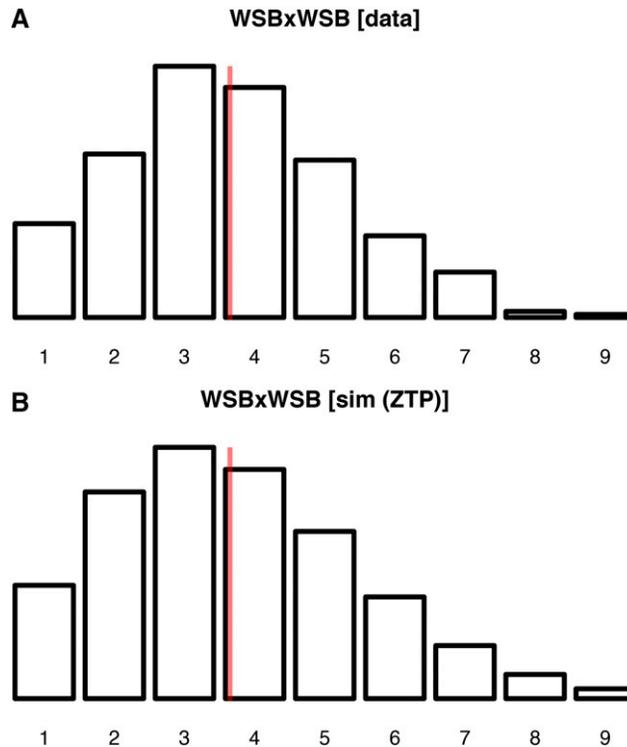


Figure 4 Comparison of observed data and (pseudo-)randomly generated data for the WSB×WSB inbreds. (A) Distribution of the litter diallel data for WSB×WSB (nlitters = 341, npups = 1,244). (B) Frequencies expected from a zero-truncated Poisson distribution with mean = 3.648 (or $\lambda = 3.542$, same as in A). The vertical red lines indicate the mean values.

For the ZTP, the first two moments (mean and variance), for values $y_i > 0$, are given by:

$$E[y_i] = \frac{\lambda_i}{1 - e^{-\lambda_i}} \quad (5)$$

$$V[y_i] = \frac{\lambda_i}{1 - e^{-\lambda_i}} \left(1 - \frac{\lambda_i e^{-\lambda_i}}{1 - e^{-\lambda_i}} \right) \quad (6)$$

The density for the ZTP, for every count $x \in 1, 2, \dots$, is given by:

$$\text{ZTPois}(x|\lambda) = \frac{\text{Pois}(x|\lambda)}{1 - \text{Pois}(0|\lambda)} \quad (7)$$

The relationship between the latent, expected, and data scales of the ZTP regression model are illustrated by the toy example shown in Figure 5. ZTP frequencies were calculated using the R package *countreg* (Zeileis and Kleiber 2018).

Diallel effect estimates

Diallel effect estimates are obtained using an MCMCglmm (Hadfield 2010) implementation of BayesDiallel (BayesDiallel-glm, henceforth), with 1.515×10^6 iterations, 1.5×10^4 iterations of burn-in, and thinning by 500 (saving only every 1/500th iteration), to obtain 3000 independent samples with minimal autocorrelation, and plotted as highest posterior densities (HPDs) in the results.

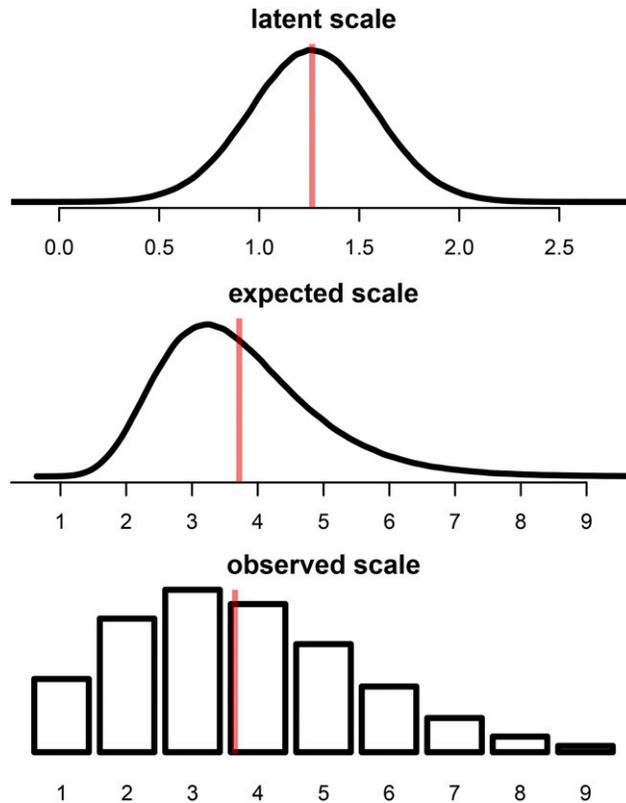


Figure 5 Latent and data scale comparison. An example illustrating zero-truncated Poisson regression of count data, with overdispersion, giving values drawn from pseudorandom normal variables drawn from $N(\text{mean} = 1.265, \text{sd} = 0.311)$, showing how the continuous *latent* scale values [top], which correspond to the scale of the linear predictor, map onto their *expected* (data) scale value (through the inverse link, $f(x) = e^x$) [middle], and how the mean *expected* (data) scale value corresponds to the *observed* (data) scale values (integers) from $ZTPois(\lambda = 3.542)$ [bottom]. These values were chosen to resemble the real distribution shown in Figure 4. The vertical red lines correspond to the mean of the values shown in each respective frame.

Conversion to the expected data scale

The effects estimated from the BayesDiallel-glm model are transformed from the latent scale to the (expected) data scale via the inverse link function, *i.e.* $\exp(a_{Aj})$, for interpretability of effects on the original data scale.

Variance Explained Using Variance Projection

To avoid the problem of interpretability in transforming variance parameter (and variance projection) estimates from the latent to the observed data scales, we instead calculate and report variance projections, as calculated using the Gaussian version of BayesDiallel. In order to account for heteroscedasticity (unequal variance) of the model residuals that arises from the approximately ZTPoisson nature of the data, we use a variance-stabilizing transformation (VST) (Yu, 2009) and run BayesDiallel again (Gaussian) to obtain Variance Projections on the modeled data. The VarPs that are calculated from these parameter estimates are an approximation of the variance contributions that we would observe in the GLMM BayesDiallel model.

APPENDIX B

DIALLEL MODEL FOR SEX RATIO

We model the male pup counts and female pup counts jointly, using the BayesDiallel linear model, formulated for binomial GLMM regression. This model directly considers genetic effects on the imbalance in male *vs.* female pups by parameterizing the number of males and number of total pups (or the total, and the fraction of males in the total). The model is elaborated in the methods section of the main manuscript.

The proportion of weaned pups that are male, or equivalently, the proportion of weaned pups that are female, is approximated by a binomial distribution. In our data, the mean and the variance of male ratio are 0.494 and 0.063, respectively.

To generate the upper and lower 95% boundaries, as shown in Figure S6, for the expected phenotype under the null hypothesis of male pup proportion = 0.5, we used the `qbinom` function in the stats package in R.

To generate the upper and lower 95% boundaries, as shown in Figure S6, for the expected phenotype under the null hypothesis of male pup proportion = 0.5, we used the `qbinom` function in the stats package in R.