© 2019. Published by The Company of Biologists Ltd | Journal of Experimental Biology (2019) 222, jeb202234. doi:10.1242/jeb.202234

RESEARCH ARTICLE

# The effect of environmental enrichment on behavioral variability depends on genotype, behavior, and type of enrichment

**Jamilla Akhund-Zade*, Sandra Ho*, Chelsea O'Leary and Benjamin de Bivort‡**

## ABSTRACT

Non-genetic individuality in behavior, also termed intragenotypic variability, has been observed across many different organisms. A potential cause of intragenotypic variability is sensitivity to minute environmental differences during development, which are present even when major environmental parameters are kept constant. Animal enrichment paradigms often include the addition of environmental diversity, whether in the form of social interaction, novel objects or exploratory opportunities. Enrichment could plausibly affect intragenotypic variability in opposing ways: it could cause an increase in variability due to the increase in microenvironmental variation, or a decrease in variability due to elimination of aberrant behavior as animals are taken out of impoverished laboratory conditions. In order to test these hypothesis, we assayed five isogenic *Drosophila melanogaster* lines raised in control and mild enrichment conditions, and one isogenic line under both mild and intense enrichment conditions. We compared the mean and variability of six behavioral metrics between our enriched fly populations and the laboratory housing control. We found that enrichment often caused a small increase in variability across most of our behaviors, but that the ultimate effect of enrichment on both behavioral means and variabilities was highly dependent on genotype and its interaction with the particular enrichment treatment. Our results support previous work on enrichment that presents a highly variable picture of its effects on both behavior and physiology.

**KEY WORDS: Quantitative genetics, Individuality, Intragenotypic variability, Microenvironmental sensitivity, *Drosophila melanogaster*, High-throughput assay**

## INTRODUCTION

Stable behavioral differences among conspecifics are seen in a wide array of species. These differences, caused (by definition) by some combination of genetic and environmental factors, are commonly referred to as individuality (Dall et al., 2004; Sih et al., 2004; Wolf and Weissing, 2010). Yet, even after experimentally homogenizing genotype and environment, individuality still persists (Gärtner, 2012), often undiminished or even increased (Ayroles et al., 2015; Buchanan et al., 2015). Multiple studies across different organisms demonstrate non-genetic individuality, which we refer to as intragenotypic variability (fruit flies: Ayroles et al., 2015; Kain et al., 2012; pea aphid: Schuett et al., 2011; nematodes: Stern et al., 2017; fish: Bierbach et al., 2017; crayfish: Vogt et al., 2008; mice: Freund et al., 2013; Kurikawa et al., 2018).

This intragenotypic variability may originate in sensitivity to stochastic microenvironmental effects that persist even when large-scale differences in environment across individuals are removed (Debat and David, 2001; Honegger and de Bivort, 2018; Willmore et al., 2007). Along these lines, environmental causes of phenotypic differences can be subdivided into deterministic (macro) and stochastic (internal or micro) aspects (Clarke, 1998; Willmore et al., 2007). Examples of macroenvironmental effects are different levels of fertilizer or different temperatures across treatments. Examples of microenvironmental effects include whether an individual animal ate more food in the morning or evening, or (in the case of flies) whether they pupated on the plastic vial or food media surface. Generally, microenvironmental effects exist within a treatment regime (Debat and David, 2001) and are hard to measure (Honegger and de Bivort, 2018). For individuals of the same genotype raised in a homogenous experimental environment, trait differences would be primarily due to microenvironmental effects, and the propensity to this variation is known as microenvironmental plasticity (Morgante et al., 2015). For a given trait, it would seem that intragenotypic variability would be maladaptive since some individuals are far from the trait optimum. Yet, in unpredictable and/or fluctuating environments, having intragenotypic variability can be advantageous (Simons, 2011). In such so-called diversifying bet-hedging strategies, variability can protect against sudden environmental changes, by increasing the likelihood that at any time a subset of the population has high fitness (Hopper, 1999; Simons, 2009).

Many studies have focused on characterizing the intragenotypic variability in morphological, physiological and behavioral traits (Abley et al., 2016; Ayroles et al., 2015; Bierbach et al., 2017; Blasco et al., 2017; Dworkin, 2005; Freund et al., 2013; Kain et al., 2012, 2015; Mellert et al., 2016; Morgante et al., 2015; Sørensen et al., 2015; Sztepanacz et al., 2017; Tonsor et al., 2013). In *Drosophila melanogaster*, intragenotypic variability was found in chill coma recovery time, starvation resistance, sternopleural bristles, wing traits and neuronal morphology in the larval ventral nerve cord (Dworkin, 2005; Mellert et al., 2016; Morgante et al., 2015; Sørensen et al., 2015; Sztepanacz et al., 2017). Morphological variations present in the ventral nerve cord and optic lobes are of particular note as they correlate with the timing of flight initiation and visually-guided locomotor biases, respectively, providing a link between morphological and behavioral intragenotypic variability (Linneweber et al., 2019 preprint; Mellert et al., 2016). Our research has identified intragenotypic variability in isogenic lines of *D. melanogaster* for turning bias, phototaxis and thermotaxis (Ayroles et al., 2015; Kain et al., 2012, 2015). Outside flies, intragenotypic variability in behavior has been

Department of Organismic and Evolutionary Biology and Center for Brain Science, Harvard University, Cambridge, MA 02138, USA.
*These authors contributed equally to this work

‡Author for correspondence (debivort@oeb.harvard.edu)

J.A., 0000-0001-5589-8258; S.H., 0000-0001-9621-8781; B.d.B., 0000-0001-6165-7696

**Received 22 February 2019; Accepted 9 August 2019**

studied in inbred mice and clonal fish (Amazon molly), with mice showing variation in exploratory behavior and fish showing variation in activity (Bierbach et al., 2017; Freund et al., 2013). If these examples of intragenotypic variability have their basis in microenvironmental differences, it may be hard to attribute the behavioral outcomes of specific individuals to their micro-causal underpinnings. It is, however, possible to test whether changing the degree of microenvironmental variation predicts changes in the amount of intragenotypic variability.

As most lab organisms are already raised in heavily standardized environments where microenvironmental variation is minimized, it is feasible to increase microenvironmental variation and examine the effects on behavior. 'Enrichment' treatments include a variety of modifications to regular laboratory housing, such as opportunities for exercise, novel object interaction and socialization (van Praag et al., 2000). Enrichment may add microenvironmental variation to a particular treatment, potentially affecting both the mean and variance of phenotypic traits (Körholz et al., 2018). Typically, enrichment treatments are hypothesized to more closely match an animal's natural habitat, increasing mean well-being and cognition (while perhaps increasing intragenotypic variability). For mice and rats, enrichment has been shown to enhance mean gliogenesis, neurogenesis, and synapse formation in the cortex, hippocampus and cerebellum, leading to improved memory and cognition (Bruel-Jungerman et al., 2005; Garthe et al., 2016; Leger et al., 2015; Mohammed et al., 2002; van Praag et al., 2000). While there are some consistent enrichment effects in rats and mice, physiological and behavioral strain differences in response to enrichment have been observed (Konkle et al., 2010; Toth et al., 2011; van de Weerd et al., 1994). There is conflicting evidence for the effect of enrichment on brain size in fish, with enrichment having no effect on three-spined sticklebacks but decreasing brain size in eastern mosquitofish (Toli et al., 2017; Turschwell and White, 2016). Early studies in *D. melanogaster* have found that changing the social milieu affects the size of brain structures, with social isolation leading to decreased sizes and numbers of Kenyon cell fibers (Barth and Heisenberg, 1997; Heisenberg et al., 1995; Technau, 1984). In addition, social isolation in *D. melanogaster* leads to faster cancer progression, suggesting that a stimulating social environment buffers against stresses (Dawson et al., 2018). In crickets, mushroom body neurogenesis is higher in enriched environments with complex visual, olfactory and auditory stimuli as compared to impoverished environments (Scotto Lomassese et al., 2000). On the other hand, enriched olfactory environments did not change mushroom body calyx size or affect odor learning in *D. melanogaster* (Wang et al., 2018).

The effect of enrichment on trait variability has been primarily studied in mice and rats, with researchers asking whether enriched rearing and housing conditions would decrease the statistical power to detect treatment effects by increasing within-sample variance (Toth et al., 2011). The evidence presented from behavioral and physiological studies has been conflicting: studies have shown that enrichment can increase, decrease or have no effect on variability depending on the trait in question (Freund et al., 2015; Toth et al., 2011; van de Weerd et al., 1994, 1997, 2002; Wolfer et al., 2004). A recent study by Körholz et al. (2018) chose to focus more directly on whether intragenotypic variability in behavior and brain plasticity is influenced by the diversity of experiences that results from an enriched environment. They found that enrichment increases variation in specific domains: mice from enriched environments showed higher variation in exploratory behavior (object interaction times, habituation, but not locomotion), adult neurogenesis and motor cortex thickness (Körholz et al., 2018). They attributed this increase in variation directly to the diversity of experiences or diversity of microenvironments that individuals could explore in an enriched environment.

Given the conflicting evidence for the effects of enrichment on variability, we propose two hypotheses for how enrichment may influence variation in traits. The first hypothesis is that enrichment introduces microenvironmental differences, which, in turn, increase trait differences through microenvironmental plasticity. Our second hypothesis is that by more closely matching natural conditions to which organisms are adapted, enrichment increases the robustness of development and somatic maintenance, with a corresponding reduction in variation (due to the removal of aberrant phenotypes that may appear in impoverished laboratory conditions). Even though they predict opposite outcomes, both of these hypotheses are intuitive, and have some support in the literature. We chose to test them by measuring intragenotypic variability in *D. melanogaster* under control and enriched treatments. This species is suitable for this work because of the ease of rearing large experimental groups from isogenic lines, and its suitability for automated behavioral phenotyping. We were also interested in testing whether the observed effects of enrichment were dependent on genotype. We measured a variety of behavioral metrics associated with spontaneous locomotion and phototaxis (Ayroles et al., 2015; Buchanan et al., 2015; Kain et al., 2012) in one isogenic line across two enrichment treatments and five isogenic lines in a single enrichment treatment to examine the effects of enrichment and the interaction of genotype and enrichment on behavioral variability. We found that while enrichment often caused a small increase in intragenotypic variability, the predominant determinants of behavioral means and variabilities were genotype and its interaction with the particular enrichment treatment.

## MATERIALS AND METHODS

### Behavior and enrichment protocols

Fly stocks were cultured in vials on Caltech formula medium at 25°C in temperature-controlled incubators on a 12 h:12 h light:dark cycle. For our isogenic populations, we used Canton-S and four lines from the *Drosophila* Genetic Reference Panel (DGRP; Mackay et al., 2012). These lines (numbers 45, 105, 535 and 796) were derived from different wild-caught gravid females and then inbred for 20 generations. Thus, there is significant genetic variation between the lines, but not within them. We chose to work with these particular lines because we have previously observed that they vary in intragenotypic variability in Y-maze turn bias (Ayroles et al., 2015). Flies were subjected to two enrichment treatments: mild enrichment (the addition of a small 'jungle gym' to each culture vial) and intense enrichment (growth of the flies in a 1 $m^3$ cage filled with many rotting fruit substrates, plants, rocks etc.) (Fig. 1).

For mild enrichment, in each vial of medium where experimental animals were to develop, 3 female and 2 male parental flies were housed for 3–5 days. The parents were removed and the jungle gym enrichment was inserted. The enrichment object consisted of plastic tubing, pipe cleaners and fuzzy pom-poms that were hot-glued to a balsa wood applicator stick that was inserted into the medium (Fig. 1C). These were identically constructed for ∼30 vials, with the exception of the pom-pom color, which in some vials was white and in the others pink at random. $F_1$ experimental progeny developed in this enriched environment for around 10 days in incubators. Once they started eclosing, they were allowed to accumulate for 1–2 days, after which they were removed and mixed under cold anesthetization with other flies from the same genotype. They were then sorted into cohorts of 40 males and 40 females and placed

Journal of Experimental Biology (2019) 222, jeb202234. doi:10.1242/jeb.202234

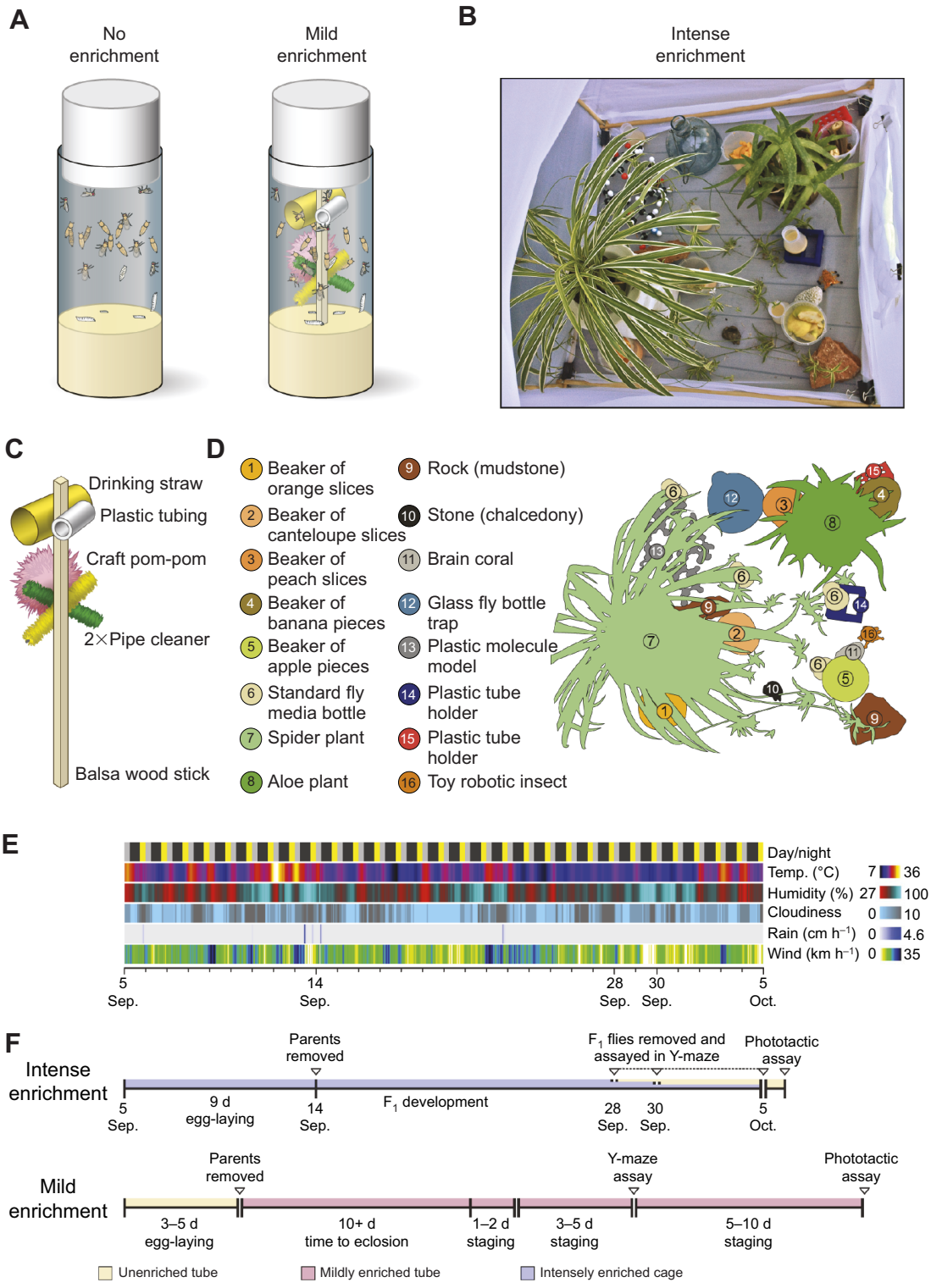


**Fig. 1. Illustration of enrichment paradigms used.** (A) Control vials and mild enrichment vials. (B) Photo of the inside of the intense enrichment cage. (C) Diagram of the mild enrichment jungle gym components. (D) The intense enrichment cage components. (E) Weather conditions that the intense enrichment cage was subject to for the experimental period. For the daylight timeline, yellow indicates potential direct sunlight on the cage, gray periods where the cage was shaded by our building and black is night time. The cloudiness timeline reflects the NOAA 10-point scale where 0 is clear skies and 10 is full cloud cover. (F) Timelines of experiments showing the development, staging and behavioral testing of the experimental animals for both mild and intense enrichment treatments. Each contiguous horizontal line indicates the time spent in a fresh container. The unenriched control was the same as the mild enrichment, except all vials used were unenriched.

in mildly enriched vials for 3–5 days. At this point, their behavior was measured in the Y-mazes for 2 h according to the methods in Buchanan et al. (2015) (although here we loaded anesthetized experimental animals on ice to transfer them into the Y-mazes, rather than $CO_2$). After the Y-maze assay, flies were anesthetized on ice and returned to their mildly enriched vials for 5–10 days at

Journal of Experimental Biology (2019) 222, jeb202234. doi:10.1242/jeb.202234

which point their phototactic preferences were measured using FlyVac according to the methods in Kain et al. (2012). This mild enrichment procedure was used for flies of all genotypes. See Fig. 2 for representations of the six behavioral metrics we acquired in these two assays, and their distributions across Canton-S flies.

For intense enrichment, we prepared a population cage using 1 m wooden dowels to make a cubic frame, with sheer white polyester drapery material as walls (Fig. 1B). A tube of this material, normally held closed by binder clips, provided access to the inside of the cage. The items shown in Fig. 1D were introduced to the cage at the time of its construction: six kinds of fly food (a variety of decomposing fruits as well as bottles of standard cornmeal medium), houseplants, stones, varied plastic objects etc. The cage was placed outside on a deck where it experienced natural fluctuations in luminance, temperature, rainfall, wind, humidity, etc. during the course of our experiment (Fig. 1E). For the experiment, a parental Canton-S population of 200 males and 200 females was placed in the cage on 5 September 2013, and removed 9 days later. $F_1$ flies were collected and assayed in the Y-mazes on 28 and 30 September, and 5 October 2013. Flies were recovered from the Y-mazes using cold anesthetization and stored in unenriched standard medium tubes in groups of ∼30 individuals until testing with FlyVac on 6 October 2013 (Fig. 1F). For all assays, males and females were tested in equal proportions.

For Y-maze enrichment experiments with Canton-S, we assayed 151 control flies, 203 mildly enriched flies and 206 intensely enriched flies. For Canton-S FlyVac experiments, we assayed 175 control flies, 140 mildly enriched flies and 86 intensely enriched flies. For Y-maze enrichment experiments with the DGRP lines, we assayed: DGRP 45: 166 control, 133 mildly enriched; DGRP 105: 130 control, 148 mildly enriched; DGRP 535: 113 control, 111 mildly enriched; DGRP 796: 132 control, 128 mildly enriched. For FlyVac enrichment experiments with DGRP lines, we assayed: DGRP 45: 157 control, 144 mildly enriched; DGRP 535: 122 control, 140 mildly enriched.

### Behavior measures and null model distributions

Behavior measures from the Y-maze assay (turn bias, number of turns, turn-direction switchiness and turn-timing clumpiness) were calculated from the vectors of turn directions and times that each fly produced in experiment. Behavior measures from FlyVac (light-choice probability and inter-choice interval) were calculated from the FlyVac data output file (Kain et al., 2012). These measures were computed and/or collected into a common data structure in MATLAB 2013a (The MathWorks, Inc., Natick, MA). With Y-maze arrays (Buchanan et al., 2015), we measured the left-vs.-right free locomotion turning bias of individual flies. With FlyVac, we measured the locomotory response to light cues of agitated flies (‘fast phototaxis’; Scott, 1943). We have previously used both of these assays to detect genetic and neural circuit regulators of intragenotypic variability (Ayroles et al., 2015; Buchanan et al., 2015; Kain et al., 2012), and between them, we examined both spontaneous and stimulus-evoked behaviors. In addition to turn bias in the Y-maze assay, we assessed: (1) the number of turns completed by individual flies within the 2 h trials, (2) flies’ tendencies to alternate between left and right turns successively (‘switchiness’), and (3) the extent to which their turning events were clustered in time (‘clumpiness’). For the FlyVac dataset, we measured the average interval between phototactic choices in addition to the light-choice probability. While the Y-maze arrays and FlyVac were primarily designed to measure locomotor turning and phototaxis, respectively, they also produce precise estimates of these other individual behavioral measures, so for the purpose of this study, we do not emphasize any of these measures over the others.

Null hypothesis distributions (Fig. 2) were generated in MATLAB 2013a by resampling (with replacement) a million

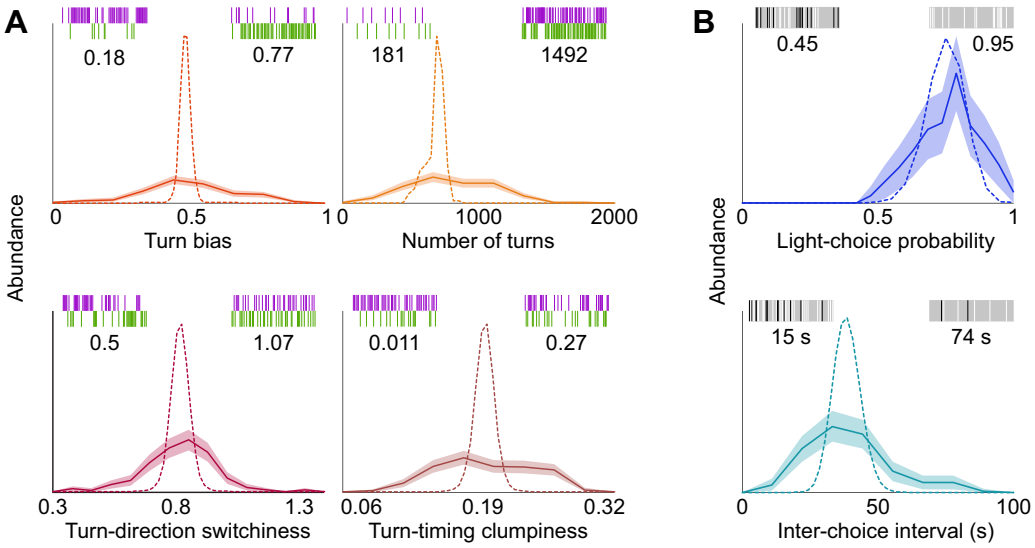


**Fig. 2. Observed and null hypothesis distributions of Y-maze and FlyVac behavioral measures for Canton-S (wild type) flies.** Dotted lines represent the distributions expected under null hypotheses in which all individuals exhibit behaviors drawn from identical distributions. The solid line represents the observed distribution, with the shaded region representing ±1 standard error of the distribution, as estimated by bootstrap resampling. Insets show 10 min of representative data of the original behavior traces of extreme individuals and the corresponding value of that metric. (A) Metrics from the Y-maze assay: turn bias is the fraction of turns made to the right, number of turns is the number of left-right choices made in the 2 h test, turn direction switchiness is a turn bias-normalized measure of the mutual information between successive turns (for higher values, left turns are more predictive of subsequent left turns and vice versa) and turn-timing clumpiness is a normalized measure of the irregularity of turns (the mean absolute deviation of the inter-turn intervals divided by the mean inter-turn interval). Purple lines represent left turns and green lines represent right turns. (B) Metrics from the FlyVac phototaxis assay: light-choice probability is the fraction of choices toward light, inter-choice interval is the mean time between choices. White lines indicate a light choice, black lines indicate a dark choice and shaded gray areas represent regions of time where no choice is made. 151 flies were analyzed for Y-maze behaviors, and 175 flies for FlyVac behaviors.

values for each distribution as follows. (1) For turn bias and light-choice probability: (i) all observed choice values (i.e. left versus right and light versus dark) were pooled across individuals, (ii) an individual was chosen at random from all tested, and a vector of length equal to the number of behavioral choices performed by that individual during the experiment was populated randomly with values from the pool, and (iii) the turn bias or light-choice probability for that vector was recorded. (2) For number of turns: (i) the observed inter-turn intervals (ITIs) were pooled across individuals, (ii) ITIs were chosen randomly one at a time until their cumulative sum exceeded 7,200,000 ms, the length of an experiment, and (iii) the number of turns in this sequence was recorded. The moderate discrepancy in mean between the null hypothesis distribution and experimental distribution in this analysis arises from the disproportionate number of short intervals contributed to the total pool by more active animals. (3) For switchiness, which arises from slight dependence between consecutive turns in the L–R (left–right) turn sequence (Ayroles et al., 2015): (i) we implemented a Markov chain in which the L–L (=R–R) transition probabilities yielded L–R sequences with mutual information between successive turns equal to the observed mutual information [0.018 bits, $P$(L–L)=0.592]; (ii) an individual was chosen at random from all tested, and a choice sequence of length equal to the number of choices performed by that individual during the experiment was generated using the Markov chain, and (iii) the switchiness of this sequence [=(#L–R+#R–L)/2×turn bias×(1–turn bias)×no. turns] was recorded. (4) For clumpiness: (i) the observed ITIs were pooled across individuals, (ii) an individual was chosen at random from all tested, and a vector of length equal to the number of choices performed by that individual during the experiment was populated randomly with values from the ITI pool, and (iii) clumpiness [=MAD(ITIs)/mean(ITIs)] was recorded, where MAD is the median absolute deviation from the median. Thus, this approach reflects sampling variation in number of turns and mean ITI as well as clumpiness. (5) For inter-choice interval (ICI): (i) the observed ICIs were pooled across individuals, (ii) an individual was chosen at random from all tested, and a vector of length equal to the number of choices performed by that individual during the experiment was populated randomly with values from the ICI pool, (iii) the mean ICI across this vector was recorded.

### Bayesian inference of mean and variance effects

To get the estimates of the posterior distributions of behavioral mean and variance, and the effects on the observed distributions of enrichment treatment, genotype, and their interactions, we employed linear and generalized linear models in the R Stan interface v.2.18.2 (https://mc-stan.org/users/interfaces/rstan.html). The Stan platform allows the user to specify desired models and performs full Bayesian inference using Hamiltonian Monte Carlo with the No U-Turn sampler (Carpenter et al., 2017). To get the posterior distributions of the mean and variance for turn bias, light-choice probability, switchiness, and clumpiness under different enrichment-genotype conditions, we specified the following model:

$$\begin{aligned} y_n &\sim \text{Normal}(\mu_n,\ \sigma_n).\\ \mu &= a + \mathbf{X}\cdot b,\\ \sigma^2 &= v_0 + \mathbf{X}\cdot v, \end{aligned} \tag{1}$$

where $y_n$ is the behavioral outcome of an individual $n$ that comes from a normal distribution with parameters $\mu_n$ (mean) and $\sigma_n$ (s.d.). $\mu$ and $\sigma^2$ are vectors specified via linear models, where $a$ and $v_0$ are intercepts, $\mathbf{X}$ is a logical predictor matrix specifying the genotype and/or enrichment treatment for each individual, and $b$ and $v$ are vectors of coefficients of the linear model.

Since the distribution of the number of turns was right-skewed, bounded to real positive integers, and overdispersed compared to a Poisson distribution, we chose to model this measure with a negative binomial as follows:

$$\begin{aligned} y_n &\sim \text{NegBinomial}(\mu_n,\ \phi_n),\\ \mu &= \exp(\mathbf{X}\cdot\beta),\\ \phi &= \exp(\mathbf{X}\cdot\gamma),\\ \sigma^2 &= \frac{\mu + \mu^2}{\phi}. \end{aligned} \tag{2}$$

Here, $y_n$ is the vector of number of turns made by individual $n$ modeled by a negative binomial distribution with parameters $\mu_n$ (mean) and $\phi_n$ (dispersion). Both parameter vectors are related to the coefficients of a generalized linear model (vectors $\beta$ and $\gamma$) via a log-link function. $\mathbf{X}$ is the experimental design matrix, as above; $\sigma^2$ is the variance vector calculated from the mean and dispersion parameter vectors.

To model inter-choice intervals, we chose a gamma distribution since the data is right-skewed and positive continuous:

$$\begin{aligned} y_n &\sim \Gamma(a_n,\ b_n),\\ \mu &= \exp(\mathbf{X}\cdot\beta),\\ \sigma^2 &= \exp(\mathbf{X}\cdot\gamma),\\ a &= \frac{\mu^2}{\sigma^2},\\ b &= \frac{\mu}{\sigma^2}, \end{aligned} \tag{3}$$

where $y_n$ is the inter-choice interval of individual $n$, and $a_n$ and $b_n$ are the shape and rate parameters of the gamma distribution, respectively, and the rest of the parameters are as above.

For the estimation of all posterior distributions, we set our priors on the coefficients to broad Cauchy distributions centered at 0 to allow them to be weakly informative (Gelman et al., 2008). Our qualitative findings were robust to the choice of prior, as uninformative uniform told the same story (Fig. S1). To sample posteriors, we used four chains and 50,000–100,000 iterations per chain, with the target average proposal acceptance probability of 0.8–0.9 and maximum tree depth of 10–15, to generate a posterior distribution of 100,000–200,000 samples (50% of chain iterations were used for tuning the Hamiltonian Monte Carlo sampler parameters and were discarded as the burn-in period). The ratio of chain effective sample size to sample size was in the range of 0.7–1.2, indicating that posterior estimate error due to autocorrelation was minimal. To get posterior distributions for the coefficients of variation, we took the square root of the variance and divided it by the mean at each step in the chain. To check our model fits, we carried out graphical posterior predictive checks (Fig. S2), and found that our models fit the data well.

We adapted the methodology used by Kruschke's BEST method (Kruschke, 2013) to our own posterior distributions in order to determine which effects were inferred to differ from zero. To estimate the posterior distribution of a treatment effect, we subtracted the parameter values of one treatment condition from the control (or the other treatment condition) at each step in the chain and took the distribution of that difference. We calculated the 99% highest density interval (the credible interval) of the posterior of treatment effects to evaluate whether the treatments had an effect: if the 99% highest density interval excluded 0, we inferred an effect between the treatments. Since this approach is subject to multiple

comparisons concerns, we chose the 99% credible interval (rather than e.g. 95%) as a more stringent indicator of effects. While we think that the 99% credible interval is a useful guide to pulling out the strongest effects we observe, we believe that the strength of Bayesian inference lies in being able to examine the posterior distributions as a whole and observing their trends (rather than applying a threshold to identify effects).

To determine the contribution of genotype, mild enrichment, and genotype-by-mild enrichment effects to the variability (coefficient of variance) and mean of each behavior, we used the following formulas:

$$CV_{ij} = CV_0 + G_i + E_j + G_i \times E_j, \quad (4)$$

$$\mu_{ij} = \mu_0 + G_i + E_j + G_i \times E_j, \quad (5)$$

where $CV_{ij}$ and $\mu_{ij}$ are the variability and mean, respectively, of genotype $i$ in treatment group $j$. $CV_0$ and $\mu_0$ are the grand variability and mean, averaged over all genotypes and treatments. $G_i$ is the deviation of the variability or mean of genotype $i$ from the grand parameter in question, calculated over all treatment groups. $E_j$ is the deviation of treatment group $j$ from the grand parameter, calculated over all genotypes. The treatment groups in this experiment were mild enrichment or control vials. $G_i \times E_j$ is the specific deviation of genotype $i$ in treatment group $j$ after accounting for the main effects of genotype $i$ and treatment group $j$. All deviations were standardized by dividing them by the grand parameter value in order to interpret them as effect sizes.

## RESULTS

### Intragenotypic variability is evident in locomotor and phototactic behaviors

To measure intragenotypic variability, we employed two automated assays (which measure spontaneous locomotion and locomotor responses to light) to rapidly collect many behavioral observations from many individual flies. We first confirmed that intragenotypic variability was present in a standard lab wild-type strain, Canton-S, in left–right turn bias and light-choice probability (Fig. 2). Indeed, the observed distributions of these measures were significantly broader ($P<0.001$ by bootstrapping, $\chi^2$ and Kolmogorov–Smirnov tests) than expected under null models in which all individuals behaved identically, i.e. sequences of behavior drawn from identical distributions (see Materials and Methods).

We next asked if there was evidence of intragenotypic variability in other measurements taken while measuring turn bias and phototactic preference in these assays. As with turn bias, the observed distributions of number of turns, turn switchiness and turn clumpiness were significantly broader than expected under null models in which all flies behaved identically (Fig. 2A). Using FlyVac data, we observed that the distribution across flies of the average interval between phototactic choices also was broader than expected if all flies were behaving identically (Fig. 2B). Thus, intragenotypic variability was evident in all six behavioral traits examined.

### Enrichment affects behavioral means in a genotype-, measure- and enrichment-dependent manner

We developed enrichment protocols that were either ‘mild’ or ‘intense’ (Fig. 1). Our mild enrichment treatment, the fly jungle gym, was designed to provide a variety of textures, colors and light conditions, whereas the intense cage enrichment was designed for flies to experience natural weather conditions, in addition to several different foods and a greater variety of biotic and abiotic substrates.

To confirm that the enrichment treatments had an effect on our flies, we examined the mean values of our six behavioral phenotypes under each treatment. We used a Bayesian framework with a weakly informative prior to estimate the posterior distributions of the means and variances of each behavioral metric under the mild and intense enrichment treatments. We used the 99% highest density interval, also termed the credible interval (Kruschke, 2013; see Materials and Methods) to assess whether the posterior distributions of the means of each behavioral metric were different from each other. Intense enrichment caused strong decreases in the mean of number of turns and inter-choice interval, and a strong increase in the turn switchiness when compared with mild enrichment and the control (Fig. 3A). For these behaviors, intense enrichment had a larger effect on the mean than mild enrichment. Mild enrichment had a less pronounced effect on the mean number of turns and turn clumpiness. There was no apparent effect of enrichment on turn bias and light-choice probability, although the FlyVac assay has lower power than the Y-maze assay. We viewed these observed mean changes as a positive control that the flies were sensitive to our enrichment treatments. Our results were supported by a non-parametric test of mean differences (Fig. S3A).

We estimated the effects of genotype, mild enrichment and genotype×mild enrichment on behavioral means using four DGRP lines (45, 105, 535, 796) and Canton-S (Fig. 4). Genotype had an effect (i.e. zero was not in the 99% credible interval of the posterior distribution) on all behaviors except turn bias. Mild enrichment caused a genotype-independent increase in number of turns and switchiness, but had no effect on any of the other behaviors. There were genotype×mild enrichment effects on number of turns, switchiness and clumpiness, and the direction of those effects was variable.

### Genotype, behavioral measure, enrichment and their interactions determine intragenotypic variability

We examined the effect of mild and intense enrichment on intragenotypic variability in our behavioral measures (Fig. 3B). We chose to look at the coefficient of variation as our measure of intragenotypic variability in order to standardize it across multiple types of measures and control for mean effects (estimates of the posterior distributions of variance effects, not normalized by the treatment means, are included in Figs S4, S5). For nearly all behaviors, intense enrichment had a larger effect on variability than mild enrichment, but these effects were not all in the same direction (results of a non-parametric test of differences in variance are shown in Fig. S3B). Intense enrichment decreased the variability of turn bias and turn direction switchiness but increased the variability of number of turns and inter-choice interval, when compared with the control and mild enrichment treatments. Intense enrichment increased variability in clumpiness, although the effect was more pronounced upon comparison to mild enrichment as opposed to the control.

Mild enrichment caused small or no differences (zero effect was within the 99% credible interval) when compared with the control treatment for all the behavioral measures from both assays, with turn-direction switchiness and turn-timing clumpiness the most likely behavioral measures to be affected by mild enrichment. Variability in turn bias and number of turns probably increased slightly under mild enrichment, whereas clumpiness and switchiness probably decreased slightly. In two of these cases, the direction of the effect matched the direction of the intense enrichment effect; in the other two cases, it did not. To summarize, intense enrichment had stronger effects on variability than mild enrichment and the direction of these effects was behavior dependent.

Journal of Experimental Biology (2019) 222, jeb202234. doi:10.1242/jeb.202234

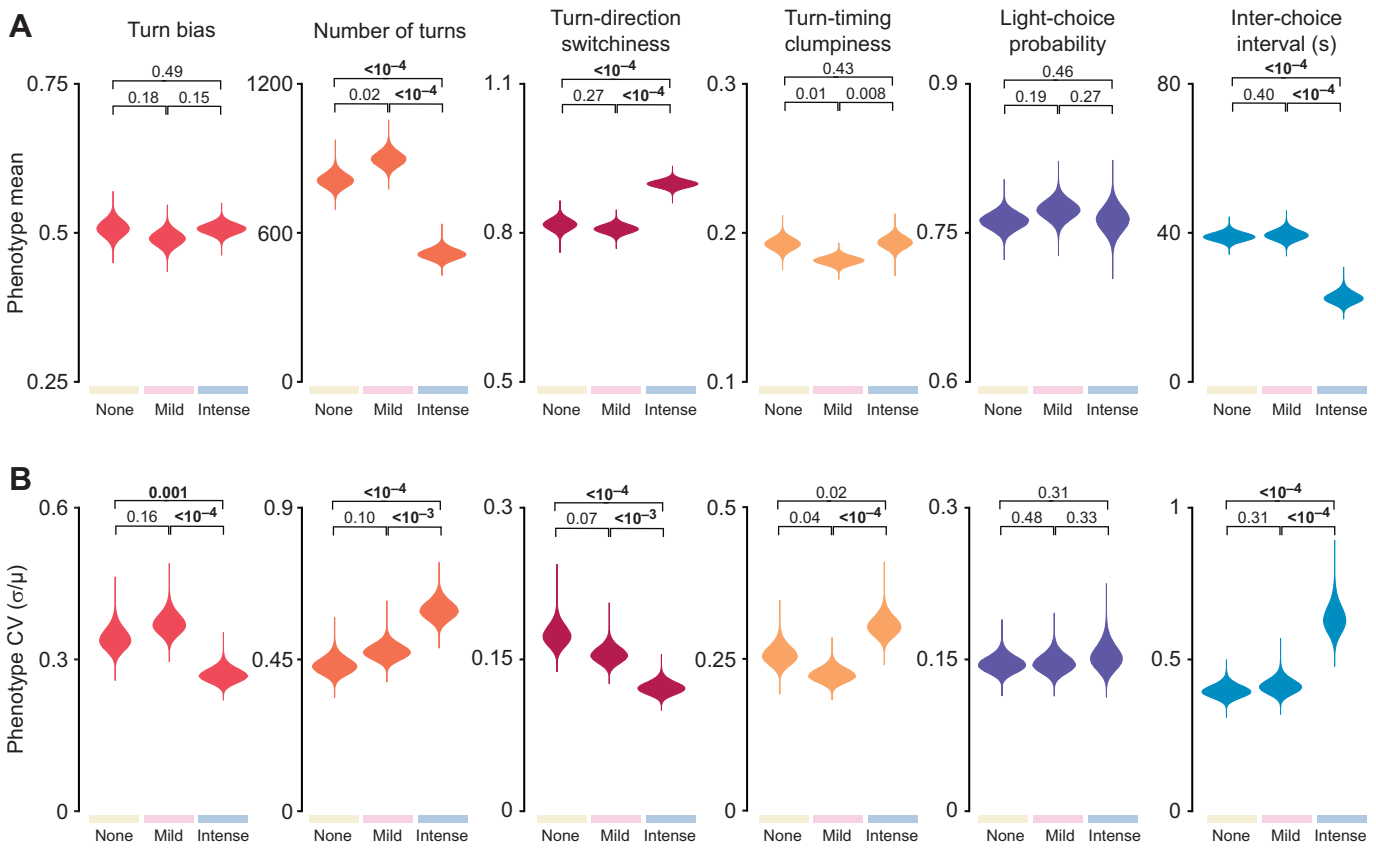


**Fig. 3. Posterior distributions of mean and intragenotypic variability (coefficient of variation; CV) for Canton-S flies under different enrichment treatments.** Values shown at the top of each plot are the fraction of the posterior distribution of differences in behavior means (A) and variabilities (B) between two treatments (e.g. control and mild enrichment) that lies either below or above zero, depending on the direction of change. Given our finite posterior sampling, we cannot estimate fractions of distributions accurately below approximately $10^{-4}$. Bold values indicate treatments for which the 99% credible interval of the treatment effect does not include 0. Sample sizes for each experiment are provided in the Materials and Methods.

In our analysis of five isogenic lines (four DGRP and Canton-S) under unenriched and mildly enriched conditions, we found that genotype, mild enrichment and genotype×mild enrichment all had effects on intragenotypic variability (Fig. 5). We found that the variability of all behavioral measurements, except the number of turns, were affected by genotype. The variability of number of turns, clumpiness, and light choice increased in a genotype-independent manner under mild enrichment. Variability of switchiness and inter-choice interval were probably also increased in a genotype-independent manner by mild enrichment (a large majority of their respective posterior distributions was above zero). We observed genotype×mild enrichment effects for number of turns, switchiness, clumpiness and light-choice probability. Of all the behavioral measures, switchiness showed the most variable and the strongest genotype×enrichment effects. To summarize, mild enrichment often increased variability in a genotype-independent fashion, but there were also frequently genotype×mild enrichment effects.

Interestingly, we found that the average magnitudes of mean effects were smaller than the average magnitudes of variability effects (Fig. 6A,B). For both mean and variability, genotype effects tended to be larger than the mild enrichment or genotype×mild enrichment effects. This pattern was especially prominent for mean effects. We also found that the size of the effects on behavioral means was uncorrelated with the size of the effects on variability (Fig. 6C).

## DISCUSSION

The goal of our study was to test two opposing hypotheses about the effects of enrichment on intragenotypic behavioral variability. We hypothesized that enrichment could increase variability due to the increase in microenvironmental diversity or decrease variability due to enriched environments more closely mimicking natural conditions (resulting in more robust development of behaviors and elimination of extremes that can occur in impoverished conditions) (Körholz et al., 2018). To this end, we examined six behaviors in *D. melanogaster* across several genotypes and employed two levels of enrichment. We found that for five of the six behaviors, when examined across several genotypes, mild enrichment via a fly jungle gym probably led to an increase in intragenotypic variability, supporting the hypothesis that enrichment causes an increase in behavioral variability as a result of the increase in microenvironmental diversity. However, these genotype-independent effects were generally smaller than the effects of genotype or genotype×enrichment interactions. Therefore, the effect of enrichment on intragenotypic variability appears to depend on the particular genotype and behavior being assayed. When we examined the effects of enrichment on variability within a genotype, we found support for both hypotheses depending on the behavioral measure and enrichment treatment examined (mild or intense). Therefore, while it was broadly true that mild enrichment caused a small increase in the intragenotypic variability, a more granular look at the effects of enrichment revealed both

Journal of Experimental Biology (2019) 222, jeb202234. doi:10.1242/jeb.202234

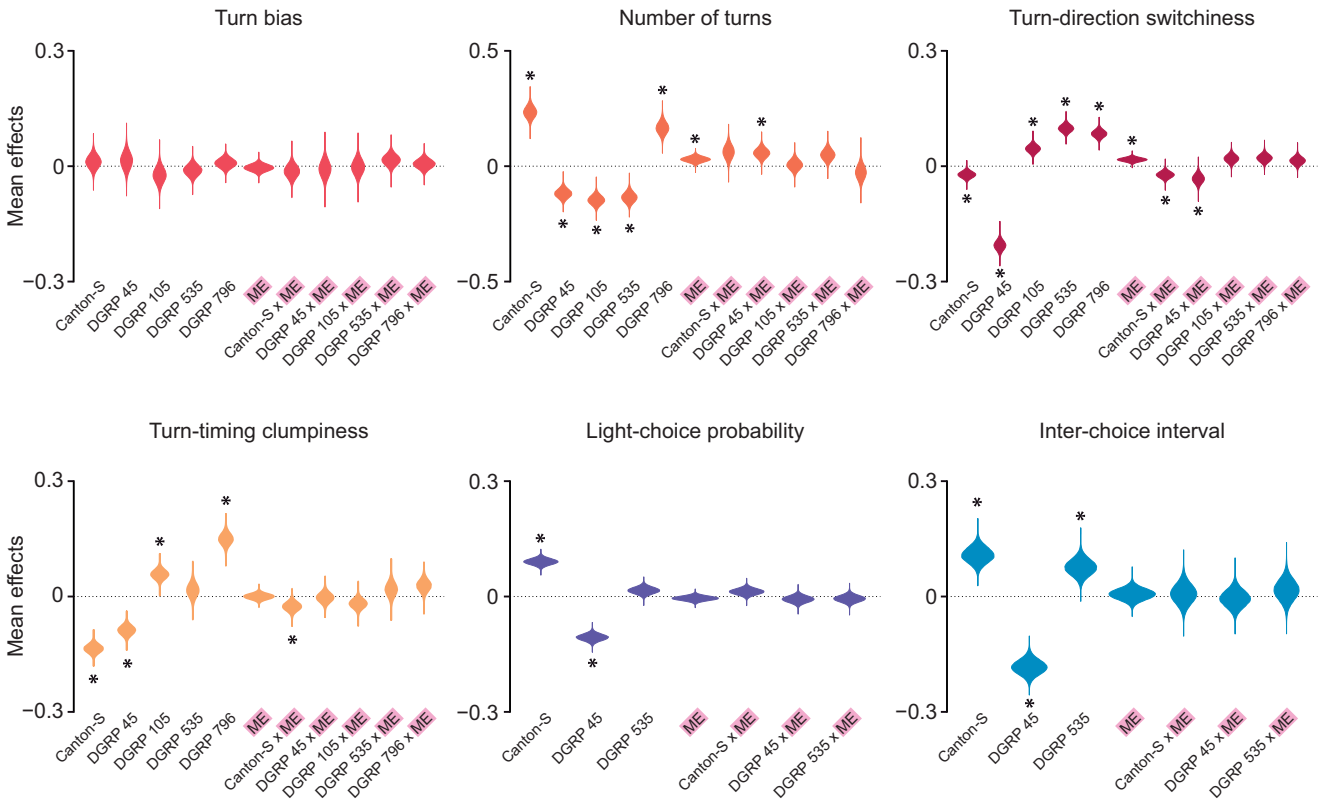


**Fig. 4. Genotype, mild enrichment and genotype×mild enrichment effects on behavioral measure means.** Asterisks mark those effects whose 99% credible interval does not include zero. All effects were normalized by the grand mean of all treatments and genotypes, so these values can be interpreted as effect sizes of each condition on the mean. Sample sizes for each experiment are provided in the Materials and Methods. ME, mild enrichment.

strong increases and decreases in variability. We also found that the effects of enrichment on behavioral means and variabilities were largely independent of each other, with variability effects having larger magnitudes than mean effects (Fig. 6). This finding confirms that our enrichment paradigm was able to affect both mean and variability, and that these effects are potentially independent. From our experiments, it remains uncertain which aspects of enrichment influence mean and which influence variability, or indeed, if these aspects are one and the same. Behavioral variability being more strongly affected by enrichment may underscore a biological flexibility that is not present in determining mean behavior.

With respect to both mean and variability, we found that genotype usually had a larger effect than the mild enrichment. This was especially obvious when looking at the genotype effects on behavioral means, where all behaviors except turn bias showed large genotype effects (Figs 4 and 6). The lack of effect of genotype on turn bias is consistent with previous work that found no differences in the mean turn biases of 159 DGRP lines (Ayroles et al., 2015). Genotype also had strong effects on intragenotypic variability (Fig. 5), as expected (Ayroles et al., 2015; Buchanan et al., 2015; Kain et al., 2012).

We found evidence of interactions between genotype and enrichment for practically all the behavioral measures examined, although the magnitude of these interactions was behavior dependent (Fig. 6). For example, turn bias and inter-choice interval showed very little genotype-by-enrichment effect for variability, but large effects were seen for turn switchiness. Dependence of variability on the particular parameter measured was previously noted in mouse enrichment studies (Körholz et al., 2018; van de Weerd et al., 2002). Behavioral parameters may fall into different categories with respect to their response to enrichment. For example, switchiness is a measure of intra-individual variability (Fig. 2), hinting at a link between the biological mechanism controlling variability from trial to trial and individual to individual (Stamps et al., 2013). Our results also make it clear that in assessing the effects of enrichment on a particular measure of behavior, genotype cannot be ignored. These interactions are consistent with previous findings in rats and mice (Konkle et al., 2010; Toth et al., 2011; van de Weerd et al., 1994), where the effects of enrichment differed between strains. Our results also imply that there is genotype-dependent plasticity in variability. In essence, phenotype variability is not a static feature of a genotype, but depending on the trait measured, the environment can have a large effect. Evolution of plasticity has usually been examined in trait means. Our findings suggest such inquiries should extend to variability. In the future, it may be necessary to consider variability as a flexible, evolvable trait to understand how phenotype distributions arise.

Our fly jungle gym enrichment (mild enrichment) featured an array of perching sites and materials. The goal was that flies on the jungle gyms would experience a diversity of perching sites and textures, as well as be forced to navigate a more complex environment. Flies would therefore be subject to a diversity of experiences closer to what they might have in the wild while still under the constraints of laboratory conditions. We expected that the mild enrichment would mostly affect locomotion and activity behaviors, such as turn bias and inter-choice interval. We found that mild enrichment caused an increase in the mean of number of turns and switchiness, as well as an increase in the variability of all behaviors except turn bias (Figs 4 and 5). Surprisingly, we saw that

Journal of Experimental Biology (2019) 222, jeb202234. doi:10.1242/jeb.202234

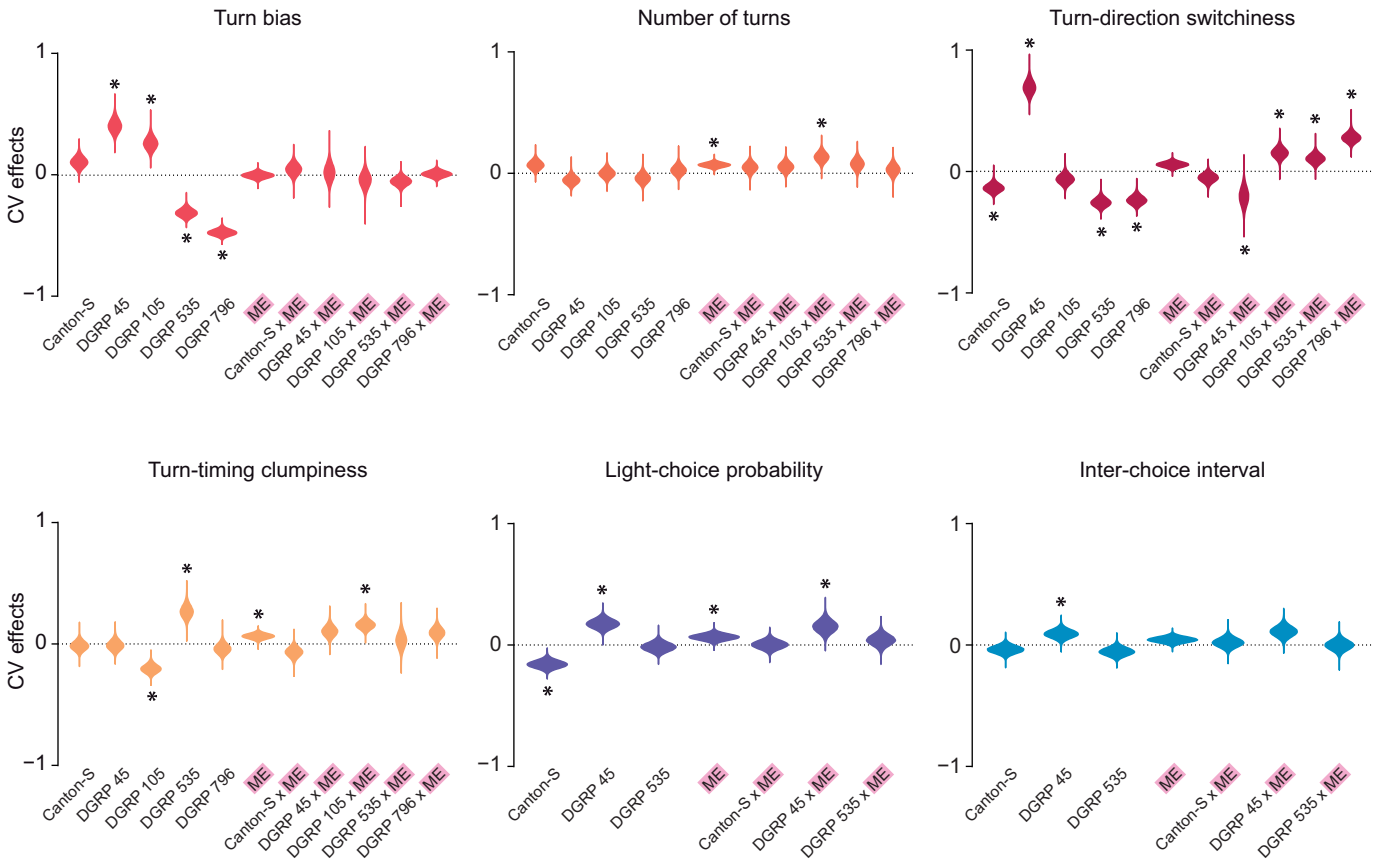


**Fig. 5. Genotype, mild enrichment and genotype×mild enrichment effects on measures of behavioral intragenotypic variability.** Asterisks mark those effects whose 99% credible interval does not include zero. All effects were normalized by the grand variability of all treatments and genotypes, so these values can be interpreted as effect sizes of each condition on intragenotypic variability. Sample sizes of each experiment are provided in the Materials and Methods. ME, mild enrichment.

variability in light-choice probability also increased under the mild enrichment treatment, which led us to believe that our jungle gym construction may have also created differential light conditions in the vial or stimulated phototaxis variability via more indirect means.

The six behavior measures examined were chosen largely because they could be measured at scale across many individuals, which is a requirement for measuring effects on variability. However, we can speculate on the ecological relevance of variability in several of the phenotypes measured. We suspect that variability in turn bias could be potentially advantageous for exploration, dispersal and/or foraging via a bet-hedging mechanism. Individuals with stronger turning biases move through the environment with lower effective diffusion constants. If the spatial scale of resources in the environment fluctuates

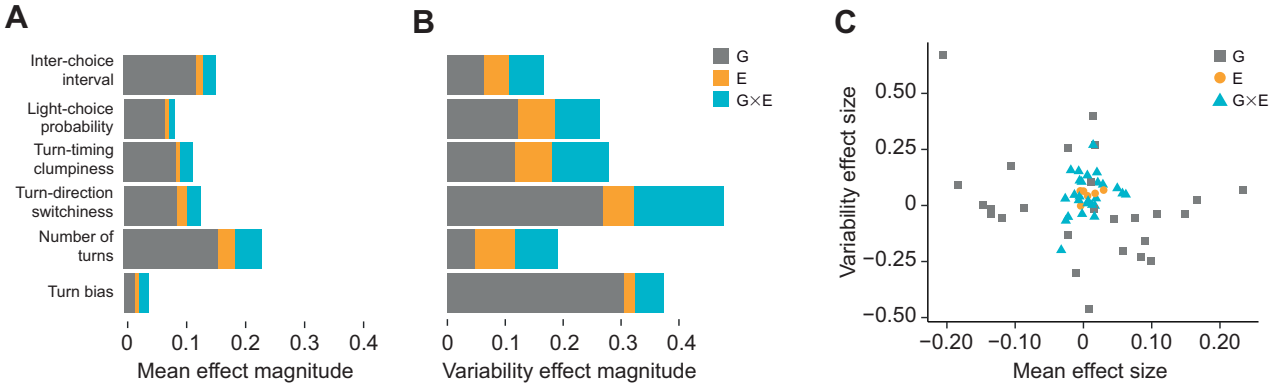


**Fig. 6. Summary of the genotype, mild enrichment and genotype×mild enrichment effects on measures of behavioral mean and intragenotypic variability.** (A,B) Length of the bar represents the average magnitude of the effect; G, genotype effect; E, mild enrichment effect; G×E, genotype×mild enrichment effect. (C) Scatter plot of variability effect sizes and mean effect sizes across all behaviors, separated into genotype, mild enrichment and genotype×mild enrichment effects. All effects were normalized by either the grand mean or variability of all treatments and genotypes, so these values can be interpreted as effect sizes of each condition.

unpredictably, variation in turn bias could reflect a matched strategy of diversifying diffusion constants. Light-choice bias in our assay may be reflective of an escape response since the fly is startled prior to the light choice (Kain et al., 2012). Variability in moving toward light upon being startled could reflect a bet-hedging mechanism as well, if the threats faced by flies are variable, i.e. if it is alternatively advantageous to seek light or dark after a startle. Predator escape behavior has been considered to be a possible bet-hedging trait; for example, clonal pea aphids show variability in predator escape behavior among individuals (Schuett et al., 2011). We have also found that variability in light preference under non-startled conditions can influence the thermal experience of a fruit fly in nature and therefore could also be part of a bet-hedging strategy (Kain et al., 2015).

We examined whether the effects on variability would change with a different type of enrichment. We raised one cohort of flies in a naturalistic setting, subject to the environmental fluctuations of the outdoors and with access to numerous organic and inorganic substrates (Fig. 1). Compared with the jungle gyms of the mild enrichment, this intense enrichment treatment had more structural complexity and diversity of biotic (fruits, plants, spider predators) and abiotic (sunlight, temperature) factors. By increasing the microenvironmental diversity along several different axes, we expected to observe stronger effects on variability on flies reared in this treatment. In general, this intense enrichment did have stronger effects on our behavioral measures than the mild enrichment. Even though the effects of intense enrichment were more pronounced, the direction of these effects was behavior dependent (Fig. 3). For example, we saw a decrease in intragenotypic variability for turn bias and turn switchiness under intense enrichment, but an increase in the variability for number of turns and turn clumpiness. The directions of these effects varied by behavior, even relative to the direction of the mild enrichment effect.

The intense enrichment treatment was created to have a higher level of microenvironmental diversity than the mild enrichment treatment, yet we observed that the behavioral variability did not always change in the same direction between these treatments. For example, we saw a small increase in variability under mild enrichment but a large decrease under intense enrichment, such as for turn bias (Fig. 3). This leads us to believe that the relationship between the mild and intense enrichment is not just a simple increase in 'enrichment intensity'. However, we recorded some direct evidence that flies in the intense enrichment indeed experienced at least one dimension of increased microenvironmental variation compared with the mild enrichment: flies recovered from the intense enrichment cage exhibited variation in the color of their gut contents, consistent with their having recently fed on different food sources (Fig. S6). This variation was absent in flies subject to mild enrichment. One of our predictions was that a more naturalistic enrichment treatment could lead to a decrease in variability because of an increase in robustness, but it could also be that naturalistic enrichments cause fly populations to exhibit more natural behaviors in general, whether or not that corresponds to a decrease in behavioral variability. Future studies could address what constitutes natural fly behavior in more detail, whether by making field-deployable assays or bringing wild flies directly to the lab for testing, although any comparisons with our current enrichment paradigm would need to carefully consider population genotypic variance.

Overall, our results support the hypothesis that enrichment increases intragenotypic variability, although this effect is highly dependent on the particular genotype, enrichment and behavior in question. We also conclude that genotype is likely the main determinant of intragenotypic variability. Our findings make it apparent that the genotype used and behavior measured will affect the inferred relationship between environmental variability and behavioral variability. Moreover, the type of enrichment (e.g. mild vs. intense enrichment) can qualitatively and quantitatively alter this relationship. This, and the effects of genotype and behavioral measure, could be why effects observed in one enrichment study may not be replicated in another (Toth et al., 2011). While the multifactorial nature of enrichment provides challenges, its specific effects continue to be of great interest in behavioral research, and high-throughput data-driven approaches have the potential to illuminate the complex relationships between environmental variability and behavioral variability.

#### Acknowledgements
We are grateful to Kyle Honegger for discussions of Bayesian inference and Nico O. Wagner for feedback on the manuscript. We also thank Carolyn Elya for moral support when the sampler misbehaved.

#### Competing interests
The authors declare no competing or financial interests.

#### Author contributions
Conceptualization: S.H., B.d.B.; Methodology: J.A., S.H., C.O., B.d.B.; Software: J.A., B.d.B.; Formal analysis: J.A.; Investigation: S.H., C.O., B.d.B.; Resources: B.d.B.; Data curation: J.A., B.d.B.; Writing - original draft: J.A., B.d.B.; Writing - review & editing: J.A., S.H., C.O., B.d.B.; Visualization: J.A., B.d.B.; Supervision: B.d.B.; Project administration: B.d.B.; Funding acquisition: B.d.B.

#### Funding
S.H. was supported by the Harvard PRISE program. J.A. was supported by the National Science Foundation PoLS-HF-SRN (NSF-1806818). B.d.B. was supported by the Alfred P. Sloan Foundation, Esther A. and Joseph Klingenstein Fund, the Susan Smith Family Foundation and the National Science Foundation (IOS-1557913).

#### Data availability
All of the measures of individual behavior and all analysis scripts are publicly available at https://zenodo.org/record/2573158. They are also hosted on our lab website at http://lab.debivort.org/enrichment.

#### Supplementary information
Supplementary information available online at
http://jeb.biologists.org/lookup/doi/10.1242/jeb.202234.supplemental

#### References
**Abley, K., Locke, J. C. W. and Leyser, H. M. O.** (2016). Developmental mechanisms underlying variable, invariant and plastic phenotypes. *Ann. Bot.* **117**, 733-748. doi:10.1093/aob/mcw016

**Ayroles, J. F., Buchanan, S. M., O'Leary, C., Skutt-Kakaria, K., Grenier, J. K., Clark, A. G., Hartl, D. L. and de Bivort, B. L.** (2015). Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *Proc. Natl. Acad. Sci. USA* **112**, 6706-6711. doi:10.1073/pnas.1503830112

**Barth, M. and Heisenberg, M.** (1997). Vision affects mushroom bodies and central complex in *Drosophila melanogaster*. *Learn. Mem.* **4**, 219-229. doi:10.1101/lm.4.2.219

**Bierbach, D., Laskowski, K. L. and Wolf, M.** (2017). Behavioural individuality in clonal fish arises despite near-identical rearing conditions. *Nat. Commun.* **8**, 15361. doi:10.1038/ncomms15361

**Blasco, A., Martínez-Álvaro, M., García, M.-L., Ibáñez-Escriche, N. and Argente, M.-J.** (2017). Selection for environmental variance of litter size in rabbits. *Genet. Sel. Evol. GSE* **49**, 48. doi:10.1186/s12711-017-0323-4

**Bruel-Jungerman, E., Laroche, S. and Rampon, C.** (2005). New neurons in the dentate gyrus are involved in the expression of enhanced long-term memory following environmental enrichment. *Eur. J. Neurosci.* **21**, 513-521. doi:10.1111/j.1460-9568.2005.03875.x

**Buchanan, S. M., Kain, J. S. and de Bivort, B. L.** (2015). Neuronal control of locomotor handedness in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **112**, 6700-6705. doi:10.1073/pnas.1500804112

**Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P. and Riddell, A.** (2017). Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1-32. doi:10.18637/jss.v076.i01

**Clarke, G. M.** (1998). The genetic basis of developmental stability. V. Inter- and intra-individual character variation. *Heredity* **80**, 562-567. doi:10.1046/j.1365-2540.1998.00294.x

**Dall, S. R. X., Houston, A. I. and McNamara, J. M.** (2004). The behavioural ecology of personality: consistent individual differences from an adaptive perspective. *Ecol. Lett.* **7**, 734-739. doi:10.1111/j.1461-0248.2004.00618.x

**Dawson, E. H., Bailly, T. P. M., Dos Santos, J., Moreno, C., Devilliers, M., Maroni, B., Sueur, C., Casali, A., Ujvari, B., Thomas, F. et al.** (2018). Social environment mediates cancer progression in *Drosophila*. *Nat. Commun.* **9**, 3574. doi:10.1038/s41467-018-05737-w

**Debat, V. David, P.** (2001). Mapping phenotypes: canalization, plasticity and developmental stability. *Trends Ecol. Evol.* **16**, 555-561. doi:10.1016/S0169-5347(01)02266-2

**Dworkin, I.** (2005). A Study of canalization and developmental stability in the sternopleural bristle system of *Drosophila melanogaster*. *Evolution* **59**, 1500-1509. doi:10.1111/j.0014-3820.2005.tb01799.x

**Freund, J., Brandmaier, A. M., Lewejohann, L., Kirste, I., Kritzler, M., Krüger, A., Sachser, N., Lindenberger, U. and Kempermann, G.** (2013). Emergence of individuality in genetically identical mice. *Science* **340**, 756-759. doi:10.1126/science.1235294

**Freund, J., Brandmaier, A. M., Lewejohann, L., Kirste, I., Kritzler, M., Krüger, A., Sachser, N., Lindenberger, U. and Kempermann, G.** (2015). Association between exploratory activity and social individuality in genetically identical mice living in the same enriched environment. *Neuroscience* **309**, 140-152. doi:10.1016/j.neuroscience.2015.05.027

**Garthe, A., Roeder, I. and Kempermann, G.** (2016). Mice in an enriched environment learn more flexibly because of adult hippocampal neurogenesis. *Hippocampus* **26**, 261-271. doi:10.1002/hipo.22520

**Gärtner, K.** (2012). A third component causing random variability beside environment and genotype: a reason for the limited success of a 30 year long effort to standardize laboratory animals? *Int. J. Epidemiol.* **41**, 335-341. doi:10.1093/ije/dyr219

**Gelman, A., Jakulin, A., Pittau, M. G. and Su, Y.-S.** (2008). A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2**, 1360-1383. doi:10.1214/08-AOAS191

**Heisenberg, M., Heusipp, M. and Wanke, C.** (1995). Structural plasticity in the *Drosophila* brain. *J. Neurosci.* **15**, 1951-1960. doi:10.1523/JNEUROSCI.15-03-01951.1995

**Honegger, K. and de Bivort, B.** (2018). Stochasticity, individuality and behavior. *Curr. Biol.* **28**, R8-R12. doi:10.1016/j.cub.2017.11.058

**Hopper, K. R.** (1999). Risk-spreading and bet-hedging in insect population biology. *Annu. Rev. Entomol.* **44**, 535-560. doi:10.1146/annurev.ento.44.1.535

**Kain, J. S., Stokes, C. and de Bivort, B. L.** (2012). Phototactic personality in fruit flies and its suppression by serotonin and white. *Proc. Natl. Acad. Sci. USA* **109**, 19834-19839. doi:10.1073/pnas.1211988109

**Kain, J. S., Zhang, S., Akhund-Zade, J., Samuel, A. D. T., Klein, M. and Bivort, B. L. and de** , (2015). Variability in thermal and phototactic preferences in *Drosophila* may reflect an adaptive bet-hedging strategy. *Evolution* **69**, 3171-3185. doi:10.1111/evo.12813

**Konkle, A. T., Kentner, A. C., Baker, S. L., Stewart, A. and Bielajew, C.** (2010). Environmental-enrichment–related variations in behavioral, biochemical, and physiologic responses of Sprague–Dawley and Long Evans rats. *J. Am. Assoc. Lab. Anim. Sci.* **49**, 427-436.

**Körholz, J. C., Zocher, S., Grzyb, A. N., Morisse, B., Poetzsch, A., Ehret, F., Schmied, C. and Kempermann, G.** (2018). Selective increases in inter-individual variability in response to environmental enrichment in female mice. *eLife* **7**, e35690. doi:10.7554/eLife.35690

**Kruschke, J. K.** (2013). Bayesian estimation supersedes the *t*-test. *J. Exp. Psychol. Gen.* **142**, 573-603. doi:10.1037/a0029146

**Kurikawa, T., Haga, T., Handa, T., Harukuni, R. and Fukai, T.** (2018). Neuronal stability in medial frontal cortex sets individual variability in decision-making. *Nat. Neurosci.* **21**, 1764-1773. doi:10.1038/s41593-018-0263-5

**Leger, M., Paizanis, E., Dzahini, K., Quiedeville, A., Bouet, V., Cassel, J.-C., Freret, T., Schumann-Bard, P. and Boulouard, M.** (2015). Environmental enrichment duration differentially affects behavior and neuroplasticity in adult mice. *Cereb. Cortex NYN* **25**, 4048-4061. doi:10.1093/cercor/bhu119

**Linneweber, G., Andriatsilavo, M., Dutta, S., Hellbruegge, L., Liu, G., Ejsmont, R., Fenk, L., Straw, A., Wernet, M., Hiesinger, P. R. et al.** (2019). A neurodevelopmental origin of behavioral individuality. *bioRxiv* 540880. doi:10.1101/540880

**Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., Casillas, S., Han, Y., Magwire, M. M., Cridland, J. M. et al.** (2012). The *Drosophila melanogaster* genetic reference panel. *Nature* **482**, 173. doi:10.1038/nature10811

**Mellert, D. J., Williamson, W., Shirangi, T. R., Card, G. M. and Truman, J. W.** (2016). Genetic and environmental control of neurodevelopmental robustness in *Drosophila*. *PLoS ONE* **11**, e0155957. doi:10.1371/journal.pone.0155957

**Mohammed, A. H., Zhu, S. W., Darmopil, S., Hjerling-Leffler, J., Ernfors, P., Winblad, B., Diamond, M. C., Eriksson, P. S. and Bogdanovic, N.** (2002). Environmental enrichment and the brain. *Prog. Brain Res.* **138**, 109-133. doi:10.1016/S0079-6123(02)38074-9

**Morgante, F., Sørensen, P., Sorensen, D. A., Maltecca, C. and Mackay, T. F. C.** (2015). Genetic architecture of micro-environmental plasticity in *Drosophila melanogaster*. *Sci. Rep.* **5**, 9785. doi:10.1038/srep09785

**Schuett, W., Dall, S. R. X., Baeumer, J., Kloesener, M. H., Nakagawa, S., Beinlich, F. and Eggers, T.** (2011). Personality variation in a clonal insect: the pea aphid, *Acyrthosiphon pisum*. *Dev. Psychobiol.* **53**, 631-640. doi:10.1002/dev.20538

**Scott, J. P.** (1943). Effects of single genes on the behavior of *Drosophila*. *Am. Nat.* **77**, 184-190. doi:10.1086/281116

**Scotto Lomassese, S., Strambi, C., Strambi, A., Charpin, P., Augier, R., Aouane, A. and Cayre, M.** (2000). Influence of environmental stimulation on neurogenesis in the adult insect brain. *J. Neurobiol.* **45**, 162-171. doi:10.1002/1097-4695(20001115)45:3<162::AID-NEU4>3.0.CO;2-X

**Sih, A., Bell, A. M., Johnson, J. C. and Ziemba, R. E.** (2004). Behavioral syndromes: an integrative overview. *Q. Rev. Biol.* **79**, 241-277. doi:10.1086/422893

**Simons, A. M.** (2009). Fluctuating natural selection accounts for the evolution of diversification bet hedging. *Proc. R. Soc. B Biol. Sci.* **276**, 1987-1992. doi:10.1098/rspb.2008.1920

**Simons, A. M.** (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proc. R. Soc. B Biol. Sci.* **278**, 1601-1609. doi:10.1098/rspb.2011.0176

**Sørensen, P., de los Campos, G., Morgante, F., Mackay, T. F. C. and Sorensen, D.** (2015). Genetic control of environmental variation of two quantitative traits of *Drosophila melanogaster* revealed by whole-genome sequencing. *Genetics* **201**, 487-497. doi:10.1534/genetics.115.180273

**Stamps, J. A., Saltz, J. B. and Krishnan, V. V.** (2013). Genotypic differences in behavioural entropy: unpredictable genotypes are composed of unpredictable individuals. *Anim. Behav.* **86**, 641-649. doi:10.1016/j.anbehav.2013.07.012

**Stern, S., Kirst, C. and Bargmann, C. I.** (2017). Neuromodulatory control of long-term behavioral patterns and individuality across development. *Cell* **171**, 1649-1662.e10. doi:10.1016/j.cell.2017.10.041

**Sztepanacz, J. L., McGuigan, K. and Blows, M. W.** (2017). Heritable micro-environmental variance covaries with fitness in an outbred population of *Drosophila serrata*. *Genetics* **206**, 2185-2198. doi:10.1534/genetics.116.199075

**Technau, G. M.** (1984). Fiber number in the mushroom bodies of adult *Drosophila melanogaster* depends on age, sex and experience. *J. Neurogenet.* **1**, 113-126. doi:10.3109/01677068409107077

**Toli, E. A., Noreikiene, K., DeFaveri, J. and Merilä, J.** (2017). Environmental enrichment, sexual dimorphism, and brain size in sticklebacks. *Ecol. Evol.* **7**, 1691-1698. doi:10.1002/ece3.2717

**Tonsor, S. J., Elnaccash, T. W. and Scheiner, S. M.** (2013). Developmental instability Is genetically correlated with phenotypic plasticity, constraining heritability and fitness. *Evolution* **67**, 2923-2935. doi:10.1111/evo.12175

**Toth, L. A., Kregel, K., Leon, L. and Musch, T. I.** (2011). Environmental enrichment of laboratory rodents: the answer depends on the question. *Comp. Med.* **61**, 314-321.

**Turschwell, M. P. and White, C. R.** (2016). The effects of laboratory housing and spatial enrichment on brain size and metabolic rate in the eastern mosquitofish, *Gambusia holbrooki*. *Biol. Open* **5**, 205-210. doi:10.1242/bio.015024

**van de Weerd, H. A., Baumans, V., Koolhaas, J. M. and van Zutphen, L. F.** (1994). Strain specific behavioural response to environmental enrichment in the mouse. *J. Exp. Anim. Sci.* **36**, 117-127.

**van de Weerd, H. A., van Loo, P. L., van Zutphen, L. F., Koolhaas, J. M. and Baumans, V.** (1997). Nesting material as environmental enrichment has no adverse effects on behavior and physiology of laboratory mice. *Physiol. Behav.* **62**, 1019-1028. doi:10.1016/S0031-9384(97)00232-1

**van de Weerd, H. A., Aarsen, E. L., Mulder, A., Kruitwagen, C. L. J. J., Hendriksen, C. F. M. and Baumans, V.** (2002). Effects of environmental enrichment for mice: variation in experimental results. *J. Appl. Anim. Welf. Sci.* **5**, 87-109. doi:10.1207/S15327604JAWS0502_01

**van Praag, H., Kempermann, G. and Gage, F. H.** (2000). Neural consequences of environmental enrichment. *Nat. Rev. Neurosci.* **1**, 191-198. doi:10.1038/35044558

**Vogt, G., Huber, M., Thiemann, M., van den Boogaart, G., Schmitz, O. J. and Schubart, C. D.** (2008). Production of different phenotypes from the same genotype in the same environment by developmental variation. *J. Exp. Biol.* **211**, 510-523. doi:10.1242/jeb.008755

**Wang, X., Amei, A., de Belle, J. S. and Roberts, S. P.** (2018). Environmental effects on Drosophila brain development and learning. *J. Exp. Biol.* **221**, jeb169375. doi:10.1242/jeb.169375

**Willmore, K. E., Young, N. M. and Richtsmeier, J. T.** (2007). Phenotypic variability: its components, measurement and underlying developmental processes. *Evol. Biol.* **34**, 99-120. doi:10.1007/s11692-007-9008-1

**Wolf, M. and Weissing, F. J.** (2010). An explanatory framework for adaptive personality differences. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 3959-3968. doi:10.1098/rstb.2010.0215

**Wolfer, D. P., Litvin, O., Morf, S., Nitsch, R. M., Lipp, H.-P. and Würbel, H.** (2004). Laboratory animal welfare: cage enrichment and mouse behaviour. *Nature* **432**, 821-822. doi:10.1038/432821a