Behavioral idiosyncrasy reveals genetic control of phenotypic variability

Julien F. Ayroles[a,b,c,1], Sean M. Buchanan[d], Chelsea O'Leary[a,d,e], Kyobi Skutt-Kakaria[a,e], Jennifer K. Grenier[c], Andrew G. Clark[c], Daniel L. Hartl[a,1], and Benjamin L. de Bivort[a,d,e]

[a]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; [b]Harvard Society of Fellows, Harvard University, Cambridge, MA 02138; [c]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853; [d]Rowland Institute at Harvard, Cambridge, MA 02142; and [e]Center for Brain Science, Harvard University, Cambridge, MA 02138

Contributed by Daniel L. Hartl, March 20, 2015 (sent for review January 23, 2015; reviewed by Gholson J. Lyon and Mark L. Siegal)

Quantitative genetics has primarily focused on describing genetic effects on trait means and largely ignored the effect of alternative alleles on trait variability, potentially missing an important axis of genetic variation contributing to phenotypic differences among individuals. To study the genetic effects on individual-to-individual phenotypic variability (or intragenotypic variability), we used *Drosophila* inbred lines and measured the spontaneous locomotor behavior of flies walking individually in Y-shaped mazes, focusing on variability in locomotor handedness, an assay optimized to measure variability. We discovered that some lines had consistently high levels of intragenotypic variability among individuals, whereas lines with low variability behaved as although they tossed a coin at each left/right turn decision. We demonstrate that the degree of variability is itself heritable. Using a genome-wide association study (GWAS) for the degree of intragenotypic variability as the phenotype across lines, we identified several genes expressed in the brain that affect variability in handedness without affecting the mean. One of these genes, *Ten-a*, implicates a neuropil in the central complex of the fly brain as influencing the magnitude of behavioral variability, a brain region involved in sensory integration and locomotor coordination. We validated these results using genetic deficiencies, null alleles, and inducible RNAi transgenes. Our study reveals the constellation of phenotypes that can arise from a single genotype and shows that different genetic backgrounds differ dramatically in their propensity for phenotypic variabililty. Because traditional mean-focused GWASs ignore the contribution of variability to overall phenotypic variation, current methods may miss important links between genotype and phenotype.

variability | variance QTL | DGRP | ten-a | personality

Quantitative genetics was founded on the assumption that phenotypic variation is explained solely by differences in mean phenotypes among genotypes. Under this model, intragenotypic variability is assumed to be attributable to nongenetic environmental perturbations (1). There is, however, growing evidence for the importance of genetic control of variance (2–4) and that variance itself is a quantitative trait. Although studies of morphology (5–7) and animal breeding (8, 9) have long noted the heterogeneity of variance among genotypes, this axis of variation has received little attention compared with the effect of genetic variation on trait means. As a result, the mechanisms by which variable phenotypes arise from a uniform genetic background are still poorly understood, particularly in the context of behavior, where variability may be a critical determinant of phenotypic differences (10, 11). Most recently, with the advent of genome-wide association studies, several groups (3, 4, 12, 13) have mapped quantitative trait loci affecting variance (vQTLs) by comparing phenotypic variances among individuals that share alleles. These studies examine the average effect of QTL alleles across genetic backgrounds and heterogeneous environments across individuals (14), in the process losing any specific effects intrinsic to each individual.

Here, we examine diversity that is typically hidden in population averages by examining phenotypic variability among individuals with the same genotype. This diversity is the variation that we would observe if we could generate a large number of copies of individuals of the same genotype in a common environment and measure a trait across them [an experiment for which isogenic lines (5–7, 14, 15) are especially suited]. In this case, phenotypic differences among genetically identical individuals result from subtle microenvironmental perturbations and stochasticity in development, whereas differences in variability among genotypes reflect genetic differences in developmental stability (7). Although intragenotypic variability contributes to phenotypic variation in a population, this source of variation is not usually estimable because, with few exceptions, each individual in an outbred diploid population is a unique instance of its genotype (Fig. 1*A*). As a consequence we have little understanding of the causes and consequences of interindividual intragenotypic variability. This phenotypic variance nevertheless has wide ranging implications. In evolutionary biology, variability offers an adaptive solution to environmental changes (15, 16). In medical genetics, many diseased states emerge beyond a phenotypic threshold, and high variability genotypes will produce a larger proportion of individuals exceeding that threshold than low variability genotypes, even if each genotypic class has the same mean. Although intragenotypic variability has been discussed in animal behavior, particularly in the context of the emergence of personality (10, 17),

Significance

If we could rear genetically identical individuals from a variety of genetic backgrounds and rear them in the same environment, how much phenotypic variation between individuals of the same genotype would we see? Would different genetic backgrounds differ in their degree of variability? What would account for these differences? We used *Drosophila* inbred lines to address these questions focusing on variability in locomotor handedness. We show that different genotypes vary dramatically in their propensity for variability, that phenotypic variability itself, as a trait, can be heritable, and that loci affecting variability can be mapped. The genetic control of variability has received little attention in quantitative genetics despite the important role variability plays in explaining phenotypic variation between individuals.

Author contributions: J.F.A. and B.L.d.B. designed research; J.F.A., S.M.B., C.O., K.S.-K., J.K.G., and B.L.d.B. performed research; J.F.A., A.G.C., D.L.H., and B.L.d.B. contributed new reagents/analytic tools; J.F.A., S.M.B., and B.L.d.B. analyzed data; and J.F.A., S.M.B., J.K.G., A.G.C., D.L.H., and B.L.d.B. wrote the paper.

Reviewers: G.J.L., Cold Spring Harbor Laboratory; and M.L.S., New York University.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. Email: ayroles@fas.harvard.edu or dhartl@oeb.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1503830112/-/DCSupplemental.

to date no genes have been associated with behavioral variability that do not also affect the mean.

To study phenotypic variability, we used a panel of wild-derived *Drosophila* inbred lines. These inbred lines are an ideal tool because the genetic variation that was present between individual flies in their natural population is now captured between lines in the panel. For each line, this allows us to measure any phenotype on a large number of individuals of the same genetic background, age, and rearing environment, thus empirically estimating the magnitude of intragenotypic variability (Fig. 1*A*). Specifically, we measured the spontaneous locomotor behavior of flies walking individually in Y-shaped mazes (18), focusing on the variability in locomotor handedness (left-right turning bias). The precision and high-throughput nature of our assays allows a large number of flies to be measured per genotype and permits robust estimates of the sampling error on variance itself.

Results

We tracked 2 h of locomotor behavior of 110 individuals (on average) from each of 159 lines from the *Drosophila* Genetic Reference Panel (DGRP) in a randomized block design. For each individual fly, we recorded the time and left-right direction of each turn in the maze (Fig. 1*B*), estimating a turn bias score as the fraction of turns that were to the right. Flies performing more than 50 turns were analyzed and completed 413 turns per trial on average. We began by comparing the mean turning bias and found no significant genetic variation across lines (Fig. S1). In other words, averaged across individuals within a line, each line is unbiased, making an equal proportion of left and right turns (with the modal fly being unbiased; Fig. 1). We verified the lack of genetic variation for turning bias within lines by crossing pairs of males and females with matched turning biases (e.g., two strongly right-biased parents). For all crosses, the phenotypic mean and variance of the distribution of the F_1 generation was statistically indistinguishable from the distribution of the parental line (Fig. S2). Handedness therefore provides an ideal framework to study the genetics of variability because genetic effects on variability are not confounded by mean effects.

Next, using parametric (ANOVA) and nonparametric (bootstrapping) statistical approaches, we compared levels of intragenotypic variability across lines and found highly significant among-line differences in variability, implying that the abundance of individuals that were either strongly left- or right-biased was itself variable among lines. This observation indicates that the degree of intragenotypic variability itself is under genetic control in these lines (Fig. 1*C* and Table S1). To obtain further evidence that intragenotypic variability is heritable, we mated two high-variance and two low-variance lines to each other and measured turning bias in the resulting progeny (phenotyping an average of 183 individuals per cross). Intercrosses between high-variance lines led to high variance F_1 progeny and crosses with low-variance lines yielded low variance F_1 progeny (Fig. 2 *A* and *B*). In both cases, the variability in the F_1 progenies was statistically indistinguishable from that of the parents.

It is conceivable that some lines might be better than others at buffering microenvironmental perturbations, in which case the degree of intragenotypic variability among lines would be correlated across traits. To test this possibility, we scored additional phenotypes from our Y-maze data, namely, the total number of turns (a measure of overall activity); the left-right mutual information between successive turns; and the regularity of turn timing. We also analyzed other phenotypes previously measured on the DGRP at the individual level [starvation resistance (19), chill coma recovery (19), startle response (19), and night sleep (20)]. We found significant genetic variation for variability in all these phenotypes, confirming that genetic control of variability is ubiquitous across phenotypes. On the other hand, we found no evidence that the variances of these traits are correlated across

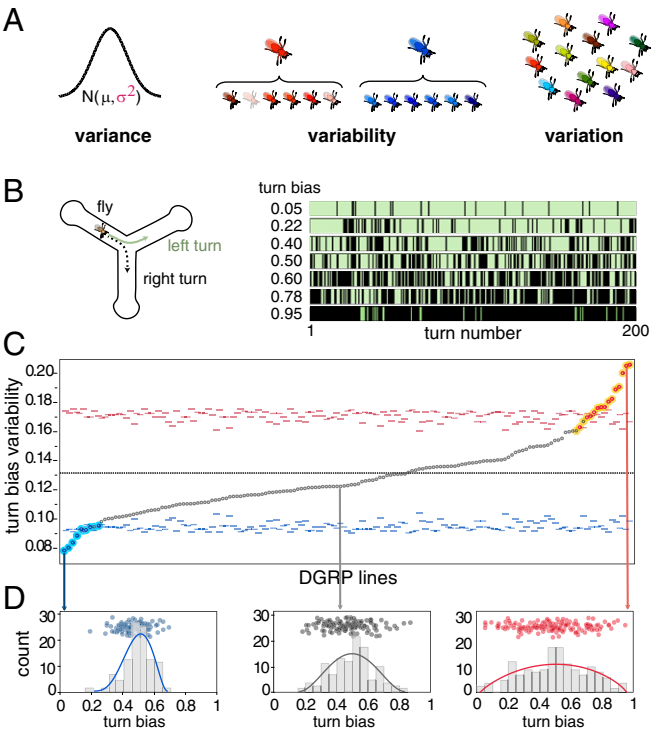

Fig. 1. Intragenotypic variability of locomotor handedness varies across DGRP lines. (*A*) The similarity between concepts of variance, variation, and variability may lead to some confusion. Variance is used to describe the standard statistical dispersion parameter (σ^2) or estimates of it derived from observations (s^2). Variability refers to the potential of an organism or genotype to vary phenotypically, phenotypic differences we could observe across clones of the same genotype (i.e., red fly = high variability genotype, blue fly = low variability genotype). Variation refers to the realized (observable) differences between individuals or genotypes. (*B*) Diagram of the Y-maze used to quantify individual locomotor behavior. Plot at right illustrates 200 sequential turns for seven representative individual flies. A turn bias of 0.05 indicates that this particular fly turned right 5% of the time (black stripes indicate right turns and green stripes left turns). (*C*) Sorted distribution of the SDs of within-line individual turn bias for 159 DGRP lines. Red and blue filled dots are significant, exceeding their corresponding tick-marked 99.9% Cis, estimated by permutation. See Table S1 for experimental sample sizes. Cyan and yellow highlighted dots are significant at $P < 0.001$ based on nonparametric bootstrap. (*D*) Distributions of turning bias across individuals for three representative DGRP lines with low, intermediate, and high intragenotypic variability. Each dot represents the turning bias of a single fly within that line. Lines are β distribution fits, chosen because they model overdispersed binomial distributions.

phenotypes [with the sole exception of mean absolute deviations (MADs) of turn bias and switchness; Fig. S3]. This result suggests that the genetic basis for intragenotypic variability is trait specific (and implicates many independent loci controlling these often-ignored traits).

The DGRP lines have been fully sequenced (19), allowing for genome-wide association mapping using the variability (i.e., MAD) of turning bias as a trait. Although the DGRP is underpowered to study the architecture of complex traits due to the relatively small number of lines ($n = 159$ in our study), it is a good resource to identify candidate genes for experimental follow-up (21, 22). To that end, we performed an association study using a series of locus-specific mixed linear models (accounting for relatedness between lines and experimental block effect) and found 36 polymorphisms in 22 genes associated with variability in turning bias using a nominal P value (19, 21) of 5×10^{-6} (Fig. S4 and Table S2). These genes are enriched with high significance for expression in the CNS both in adults and in larvae [adult CNS enrichment in

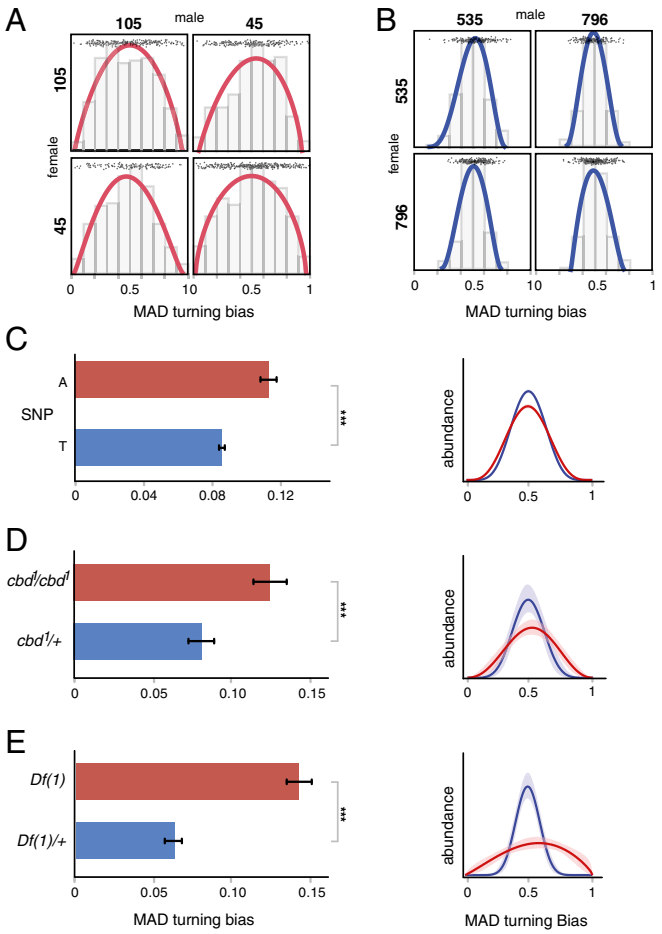

Fig. 2. Intragenotypic variability for turning bias is heritable. Effect of a Ten-a mutation on intragenotypic variability. (*A*) Distribution of F_1 turn biases resulting from high variance line 105 reciprocally crossed to high variance line 45 (Brown-Forsythe, $P = 0.08$; $n_{105 \times 105} = 235$; $n_{45 \times 45;} = 315$; $n_{105 \times 45} = 223$; $n_{45 \times 105} = 135$). (*B*) Distribution of F_1 turn biases resulting from low variance line 535 reciprocally crossed to low line variance line 796 (Brown-Forsythe, $P = 0.02$; $n_{535 \times 535} = 197$ $n_{796 \times 796} = 265$; $n_{796 \times 535} = 160$; $n_{535 \times 796} = 234$). In both panels, the progeny are presented on the off diagonal. Lines are β distribution fits. Points are individual flies. For both *A* and *B*, *P* values comparing F_1 to parents ranged from 0.14 to 0.99, uncorrected for multiple comparisons. (*C*) Intragenotypic variability (MAD) in turn bias of flies harboring alternative alleles of the Ten-a SNP identified in our GWAS ($n = 159$; GWAS, $P < 3 \times 10^{-6}$; phenotypic variance explained by this polymorphism: $R^2 = 19.5\%$). (*D*) Turn bias MAD of a homozygous *Ten-a* null allele (*cbd*^{1}; red) and heterozygous control (blue). *bk* indicates the *Ten-a*$^+$ genetic background Berlin-K. $n_{cbd1/bk} = 59$, $n_{cbd1/cbd1} = 99$; Brown-Forsythe, $P = 0.0074$; bootstrapping, $P < 0.001$. (*E*) Turn bias MAD of a line bearing a homozygous deficiency overlapping *Ten-a* (red) and heterozygous control (blue). $n_{Df(1)-bk} = 100$, $n_{Df(1)Ten-a} = 97$; Brown-Forsythe, $P = 1.5^{-11}$; bootstrapping, $P < 0.001$. ***$P < 0.001$. Right plots in all panels are corresponding β distribution fits of the distribution of turn bias scores within each experimental group. Shaded regions are 95% CIs on the β fits, estimated by bootstrap resampling; CIs in *A* are small compared with line thickness. Error bars are ±SE estimated by bootstrap resampling.

adult: Fisher exact test, $P < 0.001$; in larvae: Fisher exact test, $P < 0.01$; data from FlyAtlas (23)]. Among these, the synaptic target recognition gene *Tenascin accessory* [*Ten-a*; genome-wide association study (GWAS), $P < 3 \times 10^{-6}$; Fig. 2C] caught our attention. Ten-a is a transmembrane signaling protein involved in synapse formation (24, 25), typically expressed presynaptically. In the antennal lobe, Ten-a supports an expression-level matching code with high-expression neurons partnering with other high expression neurons (and low with low) (24). Teneurin impairment causes profound neuromuscular junction disruption (25). Within the central brain, Ten-a mutation causes midline fusion defects within the central complex), a brain structure implicated in sensory integration and locomotion (26). *Ten-a* is highly conserved from insects to mammals (27). To validate the role of *Ten-a* in modulating variability in turning bias, we used a null allele (*Ten-a*$^{cbd\text{-}KS96}$; Fig. 2*D*), a deficiency overlapping *Ten-a* (*Df(1)Ten-a* (24) (Fig. 2*E*), and expression knockdown using inducible RNAi (*tub-Gal4;tub-Gal80*^{ts} > UAS-TRiP.JF03375; Fig. 3). In all cases, disrupting *Ten-a* increased the variability in turning bias with no effect on the mean. The effect of RNAi knockdown suggests a quantitative relationship between *Ten-a* mean expression and variance in turning bias.

The bias in handedness of a given fly is a fixed property of that individual (e.g., a young adult with a strong left bias will display this bias throughout its life) (18). The persistence of this bias suggests that handedness may be wired during development. To determine whether there is a critical developmental period when *Ten-a* expression is required to regulate variability, we used temperature-inducible RNAi to knock down *Ten-a* in sliding 3-d windows (Fig. 3 *A* and *B*). We found that knocking down *Ten-a* expression in midpupae increases the resulting adults' variability (Fig. 3*D*). This stage coincides with a spike in *Ten-a* expression (Fig. 3*C*) and the formation of the central complex (28, 29).

Discussion

In this study, we used a simple behavioral trait to show that individual genotypes vary considerably in their degree of intragenotypic variability (7, 15) and found that this variation is heritable. Similar to work on fluctuating asymmetry (30), such experiments allow us to estimate how robust development is to microenvironmental perturbation and highlight the consequences of this variation for an individual's phenotype. Our use of inbred lines enables the estimation of a parameter (intragenotypic variability) that otherwise could not be observed and uncovers the spectrum of phenotypes a given genotype can produce in a given environment. Furthermore, using association mapping uncovered a gene, ten-a, which implicates the central complex of the brain.

In a companion study, Buchanan et al. (18) mapped a set of neurons within the central complex (i.e., protocerebral bridge columnar neurons) that regulates the magnitude of left-right turn bias and therefore the magnitude of intragenotypic variability. Together these studies constitute a rare example linking natural genetic variation for a complex behavioral trait, to mutants implicating a brain region, to a specific subcircuit within this region. Thus, we can begin to paint the path from genetic variation to behavioral individuality.

One of the great challenges in modern biology is to understand the functional consequences of genomic variation and to determine how and when it contributes to phenotypic differences among individuals. During the last decade, we have made remarkable progress in understanding the genetic basis of complex traits and diseases, thanks in part to the application of GWAS to large cohorts. Unfortunately, we have fallen short of the goal of explaining heritability for complex traits in terms of allelic effects (31, 32). The traditional framework used to map QTLs focuses on the average effect of alternative alleles averaged in a population. However, as we have shown in this study, when phenotypic variation results from alleles that modify phenotypic variance rather than the mean, this link between genotype and phenotype will be not be detected. The case of locomotor handedness is an extreme example, where there is virtually no heritability for mean handedness and all of the phenotypic variation in this population is attributable to intragenotypic variability. Nevertheless, it highlights the important contribution genetic control of variability can play in our understanding of the cause of phenotypic variation.

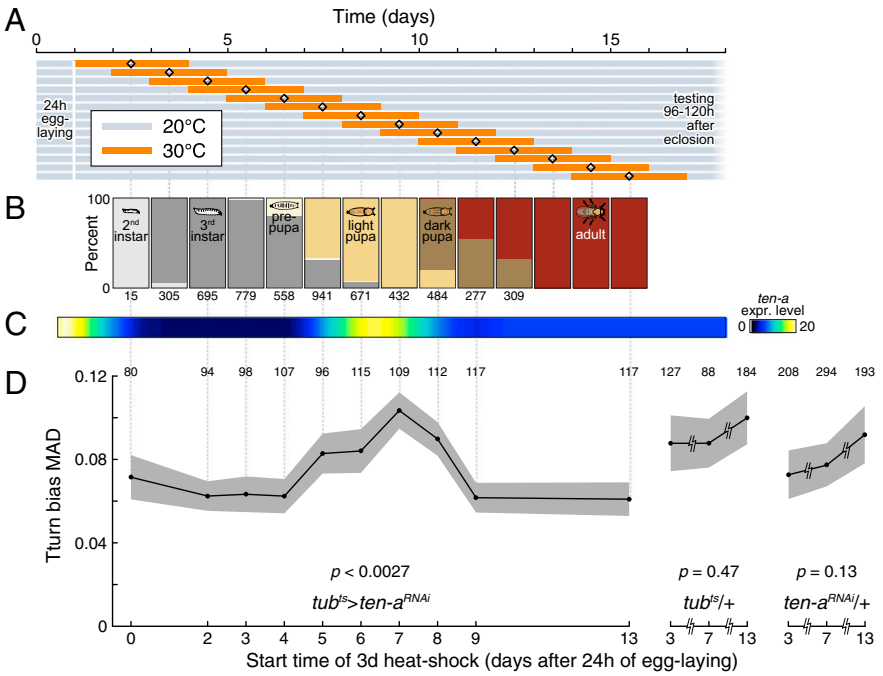

Fig. 3. Disruption of *Ten-a* expression in midpupa affects behavioral variance. (*A*) Time courses of sliding window *Ten-a* RNAi induction. Flies laid eggs for 24 h prior the start of the experiment and were reared at 20 °C (gray) until 3 d of RNAi induction at 30 °C (orange). Flies were then returned to 20 °C until they were tested 3–5 d after eclosion. (*B*) Fraction of flies at any developmental stage during the course of the experiment. Numbers indicate sample sizes. (*C*) *Ten-a* expression level over development. Expression level derived from modENCODE. (*D*) Effect of temperature-inducible *Ten-a* RNAi on the variability of turning bias over development. Knockdown effect varied significantly with the timing of the induction window ($P = 0.0027$) estimated by a bootstrapping omnibus test (*SI Text*), with a knockdown starting on day 7 greatly increasing variability. This knockdown window coincides with the peak of *Ten-a* expression during pupation. Gray regions represent ±SE, estimated by bootstrapping. To the right, the controls, tub^{ts}/+ and $Ten\text{-}a^{RNAi}$/+, measured after 3-d, 30 °C windows starting on days 3, 7, and 13, show no effect ($P < 0.47$ and $P < 0.13$, respectively). Numbers above data indicate sample sizes. Vertical guide lines associate data points across panels.

If, in a common macroenvironment, different genetic backgrounds vary in their propensity for phenotypic variability, individuals drawn from a high-variability genetic background have the potential to explore a wider range of phenotypic space than those drawn from a low-variability background (sometimes far beyond what may be determined by the mean effect alone). We observe intragenotypic variability for every phenotype we investigate, ranging from behavioral to metabolic, indicating that variability is ubiquitous. Maintaining variability could be advantageous in the context of evolutionary adaptation, but in human genetics, it could be deleterious when an extreme phenotype enhances disease risk. The implications for medical genetics are far-reaching (14, 31, 32), specifically for attempts to predict phenotypes from genotypes. This point is illustrated in Fig. 4: if we consider that each individual is a random draw from a distribution determined (in part) by its genotype, then we should not think of genotypes as determining the phenotypic value of that individual; rather, we should think of genotypes as determining the probability of that individual having a particular phenotypic value (33). This model requires the development of additional experimental and statistical approaches for mapping QTLs, and several are already being developed (4, 12, 14, 33–35).

Our work does not address the adaptive significance of intragenotypic variability or the evolutionary forces maintaining variation at alleles affecting variability. Addressing such questions would, for example, require additional information on the fitness consequences of variability in handedness. Although the allele frequencies of the most significant SNPs are relatively high in the DGRP, we did not detect any significant deviation from neutrality for the genes harboring these SNPs. It should also be emphasized that differences in variability across lines could emerge from a neutral process. The nature of the forces influencing the evolution of alleles that affect variability has, more generally, been the focus of a rich theoretical literature explaining this phenomenon from a game theory perspective (36) or in terms of bet-hedging (37). Under various scenarios, increased phenotypic variability may allow some individuals in a population to explore a broader range of phenotypic space, thus maintaining this population at, or close to, some fitness optimum over time. This scenario should be particularly favored in fluctuating environments (38, 39). We point out, however, that it is still an open question whether genetic mechanisms leading to variation in intragenotypic variability are also associated with those underlying phenotypic plasticity (i.e., genotype-by-environment interactions in response to macroenvironmental variation, described, for example, through reaction norms).

Although inbred lines are an ideal way to study the genetic basis of intragenotypic variability (because variation between individuals within a line is caused primarily by microenvironmental effects), not all systems are amenable to this approach. In many circumstances, alternate designs are available. For example, if the phenotype of interest is molecular, recent progress in single-cell technology now makes it possible to measure cell to cell variation within an individual genotype (14), enabling the study of intragenotypic variability in natural populations, including humans (33). At the organismal level, humans also have the experimental confound of being outbred. Approaches in this case range from the use of twin studies to family-based analyses (34). In systems where controlled crosses can be carried out, a wider range of options is possible (2, 12, 13, 40). These approaches have been particularly effective in breeding programs, where intragenotypic variability is not desirable (2). In fact, the idea that there may be genetic variation underlying phenotypic variability dates back to the 1950s (41–43), but the actual estimates

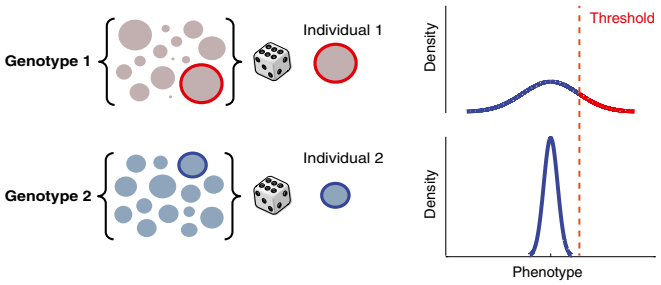

Fig. 4. Consequences of intragenotypic variability on the fraction of a hypothetical population exceeding a disease threshold. Visual representation of the effects of variance on the prevalence of phenotypes exceeding a threshold, such as a disease state. Genotypes 1 and 2 differ in their degree of intragenotypic variability. The sets of circles at the left represent the range of possible outcomes for each genotype. Generally, each individual in an outbred diploid organism is a unique instance of its genotype. By contrast, our experiments with inbred lines allow us to consider multiple individuals from the same distribution. An individual drawn at random from genotype 1 (high variability) may land in the tail of the distribution, potentially in disease space. On the other hand, an individual drawn randomly from genotype 2 never gets a chance to explore the phenotypic space explored by genotype 1, even if it is just as much of an outlier within its respective distribution.

of the heritability of this component are more recent and primarily derived from outbred organisms, using family-based analyses in agricultural species ranging from rabbits to dairy cows (reviewed in ref. 2). By using a model organism to study the mechanisms underlying variability, our study adds to a growing body of literature recognizing the importance of variance control in complex trait genetics.

Materials and Methods

***Drosophila* Stocks.** The DGRP consists of a collection of isofemale lines derived from a single field collection from the Raleigh, NC, farmers' market, followed by 20 generations of full-sib mating that rendered most loci homozygous within lines (expected $F = 0.986$) (19). The DGRP lines are available from the *Drosophila* Stock Center (flystocks.bio.indiana.edu). Stocks used for *Ten-a* validation were Berlin-K, *central-body-defect*^{KS96}, *Df1-Ten-a*, and RNAi TRiP. JF03375 (*tub-Gal4;tub-Gal80*^{ts} > UAS-TRiP.JF03375). All flies were reared on standard fly media (Scientiis and Harvard University BioLabs Fly Food Facility), in a single 25 °C incubator at 30–40% relative humidity with a 12/12-h light/dark cycle. Before each assay, flies were fully randomized across blocks, lines, Y-maze arrays, and position on the array. At least three strains were assayed simultaneously on each array.

Phenotypic Assay. Each experiment examines one array of 120 Y-mazes (refered to as maze-array). Mazes were illuminated from below with white LEDs and imaged with 2MP digital cameras, and the X-Y positions of each fly centroid were automatically tracked and recorded with custom written software. Further details about the assay are provided in ref. 18; the code is available at lab.debivort.org/neuronal-control-of-locomotor-handedness/. We estimated the degree of variability of each line using the MAD (4, 13). It is defined as the median of the absolute deviation from each observation's median: MAD = median $[|X_i - \text{median}(X_i)|]$, where X_i is the phenotypic score of an individual fly within a line. MAD scores were computed for each line for each phenotype. Only females were used in this experiment, and only lines yielding data from a minimum of 75 individuals were included. Fly behavior in the mazes was monitored for 2 h. This assay generated four phenotypes: (*i*) the handedness or left/right turning bias in the arms of the maze summed over all left/right decisions; (*ii*) the number of turns over the 2 h period, an estimate of overall locomotor activity; (*iii*) the "switchiness" of the right/left turn sequence, which is related to the mutual information between successive turns (e.g., LLLLLRRRRR: low switchiness, high mutual information; LLRLLRRRLR: moderate switchiness, low mutual information; LRLRLRLRLR: high switchiness, high mutual information) defined as $(N_{<L,R>} + N_{<R,L>})/(2N_RN_L/N)$, where $N_{<L,R>}$ is the number of left turns followed by right turns, $N_{<R,L>}$ is the number of right turns followed by left turns N_R is the number of right turns, N_L is the number of left turns, and N is the total number of turns; and (*iv*) the regularity of turn timing: a fly with a high score makes turns uniformly throughout the experiment, whereas a low score would characterize a fly making a small number of dense streaks of turns but is inactive for dozens of minutes at a time. It is defined as MAD(ITIs)/(7,200/N), where ITIs is the vector of interturn intervals in seconds.

Quantitative Genetic Analysis.

Analysis of means. To determine whether there was genetic variation segregating in the DGRP affecting the mean turning bias, we partitioned the variance for line means using the following ANOVA model: $Y = \mu + L_{random} + B_{random} + L \times B_{random} + A + X + A \times X + e$, where Y is turning bias score of each fly, L is the effect of line treated as random, B is the effect of block treated as random, X is the box effect, A is the maze-array effect, and e is the error variance (Table S1). ANOVA was implemented using PROC MIXED in SAS 9.3.

Variance heterogeneity. We used several statistical approaches to estimate heterogeneity of variance for turning bias between lines (Table S1): (*i*) the Brown–Forsythe test, which is based on a one-way ANOVA and relies on the absolute deviation from the median (4, 44); (*ii*) nonparametric bootstrapping in which we first pooled all of the turn bias scores for all individual flies across lines and then resampled each line experimental group from this pool, matching the sample size (lines in which the MAD of the resampled group was closer to the MAD of the pooled data, in fewer than 10 of 10,000 resamples, were taken as significant; this tests the null hypothesis that each group is drawn from an identical distribution of observations, using MAD as a test statistic); (*iii*) a nonparametric version of the analysis of mean for variances (ANOMV) (45) [this approach compares the group means of the MAD to the overall mean MAD under the null hypothesis that the group MAD means equal each line specific MAD (results in Table S1), implemented in SAS 9.3]; and finally (*iv*) we used the same ANOVA model described above for the analysis of means but used the absolute deviation from the median (4, 5) as a measure for each fly as the dependent variable. This test was implemented using PROC MIXED in SAS 9.3.

Phenotypic correlation between traits. We assessed four traits as measured in our study and four additional traits gathered from the literature (SD for starvation, startle response, chill coma recovery, coefficient of environmental variation for night sleep). Data are from refs. 19 and 20. Phenotypic correlation between each trait pair was computed as the Pearson product-moment correlation (implemented using PROC GLM in SAS 9.3). P values were not corrected for multiple comparison.

High and Low Variance Lines Intercrosses. To confirm that variability was heritable, we crossed high variability lines 45 and 105 together and low variability lines 796 and 535 together. We assessed statistical significance between parental lines and their progeny using the Brown–Forsythe test and a bootstrapping two-tailed z-test (with n = 10,000 resamples). We resampled the turn bias of the parents and for each iteration calculated the MAD of turning bias and then compared the MAD for the F_1 progeny to their parents.

Genome-Wide Association Mapping. GWAS was performed using the code and approach described in ref. 19 (dgrp2.gnets.ncsu.edu). We fitted a series of loci-specific mixed linear models using the following model: $Y = \mu + Sb + Iu + e$, where Y is the MAD of turning bias of each DGRP lines, S is the design matrix for the fixed SNP effect b, I is the incidence matrix for the random polygenic effect u, and e is the residual (19). A total of 1,931,250 SNPs and indels were used in our analyses, with the minor alleles present in at least seven DGRP lines, using only biallelic sites. For each tissue, we used FlyAtlas AffyCalls (23) to determine which genes were expressed in which tissue. To determine significance, we used a Fisher's exact test comparing the expected number of gene expressed in each tissue across the entire genome to the observed number of genes expressed in each tissue in our gene list.

Validation of *Ten-a* Effect on Variability.

Ten-a null and deficency. The turning bias and MAD of turning bias of homozygotes of both the null allele *Ten-a*$^{cbd\text{-}KS96}$ (28) and deficiency overlapping *Ten-a Df(1)Ten-a* (29) were compared with heterozygous animals over their genetic background, Berlin-K.

Time course knockdown of Ten-a RNAi. Ten adult *Ptub-Gal80ts;Ptub-Gal4/Sb* females were crossed to three *UAS-Ten-a RNAi y1,v1;P(TRiP.JF03375)attP2* males for RNAi induction. Flies were allowed to mate for 24 h at 20 °C, at which point the parents were removed, and the bottles containing F_1 eggs were returned to 20 °C until the beginning of their heat shock window. Flies were exposed for 72 h to 30 °C temperature in a sliding window each day over 14 windows (Fig. 3*A*). All flies assayed were between 3 and 5 d after

eclosion. In parallel, each day, developing flies of the same genotype were examined and counted to determine the fraction of flies in each developmental stage at the time of RNAi induction (Fig. 3*B*). Controls were performed using *Ptub-Gal80ts;Ptub-Gal4/Sb* females crossed to Canton-S males and Canton-S females crossed to *UAS-Ten-a RNAi y1,v1;P(TRiP.JF03375)attP2* males (Fig. 3*D*); otherwise, they were treated identically.

Ten-a expression. Data for *Ten*-a expression over developmental time (Fig. 3*C*) were downloaded from FlyBase and derived from ModEncode (modENCODE DDC ids: modENCODE_4433, _4435 and _4439 through _4462). These data reflect animals synchronized by developmental stage to within 2 h. To make these data comparable to our experimental groups, in which egg laying occurred over 24 h, we corresponded the developmental stages of the FlyBase data to our developmental stage time course (Fig. 3*B*), linearly interpolated the expression values, and applied a 24-h sliding window average to the interpolated data, mimicking the dispersion effects of our longer egg collection window.

ACKNOWLEDGMENTS. We thank Ian Dworkin, Mia Levine, Noah Zaitlen, and Eric Stone for comments, discussion, and helpful feedback on the manuscript. This work was supported by National Institutes of Health Grant R01 AI064950 (to A.G.C.), a Harvard Society of Fellows fellowship and Harvard Milton funds (to J.F.A.), and the Rowland Junior fellowship (to B.L.d.B.).

1. Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits* (Sinauer Associates, Sunderland, MA), 1st Ed.
2. Hill WG, Mulder HA (2010) Genetic analysis of environmental variation. *Genet Res* 92(5-6):381–395.
3. Yang J, et al. (2012) FTO genotype is associated with phenotypic variability of body mass index. *Nature* 490(7419):267–272.
4. Shen X, Pettersson M, Rönnegård L, Carlborg Ö (2012) Inheritance beyond plain heritability: Variance-controlling genes in *Arabidopsis thaliana*. *PLoS Genet* 8(8): e1002839.
5. Dworkin I (2005) A study of canalization and developmental stability in the sternopleural bristle system of *Drosophila melanogaster*. *Evolution* 59(7):1500–1509.
6. Mackay TF, Lyman RF (2005) *Drosophila* bristles and the nature of quantitative genetic variation. *Philos Trans R Soc Lond B Biol Sci* 360(1459):1513–1527.
7. Willmore KE, Young NM, Richtsmeier JT (2007) Phenotypic variability: Its components, measurement and underlying developmental processes. *Evol Biol* 34(3-4):99–120.
8. SanCristobal-Gaudy M, Bodin L, Elsen JM, Chevalet C (2001) Genetic components of litter size variability in sheep. *Genet Sel Evol* 33(3):249–271.
9. Sorensen D, Waagepetersen R (2003) Normal linear models with genetically structured residual variance heterogeneity: A case study. *Genet Res* 82(3):207–222.
10. Stamps JA, Groothuis TG (2010) Developmental perspectives on personality: Implications for ecological and evolutionary studies of individual differences. *Philos Trans R Soc Lond B Biol Sci* 365(1560):4029–4041.
11. Stamps JA, Saltz JB, Krishnan VV (2013) Genotypic differences in behavioural entropy: Unpredictable genotypes are composed of unpredictable individuals. *Anim Behav* 86(3):641–649.
12. Rönnegård L, Valdar W (2011) Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics* 188(2):435–447.
13. Rönnegård L, Valdar W (2012) Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genet* 13:63.
14. Geiler-Samerotte KA, et al. (2013) The details in the distributions: Why and how to study phenotypic variability. *Curr Opin Biotechnol* 24(4):752–759.
15. Queitsch C, Sangster TA, Lindquist S (2002) Hsp90 as a capacitor of phenotypic variation. *Nature* 417(6889):618–624.
16. Gibson G, Wagner G (2000) Canalization in evolutionary genetics: A stabilizing theory? *BioEssays* 22(4):372–380.
17. Kain JS, Stokes C, de Bivort BL (2012) Phototactic personality in fruit flies and its suppression by serotonin and white. *Proc Natl Acad Sci USA* 109(48):19834–19839.
18. Buchanan SM, Kain JS, de Bivort BL (2015) Neuronal control of locomotor handedness in *Drosophila*. *Proc Natl Acad Sci USA*, 10.1073/pnas.1500804112.
19. Huang W, et al. (2014) Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res* 24(7):1193–1208.
20. Harbison ST, McCoy LJ, Mackay TF (2013) Genome-wide association study of sleep in *Drosophila melanogaster*. *BMC Genomics* 14:281.
21. Mackay TFC, et al. (2012) The Drosophila melanogaster Genetic Reference Panel. *Nature* 482(7384):173–178.
22. Turner TL, Miller PM (2012) Investigating natural variation in *Drosophila* courtship song by the evolve and resequence approach. *Genetics* 191(2):633–642.
23. Robinson SW, Herzyk P, Dow JAT, Leader DP (2013) FlyAtlas: Database of gene expression in the tissues of *Drosophila melanogaster*. *Nucleic Acids Res* 41(Database issue):D744–D750.
24. Hong W, Mosca TJ, Luo L (2012) Teneurins instruct synaptic partner matching in an olfactory map. *Nature* 484(7393):201–207.
25. Mosca TJ, Hong W, Dani VS, Favaloro V, Luo L (2012) Trans-synaptic Teneurin signalling in neuromuscular synapse organization and target choice. *Nature* 484(7393): 237–241.
26. Strauss R (2002) The central complex and the genetic dissection of locomotor behaviour. *Curr Opin Neurobiol* 12(6):633–638.
27. Tucker RP, Beckmann J, Leachman NT, Schöler J, Chiquet-Ehrismann R (2012) Phylogenetic analysis of the teneurins: Conserved features and premetazoan ancestry. *Mol Biol Evol* 29(3):1019–1029.
28. Cheng X, et al. (2013) Ten-a affects the fusion of central complex primordia in *Drosophila*. *PLoS ONE* 8(2):e57129.
29. Young JM, Armstrong JD (2010) Building the central complex in *Drosophila*: The generation and development of distinct neural subsets. *J Comp Neurol* 518(9): 1525–1541.
30. Carter AJ, Houle D (2011) Artificial selection reveals heritable variation for developmental instability. *Evolution* 65(12):3558–3564.
31. Lyon GJ, O'Rawe J (2015) Human genetics and clinical aspects of neurodevelopmental disorders. *The Genetics of Neurodevelopmental Disorders*, ed Mitchell K (Wiley, Hoboken, NJ).
32. Gibson G (2010) Hints of hidden heritability in GWAS. *Nat Genet* 42(7):558–560.
33. Yvert G (2014) 'Particle genetics': Treating every cell as unique. *Trends Genet* 30(2): 49–56.
34. Conley D, Rauscher E, Siegal ML (2013) Beyond orchids and dandelions: Testing the 5-HTT "risky" allele for evidence of phenotypic capacitance and frequency-dependent selection. *Biodemography Soc Biol* 59(1):37–56.
35. Rönnegård L, Felleki M, Fikse F, Mulder HA, Strandberg E (2010) Genetic heterogeneity of residual variance: Estimation of variance components using double hierarchical generalized linear models. *Genet Sel Evol* 42:8.
36. Smith JM, Price GR (1973) The logic of animal conflict. *Nature* 246:15–18.
37. Seger J, Brockmann HJ (1987) What is bet-hedging? *Oxf Surv Evol Biol* 4:182–211.
38. Hoffmann AA, Merilä J (1999) Heritable variation and evolution under favourable and unfavourable conditions. *Trends Ecol Evol* 14(3):96–101.
39. Zhang XS (2005) Evolution and maintenance of the environmental component of the phenotypic variance: Benefit of plastic traits under changing environments. *Am Nat* 166(5):569–580.
40. Mulder HA, Rönnegård L, Fikse WF, Veerkamp RF, Strandberg E (2013) Estimation of genetic variance for macro- and micro-environmental sensitivity using double hierarchical generalized linear models. *Genet Sel Evol* 45(1):23.
41. Hogben LT (1933) *Nature and Murture* (W.W. Norton Company, New York).
42. Grüneberg H (1954) Variation within inbred strains of mice. *Nature* 173(4406): 674–676.
43. Reeve ECR, Robertson FW (1953) Analysis of environmental variability in quantitative inheritance. *Nature* 171(4359):874–875.
44. Brown MB, Forsythe AB (1974) Robust tests for the equality of variances. *J Am Stat Assoc* 69(346):364–367.
45. Nelson PR, Wludyka PS, Copeland K (2005) *The Analysis of Means: A Graphical Method for Comparing Means, Rates, and Proportions* (SIAM, Philadelphia).

Supporting Information

Ayroles et al. 10.1073/pnas.1503830112

SI Text

All raw data and analysis scripts are available at lab.debivort.org/genetic-control-of-phenotypic-variability.

Drosophila Stocks. The DGRP was created as a community resource for the genetic mapping of complex traits (1). It consists of a collection of isofemale lines derived from a single field collection from the Raleigh, NC, farmers market, followed by 20 generations of full-sib mating that rendered most loci homozygous within lines (expected $F = 0.986$) (2). As a result, the genetic variation that was present between individual flies in the natural population is now captured between lines in the panel. This property of wild-derived inbred lines allows us to measure any phenotype on a given genotypic background and phenotype the same genotype a large number of times in any environment. Completion of the genome sequencing for all lines combined with *Drosophila*'s generally rapid decay in linkage disequilibrium between polymorphic sites makes the DGRP a powerful tool to identify genetic polymorphisms that affect quantitative phenotypes (3). The DGRP lines are available from the *Drosophila* Stock Center (flystocks.bio.indiana.edu). We used a total of 159 lines in this study (lines with the highest inbreeding coefficient); a list is provided along with data at lab.debivort.org/genetic-control-of-phenotypic-variability. Stocks used for *Ten-a* validation were Berlin-K, *central-body-defect*^{KS96} (4), *Df1-Ten-a* (5), and RNAi TRiP.JF03375 (6). All flies were reared on standard fly media (Scientiis and Harvard University BioLabs Fly Food Facility), in a single 25 °C incubator at 30–40% relative humidity with a 12/12-h light/dark cycle. Before each assay, flies were fully randomized across blocks, lines, Y-maze arrays, and position on the array. At least three strains were assayed simultaneous on each array.

Variance, Variation, Variability. The similarity between concepts of variance, variation, and variability may lead to some confusion. The meanings of these terms are reviewed in Wagner and Altenberg (7). In accordance with their definition, we used the term variance (8, 9) to describe the standard statistical dispersion parameter (σ^2) or estimates of it derived from observations (s^2). Variability refers to the potential of an organism or genotype to vary phenotypically. Variation refers to the realized (observable) differences between individuals or genotypes.

Phenotypic Assay. Studying variance as a trait poses a number of challenges including the large sample size required (precise estimates of variance requires a larger number of observations than needed to estimate means), the experimental design (as to not confound sources of error), and potential measurement error of the phenotype itself (9). It is with these considerations in mind that we developed a high-throughput assay aimed at monitoring the behavior of individual flies placed into individual Y-mazes (10) (Fig. 1*A*). Each experiment examines one array of 120 Y-mazes (refered to as maze-arrray). Mazes were illuminated from below with white LEDs (5500K; LuminousFilm), imaged with 2MP digital cameras (Logitech), and the X-Y positions of each fly's centroids were automatically tracked and recorded with software custom written in LabView (National Instruments). Further details about the assay are provided in ref. 10; the code is available at lab.debivort.org/neuronal-control-of-locomotor-handedness/. Although various statistics can be computed to estimate the degree of variability of a distribution, in this study, we use one the most robust metrics, the MAD (11, 12). It is defined as the median of the absolute deviation from each observation's median: MAD = median $[|X_i - \text{median}(X_i)|]$, where X_i is the phenotypic score of an individual fly within a line. MAD scores were computed for each line for each phenotype. Only females were used in this experiment, and only lines yielding data from a minimum of 75 individuals were included. Before each assay, flies were very lightly anesthetized, rapidly transferred to an individual Y-maze, and given a recovery period of 20 min before the start of the assay. Fly behavior in the mazes was monitored for 2 h. This assay generated four phenotypes. (*i*) The handedness or left/right turning bias in the arms of the maze summed over all left/right decisions. A turning bias score of 0.8 for a given fly would indicate that this individual made left turns 80% of the time at the maze's junction over the 2-h period. This simple phenotype is particularly well suited for this study given that it is measured without error, and the high number of turns for any given fly ensures a robust estimate of the turning bias and it variance for each fly. (*ii*) The number of turns over the 2-h period, an estimate of overall locomotor activity. (*iii*) The switchiness or the mutual left-right information between successive turns right/left turn sequence (e.g., LLLLLRRRRR: low switchiness, high mutual information; LLRLLRRRLR: moderate switchiness, low mutual information; LRLRLRLRLR: high switchiness, high mutual information) defined as $(N_{<L,R>}+N_{<R,L>})/(2N_RN_L/N)$, where $N_{<L,R>}$ is the number of left turns followed by right turns, $N_{<R,L>}$ is the number of right turns followed by left turns, N_R is the number of right turns, N_L is the number of left turns, and N is the total number of turns. (*iv*) The regularity of turn timing: a fly with a high score makes turns uniformly throughout the experiment, whereas a low score would characterize a fly making a small number of dense streaks of turns but is inactive for dozens of minutes at a time. It is defined as MAD(ITIs)/(7,200/N), where ITIs is the vector of interturn intervals in seconds. The left/right turning bias is the main focus of this study; additional traits were measured to illustrate that the degree of variability across traits is not correlated between lines.

Quantitative Genetic Analysis.

Analysis of means. To determine whether there was genetic variation segregating in the DGRP affecting the mean turning bias, we partitioned the variance for line means using the ANOVA model $Y = \mu + L_{random} + B_{random} + L \times B_{random} + A + X + A \times X + e$, where Y is turning bias score of each fly; L is the effect of line treated as random, B is the effect of block treated as random, X is the box effect, A is the maze-array effect, and e is the error variance (Table S1). ANOVA was implemented using PROC MIXED in SAS 9.3 (13).

Variance heterogeneity. We used several statistical approaches to estimate heterogeneity of variance for turning bias between lines (Table S1). (*i*) The Brown–Forsythe test, which is based on a one-way ANOVA and relies the absolute deviation from the median (8). (*ii*) Nonparametric bootstrapping in which we first pooled all of the turn bias scores for all individual flies across lines and then resampled each line experimental group from this pool, matching the sample size. Lines in which the MAD of the resampled group was closer to the MAD of the pooled data in fewer than 10 of 10,000 resamples were taken as significant. This analysis tests the null hypothesis that each group is drawn from an identical distribution of observations, using MAD as a test statistic. (*iii*) A nonparametric version of the ANOMV (14, 15). This approach compares the group means of the MAD to the overall mean MAD under the null hypotheses that the group MAD means equals each line specific MAD (results in Table S1), implemented

in SAS 9.3 (13, 15). (*iv*) Finally, we used the same ANOVA model described above for the analysis of mean but used the absolute deviation from the median (11, 16) as a measure for each fly as the dependent variable, implemented using PROC MIXED in SAS 9.3 (13).

Phenotypic correlation between traits. We assessed four traits as measured in this study and four additional traits gathered from the literature (SD for starvation, startle response, chill coma recovery, coefficient of environmental variation for night sleep). Data are from refs. 2 and 17. The phenotypic correlation between traits was computed as the Pearson product-moment correlation (implemented using PROC GLM in SAS 9.3). *P* values are not corrected for multiple comparison (18).

High and Low Variance Lines Intercrosses. To confirm that variability was heritable, we crossed high variability lines 45 and 105 together and low variability lines 796 and 535 together. Ten females and five males were used for each cross. Flies were reared and phenotyped using the same protocoled described above. Note that parental behavior was remeasured concurrently with F_1 behavior following a corresponding self-cross (e.g., 45 × 45). We assessed statistical significance between parental lines and their progeny using the Brown–Forsythe test and a bootstrapping two-tailed *z*-test (with $n = 10{,}000$ resamples). We resampled the turn bias of the parents and for each iteration calculated the MAD of turning bias and then compared the MAD for the F_1 progeny to their parents.

Genome-Wide Association Mapping. GWAS was performed using the code and approach described in ref. 2 (dgrp2.gnets.ncsu.edu). In a first step, phenotypic stores were adjusted for the potential effect of *Wolbachia* and known large inversions segregating in this panel [namely: *In(2L)t*, *In(2R)NS*, *In(3R)P*, *In(3R)K*, and *In(3R)Mo*]; none of them were associated with variability turning bias. We then fitted a series of loci-specific mixed linear model using the model: $Y = \mu + Sb + Iu + e$, where Y is the MAD of turning bias of each DGRP lines, S is the design matrix for the fixed SNP effect b, I is the incidence matrix for the random polygenic effect u, and e is the residual (2). A total of 1,931,250 SNPs and indels were used in these analyses with the minor alleles present in at least seven DGRP lines, using only biallelic sites. Polymorphisms segregating within lines were discarded and for each SNP at least 60 DGRP lines had to have been genotyped to be analyzed. Given the number lines available in the DGRP, GWAS will generally be underpowered (19); however, our goal is not to describe the overall genetic architecture of each of these phenotypes but rather to identify interesting candidate genes that would provide some insight into the genetic basis of variance control. For this reason, we used a liberal threshold of $P < 10^{-6}$. The analysis for tissue enrichment was based on FlyAtlas data, which are publically available (20). For each tissue, we used FlyAtlas AffyCalls (21) to determine which genes were expressed in which tissue (using a conservative filter of four of four present calls). To determine significance, we used Fisher's exact test comparing the expected number of gene expressed in each tissue across the entire genome to the observed number of gene expressed in each tissue in our gene list.

Validation of *Ten-a* Effect on Variability.

Ten-a null and deficency. The turning bias and MAD turning bias of homozygotes of both the null allele *Ten-a*$^{cbd\text{-}KS96}$ (5) and deficiency overlapping *Ten-a Df(1)Ten-a* (10) were compared with heterozygous animals over their genetic background, Berlin-K.

Time course knockdown of Ten-a RNAi. Ten adult *Ptub-Gal80ts;Ptub-Gal4/Sb* females were crossed to three *UAS-Ten-a RNAi y1,v1; P(TRiP.JF03375)attP2* males for RNAi induction. Flies were allowed to mate for 24 h at 20 °C, at which point the parents were passaged out, and the bottles containing F_1 eggs were returned to 20 °C until the beginning of their heat shock window. Flies were exposed for 72 h to 30 °C temperature, in a sliding window each day over 14 windows (Fig. 3*A*). All flies assayed were between 3 and 5 d after eclosion. In parallel, each day, developing flies of the same genotype were examined and counted to determine the fraction of flies in each developmental stage at the time of RNAi induction (Fig. 3*B*). Stages containing larval animals were microwaved to melt the media and poured through a sieve, and larval carcasses were counted under a dissecting scope. Controls were performed using *Ptub-Gal80ts;Ptub-Gal4/Sb* females crossed to Canton-S males and Canton-S females crossed to *UAS-Ten-a RNAi y1,v1;P(TRiP.JF03375)attP2* males (Fig. 3*D*); otherwise, they were treated identically. Data for *Ten*-a expression over developmental time (Fig. 3*C*) were downloaded from FlyBase (22) and derived from ModEncode (23) (modENCODE DDC ids: modENCODE_4433, _4435 and _4439 through _4462). These data reflect animals synchronized by developmental stage to within 2 h. To make these data comparable to our experimental groups, in which egg laying occurred over 24 h, we corresponded the developmental stages of the FlyBase data to our developmental stage time course (Fig. 3*B*), linearly interpolated the expression values, and applied a 24-h sliding window average to the interpolated data, mimicking the dispersion effects of our longer egg collection window.

1. Yang J, et al. (2012) FTO genotype is associated with phenotypic variability of body mass index. *Nature* 490(7419):267–272.
2. Huang W, et al. (2014) Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res* 24(7):1193–1208.
3. Mackay TFC, et al. (2012) The Drosophila melanogaster Genetic Reference Panel. *Nature* 482(7384):173–178.
4. Strauss R (2002) The central complex and the genetic dissection of locomotor behaviour. *Curr Opin Neurobiol* 12(6):633–638.
5. Cheng X, et al. (2013) Ten-a affects the fusion of central complex primordia in *Drosophila*. *PLoS ONE* 8(2):e57129.
6. Carter AJ, Houle D (2011) Artificial selection reveals heritable variation for developmental instability. *Evolution* 65(12):3558–3564.
7. Wagner G, Altenberg L (1996) Complex adaptations and the evolution of evolvability. *Evolution* 50(3):967–976.
8. Hill WG, Mulder HA (2010) Genetic analysis of environmental variation. *Genet Res* 92(5-6):381–395.
9. Geiler-Samerotte KA, et al. (2013) The details in the distributions: Why and how to study phenotypic variability. *Curr Opin Biotechnol* 24(4):752–759.
10. Buchanan SM, Kain JS, de Bivort BL (2015) Neuronal control of locomotor handedness in *Drosophila*. *Proc Natl Acad Sci USA*, 10.1073/pnas.1500804112.
11. Shen X, Pettersson M, Rönnegård L, Carlborg Ö (2012) Inheritance beyond plain heritability: Variance-controlling genes in *Arabidopsis thaliana*. *PLoS Genet* 8(8):e1002839.
12. Rönnegård L, Valdar W (2012) Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genet* 13:63.
13. Littell RC (2006) *SAS for Mixed Models* (SAS institute, Cary, NC).
14. Brown MB, Forsythe AB (1974) Robust tests for the equality of variances. *J Am Stat Assoc* 69(346):364–367.
15. Nelson PR, Wludyka PS, Copeland K (2005) *The Analysis of Means: A Graphical Method for Comparing Means, Rates, and Proportions* (SIAM, Philadelphia).
16. Dworkin I (2005) A study of canalization and developmental stability in the sternopleural bristle system of *Drosophila melanogaster*. *Evolution* 59(7):1500–1509.
17. Harbison ST, McCoy LJ, Mackay TF (2013) Genome-wide association study of sleep in *Drosophila melanogaster*. *BMC Genomics* 14:281.
18. Seger J, Brockmann HJ (1987) What is bet-hedging? *Oxford Surveys in Evolutionary Biology*, eds Harvey PH, Partridge L (Oxford Univ Press, London), pp 182–211.
19. Turner TL, Miller PM (2012) Investigating natural variation in *Drosophila* courtship song by the evolve and resequence approach. *Genetics* 191(2):633–642.
20. Chintapalli VR, Wang J, Herzyk P, Davies SA, Dow JA (2013) Data-mining the FlyAtlas online resource to identify core functional motifs across transporting epithelia. *BMC Genomics* 14:518.
21. Robinson SW, Herzyk P, Dow JAT, Leader DP (2013) FlyAtlas: Database of gene expression in the tissues of *Drosophila melanogaster*. *Nucleic Acids Res* 41(Database issue):D744–D750.
22. St. Pierre SE, Ponting L, Stefancsik R, McQuilton P, the FlyBase Consortium (2014) FlyBase 102—advanced approaches to interrogating FlyBase. *Nucleic Acids Res* 42(D1): D780–D788.
23. modENCODE Consortium (2010) Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science* 330(6012):1787–1797.

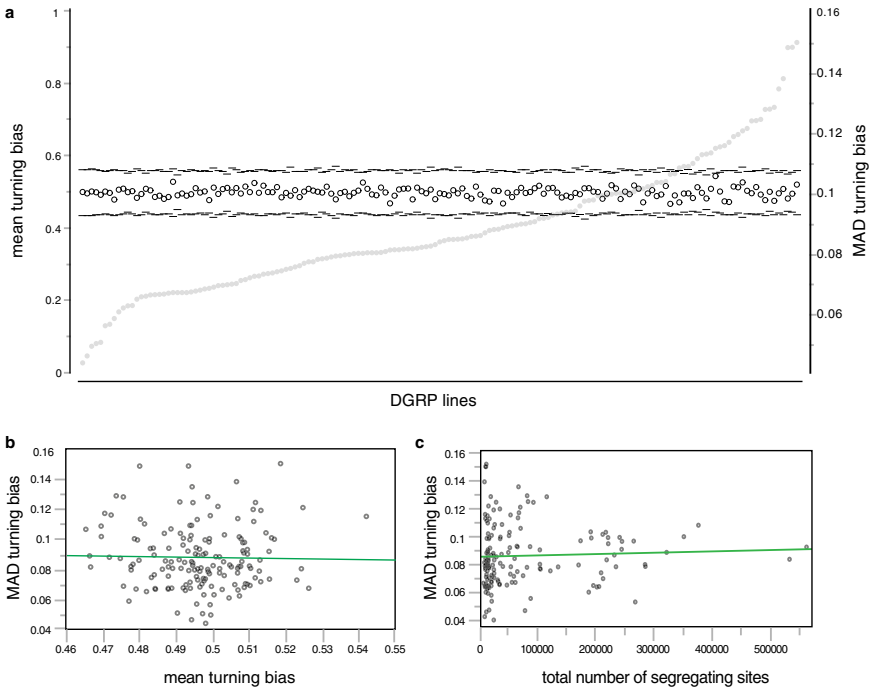

Fig. S1. Turn bias variability is correlated neither with the mean of turning bias nor residual genetic variation within line. (*A*) Distribution of the mean turning bias for each line, ranked by the MAD of turning bias (gray dots, scale on the right axis). Tick marks represent a 99.9% CI around the mean. There is no significant difference between lines in mean turning bias. Each lines are on average unhanded, making equal portions of left and right turns. ANOVA, $P < 0.87$. (*B*) No relationship between the mean turning bias and intragenotypic variability ($r^2 = 0.0004$, $P = 0.80$). (*C*) No relationship between intragenotypic variability and residual genetic variation segregating within line ($r^2 = 0.0005$, $P = 0.78$).

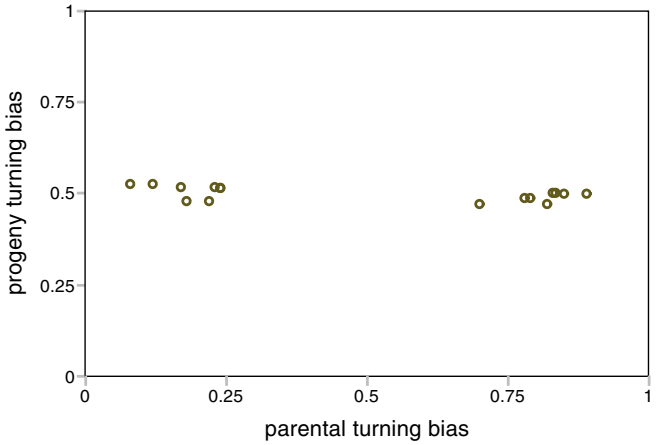

Fig. S2. No genetic variation segregating within lines is associated with turning bias. Crosses between extreme left turning parents (*Left*) and crosses between extreme right turning parents (*Right*) produce unbiased F_1 with a 50/50 left-right turning bias. Data modified from ref. 4.

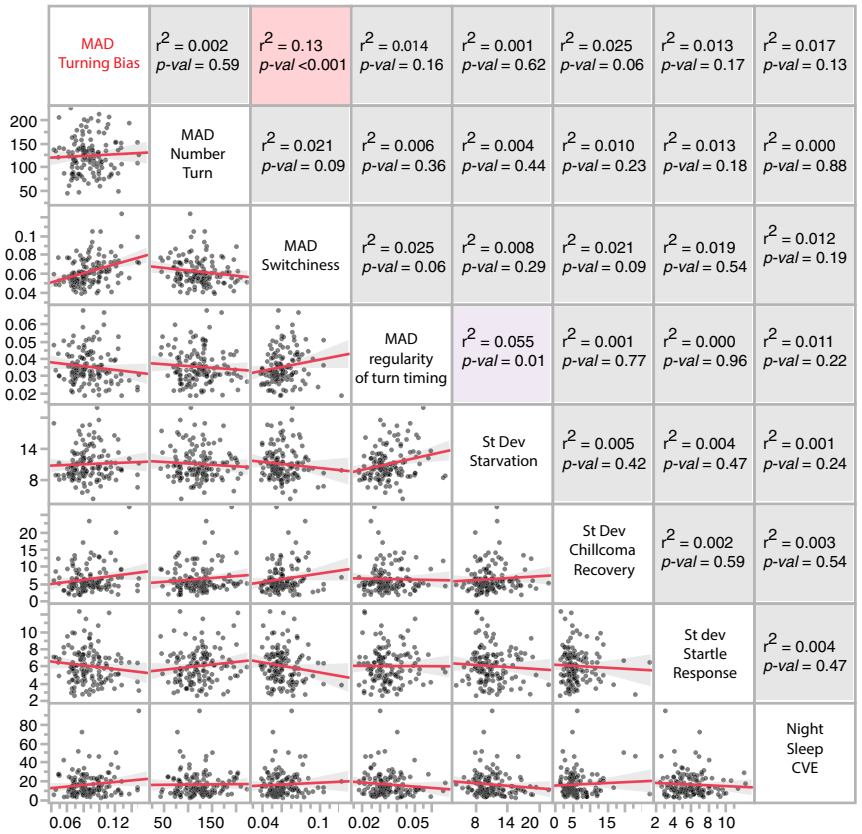

Fig. S3. Intragenotypic variability in various phenotypes is predominantly uncorrelated. Scatter plots of pairs of measures of intragenotypic variability. Points are DGRP lines. Red line is linear fit with 95% CI in gray. SDs for starvation resistance, chill coma recovery, and startle response were calculated based on data from ref. 22. CVE for night sleep data from ref. 23. For starvation resistance, *n* per line = 40; chill coma, *n* per line = 50; startle response, *n* per line = 40; night sleep, *n* per line = 32.

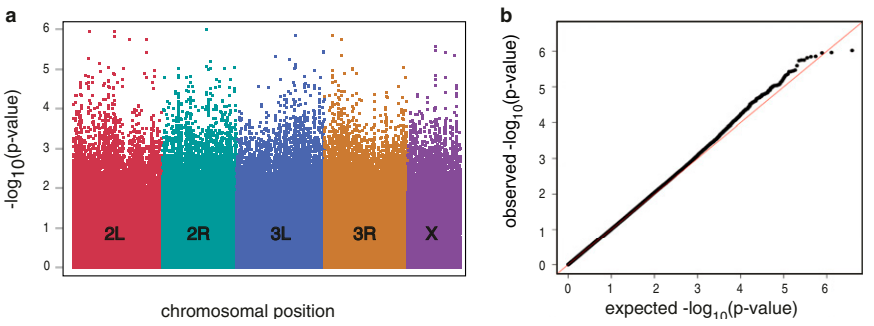

Fig. S4. GWAS *P* value distributions. (*A*) $-\log_{10}(P \text{ value})$ plotted along each chromosomal position for all SNPs. Colors and letters indicate chromosome arms. (*B*) QQ plot comparing observed *P* value to a uniform distribution of expected *P*- value ($\lambda = 1.07$).

Table S1. Statistics for analysis of mean and variance across DGRP lines for turning bias

Chrs	Position	Variant	Minor allele	Major allele	Minor allele frequency	Minor allele count	Major allele count	Single *P* value	Mixed model *P* val	Flybase ID	Gene ID	Genomic annotation
2R	13909827	SNP	C	T	0.10	15	141	2.63E-06	9.50E-07	FBgn0034289	CG10910	UTR_3_PRIME
2L	9129764	SNP	C	G	0.05	8	145	9.72E-07	1.09E-06	FBgn0052982	CG32982	INTRON
2L	3811210	SNP	T	C	0.11	17	140	1.10E-06	1.11E-06	FBgn0031573	CG3407	SYNONYMOUS_CODING
3R	5804797	DEL	T	TT	0.28	40	105	1.29E-06	1.39E-06	FBgn0263097	Glut4EF	INTRON
3L	13006450	SNP	C	A	0.17	27	129	1.39E-05	1.42E-06	FBgn0036333	MICAL-like	SYNONYMOUS_CODING
2L	9121721	SNP	T	G	0.07	11	146	5.92E-06	1.53E-06			
2L	16392736	SNP	T	C	0.37	52	87	2.68E-06	1.77E-06			
3R	9419254	SNP	C	G	0.22	33	119	1.50E-06	1.77E-06	FBgn0038159	CG14369	UPSTREAM
2L	12447996	SNP	A	T	0.40	56	85	1.13E-06	1.84E-06	FBgn0032434	CG5421	INTRON
X	12153076	SNP	T	A	0.08	12	141	4.61E-06	2.54E-06	FBgn0259240	Ten-a	INTRON
X	12153062	SNP	A	C	0.08	12	143	6.21E-06	3.37E-06	FBgn0259240	Ten-a	INTRON
3L	23001316	SNP	T	C	0.08	12	144	8.36E-07	3.54E-06	FBgn0262509	nrm	INTRON
2L	16392768	SNP	G	A	0.34	49	94	4.64E-06	3.64E-06			
3L	23001309	SNP	A	G	0.08	12	143	9.33E-07	3.74E-06	FBgn0262509	nrm	INTRON
X	16161192	SNP	T	A	0.35	54	100	6.95E-07	3.75E-06	FBgn0040207	kat80	UTR_3_PRIME
X	19449711	SNP	G	A	0.06	9	149	1.44E-05	4.38E-06			
3R	9414756	SNP	G	T	0.25	38	112	9.86E-07	4.58E-06	FBgn0038158	CG14370	DOWNSTREAM
3L	8951894	SNP	C	G	0.22	35	121	2.80E-06	4.77E-06	FBgn0035941	CG13313	UPSTREAM
2L	3572163	SNP	A	T	0.20	30	120	3.06E-05	5.26E-06			
3L	11733450	SNP	A	G	0.20	30	118	1.46E-05	5.81E-06	FBgn0036202	CG6024	INTRON
2R	14651257	DEL	AA	A	0.20	30	122	2.08E-06	6.44E-06	FBgn0034389	Mctp	INTRON
2L	6209462	SNP	A	G	0.05	8	148	6.11E-06	7.03E-06	FBgn0085409	CG34380	UTR_3_PRIME
3L	13006461	SNP	A	C	0.16	25	128	7.03E-05	7.33E-06	FBgn0036333	MICAL-like	NON_SYNONYMOUS_CODING
2R	13909829	SNP	A	G	0.10	16	140	1.87E-05	8.09E-06	FBgn0034289	CG10910	UTR_3_PRIME
2L	8003155	SNP	G	A	0.18	24	110	1.30E-05	8.58E-06	FBgn0031972	Wwox	INTRON
3R	26478454	SNP	C	T	0.46	67	80	2.89E-05	8.69E-06			
3L	12859591	SNP	A	G	0.17	27	128	7.59E-05	8.82E-06			
2R	20092066	SNP	A	T	0.12	19	136	2.54E-05	9.08E-06	FBgn0034990	CG11406	INTRON
3R	13031950	SNP	G	T	0.43	62	83	4.05E-05	9.13E-06			
2R	17929411	SNP	G	A	0.09	14	141	1.36E-06	9.16E-06	FBgn0005778	PpD5	SYNONYMOUS_CODING
2L	10873583	DEL	T	TGA	0.20	29	117	7.57E-06	9.21E-06			
2R	14823241	SNP	A	G	0.07	11	144	6.60E-06	9.26E-06	FBgn0034408	sano	INTRON
3L	12688628	SNP	T	C	0.11	17	141	8.66E-07	9.29E-06	FBgn0014343	mirr	INTRON
2R	7130275	SNP	T	G	0.17	26	129	5.03E-05	9.52E-06	FBgn0033593	Listericin	UPSTREAM
3R	13031960	SNP	C	A	0.42	61	85	3.98E-05	9.85E-06			
2R	14548570	SNP	T	C	0.48	75	81	1.49E-05	9.89E-06	FBgn0259202	CG42306	INTRON

Chrs, chromosome; *F*, *F* ratio statistic; *P*, *P* value for *F* ratio statistic

Table S2. Top GWAS hits for MAD of turning bias

Analysis category	df	*F*	*P*
Analysis of variance for mean turning bias			
$Line_{random}$	158	0.88	0.85
$Block_{random}$	28	1.12	0.29
Line × $block_{random}$	772	1	0.49
Box	5	0.41	0.84
Maze-array	11	0.52	0.88
Box × maze-array	49	1.13	0.26
Analysis of variance for the absolute median deviation of turning bias			
$Line_{random}$	158	4.31	<0.00001
$Block_{random}$	28	0.82	0.74
Line × $block_{random}$	772	1.04	0.2
Box	5	1.76	0.11
Maze-array	11	0.67	0.76
Box × maze-array	49	1.16	0.22
Alternative test for heterogeneity of variance between DGRP lines for turning bias			
O'Brien	158	8.5953	<0.00001
Brown–Forsythe	158	7.567	<0.00001
Levene	158	7.701	<0.00001
Bootstrap		Results in Table S1	
ANOMV		Results in Table S1	

F, *F* ratio statistic; *P*: *P* value for *F* ratio statistic.