bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The structure of behavioral variation within a genotype

Zach Werkhoven[1], **Alyssa Bravin**[1], **Kyobi Skutt-Kakaria**[1,2], **Pablo Reimers**[1,3], **Luisa Pallares**[4], **Julien Ayroles**[4], **Benjamin de Bivort**[1,*]

1 — Center for Brain Science and Department of Organismic and Evolutionary Biology, Cambridge, MA 02138

2 — Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

3 — Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147

4 — Department of Ecology and Evolutionary Biology, Princeton University, 08540 Princeton, NJ, USA

* — correspondence: debivort@oeb.harvard.edu

Abstract

Individual animals vary in their behaviors. This is true even when they share the same genotype and were reared in the same environment. Clusters of covarying behaviors constitute behavioral syndromes, and an individual's position along such axes of covariation is a representation of their personality. Despite these conceptual frameworks, the structure of behavioral covariation within a genotype is essentially uncharacterized and its mechanistic origins unknown. Passing hundreds of isogenic *Drosophila* individuals through an experimental pipeline that captured hundreds of behavioral measures, we found correlations only between sparse pairs of behaviors. Thus, the space of behavioral variation has many independent dimensions. Manipulating the physiology of the brain, and specific neural populations, altered specific correlations. We also observed that variation in gene expression can predict an individual's position on some behavior axes. This work represents the first steps in understanding the biological mechanisms determining the structure of behavioral variation within a genotype.

Keywords: *high-throughput behavior, individuality, personality, covariation, neural circuits, isogenic animals, Drosophila melanogaster*

Introduction

Individuals display idiosyncratic differences in behavior that often persist through time and are robust to situational context. Some persistent individual behavioral traits commonly occur in correlated groups and can therefore be said to covary. Behavioral ethologists have long understood that types of human and animal personalities often fall on multivariate axes of variation. A five dimensional model (Lewis, 1993) known as The Big Five personality traits is frequently used by psychologists to describe the range of human personality, and a similar model has been used to describe personality in fish (Réale et al., 2007). For example, human propensity for behaviors such as assertive- ness, talkativeness, and impulsiveness, are collectively described as extraversion and are thought to be anticorrelated with behaviors such as passivity, shyness, and deliber- ateness, all behaviors associated with introversion (Matthews et al., 2003).

In animal species, correlated suites of behaviors are described as behavioral syndromes. Aggressive behaviors such as fighting over mates or food are frequently correlated with exploratory behaviors such as foraging, and social interaction, have been observed in insects (Jeanson and Weidenmüller, 2014), arachnids (Grinsted et al., 2013), fish (Huntingford, 1976), and birds (Oers et al., 2004). The prevalence of behavioral correlation in so many species suggests that covariation is likely a universal feature of behavior. Although correlated individual differences in behavior are commonly observed, the struc- ture and mechanisms of behavioral covariation are not well understood. Typically, where an individual lands on these behavioral axes is thought to be established by a deterministic confluence of genetic and environmental effects. But there is increasing evidence that substantial individual behavioral variation is rooted in intragenotypic variation (Honegger and

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

de Bivort, 2018; Akhund-Zade et al., 2019). The extent to which intragenotypic behavioral variation is organized into syndromes or axes is essentially uncharacterized.

Substantial variation in specific behavioral measures, even in inbred lines raised in standardized conditions has been observed in several clonal animals, including: geckoes (of ethology, 2018), amazonian mollies (Bierbach et al., 2017), aphids and nematodes (Schuett et al., 2011; Stern et al., 2017). Genetic model systems hold particular promise for the mechanistic dissection of this variation, and intragenotypic variability (IGV) in behavior has been characterized in mice (Freund et al., 2013), zebrafish (Pantoja et al., 2016) and *Drosophila*. In flies, IGV of many behaviors has been studied, including: phototaxis (Kain et al., 2012), locomotor handedness and wing-folding (Buchanan et al., 2015), spontaneous microbehaviors (Kain et al., 2013; Todd et al., 2017), thermal preference (Kain et al., 2015) and object-fixated locomotion (Liu et al., 2018). Mechanistic studies of these behavioral phenomena have addressed two major questions: 1) what biological mechanisms underlie the magnitude of behavioral variability (e.g., genetic variation (Ayroles et al., 2015), or neural state variation (Kain et al., 2012; Buchanan et al., 2015)), and 2) what specific differences within individual nervous systems predict individual behavioral biases (Liu et al., 2018; Mellert et al., 2016). The mechanistic basis of individuality is an exciting new field, but no study to date has focused on characterizing the large-scale variance-covariance structure of IGV in behavior.

Behavioral correlations within a genotype could arise through a number of biological mechanisms including cell-to-cell variation in gene expression or individual differences in neural circuit wiring or synaptic weights. For example, stochastic variation during developmental critical windows has the potential to impart lasting differences between individuals in the absence of conspicuous genetic or environmental differences. Any such differences affecting nodes common to multiple behaviors in neural or molecular pathways may result in correlated shifts in behavior. Here, we study this directly by focusing on the correlation structure of behavioral variation when genetic and environmental variation are minimized.

This is an important biological question for several reasons. The structure of intragenotypic behavioral variability: 1) will shape the distribution and kinds of personalities that a population of organisms displays, even when they have matched genomes and environments, 2) is the product of stochastic biological outcomes, and its organization reveals how stochasticity drives variation, 3) constrains the evolution of behavior and adaptive phenotypic strategies like bet-hedging (Hopper, 1999), 4) is a relatively uncharacterized component of neural diversity and its manifestations in behavior and disease, and 5) may shed light on how the nervous system orchestrates behavior as a whole. In flies we have a suitable experimental system for directly characterizing this structure, as we can produce large numbers of individuals with nearly identical genomes, reared in the same environment, and collect many behavioral measures per individual. We performed this experiment in wild type inbred flies as well as wild type outbred flies and collections of transgenic lines manipulating neural activity. This approach let us contrast the structure of intragenotypic behavioral variability in animals where the source of variability is, respectively, stochastic fluctuations, genetic differences + stochastic fluctuations, and systematic perturbations of the nervous system + stochastic fluctuations. We found that in all cases, behavioral variation has high dimensionality, i.e., many independent axes of variation. The addition of variation from genetic differences and neural perturbations did not fundamentally alter this qualitative result, suggesting that stochastic fluctuations and genetic differences may structure behavior through common biological mechanisms.

Results

A high-throughput pipeline for measuring behavioral covariation

The first step in revealing the structure of behavioral variation within a genotype is to devise an experimental pipeline that produces a data set of many (200+) individual flies, with many behavioral measurements each. We first developed a number of behavioral assays, measuring both spontaneous and stimulus-evoked responses of individual flies, which could be implemented in a common experimental platform (Figure 1A; (Werkhoven et al., 2019)). This instrument features an imaging plane, within which flies moved in arenas of various geometries. Fly position was tracked with digital cameras using diffused infrared illumination invisible to the flies. Visual stimuli were presented to the animals using DLP projectors or LEDs embedded in the arena walls. We implemented six assays in this style, assessing 1) spontaneous walking in circular arenas, 2) preference to rest in brighter or dimmer light (in an environment of spatially structured illumination), 3) preference to rest in higher or lower light levels (in a fictive, temporally-modulated light environment), 4) optomotor responses to rotating visual stripes, 5) spontaneous left-right decision making in Y-mazes, and 6) phototaxis in Y-mazes, where flies are given a choice of walking toward or away from a lit LED (Figure 1B).

To these assays, we added three more, assessing 7) odor sensitivity in linear chambers (Claridge-Chang et al., 2009; Honegger & Smith et al., 2019) in which half of the compartment is filled with an aversive odorant, 8) spontaneous behavior, acquired via high- resolution 100Hz video and suitable for pixel-based unsupervised classification (Berman et al., 2014), and 9) circadian activity and spontaneous locomotion in 96-well plates with access to food. Each of these assays produced multiple behavioral metrics for each individual fly. For example, flies behaving in the phototactic "LED Y-maze" (assay 6) are performing phototaxis and exploratory locomotion but yield several different behavioral measures, including: the number of choices made by passing through the choice-point of the Y-maze (a measure of total activity), the fraction of turns that are to the right, the fraction of turns that are toward the lit LED, the number of pauses in which the animal did not move, the average duration of pauses, etc. Thus, the total collection of behavioral measures across all assays per fly was quite large (up to 121), constituting a diverse, inclusive characterization of individual behavior. Each assay has a particular measure that captures the behavior it is primarily designed to assess (e.g., the fraction of turns toward the lit LED in the LED Y-maze). In control experiments, we confirmed that these primary measures are consistent across days within an in-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

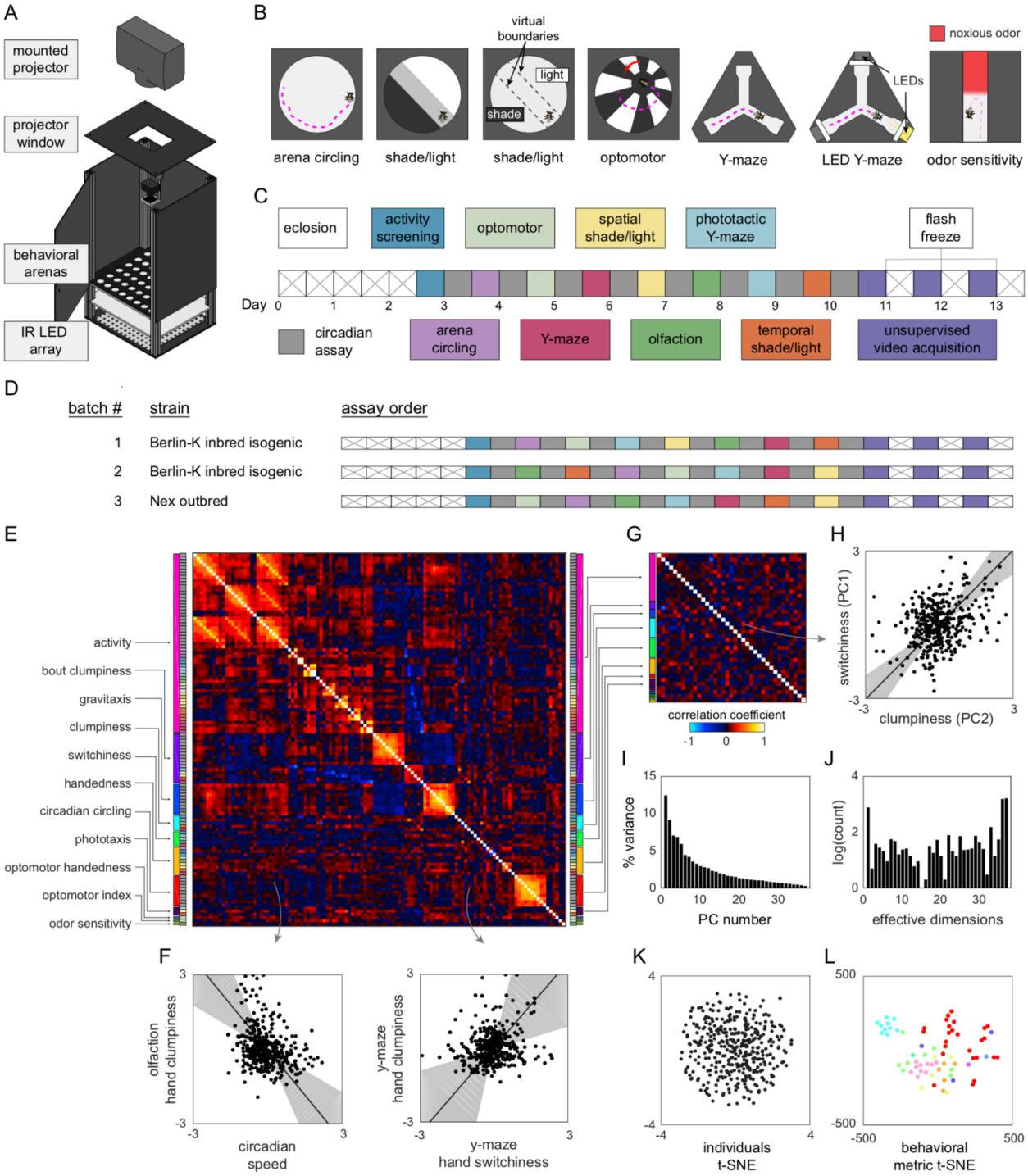


Figure 1 — Decathlon experimental design and structure of intragenotypic behavioral variation. A) Schematic of the imaging rig used for most Decathlon experiments. B) Schematics of the behavioral assays, illustrating the geometry of the arenas and stimulus structure. C) Timeline of the Decathlon experiment. Colors indicate the assays conducted on each day, half-black half-white blocks indicate the circadian assay and storage in 96-well plates. D) Timelines of the three Decathlon experiments, indicating the randomized order of assays 2-8. E) Full correlation matrix of all raw behavioral measures taken in the Decathlon. Colored blocks indicate blocks of measures we thought a priori might be correlated (outer blocks, text labels). Inner blocks indicate assay. F) Example scatter plots associated with measure correlations. Points are individual flies. Line is the best fit (PC1 of these points), grey region is the 95% confidence interval of the fit, as determined by bootstrap resampling. G) Distilled correlation matrix in which all correlated metrics represent unexpected relationships. H) Example scatter plot from the distilled correlations. Plot elements as in F. I) Scree plot of the ranked, normalized eigenvalues, i.e., the % variance explained by each PC, of the distilled behavior matrix, versus PC #. K) Effective dimensionality spectrum (See text and Figure S9) for the distilled matrix. Height of bars indicates organization at that dimensionality. K) Points corresponding to individual flies non-linearly embedded using t-SNE from the 121-dimensional raw measure space to two dimensions. L) Points corresponding to behavioral measures non-linearly embedded using t-SNE from the 384-dimensional space of flies to two dimensions. Colors indicate groups of measures we expected a priori to be related.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

dividual (Figure S1) (i.e., they reflect persistent idiosyncrasies (Kain et al., 2012; Buchanan et al., 2015; Honegger & Smith et al., 2019)).

In order to obtain all of the behavior measures from each experimental animal, we combined in a serial experimental pipeline lasting 13 consecutive days (Figure 1C), generally with one unique assay per day and continuous circadian imaging (assay 9) between assays. This pipeline begins with 3 day-old flies being loaded into the 96-well circadian imaging plates. Using a common behavioral platform, as much as possible, and storing flies between experiments in 96-well plates made maintaining the errorless identity of flies over the whole 13 day experiment substantially easier. Starting on day 3, daily assays began. On each day, flies were lightly anesthetized on an ice-chilled plate and aspirated, maintaining their identity, into the assay arrays. After the assay was completed (typically after 2hrs of recording) flies were again lightly anesthetized and returned to 96-well plates for renewed circadian imaging. On the first such day, flies were loaded into an array of circular arenas and imaged for total activity (in a version of assay 1). At this point, the most active 192 flies were retained for further testing. In preliminary experiments, we found that flies that were inactive at the beginning of the pipeline were very unlikely to produce substantial amounts of data over the rest of the pipeline. With the addition of this activity-screening assay, the total number of experiments was ten, and as each fly "competes" in all ten events, we refer to the entire pipeline as a Decathlon.

It is possible that the assay order has some effect on the recorded behavior measures. So we randomized the assay order between Decathlon implementations as much as possible (Figure 1D), subject to two restrictions: activity-screening was always the first assay, and high-resolution imaging for unsupervised analysis (assay 8) was always the last assay. (This assay has lower throughput, and three days were required to complete all 168 remaining flies. If this assay were performed earlier in the pipeline, it might introduce heterogeneity across subsequent assays.) When each fly completed its run through all Decathlon assays (i.e., over the three days of assay 8 imaging), it was flash-frozen in liquid nitrogen for RNA sequencing.

Behavioral covariation in inbred wild type animals

To collect data that would reveal the structure of behavioral variation within a genotype, we conducted two Decathlons using highly inbred, nearly isogenic flies derived from the wild type strain Berlin-K (BSC8522, (Nöthel, 1981)). We confirmed that this strain was, indeed, highly isogenic with genomic sequencing of individual animals, finding 75 SNPs in the population across the entire genome (Figure S2). 115 flies completed the first Decathlon, and 176 the second. While we aimed to collect 121 measures per fly, a portion of values were missing, typically because flies did not meet assay-specific activity cutoffs. For subsequent analyses (see Figure S3 for a schematic of all analysis pipelines), it was sometimes necessary to have a complete data matrix. So we infilled missing values using the alternating least squares method, which, as judged by analyses of toy ground-truth data, performed better than mean-infilling (Figure S4). For the sake of maximal statistical power, we wanted to merge the data sets from the two Berlin-K^{iso} Decathlons. The correlation matrices of these two data sets were not identical, but were substantially more similar than expected by chance (Figure S5), implying that while there were inter-Decathlon effects, much of the same structure was present in each and merging them justified. To do this, we z-score normalized the data points from each arena array/batch (within each Decathlon) across flies, thus eliminating any arena, assay, and Decathlon effects and enriching the data for contrasts between individuals. A grand data matrix was made by concatenating these batches (382 individuals x 121 behavior measures).

The full correlation matrix of this Berlin-K^{iso} data set is shown in Figure 1E. It contains a substantial amount of structure, indicating that large groups of behavioral measures covary. But the covariance of many pairs of measures in the matrix is not surprising. For example, almost all our assays generate some measure of locomotor activity (meanSpeed in circular arenas, number of turns in Y-mazes, meanSpeed in the olfactory tunnels, etc.), and one might expect that especially active flies in one assay will be especially active in another assay. Additional unsurprising structure in this matrix comes identical measures recorded in each of the 11-13 circadian assays each fly completed. But, even in this first analysis, surprising correlations were evident. For example, flies with higher variation in the inter-turn interval in the olfactory assay ("clumpiness" in their olfactory turning) exhibited higher mean speed in the circadian assays, and flies with higher variation in inter-turn intervals in the Y-maze (clumpiness in their Y-maze turning) exhibited lower mutual information in the direction of subsequent turns in the Y-maze ("switchiness" in their Y-maze handedness) (Figure 1F). See below, Buchanan et al., 2015, and Akhund-Zade et al., 2019 for more about these measures.

To produce an exhaustive list of such non-trivial correlations, we distilled the grand correlation matrix to a smaller matrix (the "distilled matrix"; Figure 1G) in which two kinds of interesting relationships were revealed: 1) uncorrelated dimensions among measures for which we had a prior expectation of correlation (e.g., if meanSpeed in circular arenas is found to be uncorrelated with meanSpeed in olfactory tunnels), and 2) correlated dimensions among measures for which we had no prior expectation of correlation. Relationships of the former class were identified by enumerating, before we ran any correlation analyses, groups of measures we expected to be correlated ("a priori groups"; Figure 1E,G). We looked for surprising independence within such groups by computing the principal components of data submatrices defined by the grouping (e.g., for the "activity" a priori group, by running PCA on the data set consisting of 382 individual flies, and 57 nominal measures of activity). We then replaced each a priori group submatrix with its projection onto its statistically significant PCs, as determined by a reshuffling analysis (see Methods, Figure S7). Some a priori groups largely matched our expectations, with relatively few independent dimensions among many measures (e.g., the gravitaxis group which had 10 measures and only 2 significant PCs), while others exhibited relatively many independent dimensions (e.g., the clumpiness group which had 5 measures and 5 significant PCs; see Figure S7 for all a priori group PCA analyses).

With a priori group submatrices represented in their respective significant PCs, the grand data matrix now contained 38 behav-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

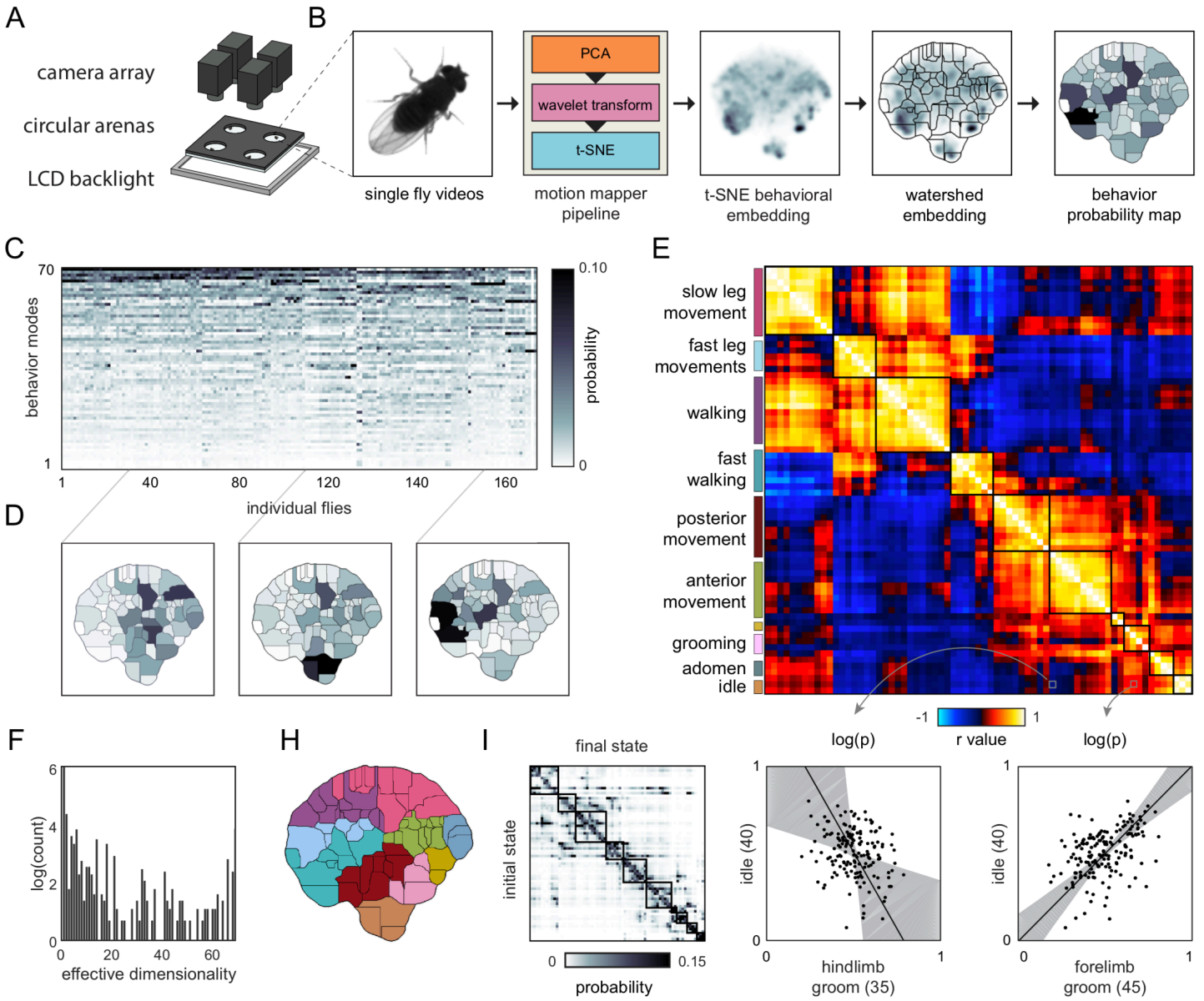


Figure 2 — Correlation structure of unsupervised behavioral classifications. A) Schematic of the four camera imaging rig used to acquire single fly videos. B) Overview of the data processing pipeline from single fly videos to behavioral probability maps. C) Behavioral classification probability density function (PDF) matrix. Columns correspond to behavioral PDFs for individual flies. D) Sample individual PDFs mapped to locations in tSNE space. Discrete regions correspond to watersheds of the tSNE embedded probability densities. E) Correlation matrix (top) for individual PDFs with rows and columns hierarchical clustered. Colored blocks indicate supervised labels applied to classifications post-hoc. Example scatter plots (bottom) of individual behavioral probabilities. Points correspond to probabilities for individual flies. Line is the best fit (PC1 of these points), grey region is the 95% confidence interval of the fit, as determined by bootstrap resampling. F) Effective dimensionality of the unsupervised behavioral classifications as calculated by the scree plot intersection of the observed and shuffled PDF matrices (see methods). H) Discrete behavioral map with individuals zones colored by supervised labels as in E. I) Transition probability matrix for behavioral classifications. Entries in the *i*th row and *j*th column correspond to the probability of transitioning from state *i* to state *j*. Blocks on the diagonal indicate clusters of supervised labels is in E.

ioral metrics. Every significant correlation between behavioral metrics at this point represents an unexpected element of structure of behavioral variation (Figure 1G). The first impression of this correlation matrix is that it is sparse. Most behavioral metrics are uncorrelated or weakly correlated, meaning there are many independent dimensions of behavioral variation. However, 176 pairs of behaviors were significantly correlated at a false discovery rate of 38%, and the distribution of *p*-values for the entries in this matrix exhibits a clear enrichment of low values (Figure S7) indicating an enrichment of significant correlations. As an example, flies with high values in the third PC of the switchiness a priori group tend to have high values in the second PC of the clumpiness a priori group (Figure 1H; a relationship that is built, in part, on the positive correlation between the ymazeTurnClumpiness and ymazeHandSwitchiness, Figure 1F). Interpreting the loadings (Figure S6) of these PCs indicates that this is a correlation between olfactory tunnel turn direction switchiness and olfactory tunnel turn timing clumpiness. We

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

detected a substantial number of correlations between different dimensions of switchiness and clumpiness (Figure S8), suggesting there are multiple couplings between these suites of traits.

Stepping back from specific pairwise correlations, we examined the overall geometry of behavioral variation. The full matrix contained 22 significant PCs, with PCs 1-3 explaining 9.3, 6.9 and 5.8% of the variance, respectively (Figure 1I). But the amount of variance explained across PCs does not provide the full picture of how many independent dimensions of variation are present in a data set. For example, a correlation matrix can be organized at different scales/hierarchically, so there need not be a single number that characterizes effective dimensionality. We developed an "effective dimensionality spectral analysis" that assessed the continuous degree of organization of a data set across the continuous range of dimensionalities from 1 to *d*, the dimensionality of the data. Briefly, we thresholded the correlation matrix across a wide range of thresholds, at each identifying the number of connected components in the matrix and recorded how often *n* connected components were observed. See Methods and Figure S9. These spectra can be interpreted as follows: peaks at dimensionality = 1 indicate that all measures are coupled in a network of at least weak correlations; peaks at dimensionality = *d* indicate that all measures have some degree of independence; peaks in between these values indicate intermediate scales of organization. Multiple peaks are possible because these kinds of organization are not mutually exclusive. The effective dimension spectrum (Figure 1J) of the distilled Decathlon data set had peaks at 1 and d (37). There was also evidence for structure over the full range of intermediate dimensionalities. Overall, the organization is one of predominantly independent behaviors, with some sparse sets of behaviors correlated with continuously varying strengths.

To assess how individual flies are distributed in behavior space, we embedded them from the 384 dimensional space into two dimensions using t-SNE (van der Maaten and Hinton, 2008). There appear to be no discrete clusters corresponding to "types" of flies. Instead, variation among flies appears continuously distributed around a single mode (Figure 1K). We also embedded behavioral metrics as points from the 121 dimensional space of flies into two dimensions (Figure 1L). This confirmed that while our intuition for which sets of metrics would be similar (the a priori groups) was right in many cases, metrics we thought would be similar were often dissimilar across flies, and sometimes metrics we did not anticipate being similar were (e.g., phototaxis and activity level).

The structure of covariation in unsupervised behavioral clusters

With data from the second Decathlon, we characterized the structure of variation in a set of behaviors that was potentially exhaustive for one behavioral condition (free walking/motion in a 2d arena; Figure 2A). High-speed, high-resolution video was acquired for four flies simultaneously in each of two rigs. Over three days, we acquired 13.5 GB of 200x200 px 100Hz videos centered on each fly as they behaved spontaneously over the course of 60 minutes. These frames were fed into an unsupervised analysis pipeline (Berman et al., 2014) that computed high-dimensional representations of these data in the time-frequency domain before embedding them in two dimensions and demarcating boundaries between 70 discrete modes of behavior (Figure 2B). The behavior of each fly was thus represented as one of 70 values at each frame. Flies exhibited a broadly similar probability distribution of performing each of these behaviors (Figure 2C), though there were conspicuous differences among individual patterns of behavior (Figure 2D).

The correlation matrix of behavioral modes identified in the unsupervised analysis was highly structured (Figure 2E; like the correlation matrix of the rest of the decathlon metrics, Figure 1E), appearing to have approximately 8 independent dimensions of variation (Figure 2F). For this analysis, there was no equivalent of a priori groups of behavioral measures, as metrics were not defined prior to the analysis. But, in examining sample movies of flies executing each of the 70 unsupervised behavioral modes (Movies S1-4), it was clear that highly correlated behavioral modes tended to reflect variations on the same type of behavior (e.g., walking) or behaviors performed on the same region of the body (e.g., anterior movements including eye and foreleg grooming; Figure 2G). In other words, individual flies that perform more eye grooming tend to perform more of other anterior behaviors. There were some correlations between behaviors implemented by disparate parts of the body. For example, flies that spent more time performing anterior grooming also spent more time performing slow leg movements (Figure 2G). The overall similarity of covarying behaviors was confirmed by defining groups of covarying behaviors and observing that they were associated with contiguous regions of the embedded behavioral map (Figure 2I). That is, behaviors whose prevalence covaries across individuals have similar time-frequency patterns across the body. Moreover, these clusters of covarying, contiguously embedded behaviors exhibited similar temporal transitions; behaviors that covary across individuals tend to precede specific sets of subsequent behaviors (Figure 2J). Thus, there appear to be couplings between the dimensions of behavioral variation across individuals, the domains of the body implementing behavior, and the temporal patterning of behaviors.

Neural perturbation of behavioral covariation

To 1) confirm that the Decathlon experiments revealed biologically meaningful couplings between behaviors and 2) probe biological mechanisms potentially giving rise to behavioral correlations, we treated correlations in the Decathlon matrices as hypotheses to test in a thermogenetic neural circuit perturbation screen. Specifically, we focused on the many correlations between measures of turn timing clumpiness and turn direction switchiness (Figures 1H, S8). Before the Decathlon experiment, we had no reason to think these measures would be correlated as one describes higher order structure in the timing of locomotor turns (clumpiness) and one describes higher order structure in the direction of sequential turns (switchiness). Our Decathlon-derived prediction was that if perturbing a circuit element caused a change in clumpiness, it would tend to also cause a change in switchiness, in a consistent direction. We looked for such correlated changes when we inactivated or activated neurons in the Central Complex, a cluster of neuropils involved in locomotor behaviors (Buchanan et al., 2015; Ofstad et al., 2011; Kottler et al., 2019) and heading estimation (Seelig and Nature, 2015; and Shah et al., 2017; Kakaria and de Bivort, 2017). We used a set of Gal4 lines (Wolff and Rubin, 2018), each of which targets a sin-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

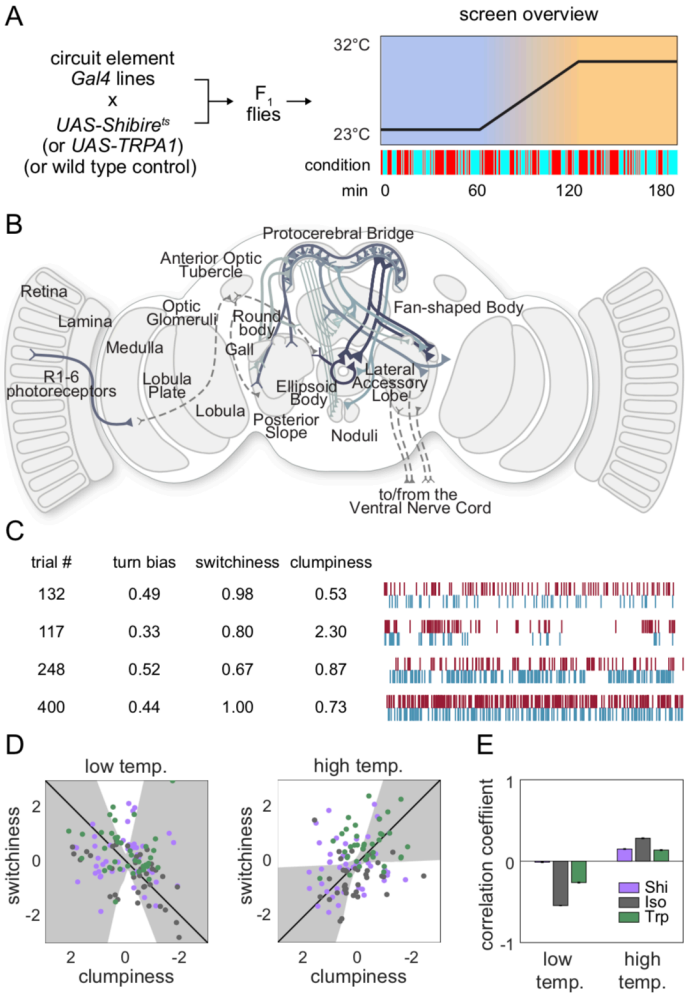


Figure 3 — Effect of thermogenetic neural perturbation on clumpiness and switchiness. A) Schematic of the experiment for testing the effects of neural perturbation on switchiness and clumpiness. Plot depicts the schema used to activate or silence neurons. Colored hash marks indicate periods when the light was on (cyan) or off (red). B) Diagram of the various neurons (solid lines) target by lines tested in the screen. See Table S3 for list. C) Example turn data from individuals exhibiting low and high clumpiness (upper rows) and low and high switchiness (bottom rows). D) Scatter plots of line average clumpiness and switchiness at the permissive (left) and restrictive (right) temperatures. Plots on the top row include all lines screened. Plots bottom row exclude the 8 lines with the lowest clumpiness and switchiness measures at the restrictive temperature. Lines indicate the line of best fit and shaded regions indicate the 95% confidence of the fit as determined by bootstrap resampling. D) Correlation coefficients for each effector type at the permissive and restrictive temperatures for all lines (top) and a subset of lines (bottom) as in C. Error bars indicate the 95% confidence interval as determined by bootstrap resampling.

gle cell type and that tile the entire Protocerebral Bridge (a Central Complex neuropil) (See Table 1), to express Shibire^{ts} (Kitamoto, 2001) or dTRPA1 (Hamada et al., 2008), thermogenetic reagents that block vesicular release and depolarize cells respectively. As controls, we used flies heterozygous for the Gal4 lines, and lacking the effector transgenes.

Flies with these genotypes were loaded into Y-mazes for behavior imaging before, during and after a temperature ramp from the permissive temperature (23°C) to the restrictive temperature (32°C for dTRPA1 and 29°C for Shibire^{ts}) (Figure 3A). At the permissive temperature, we observed significant negative correlations in the average line clumpiness and switchiness of control and dTrpA1-expressing lines (Figure 3B,D). This suggests that the mechanisms which couple variation in switchiness and clumpiness within a genotype may also be at play across genotypes. Surprisingly, at the restrictive temperature, we saw significant positive correlations between clumpiness and switchiness in all three experimental treatments: control (Gal4/+), Gal4/Shibire^{ts} and Gal4/dTRPA1 lines (Figure 3 C,D). That this correlation appeared in controls suggests that temperature alone can selectively alter the function of circuit elements regulating both clumpiness and switchiness, effectively reversing their coupling. The dTrpA1- and Shibire^{ts}-expressing lines also showed this reversal, but to a lesser extent, suggesting that perturbing neurons in the central complex can block temperature-induced changes in the coupling of clumpiness and switchiness.

Gene expression correlates of individual behavioral biases

That thermogenetic manipulation can disrupt correlation between behavioral measures suggests that specific patterns of neural activity underlie the structure of behavioral variation in wild type flies. Such physiological variation could arise in stochastic variation in gene expression (Lin et al., 2016) in circuit elements. To test this hypothesis, we performed RNA sequencing on the heads of the flies at the end of the first Decathlon experiment (Figure 4A). We used Tm3'seq (Pallares et al., 2019) to make 3'-biased libraries for each individual animal. We quantified the expression of 17,470 genes in 48 flies. The expression profiles were strongly correlated across individuals (Figure 4B), but there was some evident variation across individuals (Figure 4C). To assess whether this variation was meaningful with respect to behavioral variation, we trained linear models to predict an individual's position on the major axes of behavioral variation from its transcriptional idiosyncrasies. Specifically, we fit models to predict flies' values on the principal components of the distilled behavioral matrix from their values on the first 8 principal components of transcriptomic variation. Cross-validated models were significantly predictive for four different behavioral components (Figure 4D). When we compared the significance of these models to that of models fit to shuffled distilled behavioral matrix principal components, we found that at least three dimensions of behavioral variation (PCs 6, 8, 9 and 12 of the distilled behavior matrix) could be predicted with higher performance (cross-validated -log(p) between 1.7 and 2.6) than any of the models fit to shuffled data.

After identifying behavioral components that could be significantly predicted from variation in gene expression, we assessed whether the genes most predictive in each model shared functional characteristics. We constructed a list of genes predictive of each behavioral component for use in gene ontology (GO) analysis by restricting the total list to those with the largest coefficients. Specifically, we computed an "importance coefficient" for each transcript as the sum, over all behavior PCs going into the model, of its loading on that PC times the linear coefficient assigned to that PC in the model. We then retained genes with importance coefficient magnitudes in the 95th percentile (874 genes per behavior component). We found that these gene lists were highly overlapping, with 44% of the genes being common

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

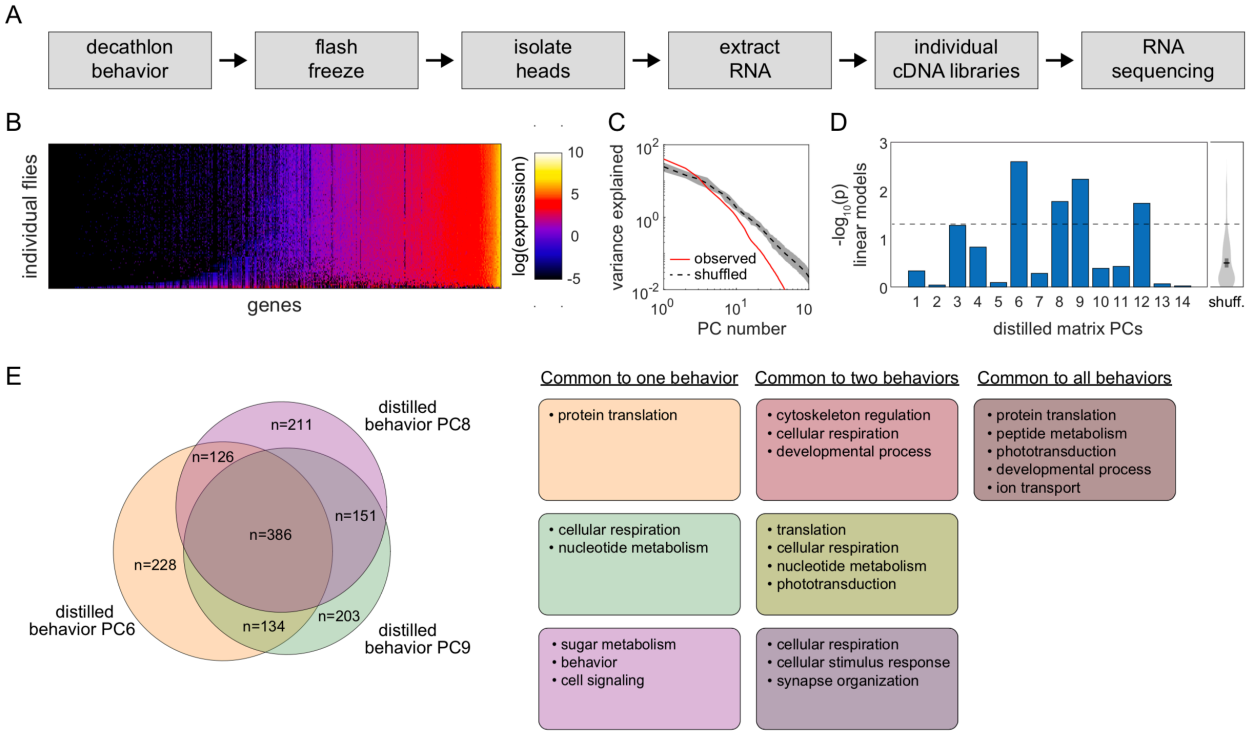


Figure 4 — Correlation between individual transcriptomes and behavioral biases. A) Experimental stages for collecting transcriptomic data for flies that have completed the Decathlon. B) Data matrix of individual head transcriptomes. Rows are individual flies (n = 48). Columns are 17,470 genes sorted by their mean expression across individuals. Rows are cluster ordered. C) Log scree plots of the variance explained for ranked principal components of gene expression variation for the observed data and shuffled data. Color indicates observed or shuffled data. Shaded regions correspond to 95% confidence intervals as calculated by bootstrap resampling. D) Performance (-log cross-validated p) of linear models predicting the behavior of individual flies on the 14 significant PCs of the distilled behavioral matrix. Dashed line indicates p = 0.05. Distribution at right is the kernel density estimate of the distribution of p-values observed when training comparable models on gene-wise shuffled expression data. The models trained on PCs 6, 8, 9 and 12 outperform this shuffled control. Black line and bars indicate the mean and +/-1 SEM of the shuffled model p values. E) Gene Ontology (GO) analysis of the 95th percentile of genes contributing the most to behavioral predictions for models of three behavioral PCs. Numbers in Venn diagram indicate the count of genes contributing, at this level, to each set of models. Biological process descriptors at right indicate the categories of GO labels that are enriched in each set of model predictors.

to all three behavioral models. The gene lists were therefore further subdivided by their intersection with one another, resulting in seven lists for GO analysis. Although each list contained a unique set of genes, we found a high degree of correspondence between the functional clusters enriched in each gene list (Figure 4E). Genes related to cellular respiration, protein translation, developmental processes, phototransduction, and ion transport were significantly enriched in the gene lists for all behavioral models. Cellular respiration was the most commonly enriched functional class, appearing in five of the seven lists, suggesting that overall metabolic rate may be broadly predictive of many behaviors. Although similar functional groups were enriched in total the gene list for each behavior, we found contrasts between them, such as synaptic organization and cell-to-cell signaling, which respectively are enriched in genes predictive of two and one PC of behavioral variation.

Comparing intra- and inter-genotypic behavioral covariation

If transcriptomic differences predict individual behavioral differences within a genotype, then the structure of behavioral variability might be very different in outbred populations, where transcriptomic differences are (presumably) much more substantial. We tested this by conducting a Decathlon experiment on outbred flies (n = 192) from a synthetic genetic mapping population (Long et al., 2014). These animals were from a high (~100)-generation intercross population ("NEX"; seeded in the first generation by eight kinds of F_1 heterozygotes produced by round-robin cross from eight inbred wild strains). A distilled correlation matrix of behavior metrics (Figure 5A) was produced by the same method as above. At first glance, it appears qualitatively similar to the distilled correlation matrix from isogenic animals (Figure 1E). This impression was confirmed in more formal comparisons of the structure of behavior in inbred and outbred populations. In isogenic and outbred populations: 1) individuals do not fall into discrete clusters, as determined by t-SNE embedding of individuals as points (Figure 5B, left). Moreover inbred and outbred flies appear to lie on the same manifold in behavior metric space; 2) behavioral measures cluster according to their membership in a priori groups similarly in outbred (Figure 5B, right) and inbred (Figure 1K) populations; 3) the distribution of percent variance explained by principal component was similar (Figure 5C, left); and 4) there is a similar spectrum of covariance structure, with most metrics independent, and a sparse network of correlations of varying strength (Figure 5C, right).

After determining that the overall structure of behavioral variation in isogenic and outbred populations is similar, we asked whether it was also similar in specific correlations. There appears to be some similarity at this level (Figure 5D); the correlation coefficient between the isogenic and outbred populations in

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

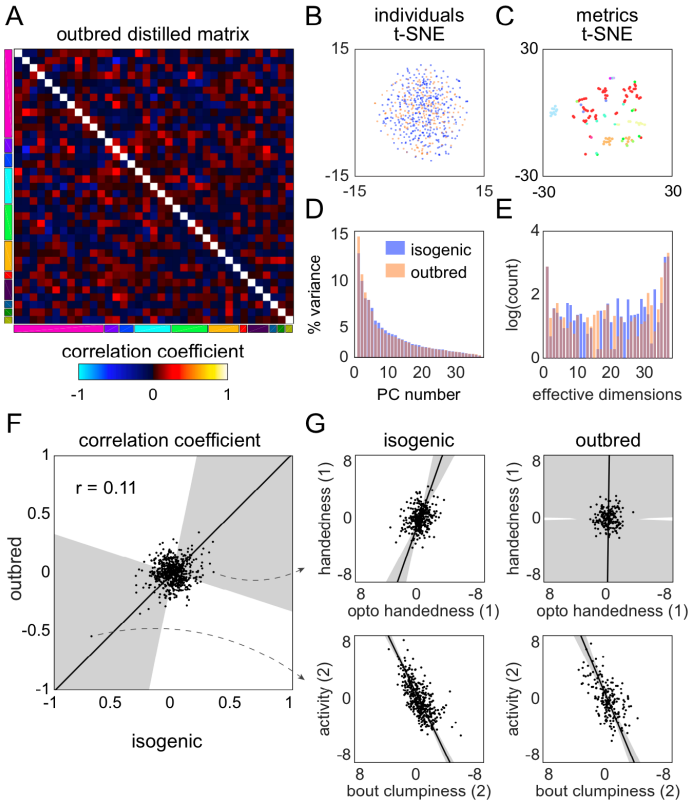


Figure 5 — Structure of behavioral variation in outbred flies. A) Distilled correlation matrix for outbred NEX flies. Colored blocks indicate a priori groups as described in figure 1. B) Points corresponding to individual flies non-linearly embedded using t-sne from the 121-dimensional raw measure space to two dimensions. Color indicates whether the genotype is isogenic (blue) or outbred (orange). C) Points corresponding to behavioral measures non-linearly embedded using t-SNE from the 192-dimensional space of flies to two dimensions. Colors correspond to a priori group. D) Scree plot of the ranked, normalized eigenvalues, i.e., the % variance explained by each PC, of the distilled covariance matrix, versus PC #. E) Log histogram of the effective dimensionality of the distilled matrix, as calculated by the number of connected components in the thresholded graph covariance matrix over 5,000 linearly spaced covariance thresholds (see methods). F) Scatter plot of the distilled matrix correlation coefficients for isogenic and outbred flies. Points correspond to distilled matrix metric pairs. G) Example scatter plots of distilled matrix metric pairs for inbred (left) and outbred (right) flies. The rows of plots highlight a pair of metrics with qualitatively different (top) and similar (bottom) correlations in inbred and outbred flies.

the pairwise correlations between behavior metrics is statistically significant ($p = 0.001$), but low in magnitude ($r = 0.12$). Examining specific pairs of scatter plots, it is clear that the correlations between specific behaviors are sometimes the same between the isogenic and outbred animals, but sometimes not (Figure 5F-G). A caveat in interpreting apparent differences between the isogenic and outbred matrices is that two qualities are different between the animals used in the respective experiments: the degree of genetic diversity, but also (necessarily) the genetic background of the flies.

Behavioral covariation under different sources of variation

Lastly, we examined how the correlation structure of behavior compared between sets of flies with variation coming from different sources. Specifically, we looked at four data sets: 1) the BABAM data set (Robie et al., 2017), in which measures were acquired from groups of flies behaving in open arenas, and variation came from the thermogenetic activation of 2,381 different sets of neurons (the first generation FlyLight Gal4 lines (Jenett et al., 2012); 2) a *Drosophila* Genome Reference Panel (DGRP; (Mackay et al., 2012)) behavioral data set, in which measures were acquired in behavioral assays similar to the Decathlon experiments (sometimes manually, sometimes automatically), and variation came from the natural genetic variation between lines in the DGRP collection; 3) a DGRP physiological data set, in which measures are physiological or metabolic (e.g., body weight and glucose levels) and variation came from the natural genetic variation between lines in the DGRP collection; and 4) the split-Gal4 Descending Neuron (DN) data set (Cande et al., 2018) in which measures came from the same unsupervised cluster approach as Figure 2, and variation comes from the optogenetic activation of specific sets of descending neurons projecting from the brain to the ventral nerve cord (Namiki et al., 2018). We analyzed these data sets with the same tools we used to characterize the structure of behavioral variation in the Decathlon experiments.

All of these data sets show substantial structure in their correlation matrices (Figure 6A, S10). The BABAM and especially the DN correlation matrices contain numerous high correlation values, indicative of strong couplings between behaviors under these neuronal manipulations. The DGRP correlation matrices, especially the DGRP behavioral matrix, look more qualitatively similar to the Decathlon matrix, with lower, sparser correlations. This suggests that behavioral variation has coarsely similar structure whether variation arises intragenotypically (e.g., through stochastic variation in transcription [Figure 4B]; Figure 1G), intergenotypically among outbred individuals (Figure 5A), or intergenotypically among inbred lines derived from wild populations (Figure 6A2). A caveat of this conclusion is that sparse correlation matrices can arise either from true, biological independence of behavior measures or from measurement error.

The effective dimensionality spectra of these matrices (Figure 6B) are similar in offering evidence of organization over a wide range of dimensionalities, including high dimensionalities. Only the BABAM spectrum has no power at the dimensionality of its raw count of measurements. The BABAM and DN spectra have a single predominant peak (at dimensionality = 1), suggesting that most measures belong to a single network of at least weak couplings. This is especially true in the BABAM data, and more true in the optogenetic experimental animals than controls in the DN data. The DGRP physiology data exhibits weak peaks at dimensionality = 1 and d, but also a broad peak between 20 and 40, suggesting that intergenotypic variation in physiology may have an intrinsic dimensionality in that range. The spectrum of the DGRP behavior data looks similar to that of the Decathlons, with peaks at dimensionality = 1 and d, and evidence for structure over the full range of intermediate dimensionalities. The distribution of individual lines in the space of DGRP behavior (and physiology) measures appears to be distributed single mode (Figure 6C), like individual flies in the Decathlon (Figure 1J). In contrast, there is some organization of individual lines in the BABAM and DN data sets, likely reflecting neuronal perturbations affecting multiple circuit elements mediating the same behavior(s) (e.g., multiple lines targeting the same neuropil). Mea-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

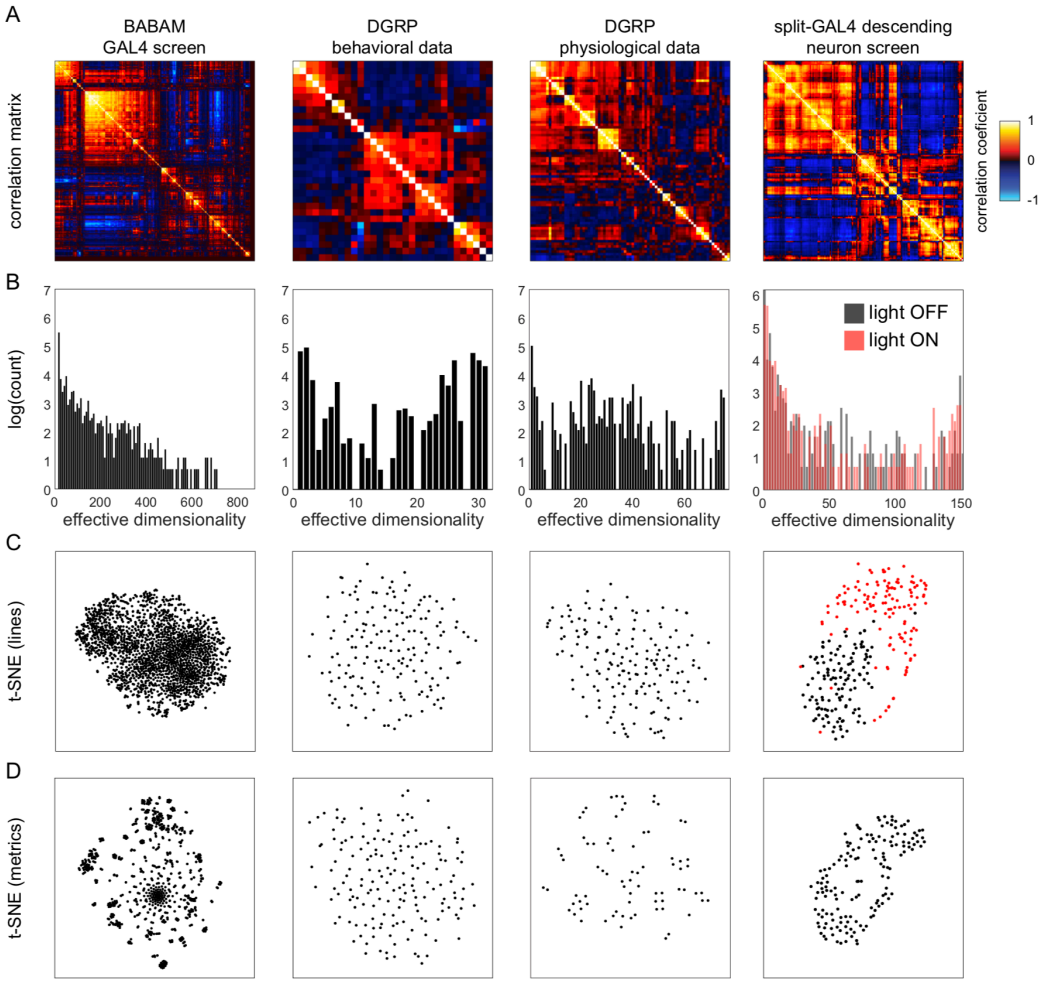


Figure 6 — Meta analysis of *Drosophila* covariation. A) Correlation matrices of previously published data sets. Columns correspond to analyses performed on each data set. From left to right, the data sets are as follows: line averages of supervised behavioral classifications following thermogenetic inactivation in the fly olympiad screen (Robie et al., 2017), line averages of behavioral phenotypic data from wild-type inbred lines in the *Drosophila* genomic reference panel (DGRP) database, line averages of physiological phenotypic data from the DGRP database, line averages of the fold change in unsupervised behavioral classifications following optogenetic activation of descending neurons (Cande et al., 2018). B) Log histograms of the effective dimensionality of all data sets, as calculated by the number of connected components in the thresholded graph covariance matrix over 1,000 linearly spaced covariance thresholds (see methods). Color in the rightmost plots (B-D) indicates either control (Gal4 driver only) or experimental animals (Gal4 x dTrapA1). C) Points corresponding to lines non-linearly embedded using t-SNE from the D-dimensional raw measure space to two dimensions (from left to right, d = 871, 31, 77, 151). D) Points corresponding to lines non-linearly embedded using t-SNE from the n-dimensional raw measure space to two dimensions (from left to right, n = 2083, 169, 169, 176).

sures fall into clusters in all of these data sets except the DGRP behavior measures (Figure 6D), which appear distributed around a single mode, perhaps reflecting the high dimensionality of behavior itself.

Discussion

Individuals exhibit different behaviors, even when they have the same genotype and have been reared in the same environment. These differences might covary, or lie on a manifold of specific geometry in behavior variation-space, but the structure of intragenotypic behavioral variation is uncharacterized. We designed a pipeline of ten behavioral assays (Figure 1), which collectively yielded up to 121 behavioral measures per individual animal. We also used unsupervised clustering to identify an additional 70 measures per individual based on a time-frequency analysis of high resolution video of the flies behaving spontaneously (Figure 2). These measures were the fly-specific rates of exhibiting each of the 70 unsupervised behavioral modes. All in all, across three 15 day Decathlon experiments, we collected 191 behavior measures from 576 flies. This allowed us to produce a full correlation matrix for all of the behavioral measures for the variation present in isogenic animals grown in the lab (Figure 1E).

There is a well-developed theoretical framework for understanding the multivariate correlation structure of phenotypes. In quantitative genetics, G-matrices characterize the variance and covariance structure of phenotypes (be they behavioral, physiological, morphological) stemming from genetic differences among individuals or strains (Lynch and Walsh, 1998; Bruijning et al., 2019). These representations allow the quantitative prediction of responses to selection and constrain the combinations of phenotypes individuals can exhibit. As such, these representations are

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

a key part of predicting the future trajectories of evolution. Just as the phenotypic variance can be parsed into genetic variance, environmental variance, GxE interaction variance etc., covariance can be similarly dissected (Charmantier et al., 2014; Berdal and Dochtermann, 2019). For example the classic model of phenotypic variance $V_P = V_G+V_E$ has a direct phenotypic covariance analogue: $Cov_P = Cov_G + Cov_E$. The last term (Cov_E) is further broken down into Temporary Environmental covariances and Permanent Environmental covariances (Cov_{PE}) that endure for the duration of observations (exactly like our measurements of individual behavioral bias). In flies, we have the potential to directly measure Cov_{PE} by rearing isogenic animals in standardized lab environments, profiling their individual biases over a wide range of behavioral measures, and directly quantifying the variance and covariance of behavioral bias. This is significant for quantitative geneticists because a meta-analysis of behavioral traits indicates that across behaviors, 23% of variance can be attributed to heritable factors (Dochtermann et al., 2019). This means that environmental factors, which include both deterministic effects and stochastic intragenotypic effects, explain 77% of behavioral variance. Thus, characterizing the structure of Cov_{PE} will contribute to closing a significant gap in our understanding of the basis of behavioral diversity.

For ethologists and behavioral neuroscientists, this work yields a view of the geometry of intragenotypic behavioral variation, which can be thought of as emergent product of developmental biological processes and the dynamic interaction of neural activity and animals' environment. From the full behavioral matrix, we made a so-called "distilled" matrix in which any significant correlation indicates a surprising new relation between behaviors (Figure 1G and S8). This form of the data minimizes duplicated measures of the same behavior, allowing us to cleanly analyze the geometry of behavioral variation. We found that behavioral measures were largely independent of each other, so that the main effective dimensionality of the behavioral space matched the number of behaviors measured. But a single number cannot fully characterize the dimensional organization of a correlation matrix; so we developed a spectral approach that examined the degree of organization across all possible dimensionalities in the data (Figure 1J, 3E, 6B, S10). This revealed a degree of organization at intermediate dimensionality corresponding to sparse correlations between specific pairs of behaviors (Figure 1G,H, S9). We found no evidence of discrete types of flies. Embedding data points corresponding to individual flies from the high dimensional space of individual biases into two dimensions produced a broad distribution around a single mode (Figure 1J).

One of the specific, surprising correlations we discovered was between "clumpiness" and "switchiness" (Figures 1F,H, 3, S8). These are slightly abstract, higher-order behavior measures, corresponding respectively to the burstiness of turn/action/decision timing and the degree of independence between consecutive binary choices. We had no a priori reason to expect these measures would be correlated, since one pertains to the structure of actions in time, and one pertains to the persistence of trial-to-trial biases. However, their linkage may reflect a shared role in controlling the higher order statistics of exploration. Sequences of behavior with clumps of bouts (either in time or in space) might contribute to fat-tailed distributions of dispersal that are advantageous over Brownian motion for foragers in environments of sparse resources (Bartumeus et al., 2002). Variation in switchiness and clumpiness across individuals might therefore reflect variation in multibehavioral navigational strategies, perhaps as part of bet-hedging evolutionary strategy (Hopper, 1999; Kain et al., 2012). From a perspective of biological mechanism, the correlation between these two behaviors (or other pairs we discovered) could be established during development. Individual wiring (Mellert et al., 2016; Linneweber et al., 2019) or physiological variations in neurons that mediate more than one behavior could impart coupled changes to all such behaviors.

If such an explanation accounts for the correlation of clumpiness and switchiness, there may be shared neural circuit elements in the circuits controlling decision-timing and decision-bias. We tested this idea in a thermogenetic screen of circuit elements in the central complex, a brain region where heading-direction is represented (Seelig and Nature, 2015) in a ring-attractor circuit (Kakaria and de Bivort, 2017). We found that increasing the temperature changed the sign of the correlation between switchiness and clumpiness (Figure 3), suggesting that changes to brain physiology can alter correlation structure, though these effects might instead be caused by any of the many changes in neural state that accompany a temperature change. Interestingly, we also found that the effector-inducing temperature manipulation alone concomitantly changed clumpiness and switchiness in some lines. This suggests that potentially subtle alterations of circuit physiology (e.g., temperature shifts well within physiological limits; Haddad, 2018) can affect the function of circuit elements governing multiple behaviors. In contrast, we found that thermogenetic manipulation of neurons in the central complex can diminish the effect of temperature shift on the correlation between clumpiness and switchiness. At the same time, the correlation of these behaviors went up significantly for the lines expressing Shibirets at the restrictive temperature. Thus extreme neural perturbations (such as those induced by thermogenetic manipulation) can apparently couple or decouple behaviors.

We included behavioral assays in the Decathlon pipeline under a number of constraints. The assays had to be high throughput, both in the number of flies that could be assayed and in measures being automatically acquired. Flies had to survive at high rates, and the measures had to be stable over multiple days (Figure S1), because the whole experiment lasted 15 days. Because not all behavior measures showed robust stability for this duration (all showed at least some day-to-day stability), the distilled Decathlon matrices likely represent an underestimate of the behavioral couplings exhibited over short periods (and perhaps an overestimate of life-long couplings, as flies can live more than a month (Linford et al., 2013)). In the end, we employed a number of spontaneous locomotion assays and simple stimulus-evoked assays like odor-avoidance and phototaxis.

Light responses were measured in a number of assays (Table S1), specifically: the LED Y-maze (in which flies turned toward or away from lit LEDs in a rapid trial-by-trial format; Werkhoven et al., 2019), the Spatial Shade-Light assay (in which flies chose to stand in lit or shaded regions of an arena that only changed every four minutes), and Temporal Shade-Light (in which the same luminance levels were used as the previous assay, but a fly experienced them by traveling into virtual zones which triggered the illumination of the whole arena at a particu-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

lar luminance) (Figure 1B). These assays were potentially redundant, and we included this cluster of phototaxis assays in part as a positive control. However, the three phototactic measures we thought would be correlated a priori were, in fact, independent of each other, each being represented in the distilled correlation matrix (Figure 1G). This may reflect flies using different behavioral algorithms (Krakauer et al., 2017), implemented by non-overlapping circuits, to implement these behaviors. Indeed, independence between behavior measure was the typical observation. This also suggests that we have not come close to sampling the full dimensionality of intragenotypic behavioral variation; if we were able to add more measurements to the experiment, they too would likely be independent.

To address potential biases in our sampling of assay and measure space, we performed an unsupervised analysis (Berman et al., 2016; Cande et al., 2018) of flies walking spontaneously in arenas. This approach has the potential to identify all the modes of behavior exhibited in that context. Moreover, because the unsupervised algorithm is fundamentally a clustering algorithm, it does not necessarily return a definitive number of clusters/behavioral modes (with more data, it can find increasingly more clusters). Because we can also extract second-order behavioral measures from this approach, such as Markov-transition rates between modes, this approach has the potential to yield a huge number of measures. In the end we were conservative in the number of measurements we chose to work with, matching it to the same order of magnitude as the number of flies we tested. The correlation matrix for the unsupervised behavior measures featured stronger correlations than the distilled Decathlon matrices (Figure 2E). Yet it had a similar effective dimensionality spectrum, indicating many independent dimensions of variation and that not all behavior modes had been sampled (Figure 2F).

Interestingly, the blocks of structure in the correlation matrix aligned, to some extent, with the blocks of structure in the Markov transition matrix of these behavioral measures. This suggests that behaviors mediated by non-overlapping circuits (those that vary independently across individuals) more rarely transition to each other over time. Conversely, behaviors mediated by overlapping circuitry are likely to follow each other sequentially. This may reflect the influence internal states (Calhoun et al., 2019), with an internal state jointly determining what subset of overlapping neurons drives behaviors that are appropriate to string together in succession (e.g., Seeds et al., 2014). We did not assess the day-to-day persistence of behavioral modes identified in the unsupervised analysis, so the observed variation across flies could reflect moods rather than permanent personalities. However, previous supervised (Kain et al., 2013) and unsupervised (Todd et al., 2017) analyses of spontaneous microbehaviors similar to those identified by unsupervised approaches have found such behaviors to persist across days.

We investigated whether individual variation in transcript abundance would predict individual biases on the axes of intragenotypic variation. At the end of the Decathlon, flies were flash frozen and their heads were RNA sequenced. We used linear models to predict PCs of the distilled behavioral matrix from PCs of gene expression variation. This dimensionality reduction was needed to bring the number of predictors in line with the number of observations we had, and meant that we might identify relationships between brain-wide fluctuations in hundreds of genes and behavioral variation. In the case of four behavior PCs, these models outperformed shuffled controls, indicating that some behavioral variation is correlated with expression variation. Gene Ontology analysis revealed fluctuation in the expression of genes involved in respiration was a key predictor in all the significant behavioral models. This may reflect the high metabolic cost of neural tissue, and could reflect individual differences in overall activity, feeding, or metabolic efficiency. Notably, most behavior axes were not predictable from linear models of expression PCs. Variation on those axes may not have its basis in transcriptional variation, or the genes whose transcriptional variation does determine these behaviors may be not expressed in adults, or expressed at low levels or in a small number of specific cells and therefore undetectable in bulk head tissue.

Increasing transcriptional variation by adding genetic variation had the potential to change the correlation structure of behavioral metrics. To our surprise, the distilled correlation matrix of outbred flies was qualitatively similar to that of our original isogenic Decathlon (Figure 5). Both outbred and isogenic matrices were dominated by independent axes of variation and sparse correlations between axes, with rough agreement in the specific pairwise correlations between these two data sets. These two data sets may have differed in their absolute variances (while appearing qualitatively similar in their covariances), but the normalization steps we took to resolve inter-Decathlon and assay batch effects precluded easy assessment of this possibility. That outbred and isogenic variation had qualitatively similar structures raises the possibility that the same kinds of biological fluctuations underlie behavioral variation in populations of each of these kinds.

We finally examined the structure of behavioral variation in collections of flies where variation came from three additional sources: thermogenetic activation of 2,205 sets of neurons across the brain (Robie et al., 2017), optogenetic activation of 176 sparse populations of descending neurons connecting the brain to the motor centers of the ventral nerve cord (Cande et al., 2018), and variation in genotype across 200 inbred strains derived from wild flies (Mackay et al., 2012). Overall, all of these datasets exhibited high degrees of independence in their behavioral measures. The structure of behavioral variation in the neural activation data sets was somewhat different from that of isogenic and outbred flies, with a dimensionality smaller than that of the number of measures, and substantially more organization in lower effective dimensionalities (Figure 6A, B). These data sets also showed clustering of individual flies in behavior space (Figure 6C). Behavioral variation across the inbred strains derived from wild flies was organized qualitatively similarly to the variation across individual flies in isogenic and outbred populations, again suggesting that the biological fluctuations across genotypes mirror those within a genotype as respects the coordination of behavior. This work represents the most complete characterization to date of the structure of behavioral variation within a genotype. We found that there are not discrete types of flies, and there are many independent dimensions of behavioral variation. Moreover, the similar organization of biological variation within and among genotypes suggests that fluctuations in the same biological processes underpin behavioral variation at both of these levels. Elucidating the causal molecular and genet-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ic fluctuations underpinning intragenotypic variation and covariation will be a challenging and revealing future research direction.

Conflict

BdB is a scientific advisor of FlySorter, LLC of Seattle, WA.

Acknowledgements

We thank Ed Soucy, Brett Graham, Adam Bercu and Joel Greenwood of Harvard's CBS Neuroengineering core for help fabricating our instruments. Julie Peng helped with the genomic experiments. Gordon Berman was generous in consulting on the analysis and sharing data. Joshua Shaevitz kindly shared code and expertise for the unsupervised analysis imaging rig. Tanya Wolff and Gerry Rubin kindly shared most of the Gal4 lines from the circuit screen. James Crall and Jennifer Erickson provided helpful feedback on the manuscript. ZW and KSK were supported by NSF Graduate Research Fellowships DGE-1144152 and #2013170544; JFA was supported by the NIH under grant no. GM124881. BdB was supported by a Sloan Research Fellowship, a Klingenstein-Simons Fellowship Award, a Smith Family Odyssey Award, a Harvard/MIT Basic Neuroscience Grant, the NSF under grant no. IOS-1557913, and the NIH under grant no. MH119092.

Methods

Lead Contact and Materials Availability

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Benjamin de Bivort (debivort@oeb.harvard.edu).

Analysis code is available at *https://github.com/de-Bivort-Lab/decathlon* and individual behavioral and gene expression data is available at *http://lab.debivort.org/structure-of-behavioral-variability/*.

Experimental Model and Subject Details

Fly strains

All decathlon experiments were performed on virgin female fruit flies derived from inbred wild type Berlin-K (BSC# 8522) isogenic flies or a custom outbred flies (NEX) (REFEF). Prior to decathlon testing, a lineage of Berlin-K isogenic flies (BK-iso) were selected for robustness from pool of inbred lineages as the result of a short screen designed to mimic the stresses of the decathlon assay battery. All inbred lineages tested were derived from three wild type strains: Berlin-K (BK), Canton-S (CS), and Oregon-R (OR). For each wild type strain, six virgin females were dispensed into separate vials and paired with a single male fly to establish separate lineages. Parental flies were removed from their vials after 2 days to separate them their progeny. For all successive generations after the first generation, the six virgin females were picked from the three vials (2 from each) with the highest number of progeny to select for fecundity and overall ease of maintenance. Virgins were then dispensed into separate vials with full sibling males for inbreeding. After 13 generations of inbreeding this way the resulting lineages (6 total for each strain) were screened for overall activity level via behavioral testing on the Y-maze assay in cohorts of 120 flies from each lineage. For each strain, the lineage with the highest number of choices (turns) in the Y-maze from was kept for further testing to improve overall sampling in behavioral experiments, resulting in a single lineage from each strain: BK-iso, CS-iso, and OR-iso.

To screen robustness to the stresses of multiple days of behavioral testing such as periodic food deprivation and repeated cold anesthetization, cohorts of 96 flies from the resulting inbred lines were singly housed and repeatedly tested on the Y-maze assay over a 3 day period. BK-iso and CS-iso were selected as the lines with lowest and second lowest (respectively) mortality after 3 days of testing. To introduce hybrid vigor and reduce the deleterious effects of inbreeding, four F_1 hybrid isogenic strains were generated by crossing each combination of parental sex and isogenic strain (e.g., ♂ BK-iso x ♀ CS-iso, ♀ BK-iso x ♂ CS-iso etc) and screened for mortality in the 3 day experiment described above. Contrary to our expectation that increased heterozygosity in the F_1 hybrid lines would result in more robust flies and lower mortality, F_1 flies displayed intermediate mortality and activity level (choice number) between BK-iso and CS-iso parentals. Thus BK-iso flies were ultimately selected for the decathlon screen.

Fly handling

Parental flies and decathlon flies (prior to individual housing) were raised on CalTech formula cornmeal media under 12 h/12 h light and dark cycle in an incubator at 25˚C and 70% humidity. A single cohort of virgin decathlon flies were collected over an 8hr period (10AM-6PM) and stored in group housing up to a maximum of 288 flies. On the following day, flies were anesthetized using carbon dioxide (CO2) and were aspirated into custom multiwell plates for individual housing (FlySorter LLC, Seattle WA) on cornmeal media. Individually housed flies were stored in a custom imaging box for circadian activity measurements (circadian chamber) on a 12 h light and dark cycle (10AM-10PM) which was housed inside of an environmentally controlled room at 23˚C and 40% humidity. Food media was replaced and individual housing trays were cleaned every 2 days while flies were in behavioral testing to keep food moisturized and prevent microbial buildup.

Following the start of decathlon (3 days post-eclosion), flies were removed from the circadian chamber each day for behavioral testing between 11AM and 2PM. Prior to all behavioral experiments (excluding olfaction) flies were cold anesthetized by transferring individual housing plates and food trays to an ice-chilled aluminum block in a 4˚C for 10 minutes. This time was minimum time required to reliably induce chill coma in most flies given the mass and low thermal conductivity of the cornmeal media separating the flies and the chill block. Once anesthetized, flies were individually transferred via aspiration to room-temperature custom behavioral arenas (without food)

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

where they quickly awoke (typically 1-10 sec) after transfer. After all flies were transferred, they were given an additional 20 minutes to recover prior to behavioral testing. Flies were then tested in custom behavioral platforms, lasting as little as 15 min (olfaction only) and as much as 2 hrs (typical). Following testing, flies were returned to individual housing via aspiration either directly (olfaction only) or after re-anesthetizing by transferring the custom behavioral arenas to an ice-chilled aluminum block for 10 minutes. Once all flies were returned, the individual housing plates were returned to the circadian chamber between 2-5PM.

Method Details

Custom behavioral experiments

Assay fabrication

Unless otherwise specified, all tracking was conducted in custom imaging boxes constructed with laser-cut acrylic and aluminum rails. Schematics of custom behavioral arenas and behavioral boxes were designed in AutoCAD. Arena parts were laser-cut from black and clear acrylic and were formed as stacked layers joined with Plastruct plastic weld. Arena floors were made from sandblasted clear acrylic and arena lids were made from custom-cut clear eighth inch acrylic. Schematics for behavioral boxes and behavioral arenas can be found on the de Bivort Lab schematics github repository. Illumination was provided by dual-channel white and infrared LED array panels mounted at the base (Part BK3301, Knema LLC, Shreveport, LA). Adjacent pairs of white and infrared LEDs were arrayed in a 14×14 grid spaced 2.2cm apart. White and infrared LEDs were wired for independent control by MOSFET transistors and a Teensy 3.2 microcontroller. Two sand-blasted clear acrylic diffusers were placed in between the illuminator and the behavioral arena for smooth backlighting.

For circadian experiments, flies were housed and imaged in individual fly storage units (FlyPlates) from FlySorter, LLC. Circadian imaging boxes consisted of a small (6”x6”x18”) enclosed behavioral box with a dual-channel LED illuminator on a 12:12hr light and dark cycle programmatically controlled via MARGO behavioral tracking software. A heat-sink was affixed to the underside of the illuminator panel (outside of the box) and fan were used to ensure the interior of the boxes remained consistent with the temperature of the environmental room.

The olfactory sensitivity assay was performed in a custom behavioral chamber (Honegger & Smith et al., 2019) modeled on Claridge-Chang et al., 2009. The apparatus consisted of 15 parallel tunnels constructed from Accura 60 plastic using stereolithography (In’Tech Industries) fabrication. Stainless steel hypo tubing (Small Parts) was used to connect the apparatus with (ID: 0.7mm). Odorized or clean air was delivered via teflon odor tubes to inlet ports at each end of the tunnel and streams vented to the room through exhaust ports in the center forming a sharp choice zone. An active vacuum was not applied to the exhaust ports, and the tunnels operated close to atmospheric pressure throughout the experiment. A clear acrylic lid was clamped in place above the apparatus to ensure an air-tight seal during odor presentation. Air dilutions could be made independently for each side of the apparatus. A custom 15-way PEEK manifold was used on each side to split the odorized flow equally between 15 tunnel inlets. A final valve (SH360T041; NResearch) was used immediately upstream of each manifold to quickly switch between pure dehumidified air and the odorized stream.

Phototactic stimuli were delivered with a custom 12”x12” PCB designed in Express PCB CAD software and manufactured by ExpressPCB. Briefly, the PCB was designed to power and independently control 216 white light LEDs (11.34mW±4.3µW at 20ma) geometrically arranged to match the maze arm ends of an array of Y-shaped behavioral arenas. LED intensity control was provided over USB via board interface with a teensy 3.2 microcontroller and constant current driver boards (Adafruit 24-Channel TLC5947 LED Driver). Holes in the footprint shape of each individual arena were custom cut into the PCB with water jet cutter to provide infrared backlighting.

Visual stimuli in the spatial shade-light, temporal shade-light, and optomotor assays were delivered in a behavior box modified to accommodate stimulus delivery with an overhead projector (Optoma S310e DLP). The DLP color wheel was removed reduce apparent flicker to the flies caused by low sensitivity to red light in flies. Stimuli were accurately targeted to individual flies by creating a registration mapping between the projector and the tracking camera using previously described method using MARGO tracking software. Briefly, a 2D polynomial registration model of a projector targeting coordinates of a small (10px) was fitted to coordinates obtained by tracking the dot as it was rastered over the projector display field. All visual stimuli were crafted and displayed with PsychToolbox.

Single fly videos used in the motionmapper unsupervised classification pipeline were acquired in behavioral boxes as described above. We replaced the dual channel LED illuminator with an LCD screen backlight to reduce apparent non-uniformities in the videos. Behavioral arenas consisted of custom thermoformed clear PTEG lid with sloping sides to prevent wall walking. Lids were coated with sigmacote (Sigma-Aldrich SL2-100ML) to prevent flies from walking on the ceiling and were clamped in place to a 0.25” tempered glass base via custom acrylic 2x2 array of arenas.

Behavioral tracking

Flies were imaged with overhead tracking cameras at varying resolutions and frame rates. Unless otherwise specified, images for behavior tracking were acquired at 10Hz with a 1280 x 1024 pixel BlackFly GigE camera (PointGrey BFLY-U3-13S2M-CS) fitted with a Fujinon YV2.8×2.8SA-2, 2.8mm-8mm, 1/3", CS mount lens. Circadian and optomotor experiments were conducted with 3Hz and 60Hz imaging respectively. Tracking images for the odor sensitivity assay were acquired at 1328 x 1048 pixel at 23Hz (PointGrey FMVU-13S2C). Single fly videos for unsupervised classification were acquired.

Fly tracking and stimulus control for all experiments (excluding unsupervised single fly imaging) were programmed in MATLAB and implemented with MARGO (Werkhoven et al., 2019). The MARGO tracking algorithm has been previously described in detail. Briefly, binary foreground blobs were segmented from a thresholded difference image computed by subtracting each frame from an estimate of the background. A rolling model of

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the background was computed as the median image of rolling stack of background sample images. Every 2min, a new sample image of the background was acquired and replaced the oldest image in the stack. Regions of interest (ROIs) were defined to encompass only a single behavioral arena (i.e., one fly), and the tracking algorithm proceeded independently for each ROI. The tracking threshold used to segment binary foreground blobs from the background was manually set prior to tracking and was independent for each experiment. Blobs below a minimum or maximum area (also defined manually for each experiment) were excluded from tracking in each frame. After filtering, the centroid trace of each ROI was updated with the position of the blob with shortest distance to the last known (non-NaN) position of each trace. The centroid trace of any ROI with no detected blobs in any given frame received no position (i.e., NaN). Output measurements of speed and area were converted from pixels/sec and pixels^2 to mm/sec and mm^2, respectively, by measuring the length a known landmark size (e.g., arena diameter) prior to tracking. Fisheye distortion of camera lenses were modeled and corrected via MATLAB's camera calibrator app.

For the acquisition of single fly videos used in the motionmapper unsupervised classification pipeline, flies were tracked in real-time to reduce video file size by extracting a 200x200 pixel region around the fly centroid. Tracking performed by computing the centroid of the entire difference image acquired calculated via the method described above. Tracking was coordinated simultaneously for four 1280x1024 pixel resolution cameras (PointGrey CM3-U3-13Y3M-CS) at 100Hz with a custom LabView script obtained from the J. Shaevitz lab (personal communication) which was lightly modified to add background subtraction to the tracking.

Decathlon experimental design

Flies underwent behavioral testing from 3 days post-eclosion up to a maximum of 14 days post-eclosion depending on when unsupervised classification was performed. For each decathlon experiment, behavioral testing began with activity screen of 288 flies in the Y-maze assay. Flies were then sorted by number turns made during the activity screen. The top 192 flies with the highest activity level were selected for testing in the remainder of the decathlon and were transferred to individual housing on day 3 post eclosion for overnight circadian behavior profiling. On subsequent days, all flies were anesthetized as described above and transferred from individual housing into a behavioral assay each day (from days 3-11 post-eclosion) and were returned individual housing for overnight circadian behavioral measurement. Flies were tested in cohorts dependent on the throughput capacity (described for each assay below) of the assay run on any given day. Although the range of testing times for all animals on some days was relatively broad (all experiments were conducted between 11AM-5PM), no behavioral experiment lasted more than 2hrs in duration, meaning that individual cohorts typically spent more than 3h (2h experiment + 2x anesthetization and transferring) off of food media. The order of individual assays (shown in figure 1) was randomized at the start of each decathlon with the exception of activity screening and unsupervised imaging which occurred at the beginning and end of each decathlon (respectively) for logistical feasibility. Descriptions and implementation details of the decathlon behavioral assays are provided below and are summarized in supplemental Table 1.

After day 11 post-eclosion, remaining living flies were split into three roughly equal-sized cohorts (approximately 56 flies each) for unsupervised behavioral imaging which were imaged in groups of 4 flies between 10AM-6PM on subsequent days. Cohorts of flies were flash-frozen on liquid nitrogen immediately following unsupervised behavioral imaging and prepped for RNA-seq as described above.

Arena circling and circadian assays

Both the arena circling and circadian assays consisted of exploration of a circular arena. Although a near identical list of measures of locomotor handedness, speed, and movement bout dynamics were recorded in each assay (circadian includes gravitaxis also), the assays are distinguished by their background light level, duration, behavioral arena, and access to food. The arena circling assay was conducted over 2 h with constant light in a wide (28 mm diameter) and shallow (1.6 mm depth) arena without food. The circadian assay was conducted overnight (20-21 h) on a 12:12 h light and dark cycle (10AM:10PM) with each fly in a narrow (6.8 mm) and deep (10.6 mm) well of a 96 multiwell plate. The depth of the arena in addition to the difference in arena ceiling (sigmacoted acrylic in arena circling and uncoated plastic mesh circadian) add a vertical and body orientation dimension to the circadian assay not present in arena circling. This feature not only adds a measure of floor vs. ceiling preference (i.e., gravitaxis) to the circadian assay but also potentially confounds measures of circling directionality due to the fact that flies can circle the arena right-side up on the floor, upside-down on the ceiling, or sideways on the walls.

LED Y-maze assay

Individual flies explored symmetrical Y-shaped arenas with LEDs at the end of each arm. For all arenas in parallel, real-time tracking detected which arm the fly was in at each frame. At the start of each trial, an LED was randomly turned on in one of the unoccupied arms. Once flies walked into one of these two new arms, all the LEDs in the arena were turned off and a new trial was initiated by randomly turning on an LED in one of the newly unoccupied arms. This process is repeated for each fly independently over 2 h. Turns were scored as positively (toward a lit LED) or negatively (toward an unlit LED) phototactic. LEDs were either fully off (OFF) or were driven at maximum intensity (ON). No intermediate LED intensity values were tested. In addition to the phototactic bias, measures of locomotor handedness peed, choice timing, and choice sequence were also recorded.

Odor sensitivity assay

In the odor sensitivity assay, flies explored linear tunnels where half of the tunnel was contained an odor (typically noxious) and the other contained no odor. Odor ports positioned at either end of the tunnels delivered either odor or odorless air. Vents positioned at the midpoint of each tunnel form a sharp boundary between the two air streams. Flies were presented with a 3min baseline period with no odor on either side followed by a 12min experimental period with half odor and half no odor. Sensitivity to the odor was estimated by measuring the fraction of occupan-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

cy time spent in the odor half of the tunnel. Flies showed a weak aversion to the designated odor side during baseline measurement, prior to the presentation of odor (0.47±0.30 mean pre-odor occupancy). Flies displayed strong initial aversion (0.17±0.22 mean first half odor occupancy) to the side of the tunnels with the odor which slightly diminished (0.22±0.23 mean second half occupancy) over the stimulus duration. In addition to odor occupancy, locomotor handedness was scored as the average direction of heading direction reversals on the long axis of the tunnels. In addition to the above measures, of speed, turn timing, and turn direction sequence were also recorded.

Optomotor assay

In this assay, optomotor stimuli were centered on the bodies of flies by projecting them onto the floor of their arenas in which they are walking freely. These stimuli consisted of a maximally-contrasting (black = 0, white = 1, min. power = 79µW±61nW, max. power = 2.4mW±4.7µW) rotating pinwheel (spatial frequency = 0.18 cycles/deg, rotational speed = 320 deg/sec) and typically evoked a turn in the direction of the rotation to stabilize the visual motion. The pinwheel center followed the position of the fly as it moved so that the only apparent motion of the stimulus is rotational motion around the body. Therefor, the stimulus was closed-loop with respect to position and open-loop with respect to rotation velocity. We observed that optomotor responses could be reliably elicited, provided individuals were already moving when the pinwheel was initiated. We therefore only presented the pinwheel when: 1) flies were moving 2), a minimum inter-trial interval (2s) had passed to prevent behavioral responses from adapting, and 3) flies exceeded a minimum distance from the edge of the arena to ensure that the stimulus occupied a significant portion of the animal's FoV. Over 2hrs of testing, an optomotor index was calculated for each fly as the fraction (normalized to [-1,1]) of body angle change that occurred in the same direction as the stimulus rotation over the duration of the stimulus. On average, flies displayed reliable optomotor responses (mean = 0.49±0.16) when stimulated. In addition to optomotor index measures of fly activity, movement bout dynamics, and locomotor handedness were also recorded.

Shade-light assay

Ambient light preference is estimated in this assay by projecting a circular stimulus with one side fully bright (intensity = 1) and the other side fully dark (intensity = 0), with the two regions separated by a small (10% of the arena diameter) boundary zone of intermediate brightness (intensity = 0.5) (min. power = 79µW ±61nW, max. power = 2.4mW±4.7µW). A light occupancy measure was calculated for each fly as the fraction of time spent in the lit region divided by the total time spent in both the lit and unlit regions. Time spent in the intermediate boundary zone was scored as no preference and was excluded from the calculation. Each 2 hr experiment was divided into 15 stimulus cycles with each cycle consisting of an alternating 4min baseline block where each arena was unlit and a 4min experimental block where shade-light stimuli were targeted to each arena. At the start of each block, each stimulus was rotated to center the intermediate zones (or virtual intermediate zone in the baseline block) over the flies bodies. In effect, this required flies to move from intermediate zone during the block to express any preference. In addition to light occupancy measures of fly activity, movement bout dynamics, and locomotor handedness were also recorded.

Temporal light-shade assay

Ambient light preference is estimated in this assay by projecting a fully bright (intensity = 1) or fully dark (intensity = 0) stimulus to the entire arena when flies crossed an invisible virtual boundary in the center (min. power = 79µW±61nW, max. power = 2.4mW±4.7µW).. Therefore, this stimulus paradigm was distinct from the spatial light-shade assay in that flies could not express a preference by navigating with respect to an apparent spatial pattern of light. A light occupancy measure was calculated for each fly as the fraction of time flies received the lit stimulus divided by the total duration of the experiment (2hrs). The invisible boundary zone was always positioned in the center of the arena such that a randomly exploring fly would receive both stimuli in roughly equal amounts. The angular orientation of the boundary zone was randomly initialized for each arena independently at the start of the experiment. To avoid rapid switching or flickering of the stimulus for flies sitting directly on the boundary, flies were required to cross a small buffer zone (5% of the arena diameter) beyond the arena center into the other zone before switching the stimulus.

Y-maze assay

In the Y-Maze assay, locomotor handedness was estimated as flies explored symmetrical Y-shaped arenas. Each time flies changed position from one arm of the maze to another turns were scored as left-handed or right-handed depending on whether the chosen arm was to the left of right on the choice point (i.e., the arena center). To avoid scoring centroid estimation errors around the choice point or small forays into new arms as turns, the flies were required to traverse a minimum distance (60% the length of the arm) into any arm before a turn was scored. In addition to locomotor handedness, measures of speed, movement bout dynamics, turn choice timing, and turn choice sequence were also recorded.

Central complex thermogenetic screen

We extracted measure of turn timing clumpiness (choiceClumpiness) and turn direction switchiness (handSwitchiness) from a previously conducted thermogenetic screen for light dependent modulation of locomotor handedness (kakaria_preprint_2019). Gal4 driver and neuronal effector parental lines were crossed to generate F_1 progeny with sparse neuronal expression of effectors that could be thermogenetically activated (dTrpA1) or silenced ($Shibire^{ts}$). For all screen experiments, fly locomotor handedness was assayed for 4 hrs in the Y-Maze with the following temperature program: 1 h at 23°C (permissive temperature), 1 h temperature ramp up to the restrictive temperature ($Shibire^{ts}$ = 29°C, dTrpA1 = 32°C), 1 h at the restrictive temperature, and 1 h temperature ramp down from the restrictive temperature to 23°C. Although activity level varied over the duration of the experiment, flies made many turns throughout all temperature blocks: completing an average of 193±119 and 757±311 turns in the permissive and restrictive periods respectively. The background light was repeatedly switched on and off in time blocks (min = 5 seconds, other times = 10 and 30, max = 60 seconds , each hour

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

had the same repeated sequence), in a random order shared between all screen experiments. Because we were interested only in temperature dependent modulation of turn dynamics, we computed individual clumpiness and switchiness scores for all turns within each temperature condition (permissive and restrictive) regardless of the light condition.

RNAseq and transcriptomic analyses

Single fly RNAseq preparation and sequencing was performed using TM3'seq, a high-throughput low-cost RNA protocol previously described in (Pallares et al., 2019). An overview of the method is provided below, but more detailed descriptions are available in an online protocol (*https://lufpa.github.io/TM3Seq-Pipeline/tm3seq_protocol*) and in the original publication.

Fly tissue preparation

Following decathlon behavioral testing, flies were briefly anesthetized on CO_2 and transferred to a 96-well plate via aspiration. Immediately following, flies were flash frozen on liquid nitrogen and were decapitated to separate heads and bodies. Well plates were then transferred to a dry ice and ethanol cold bath to keep the samples frozen while heads were individually transferred to a separate 96-well plate with a cold probe needle. Samples were then stored at -80°C. After storage, tissue was ground in 100µl of lysis buffer with a 2.8mm stainless steel grinding bead steel grinding bead for 10 minutes at maximum speed with a homogenizer. CyBio® FeliX liquid handling robot (Analitik Jena) was used to perform mRNA extraction from the resulting lysate using a Dynabeads™ mRNA DIRECT™ Purification Kit (ThermoFisher, #61012) and a custom protocol (REFREF) optimized for low cost, yielding approximately 10 to 20 ng mRNA per head.

Library preparation

RNA (10µl of 1ng/µl mRNA) was added to 1µl of 0.83uM oligo (Tn5ME-B-30T) for a 3 minute incubation at 65°C immediately prior to reverse transcription. The first strand of cDNA was synthesized by reverse transcription of the mRNA via a 1hr incubation at 42°C by adding the following to the reaction mixture above: 1µl SMARTScribe™ RT (Takara, #639538), 1µl dNTPs 10mM (NEB, #N0447S), 2µl DTT 0.1M (Takara, #639538), 4µl 5× First-Strand buffer (Takara, #639538), and 1µl B-tag-sw oligo. Following synthesis, the reverse transcriptase was inactivated via a 15 minute incubation 70°C. A cDNA amplification mixture was prepared by adding5µl of the resulting first strand cDNA with 7.5µl of OneTaq HS Quick-load 2× (NEB, #M0486L) and 2.5µl water. The cDNA was then amplified in a thermocycler via the following program: 68°C 3min, 95°C 30sec, [95°C 10sec, 55°C 30sec, 68°C 3min] *3 cycles, 68°C 5min.

Tn5 tagmentation was used along side universal adaptors for library amplification. The adapter-B was previously added during synthesis of the first cDNA strand. To create a Tn5 adaptor-A, an adapter annealing mixture was prepared by adding 10µl (100µM) of a forward oligo (adapter-A) and 10µl (100µM) reverse adapter-A oligo (Tn5MErev) to 80µl re-association buffer (10mM Tris pH 8.0, 50mM NaCl, 1mM EDTA). Oligos were annealed in a thermocycler with the following cycle program: 95°C for 10 minutes, 90°C for 1 minute followed by 60 cycles reducing temperature by 1°C/cycle, hold at 4°C. 5µl of 1µM annealed adapter was then anneal to 5µl of Tn5 in a thermal cycler for 30min at 37°C. 5µl of cDNA was mixed with 1µl of pre-charged Tn5. The adapter-A loaded Tn5 was diluted 7× in re-association buffer: Glycerol (1:1)., 4µl of TAPS buffer 5× pH 8.5; 50mM TAPS, 25mM MgCl2, 50% v/v DMF), and 5µl of water. The mixture was then and incubated for 7min at 55°C followed by an additional 7 minute incubation with 3.5µl of SDS 0.2% (Promega, #V6551) to ensure that Tn5 was dissociated from the cDNA. The resulting cDNA libraries were then amplified. Briefly, 10µl of OneTaq HS Quick-Load 2x (NEB, #M0486L), 1µl i5 primer 1uM, 1µl i7 primer 1µM, and 7µl of water were used to amplify 1µl of the tagmentation reaction following the program: 68°C 3min, 95°C 30sec, [95°C 10sec, 55°C 30sec, 68°C 30sec] *12 cycles, 68°C 5min.

Sequencing and expression quantification

Samples were sequenced on an Illumina HiSeq 2500 in separate runs, using dual indexes and single-end ~100 bp sequencing. Low quality bases and adapter sequences were removed from reads using Trimmomatic 0.32 (SE ILLUMINACLIP:1:30:7 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN: 20). Downstream analysis was only performed on reads at least 20 nucleotides after trimming. Reads were then mapped to the r6.14 *Drosophila melanogaster* genome assembly. Read counts were further filtered to include only reads assigned to protein coding genes with a minimum of 500k reads per individual (n.b. measuring 3' biased expression).

Genomic sequencing of inbred flies

DNA was extracted from individual flies using Zymo Quick-DNA Microprep Kit (Cat# D3020). Sequencing libraries were prepared using the transposase Tn5, using the same protocol described above (from cDNA). Samples were sequenced on an Illumina HiSeq 2500, using dual indexes, paired-end ~150 bp sequencing. After de-multiplexing, low quality bases and adapter sequences were removed from reads using Trimmomatic 0.32 (SE ILLUMINACLIP:1:30:7 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:20). Reads were then mapped to the r6.14 *Drosophila melanogaster* genome assembly using BWA -aln (Li amd Durbin, 2009). Genotypes were generated following GATK best practices (*https://software.broadinstitute.org/gatk/best-practices/*) (McKenna et al., 2010). Heterozygosity was estimated using VCFTools v0.1.16 --het (Danecek et al., 2011).

Quantification and statistical analyses

Merging Decathlon data sets

Inspection of raw metric distributions separated by batch (i.e. a single cohort of flies tested in a single behavioral box at the same time) showed batch effects on the sample means and dispersion even within the same decathlon experiment. Therefore, data was z-scored separately by batch and metric type to control for sample differences in mean and variance. Data of the same assay, metric, and day of testing (i.e. the unique measures from a single day of decathlon behavior) were then combined, resulting in an initial *nxd* data matrix for a single decathlon experiment where *n* is the number of individuals and *d* is the number of met-

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

rics. The matrix contained substantial fraction of missingness (mean=0.23±0.22) within metrics after initial construction. We used alternating least squares (ALS) to estimate a complete matrix that preserved, as accurately as possible, the covariance structure of the underlying data (Figure S4). To reduce run to run variation in ALS, we generated 200 complete data matrices with ALS and computed a final complete median matrix of all repetitions. Simulations with ground truth data (Figure S4) with covariance structure similar to that of the decathlon data sets showed that this process resulted in both lower error in the resulting matrices and lower run to run variation in the final matrix. This process resulted in three complete matrices: two *nxd* matrices for the two BK-iso decathlon experiments and one *nxd* matrix for the NEX decathlon experiment.

To combine the decathlon-1 and decathlon-2 matrices for BK-iso, we *z*-scored by behavioral metric within each matrix to adjust for decathlon experiment batch effects on mean and variance. Data from the two matrices were then combined by matching unique assay and metric combination. Because the order of assays was randomized for each decathlon experiment, day of testing was ignored when combining behavioral metrics from all non-circadian assays (e.g. olfaction odorOccupancy from day 7 of decathlon-1 was combined with olfaction odorOccupancy from day 8 of decathlon-2). Circadian metrics were matched by day of testing due to circadian measurements being collected on all days of testing (e.g. circadian meanSpeed from day 1 of decathlon-1 was combined with circadian meanSpeed from day 1 of decathlon-2). Placeholder NaN values were inserted in cases where no matching metric existed in the other matrix (e.g. temporal phototaxis for decathlon-1). The resulting full data matrix was then z-scored by metric and any residual missing values (due to lack of an existing metric match) were then infilled via the ALS method described above.

Distilled matrix generation

We created a distilled matrix to condense any covarying features for which we had a prior expectation that the metrics might be correlated for obvious or uninteresting reasons. We defined such groups of measures (a priori groups) which primarily consisted of either duplicate measures (e.g., all 10 days of circadian gravitaxis) or measures that are likely to be linked by the same underlying phenomenon (e.g., choiceNumber and meanSpeed across multiple assays). Details of the metrics in each a priori group are detailed in supplemental table 2. We performed PCA on each a priori group separately in an attempt to capture group variance with fewer dimensions. We defined an adaptive cutoff for the number of principal components (PCs) retained from each group as the highest principal component above or within the 95% confidence interval of the variance explained by PCs computed on a shuffled version of the same matrix. We reasoned that PCs above or within the variance explained for PCs fit on the shuffled matrix (a matrix with approximately independent features) represented components fit to meaningful covariance of metrics or components primarily aligned single independent metric.

Multidimensional analyses

We estimated effective dimensionality of the decathlon metrics as the number of connected components in a thresholded covariance matrix represented as a directionless graph. We swept 200 threshold values uniformly spaced between the minimum and maximum covariance. A distribution of the effective dimensionality was formed by iteratively counting number of connected components in the graph at each threshold value. To estimate a lower limit on the dimensionality and to assess the degree to which the number of connected components was dependent on specific groups of metrics, we repeated the steps above on matrices after randomly selecting of metrics to drop from the matrix, iteratively increasing the number of metrics dropped from 1 to *d*-1 (i.e., one feature remaining).

Unless otherwise specified, t-SNE embeddings of all data sets was performed on the euclidean pairwise distances of *z*-scored values with perplexity = 20. All embeddings were optimized by minimizing the KL-divergence between the original and embedded data. t-SNE performed on decathlon metrics was run with perplexity = 8 where we had an expectation that clusters would be relatively small (e.g., 3-8 data points) due to the low number of measures in our a priori groups (median=5) and low number of unique measures from each assay (median=8) and day of testing.

Unsupervised behavioral classification

As previously described (Berman et al., 2014), single fly videos were decomposed into behavioral classification time series with the motion mapper pipeline. Briefly, 200x200px frames centered on the flies were translationally and rotationally aligned with subpixel accuracy to a template fly body to restrict frame to frame variation to postural changes by the flies. The data dimensionality was reduced by restricting further analysis to the 6,700px with the highest variance. The data for all individuals was further compressed with PCA into 50 eigenmode postural time series. Principal component time series were then spectrally decomposed into 25 uniformly spaced frequency channels via Morlet wavelet transformation resulting in a high-dimensional (1,250) representation of each frame at various timescales. A representative sub-sampled training set of frames was constructed for each individual by embedding their data into two dimensions with t-SNE (van de Maaten and Hinton, 2008) and selecting frames according to proportionally to their local probability density in the embedded space. A joint two-dimensional embedding for all individuals data was then constructed by embedding the combined training set via t-SNE. Individual embeddings were generated by projecting all remaining data points into the joint embedding. As previously described, the distribution of log-speed trajectories within the embedded space was well-described by a two-component gaussian mixture model (GMM) with the majority of frames falling into a low speed mode. We used the standard deviation of the low speed GMM component to define a small gaussian kernel (σ=1.37) that was convolved with the embedded points the compute a continuous density. The embedded space was then segmented into discrete regions by computing watersheds (Najman and Schmitt, 1994) of the negative density. Behavioral classifications were then assigned to frames by the watershed region occupied. Frames were filtered by defining an embedded speed threshold (2.78) as the intersection of GMM components where the log-speed distribution was maximally separable. All frames above speed threshold received no classification. Individual occupancy within each classification

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

was then used to compute a discrete probability density function for each individual.

Human readable labels were applied to each classification by generating 8x8 tiled movies of frames corresponding to individual pauses within each classification above a minimum duration of 10 ms. Labels were created by scoring which behaviors appeared to be frequently represented after repeated viewing. The labels were composed with the following format: body part (e.g. ,wing, forelimb, abdomen), the behaviors displayed (e.g., walking, grooming, movement), and a qualitative descriptor of the speed of movement (e.g., idle, slow, fast).

Statistics

Unless otherwise stated, all reported correlations were computed as the Spearman rank correlation. *P*-values reported for correlation coefficients were calculated via two-tailed *t*-test of the null hypothesis that the regression coefficient was not significantly different from zero.

To generate bootstrapped distributions of correlation the correlation matrices, decathlon matrices (either the full or distilled matrices) were sorted to match metrics across decathlon data sets as described above. A pair of bootstrapped matrices were created by bootstrapping individuals from either the same (e.g., resampling BK-iso matrix twice) or different decathlon matrices up to the size of the original matrix. The correlation matrix was computed for each bootstrapped matrix and the unique, off-diagonal *r*-values of each matrix were stored. A single correlation of correlations was then calculated on the two sets of *r*-values. The above steps were then repeated over 100 repetitions for all unique combinations of decathlon data sets.

False discovery rate

We computed the false discovery rate (FDR) for significance of metric correlations as a function of α-value for each combination of decathlon data set (BK-iso or NEX) and matrix type (full or distilled). Each data matrix was bootstrap resampled (100 repetitions) up to the size of the original matrix. Correlation matrices were computed for each matrix and p-values were calculated correlation coefficient (see statistics). We then calculated a mean kernel density estimates for all p-values of unique metric combinations across bootstrap replicates. The above steps were then repeated for bootstrap shuffled data matrices to create a null distribution of p-values for correlation coefficients. We defined false discovery rate as the shuffled p-value density divided by the observed p-value density. For determining significance of correlation, we set α=0.05, corresponding to the following FDR: BK-iso full = 0.30, NEX full = 0.30, BK-iso distilled = 0.37, NEX distilled = 0.42.

Acknowledgements

J. Akhund-Zade, S. Ho, C. O'Leary, and B. de Bivort. The effect of environmental enrichment on behavioral variability depends on genotype, behavior, and type of enrichment. *Journal of Experimental Biology*, 2019.

J. F. Ayroles, S. M. Buchanan, O. Chelsea, S. Kyobi, J. K. Grenier, A. G. Clark, D. L. Hartl, and B. L. de Bivort. Behavioral idiosyncrasy reveals genetic control of phenotypic variability. *Proc National Acad Sci*, 112(21):6706–6711, 2015.

F. Bartumeus, J. Catalan, U. Fulco, M. Lyra, and G. Viswanathan. Optimizing the encounter rate in biological interactions: Lévy versus brownian strategies. *Phys.Rev. Lett.*, 88(9):097901, 2002.

M. Berdal and N. A. Dochtermann. Adaptive alignment of plasticity with genetic variation and selection. *Journal of Heredity*, 110(4): 514-21. 2019.

G. J. Berman, D. M. Choi, W. Bialek, and J. W. Shaevitz. Mapping the stereotyped behaviour of freely moving fruit flies. *Journal of The Royal Society Interface*, 11(99): 20140672–20140672, 2014.

G. J. Berman, W. Bialek, and J. W. Shaevitz. Predictability and hierarchy in *Drosophila* behavior. *Proc National Acad Sci*, 113(42): 11943–11948, 2016.

D. Bierbach, K. Laskowski, and M. Wolf. Behavioural individuality in clonal fish arises despite near-identical rearing conditions. *Nature communications*, 8:15361. 2017.

M. Bruijning, C. J. E. J. Metcalf, E. Jongejans, and J. F. Ayroles. The evolution of variance control. *Trends Ecol. Evol.* 2019. doi: 10.1016/j.tree.2019.08.005.

S. M. Buchanan, J. S. Kain, and B. L. de Bivort. Neuronal control of locomotor handedness in *Drosophila*. *Proc National Acad Sci*, 112(21):6700–6705, 2015.

A. Calhoun, J. Pillow, and M. Murthy. Unsupervised identification of the internal states that shape natural behavior. *bioRxiv*, 2019. doi: 10.1101/691196.

J. Cande, S. Namiki, J. Qiu, W. Korff, G. Card, J. W. Shaevitz, D. L. Stern, G. J. Berman. Optogenetic dissection of descending behavioral control in *Drosophila*. *eLife*, 2018. doi: 10.7554/eLife.34275.

A. Charmantier, D. Garant, and L. Kruuk. *Quantitative genetics in the wild*. 2014. doi:10.1093/acprof:oso/9780199674237.001.0001.

A. Claridge-Chang, R. Roorda, E. Vrontou, L. Sjulson, H. Li, and G. Miesenböck. Writing memories with light-addressable reinforcement circuitry. *Cell*, 139(2):405-415. 2009.

P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin and 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics*, 27(15): 2156-8. 2011.

N. A. Dochtermann, T. Schwab, M. Anderson Berdal, J. Dalos, and R. Royauté. The heritability of behavior: A meta-analysis. *J. Hered.*, 110(4):403–410, 2019.

J. Freund, A. M. Brandmaier, L. Lewejohann, I. Kirste, M. Kritzler, A. Krüger, N. Sachser, U. Lindenberger, and G. Kempermann. Emergence of individuality in genetically identical mice. *Science*, 340(6133):756–759, 2013.

R. L. Goldberg. The structure of phenotypic personality traits. *American Psychologist*, 48(1): 26-34. 1993.

J. Green, A. Adachi, K. K. Shah, J. Hirokawa, P. Magani, and G. Maimon. A neural circuit architecture for angular integration in *Drosophila*. *Nature*, 546: 101–106. 2017.

L. Grinsted, J. N. Pruitt, V. Settepani, and T. Bilde. Individual personalities shape task differentiation in a social spider. *Proc Biological Sci Royal Soc*, 280(1767):20131407, 2013.

S. A. Haddad and E. Marder. Circuit robustness to temperature perturbation is altered by neuromodulators. *Neuron*, 100(3):609-623.e3 2018.

F. Hamada, M. Rosenzweig, K. Kang, and S. R. Pulver, A. Ghezzi, T.J. Jegla, and P. A. Garrity. An internal thermal sensor controlling

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

temperature preference in *Drosophila*. *Nature*, 454(7201):217-20. 2008.

K. Honegger and B. de Bivort. Stochasticity, individuality and behavior. *Current Biology*, 28(1):R8-R12. 2018.

K. Honegger, M. A.-Y. Smith, M. A. Churgin, G. C. Turner, and B. L. de Bivort. Idiosyncratic neural coding and neuromodulation of olfactory individuality in *Drosophila*. *Proc National Acad Sci*, 2019. doi: 10.1073/pnas.1901623116.

K. Hopper. Risk-spreading and bet-hedging in insect population biology. *Annu. Rev. Entomol.*, 44(1):535–60, 1999.

F. Huntingford. The relationship between anti-predator behaviour and aggression among conspecifics in the three-spined stickleback, gasterosteus aculeatus. *Animal Behaviour*, 24(20: 245-60. 1976.

R. Jeanson and A. Weidenmüller. Interindividual variability in social insects - proximate causes and ultimate consequences. *Biol Rev*, 89(3):671–687, 2014.

A. Jenett, G. Rubin, T. Ngo, D. Shepherd, C. Murphy, H. Diomme, B. D. Pfeiffer, A. Cavallaro, D. Hall, J. Jeter, N. Iyer, D. Fetter, J. H. Hausenfluck, H. Pen, E. T. Trautman, R. R. Svirkas, E. W. Meyers, Z. R. Iwinski, Y. Aso, G. M. DePasquale, A. Enos, P. Hulamm, S. C. Lam, H. H. Lo, T. R. Laverty, F. Long, L. Qu, S. D. Murphy, K. Rokicki, T. Safford, K. Shaw, J. H. Simpson, A. Sowell, S. Tae, Y. Yu and C. T. Zugates. A GAL4-driver line resource for *Drosophila* neurobiology. *Cell reports*, 2(4):991-1001. 2012.

J. Kain, C. Stokes, Q. Gaudry, X. Song, J. Foley, R. Wilson, and B. de Bivort. Leg-tracking and automated behavioural classification in *Drosophila*. *Nature Communications*, 4(1), 2013.

J. Kain, C. Stokes, and B. de Bivort. Phototactic personality in fruit flies and its suppression by serotonin and white. *Proc National Acad Sci*, 109(48):19834–19839, 2012.

J. Kain, S. Zhang, A. Jamilla, A. D. Samuel, M. Klein, and B. de Bivort. Variability in thermal and phototactic preferences in *Drosophila* may reflect an adaptive bet-hedging strategy. *Evolution*, 69(12): 3171–3185, 2015.

K. Kakaria and B. de Bivort. Ring attractor dynamics emerge from a spiking model of the entire protocerebral bridge. *Frontiers in behavioral neuroscience*, 11:8. 2017.

T. Kitamoto. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *Journal of neurobiology*, 47(2):81-92. 2001.

B. Kottler, R. Faville, J. Bridi, and F. Hirth. Inverse control of turning behavior by dopamine d1 receptor signaling in columnar and ring neurons of the central complex in *Drosophila*. *Current Biology*, 29(4):567-577.e6. 2019.

J. Krakauer, A. Ghazanfar, and A. Gomez-Marin, M. A. MacIver, D. Poeppel. Neuroscience needs behavior: correcting a reductionist bias. *Neuron*, 93(3):480-490. 2017.

H. Li and R. Durbin. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):1754-60. 2009.

Y. Lin, Z. Chen, B. Oliver, S. Harbison. Microenvironmental gene expression plasticity among individual *Drosophila melanogaster*. *G3: Genes*, 2016

N. Linford, C. Bilgir, J. Ro, and S. D. Pletcher. Measurement of lifespan in *Drosophila melanogaster. JoVE (Journal of Visualized Experiments*, (71), e50068. 2013.

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

G. Liu, T. Nath, G. A. Linneweber, A. Claeys, Z. Guo, J. Li, M. Bengochea, S. Backer, B. Weyn, M. Sneyders, H. Nicasy, P. Yu, P. Scheunders, and B. A. Hassan. A simple computer vision pipeline reveals the effects of isolation on social interaction dynamics in *Drosophila*. *Plos Comput Biol*, 14(8):e1006410, 2018.

A. Long, S. J. Macdonald, and E. G. King. Dissecting complex traits using the *Drosophila* synthetic population resource. *Trends in Genetics*, 30(11):488-95. 2014.

M. Lynch and B. Walsh. *Genetics and analysis of quantitative traits.* Sinauer, 1998.

L. van der Maaten and G. Hinton.Visualizing data using t-SNE. *Journal of machine learning research*, 9:2579-2605. 2008.

T. Mackay, S. Richards, E. Stone, A. Barbadilla, J. F. Ayroles, D. Zhu, S. Casillas, Y. Han, M. M. Magwire, J. M. Cridland, M. F. Richardson, R. R. H. Anholt, M. Barrón, C. Bess, K. P. Blankenburg, M. A. Carbone, D. Castellano, L. Chaboub, L. Duncan, Z. Harris, M. Javaid, J. C. Jayaseelan, S. N. Jhangiani, K. W. Jordan, F. Lara, F. Lawrence, S. L. Lee, P. Librado, R. S. Linheiro, R. F. Lyman, A. J. Mackey, M. Munidasa, D. M. Muzny, L. Nazareth, I. Newsham, L. Perales, L.-L. Pi, C. Qu, M. Ràmia, J. G. Reid, S. M. Rollman, J. Rozas, N. Saada, L. Turlapati, K. C. Worley, Y.-Q. Wu, A. Yamamoto, Y. Zhu, C. M. Bergman, K. R. Thornton, D. Mittelman, and R. A. Gibbs. The *Drosophila melanogaster* genetic reference panel. *Nature*, 482:173–178. 2012.

G. Matthews, I. Deary, and M. Whiteman. *Personality traits*. Cambridge University Press. 2003.

A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Giramella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9): 1297-1303. 2010.

D. Mellert, W. Williamson, T. Shirangi, and G. M. Card, and J. W. Truman. Genetic and environmental control of neurodevelopmental robustness in *Drosophila*. *PloS one*, 11(5):e0155957. 2016.

L. Najman and M. Schmitt. Watershed of a continuous function. *Signal Processing*, 38(1):99-12. 1994.

S. Namiki, M. Dickinson, A. Wong, W. Korff, and G. Card. The functional organization of descending sensory-motor pathways in *Drosophila*. *eLife.* e34272. 2018.

H. Nöthel. Investigations on radiosensitive and radioresistant populations of *Drosophila melanogaster*. XVI. Adaptation to the mutagenic effects of X-rays in several experimental populations with irradiation histories. *Mutation Research,* 111(3): 325-40. 1983.

V. K. Oers, P. Drent, P. de Goede, and A. J. van Noordwijk. Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Biological Society,* 271(1534): 65-73. 2004.

O. Sakai. Comparison of personality between juveniles and adults in clonal gecko species. *Journal of ethology*, 36(3): 221-8. 2018.

T. Ofstad, C. Zuker, and M. Reiser Visual place learning in *Drosophila melanogaster*. *Nature*, 474: 204-7. 2011.

L. Pallares, S. Picard, and J. F. Ayroles. TM3'seq: a tagmentation-mediated 3'sequencing approach for improving scalability of RNA-seq experiments. *bioRxiv*, 2019. doi: 10.1101/585810.

C. Pantoja, A. Hoagland, E. Carroll, V. Karalis, A. Conner, and E. Y. Isacoff. Neuromodulatory regulation of behavioral individuality in zebrafish. *Neuron*, 91(3):587-601. 2016.

D. Réale, S. M. Reader, D. Sol, P. T. MacDougall, N. J. Dingemanse. Integrating animal temperament within ecology and evolution. *Biological Reviews of the Cambridge Philosophical Society*, 82(2): 291-318. 2007.

A. Robie, J. Hirokawa, A. Edwards, L. A. Umayam, A. Lee, M. L. Phillips, G. M. Card, W. Korff, G. M. Rubin, J. H. Simpson, M. B. Reiser, and K. Branson. Mapping the neural substrates of behavior. *Cell*, 170(2): 393-406. 2017.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

W. Schuett, S. Dall, J. Baeumer, M. H. Kloesener, S. Nakagawa, F. Beinlich, and T. Eggers. Personality variation in a clonal insect: The pea aphid, *Acyrthosiphon pisum*. *Dev Psychobiol*, 53(6):631–640, 2011.

A. Seeds, P. Ravbar, P. Chung, S. Hampel, F. M. Midgley, B. D. Mensh, and J. H. Simpson. A suppression hierarchy among competing motor programs drives sequential grooming in *Drosophila*. *eLlife*, 3:e02951. 2014.

J. Seelig and V. Jayaraman. Neural dynamics for landmark orientation and angular path integration. *Nature*, 521(7551):1866-91. 2015.

S. Stern, C. Kirst, and C. Bargmann. Neuromodulatory control of long-term behavioral patterns and individuality across development. *Cell*, 171(7):1649-662. 2017.

J. G. Todd, J. S. Kain, and B. de Bivort. Systematic exploration of unsupervised methods for mapping behavior. *Phys Biol*, 14(1): 015002, 2017.

Z. Werkhoven, C. Rohrsen, C. Qin, B. Brembs, and B. de Bivort. MARGO (Massively automated real-time GUI for object-tracking), a platform for high-throughput ethology. *BioRxiv*, 2019. doi: 10.1101/593046.

T. Wolff and G. M. Rubin. Neuroarchitecture of the *Drosophila* central complex: A catalog of nodulus and asymmetrical body neurons and a revision of the protocerebral bridge catalog. *The Journal of comparative neurology*, 526(16):2585-611. 2018.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Materials

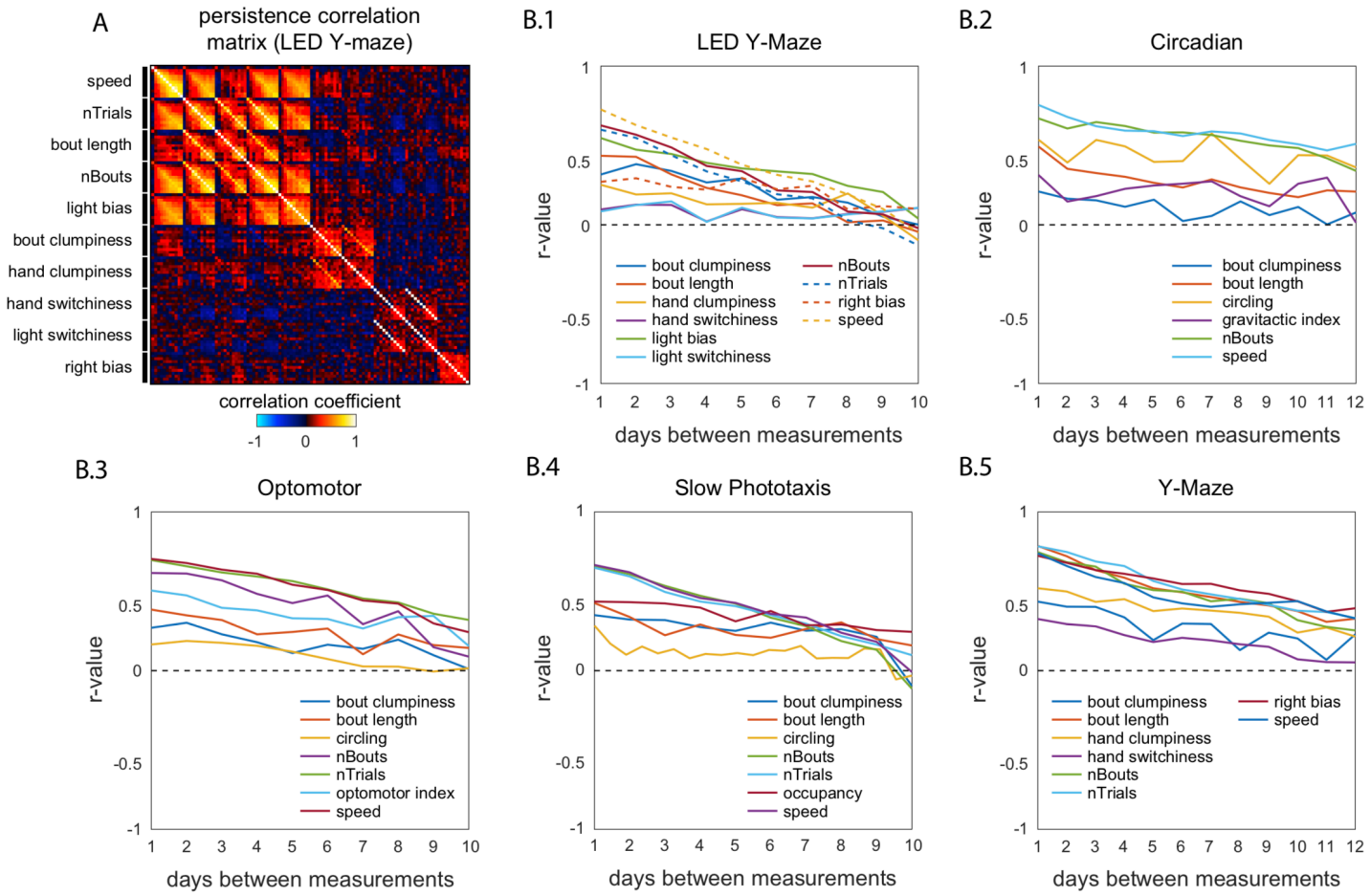


Figure S1 — Persistence of primary behavioral metrics across assays. A) Example correlation matrix for 10 retests of the same individuals on a single assay (phototactic Y-Maze) across subsequent days. Rows and columns were sorted to cluster the same measures across days. B.1-5) Plots of average correlation coefficient (*r*-value) as a function of the days between measurements. Colors indicate the metric within each plot separately (see legends). Dashed black line denotes no correlation (*r*=0). Metrics in all assays showed persistent individual variation that diminished over time.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

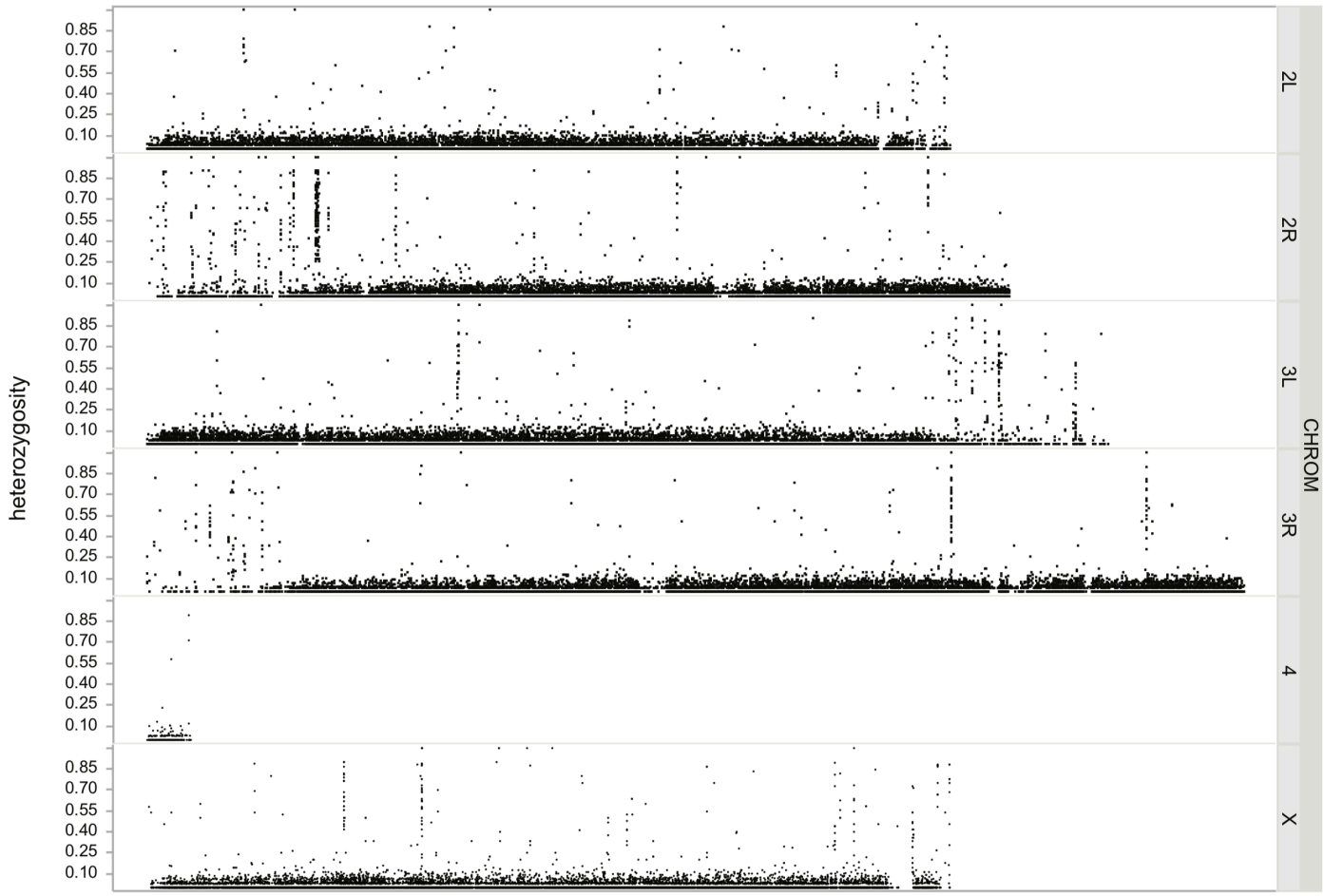


Figure S2 — Genomic sequencing to confirm isogeny of BK^{iso}. Fraction of BerlinK^{iso} flies (*n*=48) that were heterozygous at any given position in the genome. Less than 10% of flies were heterozygous at most sites. Flies showed evidence of residual heterozygosity at approximately 75 sites throughout the genome. These can be seen as distinct regions appearing as "columns" of heterozygous sites, and are most prevalent at the ends of chromosome arms, near the centromeres.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

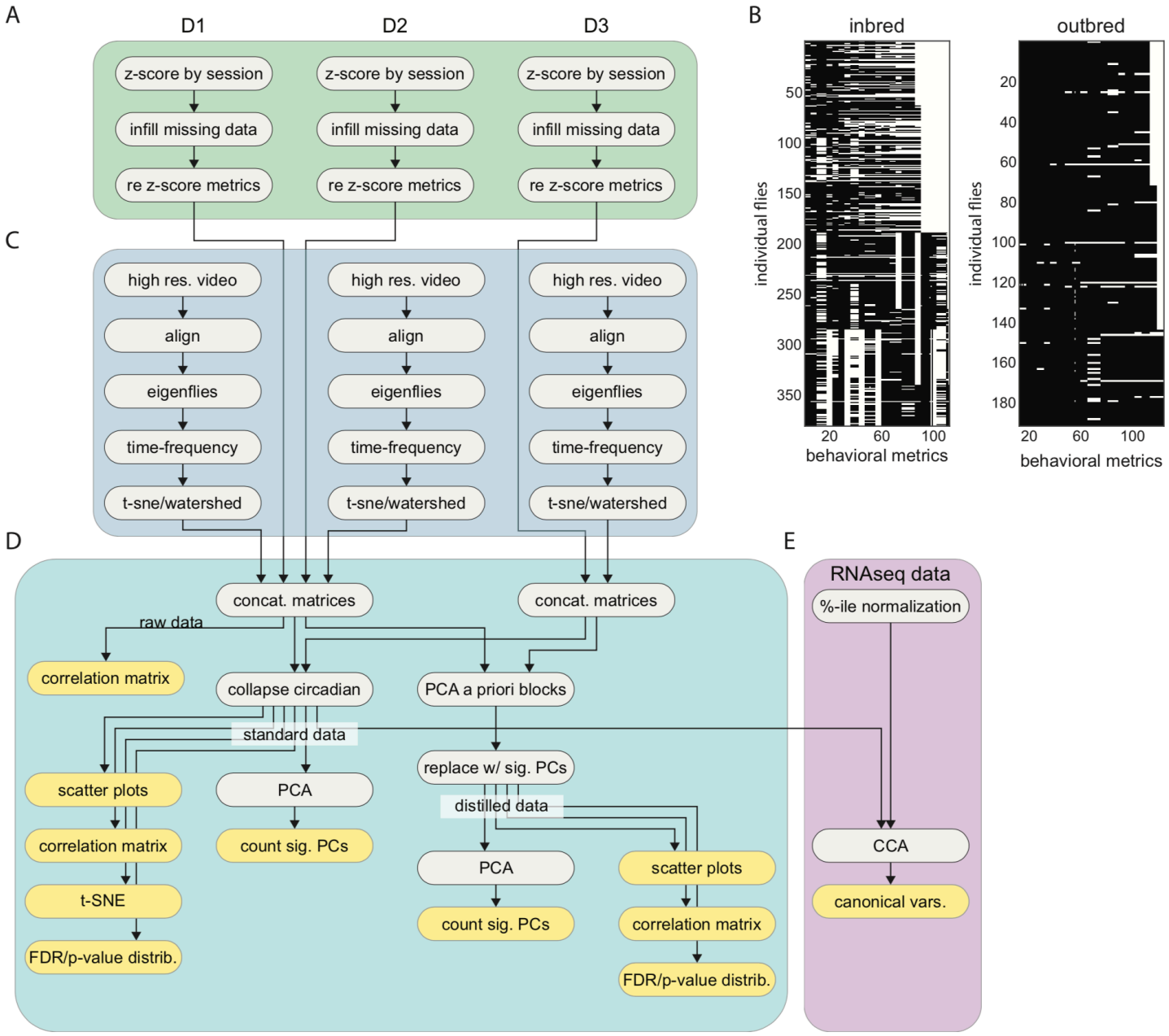


Figure S3 — Schematic of the decathlon analysis pipeline. A) Decathlon behavioral data preprocessing pipeline. Metrics are *z*-scored by imaging session to adjust for batch effects. Missing data is infilled with an ALS imputed matrix calculated as the average of 200 ALS imputations. B) Data matrices from the inbred and outbred Decathlon experiments. Rows are flies, columns are behavioral measures. White cells indicate missing data prior to infilling. C) Data acquisition and analysis for the unsupervised behavioral classification pipeline. High resolution, high frame rate single fly videos are aligned to a template fly and compressed into principal component time series (i.e., eigenflies). PC time series are then decomposed into 25 frequency spectral time series via Morlet wavelet transformation. The resulting high dimensional data is then embedded into two dimensions with t-SNE and then clustered into discrete behavioral modes via watershed transformation. D) Diagram of decathlon behavioral data analysis. Data from the decathlon assays and unsupervised behavioral classifications are combined into master behavioral data matrices for inbred (left) and outbred (right) flies separately. The covariance structure and effective dimensionality of the resulting matrices are then analyzed independently or are compressed into a "distilled" matrix with fewer dimensions. Distilled matrices are generated by retaining significant PCs within each a priori metric group above or within the variance explained by a model of *n*-independent dimensions (see methods). E). Individual flies undergo gene expression profiling via RNAseq. Gene expression and behavioral data are then analyzed with canonical correlation analysis to identify principal axes of gene expression that are correlated to principal axes of behavior.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

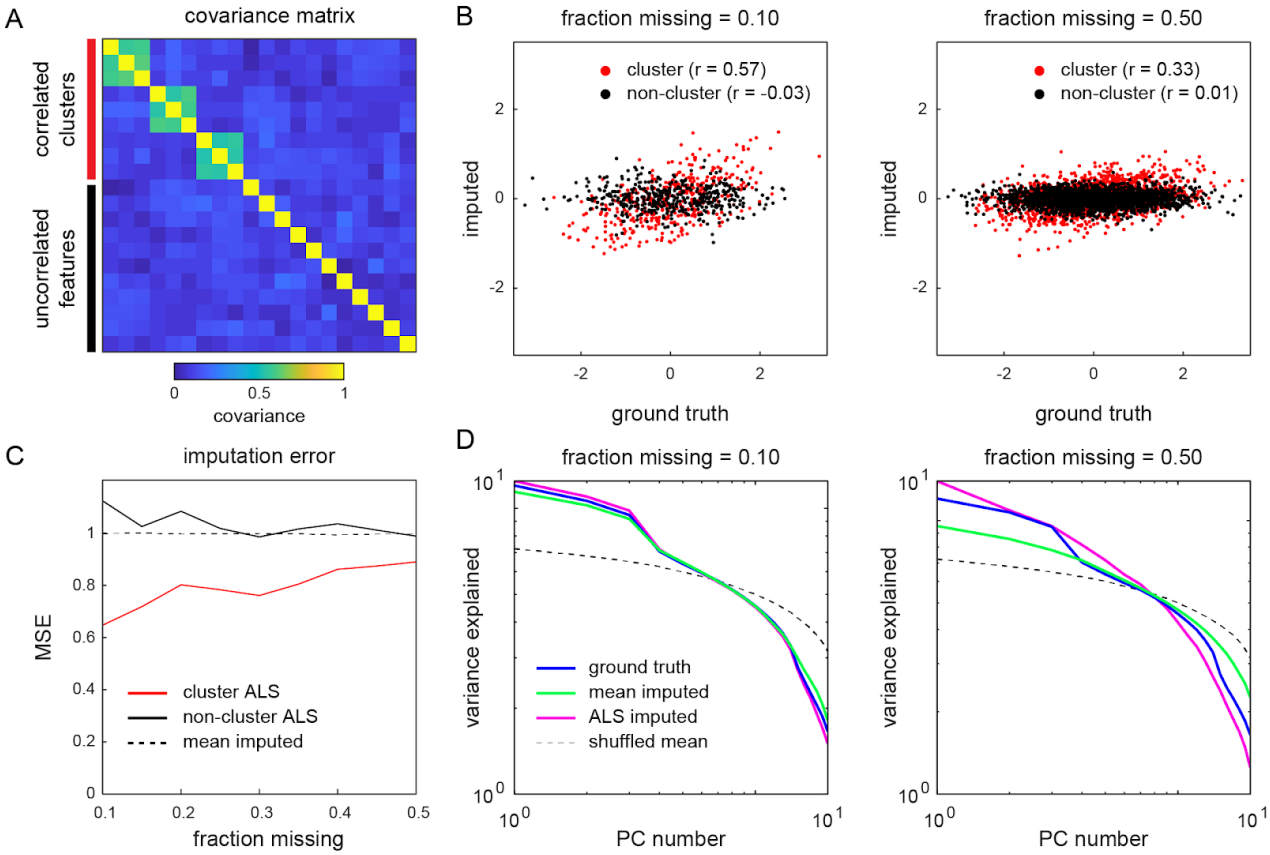


Figure S4 — Toy data comparison of matrix-infilling methods. A) Covariance matrix of a toy data set generated from a multivariate normal distribution with covariance = 0.5 for features in correlated clusters (red block) and covariance = 0 for independent features. B) Scatter plots of data from a ground truth vs the same data infilled with average ALS imputation (200 repetitions) after randomly deleting either 10% (left) and 50% (right) of the entries. Color indicates whether the data belonged to a correlated cluster as in A. C) Comparison of mean squared error (MSE) for mean infilled and ALS average infilled data. D) Log scree plots of the variance explained for ranked principal components resulting from PCA on the ground truth, mean infilled, average ALS infilled, and shuffled matrix.

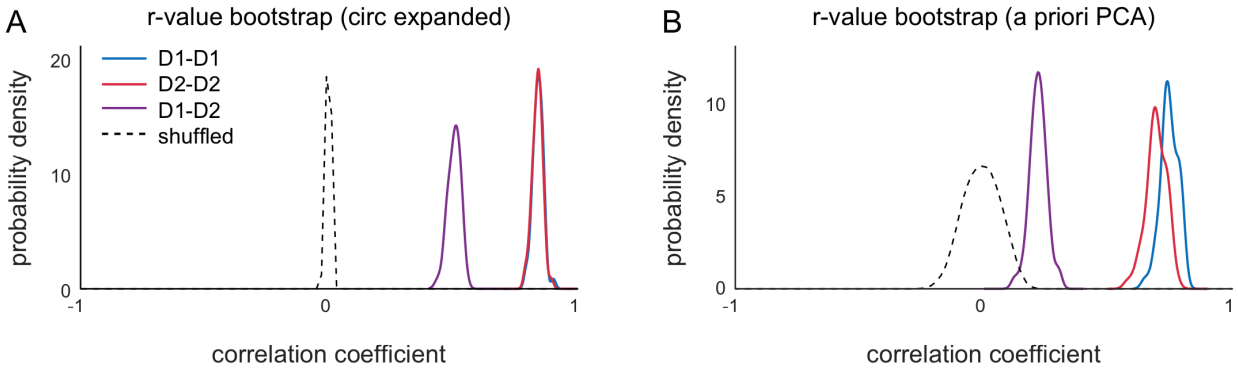


Figure S5 — Correlation of correlation matrix values between D1 vs D2 and shuffled matrices. A) Distributions of correlation of full matrix correlation matrices across bootstrap replicates. Correlation coefficients were calculated by bootstrap resampling data matrices (decathlon-1 or decathlon-2) and computing the correlation matrix. Correlation was then computed between two matrices (i.e. either the same or different matrices across resamplings). B) Distributions of correlation of distilled matrix correlation matrices across bootstrap replicates as in A.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

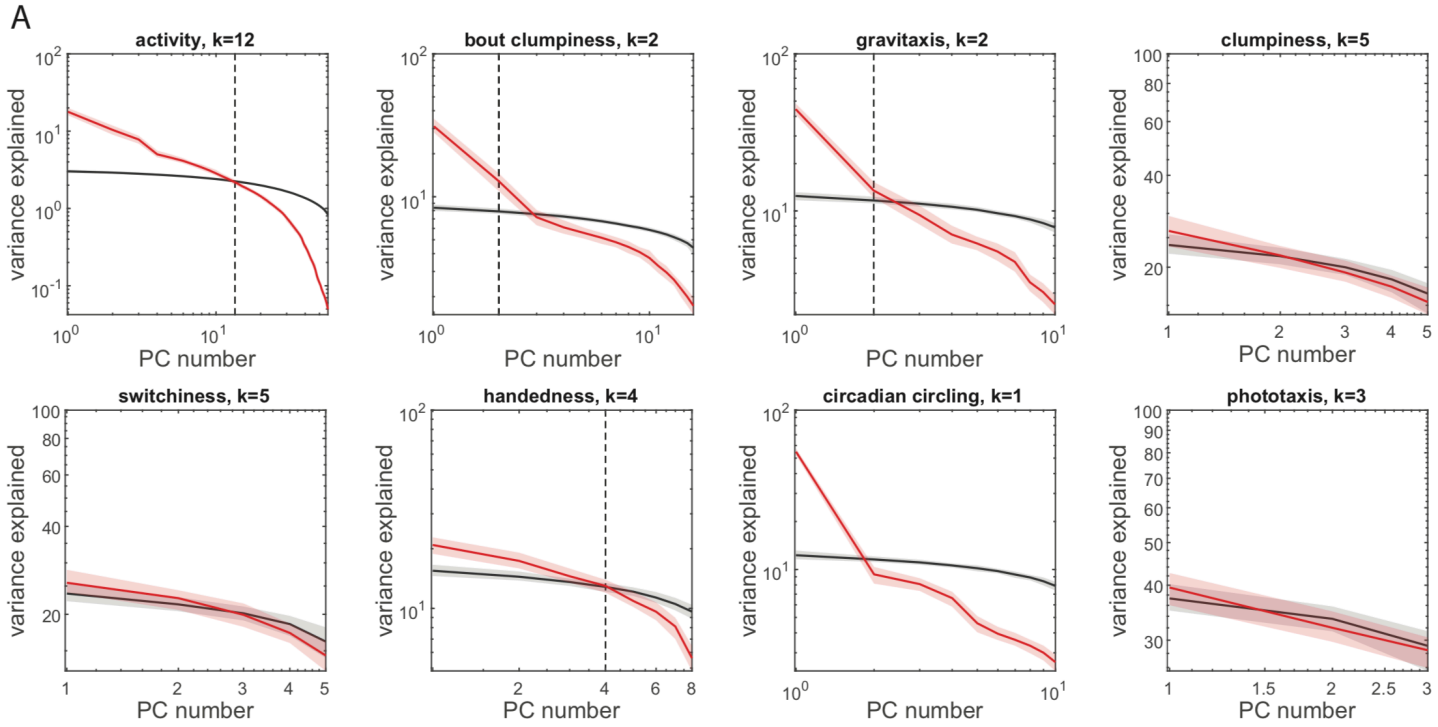
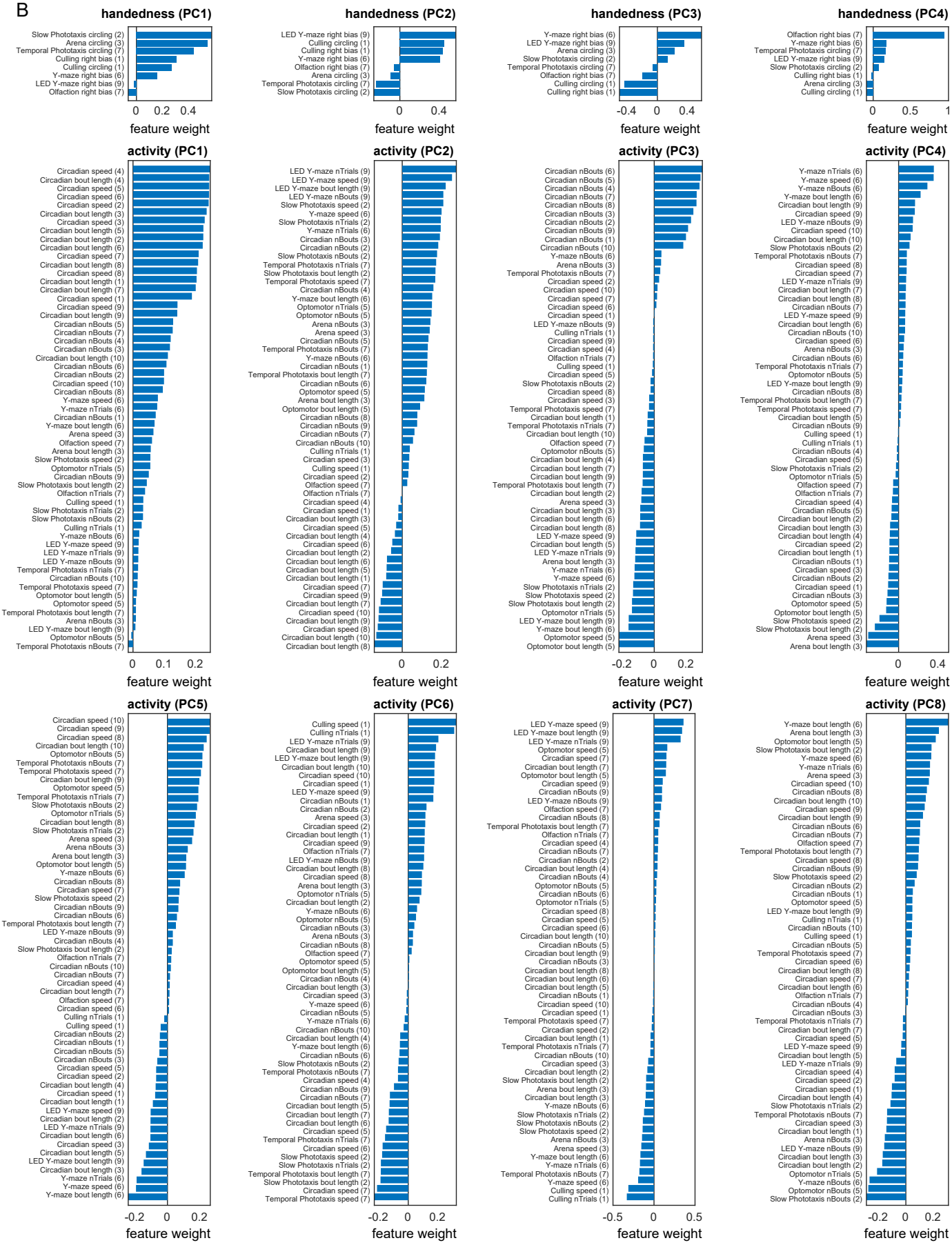


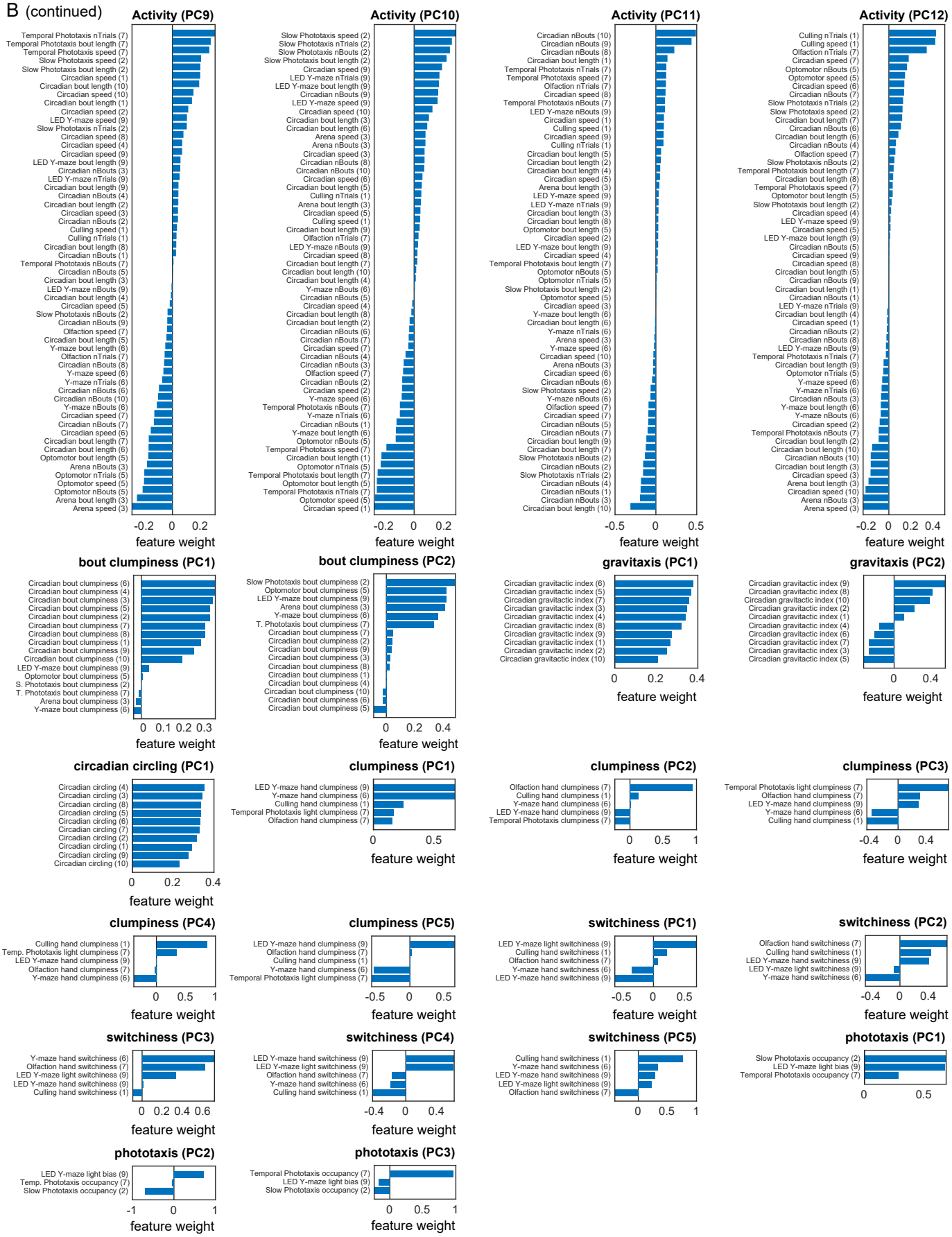
Figure S6 (continues on next two pages) — PCA of a priori groups. A) Log scree plots of the normalized ranked eigenvalues (i.e., variance explained) for PCA performed on metrics from each a priori group separately. Color indicates observed or shuffled data. Shaded regions correspond to 95% confidence intervals as calculated by bootstrap resampling. The number of significant PCs (*k*) was calculated as the highest rank PC above or within the 95% confidence interval of the shuffled matrix variance explained. B) Metric loadings of significant PCs for each a priori group.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

B (continued)



bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

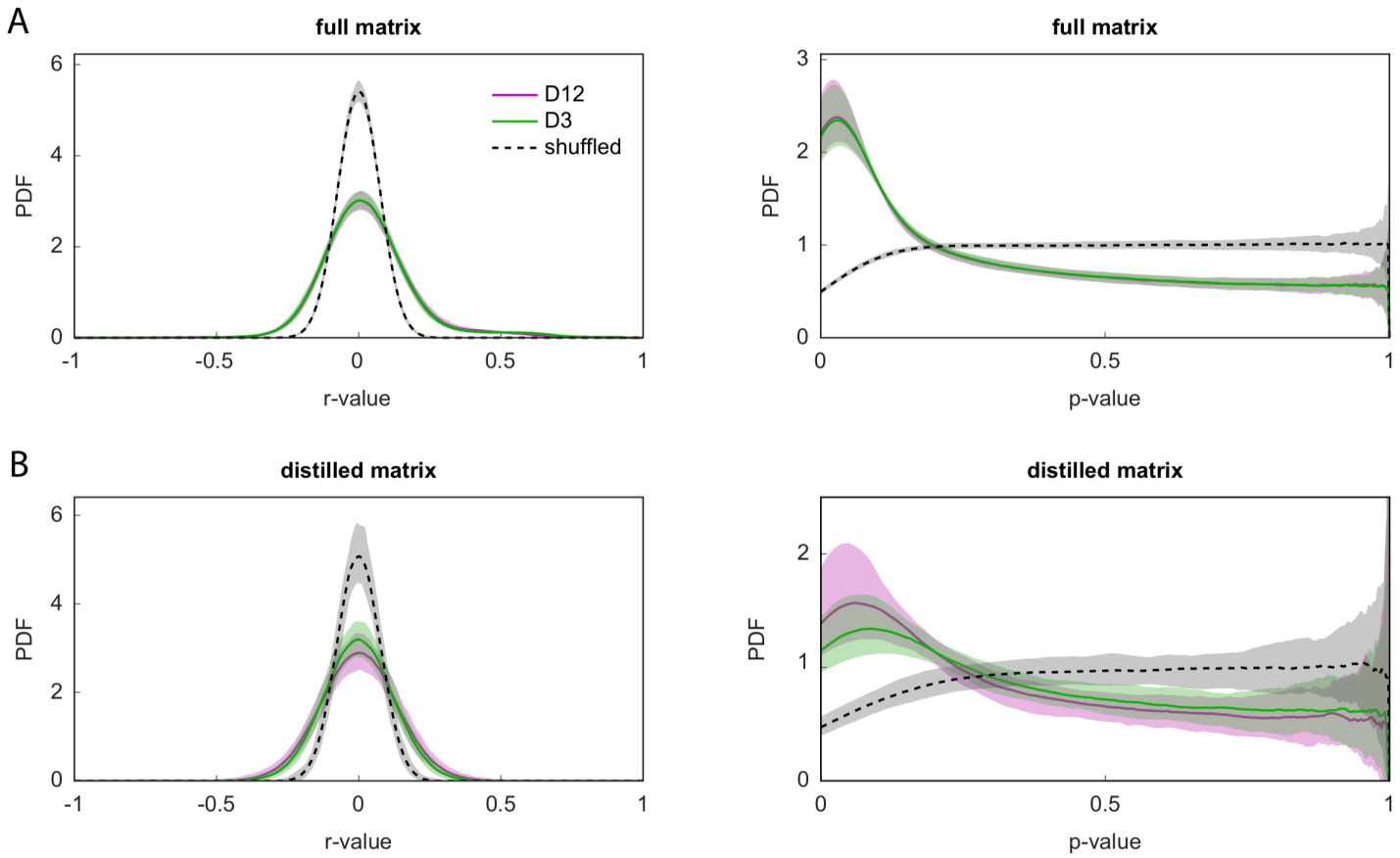


Figure S7 — Distribution of correlation coefficients and *p*-values in the full and distilled correlation matrix. A) Kernel density estimates of the unique (i.e., lower matrix triangle) correlation coefficients in the full (top) and distilled (bottom) correlation matrices. Distributions exclude duplicates pairwise and self correlations. B) Kernel density of the the unique *p*-values for the correlation coefficients in A. In all plots, dashed lines indicate distributions for column-wise (i.e., within each feature) shuffled matrices. Shaded regions correspond to 95% confidence interval calculated by bootstrap resampling.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

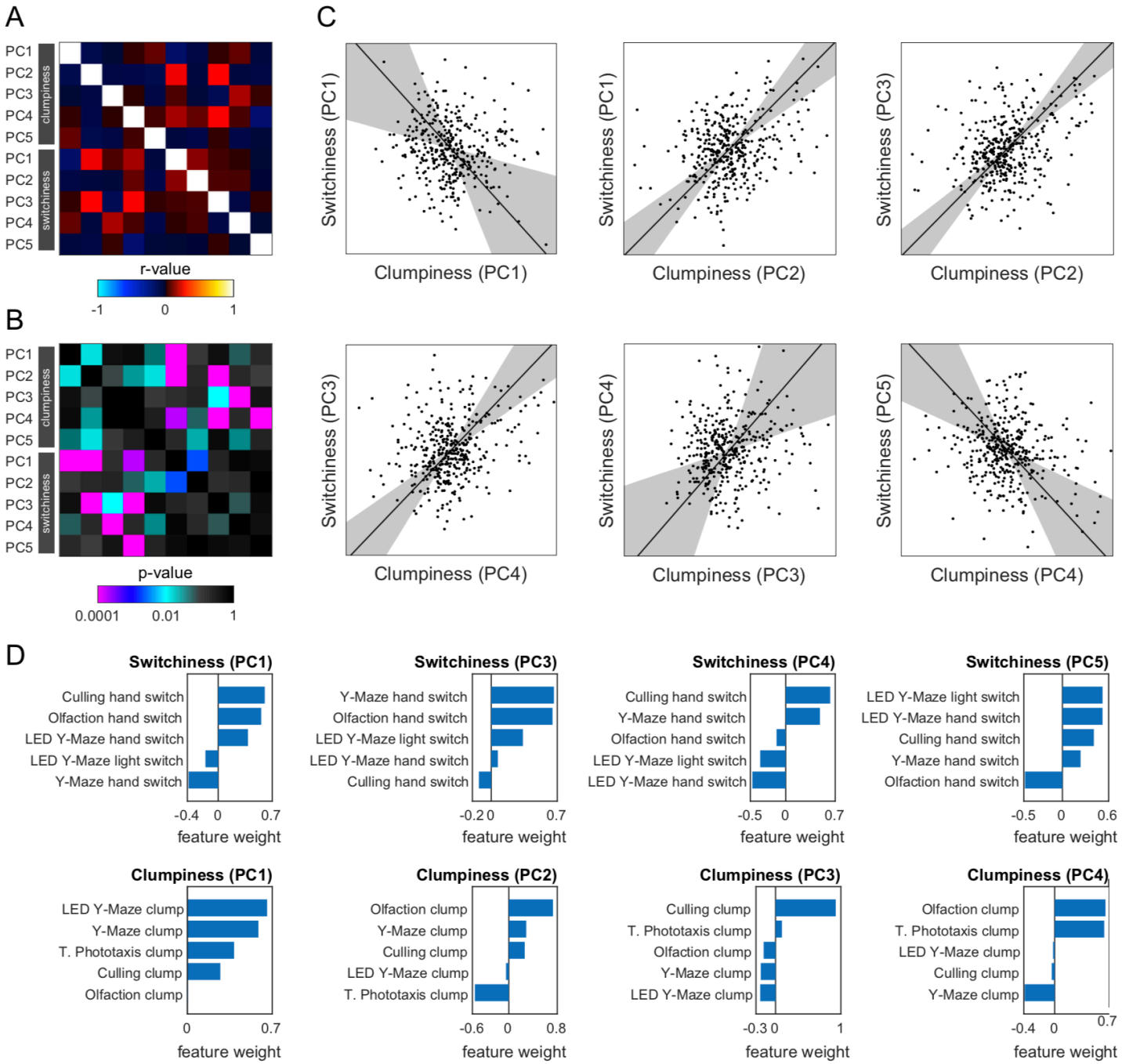


Figure S8 — Significant correlations among the principal components of switchiness and clumpiness. A) Subset of the distilled correlation matrix corresponding to the significant PCs of switchiness and clumpiness. B) *P*-value matrix for the correlation coefficients in A. C) Scatter plots of significant correlations between switchiness and clumpiness. Points correspond to individual flies. Line indicates the line of best fit and the shaded region indicates the 95% confidence interval of the fit as calculated by bootstrap resampling. D) Metric loadings for the PCs in C.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

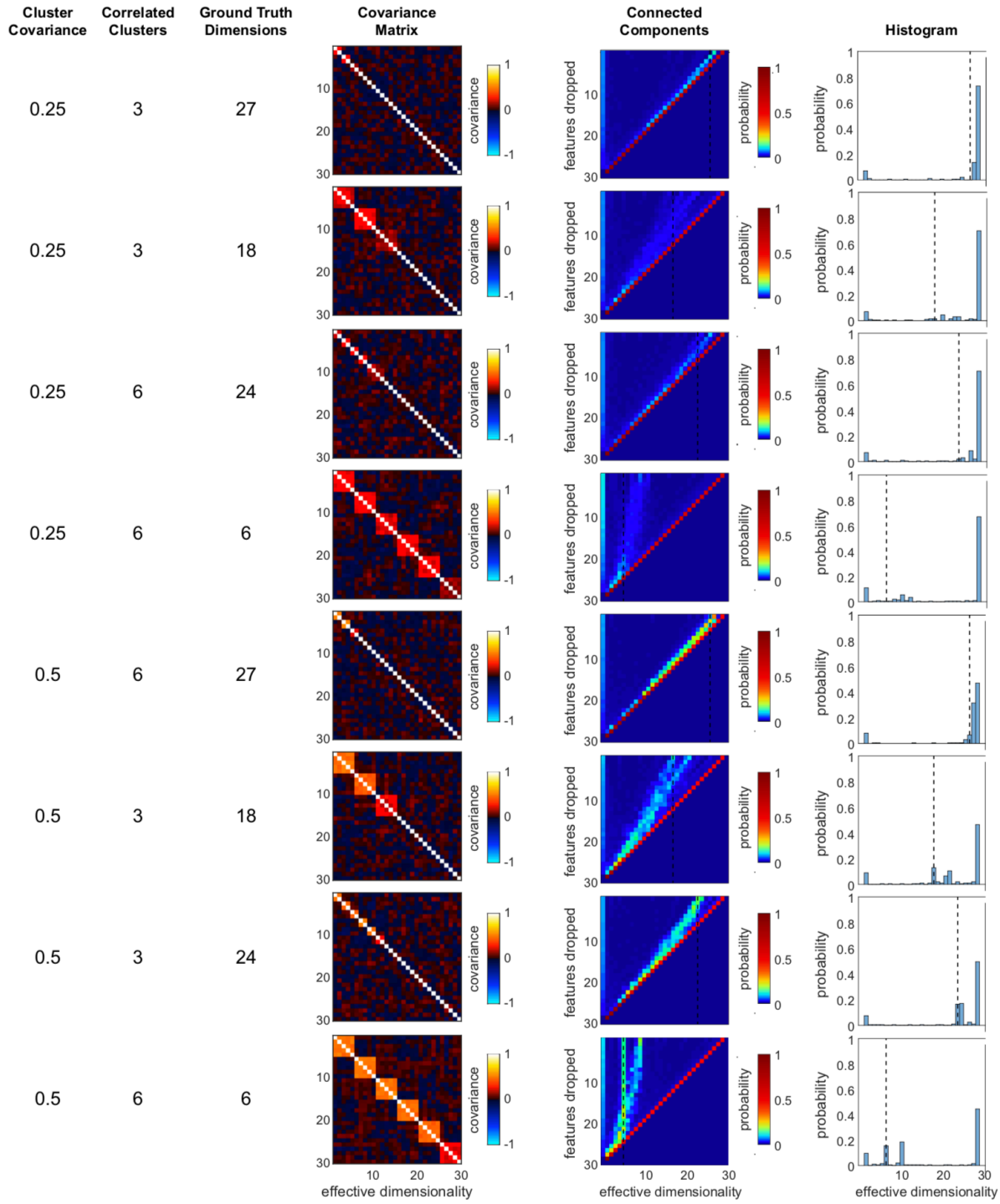


Figure S9 — Effective dimensionality simulations with toy data. A) Parameters used to generate a 200x30 ground truth matrix from a multivariate normal distribution as in Figure S4. Cluster covariance refers to the covariance between features belonging to the same correlated cluster. Correlated clusters corresponds to the total number of such clusters. Ground truth dimensionality is the sum of number of correlated clusters and the total number of independent features. B) Covariance matrices of toy data sets with correlated clusters along the diagonal. One cluster in each data set has covariance equal to half of the remaining clusters. C) Effective dimensions heatmap as a function of number of features retained in the toy data after dropping *n* random features. Rows correspond to histograms of number of connected components in the covariance matrices. D) Connected components histogram for the full covariance matrix (i.e., the top row of C).

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

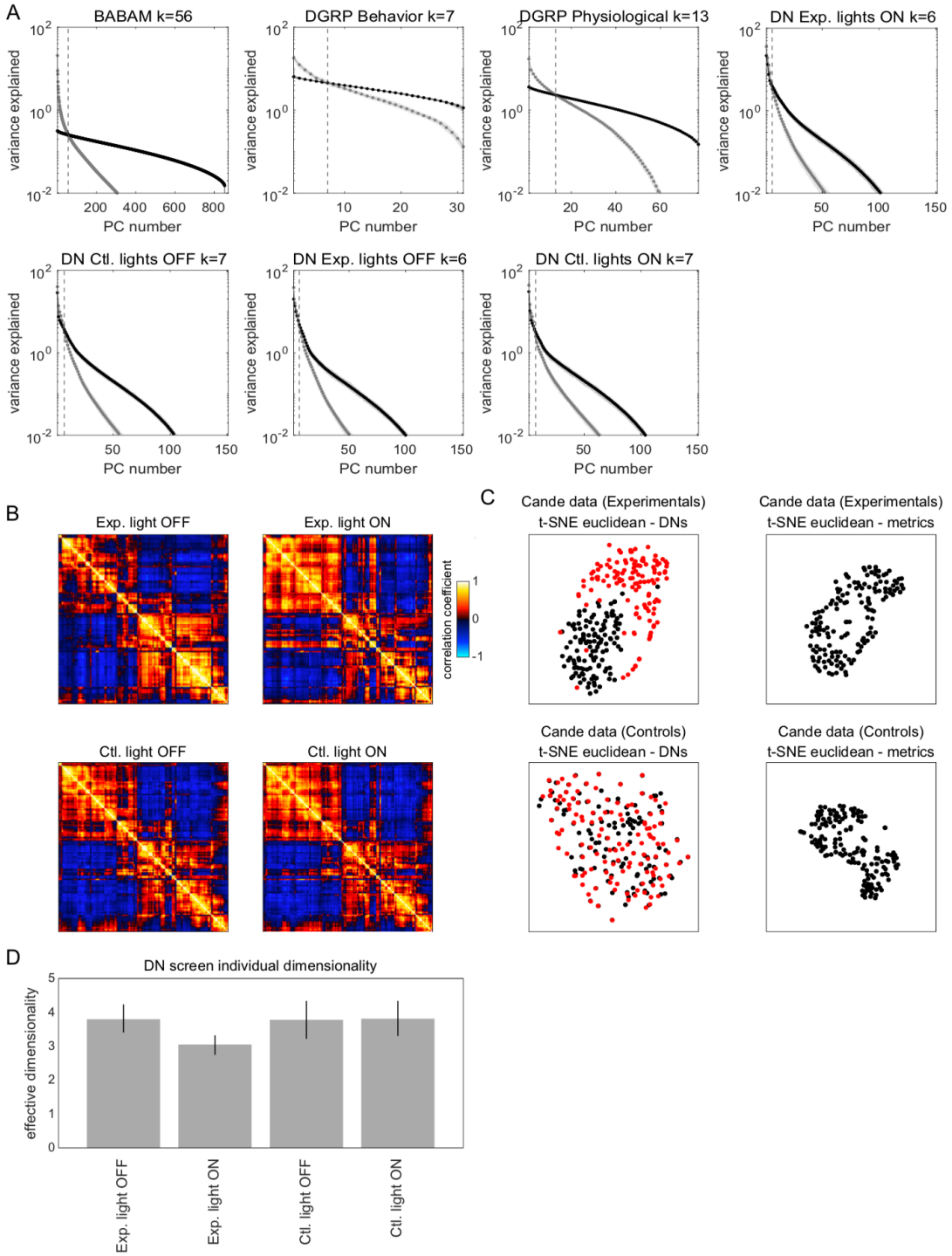


Figure S10 — Effective dimensionality of behavioral data. A) Scree plots showing the variance explained for each principal component of the BABAM Gal4 screen, DGRP (behavioral and physiological), and descending neuron screen (all experimental groups and conditions) behavioral data sets. Point colors indicate variance explained for the observed (gray) and shuffled (black) data matrices. The dashed line denotes one measure of effective dimensionality, where the variance explained of the observed principal components is below the 95% confidence interval (shaded regions) of the shuffled data. B) Correlation matrices for the combined behavioral PDF for each descending neuron set separated by experimental group and condition. C) t-SNE embeddings of the descending neurons lines (left) and unsupervised metrics (right) from the descending neuron screen. Color of the data points in the left hand plot indicates whether individuals were control (black; Gal4/+) or experimental (red; Gal4/UAS-CsChrimson). D) Average effective dimensionalit (as measured by the intersection of observed and shuffled ranked PC variances) of the individual behavioral PDFs separated by experimental group and condition. Error bars are the 95% confidence interval of the mean.

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Movies

Movie S1 — *https://www.youtube.com/watch?v=3cuSapTnSOU* — Examples of a mode of walking behavior as identified by the unsupervised analysis, from movies of single flies, made up of successive frames classified as the same behavior. Colored dots indicate whether flies are outbred (NEX; red) or inbred (Berlin-Kiso; blue).

Movie S2 — *https://www.youtube.com/watch?v=E1vIl-1nLZU* — Examples of a mode of wing grooming behavior as identified by the unsupervised analysis, from movies of single flies, made up of successive frames classified as the same behavior. Colored dots indicate whether flies are outbred (NEX; red) or inbred (BerlinKiso; blue).

Movie S3 — *https://www.youtube.com/watch?v=PY-C4NONOv7s* — Examples of a mode of head grooming behavior as identified by the unsupervised analysis, from movies of single flies, made up of successive frames classified as the same behavior. Colored dots indicate whether flies are outbred (NEX; red) or inbred (BerlinKiso; blue).

Movie S4 — *https://www.youtube.com/watch?v=lwCetEi0siA* — Examples of a mode of abdomen flexing behavior as identified by the unsupervised analysis, from movies of single flies, made up of successive frames classified as the same behavior. Colored dots indicate whether flies are outbred (NEX; red) or inbred (BerlinKiso; blue).

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Name	Description	Metrics Collected	Arena Geometry	Throughput	Duration	Activity Inclusion Criteria
Activity Screening	Activity measurement profiling at the start of the decathlon experiment used to narrow future testing from a cohort of 288 flies to the most active 192 flies	circlingBias, turnBias, choiceNumber, turnClumpiness, handSwitchiness, meanSpeed, boutNumber, boutLength, boutClumpiness	Y-shaped arena with arms equally spaced apart at 120 degrees (12.7mm length for each maze arm)	96 flies per behavioral box	2 hrs	highest 192 choiceNumber
Arena Circling	Open exploration of circular arena.	circlingBias, meanSpeed, boutNumber, boutLength, boutClumpiness	circular (28mm diameter)	48 flies per beahvioral box	2 hrs	meanSpeed > centroid error
Circadian	Overnight activity level monitoring on food.	circlingBias, meanSpeed, boutNumber, boutLength, boutClumpiness	96-multiwell plate (6.8mm diameter, 10.5mm depth)	96 flies per behavioral box	20-21 hrs	meanSpeed > centroid error
LED Y-maze Phototaxis	Discrete two-choice (light or no light) phototactic assay. Choices are scored as turns toward or away from a lit arm of the maze and as turns to the right or left of the choice point.	circlingBias, turnBias, photoBias, choiceNumber, turnClumpiness, handSwitchiness, photoSwitchiness, meanSpeed, boutNumber, boutLength, boutClumpiness	Y-shaped arena with LEDs at the end of each arm (12.7mm length for each maze arm)	72 flies per behavioral box	2 hrs	choiceNumber > 10
Olfactory Sensitivity	Odor vs no-odor adaptation assay measuring sensitivity to a low concentration (typically aversive) odorant	odorOccupancy, preodorOccupancy, turnBias, choiceClumpiness, handSwitchiness, choiceNumber, meanSpeed	linear tunnels with odorant ports at either end and decision boundary in center	15 flies per experiment	15 min (3 min preodor, 12 min odor)	meanSpeed > centroid error
Optomotor	Optomotor response assay where flies are repeatedly targeted with a high-contrast rotating pinwheel stimulus projected onto the arena floor	circlingBias, optoIndex, meanSpeed, trialNumber, boutNumber, boutLength, boutClumpiness	circular (28mm diameter)	48 flies per beahvioral box	2 hrs	meanSpeed > centroid error
Spatial Shade-Light	Phototaxis assay where dark and light occupancy are measured. Individual arenas are targeted with a half lit, half dark spatial pattern.	circlingBias, meanSpeed, boutNumber, boutLength, boutClumpiness, photoOccupancy	circular (28mm diameter)	48 flies per beahvioral box	2 hrs	meanSpeed > centroid error
Temporal Shade-Light	Phototaxis assay where dark and light occupancy are measured. Individual arenas are either fully dark or fully lit depending on fly position relative to an invisible virtual boundary.	circlingBias, meanSpeed, boutNumber, boutLength, boutClumpiness, photoOccupancy	circular (28mm diameter)	48 flies per beahvioral box	2 hrs	meanSpeed > centroid error
Y-maze	Discrete two-choice (right or left) locomotor handedness assay. Choices are scored as turns to the right or left of the choice point in a Y-shaped arena.	circlingBias, turnBias, choiceNumber, turnClumpiness, handSwitchiness, meanSpeed, boutNumber, boutLength, boutClumpiness	Y-shaped arena with arms equally spaced apart at 120 degrees (12.7mm length for each maze arm)	96 flies per behavioral box	2 hrs	choiceNumber > 10

Table S1 — Details of decathlon behavioral assays

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Name	Metrics	Metric Number	PCs Retained	% Variance Captured
Activity	**Circadian:** bout_length (1-10), nBouts (1-10), speed (1-10) **Arena Circling:** bout_length, nBouts, speed **Culling:** nTrials, speed **LED Y-Maze:** bout_length, nBouts, speed, nTrials **Olfaction:** nTrials, speed **Optomotor:** bout_length, nBouts, nTrials, speed **Shade-Light:** bout_length, nBouts, nTrials, speed **Temporal Shade-Light:** bout_length, nBouts, nTrials, speed **Y-Maze:** bout_length, nBouts, nTrials, speed	57	13	68.36
Bout Clumpiness	**Circadian:** bout_clumpiness (1-10) **Arena Circling:** bout_clumpiness **Culling:** bout_clumpiness **LED Y-Maze:** bout_clumpiness **Optomotor:** bout_clumpiness **Shade-Light:** bout_clumpiness **Temporal Shade-Light:** bout_clumpiness **Y-Maze:** bout_clumpiness	16	2	43.55
Gravitaxis	**Circadian:** gravitaxis (1-10)	10	2	57.72
Clumpiness	**Culling:** turn_clumpiness **LED Y-Maze:** turn_clumpiness **Olfaction:** turn_clumpiness **Temporal Shade-Light:** choice_clumpiness **Y-Maze:** turn_clumpiness	5	5	100
Switchiness	Culling: hand_switchiness LED Y-Maze: hand_switchiness, light_switchiness Olfaction: hand_switchiness Y-Maze: hand_switchiness	5	5	100
Handedness	**Arena Circling:** circling **Culling:** right_bias **LED Y-Maze:** right_bias **Olfaction:** right_bias **Shade-Light:** circling **Temporal Shade-Light: circling** **Y-Maze:** right_bias	8	4	65.02
Circadian Handedness	**Circadian:** circling (1-10)	10	1	55.11
Optomotor Handedness	**Optomotor:** circling	1	1	100
Phototaxis	**LED Y-Maze:** light_bias **Shade-Light: light_occupancy** **Temporal Shade-Light:** light_occupancy	3	3	100
Optomotor Index	**Optomotor:** index	1	1	100
Odor Sensitivity	**Olfaction:** odor_occupancy	1	1	100

Table S2 — Breakdown of a priori groups: group name, metrics in each group, number of metrics in each group, number of PCs kept from each group, variance explained by kept PCs).

bioRxiv preprint doi: https://doi.org/10.1101/779363; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Targeted Neurons	Gal4 Alleles	Source
P-FN3	SS00079, SS00406, SS02255, SS02221, SS00007, SS02304, SS00083	Gift from Tanya Wolff and Gerry Rubin
P-FN4	SS00161, SS00078, SS00019, SS00082, SS00083	Gift from Tanya Wolff and Gerry Rubin
P-FN5	SS02243, SS02209, SS02211, SS00082	Gift from Tanya Wolff and Gerry Rubin
P-EN	SS02241, SS02268, SS02232	Gift from Tanya Wolff and Gerry Rubin
P-EG	SS00097	Gift from Tanya Wolff and Gerry Rubin
E-PG	SS00090, SS00131, SS02258	Gift from Tanya Wolff and Gerry Rubin
E-PG9	SS04781, SS04782, SS02374	Gift from Tanya Wolff and Gerry Rubin
P-F-gall	SS02267, SS02271	Gift from Tanya Wolff and Gerry Rubin
P-F-ROB	SS02192, SS02293	Gift from Tanya Wolff and Gerry Rubin
PF-CreL	SS02252, SS02214	Gift from Tanya Wolff and Gerry Rubin
PsL-PB	SS02216, SS02200	Gift from Tanya Wolff and Gerry Rubin
PBCap	SS04775	Gift from Tanya Wolff and Gerry Rubin
PBInt	SS02231, SS02235, SS00116	Gift from Tanya Wolff and Gerry Rubin
Ps-P	SS02296, SS00155, SS00153	Gift from Tanya Wolff and Gerry Rubin
PsIb-P	SS04769	Gift from Tanya Wolff and Gerry Rubin

Table S3 — List of Gal4 lines and the cell types they target from the Shibire^{ts} and dTRPA1 screen