

Functional Validation of Candidate Genes Detected by Genomic Feature Models

Palle Duun Rohde,^{*,†,*,1,2} Solveig Østergaard,^{*,1} Torsten Nygaard Kristensen,^{§,**} Peter Sørensen,^{*}

Volker Loeschcke,[§] Trudy F. C. Mackay,^{††,**,§§} and Pernille Sarup^{*}

^{*}Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, [†]Center for Integrative Sequencing, and [§]Section for Genetics, Ecology and Evolution, Department of Bioscience, Aarhus University, 8830 Tjele, Denmark, ^{††}The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, 8000 Aarhus, Denmark, ^{**}Section for Biology and Environmental Science, Department of Chemistry and Bioscience, Aalborg University, 9220 Aalborg, Denmark, ^{††}Department of Biological Sciences, ^{**}Program in Genetics, and ^{§§}W. M. Keck Center for Behavioral Biology, North Carolina State University, Raleigh, North Carolina 27695

ORCID IDs: 0000-0003-4347-8656 (P.D.R.); 0000-0002-5838-1251 (P.S.)

ABSTRACT Understanding the genetic underpinnings of complex traits requires knowledge of the genetic variants that contribute to phenotypic variability. Reliable statistical approaches are needed to obtain such knowledge. In genome-wide association studies, variants are tested for association with trait variability to pinpoint loci that contribute to the quantitative trait. Because stringent genome-wide significance thresholds are applied to control the false positive rate, many true causal variants can remain undetected. To ameliorate this problem, many alternative approaches have been developed, such as genomic feature models (GFM). The GFM approach tests for association of set of genomic markers, and predicts genomic values from genomic data utilizing prior biological knowledge. We investigated to what degree the findings from GFM have biological relevance. We used the *Drosophila* Genetic Reference Panel to investigate locomotor activity, and applied genomic feature prediction models to identify gene ontology (GO) categories predictive of this phenotype. Next, we applied the covariance association test to partition the genomic variance of the predictive GO terms to the genes within these terms. We then functionally assessed whether the identified candidate genes affected locomotor activity by reducing gene expression using RNA interference. In five of the seven candidate genes tested, reduced gene expression altered the phenotype. The ranking of genes within the predictive GO term was highly correlated with the magnitude of the phenotypic consequence of gene knockdown. This study provides evidence for five new candidate genes for locomotor activity, and provides support for the reliability of the GFM approach.

KEYWORDS

Drosophila melanogaster
DGRP
genomic prediction
set test
locomotor activity

One of the major challenges in modern biology is to understand the link between molecular genetic variation and quantitative trait variation. For the vast majority of quantitative traits and diseases, phenotypic variation

is caused by the joint effects of multiple segregating genetic variants, their interactions, environmental effects, and genotype-environment interactions and correlations (Falconer and Mackay 1996; Lynch and Walsh 1998). Knowledge of the genetic architecture of complex traits including species-specific causal genetic variants, and the distribution of their effect sizes and frequencies is important in multiple disciplines, such as animal and plant breeding, adaptive evolution, and in the study of complex human diseases and disorders.

Technological advancements in molecular biology, in particular the development of array-based and high-throughput sequencing platforms, have enabled large scale genome-wide scans for statistical associations between single nucleotide polymorphisms (SNPs) and quantitative traits and diseases (Balding 2006; Hardy and Singleton 2009). These genome-wide association studies (GWAS) have now been conducted on a large range of human diseases and traits, livestock and

Copyright © 2018 Duun Rohde et al.

doi: <https://doi.org/10.1534/g3.118.200082>

Manuscript received January 29, 2018; accepted for publication March 7, 2018; published Early Online March 8, 2018.

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

Supplemental material available at Figshare: <https://doi.org/10.25387/g3.5951581>

¹Equal contribution

²Corresponding author: Palle Duun Rohde, Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark, E-mail: palle.d.rohde@mbg.au.dk

plant production traits (Dekkers 2012; Xiao *et al.* 2017), model organisms (Atwell *et al.* 2010; Mackay *et al.* 2012), as well as non-model species (Husby *et al.* 2015).

One of the major challenges with GWAS is the inability to detect all causal SNPs, or all SNPs correlated with the causal variants. Stringent genome-wide significance thresholds are needed in order to efficiently control the false positive rate. Because most SNP effect sizes are small to moderate, the majority of the causal variants will remain undetected (Yang *et al.* 2010, 2011; Visscher *et al.* 2012). Therefore, in recent years, methods have been developed for assessing the joint effect of multiple SNPs on trait variability. Such methods include set test approaches (Mooney *et al.* 2014; de Leeuw *et al.* 2016), regional SNP-based heritability approaches (Nagamine *et al.* 2012; Uemoto *et al.* 2013), and genomic prediction models using all available SNPs simultaneously (Meuwissen *et al.* 2001; Speed and Balding 2014). The main advantage of these methods is that they consider the contribution from SNPs whose effect sizes are too small to be classified as associated variants in a traditional GWAS.

One of the emerging themes obtained from GWAS is that top associated SNPs tend to cluster in biological pathways (Lango Allen *et al.* 2010; Lage *et al.* 2012; Maurano *et al.* 2012; O’Roak *et al.* 2012). This knowledge could be utilized more directly in the statistical models, and has the potential to increase the power to uncover the underlying biology of complex trait phenotypes. One approach is the statistical framework entitled genomic feature models, GFM (implemented as an R-package which is available at <http://psoerensen.github.io/qgg/>). We have successfully applied this modeling approach to cattle (Edwards *et al.* 2015; Fang *et al.* 2017a; Fang *et al.* 2017b), pigs (Sarup *et al.* 2016), mice (Ehsani *et al.* 2015), fruit flies (Edwards *et al.* 2016; Rohde *et al.* 2017; Sørensen *et al.* 2017) and humans (Rohde *et al.* 2016a), and have shown that these models can provide novel biological knowledge of complex traits. Some challenges with this approach still remain. First, when the genomic feature analysis is based on large gene sets, it may be useful to reduce, or restrict, the list of genes within the associated gene set, to those genes with the greatest contribution to the overall trait variability. Second, to date the results from GFM have been limited to discovery of putative causal variants, and true functional validation of the variants has been lacking.

We have previously described the SNP set test approach – the covariance association test (CVAT) – as a powerful method for associating a set of SNPs with human diseases and complex traits (Rohde *et al.* 2016a; Sørensen *et al.* 2017). Here, we propose that CVAT can be used to rank genes within a large gene set, which collectively display statistical association with the trait phenotype, according to their estimated effect sizes. In order to experimentally test this, we used *Drosophila melanogaster* as a model system. *D. melanogaster* has many advantages over other model systems, such as a short generation time, easy husbandry, limited ethical restrictions, and a vast diversity of readily available genetic tools (e.g., functional mutants, temporal/spatial gene expression knockdown/in). A particularly useful resource is the *Drosophila* Genetic Reference Panel, DGRP (Mackay *et al.* 2012). The DGRP consists of 205 genome-wide homozygous lines derived by 20 generations of consecutive full-sib mating of wild-caught flies. Genome sequence data of the DGRP lines are publicly available (Mackay *et al.* 2012; Huang *et al.* 2014). The DGRP allows researchers to investigate the genetic basis for any quantitative trait phenotype. To date the DGRP has been used to study >45 quantitative traits (Anholt and Mackay 2017; Mackay and Huang 2017).

The aim of this study was twofold: (1) to investigate the applicability of CVAT to rank genes within a set of associated genes; and (2) to provide functional validation for the findings of the genomic feature models.

First, we used the genomic feature prediction models to identify large sets of genes, here defined by gene ontology (GO) categories (The Gene Ontology Consortium 2000), that were predictive of the trait values. Next, we used CVAT to rank the genes, within the larger set of genes that when considered jointly increased the predictive performance, according to the genomic variance captured by the individual genes within the predictive GO term.

We applied these methods to the quantitative trait locomotor activity in *D. melanogaster*. Collecting data on a new phenotype instead of using published data has the advantage of allowing us to perform the functional validation in the same manner as we did in phenotyping the DGRP, as well as potentially providing new biological knowledge of a complex trait phenotype. Locomotion is an important fitness component that is central for an individual’s survival and reproduction because it allows animals to localize mates and energy resources, defend territories, and escape from predators and environmental stress elements. Locomotor activity is a complex trait, and the genetic component is governed by the joint segregation of multiple quantitative trait loci, and likely their interactions. As a measurable trait, locomotor activity encompasses a broad range of different types of activity measures, some of which are species specific. Despite species- and trait specific differences, quantitative genetic analyses have revealed abundant genetic variation for different measures of locomotor activity across species (Burnet *et al.* 1988; Swallow *et al.* 1998; Lightfoot *et al.* 2004, 2008, 2010; Turner *et al.* 2005; Jordan *et al.* 2006, 2007).

Many different aspects of *Drosophila* locomotion have been studied, including phototaxis, geotaxis (Carpenter 1905), circadian rhythms of locomotor activity (Konopka and Benzer 1971), and rover-sitter foraging behavior (Osborne 1997). *Drosophila* locomotor activity has been quantified in several different ways, such as reactivity methods (i.e., quantifying the level of activity after physical disturbance) (Gargano *et al.* 2005; Jordan *et al.* 2006), infrared monitoring systems that quantify the number of times a fly passes a certain point (Rosato and Kyriacou 2006; Pfeiffenberger *et al.* 2010; Bährndorff *et al.* 2016), and video tracking methods (Zimmerman *et al.* 2008; Colomb *et al.* 2012; Gilestro 2012; Garbe *et al.* 2015). Here, we quantified locomotor activity in the DGRP using a high-throughput video tracking method (Rohde *et al.* 2016b) to quantify the total distance covered during a five-minute trial.

METHODS

Experimental design

The workflow of this study is depicted in Figure 1. We quantified locomotor activity for 204 DGRP lines in a highly-replicated study design. Genomic feature sets were defined based on GO categories. Each feature set was used in a genomic prediction model, and the predictive performance was compared to a null model that weight all SNP markers equally. The genes within a particular GO category are likely to contribute unequally to the predictive performance, as well as to forming the trait phenotype. Therefore, we used CVAT to rank the genes within the predictive GO categories according to their contribution to the trait variation. The genes that contribute most within the predictive GO categories were selected, and used in a functional validation experiment, where expression of these genes was suppressed using the binary *UAS-GAL4* system, and the phenotypic consequence on locomotor activity was assessed.

Drosophila stocks and husbandry

The DGRP lines (Mackay *et al.* 2012; Huang *et al.* 2014) were maintained in Prof. T. F. C. Mackay’s laboratory (North Carolina State University,

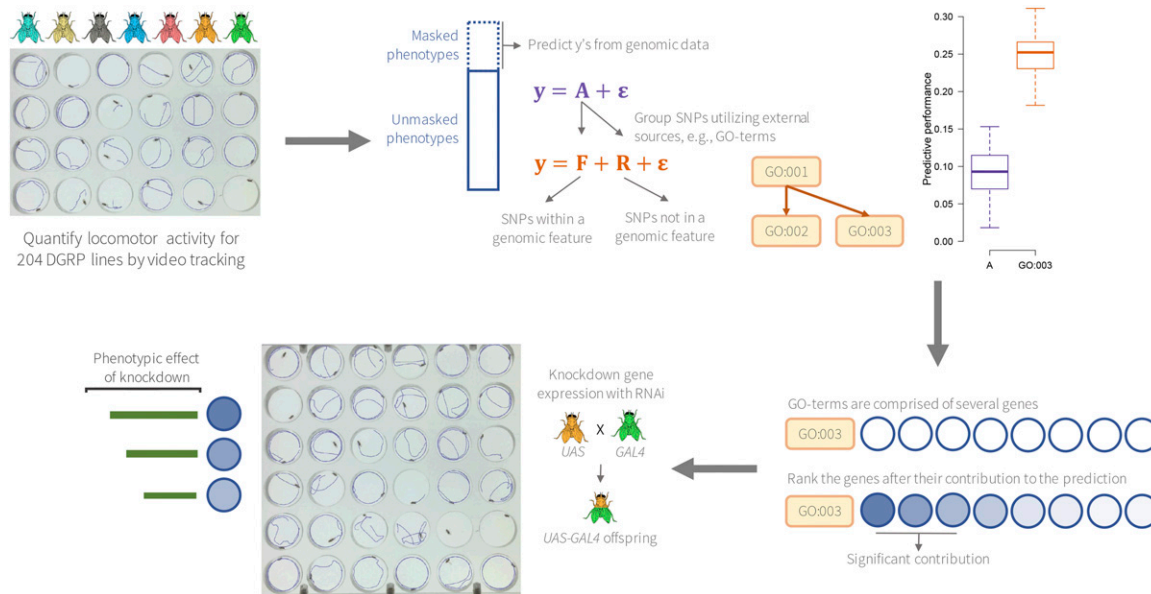


Figure 1 Outline of the study workflow starting with quantifying levels of activity for the 204 DGRP lines. A genomic feature prediction model was used to identify marker sets, defined by gene ontology (GO) terms that increased the predictive ability compared to a model where all markers were used simultaneously. The genes within a predictive GO term are likely to contribute unequally to the predictive performance, thus, the genes were ranked according to their contribution (quantified using CVAT). Top genes were functionally validated by suppressing gene expression and assessing the effect of such reduction on locomotor activity.

Raleigh, North Carolina 27695) on cornmeal-molasses-agar medium, 25°, 70% humidity, and a 12-h light-dark cycle. *UAS-RNAi* lines for gene expression knockdown (*CG10920*^{KK109327}, *CG14160*^{KK104744}, *CG15553*^{KK104514}, *CG1628*^{KK109588}, *CG17930*^{KK108912}, *CG32103*^{KK108078}, *CG33233*^{KK106897}, *Dic1*^{KK103757}, *DPCoAC*^{KK101378}, *Rim2*^{KK100807}, *Shawn*^{KK109948}) were obtained from the Vienna *Drosophila* Stock Center (<http://stockcenter.vdrc.at>) (Dietzl *et al.* 2007), and a *tubulin-GAL4* driver line (*y¹w^{*};P{tubP-GAL4}LL7/TM3,Sb¹*) was obtained from the Bloomington *Drosophila* Stock Center (<http://flystocks.bio.indiana.edu>). The *UAS-RNAi* and *GAL4* lines were maintained at the *Drosophila* laboratory at Department of Bioscience, Aarhus University (8000 Aarhus, Denmark) on oatmeal-yeast-sugar-agar medium, 25°, 70% humidity, and a 12-h light-dark cycle).

Quantifying locomotor activity

We used an assay to quantify locomotor activity that relies on video tracking software (Rohde *et al.* 2016b). The DGRP lines were phenotyped at North Carolina State University (Raleigh, North Carolina 27695), and the *UAS-GAL4* RNAi knockdown lines were phenotyped at the Department of Bioscience, Aarhus University (8000 Aarhus, Denmark). Since phenotyping was performed in two different laboratories, there were small experimental differences between the two experimental designs (see details below).

Quantifying locomotor activity for the DGRP lines: The behavioral arenas were constructed in transparent polycarbonate with 4 × 6 behavioral chambers (each of 16 mm in diameter and 6 mm in height) behavioral arena. The locomotor assays were performed in a behavioral room (25°, 70% relative humidity) between 8:00-11:00 AM. The behavioral arenas were illuminated with exogenous light sources. An iPad Mini (Apple, Cupertino, California) was mounted above the behavioral arena to obtain five minutes of video recordings from each of the 24 behavioral chambers. Locomotor activity was then quantified as the total distance traveled for individual flies, extracted from the video files

using the tracking software EthoVision XT (v. 10.0) (Noldus, Wageningen, The Netherlands).

We obtained locomotor activity measurements for approximately 24 individual males for each of 204 DGRP lines. For logistical reasons the DGRP lines were divided into eight blocks, such that each block contained approximately 28 DGRP lines, and every block was assayed over six consecutive days. Two to five day old flies were anesthetized with CO₂ and transferred to the behavioral arenas 16-18 hr prior to the assay. During this time, the flies had access to food, which was removed at the start of the assay.

Quantifying locomotor activity for the gene expression knockdown lines:

The behavioral arenas used for the gene expression knockdown experiments were likewise constructed in transparent polycarbonate, but contained 6 × 6 behavioral chambers (each of 16 mm in diameter and 6 mm in height) per behavioral arena. The behavioral arenas were illuminated from below by a light box (LP400, Dörr, Chesterfield, UK) to ensure high contrast between the flies and the background. An iPad Air (Apple, Cupertino, California) mounted above the behavioral arena was used to obtain ten minutes of video recordings. All behavioral tests were performed in a behavioral room (25°, 70% relative humidity) between 8:00-11:00 AM.

For each *UAS-RNAi* line, and the corresponding control line, approximately 20 virgin females were crossed to five *tubulin-GAL4* males. Approximately 30 F1 male offspring containing the *UAS-GAL4* construct were assayed. The flies being tested were gently moved to the individual behavioral chambers (without anesthetization using an aspirator), and the video recordings were obtained immediately after loading the flies to the behavioral arenas. The observed phenotype of the *UAS-GAL4* offspring was compared to the offspring of the control line crossed to the *GAL4* line using a standard linear model accounting for experimental effects (date and behavioral plates).

Quantitative genetic analysis

To estimate the broad sense heritability (H^2) of locomotor activity in the DGRP, we fitted the mixed model $Y = \mu + d + b + p + L + E$, where

Y_o is the phenotype, μ is the overall mean, d is the fixed effect to account for different measurement days, b is a fixed block effect, p is a fixed plate effect, L is the random line effect, and E is the residual. The broad sense heritability was estimated as $\hat{H}^2 = \frac{\hat{\sigma}_L^2}{\hat{\sigma}_L^2 + \hat{\sigma}_E^2}$, where $\hat{\sigma}_L^2$ and $\hat{\sigma}_E^2$ are the variance components for the line and residual terms. Variance components were estimated using the *lme4* package for R (Bates *et al.* 2015; R Core Team 2017).

Genomic feature models

The following section describes the workflow for genomic prediction utilizing prior biological knowledge. The first step is to link SNPs to the genomic feature classes. The second steps involve computing genomic relationship matrices and performing the prediction. Finally, the genomic variance within predictive feature sets was partitioned to minor units, such as genes. Functions and example scripts are publicly available at <http://psoerensen.github.io/qgg/>.

Genomic data and feature sets: The DGRP genotypes were obtained from <http://dgrp2.gnets.ncsu.edu/>. All genomic analyses were based on segregating biallelic SNPs obtained using the standard filtering process (Mackay *et al.* 2012; Huang *et al.* 2014): SNPs were included if the minor allele frequency ≥ 0.05 , if the Phred quality score (the sequencing quality of a given SNP) > 500 , and if the genotype call rate ≥ 0.8 . This resulted in 1,725,755 SNPs distributed across the six chromosome arms (2R, 2L, 3R, 3L, 4 and X).

The feature sets considered were genes and gene ontology (GO) categories (The Gene Ontology Consortium 2000). First, SNPs were mapped to genes (all SNPs within the open reading frame) using Fly-Base v5.49 annotations of the *D. melanogaster* reference genome (Tweedie *et al.* 2009). Second, genes were aggregated based on gene ontology categories using the BioConductor package *org.Dm.eg.db* (Carlson 2015). A total of 963,235 SNPs was mapped to 10,517 known genes and 1,134 GO terms. The number of SNPs within a single GO term varied from 23 to 163,938 SNPs.

Additive genomic relationship matrices: A central component for predicting trait values using genomic best linear unbiased prediction (GBLUP) is a matrix that captures the genetic similarity between all pairs of individuals. The additive genomic relationship matrix can be computed as $\mathbf{G} = \mathbf{W}\mathbf{W}'/m$ (VanRaden 2008), where m is the number of SNPs on which the relationship matrix is computed, and \mathbf{W} is a centered and scaled genotype matrix. Each column vector of \mathbf{W} is computed as $\mathbf{w}_i = \frac{\mathbf{a}_i - 2p_i}{\sqrt{2p_i(1-p_i)}}$, where p_i is the minor allele frequency of the i -th SNP, and \mathbf{a}_i is the i -th column vector of the allele count matrix, \mathbf{A} , that contains the genotypes coded as 0 and 2 (counting the number of minor alleles).

The common use of GBLUP models is to model a single random genomic effect, thus, assuming that all SNP effect sizes are drawn from a common Gaussian distribution. If including multiple random genomic effects, this assumption can be relaxed by allowing SNP effect sizes to have different distributions. Incorporating multiple random effects requires the computation of additional genomic relationship matrices based on a subset of SNPs, for example, those within a genomic feature (f) and the remaining SNPs not within the feature set (r): $\mathbf{G}_f = \mathbf{W}_f\mathbf{W}_f'/m_f$ and $\mathbf{G}_r = \mathbf{W}_r\mathbf{W}_r'/m_r$.

Genomic prediction: In the general case, the GBLUP (Meuwissen *et al.* 2001) model is written as

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{g} + \mathbf{e}. \text{ (Model 1)}$$

where \mathbf{y} is a vector of phenotypic observations, \mathbf{X} and \mathbf{Z} are design matrices linking fixed (\mathbf{b}) and random genomic effects (\mathbf{g}) to the observations, and the residual effects (\mathbf{e}). Under this model, it is assumed that the observed phenotype is $\mathbf{y} \sim N(\mathbf{X}\mathbf{b}, \mathbf{V})$ where $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}'\sigma_g^2 + \mathbf{I}\sigma_e^2$. A commonly used method to assess the predictive performance is to apply a cross-validation scheme, where a subset of the data are masked. To avoid undesirable data structure in the resampling, it may be helpful to adjust the phenotypes for fixed effects. Model 1 is an animal model with repeated measurements per DGRP line, thus, to retain the replicate data structure the adjusted phenotypic values for the i -th DGRP line was computed as $\tilde{y}_i = \hat{g}_i + \hat{e}_i$ (one \hat{g} per DGRP line, and one \hat{e} per DGRP line per replicate). Thus, the GBLUP model reduces to

$$\tilde{\mathbf{y}} = \mathbf{Z}\mathbf{g} + \mathbf{e}. \text{ (Model 2)}$$

When Model 2 has been fitted on the training data (t), the genomic effects in the validation set (v) can be computed using Equation 1,

$$\hat{\mathbf{g}}_v = \left(\mathbf{G}_{v,t}\hat{\sigma}_g^2 \right) \left[\mathbf{G}_{t,t}\hat{\sigma}_g^2 + \mathbf{I}_{t,t}\hat{\sigma}_e^2 \right]^{-1} \left(\tilde{\mathbf{y}}_t - \hat{\boldsymbol{\mu}}_t \right) \text{ (Equation 1)}$$

Model 2 can be extended to a genomic feature model (GFBLUP, Model 3) by dividing the total genomic effects captured by all SNPs, by the genomic effects captured by SNPs within the feature set (f), and the genomic effects captured by the SNPs not included in the feature set (r) (Edwards *et al.* 2015, 2016; Rohde *et al.* 2017; Sørensen *et al.* 2017),

$$\tilde{\mathbf{y}} = \mathbf{Z}\mathbf{f} + \mathbf{Z}\mathbf{r} + \mathbf{e}. \text{ (Model 3)}$$

The total genomic effects in the validation set can then be computed using Equation 2,

$$\hat{\mathbf{g}}_v = \left(\mathbf{G}_{f,t}\hat{\sigma}_f^2 + \mathbf{G}_{r,t}\hat{\sigma}_r^2 \right) \left[\mathbf{G}_{f,t,t}\hat{\sigma}_f^2 + \mathbf{G}_{r,t,t}\hat{\sigma}_r^2 + \mathbf{I}_{t,t}\hat{\sigma}_e^2 \right]^{-1} \times \left(\tilde{\mathbf{y}}_t - \hat{\boldsymbol{\mu}}_t \right). \text{ (Equation 2)}$$

The predictive performance (PA) of the GBLUP and GFBLUP models was quantified as Pearson's correlation between the observed and predicted genomic values. The models (Model 2 or Model 3) were fitted using 90% of the data, and the estimated genomic parameters were used to predict the genomic values in the remaining 10% of the data. This procedure was repeated 50 times on random subdivisions of the entire data set. This prediction design was chosen as similar prediction studies using the DGRP system has been found usable (Rohde *et al.* 2017; Sørensen *et al.* 2017). A genomic feature model (Model 3) was fitted for each of the genomic feature categories, and the predictive performance of each genomic feature model was compared to the null model (Model 2) by assessing if the predictive ability of the genomic feature model was increased compared to the GBLUP model using Welch's t -test of unequal variance (Welch 1947). Subsequently, all p -values were adjusted for multiple testing using the false discovery rate, and significance level was set to 0.05.

In addition to evaluating the models on predictive performance, the GBLUP and GFBLUP models were also assessed based on estimated genomic parameters. Inferences on the genomic heritability of the models were based on $\hat{h}_{\text{GBLUP}}^2 = \hat{\sigma}_g^2 / (\hat{\sigma}_g^2 + \hat{\sigma}_e^2)$, and $\hat{h}_{\text{GFBLUP}}^2 = (\hat{\sigma}_f^2 + \hat{\sigma}_r^2) / (\hat{\sigma}_f^2 + \hat{\sigma}_r^2 + \hat{\sigma}_e^2)$, as well as by partitioning the genomic variance of the GFBLUP model $\hat{h}_f^2 = (\hat{\sigma}_f^2) / (\hat{\sigma}_f^2 + \hat{\sigma}_r^2)$ and $\hat{h}_r^2 = (\hat{\sigma}_r^2) / (\hat{\sigma}_f^2 + \hat{\sigma}_r^2)$. These ratios quantify the proportion of total genomic variance captured by (\hat{h}_f^2), and not captured by (\hat{h}_r^2), the SNPs in the feature set.

Estimating the variance components in Model 2 and Model 3 was performed using the average information restricted maximum likelihood (AI-REML) procedure (Madsen *et al.* 1994; Johnson and Thompson 1995) as implemented in DMU software. We have developed an R interface that enables users to perform analysis within R that otherwise rely on DMU (DMU can be downloaded from <http://dmu.agrsci.dk/DMU/>). Our R package qgg is accessible at <http://psoerensen.github.io/qgg/>, including examples on how to perform the genomic feature analyses.

Partitioning of genomic variance to gene level: To partition the genomic variance of a predictive GO category to genomic variance at the gene level we adapted the covariance association test (CVAT) (Rohde *et al.* 2016a; Sørensen *et al.* 2017).

The CVAT method was originally developed as a set test approach that captures the covariance between the total genomic effects from all markers and the genomic effects from the markers within the feature set (Rohde *et al.* 2016a). Here, we instead considered the covariance between the genomic effects of a GO term ($\hat{\mathbf{g}}_{GO}$) and the genomic effects at gene level within that particular GO term ($\hat{\mathbf{g}}_{gene}$),

$$T_{CVAT} = \hat{\mathbf{g}}_{GO} \hat{\mathbf{g}}_{gene}^{-1} \quad (\text{Equation 3})$$

where $\hat{\mathbf{g}}_{GO}$ are the genomic feature effects estimated from Model 3, and $\hat{\mathbf{g}}_{gene} = \sum_{i=1}^{m_{gene}} \mathbf{w}_i \hat{\mathbf{s}}_{GO_i}$. The vector of SNP effects, $\hat{\mathbf{s}}_{GO_i} = \mathbf{W}_{GO}^{-1} (\mathbf{W}_{GO} \mathbf{W}_{GO}^{-1})^{-1} \hat{\mathbf{g}}_{GO_i}$, where \mathbf{W}_{GO} corresponds to the centered and scaled genotype matrix of the SNPs within one particular GO term. To determine the degree significance an empirical distribution of T_{CVAT} was obtained based on a circular permutation approach (Cabrera *et al.* 2012), where the genome was considered circular in order to retain the same order of SNPs but receive new SNP effects in each permutation. This decouples the association between the SNP and the genomic feature, but retains the correlation structure among the SNP effects. In each iteration of the permutation approach a new T_{CVAT} statistic was computed (repeated 10,000 times), and the p -value was computed as a one-tailed test of the proportion of the randomly sampled summary statistics being larger than the observed summary statistic (see Rohde *et al.* (2016a) and Sørensen *et al.* (2017) for additional details).

Marginal SNP analysis

The CVAT results were compared to a standard marginal SNP analysis. Single marker associations evaluate the association between each segregating SNP and the trait variation. In order to account for the experimental fixed effects and the genetic similarity among DGRP lines the estimated genomic effects ($\hat{\mathbf{g}}$, from Model 2) was used as response variable. The marginal SNP analysis was a t -test on the regression coefficient from the regression of $\hat{\mathbf{g}}$ on each segregating SNP in the DGRP, *i.e.*, a total of 1,725,755 regression analyses.

Data availability

The DGRP genotypes can be accessed via the DGRP2 website <http://dgrp2.gnets.ncsu.edu/>, and the phenotypic data are given in Table S1. Supplemental material available at Figshare: <https://doi.org/10.25387/g3.5951581>.

RESULTS AND DISCUSSION

We quantified male locomotor activity in 204 DGRP lines using video tracking to measure the total distance traveled per individual in the course of five minutes. We found substantial genetic variation (Table S2)

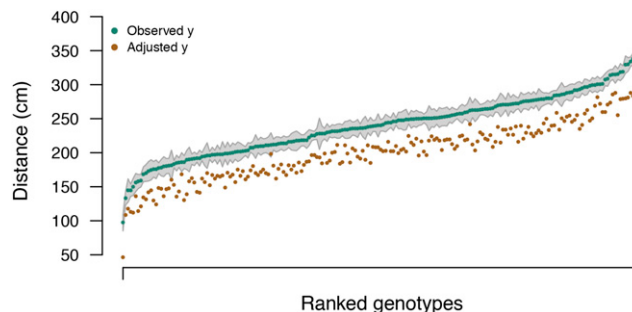


Figure 2 Rank-ordered mean for locomotor activity (green) and the corresponding mean values adjusted for experimental fixed effects (brown). Shaded area depicts the standard error of the mean.

in locomotor activity, with an approximate fourfold difference between the least active and the most active DGRP lines (Figure 2, Table S1). The broad sense heritability for male locomotor activity was $\hat{H}^2 = 0.40$ (SE = 0.03). We estimated the proportion of total phenotypic variation explained by common variants (MAF > 0.05) using the additive genomic relationship matrix as $\hat{h}^2 = 0.26$ (SE = 0.02); thus, 65% (0.26 / 0.40) of the total genetic variation was captured by common, additive variants. The estimated broad sense heritability is in the range of other estimates of *D. melanogaster* locomotor activity (Jordan *et al.* 2006, 2007).

We used the GBLUP model (Model 2) to predict the genomic values for locomotor activity by estimating genomic parameters on 90% of the data, and using those parameters to predict the genomic values in the remaining 10% of the data (Equation 1). The validation sets were chosen randomly and this procedure was repeated a total of 50 times on different training and validation sets. The GBLUP model uses all available SNPs assuming the effects are drawn from a common Gaussian distribution. The performance of the model was quantified as the correlation (r) between the predicted and observed phenotypic values in the validation set. We found low mean predictive ability (PA \pm SEM) for the GBLUP model (PA = 0.12 \pm 0.033). The maximum predictive ability of line means is $H^2 = r^2 = \text{PA}^2$ (Mrode 2005; Goddard 2009). The heritability based on line means can be approximated as $\hat{H}^2 = \hat{\sigma}_g^2 / (\hat{\sigma}_g^2 + \frac{\hat{\sigma}_e^2}{\bar{n}})$ (Mackay and Huang 2017), where $\hat{\sigma}_g^2$ and $\hat{\sigma}_e^2$, respectively, are the among-line and within-line variance of the individual data, and \bar{n} is the average number of flies scored per DGRP line (here $\bar{n} = 23$). The broad sense heritability of line means was $\hat{H}^2 = 0.94$, thus, the GBLUP model only accounts for 0.014/0.94 = 1.5% of the observed heritability of line means. Thus, assuming all SNP effects to be from a common Gaussian distribution resulted in a low proportion of the heritability explained, in agreement with a similar study on *D. melanogaster* aggressive behavior (Rohde *et al.* 2017).

The one component GBLUP model assumes all SNP effects are from a common Gaussian distribution; however, this assumption is not likely to be true (Speed and Balding 2014). A relaxation of the Gaussian assumption can be obtained by fitting multiple random components, such as the GFBLUP models (Model 3), by allowing the SNP effects within those components to have different effect sizes (small-moderate-large). An example of this was shown by Rohde *et al.* (2016a), where the SNPs were partitioned according to minor allele frequencies to obtain different distributions of SNP effects. Here, we build genomic relationship matrices based on SNPs within GO categories. For each GO term the GFBLUP model was fitted, and the predictive abilities were compared to the predictive performance of the GBLUP model. The five GO

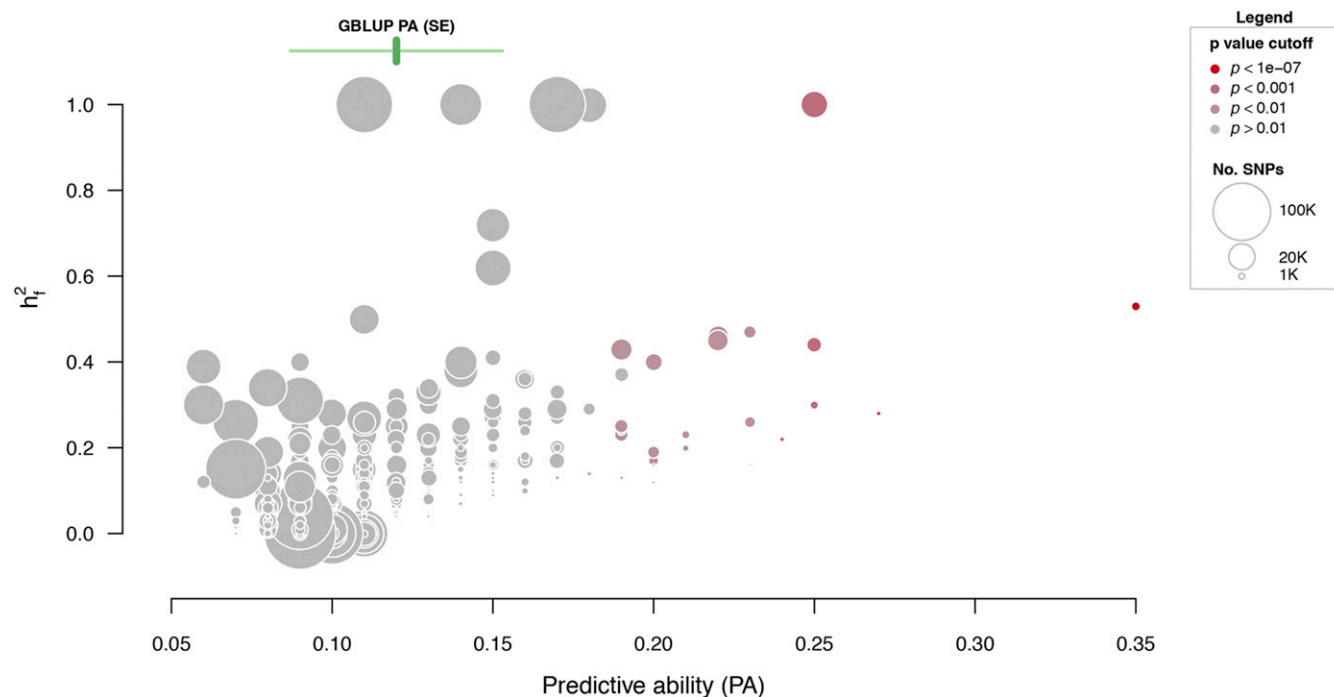


Figure 3 Results from the GFBLUP models. Each point corresponds to one GO term that is plotted within the space of genomic variance explained (h^2) and predictive ability (PA). The size of each point relates to the number of SNPs within the GO term, and the color indicates the p -value of increased predictive ability compared to the GBLUP model. The mean predictive ability \pm standard error (SE) of the GBLUP model is indicated with green vertical and horizontal lines, respectively.

terms with the highest predictive abilities are shown in Table 1 (the full list is given in Table S3).

When we jointly examined several model parameters, interesting patterns emerged (Figure 3). GO terms with high PA also tended to explain a larger fraction of the genomic variance (h^2). GO terms with many SNPs do not have higher predictive abilities, or capture more of the genomic variance. Instead, large GO categories, *i.e.*, those that contain many SNPs, tend to explain the least genomic variance (Figure 3). This is probably a consequence of too many non-causal SNPs in the feature set, which adds noise to the model. Four non-significant and one marginally significant GO term explain 100% of the genomic variance (Figure 3). Explaining 100% of the variance when the analysis is based on a small proportion of all genomic markers is naturally an overestimation. This can arise if two genomic relationship matrices are very similar, because then it is likely that parts of the genomic variance will be captured by only one of the components, thereby leading to overestimation.

Next, we considered the GO term that increased the predictive ability significantly compared to the GBLUP model in more detail (mean PA = 0.35 ± 0.026 , Table 1). The predictive GO term GO:0022857 contains

genes involved in transmembrane transport. In the DGRP genotype data the GO term GO:0022857 contained 59 genes and 2,563 biallelic SNPs (at MAF > 0.05). Partitioning the genomic variance between the SNPs within GO:0022857 and SNPs not located within genes related to GO:0022857, the GFBLUP model accounts for 13% of the heritability compared to the standard GBLUP model that only accounted for 1.5%. Thus, allowing for differential weight on the SNP effects increased the predictive performance, and therefore increased how much of the heritability the GFBLUP model accounted for. This pattern is similar to the observation in Rohde *et al.* (2017) for aggressive behavior in the DGRP. We then partitioned the genomic variance within that GO term among the 59 genes using CVAT. This method considers the covariance between the total genomic effects of the GO term and the genomic effects of the genes within the GO term (Equation 3). The resulting statistic is a p -value indicating if the proportion of genomic variance explained by the gene is larger than a randomly sampled set of SNPs containing the same number of SNPs as the gene being considered. A total of 15 genes had a p -value < 0.05, indicating that these genes capture a larger proportion of the total genomic variance within the predictive GO term than a random set

Table 1 The top five GO terms with highest predictive ability (PA). For each GO term the following information is listed: Number of genes (No. genes) and SNPs (No. SNPs) within the GO term, the mean PA with standard errors (SE), the raw (p) and adjusted p -values (by false discovery rate (FDR)) for increased predictive performance compared to the GBLUP model, and the proportion of genomic variance explained by the GO term (h^2)

GO term	No. genes	No. SNPs	PA \pm SE	p -value	FDR p -value	h^2
1. GO:0022857	59	2563	0.35 ± 0.026	6.8×10^{-6}	7.7×10^{-5}	0.53
2. GO:0006730	17	749	0.27 ± 0.029	2.1×10^{-4}	1.2×10^{-1}	0.28
3. GO:0006810	80	6893	0.25 ± 0.028	9.5×10^{-4}	3.6×10^{-1}	0.44
4. GO:0055114	368	22029	0.25 ± 0.029	1.7×10^{-3}	3.9×10^{-1}	1.00
5. GO:0030866	21	2161	0.25 ± 0.027	1.3×10^{-3}	3.7×10^{-1}	0.30

1: transmembrane transporter activity; 2: one-carbon metabolic process; 3: transport; 4: oxidation-reduction process; 5: cortical actin cytoskeleton organization.

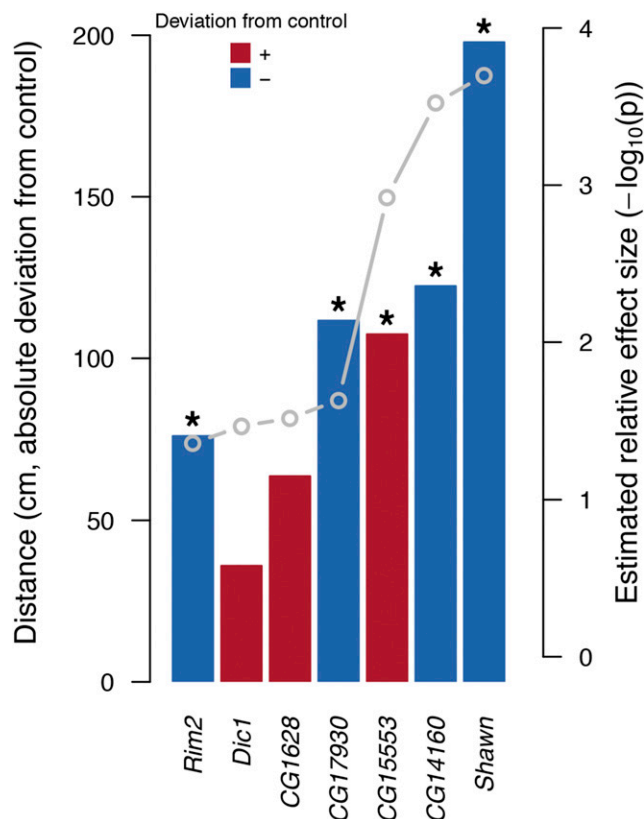


Figure 4 Effects on distance moved by gene expression knockdown expressed as absolute deviations from control line (red: increased activity; blue: decreased activity) ranked according to estimated effect size. Asterisks (*) indicate statistical significance (*i.e.*, $p < 0.05$) between knockdown strain and control strain. Gray points (belongs to the right y-axis) show the estimated relative effect size within the predictive GO term, here illustrated as $-\log_{10}(p)$ from CVAT.

of SNPs within that GO term (Table S4). We compared this result with the results from the marginal SNP analysis where no SNPs passed the genome-wide significance threshold (Figure S1). The SNP p -values of the genetic markers located within the 15 CVAT associated genes ranged from 1.45×10^{-2} to 4.35×10^{-6} . The majority of the CVAT associated genes had SNP p -values around 1×10^{-3} (Figure S1); thus, these genes would not have been identified by the marginal SNP analysis. This discrepancy in results was expected because the marginal SNP analysis picks up individual SNPs with the largest effects, whereas CVAT evaluates the joint effect of multiple genomic markers and can therefore detect SNPs that individually have small effects.

Given our list of 15 candidate genes potentially affecting locomotor activity, we set out to functionally validate these genes by investigating the phenotypic consequence of gene expression knockdown in adult flies using the bipartite *UAS-GAL4* system. Only 11/15 genes were available with the desired genetic background, of which seven lines produced viable offspring after crossing to the ubiquitous *GAL4*-driver. Thus, a total of seven genes were assessed for their effect of gene expression knockdown on locomotor activity; *CG1628*, *CG14160*, *CG15553*, *CG17930*, *Dic1*, *Rim2* and *Shawn*. Five of the seven tested knockdown lines resulted in significant locomotor deviations from the control line with the same genetic background (Figure 4). The gene expression knockdown resulted in offspring becoming both more (*CG15553*) and less active (*Rim2*, *CG17930*, *CG14160*, *Shawn*, Figure 4) than the

respective control line, indicating that the knockdown lines do not in general suffer strongly from the gene expression knockdown. Importantly, the correlation between the absolute effect size of gene expression knockdown and degree of genomic variance explained was very high, $\rho = 0.91$ (p -value = 0.005, Figure 4). Thus, we not only validated the functional effects of the candidate genes for locomotor activity, but also provided functional evidence supporting the success of our method for identifying a restricted set of important genes ranked by their effect sizes from a larger set of potential candidate genes.

The genes *CG14160*, *CG15553*, and *CG17930* have not previously been phenotypically annotated in *D. melanogaster*, thus, here we provide first evidence that the genes are involved in explaining variation in a behavioral phenotype. The gene *Shawn* is a mitochondrial carrier in the *Drosophila* nervous system (Slabbaert *et al.* 2016), and *Rim2* encodes a deoxynucleotide transporter located within the mitochondria. Both *Rim2* and *Shawn* have conserved human homologous gene sequences, *SLC25A36* and *SLC25A40*, respectively, which have been found to contain susceptibility loci for bipolar disorder (Winham *et al.* 2014) and epilepsy (Sirén *et al.* 2010).

Fruit flies and mammals have a common evolutionary origin of basic biological processes, including development of the nervous system (Adams *et al.* 2000), and approximately 75% of human disease genes have at least one homologous gene in *D. melanogaster* (Reiter *et al.* 2001). Human neurological diseases, *e.g.*, Parkinson's and Huntington's diseases, are associated with locomotor deficits, whereas some neuropsychiatric disorders, *e.g.*, attention-deficit/hyperactivity disorder and depression, are associated with changes in activity levels (American Psychiatric Association 2013). Therefore, understanding the genetic architecture of locomotor activity in model organisms might also provide an important link to human health.

For example, Parkinson's disease has been shown to be linked to degeneration of certain dopaminergic neurons (Olanow and Tatton 1999), and dopamine has been shown to affect locomotion in fruit flies (Connolly *et al.* 1971; Jordan *et al.* 2006; Riemensperger *et al.* 2011; van der Voet *et al.* 2015), and mice (Garland *et al.* 2011). The fact that both *Rim2* and *Shawn* have conserved human homologous gene sequences, *SLC25A36* and *SLC25A40*, respectively, which have been found to contain susceptibility loci for bipolar disorder (Winham *et al.* 2014) and epilepsy (Sirén *et al.* 2010) illustrate the potential of using *D. melanogaster* as a model organism to study complex human psychological and behavioral disorders.

In conclusion, we provide functional support both for the candidate genes detected by CVAT, and for the ranking of effect sizes suggested by CVAT. These results are important because they provide evidence for the two challenges relating to GFM analyses; namely the need to have an efficient method to rank genes within a larger set of associated genes, and to perform biological validation of the genomic findings from GFM. Thus, these results demonstrate that the findings from the GFM analyses not are statistical artifacts, but indeed have biological relevance.

ACKNOWLEDGMENTS

This work was partly funded by grants from the Lundbeck Foundation (R155-2014-1724), Centre for Integrative Sequencing at Aarhus University, the Danish Strategic Research Council (GenSAP: Centre for Genomic Selection in Animals and Plants, contract no. 12-132452), the Danish Natural Science Research Council (Sapere Aude grant to TNK), the National Institutes of Health (R01-AA016560 and R01-AG043490 to TFCM) and ECO-FCE which is funded by the European Union Seventh Framework Program (FP7/2007–2013) under grant agreement no. 311794. The authors thank

Doth Andersen for technical assistance in the *Drosophila* laboratory at Aarhus University, Denmark.

LITERATURE CITED

- Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. *Science* 287(5461): 2185–2195. <https://doi.org/10.1126/science.287.5461.2185>
- American Psychiatric Association, 2013 Diagnostic and Statistical Manual of Mental Disorders, Ed. 5th. American Psychiatric Publishing, Arlington, VA.
- Anholt, R. R. H., and T. F. C. Mackay, 2018 The road less traveled: From genotype to phenotype in flies and humans. *Mamm. Genome* 29: 5. <https://doi.org/10.1007/s00335-017-9722-7>
- Atwell, S., Y. S. Huang, B. J. Vilhja, G. Willems, M. Horton *et al.*, 2010 Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465(7298): 627–631. <https://doi.org/10.1038/nature08800>
- Bahrndorff, S., S. Gertsen, C. Pertoldi, and T. N. Kristensen, 2016 Investigating thermal acclimation effects before and after a cold shock in *Drosophila melanogaster* using behavioural assays. *Biol. J. Linn. Soc. Lond.* 117(2): 241–251. <https://doi.org/10.1111/bij.12659>
- Balding, D. J., 2006 A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* 7(10): 781–791. <https://doi.org/10.1038/nrg1916>
- Bates, D., M. Mächler, B. M. Bolker, and S. C. Walker, 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67(1): 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Burnet, B., L. Burnet, K. Connolly, and N. Wihiamson, 1988 A genetic analysis of locomotor activity in *Drosophila melanogaster*. *Genetics* 61: 111–119.
- Cabrera, C. P., P. Navarro, J. E. Huffman, A. F. Wright, C. Hayward *et al.*, 2012 Uncovering networks from genome-wide association studies via circular genomic permutation. *G3 Genes, Genomes. Genet.* 2: 1067–1075. <https://doi.org/10.1534/g3.112.002618>
- Carlson, M., 2015 org.Dm.eg.db: Genome wide annotation for Fly. (R package version 3.2.3).
- Carpenter, F. W., 1905 The reactions of the pomace fly (*Drosophila ampelophila* Loew) to light, gravity, and mechanical stimulation. *Am. Nat.* 39(459): 157–171. <https://doi.org/10.1086/278502>
- Colomb, J., L. Reiter, J. Blaszkiewicz, J. Wessnitzer, and B. Brembs, 2012 Open source tracking and analysis of adult *Drosophila* locomotion in Buridan's paradigm with and without visual targets. *PLoS One* 7(11): 1–12. <https://doi.org/10.1371/annotation/41b2d3fd-e816-420c-80d0-88290796b1cd>
- Connolly, K., G. Tunnicliff, and J. T. Rick, 1971 The effects of gamma-hydrobutyric acid on spontaneous locomotor activity and dopamine levels in a selected strain of *Drosophila melanogaster*. *Comp. Biochem. Physiol. Part B* 40(2): 321–326. [https://doi.org/10.1016/0305-0491\(71\)90216-1](https://doi.org/10.1016/0305-0491(71)90216-1)
- Dekkers, J., 2012 Application of genomics tools to animal breeding. *Curr. Genomics* 13: 207–212. <https://doi.org/10.2174/138920212800543057>
- Dietzl, G., D. Chen, F. Schnorrer, K.-C. Su, Y. Barinova *et al.*, 2007 A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448(7150): 151–156. <https://doi.org/10.1038/nature05954>
- Edwards, S. M., I. F. Sørensen, P. Sarup, T. F. C. Mackay, and P. Sørensen, 2016 Genomic prediction for quantitative traits is improved by mapping variants to gene ontology categories in *Drosophila melanogaster*. *Genetics* 203(4): 1871–1883. <https://doi.org/10.1534/genetics.116.187161>
- Edwards, S. M., B. Thomsen, P. Madsen, and P. Sørensen, 2015 Partitioning of genomic variance reveals biological pathways associated with udder health and milk production traits in dairy cattle. *Genet. Sel. Evol.* 47(1): 60. <https://doi.org/10.1186/s12711-015-0132-6>
- Ehsani, A., L. Janss, D. Pomp, and P. Sørensen, 2015 Decomposing genomic variance using information from GWA, GWE and eQTL analysis. *Anim. Genet.* 47(2): 165–173. <https://doi.org/10.1111/age.12396>
- Falconer, D. S., and T. F. C. Mackay, 1996 Introduction to Quantitative Genetics, Ed. 4th. Longman Group, Harlow, Essex.
- Fang, L., G. Sahana, P. Ma, G. Su, Y. Yu *et al.*, 2017a Exploring the genetic architecture and improving genomic prediction accuracy for mastitis and milk production traits in dairy cattle by mapping variants to hepatic transcriptomic regions responsive to intra-mammary infection. *Genet. Sel. Evol.* 49(1): 44. <https://doi.org/10.1186/s12711-017-0319-0>
- Fang, L., G. Sahana, G. Su, Y. Yu, S. Zhang *et al.*, 2017b Integrating sequence-based GWAS and RNA-seq provides novel insights into the genetic basis of mastitis and milk production in dairy cattle. *Sci. Rep.* 7: 45560. <https://doi.org/10.1038/srep45560>
- Garbe, D. S., W. L. Bollinger, A. Vigderman, P. Masek, J. Gertowski *et al.*, 2015 Context-specific comparison of sleep acquisition systems in *Drosophila*. *Biol. Open* 4(11): 1558–1568. <https://doi.org/10.1242/bio.013011>
- Gargano, J. W., I. Martin, P. Bhandari, and M. S. Grotewiel, 2005 Rapid iterative negative geotaxis (RING): A new method for assessing age-related locomotor decline in *Drosophila*. *Exp. Gerontol.* 40(5): 386–395. <https://doi.org/10.1016/j.exger.2005.02.005>
- Garland, T., H. Schutz, M. A. Chappell, B. K. Keeney, T. H. Meek *et al.*, 2011 The biological control of voluntary exercise, spontaneous physical activity and daily energy expenditure in relation to obesity: human and rodent perspectives. *J. Exp. Biol.* 214(2): 206–229. <https://doi.org/10.1242/jeb.048397>
- Gilestro, G. F., 2012 Video tracking and analysis of sleep in *Drosophila melanogaster*. *Nat. Protoc.* 7(5): 995–1007. <https://doi.org/10.1038/nprot.2012.041>
- Goddard, M., 2009 Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica* 136(2): 245–257. <https://doi.org/10.1007/s10709-008-9308-0>
- Hardy, J., and A. Singleton, 2009 Genomewide association studies and human disease. *N. Engl. J. Med.* 360(17): 1759–1768. <https://doi.org/10.1056/NEJMra0808700>
- Huang, W., A. Massouras, Y. Inoue, J. Peiffer, M. Ràmia *et al.*, 2014 Natural variation in genome architecture among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24(7): 1193–1208. <https://doi.org/10.1101/gr.171546.113>
- Husby, A., T. Kawakami, L. Ro, H. Ellegren, and A. Qvarnstrom, 2015 Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. *Proc. Biol. Sci.* 282(1806): 20150156. <https://doi.org/10.1098/rspb.2015.0156>
- Johnson, D. L., and R. Thompson, 1995 Restricted maximum likelihood estimation of variance components for univariate animal models using sparse matrix techniques and average information. *J. Dairy Sci.* 78(2): 449–456. [https://doi.org/10.3168/jds.S0022-0302\(95\)76654-1](https://doi.org/10.3168/jds.S0022-0302(95)76654-1)
- Jordan, K. W., M. A. Carbone, A. Yamamoto, T. J. Morgan, and T. F. C. Mackay, 2007 Quantitative genomics of locomotor behavior in *Drosophila melanogaster*. *Genome Biol.* 8(8): R172. <https://doi.org/10.1186/gb-2007-8-8-r172>
- Jordan, K. W., T. J. Morgan, and T. F. C. Mackay, 2006 Quantitative trait loci for locomotor behavior in *Drosophila melanogaster*. *Genetics* 174(1): 271–284. <https://doi.org/10.1534/genetics.106.058099>
- Konopka, R. J., and S. Benzer, 1971 Clock mutants of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 68(9): 2112–2116. <https://doi.org/10.1073/pnas.68.9.2112>
- Lage, K., S. C. Greenway, J. A. Rosenfeld, H. Wakimoto, and J. M. Gorham, 2012 Genetic and environmental risk factors in congenital heart disease functionally converge in protein networks driving heart development. *Proc. Natl. Acad. Sci. USA* 109: 14035–14040. <https://doi.org/10.1073/pnas.1210730109>
- Lango Allen, H., K. Estrada, G. Lettre, S. I. Berndt, M. N. Weedon *et al.*, 2010 Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467(7317): 832–838. <https://doi.org/10.1038/nature09410>
- de Leeuw, C. A., B. M. Neale, T. Heskes, and D. Posthuma, 2016 The statistical properties of gene-set analysis. *Nat. Rev. Genet.* 17(6): 353–364. <https://doi.org/10.1038/nrg.2016.29>
- Lightfoot, J. T., L. Leamy, D. Pomp, M. J. Turner, A. A. Fodor *et al.*, 2010 Strain screen and haplotype association mapping of wheel running

- in inbred mouse strains. *J. Appl. Physiol.* 109(3): 623–634. <https://doi.org/10.1152/jappphysiol.00525.2010>
- Lightfoot, J. T., M. J. Turner, M. Daves, A. Vordermark, and S. R. Kleeberger, 2004 Genetic influence on daily wheel running activity level. *Physiol. Genomics* 19(3): 270–276. <https://doi.org/10.1152/physiolgenomics.00125.2004>
- Lightfoot, J. T., M. J. Turner, D. Pomp, S. R. Kleeberger, and L. J. Leamy, 2008 Quantitative trait loci for physical activity traits in mice. *Physiol. Genomics* 32(3): 401–408. <https://doi.org/10.1152/physiolgenomics.00241.2007>
- Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits*, Sinauer Associates, Sunderland, MA.
- Mackay, T. F. C., and W. Huang, 2018 *Charting the genotype-phenotype map: lessons from the Drosophila melanogaster Genetic Reference Panel*. Wiley Interdiscip. Rev. Dev. Biol. 7: e289. <https://doi.org/10.1002/wdev.289>
- Mackay, T. F. C., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* Genetic Reference Panel. *Nature* 482(7384): 173–178. <https://doi.org/10.1038/nature10811>
- Madsen, P., J. Jensen, and R. Thompson, 1994 Estimation of (co)-variance components by REML in multivariate mixed linear models using average of observed and expected information, pp. 455–462 in *Fifth World Congress of Genetics Applied to Livestock Production*, Guelph, Ontario, Canada.
- Maurano, M. T., R. Humbert, E. Rynes, R. E. Thurman, E. Haugen *et al.*, 2012 Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337(6099): 1190–1195. <https://doi.org/10.1126/science.1222794>
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard, 2001 Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819–1829.
- Mooney, M. A., J. T. Nigg, S. K. McWeeney, and B. Wilmot, 2014 Functional and genomic context in pathway analysis of GWAS data. *Trends Genet.* 30(9): 390–400. <https://doi.org/10.1016/j.tig.2014.07.004>
- Mrode, R. A., 2005 *Linear Models for the Prediction of Animal Breeding Values*, CABI Publishing, Wallingford, UK. <https://doi.org/10.1079/9780851990002.0000>
- Nagamine, Y., R. Pong-Wong, P. Navarro, V. Vitart, C. Hayward *et al.*, 2012 Localising loci underlying complex trait variation using regional genomic relationship mapping. *PLoS One* 7(10): e46501. <https://doi.org/10.1371/journal.pone.0046501>
- O’Roak, B. J., L. Vives, S. Girirajan, E. Karakoc, N. Krumm *et al.*, 2012 Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485(7397): 246–250. <https://doi.org/10.1038/nature10989>
- Olanow, C. W., and W. G. Tatton, 1999 Etiology and pathogenesis of Parkinson’s disease. *Annu. Rev. Neurosci.* 22(1): 123–144. <https://doi.org/10.1146/annurev.neuro.22.1.123>
- Osborne, K. A., 1997 Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* 277(5327): 834–836. <https://doi.org/10.1126/science.277.5327.834>
- Pfeiffenberger, C., B. C. Lear, K. P. Keegan, and R. Allada, 2010 Locomotor activity level monitoring using the *Drosophila* activity monitoring (DAM) system. *Cold Spring Harb. Protoc.* 5: 1238–1242. <https://doi.org/10.1101/pdb.prot5518>
- R Core Team, 2017 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Reiter, L. T., L. Potocki, S. Chien, M. Gribskov, and E. Bier, 2001 A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11(6): 1114–1125. <https://doi.org/10.1101/gr.169101>
- Riemensperger, T., G. Isabel, H. Coulom, K. Neuser, L. Seugent *et al.*, 2011 Behavioral consequences of dopamine deficiency in the (*Drosophila*) central nervous system. *Proc. Natl. Acad. Sci. USA* 108(2): 834–839. <https://doi.org/10.1073/pnas.1010930108>
- Rohde, P. D., D. Demontis, B. C. Cuyabano *et al.*, 2016a Covariance Association Test (CVAT) identifies genetic markers associated with schizophrenia in functionally associated biological processes. *Genetics* 203(4): 1901–1913. <https://doi.org/10.1534/genetics.116.189498>
- Rohde, P. D., B. Gaertner, K. Ward, P. Sørensen, and T. F. C. Mackay, 2017 Genomic analysis of genotype-by-social environment interaction for *Drosophila melanogaster*. *Genetics* 206(4): 1969–1984. <https://doi.org/10.1534/genetics.117.200642>
- Rohde, P. D., L. S. Madsen, S. M. Neumann Arvidson, V. Loeschcke, D. Demontis *et al.*, 2016b Testing candidate genes for attention-deficit/hyperactivity disorder in fruit flies using a high throughput assay for complex behavior. *Fly (Austin)* 10(1): 25–34. <https://doi.org/10.1080/19336934.2016.1158365>
- Rosato, E., and C. P. Kyriacou, 2006 Analysis of locomotor activity rhythms in *Drosophila*. *Nat. Protoc.* 1(2): 559–568. <https://doi.org/10.1038/nprot.2006.79>
- Sarup, P., J. Jensen, T. Ostensen, M. Henryon, and P. Sørensen, 2016 Increased prediction accuracy using a genomic feature model including prior information on quantitative trait locus regions in purebred Danish Duroc pigs. *BMC Genet.* 17(1): 11. <https://doi.org/10.1186/s12863-015-0322-9>
- Sirén, A., A. Polvi, L. Chahine, M. Labuda, S. Bourgoin *et al.*, 2010 Suggestive evidence for a new locus for epilepsy with heterogeneous phenotypes on chromosome 17q. *Epilepsy Res.* 88(1): 65–75. <https://doi.org/10.1016/j.epilepsyres.2009.09.022>
- Slabbaert, J. R., S. Kuenen, J. Swerts, I. Maes, V. Uytterhoeven *et al.*, 2016 Shawn, the *Drosophila* homolog of SLC25A39/40, is a mitochondrial carrier that promotes neuronal survival. *J. Neurosci.* 36(6): 1914–1929. <https://doi.org/10.1523/JNEUROSCI.3432-15.2016>
- Speed, D., and D. J. Balding, 2014 MultiBLUP: Improved SNP-based prediction for complex traits. *Genome Res.* 24(9): 1550–1557. <https://doi.org/10.1101/gr.169375.113>
- Swallow, J. G., P. A. Carter, and T. Garland, 1998 Artificial selection for increased wheel running behavior in house mice. *Behav. Genet.* 28(3): 227–237. <https://doi.org/10.1023/A:1021479331779>
- Sørensen, I. F., S. M. Edwards, P. D. Rohde, and P. Sørensen, 2017 Multiple trait covariance association test identifies gene ontology categories associated with chill coma recovery time in *Drosophila melanogaster*. *Sci. Rep.* 7(1): 2413. <https://doi.org/10.1038/s41598-017-02281-3>
- The Gene Ontology Consortium, 2000 Gene ontology: Tool for the identification of biology. *Nat. Genet.* 25(1): 25–29. <https://doi.org/10.1038/75556>
- Turner, M. J., S. R. Kleeberger, J. T. Lightfoot, M. J. Turner, S. R. Kleeberger *et al.*, 2005 Influence of genetic background on daily running-wheel activity differs with aging Influence of genetic background on daily running-wheel activity differs with aging. 76–85.
- Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton *et al.*, 2009 FlyBase: Enhancing *Drosophila* gene ontology annotations. *Nucleic Acids Res.* 37(Database): D555–D559. <https://doi.org/10.1093/nar/gkn788>
- Uemoto, Y., R. Pong-Wong, P. Navarro, V. Vitart, C. Hayward *et al.*, 2013 The power of regional heritability analysis for rare and common variant detection: Simulations and application to eye biometrical traits. *Front. Genet.* 4: 232. <https://doi.org/10.3389/fgene.2013.00232>
- VanRaden, P. M., 2008 Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91(11): 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Visscher, P. M., M. A. Brown, M. I. McCarthy, and J. Yang, 2012 Five Years of GWAS Discovery. *Am. J. Hum. Genet.* 90(1): 7–24. <https://doi.org/10.1016/j.ajhg.2011.11.029>
- van der Voet, M., B. Harich, B. Franke, and A. Schenck, 2015 ADHD-associated dopamine transporter, latrophilin and neurofibromin share a dopamine-related locomotor signature in *Drosophila*. *Mol. Psychiatry* 10: 1–9.
- Welch, B. L., 1947 The generalization of “Student’s” problem when several different population variances are involved. *Biometrika* 34: 28–35.
- Winham, S. J., A. B. Cuellar-Barboza, A. Oliveros, S. L. Mcelroy, S. Crow *et al.*, 2014 Genome-wide association study of bipolar disorder

- accounting for effect of body mass index identifies a new risk allele in TCF7L2. *Mol. Psychiatry* 19(9): 1010–1016. <https://doi.org/10.1038/mp.2013.159>
- Xiao, Y., H. Liu, L. Wu, M. Warburton, and J. Yan, 2017 Genome-wide association studies in maize: Praise and stargaze. *Mol. Plant* 10(3): 359–374. <https://doi.org/10.1016/j.molp.2016.12.008>
- Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders *et al.*, 2010 Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42(7): 565–569. <https://doi.org/10.1038/ng.608>
- Yang, J., T. A. Manolio, L. R. Pasquale, E. Boerwinkle, N. Caporaso *et al.*, 2011 Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* 43(6): 519–525. <https://doi.org/10.1038/ng.823>
- Zimmerman, J. E., D. M. Raizen, M. H. Maycock, G. Maislin, and A. I. Pack, 2008 A video method to study *Drosophila* sleep. *Sleep* 31(11): 1587–1598. <https://doi.org/10.1093/sleep/31.11.1587>

Communicating editor: Y. Kim