Testing implications of the polygenic model for the genetic analysis of loci identified through genome-wide association

Wenyu Zhang, R. Guy Reeves, Diethard Tautz#

Address for all authors:

Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany

#corresponding author:

Email: tautz@evolbio.mpg.de

Phone: +49 4522763390

Running title: Test of the polygenic model

Key words: *Drosophila melanogaster*, morphology, automated phenotyping, DGRP lines, genome wide association mapping, omnigenic model

Impact statement:

Quantitative traits are similarly likely influenced by randomly picked loci as by loci identified in a genome-wide association study.

Abstract

It has been proposed that many loci with no significant association in GWA studies can nonetheless contribute to the phenotype through modifier interactions with the core genes, implying a polygenic determination of quantitative traits. We have tested this hypothesis by using *Drosophila* pupal phenotypes. We identified candidate genes for pupal length determination in a GWA and show for disrupted versions of the genes that most are indeed involved in the phenotype, presumably forming a core pathway. We then randomly chose genes below the GWA threshold and found that three quarters of them had also an effect on the trait. We further tested the effects of these knockout lines on an independent behavioral pupal trait (pupation site choice) and found that a similar, but non-correlated fraction of them had a significant effect as well. Our data thus confirm the prediction that a large number of genes can influence independent quantitative traits.

Introduction

Organismal phenotypes, *i.e.*, traits measured in whole organisms usually have a quantitative distribution and their genetic architecture can be studied by genome-wide association (GWA) approaches. In the past years, these approaches have suggested that such phenotypes have a polygenic architecture in the sense that many genes of moderate or small effect contribute to the phenotype (Wood et al., 2014; Yang et al., 2011, 2010). Hence, the attention has turned towards the loci falling below this cut-off. A study focussing on these small effect loci has suggested that all, or at least almost all, genes can be expected to contribute to the phenotype. This has led to the suggestion of an omnigenic model for quantitative traits (Boyle et al., 2017; Liu et al., 2019). It assumes that there could be a set of core genes forming pathways with special relevance for the phenotype, which are modified by many, if not all other genes expressed in the same cells or developmental stage. Although the effect sizes of these other genes are expected to be smaller than those of the core genes, in sum they explain more of the genetic variance or heritability than the set of core genes.

A general strategy to verify the genetic function of a locus that was identified through a GWA is to study the effect of an artificial knockout allele of the locus in an isogenic background. If this results in a phenotype that is related to the phenotype that was originally studied in the GWA, one considers the association as genetically confirmed. However, in a typical GWAS, one studies only naturally segregating alleles, which are usually not knockout alleles. Hence, one is asking very different genetic questions with these approaches, but these can potentially complement each other. A gene with an allele that segregates as a modifier of other pathways in natural populations may have a much stronger effect when it is disrupted by a major mutation, thus revealing its connections to the phenotype much more clearly. Intriguingly, under the omnigenic model, one would predict that basically every randomly chosen gene that is expressed in the relevant developmental stage could show an effect on the phenotype studied, when it is disrupted by a major mutation. We have tested this hypothesis here, by comparing the genetic effects of gene disruption lines for loci that were identified through a GWA versus loci that were randomly chosen.

We have chosen *Drosophila* pupal size and pupation site choice as quantitative phenotypes, for which we have previously developed automated phenotyping pipelines and have shown their polygenic basis (Reeves and Tautz, 2017; Zhang et al., 2020). The pupal stage is an immobile life history stage of transformation between larval and adult stages found in holometabolous insects (Jones and Reiter, 1975; Price, 1970). Pupal size (or pupal case length) can be considered as an analogue to human height, in terms of morphological measure and various biological properties, such as equal paternal and maternal impact on offspring trait variance (Reeves and Tautz, 2017). Pupation site choice (or pupation height) is measured as the distance of the pupation site above the surface of the food. It reflects an ecologically relevant trait that is behaviourally determined and where a core network could be identified (Zhang et al., 2020). The phenotypes are not directly related to each other (Zhang et al., 2020).

For assessing the genetic contribution to these quantitative traits, we use the extensive genetic resources available for *Drosophila*. This includes the *Drosophila* Genetic Reference Panel (DGRP) as a resource for association mapping, as well as the gene disruption lines maintained at the Bloomington *Drosophila* Stock Centre (BDSC). The DGRP consists of a population of more than 200 fully sequenced highly inbred *Drosophila melanogaster* strains derived from a wild population in the Raleigh, US (Huang et al., 2014; MacKay et al., 2012), which has been successfully applied to detect the association between a broad range of complex traits and their underlying genetic basis (Dembeck et al., 2015; Durham et al., 2014; Huang et al., 2014). Together with the well-resolved richness in genetic polymorphisms and rapid decay in linkage disequilibrium (LD) in these strains (Huang et al., 2014), these make the DGRP an excellent resource to study the genetic architecture of pupal phenotypes in *Drosophila melanogaster* at a fine scale resolution. The BDSC gene disruption lines were created through systematic transposon mutagenesis experiments (Bellen et al., 2011). They can be used to test GWA candidate genes for a possible involvement in the trait, but also as a source for randomly chosen genes to study their effect on a quantitative trait.

In the present study, we conduct a standard GWA for pupal size based on the DGRP panel and use gene expression profiling analysis to reveal the relevant developmental stages for this trait. Following the predictions of the omnigenic model, we explore also the effects of genes randomly chosen among genes expressed in the respective life stage and with no significant GWA score. We find that not only a large fraction of the GWA predicted genes, but also of the randomly chosen genes have an effect on the pupal case length phenotype. Moreover, we find that many of the same genes affect also the pupation site choice phenotype, although the effect sizes between these two phenotypes are not correlated. We conclude that these findings support predictions of an omnigenic genetic architecture for quantitative traits and proof that association studies versus classic genetic studies reveal indeed different aspects of the genetic architecture.

Results

Phenotyping

We used pupal case length, defined as the major axis length in millimetre (mm) of the pupa as an ellipse-shape object (Figure 1), as a measure of pupal size in *Drosophila melanogaster.* We adapted and modified the earlier image-analysis

based phenotyping pipeline (Reeves and Tautz, 2017) for the high-throughput measurement of pupal case length. The same approach allows also to score the independent phenotype of pupation site choice (pupation height) and for both we have shown that a very large number of phenotypes can be measured with high precision (Reeves and Tautz, 2017; Zhang et al., 2020).

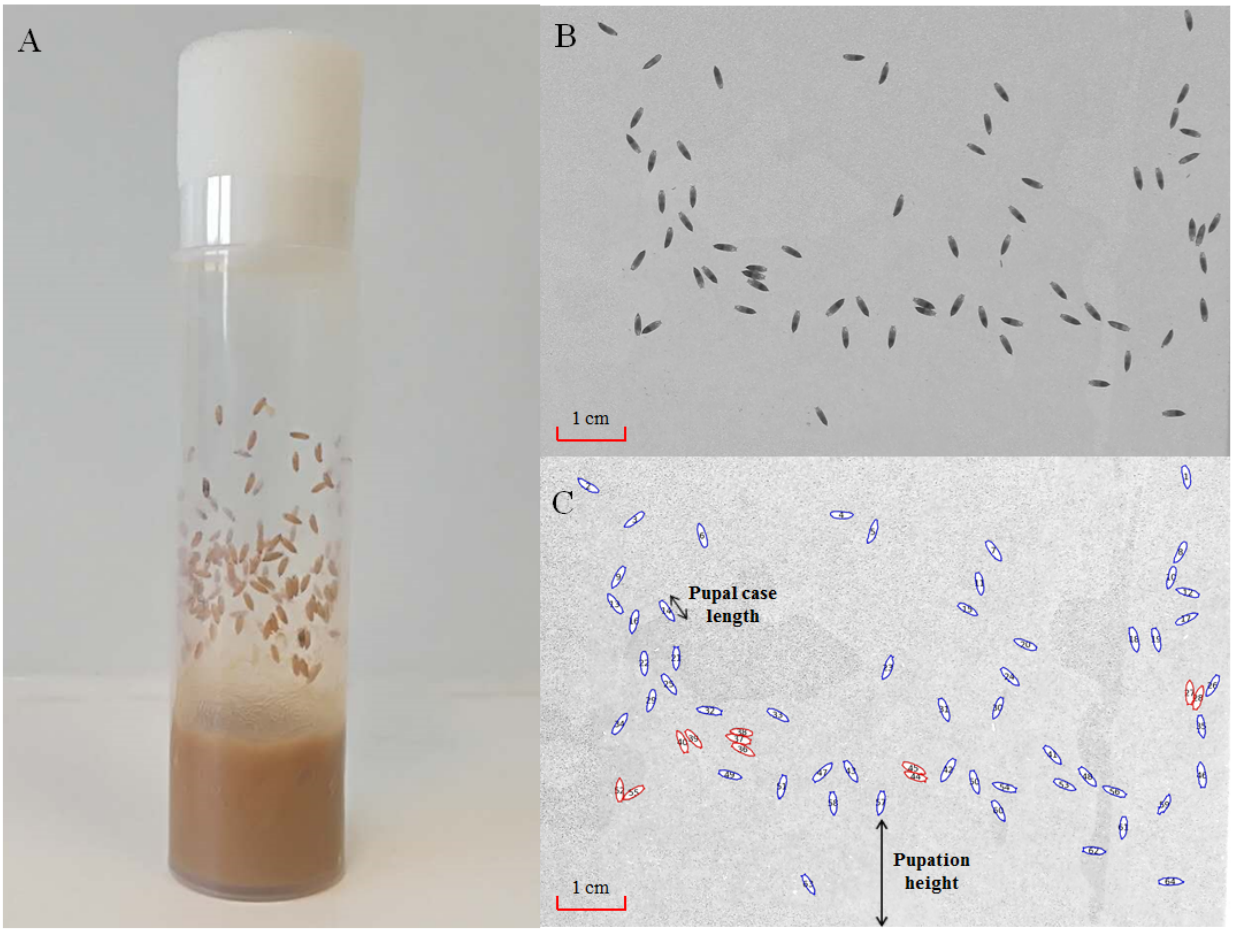

Figure 1: Automated measurement of pupal case length and pupation height. (A) shows a vial with sufficient number of pupae attached in the wall. A plastic sheet that lines the wall is included inside the vial, and is ready to be taken out for being photographed (B). The image analysis software CellProfiler (Carpenter et al., 2006) is then used to identify the outlines of the pupae and measure their major axis lengths and distances between their pupation sites to the food surface (C). The pupae that can be reliably measured (singularized) are marked in blue, and the ones with lower confidence scores (aggregated) are marked in red.

To assess whether the DGRP lines reflect the scope of natural variation, we measured also the pupal case length phenotypes of 14 natural wild-type *Drosophila melanogaster* strains collected from 9 global geographic regions (strain details in Supplementary Table S1a and S1b). Two of these wildtype stocks (S-317 and S-314) were continually re-measured to act as controls, and showed consistent measurement on pupal case length throughout all experiments (Figure 2–figure supplement 1).

The profiles of pupal case length from the wildtype and of 198 inbred lines from the *Drosophila* Genetic Reference Panel (DGRP) (Huang et al., 2014; MacKay et al., 2012) are shown in Figure 2. We observed a large variation of

average pupal case length among strains, ranging from 2.9 mm to 3.4 mm for natural wild-type strains, and ranging from 2.7 mm to 3.5 mm for DGRP inbred stocks. The spread of pupal case length among the DGRP inbred lines exceeded that of the DSPR recombinant inbred stocks (ranging from 2.9 mm to 3.5 mm) (Reeves and Tautz, 2017), suggesting that DGRP lines capture indeed the existing variation of pupal case length in *Drosophila melanogaster*. The same conclusion applies also to the pupation site choice phenotype, as we have shown previously (Zhang et al., 2020).

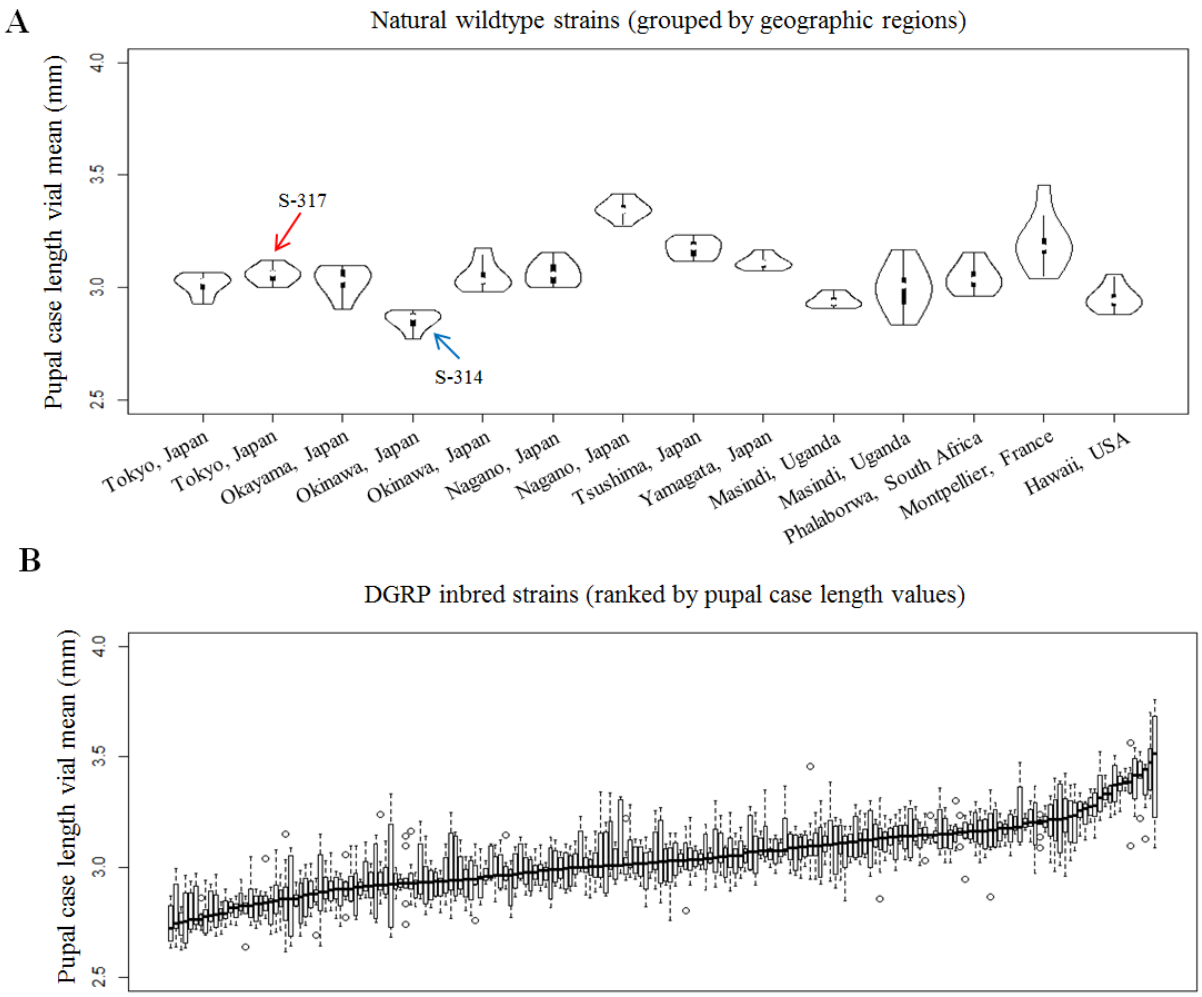

Figure 2: Pupal case length distribution for different *Drosophila melanogaster* strains. The distribution of natural wild derived stocks is grouped by geographic regions (A), and that of DGRP inbred strains is ranked by pupal case length values. The two natural stocks (S-317 and S-314) acting as controls are labelled in the panel (A).

Figure 2–figure supplement 1: Repeated measurement of pupal case length of two control stocks.

Heritability and chromosome effects

The estimate of the total genetic component of pupal case length, *i.e.*, broad sense heritability (H^2), was determined as the proportion of total variance in the mean strain pupal case length measurements compared to the average within

each strain (Schmidt et al., 2017). Our estimates of values of H^2 (0.71 for wild-type strains; 0.52 for DGRP inbred strains) were similar to the ones analysed with the same methodology, but based on another set of *Drosophila melanogaster* strains (Table 1) (Reeves and Tautz, 2017). Combining with the previous estimates on narrow sense heritability (h^2, ranging from 0.44-0.50, Table 1) (Reeves and Tautz, 2017), our analysis suggested that the additive genetic impact (*i.e.*, narrow sense heritability) contributed to a major part (~78%) of total genetic component of pupal case length in *Drosophila melanogaster.*

Table 1 Statistics for the heritability estimates of pupal case length

Heritability	Method	Dataset	Heritability Estimate
H^2	IBM SPSS (Variance component)	[a]wild-type strains	0.71
		[b]DGRP inbred strains	0.52
		[c]Four-Way	0.58
		[d]Eight-Way (DSPR)	0.61
h^2	Mid-parent regression	[e]Four-Way	0.44 ± 0.04 SE
		[f]Eight-Way (DSPR)	0.50 ± 0.09 SE

[a] Number of stocks = 14; Average number of replicates per stock = 7.8 ± 0.6 SD; Average number of measured pupae per vial = 69 ± 24 SD;
[b] Number of ILs = 198; Average number of replicates per stock = 8.2 ± 1.6 SD; Average number of measured pupae per vial = 40 ± 17 SD;
[c] Number of RILs = 81; Average number of replicates per stock = 5.5 ± 0.8 SD; Average number of measured pupae per vial = 59 ± 24 SD;
[d] Number of RILs = 195; Average number of replicates per stock = 6.7 ± 2.4 SD; Average number of measured pupae per vial = 70 ± 34 SD;
[e] Number of single-pair crosses = 363; Average number of measured pupae per vial = 62 ± 17 SD;
[f] Number of single-pair crosses = 67; Average number of measured pupae per vial = 47 ± 15 SD;
[c-f] Datasets were retrieved from (Reeves and Tautz, 2017);
ILs: Inbred lines; RILs: recombinant inbred lines; DSPR: *Drosophila* Synthetic Population Resource;
SD: standard deviation; SE: standard error;

Table 1-figure supplement 1: Chromosomal effects on explained pupal case length variance.

We partitioned the explained phenotypic variance by chromosome, via calculating the "SNP heritability" in the DGRP stocks (Wray et al., 2013), *i.e.*, the estimate of the proportion of phenotypic variance explained by all available SNPs (or genetic variants) within each chromosome. Our results show a much higher contribution of autosomes to the variance of pupal case length (Table 1-figure supplement 1), with the exception of chromosome 4, mostly likely due to the limited number of genetic variants within this chromosome. Hence, pupal case length is a highly heritable phenotype and the DGRP lines include the genetic variation suitable for mapping, similarly as we have previously shown for the pupation site choice phenotype (Zhang et al., 2020).

Genome-wide association analysis

We used the genetic variants of DGRP freeze 2 (Huang et al., 2014), and excluded the variants with missing values above 20% and minor allele frequency below 5% from the further analysis. Several possible covariates were assessed, in order to optimize the model for GWA mapping analysis (Figure 3).

Around half of the DGRP strains are infected by *Wolbachia pipientis*, a maternally transmitted endosymbiotic bacterium (Huang et al., 2014) (Supplementary Table S1b). We examined the possible effects of *Wolbachia* infection on pupal case length through two different approaches. We first compared the pupal case length values between infected strains and non-infected strains, and found no significant difference between these two groups (Wilcoxon rank-sum test: p-value=0.90, (Figure 3- table supplement 1). Second, we randomly selected three DGRP strains with *Wolbachia* infection, and tested the phenotypic change on pupal case length after the removal of *Wolbachia* infection by using tetracycline treatment. No significant statistical differences on pupal case length were observed for all tested strains (Figure 3-figure supplement 1). Based on these analyses, we concluded that the *Wolbachia* infection on DGRP strains had only minor, if any, influence on pupal case length, and consequently we did not incorporate the *Wolbachia* infection status in the association analysis below.

The DGRP lines were generated in a way that population structure impacts should be minimized, but some genetic relatedness leading to cryptic population structure might still exist (Mackay and Huang, 2018). To examine whether any cryptic population structure could contribute to the observed pupal case length variation of DGRP inbred stocks, we exploited PLINK (Purcell et al., 2007) to identify major principal components (PCs) of genetic variants in the DGRP strains. Based on the projection analysis of the top 2 PCs, we found no obvious clusters on the basis of pupal case length value classes (Figure 3-figure supplement 2). However, further correlation analysis revealed three out of the top 20 PCs to be significantly associated with pupal case length values (Figure 3-figure supplement 3, Pearson's correlation test, p-value$\leqslant$0.05), suggesting a possible impact on pupal case length from cryptic population structure. In order to correct any potential influence from the cryptic population structure, we decided to use a linear mixed model implemented in fastlMM (Lippert et al., 2011) program (version 0.2.32) for the GWA mapping analysis.

Approximately 45% of the DGRP strains harbour at least one type of major genomic inversion (Huang et al., 2014) (Supplementary Table S1b), and these major genomic inversions might contribute to the observed population structure and have impact on pupal case length as well. We systematically tested the correlations between genomic inversion status in DGRP strains and the top 2 PCs based on genetic variants as mentioned above, and found significant effects from In(2L)t and In(3R)Mo (Pearson's correlation test, p-value$\leqslant$0.05, Figure 3-table supplement 2), hinting their potential roles in population divergence (Hoffmann and Rieseberg, 2008). Moreover, we also observed a significant association between pupal case length and In(3R)Mo (Pearson's correlation test, p-value=0.0002). Accordingly, the presence status of In(3R)Mo in DGRP lines was incorporated as a covariate of the linear mixed model used for GWA mapping analysis on pupal case length. We used a significance cut-off of p-value $< 1 \times 10^{-5}$, which is a nominal threshold frequently used in *Drosophila* quantitative trait genetic studies (Dembeck et al., 2015; Lee et al., 2017; Zhang

et al., 2020). The mapping revealed a number of associations above the chosen threshold (Figure 3A) and the qq-plot shows that there is substantial genetic variation above the random expectation (Figure 3B).

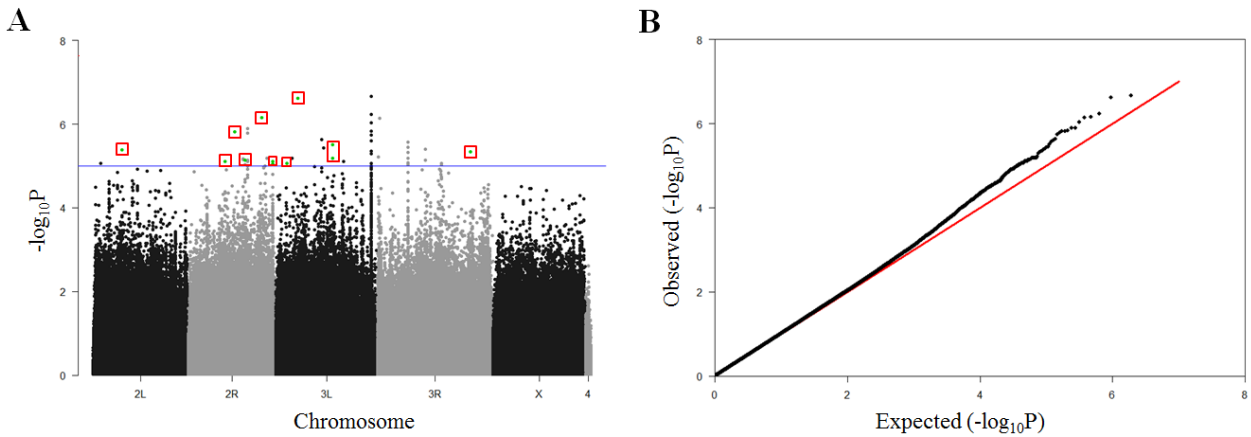

Figure 3: GWA mapping results on pupal case length. (A) Manhattan plot for GWA results. The p-values in $-\log_{10}$ transformation are shown on the y-axis. The blue horizontal line marks the nominal p-value threshold (1×10^{-5}) used in this study. The genetic variants in red rectangles are the ones selected for experimental validation as shown in Figure 4. (B) Quantile-quantile (Q-Q) plot for GWA results. The expectation and observation of GWA p-values in $-\log_{10}$ transformation are represented by the x-axis and y-axis, respectively.

Figure 3-figure supplement 1: Effect of *Wolbachia* infection on pupal case length.

Figure 3-figure supplement 2: Correlation between pupal case length and population structure.

Figure 3-figure supplement 3: Correlation between pupal case length and top 20 PCs.

Figure 3-table supplement 1: Comparison of pupal case length between Wolbachia-free and Wolbachia-infected stocks.

Figure 3-table supplement 2: Correlation between pupal case length and inversion status in DGRP strains.

Candidate genes from the GWA analysis

We found 50 significant SNPs to be associated with pupal case length in the DGRP strains, corresponding to 67 associating genes that locate within 5kb up/down-stream (default setting in SnpEff (Cingolani et al., 2012)) of these genetic variants (Table 2). The effect sizes of individual SNPs range from 0.5% to 3.3% of the total phenotypic variation (Table 2-figure supplement 1).

To identify possible additional candidate genes associated with the variants, we examined the long-range linkage disequilibrium (LD) between pairs of detected candidate variants and with other genetic variants found in the DGRP strains. LD blocks were then calculated for each significant genetic variant with a commonly used threshold $r^2 = 0.8$ (Pallares et al., 2014), and 17 significant LD blocks were found with average block size of 19.3 kb (Table 2-table supplement 1). This finding is in line with the observation of a rapid decay of LD in the panel strains reported in the original DGRP resource reference (Huang et al., 2014). Combining the additional genes identified in the above LD blocks, we identified in total 90 candidate genes associating with pupal case length variation in *Drosophila melanogaster*.

226 **Table 2: Genome-wide association results on pupal case length**

Genetic variant	MAF	[1]Effect size	Association p-value	[2]Associating genes (location)	Tester (BDSC ID)	Progenitor (BDSC ID)
2L_1830648_SNP	0.10	0.8%	8.78×10^{-6}	CG31664(Upstream); c-cup(Downstream)		
2L_6980934_SNP	0.08	0.7%	4.08×10^{-6}	***uif*** (Exon)	27826 (***)	5905
2R_8938928_SNP	0.07	0.6%	7.76×10^{-6}	CG13323,CG13324(Upstream); ***Drl-2***(Intergenic)	12655/25627 (***)	3605/5905
2R_11192893_SNP	0.09	0.9%	1.52×10^{-6}	Trpm(Intron); ***CG8152***(Downstream); ND-51L2(Intergenic)	18334 (*)	6326
2R_11192898_SNP	0.09	0.9%	1.52×10^{-6}	Trpm(Intron); ***CG8152***(Downstream); ND-51L2(Intergenic)		
2R_13228392_SNP	0.19	1.5%	7.07×10^{-6}	mbl(Intron)		
2R_13228444_SNP	0.17	1.4%	6.66×10^{-6}	mbl(Intron)		
2R_13629116_SNP	0.08	0.7%	9.37×10^{-6}	Vps50(Exon); Hyccin,PIG-A,Ir54a(Upstream); sub(Downstream)		
2R_13635814_SNP	0.38	2.4%	7.35×10^{-6}	sub,CG10931,***CG5002***(Upstream); Vps50,Ir54a(Downstream)	26108 (***)	5905
2R_14269079_SNP	0.05	0.5%	1.28×10^{-6}	BomT1(Exon); IM1,BomS4(Upstream); BomBc1,BomBc2(Downstream)		
2R_14269225_SNP	0.05	0.5%	1.29×10^{-6}	BomT1(Exon); IM1,BomS4(Upstream); BomBc1,BomBc2(Downstream)		
2R_14269259_SNP	0.05	0.5%	1.64×10^{-6}	BomT1(Exon); IM1,BomS4(Upstream); BomBc1,BomBc2(Downstream)		
2R_14270372_SNP	0.05	0.5%	7.78×10^{-6}	BomBc1(Exon); BomT1,IM1,BomS4,IM2(Upstream); BomBc2(Downstream)		
2R_14270389_SNP	0.09	0.8%	7.38×10^{-6}	BomBc1(Exon); BomT1,IM1,BomS4,IM2(Upstream); BomBc2(Downstream)		
2R_17777992_SNP	0.23	2.0%	7.04×10^{-7}	***Fili***(Intron)	23968 (ns)	5905
2R_18442616_SNP	0.08	0.7%	9.94×10^{-6}	Px(Intron); gas(Downstream)		
2R_18984873_SNP	0.08	0.7%	6.47×10^{-6}	CG9876(Upstream); RpL37b,CG42703(Downstream)		
2R_20362715_SNP	0.11	1.0%	9.01×10^{-6}	CG3494(3'UTR); CG16837(Upstream); CG3492(Downstream); ***CG13579***(Intron)	25600 (***)	5905
2R_20365015_SNP	0.15	1.2%	7.68×10^{-6}	CG3492,CG3494,CG16837(Downstream); ***CG13579***(Intron)		
3L_2625079_SNP	0.32	2.1%	8.62×10^{-6}	***Pxn***(Intron)	19060 (***)	6326
3L_3944957_SNP	0.28	2.0%	6.57×10^{-6}	CG14982(Exon); CG12766,CG10863,CG10853(Upstream); CG14983(Downstream)		
3L_5381932_SNP	0.16	1.7%	2.38×10^{-7}	***Ir64a***, DIP-delta(Intergenic)	24610 (**)	5905
3L_11037284_SNP	0.06	0.6%	2.35×10^{-6}	CG43693(Intron)		
3L_11501976_SNP	0.12	1.1%	3.72×10^{-6}	CG33500(Exon); Sgs7,Sgs3,CG7512(Upstream); Sgs8,CG33272(Downstream)		
3L_13654966_SNP	0.26	2.0%	3.09×10^{-6}	***bru3***(Intron)	27747 (***)	5905
3L_13654988_SNP	0.34	2.6%	6.45×10^{-6}	***bru3***(Intron)		
3L_16391306_SNP	0.29	2.1%	7.80×10^{-6}	CG13035(Intron); Grp,fax(Downstream)		

3L_23049662_SNP	0.18	1.8%	5.78×10^{-7}	Nrm,RpL10(Intergenic)		
3L_23056314_SNP	0.20	2.2%	2.16×10^{-7}	Nrm,RpL10(Intergenic)		
3L_23057931_SNP	0.18	1.7%	2.55×10^{-6}	Nrm,RpL10(Intergenic)		
3L_23062396_SNP	0.18	1.8%	1.50×10^{-6}	Nrm,RpL10(Intergenic)		
3L_23064396_SNP	0.18	1.8%	9.29×10^{-7}	Nrm,RpL10(Intergenic)		
3L_23064489_SNP	0.18	1.6%	5.02×10^{-6}	Nrm,RpL10(Intergenic)		
3L_23064760_SNP	0.17	1.5%	8.72×10^{-6}	Nrm,RpL10(Intergenic)		
3L_23064763_SNP	0.18	1.7%	1.81×10^{-6}	Nrm,RpL10(Intergenic)		
3L_23073783_SNP	0.19	1.6%	4.49×10^{-6}	Nrm,RpL10(Intergenic)		
3R_184071_SNP	0.10	0.9%	6.17×10^{-6}	CG1090(Upstream); hkb(Intergenic)		
3R_541976_SNP	0.44	3.3%	7.33×10^{-7}	eIF3f1(Upstream); smash,Nep2(Downstream)		
3R_7334295_SNP	0.14	1.2%	6.52×10^{-6}	dpr5,dpr4(Intergenic)		
3R_7334392_SNP	0.16	1.3%	6.62×10^{-6}	dpr5,dpr4(Intergenic)		
3R_7334511_SNP	0.13	1.2%	2.69×10^{-6}	dpr5,dpr4(Intergenic)		
3R_7334913_SNP	0.27	2.0%	8.35×10^{-6}	dpr5,dpr4(Intergenic)		
3R_7335569_SNP	0.14	1.3%	3.61×10^{-6}	dpr5,dpr4(Intergenic)		
3R_7336708_SNP	0.14	1.2%	4.75×10^{-6}	dpr5,dpr4(Intergenic)		
3R_7343055_SNP	0.13	1.1%	9.34×10^{-6}	dpr5,dpr4(Intergenic)		
3R_11590987_SNP	0.11	1.1%	3.98×10^{-6}	CG12784,CG5623(Downstream)		
3R_11591060_SNP	0.10	1.0%	7.33×10^{-6}	CG12784,CG5623(Downstream)		
3R_15455022_SNP	0.25	2.0%	8.72×10^{-6}	CG6231(Intron); subdued(Downstream)		
3R_15455023_SNP	0.24	1.9%	9.34×10^{-6}	CG6231(Intron); subdued(Downstream)		
3R_22486221_SNP	0.06	0.6%	4.59×10^{-6}	Tb(Downstream); ***beat-VII***(Intron)	24707 (***)	5905

MAF: minor allele frequency; BDSC: Bloomington Drosophila Stock Centre;

[1] the GWA effect size is the percentage of the total phenotypic variation explained by the focal SNP;

[2] associating genes are defined as the ones locate within 5kb up/down-stream of target genetic variant;

Highlighted in bold and italic are the ones applied for experimental confirmations as shown in Figure 4;

***: p-value⩽0.001; **: p-value⩽0.01; *: p-value⩽0.05; ns.: p-value>0.05.

Table 2–figure supplement 1: Effect size distribution of GWA significant genetic variants.

Table 2-table supplement 1: Additional associated genes from LD blocks.

Phenotype confirmations

To test the phenotypic effects from different alleles on pupal case length at single gene scale, we used transposon insertion mutagenesis lines that have been constructed in a common co-isogenic background (Bellen et al., 2011). Eleven gene disruption lines corresponding to 10 of the GWA identified associated genes were available for this experiment (Table 2). All experimental tests were done via replicated phenotypic comparisons between co-isogenic stocks and homozygous mutant lines, except for the gene of *Pxn,* which was tested as heterozygote (Figure 4), as the homozygous disruption of this gene is lethal (see Methods).

Figure 4 shows the profile on the measured pupal case length differences between mutant lines and the respective progenitor stocks. All apart one of the lines showed a significant difference (Wilcoxon rank-sum test, p-value ⩽ 0.05), interestingly all towards an increased pupal case length, implying that the effect is not just due to a general physiological disruption of pupal growth. No significant correlation was found between GWA predicted effect sizes and the phenotypic effect sizes from the above gene disruption tests (Pearson's correlation test, r value = 0.28, p-value = 0.44), confirming the notion that naturally segregating alleles have usually different genetic effect sizes when compared to disruption alleles. It should be noted that two independent experimental tests were conducted for gene *Drl-2,* with transposon insertions landing in the different regulatory regions and from different co-isogenic backgrounds (Supplementary Table S1c). Both showed consistent phenotypic effects, affirming the robustness of the gene disruption assays.

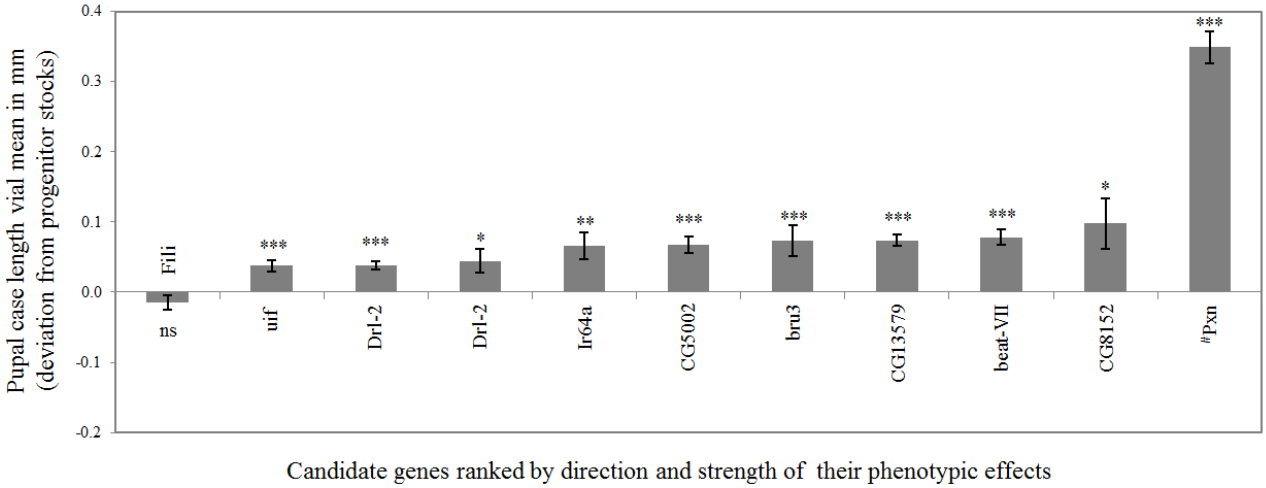

Figure 4: Gene disruption test results on the pupal case length phenotypic effects of GWA candidate genes. Phenotypic effects of candidate genes were from transposon insertion tests. The phenotypic effects were measured as the deviation of pupal case length of stocks with gene disruption compared with that from corresponding progenitor stocks. The error bars show the standard error of mean (SEM) values. Statistical p-values were computed via Wilcoxon rank-sum tests. ***: p-value⩽0.001; **: p-value⩽0.01; *: p-value⩽0.05; ns: p-value>0.05. Stock marked with "#" was tested as heterozygotes; the remaining were stocks tested as homozygotes.

Expression analysis

Genes involved in determining pupal case length are expected to be expressed at developmental stages close to the pupation process. To explore this, we analysed a developmental RNA-Seq dataset from *Drosophila melanogaster* (Graveley et al., 2011), which included the transcriptome of 27 distinct developmental stages covering all the major phases of the *Drosophila* life cycle. For the ease of presentation, we further grouped these 27 distinct developmental stages into six super-developmental stages: early embryo, middle embryo, late embryo, larval stage, pupal stage, and adult stage (Figure 5).

Compared with total annotated coding genes as a control, we observed an enrichment of GWA candidate genes with expression (FPKM >0) in all 27 tested distinct developmental stages (Figure 5A and Figure 5-table supplement 1), suggesting that GWA tends to pick up gene candidates with broad expression. More intriguingly, we found these GWA candidate genes are mostly significantly enriched in the developmental period between late larval stage and pupal stage (Fisher's exact test, p-value⩽0.01, Figure 5B). As the pupation process occurs from the late larval stage, and the genesis of pupal case is completed before eclosion, our results imply that the developmental period between late larval and pupal stage might be the most relevant developmental stages for pupal case length morphogenesis in *Drosophila melanogaster*.

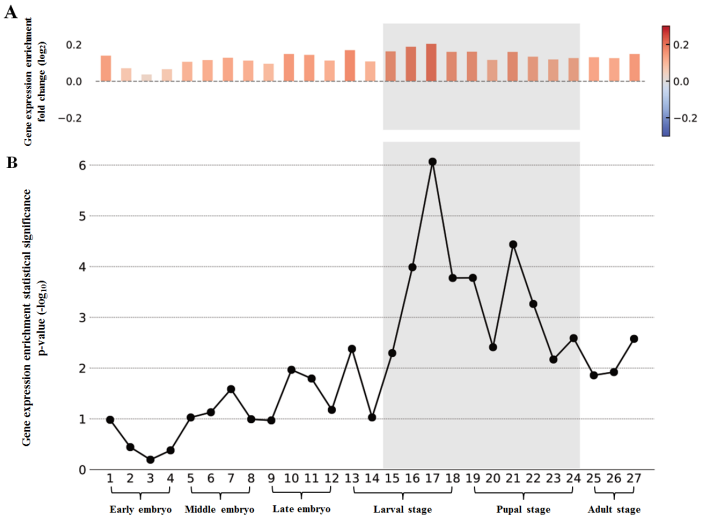

Figure 5: Gene expression enrichment patterns of GWA candidate genes across different developmental stages. The expression enrichment statistics of GWA candidate genes were calculated with total annotated protein coding genes as a control. The panel (A) shows the gene expression enrichment fold changes (*i.e.*, the ratios between the fractions of expressed GWA candidate genes and total coding genes, in $\log_2$ scale) between GWA candidate genes and total coding genes, and the panel (B) shows the statistical significances (p-values in $-\log_{10}$ scale, Fisher's exact test) of the above gene expression enrichment patterns in 27 developmental stages. *Early embryo*: 1) 0-2 hrs. embryo; 2) 2-4 hrs. embryo; 3) 4-6 hrs. embryo; 4) 6-8 hrs. embryo; *Middle embryo*: 5) 8-10 hrs. embryo; 6) 10-12 hrs. embryo; 7) 12-14 hrs. embryo; 8) 14-16 hrs. embryo; *Late embryo*: 9) 16-18 hrs. embryo; 10) 18-20 hrs. embryo; 11) 20-22 hrs. embryo; 12) 22-24 hrs. embryo; *Larval stage:* 13) L1 stage; 14) L2 stage; 15) L3 stage of 12 hr. post-molt; 16) L3 stage of dark blue gut (PS1-2); 17) L3 stage of light blue gut (PS3-6); 18) L3 stage of clear blue gut (PS7-9); *Pupal stage:* 19) white prepupa (WPP), 20) 12 hr. after WPP (P5), 21) 24 hr. after WPP (P6); 22) 2 days after WPP (P8), 23) 3 days after WPP (P9-10), 24) 4 days after WPP (P15); *Adult stage:* 25) male/female after 1-day eclosion; 26) male/female after 5-day eclosion; 27) male/female after 30-day eclosion. For adult developmental stages, the enrichment results for male and female were averaged for presentation. The shaded area indicates the "relevant" developmental stages of pupal case length morphogenesis.

Figure 5-table supplement 1: Summary on gene expression patterns across different developmental stages.

292

293 *Phenotypic effects of randomly chosen genes*

294 The omnigenic model (Boyle et al., 2017) predicts that most, if not all genes can affect a given quantitative trait, when
295 they are expressed in the relevant developmental stage. We have therefore set out to use our phenotyping pipeline to
296 test this prediction. We selected 47 gene disruption lines (corresponding to 45 disrupted genes, all tested as
297 homozygotes) constructed in a co-isogenic background from the panel of *Drosophila* gene disruption stocks (Bellen
298 et al., 2011) using the following criteria: 1) the GWA p-value of the focal gene falls below the significance threshold
299 (*i.e.*, $> 1 \times 10^{-5}$); 2) the corresponding gene has detectable expression in the above defined relevant developmental
300 stages for pupal case length (FPKM >0), and 3) homozygous disruption is viable. The latter criterion biases against
301 essential genes (approximately 30% are not homozygous viable).

302 Figure 6A compares the phenotypic effect sizes of all strains tested in this study with their respective pupal case length
303 GWAS p-values. It shows that the genes picked because of their GWA significance have not necessarily the largest
304 phenotypic effects as disruption lines, and that there is no significant difference on the overall phenotypic effects
305 between GWA candidate genes and randomly chosen genes (Wilcoxon rank-sum test p-value=0.39, Figure 6-figure
306 supplement 1A). The conclusion remained the same when the only one GWA predicted gene that was tested as
307 heterozygotes (*Pxn,* Figure 4) was removed from the analysis (Wilcoxon rank-sum test p-value=0.14). Overall, around
308 76% of random genes (34 of 45) showed significant effects (p-value ≤ 0.05) on the pupal case length phenotype (Figure
309 6B), in contrast with the 90% of GWA candidate genes (9 out of 10), but these fractions are not significantly different
310 (p-value = 0.43, Fisher's exact test).

311 It should also be noted that these strains may 1) have different co-isogenic backgrounds, 2) have disruption in different
312 regions, either coding regions (disruption on the structure of coded proteins) or regulatory regions (alteration of the
313 gene expression levels). However, these two factors seemed to play only minor roles in the phenotype differences,
314 given the observation that similar extents phenotypic effects on pupal case length can be observed for the disruption
315 of the same gene from difference genetic background and/or different gene regions (*Drl-2*, *CG14007*, and *CG42260*.
316 See Figure 6-table supplement 1). Moreover, systematic comparison on the effect size on pupal case length also showed
317 no significant difference either on the disruption of different gene regions (Figure 6-figure supplement 1B), or the
318 percentage of disruption on the gene's CDS regions (Figure 6-figure supplement 1C).

319

320 *Phenotypic effects on an independent phenotype*

321 If most genes have an effect on one quantitative trait, one would expect that the same is true for a second trait. As
322 discussed above, pupation site choice is such a second independent trait, that we measured with the same setup. Given
323 the sharing of gene expressions in the common relevant developmental stages (Figure 6-table supplement 1), all the
324 disrupted genes (both the 10 tested GWA genes and 45 random genes) on the basis of pupal case length can therefore
325 serve as random gene disruption tests for pupation site choice. Indeed, we observed that around 67% of the randomly
326 tested genes (38 of 55) actually showed significant effects (p-value$\leqslant$0.05) on the pupation site choice trait (Figure

6C). This fraction of effects is not significantly different from the fraction of effects of randomly tested genes for the pupal case length phenotype reported above (p-value = 0.51, Fisher's exact test). This supports the notion that the majority of genes with expression in the relevant developmental stages can have effects on a given quantitative trait at this stage. Interestingly, the effect sizes do not correlate for the two phenotypes tested (Figure 6D), implying that the effects are channelled through different networks.

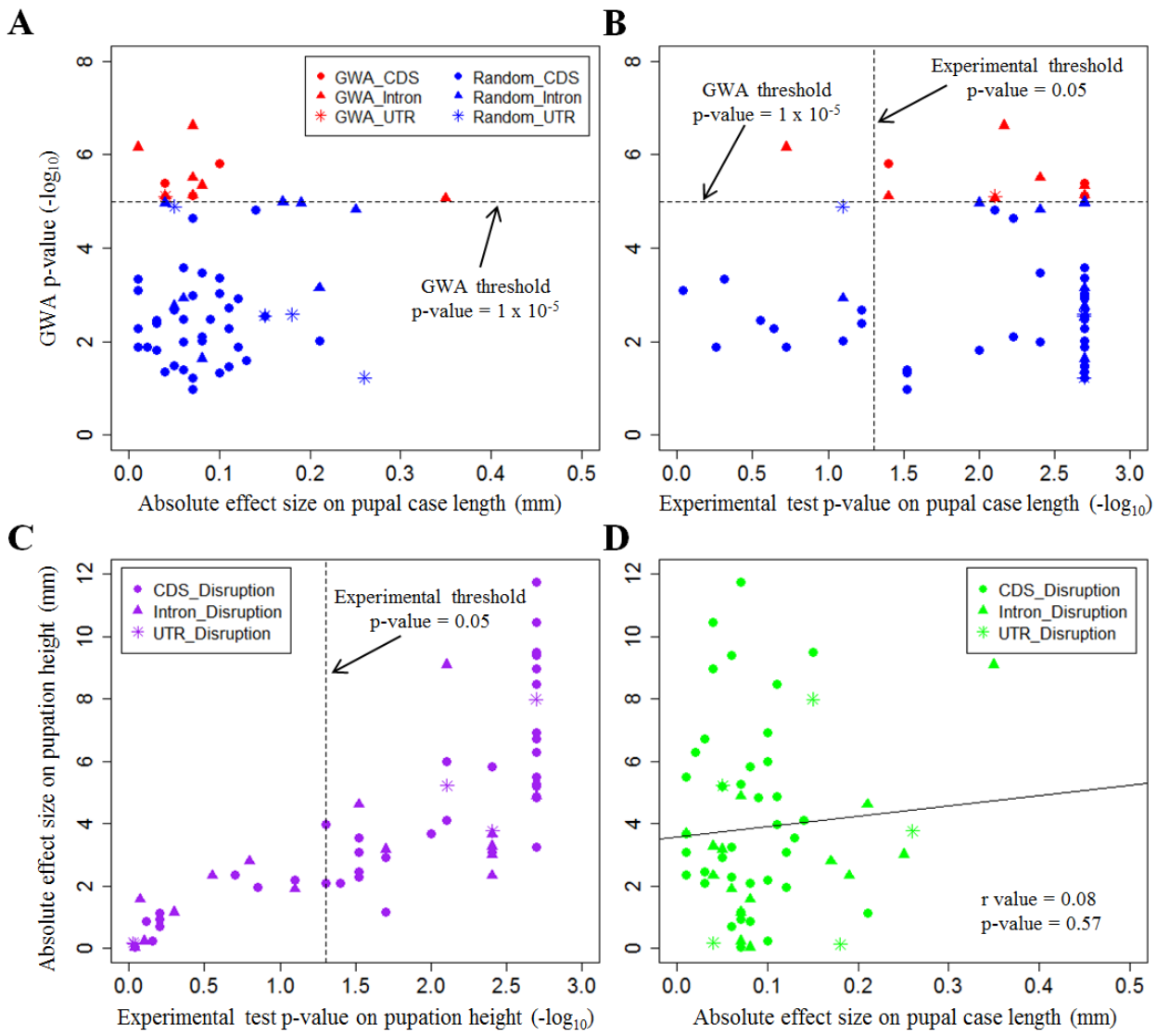

Figure 6: The impact of gene disruption on pupal case length and pupation height. Comparison between GWA p-values and absolute experimental effect sizes and statistical p-values for gene disruption of GWA candidate genes (labelled in red) and randomly selected genes (labelled in blue) on pupal case length are shown in (A) and (B), respectively. The GWA p-value is defined as the lowest p-value from all the genetic variants within the 5kb up/down-stream of target gene (including the gene itself). (C) Pupation site choice (height) comparison between absolute experimental effect sizes and statistical p-values for gene disruption of all tested genes. Experimental statistical p-values were computed via Wilcoxon rank-sum tests. The horizontal dashed line marks the GWA nominal p-value threshold (1 x 10^{-5}), and the vertical dashed line indicates the p-value threshold for experimental tests (0.05). (D) Correlation between the absolute experimental effect sizes on pupation height and pupal case length. The statistics on the correlation was computed by using Pearson's correlation test. Each dot represents one gene disruption line. Genes with transposon insertion mutagenesis in different genic regions (CDS, Intron, or UTR) are labelled with various dot shapes. The absolute experimental effect size was measured as the absolute deviation of pupal case length/pupation height of the stock with the target gene disruption compared with that from corresponding progenitor stock.

Figure 6-figure supplement 1: Gene disruption effect sizes on pupal case length based on their categories.

Figure 6-figure supplement 2: Genetic variant density comparison between GWA genes and random genes with significant pupal case length phenotypic effects.

Figure 6-table supplement 1: Detailed information on all the tested gene disruption lines.

Discussion

We have here tested predictions of the omnigenic model for quantitative traits, especially the question of the possible role of modifier loci without significant GWA score (Boyle et al., 2017; Liu et al., 2019). Although it is debated whether the term "omnigenic" is more useful than the long established terms "polygenic" and "infinitesimal" (Wray et al., 2018), the analysis by (Boyle et al., 2017) has certainly sparked new interest in this almost century-old question. But we are only now coming into a phase where predictions from these models can be directly tested. Most of the evidence has so far come from human studies, which are often focussed on disease questions and their associated special considerations and limitations (Wray et al., 2018). But for well-developed genetic model systems, such as *Drosophila*, one can do direct genetic experiments.

The *Drosophila* system allows to systematically using both, a GWA approach, as well as the classic knock-out approach to assess the effect of loci on given phenotypes. Furthermore, the availability of automated phenotyping pipelines allows to score a large number of individuals under highly controlled conditions and thus get a statistically sound resolution even for small effects. We have explored this here for two pupal phenotypes that are particularly amenable to analysis through automatic image processing.

Using the DGRP lines that represent segregating wildtype alleles and the commonly used corrections for confounding factors, we were able to identify a set of 90 candidate genes within 50 associated genetic loci for the pupal size phenotype (Table 2 and Table 2-table supplement 1). As expected for a typical polygenic trait, each genetic variant could only explain a small fraction (< 4%) of total phenotypic variance, implying that most of the heritability of the trait comes from a large number of loci with GWA scores below the chosen threshold. Using the transposon insertion mutagenesis lines that have been constructed in a common co-isogenic background (Bellen et al., 2011), we found that nine out of ten tested GWA candidate genes showed significant effects (p-value$\leqslant$0.05) on pupal size. These include the well-studied gene *Tubby (Tb)*, which was reported to be involved in body morphogenesis. *Tubby* is actually also a phenotypic pupal marker for a balancer chromosome (Lattao et al., 2011). Other clear candidate genes include two loci known to be involved in developmental processes, such as *uninflatable* (*uif)*, which modulates Notch signaling (Loubéry et al., 2014) and the receptor tyrosine kinase *Derailed 2* (*Drl-2*) that transduces Wnt5 signaling (Sakurai et al., 2009). Accordingly, the GWA did identify loci that could be part of core pathways of the phenotype, similar as we showed it for the pupation site choice phenotype (Zhang et al., 2020).

Gene expression analysis on GWA candidate genes revealed their enrichment in the developmental interval from late larval to the pupal stage, suggesting these developmental stages to be relevant of pupal case length phenotype. The omnigenic model (Boyle et al., 2017) predicts that most, if not all genes can affect a given quantitative trait, when they

are expressed in the relevant developmental stage and organ(s). We thus set out to use gene disruption experimental tests on random genes with expression in the relevant developmental stages. The GWA p-value for these 45 randomly chosen genes is below any threshold that one would normally consider using. Accordingly, none of them would have been identified as GWA candidate genes. However, our results showed that approximately three quarters of them actually had significant effects (p-value$\leqslant$0.05) on the pupal case length phenotype. Intriguingly, a similar fraction, but only partially overlapping loci, showed also an effect on the behavioural pupation site choice phenotype.

One can ask why these randomly chosen genes, especially the ones showing significant phenotypic effects on pupal phenotypes, were not picked up by GWA. Actually, there is no observed significantly lower genetic density or allele frequency of random chosen genes with significant pupal case length phenotypic effects compared with GWA candidate genes (Figure 6-figure supplement 2). Therefore, the reason why they have not shown up in the GWA has to be that they do not include segregating variants of sufficient effect size in the population from which the DGRP was derived. In fact, one can expect that most genetic variants present in a natural *Drosophila* population are unlikely to represent gene disabling mutations. And the variants that are gene disabling should be rare because they should be selected against, *i.e.*, they should also only seldomly occur as homozygotes. Hence, to test these genes as homozygous gene disruptions is rather unnatural. Still, it is an indication that these genes are involved in some form in the phenotype. This is a general issue when comparing gene effects from a GWA analysis with those from classic genetic analyses. The former trace the effects of naturally occurring variants, the latter the ones of gene disruptions. These converge only for human genetic disease studies, but not for studies on natural genetic variation of quantitative traits. One has to keep this dichotomy of genetic views in mind when placing GWA results in the context of classic genetic results.

Conclusion

Our data confirm the major prediction of the polygenic/omnigenic model, namely that most randomly chosen genes can potentially influence any given trait. It should be possible to apply this test also to other phenotypes or genetic systems where the necessary stocks or experimental procedures are available. Evidently, if almost any random gene is involved in a given phenotype, why should one then do a GWA in the first place? Hence, it should be of special interest to assess whether GWA p-values provide generally a guide towards the core networks required to generate the phenotype.

Materials and Methods

Drosophila stocks and flies rearing

All flies were reared in vials containing cornmeal–molasses–agar medium at 24°C, 55% -78% relative humidity, and a 12:12-h light-dark cycle. A HOBO® data logger was placed in the incubator to monitor and record any potential environmental changes, including temperature, light and humidity. Given the only minor impact on pupal case length from the change of relative humidity across experiments (Figure 2-figure supplement 1B), no special action was taken to control this factor.

In total, 14 wild type stocks, 198 DGRP inbred lines, 58 transposon insertion mutagenesis stocks (disruption for 55 annotated protein coding genes) and 3 corresponding progenitor stocks were assayed in this study. The detailed description and measurement data of these stocks can be found in Supplementary Table S1.

Phenotyping of pupal case length

We conducted the measurement of pupal case length by adopting an earlier established image-analysis based automated phenotyping pipeline as described in (Reeves and Tautz, 2017; Zhang et al., 2020). The entire phenotyping pipeline is implemented with five successive procedures:

1) Vial preparation. Standard cornmeal–molasses–agar food was dispensed into standard 28.5 mm diameter, 95 mm height vials (Genesee Scientific) to a depth of approximately 20mm. Once the food vials had fully cooled, 10.1 cm x 10.5 cm squares of transparent film (nobo, plain paper copier film, 33638237) were slid into the bottom of each vial lining their entire vertical wall. Inserting the film does not change the storage properties of the food. Prior to introducing adult flies, a very small amount of live yeast paste was dotted on the food surface. A barcode with custom printed semi-transparent label was affixed to the outside of each vial as the unique identifier.

2) Mating flies. Approximately ten 2-5 days old healthy female flies (fifteen for inbred stocks to compensate the reduced fertility) and five similar-aged male flies were introduced into each vial, under the incubation condition as mentioned above. These adult flies were cleared from the vials after 1-2 days introduction and vials were kept in the same incubation condition for another 8-9 days to allow them to reach pupation stage. In case of observing that most of offspring in the vials were present as pupae attached to the transparent film, the film was then gently taken out from each vial, and the food from the lower part was scrapped away and any viable larvae were also removed.

3) Film Photographing. Once removed from the vial, the film (with the barcode label affixed) was then placed into a custom plastic frame, which holds the film flat for further photographing. Frames were then photographed in a light tight box with a sliding door, which provided illumination only from underneath the frame, effectively silhouetting the pupae while minimizing tangential shadows. Every photograph included a 1 euro cent coin (16.25 mm diameter), for the control of camera coordinate changes and the conversion of measurements from pixels to millimetres (mm). Batches of the resulting images were then introduced into the following image analysis procedure.

442 *4) Image analysis.* We applied the open-source image analysis software CellProfiler (v2.1.0) (Carpenter et al., 2006)
443 for the recognition of pupae and measurements of a variety of attributes, with a customized pipeline adopted from
444 (Zhang et al., 2020). Firstly, by using the CellProfiler module "identify primary objects", we identified any "primary
445 object" with significant distinction from the background without restriction on their sizes. Secondly, the above
446 identified objects composing of multiple touching pupae were disentangled into distinct pupae (module "Untangle
447 Worms"). Thirdly, the resulting candidate pupae were shrunk and re-propagated outwards for a more precise detection
448 of the edges of each pupa based on boundary changes in pixel intensity (module "Identify secondary objects"). Lastly,
449 distinct attributes for the pupae were calculated and a specific confidence class was assigned for each pupa based on
450 its size attribute. Based on manual curation on 40 randomly selected films, we found that the above CellProfiler pipeline
451 can reach to a 96% of sensitivity (fraction of identified true pupae), but a modest false discovery rate of 19% for
452 identified putative pupae (Zhang et al., 2020). To further reduce the false discovery rate, we refined an additional
453 criterion setting based on the size attributes of "true" pupae based on manual curation (Zhang et al., 2020). Applying
454 this setting of new criteria, the false discovery rate for pupae detection dropped to around 0.15%, with only a tiny
455 fraction (< 0.7%) of loss for true positive results.

456 *5) Pupal case length measurement.* The pupal case length is defined as the length of the major axis of the ellipse that
457 has the same normalized second central moments as the region of identified pupae, measured with the
458 "Areashape_MajorAxisLength" index in CellProfiler. Based on its ratio to the diameter measurement of 1 euro cent
459 coin (16.25 mm diameter) included in the photograph, the measurement of case length of each pupa was converted
460 from pixels to mm. The mean of all the measurements of pupal case length for all the pupae in the food vial was taken
461 as representative for the focal vial. To further reduce the potential bias from low sampling effect, we only included
462 vials with a pupal density of a minimum of 15, and required each stock should include at least 6 such reliable vial
463 measurements.

464

465 *Repeated measurements of control stocks*

466 Throughout the experiments for wildtype and DGRP inbred strains measurements, two wild-type stocks (S-314 and S-
467 317) were continually re-measured in the manner described above to control for environmental effects, especially small
468 fluctuations in humidity. Such re-measurement of control stocks were not applied for the functional validation
469 experiments as shown below, as the experimental tests on each pair of gene disruption stock and progenitor stock were
470 conducted in the same experimental condition. The impacts of the incubator relative humidity changes on the pupal
471 case length measurement (across all rounds of DGRP inbred stock experiments) were directly examined by using
472 Pearson's correlation tests for two control stocks separately.

473

474

475

476 *Wolbachia infection effect test*

477 *Wolbachia pipientis* is a maternally transmitted endosymbiotic bacterium that was reported to infect around 53% of
478 DGRP inbred stocks (Supplementary Table S1b) (Huang et al., 2014). We used two different approaches to examine
479 whether there is any significant influence on pupal case length from *Wolbachia* infection. Firstly, we directly compared
480 the pupal case length measurement between *Wolbachia*-infected stocks and *Wolbachia*-uninfected stocks, and
481 computed the statistical significance on the pupal case length difference between these two groups via a Wilcoxon
482 rank-sum test.

483 Secondly, we conducted a direct experimental test to compare the changes of pupal case length measurement between
484 *Wolbachia* infected stocks and *Wolbachia*-free stocks, which were created though two generations of tetracycline
485 treatment on the infected stocks (by adding an appropriate volume of 100 mg/ml of tetracycline suspended in 99%
486 ethanol to the surface of the solid prepared food) and then reared for at least another two generations with standard
487 food to avoid any detrimental parental effects (Zeh et al., 2012). We firstly ruled out the possibility that tetracycline
488 treatment could have an influence on pupal case length, by comparing the pupal case length measurement changes
489 between three randomly selected *Wolbachia*-free stocks (DR_14, DR_45, DR_106), and the same stocks with the
490 above mentioned tetracycline treatment procedure. Then we randomly selected three *Wolbachia* infected stocks
491 (DR_16, DR_21, DR_67), and half of the files were treated with tetracycline as mentioned to create *Wolbachia*-free
492 stocks, and the rest half the flies from same three strains were also reared with standard food across the experiment as
493 controls. We extracted genomic DNA from the above 6 stocks individually by using DNeasy blood and tissue kit
494 (Qiagen), and measured the purity and concentration of the resulting DNA with NanoDrop ND-1000
495 spectrophotometer (Thermofisher). A diagnostic PCR to test for the presence of the *Wolbachia wsp* gene was
496 performed by using the primers wsp81F (5′-tggtccaaaatgtgagaaac-3′) and wsp691r (5′-aaaattaaacgctactcca-3′)
497 (Richardson et al., 2012), under the reaction condition of 35 cycles of 94°C for 15 seconds, 55°C for 30 seconds and
498 72°C for 1 minute. A standard (1%) agarose gel electrophoresis was used to test for the presence of the PCR product
499 (~630 bp), with the broad range Quick DNA Marker (NEB #N0303) as loading ladder. Pupal case length measurement
500 between the three *Wolbachia*-infected and created *Wolbachia*-free lines were then measured and compared by using
501 Wilcoxon rank-sum tests.

502

503 *Principal component analysis and genomic inversion effect test*

504 All the genetic variant calling data and major genomic inversion status were retrieved from DGRP freeze 2 (Huang et
505 al., 2014), for which the coordinates were based on Flybase version 5 (Attrill et al., 2016). Only the genetic variants
506 with <20% missing values and ≥5% minor allele frequency (MAF) were kept (corresponding to 1,903,028 genetic
507 variants) for further analysis. We assessed the possible influence of population structure on the pupal case length
508 measurement in DGRP inbred lines, by using the PCA module from PLINK v1.90 (Purcell et al., 2007) to identify
509 principal components (PCs) from the above filtered genetic variant data. We firstly checked the clustering status on
510 the three equal-sized groups of DGRP inbred strains with ranked pupal case length values (*i.e.*, low pupal case length
511 value, medium pupal case length value and high pupal case length value) based on the projection analysis of the top

512 two PCs, and then examined the correlation between pupal case length values and the top 20 PCs by using Pearson's
513 correlation tests.

514 To explore the potential contribution of major genomic inversions to the population structure within DGRP strains, we
515 used Pearson's correlation tests to calculate the correlations between genomic inversion status and the top 2 PCs from
516 the above principal component analysis. Moreover, we also tested the correlation between pupal case length values
517 and the status of each genomic inversion within DGRP strains. The resultant significant correlated major genomic
518 inversion (*i.e.*, In(3R)Mo) was included as a covariate for the following chromosomal effect partitioning and genome-
519 wide association analysis.

520

521 *Estimates of heritability and chromosomal effect partitioning*

522 We estimated the broad sense heritability (H^2) of pupal case length with the variance components of a linear model of
523 the form: Phenotype = Population mean + Line effect + error (Schmidt et al., 2017). We computed the total phenotypic
524 variance as Genetic Variance (G_v) + Environmental Variance (E_v), and the H^2 as G_v/G_v+E_v. This calculations for both
525 wild-type and DGRP inbred stocks were implemented by using IBM SPSS Statistics (version 22) (SPSS Inc, 2007),
526 with pupal case length measurement as the dependent variable and DGRP stock names as a random factor. Meanwhile,
527 we compared the estimates of H^2 computed with the same methodology, but from different *Drosophila melanogaster*
528 strains datasets (Reeves and Tautz, 2017). The estimates on the narrow sense heritability (h^2) of pupal case length
529 based on mid-parent regression were directly retrieved from (Reeves and Tautz, 2017).

530 We partitioned the fraction of pupal case length phenotypic variance explained by the genetic variants from each
531 *Drosophila melanogaster* chromosome/arm (*i.e.*, 2L, 2R, 3L, 3R, 4, X) by using GEMMA (version 0.96) (Zhou and
532 Stephens, 2012). GEMMA implemented a univariate linear mixed model to estimate the variance (denoted as PVE)
533 explained by each chromosome/arm, with the controlling for population structure effect via incorporating a genetic
534 relationship matrix (GRM) built from DGRP strains and the major genomic inversion In(3R)Mo status as a covariate
535 in the analysis. Even though the absolute estimates are inflated due to the individual relatedness, the relative effects of
536 chromosomes/arms are still informative for comparison, as the overestimation is uniformly spread across the whole
537 genome (Pallares et al., 2014; Yang et al., 2011).

538

539 *Genome-wide association analysis*

540 We performed genome-wide association (GWA) analysis on pupal case length for the above filtered genetic variants
541 from DGRP freeze 2 (Huang et al., 2014), by using the FastLMM (Lippert et al., 2011) program (version 0.2.32), with
542 the major genomic inversion In(3R)Mo status being included as a covariate for the GWA analysis. This program fitted
543 a linear mixed model that could control for population stratification effects.

544 We defined GWA significant associated genetic variants by using p-value < 1 x 10^{-5}, which is a nominal threshold
545 frequently used in *Drosophila* quantitative trait genetic studies (Dembeck et al., 2015; Lee et al., 2017; Zhang et al.,
546 2020). The R package "qqman" (Turner, 2014) was exploited for the visualization of GWA results in a Manhattan plot
547 and QQ plot. We calculated the effect size of each significant genetic variant in the following way (Pallares et al.,
548 2014):

549 $$\text{Effect size} = (\beta^2 * G_v) / P_v \quad (1)$$

550 Where G_v is the variance of the genotype at the focal genetic variant, and P_v is the total variance of the pupal case
551 length phenotype. β is reported as "SNPWeight" for each genetic variant in the FastLMM output.

552 We predicted the effects of GWA significant genetic variants by using SnpEff (Cingolani et al., 2012) with default
553 parameters, and taken all the corresponding affected protein-coding genes as associating genes. In short, all the protein-
554 coding genes within 5 kb up/down-stream of focal genetic variant were taken as its associating genes. We also tested
555 the genotypic linkage disequilibrium (LD) for each GWA significant genetic variant and other genetic variants by
556 calculating the squared correlation estimator r^2 with PLINK v1.90 (Purcell et al., 2007). A significant genetic region
557 (QTL) was defined by the position of the most distant downstream and upstream genetic variants showing a minimum
558 r^2 of 0.8 to the focal GWA significant genetic variant. The combined associating genes from SnpEff predictions and
559 the genes within the QTL regions were considered as GWA candidate genes.

560

561 *Expression analysis*

562 We downloaded the *Drosophila melanogaster* developmental raw Illumina paired-end RNA-Seq data from Graveley
563 et al (Graveley et al., 2011) (NCBI SRA accession: SRP001065). This dataset included the transcriptome of 27 distinct
564 developmental stages covering all the major steps in *Drosophila* life cycle, which were further collapsed into six super-
565 developmental stages: *1) Early embryo*: including 0-2 hrs. embryo, 2-4 hrs. embryo, 4-6 hrs. embryo, and 6-8 hrs.
566 embryo; *2) Middle embryo:* including 8-10 hrs. embryo, 10-12 hrs. embryo, 12-14 hrs. embryo, and 14-16 hrs. embryo;
567 *3) Late embryo:* including 16-18 hrs. embryo, 18-20 hrs. embryo, 20-22 hrs. embryo, and 22-24 hrs. embryo; *4) Larval*
568 *stage:* including L1 stage, L2 stage, L3 stage of 12 hr. post-molt, L3 stage of dark blue gut (PS1-2), L3 stage of light
569 blue gut (PS3-6), and L3 stage of clear blue gut (PS7-9); *5) Pupal stage:* including white prepupa (WPP), 12 hr. after
570 WPP (P5), 24 hr. after WPP (P6), 2 days after WPP (P8), 3 days after WPP (P9-10), and 4 days after WPP (P15); *6)*
571 *Adult stage:* male/female after 1-day eclosion, male/female after 5-day eclosion, and male/female after 30-day eclosion.

572 We trimmed and filtered the low-quality raw Illumina RNA-Seq reads sequenced from the sample of each
573 developmental stage by using Fastp program v0.20.0 (Chen et al., 2018), and only included the sequencing reads with
574 minimum length of 20bp and average quality score of 20 for further analysis. The filtered RNA-Seq reads were aligned
575 to *Drosophila melanogaster* v6 reference genome sequence with HISAT2 v2.1.0 (Kim et al., 2019), taking advantage
576 of the protein-coding gene annotation in Flybase v6.32 (Attrill et al., 2016) by using the --exon option of the hisat2-
577 build. Then we counted the fragments mapped to the annotated genes with featureCounts v1.6.3 (Liao et al., 2014),

578 and calculated the expression level of each annotated coding gene from the sample of each developmental stage in the
579 unit of FPKM (Fragments Per Kilobase of transcript per Million mapped reads).

580 We calculated the gene expression enrichment fold changes across these 27 developmental stages as the ratios between
581 the fractions of genes with expression (FPKM >0) for GWA candidate genes and that of total protein coding gene
582 annotated in Flybase v6.32 (Attrill et al., 2016). We took the average expression enrichment statistics of males and
583 females in the adult stages as for representation. The statistical significances (p-values) of expression enrichment
584 patterns were computed by using Fisher's exact tests.

585

586 *Automatic measurement of pupation site choice*

587 We also assessed the pupation height (an indicator of pupation site choice) for the gene disruption stocks used in the
588 experimental validation tests as shown below. The measurement of pupation height follows the procedure described
589 in (Zhang et al., 2020). In brief, the pupation height was defined as the distance from the vertical coordinate of pupation
590 site (pupal centre) to the food surface in the vial in millimetre (mm). As the pupation height measurements are sensitive
591 to the pupal density in the vial, the raw measurements for pupation height were further adjusted according to the
592 equations described in (Zhang et al., 2020).

593

594 *Functional validation experiments*

595 We validated the phenotypic effects of 10 GWA candidate genes from 11 gene disruption mutagenesis stocks (The
596 gene *Drl-2* was tested twice with different disruption locations), based on their stock availability in the Drosophila
597 gene disruption project (Bellen et al., 2011) (Supplementary Table S1c). All these mutagenesis stocks were conducted
598 with transposon insertion disruption, for which each focal gene was disrupted by a transposon (Minos, P-element or
599 PBac) insertion in the gene region (either coding or regulatory region). The exact landing site of the transposon was
600 determined based on the sequencing reads of both sides of flanking regions (Bellen et al., 2011). The detailed
601 information about these gene mutagenesis stocks and their co-isogenic progenitors is provided in Supplementary Table
602 S1c. Except one disruption stock (Blooming stock ID: 19060; disrupted gene: *Pxn*), all the rest 10 stocks are
603 homozygous insertion viable, and thus were tested as homozygotes (In case of semi-lethal, only the adult flies with no
604 balancer maker were used for experimental tests). For each test, at least 8 vials (15 females and 5 males introduced to
605 each vial) were set up for each insertion and its progenitor stock, and their phenotypic values (pupal case length and
606 pupation height) were measured by following the above phenotyping pipeline. The experimental tests on each pair of
607 gene disruption stock and progenitor stock were conducted in the same experimental round, thus any potential influence
608 (food preparation, incubator environment, et al) on the phenotype measurements can be avoided. The experimental
609 effect size of each gene disruption was calculated as the average deviation of phenotypic values of gene disruption
610 stock from that of its progenitor stock. The statistical differences of the phenotypic measurements between gene
611 disruption stocks and their progenitor stocks were performed by using Wilcoxon rank-sum tests.

612 We developed a distinct approach to test the phenotypic effect of the above homozygous complete-lethal stock (Zhang
613 et al., 2020), which is segregating balancer chromosomes with Tubby (Tb) as visible marker at pupal case (short
614 rounded pupae). We crossed this gene disruption stock (virgin females) with its progenitor stock (males) to generate
615 F1 generation, and only measured the phenotypic values of hemizygous insertion individuals without Tb maker (no
616 presence of balancer chromosome), and then compared with that of its progenitor stock. The presence/absence of Tb
617 maker for individual pupae was distinguished by their pupal case length (>73 pixels for Areashape_major_len from
618 the output of CellProfiler), on the observation of the clear distinction between individual pupae with and without Tb
619 marker (Zhang et al., 2020). Moreover, a manual curation was performed to further separate ambiguous individuals. It
620 should be noted that the phenotypic effect of this homozygous complete-lethal stock was tested in hemizygote manner
621 (Blooming stock ID: 19060; disrupted gene: *Pxn*).

622 Additionally, we also experimentally tested the phenotypic effects of 45 selected genes, on the basis of the following
623 criteria: 1) the GWA p-value on the focal gene falls out the significance threshold (1×10^{-5}); 2) the focal gene has
624 detectable expression in relevant development stages for pupal case length (late larval and pupal stage; FPKM >0),
625 and 3) the focal gene has homozygous transposon insertion disruption viable stock(s) from the panel of *Drosophila*
626 gene disruption project (Bellen et al., 2011). The latter criterion biases against essential genes (approximately 30% are
627 not homozygous viable), but otherwise the selection was essentially random. Two of the above chosen genes (*CG14007*,
628 and *CG42260*) have 2 qualified transposon insertion stocks with disruption in different locations (while one disruption
629 stock for each of the rest 43 genes), and thus in total 47 transposon insertion stocks were applied for the experimental
630 tests. The phenotyping test experiments were conducted with the same procedure as stated above. The GWA
631 association p-value was taken as the lowest p-values from all the genetic variants within the gene and the 5kb up/down-
632 stream of target gene.

633 Furthermore, we analysed the pattern on absolute pupal case length phenotypic effect size (absolute deviation of the
634 phenotypic values of gene disruption stock from that of progenitor stock) of the disruption of the genes, based on their
635 categories: 1) GWA candidate gene or randomly chosen gene; 2) the gene region of the transposon insertion landing
636 site (coding region, intron or UTR); 3) the percentage of disruption on gene's coding region. For the latter one, only
637 the gene disruptions in the coding region were included for analysis, and the average percentage of disruption was
638 taken as representative, in case of multiple transcripts for a focal gene. The statistical difference on the absolute pupal
639 case length phenotypic effect size between two categorized groups were tested by using Wilcoxon rank-sum tests, and
640 the correlation between the absolute pupal case length phenotypic effect size and percentage of gene's coding region
641 disruption with Pearson's correlation test.

642

643 *Genetic variant density analysis*

644 We compared the genetic variant density (# of genetic variants per kb region) pattern between GWA candidate genes
645 and randomly chosen genes (the subset which have significant phenotypic effects on pupal case length, p-value$\leqslant$0.05),
646 based on the information from of DGRP freeze 2 (Huang et al., 2014). The genetic variants within 5kp up/downstream
647 the focal gene (include the gene itself) were included for the analysis. Two types of genetic variants were analysed: 1)

648 the total genetic variants; 2) the "good" genetic variants that passed the criteria for GWA analysis, *i.e.*, with missing
649 values below 20% and minor allele frequency above 5%. For the latter one, we also compared the minor allele
650 frequency (MAF) pattern between these two groups of genes. Overall comparisons of genetic variant density between
651 two groups of genes (whether GWA candidate genes have on average higher variant density/MAF than random genes)
652 were performed with one-sided Wilcoxon rank-sum test.

653

654 *Data availability*

655 All the *Drosophila melanogaster* strains used in this study are public available through either Bloomington *Drosophila*
656 Stock centre (https://bdsc.indiana.edu/) or EHIME *Drosophila* stock centre (https://kyotofly.kit.jp/cgi-
657 bin/ehime/index.cgi). The detailed information about the global wildtype, DGRP inbred lines, gene disruption stocks
658 and their progenitors can be found in the Supplementary Table 1. The primers used for *Wolbachia* infection detection
659 are listed in the above text of this section.

660

661 **Acknowledgements**

662 We thank the lab members for helpful discussions and suggestions. We thank Anita Moeller, Elke Blohm Sievers, and
663 Michaela Schwarz for their excellent technical help with conducting this experiment. This work was supported by
664 institutional funding through the Max-Planck Society.

665

666 **Competing interests**

667 DT: Senior editor, eLife. The other authors declare that no competing interests exist.

668

669 **Supporting information**

670 **Supplementary Table 1:** Information for the fly stocks used in this study

671

672

673 **References**

674 Attrill H, Falls K, Goodman JL, Millburn GH, Antonazzo G, Rey AJ, Marygold SJ. 2016. Flybase: Establishing a
675 gene group resource for Drosophila melanogaster. *Nucleic Acids Res* **44**:D786-792. doi:10.1093/nar/gkv1046

676 Bellen HJ, Levis RW, He Y, Carlson JW, Evans-Holm M, Bae E, Kim J, Metaxakis A, Savakis C, Schulze KL,
677 Hoskins RA, Spradling AC. 2011. The Drosophila gene disruption project: Progress using transposons with
678 distinctive site specificities. *Genetics* **188**:731–743. doi:10.1534/genetics.111.126995

679 Boyle EA, Li YI, Pritchard JK. 2017. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*
680 **169**:1177–1186. doi:10.1016/j.cell.2017.05.038

681 Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA,
682 Moffat J, Golland P, Sabatini DM. 2006. CellProfiler: Image analysis software for identifying and quantifying
683 cell phenotypes. *Genome Biol* **7**:R100. doi:10.1186/gb-2006-7-10-r100

684 Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**:i884–
685 i890. doi:10.1093/bioinformatics/bty560

686 Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for
687 annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of
688 Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**:80–92. doi:10.4161/fly.19695

689 Dembeck LM, Böröczky K, Huang W, Schal C, Anholt RR, Mackay TF. 2015. Genetic architecture of natural
690 variation in cuticularhydrocarbon composition in Drosophila melanogaster. *Elife* **4**:e09861.
691 doi:10.7554/eLife.09861

692 Durham MF, Magwire MM, Stone EA, Leips J. 2014. Genome-wide analysis in Drosophila reveals age-specific
693 effects of SNPs on fitness traits. *Nat Commun* **5**:4338. doi:10.1038/ncomms5338

694 Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, Van Baren MJ, Boley N, Booth
695 BW, Brown JB, Cherbas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D,
696 Tuch BB, Zaleski C, Zhang D, Blanchette M, Dudoit S, Eads B, Green RE, Hammonds A, Jiang L, Kapranov
697 P, Langton L, Perrimon N, Sandler JE, Wan KH, Willingham A, Zhang Y, Zou Y, Andrews J, Bickel PJ,
698 Brenner SE, Brent MR, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Oliver B, Celniker SE. 2011. The
699 developmental transcriptome of Drosophila melanogaster. *Nature* **471**:473–479. doi:10.1038/nature09715

700 Hoffmann AA, Rieseberg LH. 2008. Revisiting the Impact of Inversions in Evolution: From Population Genetic
701 Markers to Drivers of Adaptive Shifts and Speciation? *Annu Rev Ecol Evol Syst* **39**:21–42.
702 doi:10.1146/annurev.ecolsys.39.110707.173532

703 Huang W, Massouras A, Inoue Y, Peiffer J, Ràmia M, Tarone AM, Turlapati L, Zichner T, Zhu D, Lyman RF,
704 Magwire MM, Blankenburg K, Carbone MA, Chang K, Ellis LL, Fernandez S, Han Y, Highnam G, Hjelmen
705 CE, Jack JR, Javaid M, Jayaseelan J, Kalra D, Lee S, Lewis L, Munidasa M, Ongeri F, Patel S, Perales L,
706 Perez A, Pu LL, Rollmann SM, Ruth R, Saada N, Warner C, Williams A, Wu YQ, Yamamoto A, Zhang Y,
707 Zhu Y, Anholt RRH, Korbel JO, Mittelman D, Muzny DM, Gibbs RA, Barbadilla A, Johnston JS, Stone EA,
708 Richards S, Deplancke B, Mackay TFC. 2014. Natural variation in genome architecture among 205 Drosophila
709 melanogaster Genetic Reference Panel lines. *Genome Res* **24**:1193–1208. doi:10.1101/gr.171546.113

710 Jones MDR, Reiter P. 1975. Entrainment of the pupation and adult activity rhythms during development in the
711 mosquito Anopheles gambiae. *Nature* **254**:242–244. doi:10.1038/254242a0

712 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with
713 HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**:907–915. doi:10.1038/s41587-019-0201-4

714 Lattao R, Bonaccorsi S, Guan X, Wasserman SA, Gatti M. 2011. Tubby-tagged balancers for the drosophila X and
715 second chromosomes. *Fly (Austin)* **5**:369–370. doi:10.4161/fly.5.4.17283

716 Lee YCG, Yang Q, Chi W, Turkson SA, Du WA, Kemkemer C, Zeng ZB, Long M, Zhuang X. 2017. Genetic

717 architecture of natural variation underlying adult foraging behavior that is essential for survival of Drosophila
718 melanogaster. *Genome Biol Evol* **9**:1357–1369. doi:10.1093/gbe/evx089

719 Liao Y, Smyth GK, Shi W. 2014. FeatureCounts: An efficient general purpose program for assigning sequence reads
720 to genomic features. *Bioinformatics* **30**:923–930. doi:10.1093/bioinformatics/btt656

721 Lippert C, Listgarten J, Liu Y, Kadie CM, Davidson RI, Heckerman D. 2011. FaST linear mixed models for genome-
722 wide association studies. *Nat Methods* **8**:833–835. doi:10.1038/nmeth.1681

723 Liu X, Li YI, Pritchard JK. 2019. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell*
724 **177**:1022–1034. doi:10.1016/j.cell.2019.04.014

725 Loubéry S, Seum C, Moraleda A, Daeden A, Fürthauer M, Gonzalez-Gaitan M. 2014. Uninflatable and notch control
726 the targeting of sara endosomes during asymmetric division. *Curr Biol* **25**:817–818.
727 doi:10.1016/j.cub.2014.07.054

728 Mackay TFC, Huang W. 2018. Charting the genotype–phenotype map: lessons from the Drosophila melanogaster
729 Genetic Reference Panel. *Wiley Interdiscip Rev Dev Biol* **7**:e289. doi:10.1002/wdev.289

730 MacKay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y, Magwire MM, Cridland
731 JM, Richardson MF, Anholt RRH, Barrón M, Bess C, Blankenburg KP, Carbone MA, Castellano D, Chaboub
732 L, Duncan L, Harris Z, Javaid M, Jayaseelan JC, Jhangiani SN, Jordan KW, Lara F, Lawrence F, Lee SL,
733 Librado P, Linheiro RS, Lyman RF, MacKey AJ, Munidasa M, Muzny DM, Nazareth L, Newsham I, Perales
734 L, Pu LL, Qu C, Ràmia M, Reid JG, Rollmann SM, Rozas J, Saada N, Turlapati L, Worley KC, Wu YQ,
735 Yamamoto A, Zhu Y, Bergman CM, Thornton KR, Mittelman D, Gibbs RA. 2012. The Drosophila
736 melanogaster Genetic Reference Panel. *Nature* **482**:173–178. doi:10.1038/nature10811.

737 Pallares LF, Harr B, Turner LM, Tautz D. 2014. Use of a natural hybrid zone for genomewide association mapping
738 of craniofacial traits in the house mouse. *Mol Ecol* **23**:5756–5770. doi:10.1111/mec.12968

739 Price GM. 1970. Pupation Inhibiting Factor in the Larva of the Blowfly Calliphora erythrocephala. *Nature* **228**:876–
740 877. doi:10.1038/228876a0

741 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly
742 MJ, Sham PC. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage
743 Analyses. *Am J Hum Genet* **81**:559–575. doi:10.1086/519795

744 Reeves RG, Tautz D. 2017. Automated Phenotyping Indicates Pupal Size in Drosophila Is a Highly Heritable Trait
745 with an Apparent Polygenic Basis. *G3 (Bethesda)* **7**:1277–1286. doi:10.1534/g3.117.039883

746 Richardson MF, Weinert LA, Welch JJ, Linheiro RS, Magwire MM, Jiggins FM, Bergman CM. 2012. Population
747 Genomics of the Wolbachia Endosymbiont in Drosophila melanogaster. *PLoS Genet* **8**:e1003129.
748 doi:10.1371/journal.pgen.1003129

749 Sakurai M, Aoki T, Yoshikawa S, Santschi LA, Saito H, Endo K, Ishikawa K, Kimura KI, Ito K, Thomas JB, Hama
750 C. 2009. Differentially expressed Drl and Drl-2 play opposing roles in Wnt5 signaling during drosophila
751 olfactory system development. *J Neurosci* **29**:4972–4980. doi:10.1523/JNEUROSCI.2821-08.2009

752 Schmidt JM, Battlay P, Gledhill-Smith RS, Good RT, Lumb C, Fournier-Level A, Robin C. 2017. Insights into DDT
753 resistance from the drosophila melanogaster genetic reference panel. *Genetics* **207**:1181–1193.
754 doi:10.1534/genetics.117.300310

755 SPSS Inc. 2007. SPSS Advanced Statistics 17.0. *Statistics (Ber)*.

756 Turner SD. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots, bioRxiv.
757 doi:10.1101/005165

758 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, Amin N,
759 Buchkovich ML, Croteau-Chonka DC, Day FR, Duan Y, Fall T, Fehrmann R, Ferreira T, Jackson AU,

760 Karjalainen J, Lo KS, Locke AE, Mägi R, Mihailov E, Porcu E, Randall JC, Scherag A, Vinkhuyzen AAE,
761 Westra HJ, Winkler TW, Workalemahu T, Zhao JH, Absher D, Albrecht E, Anderson D, Baron J, Beekman M,
762 Demirkan A, Ehret GB, Feenstra B, Feitosa MF, Fischer K, Fraser RM, Goel A, Gong J, Justice AE, Kanoni S,
763 Kleber ME, Kristiansson K, Lim U, Lotay V, Lui JC, Mangino M, Leach IM, Medina-Gomez C, Nalls MA,
764 Nyholt DR, Palmer CD, Pasko D, Pechlivanis S, Prokopenko I, Ried JS, Ripke S, Shungin D, Stancáková A,
765 Strawbridge RJ, Sung YJ, Tanaka T, Teumer A, Trompet S, Van Der Laan SW, Van Setten J, Van Vliet-
766 Ostaptchouk J V., Wang Z, Yengo L, Zhang W, Afzal U, Ärnlöv J, Arscott GM, Bandinelli S, Barrett A, Bellis
767 C, Bennett AJ, Berne C, Blüher M, Bolton JL, Böttcher Y, Boyd HA, Bruinenberg M, Buckley BM, Buyske S,
768 Caspersen IH, Chines PS, Clarke R, Claudi-Boehm S, Cooper M, Daw EW, De Jong PA, Deelen J, Delgado G,
769 Denny JC, Dhonukshe-Rutten R, Dimitriou M, Doney ASF, Dörr M, Eklund N, Eury E, Folkersen L, Garcia
770 ME, Geller F, Giedraitis V, Go AS, Grallert H, Grammer TB, Gräßler J, Grönberg H, De Groot LCPGM,
771 Groves CJ, Haessler J, Hall P, Haller T, Hallmans G, Hannemann A, Hartman CA, Hassinen M, Hayward C,
772 Heard-Costa NL, Helmer Q, Hemani G, Henders AK, Hillege HL, Hlatky MA, Hoffmann W, Hoffmann P,
773 Holmen O, Houwing-Duistermaat JJ, Illig T, Isaacs A, James AL, Jeff J, Johansen B, Johansson Å, Jolley J,
774 Juliusdottir T, Junttila J, Kho AN, Kinnunen L, Klopp N, Kocher T, Kratzer W, Lichtner P, Lind L, Lindström
775 J, Lobbens S, Lorentzon M, Lu Y, Lyssenko V, Magnusson PKE, Mahajan A, Maillard M, McArdle WL,
776 McKenzie CA, McLachlan S, McLaren PJ, Menni C, Merger S, Milani L, Moayyeri A, Monda KL, Morken
777 MA, Müller G, Müller-Nurasyid M, Musk AW, Narisu N, Nauck M, Nolte IM, Nöthen MM, Oozageer L, Pilz
778 S, Rayner NW, Renstrom F, Robertson NR, Rose LM, Roussel R, Sanna S, Scharnagl H, Scholtens S,
779 Schumacher FR, Schunkert H, Scott RA, Sehmi J, Seufferlein T, Shi J, Silventoinen K, Smit JH, Smith AV,
780 Smolonska J, Stanton A V., Stirrups K, Stott DJ, Stringham HM, Sundström J, Swertz MA, Syvänen AC, Tayo
781 BO, Thorleifsson G, Tyrer JP, Van Dijk S, Van Schoor NM, Van Der Velde N, Van Heemst D, Van Oort FVA,
782 Vermeulen SH, Verweij N, Vonk JM, Waite LL, Waldenberger M, Wennauer R, Wilkens LR, Willenborg C,
783 Wilsgaard T, Wojczynski MK, Wong A, Wright AF, Zhang Q, Arveiler D, Bakker SJL, Beilby J, Bergman
784 RN, Bergmann S, Biffar R, Blangero J, Boomsma DI, Bornstein SR, Bovet P, Brambilla P, Brown MJ,
785 Campbell H, Caulfield MJ, Chakravarti A, Collins R, Collins FS, Crawford DC, Cupples LA, Danesh J, De
786 Faire U, Den Ruijter HM, Erbel R, Erdmann J, Eriksson JG, Farrall M, Ferrannini E, Ferrières J, Ford I,
787 Forouhi NG, Forrester T, Gansevoort RT, Gejman P V., Gieger C, Golay A, Gottesman O, Gudnason V,
788 Gyllensten U, Haas DW, Hall AS, Harris TB, Hattersley AT, Heath AC, Hengstenberg C, Hicks AA, Hindorff
789 LA, Hingorani AD, Hofman A, Hovingh GK, Humphries SE, Hunt SC, Hypponen E, Jacobs KB, Jarvelin MR,
790 Jousilahti P, Jula AM, Kaprio J, Kastelein JJP, Kayser M, Kee F, Keinanen-Kiukaanniemi SM, Kiemeney LA,
791 Kooner JS, Kooperberg C, Koskinen S, Kovacs P, Kraja AT, Kumari M, Kuusisto J, Lakka TA, Langenberg C,
792 Le Marchand L, Lehtimäki T, Lupoli S, Madden PAF, Männistö S, Manunta P, Marette A, Matise TC,
793 McKnight B, Meitinger T, Moll FL, Montgomery GW, Morris AD, Morris AP, Murray JC, Nelis M, Ohlsson
794 C, Oldehinkel AJ, Ong KK, Ouwehand WH, Pasterkamp G, Peters A, Pramstaller PP, Price JF, Qi L, Raitakari
795 OT, Rankinen T, Rao DC, Rice TK, Ritchie M, Rudan I, Salomaa V, Samani NJ, Saramies J, Sarzynski MA,
796 Schwarz PEH, Sebert S, Sever P, Shuldiner AR, Sinisalo J, Steinthorsdottir V, Stolk RP, Tardif JC, Tönjes A,
797 Tremblay A, Tremoli E, Virtamo J, Vohl MC, Amouyel P, Asselbergs FW, Assimes TL, Bochud M, Boehm
798 BO, Boerwinkle E, Bottinger EP, Bouchard C, Cauchi S, Chambers JC, Chanock SJ, Cooper RS, De Bakker
799 PIW, Dedoussis G, Ferrucci L, Franks PW, Froguel P, Groop LC, Haiman CA, Hamsten A, Hayes MG, Hui J,
800 Hunter DJ, Hveem K, Jukema JW, Kaplan RC, Kivimaki M, Kuh D, Laakso M, Liu Y, Martin NG, März W,
801 Melbye M, Moebus S, Munroe PB, Njølstad I, Oostra BA, Palmer CNA, Pedersen NL, Perola M, Pérusse L,
802 Peters U, Powell JE, Power C, Quertermous T, Rauramaa R, Reinmaa E, Ridker PM, Rivadeneira F, Rotter JI,
803 Saaristo TE, Saleheen D, Schlessinger D, Slagboom PE, Snieder H, Spector TD, Strauch K, Stumvoll M,
804 Tuomilehto J, Uusitupa M, Van Der Harst P, Völzke H, Walker M, Wareham NJ, Watkins H, Wichmann HE,
805 Wilson JF, Zanen P, Deloukas P, Heid IM, Lindgren CM, Mohlke KL, Speliotes EK, Thorsteinsdottir U,
806 Barroso I, Fox CS, North KE, Strachan DP, Beckmann JS, Berndt SI, Boehnke M, Borecki IB, McCarthy MI,
807 Metspalu A, Stefansson K, Uitterlinden AG, Van Duijn CM, Franke L, Willer CJ, Price AL, Lettre G, Loos
808 RJF, Weedon MN, Ingelsson E, O'Connell JR, Abecasis GR, Chasman DI, Goddard ME, Visscher PM,
809 Hirschhorn JN, Frayling TM. 2014. Defining the role of common variation in the genomic and biological
810 architecture of adult human height. *Nat Genet* **46**:1173–1186. doi:10.1038/ng.3097

811 Wray NR, Wijmenga C, Sullivan PF, Yang J, Visscher PM. 2018. Common Disease Is More Complex Than Implied
812 by the Core Gene Omnigenic Model. *Cell*. doi:10.1016/j.cell.2018.05.051

813 Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. 2013. Pitfalls of predicting complex traits from
814 SNPs. *Nat Rev Genet* **14**:507–515. doi:10.1038/nrg3457

815 Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG,
816 Montgomery GW, Goddard ME, Visscher PM. 2010. Common SNPs explain a large proportion of the
817 heritability for human height. *Nat Genet* **42**:565–569. doi:10.1038/ng.608

818 Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, De Andrade M, Feenstra B,
819 Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA,
820 Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME, Visscher PM. 2011. Genome partitioning of genetic
821 variation for complex traits using common SNPs. *Nat Genet* **43**:519–525. doi:10.1038/ng.823

822 Zeh JA, Bonilla MM, Adrian AJ, Mesfin S, Zeh DW. 2012. From father to son: Transgenerational effect of
823 tetracycline on sperm viability. *Sci Rep* **2**:375. doi:10.1038/srep00375

824 Zhang W, Reeves GR, Tautz D. 2020. Identification of a genetic network for an ecologically relevant behavioural
825 phenotype in Drosophila melanogaster. *Mol Ecol* **29**:502–518. doi:10.1111/mec.15341

826 Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* **44**:821–
827 824. doi:10.1038/ng.2310

828