Genes sharing the same *D. melanogaster* homolog

One potential source of bias in determining the level of genetic covariance within functional groups compared that among random sets of genes was that in some of the selected groups, *D. melanogaster* homologs corresponded to more than one gene in a functional group. These genes may have had a higher chance of genetically covarying because they potentially shared recent common ancestry. We tested if the level of genetic covariance between pairs of genes sharing the same *D. melanogaster* homolog was higher than the average level of genetic covariance between all pairs of selected genes. We performed a Wilcoxon test comparing the median of the bivariate covariances values between genes that shared their *D. melanogaster* homolog and all the other bivariate covariances calculated between pairs of genes in the selected groups. We found that the genes sharing a *D. melanogaster* homolog did not have higher genetic covariance that other pairs of genes ($W = 497130$, $p = 0.77$).

Pleiotropy within functional groups

To determine whether levels of mutational and standing genetic covariances in the functional groups were larger than expected due to sampling error alone, we implemented the randomisation test used previously in analyses of these data (McGUIGAN *et al.* 2014a; McGUIGAN *et al.* 2014b; BLOWS *et al.* 2015). For each functional group, we ran the mixed model (2) on data where the phenotypic scores for each gene were shuffled among lines, while keeping the two replicate measures for the five probes within each line together. This shuffling approach disrupted the covariances between genes at the among-line level, whilst maintaining the estimates of univariate variance for each gene at the among-line, within-line and residual levels (See details in McGUIGAN *et al.* 2014b). We created 50 such independent datasets each for the M and G lines to obtain 50 $\mathbf{M}_{se}$ and 50 $\mathbf{G}_{se}$ covariance matrices that represented sampling error (Table 2). Where a model did not converge in the shuffled data, we used the best unconverged model. In 15 of 1400 cases, the returned eigenvalues for $\mathbf{M}_{se}$ or $\mathbf{G}_{se}$ were extreme outliers ($>10^{10}$). We removed these outliers and re-ran the model on an additional shuffled data set to obtain the 50 required sets of eigenvalues based solely on sampling error. Significance of an observed eigenvalue was established if the value was larger than the upper 95% CI obtained from the distribution of sampling error eigenvalues.

The functional groups based on GO term enrichment typically contained more mutational covariance than expected by chance with all but one functional group having at least one of eigenvalue above the level of mutational covariance based on sampling error alone (Table S1). Furthermore, 9 of the 14 functional groups had a second eigenvalue that was significantly above that expected from sampling error alone (Table S1).

Function-specific standing genetic covariance within functional groups was somewhat less common than function-specific mutational covariance; 10 functional groups showed genetic covariance above the level of sampling error (captured by $\mathbf{G}_{se}$), and five functional groups showed genetic covariance above $\mathbf{G}_{se}$ in two axes ($\boldsymbol{g}_{max}$ and $\boldsymbol{g}_2$) (Table S1). It should be noted that there were 30 G-lines compared to 41 M-lines, and therefore the power of our tests for genetic covariance above sampling error or biological background was lower than for mutational covariance.

Altogether, only one of the selected functional group (GO:0007379, "segment specification") was not associated with mutational or standing genetic covariance above levels of covariances expected due to sampling error alone. As there was no sign of covariance (i.e., variational modularity) in this group, we did not consider it further. All other 13 functional groups typically contained more mutational covariance than expected by chance (Table S1).

Table S1: Mutational and genetic variance in the first and second eigenvectors of **M** and **G**. The 95% confidence intervals (CI) refer to the magnitude of eigenvalues due solely to sampling error, estimated from the 50 $\mathbf{M}_{se}$ and 50 $\mathbf{G}_{se}$ matrices. Significance levels are * > 95%, ** > 99%, and *** outside confidence intervals.

	λm_{max} [CI]	λm_2 [CI]	λg_{max} [CI]	λg_2 [CI]
Chorion	0.58 *** [0.38; 0.54]	0.33 [0.21; 0.36]	0.45 [0.40; 0.65]	0.33 [0.12; 0.41]
Amino A	0.64 *** [0.29; 0.49]	0.27 [0.19; 0.29]	1.19 * [0.92; 1.19]	0.72 [0.57; 0.82]
Segment	0.51 [0.37; 0.55]	0.29 [0.27; 0.40]	1.00 [0.93; 1.13]	0.40 [0.37; 0.54]
NeuroT	1.05 *** [0.52; 0.68]	0.31 *** [0.38; 0.51]	0.86 [0.69; 1.00]	0.62 [0.53; 0.71]
GT	1.16 *** [0.60; 0.83]	1.04 *** [0.45; 0.67]	1.48 *** [1.11; 1.47]	1.19 *** [0.84; 1.12]
Pro-DNA	1.11 *** [0.69; 0.84]	0.71 *** [0.39; 0.64]	1.25 * [0.85; 1.23]	0.60 * [0.60; 0.89]
Bacterium	2.54 *** [0.85; 1.11]	0.94 *** [0.70; 0.91]	2.44 *** [1.33; 1.77]	1.78 *** [1.10; 1.43]
Chitin	1.81 *** [0.77; 1.01]	1.08 *** [0.60; 0.83]	1.63 * [1.14; 1.60]	1.28 [1.03; 1.35]
Sensory	1.94 *** [0.94; 1.21]	0.89 [0.75; 0.95]	2.12 *** [1.35; 2.04]	1.85 *** [1.26; 1.54]
Ion Tsp	1.51 *** [0.68; 0.92]	0.90 *** [0.55; 0.77]	1.84 [1.28; 1.87]	1.42 [1.05; 1.44]
Heme	2.54 *** [0.98; 1.23]	2.20 *** [0.75; 0.93]	2.65 *** [1.84; 2.41]	1.97 [1.60; 1.98]
Cuticle	4.71 *** [1.29; 1.61]	1.62 *** [1.04; 1.33]	3.09 *** [1.71; 2.27]	1.85 *** [1.52; 1.82]
Cell Fate	2.41 *** [1.10; 1.41]	1.41 *** [0.94; 1.22]	2.11 *** [1.64; 1.96]	1.61 [1.37; 1.71]
Endopep	3.01 *** [1.42; 1.84]	2.10 *** [1.22; 1.48]	4.60 *** [2.78; 3.47]	3.70 *** [2.49; 3.01]

Enrichment in KEGG pathways

We identified two independent KEGG pathways significantly enriched in genes with mutational variance. These two pathways, Starch and sucrose metabolism (dme00500, p=0.002) and limonene and pinene degradation (dme00903, p=0.039), have been inferred to be present in *Drosophila* through homology with plant species. There was substantial overlap between these two KEGG pathways and two of the selected GO terms. Thirteen of the 24 genes present in the starch and sucrose metabolism pathway were part of the glucuronosyltransferase activity GO term (group E), and 24 of the 26 genes in limonene and pinene degradation pathways could also be found in the heme binding GO term (group K). In all analyses, the KEGG pathways groups gave the same qualitative results as the GO term groups they overlap with, and we therefore only present the results for the GO terms analysis.