

Sex-dependent and -independent regulatory systems of size variation in natural populations

Hirokazu Okada, Ryohei Yagi, Vincent Gardeux, Bart Deplancke and Ernst Hafen.

Review timeline:

Submission date:	20 th May 2019
Editorial Decision:	26 th June 2019
Revision received:	23 rd September 2019
Editorial Decision:	24 th October 2019
Revision received:	29 th October 2019
Accepted:	30 th October 2019

Editor: Hirokazu Okada

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision

26th June 2019

Thank you for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who agreed to evaluate your manuscript. You will see from the comments below that the referees find the manuscript to be of interest. They raise, however, several issues, which should be convincingly addressed in a revision.

Without reiterating all the points raised in the reviews below, some of the more substantial issues are the following:

- The comments from referee #1 regarding the data analysis and statistics should be addressed.
- Further experimental validations suggested by referee #2 need to be included to better support the conclusions of the study.

REFeree REPORTS.

Reviewer #1:

This is a very well written paper that is data dense. The topic fits the scope of the journal very well and would likely appeal to the readership. The paper uses the *Drosophila* genetic reference panel (DGRP) as a reference population to dissect the genetics of wing size using a systems biology approach. The data contain several layers. 1) DNA sequences collected by the DGRP; 2) gene expression of both sexes in 32 extreme lines in imaginal discs. 3) protein expression from a previous study in 30 lines; I believe not all of these lines overlap with the 32 RNA expression lines but it's not very clear from the paper. 4) wing size trait data from these selected lines and other lines, also collected from a previous study. Then the authors performed a series of analyses to correlate these different levels of variation and interpret their data using analysis such as pathway enrichment, etc..

The analyses are sound and comprehensive. They found that canonical growth genes don't contribute much to the natural variation. Instead, they found a bunch of non-canonical genes that seem to regulate wing size variation within each sex. The most interesting and valuable part of the paper is the prediction piece, in which they are able to predict the phenotype of wing size by summing up allelic effects (+ or -) of a handful of SNPs that are associated with the trait and RNA expression. There are several main comments I have regarding the analyses and interpretation, which I think are important.

1) Figure 2E: The y axis of this figure is association P values for type 3. However, my understanding is that your type 3 associations were identified through type 1 analysis. So the P values are not for type 3 but type 1. In addition, I think it would make more sense to make a plot of the same x, and y but for all genes (without selecting canonical growth genes). If all genes have the same pattern, then it's not surprising that a selected subset also shares the same pattern. Based on Figure 2A, the majority of genes are type 3, I suspect this to be true. This is very important and affects some of the interpretations.

2) Figure 4A: the x-axis is SNP effect, but the y axis is minimum p value for a gene. Can you plot the P values of SNPs on the y axis instead? I don't see the necessity to use a gene minimum p value.

3) Figure 5A: I don't understand why this figure allows you to display "mediators". Some graph to explain what your definition of mediators is would be helpful.

4) One of the main conclusions of the paper is that a few trait associated SNPs are able to predict phenotypes with good accuracy. $R^2 = 0.16$ translates to correlation of 0.40, which is pretty good compared to genomic prediction in *Drosophila*. However, it's not clear that when prediction was made on 143 lines whether these lines include the 32 lines used in the present study. The correct analysis is to run this with an independent set of lines. If the 32 lines were among the 143 lines, please remove them and re-evaluate the correlation between the predicted and observed wing sizes. This is very important because including the lines in the test set can significantly bias the prediction accuracy upwards.

Minor comments:

Bottom of page 3: "it is fully unknown" -> "it is not fully known"

Page 4: The contrast between

Page 7, Figure 1E: these RNAs have the smallest sd, do they also have the highest mean expression? I wonder what this means.

Reviewer #2:

Okada et al.

This is an interesting manuscript that attempts to analyse the genetic basis for natural wing size variation between *Drosophila* sexes and between different natural strains. Overall, the authors have done a good job of their analysis and presentation of the findings. I recommend publication.

Minor comments:

1. In the abstract, the authors should mention that the Ecdysone Receptor was identified as being differently expressed between fly lines of different wing size, along with other novel genes not previously implicated in size control. The EcR result is important and should be highlighted rather than be buried. It is interesting that size differences between mammals (including sex differences) are also strongly governed by steroid hormones.

2. Is it also worth mentioning that not a single component of the Insulin-IGF1 pathway was identified in this analysis, despite the fact that IGF1R is crucial in variation between species of dogs (artificially selected)? This might be because IGF1R-variation leads to large dogs with short life spans, while natural variation leads to large mammals with long life spans.

3. Did the authors attempt to RNAi their HINT homolog in the wing? Is there any chance that the variation in HINT might also affect the neighbouring downstream gene eIF2a?

Reviewer #3:

A number of genetic studies in the past have identified genes that have the capability to affect tissue size during development if genetically perturbed. Importantly, however, this does not mean that these genes actually do regulate tissue size during development. They might simply be required for growth, but not be the regulatory node that the biological system uses to control growth physiologically. Hence an important open question is what genes control tissue and animal size in vivo in non-experimentally-perturbed conditions. This interesting study starts addressing this issue by looking at the association of mRNA levels in the developing *Drosophila* wing with final wing size, across fly lines and between males and females. The manuscript is well written. The data are interesting. For the most part, the conclusions are solid and well founded (with the two exceptions mentioned below). The approach presented here seems to have been quite successful, identifying 13 new genes (out of 14 tested) that affect organ size. I think this study will be of interest to a wide range of people interested in development, organ size and growth, patterning, metabolic regulation, and signaling pathways that control cell growth and proliferation.

The following results would need to be strengthened:

1. The size phenotypes presented in Fig 3C would need to be confirmed/repeated with independent RNAi lines. Especially undergrowth phenotypes may reflect unhappy cells, and there is a good chance of them being due to off-target effects of the RNAis.
2. The authors identify CG34015 as an important regulator of wing size, but this is not functionally tested. Does modulation of CG34015 expression (either RNAi or overexpression) affect wing size? Since the nearby SNPs are actually in the neighboring gene *Rbp2*, it would be good to also test this one.

Minor Issues

1. Fig 1F - it is not completely clear from the main text and figure legend how this analysis was done. Were RNA and protein levels correlated within one fly strain, thereby yielding a correlation coefficient, and this was done for all 32 fly strains? (In which case I would expect the total sum of correlation coefficients in the curve in Fig 1F to be 32. However, it looks like the distribution of roughly 1000 correlation coefficients is shown in 1F.) So instead, was the correlation performed comparing RNA and protein for one gene across the 32 fly strains, thereby yielding a correlation coefficient for each gene ?) This should be spelled out a bit more.
2. It is not very clear from the text how the 20 size-associated genes to test were selected. For instance, *yip2* appears at position 392 on the Suppl. Table 3 list when sorted by FDR. It would be good to explain in more detail how these 20 genes were selected, because by showing they were not cherry-picked would strengthen the statement that "This high rate of true positives (13/14=93%) suggests other untested, size-associated genes to be novel growth regulators."
3. Page 18 - should the reference to Fig 2B actually reference Fig 6B?
4. Fig 2F: The literature suggests Stat92E promotes growth. e.g. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2914209/> Why does the association between Stat92E mRNA and size appear to be negative (ie higher levels associated with smaller wings) in Fig 2F? The authors should discuss this.
5. Likewise, in Suppl. Fig 3, higher *Tkv* expression is associated with smaller wings, but in the literature *Tkv* is thought to promote growth?
6. Can the authors discuss the advantages of looking at mRNA levels (as done here) compared to the

protein levels (already published), given that the two don't correlate, and protein levels are what matter at the end of the day? e.g. one can also see lowly-expressed genes, and non-coding RNAs?

1st Revision - authors' response

23rd September 2019

Responses to the reviewer's comments

Reviewer #1:

This is a very well written paper that is data dense. The topic fits the scope of the journal very well and would likely appeal to the readership. The paper uses the *Drosophila* genetic reference panel (DGRP) as a reference population to dissect the genetics of wing size using a systems biology approach. The data contain several layers. 1) DNA sequences collected by the DGRP; 2) gene expression of both sexes in 32 extreme lines in imaginal discs. 3) protein expression from a previous study in 30 lines; I believe not all of these lines overlap with the 32 RNA expression lines but it's not very clear from the paper. 4) wing size trait data from these selected lines and other lines, also collected from a previous study. Then the authors performed a series of analyses to correlate these different levels of variation and interpret their data using analysis such as pathway enrichment, etc. The analyses are sound and comprehensive. They found that canonical growth genes don't contribute much to the natural variation. Instead, they found a bunch of non-canonical genes that seem to regulate wing size variation within each sex. The most interesting and valuable part of the paper is the prediction piece, in which they are able to predict the phenotype of wing size by summing up allelic effects (+ or -) of a handful of SNPs that are associated with the trait and RNA expression. There are several main comments I have regarding the analyses and interpretation, which I think are important.

1) Figure 2E: The y axis of this figure is association P values for type 3. However, my understanding is that your type 3 associations were identified through type 1 analysis. So the P values are not for type 3 but type 1. In addition. I think it would make more sense to make a plot of the same x, and y but for all genes (without selecting canonical growth genes). If all genes have the same pattern, then it's not surprising that a selected subset also shares the same pattern. Based on Figure 2A, the majority of genes are type 3, I suspect this to be true. This is very important and affects some of the interpretations.

Response: The P values for type 3 association, plotted in the y axis of the Figure 2E, were evaluated through type 3 analysis (ANOVA on sex), not through type 1 analysis (regression on size trait). This is now clarified in the figure legend for Figure 2E.

As the reviewer points out, if we make the same plot using all the associated genes (without selecting canonical growth genes), the plot should make a similar pattern to Figure 2E. This is obvious since the majority of the size-associated genes are genes with type 1-exclusive association, as shown in Figure 2A, and the genes with type 1-exclusive association are mostly genes with type 3 association, as shown in Appendix Figure S1. Based on these results, we mentioned in page 8 (line 15) that a large part of the size-associated genes is involved in sexual size dimorphism. However, what is striking is that canonical growth genes followed this pattern even though we have long expected that canonical growth genes should show good associations with size through both type 2 and type 3. The finding is against our long-held thought that canonical genes are main regulators responsible for between-line size variation.

2) Figure 4A: the x-axis is SNP effect, but the y axis is minimum p value for a gene. Can you plot the P values of SNPs on the y axis instead? I don't see the necessity to use a gene minimum p value.

Response: Following the reviewer's suggestion, Figure 4A has been modified so that P values of all the eQTLs (SNPs) are plotted on the y axis. The top SNPs are now labeled with the corresponding gene names.

3) Figure 5A (now as Figure 5B): I don't understand why this figure allows you to display "mediators". Some graph to explain what your definition of mediators is would be helpful.

Response: We defined "mediators" as genes that exhibited triangular associations among SNP variation, the expression level and wing size phenotype. We identified mediator genes, first by selecting RNAs whose expression is doubly associated with both variations in SNPs and wing size traits (absCS/relCS) at 20% FDR, and second by selecting RNAs whose associated SNPs are also associated with wing size traits. We have now added a panel (new Figure 5A) that explains the definition and selection process of mediators. The main text also explains this (page 16-17).

4) One of the main conclusions of the paper is that a few trait associated SNPs are able to predict phenotypes with good accuracy. $R^2 = 0.16$ translates to correlation of 0.40, which is pretty good compared to genomic prediction in *Drosophila*. However, it's not clear that when prediction was made on 143 lines whether these lines include the 32 lines used in the present study. The correct analysis is to run this with an independent set of

lines. If the 32 lines were among the 143 lines, please remove them and re-evaluate the correlation between the predicted and observed wing sizes. This is very important because including the lines in the test set can significantly bias the prediction accuracy upwards.

Response: As the reviewer points out, our population-level prediction included the 32 lines. We checked predictability using the rest of 111 lines only without the 32 lines, and found the correlation with wing size was not good (Pearson correlation coefficient $\rho=0.056/0.041$ (P value=0.28/0.33) for female absolute/relative wing size predictions). This probably occurred because, without the 32 lines with extreme wing sizes, the wing size variation of the 111 lines was much smaller than the population-level size variation. Indeed, the majority of the wing sizes from the 111 lines is concentrated around the population mean as shown in Appendix Figure S10A).

We reasoned that the limited predictability came from the fact that we applied the same gene and SNP sets in the model to the 4 different biological situations (male/female, absolute/relative wing size), even though we identified different sets of mediator genes and SNPs for each situation. Another factor that possibly compromised the predictability is SNP numbers we included in the model. We only used the top-associated SNP for each mediator gene so that the model did not consider the relative importance of genes that may be represented by the number of associated SNPs.

We have reconstructed a modified model that now considers only genes relevant for each situation (sex/wing size trait), and consider all SNPs corresponding to the mediator genes in each situation (as selected in Figure 5B). The genes and SNPs used for each situation are described in Dataset EV8. Now the new model successfully predicts the wing sizes of the medium-ranged 111 lines at the Pearson correlation of 0.14 – 0.20 with statistical significance ($p<0.05$) for all situations (except a case of male and relative wing size, $p=0.078$) (as shown in Appendix Figure S10B). The model now reproduces wing size variation of the 32 extreme lines at the Pearson correlation of 0.58 – 0.72 (as shown in Figure 6B) for all 4 situations, and predicts the population-wide wing size (143 lines) at the Pearson correlation of 0.29 – 0.36 (as shown in Figure 6C).

The main text now describes the new model (page 18 and 19).

Minor comments:

1. Bottom of page 3: "it is fully unknown" -> "it is not fully known"

Response: Corrected as suggested.

2. Page 4: The contrast between

Response: We could not find the expression “The contrast between” in page 4 of the manuscript.

3. Page 7, Figure 1E: these RNAs have the smallest sd, do they also have the highest mean expression? I wonder what this means.

Response: As suggested by the reviewer, we have confirmed that the RNAs have the highest mean expressions (as shown in new Figure 1E). This suggests that the RNAs co-quantified at the protein level are mostly abundant RNAs that probably make them stable (invariable) among lines. Presumably, their functions are vital so that their expression cannot be highly variable among individuals in the population.

We now describe the analysis of the mean expression in the manuscript (page 7, lines 5-7)

Reviewer #2:

This is an interesting manuscript that attempts to analyze the genetic basis for natural wing size variation between *Drosophila* sexes and between different natural strains. Overall, the authors have done a good job of their analysis and presentation of the findings. I recommend publication.

Minor comments:

1. In the abstract, the authors should mention that the Ecdysone Receptor was identified as being differently expressed between fly lines of different wing size, along with other novel genes not previously implicated in size control. The EcR result is important and should be highlighted rather than be buried. It is interesting that size differences between mammals (including sex differences) are also strongly governed by steroid hormones.

Response: We have added the following sentence in the abstract: “Only a few growth genes including a receptor of steroid hormone ecdysone exhibit association with between-line size differences.”

We now have discussed EcR in the “Discussion” section of the manuscript (page 21-22) as follows:

We should note that EcR, one of the canonical growth regulators, exceptionally exhibited a significant association with between-line wing size at $FDR < 20\%$. A molting hormone, ecdysone, and its nuclear receptor EcR have been thought to provide a substantial contribution to *Drosophila* body and organ size (Mirth and Shingleton, 2012). EcR forms a heterodimer with Ultraspiracle (*usp*) and, without ecdysone, binds to the promoter of target genes and represses transcription (Schubiger, 2005). Once EcR binds to ecdysone, the transcriptional repression by the unliganded EcR-*usp* complex is eliminated and EcR, recruiting transcriptional activators and co-factors, activates transcription, thus allowing developmental transitions to proceed. Knockdown of EcR has been shown to derepress transcription, which also allows developmental programs to proceed, mimicking ecdysone action (Mirth et al., 2009). We now observed the negative correlation of EcR levels with between-line wing size variation. A possible explanation would be that higher levels of EcR requires a smaller amount of ecdysone for developmental transitions through transcriptional activation and therefore, higher levels of EcR expedite the transitions, resulting in a shorter growth period and smaller tissues. It is also possible that the variability in ecdysone titers secreted from prothoracic glands and its temporal profiles affect between-line wing size non-autonomously. The relative contribution of EcR compared to that of novel regulators that were identified in the study remains to be determined.

2. Is it also worth mentioning that not a single component of the Insulin-IGF1 pathway was identified in this analysis, despite the fact that IGF1R is crucial in variation between species of dogs (artificially selected)? This might be because IGF1R-variation leads to large dogs with short life spans, while natural variation leads to large mammals with long life spans.

Response: We discuss this issue in the “Discussion” section of the manuscript (page 22):

Insulin-like signaling and mTOR (insulin/TOR) pathway have been thought as the key players in the environmental and physiological control of body/organ size in *Drosophila* (Mirth & Shingleton, 2012). Our list of canonical growth genes includes several key components from the insulin/TOR pathway such as *Ilp2* (insulin-like peptide), *InR* (insulin-like receptor), *Pi3K92E* (catalytic subunit of phosphatidylinositol 3-kinase), *Pten* (phosphatase and tensin homolog), *Akt1* (protein kinase B) and *foxo* (forkhead box, sub-group O transcription factor). Among these, *InR* was identified to be associated with wing size at $FDR < 20\%$ (Fig 2E). The positive association to wing size was through type 3, suggesting its contribution to sexual size dimorphism. GWAS and eQTL mapping in our studies did not find any SNPs associated in the *InR* region. However, GWAS of domestic dogs that experienced intensive selection for size has found that alleles at insulin-like growth factor 1 (*IGF1*) and its receptor (*IGF1R*: mammalian ortholog of *InR*) are critical to the size variation (Sutter et al, 2007; Hoopes et al, 2012). This may indicate that very rare alleles that usually do not survive in natural settings were artificially selected in the domestic dog breeding.

3. Did the authors attempt to RNAi their HINT homolog in the wing? Is there any chance that the variation in HINT might also affect the neighbouring downstream gene *eIF2a*?

Response: We have found a single UAS-RNAi line for CG34015, the HINT homolog, available from Bloomington *Drosophila* stock center. We crossed this line with a hedgehog-Gal4 driver line, as we performed in Figure 3, and observed a significant reduction of size in the P-compartment of the wing in both sexes (as shown in Figure 5C and D). We have thus revealed that CG34015 positively regulates wing size. The experiment is now described in the main text (page 17, lines 11-13).

eIF-2alpha, immediately downstream of CG34015, shows a strong sexual dimorphic expression pattern at both RNA and protein levels (P value= 3×10^{-11} and 4×10^{-6} , respectively, for type 3 association), but does not exhibit between-line size association (P value=0.13 and 0.82 for type 2 association) as shown in Appendix Figure S11. Since RNA expression of CG34015 shows the strongest association only with between-line size differences (P value= 2×10^{-6} for type 2, 0.28 for type 3), the genetic control of *eIF-2alpha* should be unrelated to the one of CG34015. This is discussed in the “Discussion” section of the manuscript (page 23).

Reviewer #3:

A number of genetic studies in the past have identified genes that have the capability to affect tissue size during development if genetically perturbed. Importantly, however, this does not mean that these genes actually do regulate tissue size during development. They might simply be required for growth, but not be the regulatory node that the biological system uses to control growth physiologically. Hence an important open question is what genes control tissue and animal size in vivo in non-experimentally-perturbed conditions. This interesting study starts addressing this issue by looking at the association of mRNA levels in the developing *Drosophila* wing with final wing size, across fly lines and between males and females. The manuscript is well written. The data are interesting. For the most part, the conclusions are solid and well founded (with the two exceptions

mentioned below). The approach presented here seems to have been quite successful, identifying 13 new genes (out of 14 tested) that affect organ size. I think this study will be of interest to a wide range of people interested in development, organ size and growth, patterning, metabolic regulation, and signaling pathways that control cell growth and proliferation.

The following results would need to be strengthened:

1. The size phenotypes presented in Fig 3C would need to be confirmed/repeated with independent RNAi lines. Especially undergrowth phenotypes may reflect unhappy cells, and there is a good chance of them being due to off-target effects of the RNAis.

Response: To address the reproducibility of the RNAi results, we obtained two more independent RNAi lines for 9 genes (out of 14 tested) in addition to the lines we used for RNAi. For the rest of 5 genes, we could not find two more independent RNAi lines available in public but could obtain a single independent RNAi line for each of the genes. We crossed them with the hedgehog-Gal4 driver line and tested RNAi effect on the size of wing P-compartment. As shown in the new Figure 3C that now contains the results from the original and new experiments, all the 4 (negative) control cases resulted in the same range of P/A compartment size ratio, indicating a high reproducibility of the experiments. Consistent with the original result, new 2 cases of CG14207 knockdowns resulted in the same range of the P/A size ratio as the controls (even though the P value of one case claims a significant change), supporting the observation in the original experiment that CG14207 may not contribute to wing size variation. In contrast, all other, new RNAi cases on 13 genes exhibited significant size changes, supporting the original experiment. Knockdown of 10 genes (out of the 13) led to reduced P-compartment sizes in all RNAi cases (even though the strength of the effects varies for some genes), suggesting that these genes are positive growth regulators. However, knockdown of the other 3 genes (CG7173, Ppn and yip2) showed significant size changes in opposite directions, dependent on lines used.

To see the direction of effect associated with wing size variation in natural populations, we have plotted RNA/protein levels of 13 genes against wing size variation in DGRP lines (Figure EV4). The plots revealed that out of the 13 genes, all genes except Ppn are positively correlated with wing size, suggesting that Ppn only is a negative growth regulator and the others are positive regulators. This supports that the 10 genes with consistent size reductions in all RNAi cases function as positive regulators. The plots support CG7173 and yip2 to be positive regulators and Ppn to be negative regulators, but more detailed assessments are needed to determine the direction in the size regulation of the 3 genes.

2. The authors identify CG34015 as an important regulator of wing size, but this is not functionally tested. Does modulation of CG34015 expression (either RNAi or overexpression) affect wing size? Since the nearby SNPs are actually in the neighboring gene Rbp2, it would be good to also test this one.

Response: As described above (in the response to the comments from reviewer 2), a single line is available in public for CG34015 perturbation. We crossed the RNAi line with a hedgehog-Gal4 driver line, as we performed in Figure 3, and observed a significant reduction of size in the P-compartment of the wing in both sexes (as shown in Figure 5C and D). We have thus revealed that CG34015 positively regulates wing size. The experiment is now described in the main text (page 17, lines 11-13).

We have already tested Rbp2 functionally before, as described in our previous GWAS paper. The GWA study tested SNPs located around the Rbp2 gene region but did not identify any single SNPs that were associated with wing size traits. In the current study, eQTL mapping also did not find any SNPs cis-associated with Rbp2 expression at FDR < 20%. However, in the GWAS paper, in addition to the single-marker GWAS, we also performed gene-based GWAS using the sum of chi-squares VEGAS method (Liu *et al.*, 2010), which evaluates P values considering all SNPs within each gene. The gene-based analysis identified Rbp2 within the top 20 associations and we functionally tested Rbp2 by RNAi. We knocked down the gene level in the whole wing using nubbin-Gal4 driver by crossing with two independent RNAi lines. Rbp2 was among genes with the best reductions in size (~10%) in both sexes. In the current study, the Rbp2 RNA level exhibited a weak, positive correlation with between-line wing size variation (P-value=0.05 for type-2 association) but the protein level rather showed a sexual dimorphic expression pattern (P-value=0.78 for type-2, 0.0055 for type-3 association) as shown in Appendix Figure S11. Considering the CG34015's strong association to between-line size variation, it seems that Rbp2 is under a different genetic control and rather contribute to sexual size dimorphism. This is discussed in the "Discussion" section of the manuscript (page 23).

Minor Issues

1. Fig 1F - it is not completely clear from the main text and figure legend how this analysis was done. Were RNA and protein levels correlated within one fly strain, thereby yielding a correlation coefficient, and this was done for all 32 fly strains? (In which case I would expect the total sum of correlation coefficients in the curve in Fig 1F to be 32. However, it looks like the distribution of roughly 1000 correlation coefficients is shown in 1F.)

So instead, was the correlation performed comparing RNA and protein for one gene across the 32 fly strains, thereby yielding a correlation coefficient for each gene?) This should be spelled out a bit more.

Response: The correlation analysis was performed comparing RNA and protein levels for each gene across 56 overlapping biological conditions (28 fly lines x 2 sexes). We revised the description in the manuscript as follows (page 7, lines 7-9): “We next examined the distribution of the gene-based correlation between the levels of RNAs and proteins by comparing both levels across the 56 biological conditions for each gene”.

2. It is not very clear from the text how the 20 size-associated genes to test were selected. For instance, *yip2* appears at position 392 on the Suppl. Table 3 list when sorted by FDR. It would be good to explain in more detail how these 20 genes were selected, because by showing they were not cherry-picked would strengthen the statement that “This high rate of true positives (13/14=93%) suggests other untested, size-associated genes to be novel growth regulators.”

Response: The 20 genes were picked without considering their (if any) putative functions. We considered 2 factors: (1) the genes were highly associated with wing size traits at either RNA or protein level. The selection was biased toward genes with type 2 association. (2) Availability of RNAi lines in public. The RNAi line for *yip2* was found available in the lab and was picked as a non-type 2 gene candidate. Note that we have never performed RNAi on *yip2* before the study.

3. Page 18 - should the reference to Fig 2B actually reference Fig 6B?

Response: Corrected.

4. Fig 2F: The literature suggests **Stat92E** promotes growth.
e.g. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2914209/>
Why does the association between Stat92E mRNA and size appear to be negative (ie higher levels associated with smaller wings) in Fig 2F? The authors should discuss this.

Response: We discuss this in the “Discussion” section of the manuscript (page 20, line 6) as follows:
It appears that RNA/protein level differences between sexes seem to agree with the direction of sexual size difference (female wing > male wing) for canonical growth genes whose effect on size is evidently known: For instance, positive growth regulators such as *dm* (Myc proto-oncogene ortholog) have higher levels in females that have larger wings, but negative growth regulators such as *Gap1* (a GAP protein for Ras) and *aos* (EGF antagonist) have lower levels in females. However, as the reviewer points out, pro-growth Stat92E has lower levels in females. These facts suggest that the final sexual size difference is determined by balancing the outcomes of system-wide interactions between positive and negative regulations controlled by many growth genes and pathways.

5. Likewise, in Suppl. Fig 3, higher **Tkv** expression is associated with smaller wings, but in the literature *Tkv* is thought to promote growth?

Response: The same discussion goes as above.

6. Can the authors discuss the advantages of looking at mRNA levels (as done here) compared to the protein levels (already published), given that the two don't correlate, and protein levels are what matter at the end of the day? e.g. one can also see lowly-expressed genes, and non-coding RNAs?

Response: We discuss in the “Discussion” section of the manuscript (page 24-25) as follows:
A recent technological advancement in omics field provides us with tools by which we can investigate on molecular pathways that bridge genetic variation and phenotypic variation. Protein is expected to be the final determinants of phenotypes and therefore, protein abundance is functionally more relevant than RNA abundance. Recent large-scale omics studies including the current study have demonstrated a relatively weak correlation between abundance of proteins and RNAs, which implicates distinct genetic architectures between the two molecular layers along the information paths from genotypes to phenotypes. Based on the central dogma of molecular biology, RNA is situated more closely to genomic information than protein. It can be, therefore, expected that genomic variants may more closely associate with RNA expression. Indeed, our study has shown concentrated eQTLs at the TSS and TES of transcripts, suggesting genetic controls of RNA metabolism. In the study, by adding the transcriptomic analysis to the previous proteomics analysis, our data add information on non-coding genes and lowly-expressed genes and now provide a more comprehensive, detailed view on the genetic information paths from genotypes through molecular layers to size phenotypes.

Thank you for sending us your revised manuscript. We have now heard back from two of the three reviewers who were asked to evaluate your study. Since their recommendations are quite similar, I prefer to make a decision now rather than further delaying the process. The reviewers are now overall supportive. I am pleased to inform you that your manuscript will be accepted in principle pending the following essential amendments.

REFEREE REPORTS

Reviewer #2:

The authors have responded very well to all three reviewers. I recommend publication without further delay.

Reviewer #3:

The authors have addressed the issues raised in my original review.

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Hirokazu Okada
Journal Submitted to: Molecular Systems Biology
Manuscript Number: MSB-19-9012

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if $n \leq 5$, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
 - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
 - are tests one-sided or two-sided?
 - are there adjustments for multiple comparisons?
 - exact statistical test results, e.g., P values = x but not P values < x;
 - definition of 'center values' as median or average;
 - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	We used 16 small wing lines and 16 big wing lines for the transcriptome-wide association study (TWAS). We knew the sample size would have enough statistical power since our previous proteome-wide association study (PWAS) used similar sample sizes.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	We collected at least 30 wing imaginal discs to make RNA samples for the RNAseq experiment based on our prior knowledge.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	We performed RNAseq on 128 RNA samples and 6 samples were removed from the subsequent analyses due to their reduced counts.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	To minimize effect of environmental confounders, dissection of fly larvae was performed at different dates and cultures and the collected wing discs at each dissection was combined later to make a single RNA sample.
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Non-parametric tests that are usually conservative were performed when applicable. For every figure the statistical test was indicated.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Non-parametric tests were performed when applicable.
Is there an estimate of variation within each group of data?	Various aspects of data variability were evaluated such as variations between replicates, sex, lines and expression levels (RNA / protein), which are described in Figure 1.

USEFUL LINKS FOR COMPLETING THIS FORM

<http://www.antibodypedia.com>
<http://1degreebio.org>
<http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor>

<http://grants.nih.gov/grants/olaw/olaw.htm>
<http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm>
<http://ClinicalTrials.gov>
<http://www.consort-statement.org>
<http://www.consort-statement.org/checklists/view/32-consort/66-title>

<http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumc>

<http://datadryad.org>

<http://figshare.com>

<http://www.ncbi.nlm.nih.gov/gap>

<http://www.ebi.ac.uk/ega>

<http://biomodels.net/>

<http://biomodels.net/miriam/>
<http://jii.biochem.sun.ac.za>
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
<http://www.selectagents.gov/>

Is the variance similar between the groups that are being statistically compared?	Biological replicates exhibit a very high Spearman correlation (median=0.97) while non-replicates between distinct lines show significantly lower correlations.
---	---

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	Details on fly lines and the culturing are provided in the Materials and Methods section.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLOS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	This study is compliant with the ARRIVE guidelines.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18. Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Data availability is included in the Data Availability section in the manuscript; All the raw RNA-seq data are available at the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-7662. All the raw MS data are stored at the Center for Computational Mass Spectrometry (http://proteomics.ucsd.edu) with MassIVE ID: MSV000079202 and MSV000079208.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	Processed transcriptomes data are attached as Dataset EV2.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biocompare (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA
---	----