molecular systems biology

Quantitative Measurement of Allele-Specific Protein Expression in a Diploid Yeast Hybrid by LC-MS

Zia Khan, Joshua S. Bloom, Sasan Amini, Mona Singh, David H. Perlman, Amy A. Caudy, Leonid Kruglyak

Corresponding author: Zia Khan, University of Chicago

Review timeline:

Submission date:	13 September 2011
Editorial Decision:	28 October 2011
Revision received:	26 March 2012
Editorial Decision:	03 May 2012
Revision received:	24 May 2012
Editorial Decision:	19 June 2012
Revision received:	26 June 2012
Accepted:	01 July 2012

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 28 October 2011

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees whom we asked to evaluate your manuscript. As you will see from the reports below, the referees find the topic of your study of potential interest. However, they raise substantial concerns on your work, which, I am afraid to say, preclude its publication in its present form.

Thus, the postulated pervasive compensation of cis-differences by trans-effects should be demonstrated much more convincingly with the appropriate rigorous experimental controls that exclude unambiguously the possibility of a systematic bias. Second, allele-specific protein expression levels should be carefully compared to ASE at the mRNA steady state level to highlight allelic differences that are not revealed at the transcript level.

Please note our policy on data deposition: http://www.nature.com/msb/authors/index.html#a3.5

*** PLEASE NOTE *** As part of the EMBO Publications transparent editorial process initiative (see our Editorial at http://www.nature.com/msb/journal/v6/n1/full/msb201072.html), Molecular Systems Biology will publish online a Review Process File to accompany accepted manuscripts. When preparing your letter of response, please be aware that in the event of acceptance, your cover letter/point-by-point document will be included as part of this File, which will be available to the scientific community. More information about this initiative is available in our Instructions to

Authors. If you have any questions about this initiative, please contact the editorial office msb@embo.org.

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favourable.

Yours sincerely,

Editor
Molecular Systems Biology

http://www.nature.com/msb

Referee report:

Reviewer #1 (Remarks to the Author):

This manuscript describes a novel method for determining allele-specific protein expression and using it to infer the contributions of cis- and trans-acting factors influencing the divergence of protein levels. It is a creative approach to generating an important and unique type of data. As with any new approach, validation is critical and the validation provided by comparing different allele-specific fragments from the same protein is convincing. I find the second half of the paper showing differences in types of changes between different types of groups much less compelling, but its inclusion will be helpful to stimulate further work. Despite my overall enthusiasm for the work, I have the following questions and comments for the authors to consider:

1. The methods used for identifying fragments and determining that different fragments come from the same protein should be spelled out in the main text. What is the error rate of these methods and how might misidentification of one or more fragments impact the results?

2. I found the control for cell concentration differences in the samples containing both one species and hybrid genotypes unsatisfying. It is nearly impossible to get these exact by controlling for OD. The methods describe normalizing so that the difference in abundance between the species and the hybrid is centered on zero. How large of a correction was required? This is a reasonable starting point, but if this is not the true relationship between genotypes will affect all other genotypes. In future work (not reasonable to request this here and now), looking at something like relative genomic DNA content of the two genotypes format the same sample after mixing might allow a more accurate adjustment for this important factor.

3. The statement on pg 7 - "over half of the protein expression levels measured did not differ between the hybrid and parental species and showed the same level in the hybrid" is confusing. Do you mean the same ASE in the hybrid at the end of the sentence?

4. Can you better justify the assumption on the bottom of pg 7 for inferring mRNA half-life in S. bayanus based on S. cerevisiae?

5. I kept expecting to see a comparison of allele-specific protein expression in Scer/Sbay hybrids to allele-specific transcript abundance in these hybrids published by Bullard et al. 2010 PNAS. This seems like a glaring omission and would in my opinion be more interesting than the other comparisons between expected targets and non-targets for different proteins.

6. Is the binding site sequence for PUF3 or PUF 4 known? If so, it would be interesting to see whether these sites are mutated in S. Bayanus - and a relatively easy way to explore the hypothesis presented.

7. How many different factors were tested for whether their targets were enriched for differences in

regulation and/or expression? How often are these relationships expected by chance?

8. The discussion section is not really a discussion section. I suggest making the main section Results and Discussion and the final section something about extensions of the method (e.g., using two labels instead of one). It is also worth pointing out somewhere what fraction of all proteins you were able to make measurements on (because they had unique peptides) and discussing the potential ascertainment bias resulting from the fact that orthologous proteins must have sufficient divergence to be resolved with the assay to be tested. Prior work has shown (although not really provided a reason for) a correlation between protein sequence divergence and divergence of transcript levels.

9. Were controls done to make sure the labeling had no effect on the results? For example, splitting a single sample with a single genotype into two, labeling half and comparing to make sure the same results are observed? Or swapping which sample is labelled in the mixtures of single species and hybrid cells?

10. Legend for Figure 5 should come before that for Table 1 (should this be called a table or figure?). Grey points in some supplemental plots hard to see.

Reviewer #2 (Remarks to the Author):

Differential expressions between alleles of a gene have been exploited for several years to identify cis-regulatory variations at the RNA level. However, measuring differential allelic expression at the protein level has never been addressed so far. Allelic differences at the protein level are arguably more relevant for being closer to the physiological phenotypes. Khan et al. are the first to measure protein allelic-differential expression. Their LC-MS approach makes use of the classical cis-trans experimental design, in which differential expression between two isogenic strains are compared against allelic differential expression in their hybrid cross. A major finding seems to be global buffering of cis-differences by opposite trans-regulation, with a remarkably good correlation pattern across >600 proteins (Fig3a). However, no relevant control experiment has been done and thus the results cannot be accepted yet. Indeed, The estimators proposed are not guaranteed to be unbiased and uncorrelated. Hence, the correlation reported in Fig. 3a could be a consequence of correlation between the estimators. For the same reason, the agreement in cis-effects within complexes (fig. 4) could be explained by a genuine trans-effect and artifactual estimates of hybrid ratios. It is therefore important to perform control experiments. Moreover, it would be very interesting for the field, which has so far had access to steady-state RNA levels only, to see how much protein level allelic differences are not captured by transcript-level data. Hence, the article would gain in impact if a comparison with allele-specific expression of steady state RNA level was done, akin to the analysis by Tirosh et al. with S. paradoxus and S. cerevisiae.

Major mandatory changes:

Experimental controls must be provided. To this end, mixtures of the parental strains can be profiled. In a 1:1 mixture of parental strains for example (Ronald et al., Gen. Res. 2005) one would expect data along the diagonal when reproducing Fig. 3a. A more informative control experiment includes a whole dilution series and the check for linear behavior in the range of interest (Gagneur et al. Mol Sys Bio 2010). At least one of this control is mandatory because the estimates of the interspecies ratio and of the hybrid ratio might turn out to be correlated in control experiments. The trend in Fig. 3a would then be an artifact. Fig. 2 is testing some assumptions of the model but not explicitly the algorithm as a whole (i.e. the interspecies ratio and the hybrid ratio inference). Fig. S1c addresses the reproducibility of the method but not its correctness.

Main text: "the trans-acting component, the difference between the parental interspecies species ratio and the hybrid ratio, had a similar effect in subunits, as might be expected for a trans regulatory factor" This is not what fig. 4a displays since the interspecies difference, rather than the trans-effect, is plotted. The trans-effect should be plotted.

Moreover, the formulae should be provided in methods and written with proper notations instead of only a figure version with cryptic labels ("A in AB") and color-codes. Precise notations will makes things clearer and not more complex. For example, one can simply denote the area under the peak

for a common peptide X, the two specific ones Y and Z, and index each variable with the genotype of interest: e.g., X_{A,B}.

Discussion of allele-specific protein levels in the context of the recent pQTL report by Foss et al. PLoS Biology, 2011 should be done. This report did not find many local pQTLs. Is it in contradiction with the amount of cis-effect reported here?

Major optional changes:

The post-transcriptional regulation is an important point. The community has worked so far mostly with steady-state RNA levels. This article could be very informative if the amount of allelic differences at the protein level yet not detected at the steady-state RNA level was estimated. RNA stability is already covered by RNA-steady state, therefore the analysis of RNA half-life presented here is not very informative. It might be not out of reach for the authors to generate allele-specific RNA level data and to perform an integrated analysis by removing the RNA-level effect (similar in essence to what Foss et al. PLoS Biology, 2011. have done with eQTL and pQTL data).

I strongly advise the authors to write a statistical model. The formulae are now derived in absence of noise. However, areas under peaks are stochastic variables. The estimators proposed (for interspecies ratio and for hybrid ratio) are not guaranteed to be unbiased and uncorrelated in presence of noise. Note for example that the estimator of the hybrid ratio is explicitly a function of the estimator of the interspecies ratio. Hence, the correlation reported in Fig. 3a and fig 4 could be a consequence of correlation between the estimators. Using a noise model would give insight into which noise assumptions are reasonable fitting the data (additive, multiplicative or other, which can be checked using the replicate data), whether the estimators are unbiased and whether they are uncorrelated under these noise models. Also simulations can then be done to check the correctness of the inference algorithms. Hence a proper statistical model would give useful insights into the method.

Minor changes:

Supplementary tables should be provided as flat text file and not as pdf.
It is assumed that the coefficient of proportion for the common peptide is the same for both alleles. This assumption is not evident. Indeed, there can be polymorphisms outside the peptide that influences its visibility. A clear case would be a polymorphism on the amino acid preceding the peptide. At this location, the polymorphism affects the enzyme cutting site but is not part of the peptide itself. Such "out of the reads" sequences has been shown to induce biases in genome sequencing and there is no reason to believe that proteomics is exempt of such problems. This should be discussed.

The discussion should touch on how the method could be adapted to the human case where homozygotes cell lines are not available.

Reviewer #3 (Remarks to the Author):

This work describes an elegant method to quantitatively measure allele-specific protein expression and uses this method to describe cis and trans differences in protein expression between two yeast species. The basic question and approach are clearly interesting and important. However, I am concerned with the data presented in Fig. 3a which, in my mind, casts serious doubt on the validity of the entire analysis.

Figure 3a suggests that only a handful of genes have "pure" cis-only or trans-only variations, while the vast majority of genes that differ between species/alleles (probably ~90%) have both cis and trans variations acting in conflicting directions, such that the hybrid allele-specific variations are larger than the interspecies variations. This effect is interpreted as widespread compensatory cis/trans effects, possibly driven by stabilizing selection. Such compensatory effects have indeed been shown previously in several studies, but only in a small proportion of the genes examined (e.g. Wittkopp et al. 2004, Tirosh et al. 2009), or in a specific subset of co-regulated genes (e.g. Kuo et al. Genome Res. 2010). The results presented here argue for systematic cis-trans (partial)

compensation, which is a highly unexpected result that deserves extensive validation. Without additional evidences for such widespread compensation I suspect that this result may reflect a technical artifact and therefore that the entire dataset is inherently biased. For example, since the allele-specific ratio are calculated as the product of multiple measurements it may somehow inflate the effects relative to the interspecies ratios which are based on fewer measurements.

The only related evidence that is presented is that essential genes evolve less than non-essential genes, supporting the impact of stabilizing selection. Similar results have been previously shown (multiple times) in analysis of divergence in mRNA levels, although a widespread cis-trans compensation was not reported for mRNA levels. It is therefore not surprising to see the essentiality effect also for protein levels and this does not make the widespread compensation hypothesis any less unexpected.

Additional comments:

1. The fact that ASE is not measured directly but instead calculated as a product of many different measurements suggests that it might suffer from high noise, including both technical and biological noise. This should be acknowledged and discussed.
2. Fig. 2 seems to demonstrate that many measurements are highly reproducible (on the diagonal), but those that are not on the diagonal are often very far from it (outliers), perhaps implying that this analysis should be used to first determine which genes can at all be analyzed for ASE.
3. It would be informative to compare these measurements of protein ASE with previous measurements of mRNA ASE in the same hybrid (e.g. Bullard et al., PNAS 2010).
4. The enriched (and directed) ASE of PUF3,4 targets may not be a consequence of the direct regulation by these factors, but instead could arise from other mechanisms. This is especially likely as these target-sets are associated with particular annotations and at least one of these enriched annotations (mitochondrial genes) has already been shown to diverge among yeast species through other mechanisms (e.g. Ihmels et al. Science 2005). Is the higher S. bayanus expression of mitochondrial genes specific to PUF3 targets or is it also observed for non-PUF3 targets with similar functional annotations? (and similarly for PUF4). Furthermore, as consensus RNA binding motifs have been proposed for both PUF3 and PUF4 the authors can examine the conservation of these motifs and test their prediction that cis-divergence of these target genes reflects changes in PUF3,4 binding sites.
5. The analysis of essential vs. non-essential genes could be confounded by absolute expression levels: as essential genes are, on average, expressed more than non-essential genes their protein levels may be determined at higher precision, thus resulting in less differences between species/alleles.
6. The analysis of mRNA decay is based on very speculative assumptions and although it suggests SOME effect of differential mRNA decay on protein levels this analysis is difficult to interpret.
7. "Our result suggests that uORF containing genes may have an increased sensitivity to trans-acting regulatory divergence". It is not clear what this sentence means exactly. I would argue for a more specific prediction: that PUB1 activity has diverged among the species. Is there any evidence for that? (e.g. differential interspecies mRNA or protein levels of PUB1 or associated factors).
8. It is not clear how to interpret the LOW trans-divergence of BFR1 targets.
9. The authors note that cis-divergence of PUB1 and BFR1 targets was not significantly different than the rest of the genes, but the quoted p-values are in fact significant as this analysis does not involve multiple hypotheses (i.e. a p-value of 0.006 is presented as evidence for lack of a significant effect of BFR1; however, this analysis is not aimed at finding a few significant effects among a large set of potential hypotheses, but instead is focused on one specific hypothesis).
10. The basic approach for measuring ASE may deserve a longer description as it currently might be difficult for readers to understand.
11. The placement of figure labels below (instead of above) the subplots is confusing in some instances(e.g. Fig. 1).

1st Revision - authors' response 26 March 2012

We thank the referees for the insightful and thoughtful comments and criticisms. Below, we have attempted to address all of their concerns, point-by-point, and have made substantial changes to the manuscript to include additional control experiments and reflect additional and revised analyses. We start with the remarks of Reviewer 3.

Reviewer 3 (Remarks to the Author):

This work describes an elegant method to quantitatively measure allele-specific protein expression and uses this method to describe cis and trans differences in protein expression between two yeast species. The basic question and approach are clearly interesting and important.

However, I am concerned with the data presented in Fig. 3a which, in my mind, casts serious doubt on the validity of the entire analysis. Figure 3a suggests that only a handful of genes have "pure" cis-only or trans-only variations, while the vast majority of genes that differ between species/alleles (probably ~90%) have both cis and trans variations acting in conflicting directions, such that the hybrid allele-specific variations are larger than the interspecies variations. This effect is interpreted as widespread compensatory cis/trans effects, possibly driven by stabilizing selection. Such compensatory effects have indeed been shown previously in several studies, but only in a small proportion of the genes examined (e.g. Wittkopp et al. 2004, Tirosh et al. 2009), or in a specific subset of co-regulated genes (e.g. Kuo et al. Genome Res. 2010). The results presented here argue for systematic cis-trans (partial) compensation, which is a highly unexpected result that deserves extensive validation. Without additional evidences for such widespread compensation I suspect that this result may reflect a technical artifact and therefore that the entire dataset is inherently biased. For example, since the allele-specific ratio are calculated as the product of multiple measurements it may somehow inflate the effects relative to the interspecies ratios which are based on fewer measurements. The only related evidence that is presented is that essential genes evolve less than non-essential genes, supporting the impact of stabilizing selection. Similar results have been previously shown (multiple times) in analysis of divergence in mRNA levels, although a widespread cis-trans compensation was not reported for mRNA levels. It is therefore not surprising to see the essentiality effect also for protein levels and this does not make the widespread compensation hypothesis any less unexpected.

We would like to thank the reviewer for the insightful observation. On further investigation, we found that we reversed a sign during the calculation of the $\log_2$ protein allele specific expression ratios (ASE) in the script that generated this plot. Our results, now shown in Figure 4a, have been revised with the correct sign during the calculation. Consistent with previous mRNA ASE measurements (e.g. Wittkop et al 2004, Tirosh et al 2009, and Kuo et al Genome Res 2010), compensatory effects are observed in a small portion of proteins.

To rule out any possibility of any technical artifact or systematic bias, we conducted a rigorous control experiment in which we created a "synthetic hybrid", a protein sample that consisted of a 1:1 mixture of the parental strains, and compared the resulting protein ASE measurements to "ground-truth" measurements obtained by comparing the strains directly. To generate the "ground-truth" measurements, we independently cultured *S. cerevisiae* using ^{15}N heavy labeled minimal medium and *S. bayanus* in unlabeled minimal medium. We used an unlabeled sample where *S. cerevisiae* was grown in unlabeled (L) minimal media with glucose as a carbon source and *S. bayanus* was grown in rich media with an acetate carbon source (YPA) as internal standards to deliberately disrupt any correlation interspecies ratios and our synthetic hybrid protein ASE ratios. As illustrated in Figure 3d, our design disrupted the correlation between interspecies ratios and the computed protein ASE ratios. More importantly, high correlations in Figure 3e (and Figure 3f) between of our calculated "synthetic hybrid" protein ASE ratios to directly measured "ground-truth" interspecies

ratios confirmed the accuracy of our method. We have expanded and revised the text to include a detailed description of the rationale of this control experiment and the corresponding results in the section "Accuracy of Protein Allele-Specific Measurements." We have added an additional figure (Figure 3) that illustrates the approach and reports the corresponding results.

Reviewer 3 (Additional Comments):

1. The fact that ASE is not measured directly but instead calculated as a product of many different measurements suggests that it might suffer from high noise, including both technical and biological noise. This should be acknowledged and discussed.

We have acknowledged and discussed this issue in the section titled, "Accuracy of Protein Allele-Specific Measurements."

2. Fig. 2 seems to demonstrate that many measurements are highly reproducible (on the diagonal), but those that are not on the diagonal are often very far from it (outliers), perhaps implying that this analysis should be used to first determine which genes can at all be analyzed for ASE.

We would like to thank the reviewer for this suggestion. In the revised manuscript, we now use the results expressed in Fig. 2 to create a "high-confidence" set of proteins with protein ASE measurements. We use two main criteria: (1) We restricted this high-confidence set to proteins where at least two distinct shared peptides, two distinct variant peptides from *S. cerevisiae*, and two distinct variant peptides from *S. bayanus* were identified and quantified. (2) We eliminated proteins where the absolute difference between $\log_2$ ratios for distinct variant and distinct shared peptides deviated by more than 0.85, to eliminate far outliers.

3. It would be informative to compare these measurements of protein ASE with previous measurements of mRNA ASE in the same hybrid (e.g. Bullard et al., PNAS 2010).

We have additionally added a correlation of our protein ASE measurements to mRNA ASE in the same hybrid using the data from Bullard et al, PNAS 2010. We observe a Spearman's 0.373 and a Pearson's correlation of 0.331. When we consider each technical replicate separately, the Spearman's correlation of technical replicate 1 is 0.324 and the Pearson's correlation is 0.306. The Spearman's correlation of technical replicate 2 is 0.407 and the Pearson's correlation is 0.354.

4. The enriched (and directed) ASE of PUF3,4 targets may not be a consequence of the direct regulation by these factors, but instead could arise from other mechanisms. This is especially likely as these target-sets are associated with particular annotations and at least one of these enriched annotations (mitochondrial genes) has already been shown to diverge among yeast species through other mechanisms (e.g. Ihmels et al. Science 2005). Is the higher S. bayanus expression of mitochondrial genes specific to PUF3 targets or is it also observed for non-PUF3 targets with similar functional annotations? (and similarly for PUF4). Furthermore, as consensus RNA binding motifs have been proposed for both PUF3 and PUF4 the authors can examine the conservation of these motifs and test their prediction that cis-divergence of these target genes reflects changes in PUF3,4 binding sites.

The same sign reversal affected the calculation of the $\log_2$ protein ASE ratios for our analysis of *cis*/*trans* divergence also occurred in our script for the analysis of PUF3,4 targets. When we corrected the sign reversal, we found a significant signal indicating *trans*-divergance of *S. cereivisaie* PUF3,4 targets. This result indicates that analysis of divergent activity of PUF3,4 between the species would be more productive. As the current data available, collected only in *S. cerevisiae*, prevents us from moving further with this analysis, we have moved this result to a supplemental results and discussion section. We hope that this result will spur further work in to possible divergence in these RNA binding proteins between species. We have also qualified this analysis to indicate that it might be confounded by divergence in other mechanisms.

5. The analysis of essential vs. non-essential genes could be confounded by absolute expression levels: as essential genes are, on average, expressed more than non-essential genes their

protein levels may be determined at higher precision, thus resulting in less differences between species/alleles.

We used the absolute protein expression measurements from the following manuscript in which absolute expression levels were estimated using GFP fluorescence of tagged proteins

Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS (2006) Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441:** 840-846

As shown in Fig. 3d, we find that the absolute expression level of essential proteins is no different from that of non-essential proteins. This rules out the possibility that the higher precision of divergence estimates of highly expressed proteins explains our observation.

6. The analysis of mRNA decay is based on very speculative assumptions and although it suggests SOME effect of differential mRNA decay on protein levels this analysis is difficult to interpret.

We see the reviewer's point. To address this concern, we have moved this section to the supplement and presented all of these results as supplemental figures. We have also added a sentence at the end of this new supplemental results and discussion section that states that without the corresponding measurements in *S. bayanus* our results are difficult to interpret. Our intent with the analysis was only to help stimulate future work.

7. "Our result suggests that uORF containing genes may have an increased sensitivity to trans-acting regulatory divergence". It is not clear what this sentence means exactly. I would argue for a more specific prediction: that PUB1 activity has diverged among the species. Is there any evidence for that? (e.g. differential interspecies mRNA or protein levels of PUB1 or associated factors).

As described in the new supplemental results and discussion section, we additionally scanned our raw protein expression measurements for shared peptides and our direct measurements found that the median interspecies protein expression level was 31.2% higher in *S. cerevisiae* than in *S. bayanus* for PUB1. This result provides additional evidence that PUB1 activity has diverged.

8. It is not clear how to interpret the LOW trans-divergence of BFR1 targets. The that cis-divergence of PUB1 and BFR1 targets was not significantly different than the rest of the genes, but the quoted p-values are in fact significant as this analysis does not involve multiple hypotheses (i.e. a p-value of 0.006 is presented as evidence for lack of a significant effect of BFR1; however, this analysis is not aimed at finding a few significant effects among a large set of potential hypotheses, but instead is focused on one specific hypothesis).

We have revised this analysis (which is now in the supplement) to use a fixed p-value cutoff of 0.01.

9. The basic approach for measuring ASE may deserve a longer description, as it currently might be difficult for readers to understand.

We have expanded this description. We hope the revised version is easier for readers to understand.

10. The placement of figure labels below (instead of above) the subplots is confusing in some instances (e.g. Fig. 1).

We have modified the figure labels in Figure 1 and elsewhere so that they are positioned above the subplots.

Reviewer 2 (Remarks to the Authors):

Differential expressions between alleles of a gene have been exploited for several years to identify cis-regulatory variations at the RNA level. However, measuring differential allelic expression at the protein level has never been addressed so far. Allelic differences at the protein level are arguably more relevant for being closer to the physiological phenotypes. Khan et al. are the first to measure protein allelic-differential expression. Their LC-MS approach makes use of the classical cis-trans experimental design, in which differential expression between two isogenic strains are compared against allelic differential expression in their hybrid cross.

A major finding seems to be global buffering of cis-differences by opposite trans-regulation, with a remarkably good correlation pattern across >600 proteins (Fig3a). However, no relevant control experiment has been done and thus the results cannot be accepted yet. Indeed, The estimators proposed are not guaranteed to be unbiased and uncorrelated. Hence, the correlation reported in Fig. 3a could be a consequence of correlation between the estimators. For the same reason, the agreement in cis-effects within complexes (fig. 4) could be explained by a genuine trans-effect and artifactual estimates of hybrid ratios. It is therefore important to perform control experiments.

Please see initial comment above to reviewer 3.

Moreover, it would be very interesting for the field, which has so far had access to steady-state RNA levels only, to see how much protein level allelic differences are not captured by transcript-level data. Hence, the article would gain in impact if a comparison with allele-specific expression of steady state RNA level was done, akin to the analysis by Tirosh et al. with S. paradoxus and S. cerevisiae.

We have additionally provided correlations to steady-state mRNA ASE expression measurements from the following study:

Bullard JH, Mostovoy Y, Dudoit S, Brem RB (2010) Polygenic and directional regulatory evolution across pathways in Saccharomyces. *Proceedings of the National Academy of Sciences* **107:** 5058-5063

Reviewer 2 (Major mandatory changes):

Experimental controls must be provided. To this end, mixtures of the parental strains can be profiled. In a 1:1 mixture of parental strains for example (Ronald et al., Gen. Res. 2005) one would expect data along the diagonal when reproducing Fig. 3a. A more informative control experiment includes a whole dilution series and the check for linear behavior in the range of interest (Gagneur et al. Mol Sys Bio 2010). At least one of this control is mandatory because the estimates of the interspecies ratio and of the hybrid ratio might turn out to be correlated in control experiments. The trend in Fig. 3a would then be an artifact. Fig. 2 is testing some assumptions of the model but not explicitly the algorithm as a whole (i.e. the interspecies ratio and the hybrid ratio inference). Fig. S1c addresses the reproducibility of the method but not its correctness.

Please see initial comment above to reviewer 3.

Main text: "the trans-acting component, the difference between the parental interspecies species ratio and the hybrid ratio, had a similar effect in subunits, as might be expected for a trans regulatory factor" This is not what fig. 4a displays since the interspecies difference, rather than the trans-effect, is plotted. The trans-effect should be plotted.

We have plotted the *trans*-effect in this figure (now Figure 5a).

Moreover, the formulae should be provided in methods and written with proper notations instead of only a figure version with cryptic labels ("A in AB") and color-codes. Precise notations will makes things clearer and not more complex. For example, one can simply

denote the area under the peak for a common peptide X, the two specific ones Y and Z, and index each variable with the genotype of interest: e.g., X_{A,B}.

We have updated Figure 1b and the corresponding caption to include more precise mathematical notations. We retained the color codes in illustrative histograms under the mathematical notation to help convey which peptides are used to derive the ratios. We hope that this approach is much clearer than our previous version.

Discussion of allele-specific protein levels in the context of the recent pQTL report by Foss et al. PLoS Biology, 2011 should be done. This report did not find many local pQTLs. Is it in contradiction with the amount of cis-effect reported here?

We have revised the manuscript to discuss this difference. Our result is consistent with previous studies of mRNA expression divergence in both yeast and in fly, which have also observed a larger contribution of *cis*-regulatory differences between species than within species:

(1) Wittkopp PJ, Haerum BK, Clark AG (2008) Genetic basis of regulatory variation within and between Drosophila species. *Nature Genetics* **40:** 346-350
(2) Emerson JJ, Hsieh L-C, Sung H-M, Wang T-Y, Huang C-J, Lu HH-S, Lu M-YJ, Wu S-H, Li W-H (2010) Natural selection on cis and trans regulation in yeasts. *Genome Research* **20:** 826-836

Reviewer 2 (Major optional changes):

The post-transcriptional regulation is an important point. The community has worked so far mostly with steady-state RNA levels. This article could be very informative if the amount of allelic differences at the protein level yet not detected at the steady-state RNA level was estimated. RNA stability is already covered by RNA-steady state, therefore the analysis of RNA half-life presented here is not very informative. It might be not out of reach for the authors to generate allele-specific RNA level data and to perform an integrated analysis by removing the RNA-level effect (similar in essence to what Foss et al. PLoS Biology, 2011. have done with eQTL and pQTL data).

To address this concern, we have correlated our results to mRNA ASE measurements obtained in the following study of the same interspecies hybrid we used in our study.

Bullard JH, Mostovoy Y, Dudoit S, Brem RB (2010) Polygenic and directional regulatory evolution across pathways in Saccharomyces. *Proceedings of the National Academy of Sciences* **107:** 5058-5063

Our results are now presented as a main figure, Figure 6. We have also moved the results on mRNA decay to a supplemental results and discussion section.

In addition, while we agree that the community would benefit greatly from an integrated analysis, a complete, integrated analysis that also considers hybrid RNA ASE, interspecies mRNA expression, and allele-specific and interspecies mRNA decay on the same strains and source samples remains out of the scope of the current manuscript. Furthermore, our coverage of the proteome is only a small fraction of what can be obtained by RNA-seq, and as we describe in the last section, we believe that this initial study will spur efforts to develop new mass spectrometry techniques that to extend coverage of our protein ASE measurements to possibly all of the yeast proteome.

Reviewer 2 (Minor changes):

Supplementary tables should be provided as flat text file and not as pdf.

We originally provided the tables as Excel spreadsheets, but the online submission system converted these to .PDF format. We apologize for the inconvenience. We have also provided the data in tab

delimited tables and the all of the analysis scripts used to generate each figure in this manuscript on the web-link included in the manuscript http://compbio.cs.princeton.edu/pview/proteinASE.

It is assumed that the coefficient of proportion for the common peptide is the same for both alleles. This assumption is not evident. Indeed, there can be polymorphisms outside the peptide that influences its visibility. A clear case would be a polymorphism on the amino acid preceding the peptide. At this location, the polymorphism affects the enzyme cutting site but is not part of the peptide itself. Such "out of the reads" sequences has been shown to induce biases in genome sequencing and there is no reason to believe that proteomics is exempt of such problems. This should be discussed.

We did not mean to imply that the proportion of shared peptides would be the same for both alleles. In fact, shared peptides may originate from two locations in one allele's sequence and a single location in the other allele's sequence. In the materials and methods we describe how these peptides can be filtered out from subsequent processing. Errors within the predicted proteome sequence of each species are a more pernicious problem, resulting in shared peptides being labeled variant and vice versa. We have discussed these peptides in the revised manuscript and have taken additional filtering steps, as suggested by reviewer 3, to mitigate their effect on our protein ASE measurements.

The discussion should touch on how the method could be adapted to the human case where homozygotes cell lines are not available.

We propose one simple approach in the last section: use the results of exome sequencing or similar re-sequencing efforts to identify the necessary internal standards on a per protein basis.

Reviewer 1 (Remarks to the Authors):

This manuscript describes a novel method for determining allele-specific protein expression and using it to infer the contributions of cis- and trans-acting factors influencing the divergence of protein levels. It is a creative approach to generating an important and unique type of data. As with any new approach, validation is critical and the validation provided by comparing different allele-specific fragments from the same protein is convincing. I find the second half of the paper showing differences in types of changes between different types of groups much less compelling, but its inclusion will be helpful to stimulate further work. Despite my overall enthusiasm for the work, I have the following questions and comments for the authors to consider:

1. The methods used for identifying fragments and determining that different fragments come from the same protein should be spelled out in the main text. What is the error rate of these methods and how might misidentification of one or more fragments impact the results?

We have expanded our discussion of how the peptide identifications and quantifications are obtained and how variant and shared peptides are identified at the end of the section titled "Measuring Protein Allele-Specific Expression by LC-MS." In the following section, where we discuss the results pertaining to the accuracy of the peptide ratios, we have provided the error rate (set through an FDR cutoff) and describe how misidentification impacts results. We also now apply further filtering steps as suggested by reviewer 3 to further reduce the impact of misidentifications.

2. I found the control for cell concentration differences in the samples containing both one species and hybrid genotypes unsatisfying. It is nearly impossible to get these exact by controlling for OD. The methods describe normalizing so that the difference in abundance between the species and the hybrid is centered on zero. How large of a correction was required? This is a reasonable starting point, but if this is not the true relationship between genotypes will affect all other genotypes. In future work (not reasonable to request this here and now), looking at something like relative genomic DNA content of the two genotypes format the same sample after mixing might allow a more accurate adjustment for this important factor.

We now report the adjustments required. They range from 19.8% to 25.7%. We have also suggested an alternative approach: Lyse the cells separately and combine the lysates based on separate total protein quantifications. Preview LC-MS runs on unfractionated protein samples can then be used to find the correction factor to reduce any remaining error further. However, this approach requires additional instrument time, which, at present, remains a limited resource and accounting for the additional variability introduced, if any, of lysing and extracting protein from the cells separately.

3. The statement on pg 7 - "over half of the protein expression levels measured did not differ between the hybrid and parental species and showed the same level in the hybrid" is confusing. Do you mean the same ASE in the hybrid at the end of the sentence?

We were referring to protein expression differences between the parental species and the hybrid, which can be used to evaluate the dominance of expression levels in the F_1 hybrid, not ASE measurements. We have modified this sentence to read, "Over half of the relative expression differences measured between the parental species and the hybrid did not differ." We apologize for any confusion.

4. Can you better justify the assumption on the bottom of pg 7 for inferring mRNA half-life in S. bayanus based on S. cerevisiae?

Due to reviewer concerns, in the revised manuscript, we have moved our results relating to mRNA half-life to a supplemental results and discussion section. The corresponding figure has also been moved to the supplement. We did, however, conduct an additional analysis using mRNA data comparing *S. cerevisiae* to *S. paradoxus* from a recent study Tirosh et al 2009 *Science*. We see correlations between steady stead mRNA expression and mRNA decay determined in *S. cerevisiae* that indicate the correlations to decay we see are not limited to our protein expression measurements.

5. I kept expecting to see a comparison of allele-specific protein expression in Scer/Sbay hybrids to allele-specific transcript abundance in these hybrids published by Bullard et al. 2010 PNAS. This seems like a glaring omission and would in my opinion be more interesting than the other comparisons between expected targets and non-targets for different proteins.

We now include a correlation of our protein ASE measurements to the mRNA ASE measurements published by Bullard et al 2010, PNAS as a main figure (Fig. 6). The comparisons between the targets and non-targets of RNA binding proteins and the corresponding figure have been moved to the supplement.

6. Is the binding site sequence for PUF3 or PUF 4 known? If so, it would be interesting to see whether these sites are mutated in S. bayanus - and a relatively easy way to explore the hypothesis presented.

Please see comment 4 to reviewer 3 above.

7. How many different factors were tested for whether their targets were enriched for differences in regulation and/or expression? How often are these relationships expected by chance?

In the supplemental discussion we provide the number of factors we tested, 46 RNA binding proteins from Hogan et al 2008, PloS Biol. To control for multiple hypothesis testing, we originally applied a Bonferonni correction to our significance cutoff in our previous manuscript. However, reviewer 3 argued we were testing separate hypotheses and a fixed p-value threshold should be used. We now use a fixed p-value cutoff of 0.01 to find these associations between targets and non-targets of RNA binding proteins.

8. The discussion section is not really a discussion section. I suggest making the main section Results and Discussion and the final section something about extensions of the method (e.g., using two labels instead of one).

We have revised the manuscript, relabeling the "Results" section as the "Results and Discussion" section. The last section has been renamed "Future Directions and Challenges."

8b. It is also worth pointing out somewhere what fraction of all proteins you were able to make measurements on (because they had unique peptides) and discussing the potential ascertainment bias resulting from the fact that orthologous proteins must have sufficient divergence to be resolved with the assay to be tested. Prior work has shown (although not really provided a reason for) a correlation between protein sequence divergence and divergence of transcript levels.

Table S1 contains a summary of the peptides detected as well as the number of proteins they allow for quantifications to be derived. We now discuss this ascertainment bias resulting from varying degrees of sequence divergence of individual proteins. We also propose two technologies that might mitigate these challenges in the last section of the manuscript.

9. Were controls done to make sure the labeling had no effect on the results? For example, splitting a single sample with a single genotype into two, labeling half and comparing to make sure the same results are observed? Or swapping which sample is labeled in the mixtures of single species and hybrid cells?

In a recently published manuscript that details our protein quantification methods, we describe the results of a control experiment in which we demonstrate that our methods are resilient to the labeling status of the sample:

Khan Z, Amini S, Bloom J, Ruse C, Caudy A, Kruglyak L, Singh M, Perlman D, Tavazoie S (2011) Accurate proteome-wide protein quantification from high-resolution ^{15}N mass spectra. *Genome Biology* **12:** R122
http://genomebiology.com/2011/12/12/R122/abstract

We have revised the manuscript to include this reference.

In addition, our control experiment with the "synthetic hybrid" uses independent measurements of interspecies expression ratios (our "ground-truth" measurements) where the labeling status of the *S. cerevisiae* sample differs. In both cases, we observe a high correlation between our "mock" protein ASE ratios and the "ground-truth" interspecies expression ratios.

10. Legend for Figure 5 should come before that for Table 1 (should this be called a table or figure?). Grey points in some supplemental plots hard to see.

We have increased the size of the points in the revised figures and repositioned the legends.

2nd Editorial Decision 03 May 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who accepted to evaluate the study. As you will see, the referees acknowledge that the revised manuscript has been improved. While the conclusions have changed, the reviewers are now more supportive. They raise however a series of remaining important concerns with regard to the conclusiveness and completion of the data analysis. These concerns should be convincingly addressed in a second and exceptional round of revision.

Reviewer #2 requests that the statistical significance of the three main claims of the study should be rigorously tested. Reviewer #3 makes furthermore three very constructive suggestions that could bring more insights into protein ASE.
Thank you for submitting this paper to Molecular Systems Biology.

Yours sincerely,

Editor
Molecular Systems Biology

Referee reports:

Reviewer #1 (Remarks to the Author):

The authors have done an outstanding job addressing the reviewer concerns in this revised manuscript. Most commendable are (1) the reanalysis of a high-confidence gene set, (2) the "synthetic" hybrid control, (3) the comparison to Bullard et al., and (4) the restructuring of the manuscript that included moving more speculative points to the supplementary materials. Modifications to the text have also made the description much more clear.

I have only one minor comment:
In the comparison to Bullard et al. on pg 13, it is pointed out that the RNA and protein data are derived from cells grown under different conditions. This is an important point and can influence the amount of post-transcriptional regulation that is inferred. Also, do you mean posttranscriptional when you say posttranslational regulatory divergence in this section? This section should be modified slightly to more explicitly allow for the effect of environmental factors in reducing the correlation coefficient between the two studies.

Reviewer #2 (Remarks to the Author):

The authors have answered to the major concern referees 3 and myself had by identifying a sign error in their implementation. They have moreover added a control experiment that demonstrates that the inferred allelic expression ratios tend to be overall consistent. The main findings are now different.

The major issue now is that none of the three main claims (1: accuracy of the method, 2: contribution of cis and trans effects to protein expression divergence and 3: significant contribution of "post-translational" cis-acting factors - see synopsis) is statistically assessed. This must be done. These findings are expected but, if proven, are still of interest to the field.

To this end, a statistical test to detect significant allelic differential expression across biological replicates would be instrumental:

Claim 1 (fig 3f). The sensitivity of the method could be demonstrated by reporting how many of the proteins differentially expressed between the parents are found significantly differentially expressed in the synthetic hybrid. Moreover, the control experiment should be done for more proteins to be informative. Indeed, few (47) proteins have been measured. Among these, the rare proteins with expression ratios greater than 2-fold (log2 >1) are also the noisiest (away from diagonal) casting doubt about the reliability of the classification of cis/trans proteins in Fig 4a (based on a log2 > 1 cutoff).

Claim 2 (fig 4a). The classification of proteins into cis, trans or both appears to be based on a fold change cutoff, which is sensitive to noise. A statistical test would provide trustable classifications.

Claim 3 (fig 6). Again, statistical testing would allow identifying significant cases with significantly larger protein expression differences than RNA expression differences.

Open minor issues regarding the responses to my former questions are listed below. I found another sign error that could have been easily noticed by the authors. In general, I would recommend the authors more thoroughly inspect and rigorously analyze their data before sending this manuscript to the journal.

Figure 1 has improved. Nevertheless, a formal method section describing unambiguously how the ratios are computed is needed.

Fig 5a) is wrong. There is a sign error in the computation of trans-effect (see the provided script complexes.R at line 7) explaining why the interspecies difference was surprisingly opposite to the trans effect. Also, the interpretation is misleading. The cis-effects appear to NOT be concordant across proteins of the same complex (yellow points). The text should be corrected.

Fig 5b is based on two proteins per complexes. Displaying boxplots for two data points is very misleading. Moreover, all the effects are small (below 2 fold). Hence, no strong conclusion of this case can be drawn. Fig 5b should thus be removed.

Reviewer #3 (Remarks to the Author):

In the revised version the authors corrected an analysis error that removed the peculiar pattern that made me question the data in the first draft. After this correction and additional filtering and validations steps, I trust that the data is probably of high (or at least adequate) quality, and therefore that the manuscript presents an elegant, largely accurate, and potentially very useful method for quantifying protein ASE.

However, due to the nature of the approach and to several important filtering steps the scale of the approach is quite limited. Protein ASE is computed for 400 proteins, and of these only 33 proteins have ASE of at least 2-fold. Only 84 proteins have ASE of at least 1.5-fold. Unfortunately, despite the extensive analysis that the authors have conducted, the limited scale seems to restrict the depth by which the authors can address the relevant questions, as I detail below, and therefore I believe that the true benefit of this approach will come only when the scale is improved.

First, one of the main questions that the authors attempt to address is the frequency of cis and trans changes. As seen in Fig. 4 only 20-30 protein are classified as cis and a similar number as trans, and these classifications are clearly very threshold dependent: some proteins classified as cis and trans have similar values, and very few proteins (10-20) are close enough to the expected behavior of cis and trans that we can reliably call them cis only or trans only. The results therefore provide some individual examples of cis and trans changes in protein expression and raise the prediction that there might be more or less a similar contribution of the two types. But this data appears insufficient to estimate the global contribution of cis and trans changes and, more importantly in my mind, to delineate the properties of these different types of changes.

Another expected benefit from having protein ASE is to address the role and mechanisms of post-transcriptional changes. Protein ASE has been compared to mRNA ASE, demonstrating a partial correlation that suggests the presence of post-transcriptional changes. However, beyond this expected and very initial observation there are no specific follow up analysis and insights.
1. S.cer-S.bay interspecies mRNA expression differences have recently been published in another work (Busby MA et al. BMC genomics 2011) and this independent data can also be used to estimate post-transcriptional changes; perhaps the protein ASE that differ from both Busby et al. and Bullard et al. can be examined further as a subset of the most reliable post-transcriptional changes.
2. Can the authors link the predicted cis-dependent post-transcriptional changes to any specific mechanism? The supplementary information includes some analysis of RNA-binding proteins in the context of protein ASE, but are these enhanced or reduced when considering specifically the estimated post-transcriptional changes? Are the estimated post-transcriptional changes correlated with changes in codon usage?
3. Another question that comes up from comparison of protein and mRNA divergence is whether there is a global trend of less divergence in protein than in mRNA levels, as observed previously (Laurent et al. Protein abundances are more conserved than mRNA abundances across diverse taxa), and if so then what are the associated mechanisms. Is there some hint of such effect here?

The authors present additional analysis that are largely expected or limited in scope and depth. These include decreased expression divergence of essential genes, similar divergence of complex subunits, apparent cis-trans compensation of proteosome subunits and hybrid misexpression. These results indeed extend the description of the method, but I still get the feeling that the main contribution of this manuscript is the method itself and that analysis of the resulting dataset only

scratches the surface of the relevant questions.

2nd Revision - authors' response 24 May 2012

We thank the referees again for the insightful and thorough comments and criticisms. Below, we address all of their concerns, point-by-point.

Reviewer #1 (Remarks to the Author)

The authors have done an outstanding job addressing the reviewer concerns in this revised manuscript. Most commendable are (1) the reanalysis of a high-confidence gene set, (2) the "synthetic" hybrid control, (3) the comparison to Bullard et al., and (4) the restructuring of the manuscript that included moving more speculative points to the supplementary materials. Modifications to the text have also made the description much more clear.

I have only one minor comment:

In the comparison to Bullard et al. on pg 13, it is pointed out that the RNA and protein data are derived from cells grown under different conditions. This is an important point and can influence the amount of post-transcriptional regulation that is inferred. Also, do you mean posttranscriptional when you say posttranslational regulatory divergence in this section? This section should be modified slightly to more explicitly allow for the effect of environmental factors in reducing the correlation coefficient between the two studies.

We have modified the text in this section to more explicitly state that environmental factors may contribute to reducing the correlation coefficient between the two studies. In regard the second point, as regulatory divergence may affect the efficiency of translation itself, resulting in differences in protein ASE measurements that are not present in mRNA ASE measurements, we have modified the section slightly to use posttranscriptional throughout.

Reviewer #2 (Remarks to the Author)

The authors have answered to the major concern referees 3 and myself had by identifying a sign error in their implementation. They have moreover added a control experiment that demonstrates that the inferred allelic expression ratios tend to be overall consistent. The main findings are now different.

The major issue now is that none of the three main claims (1: accuracy of the method, 2: contribution of cis and trans effects to protein expression divergence and 3: significant contribution of "post-translational" cis-acting factors - see synopsis) is statistically assessed. This must be done. These findings are expected but, if proven, are still of interest to the field.

To this end, a statistical test to detect significant allelic differential expression across biological replicates would be instrumental:

Claim 1 (fig 3f). The sensitivity of the method could be demonstrated by reporting how many of the proteins differentially expressed between the parents are found significantly differentially expressed in the synthetic hybrid. Moreover, the control experiment should be done for more proteins to be informative. Indeed, few (47) proteins have been measured. Among these, the rare proteins with expression ratios greater than 2-fold (log2 >1) are also the noisiest (away from diagonal) casting doubt about the reliability of the classification of cis/trans proteins in Fig 4a (based on a log2 > 1 cutoff).

To statistically assess our first claim, we used a permutation test to ask whether the agreement between our "mock" protein ASE measurements from the synthetic hybrid and the directly measured "ground truth" ratios in Figure 3e and Figure 3f occurred by chance. We used the sum of the absolute differences, the L_1 norm, between our $\log_2$ "mock" protein ASE ratios and the directly measured "ground truth" measurements as a test statistic. We chose this test statistic because it provided a more stringent measure of the agreement between the two measurements than both Pearson and Spearman's correlation, which only capture linear and rank-based dependence between variables respectively. We generated a null distribution over this statistic by permuting the ratios used to compute our protein ASE ratios. A total of 100,000 permutations allowed us to assign a $p < 10^{-5}$ to the observed L_1 norm. This additional statistical analysis confirmed that it is highly unlikely the agreement we observed in our control experiment was due to chance. We detail this additional analysis at the end of the section titled "Accuracy of Protein Allele-Specific Expression Measurements."

Claim 2 (fig 4a). The classification of proteins into cis, trans or both appears to be based on a fold change cutoff, which is sensitive to noise. A statistical test would provide trustable classifications.

To control for misclassification of proteins into the categories *cis* and *trans*, we used the $\log_2$ "mock" protein ASE ratios and the $\log_2$ directly measured "ground truth" ratios from our control experiment. Because the measurements captured noise when agreeing ratios were measured using similar methods as the proteins in Figure 4a, we used them to estimate of the misclassification rate of *cis*-labeled proteins (where both the interspecies ratio and the protein ASE ratio should agree). To avoid selection of a cutoff that was sensitive to noise in these data, we used the average of a distribution derived by bootstrap over the following statistic: the cutoff at which the misclassification rate, estimated from the control experiment data, was limited to 5%. Using 10^5 bootstrap iterations, we derived a $\log_2$ cutoff of 1.0286, which we applied to generate a slightly revised Figure 4a. We describe his statistical approach in the section titled "Protein Expression Divergence."

Claim 3 (fig 6). Again, statistical testing would allow identifying significant cases with significantly larger protein expression differences than RNA expression differences.

To add statistical support to our third claim, we asked whether the correlations between our protein ASE measures and the previous mRNA ASE measurements from Bullard et al 2010 PNAS were statistically significant. To this end, we permuted these two measurements to generate a null distribution over the Pearson's correlation and, separately, the Spearman's correlation. The results of this permutation test confirmed that correlations we observed were highly significant ($p < 10^{-5}$).

To identify proteins with a significant posttranscriptional ASE component, we focused on the difference between our protein ASE and the mRNA ASE measurements from Bullard et al 2010 PNAS

$\log_2$(posttranscriptional ASE) = $\log_2$(protein ASE) - $\log_2$(mRNA ASE)

We used bootstrap to estimate a threshold such that the less than 5% of the absolute differences between replicate measurements of $\log_2$(posttranscriptional ASE) (across our two replicates) were mislabeled as differing. We used the average of this distribution to a select a cutoff that is less sensitive to noise to identify proteins with a significant posttranscriptional *cis* effects in a revision of Figure 6.

Open minor issues regarding the responses to my former questions are listed below. I found another sign error that could have been easily noticed by the authors. In general, I would recommend the authors more thoroughly inspect and rigorously analyze their data before sending this manuscript to the journal.

Figure 1 has improved. Nevertheless, a formal method section describing unambiguously how the ratios are computed is needed.

We have added a formal methods section titled "Computation of Protein ASE Ratios" that describes unambiguously how we computed the protein ASE ratios in our study.

Fig 5a) is wrong. There is a sign error in the computation of trans-effect (see the provided script complexes.R at line 7) explaining why the interspecies difference was surprisingly opposite to the trans effect. Also, the interpretation is misleading. The cis-effects appear to NOT be concordant across proteins of the same complex (yellow points). The text should be corrected.

We have addressed the sign error in the complexes.R and revised Figure 5a, now a single Figure 5 panel. We addressed the concern regarding the concordance of *cis*-effects, *trans* effects, and interspecies differences by statistical analysis. We quantified the degree of coordination using the standard deviation of the $\log_2$ expression ratios. To determine if the expression divergence measurements of the complexes were significantly coordinated, we permuted the expression ratios to compute a null distribution over the standard deviation of a complex with the same number of measured subunits. In agreement with the reviewer's observation, our statistical analysis revealed that *cis*-effects are not significantly coordinated. In contrast, *trans* effects were significantly coordinated for 9 out of 12 of complexes analyzed.

Fig 5b is based on two proteins per complexes. Displaying boxplots for two data points is very misleading. Moreover, all the effects are small (below 2 fold). Hence, no strong conclusion of this case can be drawn. Fig 5b should thus be removed.

The reviewer raises a valid point. We have removed this figure and the corresponding text.

Reviewer #3 (Remarks to the Author):

In the revised version the authors corrected an analysis error that removed the peculiar pattern that made me question the data in the first draft. After this correction and additional filtering and validations steps, I trust that the data is probably of high (or at least adequate) quality, and therefore that the manuscript presents an elegant, largely accurate, and potentially very useful method for quantifying protein ASE.

However, due to the nature of the approach and to several important filtering steps the scale of the approach is quite limited. Protein ASE is computed for 400 proteins, and of these only 33 proteins have ASE of at least 2-fold. Only 84 proteins have ASE of at least 1.5-fold. Unfortunately, despite the extensive analysis that the authors have conducted, the limited scale seems to restrict the depth by which the authors can address the relevant questions, as I detail below, and therefore I believe that the true benefit of this approach will come only when the scale is improved.

We agree with the reviewer, and to this end, we have proposed several approaches in the "Future Directions and Challenges" that promise to improve the scale and depth of the method. Our method is only limited by the sensitivity and throughput of quantitative mass spectrometry, which continues advance rapidly. For instance, two recent studies in Molecular Systems Biology demonstrated that approximately 10,000 proteins could be quantified in a human cell line:
http://www.nature.com/msb/journal/v7/n1/full/msb201181.html
http://www.nature.com/msb/journal/v7/n1/full/msb201182.html

First, one of the main questions that the authors attempt to address is the frequency of cis and trans changes. As seen in Fig. 4 only 20-30 protein are classified as cis and a similar number as trans, and these classifications are clearly very threshold dependent: some proteins classified as cis and trans have similar values, and very few proteins (10-20) are close enough to the expected behavior of cis and trans that we can reliably call them cis only or trans only. The results therefore provide some individual examples of cis and trans changes in protein expression and raise the prediction that there might be more or less a similar contribution of

the two types. But this data appears insufficient to estimate the global contribution of cis and trans changes and, more importantly in my mind, to delineate the properties of these different types of changes.

The reviewer raises a valid concern. We believe that subsequent studies will address in greater depth these very questions. We expect our initial observations will follow the same progression as studies of mRNA-levels. Initial studies of the contribution of cis and trans changes to mRNA-levels measured only 34 allele specific and interspecies mRNA ratios (Witkopp et al 2004, Nature). Only in recent years, using both arrays and more recently RNA-seq, have these measurements have been extended genome wide (e.g. Bullard et al 2010 PNAS; McManus et a 2010 Genome Research; Tirosh et al 2009 Science, etc.), allowing questions regarding the global contribution and properties of these changes and their effects on mRNA-levels to be explored in greater depth.

Another expected benefit from having protein ASE is to address the role and mechanisms of post-transcriptional changes. Protein ASE has been compared to mRNA ASE, demonstrating a partial correlation that suggests the presence of post-transcriptional changes. However, beyond this expected and very initial observation there are no specific follow up analysis and insights.

1. S.cer-S.bay interspecies mRNA expression differences have recently been published in another work (Busby MA et al. BMC genomics 2011) and this independent data can also be used to estimate post-transcriptional changes; perhaps the protein ASE that differ from both Busby et al. and Bullard et al. can be examined further as a subset of the most reliable post-transcriptional changes.

Please see response to comment 3 below.

2. Can the authors link the predicted cis-dependent post-transcriptional changes to any specific mechanism? The supplementary information includes some analysis of RNA-binding proteins in the context of protein ASE, but are these enhanced or reduced when considering specifically the estimated post-transcriptional changes? Are the estimated post-transcriptional changes correlated with changes in codon usage?

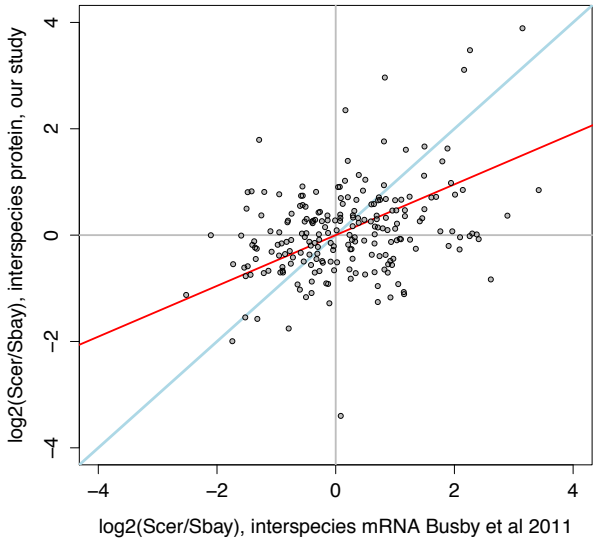
Yes we can link our data to *cis*-regulation by uAUGs, start codons in the 5'-untranslated region (5'UTR) of a transcript. Recent studies have suggested that uAUGs act as potent and widespread regulators of mRNA and protein expression levels (Calvo et al 2009 PNAS and Yun et al 2012 Genome Research). We found, as illustrated in a new Fig. 6b and described in the section "Comparative Analysis of mRNA and Protein ASE Measurements," that uAUG containing transcripts may have greater sensitivity to *cis*-acting regulatory divergence.

We scanned through targets of RNA binding proteins for significant associations between a magnitude and directional estimate of the post-transcriptional component of *cis*-acting regulatory variation, estimated as $\log_2$(posttranscriptional ASE) = $\log_2$(protein ASE) - $\log_2$(mRNA ASE), the off diagonal component of a scatter plot of mRNA ASE and protein ASE measurements. We found that magnitude of this quantity was significantly less for targets of BFR1 and Scp160 than non-targets of these proteins. The presence of the additional association with Scp160 may reflect the distinct functional roles of these proteins. We now include this additional result in the supplemental results and discussion.

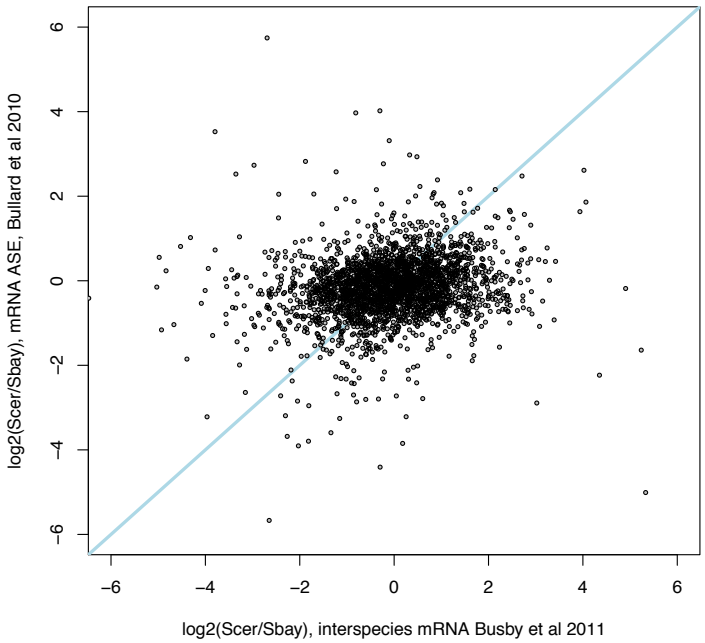
We conducted an additional analysis using differing estimates of translation efficiency, based adaptation of codon usage the tRNA pool from Man and Pilpel 2007 Nature Genetics. This additional analysis is now included in the supplemental results and discussion section.

3. Another question that comes up from comparison of protein and mRNA divergence is whether there is a global trend of less divergence in protein than in mRNA levels, as observed previously (Laurent et al. Protein abundances are more conserved than mRNA abundances across diverse taxa), and if so then what are the associated mechanisms. Is there some hint of such effect here?

The reviewer provides an excellent suggestion for additional analysis. We started by using the data from Busby MA et al BMC Genomics 2012 to address this question.



We were initially pleased to see that our data reproduced the global trend observed in Laurent et al. Protein abundance differences seem to be significantly less than the mRNA differences between the two species. Prior to measuring post-transcriptional effects using this data set, we decided to examine the *trans*-effects on mRNA levels due to the differing growth conditions between Bullard et al 2010 PNAS and Busby et al 2011 BMC Genomics by plotting the mRNA ASE measurements in Bullard et al with the interspecies mRNA measurements in Busby et al 2011 BMC Genomics. We were disturbed to see the plot below for 2,615 transcripts, generated after filtering transcripts where the $\log_2$-fold difference between the biological replicates in Busby et al 2011 BMC Genomics was greater than 1.5. A nearly identical plot (not shown) was obtained using the uniquely mapping read counts, after filtering on the sum of the total number of reads mapped to each ortholog, or using each biological replicate independently.



The plot suggests an unlikely pervasive *trans* effect on mRNA levels due to the differing conditions. We believe this result could potentially reflect a systematic bias in the 454 sequencing technology used in the Busby et al 2011 BMC Genomics study. For this reason, we did not include any results derived from the data from Busby et al 2011 BMC Genomics in the revised manuscript.

The authors present additional analysis that are largely expected or limited in scope and depth. These include decreased expression divergence of essential genes, similar divergence of complex subunits, apparent cis-trans compensation of proteosome subunits and hybrid misexpression. These results indeed extend the description of the method, but I still get the feeling that the main contribution of this manuscript is the method itself and that analysis of the resulting dataset only scratches the surface of the relevant questions.

Indeed, we anticipate that subsequent studies that use the data provided here or leverage more advanced mass spectrometry techniques to replicate this work with higher proteome coverage will delve deeper into the relevant questions. By providing a novel method that allows access to measurements not readily obtainable in any prior study and several examples of insights that can be gained from this new type of data, we hope that our manuscript sets the basis for a broad range of future work.

3rd Editorial Decision 19 June 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from referee #1. I am afraid that this reviewers still has valid points of criticisms with regard to the study:

- 'claim 1': the concerns refer to the presentation and the need for clarification. This seem thus addressable.

- 'claim 2': while a 5% significance threshold is used the % of proteins that display cis or trans-effects are not provided. Given that the analysis is performed on 399 proteins and looking at figure 4a it seems that this concern should be addressable as well. If this conclusions cannot be supported, I am afraid that it would considerably detract from the overall conclusiveness of the study.

- 'claim 3': this issue seems more difficult to resolve. If there is a way to take into account the uncertainty at the RNA level and the uncertainty due to differences in experimental conditions, the analysis should be performed appropriately. Otherwise, this claim should be removed from the synopsis and toned down and clarified. In particular, on page 13, "the modest correlation...suggets that extensive degree of cis-acting..." should be reworded to emphasize that "the observation that the correlation is modest suggests that posttranscrtional cis-acting divergence may exist..." On p14 "we estimate that 98 out of 358 proteins reflect posttranscriptional effect" should be toned down appropriately.

- With regard to the exploratory analyses including the analysis of uAUG, codon usage, RNA-binding proteins and mRNA stability, should all be described in supplementary information and only briefly mentioned in the main Results section. The respective supplementary figures should be explicitly cited in the main Results section. The outcome of these analyses should also be summarized with a short statement in the main Result section, otherwise the purpose and outcome of these analysis, even if negative, are not understandable.

Please not that we won't be able to offer any additional round of revision beyond this one.

Please resubmit your revised manuscript online, with a covering letter listing amendments and responses to each point raised by the referees. Please resubmit the paper **within one month** and ideally as soon as possible. If we do not receive the revised manuscript within this time period, the

file might be closed and any subsequent resubmission would be treated as a new manuscript. Please use the Manuscript Number (above) in all correspondence.

Thank you for submitting this paper to Molecular Systems Biology.

Yours sincerely,

Editor
Molecular Systems Biology

Referee reports:

Reviewer #2 (Remarks to the Author):

The authors have established a fold-change cutoff, one common to all proteins, inferred from the distribution of allelic ratios errors observed in a negative control experiment. There is room for improvement for this test for allelic differential expression because all proteins do not show the same variance across biological replicates. Nevertheless this simple approach can suffice to support in part the three claims. Claim 1 (quantitative assessment of the accuracy of the method) is now established. Claim 2 (significant contribution of cis and trans) and claim 3 (significant "post-transcriptional" effects) remain to be proven as explained below. I believe the authors can address them. If only claim 1 can be supported, the contribution of this study will be a method of interest to the field.

Reviewer 2: Claim 1 (fig 3f). The sensitivity of the method could be demonstrated by reporting how many of the proteins differentially expressed between the parents are found significantly differentially expressed in the synthetic hybrid. Moreover, the control experiment should be done for more proteins to be informative. Indeed, few (47) proteins have been measured. Among these, the rare proteins with expression ratios greater than 2-fold (log2 >1) are also the noisiest (away from diagonal) casting doubt about the reliability of the classification of cis/trans proteins in Fig 4a (based on a log2 > 1 cutoff).
Authors: To statistically assess our first claim, we used a permutation test to ask whether the agreement between our "mock" protein ASE measurements from the synthetic hybrid and the directly measured "ground truth" ratios in Figure 3e and Figure 3f occurred by chance. We used the sum of the absolute differences, the L1norm, between our log2 "mock" protein ASE ratios and the directly measured "ground truth" measurements as a test statistic. We chose this test statistic because it provided a more stringent measure of the agreement between the two measurements than both Pearson and Spearman's correlation, which only capture linear and rank-based dependence between variables respectively. We generated a null distribution over this statistic by permuting the ratios used to compute our protein ASE ratios. A total of 100,000 permutations allowed us to assign a $p < 10\text{-}5$ to the observed L1 norm. This additional statistical analysis confirmed that it is highly unlikely the agreement we observed in our control experiment was due to chance. We detail this additional analysis at the end of the section titled "Accuracy of Protein Allele- Specific Expression Measurements."

R:
The statistical assessment shows that the method is overall consistent: The difference between true and estimated ratios over all 47 proteins are smaller than "by chance". The contrary would have been worrisome. What is actually needed and more informative is an estimate of what true fold change the method is able to detect. The authors have better addressed this issue by answering to my 2nd point: The relative error is about 2-fold.

Moreover, the resampling procedure is not enough detailed (no text in methods!). In particular it is not clear across what entities (across the proteins?) the resampling of the ratios is performed.

Thus, I suggest removing this resampling-based test. Instead, plotting the distribution of the allelic ratio errors (estimated - "ground truth") will be more informative. The distribution of these errors are the basis for the fold-change cut-off (see point 2).

A quantitative claim of the accuracy of the method can then be that "allelic ratios are estimated at {plus minus}2-fold precision in ~ 95% of the cases".

R2: Claim 2 (fig 4a). The classification of proteins into cis, trans or both appears to be based on a fold change cutoff, which is sensitive to noise. A statistical test would provide trustable classifications.

A: To control for misclassification of proteins into the categories cis and trans, we used the log2 "mock" protein ASE ratios and the log2 directly measured "ground truth" ratios from our control experiment. Because the measurements captured noise when agreeing ratios were measured using similar methods as the proteins in Figure 4a, we used them to estimate of the misclassification rate of cis-labeled proteins (where both the interspecies ratio and the protein ASE ratio should agree). To avoid selection of a cutoff that was sensitive to noise in these data, we used the average of a distribution derived by bootstrap over the following statistic: the cutoff at which the misclassification rate, estimated from the control experiment data, was limited to 5%. Using 105 bootstrap iterations, we derived a log2 cutoff of 1.0286, which we applied to generate a slightly revised Figure 4a. We describe his statistical approach in the section titled "Protein Expression Divergence."

R2: The control experiments provide an empirical null distribution that is then used to set a cut-off at 5% type I error. Bootstrapping here is likely an overhead. Is the estimation of the cut-off when done on the primary original data fundamentally different? If kept, the bootstrap procedure should be described in methods with details about the resampling procedure (what is resampled, is that done with replacement?).

The claim (synopsis) that there is a significant amount of proteins with trans effects holds if more than 5% of the proteins are above the 5% type I error cut-off (multiple testing issue). This percentage should be given.

The same applies for the number of proteins with cis-effects.

R: Claim 3 (fig 6). Again, statistical testing would allow identifying significant cases with significantly larger protein expression differences than RNA expression differences.

A:To add statistical support to our third claim, we asked whether the correlations between our protein ASE measures and the previous mRNA ASE measurements from Bullard et al 2010 PNAS were statistically significant. To this end, we permuted these two measurements to generate a null distribution over the Pearson's correlation and, separately, the Spearman's correlation. The results of this permutation test confirmed that correlations we observed were highly significant ($p < 10\text{-}5$).

R: Why using bootstrapping when statistical tests are available for both Pearson and Spearman's correlation? Simply have the p-value of these classical tests after each reported correlation.

From this analysis it can be claimed that allele-specific variations at the mRNA level are significantly propagated at the protein level. However, it cannot be claimed that there is significant amount of "post-transcriptional" allelic variations. This claim should be removed (synopsis).

A: To identify proteins with a significant posttranscriptional ASE component, we focused on the difference between our protein ASE and the mRNA ASE measurements from Bullard et al 2010 PNAS log2(posttranscriptional ASE) = log2(protein ASE) - log2(mRNA ASE). We used bootstrap to estimate a threshold such that the less than 5% of the absolute differences between replicate measurements of log2(posttranscriptional ASE) (across our two replicates) were

mislabeled as differing. We used the average of this distribution to a select a cutoff that is less sensitive to noise to identify proteins with a significant posttranscriptional cis effects in a revision of Figure 6.

R: This analysis controls for variations at the proteomic level. However, it does not control for uncertainty at the RNA level. In the text the authors correctly recognize this drawback together with the issue that the RNA and the protein levels are not measured in the same growth conditions. Thus the amount of proteins off the diagonal suggests but does not demonstrate that there is significant post-transcriptional allelic variations.

Finally, the newly included uAUG analysis does not demonstrate significant post-transcriptional variations neither. Indeed, significant allelic expression variation is seen at the mRNA level, and at the total protein level. The latter however is correlated with mRNA level and thus does not add information. To make this claim, significant differences should be seen for the "post-transcriptional" statistic (i.e the difference). Without this analysis, the uAUG section does not add relevant information and should be removed.

3rd Revision - authors' response 26 June 2012

Reviewer #2 (Remarks to the Author): The authors have established a fold-change cutoff, one common to all proteins, inferred from the distribution of allelic ratios errors observed in a negative control experiment. There is room for improvement for this test for allelic differential expression because all proteins do not show the same variance across biological replicates. Nevertheless this simple approach can suffice to support in part the three claims. Claim 1 (quantitative assessment of the accuracy of the method) is now established. Claim 2 (significant contribution of cis and trans) and claim 3 (significant "post-transcriptional" effects) remain to be proven as explained below. I believe the authors can address them. If only claim 1 can be supported, the contribution of this study will be a method of interest to the field.

Reviewer 2: Claim 1 (fig 3f). The sensitivity of the method could be demonstrated by reporting how many of the proteins differentially expressed between the parents are found significantly differentially expressed in the synthetic hybrid. Moreover, the control experiment should be done for more proteins to be informative. Indeed, few (47) proteins have been measured. Among these, the rare proteins with expression ratios greater than 2-fold (log2 >1) are also the noisiest (away from diagonal) casting doubt about the reliability of the classification of cis/trans proteins in Fig 4a (based on a log2 > 1 cutoff).

Authors: To statistically assess our first claim, we used a permutation test to ask whether the agreement between our "mock" protein ASE measurements from the synthetic hybrid and the directly measured "ground truth" ratios in Figure 3e and Figure 3f occurred by chance. We used the sum of the absolute differences, the L1norm, between our log2 "mock" protein ASE ratios and the directly measured "ground truth" measurements as a test statistic. We chose this test statistic because it provided a more stringent measure of the agreement between the two measurements than both Pearson and Spearman's correlation, which only capture linear and rank-based dependence between variables respectively. We generated a null distribution over this statistic by permuting the ratios used to compute our protein ASE ratios. A total of 100,000 permutations allowed us to assign a $p < 10\text{-}5$ to the observed L1 norm. This additional statistical analysis confirmed that it is highly unlikely the agreement we observed in our control experiment was due to chance. We detail this additional analysis at the end of the section titled "Accuracy of Protein Allele- Specific Expression Measurements."

R: The statistical assessment shows that the method is overall consistent: The difference between true and estimated ratios over all 47 proteins are smaller than "by chance". The contrary would have been worrisome. What is actually needed and more informative is an estimate of what true fold

change the method is able to detect. The authors have better addressed this issue by answering to my 2nd point: The relative error is about 2-fold. Moreover, the resampling procedure is not enough detailed (no text in methods!). In particular it is not clear across what entities (across the proteins?) the resampling of the ratios is performed.

We have expanded the methods text on page 21 to include the precise entities used in the bootstrapping procedure. The bootstrapping procedure is also included in the R scripts (protein_control.R), which accompany the raw data on the supplementary web site.

Thus, I suggest removing this resampling-based test. Instead, plotting the distribution of the allelic ratio errors (estimated - "ground truth") will be more informative. The distribution of these errors are the basis for the fold-change cut-off (see point 2). A quantitative claim of the accuracy of the method can then be that "allelic ratios are estimated at {plus minus}2-fold precision in ~ 95% of the cases".

We used the allelic ratio errors as a null distribution, as suggested. We sorted the absolute log2 differences |estimated – "ground truth"| and selected the difference at the 97.5th percentile as the cutoff. We used the 97.5th percentile as the cutoff is for a two-tailed distribution. We derived a slightly less conservative log2 cutoff primary data of 0.9560. The result is not appreciably different from our bootstrap derived cutoff. We feel that the bootstrap adds additional stability to our estimation procedure and provides a slightly more conservative cutoff. For this reason, we have decided to leave it as is and have provided details in the Materials and Methods on page 21. As the reviewer points out, the analysis relates to the precision of the protein ASE measurements. We have moved this analysis to the end of the section now titled "Accuracy and Precision of Protein Allele-Specific Expression Measurements" on page 10 of the revised manuscript.

R2: Claim 2 (fig 4a). The classification of proteins into cis, trans or both appears to be based on a fold change cutoff, which is sensitive to noise. A statistical test would provide trustable classifications.

A: To control for misclassification of proteins into the categories cis and trans, we used the log2 "mock" protein ASE ratios and the log2 directly measured "ground truth" ratios from our control experiment. Because the measurements captured noise when agreeing ratios were measured using similar methods as the proteins in Figure 4a, we used them to estimate of the misclassification rate of cis-labeled proteins (where both the interspecies ratio and the protein ASE ratio should agree). To avoid selection of a cutoff that was sensitive to noise in these data, we used the average of a distribution derived by bootstrap over the following statistic: the cutoff at which the misclassification rate, estimated from the control experiment data, was limited to 5%. Using 105 bootstrap iterations, we derived a log2 cutoff of 1.0286, which we applied to generate a slightly revised Figure 4a. We describe his statistical approach in the section titled "Protein Expression Divergence."

R2: The control experiments provide an empirical null distribution that is then used to set a cut-off at 5% type I error. Bootstrapping here is likely an overhead. Is the estimation of the cut-off when done on the primary original data fundamentally different?

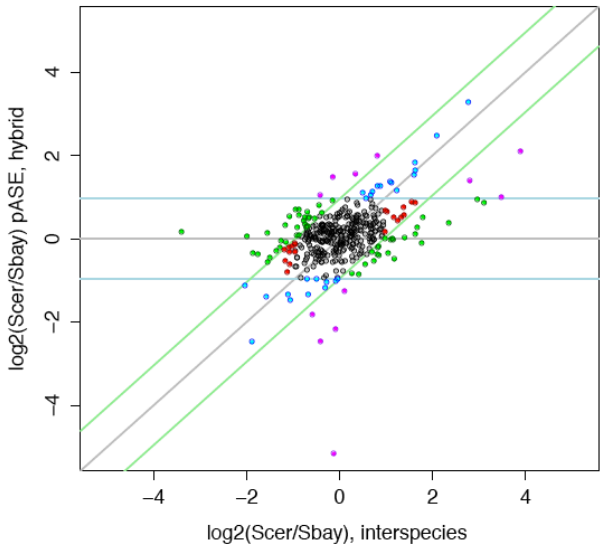
As mentioned in our response above, the estimation cutoff using the primary data is 0.9560, which is not fundamentally different from our bootstrap-derived cutoff.

If kept, the bootstrap procedure should be described in methods with details about the resampling procedure (what is resampled, is that done with replacement?).

We originally performed sampling with replacement and have expanded the methods text. We also provided R scripts that assure that the analyses can be repeated without any ambiguity.

The claim (synopsis) that there is a significant amount of proteins with trans effects holds if more than 5% of the proteins are above the 5% type I error cut-off (multiple testing issue). This percentage should be given. The same applies for the number of proteins with cis-effects.

The claim in the synopsis holds when we use the more conservative bootstrap derived cutoff. We find that 8.0% (21+11)/398 proteins showed cis-effects (blue + purple points in Fig 4a.) and 13.8% (44+11)/398 proteins showed trans-effects (green + purple points in Fig 4a). Note 11 proteins showed both cis and trans effects (purple points in Fig. 4a). In both cases the number of proteins is above 5%.
The numbers increase for primary data cutoff as well 9.5% (26+12) / 398 for proteins with cis-effects and 16.8% (55+12)/398 for proteins with trans effects. Note for this cutoff 12 proteins show both cis effects and trans effects. Figure 4a for this cutoff is provided below. Compare to figure 4a in the manuscript.



We have highlighted this result in the section text on page 11. As mentioned above, we have described the cis/trans classification procedure and cutoff selection in the methods on page 21. We have also provided an R script with the classification procedure cistrans.R on the supplementary web site to assure every detail is reproducible without ambiguity.

R: Claim 3 (fig 6). Again, statistical testing would allow identifying significant cases with significantly larger protein expression differences than RNA expression differences.

A: *To add statistical support to our third claim, we asked whether the correlations between our protein ASE measures and the previous mRNA ASE measurements from Bullard et al 2010 PNAS were statistically significant. To this end, we permuted these two measurements to generate a null distribution over the Pearson's correlation and, separately, the Spearman's correlation. The results of this permutation test confirmed that correlations we observed were highly significant ($p < 10\text{-}5$).*

R: Why using bootstrapping when statistical tests are available for both Pearson and Spearman's correlation? Simply have the p-value of these classical tests after each reported correlation.

We now report the p-values of these classical tests instead of performing permutations.

From this analysis it can be claimed that allele-specific variations at the mRNA level are significantly propagated at the protein level. However, it cannot be claimed that there is significant amount of "post-transcriptional" allelic variations. This claim should be removed (synopsis).

We have removed this claim from the synopsis.

A: To identify proteins with a significant posttranscriptional ASE component, we focused on the difference between our protein ASE and the mRNA ASE measurements from Bullard et al 2010

PNAS log2(posttranscriptional ASE) = log2(protein ASE) - log2(mRNA ASE). We used bootstrap to estimate a threshold such that the less than 5% of the absolute differences between replicate measurements of log2(posttranscriptional ASE) (across our two replicates) were mislabeled as differing. We used the average of this distribution to a select a cutoff that is less sensitive to noise to identify proteins with a significant posttranscriptional cis effects in a revision of Figure 6.

R: This analysis controls for variations at the proteomic level. However, it does not control for uncertainty at the RNA level. In the text the authors correctly recognize this drawback together with the issue that the RNA and the protein levels are not measured in the same growth conditions. Thus the amount of proteins off the diagonal suggests but does not demonstrate that there is significant post-transcriptional allelic variations.

We have qualified our claims in the manner suggested by the editor.

Finally, the newly included uAUG analysis does not demonstrate significant post-transcriptional variations neither. Indeed, significant allelic expression variation is seen at the mRNA level, and at the total protein level. The latter however is correlated with mRNA level and thus does not add information. To make this claim, significant differences should be seen for the "posttranscriptional" statistic (i.e the difference). Without this analysis, the uAUG section does not add relevant information and should be removed.

As per request of the editor, we have moved the additional analysis of the uAUG containing transcripts to the supplementary results and discussion section.

