molecular systems biology

Inferring transcriptional and microRNA-mediated regulatory programs in glioblastoma

Manu Setty, Karim Helmy, Aly A. Khan, Joachim Silber, Aaron Arvey, Frank Neezen, Phaedra Agius, Jason T. Huse, Eric C. Holland, and Christina S. Leslie

Corresponding author: Christina Leslie, Sloan Kettering Institute

Review timeline:

Submission date:	24 October 2011
Editorial Decision:	06 December 2011
Revision received:	12 April 2012
Editorial Decision:	18 May 2012
Revision received:	12 June 2012
Editorial Decision:	12 July 2012
Revision received:	17 July 2012
Accepted:	25 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 06 December 2011

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

The editor would like to highlight two apparently fundamental issues that were raised by the reviewers. These points will need to be conclusively addressed, possibly with additional experiments and analyses, before this work would be appropriate for publication in Molecular Systems Biology.

1. Regarding the ability of this method to identify novel factors driving the GBM subtype differences, the first reviewer found the validation experiments with miR-132 and miR-124 less than conclusive, and was not convinced that this method was genuinely superior at identifying drivers compared to previous works (e.g. Kim et al. 2011, Genome Res). This reviewer felt that it would be important to directly demonstrate phenotypic changes in response to miR-132 or miR-124.

2. The reviewers felt that additional work was needed to demonstrate the value of the integrated model over simpler classifiers. This is best expressed in reviewer #2's point #1, but seems closely related the reviewer #1's point #2.

In addition, the reviewers had a series of more specific concerns that will also need to be addressed

© European Molecular Biology Organization

with additional analyses or clarification.

PLEASE NOTE As part of the EMBO Publications transparent editorial process initiative (see http://www.nature.com/msb/journal/v6/n1/full/msb201072.html), Molecular Systems Biology now publishes online a Review Process File with each accepted manuscript. 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. Authors may opt out of the transparent process at any stage prior to publication (contact us at msb@embo.org). More information about this initiative is available in our Instructions to Authors.

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 favorable.

Yours sincerely,
Editor - Molecular Systems Biology
msb@embo.org

Reviewer #1 (Remarks to the Author):

The authors describe an elaborate machine learning scheme to reverse engineer integrated regulatory networks in GBM TCGA samples, including predicted protein-DNA and miRNA-mRNA interactions. They include CNV, DNase, promoter methylation data to improve their ability to predict variations in gene expression across tumors. To integrate these data, the authors use a linear regression scheme with individual weights for copy number, and each considered miRNA and TF. Predicted TF interactions were filtered using epigenetic data, and only differentially expressed miRNAs were considered. The authors make the following claims:
(1) Their models predict normal vs. tumor expression changes (r^2 < 0.18; p<2-16); CNV data helped (p<3e-16), DNase data helped (p<??), and methylation data helped (p<??).
(2) A combination of miRNAs and genes could be used to recover previously identified GBM classes (Verhaak et al.), and especially classify proneural and mesenchymal GBM subtypes.
(3) The authors identified YY1, GABP, miR-124 and miR-132 as drivers of proneural GBM.
(4) They find similarities between their driver coding genes and known drivers of proneural progenitors. Thus, predicted gene drivers of the proneural subtype have literature evidence.
(5) Up regulation of miR-124 and miR-132 in proneural neurospheres lead to changes in expression of their direct and indirect target genes. These genes were enriched for proneural vs. mesenchymal differentially expressed genes. Suggesting that these miRs are drivers of proneural GBM and that their overexpression reverses expression changes in proneural tumors. The authors use miR-380 and miR-448, which are under expressed in proneural tumors (vs. normal?) as controls.
(6) The authors suggest that miR-132, NFYB and YY1 may have common proneural-specific targets, and may regulate chromatin-related functions.

While I appreciate the complexity of the task undertaken here and the effort by the authors, I don't believe that this work is ready for publication at MSB. Here, the authors appear to have two distinct goals and neither one of them is reached.

(1)The authors would like to use all available tools to identify drivers of GMB subtypes. Their conclusion is that YY1, GABP, miR-124 and miR-132 may drive its proneural subtype. However, this is not supported by their validation experiments. The authors selected miR-132 and mir-124 (also selected by Tae-Min Kim, Wei Huang, Richard Park, et al., Cancer Res 2011;71:3387-3399) as candidate drivers. But these are not expressed in proneural tumors and are selected because of their targets across tumors. The authors discovery that their targets are differentially expressed is at best validation that the miR target prediction method they used is better than random. To show that miR-132 and mir-124 really drive the subtype, the authors should demonstrate phenotypic changes. Moreover, to further increase confidence in their driver prediction, the authors should use other miRNAs reported by Kim et al. as classifiers of the proneural subtype. Moreover, it is now commonly believed that the proneural subtype described by the authors is composed of at least 2

distinct types, as described by Noushmehr et al. and cited in this manuscript. One of these types is strongly associated and may be caused by promoter hyper-methylation which results from a mutation in IDH1. The authors should treat the two subtypes separately.

(2)The authors would like to establish the improved strength of an integrated approach to tumorigenesis driver discovery, as was done before by other groups. To do this successfully, an integrated module should be discovered and demonstrated. The authors start on this path but come short of producing convincing evidence. Such evidence should demonstrate that without miRNAs or genes, our ability to explain subtype initiation is impaired.
Moreover, it should be pointed out that GBM tumors are very different from normal brain tissue and are easy to classify. CNV and methylation data are especially predictive, as has been previously shown by several researchers and cited in this manuscript. Moreover, GBM classes and especially proneural and mesenchymal GBM subtypes have been previously identified and classified by genes and miRs and supported by CNVs (Verhaak et al.; Carro et al.; Kim et al.)

In conclusion, I believe that the authors are doing important work, but I don't believe that this work is ready for publication at MSB.

Reviewer #2 (Remarks to the Author):

This manuscript from Setty et al. utilizes TCGA data on glioblastoma and outlines a computational framework to model transcription factor and microRNA regulated gene expression. Using this model, the authors show that the model parameters are capable to differentiating separate classes of glioblastoma samples, and important regulators can be predicted by computationally removing one regulator at a time from the model. Experimental overexpression of two of the miRNAs, miR-132 and miR-124, induces changes in messenger RNA expression consistent with the model.
Overall, this beautifully outlined manuscript is one of the attempts to address the challenge on computational analysis of multiple datasets produced by large genomic efforts such as TCGA and ENCODE. It will be of interest not only to computational biologists, but to the broader community of cancer biologists and miRNA researchers. In this regard, although the manuscript is highly readable by bench scientists with sophisticated sense of computational techniques, I suggest the authors to use plain language to briefly explain seemingly standard computational techniques whenever possible.
The following can further enhance this manuscript.
1. It is interesting that the parameters for the model can be used to cluster and predict tumor subtypes. This is a piece of the evidence that the model itself is valid. However, it is not clear (1) whether the model is providing more information than the direct expression levels of the miRNAs and TFs, and (2) whether the two subtypes are so distinct so that by random chance this classification can be achieved. Please provide analysis (1) to cluster the samples by the expression levels of the indicated TFs and miRNAs, and (2) to cluster the samples with model parameters by randomly picking the same number of TFs and miRNAs as shown in Fig 2B, from TFs and miRs that pass the selection criteria (i.e. differentially expressed etc.)
2. It will be nice to somehow show a distribution of the model parameters, for modeling on the single sample level. Particularly it will be interesting to know the variation of coefficients for the same miRNA on the same gene. Maybe a figure for a representative miRNA:target pair and a TF:target pair, plus a table to summarize more data.
3. For Fig3c, it will be nice to also show the survival difference separated by the expression of miR-132, in addition to using the model coefficients of miR-132.
4. The authors claim that miR-380, miR-448 and miR-443 did not have effects on differential gene expression, unlike miR-124 and miR-132. But there is no visual data associated with this claim. Please show data, in similar format as in Fig 5a and Fig 5b.
5. It is not clear to me from Supp Fig 1b how the authors concluded that "methylation consistently entering the model as a negative regulator of gene expression", although biologically it does make sense. Please explain in fig legend or in text if possible.
Minior:
1. In a number of figures, including supplement, the fonts are too small on printed version of the manuscript.
2. Supplemental Fig 1, please describe in figure legend what the colors mean and what the dots mean (I assumed the color means class label for the tumor subtypes and dots are samples).
3. On page 2 of text, trascriptomic subtypes "are may" arise ... should be fixed.

4. In text describing Figure 4c, MYB should be MYBL2? MYB is not in the table.
5. I suggest the authors to label their main figures with figure number, text with page number, and supplemental tables (the file in pdf format) with table numbers and page numbers, unless explicitly instructed not to by the journal.

Reviewer #3 (Remarks to the Author):

In the manuscript entitled "Inferring transcriptional and microRNA-mediated regulatory programs in glioblastoma", the authors proposed an integrative algorithm combining mRNA, copy number and promoter methylation profiles, along with regulatory sequence information of TF and miRNA to decipher gene expression changes in GBM. The model was trained either sample-by-sample or jointly with subtype assignments to allow the authors to identify both common and subtype-specific regulators. They subsequently focused on regulators specific for the proneural subtype. They confirmed that predicted gene expression signatures for proneural subtype regulators were consistent with in vivo expression changes in a PDGF-driven mouse model. They also tested two predicted proneural drivers, miR-124 and miR-132, which were underexpressed in proneural GBMs, by overexpression in neurospheres. This is an interesting approach to identify novel TF/miRNA for targeted therapy and for understanding mechanisms important for GBM biology given the massive amount of publicly available data. The approach led to some novel insights into the role of two miRNAs (i.e. miR-124 and miR-132) in GBM.

Major points:
1. The authors built their regression model simply based on the counts of TF and miRNA binding sites in the gene's regulatory regions, unfortunately neglecting the fact that the miRNA expression data is available for more than 400 GBM cases in TCGA database. The computational prediction of miRNA binding targets is known to have high false positive rate. Although the traditional pair-wise anti-correlation between miRNAs and predicted targets also introduces false discoveries, as authors mentioned in discussion, the prediction considering both regulatory sequence and inverse correlation between miRNA and target mRNA expression, has been shown by many studies to greatly improve the prediction accuracy. The author should consider this in their model.
2. The proposed linear regression model also integrated the promoter methylation data in samples where array-based DNA methylation data was available. However, unlike the mRNA, copy number, and miRNA profiles, the values in TCGA DNA methylation data have an unsymmetrical beta distribution. The author should discuss whether they linear regression model is robust to the different distributions among different variables. In addition, since there are ~15% missing values in the DNA methylation data, the author should also make available their preprocessing procedures (e.g. imputation algorithm) of methylation data.
3. In result section "Joint learning of tumor models captures subtype-specific regulatory programs in GBM", the author identified miR-132 as a specific regulator for GBM proneural subtype. They next showed that samples with high model coefficients for miR-312 show a significant beneficial overall survival than samples with low model coefficients. The biological meaning of model coefficient of miR-132 in this analysis is ambiguous, although the authors seem to use it to represent the degree of dysregulation of miR-132 in each sample (Page 10, 1st sentence). Shouldn't the authors use a more intuitive and direct way by examining the association of miR-312 expression with overall survival?
Minor points:
1. Page 7, 1st paragraph, there is no "proliferative" GBM subtype in TCGA.
2. Page 10, 1st paragraph, in order to demonstrate miR-132 is an independent predictor of survival from G-CIMP status, the author can build Cox proportional hazards model including G-CIMP status and miR-132 as covariates.
3. There seems no distinct difference between the clustering results of real data (Figure 2b) and randomized motif data (supplementary Figure 2b). In both figures, the mesenchymal cases (red) are prone to be clustered together. The author may want to use some quantitative statistics (e.g. adjusted Rand Index) to compare the clustering results.

1st Revision - authors' response 12 April 2012

10 April 2012

Dr. Andrew L. Hufton, PhD
Editor
Nature/EMBO Molecular Systems Biology

Dear Dr. Hufton:

We thank the reviewers for their detailed comments on the original submission of our manuscript, "Inferring transcriptional and microRNA-mediated regulatory programs in glioblastoma." We have made substantial revisions in order to address all of the reviewers' comments, including additional requested computational analyses as well as further experimental validation to confirm the predicted phenotype of one of the identified microRNAs. We have also clarified the distinction between our integrative approach, which identifies "driver" microRNAs and TFs that account for global tumor vs. normal expression changes in target genes, and classifier-based methods, which find "biomarkers" that discriminate between specific subtypes but may not be of central biological importance. A point-by-point response to reviews is given below.

We hope that the revised version of the manuscript is now suitable for publication in *Molecular Systems Biology*.

Sincerely,

Christina Leslie
Associate Member and Lab Head
Computational Biology Program
Memorial Sloan-Kettering Cancer Center

Editor's summary

> *The editor would like to highlight two apparently fundamental issues that were raised by the reviewers. These points will need to be conclusively addressed, possibly with additional experiments and analyses, before this work would be appropriate for publication in Molecular Systems Biology.*
>
> *1. Regarding the ability of this method to identify novel factors driving the GBM subtype differences, the first reviewer found the validation experiments with miR-132 and miR-124 less than conclusive, and was not convinced that this method was genuinely superior at identifying drivers compared to previous works (e.g. Kim et al. 2011, Genome Res). This reviewer felt that it would be important to directly demonstrate phenotypic changes in response to miR-132 or miR-124.*

The previous work of Kim et al., *Cancer Research 2011*, gives a careful application of standard analyses – including clustering of microRNA expression profiles to define subtypes and examination of pairwise miRNA-mRNA correlations to try to find regulatory relationships – to study the potential role of microRNAs in tumor initiation in GBM. We describe both in the Discussion and later in our response some of the statistical advantages of our integrative approach.

We would argue that the follow-up experiments we included in the paper do show the value of the integrative approach over standard analyses, e.g. prioritizing microRNAs by their differential

expression: our predicted microRNAs (miR-124 and miR-132) as well as our controls (miR-433, miR-380, miR-448) are all expressed at lower levels in proneural tumors than in normal brain, but only the predicted microRNAs drove expression changes that were concordant with those in proneural tumors. Namely, when miR-124 or miR-132 expression was restored in a neurosphere model of proneural GBM, the set of downregulated targets in the neurosphere showed upregulation in proneural tumors, and the upregulated genes (secondary effects) in the neurosphere were downregulated in proneural tumors; these statements were not true for the control microRNAs.

However, we acknowledge that demonstrating a phenotype is stronger evidence. For miR-124, we observed that genes downregulated in the overexpression experiment were enriched for DNA replication annotations, including cell cycle, cell proliferation, and DNA repair. Therefore, we carried out a cell proliferation assay 48 hours after miR-124 transfection in PDGF-driven neurospheres and confirmed a significant reduction in the number of cells in S-phase ($p < 2e - 5$, t-test) accompanied by a similarly significant increase in the number of cells in G0/G1 phase ($p < 2e - 5$, t-test) compared to negative controls (Figure 5d). This additional data, included in the revised manuscript, establishes a cell proliferation/cell cycle phenotype for miR-124 in a proneural neurosphere model.

For miR-132, the annotations associated to its regulated target set in the model (Figure 3c) and to the differentially expressed genes in the overexpression experiment were all related to epigenetic regulation of gene expression, chromosome organization, and chromatin-related functions. Unfortunately, there is no obvious simple assay to check for global changes in chromatin state, and certainly not one that is likely to work in the setting of a transient microRNA transfection, since chromatin remodeling is expected to take time. For example, the recent result showing that the IDH1 mutation is sufficient to lead to the G-CIMP phenotype used cells stably expressing mutant IDH1 at passage 15 to demonstrate methylation changes [4]; this means we would probably need a stable miRNA expression system rather than a transient transfection. Thus we concluded that establishing a chromatin-related phenotype for miR-132 would require more involved follow-up experiments than could be reasonably carried out within the time before revision and in the context of a mainly computational paper. Instead, we have included additional discussion of miR-132's confirmed targets KDM5A (a histone demethylase) and EP300, which are both involved in epigenetic regulation of gene expression, and added evidence that KDM5A's targets are dysregulated in proneural tumors (page 16, main text and Supplementary Figure 11).

> *2. The reviewers felt that additional work was needed to demonstrate the value of the integrated model over simpler classifiers. This is best expressed in reviewer #2's point #1, but seems closely related the reviewer #1's point #2.*

It is important to clarify that our integrated approach does not solve the same problem as a classification approach. In a classification problem, one accepts that the class labels – here, proneural vs. classical vs. mesenchymal – are valid, and one trains a model that can accurately predict the class from features of one or more kinds of molecular profiling experiments (e.g. mRNA expression levels, microRNA expression levels, etc.). It is standard to perform feature selection to extract a smaller number of features that still accurately predict the class label, but typically many different "signatures" containing different sets of features have similar prediction accuracy, so inclusion of a gene/microRNA a discriminative signature is no guarantee that this "biomarker" is intrinsically important to the biology of a subtype. Moreover, in the setting of cancer subtypes, we have a problem of circularity: GBM classes are defined by clustering of expression profiles, and classifiers that successfully discriminate between these subtypes mainly show that whatever (possibly subtle) expression differences the clustering algorithm uses to make cluster assignments

can be learned and reproduced by the classifier. We do not know *a priori* that these expression clusters represent biologically distinct subtypes; they may arise for more mundane reasons, such as degree of stromal contamination.

In our integrated approach, we can ask: how biologically distinct are the different expression subtypes in terms of transcriptional and microRNA-mediated regulation? Rather than identifying "biomarkers" that may predict a subtype without necessarily being intrinsically important to the biology of the subtype, we find microRNAs and TFs that can explain global tumor vs. normal changes in gene expression within the subtype. As an example of the distinction between our approach and the classifier approach, we find strong differences in the gene regulation models for proneural and mesenchymal tumors, but we find that the classical tumors appear to be a mixture of the other subtypes, at least in terms of gene regulation. Even when we encode the three classes in our group lasso approach, we find no microRNA/TF regulators that are unique to the classical subtype. Our results are consistent with a recent meta-analysis of GBM expression subtypes by several pioneers of GBM classification, who argued that the proneural and mesenchymal subtypes are the two biological meaningful subtypes [5]. Nevertheless, it is no doubt possible to train a classifier to discriminate between tumors in the classical cluster and the other clusters and to extract classical "biomarkers" from this classifier. (Perhaps, if instead of using clustering, one tried a mixture model on gene expression profiles, one could directly identify proneural and mesenchymal subtypes as the two mixture components.)

Reviewer #1

> *The authors describe an elaborate machine learning scheme to reverse engineer integrated regulatory networks in GBM TCGA samples, including predicted protein-DNA and miRNA-mRNA interactions. They include CNV, DNase, promoter methylation data to improve their ability to predict variations in gene expression across tumors. To integrate these data, the authors use a linear regression scheme with individual weights for copy number, and each considered miRNA and TF. Predicted TF interactions were filtered using epigenetic data, and only differentially expressed miRNAs were considered. The authors make the following claims: (1) Their models predict normal vs. tumor expression changes ($r^2 < 0.18; p < 2-16$); CNV data helped ($p < 3e-16$), DNase data helped ($p <??$), and methylation data helped ($p <??$). (2) A combination of miRNAs and genes could be used to recover previously identified GBM classes (Verhaak et al.), and especially classify proneural and mesenchymal GBM subtypes. (3) The authors identified YY1, GABP, miR-124 and miR-132 as drivers of proneural GBM. (4) They find similarities between their driver coding genes and known drivers of proneural progenitors. Thus, predicted gene drivers of the proneural subtype have literature evidence. (5) Up regulation of miR-124 and miR-132 in proneural neurospheres lead to changes in expression of their direct and indirect target genes. These genes were enriched for proneural vs. mesenchymal differentially expressed genes. Suggesting that these miRs are drivers of proneural GBM and that their overexpression reverses expression changes in proneural tumors. The authors use miR-380 and miR-448, which are under expressed in proneural tumors (vs. normal?) as controls. (6) The authors suggest that miR-132, NFYB and YY1 may have common proneural-specific targets, and may regulate chromatin-related functions.*

This is a correct summary of our findings, with the following clarification: the TF/microRNA "drivers" that we find are statistical, in that they explain global tumor vs. normal gene expression

changes via presence of their binding sites in promoters/3′UTRs in a regularized regression model. We cannot claim that the regulators we find are tumor initiators based on this statistical evidence. However, if a miRNA/TF does not explain the differential expression of its predicted targets, it is unlikely to play a causal role in tumor initiation. Therefore, our analysis finds good candidates for causal regulators of GBM expression programs.

With respect to point (1) above, as shown in Supplementary Figure 1, the p-values for the performance improvement from adding DNase data and methylation data are $p < 2e - 16$ and $p < 9e - 9$ respectively using the Wilcoxon signed rank test.

> *While I appreciate the complexity of the task undertaken here and the effort by the authors, I don't believe that this work is ready for publication at MSB. Here, the authors appear to have two distinct goals and neither one of them is reached.*
>
> *(1)The authors would like to use all available tools to identify drivers of GBM subtypes. Their conclusion is that YY1, GABP, miR-124 and miR-132 may drive its proneural subtype. However, this is not supported by their validation experiments. The authors selected miR-132 and mir-124 (also selected by Tae-Min Kim, Wei Huang, Richard Park, et al., Cancer Res 2011;71:3387-3399) as candidate drivers. But these are not expressed in proneural tumors and are selected because of their targets across tumors.*

We thank the reviewer for pointing out the Kim et al. paper. We have now included the key finding of the paper in our introductory section about miRNAs in glioblastoma (page 4, main text). We would like to clarify that miR-132 and miR-124 are both underexpressed in proneural tumors relative to normal brain. All differentially expressed miRNAs in a tumor sample are considered as potential regulators for that tumor's regression model. miR-132 and miR-124 are identified as regulators for the proneural subtype by our analysis because these are consistently included in the regression models of proneural tumors. In other words, they help explain the differential expression of genes containing binding sites for these microRNAs in their 3′UTRs; moreover, the regression coefficient is consistent with microRNA-mediated regulation (these miRs are down and presence of their sites correlates with upregulation of targets).

> *The authors discovery that their targets are differentially expressed is at best validation that the miR target prediction method they used is better than random.*

It is not true that all differentially expressed microRNAs help explain target expression change; indeed, most such microRNAs are not included in the model. Therefore, the quality of our target prediction approach does not explain why miR-124 and miR-132 are chosen and many other differentially expressed miRNAs are not. Rather, our approach assumes that some microRNAs may be important in driving expression changes in the tumor, while other microRNAs may be differentially expressed as a downstream effect of important transcriptional/microRNA-mediated regulatory changes but are not themselves important to tumor etiology. (Possibly, some microRNAs are significantly differentially expressed but their absolute cellular abundance does not change enough to produce a noticeable effect.)

> *To show that miR-132 and mir-124 really drive the subtype, the authors should demonstrate phenotypic changes.*

To confirm the predicted cell proliferation role of miR-124, we carried out a cell proliferation assay after miR-124 transfection in proneural neurospheres and confirmed a significant reduction

in number of cells in S-phase ($p < 2e - 5$, t-test) and increase in number of cells in G0/G1 phase ($p < 2e - 5$, t-test) compared to negative controls (Figure 5d). This additional data establishes a cell proliferation/cell cycle phenotype for miR-124 in a proneural neurosphere model.

As explained in the response to the editorial letter above (response to point 1), the predicted role of miR-132 was epigenetic regulation of gene expression, which we did not feel could be established in the context of transient microRNA transfection experiments; the complex follow-up experiments that might confirm an epigenetic phenotype would not be possible to accomplish within several months and are out of scope for a mainly computational paper. Since proneural GBM has been connected with epigenetic changes through the discovery of G-CIMP, we do find the possible epigenetic connection between miR-132 and the proneural subtype to be tantalizing, even though we are not ourselves in the position to confirm this role. We have added additional data on confirmed miR-132 targets with a known role in epigenetic regulation, in particular KDM5A (page 16, main text). We have also made the analysis of gene ontology annotations associated with miR-132 more statistically stringent, so that only the most significant terms are described in the text.

> *Moreover, to further increase confidence in their driver prediction, the authors should use other miRNAs reported by Kim et al. as classifiers of the proneural subtype.*

Please see the explanation of the distinction between the classifier/biomarker approach and our integrated approach in the response to point 2 in the editorial letter. Our "driver" microRNAs are not features that classify the proneural subtype per se but rather are regulators whose inferred activity in the regression model explains tumor vs. normal expression changes of target genes.

Please note that binding sites for all of the miRNAs reported by Kim et al. – indeed, all differentially expressed microRNAs in each tumor (relative to normal brain) – are used as features in the regression model and therefore have the opportunity to be selected by the model. The microRNAs that were reported by Kim et al. but not found by our approach are those that do not help explain target expression changes, at least once other microRNA features are taken into account. Interestingly, we found that miR-9, a key regulator identified by Kim et al., is just below the significance threshold in the mesenchymal subtype.

> *Moreover, it is now commonly believed that the proneural subtype described by the authors is composed of at least 2 distinct types, as described by Noushmehr et al. and cited in this manuscript. One of these types is strongly associated and may be caused by promoter hyper-methylation which results from a mutation in IDH1. The authors should treat the two subtypes separately.*

We did examine this question by retraining the proneural models using G-CIMP as a separate subtype. However, methylation data was available only for 53 samples of the 161 samples in our training set and only 16/54 proneural samples. Therefore, we used the proneural G-CIMP ($N = 22$) and non-G-CIMP samples ($N = 35$) in the test set to learn both sample-by-sample and group models. Please note that while methylation data shows a clear separation between G-CIMP and non-G-CIMP subtypes (Figure R1a), gene expression clustering does not identify the same split (Figure R1b) or indeed support two transcriptomic subclasses. We observed that the regression models do not recover the G-CIMP vs. non-CIMP subtypes (Figure R1c), and the selected regulators for both subtypes are similar. Since our regression models are trained to predict gene expression changes, it is perhaps unsurprising that we do not see significant differences in regulatory programs.

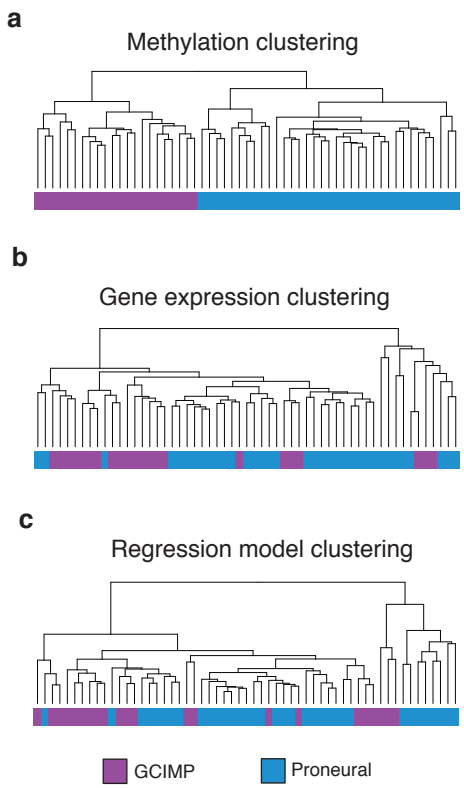


Figure R1: **Comparison of G-CIMP and non-G-CIMP proneural samples**. (a) Hierarchical clustering using methylation data separates G-CIMP and non-GCIMP proneual samples into two distinct clusters. (b) Gene expression data does not show a similar clustering. (c) As a result, the regression models do not show significant differences between the subtypes; rather, the clustering of regression models is similar to that obtained from clustering gene expression profiles.

> *(2)The authors would like to establish the improved strength of an integrated approach to tumorigenesis driver discovery, as was done before by other groups. To do this successfully, an integrated module should be discovered and demonstrated. The authors start on this path but come short of producing convincing evidence. Such evidence should demonstrate that without miRNAs or genes, our ability to explain subtype initiation is impaired. Moreover, it should be pointed out that GBM tumors are very different from normal brain tissue and are easy to classify. CNV and methylation data are especially predictive, as has been previously shown by several researchers and cited in this manuscript. Moreover, GBM classes and especially proneural and mesenchymal GBM subtypes have been previously identified and classified by genes and miRs and supported by CNVs (Verhaak et al.; Carro et al.; Kim et al.)*

We agree with the reviewer that the expression profiles of normal brain tissues are very different from GBM tumors. However, we use the normal profiles as a reference to determine tumor vs. normal gene expression changes in GBM – i.e. our y_g values in the regression model are log gene expression ratios of tumor vs. median normal rather than log intensity values from the tumor profile only. Since these normal samples may not be the most appropriate reference, we demonstrated the consistency of our predicted regulators in the proneural subtype using *in vivo* data from proneural

mouse models (Figure 4).

We would like to reiterate the difference between our approach and the more typical classifier/biomarker approach (please see text in response to point 2 of the editorial letter for a detailed description). Briefly, we are not identifying miRNAs/TFs that discriminate between subtypes but rather finding regulators that can account for tumor vs. normal expression changes either across GBM subtypes or in specific subtypes. For example, miR-124 is a significant regulator across subtypes – it is important in proneural GBM, but it does not discriminate proneural GBM from other subtypes. Meanwhile, some microRNAs are differentially expressed in proneural GBM (e.g. our controls from our follow-up transfection experiments, like miR-448) and may be useful for a classifier, but they are not included in the model because they do not significantly explain target gene expression (and in our transfection experiments, we found that the expression changes caused by miR-448 and other controls are not concordant with expression changes in tumors).

We agree with the reviewer that different copy number aberrations and DNA methylation changes are enriched in different expression subtypes. Note, however, that we are using copy number and methylation changes only as two mechanistic features to help account for local gene expression changes; microRNA and TF binding sites explain the residual expression changes after accounting for copy number and methylation. Again, this approach is different from using specific copy number/methylation changes as biomarkers for subtype classification.

To summarize, the main difference between our work and many other computational efforts is precisely that we are not clustering molecular profiles to find subtypes or using classifiers to predict these subtypes. Instead, we learn a mechanistic model for explaining gene expression changes, allowing us to ask whether transcriptional/microRNA-mediated regulation underlies common and subtype-specific expression changes in GBM. In particular, we find that proneural and mesenchymal subtypes have distinct regulatory programs, while the tumors assigned to the classical subtype seemingly represent an intermediate state in between these other two models; we find no TF/microRNA regulators unique to the classical subtype. These finding are consistent with a recent meta-analysis of GBM classification schemes, which also identifies proneural and mesenchymal as the key GBM subtypes [5]. Rather that take the "biomarker" approach, which finds genes/microRNAs that discriminate between subtypes but may not be essential to the biology of the subtype, our identified regulators and their associated gene sets provide biological insights into GBM expression programs.

> *In conclusion, I believe that the authors are doing important work, but I don't believe that this work is ready for publication at MSB.*

We hope that with the additional analyses and explanations included in the revision and response, we are able to convince the reviewer of the value and maturity of the work.

Reviewer #2

> *This manuscript from Setty et al. utilizes TCGA data on glioblastoma and outlines a computational framework to model transcription factor and microRNA regulated gene expression. Using this model, the authors show that the model parameters are capable to differentiating separate classes of glioblastoma samples, and important regulators can be predicted by computationally removing one regulator at a time from the model. Experimental overexpression of two of the miRNAs, miR-132 and miR-124, induces changes in messenger RNA expression consistent with the model. Overall, this beautifully outlined manuscript is one of the attempts to address the challenge on computational analysis of multiple datasets produced by large genomic efforts such as TCGA and ENCODE.*

It will be of interest not only to computational biologists, but to the broader community of cancer biologists and miRNA researchers. In this regard, although the manuscript is highly readable by bench scientists with sophisticated sense of computational techniques, I suggest the authors to use plain language to briefly explain seemingly standard computational techniques whenever possible.

We thank the reviewer for appreciating the novelty and potential impact of the work. We have added a few explanatory sentences to introduce computational techniques at several places in the Results section.

The following can further enhance this manuscript. 1. It is interesting that the parameters for the model can be used to cluster and predict tumor subtypes. This is a piece of the evidence that the model itself is valid. However, it is not clear (1) whether the model is providing more information than the direct expression levels of the miRNAs and TFs, and (2) whether the two subtypes are so distinct so that by random chance this classification can be achieved. Please provide analysis (1) to cluster the samples by the expression levels of the indicated TFs and miRNAs, and (2) to cluster the samples with model parameters by randomly picking the same number of TFs and miRNAs as shown in Fig 2B, from TFs and miRs that pass the selection criteria (i.e. differentially expressed etc.)

We performed hierarchical clustering shown in Figure 2b using all the transcription factors and microRNAs. As the reviewer points out, we use this clustering and the prediction of subtypes shown in Figure 2c to show that the regression models are behaving as expected, i.e. they are able to capture the differences between transcriptomic subtypes (in particular, proneural vs. mesenchymal subtypes).

We also demonstrate that the clusters are not recovered by random chance by learning the models on randomized features that pass the selection criteria (Supplementary Figure 2b). We have added the hierarchical clustering figure of TFs and miRNAs expression in Supplementary Figure 2c. This figure shows that the expression of the selected features does not recover the proneural and mesenchymal subtypes.

2. It will be nice to somehow show a distribution of the model parameters, for modeling on the single sample level. Particularly it will be interesting to know the variation of coefficients for the same miRNA on the same gene. Maybe a figure for a representative miRNA:target pair and a TF:target pair, plus a table to summarize more data.

Our regression models learn model coefficients for regulators for each sample, and we do not learn a different parameter for each regulator-target pair. However, we can use a target gene's error change when the regulator coefficient is set to zero as a measure of the extent of regulation of a particular target by the regulator. We have included plots to show the variance of model coefficient and error change for two regulator target pairs: (1) miR-132 and SOX11; (2) REST and SST. These plots are shown in Supplementary Figure 5, and a discussion has been included in page 10 of main text. A summary of the model coefficients in different subtypes has been tabulated in Supplementary Table 4.

3. For Fig3c, it will be nice to also show the survival difference separated by the expression of miR-132, in addition to using the model coefficients of miR-132.

We have added the survival difference separated expression of miR-132 in Supplementary Figure 6. As shown in the figure, expression of miR-132 does not show the same significance in survival difference as the corresponding plot for miR-132's regression coefficient. This demonstrates the value of using the extent of regulation as a factor in measuring the survival difference.

> *4. The authors claim that miR-380, miR-448 and miR-443 did not have effects on differential gene expression, unlike miR-124 and miR-132. But there is no visual data associated with this claim. Please show data, in similar format as in Fig 5a and Fig 5b.*

We had planned to include these figures in the original supplementary materials document and apologize for the oversight. We have included these figures in Supplementary Figure 9.

> *5. It is not clear to me from Supp Fig 1b how the authors concluded that "methylation consistently entering the model as a negative regulator of gene expression", although biologically it does make sense. Please explain in fig legend or in text if possible.*

In Supplementary Figure 1b, we demonstrate that including methylation as a feature in our regression model significantly increases cross-validation performance as measured by Spearman correlation. This is similar to the increase in performance obtained by including copy number as a feature (Figure 2a, red line vs. grey line). We then examined the regression models learned after including methylation as a feature and determined that the model coefficient associated with methylation is a high negative value in all the samples. This is analogous to copy numbers, which enter the model with a high positive value. Hence, we conclude that methylation is a negative regulator of gene expression. This has been clarified in the text in page 6.

> *Minor: 1. In a number of figures, including supplement, the fonts are too small on printed version of the manuscript.*

We apologize for this inconvenience. We have increased the font size of legends in a number of figures.

> *2. Supplemental Fig 1, please describe in figure legend what the colors mean and what the dots mean (I assumed the color means class label for the tumor subtypes and dots are samples).*

The reviewer is correct: the colors do mean the different subtypes. We have now added a legend indicating the same in Supplementary Figure 1.

> *3. On page 2 of text, trascriptomic subtypes "are may" arise ... should be fixed.*

We thank the reviewer for pointing out this mistake. We have fixed this in the main text.

> *4. In text describing Figure 4c, MYB should be MYBL2? MYB is not in the table.*

The reviewer is again correct that MYB should read MYBL2. This has been fixed in Figure 4c.

> *5. I suggest the authors to label their main figures with figure number, text with page number, and supplemental tables (the file in pdf format) with table numbers and page numbers, unless explicitly instructed not to by the journal.*

We have now added page numbers in the main text and supplementary text. We apologize for not including this information in our initial submission.

Reviewer #3

In the manuscript entitled "Inferring transcriptional and microRNA-mediated regulatory programs in glioblastoma", the authors proposed an integrative algorithm combining mRNA, copy number and promoter methylation profiles, along with regulatory sequence information of TF and miRNA to decipher gene expression changes in GBM. The model was trained either sample-by-sample or jointly with subtype assignments to allow the authors to identify both common and subtype-specific regulators. They subsequently focused on regulators specific for the proneural subtype. They confirmed that predicted gene expression signatures for proneural subtype regulators were consistent with in vivo expression changes in a PDGF-driven mouse model. They also tested two predicted proneural drivers, miR-124 and miR-132, which were underexpressed in proneural GBMs, by overexpression in neurospheres. This is an interesting approach to identify novel TF/miRNA for targeted therapy and for understanding mechanisms important for GBM biology given the massive amount of publicly available data. The approach led to some novel insights into the role of two miRNAs (i.e. miR-124 and miR-132) in GBM.

We thank the reviewer for this positive assessment of the manuscript.

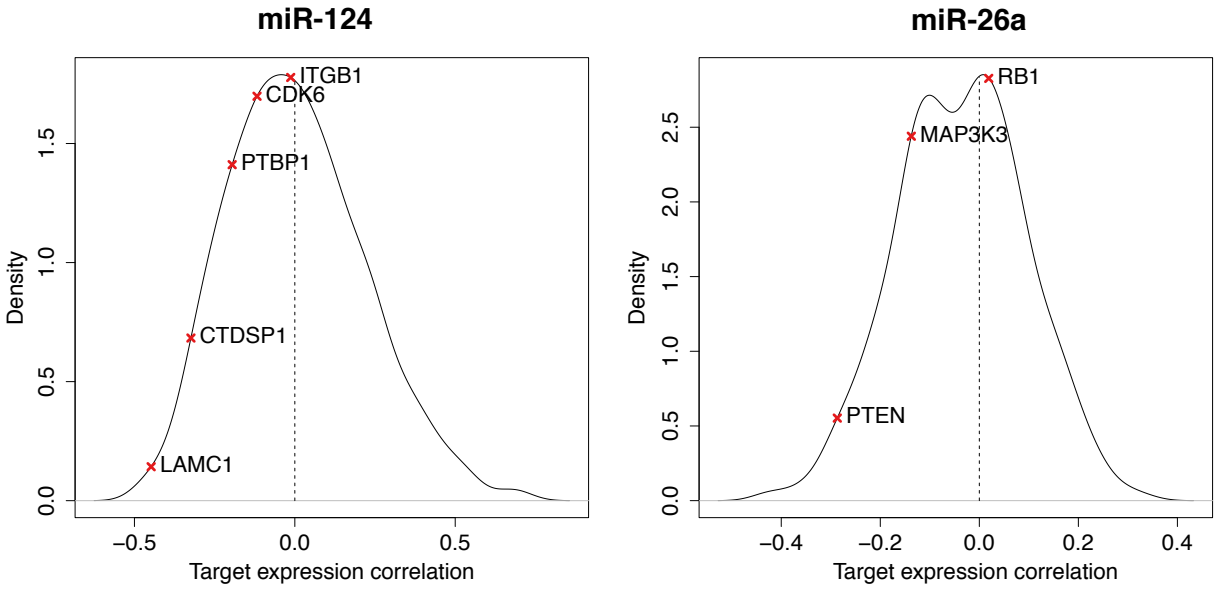
Major points: 1. The authors built their regression model simply based on the counts of TF and miRNA binding sites in the gene's regulatory regions, unfortunately neglecting the fact that the miRNA expression data is available for more than 400 GBM cases in TCGA database. The computational prediction of miRNA binding targets is known to have high false positive rate. Although the traditional pair-wise anti-correlation between miRNAs and predicted targets also introduces false discoveries, as authors mentioned in discussion, the prediction considering both regulatory sequence and inverse correlation between miRNA and target mRNA expression, has been shown by many studies to greatly improve the prediction accuracy. The author should consider this in their model.

As the reviewer points out, we do not use the microRNA expression directly in the learning process, but we do use this data to determine differentially expressed miRNAs (Figure 1, Materials and methods). We only consider differentially expressed miRNAs as features for both sample and group regression models.

Similar to the suggestion, we initially explored using the pairwise anti-correlation between expression of miRNAs and predicted targets to reduce the set of predicted targets. We decided not to use this in our models for the following reasons:

1. We observed that the distribution of these correlations is different for individual microRNAs, and hence we would need to estimate a cutoff for each miRNA separately. Moreover, analyzing the expression for each subtype separately in the group models would further increase the parameters.

2. Some of the miRNA targets that have been validated in the context of glioblastoma do not show strong anti-correlation with miRNA expression. Figure 1a and 1b show that the correlations of validated targets of miR-124 (Figure R2a) and miR-26a (Figure R2b) tend to be distributed across the correlation spectrum. We believe this lack of anti-correlation is due to combinatorial regulation by multiple TFs/miRNAs, which is conveniently handled by our regression framework but cannot be modeled using pairwise statistics.



(a) miR-124

(b) miR-26a

Figure R2: **miRNA target correlations**. Density plot of the predicted target correlation for miR-124(a) and miR-26a (b). Validated targets in glioblastoma context are highlighted and show that these correlations are distribution across the spectrum. Validated targets were obtained from [1, 2]

.

3. We perform miRNA target prediction after grouping the miRNAs by their seed families. In this situation, it is not clear which family member should be used for determining the correlations.

Instead, our regression approach has the *representational* advantage that we can model combinatorial regulation (same target gene can be regulated by multiple TFs and/or miRNAs) and the *statistical* advantage of using a supervised learning on 1000s of training examples (gene motif counts and corresponding expression changes).

Furthermore, our regression models associate a target gene set with each predicted regulator (please see "Materials and methods: Feature dependency analysis"). This associates a biological context to the predicted targets and functionality to the regulator. Finally, we note that one of the high ranking regulators is miR-218, which is in fact significantly downregulated but is still inferred as a negative regulator of its predicted targets. While we do not have a mechanistic explanation for this result, the statistical signal for miR-218 is clear (see Supplementary Figure 12). This finding would have been missed if we had restricted our analysis to miRNAs whose expression is anti-correlated with target expression levels.

2. The proposed linear regression model also integrated the promoter methylation data in samples where array-based DNA methylation data was available. However, unlike the mRNA, copy number, and miRNA profiles, the values in TCGA DNA methylation data have an unsymmetrical beta distribution. The author should discuss whether they linear regression model is robust to the different distributions among different variables. In

addition, since there are 15% missing values in the DNA methylation data, the author should also make available their preprocessing procedures (e.g. imputation algorithm) of methylation data.

We use methylation as an additional feature for a subset of the training samples (53/161). Methylation enters the model as one feature similar to copy number. As demonstrated in Supplementary Figure 1, there is a significant improvement in the regression performance using methylation as an additional feature ($p < 9e - 9$, Wilcoxon signed rank test). Additionally, the coefficient associated with methylation feature is always negative and thus consistent with the repressive role of methylation on gene expression. We concur with the reviewer that there might be better approaches to account for the asymetric beta distribution of methylation data, but these observations lead us to conclude that using methylation as a linear covariate is a reasonable first approximation. Finally, we do not impute missing values in DNA methylation data and set the feature value to 0 if methylation is not available.

3. In result section "Joint learning of tumor models captures subtype-specific regulatory programs in GBM", the author identified miR-132 as a specific regulator for GBM proneural subtype. They next showed that samples with high model coefficients for miR-312 show a significant beneficial overall survival than samples with low model coefficients. The biological meaning of model coefficient of miR-132 in this analysis is ambiguous, although the authors seem to use it to represent the degree of dysregulation of miR-132 in each sample (Page 10, 1st sentence). Shouldn't the authors use a more intuitive and direct way by examining the association of miR-132 expression with overall survival?

We have now included survival difference based on miR-132 expression in Supplementary Figure 6. As shown in the figure, expression of miR-132 does not show the same significance in survival difference. This demonstrates the value of using the extent of regulation as a factor in measuring the survival difference. Since a tumor is a complex system with many interacting components, its possible the extent of dysregulation of a particular factor is not fully reflected in its expression changes; moreover, noise in microarray data means that there is uncertainty in individual measurements of a microRNA's expression level. Hence using the extent of dysregulation that can attributed to miR-132 appears to be a better predictor of survival in proneural glioblastoma patients than its expression level.

Minor points: 1. Page 7, 1st paragraph, there is no "proliferative" GBM subtype in TCGA.

The proliferative subtype was defined in an earlier GBM study [3]. This has been clarified in the text.

2. Page 10, 1st paragraph, in order to demonstrate miR-132 is an independent predictor of survival from G-CIMP status, the author can build Cox proportional hazards model including G-CIMP status and miR-132 as covariates.

We do not have methylation data for all the samples in our training set. Of the 54 proneural samples, methylation data was available for 16 samples of which 7 were classified as G-CIMP. Thus its possible that some of the remaining 38 samples with no methylation data might also

be potentially G-CIMP. We tried training Cox proportional hazard models using (i) only miR-132 as covariate and (ii) miR-132 and G-CIMP as covariates. We found that in both cases the coefficient associated with miR-132 was significant ($p < 0.003$), whereas the coefficient associated with G-CIMP status in (ii) was not significant ($p < 0.11$).

> *3. There seems no distinct difference between the clustering results of real data (Figure 2b) and randomized motif data (Supplementary Figure 2b). In both figures, the mesenchymal cases (red) are prone to be clustered together. The author may want to use some quantitative statistics (e.g. adjusted Rand Index) to compare the clustering results.*

We thank the reviewer for pointing out the adjusted Rand index statistic for quantifying the hierarchical clusters. We determine the adjusted Rand index of the clustering of the models show in Fig 2b to be 0.8233 and the adjusted Rand index of the clustering obtained by randomized features shown in Supplementary Figure 2b to be 0.231. This demonstrates that the clustering obtained by learning models cannot be reproduced using randomized features. This information has been included in page 7 of the main text.

References

[1] MicroRNAs and glioblastoma; the stem cell connection. Godlewski J, Newton HB, Chiocca EA, Lawler SE. Cell Death & Differentiation. June 2009.

[2] Integrative genome analysis reveals an oncomir/oncogene cluster regulating glioblastoma survivorship. Kim H, Huang W, Jiang X, Pennicooke B, Park PJ, Johnson MD. PNAS. Feb 2010.

[3] Molecular subclasses of high-grade glioma prediction prognosis, delinate a pattern of disease progression, and resemble stages in neurogenesis. Phillips HS et. al. Cancer Cell. March 2006.

[4] IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, et al. Nature. Feb 2012.

[5] Molecular subclassification of diffuse gliomas: seeing order in the chaos. Huse JT, Phillips HS, Brennan CW. Glia. Mar 2011.

2nd Editorial Decision 18 May 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who agreed to evaluate your revised manuscript. As you will see from the reports below, the referees did feel that the changes made to this work, and the new experimental evidence presented, improved this work. The reviewers, however, have important remaining concerns, which, I am afraid to say, must preclude its publication at present.

In most cases, Molecular Systems Biology only allows a single round of major revision. Nonetheless, the first reviewer seemed to feel that at least some of his/her concerns were related to issues of presentation, and the second reviewer felt that additional independent validation experiments could be conducted with existing public datasets. As such, I would like to offer you the exceptional opportunity to prepare a second and final revision of this work. Below I outline the most important outstanding issues as I see them:

1. Circularity, Classification, and Independent Validation

The reviewers are still troubled by the validation for the classification accuracy of this approach, with both reviewers raising concerns regarding the independence of the current validation.

-- The ability of miR-132's regulatory coefficient to predict survival is tested on the same of samples that identified it as a key regulator, so the validation is not fully independent. The second reviewer suggests that it may be possible perform a similar analysis on an independent GBM expression dataset (even in the absence of copy number data). Some type of fully independent validation of the predictive ability of these regulatory coefficients, as suggested, may be the most direct way to address the reviewers' concerns.

-- The tests of subtype classification accuracy (page 7-8) are to some degree circular because the subtype labels were originally assigned by gene expression signatures. You acknowledge this issue in your response to the reviewers, and then emphasize that the main aim of your work was not subtype classification, but rather functional identification of regulatory drivers that underlie subtype differences. Nonetheless, the first reviewer clearly felt that this issue was not clearly presented in the main manuscript. The editor feels that the main aims of this work must be made much more apparent in the manuscript, and the claims regarding the classification accuracy of your models must be reduced and associated with clear caveats, i.e. the classification analyses only provide technical validation showing that the more complex regression models used here can recover existing expression-based subtypes.

On a somewhat related note, the editor feels that the underlying TCGA GBM datasets used in your work should be described in more detail, so that others can replicate these analyses. It would be useful if a table is provided listing the samples, and clearly noting which were included in the initial training set of 161 samples and which were in the independent test set of 160 samples. Please also list the subtype classification for each sample and the source for this classification (i.e. were the same expression signatures used consistently in classifying all samples?).

2. Reviewer #3 still feels that the effect of microRNA expression on the regression models should be tested. Even if you chose not to use these data in subsequent analyses for the reasons outlined in your previous response letter, the editor agrees that it would be useful to directly test whether microRNA expression can improve the regression models, and to report the model improvement (or lack thereof) as compared to the other data types.

In addition, the editor asks that you to address the following format and data issues when preparing your revised work:

1. Please deposit all new expression datasets (mouse OPC and neurosphere profiles) in a public repository, such as GEO or ArrayExpress, and include a confidential reviewer login with the revised manuscript.

2. Please state the number of independent biological replicates that underlie Fig. 5d. In addition, we would strongly encourage you to provide the underlying numeric data as "figure source data", (e.g. <http://tinyurl.com/365zpej>). Please see our Instructions of Authors for more details on preparation and formatting of figure source data (<http://www.nature.com/msb/authors/index.html#a3.4.3>).

Please also include all of the items in the standard revision checklist (below), when submitting your revised work.

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 favorable.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Reviewer #1 (Remarks to the Author):

While I appreciate the authors' efforts to clarify and improve the original manuscript, I remain confused when trying to identify the key contributions made here. I find that in order to discuss the manuscript, I have to reiterate its accomplishments as I see them.

Using established machine learning techniques, the authors train a predictive model of differential gene expression in glioblastoma. Specifically, the model uses gene and miRNA expression data, and CNV and methylation array data within a regression scheme to predict tumor/normal differential expression. The choice to learn differential expression rather than explicit tumor expression is interesting and note worthy and is important for both conceptual and technical reasons. Having selected predictive features, when possible, the authors proceeded to learn associated high-quality target sets and to infer activity in GBM subtypes. The authors extended the finding by integrating multiple regulators to infer activity of "regulator modules". Finally, using statistical and experimental methodology the authors provide evidence that their inferred target sets are enriched with true targets.
These results, as I see them, are interesting, but there are fundamental disagreements between my list and that claimed by the authors. I'll try to enumerate my reservations.
(1) Identification of true drivers of disease vs. differentially expressed regulators that do not drive disease - the authors claim that the identification of differentially expresses regulators with differential activity is evidence for their being drivers of disease, while differentially expressed regulators that are not associated with differential activity may be passengers. An alternative explanation is that differentially expressed regulators whose targets (activity) are not differentially regulated are associated with a poorly inferred target set or inferred activity.
(2) Approaches based on regression and higher-order correlation are superior to standard expression clustering analysis. The work by Kim et al. is rooted in expert knowledge and its use of simple methodology makes it less prone to exaggerating predictive ability. The problems of circularity cited by the authors in the reply to reviewers are present in their analysis too. While their FDR cutoff may not lead to predicting regulators of the classical GBM subtype, setting appropriate cutoff using simple clustering leads to the same conclusions. Moreover, given that hundreds if not thousands of genes (depending on cutoff) are both differentially expressed across tumor/normal and mesenchymal/proneural GBM, it is not surprising that features derived from tumor/normal differential expression data will also be predictive of mesenchymal/proneural differential expression. The circularity problem is transformed but it is not eliminated. Indeed, we expect that regulators that are differentially expressed across mesenchymal/proneural GBM will have mesenchymal/proneural specific activity. Moreover, because the two subtypes have such a large distinct differential program, simple cross validation will not help resolve the circularity problem either. The test set was also formulated using global gene expression patterns and it is informed

about expression patterns in the training set.
(3) Because of the special structure of the both dataset and the predictive features derived by the author, Sup. Fig. 2b is not a sufficient null model. Ideally the null model will retain key features - regulators that are differentially expressed and features that are as predictive of its activity in the entire set or the remainder of the set. Selecting random features is not a fair comparison here.
(4) The revision includes validation of miR-124 as a regulator of cell cycle in proneural GBM. While the introduction of miR-124 may significantly slow cell proliferation, the change is no greater than 15%. Moreover, will inhibition of miR-124 in mesenchymal GBM have a different effect? It's hard for me to directly relate the presented result to the idea that miR-124 is a driver regulator of proneural GBM. Isn't this additional evidence that its predicted target set is enriched with true targets?
(5) My main issue with the original manuscript was that it claimed to identify sub networks, composed of miRNAs and TFs, that may drive GBM subtype, but offers little direct evidence to this effect. This matter remains unresolved.

In short, I applaud the authors for embarking on this study and I do believe that they may be able to point to enriched target sets for certain regulators, but I don't agree with some of their far reaching conclusions. I see this as an integrative approach to improve the prediction of functional TF and miRNA targets. As such, this would be a solid contribution, but to be presented as such, the manuscript will need substantial modification.

Reviewer #2 (Remarks to the Author):

The authors have addressed all my concerns in the initial review.

Just a brief and minor point to improve data representation of the new data figure Fig 5d: it will be nice to show representative flow cytometry plots for the cell cycle analysis, possibly in supplements, in addition to the bar plot of Fig 5d.

Reviewer #3 (Remarks to the Author):

In their revised manuscript, Setty et al have performed some additional analyses and experimental validation to confirm the predicted phenotype of one of the identified microRNAs. My major concern was that the author claimed that they took the advantage of integrated TCGA data to indentify the driver miRNA/TF, however, the miRNA/TF TCGA data was the least they explored. In their regression model, they considered each gene's copy number and promoter methylation status from TCGA, but only considered the number of predicted binding site of microRNA/TF. The authors argued in their revised manuscript by showing some false negative cases that the miRNA expression data may introduce. But the authors should further support their argument by including miRNAs expression in the model and evaluating how that would improve the model. This would be similar to what authors showed in Figure 1a that the inclusion of copy number data can greatly improvement the performance of model.
My second concern is that the authors tried to use the model coefficient of each regulator in each sample to represent the "extent of regulation" of candidate regulator. They further use this coefficient to classify the tumor subtypes and to associate with patient survival. In their revised manuscript, they showed that this coefficient is superior to the actual expression of miRNA/TF in terms of subtype classification and association with survival (e.g. miR-132). This is an interesting result. However, it is not clear whether this can be applied to independent data. The manuscript will be largely enhanced if the author can apply their method in an independent and public accessible gene expression data and validate their discovery. This is feasible, because the authors' model relies largely on gene expression, miRNA and TF binding sites. The copy number and miRNA expression are not necessary to infer the model coefficients. A minor point is that the authors may want to consider changing Figure 4a, b, Figure 5a and b into GSEA figures.

2nd Revision - authors' response 12 June 2012

11 June 2012

Dr. Andrew L. Hufton, PhD
Editor
Nature/EMBO Molecular Systems Biology

Dear Dr. Hufton:

We thank the reviewers for their further comments on the revised version of our manuscript, "Inferring transcriptional and microRNA-mediated regulatory programs in glioblastoma," and we are appreciative of the opportunity to submit a second and final revision in order to address remaining issues.

We agree with your assessment that most of the concerns of Reviewer #1 are related to presentation, including the need for a clearer description of the goals of our algorithmic approach. In our previous response to reviews, we tried to clarify the distinction between our integrative approach, which identifies "driver" microRNAs and TFs that account for global tumor vs. normal expression changes in target genes, and classifier-based methods, which find "biomarkers" that discriminate between specific subtypes but may not be of central biological importance. We now make note of this distinction in a section in the conclusion, and in the introduction we clarify that classifying tumors into subtypes is not the goal of the study. We have also removed one figure panel (Figure 2C) which contained a sanity check that the regression models could be used to predict subtype and placed this result in the supplement; we felt that this sanity check result might be confusing readers about the goals of the study by blurring the distinction between classification approaches and modeling regulation of expression changes.

We have also clarified the term "driver" as applied to TFs and microRNAs, as we use it in a statistical sense to denote regulators whose inferred dysregulated activity can account for widespread tumor vs. normal expression changes. We do not assert that the dysregulation of these TFs/microRNAs represents the set of key events in tumor development. Finally, we have included the additional survival analysis and results using microRNA expression levels in the model, as requested by the Reviewer #3, and we have fulfilled all editorial requests for description of data sets and depositing of new experimental data in public repositories. A point-by-point response to reviews is given below.

We hope that the revised version of the manuscript is now suitable for publication in *Molecular Systems Biology*.

Sincerely,

Christina Leslie
Associate Member and Lab Head
Computational Biology Program
Memorial Sloan-Kettering Cancer Center

Editor's summary

In most cases, Molecular Systems Biology only allows a single round of major revision. Nonetheless, the first reviewer seemed to feel that at least some of his/her concerns were related to issues of presentation, and the second reviewer felt that additional independent

validation experiments could be conducted with existing public datasets. As such, I would like to offer you the exceptional opportunity to prepare a second and final revision of this work. Below I outline the most important outstanding issues as I see them:

We very much appreciate the opportunity to submit a second and final revision. As described below, we have addressed all reviewer comments and editorial requests.

1. Circularity, Classification, and Independent Validation

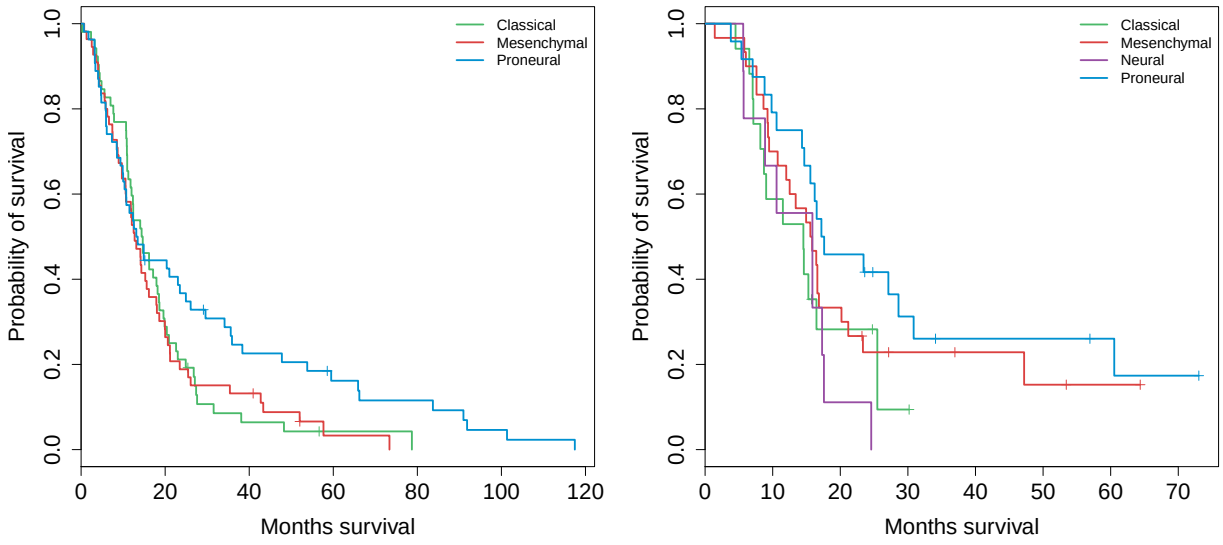
The reviewers are still troubled by the validation for the classification accuracy of this approach, with both reviewers raising concerns regarding the independence of the current validation.

– The ability of miR-132's regulatory coefficient to predict survival is tested on the same of samples that identified it as a key regulator, so the validation is not fully independent. The second reviewer suggests that it may be possible perform a similar analysis on an independent GBM expression dataset (even in the absence of copy number data). Some type of fully independent validation of the predictive ability of these regulatory coefficients, as suggested, may be the most direct way to address the reviewers' concerns.

W.r.t. "classification accuracy", please note that we learn a separate regression model for each tumor, so the model coefficients cannot be transferred to other tumors. (We can compute the average regression model over the training data set or over samples belonging to a predefined subtype, but this does not help us infer the tumor-specific activity of a microRNA for each sample in the test set.) As described above, in order to de-emphasize the use of the models to classify subtypes – a result we mainly included as a sanity check – we have moved this analysis to the supplement.

We note that we did not use any information about survival in our regression models. Therefore, the ability of miR-132 to predict survival within the TCGA data is not the result of training bias. Nevertheless, we did identify an additional data set from Murat et al. [1] which contained normal brain as well as GBM expression data, enabling us to perform an independent analysis. It should be noted that this data set contains only gene expression data and not copy number or miRNA expression data. Moreover, the data set is small in size ($N = 80$), which limits the power of survival analysis. For example, after we classified tumors in this data set into proneural, mesenchymal, classical, and neural using the gene signatures defined in [2] (see Materials and Methods), we saw a trend towards longer survival in proneural samples versus non-proneural samples (the main survival result from [2]), but the p-value was not significant (Figure R1). This unfortunately suggests that the sample size is too small to replicate survival analysis from the larger TCGA data set. We then ran our sample-by-sample regression models (removing neural samples) and measured the survival difference between patients with high and low miR-132 model coefficients. We again saw a trend that patients with high miR-132 coefficients have longer survival (Supplementary Figure 6b), but again the p-value is not significant. Given the limitations of the data set (sample sample size, no copy number or miRNA expression data), we see this result as consistent with our main survival analysis for miR-132 and the best we could do in terms of replicating the analysis with another study. Since other available GBM data sets do not include normal brain reference samples, we were not able to use them for independent replication of the analysis.

– The tests of subtype classification accuracy (page 7-8) are to some degree circular because the subtype labels were originally assigned by gene expression signatures. You acknowledge this issue in your response to the reviewers, and then emphasize that the main aim of your work was not subtype classification, but rather functional identification



(a) TCGA

(b) Murat

Figure R1: **Survival analysis in TCGA subtypes**. (a) Proneural GBM patients show a significantly better survival compared to other subtypes in TCGA dataset. (b) Proneural samples in Murat dataset show marginally better survival.

> *of regulatory drivers that underlie subtype differences. Nonetheless, the first reviewer clearly felt that this issue was not clearly presented in the main manuscript. The editor feels that the main aims of this work must be made much more apparent in the manuscript, and the claims regarding the classification accuracy of your models must be reduced and associated with clear caveats, i.e. the classification analyses only provide technical validation showing that the more complex regression models used here can recover existing expression-based subtypes.*

We have tried to clarify the goals in the introduction: previously, for compactness, we referred to regression models trained "sample-by-sample or jointly with subtype assignments"; now, we explain the sample-by-sample approach first, discuss checking how well tumor-specific regression models correlate with subtype, reiterate that our method is not primarily about subtype classification, then finally describe how we can incorporate subtype assignments to jointly learn the regression models for all tumors. We have put the former Figure 2C panel on classification accuracy into the supplement in order to deemphasize this result.

> *On a somewhat related note, the editor feels that the underlying TCGA GBM datasets used in your work should be described in more detail, so that others can replicate these analyses. It would be useful if a table is provided listing the samples, and clearly noting which were included in the initial training set of 161 samples and which were in the independent test set of 160 samples. Please also list the subtype classification for each sample and the source for this classification (i.e. were the same expression signatures used consistently in classifying all samples?).*

We have added an additional supplementary table describing the training and test datasets (Supplementary Table 1). Four transcriptomic classes in the training data set were defined using consensus clustering in [2], and samples in the test set were classified into these four categories using gene signatures derived from the training set [2]. We have added this information in the Materials and Methods section under the title "Classification of test samples". In order to further aid the reproducibility, we have created an R package called "RegulatorInference" and made it available at `http://cbio.mskcc.org/Public/Leslie/RegulatorInference` (Username: msb, Password: gbm). We will make this package publicly available upon publication of the manuscript.

> *2. Reviewer #3 still feels that the effect of microRNA expression on the regression models should be tested. Even if you chose not to use these data in subsequent analyses for the reasons outlined in your previous response letter, the editor agrees that it would be useful to directly test whether microRNA expression can improve the regression models, and to report the model improvement (or lack thereof) as compared to the other data types.*

To test the usefulness of miRNA expression levels, we ran our sample-by-sample regression using miRNA expression instead of conserved seed matches as feature values, and we found that the cross-validation results did not change significantly. (For a fairer comparison, we still prune features based on differential expression, so that the new model does not have to contend with a larger number of irrelevant features.) A discussion of this additional modified model has been included in the manuscript, and the result is presented in Supplementary Figure 1c. Since we do not see a significant change, we decided to retain all the downstream analyses based on conserved seed matches.

> *In addition, the editor asks that you to address the following format and data issues when preparing your revised work:*
>
> *1. Please deposit all new expression datasets (mouse OPC and neurosphere profiles) in a public repository, such as GEO or ArrayExpress, and include a confidential reviewer login with the revised manuscript.*

We have deposited all the expression data to Gene expression omnibus. The data can be accessed using the accession numbers GSE32876 and GSE38591 (`http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=nrmfhecakcsomxu&acc=GSE38591`). This information has been included in the "Data availability" section in Materials and methods.

> *2. Please state the number of independent biological replicates that underlie Fig. 5d. In addition, we would strongly encourage you to provide the underlying numeric data as "figure source data", (e.g. <`http://tinyurl.com/365zpej`>). Please see our Instructions of Authors for more details on preparation and formatting of figure source data (<`http://www.nature.com/msb/authors/index.html#a3.4.3`>).*

We used three independent replicates for the data shown in Fig. 5d. This has been clarified in the manuscript, and we have also included a figure source data file as part of the submission.

Reviewer #1

> *While I appreciate the authors' efforts to clarify and improve the original manuscript, I remain confused when trying to identify the key contributions made here. I find that in order to discuss the manuscript, I have to reiterate its accomplishments as I see them.*

Using established machine learning techniques, the authors train a predictive model of differential gene expression in glioblastoma. Specifically, the model uses gene and miRNA expression data, and CNV and methylation array data within a regression scheme to predict tumor/normal differential expression. The choice to learn differential expression rather than explicit tumor expression is interesting and note worthy and is important for both conceptual and technical reasons. Having selected predictive features, when possible, the authors proceeded to learn associated high-quality target sets and to infer activity in GBM subtypes. The authors extended the finding by integrating multiple regulators to infer activity of "regulator modules". Finally, using statistical and experimental methodology the authors provide evidence that their inferred target sets are enriched with true targets. These results, as I see them, are interesting, but there are fundamental disagreements between my list and that claimed by the authors. I'll try to enumerate my reservations.

We would like to provide a few clarifications to this summary of our study. First, the machine learning methods (sparse structured regression) are indeed established outside of computational biology but have not been previously been applied in the context of modeling expression changes in cancer. While we build on previous regression approaches for modeling transcriptional response in yeast, the ability to scale up to human tumor data sets, the inclusion of microRNA-mediated regulation, and many technical choices (use of DNase accessibility to define promoter regulatory regions, group lasso for sharing information across tumors while modeling subtypes) are important to the success of the approach. We therefore feel that the formulation of the learning problem and the introduction of the algorithmic approach are significant contributions of our work.

We do not use the term "regulator modules" in the paper. We do combine multiple regulators in the linear regression model – this simply means that a gene's expression change is explained by a set of TFs that target it via binding sites in the promoter and by a set of microRNAs that target it via sites in the 3′UTR. This approach is statistically more powerful than e.g. looking separately at each predicted TF or miRNA target set and testing for enrichment at the top or bottom of the distribution of gene expression changes – we account for the simple fact that genes have multiple regulators.

We disagree with the statement that we use statistical and experimental methodology to provide evidence that our "inferred target sets are enriched with true targets". We actually do not address this question, since our paper is not about improving TF or microRNA target prediction. In particular, our experimental validation is meant to show that identifying microRNAs through the regression model is more powerful than just looking for differentially expressed microRNAs – i.e. if we do microRNA overexpression experiments to see what mRNA expression changes are actually caused by a microRNA, then the expression changes for our inferred microRNAs are consistent with expression changes in tumor data sets. Meanwhile, our control microRNAs – these are microRNAs that are differentially expressed in proneural tumors based on microRNA microarray data but are not chosen by the regression model – drive expression changes that seem to have nothing to do with tumors. Note that this experimental validation has almost nothing to do with the target prediction method: we overexpress the microRNA, we see which genes go down and up, and we ask whether these sets behave concordantly in the tumor expression data. We need to qualify "almost nothing to do with target prediction", in that we filter the genes that go down by requiring that they have a conserved 7-mer seed match – so that this list is closer to an experimentally defined target list – however if we remove the filter, we get the same result.

(1) Identification of true drivers of disease vs. differentially expressed regulators that do not drive disease - the authors claim that the identification of differentially expresses

regulators with differential activity is evidence for their being drivers of disease, while differentially expressed regulators that are not associated with differential activity may be passengers. An alternative explanation is that differentially expressed regulators whose targets (activity) are not differentially regulated are associated with a poorly inferred target set or inferred activity.

We now clarify in the text that we use the term "driver" in a statistical sense to denote regulators whose inferred dysregulated activity can account for widespread tumor vs. normal expression changes in the model. We cannot assert that dysregulation of these TFs/microRNAs comprises the key events that led to tumor development – we simply claim that the dysregulation of these TFs/microRNAs can account for a significant component of the pathological expression patterns in tumors.

W.r.t. the reviewer's concern about the relationship between a microRNA having a good starting list of targets ("enriched with true targets") and getting chosen by the model, we can examine the logic as follows. It is true that if a microRNA truly explains expression changes but for some reason, our simple target prediction method (conserved 7-mer seeds) does especially badly for this microRNA, then we have no hope of finding that microRNA in the regression model. Therefore, poor target prediction could lead to false negatives. However, we do not identify miR-124 and miR-132 simply because the target prediction works well enough, but also because presence of their seeds helps explain differential expression. Moreover, the overexpression experiments confirm that these microRNAs drive expression changes that are concordant with expression changes seen in proneural tumors. In that sense, we establish that miR-124 and miR-132 are not false positives. Similarly, the experiments for the control miRNAs (which are differentially expressed but not chosen by the model) suggest that these controls are true negatives.

More generally, in our lab over the past 4 years we have consistently found that we could infer dysregulated miRNAs through statistical analysis of mRNA data, even when using simple target prediction methods. For example, in Khan et al. (Nature Biotechnology, 2009), we give statistical evidence that overexpression of an exogenous microRNA can lead to derepression of targets of endogenous microRNAs through competition for protein machinery like RISC. Through an unbiased regression analysis of many microRNA expression experiments in HeLa cells, the microRNAs that are identified with positive regression coefficients (targets are upregulated) coincide with the most highly expressed miRNAs in HeLa. This result suggests that by and large we are not suffering from too many false negatives due to some miRNAs having poorly predicted target sets.

(2) Approaches based on regression and higher-order correlation are superior to standard expression clustering analysis. The work by Kim et al. is rooted in expert knowledge and its use of simple methodology makes it less prone to exaggerating predictive ability. The problems of circularity cited by the authors in the reply to reviewers are present in their analysis too. While their FDR cutoff may not lead to predicting regulators of the classical GBM subtype, setting appropriate cutoff using simple clustering leads to the same conclusions. Moreover, given that hundreds if not thousands of genes (depending on cutoff) are both differentially expressed across tumor/normal and mesenchymal/proneural GBM, it is not surprising that features derived from tumor/normal differential expression data will also be predictive of mesenchymal/proneural differential expression. The circularity problem is transformed but it is not eliminated. Indeed, we expect that regulators that are differentially expressed across mesenchymal/proneural GBM will have mesenchymal/proneural specific activity. Moreover, because the two subtypes have such a large distinct differential program, simple cross validation will not help resolve the

circularity problem either. The test set was also formulated using global gene expression patterns and it is informed about expression patterns in the training set.

Our approach based on regression is different than standard expression clustering, with different goals, which we now try to clarify in the manuscript. Clustering tumors by expression profiles simply tries, without supervision, to find subpopulations ("transcriptomic subtypes") in the data set. A priori, we do not know that the different subpopulations are associated with distinct transcriptional and microRNA-mediated expression programs: the differences in expression patterns between subpopulations might be attributable to different frequencies of large-scale copy number aberrations or to different degrees of stromal contamination, to give two possibilities. In our sample-by-sample approach, we do not use subtypes at all, and we simply ask whether we can explain tumor-vs.-normal expression changes in terms of TF/miRNA regulation. We then relate the inferred regulators to the pre-existing subclasses by using the group lasso model; since the subtype definitions pre-date our models, it is interesting to ask about the relationship between the two approaches, e.g. whether there are dysregulated miRs/TFs specific to pre-defined subtypes.

It is true that mesenchymal/proneural types have very different expression profiles, but we show that we cannot learn the distinction between the two with a randomized feature map. Therefore, the fact that we learn distinct regression models for tumors in these two subtypes is not attributable just to random correlations between motif features and systematic subtype-specific expression patterns.

(3) Because of the special structure of the both dataset and the predictive features derived by the author, Sup. Fig. 2b is not a sufficient null model. Ideally the null model will retain key features - regulators that are differentially expressed and features that are as predictive of its activity in the entire set or the remainder of the set. Selecting random features is not a fair comparison here.

We are unsure of what is being requested here. In our approach, we explain differential expression in terms of regulatory elements (TF/miR binding sites). To ask whether we might be identifying regulators "at random" due to random correlations between motif hits and differential expression, we need a null model where the randomized motif matrix has the same distributional properties as the real one but the original motif-to-target-genes mapping has been lost. We do this in two ways: (1) by permuting the rows of the expression matrix, so that every gene gets another gene's motifs; (2) by randomizing the hit list for each motif, so that the corresponding regulator has the same number of targets, but the target gene list is randomized. These seem to be logical controls for our purpose.

(4) The revision includes validation of miR-124 as a regulator of cell cycle in proneural GBM. While the introduction of miR-124 may significantly slow cell proliferation, the change is no greater than 15%. Moreover, will inhibition of miR-124 in mesenchymal GBM have a different effect? It's hard for me to directly relate the presented result to the idea that miR-124 is a driver regulator of proneural GBM. Isn't this additional evidence that its predicted target set is enriched with true targets?

We did not repeat the experiment in mesenchymal tumorspheres, but note that miR-124 is selected in the model across subtypes – it is underexpressed relative to normal in all subtypes and is chosen with a positive regression coefficient (targets unregulated relative to normal, i.e. derepressed). Therefore, we would expect a cell cycle phenotype in mesenchymal tumorspheres as well. We did our experiments in proneural GBM due to the expertise of our collaborators and availability tumor and normal expression data of a related mouse model.

Again, we only claim that dysregulation of miR-124 is a significant driver of expression changes in proneural GBM, not that is drives tumor development. In all our overexpression experiments – including the controls – the downregulated genes are enriched with predicted targets. The point is that when we look at these downregulated genes with conserved 7-mer seeds – our experimentally defined "true targets" – and see how they behave in proneural tumors, we see concordant expression changes for miR-124 "true targets" and no concordance for the controls. That is, we use the overexpression experiments to see what changes the microRNA causes, and then we go back and check if these changes are consistent with changes in the tumors. We are not using the transfection experiments to confirm our target predictions.

> *(5) My main issue with the original manuscript was that it claimed to identify sub networks, composed of miRNAs and TFs, that may drive GBM subtype, but offers little direct evidence to this effect. This matter remains unresolved.*

To clarify, we do not claim that our algorithmic find "subnetworks" of miRNAs and TFs driving GBM subtypes. All the miRNAs and TFs start as independent features in our regression models, and the regularization structure encourages inclusion of a small set of non-redundant features that explain tumor-vs.-normal gene expression changes. We give statistical evidence that the miRNAs and TFs identified by the model can account for expression changes (1) by cross-validation on held-out genes within a tumor and (2) on held-out tumors, using average regression models for each subtype over tumors in the training set. We further show that two microRNAs identified for the proneural subtype drive expression changes that are concordant with their inferred role in proneural tumors. The control microRNAs do not show this concordance: these microRNAs of course drive expression changes when overexpressed, their predicted targets are enriched in the list of genes that go down, but these downregulated targets do not show concordant expression changes in tumors. We also demonstrate a predicted phenotype, namely a cell cycle phenotype for miR-124. Using our experimental results and literature evidence, we suggest a regulatory network underlying experssion changes in proneural GBM (Figure 5e) in the Discussion (which is the only place where the term "network" is used). True *in vivo* validation in a mouse model is clearly beyond the scope of a primarily computational paper.

> *In short, I applaud the authors for embarking on this study and I do believe that they may be able to point to enriched target sets for certain regulators, but I don't agree with some of their far reaching conclusions. I see this as an integrative approach to improve the prediction of functional TF and miRNA targets. As such, this would be a solid contribution, but to be presented as such, the manuscript will need substantial modification.*

We hope we have convinced the reviewer that the paper is not primarily about improving TF/miRNA target prediction, but rather a way of inferring dysregulated TFs/miRs despite noisy target prediction (we can define regulated TF/miRNA target sets based on the model, but we mainly use this to improve gene ontology analysis). We hope that the changes we have made to the manuscript have clarified our goals and terminology sufficiently.

Reviewer #2

> *The authors have addressed all my concerns in the initial review.*

Just a brief and minor point to improve data representation of the new data figure Fig 5d: it will be nice to show representative flow cytometry plots for the cell cycle analysis, possibly in supplements, in addition to the bar plot of Fig 5d.

We have added representative flow cytometry plots in Supplementary Figure 11.

Reviewer #3

In their revised manuscript, Setty et al have performed some additional analyses and experimental validation to confirm the predicted phenotype of one of the identified microRNAs. My major concern was that the author claimed that they took the advantage of integrated TCGA data to identify the driver miRNA/TF, however, the miRNA/TF TCGA data was the least they explored. In their regression model, they considered each gene's copy number and promoter methylation status from TCGA, but only considered the number of predicted binding site of microRNA/TF. The authors argued in their revised manuscript by showing some false negative cases that the miRNA expression data may introduce. But the authors should further support their argument by including miRNAs expression in the model and evaluating how that would improve the model. This would be similar to what authors showed in Figure 1a that the inclusion of copy number data can greatly improvement the performance of model.

We thank the reviewer for additional comments on the manuscript. We have now included the analysis using miRNA expression instead of number of conserved seed matches. The results are shown in Supplementary Figure 1c. We do not see a significant change in cross-validation results between miRNA expression and conserved seed matches. The results are from sample-by-sample analysis and were obtained by replacing the number of conserved seed matches in the feature matrix with the highest expressed miRNA in the family for each sample. Since the cross-validation results are more or less unaltered, we decided to retain all the downstream analyses based on conserved seed matches.

My second concern is that the authors tried to use the model coefficient of each regulator in each sample to represent the "extent of regulation" of candidate regulator. They further use this coefficient to classify the tumor subtypes and to associate with patient survival. In their revised manuscript, they showed that this coefficient is superior to the actual expression of miRNA/TF in terms of subtype classification and association with survival (e.g. miR-132). This is an interesting result. However, it is not clear whether this can be applied to independent data. The manuscript will be largely enhanced if the author can apply their method in an independent and public accessible gene expression data and validate their discovery. This is feasible, because the authors' model relies largely on gene expression, miRNA and TF binding sites. The copy number and miRNA expression are not necessary to infer the model coefficients.

We would like to clarify that a regression model is learned for each individual patient, and thus the model coefficients cannot be transferred to other patients. In order to validate this result in an independent data set, we retrained our regression models on Murat dataset [1]. It should be noted that this data set contains only gene expression data and no copy number or miRNA expression data. We first classified the patients of the Murat dataset based on the gene signature defined in [2] (see Materials and Methods). We then ran our sample-by-sample regression models and measured

survival difference between patients with high and low miR-132 model coefficients. The results are shown in Supplementary Figure 6b. This shows that patients with high miR-132 coefficients do show a trend towards higher survival, but the result is not statistically significant. As we note in the response to the Editor's comments, we believe the problem is the small sample size ($N = 80$): the survival difference between proneural and non-proneural tumors and between proneural and mesenchymal classes are not significant in this data set either (see Figure R1), although the trend is visible.

> *A minor point is that the authors may want to consider changing Figure 4a, b, Figure 5a and b into GSEA figures.*

We thank the reviewer for the suggestion to use GSEA, but we feel the cumulative distribution plots represent the shift in gene expression change patterns better than the leading edge analysis of GSEA.

References

[1] Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma. Murat A, Migliavacca E, Gorlia T, Lambiv WL, et al. (2008). Journal of Clinical Oncology. June 2010.

[2] Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1. Verhaak R, Hoadley K, et al. Cancer Cell. Jan 2010.

3rd Editorial Decision 12 July 2012

Thank you again for sending us your revised manuscript. We have now heard back from the two reviewers who agreed to evaluate this revised work. After the first round of review, Reviewer #2 was supportive of the publication of this work, and Reviewer #3 agrees that the new analysis with the Murat et al dataset is supportive. While Reviewer #3 has suggestions for further analysis, the editor feels that additional work would lie outside the scope of this manuscript, especially given the substantial supporting evidence that has been provided in the previous revisions.

Moreover, the reviewer feels that the remaining concerns raised by Reviewer #1 seem to result from a technical misunderstanding. The editor feels that this issue can be addressed with some minor final clarifications, and as such I would like to invite to prepare a final revision of this manuscript.

Reviewer #1 has some remaining conceptual concerns regarding the results presented in Fig. 5C, writing, "similarly, In Fig. 5C, the authors describe 3 miRNAs whose predicted targets do not respond to transfection. My interepretation is that these miRNAs are associated with genes that are not regulated by them." In contrast, the editors understanding is that the results in Fig 5C are measuring the concordance between the genes dysregulated after miRNA knockdown with the genes specific to proneural gene expression, not concordance between predicted targets and gene expression shifts. Indeed, Fig. S8 seems to show that all miRNAs are similarly effective at regulating their predicted targets in this assay, helping to rule out a clear bias caused by poor target prediction in the control miRNAs. Going back to Fig. 5C, it is my reading that "upregulated in neurospheres" concordance is fully independent of the computational miRNA target predictions. The wording in the results makes it somewhat unclear whether "downregulated in neurospheres" were filtered specifically for computationally predicted targets. This should be clarified, but I do not think this would affect the results, given the independence of the "upregulated" set, and Figure S8 which seems to address this Reviewer's main technical concern.

Overall, the editor feels that you, the authors, have sufficiently demonstrated that they can reliably identify miRNAs that are specifically relevant to the proneural gene expression signature. Nonetheless given the persistent misunderstandings:

- Please send us a short reply letter confirming that I correctly understood Supp. Fig. 8, and with a brief response clarifying whether the Fig. 5 results were filtered by computational miRNA target predicts.

- Please revise the manuscript text to help avoid similar misunderstandings from future readers. Specifically the Results section should mention that the control miRNAs were also shown to down-regulate their predicted targets with efficiency similar to miR-124 and miR-134, and should specifically state whether the results in Fig. 5C were filtered for predicted targets (and how this might affect the results).

In addition, I would like to acknowledge that this editorial decision took longer than usual, given some delays in receiving the reviewers reports, and the need to consider their points in detail before rendering a decision. I apologize for the delay, but I do feel that it is essential that key points of confusion are resolved before this work is accepted for publication. Every effort will be made to expedite the reevaluation of the final revised work, and assuming that the clarifications provided are satisfactory, this work will not be sent back to the reviewers.

Thank you for submitting this paper to Molecular Systems Biology.

Yours sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Reviewer #1 (Remarks to the Author):

The authors and I have a fundamental disagreements that can not be resolved by pedagogic additions and patchwork. To me, the key issue is how we treat the inferred target sets for regulators, and for miRNAs in particular. This disagreement affects our interpretation of results shown here, and below I'll specifically address results described in Fig. 5A-C.

In Figures 5A-B the authors show that inferred direct and indirect targets for miR-124,132 respond to their upregulation in Proneural cells. Unlike proteins, miRNAs are not known to be activated or inactivated by factors that change their structure. Accurate identification of a miRNA's target set will always lead to accurate prediction of the set of genes that will respond to this miRNA's transfection. Given that, statistically speaking, the authors identified gene sets that respond to the transfection of the two miRNAs suggests that their target prediction methods are better than random selection.

Similarly, In Fig. 5C, the authors describe 3 miRNAs whose predicted targets do not respond to transfection. My interepretation is that these miRNAs are associated with genes that are not regulated by them. If these genes are not true responders to the three differentially expressed miRNAs in Proneural cells, then they should not have been excluded by statistical test in the first place.

To summarize, I believe that miRNAs that are differentially expressed across tumor types will affect their targets. Some miRNAs will play a greater role in the tumor type specific program. However, if predicted targets are not responding, they may be nothing more than false positive predictions. To disprove this, the authors will have to show that tranfection in another context will lead to improved response.

As I wrote before, I believe that the main advance here is target prediction, or regulatory network construction. Not the identification of the drivers of the Proneural cellular program.

Reviewer #3 (Remarks to the Author):

I was a bit disappointed with the author's cross validation. I was expecting that they could "integrate" the miRNA expression into their model. Instead, their "replaced" the miRNA binding site information with miRNA expression data. The miRNA binding site information is obviously very important and should not be excluded from the model. In the supplementary Figure 1c, the miRNA expression data alone confer the model a similar prediction performance as the model build with binding site information.

The response to the independent validation critique is also disappointing. Authors had a marginal validation in a dataset with 80 GBM samples (Murat et al., 2008). It is not clear why the authors did not use the Rembrandt dataset, which is a lot bigger.

3rd Revision - authors' response 17 July 2012

13 July 2012

Dr. Andrew L. Hufton, PhD
Editor
Nature/EMBO Molecular Systems Biology

Dear Dr. Hufton:

Thank you for inviting us to prepare a final revision of our manuscript entitled "Inferring transcriptional and microRNA-mediated regulatory programs in glioblastoma" for *Molecular Systems Biology*.

This note is just to confirm that you have correctly interpreted the results presented in Supplementary Figure 8 and Figure 5, namely:

- Supp. Fig. 8 shows that the computationally predicted target genes of all miRNAs – i.e. the predicted driver miRNAs (miR-124, miR-132), the less confident miRNA from sample-by-sample analysis (miR-433), and control miRNAs (miR-433, miR-380, miR-448) – are significantly downregulated 24 hours after their respective transfections. This result rules out the possibility that our target prediction method was less accurate for the control miRNAs than for the predicted driver miRNAs.

- Fig. 5 indeed examines the concordance of expression changes induced by miRNAs in tumorspheres with tumor versus normal expression changes in proneural tumors. For each tested miRNA, we consider two gene sets: (i) predicted targets of the miRNA that are downregulated in the transfection experiments; and (ii) genes that are upregulated in the transfection experiment. The first set is filtered by computational target prediction, as you surmised, and intended to represent an experimentally derived signature of direct miRNA regulation; the second set is not filtered by target prediction and represents secondary effects of miRNA regulation. For the predicted driver miRNAs, miR-124 and miR-132, we show that these gene sets display opposite expression changes in proneural tumors versus normal samples, i.e. downregulated targets from the transfection are shifted up in expression in tumors, and upregulated genes from the transfection are shifted down in the tumors.

We have added a few sentences to the text (shown with highlight) so that readers will avoid technical misunderstandings. We have also clarified that for the analysis in Figure 5, we filter for predicted targets that are downregulated in the transfection experiments in order to get a signature for direct miRNA regulation. However, as suggested, to remove any dependence on miRNA target prediction in the concordance analysis, we have added analysis where we consider downregulated genes (without filtering for target sites) in each transfection experiment and ask whether these gene sets are upregulated in proneural tumors versus normal samples. Since miRNA transfections lead to very high overexpression of miRNAs and downregulation of a broad set of genes, we examined genes that were most strongly downregulated (FDR corrected $p < 0.001$) in each of the transfection experiments and found strong concordance with proneural tumor expression changes for miR-124 ($p < 0.01$), significant concordance for miR-132 (and for the less confident prediction, miR-433; $p < 0.05$), and no concordance for the controls.

Finally, we note in response to Reviewer #3 that we did not report analysis on the Rembrandt data set because it did not include normal samples; moreover, due to differences in the array platforms used in this study and in TCGA, we could not normalize the two data sets together without introducing batch effects.

We have included the "standfirst" text and highlights in the article file and have also added a thumbnail figure as part of the submission. We hope that we have addressed all concerns in a satisfactory manner.

Sincerely,

Manu Setty and Christina Leslie
Computational Biology Program
Memorial Sloan-Kettering Cancer Center