Reviewer's report

Title: Transcription factor E2F1 promotes EMT by regulating ZEB2 in small cell lung cancer

Version: 0 **Date:** 01 Mar 2017

Reviewer: Robert Cardnell

Reviewer's report:

The authors present data from a large panel of SCLC patients and a number of cell lines to suggest that E2F1 promotes EMT in SCLC via Sip1. The collection of biopsy samples from 60 SCLC patients with sufficient material for assessing the expression of multiple proteins is an impressive feat and has the potential to yield highly valuable data.

While this reviewer doesn't necessarily disagree with the hypothesis that E2F1 can promote EMT is SCLC, the data presented here - in its current form - does not adequately support the hypothesis.

Major concerns:

Figure 1A.

1. The authors report E-cadherin, beta-catenin, claudin-1, vimentin and N-cadherin expression by IHC. From the data presented - IHC from what I presume is a single patient (although this is not stated) and a statement of the proportion of samples expressing each marker - it is hard to properly assess expression of these proteins. The authors should include quantification of the extent of staining (i.e. range and median), and an indication of what score was necessary for a sample to be considered to express a given protein. If such quantification is not possible with the IHC scoring presented, the use of an alternative methodology such as the H-score should be employed (e.g. Tang et al., 2011 Clin. Can. Res). The authors may also wish to consider if the expression of the EMT proteins is bimodal.

2. The authors also state that E-cadherin and beta-catenin expression are highly consistent. It is impossible to asses this based on the data presented. Either the authors must present additional IHC images, or preferably present some statistical measure (e.g. correlation) comparing E-cadherin and beta-catenin expression. The authors should also consider validating their observations using gene expression data from one of the publicly available datasets (e.g. SCLC cell lines - Polley et al 2016, SCLC patient samples - George et al 2016).

3. The IHC images are missing a scale bar, do not indicate magnification and do not state the type of microscopy.

Figure 1B.

1. The authors speculate, but do not test, that vimentin expression is transient. This is therefore, in this reviewers opinion, better addressed in the discussion.

Figure 2A:

1. The shRNA knockdown of E2F1 in H446 cells should be quantified relative to loading control. There is a reasonable amount of E2F1 expression remaining in clone 4 at the longer exposure, especially considering the greater efficiency of the siRNA knockdown shown in Figure 2B. The shRNA knockdown should thus be validated at the mRNA level. The reviewer is also interested to know why H466 and H1688 were chosen as the in vitro models.

2. The cell images are missing a scale bar, do not indicate magnification, do not state how/if the cells were stained and do not state the type of microscopy. The cells in the two images also appear to be at different degrees of confluence and it would be easier to compare with the addition of a nuclear stain.

Figure 2B:

1. The cell images are missing a scale bar, do not indicate magnification, and do not state the type of microscopy. The images are not particularly clear and the control cell line is too confluent for a meaningful assessment of morphology.

Figure 3:

1. While E2F1 does appear to modestly reduce alpha-tubulin and beta-actin expression it is hard to fully assess in the absence of quantification and replicates. The western blots should be repeated with replicates and subsequent quantification.

2. The IF images lack a scale bar and an indication of magnification.

3. Have the authors do not appear to examined the effect of E2F1 loss on the expression of alpha-tubulin or beta-actin beyond H446 cells, a cell line specific effect cannot thus be ruled out.

Figure 4:

1. Please indicate the number of replicates used to generate the error bars and calculate the p-values.
2. Again, by testing in only H446 cells, a cell line specific effect cannot be discounted.

Figure 5A:

1. The authors show a decrease in N-cadherin mRNA expression, however N-cadherin was not detected by IHC in the patient samples. Could the authors comment on this discrepancy.
2. Please indicate the number of replicates, and what p-value * and ** represent.

Figure 5B:

1. Replicates and quantification of the western blots would aid data interpretation.

Figure 5D:

1. Please indicate number of replicates, and fold change in the mRNA data. It is also not clear that there is a meaningful increase in E-cadherin expression in the E2F1si cells from this western blot image, protein expression should be quantified. E2F1 knockdown in this experiment appears to be significantly less effective than in Figure 2B.

How does expression (mRNA) of E-cadherin compare between H446 and H1688 in the parental cell lines? Could the authors also comment on why E-cadherin increases in H1688 but not H446 following E2F1 knockdown?

Figure 6:

1. As with Figure 1, it is hard to gauge the relationship between ZEB1, Sip, Snail and Slug from the IHC images presented. The concerns raised above should also be addressed in this Figure.
2. In the ChIP-seq data, where any other known EMT transcription factors (TF's) identified as targets of E2F1. The authors should include other TF's identified, or state that Sip1 was the only EMT associated TF identified.

Figure 7:

1. Please indicate number of replicates, and fold change in the mRNA data. The western blots for Sip1 are too faint more meaningful interpretation.

2. Please see my comments on IHC images and correlations based thereon for Figure 1.

3. Please indicate the number of replicates, and what p-value * and ** represent for the luciferase assay.

Other:

The IHC data generated has the potential to be highly informative. Quantified expression of each of these markers across the 60 patients is more than adequate powered for statistical analyses. Such analyses (assuming they support the authors assertions) will drastically improve this manuscript and would provide a robust backbone to support the hypothesis.

As Rb loss (which is near universal in SCLC) disregulates E2F1, have the authors considered comparing E2F1 expression to Rb status/expression?

Minor concerns:

1. Gene names (italicized) and protein names are incorrectly formatted and are incorrectly used (e.g. when assaying mRNA "CDH1" is correct whereas "E-cadherin" is not)

2. Where applicable p-values and n-values should be added to the text.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Unable to assess

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons