

# Epstein-Barr virus derived circular RNA LMP2A induces stemness in EBV-associated gastric cancer

Li-ping Gong, Jian-ning Chen, Min Dong, Zhen-dong Xiao, Zhi-ying Feng, Yu-hang Pan, Yu Zhang, Yu Du, Jing-yue Zhang, Yuan-hua Bi, Jun-ting Huang, Jing Liang, and Chun-kui Shao

DOI: [10.15252/embr.201949689](https://doi.org/10.15252/embr.201949689)

Corresponding author(s): Chun Shao ([shaochk@mail.sysu.edu.cn](mailto:shaochk@mail.sysu.edu.cn)) , Li Gong ([1125062839@qq.com](mailto:1125062839@qq.com))

---

## Review Timeline:

Submission Date:	15th Nov 19
Editorial Decision:	14th Jan 20
Revision Received:	10th Apr 20
Editorial Decision:	26th May 20
Revision Received:	4th Jun 20
Editorial Decision:	1st Jul 20
Revision Received:	9th Jul 20
Accepted:	17th Jul 20

---

Editor: Achim Breiling

## 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. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Dear Dr. Shao,

Thank you for the submission of your research manuscript to EMBO reports. We have now received reports from the three referees that were asked to evaluate your study, which can be found at the end of this email.

As you will see, all referees think that the findings are of interest, but they also have several comments, concerns and suggestions, indicating that a major revision of the manuscript is necessary to allow publication in EMBO reports. As the reports are below, and all points need to be addressed, I will not detail them here.

Given the constructive referee comments, we would like to invite you to revise your manuscript with the understanding that all referee concerns must be addressed in the revised manuscript and in a detailed point-by-point response. Acceptance of your manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact me if a 3-months time frame is not sufficient so that we can discuss the revision further.

When submitting your revised manuscript, please also carefully review the instructions that follow below.

PLEASE NOTE THAT upon resubmission revised manuscripts are subjected to an initial quality control prior to exposition to re-review. Upon failure in the initial quality control, the manuscripts are sent back to the authors, which may lead to delays. Frequent reasons for such a failure are the lack of the data availability section (please see below) and the presence of statistics based on  $n=2$  (the authors are then asked to present scatter plots or provide more data points).

When submitting your revised manuscript, we will require:

1) a .docx formatted version of the final manuscript text (including legends for main figures, EV figures and tables), but without the figures included. Please make sure that the changes are highlighted to be clearly visible. Figure legends should be compiled at the end of the manuscript text.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure), of main figures and EV figures. Please upload these as separate, individual files upon re-submission.

The Expanded View format, which will be displayed in the main HTML of the paper in a collapsible format, has replaced the Supplementary information. You can submit up to 5 images as Expanded View. Please follow the nomenclature Figure EV1, Figure EV2 etc. The figure legend for these should be included in the main manuscript document file in a section called Expanded View Figure Legends after the main Figure Legends section. Additional Supplementary material should be supplied as a single pdf labeled Appendix. The Appendix should have page numbers and needs to include a table of content on the first page (with page numbers) and legends for all content. Please

follow the nomenclature Appendix Figure Sx, Appendix Table Sx etc. throughout the text, and also label the figures and tables according to this nomenclature.

For more details please refer to our guide to authors:

<http://www.embopress.org/page/journal/14693178/authorguide#manuscriptpreparation>

See also our guide for figure preparation:

[http://wol-prod-cdn.literatemonline.com/pb-assets/embosite/EMBOPress\\_Figure\\_Guidelines\\_061115-1561436025777.pdf](http://wol-prod-cdn.literatemonline.com/pb-assets/embosite/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf)

3) a .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point responses to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

4) a complete author checklist, which you can download from our author guidelines (<https://www.embopress.org/page/journal/14693178/authorguide>). Please insert page numbers in the checklist to indicate where the requested information can be found in the manuscript. The completed author checklist will also be part of the RPF.

Please also follow our guidelines for the use of living organisms, and the respective reporting guidelines: <http://www.embopress.org/page/journal/14693178/authorguide#livingorganisms>

5) that primary datasets produced in this study (e.g. RNA-seq, ChIP-seq and array data) are deposited in an appropriate public database. See: <http://embor.embopress.org/authorguide#datadeposition>

Please remember to provide a reviewer password if the datasets are not yet public.

The accession numbers and database should be listed in a formal "Data Availability " section (placed after Materials & Methods) that follows the model below. Please note that the Data Availability Section is restricted to new primary data that are part of this study.

# Data availability

The datasets produced in this study are available in the following databases:

- RNA-Seq data: Gene Expression Omnibus GSE46843  
(<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46843>)  
- [data type]: [name of the resource] [accession number/identifier/doi] ([URL or identifiers.org/DATABASE:ACCESSION])

\*\*\* Note - All links should resolve to a page where the data can be accessed. \*\*\*

Moreover, I have these editorial requests:

6) We strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example

scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. If you want to provide source data, please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure.

7) Our journal encourages inclusion of \*data citations in the reference list\* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at: <http://www.embopress.org/page/journal/14693178/authorguide#referencesformat>

8) Regarding data quantification and statistics, can you please specify, where applicable, the number "n" for how many independent experiments (biological replicates) were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. Please provide statistical testing where applicable, and also add a paragraph detailing this to the methods section. See: <http://www.embopress.org/page/journal/14693178/authorguide#statisticalanalysis>

9) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript. Please find instructions on how to link your ORCID ID to your account in our manuscript tracking system in our Author guidelines: <http://www.embopress.org/page/journal/14693178/authorguide#authorshipguidelines>

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

Yours sincerely,

Achim Breiling  
Editor  
EMBO Reports

-----  
Referee #1:

The SNU179 cell line was established in 1996 from a Korean EBV positive gastric carcinoma. It contains EBV and is of epithelial origin. Using this cell line repeatedly for tumor formation in mice in the presence of mutagenic 5-Fluorouracil resulted in a new cell line SNU-4th, which is more tumorigenic and has some markers characteristic of cancer stem cells from carcinomas. In these cells, a cytoplasmic circular RNA derived from some exons of the EBV LMP2 RNA is shown to act as a sponge for miR-3908, resulting in reduced levels of p53 and playing a role in the tumorigenicity of the cells. This circular RNA is found to be associated with metastasis in EBV associated gastric cancer so the novel pathway for p53 regulation offers a potential diagnostic and therapeutic opportunity.

The demonstration of the novel pathway in these SNU-4th cells connecting ebv-circLMP2A to miR-

3908, TRIM59 and p53 is clear and convincing. It fits well with the known lack of p53 mutations in EBV associated gastric cancer and it is plausible that it could provide a novel diagnostic marker.

The main question about the paper is whether cancer stem cells of the type described (CD44+ CD24-, ebv-circLMP2A +, p53 low) actually exist in EBV associated gastric cancer biopsies. The results are a comparison of a few cell lines which have been exposed to very strong selection for growth in culture over long periods of time. A direct demonstration of these cells in primary tumor material would greatly strengthen the paper.

-----  
Referee #2:

The paper entitled Epstein-Barr virus (EBV) derived circLMP2A induces cancer stemness in EBV-associated gastric carcinoma, by Gong et al. describes a comprehensive series of experiments showing that the circRNA named ebv-circLMP2A is involved in inducing and maintaining stemness phenotypes of EBV-associated gastric carcinoma (EBVaGC) through targeting the miR 3908/TRIM59/p53 axis. In addition, this paper reveals that ebv-circLMP2A can predict poor prognosis of EBVaGC patients. Overall, I think the study is well-designed and the data well-presented. I believe that the authors provide sufficient evidence to support most of their conclusions. However, some points have to be addressed before publication:

#### Major revisions

1. For the overexpression experiments the authors do not show that the product from the vector is actually circular. To verify this, Northern blotting +/- RNase R treatment should be performed. This experiment will also provide information about whether the vector produces concatemers. Also, the overexpression is very efficient. The authors should comment on whether the levels can be considered physiologically relevant.
2. In fig. 3g it can be observed that patients generally survive for a long time. Therefore, overall survival may not be the best clinical parameter to assess, as patients may die from other causes. Please consider also to include disease specific survival, time to progression or other relevant measures.
3. How did the authors select the eight genes assessed in figure 4c and 5c? In figure 1, thirteen genes were analyzed, including Twist1, Snail, Slug, Zeb1 and Mmp7, which were not analyzed in fig 4c and 5c. For consistency, I suggest that these genes are analyzed and included in fig 4c and 5c.
4. The authors use GAPDH for normalization of all RT-qPCR assays. Using only a single reference gene for RT-qPCR is not advisable (1). Also, GAPDH might be a poor choice for normalizing the data as it has previously been shown to enhance the aggressiveness and the vascularization of tumors (2) and the transcriptional levels of GAPDH are highly up-regulated in some cancers (3,4). Therefore, the authors need to provide evidence that this gene is stably expressed across the samples and not differentially expressed between groups that are compared.
5. The authors analyzed all RT-qPCR data with the  $\Delta\Delta CT$  method (5), without assessing PCR efficiencies of the assays. This quantification strategy is only valid when PCR efficiencies are approximately equal between genes of interest and reference genes (1). Therefore, the authors should assess the PCR efficiency of their RT-qPCR assays.

6. Experimental details are lacking for the RNA-seq experiments (e.g. what library preparation kit was used?).

7. According to mirBase ([http://www.mirbase.org/cgi-bin/mirna\\_entry.pl?acc=MI0016412](http://www.mirbase.org/cgi-bin/mirna_entry.pl?acc=MI0016412)), miR-3908 is supported by very few sequencing reads and no 3p arm has been annotated. Therefore, I suggest that the authors provide additional experimental support that this is actually a real miRNA through AGO-CLIP or Northern Blotting.

#### Minor revisions

1. Has ebv-circLMP2A been discovered before (e.g. is it present in any of the circRNA databases available). If so, please provide an identification number for this circRNA.

2. In fig. 1G, the authors show that ebv-circLMP2A is associated with poor patient outcome. It would be interesting to also analyze whether the amount of ebv-circLMP2A relative to LMP2A shows the same. E.g. whether this is an independent effect of the circRNA and not due to the amount of virus overall.

3. In fig 6j, why are there T's and not U's present in the sequences for the circRNAs?

4. The authors need to assess/discuss that the changes in expression levels of a single circRNA generally will not lead to significant changes in competing miRNA binding sites relative to all of the corresponding miRNA binding sites present in all mRNAs (6). Moreover, endogenous stoichiometric relations between the miRNA-binding sites of ebv-circLMP2A and the corresponding mRNA target sites of the miRNA may not be mirrored in the overexpression experiments, which therefore may lack physiological relevance (7). Alternatively, the circRNA may not function as a sponge but through target RNA-directed miRNA degradation (TDMD). However, this usually requires extensive complementarity outside of the seed sequence (8,9) as exemplified by the long non-coding RNA, Cyrano, which induce destruction of miR-7 through a single highly conserved binding site of unusually high complementarity to miR-7 (10). Please discuss this in relation to the findings that it is binding site 1 and 3, but not site 2 that were critical for ebv-circLMP2A to sponge miR-3908.

5. When using Student's t-tests, did the authors make sure that the data followed a normal distribution?

6. In the discussion section, the authors should discuss why ebv-circLMP2A seems to have a much larger effect on patient survival than miR-3908.

7. Throughout there are sentences that can be difficult to read and grammatical errors (e.g. "...but also responsible for tumor progression, metastasis and therapy- resistant", "...through long-term treatment of EBVaGC cell line SNU719 with 5-Fluorouraci in vivo passage" and "...shedding a light on the pathogenic function of ebv-circRNAs"). These sentences are from the abstract alone, but please read the entire manuscript carefully to address this.

#### References for this review

1. Bustin, S.A., et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry* 55, 611-622 (2009).
2. Chiche, J., et al. GAPDH enhances the aggressiveness and the vascularization of non-Hodgkin's B lymphomas via NF-kappaB-dependent induction of HIF-1alpha. *Leukemia* 29, 1163-1176 (2015).
3. Wang, D., Moothart, D.R., Lowy, D.R. & Qian, X. The expression of glyceraldehyde-3-phosphate

dehydrogenase associated cell cycle (GACC) genes correlates with cancer stage and poor survival in patients with solid tumors. PloS one 8, e61262 (2013).

4. Schek, N., Hall, B.L. & Finn, O.J. Increased glyceraldehyde-3-phosphate dehydrogenase gene expression in human pancreatic adenocarcinoma. Cancer research 48, 6354-6359 (1988).
5. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif.) 25, 402-408 (2001).
6. Denzler, R., Agarwal, V., Stefano, J., Bartel, D.P. & Stoffel, M. Assessing the ceRNA hypothesis with quantitative measurements of miRNA and target abundance. Molecular cell 54, 766-776 (2014).
7. Thomson, D.W. & Dinger, M.E. Endogenous microRNA sponges: evidence and controversy. Nature reviews. Genetics 17, 272-283 (2016).
8. Ghini, F., et al. Endogenous transcripts control miRNA levels and activity in mammalian cells by target-directed miRNA degradation. Nature communications 9, 3119 (2018).
9. Sheu-Gruttadauria, J., et al. Structural Basis for Target-Directed MicroRNA Degradation. Molecular cell 75, 1243-1255.e1247 (2019).
10. Kleaveland, B., Shi, C.Y., Stefano, J. & Bartel, D.P. A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain. Cell 174, 350-362.e317 (2018).

-----  
Referee #3:

This report ascribes the initiation and maintenance of EBV-associated gastric carcinomas to an EBV circular RNA derived from the LMP2A locus. The authors present studies to link the expression of circLMP2A to the development of phenotypic stem cell properties in infected cells. They propose that circLMP2A acts as a microRNA sponge to mechanistically effect this stemness by targeting the miR-3908/TRIM59/p53 axis. Studies on the functional biology of circular RNAs are rapidly accumulating and these non-coding RNAs represent an exciting, novel class of regulatory molecules in the cell. For this manuscript, however, some major points need to be addressed:

(pages are assigned by this reviewer for ease of communication starting with 1 for the title page - it is helpful for review if the authors can number the manuscript lines)

1. Figure 1: With the results that are presented in Figure 3, one would expect that cells from the third xenograft shown in 1B would also have relatively high ebv-circRNA expression levels similar to the fourth xenograft tested in 3B. Was this tested retrospectively?
2. Figure 2: Panel 2C is confusing, how can there be more than 1 backsplice junction in a circRNA? Similarly, please explain what the following text means, it is not clear: "a total of 262 distinct ebv-circRNAs candidates were found in these tissues and 144 of which contained at least two backspliced reads". In Panel D, how can the "length of ebv-circRNA" be determined without detailed identification, cloning, and sequencing of each circRNA species. Panel E looks like it wants to be a Venn diagram, except none of the areas overlap. Perhaps this should be expressed as a bar graph. Please also define "intergenic", "introns", "exon-intergenic" and "exons-intron" and how these were determined because exon-intergenic, intergenic, and intron backsplicing would be unusual. Panel F seems to say that there are at least 10 ebv genes that produce more than 1 circular RNA "isotype". What do the authors mean to show here?
3. Figure 3: The manuscript is highly focused on one particular circular "isoform" from the LMP2A

gene locus: a 429 bp circular RNA with an exon5-to-exon3 backsplice junction. In Ungerleider N et al's 2018 paper (reference #26) low numbers of an LMP2A exon5-to-exon4 backsplice junction is detected in YCCEL1 cells (which is used extensively in this study as well), but no exon5-to-exon3 circulars. Were exon5-to-exon4 circLMP2A found or looked for using divergent primers? Panel 3C is really problematic, there is not a lot of difference between Rnase R treated and untreated samples for ebv-circLMP2A by PCR in each of the three cell types looked at. This suggests that the molecule being assayed for is not a true circular RNA. The statement in the text on page 8, ".....whereas ebv-circLMP2A was resistant to RNase R digestion" is not supported by the data. Further, Panel 3B indicates ebv-cirLMP2A is highly expressed in SNU-4th cells however, the PCR gel results from Panel 3C is not consistent with high levels of ebv-circLMP2A in SNU-4th compared to SNU719 or YCCEL1. Panel E should have a control showing the efficiency of fractionation into cytoplasmic and nuclear samples.

Table1: This reviewer has serious concerns about the ability to detect ebv-circLMP2A from paraffin embedded samples (page16) particularly in samples of paraffin blocks from 2006. Table 1 indicates that the ebv-circLMP2A is found in every block and can be quantified to be low or high - this is not the experience in the field. Please also test for a cellular control circular RNA in each case.

4. Figure 4: For these series of experiments, where possible (for example 4C, 4J, 4L), the ebv-circLMP2A should also be quantitatively assay for by PCR in sh-control and sh-ebv-circLMP2A samples.

5. Figure5: For these sets of experiments it is critical that the ebv-circLMP2A over-expression system be confirmed. The circular molecule should be validated throughout its entire 429bp length to be intact by sequencing of PCR overlapping products. The sequence of the circularization junction needs to be particularly confirmed and stated. RNaseR treatment should be applied to confirm circularity of expression product.

6. Figure6: Please describe the oligo probe used as the control for the circLMP2A probe. Is the oligo control sequence the same length as the circLMP2A probe, and of an identical nucleic acid composition?

#### Other Comments

1. Page 4, bottom para: " In 2018, our study and two other reports have proven....." It is an error to attribute the authors' previous study (reference #25) as a 2018 study.

2. Abstract: the EBV-circLAMP2A is described as "highly expressed". This is not accurate based on the PCR band strength and lack of quantitative assays.

3. Introduction: Authors spend a lot of verbiage on the uncertainty of the role of EBV to the pathogenesis of EBVaGC. Actually, there is not a lot of uncertainty about this, particularly when EBV is clonal by Terminal Repeat analysis in every EBVaGC tumor cell.



Dear Editor and Reviewers:

Thank you for reviewing and providing helpful comments on our manuscript entitled “Epstein-Barr virus (EBV) derived circLMP2A induces cancer stemness in EBV-associated gastric carcinoma” (ID: EMBOR-2019-49689V1).

We have studied the comments carefully and have made revision to the original article seriously. In our responses, questions from the editors and reviewers are marked in *italics*, and our point-by-point responses are in black font. Additionally, we uploaded a clean version of the manuscript as requested and revised portion are marked in red in the revised manuscript. The answers to the comments are listed as follows.

**Reply to reviewer 1:**

Reviewer 1:

*The SNU179 cell line was established in 1996 from a Korean EBV positive gastric carcinoma. It contains EBV and is of epithelial origin. Using this cell line repeatedly for tumor formation in mice in the presence of mutagenic 5-Fluorouracil resulted in a new cell line SNU-4th, which is more tumorigenic and has some markers characteristic of cancer stem cells from carcinomas. In these cells, a cytoplasmic circular RNA derived from some exons of the EBV LMP2 RNA is shown to act as a sponge for miR-3908, resulting in reduced levels of p53 and playing a role in the tumorigenicity of the cells. This circular RNA is found to be associated with metastasis in EBV associated gastric cancer so the novel pathway for p53 regulation offers a potential diagnostic and therapeutic opportunity.*

*The demonstration of the novel pathway in these SNU-4th cells connecting ebv-circLMP2A to miR-3908, TRIM59 and p53 is clear and convincing. It fits well with the known lack of p53 mutations in EBV associated gastric cancer and it is plausible that it could provide a novel diagnostic marker.*

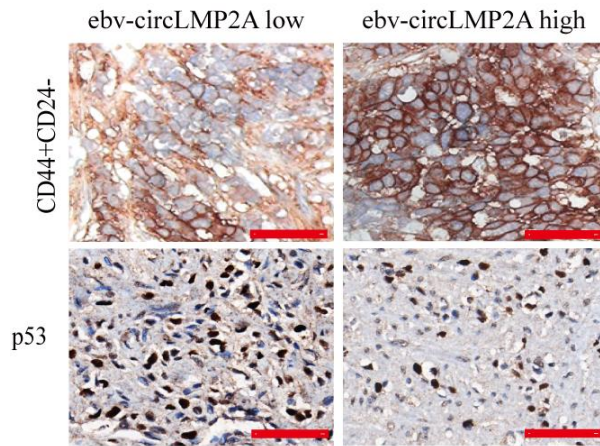
*The main question about the paper is whether cancer stem cells of the type described (CD44+ CD24-, ebv-circLMP2A +, p53 low) actually exist in EBV associated gastric*

---

*cancer biopsies. The results are a comparison of a few cell lines which have been exposed to very strong selection for growth in culture over long periods of time. A direct demonstration of these cells in primary tumor material would greatly strengthen the paper.*

**RE:** We sincerely thank the reviewer for this important comment. According to your suggestion, we studied the expression of CD44/CD24 and p53 in 69 paraffin-embedded EBV associated gastric cancer (EBVaGC) samples by immunohistochemistry and the expression of ebv-circLMP2A have been previously detected by real-time PCR in this original manuscript (Table1). According to the CD44/CD24 double immunohistochemical staining results, CD44 was expressed in cell membrane and stained brown, CD24 was located in cell membrane and cytoplasm and stained dark blue. CD44+ CD24- cells showed membranous brown staining. P53 was expressed in nucleus and stained brown. The following weighted scoring method was adopted to quantify the CD44+ CD24- and p53 expression. A mean percentage of positive tumor cells were determined in at least 5 randomly high-power microscopic fields and assigned to one of the following categories: (a) 0,  $\leq 5\%$ ; (b) 1, 6-25%; (c) 2, 26-50%; (d) 3,  $>50\%$ ; the intensity of CD44+ CD24- and p53 immunostaining was scored as follows: (a) weak, 1; (b) moderate, 2; (c) intense, 3. The percentage of positive tumor cells and the staining intensity scores were multiplied to generate a weighted score for each case: 0-1, (-); 2-3, (+); 4-6, (++);  $>6$ , (+++). Next, we further analyzed the correlation between the ebv-circLMP2A expression and CD44+ CD24- as well as p53 expression. And we observed that in ebv-circLMP2A high expression EBVaGC specimens, the CD44+ CD24- cells were significantly increased and the p53+ cells were significantly decreased (Fig. 8L, Table EV3).

The new data have been added in the Result section in the revised manuscript (see page 12, line 329 - 332).



**Fig. 8L** The expression of CD44+ CD24- and p53 were measured using IHC in 69 EBVaGC patient samples

**Table EV3.** CD44+ CD24- and p53 expression in tumor cells in EBVaGC

ebv-circLMP2A expression	cases	CD44+ CD24-*				P53**			
		-	+	++	+++	-	+	++	+++
Low	43	17	15	7	4	5	12	18	8
High	26	4	5	13	4	9	13	1	3

NOTE. *P*-values were obtained from Pearson Chi-Square tests.

\* $\chi^2 = 8.314$ , *P* = 0.04

\*\* $\chi^2 = 15.413$ , *P* = 0.001

### Replies to reviewer 2:

Reviewer 2:

*The paper entitled Epstein-Barr virus (EBV) derived circLMP2A induces cancer stemness in EBV-associated gastric carcinoma, by Gong et al. describes a comprehensive series of experiments showing that the circRNA named ebv-circLMP2A is involved in inducing and maintaining stemness phenotypes of EBV-associated gastric carcinoma (EBVaGC) through targeting the miR 3908/TRIM59/p53 axis. In addition, this paper reveals that ebv-circLMP2A can predict poor prognosis of EBVaGC patients. Overall, I think the study is well-designed*

---

*and the data well-presented. I believe that the authors provide sufficient evidence to support most of their conclusions. However, some points have to be addressed before publication:*

***Major revisions***

*1. For the overexpression experiments the authors do not show that the product from the vector is actually circular. To verify this, Northern blotting +/- RNase R treatment should be performed. This experiment will also provide information about whether the vector produces concatemers. Also, the overexpression is very efficient. The authors should comment on whether the levels can be considered physiologically relevant.*

**RE:** Thanks for your critical comment, we totally agree with your opinion. Northern blotting +/- RNase R treatment has been performed to confirm the product from the over-expression system is actually circular RNA (Fig. EV2D). We also performed Sanger sequencing to confirm the distinct circularization junction sequence of ebv-circLMP2A produced from the over-expression system (Fig. EV2E). The new data have been added in the Result section in the revised manuscript (See page 9, line 241-page 10, line 244).

Besides, your concern of whether the overexpression levels can be considered physiologically relevant is very thoughtful. Indeed, the expression level of ebv-circLMP2A in parental SNU719 cells and YCCEL1 cells is very low, but it is significantly higher in our enriched EBVaGC cancer stem cells (CSCs) SNU-4th cells. In our study, silencing of ebv-circLMP2A inhibited the stemness phenotypes of SNU-4th cells *in vitro* and *vivo* experiments. As we know, CSCs represent a unique subpopulation of cells within the tumors[1]. Although the number of CSCs is very small, they are bestowed with the capacity to self-renew and are regarded as the bad seed of tumor development and resistance[2]. So, overexpression of ebv-circLMP2A in SNU719 cells and YCCEL1 cells can help us artificially simulate the CSCs in EBVaGC and the physiological role of EBVaGC CSCs in the development of EBVaGC needs to be verified in further studies.

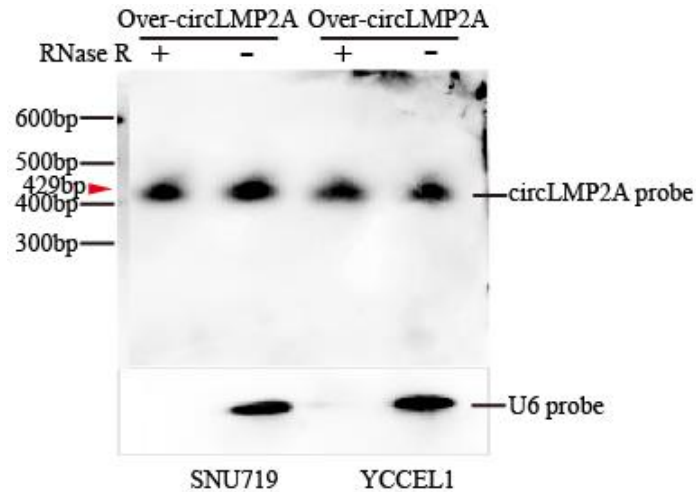


Fig. EV2D Northern blots for detecting ebv-circLMP2A in SNU719 and YCCEL1 cells (ebv-circLMP2A over-expression stable transfectants) treated with or without RNase R digestion

ebv-circLMP2A\_E5\_E3(429bp) sequence:

**TCTTTGCA**AATTTGCCTAACATGGAGGATTGAGGACCCACCTTTTAATTCTCT  
TCTGTTTGCATTGCTGGCCGCAGCTGGCGGACTACAAGGCATTTACGTTCT  
GGTGATGCTTGTGCTCCTGATACTAGCGTACAGAAGGAGATGGCGCCGTTT  
GACTGTTTGTGGCGGCATCATGTTTTTGGCATGTGTACTTGTCCATCATCGTC  
GACGCTGTTTTGCAGCTGAGTCCCCTCCTTGGAGCTGTAACGTGGTTTTCC  
ATGACGCTGCTGCTACTGGCTTTCGTCCTCTGGCTCTCTTCGCCAGGGGGC  
CTAGGTA CTCTTGGTGCAGCCCTTTTAACATTGGCAGCAGCTCTGGCACTG  
CTAGCGTCACTGATTTTGGGCACACTTAACTTGACTACAATGTTCCCTTCTCA  
TGCTCC**TATGGACACTTG**

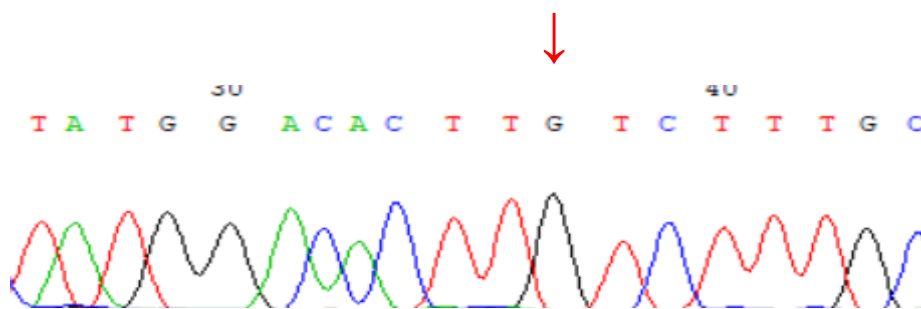


Fig. EV2E The backsplice sequence of ebv-circLMP2A which from the over-expression system was validated by Sanger sequencing. Red arrow represents the “head-to-tail” splicing sites of ebv-circLMP2A.

2. In fig. 3g it can be observed that patients generally survive for a long time. Therefore, overall survival may not be the best clinical parameter to assess, as

---

*patients may die from other causes. Please consider also to include disease specific survival, time to progression or other relevant measures.*

**RE:** We sincerely thank the reviewer for this important comment. In fact, when we contacted the patients or their families, only the patients' condition and time of death were available. Due to the regional cultural differences in China, many patients' families are reluctant to talk about the details of patients, especially those who have died, so only a few patients have data on the cause of death. Additionally, overall survival (OS) is a good indicator of cancer survival, and many clinical trials use OS as the study endpoint[3]. Therefore, since we did not acquire the cause of patients' death, we were unable to perform disease-free survival (DFS) analysis. However, as your suggested, for advanced patients, we have performed progression-free survival (PFS) and found that higher expression of ebv-circLMP2A also predicts poor PFS than those with low ebv-circLMP2A expression (15.4% vs 72.6%,  $p < 0.001$ , Fig.3H). The new data have been added in the Result section in the revised manuscript (see page8, line 201- 203).

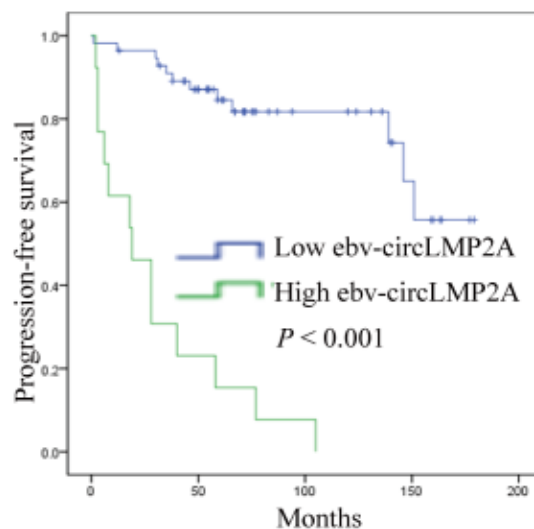


Fig.3H Kaplan-Meier survival curves analysis of the correlation between ebv-circLMP2A expression and PFS.

3. *How did the authors select the eight genes assessed in figure 4c and 5c? In figure 1, thirteen genes were analyzed, including Twist1, Snail, Slug, Zeb1 and Mmp7, which were not analyzed in fig 4c and 5c. For consistency, I suggest that these genes are*

analyzed and included in fig 4c and 5c.

**RE:** Thanks for your critical suggestion. There are many markers of stemness phenotype, so we selected the representative EMT-related markers (E-cad, vimentin), stemness markers (Sox2, Klf4, Bmi1, Oct4) and drug resistance genes (ABCG2, Mrp1) for testing. Per your advice, we have added the expression level of EMT associated transcription factors (Twist1, Snail, Slug, Zeb1) and matrix metalloproteinases7 (Mmp7) in Fig 4C and 5C. The new data have been added in the Result section (see Fig 4C and 5C in the revised manuscript).

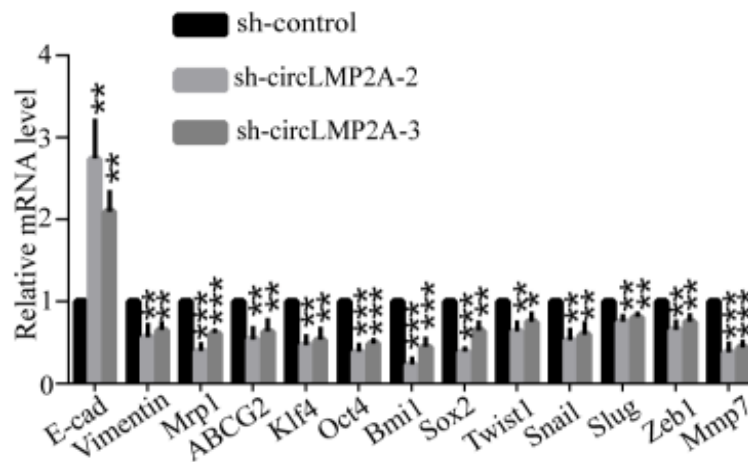


Fig 4C EMT-related markers and transcription factors, stemness markers, drug resistance genes and matrix metalloproteinases7 were evaluated by real-time PCR in SNU-4th cells transfected with sh-circLMP2A-2, sh-circLMP2A-3 or sh-control.

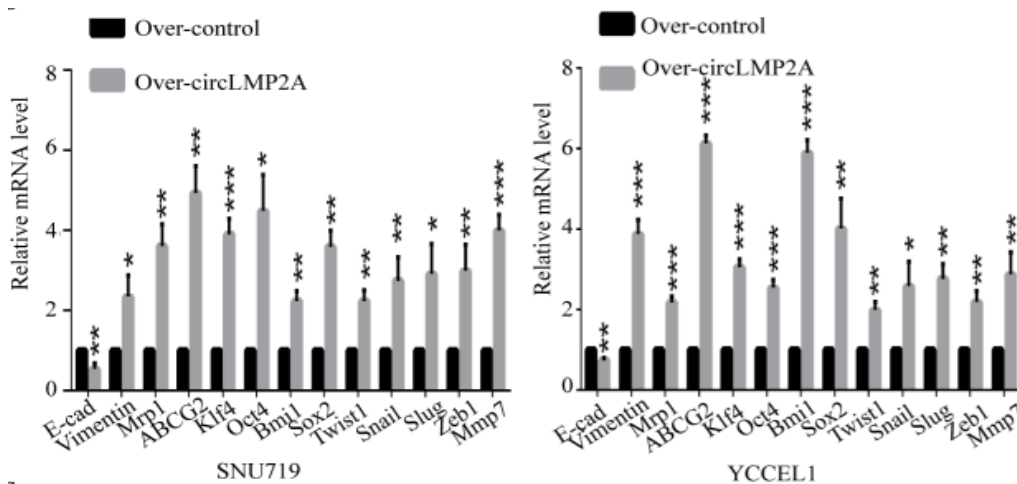
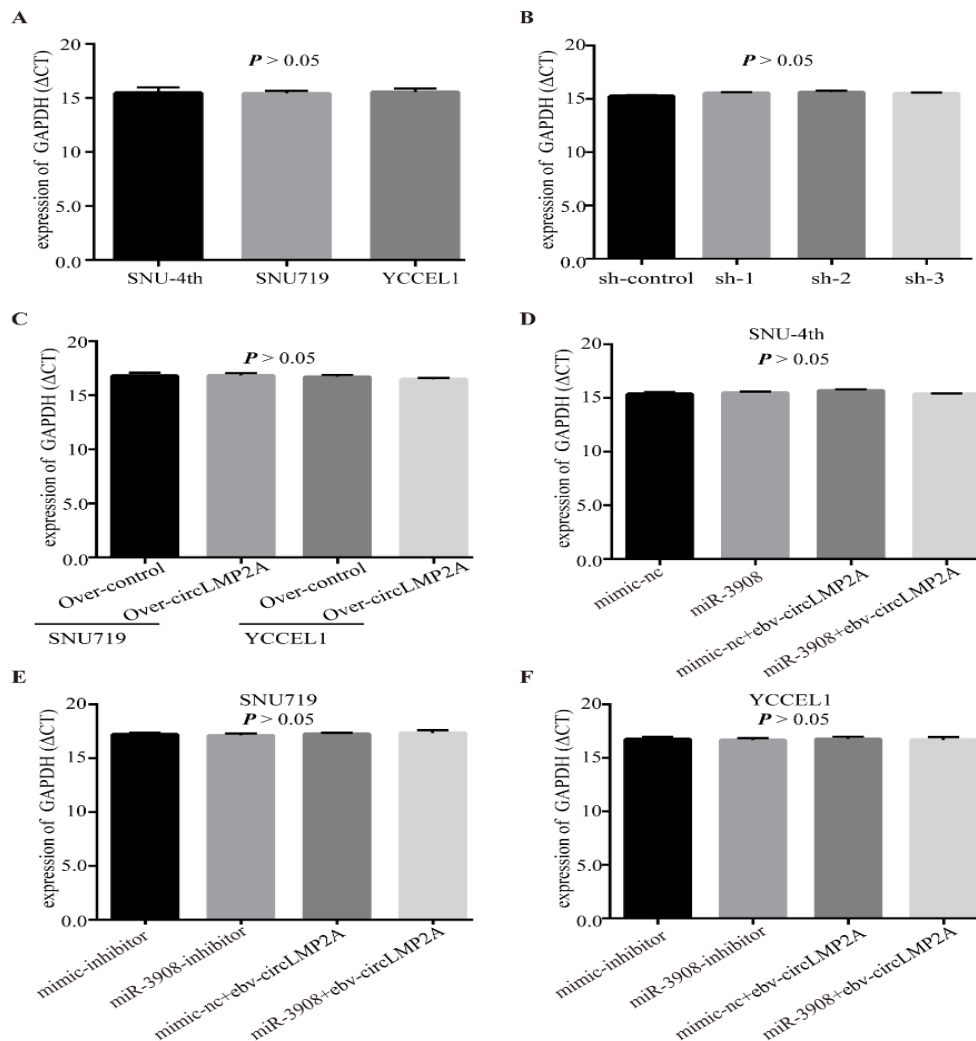


Fig 5C EMT-related markers and transcription factors, stemness markers, drug resistance genes and matrix metalloproteinases7 were evaluated by real-time PCR in SNU719 and YCCEL1 cells transfected with ebv-circLMP2A or vector.

4. The authors use GAPDH for normalization of all RT-qPCR assays. Using only a single reference gene for RT-qPCR is not advisable (1). Also, GAPDH might be a poor choice for normalizing the data as it has previously been shown to enhance the aggressiveness and the vascularization of tumors (2) and the transcriptional levels of GAPDH are highly up-regulated in some cancers (3,4). Therefore, the authors need to provide evidence that this gene is stably expressed across the samples and not differentially expressed between groups that are compared.

**RE:** Thanks for your thoughtful comments. According to your suggestion, we have re-analyzed the expression level of GAPDH between groups that are compared with RT-qPCR assay in our study and found that GAPDH is stably expressed across the samples and not differentially expressed between all groups that are compared (Fig. R1).

Fig. R1 The expression level of GAPDH in all groups that are compared



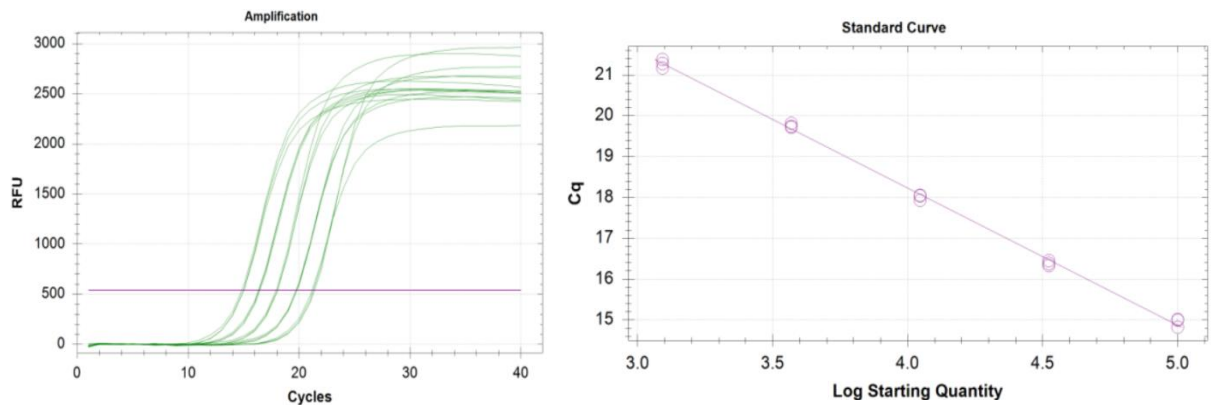


5. The authors analyzed all RT-qPCR data with the  $\Delta\Delta CT$  method (5), without assessing PCR efficiencies of the assays. This quantification strategy is only valid when PCR efficiencies are approximately equal between genes of interest and reference genes (1). Therefore, the authors should assess the PCR efficiency of their RT-qPCR assays

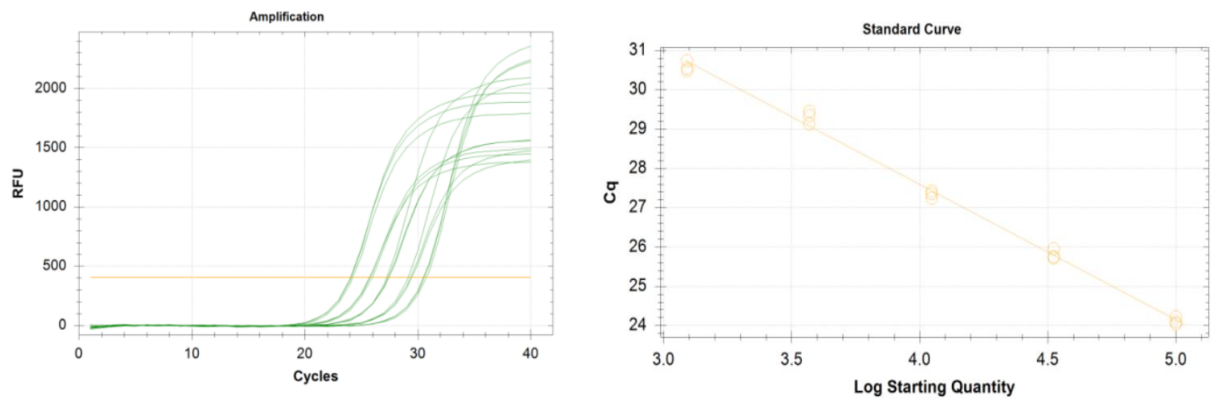
**RE:** Thanks for your thoughtful comment. According to your suggestion, we have performed the calibration curves of reference gene (GAPDH) and thirteen genes of interest to assess the PCR amplification efficiency. According to the CT value of the target genes, we successively diluted the cDNA sample three times and set five concentration gradients, then detected the expression changes of the target genes by RT-qPCR. We found that the amplification efficiency of the reference gene and thirteen genes were close to 100% (Fig. R2), suggesting the quantification strategy of RT-qPCR data is reliable.

Fig. R2 The calibration curves of the reference gene and thirteen genes of interest

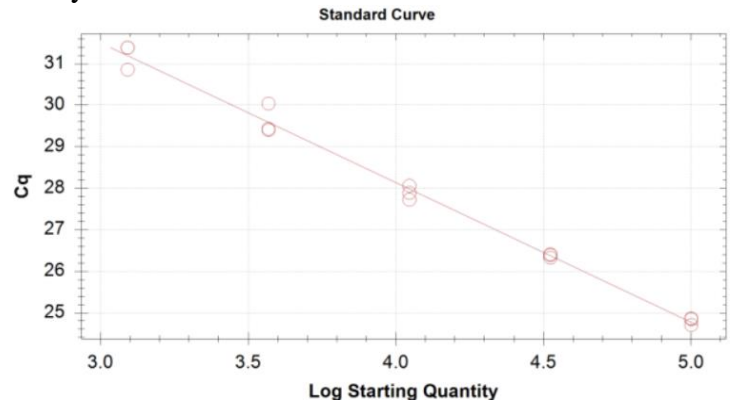
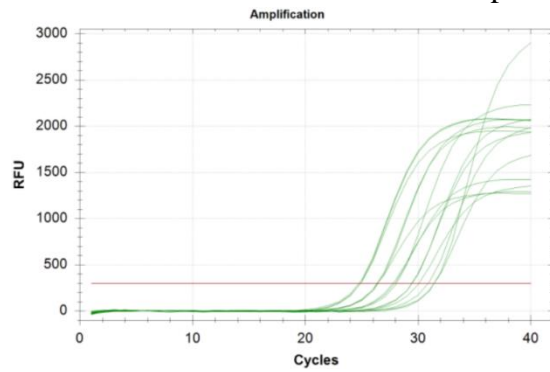
1. GAPDH E=98.5% R<sup>2</sup>=0.998 Slope= -3.359 y-int=31.666



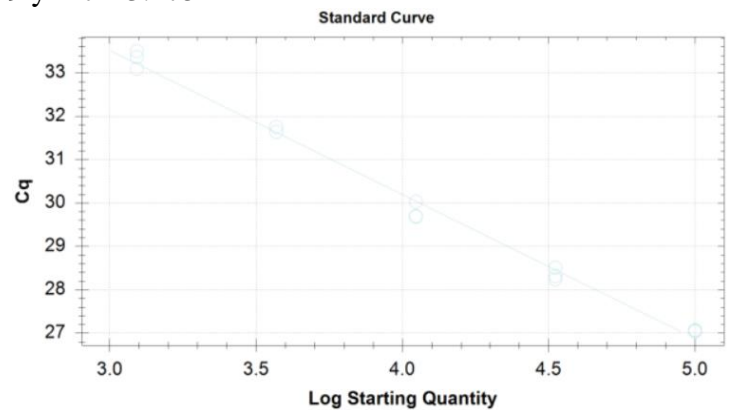
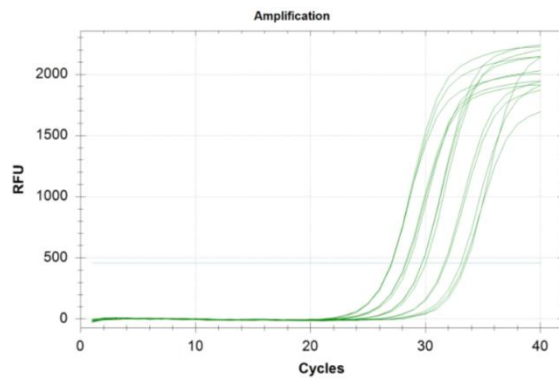
2. E-cad E=95.1% R<sup>2</sup>=0.995 Slope= -3.446 y-int=41.378



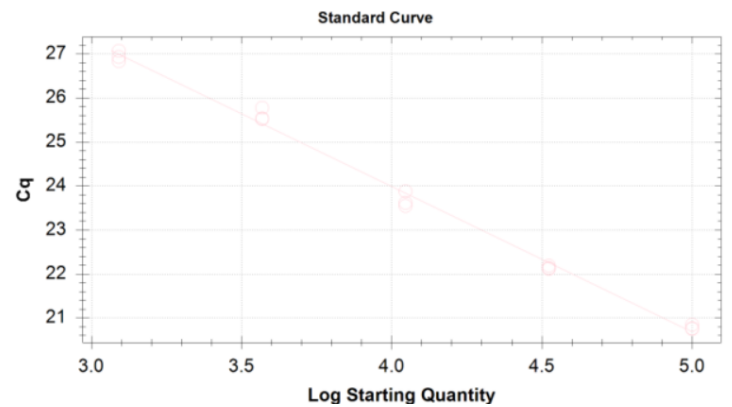
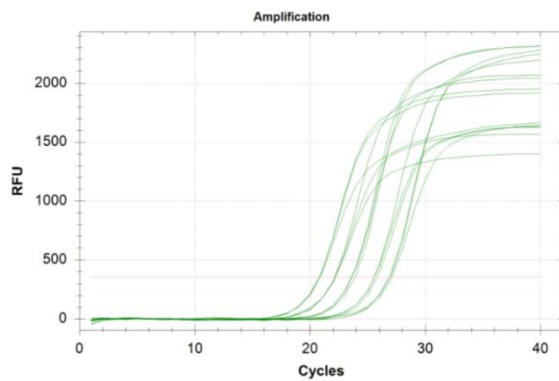
3. Vimentin E=98.2% R<sup>2</sup>=0.993 Slope= -3.366 y-int=41.598



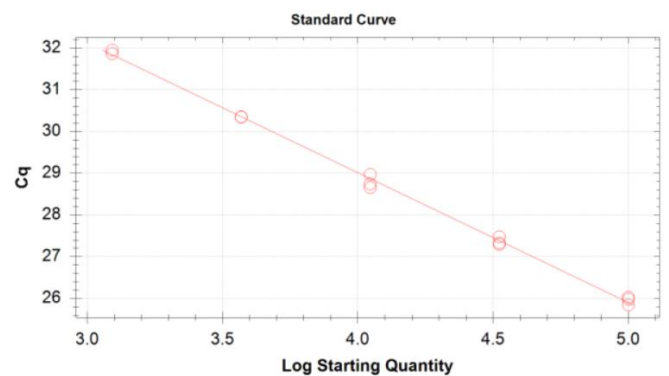
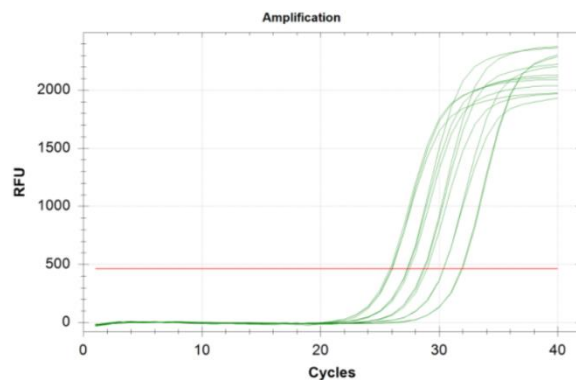
4. Mrp1 E=100.1% R<sup>2</sup>=0.992 Slope= -3.319 y-int=43.475



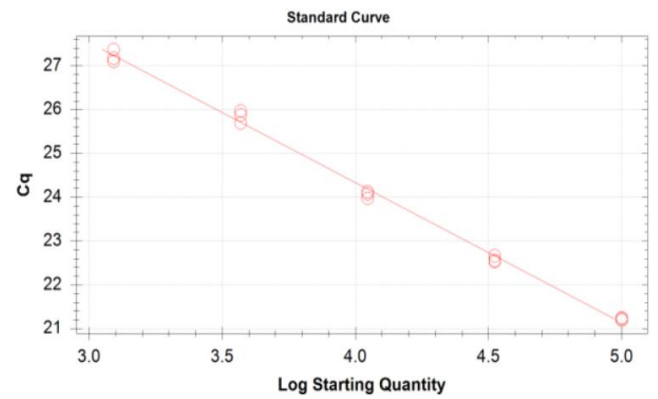
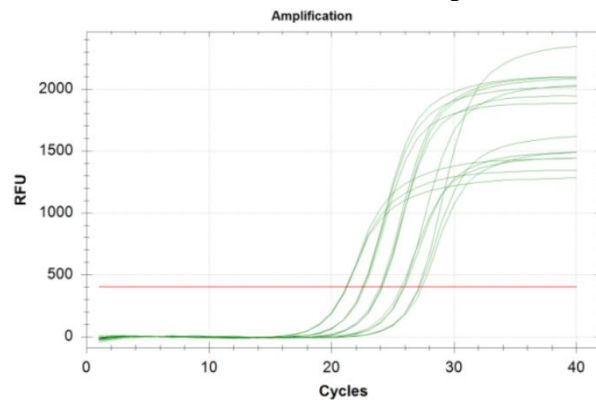
5. ABCG2 E=100.6% R<sup>2</sup>=0.995 Slope= -3.307 y-int=37.218



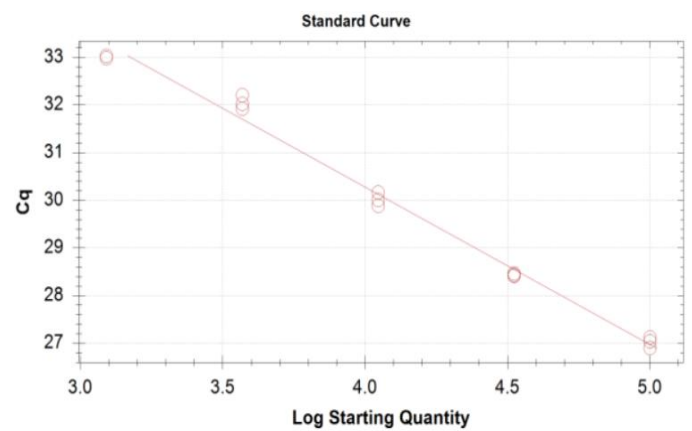
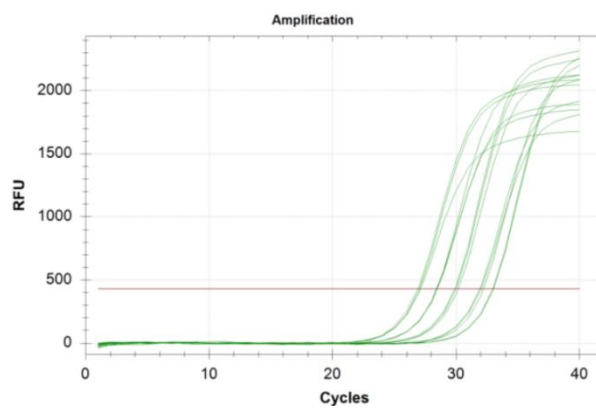
6. Klf4 E=109.5% R<sup>2</sup>=0.998 Slope= -3.113 y-int=41.465



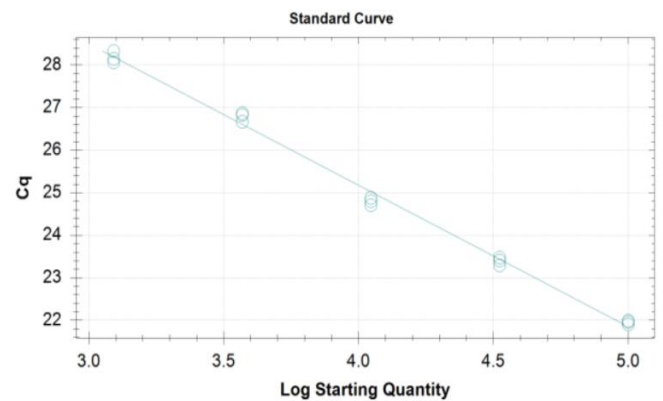
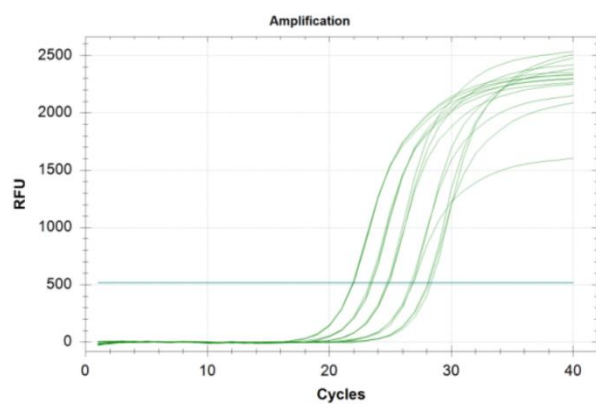
7. Oct4 E=105.5% R<sup>2</sup>=0.996 Slope= -3.197 y-int=37.121



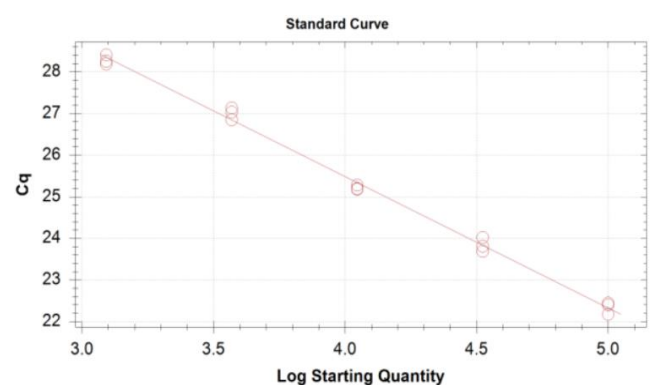
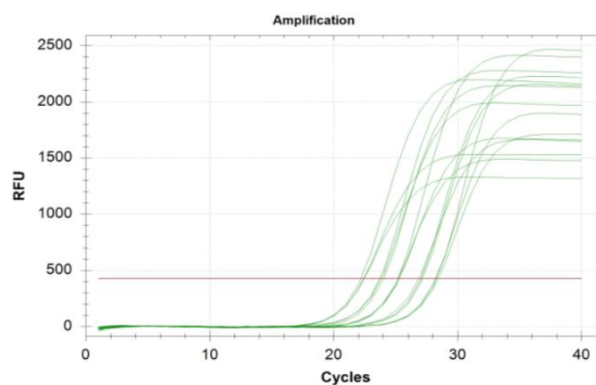
8. Bmi1 E=100.7% R<sup>2</sup>=0.989 Slope= -3.306 y-int=43.503



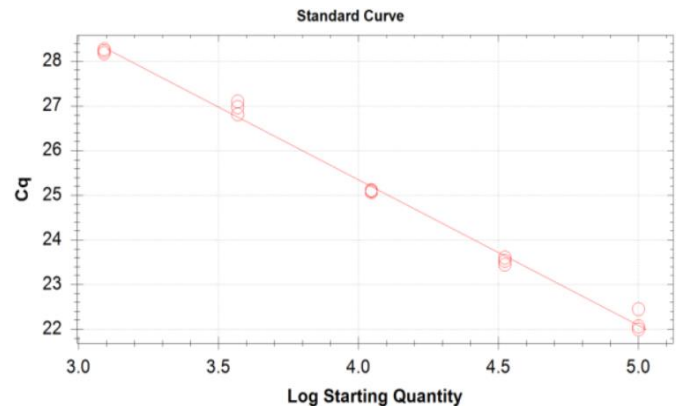
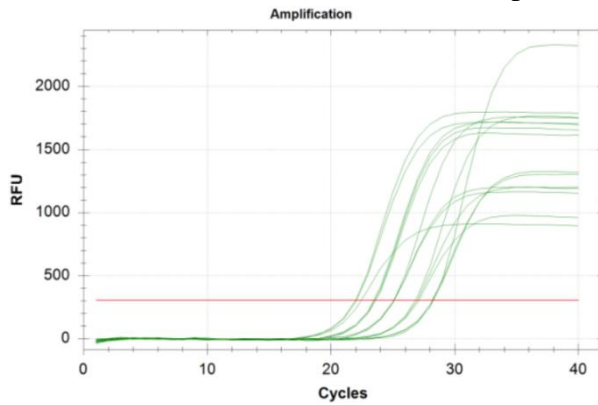
9. Sox2 E=100% R<sup>2</sup>=0.995 Slope= -3.322 y-int=38.463



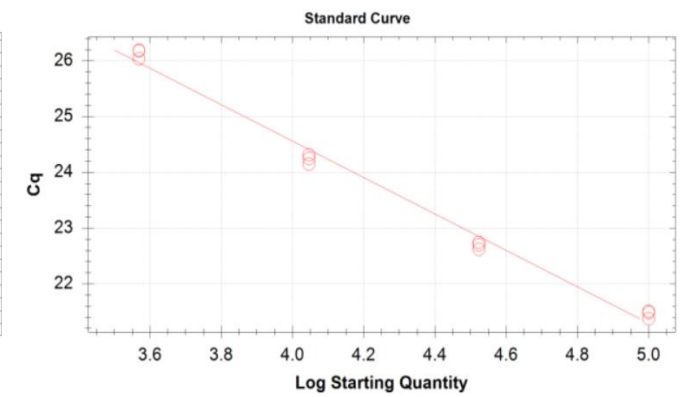
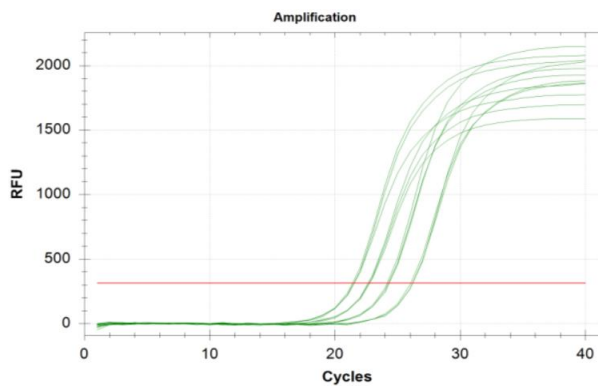
10. Twist1 E=107.4% R<sup>2</sup>=0.996 Slope= -3.156 y-int=38.108



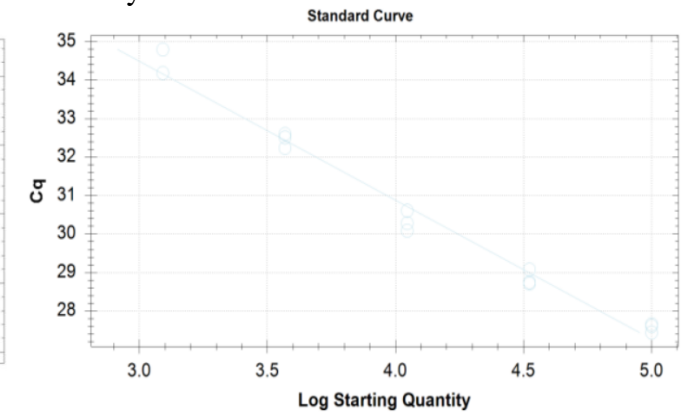
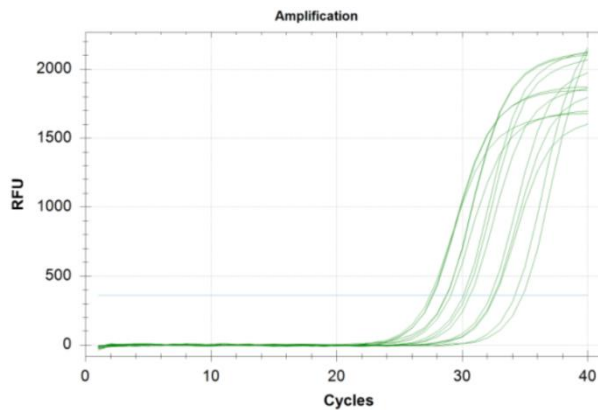
11. TRIM59 E=102.6% R<sup>2</sup>=0.994 Slope= -3.261 y-int=38.396



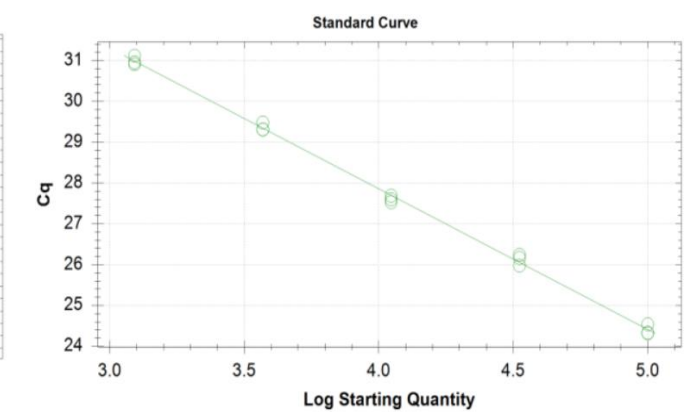
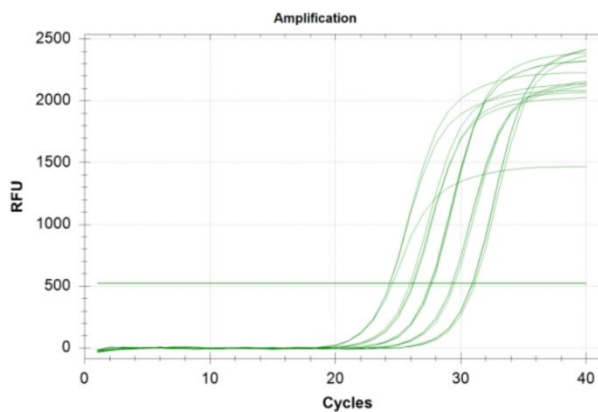
12. P53 E=102.4% R<sup>2</sup>=0.990 Slope= -3.266 y-int=37.623



13. LMP2A mRNA E=89.2% R<sup>2</sup>=0.981 Slope= -3.612 y-int=45.328



14. ebv-circLMP2A E=95.1% R<sup>2</sup>=0.998 Slope= -3.445 y-int=41.638



6. *Experimental details are lacking for the RNA-seq experiments (e.g. what library preparation kit was used?)*

**RE:** Thanks for your critical comment. We have added this information to the revised Materials and Methods section of the manuscript (page 20, line 501-511).

7. *According to miRbase ([http://www.mirbase.org/cgi-bin/mirna\\_entry.pl?acc=MI0016412](http://www.mirbase.org/cgi-bin/mirna_entry.pl?acc=MI0016412)), miR-3908 is supported by very few sequencing reads and no 3p arm has been annotated. Therefore, I suggest that the authors provide additional experimental support that this is actually a real miRNA through AGO-CLIP or Northern Blotting.*

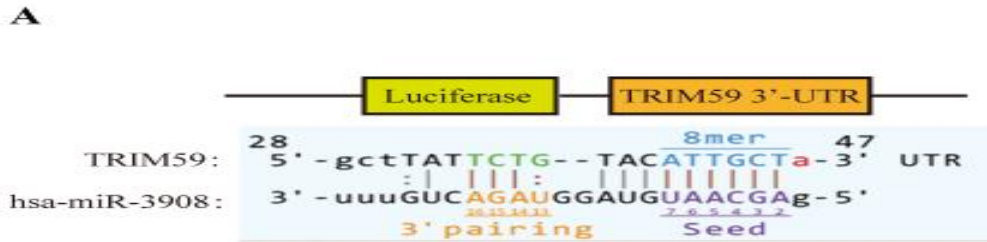
**RE:** Thanks for your critical suggestion. Although the sequencing reads of miR-3908 on miRbase is relatively low, the expression of microRNA is very tissue-specific, which may be highly expressed in specific conditions [4]. In our study, we detected the expression level of seven miRNAs in Fig 6B by RT-qPCR, and found that the CT value of miR-3908 was about 27 and the CT value of the important tumor suppressor miR-15b-5p [5] was about 26. The expression level of miR-3908 was close to the familiar microRNA miR-15b-5p, reflecting that miR-3908 should be a real miRNA. Besides, the PCR product of miR-3908 was also visualized by agarose gel electrophoresis in Fig 6H. And through literature review, we found the expression of miR-3908 in breast cancer [6] and glioma [7] was down-regulated and miR-3908 could inhibit the tumorigenicity of breast cancer and glioma by targeting AdipoR1 3'UTR, as shown below (Fig. R3), the sequence of miR-3908 in the literature was consistent with that in our results Fig 8A.

Based on the above findings, we believe that miR-3908 is actually a real miRNA.

Fig. R3 The screenshot of the reference paper.

Predicted consequential pairing of target region (top) and miRNA (bottom)	Site type	Context++ score	Context++ score percentile	Weighted context++ score	Conserved branch length	P <sub>CT</sub>
Position 28-34 of AdipoR1 3' UTR hsa-miR-3908	5' ... CAUGCAAAGAAUUGCAUUGCUG...                 3'          UUUGUCAGAUGGAUGUAACGAG	7mer-m8	-0.02	57	-0.02	0

Fig 8A The screenshot from our manuscript.



**Minor revisions**

1. Has *ebv-circLMP2A* been discovered before (e.g. is it present in any of the circRNA databases available). If so, please provide an identification number for this circRNA.

**RE:** Thanks for your thoughtful question. This *ebv-circLMP2A* with an exon5-to-exon3 backsplice junction has been reported in Erik K. Flemington ‘s paper published in 2018, titled “The Epstein Barr virus circRNAome”[8]. The author did not upload these data to the circRNA databases, but we can find this circRNA in the result section of this paper (Fig. R4).

Fig. R4 The screenshot of the reference paper.

ID	chrEBV_Akata_inverted		B-cell latency type I					B-cell latency type III				Gastric Cancer SNU719 YCCEL	
	Coord. 1	Coord. 2	Akata		Mutu I		Sav I	Mutu III	IB4	Jijoye	JY		
			Lat	React	Lat	React	Lat				Lat		
ebv-circLMP2_E8_E2	63427	65051	6	348	0	9	0	0	5	0	0	0	0
ebv-circLMP2_E7_E2	63427	64864	0	34	0	0	0	0	0	0	0	0	0
ebv-circLMP2_E5_E3	63729	64320	0	3	0	0	0	0	2	0	16	2	0
ebv-circLMP2_E8_E3	63729	65051	0	9	0	0	0	0	0	0	0	0	0
ebv-circLMP2_E5_E4	63909	64320	0	1	0	0	0	0	1	0	8	0	2
ebv-circLMP2_E4_E3	63729	64157	0	0	0	0	0	0	3	0	8	0	0
ebv-circLMP2_E1_RPMS1_E3	41172	58045	0	0	0	1	0	0	0	0	0	0	0
ebv-circLMP2_E1_A73_E3	51369	58045	0	3	0	8	0	0	0	0	0	0	0

2. In fig. 1G, the authors show that *ebv-circLMP2A* is associated with poor patient outcome. It would be interesting to also analyze whether the amount of *ebv-circLMP2A* relative to *LMP2A* shows the same. E.g. whether this is an independent effect of the circRNA and not due to the amount of virus overall.

**RE:** Thanks for your thoughtful suggestion. In 2019, two papers on circRNA studies were published simultaneously in the journal *Cell*, which showed that the expression of circRNA does not exactly correspond to the expression of corresponding linear RNA products and that circRNA functions as a non-coding RNA independent of its



linear transcript counterpart [9, 10]. Besides, within the 78 EBVaGC, there were 34 cases that exhibited positive LMP2A expression (43.6%) by immunohistochemistry in our previous study [11], which was consistent with previous report that only about half of EBVaGC cases expressed LMP2A [12]. Therefore, the expression of LMP2A was independent of the amount of Epstein-Barr virus overall. Based on the above studies, using the amount of ebv-circLMP2A relative to LMP2A for survival analysis may not be an appropriate choice.

3. In fig 6j, why are there T's and not U's present in the sequences for the circRNAs?

**RE:** Thanks for your critical comment. We are sorry that we made a mistake by writing U's as T's and we have replaced the T's with U's in revised Fig. 6J.

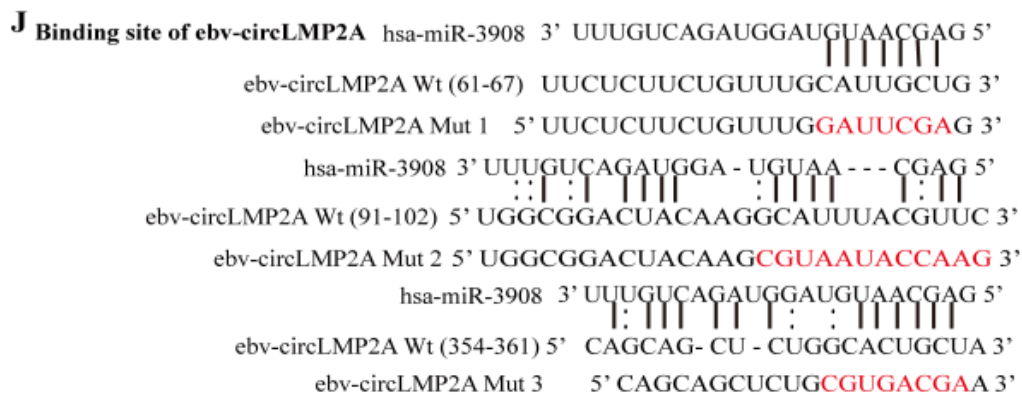


Fig. 6J Wild type and mutant ebv-circLMP2A and miR-3908 sequences were presented. Red fonts represented the mutant bases.

4. The authors need to assess/discuss that the changes in expression levels of a single circRNA generally will not lead to significant changes in competing miRNA binding sites relative to all of the corresponding miRNA binding sites present in all mRNAs (6). Moreover, endogenous stoichiometric relations between the miRNA-binding sites of ebv-circLMP2A and the corresponding mRNA target sites of the miRNA may not be mirrored in the overexpression experiments, which therefore may lack physiological relevance (7). Alternatively, the circRNA may not function as a sponge but through target RNA-directed miRNA degradation (TDMD). However, this

---

*usually requires extensive complementarity outside of the seed sequence (8,9) as exemplified by the long non-coding RNA, Cyrano, which induce destruction of miR-7 through a single highly conserved binding site of unusually high complementarity to miR-7 (10). Please discuss this in relation to the findings that it is binding site 1 and 3, but not site 2 that were critical for ebv-circLMP2A to sponge miR-3908.*

**RE:** We sincerely thank the reviewer for these important suggestions. Interestingly, although ebv-circLMP2A contained three predictive binding sites of miR-3908, only binding site1 and site3 were critical for ebv-circLMP2A to sponge miR-3908. This suggests that the presence of putative miRNA binding sites in circRNAs does not necessarily mean that the circRNA inhibits the miRNA by an absolute stoichiometric relationship at the real cell level. Indeed, the changes in expression levels of a single circRNA generally will not lead to significant changes in competing miRNA binding sites [13], and the circRNA may not inhibit the miRNA by functioning as a sponge but through target RNA-directed miRNA degradation (TDMD) which requires extensive complementarity outside of the seed sequence [14]. Here, we found that the site3 of ebv-circLMP2A and miR-3908 have extensive complementarity outside of the seed sequence, suggesting that the function of ebv-circLMP2A may also suppress miR-3908 through target RNA-directed miRNA degradation in addition to sponging it. We have added this content in the revised discussion section of the manuscript (page 15, line 383-395).

5. *When using Student's t-tests, did the authors make sure that the data followed a normal distribution?*

**RE:** Thanks for your critical comment. When using the Student's t-tests, we used SPSS to perform the Shapiro-wilk test to check and ensure that the data followed the normal distribution.

6. *In the discussion section, the authors should discuss why ebv-circLMP2A seems to have a much larger effect on patient survival than miR-3908.*

**RE:** Thanks for your thoughtful suggestion. CircRNAs are involved in the regulation



---

of cancer occurrence and development through multiple mechanisms, such as the adsorption and regulation of miRNA as a natural miRNA sponge, binding to transcription regulatory elements or interacting with proteins to regulate gene transcription [13]. A circRNA can have a tumour-suppressive or oncogenic function by acting as a miRNA sponge of multiple different miRNAs rather than by containing multiple sites for one particular miRNA. For example, the oncogenic circCCDC66 contains binding sites for several miRNAs, including miR-33b and miR-93, which both target the *MYC* oncogene [15]. In our study, we found ebv-circLMP2A exerts its function through attenuating the inhibitory function of miR-3908. However, other mechanisms of ebv-circLMP2A may also exist in promoting the progression of EBVaGC. For example, ebv-circLMP2A may also sponge other miRNAs, or regulate other genes expression by unknown mechanisms, etc. Therefore, it is reasonable that ebv-circLMP2A has a much larger effect on patient survival than miR-3908. We have added this content in the revised discussion section of the manuscript (page16, line 402-411).

7. *Throughout there are sentences that can be difficult to read and grammatical errors (e.g. "...but also responsible for tumor progression, metastasis and therapy-resistant", "...through long-term treatment of EBVaGC cell line SNU719 with 5-Fluorouraci in vivo passage" and "...shedding a light on the pathogenic function of ebv-circRNAs"). These sentences are from the abstract alone, but please read the entire manuscript carefully to address this.*

**RE:** Thanks for your critical comment. We have read the entire manuscript and corrected some sentences that can be difficult to read and grammatical errors in the revised manuscript.

**Replies to reviewer 3:**

Reviewer3:

*This report ascribes the initiation and maintenance of EBV-associated gastric carcinomas to an EBV circular RNA derived from the LMP2A locus. The authors*

---

*present studies to link the expression of circLMP2A to the development of phenotypic stem cell properties in infected cells. They propose that circLMP2A acts as a microRNA sponge to mechanistically effect this stemness by targeting the miR-3908/TRIM59/p53 axis. Studies on the functional biology of circular RNAs are rapidly accumulating and these non-coding RNAs represent an exciting, novel class of regulatory molecules in the cell. For this manuscript, however, some major points need to be addressed.*

*1. Figure 1: With the results that are presented in Figure 3, one would expect that cells from the third xenograft shown in 1B would also have relatively high ebv-circRNA expression levels similar to the fourth xenograft tested in 3B. Was this tested retrospectively?*

**RE:** Thanks for your critical comment. In this study, on the basis of the drug-resistance characteristics of cancer stem cells (CSCs), we established a method to enrich EBVaGC CSCs by a successive xenograft model under chemotherapy pressure. The growth curve of the passage xenografts was used to determine whether the xenografts model was terminated ((Fig. 1B). We defined the freshly purified single tumor cells obtained from each generation xenografts treated with 5-Fu as SNU-1st, SNU-2nd, SNU-3rd, SNU-4th cell, respectively. Then we detected the expression of ebv-circLMP2A in parental SNU719, SNU-1st, SNU-2nd, SNU-3rd, SNU-4th cells by RT-qPCR. We found that the expression of ebv-circLMP2A was gradually increased from SNU-1st to SNU-3rd cells and there was no significant difference between the SNU-3rd and SNU-4th cells (Fig. R1). Since most of SNU-3rd cells were used to construct the fourth generation of xenografts, we used the SNU-4th cells in the following study.

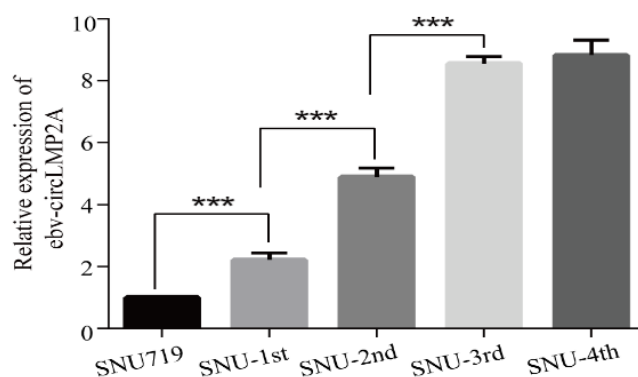


Fig. R1 The expression of ebv-circLMP2A in parental SNU719, SNU-1st, SNU-2nd, SNU-3rd, SNU-4th cells.

2. *Figure 2: Panel 2C is confusing, how can there be more than 1 backsplice junction in a circRNA? Similarly, please explain what the following text means, it is not clear: "a total of 262 distinct ebv-circRNAs candidates were found in these tissues and 144 of which contained at least two backspliced reads". In Panel D, how can the "length of ebv-circRNA" be determined without detailed identification, cloning, and sequencing of each circRNA species. Panel E looks like it wants to be a Venn diagram, except none of the areas overlap. Perhaps this should be expressed as a bar graph. Please also define "intergenic", "introns", "exon-intergenic" and "exons-intron" and how these were determined because exon-intergenic, intergenic, and intron backsplicing would be unusual. Panel F seems to say that there are at least 10 ebv genes that produce more than 1 circular RNA "isotype". What do the authors mean to show here?*

RE: Thanks for your critical comments. We are sorry for the confusion of your understanding of these results. In Figure 2C, Backsplice reads (the abscissa) represents the total number of reads that spanned backsplice junctions in circRNA sequencing, which was used as an absolute measure of circRNA abundance. The abundance of circRNA greater than 2 suggests that the circRNA is a real circular RNA rather than an error in circRNA sequencing process [16]. For example, "a total of 262 distinct ebv-circRNAs candidates were found in these tissues and 144 of which

---

contained at least two backspliced reads" means that there were 262 Epstein-Barr virus-derived circRNAs (ebv-circRNAs) with abundance greater than or equal to 1, of which 144 ebv-circRNAs had an abundance greater than 2. Similar data description also appears in Fig.1a, titled "Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs" [17]. In Panel D, strictly speaking, high-throughput sequencing methods cannot get the full length of circRNAs, but can determine the sequence of backsplice site of circRNAs. The "length of ebv-circRNA" is predicted based on the complete comparison between the reference EBV genome (Accession No. NC\_007605.1) and the sequence of backsplice site. For more detailed prediction methods, we can refer to the paper titled "Specific identification and quantification of circular RNAs from sequencing data" [18]. In Panel E, according to your suggestion, we have modified the Venn diagram to a bar graph (see Fig.2E in the revised manuscript). In addition, the full-length sequences of the circRNAs were predicted through a complete comparison between the reference EBV genome (Accession No. NC\_007605.1) and the sequence of splicing site, and "intergenic", "introns", "exon-intergenic" and "exons-intron" were defined according to the annotated information of the location of the full-length sequence of the circRNAs [16]. In Panel F, the "number of circRNAs produced from one gene" refers to that one gene could generate multiple ebv-circRNAs isotypes. The number in the Fig.2F represents the number of ebv-circRNAs isotypes derived from one EBV gene. Similar data description also appears in Fig.2a, titled "Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs" [17].

3. *Figure 3: The manuscript is highly focused on one particular circular "isoform" from the LMP2A gene locus: a 429 bp circular RNA with an exon5-to-exon3 backsplice junction. In Ungerleider N et al's 2018 paper (reference #26) low numbers of an LMP2A exon5-to-exon4 backsplice junction is detected in YCCEL1 cells (which is used extensively in this study as well), but no exon5-to-exon3 circulars. Were exon5-to-exon4 circLMP2A found or looked for using divergent primers? Panel 3C is*

---

*really problematic, there is not a lot of difference between RNase R treated and untreated samples for ebv-circLMP2A by PCR in each of the three cell types looked at. This suggests that the molecule being assayed for is not a true circular RNA. The statement in the text on page 8, ".....whereas ebv-circLMP2A was resistant to RNase R digestion" is not supported by the data. Further, Panel 3B indicates ebv-cirLMP2A is highly expressed in SNU-4th cells however, the PCR gel results from Panel 3C is not consistent with high levels of ebv-circLMP2A in SNU-4th compared to SNU719 or YCCELI. Panel E should have a control showing the efficiency of fractionation into cytoplasmic and nuclear samples.*

RE: Thanks for your critical comments. In our study, we established a xenograft model to enrich EBVaGC CSCs SNU-4th cells through long-term treatment of EBVaGC cell line SNU719 with 5-Fluorouraci (5-Fu) in mice vivo passage, and analyzed the difference of circRNAs expression between the fourth passage xenografts treated with 5-Fu and PBS by RNA-sequencing. In this RNA-sequencing results, we found only one circular "isoform" (LMP2A exon5-to-exon3) from the *LMP2A* gene locus (see Dataset EV1), consistent with the report in Ungerleider N et al's paper published in 2018 [8] that there was only one isoform (LMP2A exon5-to-exon3) from the *LMP2A* gene locus in SNU719 cells (Fig. R2). Therefore, we did not detect the expression of exon5-to-exon4 circLMP2A. But according to your suggestion, we analyzed the expression level of exon5-to-exon4 circLMP2A using the divergent primers (which from the Ungerleider N et al's 2018 paper) in SNU-4th and SNU719 cells by RT-qPCR, and no exon5-to-exon4 circLMP2A expression was detected.

In Fig.3C, we designed convergent primers to amplify *LMP2A* mRNA and divergent primers to amplify ebv-circLMP2A. There is no difference of the ebv-circLMP2A expression between the RNase R treated and untreated samples by PCR, suggesting that ebv-circLMP2A was resistant to RNase R digestion. As we know, circRNAs are formed from a covalently closed loop that can tolerate the degradation of RNase R. Therefore, this data suggests that the ebv-circLMP2A is a true circular RNA. Similar data description can also be found in Fig.2B in another paper, titled "Novel Role of

FBXW7 Circular RNA in Repressing Glioma Tumorigenesis” [19].

In Panel 3B, we detected the expression of ebv-circLMP2A in SNU-4th, SNU719, YCCEL1 cells by RT-qPCR, and analyzed the RT-qPCR data with the  $\Delta\Delta\text{CT}$  method. However, the data from Panel 3C were detected by PCR gel. Compared with PCR gel, RT-qPCR data are absolute measurement of ebv-circLMP2A abundance and have high sensitivity. Besides, the PCR gel was performed to prove the existence of ebv-circLMP2A, not to compare the expression differences, so the RNA samples of these three cells were not strictly balanced. Therefore, the expression differences of ebv-circLMP2A in SNU-4th, SNU719 and YCCEL1 cells in Panel 3C were not as obvious visually as those in Panel 3B. In Panel 3E, GAPDH and U6 were used as controls in cytoplasm and nucleus, respectively. We have added this content in the revised Figure legend of the manuscript (page 33, line 889- 890).

Fig. R2 The screenshot of the reference paper.

ID	chrEBV_Akata_inverted		B-cell latency type I						B-cell latency type III				Gastric Cancer	
	Coord. 1	Coord. 2	Akata		Mutu I		Sav I	Mutu III	IB4	Jijoye	JY	SNU719	YCCEL1	
			Lat	React	Lat	React	Lat							
ebv-circLMP2_E8_E2	63427	65051	6	348	0	9	0	0	5	0	0	0	0	
ebv-circLMP2_E7_E2	63427	64864	0	34	0	0	0	0	0	0	0	0	0	
ebv-circLMP2_E5_E3	63729	64320	0	3	0	0	0	0	2	0	16	2	0	
ebv-circLMP2_E8_E3	63729	65051	0	9	0	0	0	0	0	0	0	0	0	
ebv-circLMP2_E5_E4	63909	64320	0	1	0	0	0	0	1	0	8	0	2	
ebv-circLMP2_E4_E3	63729	64157	0	0	0	0	0	0	3	0	8	0	0	
ebv-circLMP2_E1_RPMS1_E3	41172	58045	0	0	0	1	0	0	0	0	0	0	0	
ebv-circLMP2_E1_A73_E3	51369	58045	0	3	0	8	0	0	0	0	0	0	0	

*Table1: This reviewer has serious concerns about the ability to detect ebv-circLMP2A from paraffin embedded samples (page16) particularly in samples of paraffin blocks from 2006. Table 1 indicates that the ebv-circLMP2A is found in every block and can be quantified to be low or high - this is not the experience in the field. Please also test for a cellular control circular RNA in each case.*

RE: Thanks for your critical comments. Circular RNAs have been reported exceptionally stable and can be detected from paraffin embedded samples [13]. For example, the expression of circPVT1 was detected in samples of paraffin blocks from 2007 [20]. In this study, we detected the expression level of ebv-circLMP2A in each paraffin embedded sample by RT-qPCR assays, and used the reference gene GAPDH for normalization of the ebv-circLMP2A expression. In addition, in fact, a total of 78 paraffin-embedded EBVaGC samples were collected in this study, of which follow-up

---

data were lost in 4 cases and reference gene GAPDH was not detected in 5 cases because the amount of RNA extracted from the samples was insufficient, so only 69 samples were included in clinicopathological and survival analysis. We have added this content in the Patient samples section of the revised Materials and Methods (see Page 18, line 447-453). Moreover, we also performed Sanger sequencing to confirm the distinct circularization junction sequence of ebv-circLMP2A which from RT-qPCR product (Fig. R3). Therefore, the expression level of ebv-circLMP2A in each paraffin embedded sample by RT-qPCR assays is reliable.

In Table1, ROC analysis was applied to distinguish the high and low expression of ebv-circLMP2A. The method has also been used to dichotomize the expression level of circRNA in other studies [20, 21]

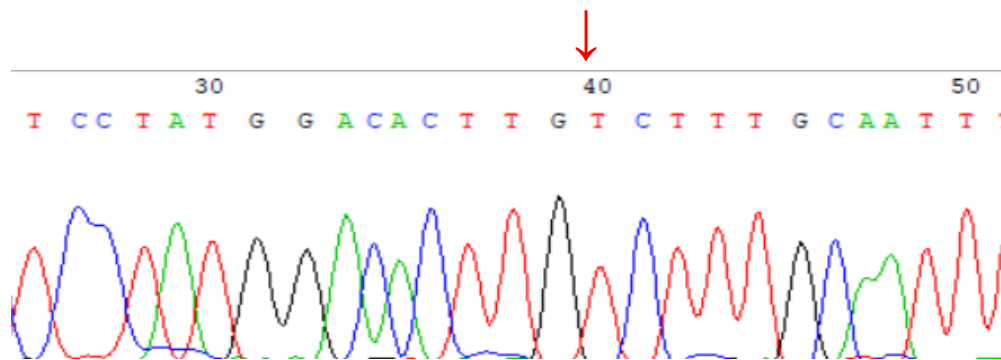


Fig. R3 The sequence of backsplice site of ebv-circLMP2A was validated by Sanger sequencing, red arrow represents the “head-to-tail” splicing sites of ebv-circLMP2A.

4. *Figure 4: For these series of experiments, where possible (for example 4C, 4J, 4L), the ebv-circLMP2A should also be quantitatively assay for by PCR in sh-control and sh-ebv-circLMP2A samples.*

**RE:** Thanks for your thoughtful suggestion. According to your suggestion, we detected the expression of ebv-circLMP2A in xenografts tissues obtained from *in vivo* tumorigenicity experiment by RT-qPCR assays (Fig. EV2F). The new data has been added in the Result section in the revised manuscript (see Page 9 line 226-228 and Page 10 line 253-255).

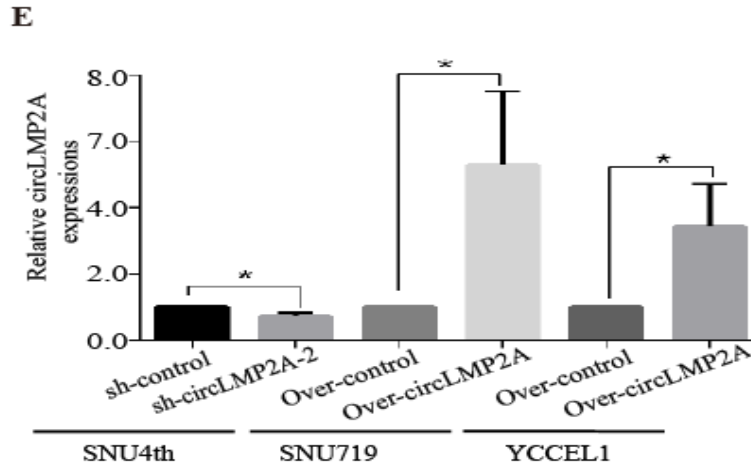


Fig. EV2F The expression of ebv-circLMP2A in the xenografts tissues.

5. *Figure5: For these sets of experiments it is critical that the ebv-circLMP2A over-expression system be confirmed. The circular molecule should be validated throughout its entire 429bp length to be intact by sequencing of PCR overlapping products. The sequence of the circularization junction needs to be particularly confirmed and stated. RNaseR treatment should be applied to confirm circularity of expression product.*

RE: Thanks for your critical comment. We agree with your opinion, northern blotting +/- RNase R treatment has been performed to confirm the product from the over-expression system is actually circular RNA (Fig. EV2D). We also performed Sanger sequencing to confirm the distinct circularization junction sequence of ebv-circLMP2A produced from the over-expression system (Fig. EV2E). The new data have been added in the Result section the revised manuscript (See page 9, line 241-page 10, line 244).



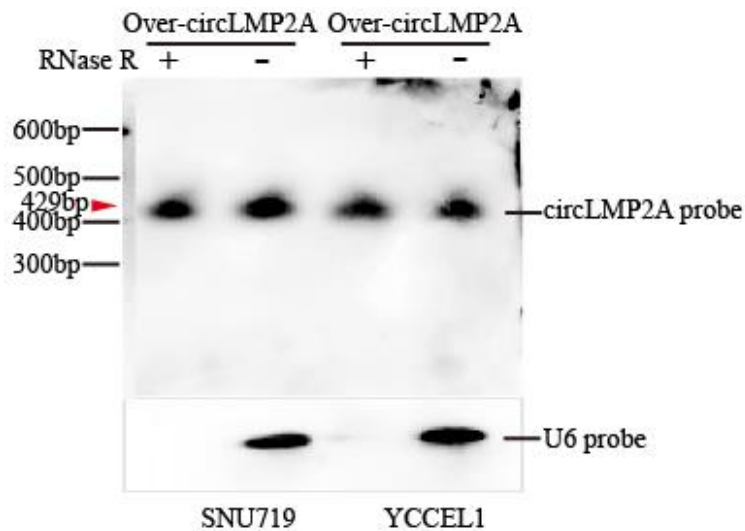


Fig. EV2D Northern blots for detecting ebv-circLMP2A in SNU719 and YCCEL1 cells (ebv-circLMP2A over-expression stable transfectants) treated with or without RNase R digestion

ebv-circLMP2A\_E5\_E3(429bp) sequence:

**TCTTTGCA**AATTTGCCTAACATGGAGGATTGAGGACCCACCTTTTAATTCTCT  
TCTGTTTGCATTGCTGGCCGCAGCTGGCGGACTACAAGGCATTTACGTTCT  
GGTGATGCTTGTGCTCCTGATACTAGCGTACAGAAGGAGATGGCGCCGTTT  
GACTGTTTGTGGCGGCATCATGTTTTTGGCATGTGTACTTGTCCATCGTC  
GACGCTGTTTTGCAGCTGAGTCCCCTCCTTGGAGCTGTAAGTGTGGTTTCC  
ATGACGCTGCTGCTACTGGCTTTCGTCTCTGGCTCTCTTCGCCAGGGGGC  
CTAGGTACTCTTGGTGCAGCCCTTTTAACATTGGCAGCAGCTCTGGCACTG  
CTAGCGTCACTGATTTGGGCACACTTAACTTGACTACAATGTTCCCTTCTCA  
TGCTCC**TATGGACACTTG**

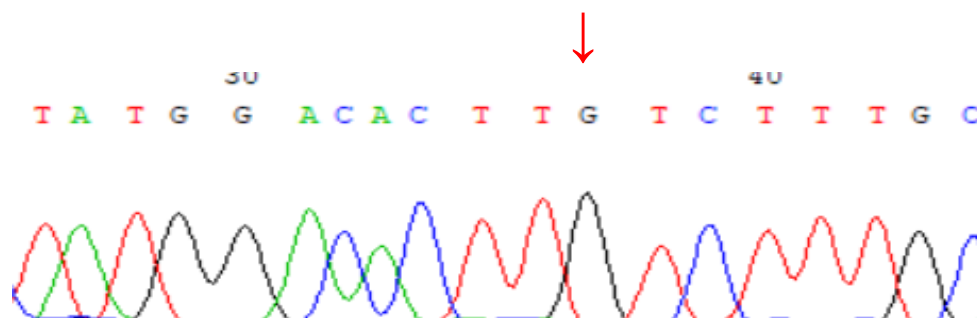


Fig. EV2E The sequence of backsplice site of ebv-circLMP2A was validated by Sanger sequencing, red arrow represents the “head-to-tail” splicing sites of ebv-circLMP2A.

6. *Figure6: Please describe the oligo probe used as the control for the circLMP2A*

---

*probe. Is the oligo control sequence the same length as the circLMP2A probe, and of an identical nucleic acid composition?*

**RE:** Thanks for your thoughtful comment. In biotin-coupled probe pull down assay, the sequence of the oligo probe against LacZ gene of bacteria used as control is: GCTGTATCGCTGGATCAAAT-biotin; and the sequence of circLMP2A probe for targeting the backsplice sequence of ebv-circLMP2A is: GATACCTGTGAACAGAAACG-biotin. The sequences of both the oligo and circLMP2A probes have the same length with different nucleic acids. We have added this information to the revised Materials and Methods section of the manuscript (See page 22, line 564-568).

## *7. Other Comments*

*1. Page 4, bottom para: " In 2018, our study and two other reports have proven....." It is an error to attribute the authors' previous study (reference #25) as a 2018 study.*

**RE:** Thanks for your critical comment. We have corrected this error in the revised Introduction section of the manuscript (see page 4, line 104-106).

*2. Abstract: the EBV-circLAMP2A is described as "highly expressed". This is not accurate based on the PCR band strength and lack of quantitative assays.*

**RE:** Thanks for your thoughtful comment. In this study, the RT-qPCR assay was performed to detect the expression of ebv-circLMP2A in SNU-4th, SNU719 and YCCEL1 cells, the expression level of ebv-circLMP2A in parental SNU719 cells and YCCEL1 cells were indeed very low, but it was significantly higher in our enriched EBVaGC CSCs SNU-4th cells (Fig.3B). We have modified this description in the revised Abstract section of the manuscript (See page 2, line 42).

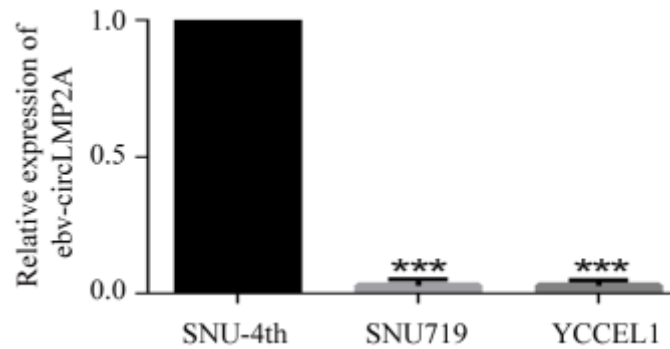


Fig.3B Real-time PCR analysis of the expression of ebv-circLMP2A in SNU-4th, SNU719 and YCCEL1 cells.

*3. Introduction: Authors spend a lot of verbiage on the uncertainty of the role of EBV to the pathogenesis of EBVaGC. Actually, there is not a lot of uncertainty about this, particularly when EBV is clonal by Terminal Repeat analysis in every EBVaGC tumor cell.*

RE: Thanks for your critical comment. We agree with your opinion and have modified this content in the revised Introduction section of the manuscript (See page 4, line 90-93).

That's all for the reply to the reviewer's comments. All the authors read and approved the revised manuscript.

Thank you very much for your attention and consideration. We are looking forward to your decision.

Sincerely yours,

Chun-kui Shao, MD, PhD

April 10, 2020

## References

1. Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. *Nature* 414, 105-111.
2. Koren, E., and Fuchs, Y. (2016). The bad seed: Cancer stem cells in tumor development and

- 
- resistance. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy* 28, 1-12.
3. Burke, H.B. (2018). Overall Survival vs Disease-Specific Survival. *JAMA oncology* 4, 586.
  4. Mishra, S., Yadav, T., and Rani, V. (2016). Exploring miRNA based approaches in cancer diagnostics and therapeutics. *Critical reviews in oncology/hematology* 98, 12-23.
  5. Sun, G., Yan, S., Shi, L., Wan, Z., Jiang, N., Li, M., and Guo, J. (2015). Decreased Expression of miR-15b in Human Gliomas is Associated with Poor Prognosis. *Cancer biotherapy & radiopharmaceuticals* 30, 169-173.
  6. Li, Y., Shan, F., and Chen, J. (2017). Lipid raft-mediated miR-3908 inhibition of migration of breast cancer cell line MCF-7 by regulating the interactions between AdipoR1 and Flotillin-1. *World journal of surgical oncology* 15, 69.
  7. Liu, X., Chen, J., and Zhang, J. (2017). AdipoR1-mediated miR-3908 inhibits glioblastoma tumorigenicity through downregulation of STAT2 associated with the AMPK/SIRT1 pathway. *Oncology reports* 37, 3387-3396.
  8. Ungerleider, N., Concha, M., Lin, Z., Roberts, C., Wang, X., Cao, S., Baddoo, M., Moss, W.N., Yu, Y., Seddon, M., et al. (2018). The Epstein Barr virus circRNAome. *PLoS pathogens* 14, e1007206.
  9. Vo, J.N., Cieslik, M., Zhang, Y., Shukla, S., Xiao, L., Zhang, Y., Wu, Y.M., Dhanasekaran, S.M., Engelke, C.G., Cao, X., et al. (2019). The Landscape of Circular RNA in Cancer. *Cell* 176, 869-881 e813.
  10. Chen, S., Huang, V., Xu, X., Livingstone, J., Soares, F., Jeon, J., Zeng, Y., Hua, J.T., Petricca, J., Guo, H., et al. (2019). Widespread and Functional RNA Circularization in Localized Prostate Cancer. *Cell* 176, 831-843 e822.
  11. Zhang, Y.W., Zhao, X.X., Tan, C., Zhang, Z.G., Jiang, Y., Chen, J.N., Wei, H.B., Xue, L., Li, H.G., Du, H., et al. (2015). Epstein-Barr virus latent membrane protein 2A suppresses the expression of HER2 via a pathway involving TWIST and YB-1 in Epstein-Barr virus-associated gastric carcinomas. *Oncotarget* 6, 207-220.
  12. Sugiura, M., Imai, S., Tokunaga, M., Koizumi, S., Uchizawa, M., Okamoto, K., and Osato, T. (1996). Transcriptional analysis of Epstein-Barr virus gene expression in EBV-positive gastric carcinoma: unique viral latency in the tumour cells. *British journal of cancer* 74, 625-631.
  13. Kristensen, L.S., Andersen, M.S., Stagsted, L.V.W., Ebbesen, K.K., Hansen, T.B., and Kjems, J. (2019). The biogenesis, biology and characterization of circular RNAs. *Nature reviews. Genetics* 20, 675-691.
  14. Ghini, F., Rubolino, C., Climent, M., Simeone, I., Marzi, M.J., and Nicassio, F. (2018). Endogenous transcripts control miRNA levels and activity in mammalian cells by target-directed miRNA degradation. *Nature communications* 9, 3119.
  15. Hsiao, K.Y., Lin, Y.C., Gupta, S.K., Chang, N., Yen, L., Sun, H.S., and Tsai, S.J. (2017). Noncoding Effects of Circular RNA CCDC66 Promote Colon Cancer Growth and Metastasis. *Cancer research* 77, 2339-2350.
  16. Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333-338.
  17. Zheng, Q., Bao, C., Guo, W., Li, S., Chen, J., Chen, B., Luo, Y., Lyu, D., Li, Y., Shi, G., et al. (2016). Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging

- 
- multiple miRNAs. *Nature communications* 7, 11215.
18. Cheng, J., Metge, F., and Dieterich, C. (2016). Specific identification and quantification of circular RNAs from sequencing data. *Bioinformatics* 32, 1094-1096.
  19. Yang, Y., Gao, X., Zhang, M., Yan, S., Sun, C., Xiao, F., Huang, N., Yang, X., Zhao, K., Zhou, H., et al. (2018). Novel Role of FBXW7 Circular RNA in Repressing Glioma Tumorigenesis. *Journal of the National Cancer Institute* 110.
  20. Chen, J., Li, Y., Zheng, Q., Bao, C., He, J., Chen, B., Lyu, D., Zheng, B., Xu, Y., Long, Z., et al. (2017). Circular RNA profile identifies circPVT1 as a proliferative factor and prognostic marker in gastric cancer. *Cancer letters* 388, 208-219.
  21. Li, Y., Zheng, F., Xiao, X., Xie, F., Tao, D., Huang, C., Liu, D., Wang, M., Wang, L., Zeng, F., et al. (2017). CircHIPK3 sponges miR-558 to suppress heparanase expression in bladder cancer cells. *EMBO reports* 18, 1646-1659.

Dear Dr. Shao,

Thank you for the submission of your revised manuscript to our editorial offices. We have now received the reports from the three referees that were asked to re-evaluate your study, you will find below. As you will see, the referees now support the publication of your study in EMBO reports. They have some further suggestions to improve the manuscript, we ask you to address in a final revised version of the manuscript.

Further, I have these editorial requests:

- As indicated also by the referees, please have the manuscript carefully proofread by a native speaker. There are still many grammatical errors. Our publisher also offers a manuscript editing service:

<https://wileyeditingservices.com/en/english-language-editing/>

- Please use this title in the submission system and in the manuscript text files:

Epstein-Barr virus derived circular RNA LMP2A induces cancer stemness in EBV-associated gastric cancer

- Please provide the abstract written in present tense.

- Please remove the bold titles from the title page (Title page, Title of paper, etc. ...), and also 'Word counts', figure and expanded view information. We do not need this here.

- The EV tables have to be uploaded as separate files, and need legends in the manuscript text file (below the EV figure legends). However, Table EV2 needs to be a dataset. Please upload this as dataset (Dataset EV2, as excel file with a legend on the first TAB). There are then 4 EV tables in total.

- Dataset EV1 also needs a legend. Please provide this on the first TAB of the excel file.

- In the reference list, please provide the journal names in italics.

- Please provide the scale bars in the microscopic images in black (or white if the background is dark). The red you used is too flashy.

- Please also add scale bars to all the images of the colony formation assays.

- Please provide the Western blot images less stretched (see e.g. 1H, 1J or 4D). Please leave the images as un-modified as possible. This just does not look natural. Please show the Western blot data as it looked after image acquisition or on the film.

- Please centre the images better in Fig. EV4E/F (no membrane edges should be visible, and the bands should show up in the middle of the boxes). See e.g. the boxes for Vimentin or BMi1 in Fig. EV4E.

- The labeling of many figures is also rather messy. See e.g. Figs. 4C/D, 5H or 6C. It should be clearly distinguishable which lane is labelled, which text belongs to which lane, and labels should not reach

into neighbouring figures. Please find a way to arrange the labels better in all the figures, maybe shortening the labels and using more abbreviations. See also:

[http://wol-prod-cdn.literatemonline.com/pb-assets/embo-site/EMBOPress\\_Figure\\_Guidelines\\_061115-1561436025777.pdf](http://wol-prod-cdn.literatemonline.com/pb-assets/embo-site/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf)

- Thank you for providing the source data, in particular of the Western blots. However, we need the source data as pdf, xls or doc files. Please provide the WB source data for one figure in one file, include size markers for scans of entire gels, label the scans with figure and panel number, and upload one PDF file per figure.

- Please indicate in the data availability section in detail which datasets have been deposited.

- Please use bigger fonts for the synopsis image (please find it attached in the size it will appear online), and do not capitalize the words (also not those in the blue box).

- Finally, please find attached a word file of the manuscript text (provided by our publisher) with changes we ask you to include in your final manuscript text, and some queries, we ask you to address. I think you already addressed these, but please double check. Please provide your final manuscript file with track changes, in order that we can see any modifications done.

In addition I would need from you:

- a short, two-sentence summary of the manuscript
- two to three bullet points highlighting the key findings of your study

I look forward to seeing the final revised version of your manuscript when it is ready. Please let me know if you have questions regarding the revision.

Kind regards,

Achim

-----  
Achim Breiling  
Editor  
EMBO Reports  
-----

Referee #1:

Authors have addressed the points raised by the previous reviews in a reasonable way and this has improved the paper. I now only have two minor points for correction:

1. Line 206 should read: expression correlated with a worse prognosis in EBVaGC patients.
2. Throughout: although the meaning is clear, there are many places where the English language would be improved by having the manuscript corrected by a native English speaker.

-----  
Referee #2:

The authors have addressed the majority of my concerns sufficiently and the manuscript has improved substantially. However, the manuscript should be corrected by a native English speaking person as there are still problems with quite a few sentences.

-----  
Referee #3:

The authors adequately address the concerns from my first review. There remain grammatical errors that should be corrected during the editorial process.



Dear Dr. Achim,

Thank you for reviewing and providing helpful comments on our manuscript entitled "Epstein-Barr virus derived circular RNA LMP2A induces cancer stemness in EBV-associated gastric cancer" (ID: EMBOR-2019-49689V2).

At your request, the manuscript has been carefully corrected by a native English speaker. You will see the editing certificate for this manuscript in the attachment below.

We also made some changes to the Western blot images to try to make them look more natural and better in all the figures. Moreover, according to your requirement, we have uploaded the source data of the Western blots for one figure in one file.

Finally, we have studied these requests carefully and have made revision to the manuscript seriously, and uploaded the final revised version of the manuscript to the submission system.

Thank you very much for your attention and consideration. We are looking forward to your decision.

Kind regards,

Li ping Gong

Dear Dr. Shao,

Thank you for the submission of your revised manuscript to our editorial offices, and for providing explanations for the duplications we observed in some of the figure panels, and revised figures. After reviewing the source data and the corrected figures, we are prepared to accept that the revised figures show the correct data, and we do not believe it is necessary to involve your institute at this time. Nevertheless, there are remaining editorial requests we ask you to address in a further revised manuscript.

- Please upload the revised figures without the duplications.
- Please provide the Western blot images less stretched and cropped in the final figures. Some panels show the Western blot data extremely stretched and too tightly cropped (in particular in Figs. 4D, 5D, 7C and EV4E/F - less so, but still too much in Fig. 1H/J and 8G/H/I). As the source data images show these blots unmodified, it is mostly impossible to compare the panels and to decide if indeed the same data is shown. Please show the Western blot panels not stretched, and not that tightly cropped. Please show the Western blot data as it looked after image acquisition or on the film, i.e. as in the source data. The images in Fig. 8K look fine. Please show all Western blots like this.
- It seems that the source data for Fig. 1H (Sox2) and Fig. 1J (E-cad) does not fit to the figure panel. Please check. Moreover, the actin panel in 1J has no source data. Please add this.
- > In general, please carefully revisit the Western blot source data and make sure that for each panel source data is provided, and that it is the correct one!
- We need ONE pdf file of source data per figure! Please combine the different source data files. Please do not provide source data as .rar files.
- In Fig. 2E the writing below the x-axis is slightly covered. Please check.
- Please separate the mouse images in Figs. 5K and EV3E with white lines. Do not splice the images together.
- It seems the blot shown in Fig. EV2D is spliced together from 3 different images. Please indicate this by black dividing lines.
- The labeling of many figures is still too messy. See e.g. Figs. 7C/D/E or Fig. EV4 E/F/G/H/I. It should be clearly distinguishable which lane is labelled, which text belongs to which lane, and labels should not reach into neighbouring figures. Please find a way to arrange the labels better in all the figures, maybe shortening the labels and using more abbreviations. See also:  
[http://wol-prod-cdn.literatunonline.com/pb-assets/embosite/EMBOPress\\_Figure\\_Guidelines\\_061115-1561436025777.pdf](http://wol-prod-cdn.literatunonline.com/pb-assets/embosite/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf)
- Please use bigger fonts for the synopsis image (please find it attached in the size it will appear online), and do not capitalize the words (also not those in the blue box). Please make sure the final file (in jpeg or tiff format) has the exact width of 550 pixels and a height of not more than 400 pixels.

I look forward to seeing the final revised version of your manuscript when it is ready. Please let me

know if you have questions regarding the revision.

Kind regards,

Achim

-----  
Achim Breiling  
Editor  
EMBO Reports  
-----

Chun Shao  
The Third Affiliated Hospital, Sun Yat-sen University  
Department of Pathology  
No.600 Tianhe Road, Guangzhou 510630, China  
Guangzhou, Guangdong 510630  
China

Dear Dr. Shao,

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. As you are aware, this File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: [emboreports@embo.org](mailto:emboreports@embo.org)]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

Should you be planning a Press Release on your article, please get in contact with [emboreports@wiley.com](mailto:emboreports@wiley.com) as early as possible, in order to coordinate publication and release dates.

Thank you again for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Yours sincerely,

Achim Breiling  
Editor  
EMBO Reports

\*\*\*\*\*

THINGS TO DO NOW:

You will receive proofs by e-mail approximately 2-3 weeks after all relevant files have been sent to

our Production Office; you should return your corrections within 2 days of receiving the proofs.

Please inform us if there is likely to be any difficulty in reaching you at the above address at that time. Failure to meet our deadlines may result in a delay of publication, or publication without your corrections.

All further communications concerning your paper should quote reference number EMBOR-2019-49689V4 and be addressed to [emboreports@wiley.com](mailto:emboreports@wiley.com).

Should you be planning a Press Release on your article, please get in contact with [emboreports@wiley.com](mailto:emboreports@wiley.com) as early as possible, in order to coordinate publication and release dates.

**YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓**

**PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER**

Corresponding Author Name: Chun-kui Shao

Journal Submitted to: EMBO reports

Manuscript Number: EMBOR-2019-49689V1

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

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

#### A- Figures

##### 1. Data

**The data shown in figures should satisfy the following conditions:**

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

##### 2. Captions

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

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

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

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

#### B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	For animal studies, a total of 36 NOD/SCID mice were randomly divided into six groups (six mice per group) .
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	For animal studies, female 4-week-old NOD/SCID mice and the weight is about 18-24g were chosen.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	For human sample studies, a total of 78 paraffin-embedded EBVaGC samples were collected in this study, of which follow-up data were lost in 4 cases and reference gene GAPDH was not detected in 5 cases because the amount of RNA extracted from the samples was insufficient, so only 69 samples were included in clinicopathological and survival analysis.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	For animal studies, firstly, we chosen female 4-week-old NOD/SCID mice and the weight is about 18-24g. Then, the mice were randomly divided into six groups (six mice per group).
For animal studies, include a statement about randomization even if no randomization was used.	For animal studies, animals are randomly grouped.
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	All experiments and data analysis were done by two investigators in a blind method.
4.b. For animal studies, include a statement about blinding even if no blinding was done	For animal studies, all experiments were done by two investigators in a blind method.
5. For every figure, are statistical tests justified as appropriate?	Yes, all statistical tests are justified as appropriate for every figure. Data differences between two groups were analyzed using the Student's t-test or Chi-square test. The Pearson's correlation coefficient analysis was carried out to analyze the correlations. ROC analysis was used to generate a cut-off value to distinguish between high and low expression of ebv-circLMP2A. Survival analysis was performed by Kaplan-Meier curves and log-rank test for significance
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	When using the Student's t-tests, we used SPSS to perform the Shapiro-wilk test to check and ensure that the data followed the normal distribution.
Is there an estimate of variation within each group of data?	Yes, we used SPSS for analysis of variance to estimate the difference in each group of data.
Is the variance similar between the groups that are being statistically compared?	Yes, the variance was similar between the groups that were statistically compared.

#### C- Reagents

#### USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

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

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

<http://biomodels.net/>

<http://biomodels.net/miriam/>  
<http://jij.biochem.sun.ac.za>  
[http://oba.od.nih.gov/biosecurity/biosecurity\\_documents.html](http://oba.od.nih.gov/biosecurity/biosecurity_documents.html)  
<http://www.selectagents.gov/>

<p>6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).</p>	<table border="1"> <thead> <tr> <th>Antibody</th> <th>Dilution</th> <th>Species</th> <th>Cat No.</th> <th>Sources</th> </tr> </thead> <tbody> <tr> <td>β-actin</td> <td>1:1000</td> <td>Rabbit</td> <td>#4970</td> <td>Cell Signaling Technology</td> </tr> <tr> <td>E-cad</td> <td>1:1000</td> <td>Rabbit</td> <td>#3195</td> <td>Cell Signaling Technology</td> </tr> <tr> <td>Vimentin</td> <td>1:1000</td> <td>Rabbit</td> <td>#5741</td> <td>Cell Signaling Technology</td> </tr> <tr> <td>Mmp7</td> <td>1:1000</td> <td>Rabbit</td> <td>ab205525</td> <td>Abcam</td> </tr> <tr> <td>Sox2</td> <td>1:1000</td> <td>Rabbit</td> <td>#3579</td> <td>Cell Signaling Technology</td> </tr> <tr> <td>Klf4</td> <td>1:1000</td> <td>Rabbit</td> <td>ab215036</td> <td>Abcam</td> </tr> <tr> <td>Bmi1</td> <td>1:10000</td> <td>Rabbit</td> <td>ab126783</td> <td>Abcam</td> </tr> <tr> <td>ABCG2</td> <td>1:5000</td> <td>Rabbit</td> <td>ab108312</td> <td>Abcam</td> </tr> <tr> <td>Oct4</td> <td>1:1000</td> <td>Rabbit</td> <td>#2750</td> <td>Cell Signaling Technology</td> </tr> <tr> <td>Mrp1</td> <td>1:500</td> <td>Mouse</td> <td>sc-365635</td> <td>Santa Cruz Biotechnology</td> </tr> <tr> <td>TRIM59</td> <td>1:500</td> <td>Rabbit</td> <td>ab69639</td> <td>Abcam</td> </tr> <tr> <td>P53</td> <td>1:1000</td> <td>Rabbit</td> <td>#2572</td> <td>Cell Signaling Technology</td> </tr> </tbody> </table>	Antibody	Dilution	Species	Cat No.	Sources	β-actin	1:1000	Rabbit	#4970	Cell Signaling Technology	E-cad	1:1000	Rabbit	#3195	Cell Signaling Technology	Vimentin	1:1000	Rabbit	#5741	Cell Signaling Technology	Mmp7	1:1000	Rabbit	ab205525	Abcam	Sox2	1:1000	Rabbit	#3579	Cell Signaling Technology	Klf4	1:1000	Rabbit	ab215036	Abcam	Bmi1	1:10000	Rabbit	ab126783	Abcam	ABCG2	1:5000	Rabbit	ab108312	Abcam	Oct4	1:1000	Rabbit	#2750	Cell Signaling Technology	Mrp1	1:500	Mouse	sc-365635	Santa Cruz Biotechnology	TRIM59	1:500	Rabbit	ab69639	Abcam	P53	1:1000	Rabbit	#2572	Cell Signaling Technology
Antibody	Dilution	Species	Cat No.	Sources																																																														
β-actin	1:1000	Rabbit	#4970	Cell Signaling Technology																																																														
E-cad	1:1000	Rabbit	#3195	Cell Signaling Technology																																																														
Vimentin	1:1000	Rabbit	#5741	Cell Signaling Technology																																																														
Mmp7	1:1000	Rabbit	ab205525	Abcam																																																														
Sox2	1:1000	Rabbit	#3579	Cell Signaling Technology																																																														
Klf4	1:1000	Rabbit	ab215036	Abcam																																																														
Bmi1	1:10000	Rabbit	ab126783	Abcam																																																														
ABCG2	1:5000	Rabbit	ab108312	Abcam																																																														
Oct4	1:1000	Rabbit	#2750	Cell Signaling Technology																																																														
Mrp1	1:500	Mouse	sc-365635	Santa Cruz Biotechnology																																																														
TRIM59	1:500	Rabbit	ab69639	Abcam																																																														
P53	1:1000	Rabbit	#2572	Cell Signaling Technology																																																														
<p>7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.</p>	<p>EBV-positive gastric cell lines SNU719 and YCCEL1 cells were purchased from Korean Cell Line Bank. We used the One step constant temperature Mycoplasma test kit(Shanghai Yise Medical Technology, Shanghai, China) to test and ensure that there is no mycoplasma contamination in the SNU4-th, SNU719 and YCCEL1 cell culture supernatant.The STR profiling of the SNU719 cell line was supplied in additional supplementary materia</p>																																																																	

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

#### D- Animal Models

<p>8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.</p>	<p>For animal studies, female 4-week-old NOD/LtSz-scid/scid (NOD/SCID) mice the weight is about 18-24g.The mice were purchased from Model Animal Research Center of Nanjing University, Nanjing, China. Details of housing and husbandry conditions:  a. Housing (type: specific pathogen free (SPF), IVC cage,3 mice/ cage companions)  b. Husbandry conditions (sterile environment with a constant temperature of 22±1°C, humidity of 55% +/- 15,12 hours of automatic lighting,autoclaved corn cob padding,drink reverse osmosis water and Co60 sterilized formula feed).  c. Welfare-related assessments:Follow the "3R" principle of animal welfare in the process of experiment, the "replacement" and "reduction" and "refinement".</p>
<p>9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.</p>	<p>NA</p>
<p>10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.</p>	<p>All animal studies were performed in accordance with the institutional ethics guidelines for the animal experiments which were approved by the Experimental Animal Ethics Committee of the Third Affiliated Hospital, Sun Yat-sen University.Follow the "3R" principle of animal welfare in the process of experiment, the "replacement" and "reduction" and "refinement".</p>

#### E- Human Subjects

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

#### F- Data Accessibility

<p>18. Provide a "Data Availability" section at the end of the Materials &amp; Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'.</p> <p>Data deposition in a public repository is mandatory for:  a. Protein, DNA and RNA sequences  b. Macromolecular structures  c. Crystallographic data for small molecules  d. Functional genomics data  e. Proteomics and molecular interactions</p>	<p>Yes, we have provided a "Data availability" section at the end of Materials &amp; Methods, as shown below.  Data availability  The data reported in this paper have been deposited in GEO: GSE145894 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE145894).</p>
<p>19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right)).</p>	<p>Yes, we have provided a datasets in the manuscript as a supplementary document (see Dataset EV1 in the manuscript)</p>
<p>20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).</p>	<p>NA</p>
<p>21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.</p>	<p>NA</p>

#### G- Dual use research of concern

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